

Octubre 2022

CODERHOUSE

INFORME PROYECTO



Finanzas Panem

TE PRESTAMOS LO QUE TE HACE FALTA



ÍNDICE

• Información de la empresa	1
• Equipo nueva área BI	1
• Evaluación de la empresa	2
• Las oficinas	2
• Análisis FODA	3
• Principales competidores	3
• Plan de acción	4
• Cliente ideal	4
• Definición de estrategias	5
• IA en cobranzas	5
• Pasos desarrollados	6
• Análisis EDA	6
• Prospectivas	7
• Situación laboral	7
• Elección del modelo	8
• Cronología del proyecto	9
• Conclusiones	10



INFORMACIÓN DE LA EMPRESA

Finanzas Panen es una empresa especializada en otorgar créditos prendarios. Desde 1999 se concentra en contribuir al mercado automotor, facilitando a los usuarios la obtención fácil y sencilla a su próximo vehículo.

Cuenta con oficinas comerciales en Martínez, provincia de Buenos Aires y comercializadores en distintos puntos del país que trabajan en forma remota.

Durante el último año y posterior a los hechos acontecidos por la pandemia, Finanzas Panem ha decidido migrar toda su operación a modalidad de trabajo remoto, lo que supuso una reingeniería en los procesos y herramientas actuales, tendiendo a la automatización y al trabajo colaborativo de todas las áreas así como la incorporación de un nuevo departamento o área BI (Business Intelligence) orientado a la inteligencia de datos y estrategia del negocio de colocación de créditos así como el recupero de mora.



EQUIPO NUEVA AREA BI



LUCIA FILGUEIRAS, VISUALIZACION DE DATOS MAPLOLIP

Encargada de los procesos de visualización de los datos y enfoque en la estandarización de los datos.



MARCELO OCAMPO, DATA ANALITICS

Enfoque en el análisis de los datos y performance de los modelos, encargado en la optimización del código. Especializado en Python.



KAREN TAMARIN, ANALISTA DE COBRANZAS

Especialista en gestión de recupero y mora, con experiencia en procesos de mejora de continua en Call center y atención al cliente.



EVALUACIÓN DE LA EMPRESA

MISIÓN

La misión de Finanzas Panem es mejorar los procesos diarios e internos lo que contribuirá a mejorar la atención de nuestros clientes. Es una misión simple, sencilla y sin pretensiones de ser una acción de marketing, simplemente se orienta a propiciar un lugar y clima de trabajo donde cada uno de los que participan encuentren un ámbito profesional y humano que nos reciba diariamente como un camino de desafíos profesionales y que sea la empresa el gran complemento para nuestras relaciones familiares, sociales y humanas.

VISIÓN

Estamos convencidos que el avance tecnológico seguirá siendo el “driver” del desarrollo de cada empresa. Es por ello que se apuesta al trabajo remoto, pero sin perder el contacto humano. En función de ello, consideramos que Finanzas Panem debe ser visualizada como una empresa con atención personalizada constante y pendiente de las necesidades tanto de los tomadores de crédito (los clientes) como de aquellos que son fundamentales en el crecimiento de la compañía; concesionarios, comerciales, agencias, proponiéndonos a desarrollar una empresa que sea la opción más simple y útil para la plaza automotor.

VALORES

Finanzas Panem pone el foco en la gestión personalizada y orientada al cliente. Entendemos que las personas que interrumpen el pago de las cuotas de su crédito deben tener un asesoramiento constante y el seguimiento necesario para coadyuvar a evitar la mora. El por ello que nuestra área de cobranzas, no solo se ocupa de la regularización y el cobro, sino también de la Bienvenida Inicial al cliente que contribuye a orientarlo acerca de su crédito, orientarlo con los medios de pago y asesorarlo en todo el trayecto de vida del crédito prendario hasta la cancelación total del mismo.



LAS OFICINAS

Transformación 2021-2022

A partir del 2022 se efectuó la migración de la gestión operativa a trabajo cien por cien remoto, en conjunto con la remodelación de las oficinas comerciales para la reincorporación de las áreas de atención al cliente, finanzas y seguros.

El resto de la compañía trabaja de manera remota a través de la plataforma integral Bitrix, que permite la conexión total de todos los colaboradores a lo largo y ancho del país.



ANÁLISIS FODA

DEBILIDADES

- Alta demanda y exceso de información que contribuye a grandes volúmenes de datos poco organizados.
- Poca segmentación y clasificación eficiente según las necesidades del negocio prendario.
- Datos demográficos dispersos.

