

# Idea de Investigación

Kevin Pérez García

UNI

10 de abril de 2024

# Sector

De acuerdo con Torres (2020)<sup>1</sup>, los casos de aplicación de Data Science en distintas industrias pueden ser:

- ▶ Servicios de Salud
- ▶ Finanzas
  - ▶ Optimización de carteras de inversión
  - ▶ Detección de fraudes
  - ▶ Predicción del tipo de cambio de divisas
  - ▶ Predicción de quiebra empresarial
  - ▶ Predicción del precio del oro
- ▶ Transporte
- ▶ E-commerce
  - ▶ Predicción de ventas
  - ▶ Análisis de opiniones o reviews de productos
  - ▶ Promociones personalizadas
  - ▶ Sistemas de recomendación de productos

# Aplicaciones de ciencia de datos en el sector financiero

De acuerdo con Ingenium (2022)<sup>2</sup>, las aplicaciones de ciencia de datos en el sector financiero pueden ser:

- ▶ Optimización de cartera
- ▶ Detección de fraude
- ▶ Modelos de valoración de activos y derivados
- ▶ Modelo predictivo de mercado
- ▶ Credit scoring

# Contexto

De acuerdo con SBS (2023)<sup>3</sup>,

- ▶ El 06 de enero del 2023, la SBS<sup>4</sup> aprobó el Reglamento de Gestión de Riesgos de Modelo.
- ▶ Este reglamento establece lineamientos mínimos para mitigar los riesgos derivados del uso de modelos empleados en la gestión de riesgos de crédito, mercado, liquidez, operacional, y lavado de activos y del financiamiento del terrorismo de las empresas del sistema financiero y de seguros.
- ▶ Su aplicación está siendo impuesta desde diciembre de 2023 hasta enero de 2026.

## Contexto

El inciso 2 del Artículo 16 de dicho reglamento menciona textualmente:

En modelos contruidos utilizando técnicas de inteligencia artificial, las empresas deben, por lo menos, realizar pruebas u otros análisis que permitan evaluar la interpretabilidad del modelo (entender el proceso de toma de decisiones y resultados del modelo), realizar la optimización de hiperparámetros como parte del proceso de aprendizaje del modelo; y realizar validaciones cruzadas u otras técnicas de similar impacto para evitar problemas de sobreajuste del modelo.

# Contexto

- ▶ Ante estas regulaciones, las entidades financieras tuvieron que:
  - ▶ formar áreas separadas
  - ▶ contar con equipos especializados
  - ▶ complementar el uso de modelos estadísticos con modelos de machine learning
- ▶ Tal es la situación de la Cooperativa de Ahorro y Crédito DILE
  - ▶ Inicio a conformar su área y equipos especializados a inicios de noviembre del año 2023.
  - ▶ Está necesitando realizar proyectos de ciencia de datos en temas financieros.

# Título y objetivos de investigación

**Título:** “Credit Scoring para la Evaluación de Riesgo Crediticio en la Cooperativa de Ahorro y Crédito DILE utilizando Aprendizaje Automático”

**Objetivo general:** identificar las variables determinantes de la probabilidad de *default*.

## **Objetivos Específicos:**

- ▶ estimar un modelo de regresión logística de admisión para obtener la probabilidad de *default* aplicable a la evaluación crediticia de los préstamos.
- ▶ estimar un ensamble de modelos de admisión para obtener la probabilidad de *default* aplicable a la evaluación crediticia de los préstamos.
- ▶ identificar las variables de naturaleza cualitativa y cuantitativa que determinan el modelo.

# Importancia

Algunas de las principales razones que hacen que este proyecto sea relevante:

- ▶ Mejora en la gestión del riesgo crediticio
- ▶ Reducción de pérdidas financieras
- ▶ Acceso a crédito más justo y equitativo
- ▶ Optimización de recursos financieros

En última instancia, estas razones contribuyen a la estabilidad financiera de DILE y a la satisfacción de sus clientes.



# Factibilidad

**Data:** DILE proporcionará los datos

**Variables:** se asume variables internas y externas

**Unidad de estudio:** Personas naturales como prestatarios

**Población objetivo:** Prestatarios

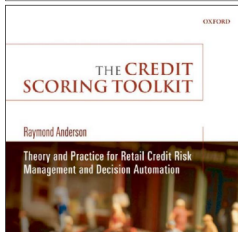
- ▶ malos: en alguno de los próximos 12 meses tuvo más de 30 días de atraso
- ▶ buenos: durante los 12 meses tuvo 0 días de atraso
- ▶ Se excluyen
  - ▶ aquellos que en alguno de los próximos 12 meses tuvo entre 1 y 30 días de atraso
  - ▶ aquellos con información incompleta en los 12 meses

# Factibilidad




**Modelos:** Regresión logística y ensamble de modelos

**Recursos computacionales:** Nube

**Bibliografía utilizada por los auditores:**



# Bibliografía

-  Ingenium. (2022). Programa de Especialización en Ciencia de Datos aplicada a Banca y Finanzas. <https://ingenium.edu.pe/management/ciencia-de-datos-aplicada-a-banca-y-finanzas/> (vid. pág. 3).
-  SBS. (2023, enero). SBS aprueba el Reglamento de Gestión de Riesgos de Modelo. <https://www.sbs.gob.pe/noticia/detallenoticia/idnoticia/2647?title=SBS%20aprueba%20el%20Reglamento%20de%20Gesti%C3%B3n%20de%20Riesgos%20de%20Modelo> (vid. pág. 4).
-  Torres, J. F. R. (2020). Fundamentos de Data Science y sus Aplicaciones en Distintas Industrias. (Vid. pág. 2).