



MODELO PREDICTIVO DE EXPORTACIONES AGRÍCOLAS EN COLOMBIA HASTA EL 2026

PROYECTO TALENTO TECH

Ricardo Andres Sanchez Sanches
Alexander González Troncoso

TABLA DE CONTENIDO

01

INTRODUCCIÓN

02

*IDENTIFICACIÓN DEL
PROBLEMA Y
OBJETIVOS*

03

ANÁLISIS DESCRIPTIVO

04

DASHBOARD

05

MODELO PREDICTIVO

06

*CONCLUSIONES Y
PERSPECTIVAS*



01

INTRODUCCIÓN

INTRODUCCIÓN

Colombia es un actor clave en el comercio agrícola global, destacándose en la exportación de café, banano, flores, cacao y palma de aceite, entre otros productos. Estas exportaciones generan empleo y fortalecen la economía del país.

Sin embargo, factores como la volatilidad del mercado, el clima y las políticas comerciales afectan la estabilidad del sector. Para enfrentar estos desafíos, este proyecto desarrolla un modelo predictivo de exportaciones hasta 2026, aplicando ciencia de datos y machine learning.





02

PROBLEMA Y OBJETIVOS

IDENTIFICACIÓN DE PROBLEMA



VARIABILIDAD

Variabilidad de los mercados internacionales y su impacto en la demanda de productos colombianos.



MEDIO AMBIENTE

Efectos del cambio climático en la producción agrícola y disponibilidad de productos para la exportación.



MODELOS

Falta de acceso a modelos predictivos que permitan planificar de manera estratégica las exportaciones.



OPTIMIZACIÓN

Necesidad de optimizar la logística y distribución de productos para minimizar pérdidas y maximizar ganancias.

OBJETIVO GENERAL

Desarrollar un modelo predictivo para la gestión de las exportaciones agrícolas tradicionales y no tradicionales en Colombia, utilizando técnicas avanzadas de análisis de datos para optimizar la toma de decisiones en el sector.

OBJETIVOS ESPECÍFICOS

- Implementar un proceso ETL con SQL y Python para limpiar, transformar y estructurar la información de exportaciones agrícolas colombianas.
- Visualizar tendencias y patrones de exportación mediante dashboards interactivos en Power BI.
- Construir un modelo predictivo basado en regresión logística, series temporales y técnicas de machine learning.
- Optimizar el modelo predictivo mediante ajuste de hiperparámetros y validación con datos recientes.
- Analizar el impacto del modelo en la planificación y gestión del comercio agrícola.

03

ANÁLISIS DESCRIPTIVO



DATOS

Se utilizó un conjunto de datos detallado sobre exportaciones agrícolas en Colombia, incluyendo productos tradicionales y no tradicionales, medidos en dólares FOB y toneladas.

- **Fuente:** Ministerio de Agricultura y Desarrollo Rural
- **Cobertura:** Nacional
- **Frecuencia de actualización:** Mensual
- **Tamaño:** 476,000 filas y 9 columnas
- **Última actualización:** 11 de septiembre de 2024

Los datos provienen del portal de datos abiertos del gobierno colombiano y permiten analizar la evolución de las exportaciones, facilitando la identificación de patrones clave para la construcción del modelo predictivo.

476 K

FILAS

9

COLUMNAS

EXPORTACIÓN

CADA FILA ES UNA

¿QUÉ HAY EN ESTE
CONJUNTO DE DATOS?

Columnas (9)

Nombre de la columna	Descripción
# Partida	Subpartida arancelaria a 10 dígitos
Tr Descripción Partida10 Dig	Descripción de la subpartida arancelaria a 10 dígitos
Tr Tradición productos	Tipo de tradicionalidad del producto
Tr Departamento	Departamento de origen
Tr País	País de destino de la exportación
# Año	Año
Tr Mes	Mes
# Exportaciones en valor (Miles USD FOB)	Valor de la exportación en miles de dólares FOB
# Exportaciones en volumen (Toneladas)	Cantidad de toneladas de la exportación

PROCESO ETL (EXTRACCIÓN, TRANSFORMACIÓN Y CARGA)

Esencial para estructurar y limpiar los datos antes del análisis y modelado.

- **Importancia de un dataset limpio:**
 - ✓ Mejora la precisión de los modelos.
 - ✓ Facilita la interpretación de resultados.
 - ✓ Optimiza el rendimiento computacional.
 - ✓ Permite una toma de decisiones basada en evidencia.
- **Características del dataset:**
 - Registros: 475,785 | Columnas: 9
 - Cobertura: 2019 - 2024
 - Exportaciones (USD FOB): Máx. 92.6M | 50% < 1,000 USD
 - Exportaciones (Toneladas): Media: 64.71 | Máx. 60,979

TRANSFORMACIÓN DEL DATASET

Se carga el Dataset por medio de pandas, en un archivo CSV.

