

**Aplicación de modelos de machine learning para predicciones de
exportaciones para el mercado floricultor colombiano**

Felipe Puentes Abril

**Trabajo de grado para optar el título de Magister en Análisis de Datos y
Sistemas Inteligentes**

Director

Oscar Fernando Becerra Angarita

Magister en Ingeniería Eléctrica

Codirector

Cesar Hernando Valencia Niño

Doctor en Ingeniería Eléctrica

Universidad Santo Tomás, Bucaramanga

División de Ingenierías y Arquitectura

Maestría en Análisis de Datos y Sistemas Inteligentes

2026

Contenido

Introducción	8
Capítulo 1. Planteamiento del problema.....	9
1.1 Contexto del problema	9
1.2 Descripción clara y delimitada de la problemática.....	11
1.3 Justificación del proyecto	12
1.4 Objetivo general	13
1.5 Objetivos específicos	14
1.6 Alcance y delimitaciones del proyecto	14
2. Marco referencial.....	15
2.1 Marco conceptual	15
2.2 Marco teórico.....	16
2.3 Estado del arte	18
2.4 Identificación de vacíos o limitaciones en la literatura	19
3. Metodología.....	21
3.1 Tipo y enfoque de la investigación	21
3.2 Descripción de las fases del proyecto	22
3.3 Fuentes de datos y criterios de selección	23
3.4 Herramientas, técnicas y tecnologías empleadas.....	24
3.5 Métodos de análisis y criterios de evaluación	25
4. Desarrollo de la solución	26

4.1 Arquitectura general de la solución	26
4.2 Diseño del modelo, sistema o metodología	27
4.3 Implementación técnica	29
4.4 Descripción del prototipo o sistema desarrollado.....	30
5. Resultados y validación	32
5.1 Resultados experimentales o de aplicación	32
5.2 Métricas de desempeño	34
5.3 Análisis e interpretación de resultados	36
5.4 Validación frente a los objetivos planteados	45
6. Conclusiones y recomendaciones	47
6.1 Conclusiones generales del proyecto.....	47
6.2 Cumplimiento de los objetivos	48
6.3 Limitaciones identificadas	49
6.4 Recomendaciones y trabajos futuros	51
Referencias	54
Apéndices	59

Lista de tablas

Tabla 1. <i>Métricas de desempeño de Random Forest y XGBoost sobre el conjunto de prueba (enero 2024 – diciembre 2025).</i>	41
Tabla 2. <i>MAPE promedio por modelo y país destino sobre el conjunto de prueba (2024–2025).</i>	42

Lista de figuras

Figura 1. <i>Evolución mensual de las exportaciones de flores colombianas por país destino, 2016–2025.</i>	39
Figura 2. <i>Valores reales frente a predichos por Random Forest y XGBoost sobre el conjunto de prueba (2024–2025).</i>	41
Figura 3. <i>Importancia de variables en el modelo Random Forest (datos mensuales).</i> ...	43
Figura 4. <i>Importancia de variables en el modelo XGBoost (datos mensuales).</i>	43
Figura 5. <i>Importancia de variables vía SHAP (media de valor SHAP , en kg) - Random Forest.</i>	45
Figura 6. <i>Importancia de variables vía SHAP (media de valor SHAP , en kg) - XGBoost.</i>	46
Figura 7. <i>Relación entre la volatilidad comercial y el error porcentual absoluto de predicción, por país destino y modelo.</i>	49

Resumen

Este proyecto de investigación aplicada implementó, evaluó y comparó dos modelos de machine learning – Random Forest y XGBoost– para predecir el volumen mensual de las exportaciones florícolas colombianas hacia sus tres principales mercados de destino: Estados Unidos, Reino Unido y Canadá. El estudio integró variables macroeconómicas de Colombia y de cada país importador (producto interno bruto, inflación y tasa de cambio), junto con variables de estacionalidad y de rezago temporal, sobre un panel país destino/mes de 396 observaciones correspondientes al periodo 2015–2025. El desarrollo se realizó en Python sobre Google Colab, siguiendo una metodología que incluyó validación cruzada para la optimización de hiperparámetros, partición cronológica de los datos y validación estadística formal mediante la prueba de Wilcoxon signed-rank. Los resultados muestran que XGBoost obtuvo un desempeño puntual ligeramente superior a Random Forest en tres de las cuatro métricas evaluadas (MAE, RMSE y R^2), aunque la prueba de Wilcoxon signed-rank, aplicada sobre un conjunto de prueba de 72 observaciones, confirmó que esta diferencia no es estadísticamente significativa. El análisis de importancia de variables, complementado con valores SHAP, evidenció de forma estadísticamente significativa ($p = 0,663$; $p = 0,0135$) que ambos modelos coinciden en que el tamaño económico del mercado importador y la estacionalidad de la demanda son los factores con mayor peso explicativo, mientras que las condiciones macroeconómicas propias de Colombia pesan comparativamente menos. El proyecto demuestra la viabilidad técnica de aplicar algoritmos de aprendizaje de conjunto a un problema de comercio exterior agrícola, y deja como principal aporte una metodología completa y reproducible que puede ser reutilizada por PROCOLOMBIA, el Ministerio de Comercio, Industria y Turismo, o por investigaciones posteriores dentro del sector floricultor colombiano.

