

UNIVERSIDAD CENTROAMERICANA “JOSÉ SIMEÓN CAÑAS”
FACULTAD DE INGENIERÍA Y ARQUITECTURA

Fundamentos de inteligencia de negocios

Ciclo 01/2025



Trabajo Ex-Aula: Período 3

Integración de Minería de Datos

Grupo: “Gaby y sus amigos”

Integrantes:

Menjivar Ayala, Oscar Ernesto	00068422
Mijango Serrano, Fabio Alberto	00156022
Mejia Oliva, David Ernesto	00004722
Quinteros Ramirez, Gabriela Sofia	00060422
Torres Campos, Miguel Andreé	00015722
Umanzor Velasquez, Rodrigo	00087022

Antiguo Cuscatlan, El Salvador

1. Resumen Ejecutivo

Este informe presenta los principales hallazgos obtenidos a través del análisis de minería de datos en una entidad bancaria. A partir del análisis de datos de clientes, cuentas, tarjetas, riesgo crediticio y comportamiento digital, se identificaron patrones de riesgo, segmentaciones de clientes y factores que permiten predecir el estado de mora. Estos resultados proporcionan información clave para optimizar estrategias de prevención de riesgo, fidelización y personalización de servicios bancarios.

2. Objetivo del Análisis

El objetivo del análisis fue descubrir patrones y relaciones relevantes entre los clientes del banco, su comportamiento financiero y su perfil de riesgo, respondiendo a preguntas como:

- ¿Qué perfiles de cliente presentan mayor riesgo de caer en mora?
- ¿Qué factores predicen mejor el estado de mora?
- ¿Cómo se relaciona el uso de canales digitales con los perfiles de riesgo?

3. Hallazgos Clave

a) Segmentación de Clientes

Mediante el uso de clustering (KMeans), se identificaron tres grupos distintos de clientes según su Score y Días en Mora:

- Cluster 0: Clientes con bajo score y alta mora.
- Cluster 1: Clientes con score medio y mora leve.
- Cluster 2: Clientes con alto score y sin mora.

Esto permite establecer estrategias diferenciadas para cada grupo, como campañas de prevención o recompensas por buen comportamiento financiero.

b) Predicción de Estado de Mora

Se entrenó un modelo de Random Forest para predecir el estado de mora utilizando como variables predictoras el Score y los Días en Mora. El modelo alcanzó una alta precisión, demostrando que es posible anticipar con fiabilidad qué clientes podrían entrar en mora. Esto habilita la generación de alertas preventivas.

c) Uso de Canales Digitales

El análisis de reglas de asociación reveló un fuerte uso conjunto de la App y la Web por parte de los clientes. Una de las reglas más frecuentes fue: 'Si el cliente usa App, probablemente también use Web'. Esto sugiere oportunidades para fortalecer la experiencia omnicanal y dirigir campañas cruzadas.

d) Análisis de Edad y Score

Se identificó que los clientes entre 30 y 45 años tienden a tener mejores scores crediticios. Además, se detectaron algunos valores atípicos (outliers) con scores extremadamente altos.

o bajos, lo que podría indicar errores, fraudes o perfiles excepcionales que merecen revisión manual.

e) Correlación entre Variables

Se encontró una correlación inversa entre el Score y los Días en Mora. A menor Score, mayor es la cantidad de días en mora, lo cual valida que el score es una variable confiable para evaluar el riesgo del cliente.

4. Conclusiones Generales

El análisis permitió establecer una segmentación clara de los clientes bancarios según su perfil de riesgo, evidenciando que el Score y los Días en Mora son determinantes clave para anticipar la mora. También se confirmó que el comportamiento digital se relaciona con dichos perfiles, abriendo la posibilidad de utilizar estos patrones para personalizar estrategias y servicios.

5. Recomendaciones

- Implementar un sistema de alertas tempranas para clientes del Cluster 0 (alto riesgo).
- Usar modelos predictivos para apoyar decisiones de crédito y cobranza.
- Diseñar campañas de fidelización digital basadas en el uso combinado de App y Web.
- Revisar de forma manual a los clientes con scores fuera del rango habitual.
- Continuar fortaleciendo la integración de datos para un análisis más detallado por canal, producto y sucursal.