Revisión y manejo de valores nulos para asegurar la calidad de los datos.

Se combinaron las columnas Año y Mes en una nueva columna Fecha con formato datetime.

Los valores en miles de USD fueron convertidos a pesos colombianos (COP) usando una API.

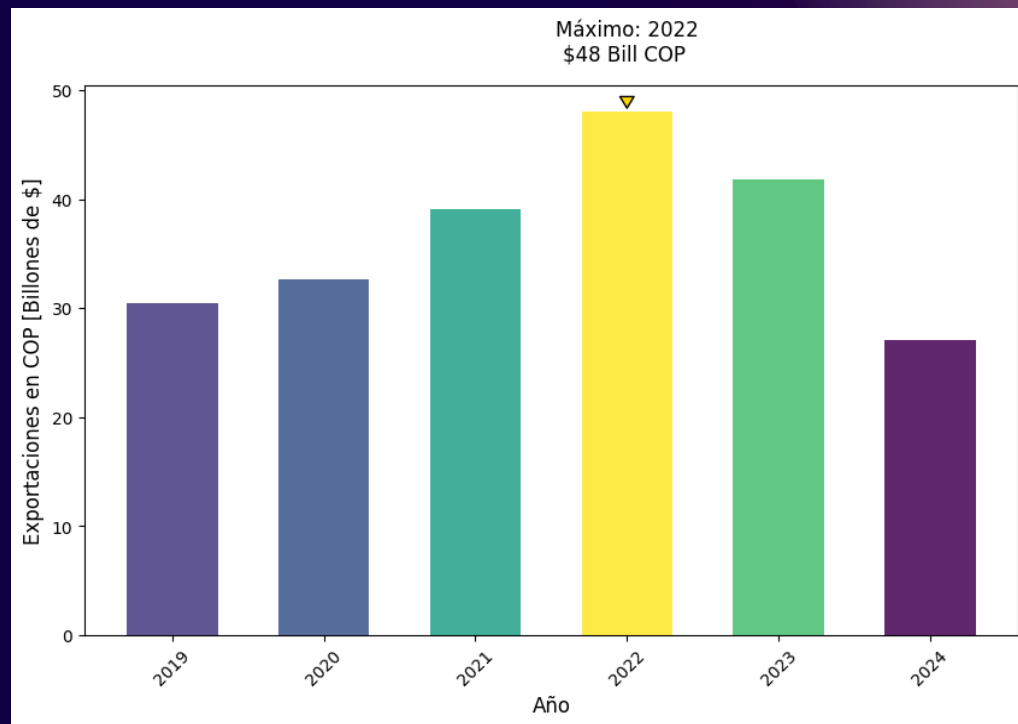
Archivo listo para los modelos.

El nuevo archivo limpio fue guardado como *Exportaciones_limpias.csv*.

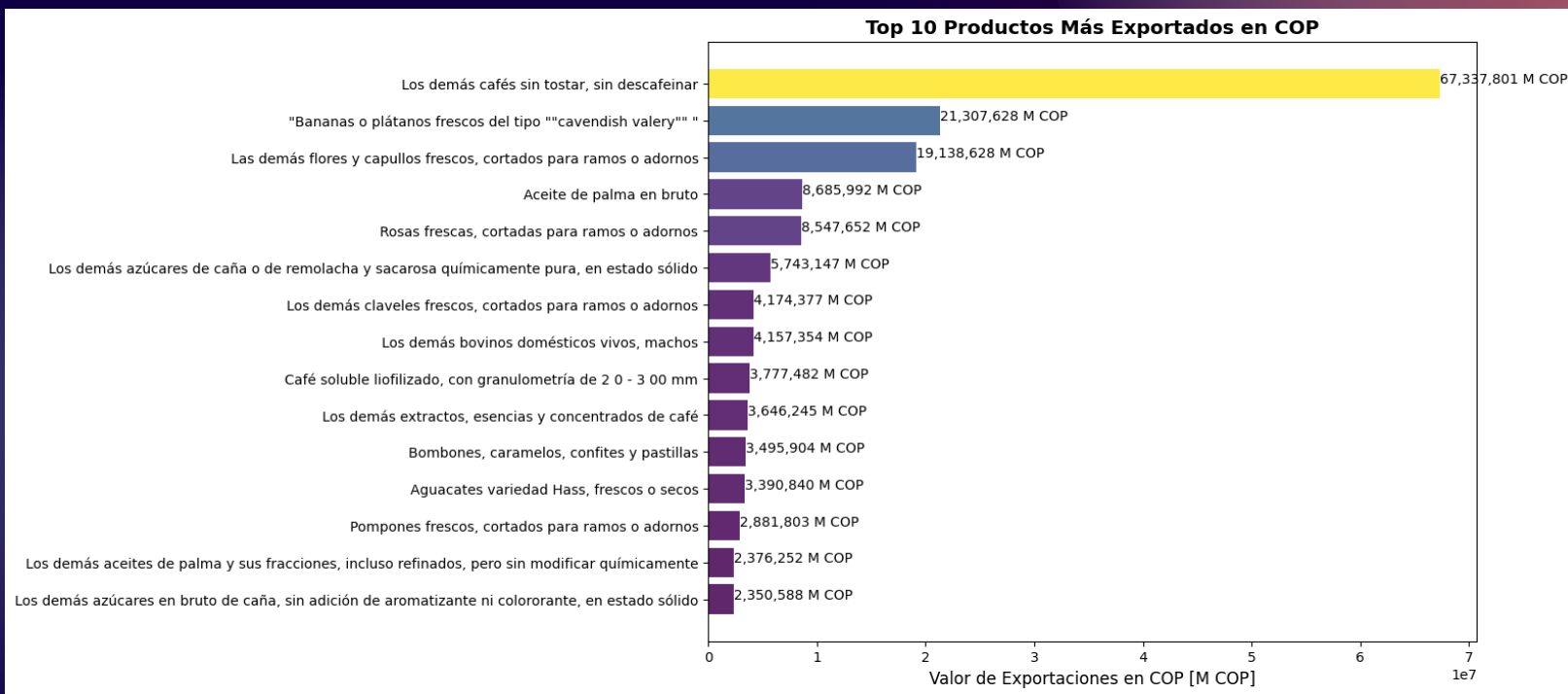
Se calculan las coordenadas de Latitud y Longitud para Departamento y Pais.

