



**FACULTAD DE INGENIERÍA, ARQUITECTURA Y
URBANISMO**

**ESCUELA PROFESIONAL DE INGENIERÍA DE
SISTEMAS**

TESIS

**COMPARACIÓN DE MÁQUINA DE SOPORTE
VECTORIAL Y REGRESIÓN LOGÍSTICA EN LA
PREDICCIÓN DE MOROSIDAD DE CUOTAS
SOCIALES DEL COLEGIO DE INGENIEROS DEL
PERÚ CONSEJO DEPARTAMENTAL LAMBAYEQUE**

**PARA OPTAR EL TÍTULO PROFESIONAL DE
INGENIERO DE SISTEMAS**

Autor:

Bach. Huamán Bernilla Jefferson Napoleón

Asesor:

Mg. Mejia Cabrera Heber Ivan

Línea de Investigación:

Infraestructura, Tecnología y Medio Ambiente

Pimentel, Perú

2020



DEDICATORIA

A:

Dios, por darme la oportunidad que cada instante de mi vida cuenta con su guía, fortaleza e iluminación y por haberme rodeado de personas de bien.

Mis padres, por haberme brindado todo su apoyo, sus enseñanzas, sus sabios consejos, por la motivación constante a lograr mis objetivos trazados, pero más que nada, por su amor.

Los emprendedores, que luchan cada día por ver realizado sus sueños.



AGRADECIMIENTO

A mis padres, por su comprensión, paciencia y ánimo.

A mi asesor, por su orientación y colaboración en el desarrollo de este trabajo de investigación.

A mi equipo de investigación, por haber contribuido al desarrollo del trabajo de investigación con proactividad y dinamismo.



RESUMEN

La complejidad de la gestión de la morosidad se ha incrementado durante los últimos años, generando en algunas organizaciones públicas y privadas, grandes pérdidas económicas. Para reducir este tipo de riesgo muchas organizaciones optaron por técnicas tradicionales como las estadísticas para optimizar su toma de decisiones, pero con la evolución de la tecnología y el incremento de datos de sus usuarios están migrando a nuevas técnicas como las de aprendizaje automático que en otros contextos han mostrado buenos resultados.

Por este motivo, en el presente trabajo se comparó una técnica de aprendizaje automático y una técnica clásica, con la finalidad de determinar que técnica es más eficiente en la predicción de la morosidad de cuotas sociales en el Colegio de Ingenieros del Perú Consejo departamental de Lambayeque, organización que afronta un incremento paulatino de la morosidad que afecta la buena toma de decisiones y no beneficia a los colegiados.

El punto de partida empieza con la selección de las técnicas a comparar para ello se efectuó un análisis exhaustivo a trabajos realizados en un contexto a fin a la morosidad como son los riesgos de crédito, con ello se elaboró un top de técnicas en función de su precisión. De acuerdo al top se seleccionaron las técnicas máquina de soporte vectorial y regresión logística.

Para efectuar la comparación de las técnicas se recopiló datos históricos de los colegiados cuya información se obtuvo de fuentes internas y externas a la organización. Posteriormente los datos recopilados por medio del proceso de extracción, transformación y carga (ETL) se limpió y estandarizó obteniéndose datos concisos y relevantes. Finalmente se aplicó las técnicas predictivas cuyos resultados son favorables para la máquina de soporte vectorial en comparación con la regresión logística. Concluyendo que la técnica máquina de soporte vectorial es más eficiente para predecir la morosidad de cuotas sociales en el Colegio de Ingenieros del Perú Consejo departamental de Lambayeque.

PALABRAS CLAVES: Técnicas, Predicción, Máquina de Soporte Vectorial, Regresión Logística, Aprendizaje automático.

ABSTRAC

The complexity of delinquency management has increased in recent years, generating large economic losses in some public and private organizations. To reduce this type of risk, many organizations opted for traditional techniques such as statistics to optimize their decision-making, but with the evolution of technology and the increase in user data they are migrating to new techniques such as machine learning than in others. Contexts have shown good results.

For this reason, in this paper we compared a machine learning technique and a classic technique, in order to determine which technique is more efficient in predicting late payments of social quotas in the College of Engineers of Peru Departmental Council of Lambayeque, an organization that faces a gradual increase in delinquency that affects good decision-making and does not benefit members.

The starting point begins with the selection of the techniques to be compared for this purpose, an exhaustive analysis was carried out on work carried out in a context to end delinquencies such as credit risks, with which a top of techniques was developed based on their precision. According to the top, the vector support machine and logistic regression techniques were selected.

To compare the techniques, historical data was collected from the members whose information was obtained from internal and external sources of the organization. Subsequently, the data collected through the process of extraction, transformation and loading (ETL) was cleaned and standardized obtaining concise and relevant data. Finally, predictive techniques were applied whose results are favorable for the vector support machine compared to logistic regression. Concluding that the vector support machine technique is more efficient for predicting late payments of social quotas at the College of Engineers of Peru Departmental Council of Lambayeque.

KEYWORD: Techniques, Prediction, Vector Support Machine, Logistic Regression, Machine Learning.



ÍNDICE

RESUMEN	4
ABSTRAC.....	5
CAPITULO I: INTRODUCCIÓN	11
1.1. Realidad problemática.....	11
1.2. Antecedentes de Estudio.	13
1.2.1. Antecedentes de la Investigación	13
1.2.2. Estado del Arte.....	18
1.3. Teorías relacionadas al tema.	21
1.3.1. Inteligencia de Negocios	21
1.3.2. Aprendizaje Automático	25
1.3.3. Técnicas de Aprendizaje Automático	27
1.3.4. Máquinas de Soporte Vectorial.....	29
1.3.5. Regresión Logística.....	37
1.3.6. Evaluación de las técnicas computacionales y estadísticas.....	40
1.3.7. Archivos .csv.....	44
1.3.8. Librería para aprendizaje automático.....	44
1.3.9. Definición de términos básicos.	45
1.4. Formulación del Problema.	45
1.5. Justificación e importancia del estudio.	45
1.6. Hipótesis.....	46
1.7. Objetivos de la investigación	46
1.7.1. Objetivo General	46
1.7.2. Objetivos Específicos.....	46
CAPITULO II: MATERIAL Y MÉTODO.....	47
2.1. Tipo y diseño de la investigación.....	47
2.2. Población y muestra.	47
2.3. Variables, Operacionalización.	48
2.4. Técnicas e instrumentos de recolección de datos.....	50
2.5. Procedimiento de análisis de datos	50
2.6. Criterios éticos.	52



2.7. Criterios de Rigor científico.....	52
2.8. Modelo Propuesto	53
2.8.1. Selección de técnicas predictivas	54
2.8.2. Procesamiento de los datos históricos de los colegiados	54
2.8.3. Proceso de Extracción, transformación y carga (ETL) para preparación de datos de entrenamiento.....	62
2.8.4. Implementación de la Máquina de Soporte Vectorial y Regresión Logística.....	75
CAPÍTULO III: RESULTADOS	90
3.1. Resultados en Tablas y Figuras.....	90
3.2. Discusión de resultados.....	97
CAPÍTULO IV: CONCLUSIONES Y RECOMENDACIONES	100
4.1. Conclusiones	100
4.2. Recomendaciones.....	101
REFERENCIAS.....	102
ANEXOS.....	106

ÍNDICE DE FIGURAS

Figura 1. Morosidad de cuotas sociales periodo 2006 - 2017 CIP-CDL..... 12

Figura 2. Arquitectura de Inteligencia de Negocios 21

Figura 3. Modelo Estrella 24

Figura 4. Modelo Copo de Nieve 25

Figura 5. Los dos pasos del aprendizaje supervisado 26

Figura 6. Diferentes técnicas de aprendizaje automático y sus datos requeridos. 27

Figura 7. Aprendizaje Supervisado. 28

Figura 8. Clasificación lineal en dos dimensiones 30

Figura 9. Distancia entre ambas categorías 31

Figura 10: Ejemplo de clasificación y función kernel radial..... 33

Figura 11. Superficies de predicción creadas por funciones kernel. 36

Figura 12 (a). Función Logística. 37

Figura 13. Curva ROC..... 43

Figura 14. Predicción bursátil usando R..... 44

Figura 15. Flujo de la realización de la investigación 54

Figura 16. Base de datos del Colegio de Ingenieros del Perú – CD Lambayeque. 55

Figura 17. Hojas de cálculo con datos de colegiados de las diversas especialidades CD Lambayeque desde el año 1994 hasta 2017..... 56

Figura 18. Modelo Entidad – Relación Base de datos CIP - CD Lambayeque..... 61

Figura 19. Consulta para el poblamiento de la Tabla Capítulo 68

Figura 20. Consulta para el poblamiento de la Tabla Especialidad..... 68

Figura 21. Consulta para el poblamiento de la Tabla Cobro 68

Figura 22. Consulta para el poblamiento de la Tabla Cajero_AtencionCliente 69

Figura 23. Consulta para el poblamiento de la Tabla Colegiado..... 69

Figura 24. Consulta para el poblamiento de la Tabla Familiar 70

Figura 25. Consulta para el poblamiento de la Tabla Pagos 70

Figura 26. Modelo Entidad Relación con dimensiones seleccionadas 73

Figura 27. Dimensión Tiempo..... 73

Figura 28. Dimensión Colegiado..... 74

Figura 29. Dimensión Cuota Mensual..... 74

Figura 30. Modelo con dimensiones implementadas 74

Figura 31. Información en formato .CSV para implementación de técnicas de predicción 75

Figura 32. Plano de dos dimensiones con variables predictivas x_1 y x_2 76

Figura 33. Posibles hiperplanos que separan las clases No Moroso y Moroso. 76

Figura 34. Vectores de soporte S_1 , S_2 y S_3 77

Figura 35. Hiperplano óptimo de separación de clases Moroso y No Moroso 80

Figura 36. Predicción de clase A y clase B. 80

Figura 37. División de datos 80 % para entrenamiento y 20 % para prueba..... 81



Figura 38. Parámetros del modelo de máquina de soporte vectorial.....	81
Figura 39. Resultados obtenidos con los datos de prueba.....	82
Figura 40. Matriz de confusión y resultado de los indicadores de eficiencia.....	82
Figura 41. Gráfico de la Curva ROC y el valor del Área bajo la curva (AUC).....	83
Figura 42. Variables predictivas x_1 y x_2	84
Figura 43. División de datos 80 % para entrenamiento y 20 % para prueba.....	86
Figura 44. Parámetros del modelo de regresión logística.....	87
Figura 45. Parámetros del modelo de regresión logística.....	87
Figura 46. Resultados obtenidos con los datos de prueba.....	88
Figura 47. Matriz de confusión y resultado de los indicadores de eficiencia.....	88
Figura 48. Gráfico de la Curva ROC y el valor del Área bajo la curva (AUC).....	89
Figura 49. Comparación de Indicadores de Eficiencia entre la técnica Máquina de Soporte Vectorial y Regresión Logística.....	97
Figura 50. Comparación del indicador área bajo la Curva (AUC) de las técnicas de Máquina de Soporte Vectorial y Regresión Logística.....	98
Figura 51. Comparación del tiempo requerido por la técnica de máquina de soporte vectorial y regresión logística en la predicción de morosidad de cuotas sociales.....	99

ÍNDICE DE TABLAS

Tabla 1 Matriz de Confusión.....	41
Tabla 2 Puntuación AUC.....	43
Tabla 3 Operacionalización de variables	49
Tabla 4 Descripción de las variables utilizadas en la investigación.....	57
Tabla 5 Descripción de las tablas que conforman la base de datos SQL de acuerdo a las variables identificadas para la predicción de morosidad de cuotas sociales del Colegio de Ingenieros CD Lambayeque.....	58
Tabla 6 Descripción de Ítems de acuerdo con la Matriz RACI.....	58
Tabla 7 Tablas seleccionadas cuya información tiene relación con las variables de predicción	59
Tabla 8 Descripción detallada de cada campo de las tablas seleccionadas de la base de datos Colegio de Ingenieros CD Lambayeque.....	60
Tabla 9 Descripción de tablas del modelo Entidad – Relación.	62
Tabla 10 Descripción de cada requerimiento que corresponde a los indicadores de predicción	63
Tabla 11 Condición y valoración en la escala de Likert.....	66
Tabla 12 Comparación del Modelo Estrella y Modelo Copo de Nieve	67
Tabla 13 Variables identificadas para la predicción	71
Tabla 14 Elección de dimensiones de acuerdo a las variables.....	72
Tabla 15 Evaluación de la técnica máquina de soporte vectorial	90
Tabla 16 Indicadores de la técnica máquina de soporte vectorial	92
Tabla 17 Evaluación de la técnica regresión logística	93
Tabla 18 Indicadores de la técnica regresión logística	95
Tabla 19 Tiempo de ejecución promedio Máquina de Soporte Vectorial Vs. Regresión Logística	96

CAPITULO I: INTRODUCCIÓN

1.1. Realidad problemática.

En la actualidad algunas organizaciones financieras y entidades de todo tamaño están inmersos en dificultades económicas a causa del creciente aumento de la morosidad. Una creciente cartera morosa, es un serio problema que no contribuye favorablemente al progreso de una organización en el mediano y largo plazo. (Aguilar, 2006).

Por ejemplo, en España, según datos publicados por el Banco de España la morosidad a causa del impago de créditos otorgados por entidades bancarias, cajas y cooperativas a personas particulares y empresas, ascienden a 97 691 millones que representan un 7,79% de un total de 1,25 billones de euros concedidos en créditos el 2017, de ello se desglosa que, un 62 % correspondió al sector empresarial y el 38% restante a las familias. (Europa Press, 2018).

De igual manera en Perú, según la Asociación de Bancos del Perú (ASBANC) la morosidad dentro de las entidades bancarias alcanzó el 2017 una tasa de 3.04% uno de los picos más elevados después de 12 años. Sin embargo, dicha tasa de acuerdo al gremio bancario se eleva a 5.77 % si se adicionan aquellos créditos que fueron reestructurados y refinanciados. Se refinancia cuando los préstamos solicitados caen en morosidad por exceder los 120 días del plazo de pago y es asumido por las entidades financieras con sus recursos, el cual no favorecen sus balances económicos. Los segmentos económicos que presentaron elevadas tasas de incumplimiento de pago fue la micro empresa y la mediana empresa con una tasa de 9% y 7,1% respectivamente. (El Comercio, 2018).

Por otro lado, el Colegio de Ingenieros del Perú Consejo Departamental de Lambayeque (CIP-CDL) organización sin fines de lucro conformado por 7 capítulos de ingeniería cuyos ingresos se generan por el pago de las cuotas sociales de los colegiados no es ajeno a la realidad antes descrita, ya que en la actualidad vienen afrontando elevadas tasas de morosidad a causa del incumplimiento de pagos pues así lo confirma el informe de situación financiera 2017 en la cual la morosidad asciende a la suma de 6,851,249.93 de la cual el 13.26 % corresponde a la morosidad por incumplimiento de pago de cuotas sociales equivalente a la suma de 908,654.00 soles, cuya incremento anual se aprecia en la figura 1. El incremento del incumplimiento de pago genera falta de liquidez lo cual repercute desfavorablemente en



la mejora de los servicios brindados a los colegiados, mejora de infraestructura, cumplimiento de pago a trabajadores y toma de decisiones. (CIP-CD Lambayeque, 2017).

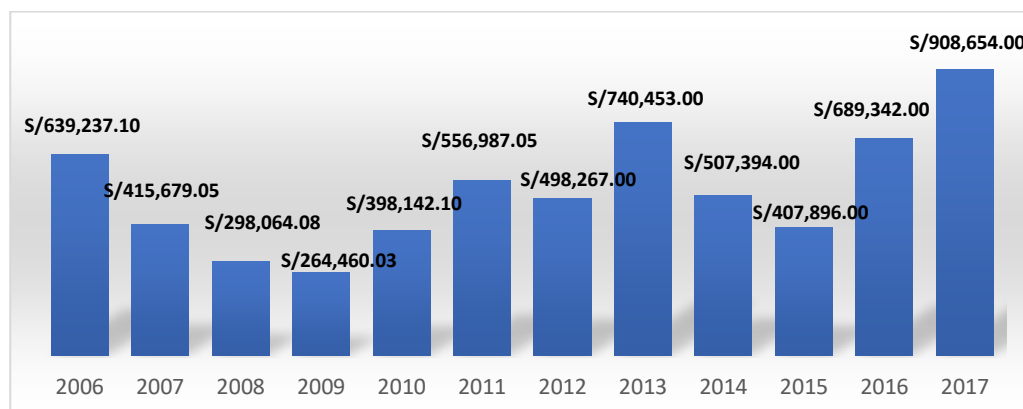


Figura 1. Morosidad de cuotas sociales periodo 2006 - 2017 CIP-CDL.

Fuente: Estadísticas CIP – Consejo departamental Lambayeque 2017.

Siendo la morosidad un problema que se incrementa conforme transcurren los años y necesita una solución, es que se realizó esta investigación utilizando técnicas que ya se vienen implementando en diferentes ámbitos como el sector financiero donde los investigadores (Zeng et al., 2017) desarrollaron su investigación en instituciones financieras que enfrentan riesgos al otorgar créditos, por ello construyeron modelos de regresión logística para predecir los incumplimientos de pago, obtuvieron el 92,55% de precisión promedio en etapa de entrenamiento y 92,53% en etapa de prueba. Por su parte (Lawi et al., 2017) mediante el método de regresión logística en conjunto con Boosting buscaron mejorar la predicción de riesgos de crédito para ello normalizaron los datos con valores 0 y 1, obtuvieron una precisión del 81 % para un conjunto de datos alemanes y un 88,4% para datos australianos , así como también (Gahlaut et al., 2017) expresan que los préstamos de crédito de los bancos muchas veces se otorgan sin saber que se devolverán o no, por ello implementaron las técnicas de árboles de decisión, bosques aleatorios, máquina de soporte vectorial, regresión lineal y redes neuronales donde identificaron que las variables que mejoran la predicción son la edad, duración y cantidad de crédito. Así mismo (Birla et al., 2016) mediante los algoritmos supervisados de aprendizaje automático buscaron predecir si un cliente que realiza una solicitud de préstamo cumplirá o no con su pago, por ello implementaron las técnicas de



árboles de decisión y regresión logística, obtuvieron el 87,73 % y 84,23 % de precisión. Como se ha descrito se vienen desarrollando investigaciones exhaustivas en diversos sectores siendo el sector financiero el más cercano al ámbito de estudio, pero los resultados obtenidos en las investigaciones varían de acuerdo al ámbito de estudio, técnica implementada y los datos usados, por ello la presente investigación se desarrolló usando las técnicas máquina de soporte vectorial y regresión logística en la predicción de morosidad en cuotas sociales del colegio de Ingenieros del Perú consejo departamental Lambayeque.

1.2. Antecedentes de Estudio.

1.2.1. Antecedentes de la Investigación

Trujillo (2017) desarrolló el trabajo de investigación “Aplicación de Metodologías Machine Learning en la Gestión de Riesgo de Crédito” en la Universidad Politécnica de Madrid, en España enfrentó la complejidad en la gestión de riesgos financieros producto de errores en los modelos de cuantificación de riesgos. Propuso la implementación de modelos de aprendizaje automático que permitan predecir si un potencial cliente incurrirá en mora luego de ser concedido el crédito. Para la implementación de los modelos usó un conjunto de datos cuya información pertenecía a clientes de una entidad financiera recopilada entre los años 2001 al 2007 conformado por 104 961 registros, que procedió a procesar para identificar los datos concisos para la construcción de los modelos de aprendizaje automático. Para crear y probar cada uno de los modelos de predicción usó el método de entrenamiento y prueba por ello dividió el conjunto de datos en 70% para entrenamiento y 30% para prueba. Los resultados que obtuvo fue 65,10 % de precisión con regresión logística, 68,00% de precisión con AdaBoost, 52,80% de precisión con Random Forest, 64,20 % de precisión con Máquina de soporte vectorial (SVM) y 67,50% de precisión con Naive Bayes. Llegó a la conclusión que las técnicas de aprendizaje automático proporcionan soluciones efectivas para la estimación de riesgos de crédito.

Tesén (2017) desarrolló el trabajo de investigación “Eficacia de los modelos de aprendizaje máquina para evaluar el riesgo crediticio de personas naturales en una institución financiera de Chiclayo” en la universidad Nacional del Santa, en Chimbote enfrentó la necesidad de establecer mecanismos para medir los riesgos de crédito. Propuso determinar las variables



significativas que expliquen el fenómeno y contribuyan a generar un modelo de medición de riesgo de crédito, a través de métodos estadísticos teóricos – prácticos. Procedió a utilizar la base de datos otorgada por una entidad financiera donde aplicó los métodos de redes neuronales, máquinas de soporte vectorial, árboles de clasificación y regresión logística cuyos resultados estimados fueron 30.58 %, 31.17 %, 32.21%, 29.63 % respectivamente. Llegó a la conclusión que los modelos de aprendizaje de máquina logran óptimos resultados de estimación de riesgo de crédito en comparación al modelo paramétrico de regresión logística.

Marin & Marulanda (2016) desarrollaron el trabajo de investigación “Modelo de aprobación de crédito en una entidad financiera” en la fundación Universitaria los Libertadores, en Colombia afrontaron el problema que aqueja a la mayoría de entidades financieras, que es encontrar métodos confiables por medio del cual evaluar de manera acertada el otorgamiento de créditos para que no caigan en un corto tiempo en impago. Proponen un modelo que de acuerdo a las variables predictivas identificadas pueda diferenciar entre un riesgo alto o bajo. Procedieron a adquirir un conjunto de datos cuya información tenía relación con las variables predictivas identificadas como el historial de préstamos, información demográfica, y comportamiento de pago de los clientes de una entidad financiera. El desarrollo del modelo lo hizo mediante el lenguaje R, para ello dividió el conjunto de datos para entrenamiento y prueba. Obtuvo 64% de exactitud, 63% de sensibilidad y 65% de especificidad para el modelo de regresión logística. Concluyeron que las variables significativas para crear el modelo de regresión logística están relacionadas a la edad, ingresos diarios, estado civil, número de hijos y número de créditos con la entidad, y las variables que resultaron no significativos fueron la variable hipoteca y forma de pago.

Carpio (2016) desarrolló el trabajo de investigación “Modelo de predicción de la morosidad en el otorgamiento de crédito financiero aplicando metodología CRISP-DM” en la Universidad Andina “Néstor Cáceres Velásquez”, en Juliaca enfrentó el crecimiento paulatino de la morosidad luego del otorgamiento de crédito. Propuso el desarrollo de un modelo de predicción de morosidad en el otorgamiento de crédito financiero. Para ello usó la metodología CRISP – DM, el cual busca el descubrimiento del conocimiento cuya estructura está formada por fases, iniciaron con la comprensión del negocio del microcrédito,



posteriormente identificaron de acuerdo a los modelos los datos de entrada y finalmente efectuaron las pruebas. Obtuvo los siguientes resultados 79% de precisión global del modelo de árboles de decisión, 82 % de precisión global del modelo bosques aleatorios (Random Forest), 80 % de precisión global del modelo de redes neuronales, 81% de precisión global para el modelo de Máquinas de soporte vectorial. Concluye que el modelo de bosques aleatorios predice mejor la morosidad de crédito con una precisión global de 82% evaluado haciendo uso de la matriz de confusión y curva ROC.

Bedor (2015) desarrolló el trabajo de investigación “Construcción de un modelo estadístico para la gestión de la cartera de cobro del Servicio de Rentas Internas Zonal 8” en el Instituto de Altos Estudios Nacionales, en Guayaquil enfrentó la cartera pendiente de cobro. Propuso definir los parámetros fundamentales para el desarrollo de un modelo estadístico que contribuya con la gestión óptima de la unidad de cobro del servicio de rentas internas Zonal 8. Para ello procedió a la recolección de datos, el cual procesó seleccionando únicamente los relacionados con las obligaciones de pagado y pendientes de pago. Obtuvo los siguientes resultados para el modelo desarrollado a través de la matriz de confusión 96,01% de sensibilidad, 60,97% de especificidad, 87,26% de clasificación correcta. Concluye que, dentro de la alta gama de modelos estadísticos para medir riesgo de impago, el método de regresión logística sobresale.

Fernández (2015) desarrolló el trabajo de investigación “Adecuación de los modelos de predicción de las crisis bancarias al caso español” en la Universidad de Sevilla, en Sevilla enfrentó la crisis bancaria española cuyas causas se agrupan en dos categorías: la macroeconómica, que están más vinculadas con la aparición de la crisis bancaria sistémica, y las microeconómicas, que se generan en fracasos bancarios aislados. Propuso comparar los modelos estadísticos y minería de datos en la predicción de la insolvencia empresarial bancaria. Para ello agrupó las técnicas en dos grupos: los estadísticos (técnicas paramétricas), entre los que destacan el análisis de regresión, el análisis factorial, el análisis discriminante, los modelos de regresión logística y los modelos de minería de datos (no paramétricos), destacando las redes neuronales, árboles de decisión, algoritmos genéticos y máquinas de vectores de soporte. Primeramente, seleccionó la muestra y definió la variable dependiente (crisis) de los años 2009 y 2010, seguidamente seleccionó las variables explicativas



predominando los indicadores CAMEL (Adecuación de capital, calidad de activo, gestión, rentabilidad, liquidez, disciplina de mercado); Obtuvo los siguientes resultados para el modelo logit 70.59 %, 89.04% de especificidad y sensibilidad cuyos errores fueron de Tipo I 6,45% y el error Tipo II del 12.10%. Y para el modelo no paramétrico árboles de clasificación 78.43% de especificidad y 94.52% de sensibilidad cuyos errores fueron Tipo I 3,23% y error Tipo II 8.87%. Concluye que los árboles de clasificación proporcionan mejores resultados con una precisión global de 87.90%.

Marcos (2015) desarrolló el trabajo de investigación “Propuesta de un modelo de Regresión Logística para analizar el riesgo crediticio en la empresa CARSA S.A.C.” en la Universidad Nacional Agraria la Molina, en Lima enfrentó el incumplimiento de pago de créditos. Propuso estudiar y diseñar modelos estadísticos que permitan hacer frente al incumplimiento de pago. Procedió al análisis de los datos para ello realizó la limpieza de base de datos y selección de variables, creó el modelo y efectuó pruebas. Identificó que la probabilidad de incumplimiento de pago es mayor en el 8% de los solicitantes. Concluye que el modelo de regresión logística aporta positivamente en la identificación del riesgo crediticio en las ventas al por menor en la empresa Carsa S.A.C

Medina & Ulfe (2015) desarrolló el trabajo de investigación “Modelo de Credit Scoring para predecir el otorgamiento de crédito personal en una cooperativa de ahorro y crédito” en la Universidad Nacional Pedro Ruiz Gallo, Lambayeque enfrentaron las dificultades que se presentan en la concesión de créditos personales, reflejados en el retraso de los pagos en las fechas establecidas que son causantes de una morosidad variable englobados en una ineficiente asignación de créditos que conllevan a la necesidad de ajustar los criterios de evaluación. Propusieron la construcción de un modelo scoring que mejore la toma de decisión al momento de otorgar un crédito personal, cuya base es clasificar a los clientes de acuerdo a la probabilidad de default. Procedieron a obtener el conjunto de datos con suficiente consistencia histórica de un total de 3000 clientes de la cooperativa de ahorro y crédito. Identificaron 30 variables que agruparon en 27 variables cuantitativas y 3 variables cualitativas, las que evaluaron para determinar las variables explicativas que permitirá la construcción de modelo. Para crear el modelo, el conjunto de datos lo dividieron en un 50% para la etapa de entrenamiento (construcción del modelo) y el otro 50% para la validación



(ajuste y comparación de los datos), las técnicas que implementaron fueron la Regresión Logística, árboles de clasificación y Redes Neuronales, una vez hallado los modelos procedieron a compararlos utilizando la matriz de confusión que permite determinar el indicador de eficiencia de cada modelo como la sensibilidad y especificidad. Obtuvieron los siguientes resultados, sensibilidad (16.00%, 16.00% y 10.40%) y especificidad (97.80%, 97.40% y 98.40%) para las técnicas de árboles de clasificación, regresión logística y redes neuronales respectivamente.

Reyes & León (2014) desarrollaron el trabajo de investigación “Capacidad predictiva de los modelos de máquina de vectores de soporte y modelo de regresión logística en el análisis de riesgo de crédito - persona” en la Universidad Nacional de Ingeniería, en Lima enfrentó el problema de los riesgos de crédito. Propuso comparar la capacidad predictiva de las técnicas estadísticas clásicas con las técnicas de Inteligencia artificial. Para ello siguió la metodología del descubrimiento del conocimiento en base de datos (Knowledge Discovering in Databases) partiendo de la definición del objetivo, selección de los datos, preparación de los datos, selección de variables y transformación, proceso de modelamiento, validación de modelo e implementación de modelo. Obtuvieron los siguientes resultados 73.80 % de sensibilidad y 70.30 % de especificidad como indicadores predictivos para el modelo logístico y 75.00 % de sensibilidad y 74.50 % de especificidad para el modelo de la máquina de soporte de vectores. Llegaron a la conclusión que la capacidad predictiva del modelo de máquina de vectores de soporte (SVM) son superiores a los indicadores del modelo logístico en el análisis de riesgo crediticio para una base de datos de banca personal.

Fernández (2013) desarrolló el trabajo de investigación “Estimación estadística de clientes morosos de Lima metropolitana y el callao que usan tarjetas de crédito para la banca de personas naturales en una entidad bancaria”, en la universidad Nacional de Ingeniería, en Lima donde enfrentó el problema de atraso en las obligaciones crediticias de los clientes de la entidad Bancaria Perú Bank. Propuso a partir de un modelo estadístico identificar las variables de mayor importancia para estimar a los clientes morosos que usan tarjetas de crédito para la banca. Procedió a obtener los datos de la Data Warehouse de la entidad Bancaria Perú Bank y el reporte crediticio consolidado de la Superintendencia de Banca, Seguros y AFP (SBS) donde definió las variables y preparó los datos considerando el 80 %



de la muestra para la construcción del modelo estadístico de regresión logística y el 20 % para la validación del modelo estadístico obteniéndose una precisión de 50.7 % y un porcentaje global de 69.6%. Llegó a la conclusión que el modelo estadístico de regresión logística da mayor importancia a las variables: uso de línea de tarjeta, zona geográfica, nivel socioeconómico, edad, ingresos y antigüedad de su tarjeta de crédito.

1.2.2. Estado del Arte.

Zeng et al. (2017) en su artículo “A new competing risks model for predicting prepayment and default using data mining” en la escuela de negocios de China, enfrentaron el riesgo crediticio en las instituciones financieras que los llevan al peligro de colapso por las pérdidas generadas. Proponen construir un modelo de riesgo competitivo para predecir el incumplimiento mediante el uso de modelos de regresión logística. Procedieron a proponer los conceptos de microfinanzas, adquirieron los datos, limpiaron los datos, seleccionaron variables que tienen un impacto significativo en el pago por adelantado y por defecto, dividieron los datos en conjunto de 90 % para entrenamiento y 10 % para pruebas, construyeron el modelo de regresión logística por adelantado y por defecto, hicieron competir los modelos y obtuvieron que las predicciones son altas con tasa de precisión promedio de 92,55% antes del percentil 16 en entrenamiento y 92,53% de tasa de precisión en prueba. Concluyeron que el rendimiento del modelo de riesgos competitivo es mucho mejor que el modelo de predicción único.

Lawi et al. (2017) en su artículo “Ensemble GradientBoost for Increasing Classification Accuracy of Credit Scoring” en la Universidad de Ciencias de la Computación, Indonesia enfrentaron los riesgos de crédito. Proponen la identificación de Credit Scoring mediante el método de regresión logística de conjunto con Boosting. Procedieron a obtener el conjunto de datos del repositorio de UCI machine learning sobre préstamos a alemanes y australianos, normalizaron los datos para la eliminación de la redundancia de los datos con valores 0 y 1, dividieron el conjunto de datos en 70% para el entrenamiento y 30% para la prueba de datos, obtuvieron una precisión de clasificación con GradientBoost del 81% para datos alemanes y 88,4% para datos australianos. Concluyeron que el algoritmo de GradientBoost mejora con



éxito el rendimiento la clasificación con regresión logística y alcanza un mayor grado de precisión en ambos conjuntos de datos.

Gahlaut et al. (2017) en su artículo “Prediction analysis of risky credit using Data minig classification models” en la Universidad de Delhi, en India enfrentaron el otorgamiento de préstamos de crédito de los bancos a sus clientes sin saber, que van a ser capaces de devolver o no. Proponen hacer uso de las técnicas de minería de datos que permita predecir y clasificar la puntuación de crédito del cliente (bueno/malo) para superar los riesgos futuros de dar préstamos a clientes que no puedan pagar. Procedieron a obtener el conjunto de datos del repositorio UCI para máquinas de aprendizaje presentado por Hans Hofmann, que contienen diversas variables tales como crédito, duración, frecuencia, edad, ocupación, etc. El conjunto de datos se repartió en 70% para entrenamiento y el 30 % para validación, seleccionaron los factores importantes de las variables (duración, cantidad, edad) que afectan al resultado final, procedieron a hacer uso de los modelos de árboles de decisión, bosques aleatorios, adaptive boosting, máquina de soporte vectorial, regresión lineal y redes neuronales, obtuvieron los siguientes resultados del área bajo la curva 0.7968, 0.7686, 0.7672 , 0.7569, 0.6885, 0.5000 correspondientes a los modelos bosques aleatorios, regresión lineal, máquina de soporte vectorial, adaptive boosting, árboles de decisión y redes neuronales. Concluyeron que entre todos factores las variables edad, duración y cantidad son los más importantes que afectan en los resultados de predicción.

Awoyemi et al. (2017) en su artículo “Credit card fraud detection using Machine Learning Techniques” en la Universidad Federal Tecnológica de Akure, en Nigeria afrontaron el fraude por el uso las tarjetas de crédito en las transacciones en línea. Proponen medir el rendimiento de detección de fraude en las transacciones de tarjetas de crédito haciendo uso de tres técnicas de aprendizaje automático: Naive Bayes, K vecino más cercanos y regresión logística. Procedieron a obtener el conjunto de datos del grupo de aprendizaje automático de la Universidad libre de Bruselas (ULB), cuyos datos contienen las transacciones de tarjetas de crédito efectuadas por titulares europeos en septiembre del 2013 y consta de 284 807 transacciones, efectúan un análisis del conjunto de datos donde seleccionan 28 componentes principales. Evaluaron el rendimiento midiendo la exactitud, sensibilidad, especificidad, precisión, coeficiente de correlación de Matthews y tasa de clasificación equilibrada para ello



dividieron el conjunto de datos 70% para formación y 30% para pruebas, además de las distribuciones 10:90, 34:66. Obtuvieron los siguientes resultados en datos no muestreados regresión logística muestras los mejores indicadores 98.24 % de Exactitud, 97,67% de sensibilidad, 98,24 % de especificidad y en datos distribuidos 10:90 la técnica K-vecinos más cercanos muestra mejores indicadores 97.15 % de Exactitud, 82.85% de sensibilidad, 100% de especificidad similares indicadores obtuvieron en la distribución 34:66. Concluyeron que la técnica K-vecino más cercanos muestra el rendimiento significativo para todas la métricas evaluadas excepto en la distribución 10:90.

Birla et al. (2016) en su artículo “Machine Learning on Imbalanced Data in Credit Risk” en el Instituto de Tecnología Vellore, India enfrentaron el riesgo de crédito, que se define como la probabilidad de impago del préstamo o crédito adquirido de una institución bancaria o financiera. Proponen desarrollar un modelo predictivo usando varios algoritmos supervisados de aprendizaje automático que puedan predecir si un nuevo cliente que realiza una solicitud de préstamo incumplirá o no con su préstamo. Procedieron a aplicar las técnicas de árboles de decisión, regresión logística al conjunto de datos desequilibrados; obtuvieron 87,73 % de precisión de la predicción probabilidades previas, 84,23% de precisión de la predicción para el método Logit. Concluyeron que el algoritmo CART y Regresión logística funcionan muy bien con datos desequilibrados, además la curva ROC es un excelente indicador para determinar el rendimiento del clasificador y es compatible con la mayoría de algoritmos.

Maher et al. (2015) en su artículo “A systematic credit scoring model base don heterogeneous classifier ensembles” en la Universidad de Brunel, en Londres enfrentaron los riesgos por la concesión de créditos, actividad clave que genera ganancias a los bancos e instituciones financieras. Proponen desarrollar un sistema de puntuación de créditos basado en conjuntos homogéneo y heterogéneos basado en tres clasificadores: Regresión Logística (LR), red Neuronal artificial (ANN) y máquina de vectores de soporte (SVM). Procedieron a obtener el conjunto de datos del repositorio para aprendizaje de máquina (UCI) de créditos otorgados a alemanes y australianos conformado por 1690 registros y 34 atributos, efectuaron la normalización de los datos con el fin de evitar el sesgo y construir el modelo, para lograr esto los atributos del conjunto de datos lo normalizaron a valores entre los intervalos 0 y 1. Evaluaron el rendimiento midiendo la precisión media, error Tipo I y error Tipo II para ello



dividieron el conjunto de datos 80% para entrenamiento y 20% para la prueba , obtuvieron los siguientes resultados regresión logística logró la mejor precisión 75,5% y SVM el menor error Tipo I para el conjunto de datos Alemán , en el conjunto de datos Australiano logró redes neuronales la mayor precisión de un 86,2% y el menor error Tipo I. Concluyeron que en la clasificación de conjuntos homogéneos la regresión logística obtiene una precisión de 76 % y 86.59% con los datos alemanes y australianos respectivamente .

1.3. Teorías relacionadas al tema.

1.3.1. Inteligencia de Negocios

La Inteligencia de Negocios consiste en la capacidad de procesar los datos en información y convertirlos en conocimiento, buscando optimizar el proceso de decisión en el negocio. Por ello la inteligencia de negocios juega un papel fundamental en la actualidad proporcionando un grupo de metodologías y aplicaciones con los que se logran recopilar, depurar y transformar los datos de sistemas transaccionales en información estructurada, permitiendo la explotación de la información (análisis On-Line Transaction Processing OLTP / On-Line Analytical Processing OLAP, informes, alertas, etc.) o su análisis y transformación en conocimiento, con lo cual da soporte a la toma de decisiones sobre el negocio (Maura et al., 2018). De forma sucinta el proceso de inteligencia de negocio se presenta a continuación.