AMENAZAS

- Incremento en gastos legales por crecimiento en ejecución de la prenda.
- Poca contactabilidad por falta de datos válidos.
- Ruidos por exceso o falta de información en campos claves para el scoring crediticio (incremento morosidad).

FORTALEZAS

- Buena red de comercializadores a lo largo del país.
- Trato personalizado que destaca como diferencial de atención.
- Créditos a tasas preferenciales y a sola firma.

OPORTUNIDADES

- Desarrollo de un área de inteligencia de datos y BI.
- Ser vanguardia en la incorporación de modelos predictivos que constituyen en una ventaja competitiva respecto de los competidores.

PRINCIPALES COMPETIDORES

- AUTOCRÉDITO
- CARFÁCIL

Ambos se caracterizan por la una agresiva fuerza de ventas que se extiende principalmente en el AMBA. Sin embargo la pandemia ha producido una merma en el esquema de trabajo de estas empresa, lo que ha favorecido al fortalecimiento de la estrategia y transformación a la omnicanalidad adoptada por Finanzas Panem.



PLAN DE ACCIÓN

ESTANDARIZACION DE PROCESOS

Consenso y ajuste de las etapas de los procesos dentro de todas las áreas de Finanzas Panem, para que éstos se asemejen a un modelo en común. Es la metodología indicada para organizar la rutina a través de pasos estandarizados, seguidos por todos los empleados y conocidos también por todos.

COLOCACIÓN DE LOS CRÉDITOS

Colocación de los créditos, mediante una eficiente segmentación demográfica, en base al scoring crediticio desarrollado mediante modelos predictivos que reporten el menor margen de error.

SEGUIMIENTO DE LA MOROSIDAD

A la vez de optimizar las herramientas como Tableros de Control, Canales de Comunicación, Personal de Cobranzas idóneo y un CRM de gestión integral, la manipulación de los datos será clave a la hora de realizar el seguimiento de la cartera morosa.

A través del desarrollo y análisis de un modelo predictivo enfocado en la evolución de la morosidad, en sus diferentes estadios, es que lograremos posibilidad soluciones eficientes para que la misma no se incremente mas allá de niveles aceptables y que permitan hacer crecer el negocio.

FIDELIZACIÓN

Una correcta clasificación de la cartera de préstamos, nos permitirá mejorar la fidelización tanto para los tomadores de créditos, los comercializadores partners y los productores de seguros, que procure una sinergia total mediante el flujo correcto de la información.

CLIENTE IDEAL

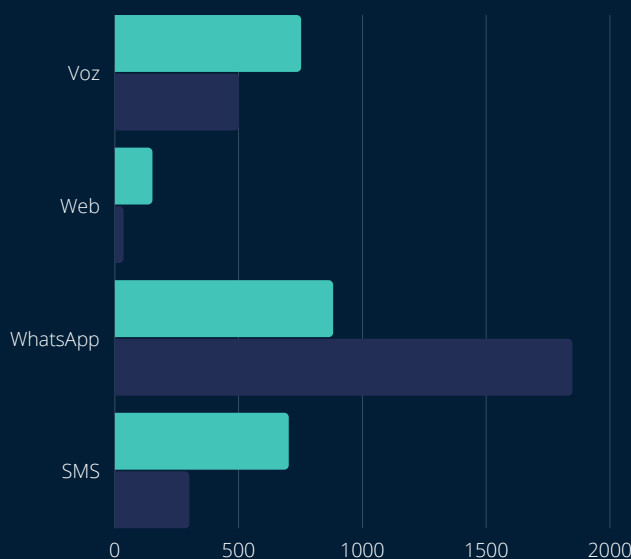


Hombre, 30-45 años

- Profesión: Padre de familia o sostén del hogar, trabajador en relación de dependencia
- Clase: Media
- Ubicación: Buenos Aires y alrededores
- Educación: Secundaria
- Estado civil: Casado o soltero
- Valores: Entusiasta de los vehículos automotores, sin deudas, embargos o créditos en mora.
- Intereses: Persona activa, siempre en movimiento.