- División por 1,000,000 para expresar los valores en millones de COP.

ANÁLISIS DESCRIPTIVO

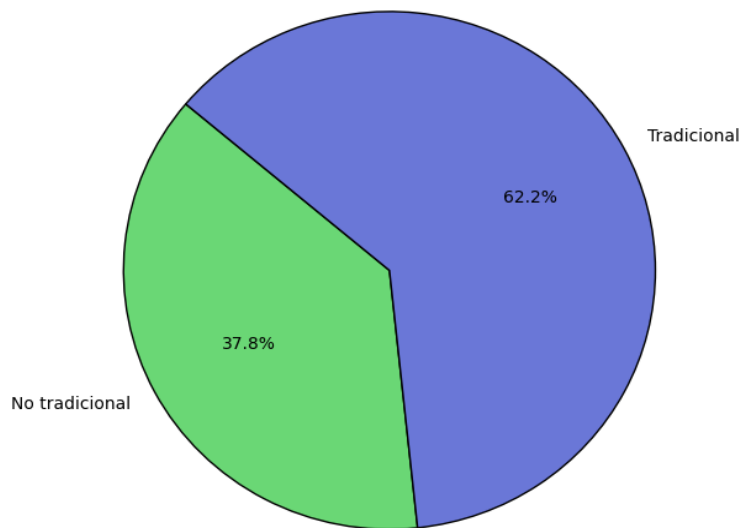


ANÁLISIS DESCRIPTIVO



ANÁLISIS DESCRIPTIVO

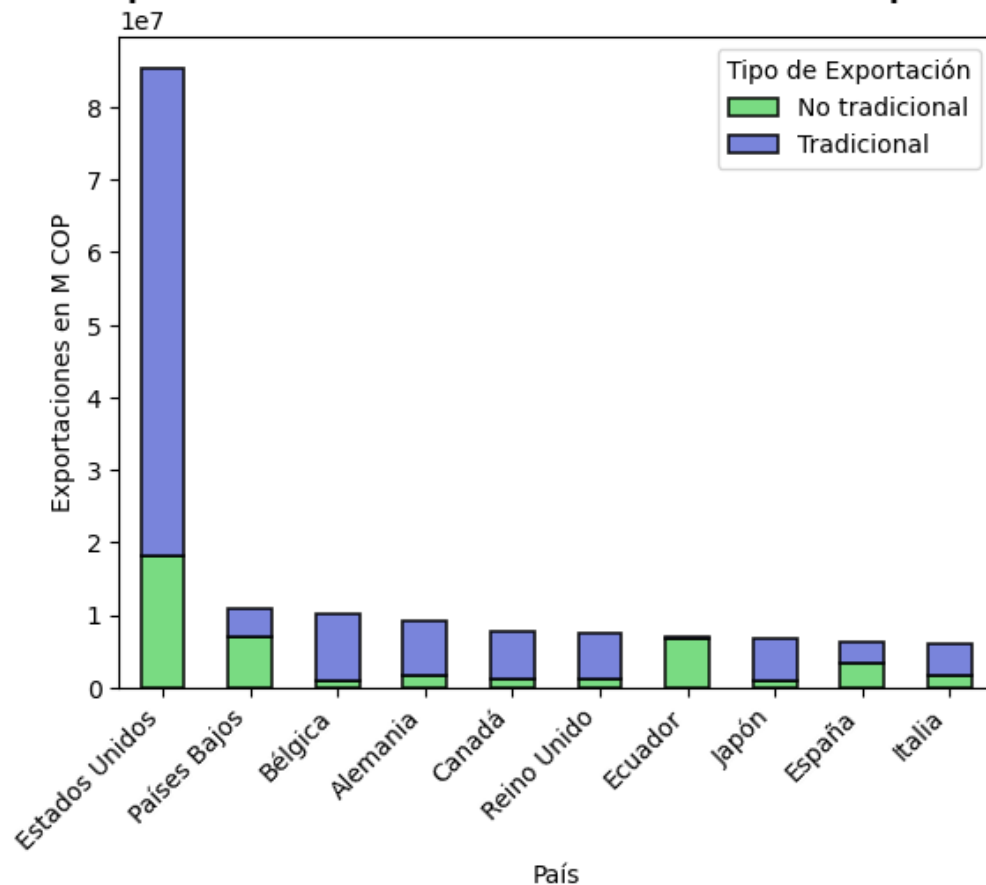
Distribución de Exportaciones según Tradición de Productos