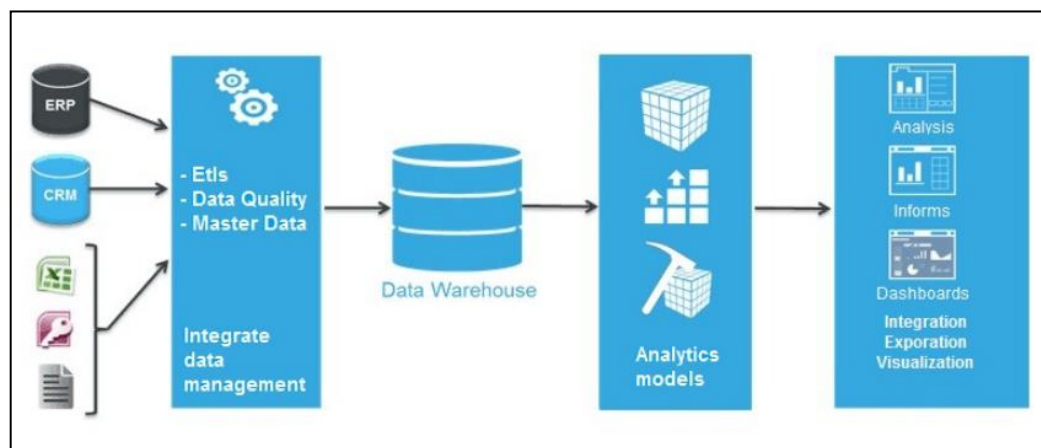


Figura 2. Arquitectura de Inteligencia de Negocios

Fuente: Maura et al. (2018)



1.3.1.1. Data Warehouse, Data Mart

Data Warehouse

Un Almacén de Datos (Data Warehouse) es una base de datos corporativa que integra y refina la información de diversas fuentes permitiendo su análisis desde una variedad de perspectivas cuyo tiempo de respuesta es óptimo. Por lo tanto, agrupa los datos de toda la empresa con el objetivo de facilitar su análisis de forma ágil. Para consolidar un almacén de datos corporativo es fundamental la integración de los datos que proceden de distintas unidades de la organización los cuales frecuentemente dispondrán de distintas estructuras. Además, un almacén de datos corporativo nos permite la separación de los datos utilizados en las operaciones diarias de aquellos datos para los propósitos de divulgación que facilitan la toma de decisiones, análisis y operaciones de control. (Chamorro, 2016).

Data Mart

Se trata de una versión condensada y más centrada de un Data Warehouse que se caracteriza por el almacenamiento de los datos de un área en particular del negocio que permiten obtener información específica como contabilidad, ventas, logística, etc. Entre los data mart de amplio uso se encuentran la Online Transactional Processing (OLTP) y Online Analytical Processing (OLAP). Los sistemas OLTP son base de datos orientadas al procesamiento de transacciones que pueden conllevar operaciones de inserción, modificación y borrado de datos. Por su parte los sistemas OLAP son base de datos orientados al procesamiento analítico. (Chamorro, 2016).

1.3.1.2. Proceso de Extracción, Transformación y Carga (ETL)

Extraer, Transformar y Cargar del inglés Extract-Transform-Load (ETL) corresponden con cada una de las etapas que comprenden este proceso cuyo objeto principal es dinamizar el traslado de los datos y su posterior transformación de los mismos, logrando la integración de los distintos sistemas y fuentes que forman la organización moderna. (Chamorro, 2016).

Proceso de Extracción

Esta etapa comprende la obtención de datos, partiendo de identificar y determinar de entre todas las fuentes de datos que conforman la organización aquellos datos que son necesarios. Una vez identificados se debe tener en cuenta si formaron parte de la carga inicial o de las actualizaciones, debido a que en toda organización se podrán determinar dos tipos de información: la información histórica y la información actual. Además, la etapa de extracción debe procurar en la menor medida posible afectar el rendimiento de los sistemas de origen. (Chamorro, 2016).

Proceso de Transformación

La etapa de transformación parte de los datos obtenidos de los sistemas operacionales, así como de los datos externos transformándolos en la estructura idónea que tiene que ir en el almacén de datos. En la etapa de transformación también se lleva a cabo la depuración e integración, en el caso de la depuración lo que se realiza es detectar y corregir aquellos valores que sean inconsistentes. En cuanto a la integración se lleva a cabo la carga de los datos hacia el almacén de datos, siendo esta inicial o una actualización. (Chamorro, 2016).

Proceso de Carga

Es la última etapa del proceso ETL, los datos extraídos y transformados son transportados de las diferentes plataformas y cargados en la base de datos correspondiente. (Chamorro, 2016).

1.3.1.3. Análisis Dimensional

Tabla de Hechos

Tabla de Hechos del inglés Fact Table viene a ser la tabla principal de un esquema dimensional (en estrella o copo de nieve) y está constituido por valores numéricos de acuerdo a las entidades del negocio. La tabla de hechos está formada por medidas y foreign keys (claves foráneas). Además, representa la relación de muchos a muchos entre dimensiones que forman el modelo dimensional. (Rosales, 2009).



Dimensiones

Las dimensiones definen objetos del negocio que permiten realizar análisis de la tendencia y el comportamiento del mismo. Por ello se valen de acuerdos, políticas y compromisos dentro de la compañía, de tal manera que permite que la organización interprete, clasifique o segmente su información el cual facilita la observación y análisis de los datos. (Rosales, 2009).

Medidas o métricas

Expresan las propiedades cualitativas o cuantitativas de los objetos a observar dentro de las organizaciones. Las medidas cuantitativas están representadas por valores o cifras porcentuales, por ejemplo, las compras en soles, cantidad de unidades compradas por producto, número de unidades en stock, ventas en soles, porcentaje de unidades vendidas. (Rosales, 2009).

Modelo Estrella

El modelo estrella está definido por un elemento central denominado tabla de hechos (Fact Table) y está relacionada a un determinado número de tablas de dimensiones. La tabla de hechos está estructurada por valores pre calculados que se determinan al totalizar valores operativos de nivel atómico de acuerdo a las distintas dimensiones, por ejemplo, colegiado, capítulo, periodos de pago, etc. (Rosales, 2009). El modelo estrella se aprecia en la Figura 3.

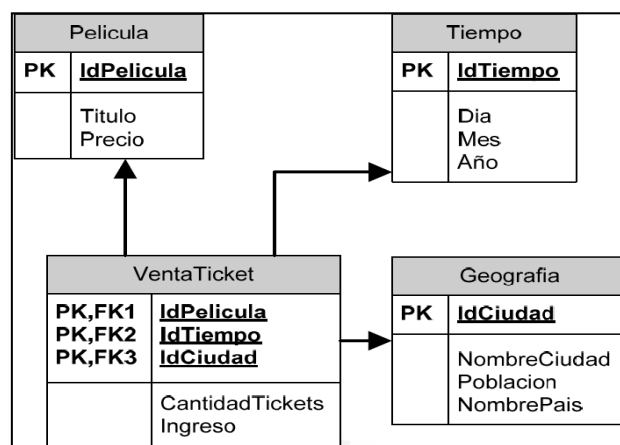


Figura 3. Modelo Estrella
Fuente: Rosales (2009)



Modelo Copo de nieve

El Modelo Copo de nieve se caracteriza porque cada tabla de dimensiones está debidamente normalizada, es decir, en su estructura pueden incluir claves que se relacionan con otras tablas de dimensión. La normalización favorece la reducción de la redundancia y tamaño en las tablas de dimensión e incrementa la versatilidad en la definición de dimensiones. A excepción de que el aumento de tablas requiere más operaciones de unión para dar respuesta a las consultas, lo que no favorece el rendimiento y el mantenimiento de las tablas adicionales (Rosales, 2009). El modelo copo de nieve se aprecia en la figura 4.

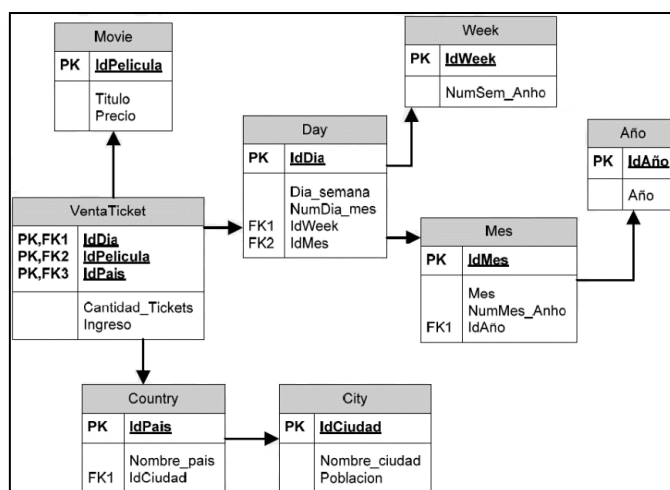


Figura 4. Modelo Copo de Nieve
Fuente: Rosales (2009)

1.3.2. Aprendizaje Automático

Aprendizaje Automático del inglés Machine Learning (ML) es una rama de la inteligencia artificial que tiene como objetivo permitir que las máquinas realicen sus trabajos hábilmente mediante el uso de software inteligente. Los métodos estadísticos de aprendizaje constituyen la columna vertebral del software inteligente que se utiliza para desarrollar inteligencia artificial. Debido a que los algoritmos de aprendizaje automático requieren datos para aprender, la disciplina debe tener conexión con la disciplina de la base de datos. (Mohammed et al., 2017).



El aprendizaje automático se define como un proceso automatizado que extrae patrones de los datos. Para construir los modelos utilizados en las aplicaciones de análisis predictivo de datos, utilizamos el aprendizaje automático supervisado. Las técnicas supervisadas de aprendizaje automático aprenden automáticamente el cual genera un modelo de la relación entre un conjunto de características descriptivas y una función objetivo basada en un conjunto de ejemplos históricos o instancias. Luego podemos usar este modelo para hacer predicciones para nuevas instancias (Kelleher et al., 2015). Estos dos pasos separados se muestran en la figura 5.

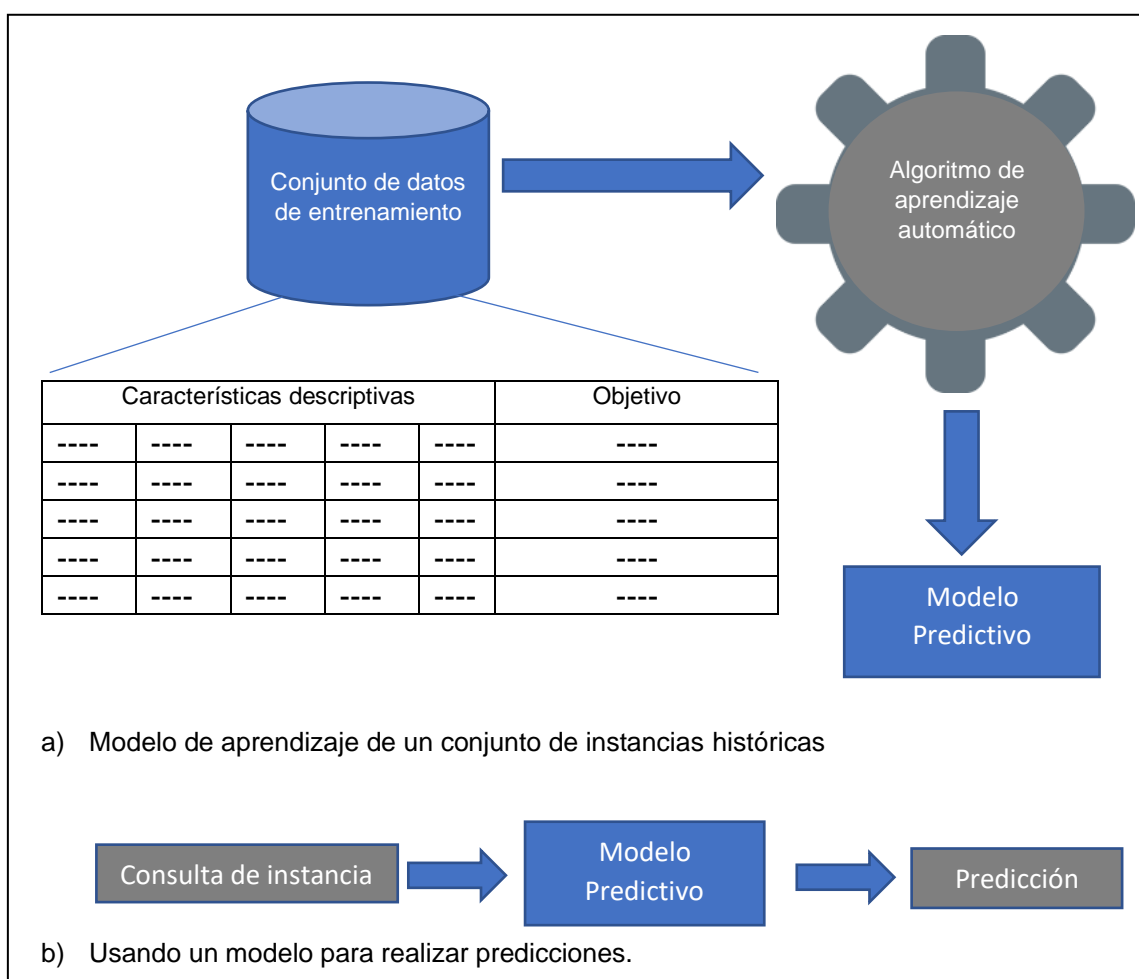


Figura 5. Los dos pasos del aprendizaje supervisado
Fuente: Kelleher et al (2015)



1.3.3. Técnicas de Aprendizaje Automático

En la figura 6 se muestran la agrupación de técnicas de aprendizaje automático.

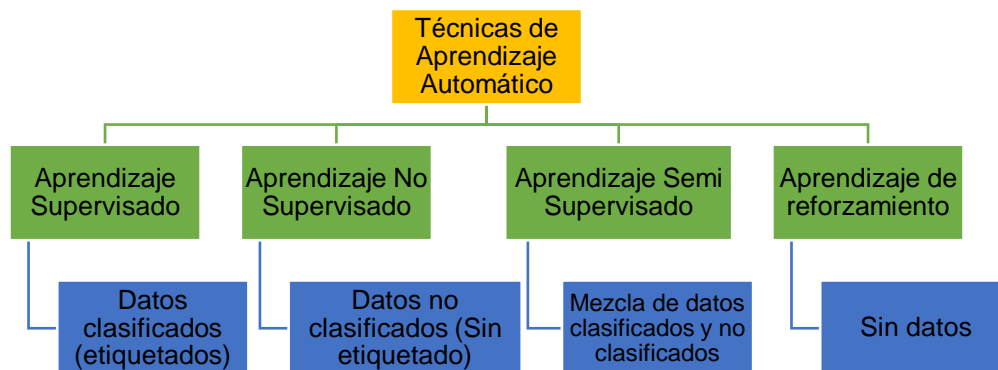


Figura 6. Diferentes técnicas de aprendizaje automático y sus datos requeridos.

Fuente: Mohammed et al (2017)

1.3.3.1. Aprendizaje Supervisado

Aprendizaje Supervisado del inglés Supervised Learning (SL) su objetivo es inferir una función o mapeo de los datos de entrenamiento que están etiquetados. Los datos de entrenamiento están formados por dos vectores: el primero representa los datos de ingreso y el segundo los resultados esperados. La función obtenida puede ser un valor numérico que se adecua a los problemas de regresión o una determinada etiqueta propio de problemas de clasificación. En la figura 7 se aprecia de forma gráfica la meta del aprendizaje supervisado el cual es obtener una función que logre predecir el valor correspondiente a cualquier nuevo objeto de entrada válido luego de ser entrenado con un conjunto de ejemplos. Para lograr predecir de manera óptima debe generalizar a partir de los ejemplos presentados al nuevo conjunto de datos no vistos. (Arrieta, 2017).



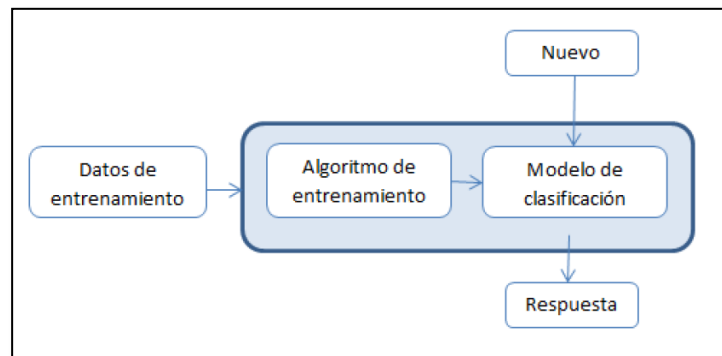


Figura 7. Aprendizaje Supervisado.

Fuente: Arrieta (2017)

El aprendizaje supervisado de acuerdo a la variable objeto, es decir lo que se está tratando de predecir, puede ser de tipo regresión o clasificación. El primer tipo es cuando la variable objeto es continua (valor numérico), y el segundo cuando la predicción esperada está definido por valores discretos (Si o No). Las técnicas más usadas son las redes neuronales, las máquinas de soporte vectorial, regresión logística, el algoritmo de K-vecinos más cercanos y árboles de decisión. (Arrieta, 2017).

1.3.3.2. Aprendizaje No Supervisado

Aprendizaje no supervisado es el tipo de aprendizaje que carece de supervisores o datos de capacitación. En otras palabras, todo lo que tenemos son datos sin etiqueta. La idea es encontrar una estructura oculta en esta información. Puede haber una serie de razones para que los datos no tengan una etiqueta. Puede deberse a la falta de fondos para pagar el etiquetado manual o la naturaleza inherente de los datos en sí. Con numerosos dispositivos de recopilación de datos, ahora los datos se recopilan a una velocidad sin precedentes. (Mohammed et al., 2017).

1.3.3.3. Aprendizaje Semi Supervisado

Según Mohammed et al. (2017) en este tipo de aprendizaje, los datos proporcionados son una mezcla de datos clasificados y no clasificados. Esta combinación de datos etiquetados y no etiquetados se usa para generar un modelo apropiado para la clasificación de datos. En la mayoría de las situaciones, los datos etiquetados son escasos y los datos no etiquetados están



en abundancia. El objetivo de la clasificación semi-supervisada es aprender un modelo que prediga las clases de datos de prueba futuros mejor que el del modelo generado al usar solo los datos etiquetados.

1.3.3.4. Aprendizaje de Reforzamiento

Según Mohammed et al. (2017) el aprendizaje de refuerzo tiene como objetivo de acuerdo a las observaciones realizadas por la interacción con el medio tomar decisiones que optimicen la recompensa o disminuyan el riesgo. Para producir programas inteligentes (también llamados agentes), el aprendizaje de refuerzo pasa por los siguientes pasos:

1. El agente observa el estado de entrada.
2. La función de toma de decisiones se usa para hacer que el agente realice una acción.
3. Después de que se realiza la acción, el agente recibe recompensa o refuerzo del entorno.
4. La información del par de acción estatal sobre la recompensa se almacena.

1.3.4. Máquinas de Soporte Vectorial

Según Castro & Pinedo (2018) las máquinas de soporte vectorial (en inglés, Support Vector Machine o SVM) es una técnica de aprendizaje automático que facilita la separación y división efectiva de un conjunto de datos haciendo uso de vectores e hiperplanos. Un conjunto de datos puede ser separado por una variedad de hiperplanos, la elección del hiperplano adecuado se verá influenciado por la distancia máxima que separa las distintas clases denominado margen, por ello si se determina un hiperplano cuyo margen entre clases es máximo la probabilidad de elección es mayor.



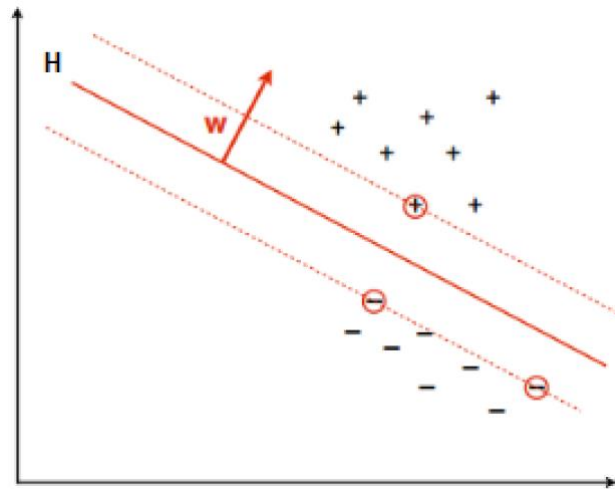


Figura 8. Clasificación lineal en dos dimensiones
Fuente: Castro & Pinedo (2018)

$$\{x \in S \mid w \cdot x + b = 0\}, w \in S, b \in R \tag{1}$$

En la figura 8 se puede apreciar una clasificación lineal de dos clases de datos, representados por un conjunto de puntos positivos ubicados en la parte superior de la diagonal H (hiperplano) y un conjunto de puntos negativos ubicados en la parte inferior de la diagonal H. Considerando que el conjunto de datos conformado por ambas clases se encuentra en el espacio S, donde “x” pertenece al espacio en estudio, los hiperplanos pueden ser descritos como el producto del vector perpendicular w y el vector que parte de las coordenadas (0,0) hasta un punto x sumado una constante b como se aprecia en la ecuación 1. (Castro & Pinedo, 2018).

Las dos regiones de clasificación que agrupan a unos como positivos y otros como negativos están definidos por el hiperplano diagonal H. Esta separación lineal divide ambas categorías, una región de positivos que se definirá como el resultado +1 y el de los negativos -1 de acuerdo a la ecuación 2. (Castro & Pinedo, 2018).

$$Clase(X_k) = \begin{cases} +1 & \text{Si } x_k + b > 0 \\ -1 & \text{Si } x_k + b < 0 \end{cases} \tag{2}$$

Haciendo uso de la definición fundamental de vectores se puede determinar la distancia que separa a los puntos fronterizos (vectores de soporte) de cada categoría, que viene hacer la diferencia de ambos vectores como se puede apreciar en la figura 9.



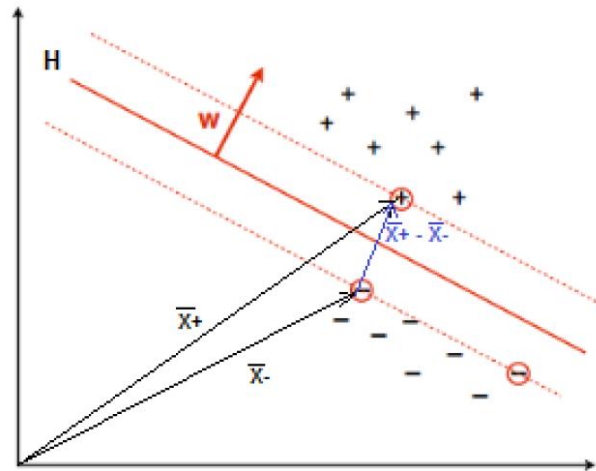


Figura 9. Distancia entre ambas categorías
Fuente: Castro & Pinedo (2018)

Lo que se busca es ampliar la distancia entre el conjunto de puntos positivos y negativos hacia el hiperplano (línea diagonal H). De acuerdo a la figura 9, la distancia a maximizar se logra al restar los vectores que se encuentran en el límite de ambos conjuntos de datos y a la misma vez más cerca al hiperplano. (Castro & Pinedo, 2018).

Se debe recordar que los vectores no representan una distancia numérica, por ello convendrá pasar a un escalar que permita determinar la longitud numérica necesaria que se logrará utilizando el vector unitario del vector w perpendicular al hiperplano y se multiplicará por la resta de vectores antes indicada, quedando la operación representada en la ecuación 3. (Castro & Pinedo, 2018).

$$D = (x_+ - x_-) \cdot \frac{w}{\|w\|} \quad (3)$$

Teniendo en cuenta las propiedades aritméticas en la ecuación 3, ambos sumandos se multiplican por el vector unitario del vector w obteniendo la ecuación 4. (Castro & Pinedo, 2018).

$$D = x_+ \cdot \frac{w}{\|w\|} - x_- \cdot \frac{w}{\|w\|} \quad (4)$$

Como el producto de x y w se conoce de acuerdo a la ecuación 2, se reemplaza los valores de 2 en 4. (Castro & Pinedo, 2018).



$$D = \frac{(1-b)}{\|w\|} - \frac{(-1-b)}{\|w\|}$$

$$D = \frac{2}{\|w\|} \tag{5}$$

Se demuestra en la ecuación 5 que para determinar la distancia debemos dividir el valor de 2 entre el módulo del vector W. Pero como se requiere que la distancia sea máxima, lo que se debe realizar es minimizar el valor de: $\| w \|$.

El vector W está sujeto a limitaciones como se aprecia en la ecuación 2, por ello se reduce a una función con restricciones para determinar el valor de una función con restricciones existen varios métodos matemáticos, uno de ellos por su uso práctico es el método de los multiplicadores de Lagrange. (Castro & Pinedo, 2018).

El método de los multiplicadores de Lagrange facilita la optimización de los valores de una función multivariable sujeta a determinadas condiciones. Por ello se puede afirmar que buscar minimizar el módulo de W, es equivalente a minimizar el módulo de W al cuadrado sobre dos, que se expresa en la ecuación 6. (Castro & Pinedo, 2018).

$$L = \frac{\|w\|^2}{2} - \sum_i \alpha [y_i * (w \cdot x_i + b) - 1] \tag{6}$$

El siguiente paso es determinar el extremo de esta función, para ello se tendrá que hacer uso de las derivadas parciales de acuerdo a la variable a derivar, es decir w. (Castro & Pinedo, 2018).

$$\frac{\partial L}{\partial w} = w - \sum_i \alpha_i [y_i * (w \cdot x_i)] = 0$$

$$w = \sum_i [y_i * (w \cdot x_i)] \tag{7}$$

A partir de la ecuación 7, determinar w se reduce a una suma lineal para ciertos casos que hay en el espacio, y con el término ciertas podemos referirnos a que alfa será cero. Pero lo mencionado se ajusta a un contexto ideal, por lo que no se adecua a todos los casos. (Castro & Pinedo, 2018).



1.3.4.1. Funciones Kernel

Las SVM, nos permiten dar solución a problemas de clasificación lineal, si el problema a clasificar no es lineal en el espacio original de estudio se recurre a las funciones kernel que transforman el espacio original a un nuevo espacio en el cual el problema se vuelve lineal. Por ello si se presenta el problema de acuerdo a la figura 10 a en el cual se distinguen tres zonas concéntricas de diferente color que a simple vista se distingue que es un problema de clasificación no lineal. El darle solución requiere hacer uso de un kernel adecuado que permita transformar el espacio original a uno nuevo como se aprecia en la figura 10 b donde se aprecia un espacio de tres dimensiones con un hiperplano capaz de clasificar adecuadamente las clases de acuerdo a un margen máximo.

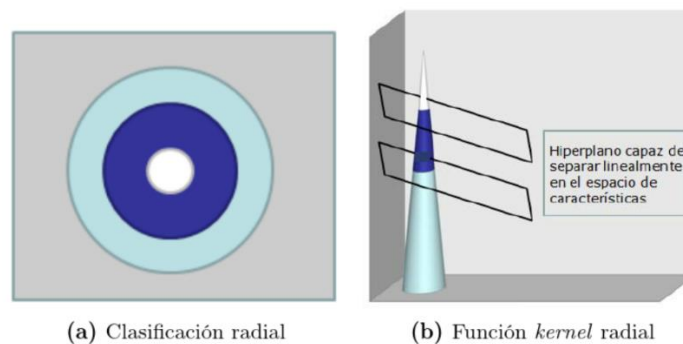


Figura 10: Ejemplo de clasificación y función kernel radial
Fuente: Gironés et al (2017)

1.3.4.2. Algoritmo SVM

En un problema de clasificación binaria cuyo espacio de características esta expresada por $x_i \in X$, así mismo cada característica pertenece a una clase binaria $c_i \in \{-1,1\}$. Para determinar si un nuevo punto pertenece a una clase negativa o positiva se necesita calcular el hiperplano que maximiza la distancia de clases positivas y negativas, con ello se afirmaría que el nuevo punto es de una clase u otra dependiendo si se encuentra a la izquierda o derecha del hiperplano de separación.



Si se presenta el caso descrito en el párrafo anterior nuestra función de clasificación quedaría expresada de acuerdo a la ecuación 8, el cual consiste en hallar la suma de las distancias afectadas por un peso w . (Gironés & Casas, 2017):

$$h(z) = \text{signo}(\sum_{i=1}^n w_i \cdot c_i \cdot k(x_i, z)) \quad (8)$$

Donde:

n cantidad de datos de entrenamiento

z punto a clasificar

$h(z) \in \{-1, 1\}$ función compacta de clasificación

$k: X \times X \rightarrow \mathbb{R}$ similitud de puntos de acuerdo a la función kernel

El grupo de pares $\{(x_i, c_i)\}_{i=1}^n$ son el conjunto de datos etiquetados. Es decir, conforman el conjunto de datos de entrenamiento, donde $c_i \in \{-1, 1\}$ es la etiqueta o clase del dato x_i .

$w_i \in \mathbb{R}$ pesos de los datos de entrenamiento

$\text{signo}()$ nos devuelve el signo de un número.

SVM trabaja en determinar los valores de w y b óptimos que logren maximizar el margen de separación de clases de acuerdo al hiperplano óptimo. (Gironés & Casas, 2017).

1.3.4.3. Tipos de funciones Kernel

La SVM entre sus funciones principales se encuentre determinar los vectores de soporte que son aquellos vectores que determinan la trayectoria del hiperplano separador. Los vectores de soporte marcan la frontera de separación máxima entre clases y garantizan la fiabilidad de a predicción. Por ello si se quiere valorar un modelo predictivo SVM se podría cuantificar aquellas predicciones que se encuentran cerca a los vectores de soporte con los que se encuentran alejado de los vectores de soporte. (Gironés & Casas, 2017).

Es por ello que a continuación se verá cada kernel de clasificación (lineal, polinomial, radial y sigmoidal).



Kernel lineal

La SVM de clasificación lineal usa la función kernel que se representa en la siguiente ecuación.

$$k(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T \cdot \mathbf{v} \quad (9)$$

donde $\mathbf{u}, \mathbf{v} \in X$.

Este tipo de kernel no requiere efectuar una transformación de espacio de características, por ello simplemente se trata del producto escalar de dos vectores de acuerdo a la ecuación 9. En la figura 11a se puede apreciar una superficie de predicción lineal. (Gironés & Casas, 2017).

Kernel polinomial

La función kernel polinomial se representa en la siguiente ecuación.

$$k(\mathbf{u}, \mathbf{v}) = (\gamma \mathbf{u}^T \cdot \mathbf{v} + b)^p \quad (10)$$

donde $\mathbf{u}, \mathbf{v} \in X$ y $\gamma > 0, b \geq 0$ y $p > 0$ son parámetros.

Es una generalización del producto escalar, cuyo espacio de trabajo es un espacio transformado de varias dimensiones del espacio original. Donde los parámetros γ, b, p , facilitan modelar la morfología de transformación del espacio original. En la figura 11b se puede apreciar la superficie de predicción de un modelo polinomial. (Gironés & Casas, 2017).

Kernel radial

La función kernel radial se representa en la siguiente ecuación.

$$k(\mathbf{u}, \mathbf{v}) = e^{-\gamma \cdot \|\mathbf{u} - \mathbf{v}\|^2} \quad (11)$$

donde $\mathbf{u}, \mathbf{v} \in X$ y $\gamma > 0$ es un parámetro.

Se conoce como kernel de Gauss, extiende la distancia de vector a vector en lugar del producto escalar. Es usada especialmente en casos complejos de índole no lineal. En la Figura 11c se muestra un gráfico que representa la superficie de predicción creada por el modelo



cuyo kernel es radial. La función es muy dependiente del parámetro γ el cual sirve para que el modelo no caiga en sobre entrenamiento (Gironés & Casas, 2017).

Kernel sigmoidal

La función kernel sigmoidal se representa en la siguiente ecuación.

$$k(u, v) = \text{tnh}(\gamma u^T \cdot v + b) \tag{12}$$

donde $u, v \in X$ y $\gamma > 0, b \geq 0$ son parámetros que permiten modelar la morfología de la transformación del espacio original. Hace uso de la tangente hiperbólica del producto escalar en el espacio original, siendo un caso que se desprende de la generalización del producto escalar. Modela relaciones complejas similares al kernel radial, en la Figura 11d se representa gráficamente el espacio de predicción creada por un modelo sigmoidal.

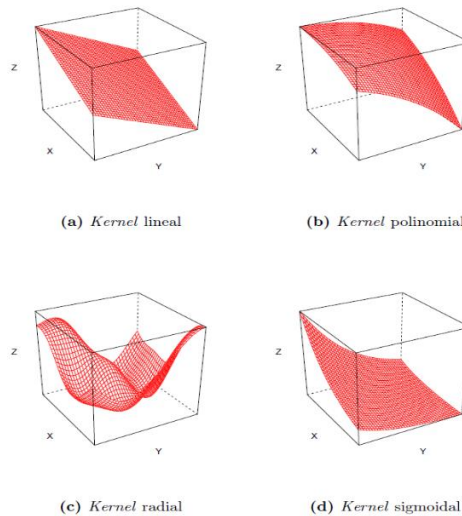


Figura 11. Superficies de predicción creadas por funciones kernel.

Fuente: Gironés et al (2017)



1.3.5. Regresión Logística

La Regresión Logística es una técnica estadística que permite estudiar la relación entre una o más variables independientes (predictoras) $X' = (x_1, x_2, \dots, x_p)$ y una variable dependiente (respuesta) Y siendo esta de tipo dicotómica ya que posee valores 1 o 0, con probabilidad π_i para $Y_i=1$ y probabilidad $1 - \pi_i$ para $Y_i = 0$, es decir la ocurrencia o no de un suceso, por ejemplo, moroso o no moroso, sano o enfermo, fumador o no fumador, spam o no spam, hipertenso o no hipertenso, etc. (Lizares, 2017).

El modelo logístico con “p” variables predictoras se representa en la siguiente ecuación.

$$\pi = \pi(x) = P(Y = 1|x) = \frac{e^{x'\beta}}{1 + e^{x'\beta}} \quad (13)$$

Donde “P” representa la probabilidad condicional que el evento $Y=1$ ocurra dada un conjunto de variables X' . La función logística tiene forma de “S”, tal como se muestra en la figura 12a la cual puede adoptar la forma inversa como muestra la figura 12b, y reduce los valores a los límites 0 y 1 ($0 \leq \pi(x) \leq 1$). (Pallarés, 2017).

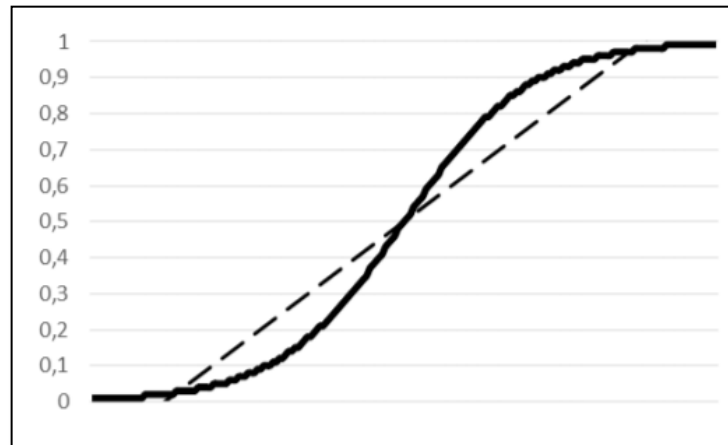


Figura 12 (a). Función Logística.

Fuente: Pallarés (2017)



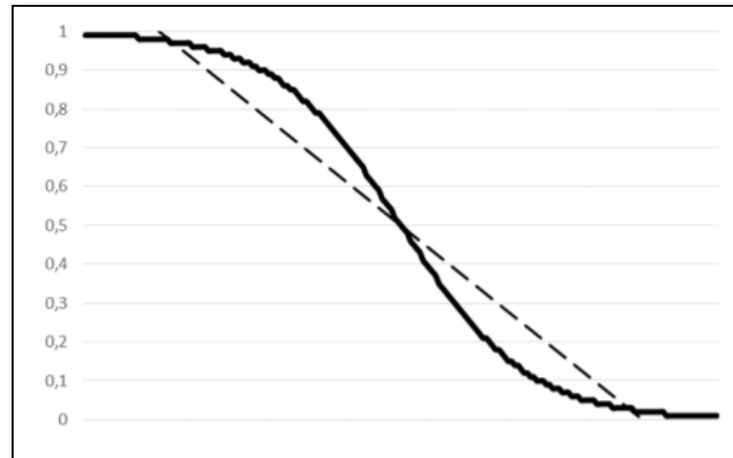


Figura 12 (b). Función Logística Inversa.
Fuente: Pallarés (2017)

Una transformación de $\pi(x)$ es fundamental dentro de la técnica de la regresión logística y es la transformación LOGIT, la cual se define en términos de $\pi(x)$ en la siguiente ecuación:

$$g(x) = \ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (14)$$

Donde β_0 es la constante y los β_i son los coeficientes de los predictores x_i del modelo. Es importante esta transformación ya que la función $g(x)$ adquiere mucha de las propiedades deseables de un modelo de regresión lineal. Los parámetros de la función LOGIT puede ser continuo, lineal y variar de $-\infty$ a $+\infty$, dependiendo del rango de x . (Lizares, 2017).

La ocurrencia de eventos (ODDS) puede expresar el modelo logístico, el cual se define como la razón expresada entre el cociente de la probabilidad de ocurrencia de un evento a la probabilidad de que no ocurra. Entonces sí:

$$\pi(x) = \text{Probabilidad de que el evento ocurra.}$$

$$1 - \pi(x) = \text{Probabilidad de que el evento no ocurra}$$

$$ODDS = \frac{\pi(x)}{1 - \pi(x)} \quad (15)$$



1.3.5.1. Regresión Logística Binaria Simple

La Regresión Logística Binaria Simple (RLBS) es el modelo más básico que presenta una sola variable predictora, y queda definida como indica la ecuación siguiente.

$$f(Y) = \log \left[\frac{P(y=1)}{1-P(y=1)} \right] = \beta_0 + \beta_1 x \tag{16}$$

Donde: f(Y)= función logística, P(y=1) probabilidad de ocurrencia, 1-P(y=1) probabilidad de no ocurrencia, β_0 es la constante y β_1 coeficiente del predictor x. La razón $P(y=1) / [1-P(y=1)]$ equivale a una ocurrencia de eventos (ODDS). Por ejemplo, cuando $P(y=1) = 0,75$, la ODDS equivale a $0,75/0,25=3,0$, lo que indica que la ocurrencia de un suceso es tres veces más probable que la no ocurrencia. (Pallarés, 2017).

La ecuación 16 utiliza el logaritmo de la ODDS, proceso denominado transformación logística (LOGIT) el cual se puede abreviar quedando la ecuación siguiente.

$$f(Y) = p(Y) = \frac{1}{1+e^{-y}} = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}} \tag{17}$$

Donde: f(Y)= función logística, p(Y) probabilidad de ocurrencia o no ocurrencia de un evento, β_0 es la constante y β_1 coeficiente del predictor x.

1.3.5.2. Regresión Logística Binaria Multinomial

Los diversos casos en estudio son bastante complejos y suelen necesitar más de una variable predictora para analizarlos, con el fin de otorgar soluciones a los problemas que se plantean. El modelo de regresión logística binaria multinomial (RLBM) está conformado por más variables predictoras que facilitan la comprensión del caso a predecir. La cual está definida por la siguiente ecuación. (Pallarés, 2017).

$$f(Y) = p(Y) = \frac{1}{1+e^{-y}} = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\dots+\beta_p X_p)}} \tag{18}$$

Donde: f(Y)= función logística, p(Y) probabilidad de ocurrencia o no ocurrencia de un evento, β_0 es la constante y los $\beta_1 \dots \beta_p$ son los coeficientes de los predictores $X_1 \dots X_p$ del modelo.