Palabras clave: machine learning, Random Forest, XGBoost, exportaciones florícolas, series de tiempo, variables macroeconómicas, predicción.

Abstract

This applied research project implemented, evaluated, and compared two machine learning models –Random Forest and XGBoost– to predict the monthly volume of Colombian flower exports to its three main destination markets: the United States, the United Kingdom, and Canada. The study integrated macroeconomic variables from Colombia and each importing country (gross domestic product, inflation, and exchange rate), together with seasonality and time-lag variables, over a destination-country/month panel of 396 observations covering the 2015–2025 period. The development was carried out in Python on Google Colab, following a methodology that included cross-validation for hyperparameter optimization, chronological data partitioning, and formal statistical validation through the Wilcoxon signed-rank test. Results show that XGBoost achieved a slightly better point performance than Random Forest on three of the four evaluated metrics (MAE, RMSE, and R^2), although the Wilcoxon signed-rank test, applied to a 72-observation test set, confirmed that this difference is not statistically significant. Variable-importance analysis, complemented with SHAP values, showed a statistically significant agreement ($\rho = 0.663$; $p = 0.0135$) that both models rank the importing market's economic size and demand seasonality as the most explanatory factors, while Colombia's own macroeconomic conditions carry comparatively less weight. The project demonstrates the technical feasibility of applying ensemble learning algorithms to an agricultural foreign-trade problem, and its main contribution is a complete, reproducible methodology that can be reused by PROCOLOMBIA, Colombia's Ministry of Commerce, Industry and Tourism, or by future research within the Colombian flower-growing sector. *Keywords: machine learning, Random Forest, XGBoost, flower exports, time series, macroeconomic variables, forecasting.*

Introducción

La floricultura colombiana se ha consolidado como uno de los sectores más dinámicos de la diversificación exportadora del país, en un contexto de pérdida de dinamismo del sector minero-energético. Anticipar con precisión el comportamiento de sus exportaciones resulta relevante tanto para los productores como para las entidades que diseñan política comercial, dado que las proyecciones disponibles actualmente se basan en su mayoría en métodos univariados que no incorporan explícitamente las condiciones macroeconómicas de los países compradores.

Este trabajo de grado presenta el desarrollo, evaluación y comparación de dos modelos de machine learning –Random Forest y XGBoost– que integran variables macroeconómicas bilaterales de Colombia y de sus tres principales países importadores (Estados Unidos, Reino Unido y Canadá) para predecir el volumen mensual de exportaciones florícolas.

El documento se organiza en seis capítulos. El Capítulo 1 plantea el problema, la justificación y los objetivos del proyecto. El Capítulo 2 presenta el marco referencial, incluyendo los conceptos, la teoría y el estado del arte relacionados con la predicción de exportaciones mediante machine learning. El Capítulo 3 describe la metodología empleada. El Capítulo 4 detalla el desarrollo de la solución, incluyendo el diseño y la implementación de los modelos. El Capítulo 5 presenta los resultados experimentales, las métricas de desempeño y su validación frente a los objetivos planteados. Finalmente, el Capítulo 6 expone las conclusiones, el cumplimiento de los objetivos, las limitaciones identificadas y las recomendaciones para trabajos futuros.

Capítulo 1. Planteamiento del problema

1.1 Contexto del problema

Durante la última década, la economía colombiana ha experimentado un proceso de transformación estructural asociado a la pérdida de dinamismo del sector minero-energético, históricamente uno de los principales motores de las exportaciones nacionales. Según la Asociación Nacional de Comercio Exterior (2025), este sector registró en 2025 una contracción del $-0,8\%$, situación que ha intensificado la necesidad de fortalecer sectores no tradicionales capaces de sostener el crecimiento externo del país.

En este contexto, la floricultura colombiana, en este panorama, se ha convertido en uno de los sectores más dinámicos de la diversificación exportadora del país. PROCOLOMBIA (2025) reportó un crecimiento del 13% en 2024 frente al año anterior, cifra que la ubica entre los productos no extractivos más relevantes de la canasta exportadora nacional. Su peso no es solo económico: el sector sostiene una cantidad considerable de empleo y desarrollo regional, lo que explica por qué se le considera estratégico dentro de la estructura productiva del país.