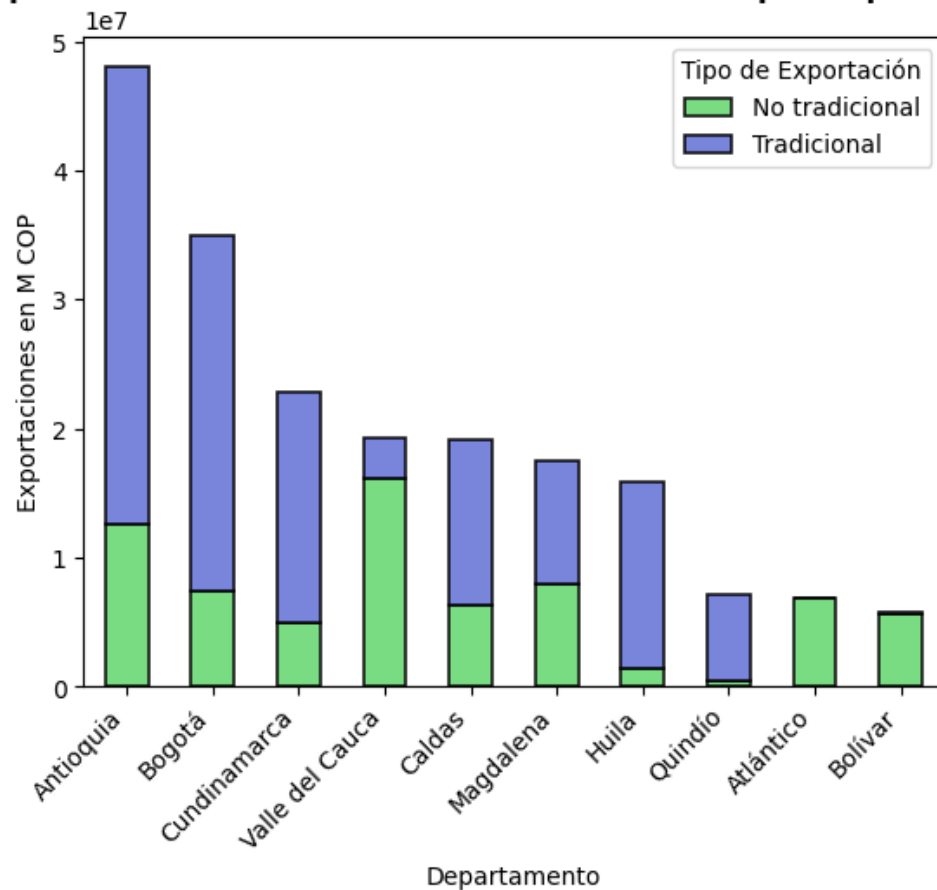
Se observa que los productos tradicionales representan el 62.2 % del total exportado, mientras que los no tradicionales abarcan el 37.8\%

Esto sugiere que, aunque hay una diversificación en los productos exportados, los sectores históricos como el café, el banano y las flores siguen dominando el mercado exterior. Este predominio indica que la economía exportadora del país aún depende en gran medida de estos productos agrícolas, reflejando su importancia histórica y estructural.

Exportaciones Tradicionales vs No Tradicionales por País



Exportaciones Tradicionales vs No Tradicionales por Departamento





04

DASHBOARD

05

MODELOS



SELECCIÓN DE VARIABLES

Conversión de fechas al tipo datetime.

- Extracción de variables temporales: Año y Mes.
- Codificación binaria de la variable Tradicional.
- Agrupación por mes, año y tipo de producto para construir una serie temporal.

Variables explicativas:

- Año: captura tendencias a largo plazo.
- Mes: modela estacionalidades.
- Tradicional: clasifica si el producto es tradicional (binaria).

Variable objetivo:

- Exportaciones en millones COP.