1.3.5.3. Estimación de los Coeficientes de la Regresión Logística

Dado un conjunto de variables predictoras $X = (x_1, x_2, \dots, x_p)$ y siendo la distribución de Y no normal, la estimación del vector $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)$ por el método de mínimos cuadrados no es óptimo, en su lugar se emplea el método de máxima verosimilitud que facilita la determinación de cada uno de los parámetros desconocidos que amplía la probabilidad de encontrar el grupo de datos observados. (Lizares, 2017).

Para estimar los parámetros desconocidos $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)$ del modelo se utilizó el método de máxima verosimilitud detallado en el párrafo anterior, con lo cual determinamos el valor de β que maximiza la función de verosimilitud $l(\beta)$, quedando definida en la siguiente ecuación.

$$l(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \tag{19}$$

Donde: $\prod_{i=1}^n$, productoria de $i = 1, \dots, n$ predictores, $\pi(x_i)^{y_i}$ probabilidad de ocurrencia de un evento, $[1 - \pi(x_i)]^{1-y_i}$ probabilidades de no ocurrencia de un evento.

Aplicando el logaritmo neperiano en la ecuación 19 la función de verosimilitud $L(\beta)$ se define como:

$$L(\beta) = \ln[l(\beta)] = \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\} \tag{20}$$

Para encontrar el valor de β se deriva $L(\beta)$ con respecto a $\beta_0, \beta_1, \dots, \beta_p$ y se iguala a cero, obteniéndose las siguientes ecuaciones:

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0 \tag{21}$$

$$\sum_{i=1}^n x_{ij} [y_i - \pi(x_i)] = 0, j = 1, \dots, p \tag{22}$$

Para dar solución a este conjunto de ecuaciones se debe recurrir al método iterativo de Newton-Raphson. (Lizares, 2017).

1.3.6. Evaluación de las técnicas computacionales y estadísticas

La evaluación de los modelos de predicción se efectúa haciendo uso de métodos apropiados como la matriz de confusión y Curvas ROC. (Carpio, 2016).



1.3.6.1. Matriz de Confusión o Matriz de error

Método de evaluación que permite en forma gráfica y compacta brindar información sobre casos de éxito y fracaso en la predicción. Donde la clase real se muestran en las filas y las predicciones de cada clase en columnas como se aprecia en la tabla 1. (Trujillo, 2017).

Tabla 1
Matriz de Confusión.

		Clase predicha	
		Clase = No	Clase = Sí
Clase real	Clase = No	TN	FP
	Clase = Sí	FN	TP

Fuente: Trujillo (2017)

Donde:

Verdadero Positivo (TP): Correctamente clasificado como positivo (moroso).

Verdadero Negativo (TN): Correctamente clasificado como negativo (no moroso).

Falso Positivo (FP): Incorrectamente clasificado como positivo. (No moroso identificado como moroso)

Falso Negativo (FN): Incorrectamente clasificado como negativo. (Moroso identificado como no moroso)

La matriz de 2x2 que se aprecia en la tabla 1 permite evaluar las predicciones en clases binarias, sin embargo, hay matrices para problemas multiclase.

A partir de la matriz de confusión se determinan las siguientes métricas:

Exactitud: La exactitud o accuracy (AC) representa el porcentaje de acierto del modelo y se determina al dividir el conjunto de casos clasificados correctamente entre todas las instancias. (Trujillo, 2017).



$$AC = \frac{TP+TN}{TP+TN+FP+FN} \quad (23)$$

Precisión: La precisión o precision (P) mide el porcentaje de casos positivos que han sido identificados, es decir, con qué frecuencia predice correctamente un caso positivo porque en el conjunto de datos que predice hay falsos positivos (no morosos identificado como moroso). (Trujillo, 2017).

$$P = \frac{TP}{TP+FP} \quad (24)$$

Sensibilidad: La sensibilidad (S) mide el porcentaje de sujetos que presentan la característica estudiada y son clasificados correctamente. (Trujillo, 2017).

$$S = \frac{TP}{TP+FN} \quad (25)$$

Especificidad: La especificidad (E) mide el porcentaje de personas que no tienen la característica estudiada y son clasificados correctamente. (Trujillo, 2017).

$$E = \frac{TN}{TN+FP} \quad (26)$$

1.3.6.2. Curva ROC

La curva ROC (Receiver Operation Characteristic) es una gráfica que se construye teniendo en cuenta en el eje horizontal los valores de la tasa de falsos positivos (1-Especificidad) y en el eje vertical los valores de la sensibilidad, donde diferentes puntos de corte indican la capacidad predictiva de un modelo orientado a la clasificación binaria, que se aprecia en la siguiente figura. (Trujillo, 2017).



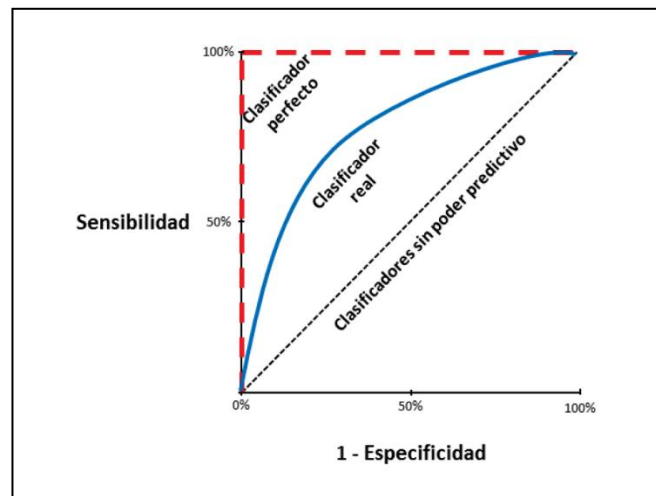


Figura 13. Curva ROC
Fuente: Trujillo (2017)

Para efectuar comparación de modelos predictivos haciendo uso de la curva ROC se debe reducir a un valor escalar apropiado que muestre el rendimiento alcanzado. Por ello el método práctico es determinar el área bajo la curva ROC (AUC), siendo esta una fracción del área de la unidad de un cuadrado su valor se encontrará entre [0,1]. De manera general el área bajo la curva (AUC) se interpreta de acuerdo a la escala que se muestra en la tabla 2 (Trujillo, 2017):

Tabla 2
Puntuación AUC

ESCALA	AUC
Excepcional	0.9 – 1.0
Excelente / bueno	0.8 – 0.9
Aceptable / razonable	0.7 – 0.8
Pobre	0.6 – 0.7
Fallido	0.5 – 0.6

Fuente: Trujillo (2017)



1.3.7. Archivos .csv

Según Rizzo (2008) los datos a menudo se suministran en formato de valores separados por comas (.csv), que es un archivo de texto que separa los datos con caracteres de texto especiales llamados delimitadores. Los archivos en formato .csv se pueden abrir en la mayoría de las aplicaciones de hoja de cálculo. Los datos de la hoja de cálculo se deben guardar en formato .csv antes de importarlos en R. En un archivo .csv, es probable que las fechas se den como cadenas, delimitadas por comillas dobles.

1.3.8. Librería para aprendizaje automático

Lenguaje R: Es un poderoso lenguaje estadístico que puede usarse para manipular y analizar datos, proporciona muchos paquetes de aprendizaje automático y funciones de visualización, que permiten a los usuarios analizar datos generados por los procesos y operaciones en marcha en cada unidad de negocio. Lo más importante, R es de código abierto y gratuito. También R simplifica enormemente el aprendizaje automático, por ello solo se necesita saber cómo cada algoritmo puede resolver su problema, y luego simplemente puede usar un paquete escrito para generar rápidamente modelos de predicción en datos con unas pocas líneas de comando. Por ejemplo, puede realizar Naïve Bayes para el filtrado de correos no deseados, realizar clusters k-means para la segmentación del cliente, usar la regresión lineal para pronosticar los precios de la vivienda o implementar un modelo oculto de Markov para predecir el mercado bursátil (Wei & Chiu, 2015), como se muestra en la figura 14.



Figura 14. Predicción bursátil usando R

Fuente: Wei & Chiu (2015)



1.3.9. Definición de términos básicos.

- a) **Modelo predictivo:** Colección de técnicas matemáticas que tienen en común el objetivo de encontrar una relación matemática entre un objetivo, respuesta o variable "dependiente" y varias variables predictoras o "independientes" con el objetivo en mente de medir valores futuros de esos predictores e insertarlos en la relación matemática para predecir los valores futuros de la variable objetivo. (Espino, 2017).
- b) **Técnicas de Predicción:** Conjunto de métodos que permiten alcanzar pronósticos en un espacio temporal por medio de información hasta el momento actual (Alea, 2005).
- c) **Eficiencia:** Logro de metas y objetivos trazados con el uso mínimo de recursos y tiempo disponible. (Hernández, 2006).
- d) **Comparación:** Se denomina comparación a la acción de fijar la atención en dos o más objetos con la finalidad de descubrir sus semejanzas o diferencias. (Real Academia Española, 2018)
- e) **Morosidad:** Se denomina morosidad cuando una persona natural o jurídica teniendo una obligación de pago no cumple en el tiempo estipulado, siendo una práctica recurrente. (Zavaleta, 2017).

1.4. Formulación del Problema.

¿Qué técnica de predicción es más eficiente para predecir la morosidad en cuotas sociales del colegio de Ingenieros del Perú consejo departamental de Lambayeque?

1.5. Justificación e importancia del estudio.

La presente investigación es de gran importancia ya que permite conocer cuál de las dos técnicas predictivas es más eficiente al predecir la morosidad. Además, es pertinente porque organizaciones públicas y privadas, al transcurrir los años no escapan de altos porcentajes de morosidad que podrían disminuir al conocer técnicas eficientes en predicción, lo que haría que sean más competitivas y líderes en su sector. La investigación en diversos sectores está en crecimiento haciendo uso de las diversas técnicas de predicción ya sean técnicas de



aprendizaje automático o los clásicos estadísticos, logrando cada vez mejores resultados de acuerdo al ámbito de estudio y los datos que emplean.

La investigación es importante para las futuras investigaciones ya que aportará bases sólidas que servirán de pautas para el desarrollo de mejoras en las técnicas de predicción de morosidad de acuerdo a sus datos.

A la sociedad científica la investigación aporta con la generación de conocimiento fruto de la comparación de dos técnicas predictivas (aprendizaje automático y estadístico) que forman parte de la inteligencia artificial cuya rama es la ingeniería de datos. Teniendo en cuenta que actualmente hay poca investigación sobre la predicción de morosidad.

1.6. Hipótesis.

La técnica máquina de soporte vectorial es más eficiente que la técnica de regresión logística en la predicción de morosidad de cuotas sociales del Colegio de Ingenieros del Perú consejo departamental de Lambayeque.

1.7. Objetivos de la investigación

1.7.1. Objetivo General

Comparar máquina de soporte vectorial y regresión logística en la predicción de morosidad de cuotas sociales en el Colegio de Ingenieros del Perú consejo departamental Lambayeque.

1.7.2. Objetivos Específicos

- a) Seleccionar las técnicas predictivas.
- b) Procesar los datos históricos de los colegiados.
- c) Realizar proceso de extracción, transformación y carga (ETL) para la extracción de datos relevantes.
- d) Aplicar las técnicas máquina de soporte vectorial y regresión logística.
- e) Evaluar los resultados obtenidos.



CAPITULO II: MATERIAL Y MÉTODO

2.1. Tipo y diseño de la investigación

Tipo de investigación

La investigación se basó en la metodología cuantitativa, porque los indicadores están representados a través de cantidades numéricas como, precisión, exactitud, sensibilidad, especificidad y área bajo la curva.

De acuerdo a González & Torres (2017) la metodología cuantitativa a través de los años de investigación se posiciona como una de las metodologías más usadas tradicionalmente por los buenos resultados que se obtienen en aquellas ciencias que se valen de aspectos observables susceptibles de medición, por medio de la estadística para el análisis de los datos.

Diseño de investigación

El diseño de la investigación fue Cuasi- Experimental, como afirma Hernández, Fernández, & Baptista (2014) que este tipo de diseño en el desarrollo de una investigación interviene en la variable independiente y visualiza el efecto sobre una o más variables dependientes; teniendo en cuenta lo antes citado, en la investigación se estimó la eficiencia de la técnica máquina de soporte vectorial y regresión logística (variable independiente), en la predicción de morosidad de cuotas sociales del colegio de Ingenieros Consejo departamental de Lambayeque (Variable dependiente).

2.2. Población y muestra.

2.2.1. Población

Se elaboró un top de 12 técnicas de predicción tomando en cuenta su precisión de predicción de acuerdo a las investigaciones relacionadas a la predicción de riesgos de crédito obtenidas de la base de datos IEEE Xplore Digital Library (IEEE, 2018) referenciado en el Anexo 2.

2.2.2. Muestra

La muestra de la investigación es elegida por conveniencia, donde los algoritmos seleccionados son máquina de soporte vectorial y regresión logística.



2.3. Variables, Operacionalización.

2.3.1. Variable independiente

Máquina de soporte vectorial y regresión logística.

2.3.2. Variable dependiente

Predicción de Morosidad de cuotas sociales consejo departamental de Lambayeque.

2.3.3. Operacionalización



Tabla 3
Operacionalización de variables.

VARIABLE INDEPENDIENTE	Dimensiones	Indicadores	Unidad de medida	Fórmula	Descripción
Técnicas predictivas de Máquina de soporte vectorial y regresión logística.	Matriz de Confusión	Precisión	%	$\frac{TP}{TP + FP}$	TP=Verdadero positivo FP=Falso positivo
		Exactitud	%	$\frac{TP + TN}{TP + TN + FP + FN}$	TP=Verdadero positivo TN=Verdadero negativo FP=Falso positivo FN=Falso negativo
		Sensibilidad	%	$\frac{TP}{(TP + FN)}$	TP=Verdadero positivo FN=Falso negativo
		Especificidad	%	$\frac{TN}{(TN + FP)}$	TN=Verdadero negativo FP=Falso positivo
		AUC	%	$\frac{1}{2}(S + E)$	S=Sensibilidad E=Especificidad
	Curva ROC (Característica Operativa del Receptor)	(área bajo la curva)			
VARIABLE DEPENDIENTE	Dimensiones	Indicadores	Unidad de medida	Fórmula	Descripción
Predicción de Morosidad de cuotas sociales del Colegio de Ingenieros del Perú consejo departamental de Lambayeque.	Tiempo	Tiempo para generar estimación (morosidad)	segundos	$T = TF - TI$	TF = Tiempo final para generar estimación (morosidad). TI = Tiempo inicial para generar estimación (morosidad). T = Tiempo en segundos.

Fuente: Elaboración propia



2.4. Técnicas e instrumentos de recolección de datos

En la presente investigación se eligió aplicar como técnicas el análisis documental, la observación y como instrumento la guía de observación.

Análisis documental. Se eligió aplicar el análisis documental debido que este tiene fines de orientación científica e informativa, como son los libros, tesis, papers y artículos científicos leídos.

Observación. Se eligió aplicar la observación científica debido a que observaremos lo que acontece en cada prueba realizada, agrupando los acontecimientos de acuerdo al problema que se estudia.

Instrumento de Recolección de Datos. Se eligió como instrumento la ficha de observación, en la cual se registró la precisión, exactitud, sensibilidad, especificidad, área bajo la curva (AUC) de las técnicas de predicción y el tiempo de estimación de la predicción de morosidad. Los datos registrados permiten efectuar cálculos estadísticos y generar gráficos que ayudan a interpretar cuán eficiente son las técnicas elegidas.

2.5. Procedimiento de análisis de datos

Para evaluar cada técnica en la presente investigación, se tomará como referencia la matriz de confusión que permite calcular las métricas de exactitud, precisión, sensibilidad, especificidad. Así como la curva ROC que permite calcular el área bajo la curva (AUC).

De acuerdo a las siguientes fórmulas:

Exactitud: Evalúa el porcentaje de acierto del modelo y se determina al dividir el conjunto de casos clasificados correctamente entre todas las instancias. (Trujillo, 2017).

$$Exactitud = \frac{Verdadero\ Positivo + Verdadero\ Negativo}{Verdadero\ Positivo + Verdadero\ Negativo + Falso\ Positivo + Falso\ Negativo} \quad (27)$$

Donde:

Verdadero Positivo: El modelo predice que el colegiado es moroso y efectivamente el colegiado es moroso.



Verdadero Negativo: El modelo predice que el colegiado no es moroso y efectivamente el colegiado no es moroso.

Falso Positivo: El modelo predice que el colegiado es moroso, pero en realidad no lo es.

Falso Negativo: El modelo predice que el colegiado no es moroso, pero en realidad es moroso.

Precisión: Mide el porcentaje de casos positivos que han sido identificados, es decir, con qué frecuencia predice correctamente un caso positivo porque en el conjunto de datos que predice hay falsos positivos (no morosos identificado como moroso). (Trujillo, 2017).

$$Precisión = \frac{Verdadero\ Positivo}{Verdadero\ Positivo + Falso\ Positivo} \quad (28)$$

Donde:

Verdadero Positivo: El modelo predice que el colegiado es moroso y efectivamente el colegiado es moroso.

Falso Positivo: El modelo predice que el colegiado es moroso, pero en realidad no lo es.

Sensibilidad: Mide la proporción de sujetos que presentan la característica estudiada y son clasificados correctamente. (Trujillo, 2017).

$$Sensibilidad = \frac{Verdadero\ Positivo}{Verdadero\ Positivo + Falso\ Negativo} \quad (29)$$

Donde:

Verdadero Positivo: El modelo predice que el colegiado es moroso y efectivamente el colegiado es moroso.

Falso Negativo: El modelo predice que el colegiado no es moroso, pero en realidad es moroso.

Especificidad: La especificidad (E) mide el porcentaje de personas que no tienen la característica estudiada y son clasificados correctamente. (Trujillo, 2017).

$$Especificidad = \frac{Verdadero\ Negativo}{Verdadero\ Negativo + Falso\ Positivo} \quad (30)$$



Donde:

Verdadero Negativo: El modelo predice que el colegiado no es moroso y efectivamente el colegiado no es moroso.

Falso Positivo: El modelo predice que el colegiado es moroso, pero en realidad no lo es.

Área bajo la Curva (AUC): Mide la capacidad predictiva de la técnica de predicción cuyo valor aceptable va de 0.8 a 1.

$$AUC = \frac{1}{2}(\text{Sensibilidad} + \text{Especificidad}) \quad (31)$$

Los resultados de las métricas se basan en el:

- a. Uso de tablas, para evaluar resultados de la técnica de aprendizaje automático y estadístico.
- b. Uso de gráficos estadísticos, para evaluar resultados de la técnica de aprendizaje automático y estadístico.

2.6. Criterios éticos.

Confidencialidad: De cada una de las personas que contribuyeron como informantes de la investigación se asegura la protección de su identidad.

Objetividad: La realidad encontrada se analiza con base en criterios técnicos e imparciales.

Originalidad: La información obtenida se citan de acuerdo a la fuente bibliográfica.

Veracidad: La información expuesta es verdadera cuidando el criterio de confidencialidad.

2.7. Criterios de Rigor científico.

Consistencia: La investigación presentó material consistente y certificado por la comunidad científica.

Fiabilidad: La investigación cumplió con los estándares y políticas en su contenido exigidos por la comunidad científica.



Validez: Los resultados obtenidos producto de la investigación fueron evaluados y analizados por especialistas en el tema para determinar su autenticidad.

Transferibilidad: La investigación proporciona información y conocimiento que puede ser transferida a investigadores que se enfoquen en contextos similares.

Neutralidad: El desarrollo minucioso de la investigación garantiza que los resultados obtenidos, no pueden ser alterados o desviados por motivaciones, intereses, y/o perspectivas del investigador.

2.8. Modelo Propuesto

La investigación inició desde la elección de las técnicas de predicción, esto consiste en elaborar un top de técnicas predictivas con base en la revisión de la literatura de distintas investigaciones realizadas con anterioridad. Posteriormente se recopiló los datos históricos de los colegiados los que se obtuvieron de fuentes internas y externas de la organización. A continuación, los datos se procesaron haciendo uso del proceso de Extracción, Transformación y Carga (ETL), esto consiste en explorar los datos, eliminar datos duplicados, incompletos y campos vacíos; consiguiendo la extracción de datos relevantes. Finalmente se aplicó la técnica de máquina de soporte vectorial y regresión logística sobre la información procesada y se verificaron los resultados.

A continuación, se muestra el flujograma de la presente investigación.

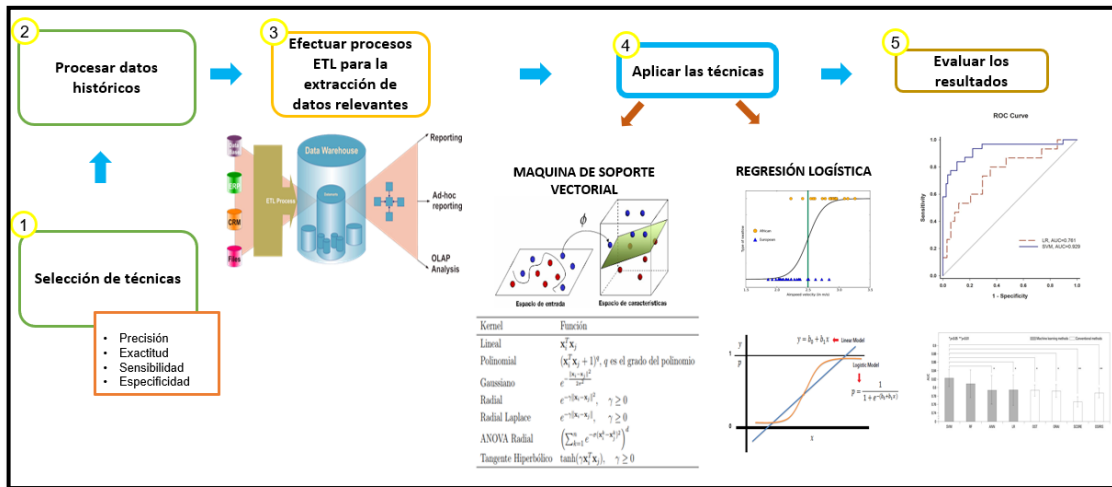


Figura 15. Flujo de la realización de la investigación

Fuente: Elaboración propia.

El presente trabajo de investigación pertenece a un grupo de investigadores los cuales son: Máximo Gabriel Silva Parraguez, José Alberto Guevara Barreto, Juan Carrasco Manay, Flor Díaz Macalopu, Elmer Anthony Monja Sandoval, Víctor Martínez Panta, Darwin Alain Vásquez Cercado y Jefferson Napoleón Huamán Bernilla. Al cual se le denominó equipo de investigación.

2.8.1. Selección de técnicas predictivas

Se realizó una revisión de artículos científicos cuyas técnicas fueron usados en un contexto próximo a la investigación que son del ámbito financiero, el cual se detalla en el anexo 1. Con ello se elaboró un top de técnicas de predicción tanto de aprendizaje automático y técnicas estadísticas tomando en cuenta la precisión, que mide la calidad de respuestas positivas del predictor, que se aprecia en el anexo 2. De acuerdo a la elevada precisión que muestran en los trabajos analizados se eligen por conveniencia las técnicas máquina de soporte vectorial y regresión logística.

2.8.2. Procesamiento de los datos históricos de los colegiados

2.8.2.1. Proceso de Adquisición

De acuerdo a Reyes & León (2014) la predicción requiere la recolección de datos de diferentes orígenes de datos, es decir internos y externos a la organización y deben tener



suficiente consistencia histórica de acuerdo a las variables predictoras relevantes del caso en estudio. Por ello en la presente investigación se recurrió a la revisión de diferentes orígenes de datos.

Se recopiló los datos históricos de los colegiados de los 7 capítulos que forman la CD Lambayeque con un total de 15000 colegiados, cuya información lo brindó el área de Informática del Colegio de Ingenieros Consejo Departamental de Lambayeque.

La información de los colegiados con sus aportes de pago de cuotas sociales, se encontró en el Gestor de Base de Datos “SQL Server” y la información complementaria se encontró en hojas de cálculo (XLSX) tal como se muestra en la figura 16 y figura 17. Además, para complementar la información requerida se optó por efectuar consultas externas al portal institucional de ESSALUD el cual se aprecia en el anexo 3.

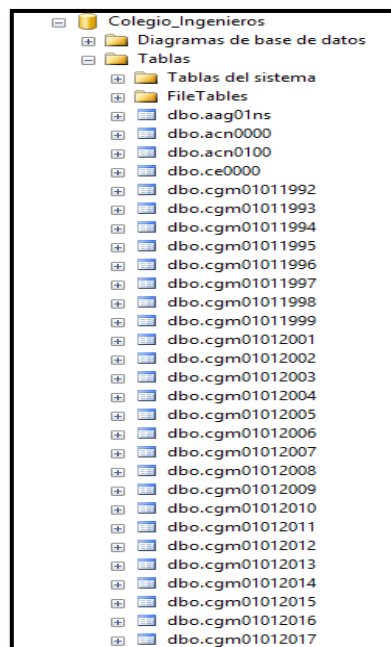


Figura 16. Base de datos del Colegio de Ingenieros del Perú – CD Lambayeque.
Fuente: Base de datos Colegio de Ingenieros CD Lambayeque.



N° CIP	Nombre	DNI	Dirección	Fec. Nacimiento	Especialidad
26816	ABAD ARRIAGA LUIS ANTONIO	68166816	JAÉN 01	30/8/1956	AGRONOMICA
93586	ABAD SÁNCHEZ JOSÉ DANIEL	16422647	CA. PABLO OLAVIDE N°200A- COSTADO REAL PLAZA - CHICLAYO	1/7/1960	MECANICO ELECTRICISTA
95553	ABANTO SÁNCHEZ DENNIS	40548607	Heros Civiles 148-Chiclayo/ 2 de Mayo 154	24/8/1979	INFORMATICO Y DE SISTEMAS
202642	ABANTO SANTA CRUZ IHUSTY ANDHERSON	71717744	LOS TERRAZOS 1973 INCA MANCO CAPAC	22/12/1992	AMBIENTAL
68136	ABASALO MEOÑO ROSA IMELDA	16621122	Incanato N° 464 - José Leonardo Ortiz	30/1/1975	QUIMICA
24739	ABÁSULO TORRES CARLOS ALBERTO	47394739	Francisco Cabrera 882	10/2/1954	PESQUERA
181292	ACEDO NECIOSUP IVAN ALEJANDRO	43549645	RICARDO PALMA	5/4/1986	SISTEMAS Y COMPUTACION
175439	ACEDO REYES JULIO CÉSAR	72553721	CALLE ALGARROBO 126	3/5/1991	AGRONOMIA
94191	ACEVEDO RODRÍGUEZ ANA DE LOS MILAGROS	40626066	Los Tumbes 133-Dtpto.202-Sta.Victoria	19/10/1980	CIVIL
193118	ACHA OSPINO ELIZABETH CECILIA	46978723	CALLE JOSE MARIA ARGUEDAS 331 C.P.M EL BOSQUE	5/12/1990	EN INDUSTRIAS ALIMENTARIAS
94232	ACHA YSIQUE JUAN RAMÓN	80276735	Calle San Antonio 135-J.LO	25/12/1977	AGRICOLA
81072	ACOSTA ACOSTA CARLOS ALBERTO	16796022	Calle Dos N°190 Ampliación 9 de Octubre	3/12/1977	CIVIL
75913	ACOSTA ACOSTA EDILBERTO	17612859	los girasoles 1270 p. j .santa rosa - Lambayeque	13/3/1973	AGRICOLA
113573	ACOSTA AGUILAR LUIS OSCAR	41837302	Andrés Razuri 218 - Lambayeque	18/3/1980	AGRICOLA
156453	ACOSTA AGURTO MARCO ANTONIO	46552168	MZ 100 LOTE 05 ANTONIO RAYMONDI	27/6/1990	CIVIL
148343	ACOSTA BARTURÉN KATHERINE GISELLE	43008095	CALLE CACIQUE CINTO N° 284 - URB. LATINA	28/3/1985	INGENIERIA DE SISTEMAS
84542	ACOSTA CASTILLO MARCO ANTONIO	17527604	Mz° c° Lt 7 Urb. Los Rosales - Lambayeque	9/10/1960	AGRICOLA
79752	ACOSTA CASTRO RONY FERNANDO	16760676	Resd. El Carmen Block F-102-Lambayeque	26/8/1976	CIVIL
97992	ACOSTA CORONADO LINSAY MILAGROS	41790980	CONDOMINIO EUCALIPTO EDIF. 31 DPTO. 201 - EL AGUSTINO - LIMA	1/7/1983	EN INDUSTRIAS ALIMENTARIAS
176815	ACOSTA CORONEL ANDRÉS ALEJANDRO	43449500	ANEXO CONCHUCOS N° 07	15/2/1986	CIVIL
203879	ACOSTA DE LA CRUZ JHONN ANGELLO	45730521	RIO GRANDE	2/5/1989	MECANICO ELECTRICISTA
187598	ACOSTA GRANADOS IRENE CAROL	46774862	CALLE ELVIRA GARCIA Y GARCIA 701-1 P JOVEN SAN MARTIN	9/12/1990	ZOOTECNISTA
211783	ACOSTA LUCERO JOSE LUIS	46569137	P JOVEN 19 DE SETIEMBRE MZ B LT 20	9/5/1990	EN INDUSTRIAS ALIMENTARIAS
30090	ACOSTA RUESTA FERNANDO AMADO	17526995		28/7/1958	ZOOTECNIA
42580	ACOSTA SANTISTEBAN JOSÉ	16403340	Callao 131- Sta. Victoria	5/9/1965	ELECTRICISTA
36193	ACOSTA SANTISTEBAN OSWALDO	16423893	Av.Sesquicentenario # 896 - Sta. Victoria	5/8/1962	CIVIL
74704	ACOSTA TORRES JIMMY OMAR	17633180	Mariano Melgar #188 - Urb.Próceres	25/10/1977	MECANICO ELECTRICISTA
39219	ACOSTA VIDAURRE ROGELIO	17545801	ELVIRA GARCIA Y GARCIA 701 -LAMBAYEQUE	10/1/1954	ZOOTECNIA
45415	ACOSTA ZEÑA LUIS ALBERTO	17548092	LOS GERANIOS N°299- URB. MIRAFLORES- LAMBAYEQUE	16/10/1965	AGRONOMO
30100	ACUÑA CARAMULLAS OSCAR	16403802	Mz. B. Lt. 8 Urb. Arturo Cabrera Fallas - Chiclayo 370	18/7/1964	CIVIL

Figura 17. Hojas de cálculo con datos de colegiados de las diversas especialidades CD Lambayeque desde el año 1994 hasta 2017.
Fuente: Base de datos CIP Lambayeque.

2.8.2.2. Selección de las variables

De acuerdo a Yeh & Lien (2009) las variables a considerar en la predicción pueden ser cuantitativas y cualitativas de acuerdo al caso en estudio. De la misma forma Trujillo (2017) cuya investigación realizó en predicción de riesgos de crédito consideró variables categóricas como numéricas. Tomando como modelo los estudios descritos, para el presente trabajo de investigación se consideró las siguientes variables Género, Estado Civil, Edad, Número de Hijos, Tipo de Seguro, Especialidad, Categoría, Monto de pago y Días de incumplimiento que se describe en la tabla 4.



Tabla 4
Descripción de las variables utilizadas en la investigación

Variable	Descripción
Género	Género del colegiado (Masculino, Femenino)
Estado Civil	Estado civil al cual pertenece el colegiado (Soltero, Casado, Viudo, Divorciado).
Edad	Edad de cada colegiado.
Número de Hijos	Número de hijos de cada colegiado.
Tipo de Seguro	Tipo de seguro al cual pertenece cada colegiado (ESSALUD, SIS, Pacífico Seguros, Rímac Seguros).
Especialidad	Especialidad del colegiado (Informático y Sistemas, Ambiental, Química, Pesquería, Sistemas y Computación, Agrícola, Agroindustrial y Comercio, Agroindustrias e Industrias Alimentarias, Agroindustrial, Ambiental y Sanitario, Automotriz, etc.).
Categoría	Categoría al cual pertenece cada colegiado (Ordinario, Vitalicio).
Monto de Pago	Monto mensual que paga cada colegiado de acuerdo con su categoría al cual pertenece.
Días de incumplimiento	Días que se retrasó cada colegiado en pagar su cuota mensual.

Fuente: Equipo de Investigación.

2.8.2.3. Descripción de las tablas que conforman la Base de Datos de acuerdo a las variables seleccionadas

De acuerdo a las variables seleccionadas se procedió a efectuar una limpieza de tablas irrelevantes que conforman la base de datos, describiendo cada campo minuciosamente como se aprecia en la tabla 5, para ello se elaboró hojas de cálculo y se hizo uso de la metodología de la matriz RACI cuyos ítems que se utilizaron son los siguientes SR, I y D que se describen en la Tabla 6. La descripción detallada de todas las tablas que conforman la base de datos se puede ver en el Anexo 4.

Tabla 5

Descripción de las tablas que conforman la base de datos SQL de acuerdo a las variables identificadas para la predicción de morosidad de cuotas sociales del Colegio de Ingenieros CD Lambayeque.

Nro	Tabla	Descripción	Nro Registros	VARIABLES DE PREDICCIÓN								
				Género	Estado Civil	Edad	Monto Pagado	Número de Hijos	Días de Incumplimien	Capítulo	Especialidad	Categoría
1	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR
2	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR
3	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR
4	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR
5	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR
6	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR
7	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR
8	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR
9	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR
10	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación

Tabla 6

Descripción de Ítems de acuerdo con la Matriz RACI.

Ítems	Descripción
SR	Sin relación
I	Relación indirecta
D	Relación Directa

Fuente: Equipo de Investigación



2.8.2.4. Selección de tablas de acuerdo a sus atributos relacionados con las variables para el proceso de análisis.

Se seleccionó las tablas que contienen información relacionada con las variables de predicción para su procesamiento como se aprecia en la tabla 7.

Tabla 7

Tablas seleccionadas cuya información tiene relación con las variables de predicción.

N°	Tabla	Descripción	Nro Registros	Género	Estado Civil	Edad	VARIABLES DE PREDICCIÓN					
							Monto de pago SR	Meses de pago SR	Número de Hijos SR	Capítulo	especialidad	Categoría
1	CAMPO	PROTEGIDO	-	I	D	I	SR	SR	SR	D	D	SR
2	CAMPO	PROTEGIDO	-	SR	SR	SR	D	D	SR	D	SR	SR
3	CAMPO	PROTEGIDO	-	SR	SR	SR	D	D	SR	D	SR	SR
4	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	I	SR	SR	SR
5	CAMPO	PROTEGIDO	-	SR	SR	SR	D	D	SR	SR	SR	SR
6	CAMPO	PROTEGIDO	-	I	D	I	SR	SR	SR	D	D	D
7	CAMPO	PROTEGIDO	-	I	SR	SR	D	D	SR	SR	SR	SR
8	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	D	SR	SR
9	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	I	D	SR
10	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	I	SR	SR	SR

Fuente: Equipo de Investigación.



2.8.2.5. Descripción detallada de cada tabla seleccionada de acuerdo a las variables a utilizar

De cada tabla seleccionada se describió sus atributos con relación a las variables de predicción, como se aprecia en la tabla 8. La descripción detallada de todas las tablas seleccionadas se puede ver el anexo 5.

Tabla 8

Descripción detallada de cada campo de las tablas seleccionadas de la base de datos Colegio de Ingenieros CD Lambayeque.

Tabla	Nro Registros	Columnas	Descripción	VARIABLES DE PREDICCIÓN										
				Género	Estado Civil	Edad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimie	Capítulo	Especialidad	Categoría	
data_web_09012014		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	I	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	I	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	D	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación.



2.8.2.6. Modelo Entidad Relación de tablas seleccionadas de la base de datos Colegio de Ingenieros CD Lambayeque.

Se procedió a construir un modelo Entidad-Relación para el mejor manejo de los datos como se muestra en la figura 18.

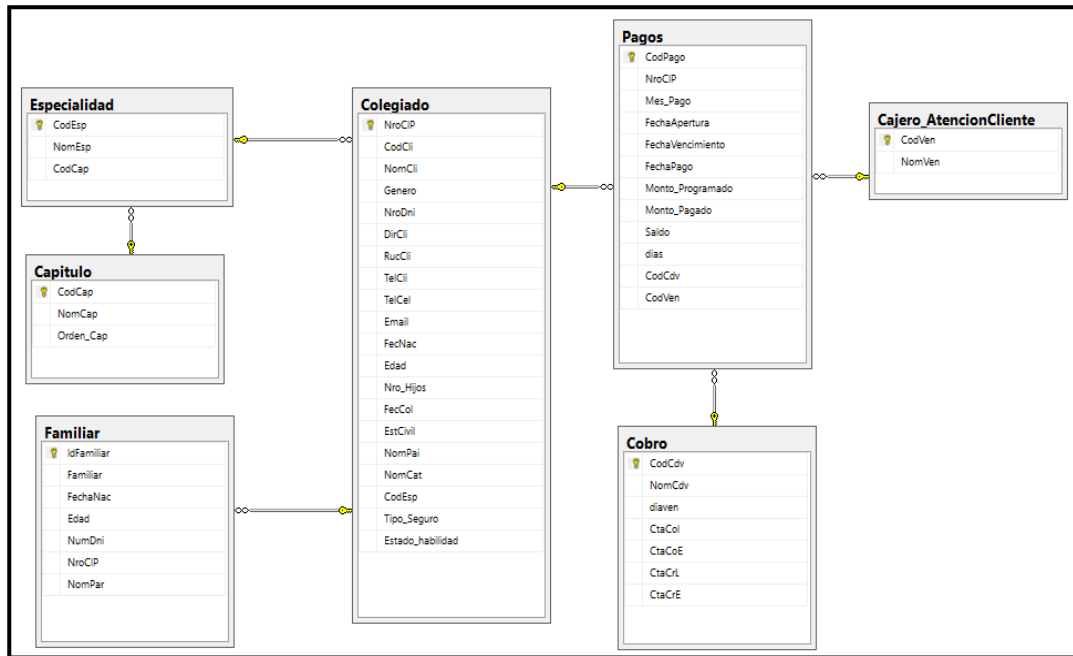


Figura 18. Modelo Entidad – Relación Base de datos CIP - CD Lambayeque.
Fuente: Equipo de Investigación

A continuación, en la Tabla 9 se describe cada tabla del modelo Entidad-Relación.



Tabla 9
 Descripción de tablas del modelo Entidad – Relación.