DEFINICIÓN DE ESTRATEGIAS

En la actualidad, la aplicación de modelos de Machine Learning e Inteligencia Artificial (IA) se usan de forma incipiente en el negocio de la cobranza, aunque hay un gran potencial de crecimiento en los próximos años. Uno de los puntos más especializados de la inteligencia artificial es el contacto con los usuarios en los Contact Centers orientados a la cobranza de deuda. Los sistemas tienen acceso a las grandes bases de datos y recogen información de cada interacción; además, pueden hacer contacto por diversos canales como páginas web, chat, llamadas de voz o correo electrónico, entre otras



En el área operativa de Atención al Cliente y Cobranzas de Finanzas Panem se cuenta con grandes volúmenes de interacciones a través de diversas plataformas de contacto como Infobip, Telefonía VoIP, Web y Comercializadores.

Estas comunicaciones que son tanto Inbound como Outbound, son una gran oportunidad de mejora para la nueva área de BI que la empresa está actualmente desarrollando.

IA EN COBRANZAS



PERSONALIZACION POR MEDIO DE BBDD

Los sistemas de inteligencia artificial tienen acceso inmediato a las grandes bases de datos sobre clientes y usuarios, algo que por su velocidad de procesamiento solo pueden realizar las máquinas.



USO DE BIG DATA

Al almacenar todas las actividades del usuario en la red se pueden combinar con los datos recogidos en interacciones anteriores.

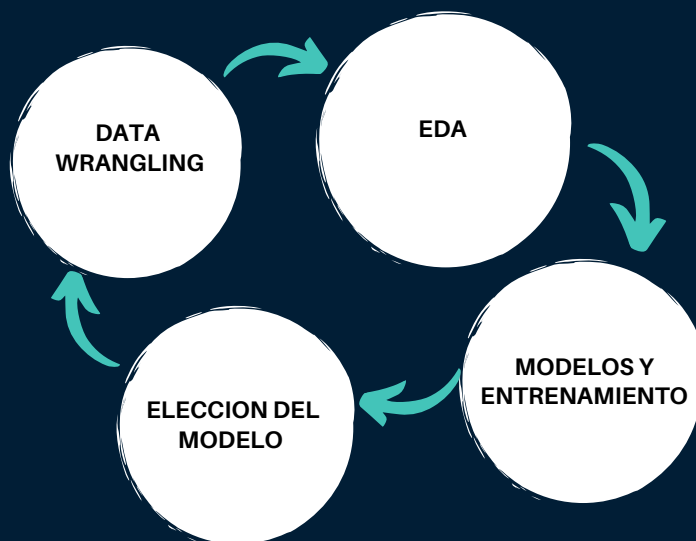


GENERACIÓN DE MODELOS PREDICTIVOS

Las máquinas inteligentes pueden analizar las interacciones con los clientes y deudores para establecer patrones de conducta y generar modelos predictivos que orienten sus propias decisiones.

➤➤➤➤ PASOS DESARROLLADOS

La prosperidad y el avance de las organizaciones dependen en gran medida de la forma en que analizan la información de sus clientes. Este panorama se puede agravar si se le suma información algunas veces desconocida: las tendencias del mercado, el panorama competitivo o los cambios en el comportamiento de los consumidores. El uso de soluciones analíticas dinámicas, construidas para obtener ventajas competitivas, puede discernir las situaciones rápidamente y ajustarlas según sea necesario para tomar decisiones más concretas. Es por ello que en este proyecto nos basamos en (4) tareas bien delimitadas en pos de producir el mejor modelo predictivo de análisis.



➤➤➤➤ ANÁLISIS EDA

En Data Science, el Análisis Exploratorio de Datos o en inglés Exploratory Data Analysis (EDA) es un método de análisis de los conjuntos de datos para resumir sus principales características. El EDA se utiliza para ver lo que los datos pueden decirnos antes de la tarea de modelación. Dado que no es una tarea sencilla, por el volumen de información o tipo de datos encontrados en la Base de Datos, determinamos para las siguientes técnicas los siguientes aspectos.

DATOS UNIVARIANTES (UNIVARIATE DATA)

Se desarrolla un análisis univariado con tres variables **Año del Auto**, **Marca de Auto**, **Situación Laboral**.