DIVISIÓN DE DATOS Y MODELOS

El conjunto de datos fue dividido en conjuntos de entrenamiento **(80 %)** y prueba **(20 %)** utilizando la función *train_test_split*. Esta división es fundamental para validar la capacidad de generalización de los modelos.

Se entrenaron cinco modelos de regresión, cada uno con características y fundamentos matemáticos distintos:

MODELOS

*REGRESIÓN LINEAL
MÚLTIPLE*

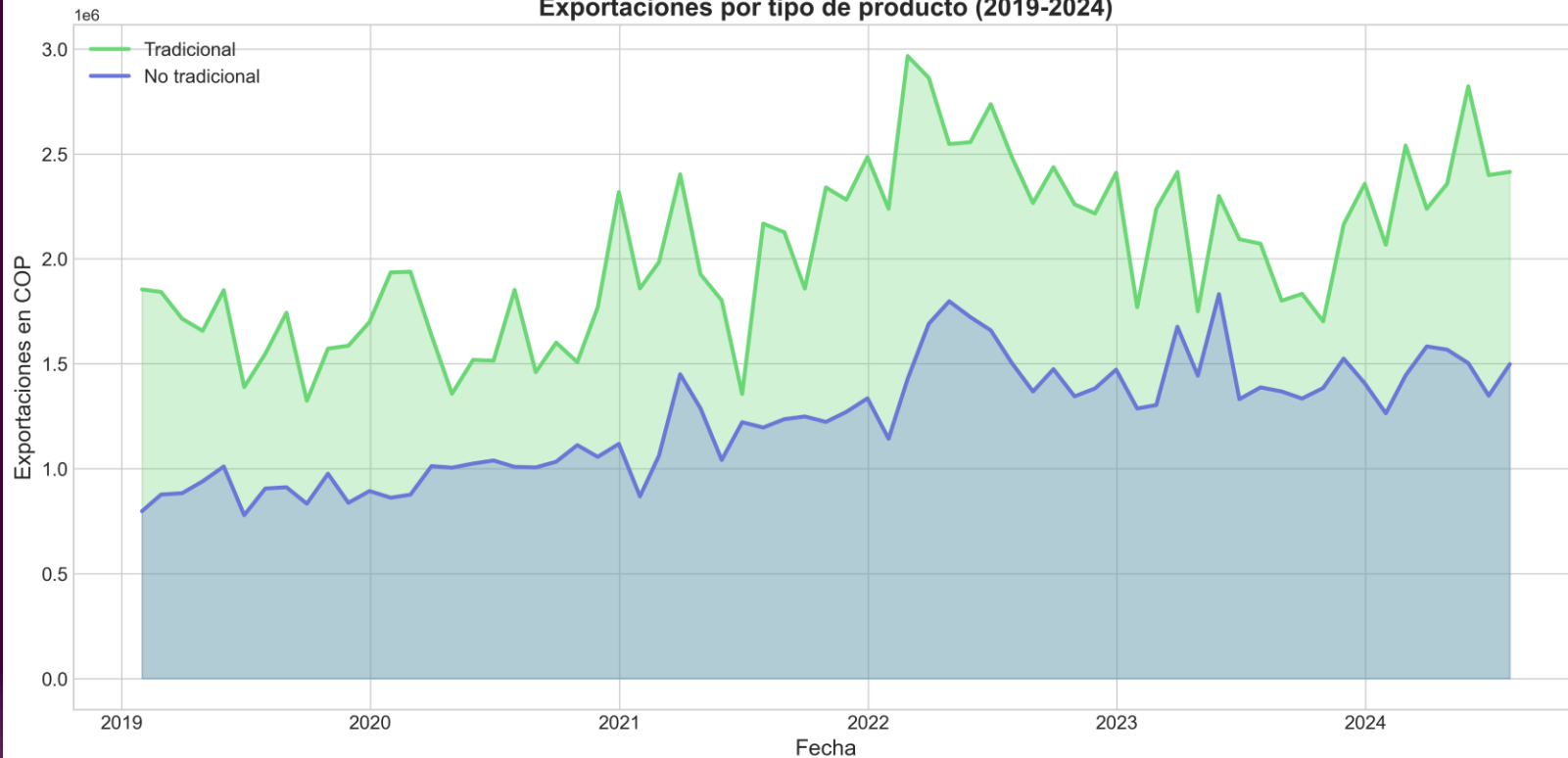
RANDOM FOREST
REGRESSOR

XGBOOST
REGRESSOR

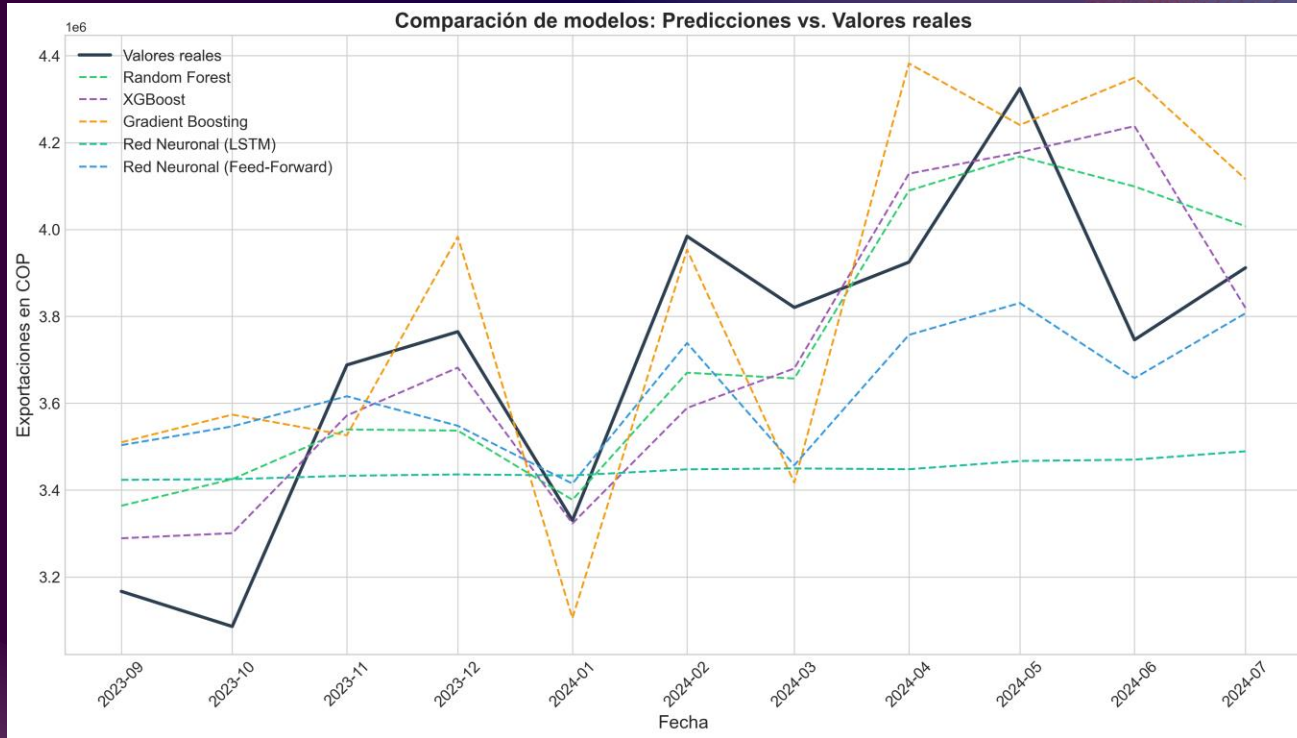
GRADIENT BOOSTING
REGRESSOR

RED NEURONAL
(MLPREGRESSOR)

Exportaciones por tipo de producto (2019-2024)