Tabla	Descripción
Especialidad	Se registran los nombres de las Especialidades de cada Capítulo.
Capitulo	Se registran los nombres los Capítulos del CIP.
Familiar	Se registran los familiares del Asociado al CIP.
Cobro	Tipo de cobro por la cual se realizarán los pagos.
Pagos	Monto a pagar del Asociado al CIP.
Cajero_AtencionCliente	Cajeros que atienden en Tesorería del CIP.
Colegiado	Asociado al CIP

Fuente: Equipo de Investigación.

El diccionario de datos de cada una de las tablas del modelo Entidad-Relación se detalla en el anexo 6.

2.8.3. Proceso de Extracción, transformación y carga (ETL) para preparación de datos de entrenamiento.

Según Chowdhury & Pal (2012) la preparación de la información para entrenamiento se lleva a cabo de acuerdo a los requerimientos de cada algoritmo. Por ello en el proceso de preparación de los datos se fijó como rango de evaluación a la información histórica de años recientes ya que el resto presenta datos muy irrelevantes que no favorecen la predicción. Por ello se consideró para la investigación la información histórica de los años 2010 al año 2017.



2.8.3.1. Requerimientos:

Los requerimientos están enfocados en mostrar información organizada de acuerdo al modelo de la investigación, por ello se optó por construir la tabla 10 con los requerimientos para la presente investigación.

Tabla 10
Descripción de cada requerimiento que corresponde a los indicadores de predicción.

Identificador:	R - 1	Nombre:	Monto Pagado por cada colegiado por mes
Tipo:	Funcionales		
Prioridad:	Alta	Necesidad:	Si
Descripción:	Permitirá Conocer el monto pagado (cuota mensual) que realizó cada colegiado.		
Identificador:	R -2	Nombre:	Cantidad de colegiados por Especialidad.
Tipo:	Funcional		
Prioridad:	Media	Necesidad:	Si
Descripción:	Nos permitirá conocer todos los colegiados que pertenecen a cada especialidad.		
Identificador:	R - 3	Nombre:	Cantidad de colegiados por Capítulo.
Tipo:	Funcional		
Prioridad:	Media	Necesidad:	Si
Descripción:	Nos permitirá conocer todos los colegiados que pertenecen a cada Capítulo.		
Identificador:	R - 4	Nombre:	Meses de Mayor Recaudación
Tipo:	Funcionales		
Prioridad:	Alta	Necesidad:	Si
Descripción:	Meses en los que se obtuvo el mayor ingreso de recaudación por el cobro de cuotas sociales de colegiados.		
Identificador:	R - 5	Nombre:	Meses de Menor Recaudación
Tipo:	Funcional		
Prioridad:	Alta	Necesidad:	Si
Descripción:	Meses en los que se obtuvo el menor ingreso de recaudación por el cobro de cuotas sociales de colegiados.		
Identificador:	R - 6	Nombre:	Cantidad de Colegiados Hábiles.
Tipo:	Funcionales		



Prioridad:	Alta	Necesidad:	Si
Descripción:	Ayudará a conocer la cantidad de Colegiados que están Hábiles.		
Identificador:	R - 7	Nombre:	Cantidad de Colegiados No Hábiles.
Tipo:	Funcionales		
Prioridad:	Alta	Necesidad:	Si
Descripción:	Ayudará a conocer la cantidad de Colegiados que no están Hábiles.		
Identificador:	R - 8	Nombre:	Pagos de Colegiados por Edad.
Tipo:	Funcionales		
Prioridad:	Alta	Necesidad:	Si
Descripción:	Permite conocer el comportamiento de pago de los colegiados por edad.		
Identificador:	R - 9	Nombre:	Cantidad de Colegiados por estado civil.
Tipo:	Funcionales		
Prioridad:	Alta	Necesidad:	Si
Descripción:	Permite conocer la cantidad de colegiados pertenecientes a cada tipo de estado civil.		
Identificador:	R - 10	Nombre:	Cantidad de Colegiados por Tipo de Seguro
Tipo:	Funcionales		
Prioridad:	Alta	Necesidad:	Si
Descripción:	Permite conocer la cantidad de colegiados por el tipo de Seguro.		
Identificador:	R - 11	Nombre:	Cantidad de Colegiados por Categoría.
Tipo:	Funcionales		
Prioridad:	Baja	Necesidad:	Si
Descripción:	Permite conocer la cantidad de colegiados pertenecientes a cada Categoría (Ordinario, Vitalicio, Fallecidos, Transferidos, Transeúntes).		
Identificador:	R - 12	Nombre:	Colegiados por Género.
Tipo:	Funcionales		
Prioridad:	Baja	Necesidad:	Si
Descripción:	Ayuda a conocer Cantidad de Colegiados por Género (Masculino y Femenino).		
Identificador:	R - 13	Nombre:	Cantidad de Hijos por cada Colegiados.
Tipo:	Funcionales		
Prioridad:	Media	Necesidad:	Si
Descripción:	Ayudará a conocer el número de hijos que tiene cada colegiado.		
Identificador:	R - 14	Nombre:	Cantidad de días de atraso de pago de la cuota mensual.
Tipo:	Funcionales		

Prioridad:	Alta	Necesidad:	Si
Descripción:	Permitirá conocer los días que se atrasó en pagar la cuota mensual los colegiados.		
Identificador:	R - 15	Nombre:	Pagos de Colegiados por Especialidad.
Tipo:	Funcionales		
Prioridad:	Alta	Necesidad:	Si
Descripción:	Permite conocer el comportamiento de pago de los colegiados por Especialidad.		
Identificador:	R - 16	Nombre:	Pagos de Colegiados por Capítulo.
Tipo:	Funcionales		
Prioridad:	Alta	Necesidad:	Si
Descripción:	Permite conocer el comportamiento de pago de los colegiados por Capítulo.		
Identificador:	R - 17	Nombre:	Pagos de Colegiados por Tipo de Estado Civil.
Tipo:	Funcionales		
Prioridad:	Alta	Necesidad:	Si
Descripción:	Permite conocer el comportamiento de pago de los colegiados por Tipo de Estado Civil.		
Identificador:	R - 18	Nombre:	Pagos de Colegiados según su número de Hijos.
Tipo:	Funcionales		
Prioridad:	Media	Necesidad:	Si
Descripción:	Permite conocer el comportamiento de pago de los colegiados según su número de hijos.		

Fuente: Equipo de Investigación.



2.8.3.2. Metodología empleada en la obtención de datos para la predicción de morosidad en cuotas sociales.

Según Chowdhury & Pal (2012) la metodología a seguir debe proporcionar herramientas que incluyan todos los requerimientos identificados. Por ello en la presente investigación de acuerdo a los modelos de predicción de morosidad se comparó los modelos dimensionales Estrella y Copo de Nieve mediante la Escala de Likert, que viene a ser un instrumento cuya escala numérica ordenada va del 1 al 5 que corresponde a una valoración determinada como se aprecia en la tabla 11. Al comparar los modelos dimensionales por sus parámetros se determinó que el modelo que se adecua al ámbito de la morosidad es el modelo Estrella ver tabla 12.

Tabla 11
Condición y valoración en la escala de Likert

Condición	Valoración
Muy bueno	(5)
Bueno	(4)
Regular	(3)
Malo	(2)
Muy malo	(1)

Fuente: Chowdhury & Pal (2012)



Tabla 12
Comparación del Modelo Estrella y Modelo Copo de Nieve

N°	Parámetros	Estrella	Copo de Nieve	Estrella	Copo de Nieve
1	Facilidad de uso	Consultas menos complejas, por lo tanto, fáciles de entender.	Consultas más complejas, por lo tanto, menos fáciles de entender.	5	2
2	Rendimiento de consulta o tiempo de ejecución de consulta	Menor número de claves externas, por lo tanto, menos complejidad y menor tiempo de ejecución de consultas	Más claves foráneas, por lo tanto, más complejidad y más tiempo de ejecución de consultas	4	2
3	Normalización	Contiene tablas no normalizadas.	Contiene tablas normalizadas.	4	3
4	Número de uniones (joins)	Menor número de joins (uniones)	Mayor número de joins (uniones)	4	3
5	Tablas de dimensiones	El esquema contiene solo una tabla de dimensiones solitarias para cada dimensión	El esquema puede tener más de una tabla de dimensiones para cada dimensión	4	3
6	Mantenimiento	El esquema contiene datos redundantes y, por lo tanto, no es tan fácil de mantener	No hay datos redundantes, y dado que las tablas están normalizadas, por lo tanto, es más fácil de mantener y actualizar	2	4
7	Tipo de Data Warehouse	Data Mart	Data Warehouse	4	4
PUNTAJE				27	21

Fuente: Equipo de investigación

2.8.3.3. Poblamiento del Modelo Entidad Relación Base de Datos CIP CD Lambayeque

El poblamiento de las tablas que forman el modelo entidad relación de la Base de datos CIP CD Lambayeque se pobló de acuerdo a las siguientes consultas.

```
merge [CIP].[dbo].[Capitulo] as cap2
using (select * from [Colegio_Ingenieros].[dbo].[tbl_capitulo]) as
cap1
on cap1.CodCap collate Modern_Spanish_CI_AS = cap2.CodCap collate
Modern_Spanish_CI_AS
when not matched then
insert (CodCap, NomCap, Orden_Cap)
values (CodCap, NomCap, orden_cap) ;
```

Figura 19. Consulta para el poblamiento de la Tabla Capítulo
Fuente: Equipo de Investigación

```
merge [CIP].[dbo].[Especialidad] as esp2
using (select * from [Colegio_Ingenieros].[dbo].[tbl_Especialidad])
as esp1
on esp1.CodEsp collate Modern_Spanish_CI_AS = esp2.CodEsp collate
Modern_Spanish_CI_AS
when not matched then
insert (CodEsp, NomEsp, CodCap)
values (CodEsp, NomEsp, CodCap) ;
```

Figura 20. Consulta para el poblamiento de la Tabla Especialidad
Fuente: Equipo de Investigación

```
merge [CIP].[dbo].[Cobro] as cob2
using (select * from [Colegio_Ingenieros].[dbo].[tbl01cdv]) as cob1
on cob1.codcdv collate Modern_Spanish_CI_AS = cob2.CodCdv collate
Modern_Spanish_CI_AS
when not matched then
insert (CodCdv, NomCdv, diaven, CtaCol, CtaCoE, CtaCrL, CtaCrE)
values (codcdv, nomcdv, diaven, CtaCoL, CtaCoE, CtaCrL, CtaCrE) ;
```

Figura 21. Consulta para el poblamiento de la Tabla Cobro
Fuente: Equipo de Investigación



```
merge [CIP].[dbo].[Cajero_AtencionCliente] as caj2
using (select * from [Colegio_Ingenieros].[dbo].[tbl01ven]) as caj1
on caj1.codven collate Modern_Spanish_CI_AS = caj2.CodVen collate
Modern_Spanish_CI_AS
when not matched then
insert (CodVen,NomVen)
values (codven,nomven);
```

Figura 22. Consulta para el poblamiento de la Tabla Cajero_AtencionCliente
Fuente: Equipo de Investigación

```
merge [CIP].[dbo].[Colegiado] as col2
using(select distinct
(c.nroCIP),c.codcli,c.nomcli,c.nrodni,c.ruccli,c.dircli,c.telcli,c.T
elCel,c.email,c.fecnac,c.feccol,
c.estciv,p.nompai,cat.nomcat,e.CodEsp,h.estado,
DATEDIFF(YEAR,c.FecNac,GETDATE()) as Edad from
Colegio_Ingenieros.dbo.mst01cli as c
inner join Colegio_Ingenieros.dbo.tbl01pai as p on c.codpai =
p.codpai
inner join Colegio_Ingenieros.dbo.tbl01cac as cat on c.codcat =
cat.codcat
left join Colegio_Ingenieros.dbo.tbl_especialidad as e on c.codesp =
e.CodEsp
left join Colegio_Ingenieros.dbo.habil2204201 as h on c.nroCIP =
h.cip
where c.nroCIP not like '' and c.nroCIP <> '72288' and c.nroCIP <>
'94001' and c.nroCIP != '.' and c.NroCIP != 'G-001') as coll
on coll.nroCIP collate Modern_Spanish_CI_AS = col2.NroCIP collate
Modern_Spanish_CI_AS
when not matched then
insert (NroCIP,CodCli,NomCli,NroDni,DirCli,RucCli,TelCli,TelCel,Email
,FecNac,Edad,FecCol,EstCivil,NomPai,NomCat,CodEsp,Estado_Habilidad)
values (nroCIP,codcli,nomcli,nrodni,dircli,ruccli,telcli,telcel,email
,fecnac,Edad,feccol,estciv,nompai,nomcat,codesp,estado);
```

Figura 23. Consulta para el poblamiento de la Tabla Colegiado
Fuente: Equipo de Investigación



```
merge [CIP].[dbo].[Familiar] as fam2
using(SELECT a.Codn, a.NomCon,a.FecNac,
DATEDIFF(YEAR,a.FecNac,GETDATE()) as Edad,
a.NroDni, c.nroCIP, b.NomPar
FROM Colegio_Ingenieros.dbo.Dtl01Con AS a
INNER JOIN Colegio_Ingenieros.dbo.mst01cli AS c ON a.Codn = c.codcli
INNER JOIN Colegio_Ingenieros.dbo.Tbl_Parentesco AS b ON a.CodPar =
b.CodPar
right JOIN CIP.dbo.Colegiado as cc on c.nroCIP collate
Modern_Spanish_CI_AS = cc.NroCIP collate Modern_Spanish_CI_AS
where cc.NroCIP collate Modern_Spanish_CI_AS = c.nroCIP collate
Modern_Spanish_CI_AS
)as fam1
on fam1.NomPar collate Modern_Spanish_CI_AS = fam2.NomPar collate
Modern_Spanish_CI_AS
when not matched then
insert (Familiar, FechaNac, Edad, NumDni, NroCIP, NomPar)
values (NomCon, FecNac, Edad, NroDni, nroCIP, NomPar);
```

Figura 24. Consulta para el poblamiento de la Tabla Familiar
Fuente: Equipo de Investigación

```
merge [CIP].[dbo].[Pagos] as pag2
using(select
v.NroCIP,v.mes,v.fecha,v.fven,v.uabo,v.monto,v.dias,v.codcdv,v.codve
n from V_PagosM as v
right join colegiado as c on v.NroCIP collate Modern_Spanish_CI_AS=
c.NroCip collate Modern_Spanish_CI_AS
where c.NroCip collate Modern_Spanish_CI_AS = v.nroCip collate
Modern_Spanish_CI_AS
)as pag1
on pag1.NroCIP collate Modern_Spanish_CI_AS = pag2.NroCIP collate
Modern_Spanish_CI_AS
when not matched then
insert
(NroCIP,Mes_Pago, FechaApertura, FechaVencimiento, FechaPago, Monto, dias
, CodCdv, CodVen)
values ( NroCIP,mes, fecha, fven, uabo, monto, dias, codcdv, codven);
```

Figura 25. Consulta para el poblamiento de la Tabla Pagos
Fuente: Equipo de Investigación



2.8.3.4. Diseño del Modelo Dimensional

El diseño del modelo dimensional se elaboró luego de poblar las tablas del modelo entidad relación de la Base de Datos CIP CD Lambayeque, así como la elección de las variables de predicción que se aprecia en la tabla 13.

Tabla 13
Variables identificadas para la predicción

Variables de predicción
Pago – Fecha de pago
Pagos – Fecha_vencimiento
Colegiado – NomCli
Colegiado – Edad
Colegiado – EstCivil
Colegiado – Tipo_Seguro
Colegiado – NomCat
Colegiado – Género
Especialidad – NomCap
Colegiado – Nro_Hijos

Fuente: Equipo de investigación

Las dimensiones se eligieron de acuerdo a las variables de predicción cuyas características guardan afinidad como se aprecia en la tabla 14.



Tabla 14
Elección de dimensiones de acuerdo a las variables

	Pago - Fecha de pago
Cuota_Mensual	Pagos - Fecha_vencimiento
	Año
Tiempo	Mes
	Día
	Colegiado - NomCli
	Colegiado - Estado_habilidad
Dimensión	Colegiado - Edad
	Colegiado - EstCivil
	Colegiado - Tipo_Seguro
Colegiado	Colegiado - NomCat
	Colegiado - Género
	Especialidad - NomEsp
	Capítulo - NomCap
	Colegiado - Nro_Hijos
Medidas	Monto
	Días Retrasados de Pago

Fuente: Equipo de Investigación.

Luego de la elección de dimensiones se procedió a construir el modelo Entidad Relación del modelo dimensional como se muestra en la figura 26.



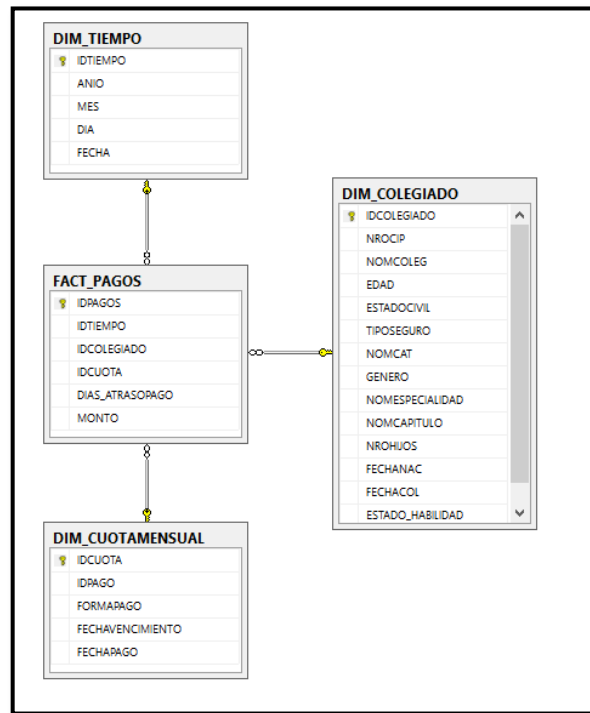


Figura 26. Modelo Entidad Relación con dimensiones seleccionadas

Fuente: Equipo de Investigación.

La descripción de atributos que forman cada dimensión se puede ver en detalle en el anexo 7.

Se pobló las dimensiones seleccionadas con consultas SQL, que se pueden apreciar desde la figura 27 a la figura 29.

```
SELECT DISTINCT FechaApertura, DATEPART(yyyy, FechaApertura) AS
Anio, DATEPART(mm, FechaApertura) AS Mes, DATEPART(dd,
FechaApertura) AS dias
FROM Pagos
WHERE (DATEPART(yyyy, FechaApertura) NOT LIKE '7') AND
(DATEPART(yyyy, FechaApertura) NOT LIKE '15') AND (DATEPART(yyyy,
FechaApertura) NOT LIKE '1195')
ORDER BY FechaApertura
```

Figura 27. Dimensión Tiempo
Fuente: Equipo de Investigación.



```

SELECT      c.NroCIP, c.NomCli, c.Edad, c.EstCivil,
c.Tipo_Seguro, c.NomCat, c.Genero, e.NomEsp, ca.NomCap,
c.Nro_Hijos, CONVERT(date, c.FecNac) AS Fecnac, CONVERT(date,
c.FecCol) AS FecCol, c.Estado_habilidad

FROM          Colegiado AS c LEFT OUTER JOIN

                Especialidad AS e ON c.CodEsp = e.CodEsp

LEFT OUTER JOIN

                Capitulo AS ca ON e.CodCap = ca.CodCap

LEFT OUTER JOIN

                ColegiadosActuales.dbo.ColegiadosActuales
AS caa ON c.NroCIP = caa.[N° CIP]

WHERE        (caa.[N° CIP] = c.NroCIP) AND (c.Edad > 20)
    
```

Figura 28. Dimensión Colegiado
Fuente: Equipo de Investigación.

```

SELECT      p.CodPago, c.NomCdv, p.FechaVencimiento, p.FechaPago

FROM          Pagos AS p INNER JOIN

                Cobro AS c ON p.CodCdv = c.CodCdv
    
```

Figura 29. Dimensión Cuota Mensual.
Fuente: Equipo de Investigación.

Con las consultas descritas se procesan la información a las dimensiones para la implementación de los algoritmos como se aprecia en la figura 30.

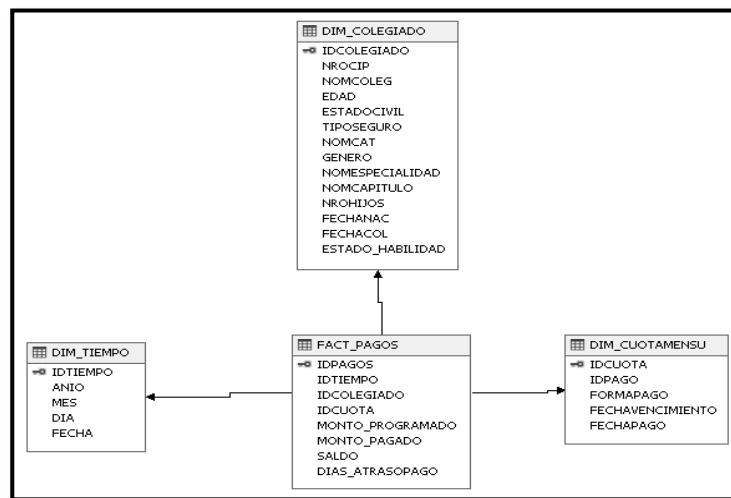


Figura 30. Modelo con dimensiones implementadas
Fuente: Equipo de Investigación.



2.8.3.5. Formato de archivo .CSV para la implementación de Técnicas:

Las técnicas requieren de archivos con información en formato .csv para su predicción, por ello para el presente trabajo se elaboró archivos en formato .csv ver figura 31.



Figura 31. Información en formato .CSV para implementación de técnicas de predicción
Fuente: Elaboración propia.

2.8.4. Implementación de la Máquina de Soporte Vectorial y Regresión Logística

2.8.4.1. Máquina de Soporte Vectorial

Se consideró los datos de la investigación y las variables predictivas para explicar la técnica de máquina de soporte vectorial. Para ello se tomó dos variables predictivas identificadas como x_1 y x_2 el cual se puede observar en la figura 32.



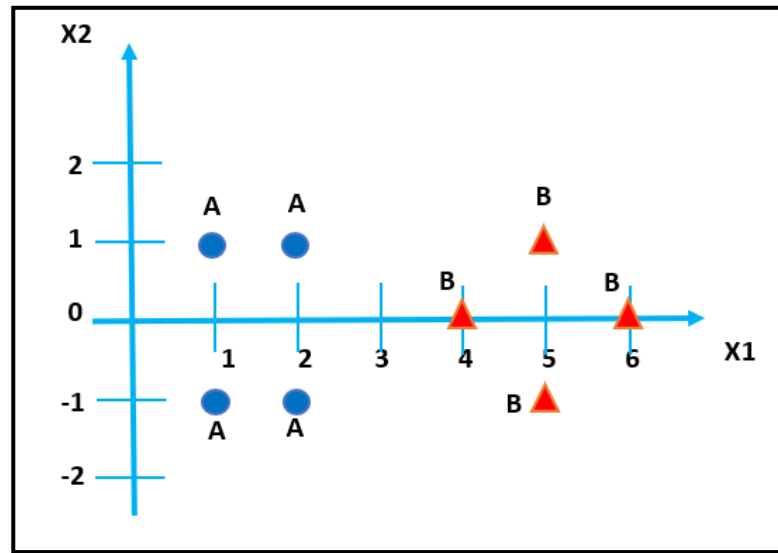


Figura 32. Plano de dos dimensiones con variables predictivas x_1 y x_2
 Fuente: Elaboración propia.

Las clases A y B que representan a los colegiados no morosos y morosos, de acuerdo a la técnica de máquina de soporte vectorial se debe encontrar un hiperplano que maximice la separación entre ambas. Por ello primero se trazan los posibles hiperplanos que separan a las clases como se aprecia en la figura 33.

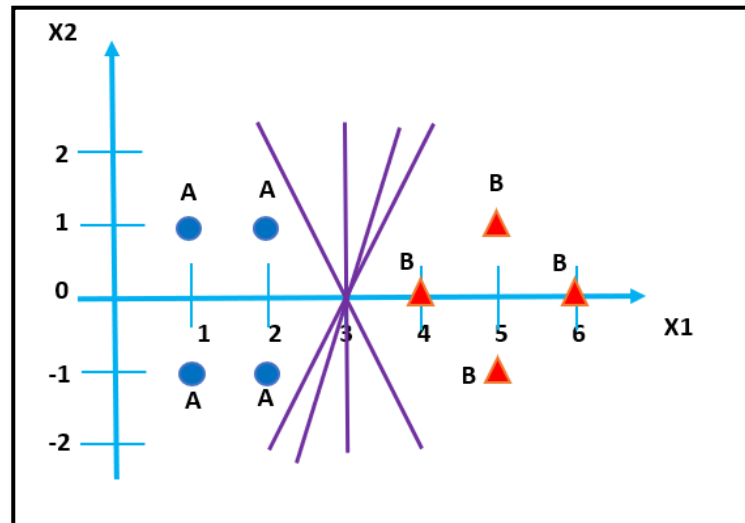


Figura 33. Posibles hiperplanos que separan las clases No Moroso y Moroso.
 Fuente: Elaboración propia.



Una vez trazado los hiperplanos se deben identificar los vectores de soporte que vienen a ser aquellas clases ubicadas en la frontera de ambas clases con margen máximo de separación representadas por S1, S2 y S3 que se aprecia en la figura 34.

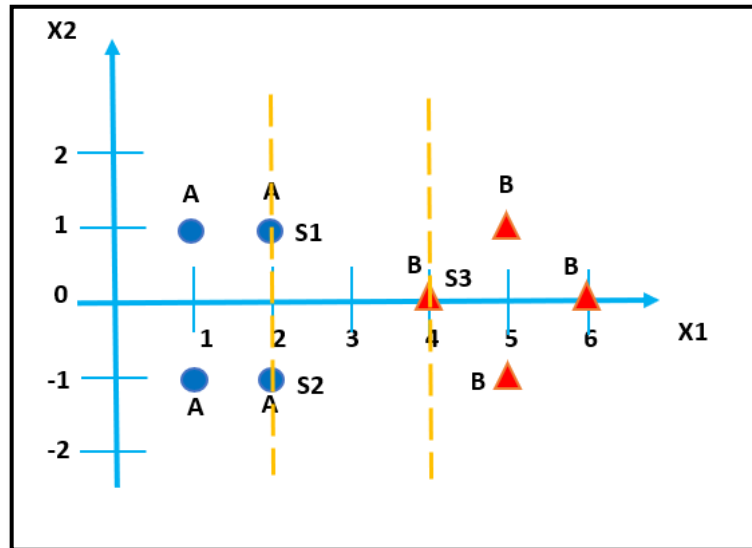


Figura 34. Vectores de soporte S1, S2 y S3
Fuente: Elaboración propia.

Para determinar el valor de cada uno de los vectores de soporte identificados se siguió los siguientes pasos:

Paso 1: Se representa cada vector de soporte por su coordenada.

$$S_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$S_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

$$S_3 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$$

Paso 2: Agregamos el valor uno para el manejo del sesgo de cada vector de soporte.

$$\hat{S}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$



$$\hat{S}_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$$

$$\hat{S}_3 = \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$$

Paso 3: Se expresa cada clase por las siguientes ecuaciones lineales.

$$\alpha_1 \hat{S}_1 \cdot \hat{S}_1 + \alpha_2 \hat{S}_2 \cdot \hat{S}_1 + \alpha_3 \hat{S}_3 \cdot \hat{S}_1 = -1 \quad (32)$$

$$\alpha_1 \hat{S}_1 \cdot \hat{S}_2 + \alpha_2 \hat{S}_2 \cdot \hat{S}_2 + \alpha_3 \hat{S}_3 \cdot \hat{S}_2 = -1 \quad (33)$$

$$\alpha_1 \hat{S}_1 \cdot \hat{S}_3 + \alpha_2 \hat{S}_2 \cdot \hat{S}_3 + \alpha_3 \hat{S}_3 \cdot \hat{S}_3 = +1 \quad (34)$$

Paso 4: Se procedió a sustituir en las ecuaciones 32, 33 y 34 los valores de \hat{S}_1 , \hat{S}_2 y \hat{S}_3 para determinar α_1 , α_2 y α_3 .

$$\alpha_1 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = -1 \quad (35)$$

$$\alpha_1 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = -1 \quad (36)$$

$$\alpha_1 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \cdot \hat{S}_3 + \alpha_2 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} = +1 \quad (37)$$

Paso 5: Se simplifica da ecuación quedando de la siguiente manera.

$$6\alpha_1 + 4\alpha_2 + 9\alpha_3 = -1 \quad (38)$$

$$4\alpha_1 + 6\alpha_2 + 9\alpha_3 = -1 \quad (39)$$

$$9\alpha_1 + 9\alpha_2 + 17\alpha_3 = +1 \quad (40)$$

Paso 6: Se resuelve el sistema de ecuaciones obteniendo los valores de α_1 , α_2 y α_3 que son los siguientes:

$$\alpha_1 = -3.25$$



$$\alpha_2 = -3.25$$

$$\alpha_3 = 3.5$$

Luego de determinar los vectores de soporte se procede a calcular el hiperplano óptimo que discrimine ambas clases de acuerdo a la ecuación siguiente.

$$\hat{W} = \sum_i \alpha_i \hat{S}_i \tag{41}$$

Como se determinó tres vectores de soporte la ecuación 41 queda expresada como sigue.

$$\hat{W} = \alpha_1 \cdot \hat{S}_1 + \alpha_2 \cdot \hat{S}_2 + \alpha_3 \cdot \hat{S}_3 \tag{42}$$

Para determinar el valor de \hat{W} se reemplazó en la ecuación 42 los valores de α_1 , α_2 y α_3 así como los valores de \hat{S}_1 , \hat{S}_2 y \hat{S}_3 .

$$\hat{W} = -3.25 \cdot \hat{S}_1 + -3.25 \cdot \hat{S}_2 + -3.25 \cdot \hat{S}_3 \tag{43}$$

$$\hat{W} = (-3.25) \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + (-3.25) \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} + (3.5) \cdot \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} \tag{44}$$

$$\hat{W} = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}$$

Por lo tanto, la ecuación del hiperplano óptimo de separación se expresa como $y = wx + b$, con $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y un desplazamiento de $b=-3$. Gráficamente la superficie de decisión con el hiperplano óptimo se aprecia en la figura 35.



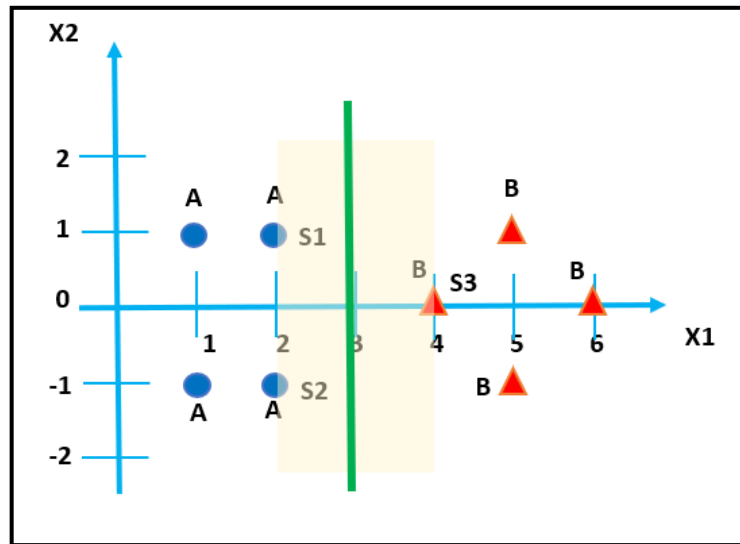


Figura 35. Hiperplano óptimo de separación de clases Moroso y No Moroso
Fuente: Elaboración propia.

De acuerdo al hiperplano óptimo las nuevas clases a predecir son de clase B por encontrarse a la derecha del hiperplano y de clase A por encontrarse a la izquierda del hiperplano. Como se aprecia en la figura 36.

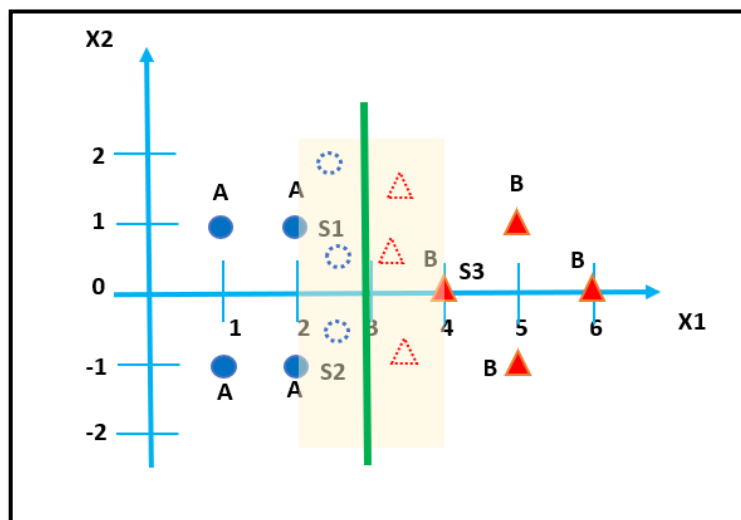


Figura 36. Predicción de clase A y clase B.
Fuente: Elaboración propia.



Se hace uso de la herramienta RStudio para la evaluación de la técnica máquina de soporte vectorial en la predicción de morosidad de cuotas sociales durante todos los meses del año 2017. Se puede apreciar en el anexo 8 el código completo implementado en RStudio.

Se dividió el archivo .csv que contiene 9600 registros en 80% para entrenamiento (7680 registros) y el 20% para prueba (1920 registros). Ver figura 37.

```

dataset          9600 obs. of 42 variables
test_set         1920 obs. of 42 variables
training_set     7680 obs. of 42 variables
Genero : Factor w/ 2 levels "F","M": 2 2 1 2 2 1 1 2 2 2 ...
Estado_civil : Factor w/ 3 levels "c","d","s": 3 2 3 3 2 3 2 3 1 3 ...
Edad : int 62 26 43 64 32 38 28 41 41 45 ...
Numero_de_Hijos: int 0 0 0 0 0 0 0 0 1 2 ...
Tipo_Seguro : Factor w/ 14 levels "EPS","ESSALUD",...: 13 5 13 13 2 2 2 14 5 2 ...
Especialidad : Factor w/ 35 levels "AGRICOLA","AGROINDUSTRIAL Y COMERCIO EXTERIOR",...: 4 6 31 30 33...
Categoria : Factor w/ 2 levels "ORDINARIO","VITALICIO": 2 1 1 2 1 1 1 1 1 1 ...
P.2000 : num 0 0 0 0 0 0 0 0 0 0 ...
P.2001 : num 0 0 0 0 0 0 0 0 0 0 ...
P.2002 : num 0 0 0 0 0 0 0 0 18 0 ...
P.2003 : num 0 0 0 0 0 0 0 0 0 0 ...
P.2004 : num 0 0 0 0 0 0 0 0 0 0 ...
P.2005 : num 0 0 0 0 0 0 0 0 0 0 ...
P.2006 : num 0 0 0 0 0 0 0 0 18 0 ...
P.2007 : num 0 0 0 0 0 0 0 0 18 0 ...
P.2008 : num 0 0 0 0 0 18 0 0 18 0 ...
P.2009 : num 0 0 0 0 0 18 0 0 0 0 ...
P.2010 : num 0 0 0 0 0 18 0 0 0 0 ...
P.2011 : num 0 0 0 0 0 0 18 18 18 ...
P.2012 : num 0 0 0 0 0 0 0 18 18 18 ...
P.2013 : num 0 0 0 0 0 0 0 18 18 ...
P.2014 : num 0 0 0 0 0 0 0 18 18 ...
P.2015 : num 0 0 0 0 0 0 0 18 18 ...
P.2016 : num 0 0 0 0 18 0 0 0 18 18 ...
T.2000 : Factor w/ 3 levels "NO PAGO","PAGO DESTIEMPO",...: 1 1 1 1 1 1 1 1 1 1 ...
    
```

Figura 37. División de datos 80 % para entrenamiento y 20 % para prueba
Fuente: Elaboración propia.

Se creó el modelo de la máquina de soporte vectorial con el conjunto de datos de entrenamiento cuyos parámetros se pueden ver en la figura 38.

```

Call:
svm(formula = MOROSO ~ ., data = training_set, type = "c-classification", kernel = "linear")

Parameters:
  SVM-Type: C-classification
 SVM-Kernel: linear
      cost: 1

Number of Support Vectors: 2720

( 1224 1496 )

Number of classes: 2

Levels:
NO SI
    
```

Figura 38. Parámetros del modelo de máquina de soporte vectorial
Fuente: Elaboración propia.



De acuerdo a la figura 38 los parámetros usados son kernel lineal y costo de 1, que al procesar los datos de entrenamiento se obtienen un total de 2720 vectores de soporte de los cuales 1224 pertenecen a la clase no moroso y 1496 a la clase moroso.

Se probó el modelo con el conjunto de datos de prueba (1920 registros) cuyos resultados se aprecian en la figura 39.

```
> pred_valid_svm
  2  3  8  14  29  33  34  38  44  45  49  55  64  68  71  79  80  87  97 101
NO NO NO NO SI SI NO NO NO NO NO NO SI SI SI NO SI NO NO
103 106 107 112 125 134 136 147 148 149 159 164 165 166 174 179 189 193 201 207
NO SI NO NO NO NO NO NO NO NO SI SI NO NO NO NO SI NO NO
208 215 217 220 227 229 231 236 237 240 244 247 287 293 294 302 306 308 309 321
SI NO NO SI NO NO SI NO SI NO NO SI SI SI NO SI NO SI NO
323 325 329 331 335 344 351 352 354 355 374 375 379 384 386 388 391 395 399 404
NO NO NO SI SI SI NO SI NO SI NO SI NO SI NO SI NO SI NO
410 416 417 426 428 433 435 436 444 461 465 467 468 477 478 487 489 490 491 497
NO NO NO SI NO NO NO NO NO NO SI NO NO SI SI SI SI SI
500 516 518 522 527 535 537 538 541 544 550 551 554 556 564 566 570 575 580 583
SI SI SI NO NO NO NO NO NO SI SI NO NO NO NO NO SI SI NO SI
SI SI SI NO NO NO NO NO NO NO SI NO NO NO NO NO SI SI NO SI
601 606 608 614 623 629 634 640 645 647 649 650 653 656 658 661 665 689 690 691
SI NO NO SI SI NO SI NO NO NO NO NO NO NO NO NO SI SI SI NO
720 724 731 736 737 739 743 745 761 762 768 769 770 771 773 776 779 789 805 809
SI SI NO NO NO NO NO SI NO SI NO NO NO NO NO SI NO SI SI SI
812 813 815 820 829 838 840 841 848 851 855 856 857 859 860 864 865 867 869 875
SI NO NO NO NO SI SI NO NO SI NO SI NO SI NO SI NO NO NO NO
877 886 889 890 893 895 897 898 909 916 917 918 922 925 928 929 931 937 945 948
NO NO NO NO NO SI SI NO SI NO SI SI NO NO NO NO NO SI SI NO
962 965 967 971 981 982 984 992 996 998 1001 1010 1019 1023 1031 1035 1039 1040 1041 1046
NO NO NO NO SI SI NO SI NO SI NO SI NO SI NO NO NO SI NO NO
1053 1054 1058 1073 1075 1076 1078 1079 1084 1098 1101 1110 1117 1119 1121 1123 1130 1132 1139 1150
SI NO SI SI NO NO SI NO SI NO SI NO SI NO SI SI SI SI NO SI SI
1151 1152 1155 1159 1160 1162 1165 1167 1173 1178 1179 1195 1202 1207 1211 1224 1225 1247 1259 1260
SI NO NO NO NO NO NO NO NO NO NO SI NO SI NO SI SI SI SI NO
1261 1267 1269 1271 1275 1282 1285 1287 1289 1293 1299 1300 1301 1302 1317 1322 1323 1326 1327 1346
NO SI NO NO NO NO NO NO SI SI SI SI SI NO SI NO SI NO SI NO SI
1352 1359 1364 1374 1375 1377 1378 1379 1381 1385 1388 1390 1391 1393 1398 1399 1402 1405 1407 1408
NO SI NO SI SI SI SI NO SI SI NO SI SI NO NO SI SI SI SI SI
1425 1438 1444 1445 1448 1458 1459 1460 1465 1470 1474 1487 1495 1503 1504 1511 1522 1523 1525 1533
SI NO SI SI NO NO SI SI SI SI SI NO NO NO NO SI NO NO NO NO
```

Figura 39. Resultados obtenidos con los datos de prueba
Fuente: Elaboración propia.