DATOS BIVARIADOS (BIVARIATE DATA)

Año y Valor Actual del auto según **Estado Civil**

DATOS MULTIVARIADOS (MULTIVARIATE DATA)

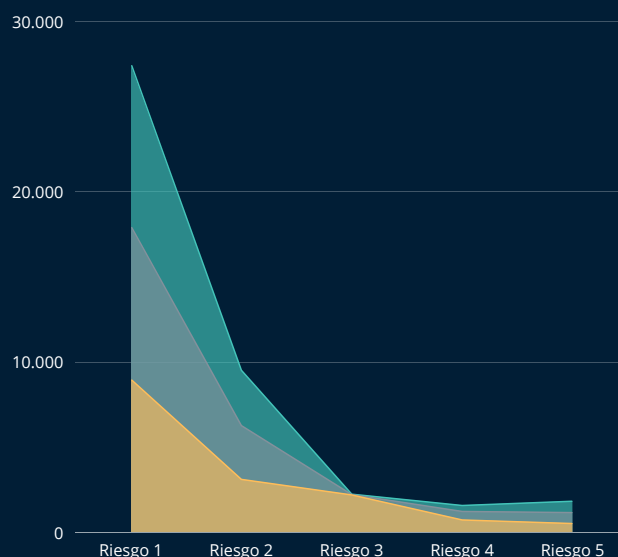
Valor Actual del Auto y Score del Cliente según **Estado Civil**. En el análisis se prueban diferentes variables combinándolas de manera tal de poder comprender la incidencia demográfica en la base de clientes de Finanzas Panem.

PROSPECTIVAS

Los autos mas vendidos se corresponden con los años 2006 y 2005, siendo Chevrolet y VW de los más elegidos en cuanto a la compra a crédito, y la mayoría de las personas que acceden a los prestamos se encuentran con trabajo en blanco, o relación de dependencia.

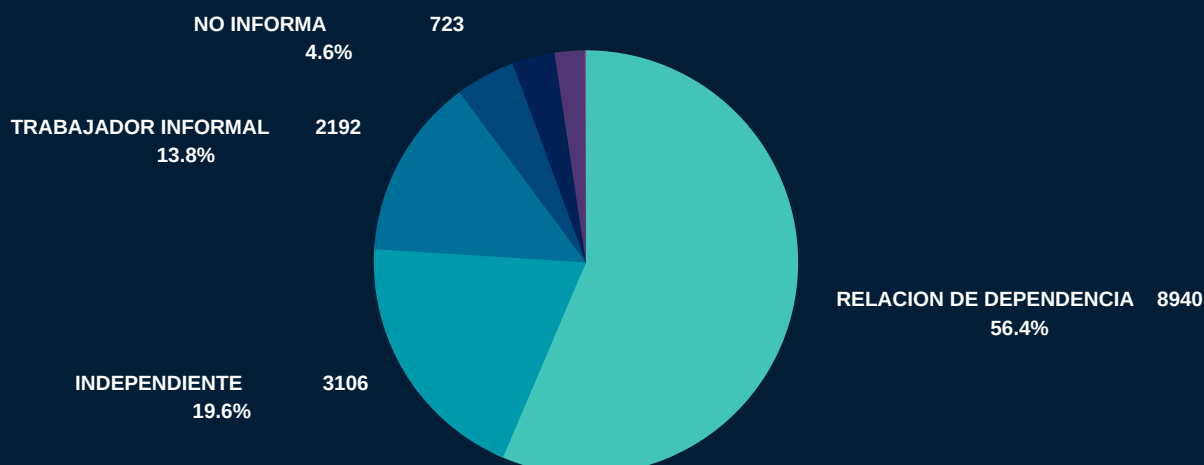
Se analizaron aspectos como la soltería, como rasgo en la mayor valoración de los autos tomados a crédito y menor riesgo crediticio.

- Menos de 619 puntos: Riesgo alto
- Entre 620 y 649 puntos: Riesgo medio
- Entre 650 y 749 puntos: Riesgo bajo
- Más de 750 puntos: Riesgo nulo



SITUACIÓN LABORAL

Respecto de la Situación Laboral, luego de la depuración de los datos, se pudo observar que es fundamental estandarizar la captura de los mismos a los fines de reducir las clasificaciones a solo siete (7) categorías. A su vez se pudo determinar que la mayoría de las personas (más de la mitad de la cartera) que acceden a los prestamos, estan en Relación de Dependencia, por lo tanto la situación en cuanto a ingresos formales, es un aspecto más que relevante a la hora de valorar el portafolio de mora, orientados en la cobranza preventiva y el scoring positivo del tomador de crédito.





ELECCIÓN DEL MODELO

Las nuevas herramientas de Big Data y Machine Learning habilitan las empresas de recobro a optimizar sus campañas a través de los Call Centers. Los datos predicen que para la compañía supone un aumento en los ingresos del 30 % y un ahorro de costes de operación del 25 %; además, reducen sustancialmente los litigios por morosidad, el último recurso de las compañías. A los fines del presente informe, presentaremos los modelos que hemos entrenado y probado, para luego escoger el mas eficiente y con mejores resultados. Con toda la materia prima recopilada –los datos– es posible desarrollar modelos de Machine Learning que sirvan como ayuda para mejorar la cobranza. Como primera iteración de esta metodología se hicieron prueba con tres modelos diferentes.