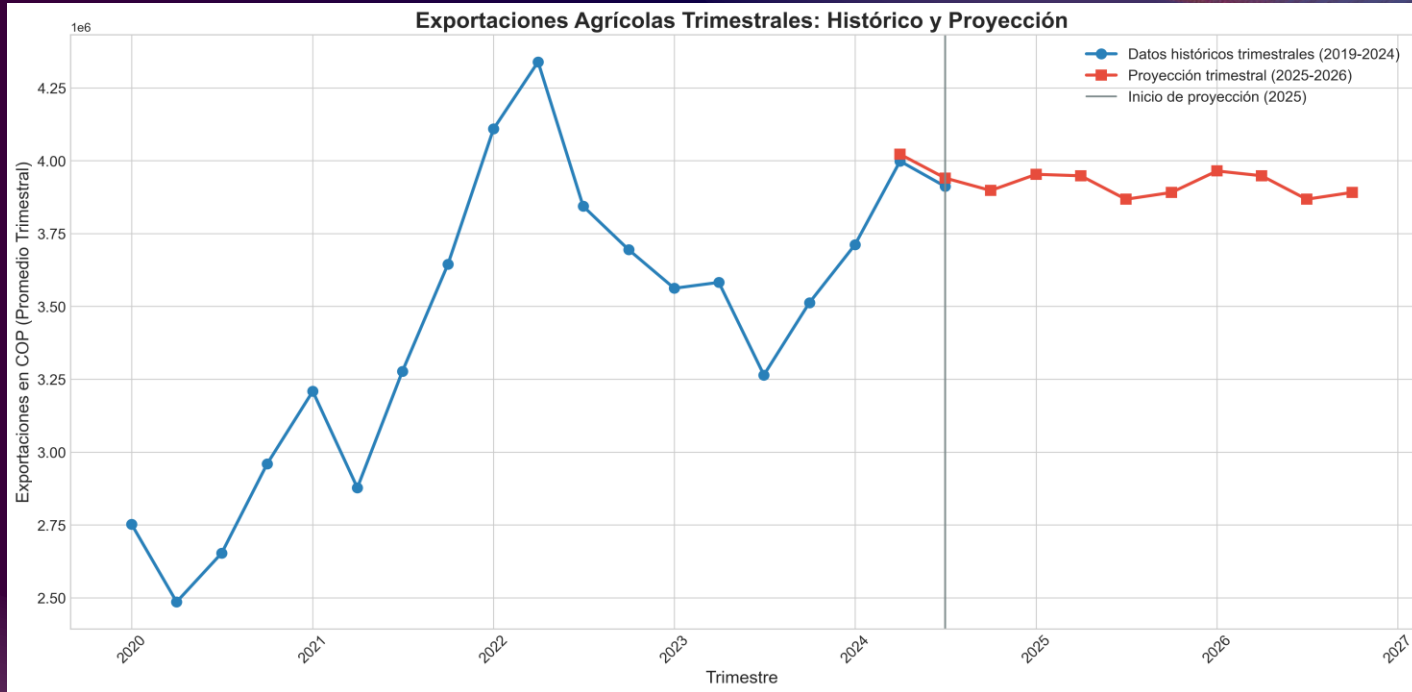
COMPARACIONES DE MODELOS



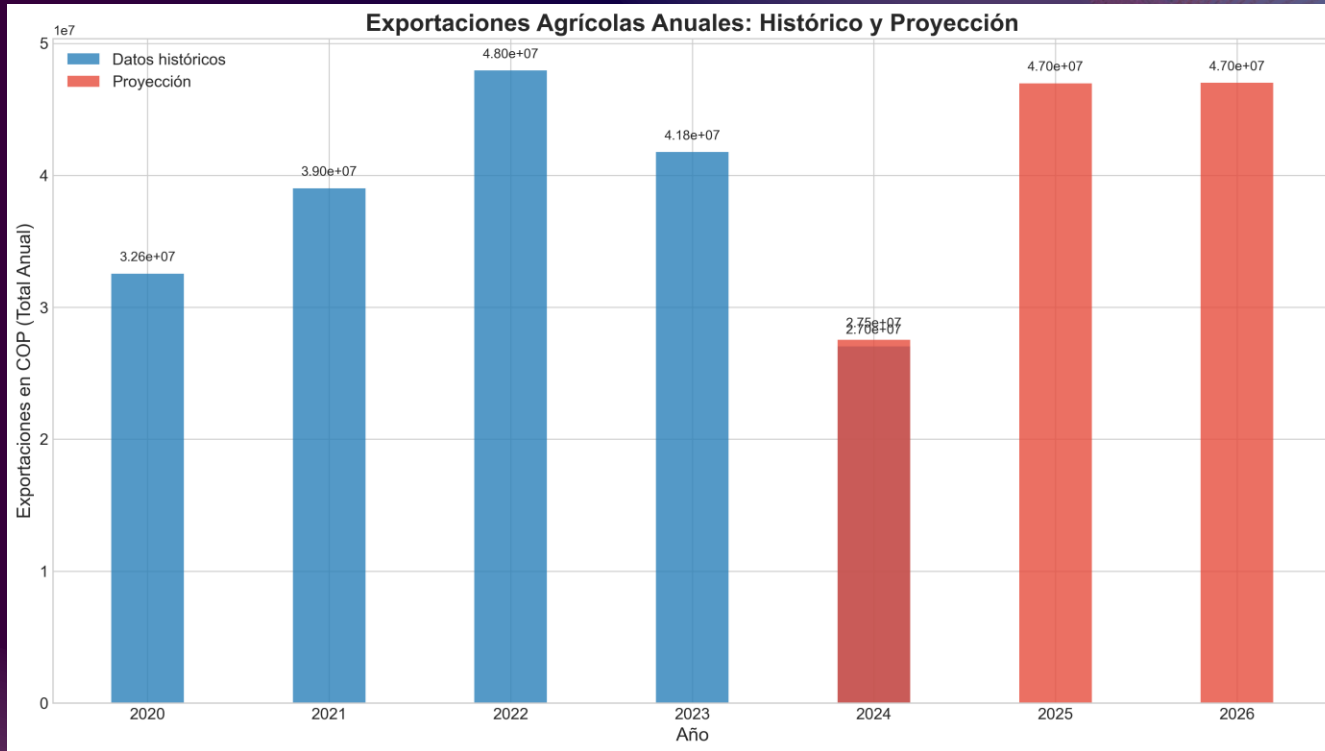
COMPARACIONES DE MODELOS

Modelo	MSE	RMSE	R ²
Random Forest	$4,92 \times 10^{10}$	221,715,52	0,6096
XGBoost	$5,19 \times 10^{10}$	227,875,83	0,5876
Gradient Boosting	$1,15 \times 10^{11}$	339,270,94	0,0859
Red Neuronal (LSTM)	$1,82 \times 10^{11}$	427,197,73	-0,4493
Red Neuronal (Feed-Forward)	$7,89 \times 10^{10}$	280,895,24	0,3734