Se determinó la matriz de confusión y los indicadores de eficiencia cuyos resultados se aprecia en la figura 40.

```
Reference
Prediction NO SI
NO 1019 142
SI 119 640

Accuracy : 0.8641
95% CI : (0.8479, 0.8791)
No Information Rate : 0.5927
P-value [Acc > NIR] : <2e-16

Kappa : 0.7171

McNemar's Test P-Value : 0.1733

Sensitivity : 0.8432
Specificity : 0.8777
Pos Pred value : 0.8184
Neg Pred value : 0.8954
Prevalence : 0.4073
Detection Rate : 0.3333
Detection Prevalence : 0.3953
Balanced Accuracy : 0.8569

'Positive' Class : SI
```

Figura 40. Matriz de confusión y resultado de los indicadores de eficiencia
Fuente: Elaboración propia



Con los datos de prueba del mes de enero se obtuvo la matriz de confusión 2x2 figura 40 donde las filas horizontales son los valores reales y los verticales los predichos por el modelo donde los verdaderos positivos (SI Morosos) = 640, verdaderos negativos (No Morosos) = 1019, falsos positivos (No morosos predichos como morosos) =142 y falsos negativos (Morosos predichos como no morosos) =119. Tomando los datos de la matriz de confusión se determinan los indicadores de eficiencia cuyos valores son los siguientes Exactitud 86.41%, Precisión 81.84%, Sensibilidad 84.32% y especificidad 87.77%. Los resultados obtenidos por mes se registraron en fichas de observación que se puede ver a detalle en el anexo 9.

Se graficó la Curva ROC y se determinó el valor del Área bajo la curva (AUC) indicador que resume la eficiente de la técnica en estudio ver figura 41. La curva ROC calculado por cada mes se puede apreciar en el anexo 10.

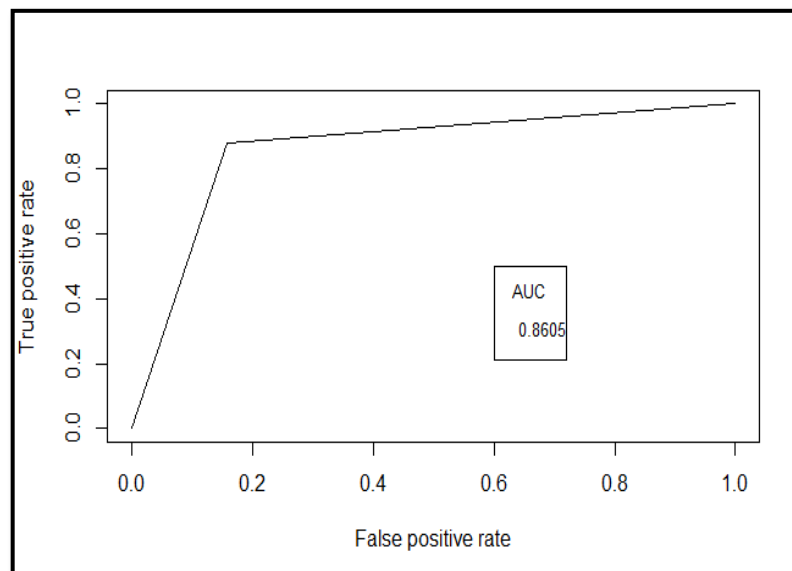


Figura 41. Gráfico de la Curva ROC y el valor del Área bajo la curva (AUC)
Fuente: Elaboración propia.

2.8.4.2. Regresión Logística

Se consideró los datos de la investigación y las variables predictivas para explicar la técnica de regresión logística. Primero se eligió dos variables predictivas x1 y x2 que se aprecia en la figura 42.



1	x1	x2	y
2	Género	Número_de_Hijos	MOROSO
3	M	0	SI
4	M	0	NO
5	M	0	NO
6	M	0	SI
7	F	0	SI
8	M	0	SI
9	M	0	NO
10	M	0	SI
11	F	0	SI
12	F	0	NO
13	M	0	SI
14	M	1	NO
15	M	2	NO
848	M	0	NO
849	M	0	SI
850	M	0	NO
851	M	0	SI
852	F	0	SI

Figura 42. Variables predictivas x1 y x2
Fuente: Elaboración propia.

Como se tienen dos variables elegidas la función logística a usar es la siguiente.

$$f(Y) = p(Y) = \frac{1}{1+e^{(\beta_0+\beta_1X_1+\dots+\beta_pX_p)}} \tag{45}$$

Donde: f(Y)= función logística, p(Y) probabilidad de ocurrencia o no ocurrencia de un evento, β_0 es la constante y los $\beta_1 \dots \beta_p$ son los coeficientes de los predictores $X_1 \dots X_p$ del modelo.

Se reemplazó las variables predictivas en la ecuación 45.

$$p(Y) = \frac{1}{1+e^{(\beta_0+\beta_1.género+\beta_2.número.de.hijos)}} \tag{46}$$

Para determinar la constante β_0 y los coeficientes β_1 y β_2 se emplea el método de máxima verosimilitud por medio de la ecuación 47.

$$l(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \tag{47}$$

Donde: $l(\beta)$ función de verosimilitud, $\prod_{i=1}^n$, productoria de $i = 1, \dots, n$ predictores, $\pi(x_i)^{y_i}$ probabilidad de ocurrencia de un evento, $[1 - \pi(x_i)]^{1-y_i}$ probabilidades de no ocurrencia de un evento.

Se aplica el logaritmo neperiano en la ecuación 47 de tal manera que la función de verosimilitud $L(\beta)$ se define como:



$$L(\beta) = \ln[l(\beta)] = \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\} \quad (48)$$

Para encontrar el valor de β se deriva $L(\beta)$ con respecto a $\beta_0, \beta_1, \dots, \beta_p$ y se iguala a cero, obteniéndose las siguientes ecuaciones:

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0 \quad (49)$$

$$\sum_{i=1}^n x_{ij} [y_i - \pi(x_i)] = 0, j = 1, \dots, p \quad (50)$$

Para dar solución a este conjunto de ecuaciones se debe recurrir al método iterativo de Newton-Raphson, cuyos resultados obtenidos son los siguientes.

$$\beta_0 = 0.3610$$

$$\beta_1 = -0.6221$$

$$\beta_2 = -0.8802$$

Se reemplazó los valores de la constante β_0 y los coeficientes β_1 y β_2 en la ecuación 46, resultando la ecuación 51.

$$p(Y) = \frac{1}{1 + e^{(0.3610 - 0.6221 \cdot \text{género} - 0.8802 \cdot \text{número_de_hijos})}} \quad (51)$$

En la técnica de regresión logística los valores de probabilidad de ocurrencia de un evento se reducen a los límites 0 y 1 ($0 \leq p(Y) \leq 1$), por ello aquel resultado superior o igual a 0,5 se toma como clase positiva (colegiado moroso), y resultados inferiores como clase negativa (colegiado no moroso).

Por ello si tomamos los valores para las variables predictivas género=0 y número de hijos =0 y se reemplazan en la ecuación 52 se obtiene como resultado que $p(Y) = 0.41$, el valor es inferior a 0,5 por ello es probable que él colegiado no sea moroso.

$$p(Y) = \frac{1}{1 + e^{(0.3610 - 0.6221 \cdot \text{género} - 0.8802 \cdot \text{número_de_hijos})}} \quad (52)$$

$$p(Y) = \frac{1}{1 + e^{(0.3610 - 0.6221 \cdot (0) - 0.8802 \cdot (0))}} \quad (53)$$

$$p(Y) = 0.41$$

$$p(Y) = 0.41 < 0,5 \text{ (Colegiado no Moroso)}$$



Si se toman nuevos valores de tal manera que género=1 y número de hijos=3 y se reemplaza en la ecuación 54 se obtiene como resultado para $p(Y) = 0.95$, el valor es superior a 0,5 por ello es más probable que el colegiado sea moroso.

$$p(Y) = \frac{1}{1+e^{(0.3610-0.6221.género-0.8802.número.de.hijos)}} \tag{54}$$

$$p(Y) = \frac{1}{1+e^{(0.3610-0.6221.(1)-0.8802.(3))}} \tag{55}$$

$$p(Y) = 0.95$$

$$p(Y) = 0.95 > 0,5 \text{ (Colegiado Moroso)}$$

Se hace uso de la herramienta RStudio para la evaluación de la técnica regresión logística en la predicción de morosidad de cuotas sociales durante todos los meses del año 2017. Se puede apreciar en el anexo 13 el código completo implementado en RStudio.

Se dividió el archivo .csv que contiene 9600 registros en 80% para entrenamiento (7680 registros) y el 20% para prueba (1920 registros). Ver figura 43.

```

dataset          9600 obs. of 42 variables
test_set         1920 obs. of 42 variables
training_set     7680 obs. of 42 variables
 Género : Factor w/ 2 levels "F","M": 2 2 1 2 2 1 1 2 2 2 ...
 Estado_civil : Factor w/ 3 levels "C","D","S": 3 2 3 3 2 3 2 3 1 3 ...
 Edad : int 62 26 43 64 32 38 28 41 41 45 ...
 Numero_de_Hijos: int 0 0 0 0 0 0 0 0 1 2 ...
 Tipo_Seguro : Factor w/ 14 levels "EPS","ESSALUD",...: 13 5 13 13 2 2 2 14 5 2 ...
 Especialidad : Factor w/ 35 levels "AGRICOLA","AGROINDUSTRIAL Y COMERCIO EXTERIOR",...: 4 6 31 30 33 7 16
 Categoría : Factor w/ 2 levels "ORDINARIO","VITALICIO": 2 1 1 2 1 1 1 1 1 1 ...
 P.2000 : num 0 0 0 0 0 0 0 0 0 0 ...
 P.2001 : num 0 0 0 0 0 0 0 0 0 0 ...
 P.2002 : num 0 0 0 0 0 0 0 0 18 0 ...
 P.2003 : num 0 0 0 0 0 0 0 0 0 0 ...
 P.2004 : num 0 0 0 0 0 0 0 0 0 0 ...
 P.2005 : num 0 0 0 0 0 0 0 0 0 0 ...
 P.2006 : num 0 0 0 0 0 0 0 0 18 0 ...
 P.2007 : num 0 0 0 0 0 0 0 0 18 0 ...
 P.2008 : num 0 0 0 0 0 18 0 0 18 0 ...
 P.2009 : num 0 0 0 0 0 18 0 0 0 0 ...
 P.2010 : num 0 0 0 0 0 18 0 0 0 0 ...
 P.2011 : num 0 0 0 0 0 0 0 18 18 18 ...
 P.2012 : num 0 0 0 0 0 0 0 18 18 18 ...
 P.2013 : num 0 0 0 0 0 0 0 0 18 18 ...
 P.2014 : num 0 0 0 0 0 0 0 0 18 18 ...
 P.2015 : num 0 0 0 0 0 0 0 0 18 18 ...
 P.2016 : num 0 0 0 0 18 0 0 0 18 18 ...
 T.2000 : Factor w/ 3 levels "NO PAGO","PAGO DESTIEMPO",...: 1 1 1 1 1 1 1 1 1 1 ...
    
```

Figura 43. División de datos 80 % para entrenamiento y 20 % para prueba
Fuente: Elaboración propia.

Se creó el modelo de la regresión logística con el conjunto de datos de entrenamiento cuyos parámetros se pueden ver en la figura 44 y figura 45.



```
Call:
glm(formula = MOROSO ~ ., family = binomial, data = training_set)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.5386  -0.4526  -0.2457   0.4148   3.2399

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.971111   0.452650   2.145 0.031922 *
GeneroM      -0.182251   0.100698  -1.810 0.070314 .
Estado_Civil -0.041629   0.256585  -0.162 0.871114 .
Estado_Civils 0.366658   0.249707   1.468 0.142008 .
Edad         0.030920   0.004634   6.673 2.51e-11 ***
Numero_de_Hijos -0.488037   0.054508  -8.953 < 2e-16 ***
Tipo_SeguroESSALUD -0.462544   0.296348  -1.561 0.118568 .
Tipo_SeguroMAPFRE EPS -0.882346   0.516909  -1.707 0.087828 .
Tipo_SeguroOANCOSALUD -0.402577   0.381663  -1.055 0.291518 .
Tipo_SeguroPACIFICO EPS -0.580093   0.348273  -1.666 0.095788 .
Tipo_SeguroPACIFICO SEGUROS -0.569016   0.887212  -0.641 0.521293 .
Tipo_SeguroPOSITIVA SEGUROS -0.448167   0.657145  -0.682 0.495244 .
Tipo_SeguroRIMAC EPS -0.568492   0.388237  -1.464 0.143115 .
Tipo_SeguroRIMAC SEGUROS 0.615451   0.628806   0.979 0.327697 .
Tipo_SeguroSANIDADES -0.431601   0.490659  -0.880 0.379057 .
Tipo_SeguroSANITAS EPS -1.568856   0.807088  -1.944 0.051914 .
Tipo_SeguroSANITAS PERU EPS -0.110610   0.778069  -0.142 0.886954 .
Tipo_SeguroSIN REGISTROS -0.209366   0.321122  -0.652 0.514414 .
Tipo_SeguroSIS -0.549103   0.317433  -1.730 0.083662 .
EspecialidadAGROINDUSTRIAL Y COMERCIO EXTERIOR -1.101611   0.453723  -2.428 0.015185 *
EspecialidadAGRONOMA 1.021951   0.209191   4.885 1.03e-06 ***
EspecialidadAGRONOMICA 0.449736   0.225530   1.994 0.046138 *
EspecialidadAGRONOMO 0.026540   0.222407   0.119 0.905014 .
EspecialidadAMBIENTAL 0.606769   0.488833   1.241 0.214509 .
EspecialidadCIVIL -0.701605   0.150544  -4.660 3.15e-06 ***
```

Figura 44. Parámetros del modelo de regresión logística
Fuente: Elaboración propia.

```
T.2005PAGO TIEMPO -0.556902 1.574234 -0.354 0.723519
T.2006PAGO DESTIEMPO 1.742916 1.600935 1.089 0.276292
T.2006PAGO TIEMPO 2.028481 1.703858 1.191 0.233841
T.2007PAGO DESTIEMPO -1.020446 1.642092 -0.621 0.534316
T.2007PAGO TIEMPO -0.587355 1.739289 -0.338 0.735591
T.2008PAGO DESTIEMPO 0.400757 1.911818 0.210 0.833964
T.2008PAGO TIEMPO 0.046653 1.983723 0.024 0.981237
T.2009PAGO DESTIEMPO 0.909179 1.866344 0.487 0.626156
T.2009PAGO TIEMPO 0.663012 1.929518 0.344 0.731136
T.2010PAGO DESTIEMPO -0.548437 1.505826 -0.364 0.715701
T.2010PAGO TIEMPO -0.509206 1.567796 -0.325 0.745339
T.2011PAGO DESTIEMPO -0.611282 0.456474 -1.339 0.180526
T.2011PAGO TIEMPO -1.015986 0.498110 -2.040 0.041382 *
T.2012PAGO DESTIEMPO 1.004276 0.902468 1.113 0.265790
T.2012PAGO TIEMPO 0.558739 0.913947 0.611 0.540970
T.2013PAGO DESTIEMPO -0.535233 0.914241 -0.585 0.558252
T.2013PAGO TIEMPO -0.655471 0.917388 -0.714 0.474920
T.2014PAGO DESTIEMPO -0.433376 0.871041 -0.498 0.618810
T.2014PAGO TIEMPO -1.040298 0.875454 -1.188 0.234717
T.2015PAGO DESTIEMPO -0.568784 0.886301 -0.642 0.521035
T.2015PAGO TIEMPO -0.823522 0.885098 -0.930 0.352148
T.2016PAGO DESTIEMPO -5.515277 0.660041 -8.356 < 2e-16 ***
T.2016PAGO TIEMPO -5.393546 0.655963 -8.222 < 2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 10306.7 on 7679 degrees of freedom
Residual deviance: 5081.5 on 7575 degrees of freedom
AIC: 5291.5

Number of Fisher Scoring iterations: 6
```

Figura 45. Parámetros del modelo de regresión logística
Fuente: Elaboración propia.

De acuerdo a la figura 44 y figura 45 al procesar los datos de entrenamiento se obtiene una constante $\beta_0 = 0.9711$ y el coeficiente de cada una de las variables predictivas, para la variable género $\beta_1 = -0.1823$, variable estado_civil $\beta_2 = -0.0416$ hasta la variable T.2016 (Tiempo de pago) cuya constante es $\beta_{41} = -5.3935$.



Se probó el modelo con el conjunto de datos de prueba cuyos resultados se aprecian en la figura 46.

```

> pred_valido
  2  3  8  14  29  33  34  38  44  45  49  55  64  68  71  79  80  87  97  101  103  106  107
NO NO NO NO SI SI NO NO NO NO NO NO SI SI SI NO SI NO NO NO NO SI NO
112 125 134 136 147 148 149 159 164 165 166 174 179 189 193 201 208 215 217 220 227 229
NO NO NO NO NO NO NO NO SI SI NO NO NO NO SI NO NO SI NO NO SI NO NO
231 236 237 240 244 247 287 293 294 302 306 308 309 321 323 325 329 331 335 344 351 352 354
SI NO SI NO NO NO SI SI SI NO SI NO SI NO NO NO SI SI SI SI NO SI NO
355 374 375 379 384 386 388 391 395 399 404 410 416 417 426 428 433 435 436 444 461 465 467
SI SI NO NO SI NO NO NO SI SI NO NO NO NO NO NO SI NO NO NO NO NO SI
468 477 478 487 489 490 491 497 500 516 518 522 527 535 537 538 541 544 550 551 554 556 564
NO NO SI SI SI SI SI SI SI SI SI NO NO NO NO NO SI SI NO NO NO NO NO
566 570 575 580 583 601 606 608 614 623 629 634 640 645 647 649 650 653 656 658 661 665 689
SI NO NO NO SI SI NO NO SI SI NO SI NO SI NO NO NO NO NO NO NO NO NO SI
690 691 720 724 731 736 737 739 743 745 761 762 768 769 770 771 773 776 779 789 805 809 812
SI NO SI SI NO NO NO NO NO SI NO SI NO NO NO NO SI NO SI SI SI NO SI
813 815 820 829 838 840 841 848 851 855 856 857 859 860 864 865 867 869 875 877 886 889 890
NO NO NO NO SI SI NO NO SI NO NO SI NO SI NO NO NO NO NO NO NO NO NO
893 895 897 898 909 916 917 918 922 925 928 929 931 937 945 948 962 965 967 971 981 982 984
NO NO SI SI NO SI SI NO NO NO NO NO SI SI NO NO NO NO NO NO SI SI NO
992 996 998 1001 1010 1019 1023 1031 1035 1039 1040 1041 1046 1053 1054 1058 1073 1075 1076 1078 1079 1084 1098
SI NO SI SI NO SI NO NO NO SI NO NO SI NO SI SI NO NO NO NO NO NO SI NO
1101 1110 1117 1119 1121 1123 1130 1132 1139 1150 1151 1152 1155 1159 1160 1162 1165 1167 1173 1178 1179 1195 1202
SI NO SI NO SI SI SI NO SI SI SI NO NO NO NO NO NO NO NO NO NO NO NO SI
1207 1211 1224 1225 1247 1259 1260 1261 1267 1269 1271 1275 1282 1285 1287 1289 1293 1299 1300 1301 1302 1317 1322
NO SI NO SI SI NO SI NO NO NO NO NO NO NO SI SI SI SI NO NO NO NO SI NO
1323 1326 1327 1346 1352 1359 1364 1374 1375 1377 1378 1379 1381 1385 1388 1390 1391 1393 1398 1399 1402 1405 1407
SI NO NO SI NO SI NO SI SI SI SI SI NO SI SI NO SI SI NO NO SI SI SI
1408 1425 1438 1444 1445 1448 1458 1459 1460 1465 1470 1474 1487 1495 1503 1504 1511 1522 1523 1525 1533 1535 1537
NO SI NO SI SI NO NO SI SI SI SI SI SI NO NO NO NO SI NO NO NO NO SI NO
1547 1553 1558 1559 1560 1566 1575 1577 1582 1583 1592 1593 1595 1600 1606 1612 1613 1616 1620 1621 1624 1632 1642
SI NO NO NO SI SI SI NO NO SI SI SI SI NO SI SI SI NO SI NO SI NO SI SI
1645 1646 1653 1663 1664 1667 1668 1673 1674 1675 1677 1680 1688 1698 1704 1718 1721 1722 1727 1729 1738 1752 1755
SI SI NO SI NO SI SI NO SI SI NO NO SI SI SI NO NO NO NO NO NO NO SI
1756 1758 1771 1776 1780 1783 1787 1788 1794 1797 1800 1805 1806 1808 1809 1812 1821 1824 1825 1830 1831 1842 1845
SI NO SI SI SI NO NO NO NO NO NO NO NO NO NO NO SI NO SI SI SI NO NO
    
```

Figura 46. Resultados obtenidos con los datos de prueba
Fuente: Elaboración propia.

Se determinó la matriz de confusión y los indicadores de eficiencia cuyos resultados se aprecia en la figura 47.

```

Reference
Prediction NO SI
NO 1030 131
SI 131 628

Accuracy : 0.8635
95% CI : (0.8474, 0.8786)
No Information Rate : 0.6047
P-Value [Acc > NIR] : <2e-16

Kappa : 0.7146

McNemar's Test P-Value : 1

Sensitivity : 0.8274
Specificity : 0.8872
Pos Pred Value : 0.8274
Neg Pred Value : 0.8872
Prevalence : 0.3953
Detection Rate : 0.3271
Detection Prevalence : 0.3953
Balanced Accuracy : 0.8573

'Positive' class : SI
    
```

Figura 47. Matriz de confusión y resultado de los indicadores de eficiencia
Fuente: Elaboración propia

Con los datos de prueba del mes de enero se obtuvo la matriz de confusión 2x2 figura 47 donde las filas horizontales son los valores reales y los verticales los predichos por el modelo donde los verdaderos positivos (SI Morosos) = 628, verdaderos negativos (No Morosos) =



1030, falsos positivos (No morosos predichos como morosos) =131 y falsos negativos (Morosos predichos como no morosos) =131. Tomando los datos de la matriz de confusión se determinan los indicadores de eficiencia cuyos valores son los siguientes Exactitud 86.35%, Precisión 82.74%, Sensibilidad 82.74% y especificidad 88.72%. Los resultados obtenidos por mes se registraron en fichas de observación que se puede ver a detalle en el anexo 14.

Se graficó la Curva ROC y se determinó el valor del Área bajo la curva (AUC) indicador que resume la eficiente de la técnica en estudio ver figura 48. La curva ROC calculado por cada mes se puede apreciar en el anexo 15.

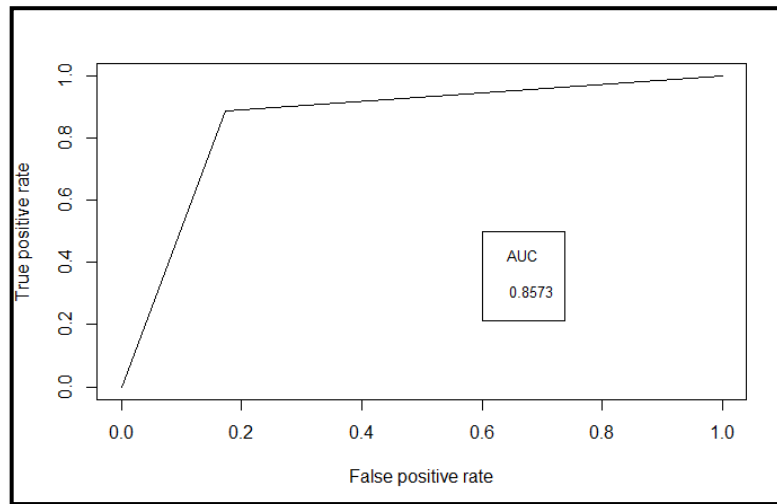


Figura 48. Gráfico de la Curva ROC y el valor del Área bajo la curva (AUC)
Fuente: Elaboración propia.

CAPÍTULO III: RESULTADOS

3.1. Resultados en Tablas y Figuras

Los resultados de la investigación se registraron en fichas de observación donde se describe la prueba realizada de cada técnica, así como los valores obtenidos para cada indicador Precisión, Exactitud, Sensibilidad, Especificidad y AUC. Con ello se estima que técnica es más eficiente para predecir la morosidad.

3.1.1. Máquina de Soporte Vectorial

Se registró cada uno de los indicadores de la técnica de máquina de soporte vectorial por cada mes de prueba efectuado, para el mes de enero se registraron los resultados que se muestra en la Tabla 15. Las fichas con los resultados por cada mes se detallan en el anexo 9.

Tabla 15
Evaluación de la técnica máquina de soporte vectorial

FICHA DE OBSERVACIÓN MÁQUINA DE SOPORTE VECTORIAL				
Computador: Computador con procesador Corei5.	Responsable: Jefferson Napoleón Huamán Bernilla. Fecha: 30 de octubre de 2018. Institución: Universidad Señor de Sipán.			
RESUMEN DE ANÁLISIS				
NÚMERO DE ANÁLISIS	1			
TIPO DE PRUEBA	Porcentaje Split (Training =80%, testing=20%)			
TIEMPO DE PREDICCIÓN (seg.)	0.71			
MES ANALIZADO	Enero			
MATRIZ DE CONFUSIÓN				
	NO	SI		
NO	1019	142		
SI	119	640		
EVALUACIÓN DE MÉTRICAS				
EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
86.41	81.84	84.32	87.77	0.8605

Fuente: Elaboración propia.



Cálculos efectuados para determinar cada métrica evaluada de acuerdo a la matriz de confusión.

Exactitud:

$$Exactitud = \frac{TP+TN}{TP+TN+FP+FN} \quad (56)$$

$$Exactitud = \frac{640+1019}{640+1019+142+119} \quad (57)$$

$$Exactitud = 0.8641 \approx 86.41\%$$

Precisión:

$$Precision = \frac{TP}{TP+FP} \quad (58)$$

$$Precision = \frac{640}{640+142} \quad (59)$$

$$Precision = 0.8184 \approx 81.84\%$$

Sensibilidad:

$$S = \frac{TP}{TP+FN} \quad (60)$$

$$S = \frac{640}{640+119} \quad (61)$$

$$S = 0.8432 \approx 84.32\%$$

Especificidad:

$$E = \frac{TN}{TN+FP} \quad (62)$$

$$E = \frac{1019}{1019+142} \quad (63)$$

$$E = 0.8777 \approx 87.77\%$$



Área bajo la curva (AUC):

$$AUC = \frac{1}{2}(S + E) \tag{64}$$

$$AUC = \frac{1}{2}(0.8432 + 0.8777) \tag{65}$$

$$AUC = 0.8777$$

Se obtuvo el promedio de los indicadores de eficiencia de la técnica máquina de soporte vectorial que se aprecia en la tabla 16, para ver a detalle los resultados de los indicadores por cada mes ver anexo 11.

Tabla 16
Indicadores de la técnica máquina de soporte vectorial

INDICADOR	PROMEDIO
EXACTITUD (%)	85.65
PRECISION (%)	82.47
SENSIBILIDAD (%)	81.51
ESPECIFICIDAD (%)	87.69
AUC	0.8497

Fuente: Elaboración propia

Los indicadores de eficiencia de la máquina de soporte vectorial que se muestran en la tabla 16, se interpretan de la siguiente manera: Con 85.65% acierta correctamente al predecir ambos casos (colegiado moroso, colegiado no moroso) del total de colegiados, con un 82.47% predice correctamente el caso positivo (colegiado moroso) del total que identifica como moroso porque en este conjunto hay falsos positivos(Colegiados no morosos identificados como morosos), con un 81.51% predice correctamente los colegiados morosos del total colegiados morosos porque en el conjunto que predice hay falsos negativos(Colegiados morosos identificados como no morosos), con un 87.69% predice aquellos colegiados que no son morosos del total de colegiados no morosos porque en el conjunto que predice hay falsos positivos(colegiados no morosos identificados como morosos). En cuanto al área bajo la curva (AUC) cuyo valor es 0.8497 indica su alta



capacidad predictiva al identificar ambos casos, ya que se encuentra por encima del rango aceptable, como afirma (Trujillo, 2017) debe encontrarse en el rango de 0.8 a 1.

3.1.2. Regresión Logística

Se registró cada uno de los indicadores de la técnica regresión logística por cada mes de prueba efectuado, para el mes de enero se registraron los resultados que se muestra en la Tabla 17. Las fichas con los resultados por cada mes se detallan en el anexo 14.

Tabla 17
Evaluación de la técnica regresión logística

FICHA DE OBSERVACIÓN REGRESIÓN LOGÍSTICA				
Computador: Computador con procesador Corei5.	Responsable: Jefferson Napoleón Huamán Bernilla.			
	Fecha: 30 de octubre de 2018.			
	Institución: Universidad Señor de Sipán.			
RESUMEN DE ANÁLISIS				
NÚMERO DE ANÁLISIS	1			
TIPO DE PRUEBA	Porcentaje Split (Training =80%, testing=20%)			
TIEMPO DE PREDICCIÓN (seg.)	0.82			
MES ANALIZADO	Enero			
MATRIZ DE CONFUSIÓN				
	NO	SI		
NO	1030	131		
SI	131	628		
EVALUACIÓN DE MÉTRICAS				
EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
86.35	82.74	82.74	88.72	0.8573

Fuente: Elaboración propia.



Cálculos efectuados para determinar cada métrica evaluada de acuerdo a la matriz de confusión.

Exactitud:

$$Exactitud = \frac{TP+TN}{TP+TN+FP+FN} \quad (66)$$

$$Exactitud = \frac{628+1030}{628+1030+131+131} \quad (67)$$

$$Exactitud = 0.8635 \approx 86.35\%$$

Precisión:

$$Precision = \frac{TP}{TP+FP} \quad (68)$$

$$Precision = \frac{628}{628+131} \quad (69)$$

$$Precision = 0.8274 \approx 82.74\%$$

Sensibilidad:

$$S = \frac{TP}{TP+FN} \quad (70)$$

$$S = \frac{628}{628+131} \quad (71)$$

$$S = 0.8274 \approx 82.74\%$$

Especificidad:

$$E = \frac{TN}{TN+FP} \quad (72)$$

$$E = \frac{1030}{1030+131} \quad (73)$$

$$E = 0.8872 \approx 88.72\%$$



Área bajo la curva (AUC):

$$AUC = \frac{1}{2}(S + E) \tag{74}$$

$$AUC = \frac{1}{2}(0.8274 + 0.8872) \tag{75}$$

AUC = 0.8573

Se obtuvo el promedio de los indicadores de eficiencia de la técnica regresión logística que se aprecia en la tabla 18, para ver a detalle los resultados de los indicadores por cada mes ver anexo 16.

Tabla 18
Indicadores de la técnica regresión logística

INDICADOR	PROMEDIO
EXACTITUD (%)	84.30
PRECISION (%)	81.38
SENSIBILIDAD (%)	79.43
ESPECIFICIDAD (%)	87.63
AUC	0.8353

Fuente: Elaboración propia

Los indicadores de eficiencia de la regresión logística que se muestran en la tabla 18, se interpretan de la siguiente manera: Con un 84.30% acierta correctamente al predecir ambos casos (colegiado moroso, colegiado no moroso) del total de colegiados, con un 81.38% predice correctamente el caso positivo (colegiado moroso) del total que identifica como moroso porque en este conjunto hay falsos positivos(Colegiados no morosos identificados como morosos), con un 79.43% predice correctamente los colegiados morosos del total colegiados morosos porque en el conjunto que predice hay falsos negativos(Colegiados morosos identificados como no morosos), con un 87.63% predice aquellos colegiados que no son morosos del total de colegiados no morosos porque en el conjunto que predice hay falsos positivos(colegiados no morosos identificados como morosos). En cuanto al área bajo la curva (AUC) cuyo valor es 0.8353 indica su alta capacidad predictiva al identificar ambos



casos, ya que se encuentra por encima del rango aceptable, como afirma (Trujillo, 2017) debe encontrarse en el rango de 0.8 a 1.

3.1.3. Tiempo para generar estimación

Cada técnica emplea un determinado tiempo para predecir en función a la información ingresada, así como el rendimiento de ejecución de la CPU para procesar la información, por ello para cada técnica se evaluó el tiempo de usuario (user time) que hace referencia al tiempo de ejecución efectuado por el usuario, tiempo del sistema (system time) tiempo que demora la CPU en la ejecución de las instrucciones para la generación de la predicción y tiempo transcurrido (elapsed time) tiempo total desde que inició el proceso.

El tiempo promedio para generar la predicción de la morosidad por ambas técnicas se aprecia en la Tabla 19. El tiempo de ejecución por prueba se puede ver a detalle en el anexo 12 y 17.

Tabla 19
Tiempo de ejecución promedio Máquina de Soporte Vectorial Vs. Regresión Logística

Técnica	User time (seg)	System time(seg)	Elapsed time(seg)
SVM	11.89	0.49	56.37
RL	4.95	0.52	45.69

Fuente: Elaboración propia.

La técnica máquina de soporte vectorial requiere de un tiempo de ejecución de 0.49 segundos para predecir el caso positivo (Colegiado moroso) y en cuanto a la técnica regresión logística requiere de un tiempo de 0.52 segundos. Por lo tanto, la técnica máquina de soporte vectorial a la hora de procesar la información para poder predecir requiere de un tiempo inferior en comparación con la técnica regresión logística.



3.2. Discusión de resultados

Con los datos de la tabla 16 y tabla 18, se comparó los indicadores de eficiencia de ambas técnicas como se muestra en la siguiente figura.

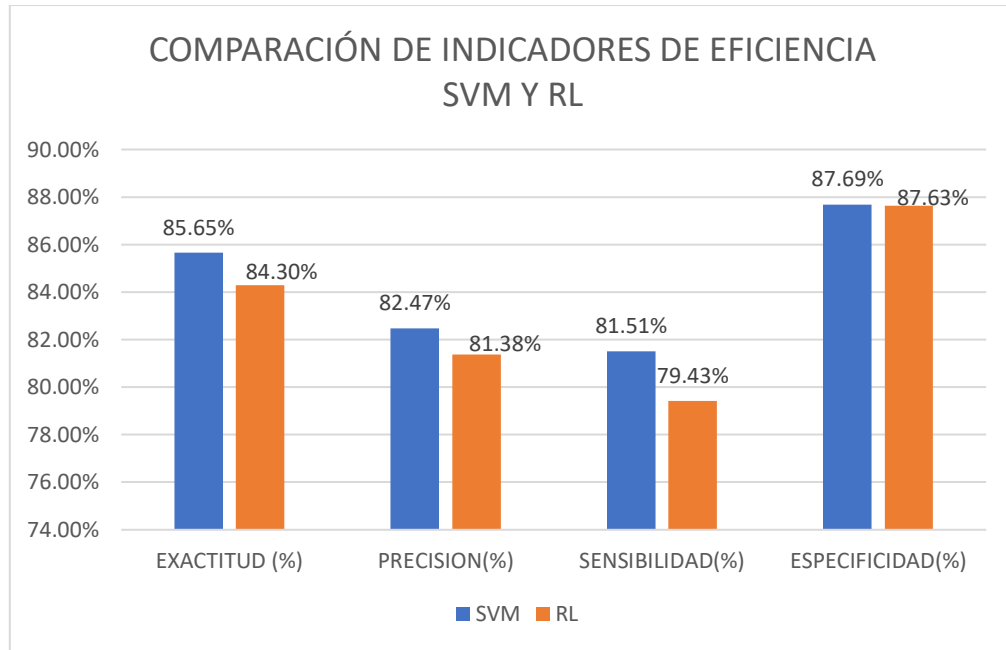


Figura 49. Comparación de Indicadores de Eficiencia entre la técnica Máquina de Soporte Vectorial y Regresión Logística
 Fuente: Elaboración propia.

Se aprecia en la figura 49 que la técnica máquina soporte vectorial en cada uno de sus indicadores de eficiencia es superior a la regresión logística, es decir predice mejor al caso positivo (colegiados morosos). Además, el indicador sensibilidad de la técnica máquina de soporte vectorial es muy superior al de la regresión logística lo que quiere decir que al usar la técnica máquina de soporte vectorial habrá muy pocos Falsos negativos (Colegiados morosos identificados como no morosos) y por el contrario con la regresión logística abran más falsos negativos (Colegiados morosos identificados como no morosos).



Con los datos de la tabla 16 y la tabla 18 se comparó las áreas bajo la curva (AUC) de ambas técnicas como se aprecia en la siguiente figura.