DESICION TREE

Algoritmo de aprendizaje supervisado no paramétrico, que se utiliza tanto para tareas de clasificación como de regresión. Tiene una estructura de árbol jerárquica, que consta de un nodo raíz, ramas, nodos internos y nodos hoja.

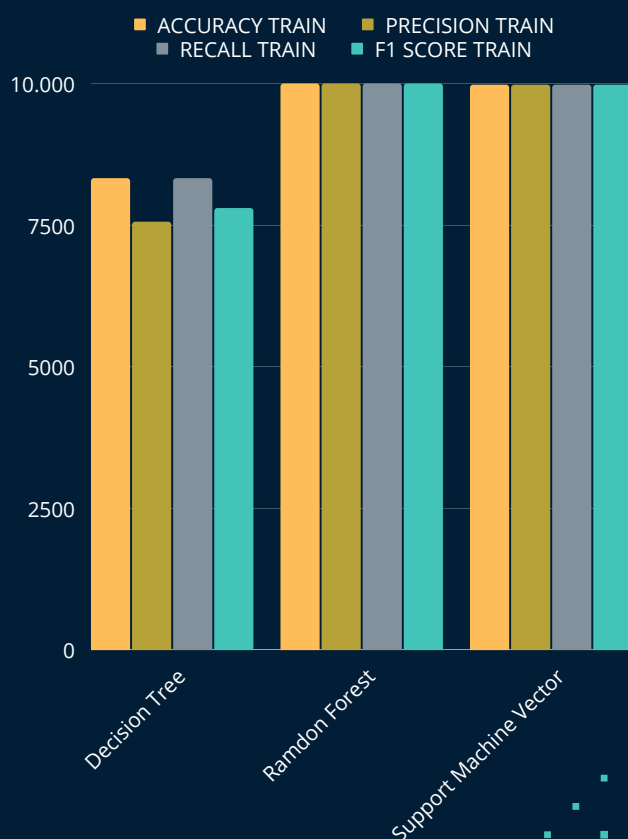
RANDOM FOREST

Modelo o técnica de aprendizaje automático muy popular, definido como conjunto de árboles. Los Random Forests tienen una capacidad de generalización muy alta para muchos problemas.

SUPPORT VECTOR MACHINE (SVR)

Es un modelo de clasificación basado en minimizar el error cuadrático de la clasificación, construyendo un hiperplano que separa los datos de la forma más precisa posible. SVM es considerado uno de los modelos más precisos y robustos posibles dentro de los algoritmos de clasificación binaria.

- **Precision (Precisión):** Con la métrica de precisión podemos medir la calidad del modelo de Machine Learning en tareas de clasificación.
- **Recall (Exhaustividad) :** La métrica de exhaustividad nos va a informar sobre la **cantidad** que el modelo de machine learning es capaz de identificar.
- **F1 :** El valor F1 se utiliza para combinar las medidas de precision y recall en un sólo valor. Esto es práctico porque hace más fácil el poder comparar el rendimiento combinado de la precisión y la exhaustividad entre varias soluciones.





CRONOLOGÍA DEL PROYECTO

PRIMER BIMESTRE

- Recuento histórico de los avances que llevaron al desarrollo de las nuevas metodologías que se aplicarán en el proyecto.
- Estudio de los conceptos básicos relacionados con Big Data, Inteligencia Artificial y Machine Learning.
- Estudio de los conceptos básicos usados en el entorno de recuperación de cartera.
- Revisión de los métodos usados por otras compañías del mismo sector para la optimización de sus procesos.

SEGUNDO BIMESTRE

- Se determina que el Excel no permite centralizar toda la información debido a que solo admite un número máximo de registros, además no permite hacer análisis estadístico avanzado.
- Excel es un formato que no es compatible con los lenguajes de programación que se van a usar para programar los modelos de optimización y de Machine Learning.
- La actualización de las bases de datos no se puede hacer de forma óptima y se requiere un trabajo manual voluminoso a principios de cada mes, es por ello que se proponen otras herramientas de captura de datos.

TERCER BIMESTRE

Durante este periodo se trabajó de manera exhaustiva, involucrándonos más a fondo en el negocio de la cobranza para consensuar mejores prácticas en pos de la recopilación de los datos de manera más eficiente:

- Exploración de los datos de forma completa.
- Limpieza de los datos.
- Arreglar los valores faltantes.
- Unificar los formatos, estandarización
- Analizar posibles sesgos en los datos.
- Comprobar la integridad de los datos.