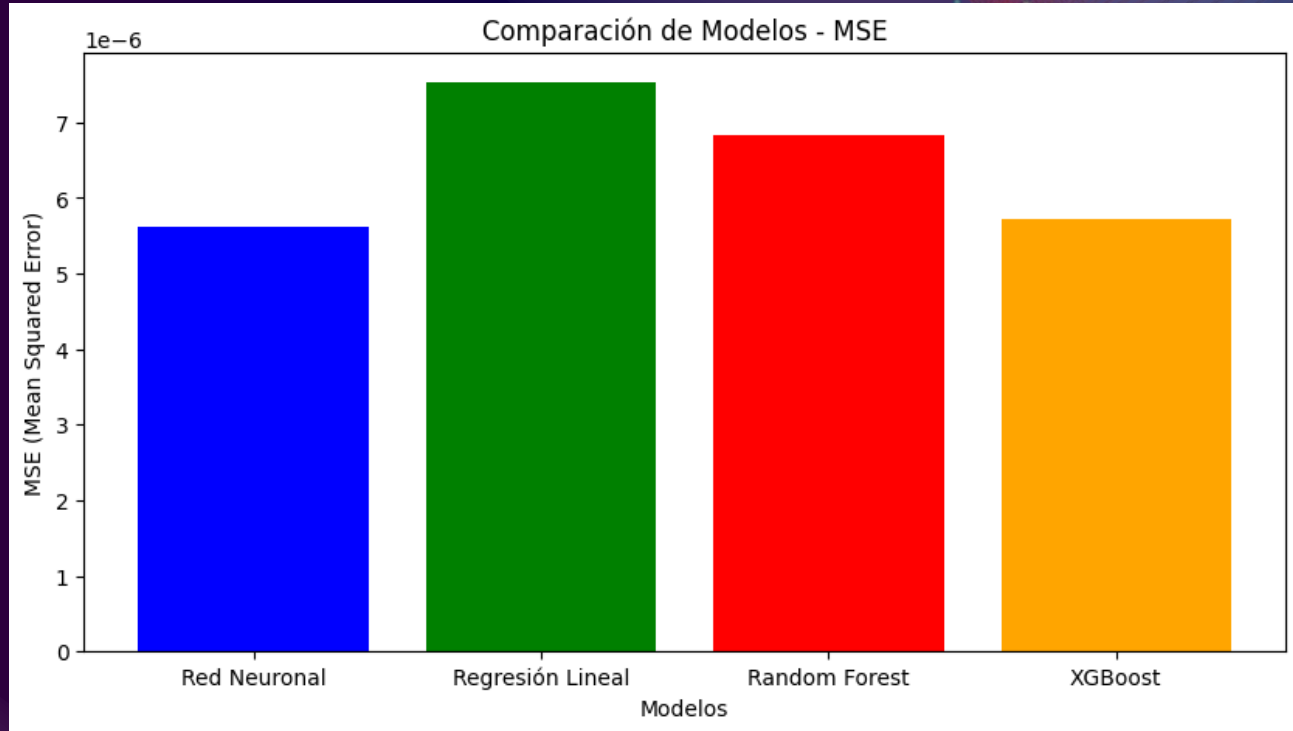
COMPARACIONES DE MODELOS



COMPARACIONES DE MODELOS



COMPARACIONES DE MODELOS





6800 s \approx 2 h
Corriendo
100 épocas



05

CONCLUSIONES Y PERSPECTIVAS

- 
- Los resultados obtenidos muestran que el modelo **Random Forest** alcanzó el mejor desempeño predictivo, superando a otros enfoques como XGBoost, Gradient Boosting y redes neuronales.
 - Este comportamiento se debe a su capacidad para manejar relaciones no lineales, su robustez frente al sobreajuste y su buen rendimiento incluso con un número reducido de variables explicativas.
 - En contraste, modelos más complejos como las redes neuronales requieren una mayor dimensionalidad y volumen de datos para desplegar su potencial. Por tanto, Random Forest se consolida como la opción más adecuada para proyectar las exportaciones agrícolas en este caso.

- 
- Importante mejorar la conectividad entre Datasets para lograr tener más variables que aporten al modelo. Esto influye en los resultados en gran medida.
 - Este proyecto sirve como una forma de visualizar de cierta forma el comportamiento de las exportaciones a futuro en el área agrícola, que es uno de nuestros fuertes como país. Tener esto presente en nuestro mercado es fundamental, para mejorar estrategias, para potenciar nuestro nivel de exportación.

GRACIAS



alexander.gonzalez@udea.edu.co

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon** and infographics & images by **Freepik**

Please keep this slide for attribution