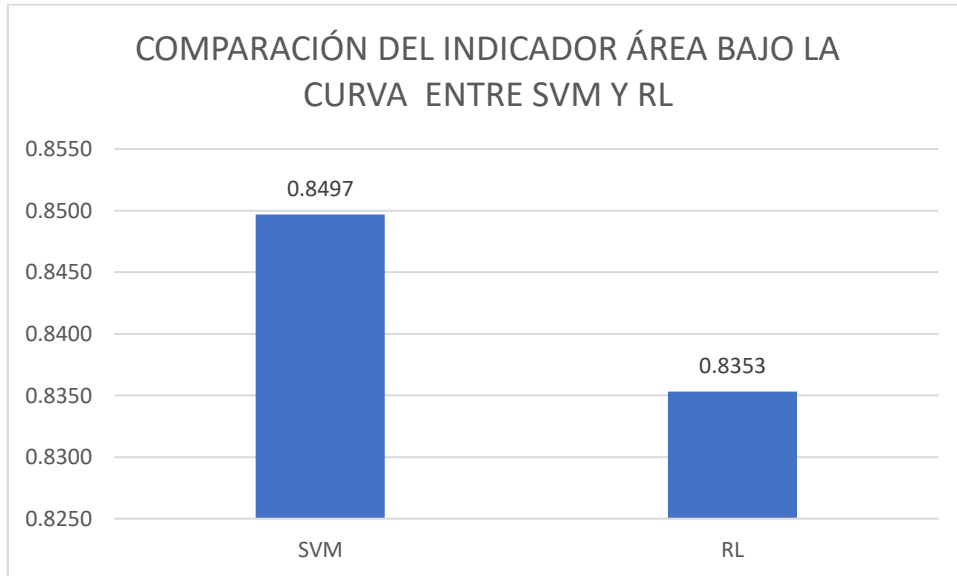


Figura 50. Comparación del indicador área bajo la Curva (AUC) de las técnicas de Máquina de Soporte Vectorial y Regresión Logística
Fuente: Elaboración propia.

En cuanto al indicador área bajo la curva (AUC), de acuerdo a (Trujillo, 2017) una técnica con un AUC entre 0.8 y 0.9 es excelente. Por ello en la figura 50 se aprecia que ambas técnicas son excelentes para predecir la morosidad, pero sobresale la técnica máquina de soporte vectorial.

De igual manera haciendo uso de los datos de la tabla 19, se comparó el tiempo requerido por ambas técnicas para efectuar la predicción, como se aprecia en la siguiente figura.



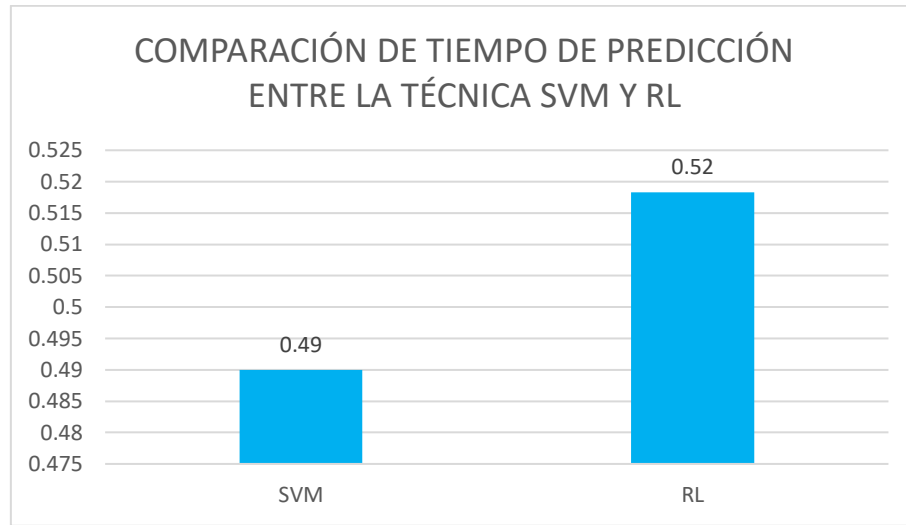


Figura 51. Comparación del tiempo requerido por la técnica de máquina de soporte vectorial y regresión logística en la predicción de morosidad de cuotas sociales
 Fuente: Elaboración propia.

De acuerdo a la figura 51 la técnica máquina de soporte vectorial requiere de menor tiempo para procesar la información y generar la predicción. Por ello, la técnica máquina de soporte vectorial frente a la técnica de regresión logística es más eficiente a la hora de predecir la morosidad de cuotas sociales.



CAPÍTULO IV: CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

- a) Se seleccionó las técnicas de predicción para ello se realizó un análisis exhaustivo de trabajos de investigación cuyo ámbito de estudio fueron los riesgos de crédito, por ser un escenario más cercano a la predicción de morosidad comparando técnicas computacionales y estadísticas por su precisión de predicción, de ello se seleccionó la técnica máquina de soporte vectorial y regresión logística por los buenos resultados alcanzados.
- b) Se procedió a realizar la recopilación de la información de los colegiados para la investigación, el cual consistió en recopilar información histórica de fuentes internas y externas a la organización que incluyan los aportes por concepto de pago de cuotas sociales de cada colegiado porque es donde se genera la morosidad.
- c) Se realizó el proceso de extracción, transformación y carga (ETL) efectuando la limpieza de los datos recopilados de acuerdo a las variables predictivas identificadas como son el Género, Estado Civil, Edad, Número de Hijos, Tipo de Seguro, Especialidad, Categoría, Monto de pago y Días de incumplimiento. Obteniendo los datos concisos y relevantes para la implementación de las técnicas de predicción.
- d) Se aplicó la técnica máquina de soporte vectorial y regresión logística para resolver el problema de la predicción de morosidad en cuotas sociales, utilizando datos conformados por un total de 9600 registros obtenidos del proceso ETL, ya que las técnicas requieren de gran cantidad de información para poder predecir. La técnica máquina de soporte vectorial obtuvo un mejor resultado en exactitud 85.65%, precisión 82.47%, sensibilidad 81.51%, especificidad 87.69% y un área bajo la curva (AUC) de 0.8497 en comparación con la regresión logística que obtuvo una exactitud 84.30%, precisión 81.38%, sensibilidad 81.38%, especificidad 87.63% y un área bajo la curva (AUC) de 0.8353.



- e) Se evaluó los resultados concluyendo que la técnica máquina de soporte vectorial es más eficiente para predecir la morosidad en cuotas sociales porque su precisión es más exacta que la regresión logística, además requiere de un tiempo inferior para predecir en comparación a la regresión logística.

4.2. Recomendaciones

- a) Se recomienda a los futuros investigadores implementar mejoras en la técnica máquina de soporte vectorial en la predicción de morosidad, incorporando variables predictivas como nivel de ingreso, tipo de vivienda, tipo de trabajo (dependiente o independiente), universidad de egreso, entre otras.
- b) Se recomienda a los futuros investigadores implementar nuevas técnicas como los de aprendizaje profundo para la predicción de morosidad, tomando en consideración el escenario del problema expresado en la presente investigación y que efectúen la comparación de resultados.
- c) Se recomienda a los futuros investigadores hacer uso de una técnica diferente para la determinación de los parámetros de la máquina de soporte vectorial, por ejemplo, validación cruzada. Así mismo que efectúen pruebas con kernels distintos y comparen sus resultados con los obtenidos en este trabajo de investigación.
- d) Se recomienda a los futuros investigadores implementar la técnica de máquina de soporte vectorial en un lenguaje diferente a R, por ejemplo, Python ya que su uso está en pleno crecimiento para la implementación de técnicas de aprendizaje automático.



REFERENCIAS

- Aguilar, G., Camargo, G., & Morales, R. (2006). Análisis de la morosidad en el sistema bancario peruano. *Economía y Sociedad*, 74-81.
- Alea, V. (2005). Análisis de datos. Universidad de Barcelona, España.
- Arrieta, E. (2017). Predicción Temprana de Morbilidad Materna Extrema Usando Aprendizaje Automático. Universidad Tecnológica de Bolívar, Cartagena.
- Awoyemi, J., Adetunmbi, A., & Oluwadare, S. (2017). Credit card fraud detection using Machine Learning Techniques: A comparative Analysis. *IEEE*, 1-9.
- Bedor, D. (2015). CONSTRUCCIÓN DE UN MODELO ESTADÍSTICO PARA LA GESTIÓN DE LA CARTERA DE COBRO DEL SERVICIO DE RENTAS INTERNAS ZONAL 8. INSTITUTO DE ALTOS ESTUDIOS NACIONALES, GUAYAQUIL.
- Birla, S., Kohli, K., & Dutta, A. (2016). Machine Learning on Imbalanced Data in Credit Risk. *IEEE*, 1-6.
- Carpio, J. (2016). MODELO DE PREDICCIÓN DE LA MOROSIDAD EN EL OTORGAMIENTO DE CRÉDITO FINANCIERO APLICANDO METODOLOGÍA CRISP-DM. UNIVERSIDAD ANDINA "NÉSTOR CÁCERES VELÁSQUEZ", JULIACA.
- Castro, H., & Pinedo, W. (2018). SISTEMA DE DIGITALIZACIÓN Y ESTRUCTURACIÓN DE INFORMÁTICA CLÍNICA CON TÉCNICAS DE RECONOCIMIENTO ÓPTICO DE CARACTERES Y PROCESAMIENTO DEL LENGUAJE NATURAL. UNIVERSIDAD DE LIMA, LIMA.
- Chamorro, J. (2016). APLICACIÓN WEB PARA LA ELABORACIÓN Y GESTIÓN DE PROCESOS ETL EN BIG DATA. UNIVERSIDAD CARLOS III DE MADRID, MADRID.
- Chowdhury, R., & Pal, B. (2012). A Data Warehouse Architectural Design Using Proposed Pseudo Mesh Schema. *ResearchGate*, 1-4.
- CIP-CD Lambayeque. (2017). Estado de situación financiera al 31 de diciembre del 2017. Lambayeque: Informática CIP.
- El Comercio. (26 de 01 de 2018). Banca: Morosidad es la más alta en 12 años, ¿bajará en 2018? Obtenido de <https://elcomercio.pe/economia/peru/banca-morosidad-alta-12-anos-bajara-2018-noticia-492242-noticia/>



- Espino, C. (2017). Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo - herramientas Open Source que permiten su uso. Universidad abierta de cataluña, Cataluña.
- Europa Press. (19 de 03 de 2018). Dos tercios de la morosidad de los créditos en 2017 fue de las empresas. Obtenido de <https://www.europapress.es/economia/macroeconomia-00338/noticia-dos-tercios-morosidad-creditos-2017-fue-empresas-20180319114015.html>
- Fernández, P. (2015). ADECUACIÓN DE LOS MODELOS DE PREDICCIÓN DE LAS CRISIS BANCARIAS AL CASO ESPAÑOL. UNIVERSIDAD DE SEVILLA, SEVILLA.
- Fernández, R. (2013). ESTIMACIÓN ESTADÍSTICA DE CLIENTES MOROSOS DE LIMA METROPOLITANA Y EL CALLAO QUE USAN TARJETAS DE CRÉDITO PARA LA BANCA DE PERSONAS NATURALES EN UNA ENTIDAD BANCARIA. UNIVERSIDAD NACIONAL DE INGENIERIA, LIMA.
- Gahlaut, A., Tushar, & Kumar, P. (2017). Prediction analysis of risky credit using Data minig classification models. IEEE, 2-7.
- Gironés, J., Casas, J., Minguillón, J., & Caihuelas, R. (2017). Minería de Datos Modelos y Algoritmos. Barcelona: UOC.
- Gonzales, J., & Torres, S. (2017). Diccionario de didáctica de Español. México: Lulu.
- Gordillo, N. (2007). Metodología, método y propuestas metodológicas en Trabajo Social. Tendencias y Retos, 119-135.
- Hernández, G. (2006). Diccionario de Economía. Colombia: Universidad Cooperativa de Colombia.
- Hernández, R., Fernández, C., & Baptista, M. (2010). Metodología de la Investigación. México: McGRAW-HILL.
- Kelleher, J., Mac, B., & Darcy, A. (2015). Fundamentals of Machine Learning for predictive Data Analytics Algorithms, Worked examples, and Case Studies. Estados Unidos: MIT Press.
- Lawi, A., Aziz, F., & Syarif, S. (2017). Ensemble GradientBoost for Increasing Classification Accuracy of Credit Scoring. 2-6.
- Lizares, M. (2017). Comparación de modelos de clasificación: regresión logística y árboles



- de clasificación para evaluar el rendimiento académico. Universidad Nacional Mayor de San Marcos, Lima.
- Maher , A., & Abbod, M. (2015). A systematic credit scoring model base don heterogeneous classifier ensembles. IEEE, 1-7.
- Marcos, E. (2015). PROPUESTA DE UN MODELO DE REGRESIÓN LOGÍSTICA PARA ANALIZAR EL RIESGO CREDITICIO EN LA EMPRESA CARSA S.A.C. UNIVERSIDAD NACIONAL AGRARIA LA MOLINA, LIMA.
- Marin, J., & Marulanda, V. (2016). MODELO DE APROBACIÓN DE CRÉDITO EN UNA ENTIDAD FINANCIERA. FUNDACION UNIVERSITARIA LOS LIBERTADORES, BOGOTA.
- Maura, G., Ramirez, A., & Shiguihara, P. (2018). Human Resources Management Model based on Business Intelligence. IEEE, 1-4.
- Medina, M., & Ulfe, H. (2015). MODELO DE CREDIT SCORING PARA PREDECIR EL OTORGAMIENTO DE CRÉDITO PERSONAL EN UNA COOPERATIVA DE AHORRO Y CRÉDITO. UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO, LAMBAYEQUE.
- Mohammed, M., Badruddin, M., & Mohammed, E. (2017). Machine Learning Algorithms and Applications. Nueva York: CRC Press.
- Pallarés, J. (2016). La metodología cuantitativa aplicada al estudio de la reincidencia en menores infractores. Universidad Jaume I, España.
- Real Academia Española. (18 de 10 de 2018). Diccionario de la lengua Española. Obtenido de <https://dle.rae.es/comparaci%C3%B3n>
- Reyes, A., & León, D. (2014). CAPACIDAD PREDICTIVA DE LOS MODELOS DE MÁQUINA DE VECTORES DE SOPORTE Y MODELO DE REGRESIÓN LOGÍSTICA EN EL ANÁLISIS DE RIESGO DE CRÉDITO - PERSONA. UNIVERSIDAD NACIONAL DE INGENIERÍA, LIMA.
- Rizzo, M. (2008). Statistical Computing with R. Nueva York: Chapman & Hall/CRC.
- Rosales, C. (2009). ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE UN DATAMART PARA EL SOPORTE DE TOMA DE DECISIONES Y EVALUACIÓN DE LAS ESTRATEGIAS SANITARIAS EN LAS DIRECCIONES DE SALUD. PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ, LIMA.



- Tesén, A. (2017). EFICACIA DE LOS MODELOS DE APRENDIZAJE DE MAQUINA PARA EVALUAR EL RIESGO CREDITICIO DE PERSONAS NATURALES EN UNA INSTITUCIÓN FINANCIERA DE CHICLAYO. UNIVERSIDAD NACIONAL DEL SANTA, CHIMBOTE.
- Trujillo, D. (2017). Aplicación de Metodologías Machine Learning en la Gestión de Riesgo de Crédito. Universidad Politécnica de Madrid, Madrid.
- Wei, Y., & Chiu, D. (2015). Machine Learning with R Cookbook. Reino Unido: Packt Publishig.
- Yeh, C., & Lien, C. (2009). The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. Expert Systems with Applications, 2473-2480.
- Zavaleta, J. (2017). Morosidad de cuentas por cobrar y afectación del cumplimiento de obligaciones corrientes. Universidad San Pedro, Chimbote.
- Zeng, J.-Y., Lin, J.-B., & Wang, T. (2017). A new competing risks model for predicting prepayment and default using data mining. International Conference on Image, Vision and Computing, 985-989.



ANEXOS

ANEXO 1. ARTÍCULOS CIENTÍFICOS REVISADOS PARA LA ELABORACIÓN DE TOP DE TÉCNICAS DE PREDICCIÓN.

N°	PAPER	¿QUÉ SE HIZO?	ALGORITMOS	RESULTADOS	CONCLUSION
1	Machine Learning Methods of Bankruptcy Prediction Using Accounting Ratios. (Li & Wang, 2018)	Realizaron una comparación de métodos estadísticos y métodos de aprendizaje automático, para predecir la bancarrota en empresas que cotizan en china. Se utilizó la base de datos Wind Financial Terminal y la base de datos económica financiera CCER.	STATISTICAL k-NEAREST NEIGHBOUR SUPPORT VECTOR MACHINES (SVM) LOGISTIC REGRESION DECISION TREE RANDOM FOREST	Precisión: 70.8% 80.5% 84.0% 79.0% 92.9% 95.9%	Los métodos empíricos muestran que los métodos de aprendizaje automático son superiores a los métodos estadísticos.
2	A PREDICTIVE SYSTEM FOR DETECTION OF BANKRUPT CY USING MACHINE LEARNING TECHNIQUE S. (Nagaraj & Sridhar, 2015)	Propusieron un sistema de predicción y para ello se evaluaron los algoritmos de aprendizaje automático y de estadística. Aplicados a un conjunto de datos cualitativo de UCI. El conjunto de datos consta de 250 instancias basadas en 6 atributos. La salida tenía dos clases de instancia como "Quiebra" y "No Quiebra". Se eliminó los atributos redundantes del conjunto de datos y para ello se utilizó la técnica de selección de características basadas en la correlación. El conjunto de datos se divide en entrenamiento (2/3 del conjunto de datos) y conjunto de pruebas (1/3 del conjunto de datos). Se aplicó validación cruzada de 10 veces para probar	LOGISTIC REGRESION RANDOM FOREST NAIVE BAYES NEURAL NETWORK SUPPORT VECTOR MACHINES (SVM)	Precisión: 97.2% 97.4% 98.3% 98.6% 99.6%	Los resultados sugieren que se pueden implementar técnicas de aprendizaje automático para la predicción de quiebra. El sistema predictivo ayuda a predecir la quiebra de un conjunto de datos del cliente, donde el modelo de máquina de soporte vectorial (SVM) alcanza mayor precisión.



la precisión. La precisión de cada algoritmo se estimó a partir de los resultados de validación cruzada.

3	Financial distress prediction: The case of French small and medium-sized firms. (Mselmi, Lahiani, & Hamza, 2017)	Tuvo como objetivo predecir las dificultades financieras de las pequeñas y medianas empresas francesas. Los datos se recogieron de la base de datos DIANE que contiene datos financieros anuales sobre empresas francesas pequeñas y medianas desde 2010 hasta 2013.	LOGIT MODEL ARTIFICIAL NEURAL NETWORK SUPPORT VECTOR MACHINES (SVM) PARTIAL LEAST SQUARES HYBRID MODEL INTEGRATING (SVM Y PARTIAL LEAST SQUARES)	Precisión:			La precisión de la predicción de ANN, SVM Y PLS-DA son estables en el tiempo. Un año antes de la quiebra, SVM supera a los otros modelos de acuerdo con la precisión general. Según la especificidad, los modelos ANN y Logit son similares y superan a SVM Y PLS-SVM en un 2.85%. Dos años antes de la quiebra, PLS-SVM supera a SVM y Logit Model, que a su vez supera a PLS-DA Y ANNs de acuerdo con la precisión general. Según la especificidad y los errores Tipo II, SVM y PLS-DA son mejores que los modelos Logit, ANN y PLS-SVM.
				UN AÑO ANTES	DOS AÑOS ANTES		
				85.71%	92.86%		
				87.14%	88.57%		
				88.57%	92.86%		
				84.29%	90%		
				85.72%	94.28%		
4	Predicting financial distress of the South Korean manufacturing industries. (Bae, 2012)	Desarrollaron un modelo de predicción de problemas financieros basado en SVM con la función radial para la predicción de las dificultades financieras de empresas manufactureras con datos financieros de 1888 empresas recopiladas por el fondo de Garantía de Crédito de Corea (KODIT). Compararon el rendimiento de precisión de la clasificación entre RSVM y las técnicas de inteligencia artificial.	SUPPORT VECTOR MACHINES (RSVM - Radial) DECISION TREE (C5.0) NEURAL NETWORK (MLP) MDA LOGIT MODEL	Precisión:			Los experimentos demuestran que RSVM siempre supera a otros modelos en el desempeño de la predicción de dificultades financieras corporativas.
				UN AÑO ANTES	DOS AÑOS ANTES		
				82.35%			
				74.47%			
				78.31%			
				77.65%			
				78.45%			



			BAYESIAN NETWORKS	70.50%	
5	Systematic review of bankruptcy prediction models: Towards a framework for tool selection. (Alaka et al., 2018)	Realizaron una revisión sistemática de 49 artículos publicados entre 2010 y 2015. Esta revisión muestra 6 herramientas de inteligencia artificial más populares y prometedoras dentro del área de investigación de modelos de predicción de bancarrota.	SUPPORT VECTOR MACHINES ARTIFICIAL NEURAL NETWORK DECISION TREE ROUGH SETS GENETIC ALGORITHMS CASE BASED REASONING	Precisión: 94.54% 93.79% 91.83% 91.10% 90.36% 59.24%	Se concluye que ninguna herramienta individual es predominantemente mejor que otras herramientas. Concluyen afirmando que solo se puede encontrar un mejor modelo de rendimiento general mediante la integración informada de herramientas para formar un modelo híbrido.
6	Predicting corporate bankruptcy: where we stand? (Aziz & Dar, 2006)	Se realizó un análisis crítico de un gran número de estudios empíricos de predicción de quiebra corporativa, basados en diversos modelos estadísticos, AIES y teóricos.	DECISION TREE CBR NEURAL NETWORKS GENETIC ALGORITHMS ROUGH SETS	EXACTITUD: 87.0% 83.0% 87.0% 88.0% 91.0%	La revisión muestra que las técnicas estadísticas (MDA y Logit) se han usado con mayor frecuencia, que el enfoque AIES es relativamente nuevo y que los modelos teóricos son relativamente infrecuentes. Se observo que la precisión predictiva era generalmente buena en todos los modelos, la revisión sugiere que el AIES y los modelos teóricos tienen una precisión predictiva promedio ligeramente mejor que los modelos estadísticos.



7	<p>C5.0 Classification Algorithm and Application on Individual Credit Evaluation of Banks.(PANG & GONG, 2009)</p>	<p>Los registros de crédito individual incluyen datos tanto numéricos como no numéricos. El árbol de decisiones es una buena solución para este tipo de problema. Este artículo se centra en la evaluación crediticia individual de un banco comercial. En este artículo, realizamos una investigación profunda sobre el algoritmo C5.0 al incorporar la tecnología de "impulso" en la matriz de costos y el árbol sensible al costo para establecer un nuevo modelo para la evaluación de crédito individual de CommercialBank. Aplicamos nuestro nuevo modelo para evaluar los registros de crédito individuales de un banco alemán y comparamos los resultados del modelo de árbol de decisión ajustado y el original.</p>	<p>DECISION TREE C5.0</p>	<p>Precisión: ENTRENAMIENTO (TIPO A) MODELO PRIMARIO:63.90% MODELO AJUSTADO:85.89% PRUEBA(TIPO A) MODELO PRIMARIO:28.81% MODELO AJUSTADO:67.80%</p>	<p>Se obtuvo una precisión del 85.89% con el modelo ajustado en comparación al modelo primario.</p>
8	<p>A practical approach to bankruptcy prediction for small businesses: Substituting the unavailable financial data for credit card sales information. (Yoon & Kwon, 2010)</p>	<p>Se desarrolló un modelo de predicción de quiebra para pequeñas empresas utilizando información de ventas con tarjeta de crédito en lugar de utilizar datos financieros insuficientes para pequeñas empresas. De este modo, se identifican una variedad de variables que están altamente relacionadas con la información de ventas de tarjetas de crédito y adoptaron una técnica de aprendizaje automático como son las máquinas de vectores de soporte (SVM), el cual compararon con otras técnicas de aprendizaje automático.</p>	<p>SUPPORT VECTOR MACHINES BACKPROPAGATION NEURAL NETWORKS (BPN) CART DECISION TREE C5.0 ANALISIS DISCRIMINANTE MULTIPLE (MDA) ANALISIS DE REGRESION LOGISTICA (LRA)</p>	<p>Precisión: ENTRENAMIENTO-VALIDACION SVM: 79.0% - 74.2% CART: 74.7% - 70.3% BPN: 78.5% - 73.1% MDA: 69.0% - 70.1% C5.0: 78.2% - 70.8% LRA: 68.9% - 70.1%</p>	<p>Se obtuvo mayor precisión con la técnica de máquina de soporte vectorial (SVM) el cual alcanzó una precisión del 79% en la etapa de entrenamiento y un 74.2% de precisión en fase de prueba.</p>



<p>9 Applying a nonparametric random forest algorithm to assess the credit risk of the energy industry in China.(Tang, Cai, & Ouyang, 2018)</p>	<p>El objetivo de este estudio es medir científicamente el riesgo crediticio de las tarjetas de crédito utilizadas en la industria energética de China y sentar las bases para una gestión integral del riesgo crediticio. Basado en un análisis de los factores que influyen en el riesgo de crédito, este estudio aplica el algoritmo de bosques aleatorios y los datos mensuales de las tarjetas de crédito utilizadas por los clientes de la industria de la energía en una sucursal del Postal Savings Bank of China desde abril de 2014 hasta junio de 2017 para crear un modelo efectivo de evaluación de riesgo crediticio y mide científicamente el riesgo crediticio en la industria energética de China.</p>	<p>RANDOM FOREST</p>	<p><u>Precisión:</u> 91.5%</p>	<p>Este estudio reveló que el método de RF tiene una mayor precisión de predicción, es más tolerante a los valores atípicos y el ruido, y es menos probable que tenga problemas de sobrealimentación. Además, se puede usar para estimar la importancia de cada característica y, por lo tanto, es más eficaz para tratar datos de alta dimensión.</p>
<p>10 Exploring the behaviour of base classifiers in credit scoring ensembles.(Marqués, García, & Sánchez, 2012)</p>	<p>Durante los últimos años, se han aplicado con éxito diferentes enfoques para los conjuntos de clasificadores a los problemas de calificación crediticia, demostrando ser más precisos que los modelos de predicción individuales. El presente documento evalúa el desempeño de siete técnicas de predicción individuales cuando se utilizan como miembros de cinco métodos conjuntos diferentes. El objetivo final de este estudio es sugerir clasificadores apropiados para cada enfoque conjunto en el contexto de la calificación crediticia.</p>	<p>VECINO MAS CERCANO CLASIFICADOR INGENUO DE BAYES (NBC) REGRESION LOGISTICA PERCEPTRON MULTICAPA REDES NEURONALES DE FUNCION MAQUINA DE VECTORES DE SOPORTE (SVM) DECISION TREE (C4.5)</p>	<p><u>Precisión:</u> Se hicieron muchas pruebas y se obtuvieron diferentes resultados, pero en general: Los mejores modelos son el árbol de decisiones, el perceptrón multicapa y la regresión logística. El vecino más cercano y Bayes son los peores modelos, independientemente del conjunto.</p>	<p>Al considerar la precisión, se concluyó que los clasificadores 1-NN y NBC tienen el peor desempeño, independientemente del método de conjunto utilizado. También es importante tener en cuenta que, para la mayoría de los enfoques conjuntos, C4.5 parece ser el mejor clasificador de base, seguido de cerca por MLP, regresión logística y SVM.</p>



11 An experimental comparison of classification techniques in debt recoveries scoring: Evidence from South Africa's unsecured lending market. (Colebatch & Engel, 1974)

El objetivo de este estudio es investigar el poder predictivo de algunas de las técnicas de clasificación más populares actualmente en uso, con atención específica para predecir la propensión de un prestatario que tiene 90 días o más de atraso en un préstamo sin garantía para pagar en un plazo fijo.

ADABOOST
REGRESION
LOGISTICA
ANALISIS
DISCRIMINANTE
CLASICO
MODELOS
ADITIVOS
GENERALIZADOS
(GAM_LR Y
GAM_GEV)

Precisión:

Los resultados son favorables para el modelo GAM con un enlace GEV (GAM_GEV) ocupa el número 1 de los 9 clasificadores considerados en este estudio.

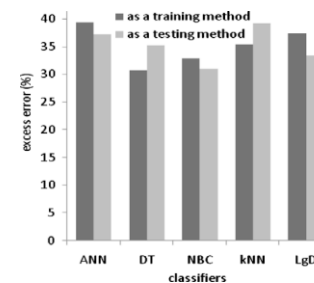
En un estudio comparativo de 9 metodologías de clasificación en 10 conjuntos de datos, hubo pruebas estadísticamente significativas de que los modelos de aditivos generalizados (GAM) con una función de enlace de valor extremo generalizado (GEV) logran mayor precisión de predicción.

12 Multiple classifier application to credit risk assessment. (Twala, 2010)

Se investiga la precisión de cinco clasificadores aplicados para la predicción de riesgo de crédito y fue probado en cuatro conjuntos de datos del mundo real.

REDES
NEURONALES
ARTIFICIALES
DECISION TREE
NBC
k-NEAREST
NEIGHBOUR
LgD

Precisión:



kNN es evidentemente el peor método global debido a la difícil elección que se tiene que hacer al determinar la métrica de distancia k que se desconoce para n finito. NBC es el más eficaz pero seguido de cerca por los métodos LgD y DT.

13 An empirical evaluation of the

En este estudio, se comparó el rendimiento predictivo de 20 clasificadores binarios en una gran muestra de datos de cambios en las calificaciones

ADABOOST

AUC:

El análisis de la puntuación AUC y H indica que los clasificadores más nuevos, como el aumento generalizado, AdaBoost y



	performance of binary classifiers in the prediction of credit ratings changes.(Jones et al., 2015)	crediticias corporativas internacionales que se producen entre 1983 y 2013. Los clasificadores se prueban empíricamente en ocho variaciones del conjunto de datos que incluyen muestras de prueba transversales y longitudinales; y la comparación del rendimiento predictivo a través de conjuntos de datos transformados y no transformados, con imputación de valor faltante y sin él.	GENERALIZED BOOSTING RANDOM FOREST RANDOM FOREST OBLIQUE NEURAL NETWORKS	AdaBoost -> 0.9436 GeneralisedBoosting -> 0.9571 RandomForests -> 0.9442 RandomForests_Oblique -> 0.9229 NeuralNetworks -> 0.8359 SVM -> 0.8385	los bosques aleatorios, superaron ampliamente a todos los demás clasificadores en las muestras de prueba de sección transversal y longitudinal, y en casi todas las variaciones de los datos muestreados. Dado que estos clasificadores parecen ser notablemente precisos y estables, pueden ser una promesa significativa para futuras investigaciones y prácticas en este campo.
14	Machine learning models and bankruptcy prediction.(Barboza, Kimura, & Altman, 2017)	Estudiaron varios modelos de aprendizaje automático y técnicas estadísticas para predecir eventos de quiebra e incumplimiento, para la gestión del riesgo de crédito. Utilizaron datos de 1985 a 2013 sobre empresas norteamericanas, integrando información de la base de datos del Centro Salomon y Compustat	BAGGING BOOSTING (VARIANTE ADABOOST) RANDOM FOREST SUPPORT VECTOR MACHINES (SVM-LIN Y SVM-RBF)	PRECISION: SVM-Linear -> 71.52% SVM-RBF -> 79.77% Boosting -> 86.65% Bagging -> 85.67% Random forest -> 87.06% Neural networks -> 72.98% Logit -> 76.29% MDA -> 52.18%	Con respecto a los índices de precisión, los resultados muestran que los modelos tradicionales (MDA, LR y ANN) tienen una capacidad predictiva más baja (entre 52% y 77%) que los modelos de aprendizaje automático (71% a 87%). Los nuevos estudios pueden adaptar estas técnicas de aprendizaje automático para otros estudios de riesgo crediticio.
15	AdaBoost based bankruptcy forecasting of Korean construction companies.(Heo & Yang, 2014)	En este documento, se muestra que el AdaBoost (impulso adaptativo) es un modelo apropiado para juzgar el riesgo financiero de las empresas de construcción coreanas. Clasificamos las empresas de construcción en tres grupos: grande, mediano y pequeño según el capital de una empresa. Analizamos la capacidad predictiva de AdaBoost y otros algoritmos para cada grupo de empresas.	ADABOOST ARTIFICIAL NEURAL NETWORK SUPPORT VECTOR MACHINES DECISION TREE Z-SCORE	PRECISION: Adaboost -> 78.5% ANN -> 77.1% SVM -> 73.3% Arbol de Desiciones -> 73.1% Z-Score -> 51.3%	Los resultados experimentales mostraron que AdaBoost tiene más poder predictivo que otros, especialmente para el gran grupo de empresas que tiene el capital más de 50 mil millones de won.



16	Classifier selection and clustering with fuzzy assignment in ensemble model for credit scoring.(Zhang, He, & Zhang, 2018)	Este estudio se basa en el conjunto de cinco de los clasificadores básicos más reconocidos en calificación crediticia: regresión logística, máquina de vectores de soporte, red neuronal, árbol de decisión de aumento de gradiente y bosque aleatorio. Se prueba en tres conjuntos de datos de puntuación de crédito (australiano, alemán, japonés) y tres de medidas de rendimiento (precisión, AUC, puntuación F).	KNN LOGISTIC REGRESION RANDOM FOREST GRADIENT BOOSTING DECISION TREE	Evaluaron 3 países: Australian, German, Japanese Precision, AUC, F-Score (0.7913-0.7095-0.8000) (0.8688-0.7560-0.8232) (0.8609-0.8649-0.7230) (0.8551-0.7530-0.8616)	Llegaron a la conclusión que las técnicas LR, SVM, RF y GBDT son los mejores clasificadores base entre los nueve que probaron.
17	A genetic algorithm approach for SMEs bankruptcy prediction: Empirical evidence from Italy.(Gordini, 2014)	Comparan el potencial de los algoritmos genéticos (AG) con los de regresión logística (LR) y máquina de vectores de soporte (SVM), en la predicción de quiebra. Se aplican a una gran muestra de 3.100 pymes manufactureras italianas, tres, dos y un año antes de la quiebra. Los resultados indican que las AG son un instrumento muy eficaz y prometedor para evaluar la probabilidad de Quiebra de PYME comparada con LR y SVM.	GENETIC ALGORITHMS SUPPORT VECTOR MACHINES LOGISTIC REGRESION	Precisión: 71.50% 69.50% 66.80%	Los resultados indican que las AG son un instrumento muy eficaz y prometedor para evaluar la probabilidad de quiebra de las PYMES comparada con LR y SVM.
18	Data mining for credit card fraud: A comparative study.(Siddhartha & Sanjeev, 2010)	En este estudio evaluaron dos técnicas avanzadas como son la máquina de soporte vectorial y bosques aleatorios junto con la conocida regresión logística en la detección de fraude de tarjetas de crédito. El conjunto de datos estuvo conformado por 50 millones de transacciones efectuadas en 13 meses de enero del 2006 hasta enero del 2017.	RANDOM FOREST (RF) LOGISTIC REGRESSION (RL) SUPPORT VECTOR MACHINES (SVM)	Precision: 96.20 % 94.70 % 93.80 %	Los resultados demuestran que bosques aleatorios para este conjunto de datos logra una precisión del 96.20 %, le sigue regresión logística con una precisión del 94.70 % y por último la máquina de soporte vectorial logra una precisión de 93.80%.
19	Credit Card Default	En este estudio se evalúa el desempeño de las técnicas de árbol de decisión, bosque aleatorio y regresión		Precision: 82.06 %	Los resultados obtenidos al emplear una división del conjunto de datos 70 % para



	Prediction using Machine Learning Techniques. (Sayjadah & Abaker, 2018)	logística en la predicción de incumplimiento de pago de tarjeta de crédito. El conjunto de datos está conformado por 30000 instancias y 24 atributos.	DECISION TREE (DT) LOGISTIC REGRESSION (RL) RANDOM FOREST (RF)	82.00 % 81.81 %	entrenamiento y 30 % para tes obtuvieron un 82.06 % de precisión con la técnica de árboles de decisión seguido de regresión logística con 82% y por último bosques aleatorios con un 81.81 % de precisión.
20	Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks. (Alfaro & Elizondo, 2008)	Se compara la precisión de predicción de las técnicas Redes Neuronales y AdaBoost en un conjunto de empresas europeas, considerando las variables de predicción usuales, como las razones financieras, así como las variables cualitativas, como el tamaño de la empresa, actividad y la estructura legal.	AdaBoost Neural Networks	Precisión: 92% 88%	Concluyen que la técnica AdaBoost es superior en la predicción con una precisión del 92% con respecto a las redes neuronales el cual alcanza una precisión del 88%.

Fuente: IEEE Xplore Digital Library (IEEE, 2018)



ANEXO 2. TOP DE TÉCNICAS PREDICTIVAS

N°	AUTOR	TÉCNICA	PRECISIÓN
1	Alaka et al. (2018)	Máquina de Soporte Vectorial	95 %
2	Alfaro & Elizondo (2008)	AdaBoost	92 %
3	Alaka et al. (2018)	Árboles de decisión	92 %
4	Alaka et al. (2018)	Algoritmos Genéticos	90 %
5	Aziz & Dar (2006)	Redes Neuronales	87 %
6	Zhang, He, & Zhang (2018)	Bosques Aleatorios	86 %
7	Mselmi et al. (2017)	Regresión Logística	93 %
8	Jones et al. (2015)	Probit GAM	91 %
9	Jones et al. (2015)	Probit Boosted	78 %
10	Bae (2012)	Análisis discriminante múltiple	78 %
11	Yoon & Kwon (2010)	CART	74 %
12	Li & Wang (2018)	KNN	70 %

Fuente: Equipo de Investigación

ANEXO 3. CONSULTA DE ASEGURADOS A ESSALUD

Consulta de Acreditación

Señor asegurado⁽¹⁾, usted y sus derechohabientes⁽²⁾ podrán ingresar sus datos e informarse si se encuentran acreditados (con vigencia de atención actualizada a la fecha de solicitud) para acceder a los servicios⁽³⁾ que brinda EsSalud.

Para que se encuentre acreditado en nuestros sistemas, su empleador debe declarar correctamente sus datos personales (apellidos y nombres, tipo y número de documento de identidad, fecha de nacimiento), los mismos que deben coincidir con su documento de identidad (Seg. Regular) así como haber efectuado el pago de aportes de acuerdo a las condiciones establecidas en cada tipo de seguro (Seg. Agrario Independiente y Seguro Potestativo).