CUARTO BIMESTRE

- Hacer un estudio de la estructura interna de los datos.
- Buscar bases de datos que puedan brindar información adicional de los clientes anteriores.
- Definir nuevas variables para un estudio óptimo e implementar modelos de clasificación que permitan hacer las predicciones adecuadas.
- Realizar un proceso de validación cruzada de los modelos para garantizar la predictibilidad.

Abril 2022

Preparación de la información

Mayo 2022

Data Wrangling

Junio 2022

Análisis Exploratorio de Datos

Julio 2022

Testeo y entrenamiento de los modelos

Agosto 2022

Entrega Preliminar 1

Septiembre 2022

Entrega Preliminar 2

Octubre 2022

Entrega Preliminar 3 - Final

Noviembre 2023

Puesta en marcha del modelo seleccionado.

Diciembre 2023

Seguimiento y Feedback



CONCLUSIONES

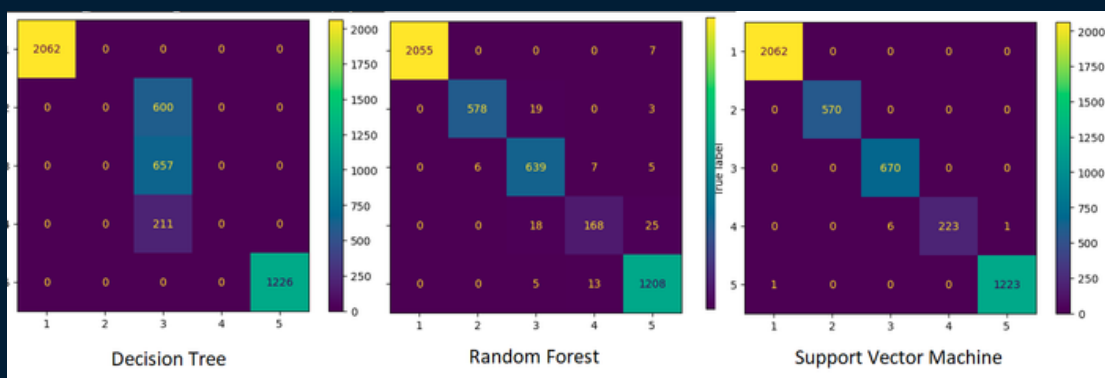
Una vez definidos los modelos y realizado los entrenamientos de los mismos, podemos llegar a concluir que:

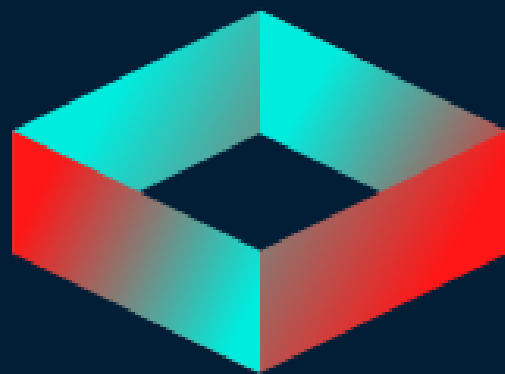
- El modelo que más errores tuvo en cuanto a clasificación, fue **Decision Tree**, donde hubo dos categorías que clasificó mal por completo; las categorías 2 (Riesgo Nulo) y 4 (Riesgo medio).
- **Random Forest** se comportó de muy buena forma, clasificando en gran parte de forma correcta, con 108 clasificaciones incorrectas dentro de todas las categorías definidas.
- El modelo con mejor comportamiento, fue **Support Vector Machine (SVM)** con muy buena clasificación, dando solo 8 registros con error dentro de todas las categorías.

Todas estas apreciaciones, se pueden ver con más detenimiento en la [*Matriz de Confusión*](#) de cada modelo que de forma más gráfica.

Es por ello que determinamos seleccionar el modelo con menores errores para la clasificación de clientes con riesgo, ya que una mala clasificación es un gran riesgo para el ciclo de vida del negocio de Créditos y Cobranzas. El modelo **Support Vector Machine** como recomendado a tomar en cuenta en la instalación de esta nueva funcionalidad para Finanzas Panem.

MATRIZ DE CONFUSIÓN COMPARATIVA





Finanzas Panem

TE PRESTAMOS LO QUE TE HACE FALTA