Asimismo, podrá consultar el centro asistencial asignado para su atención considerando la ubicación de su domicilio. En caso que desee actualizar su domicilio podrá acercarse a la oficina de EsSalud más cercana, portando su documento de identidad personal⁽⁴⁾

Buscar por

Tipo de Documento ▼

Tipo de Documento:

▼

Número de Documento

Consultar

Borrar

Ingrese el código de la imagen:

13322

Fuente: <http://ww4.essalud.gob.pe:7777/acredita/>. ESSALUD 2018



ANEXO 4. DESCRIPCIÓN DE LAS TABLAS QUE CONFORMAN LA BASE DE DATOS ORIGINAL

Nro	Tabla	Descripción	Nro Registros	VARIABLES DE PREDICCIÓN											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
1	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
2	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
3	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
4	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
5	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
6	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
7	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
8	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
9	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
10	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
11	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
12	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
13	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
14	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
15	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
16	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
17	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
18	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
19	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
20	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
			Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría		
21	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
22	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
23	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
24	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
25	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
26	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
27	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
28	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
29	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
30	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
31	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
32	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
33	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
34	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
35	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
36	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
37	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
38	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
39	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
40	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



Nro	Tabla	Descripción	Nro Registros	VARIABLES DE PREDICCIÓN											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
41	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
42	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
43	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
44	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
45	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
46	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
47	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
48	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
49	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
50	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
51	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
52	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
53	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
54	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
55	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
56	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



Nro	Tabla	Descripción	Nro Registros	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
58	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
59	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
60	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
61	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
62	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
63	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
64	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
65	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
66	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
67	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
68	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
69	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
70	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
71	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
72	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
73	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
74	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
75	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
76	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
77	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
78	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
79	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
80	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
81	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
82	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
83	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
84	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
85	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
86	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
87	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
88	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
89	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
90	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
91	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
92	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
93	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
94	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
95	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
96	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
97	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
98	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
99	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
100	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
101	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
102	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
103	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
104	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
105	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
106	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
107	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
108	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
109	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	VARIABLES DE PREDICCIÓN											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
110	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
111	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
112	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
113	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
114	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
115	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
116	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
117	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
118	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
119	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
120	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
121	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
122	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
123	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
124	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
125	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
126	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
127	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
128	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
129	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
130	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
131	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
132	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
133	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
134	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
135	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
136	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
137	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
138	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
139	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
140	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
141	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
142	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
143	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
144	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
145	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
146	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
147	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
148	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
149	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
150	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
151	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
152	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
153	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
154	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
155	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
156	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
157	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
158	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
159	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
160	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
161	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
162	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
163	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	VARIABLES DE PREDICCIÓN											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
164	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
165	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
166	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
167	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
168	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
169	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
170	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
171	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
172	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
173	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
174	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
175	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
176	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
177	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
178	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
179	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
180	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
181	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
182	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
183	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
184	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
185	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
186	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
187	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
188	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
189	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
190	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros													
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría		
191	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
192	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
193	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
194	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
195	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
196	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
197	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
198	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
199	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
200	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
201	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
202	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
203	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
204	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
205	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
206	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
207	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
208	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
209	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
210	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
211	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
212	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
213	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
214	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
215	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
216	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
217	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros													
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría		
218	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
219	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
220	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
221	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
222	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
223	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
224	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
225	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
226	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
227	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
228	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
229	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
230	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
231	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
232	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
233	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
234	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
235	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
236	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
237	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
238	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
239	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
240	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
241	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
242	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
243	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
244	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría	
245	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
246	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
247	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
248	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
249	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
250	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
251	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
252	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
253	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
254	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
255	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
256	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
257	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
258	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
259	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
260	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
261	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
262	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
263	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
264	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
265	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
266	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
267	CAMPO	PROTEGIDO	-	SR	S	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
268	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
269	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
270	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
271	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
272	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
273	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
274	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
275	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
276	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
277	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
278	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
279	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
280	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
281	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
282	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
283	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
284	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
285	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
286	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
287	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
288	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
289	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
290	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
291	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
292	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
293	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
294	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
295	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
296	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
297	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
298	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría	
299	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
300	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
301	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
302	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
303	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
304	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
305	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
306	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
307	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
308	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
309	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
310	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
311	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
312	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
313	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
314	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
315	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
316	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
317	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
318	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
319	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
320	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
321	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
322	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
323	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
324	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
325	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
326	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
327	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
328	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
329	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
330	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
331	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
332	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
333	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
334	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
335	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
336	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
337	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
338	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
339	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
340	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
341	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
342	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
343	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
344	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
345	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
346	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
347	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
348	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
349	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
350	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
351	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
352	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	VARIABLES DE PREDICCIÓN											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
353	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
354	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
355	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
356	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
357	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
358	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
359	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
360	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
361	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
362	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
363	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
364	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
365	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
366	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
367	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
368	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
369	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
370	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
371	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
372	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
373	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
374	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
375	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
376	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
377	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
378	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
379	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
380	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
381	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
382	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
383	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
384	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
385	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
386	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
387	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
388	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
389	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
390	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
391	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
392	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
393	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
394	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
395	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
396	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
397	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
398	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
399	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
400	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
401	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
402	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
403	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
404	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
405	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
406	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros													
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría		
407	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
408	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
409	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
410	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
411	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
412	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
413	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
414	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
415	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
416	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
417	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
418	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
419	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
420	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
421	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
422	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
423	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
424	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
425	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
426	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
427	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
428	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
429	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
430	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
431	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
432	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
433	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Variables de Predicción											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría	
434	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
435	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
436	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
437	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
438	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
439	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
440	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
441	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
442	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
443	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
444	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
445	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
446	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
447	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
448	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
449	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
450	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
451	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
452	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
453	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
454	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
455	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
456	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
457	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
458	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
459	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
460	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
461	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
462	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
463	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
464	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
465	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
466	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
467	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
468	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
469	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
470	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
471	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
472	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
473	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
474	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
475	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
476	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
477	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
478	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
479	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
480	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
481	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
482	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
483	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
484	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
485	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
486	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
487	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Variables de Predicción											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
488	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
489	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
490	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
491	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
492	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
493	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
494	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
495	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
496	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
497	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
498	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
499	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
500	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
501	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
502	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
503	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
504	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
505	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
506	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
507	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
508	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
509	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
510	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
511	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
512	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
513	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
514	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría	
515	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
516	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
517	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
518	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
519	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
520	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
521	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
522	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
523	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
524	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
525	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
526	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
527	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
528	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
529	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
530	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
531	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
532	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
533	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
534	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
535	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
536	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
537	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
538	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
539	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
540	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
541	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Variables de Predicción											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría	
542	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
543	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
544	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
545	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
546	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
547	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
548	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
549	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
550	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
551	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
552	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
553	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
554	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
555	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
556	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
557	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
558	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
559	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
560	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
561	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
562	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
563	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
564	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
565	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros													
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría		
566	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
567	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
568	CAMPO	PROTEGIDO	-	I	D	I	SR	SR	SR	SR	SR	SR	D	D	SR	SR
569	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
570	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
571	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
572	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
573	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
574	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
575	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
576	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
577	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
578	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
579	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
580	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
581	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
582	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	D	D	SR	I	D	SR	SR	SR	SR
583	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
584	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	D	D	SR	I	D	SR	SR	SR	SR
585	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
586	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
587	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
588	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
589	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
590	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Variables de Predicción												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría		
591	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
592	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
593	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
594	CAMPO	PROTEGIDO	-	SR	SR	SR	S	SR	SR	SR	SR	SR	SR	SR	SR	
595	CAMPO	PROTEGIDO	-	SR	SR	SR	S	SR	SR	SR	SR	SR	SR	SR	SR	
596	CAMPO	PROTEGIDO	-	SR	SR	SR	S	SR	SR	SR	SR	SR	SR	SR	SR	
597	CAMPO	PROTEGIDO	-	SR	SR	SR	S	SR	SR	SR	SR	SR	SR	SR	SR	
598	CAMPO	PROTEGIDO	-	SR	SR	SR	S	SR	SR	SR	SR	SR	SR	SR	SR	
599	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
600	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
601	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
602	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
603	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
604	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
605	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
606	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
607	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
608	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
609	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
610	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
611	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
612	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
613	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
614	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	VARIABLES DE PREDICCIÓN											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría	
615	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
616	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
617	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
618	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
619	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
620	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
621	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
622	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
623	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
624	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
625	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
626	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
627	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
628	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
629	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
630	CAMPO	PROTEGIDO	-	SR	SR	SR	S	SR	SR	SR	SR	SR	SR	SR	SR
631	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
632	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
633	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	I	SR	SR	SR	SR	SR
634	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
635	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
636	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
637	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
638	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
639	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría	
640	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
641	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
642	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
643	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
644	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
645	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
646	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
647	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
648	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
649	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
650	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
651	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
652	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
653	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
654	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
655	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
656	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
657	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
658	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
659	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
660	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
661	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
662	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
663	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
664	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
665	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
666	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Variables de Predicción											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría	
667	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
668	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
669	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
670	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
671	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
672	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
673	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
674	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
675	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
676	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
677	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
678	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
679	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
680	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
681	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
682	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
683	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
684	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
685	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
686	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
687	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
688	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
689	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
690	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
691	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
692	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
693	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
694	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
695	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
696	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
697	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
698	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
699	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
700	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
701	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
702	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
703	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
704	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
705	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
706	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
707	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
708	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
709	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
710	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
711	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
712	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
713	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
714	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
715	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
716	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
717	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
718	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
719	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
720	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
721	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
722	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
723	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
724	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
725	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
726	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
727	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
728	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
729	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
730	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
731	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
732	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
733	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
734	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
735	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
736	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
737	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
738	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
739	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
740	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
741	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
742	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	D	D	SR	I	SR	SR	SR
743	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
744	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
745	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
746	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
747	CAMPO	PROTEGIDO	-	I	D	I	SR	SR	SR	SR	SR	D	D	D



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	VARIABLES DE PREDICCIÓN										
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría
748	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
749	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
750	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
751	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
752	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
753	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
754	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
755	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
756	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
757	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
758	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
759	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
760	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
761	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
762	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
763	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
764	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
765	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
766	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
767	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
768	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
769	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
770	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
771	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
772	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
773	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Variables de Predicción												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría		
774	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
775	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
776	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
777	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
778	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
779	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
780	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
781	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
782	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	D	D	SR	I	SR	SR	SR	SR	SR
783	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
784	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
785	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
786	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
787	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
788	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
789	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
790	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
791	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Variables de Predicción												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría		
792	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
793	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
794	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
795	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
796	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
797	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
798	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
799	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
800	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	D	SR	SR	SR
801	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
802	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
803	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
804	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
805	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
806	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
807	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
808	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
809	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
810	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
811	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
812	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
813	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
814	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
815	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
816	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
817	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	VARIABLES DE PREDICCIÓN												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría		
818	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
819	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
820	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
821	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
822	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
823	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
824	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
825	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
826	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
827	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
828	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
829	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
830	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
831	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
832	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
833	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
834	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	I	D			SR
835	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
836	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Variables de Predicción												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría		
837	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
838	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
839	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
840	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
841	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
842	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
843	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
844	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
845	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
846	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
847	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	I	SR	SR	SR	SR	SR
848	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
849	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
850	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
851	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
852	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
853	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
854	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
855	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
856	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
857	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
858	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
859	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
860	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
861	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
862	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros													
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría		
863	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
864	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
865	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
866	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
867	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
868	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
869	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
870	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
871	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
872	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
873	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
874	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
875	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
876	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
877	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	I	SR	SR
878	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
879	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
880	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
881	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
882	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
883	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
884	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
885	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
886	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
887	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Variables de Predicción												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría		
888	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
889	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
890	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
891	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
892	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
893	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
894	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
895	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
896	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
897	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
898	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
899	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
900	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
901	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
902	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
903	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
904	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
905	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
906	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
907	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
908	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
909	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
910	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
911	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
912	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
913	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
914	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Variables de Predicción												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría		
915	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
916	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
917	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
918	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
919	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
920	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
921	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
922	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
923	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	D	
924	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
925	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
926	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
927	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
928	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
929	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
930	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
931	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
932	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
933	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	I	SR	SR	SR	SR	SR	SR	SR	
934	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SI	SR	SR	SR	SR	SR	
935	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
936	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
937	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
938	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
939	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
940	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
941	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
942	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
943	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
944	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
945	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
946	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
947	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
948	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
949	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
950	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
951	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
952	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
953	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
954	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
955	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
956	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
957	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
958	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
959	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
960	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
961	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
962	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
963	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
964	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
965	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
966	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
967	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
968	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
969	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
970	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
971	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
972	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
973	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
974	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	I	SR	SR	SR	SR	SR	SR	SR
975	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
976	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
977	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
978	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
979	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
980	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
981	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
982	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
983	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
984	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
985	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
986	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
987	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
988	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
989	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
990	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
991	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
992	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
993	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
994	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
995	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
996	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
997	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
998	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
999	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1000	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1001	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1002	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1003	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1004	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1005	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1006	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1007	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros	VARIABLES DE PREDICCIÓN												
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Numero de Hijos	Dias de Incumplimiento	Capítulo	Especialidad	Categoría		
1008	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1009	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1010	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1011	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1012	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1013	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1014	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1015	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1016	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1017	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1018	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1019	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1020	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1021	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1022	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1023	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1024	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1025	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1026	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1027	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1028	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1029	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1030	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



VARIABLES DE PREDICCIÓN

Nro	Tabla	Descripción	Nro Registros													
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría		
1031	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1032	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1033	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1034	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1035	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1036	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
1037	CAMPO	PROTEGIDO	-	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación



ANEXO 5. DESCRIPCIÓN DE LAS TABLAS QUE TIENEN RELACIÓN CON LAS VARIABLES DE PREDICCIÓN

Tabla	Nro Registros	Columnas	Descripción	VARIABLES DE PREDICCIÓN											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
data_web_09012014	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	I	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	I	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	D	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	D	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	D	SR	SR

Fuente: Equipo de Investigación



VARIABLES DE PREDICCIÓN

Tabla	Nro Registros	Columnas	Descripción	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
dnava	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	I	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	D	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	I	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	D	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación



VARIABLES DE PREDICCIÓN

Tabla	Nro Registros	Columnas	Descripción	VARIABLES DE PREDICCIÓN										
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
drolly	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	D	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	D	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	I	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	D	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	I	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	I	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	D	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	I	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación



VARIABLES DE PREDICCIÓN

Tabla	Nro Registros	Columnas	Descripción	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
Dtl01 Con	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	I	SR	SR
		CAMPO	PROTEGIDO	I	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	D	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	D	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación



VARIABLES DE PREDICCIÓN

Tabla	Nro Registros	Columnas	Descripcion	VARIABLES DE PREDICCIÓN											
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría	
mstf01ccc	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	D	D	SR	D	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
		CAMPO	PROTEGIDO	I		I									
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	D	SR	SR	SR	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	I	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	D	SR	SR	SR	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	
				CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	I	SR	SR	SR	SR
				CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	D	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR		
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR		



CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	I	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación



Tabla	Nro Registros	Columnas	Descripción	VARIABLES DE PREDICCIÓN										
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
mst01cli	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	D	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	I	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	D	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	D	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	D	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	I	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	I	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	D	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación



VARIABLES DE PREDICCIÓN

Tabla	Nro Registros	Columnas	Descripción	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
roly	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	I	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	I	D	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	D	SR	I	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación



VARIABLES DE PREDICCIÓN

Tabla	Nro Registros	Columnas	Descripción	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
tbl_capitulo	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	D	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación

VARIABLES DE PREDICCIÓN

Tabla	Nro Registros	Columnas	Descripción	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
tbl_especialidad	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	D	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR



Fuente: Equipo de Investigación

Tabla	Nro Registros	Columnas	Descripción	VARIABLES DE PREDICCIÓN										
				Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
Tbl_Parentesco	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	I	SR	SR	SR	SR

Fuente: Equipo de Investigación

Tabla	Nro Registros	Columnas	Descripción	VARIABLES DE PREDICCIÓN									
				Género	Estado Civil	Edad	Monto Pagado	Número de Hijos	Días de Incumplimiento	Especialidad	Categoría		
tbl01cdv	-	CAMPO	PROTEGIDO	SR	SR	SR	I	SR	SR	SR	SR		
		CAMPO	PROTEGIDO	SR	SR	SR	I	SR	SR	SR	SR		
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR		
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR		
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR		
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR		
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR		
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR		
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR		



Codcon

Sin Relevancia

SR SR SR SR SR SR SR SR

Fuente: Equipo de Investigación

VARIABLES DE PREDICCIÓN

Tabla	Nro Registros	Columnas	Descripción	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
tbl01cac	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	D
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	D
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación



VARIABLES DE PREDICCIÓN

Tabla	Nro Registros	Columnas	Descripción	Género	Estado Civil	Edad	Habilidad	Monto Pagado	Meses de Pago	Número de Hijos	Días de Incumplimiento	Capítulo	Especialidad	Categoría
tbl01ven	-	CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR
		CAMPO	PROTEGIDO	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR	SR

Fuente: Equipo de Investigación



**ANEXO 6. DICCIONARIO DE DATOS DE LA BASE DE DATOS
TRANSACCIONAL**

Anexo 6.1. Descripción de la tabla Especialidad

Nombre Columna	Tipo Dato	Null Option	Descripción	PK	FK
CodEsp	Char(6)	Not null	Código Especialidad	PK	-
NomEsp	Varchar(100)	Not null	Nombre especialidad	-	-
CodCap	Char(3)	Null	Código del capítulo	-	FK

Fuente: Equipo de Investigación

Anexo 6.2. Descripción de la tabla Capítulo

Nombre Columna	Tipo Dato	Null Option	Descripción	PK	FK
CodCap	Char(3)	Not null	Código Capítulo	PK	-
NomCap	Varchar(100)	Not null	Nombre Capítulo	-	-
Orden_Cap	Int	Null	Orden de los Capítulo	-	FK

Fuente: Equipo de Investigación



Anexo 6.3. Descripción de la tabla Familiar

Nombre Columna	Tipo Dato	Null Option	Descripción	PK	FK
IdFamiliar	Int	Not null	Código del Familiar del Asociado al CIP	PK	-
Familiar	Varchar(255)	Not null	Nombres Completos del Familiar del Asociado al CIP	-	-
FechaNac	Datetime	Null	Fecha de Nacimiento del Familiar del Asociado al CIP	-	-
Edad	Int	Null	Edad del Familiar Asociado al CIP	-	-
NumDni	Char(10)	Null	Numero de DNI del familiar Asociado al CIP	-	-
NroCIP	Char(7)	Not null	Numero del CIP del Asociado	-	FK
NomPar	Varchar(20)	Not null	Nombre del Parentesco del Asociado al CIP	-	-

Fuente: Equipo de Investigación



Anexo 6.4. Descripción de la tabla Cobro

Nombre Columna	Tipo Dato	Null Option	Descripción	PK	FK
CodCdv	Char(2)	Not null	Codigo del Cobro del Asociado al CIP	PF	-
NomCdv	Char (20)	Not null	Nombre del Cobro del Asociado al CIP	-	-
Diaven	Int	Null	Número de días a pagar	-	-
CtaCol	Char(12)	Null	Numero de Cuota a pagar	-	-
CtaCoe	Char(12)	Null	Numero de Cuota a pagar	-	-
CtaCrL	Char(12)	Null	Numero de Cuota a pagar	-	-
CtaCrE	Char(12)	Null	Numero de Cuota a pagar	-	-

Fuente: Equipo de Investigación

Anexo 6.5. Descripción de la tabla Cajero_AtencionCliente

Nombre Columna	Tipo Dato	Null Option	Descripción	PK	FK
CodVen	Char(5)	Not null	Codigo del Cajero	PK	-
NomVen	Char(30)	Not null	Nombre del Cajero	-	-

Fuente: Equipo de Investigación



Anexo 6.6. Descripción de la tabla Pagos

Nombre Columna	Tipo Dato	Option Null	Descripción	PK	FK
Cod Pago	Int	Not null	Código del Pago	PK	-
NroCIP	Char(7)	Not null	Numero CIP de los Asociados	-	-
Mes_Pago	Varchar(10)	Not null	Monto totañ del mes	-	-
FechaApertura	Date	Not null	Fecha de inicio de la apertura	-	-
FechaVencimiento	Date	Not null	Fecha de vencimiento	-	-
FechaPago	Date	Null	Fecha del Pago asociados al CIP	-	-
Monto_Programacion	Float	Not null	Monto programado en la U	-	-
Monto_Pagado	Float	Not null	Monto total para pagar de los asociados al CIP.	-	-
Saldo	Float	Floatnot null	Saldo del Asociado al CIP	-	-
Días	Int	Null		-	-
CodCdv	Char(2)	Null	Codigo de Cobro	-	FK
CodVen	Char(2)	Null	Codigo de Venta	-	FK

Fuente: Equipo de Investigación



ANEXO 7. DESCRIPCIÓN DE ATRIBUTOS DE LAS DIMENSIONES

Anexo 7.1. Descripción de atributos de la dimensión DIM_COLEGIADO

CAMPO	TIPO	LONGITUD	DESCRIPCIÓN
IDCOLEGIADO	Int	-	Clave Primaria
NROCIPI	Char	7	Número CIP del colegiado
NOMCOLEG	Varchar	100	Nombre completo del colegiado
EDAD	Char	3	Edad del colegiado
ESTADOCIVIL	Char	1	Estado Civil del colegiado
TIPOSEGURO	Varchar	30	Tipo de Seguro (ESSALUD, PACIFICO, MAPFRE) al que se encuentra afiliado el colegiado.
NOMCAT	Varchar	30	Nombre de categoría del colegiado (Ordinario, Vitalicio, transeúnte)
GENERO	Varchar	1	Género del colegiado
NOMESPECIALIDAD	Varchar	100	Nombre de especialidad del colegiado
NOMCAPITULO	Varchar	100	Nombre del capítulo al cual pertenece una especialidad
NROHIJOS	Char	3	Número de hijos del colegiado
FECHANAC	Date	-	Fecha de nacimiento del colegiado
FECHACOL	Date	-	Fecha de incorporación al colegio de Ingenieros
ESTADO_HABILIDAD	Varchar	30	Estado de habilidad

Fuente: Equipo de Investigación

Anexo 7.2. Descripción de atributos de la dimensión DIM_CUOTAMENSUAL

CAMPO	TIPO	LONGITUD	DESCRIPCIÓN
IDCUOTA	Int	-	Clave Primaria
IDPAGO	Int	-	Identificador de Pago.
FORMAPAGO	Varchar	30	Tipo de pago (Efectivo, trajeta)
FECHAVENCIMIENTO	date	-	Fecha de vencimiento de pago programado por concepto de cuota mensual.
FECHAPAGO	date	-	Fecha en el que efectúa el pago de cuota social el colegiado.

Fuente: Equipo de Investigación



Anexo 7.3. Descripción de atributos de la dimensión DIM_TIEMPO

CAMPO	TIPO	LONGITUD	DESCRIPCIÓN
IDTIEMPO	Int	-	Clave Primaria
ANIO	Varchar	4	Año en curso.
MES	Varchar	2	Mes en curso.
DIA	Varchar	2	Día en curso.
Fecha	date	-	Fecha en curso.

Fuente: Equipo de Investigación

Anexo 7.4. Descripción de atributos de la tabla hechos FACT_PAGOS

CAMPO	TIPO	LONGITUD	DESCRIPCIÓN
IDPAGOS	Int	-	Clave Primaria.
IDTIEMPO	Int	-	Clave Foránea.
IDCOLEGIADO	Int	-	Clave Foránea.
MONTO_PROGRAMADO	Float	-	Monto programado mes a mes por concepto de cuota social.
MONTO_PAGADO	Float	-	Monto pagado por el colegiado por concepto de cuota mensual.
SALDO	Float	-	Monto acumulado por atraso en pago por concepto de cuotas sociales.
DIAS_ATRASOPAGO	Int	-	Días de atraso en el pago de la cuota mensual.

Fuente: Equipo de Investigación



ANEXO 8. CÓDIGO EN RSTUDIO DE LA IMPLEMENTACIÓN DE MÁQUINA DE SOPORTE VECTORIAL

```

#Cargar los datos
dataset <-read.csv("Enero_Algoritmo.csv", header = T)
dataset

testing <- read.csv("Enero_Predecir.csv")

#Cargar librerías
library(caTools)
library(caret)
library(ROCR)
library(e1071)

#Datos de entrenamiento
set.seed(1234)
split <- sample.split(dataset$MOROSO, SplitRatio = 0.80)
training_set <- subset(dataset, split == TRUE)
test_set <- subset(dataset, split == FALSE)

#Comprobando distribución
table(training_set$MOROSO)

table(test_set$MOROSO)

#Crear modelo
clasificadorSVM <- svm(MOROSO ~ ., data = training_set,
                      type = 'C-classification', kernel = 'linear')

summary(clasificadorSVM ) # ver resumen de modelo

#Matriz de confusión
pred_valid_svm <- predict(clasificadorSVM, newdata = test_set)
pred_valid_svm
matrizConfusion <- table(test_set$MOROSO, pred_valid_svm)
matrizConfusion

confusionMatrix( test_set$MOROSO,pred_valid_svm, positive='SI' )

```



```

##Prediccion con datos de testing
str(test_set)
str(testing)

prueba <- predict(clasificadorSVM, testing)
prueba
head(prueba)
testing$MOROSO <- prueba
write.csv(testing, "Enero_datapredecidaSVM.csv", row.names=F)

proc.time()

#Curva Roc
pred1 <- prediction(as.numeric(pred_valid_svm), as.numeric(test_set$MOROSO))
perf1 <- performance(pred1, "tpr", "fpr")
plot(perf1)

#Area bajo la curva (AUC)
auc <- performance(pred1, "auc")
auc <- unlist(slot(auc, "y.values"))
auc <- round(auc, 4)
legend(.6, .5, auc, title = "AUC", cex = .8)

```

Fuente: Elaboración propia



ANEXO 9. FICHA DE OBSERVACIÓN MAQUINA DE SOPORTE VECTORIAL.

ANEXO 9.1. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE ENERO.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Corei5. **Responsable:** Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS	1
TIPO DE PRUEBA	Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN (seg.)	0.71
MES ANALIZADO	Enero

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1019	142
SI	119	640

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
86.41	81.84	84.32	87.77	0.8605

Fuente: Elaboración propia.



ANEXO 9.2. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE FEBRERO.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Corei5. **Responsable:** Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 2
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.7
 (seg.)
MES ANALIZADO Febrero

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1028	131
SI	146	615

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
85.57	82.44	80.81	88.70	0.8476

Fuente: Elaboración propia.



ANEXO 9.3. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE MARZO.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 3
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.59
 (seg.)
MES ANALIZADO Marzo

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1016	130
SI	133	641

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
86.30	83.14	82.82	88.66	0.8574

Fuente: Elaboración propia.



ANEXO 9.4. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE ABRIL.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Corei5. **Responsable:** Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 4
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.39
 (seg.)
MES ANALIZADO Abril

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1016	130
SI	114	660

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
87.29	83.54	85.27	88.66	0.8696

Fuente: Elaboración propia.



ANEXO 9.5. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE MAYO.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Corei5. **Responsable:** Jefferson Napoleón Huamán Bernilla.
 Computador con procesador **Fecha:** 30 de octubre de 2018.
 Corei5. **Institución:** Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 5
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.42
 (seg.)
MES ANALIZADO Mayo

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1025	132
SI	121	642

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
86.82	82.95	84.14	88.59	0.8637

Fuente: Elaboración propia.



ANEXO 9.6. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE JUNIO.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Corei5. **Responsable:** Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 6
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.37
 (seg.)
MES ANALIZADO Junio

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1028	129
SI	130	633

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
86.51	83.07	82.96	88.85	0.8591

Fuente: Elaboración propia.



ANEXO 9.7. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE JULIO.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Corei5. **Responsable:** Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 7
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.42
 (seg.)
MES ANALIZADO Julio

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1018	142
SI	145	615

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
85.05	81.24	80.92	87.76	0.8434

Fuente: Elaboración propia.



ANEXO 9.8. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE AGOSTO.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Corei5. **Responsable:** Jefferson Napoleón Huamán Bernilla.
 Computador con procesador **Fecha:** 30 de octubre de 2018.
 Corei5. **Institución:** Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 8
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.48
 (seg.)
MES ANALIZADO Agosto

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1028	137
SI	169	586

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
84.06	81.05	77.62	88.24	0.8293

Fuente: Elaboración propia.



ANEXO 9.9. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE SEPTIEMBRE.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 9
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.5
 (seg.)
MES ANALIZADO Septiembre

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1016	130
SI	114	660

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
87.29	83.54	85.27	88.66	0.8696

Fuente: Elaboración propia.



ANEXO 9.10. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE OCTUBRE.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Corei5. **Responsable:** Jefferson Napoleón Huamán Bernilla.
 Computador con procesador **Fecha:** 30 de octubre de 2018.
 Corei5. **Institución:** Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 10
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.4
 (seg.)
MES ANALIZADO Octubre

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1038	113
SI	143	626

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
86.67	84.71	81.40	81.40	0.8579

Fuente: Elaboración propia.



ANEXO 9.11. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE NOVIEMBRE.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Corei5. **Responsable:** Jefferson Napoleón Huamán Bernilla.
 Computador con procesador **Fecha:** 30 de octubre de 2018.
 Corei5. **Institución:** Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 11
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.51
 (seg.)
MES ANALIZADO Noviembre

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1001	141
SI	172	606

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
83.70	81.12	77.89	87.65	0.8277

Fuente: Elaboración propia.



ANEXO 9.12. RESUMEN DE ANÁLISIS DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL CORRESPONDIENTE AL MES DE DICIEMBRE.

**FICHA DE OBSERVACIÓN
MÁQUINA DE SOPORTE VECTORIAL**

Computador: Corei5. **Responsable:** Jefferson Napoleón Huamán Bernilla.
 Computador con procesador **Fecha:** 30 de octubre de 2018.
 Corei5. **Institución:** Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 12
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.39
 (seg.)
MES ANALIZADO Diciembre

MATRIZ DE CONFUSIÓN

	NO	SI
NO	976	141
SI	203	600

EVALUACIÓN DE MÉTRICAS

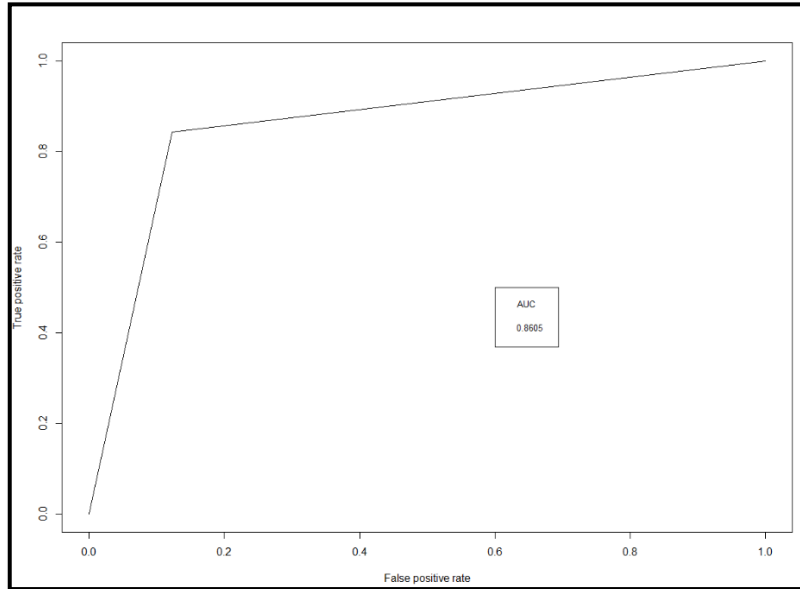
EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
82.08	80.97	74.72	87.38	0.8105

Fuente: Elaboración propia.



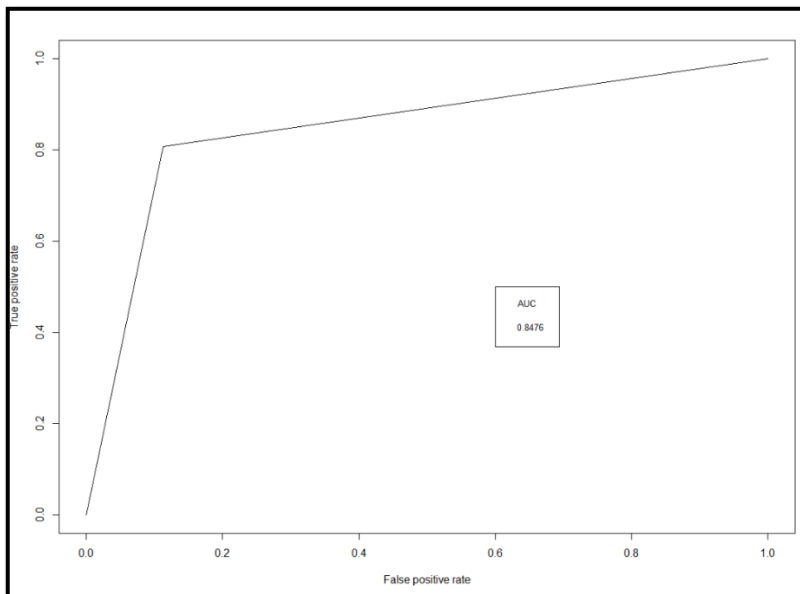
ANEXO 10. CURVA ROC MÁQUINA DE SOPORTE VECTORIAL(SVM).

ANEXO 10.1. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A ENERO



Fuente: Elaboración propia

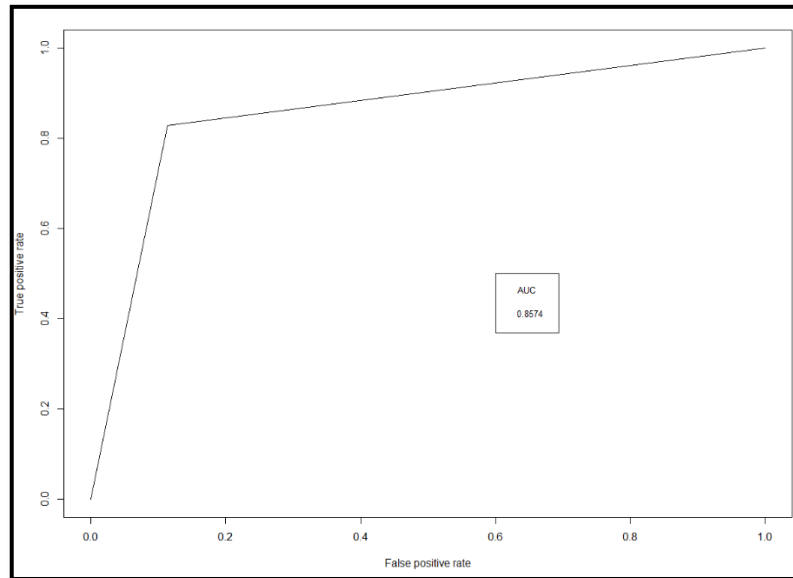
ANEXO 10.2. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A FEBRERO



Fuente: Elaboración propia

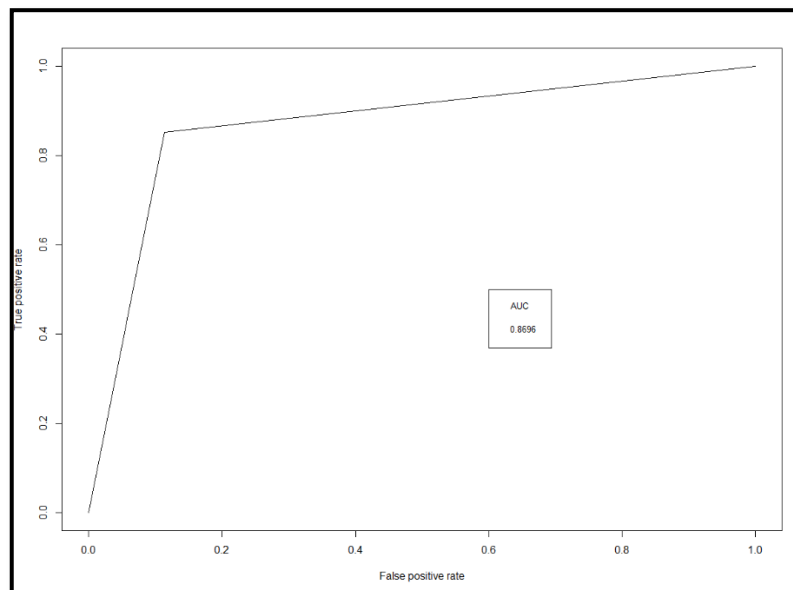


ANEXO 10.3 GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A MARZO



Fuente: Elaboración propia

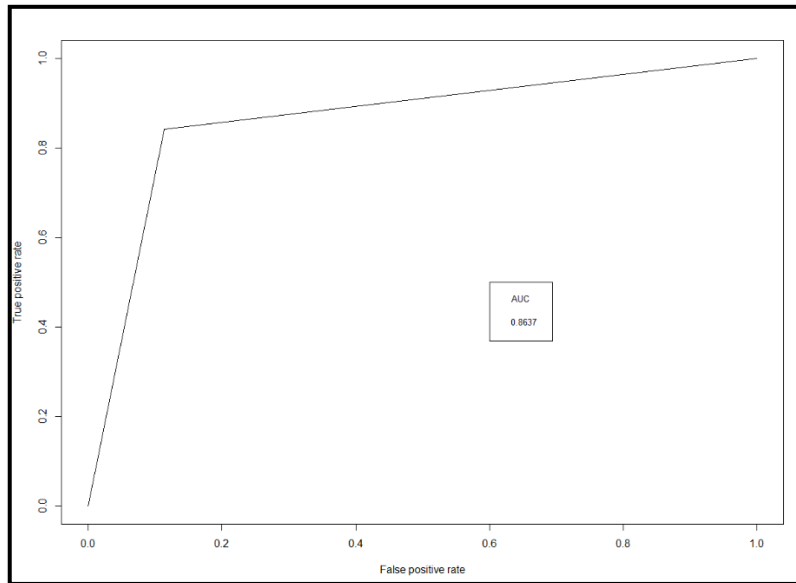
ANEXO 10.4. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A ABRIL



Fuente: Elaboración propia

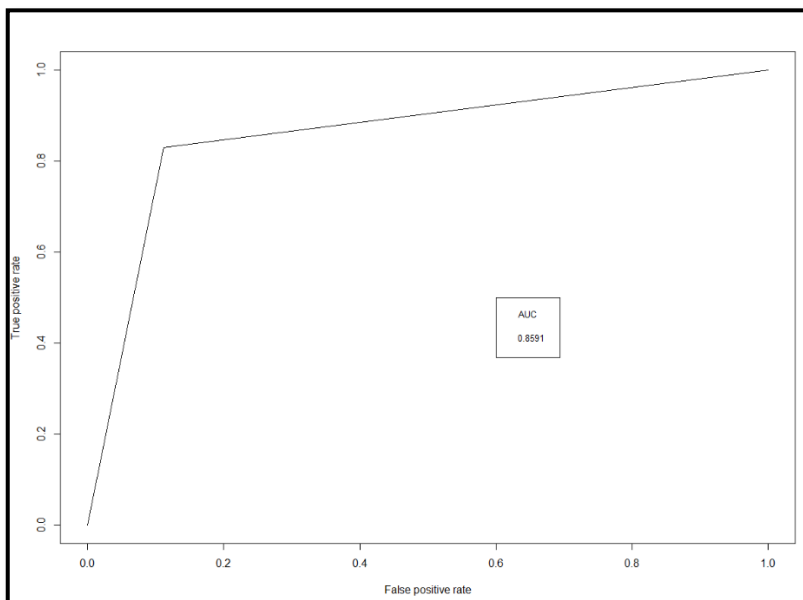


ANEXO 10.5. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A MAYO



Fuente: Elaboración propia

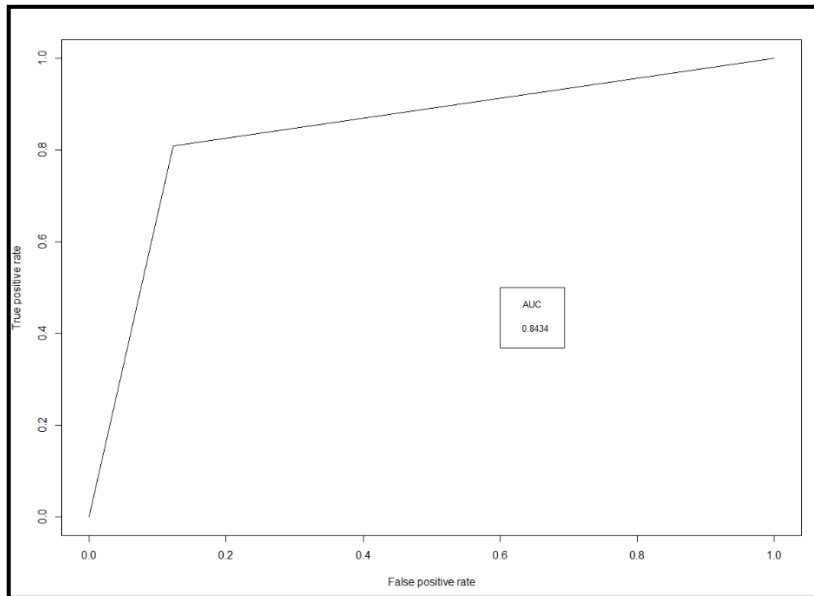
ANEXO 10.6. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A JUNIO



Fuente: Elaboración propia

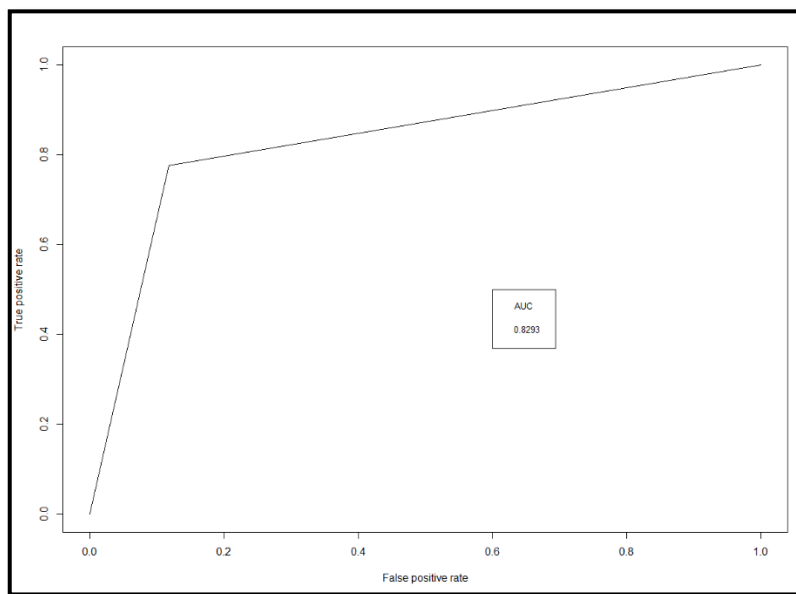


ANEXO 10.7. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A JULIO



Fuente: Elaboración propia

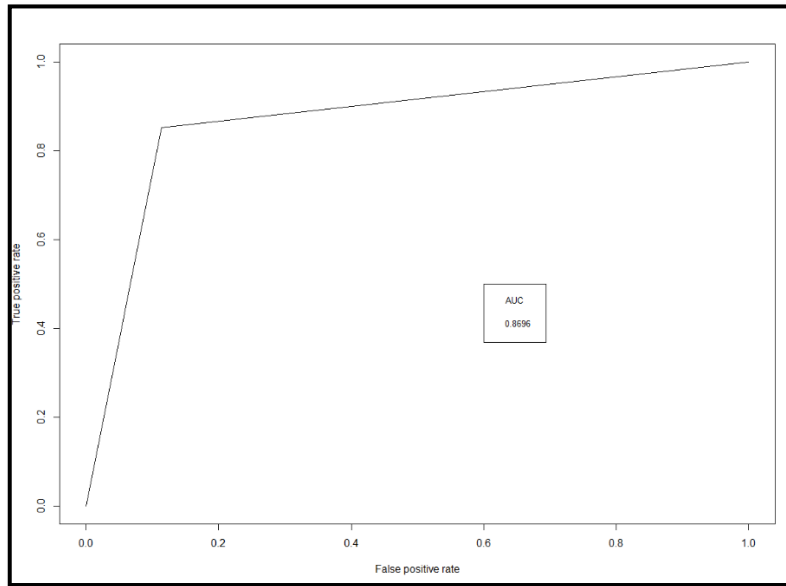
ANEXO 10.8. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A AGOSTO



Fuente: Elaboración propia

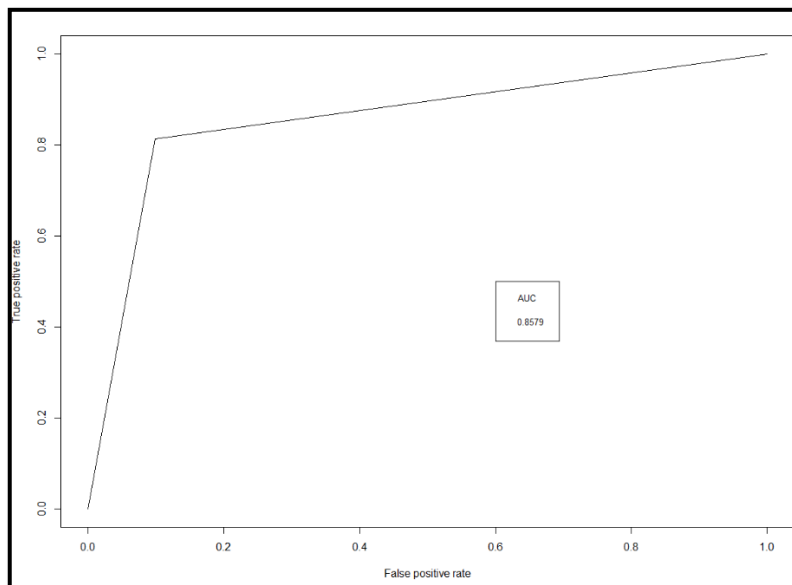


ANEXO 10.9. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A SEPTIEMBRE



Fuente: Elaboración propia

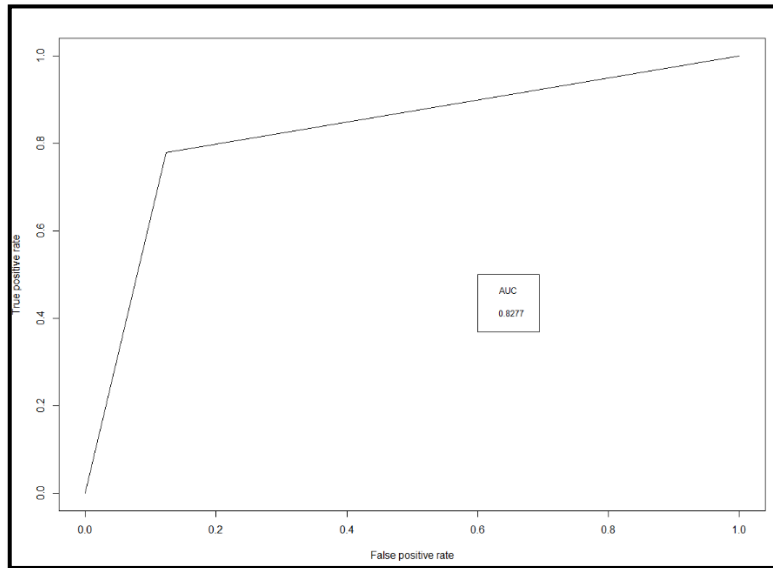
ANEXO 10.10. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A OCTUBRE



Fuente: Elaboración propia

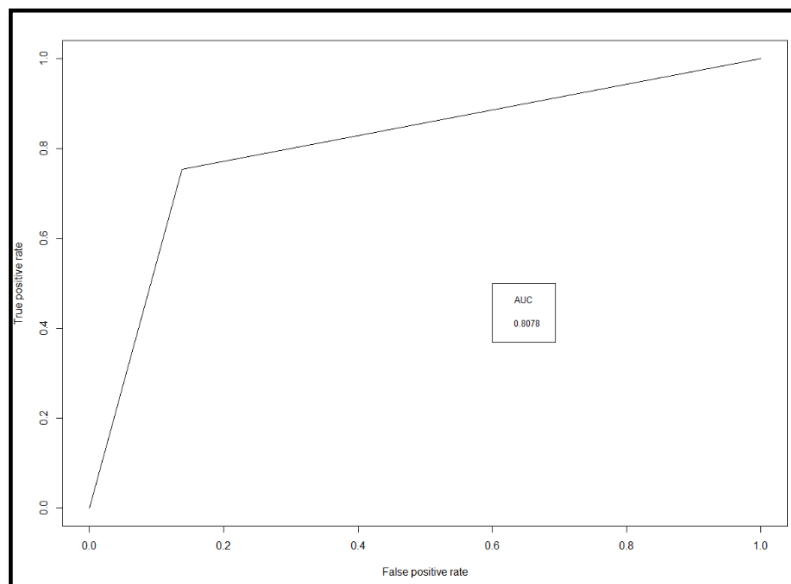


ANEXO 10.11. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A NOVIEMBRE



Fuente: Elaboración propia

ANEXO 10.12. GRÁFICO DE LA CURVA ROC SVM CORRESPONDIENTE A DICIEMBRE



Fuente: Elaboración propia



**ANEXO 11. INDICADORES DE EFICIENCIA DE LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL (SVM)
DETERMINADOS CADA MES DEL AÑO 2017**

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	PROMEDIO
EXACTITUD (%)	86.41	85.57	86.30	87.29	86.82	86.51	85.05	84.06	87.29	86.67	83.70	82.08	85.65
PRECISION (%)	81.84	82.44	83.14	83.54	82.95	83.07	81.24	81.05	83.54	84.71	81.12	80.97	82.47
SENSIBILIDAD (%)	84.32	80.81	82.82	85.27	84.14	82.96	80.92	77.62	85.27	81.40	77.89	74.72	81.51
ESPECIFICIDAD (%)	87.77	88.70	88.66	88.66	88.59	88.85	87.76	88.24	88.66	81.40	87.65	87.38	87.69
AUC	0.8605	0.8476	0.8574	0.8696	0.8637	0.8591	0.8434	0.8293	0.8696	0.8579	0.8277	0.8105	0.8497

Fuente: Elaboración propia



ANEXO 12. TIEMPO DE PRONÓSTICO DE MOROSIDAD DE CUOTAS SOCIALES CON LA TÉCNICA MÁQUINA DE SOPORTE VECTORIAL DETERMINADOS CADA MES DEL AÑO 2017

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	PROMEDIO
User time (seg)	11.87	11.51	10.18	10.85	10.87	11.03	11.51	11.29	11.00	13.34	13.78	15.45	11.89
System time(seg)	0.71	0.7	0.59	0.39	0.42	0.37	0.42	0.48	0.5	0.4	0.51	0.39	0.49
Elapsed time(seg)	69.45	76.79	49.92	65.17	53.01	45.12	49.00	49.21	54.11	53.93	48.46	62.21	56.37

Fuente: Elaboración propia



ANEXO 13. CÓDIGO EN RSTUDIO DE LA IMPLEMENTACIÓN DE REGRESIÓN LOGÍSTICA

```

#Cargar los datos
dataset <- read.csv("Enero_Algoritmo.csv", header = T)
dataset

testing <- read.csv("Enero_Predecir.csv")

#Cargar librerias
library(caTools)
library(caret)
library(ROCR)

#Datos de entrenamiento
set.seed(1234)
split <- sample.split(dataset$MOROSO, SplitRatio = 0.80)
training_set <- subset(dataset, split == TRUE)
test_set <- subset(dataset, split == FALSE)

#Comprobando distribucion
table(training_set$MOROSO)

table(test_set$MOROSO)

#Crear modelo
clasificadorRL <- glm(MOROSO ~ ., family = binomial, data = training_set)

summary(clasificadorRL) # ver resumen de modelo

#Matriz de confusion
pred_valid <- predict(clasificadorRL, type = 'response', newdata = test_set)
pred_valid <- ifelse(pred_valid > 0.5, 1, 0)
pred_valid <- factor(pred_valid, levels = c("0", "1"), labels = c("NO", "SI"))
pred_valid
matrizConfusion <- table(test_set$MOROSO, pred_valid)
matrizConfusion

confusionMatrix( test_set$MOROSO,pred_valid,positive='SI' )

```

```
##Prediccion con datos de testing
str(test_set)
str(testing)

prueba <- predict(clasificadorRL, testing,type = "response")
prueba <- ifelse(prueba > 0.5, 1, 0)
prueba <- factor(prueba, levels = c("0", "1"), labels = c("NO", "SI"))
prueba
head(prueba)
testing$MOROSO <- prueba
write.csv(testing,"Enero_datapredecidaRL.csv",row.names=F)
proc.time()

#Curva Roc
pred1 <- prediction(as.numeric(pred_valid), as.numeric(test_set$MOROSO))
perf1 <- performance(pred1, "tpr", "fpr")
plot(perf1)

#Area bajo la curva (AUC)
auc <- performance(pred1, "auc")
auc <- unlist(slot(auc, "y.values"))
auc <- round(auc, 4)
legend(.6, .5, auc, title = "AUC", cex = .8)
```

Fuente: Elaboración propia

ANEXO 14. FICHA DE OBSERVACIÓN REGRESION LOGÍSTICA.

ANEXO 14.1. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESION LOGÍSTICA CORRESPONDIENTE AL MES DE ENERO.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 1
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.82
 (seg.)
MES ANALIZADO Enero

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1030	131
SI	131	628

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
86.35	82.74	82.74	88.72	0.8573

Fuente: Elaboración propia.



ANEXO 14.2. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE FEBRERO.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 2
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.75
 (seg.)
MES ANALIZADO Febrero

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1023	136
SI	139	622

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
85.68	82.06	81.73	88.27	0.8500

Fuente: Elaboración propia.



ANEXO 14.3. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE MARZO.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 3
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.57
 (seg.)
MES ANALIZADO Marzo

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1021	125
SI	149	625

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
85.73	83.33	80.75	89.09	0.8492

Fuente: Elaboración propia.



ANEXO 14.4. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE ABRIL.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS	4
TIPO DE PRUEBA	Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN (seg.)	0.64
MES ANALIZADO	Abril

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1019	127
SI	126	648

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
86.82	83.61	83.72	88.92	0.8632

Fuente: Elaboración propia.



ANEXO 14.5. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE MAYO.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: **Responsable:** Jefferson Napoleón Huamán Bernilla.
Computador con procesador **Fecha:** 30 de octubre de 2018.
Corei5. **Institución:** Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 5
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.48
 (seg.)
MES ANALIZADO Mayo

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1025	132
SI	140	623

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
85.83	82.52	81.65	88.59	0.8512

Fuente: Elaboración propia.



ANEXO 14.6. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE JUNIO.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS	6
TIPO DE PRUEBA	Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN (seg.)	0.46
MES ANALIZADO	Junio

MATRIZ DE CONFUSIÓN

	NO	SI
NO	968	162
SI	190	600

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
81.67	78.74	75.95	85.66	0.8081

Fuente: Elaboración propia.



ANEXO 14.7. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE JULIO.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 7
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.4
 (seg.)
MES ANALIZADO Julio

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1019	141
SI	139	621

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
85.42	81.50	81.71	87.84	0.8478

Fuente: Elaboración propia.



ANEXO 14.8. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE AGOSTO.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: **Responsable:** Jefferson Napoleón Huamán Bernilla.
 Computador con procesador Corei5. **Fecha:** 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS	8
TIPO DE PRUEBA	Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN (seg.)	0.39
MES ANALIZADO	Agosto

MATRIZ DE CONFUSIÓN

	NO	SI
NO	966	164
SI	188	602

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
81.67	78.59	76.20	85.49	0.8084

Fuente: Elaboración propia.



ANEXO 14.9. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE SEPTIEMBRE.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 9
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.46
 (seg.)
MES ANALIZADO Septiembre

MATRIZ DE CONFUSIÓN

	NO	SI
NO	963	154
SI	198	605

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
81.67	79.71	75.34	86.21	0.8078

Fuente: Elaboración propia.



ANEXO 14.10. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE OCTUBRE.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS	10
TIPO DE PRUEBA	Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN (seg.)	0.35
MES ANALIZADO	Octubre

MATRIZ DE CONFUSIÓN

	NO	SI
NO	1033	118
SI	157	612

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
85.68	83.84	79.58	89.75	0.8467

Fuente: Elaboración propia.



ANEXO 14.11. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE NOVIEMBRE.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 11
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.5
 (seg.)
MES ANALIZADO Noviembre

MATRIZ DE CONFUSIÓN

	NO	SI
NO	992	150
SI	168	610

EVALUACIÓN DE MÉTRICAS

EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
83.44	80.26	78.41	86.87	0.8264

Fuente: Elaboración propia.



ANEXO 14.12. RESUMEN DE ANÁLISIS DE LA TÉCNICA REGRESIÓN LOGÍSTICA CORRESPONDIENTE AL MES DE DICIEMBRE.

**FICHA DE OBSERVACIÓN
REGRESIÓN LOGÍSTICA**

Computador: Computador con procesador Corei5.
Responsable: Jefferson Napoleón Huamán Bernilla.
Fecha: 30 de octubre de 2018.
Institución: Universidad Señor de Sipán.

RESUMEN DE ANÁLISIS

NÚMERO DE ANÁLISIS 12
TIPO DE PRUEBA Porcentaje Split (Training =80%, testing=20%)
TIEMPO DE PREDICCIÓN 0.4
 (seg.)
MES ANALIZADO Diciembre

MATRIZ DE CONFUSIÓN

	NO	SI
NO	963	154
SI	198	605

EVALUACIÓN DE MÉTRICAS

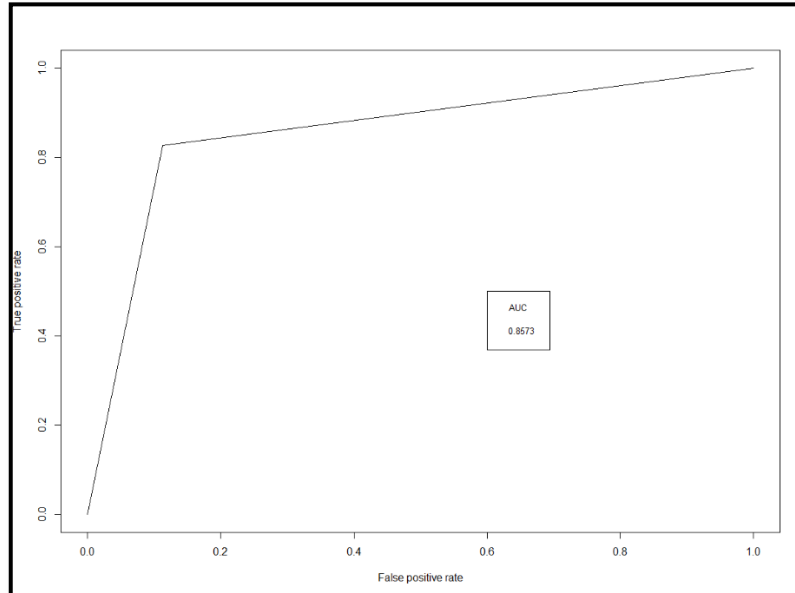
EXACTITUD	PRECISIÓN	SENSIBILIDAD	ESPECIFICIDAD	ROC(AUC)
81.67	79.71	75.34	86.21	0.8078

Fuente: Elaboración propia.



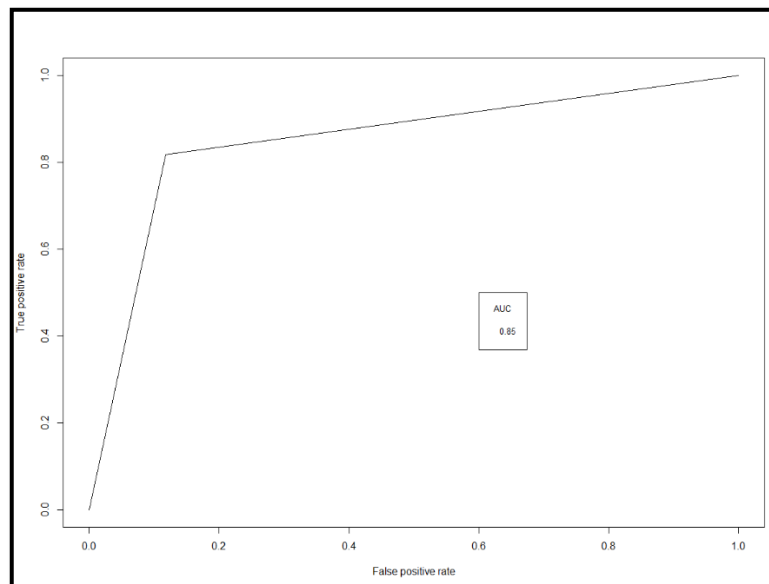
ANEXO 15. CURVA ROC REGRESIÓN LOGÍSTICA(RL).

ANEXO 15.1. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A ENERO



Fuente: Elaboración propia

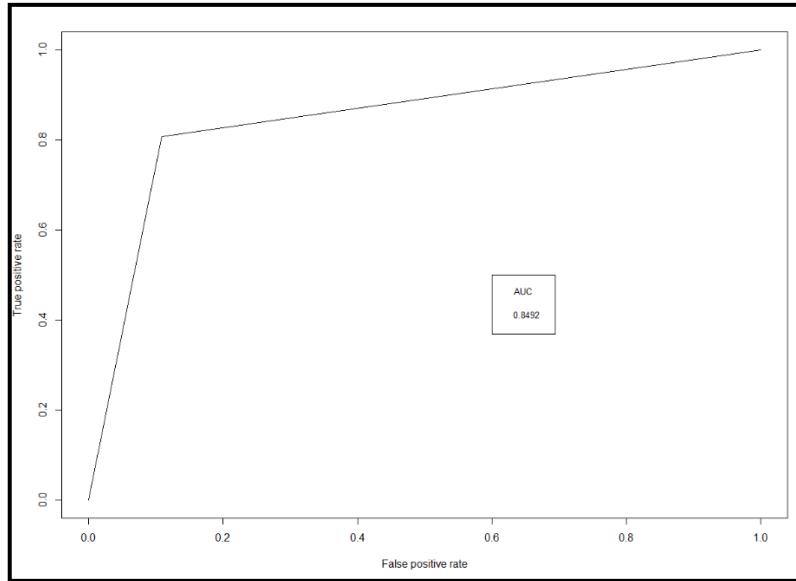
ANEXO 15.2. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A FEBRERO



Fuente: Elaboración propia

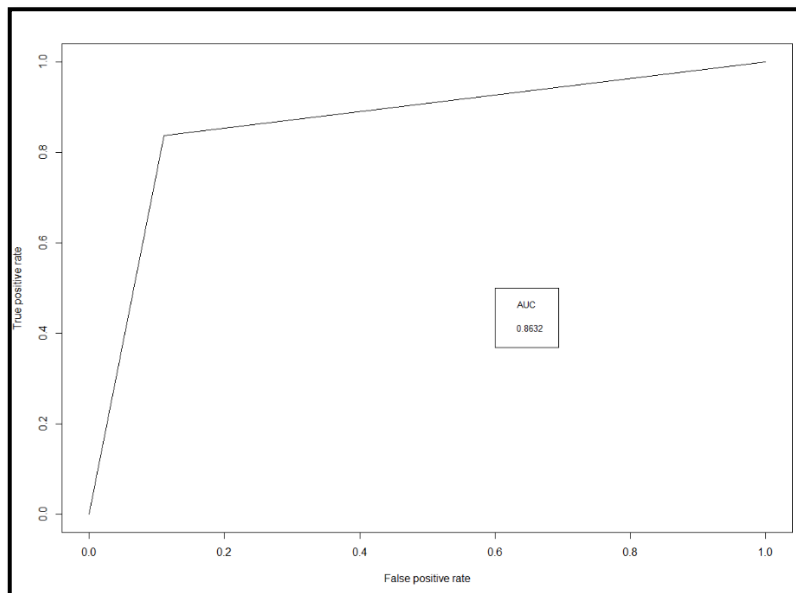


ANEXO 15.3. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A MARZO



Fuente: Elaboración propia

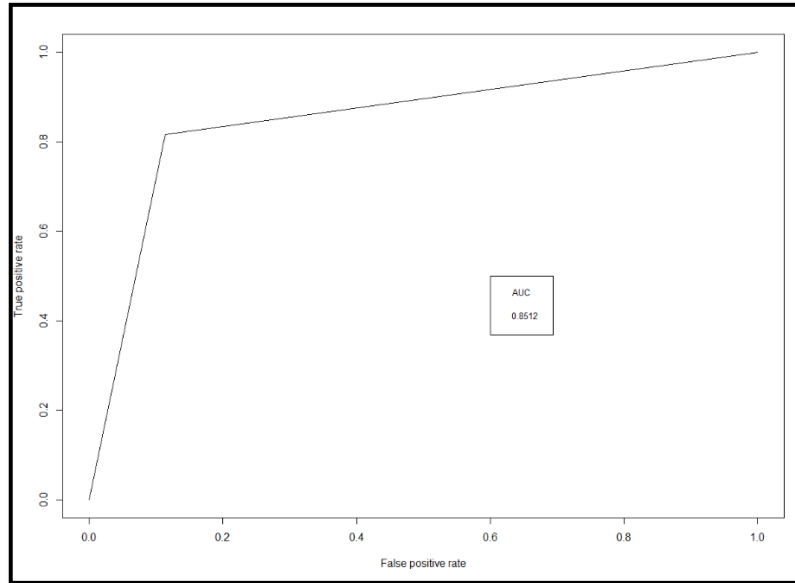
ANEXO 15.4. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A ABRIL



Fuente: Elaboración propia

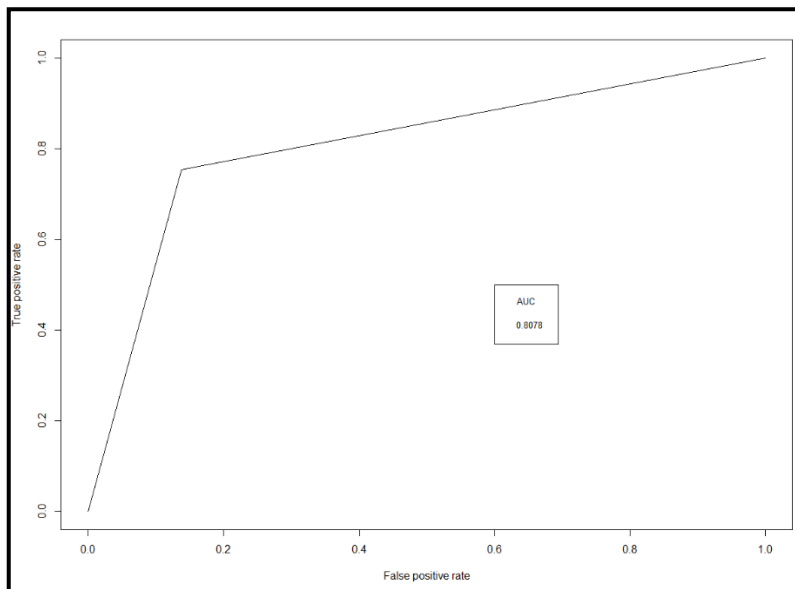


ANEXO 15.5. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A MAYO



Fuente: Elaboración propia

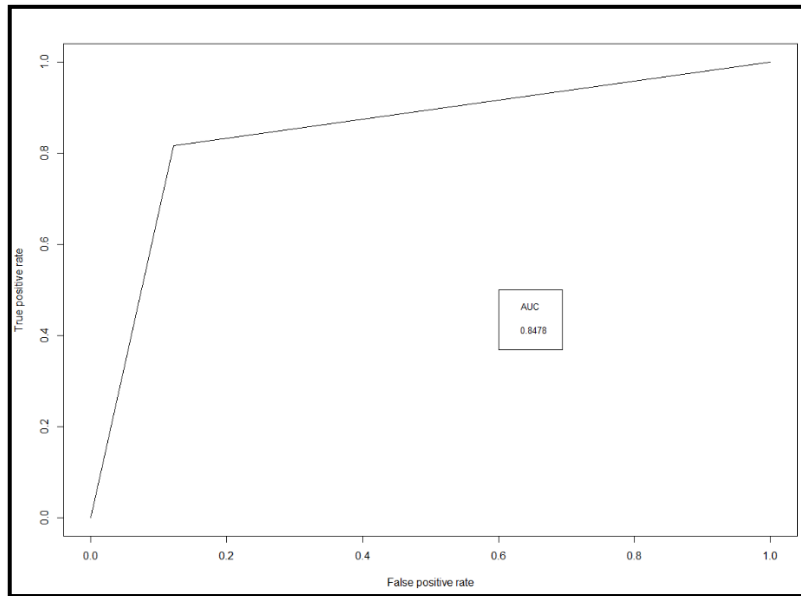
ANEXO 15.6. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A JUNIO



Fuente: Elaboración propia

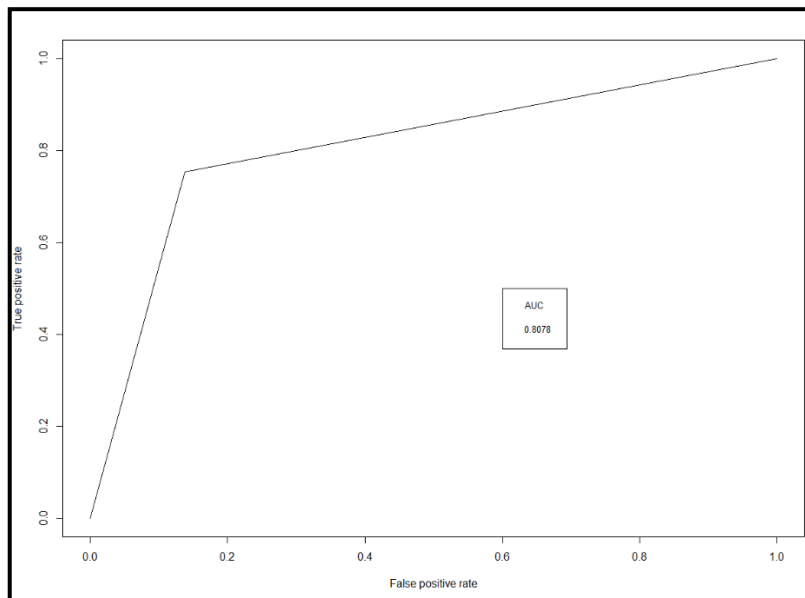


ANEXO 15.7. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A JULIO



Fuente: Elaboración propia

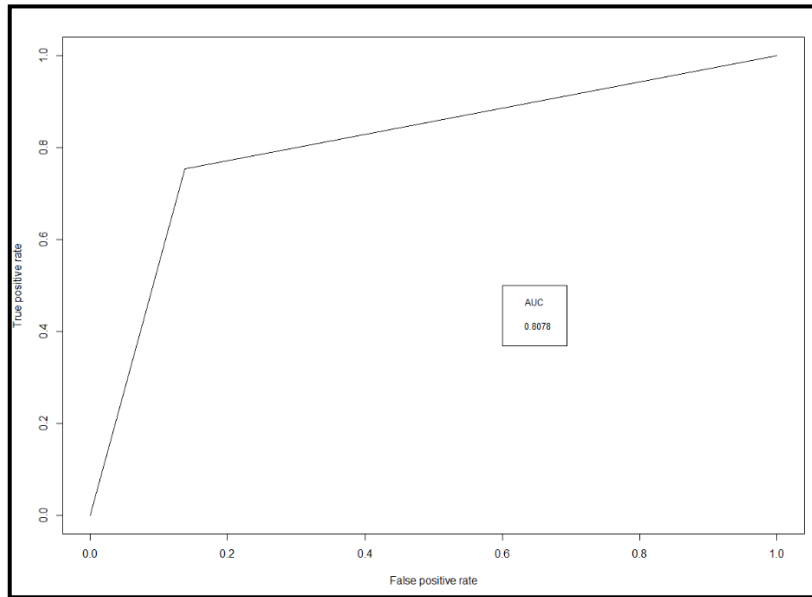
ANEXO 15.8. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A AGOSTO



Fuente: Elaboración propia

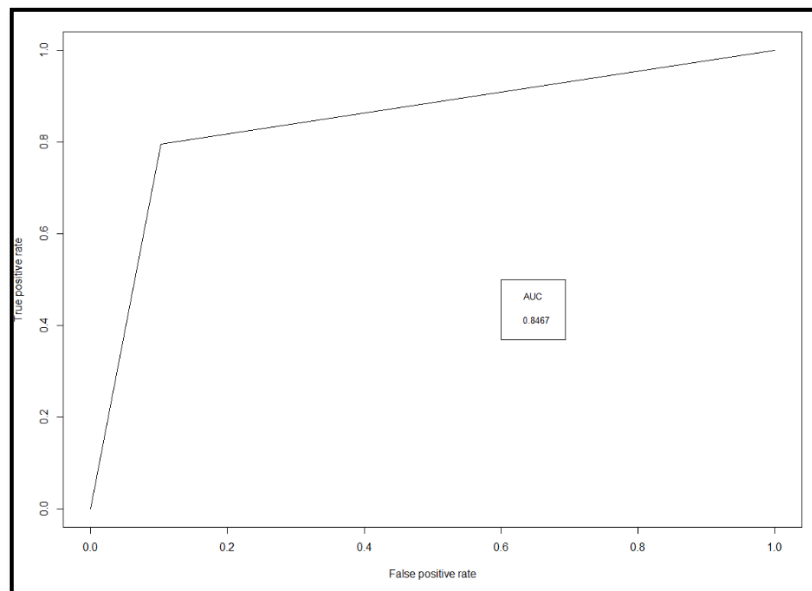


ANEXO 15.9. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A SEPTIEMBRE



Fuente: Elaboración propia

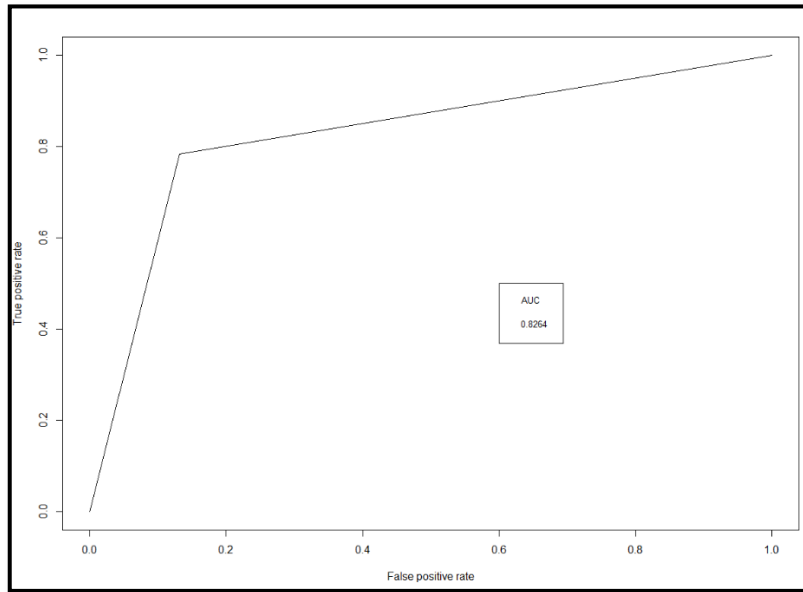
ANEXO 15.10. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A OCTUBRE



Fuente: Elaboración propia

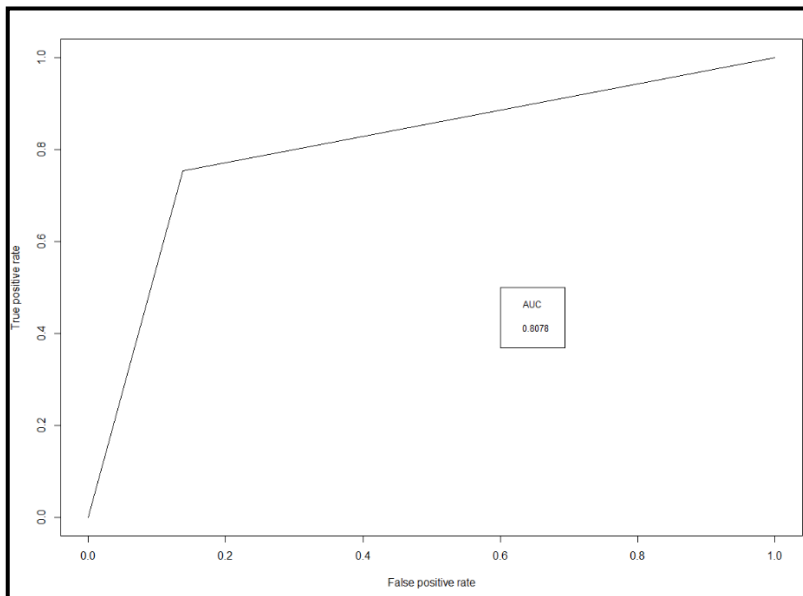


ANEXO 15.11. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A NOVIEMBRE



Fuente: Elaboración propia

ANEXO 15.12. GRÁFICO DE LA CURVA ROC RL CORRESPONDIENTE A DICIEMBRE



Fuente: Elaboración propia



**ANEXO 16. INDICADORES DE EFICIENCIA DE LA TÉCNICA REGRESIÓN LOGÍSTICA (RL) DETERMINADOS CADA
MES DEL AÑO 2017**

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	PROMEDIO
EXACTITUD (%)	86.35	85.68	85.73	86.82	85.83	81.67	85.42	81.67	81.67	85.68	83.44	81.67	84.30
PRECISION (%)	82.74	82.06	83.33	83.61	82.52	78.74	81.50	78.59	79.71	83.84	80.26	79.71	81.38
SENSIBILIDAD (%)	82.74	81.73	80.75	83.72	81.65	75.95	81.71	76.20	75.34	79.58	78.41	75.34	79.43
ESPECIFICIDAD (%)	88.72	88.27	89.09	88.92	88.59	85.66	87.84	85.49	86.21	89.75	86.87	86.21	87.63
AUC	0.8573	0.8500	0.8492	0.8632	0.8512	0.8081	0.8478	0.8084	0.8078	0.8467	0.8264	0.8078	0.8353

Fuente: Elaboración propia



**ANEXO 17. TIEMPO DE PRONÓSTICO DE MOROSIDAD DE CUOTAS SOCIALES CON LA TÉCNICA REGRESIÓN
LOGÍSTICA DETERMINADOS CADA MES DEL AÑO 2017**

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	PROMEDIO
User time (seg)	5.57	5.34	4.76	4.84	5.34	4.73	4.62	4.89	4.89	4.96	4.84	4.62	4.95
System time(seg)	0.82	0.75	0.57	0.64	0.48	0.46	0.4	0.39	0.46	0.35	0.5	0.4	0.52
Elapsed time(seg)	54.36	52.63	47.54	46.93	52.68	41.76	41.21	38.5	37.18	49.25	43.96	42.31	45.69

Fuente: Elaboración propia