Pese a esa relevancia, el desempeño del sector floricultor depende de varios factores económicos y comerciales que no siempre son fáciles de anticipar: el tipo de cambio, la demanda internacional, los costos logísticos, la estacionalidad ligada a fechas comerciales clave y, en general, la exposición a choques geopolíticos globales. La combinación de todo esto genera un nivel de incertidumbre considerable sobre el comportamiento de las exportaciones, lo que complica tanto la planificación productiva de los cultivadores como la toma de decisiones de las entidades que promueven el comercio exterior del sector.

En este escenario, poder anticipar el comportamiento del comercio internacional resulta clave para que el sector gane competitividad. El problema es que los enfoques tradicionalmente usados para proyectar exportaciones suelen partir de modelos econométricos que asumen relaciones lineales entre variables económicas, un supuesto que se queda corto cuando el sistema que se quiere analizar tiene interacciones dinámicas y no lineales entre múltiples factores (Yang et al., 2023; Sellami et al., 2024).

Frente a esto, la literatura reciente apunta al Machine Learning y al Deep Learning como herramientas más robustas para modelar relaciones de este tipo. Varios estudios han encontrado mejoras significativas en la precisión de las predicciones de comercio internacional al usar estos enfoques: Gulzar et al. (2024), por ejemplo, muestran que las redes neuronales artificiales superan a los modelos logísticos y de soporte vectorial en la predicción de exportaciones tecnológicas, mientras que Yang et al. (2023) reportan reducciones importantes en métricas de error como el RMSE y el MAPE al usar modelos híbridos de deep learning.

Con todo, estos avances rara vez se han aplicado al comercio de productos agrícolas perecederos. Hay poca evidencia empírica sobre modelos de Machine Learning aplicados a la predicción de exportaciones en sectores como la floricultura, donde las dinámicas de mercado combinan factores económicos, estacionales y logísticos que piden enfoques analíticos hechos a la medida.

1.2 Descripción clara y delimitada de la problemática

El sector floricultor gana peso cada vez más dentro de la diversificación exportadora colombiana, pero la capacidad de anticipar con precisión sus exportaciones sigue siendo limitada. Las proyecciones que manejan hoy los distintos actores del sector se apoyan en enfoques analíticos tradicionales que no siempre capturan bien las relaciones complejas entre variables macroeconómicas, comerciales y temporales que mueven el comercio internacional de flores.

Esa limitación pesa más en un entorno de alta volatilidad económica, cambios en la demanda internacional y fluctuaciones cambiarias que golpean directamente la competitividad del sector. Sin herramientas predictivas que integren varias dimensiones del comercio internacional a la vez, es difícil generar proyecciones precisas, y eso termina restringiendo la planificación estratégica tanto de los productores como de las instituciones que promueven y gestionan el comercio exterior.

Además, aunque ya está demostrado el potencial de los modelos de Machine Learning para mejorar la precisión de las predicciones económicas, la evidencia empírica disponible se concentra en sectores manufactureros o tecnológicos (Yang et al., 2023; Sellami et al., 2024; Gulzar et al., 2024). Los estudios sobre productos agrícolas perecederos como las flores siguen siendo escasos, y todavía más en economías emergentes como la colombiana.

En este sentido, se identifica una brecha en el análisis del comercio exterior del sector floricultor colombiano, relacionada con la ausencia de estudios que evalúen el potencial de los modelos de Machine Learning para mejorar la predicción de las exportaciones florales a partir de variables macroeconómicas relevantes tanto de Colombia como de sus principales socios comerciales.

Por lo tanto, la problemática de esta investigación se centra en analizar si la implementación y comparación de modelos predictivos basados en Machine Learning, específicamente Random Forest y XGBoost, puede mejorar la precisión de las proyecciones de exportaciones de flores colombianas frente a las estimaciones actualmente utilizadas por entidades del sector floricultor, contribuyendo así a fortalecer las herramientas analíticas disponibles para la toma de decisiones en un sector estratégico para la economía del país.

1.3 Justificación del proyecto

La floricultura colombiana constituye uno de los sectores más representativos del potencial exportador no tradicional del país. Con exportaciones que superan los US\$2.000 millones anuales y una participación superior al 70% del mercado estadounidense de flores cortadas, el sector no solo impulsa el crecimiento económico, sino que también genera más de 200.000 empleos directos e indirectos (PROCOLOMBIA, 2025).

Ese peso económico no exime al sector de enfrentar desafíos ligados a la volatilidad de los mercados internacionales. La variabilidad de los precios, las fluctuaciones cambiarias, los costos logísticos y la estacionalidad internacional, marcada sobre todo por San Valentín y el Día de la Madre, afectan directamente tanto la estabilidad de los ingresos como la capacidad de planificación estratégica de productores y exportadores.

Por eso contar con herramientas capaces de anticipar el comportamiento de las exportaciones cobra especial relevancia. Los modelos de Machine Learning permiten capturar relaciones no lineales y dinámicas complejas entre variables económicas, y en principio pueden ofrecer mayor precisión predictiva que los enfoques econométricos tradicionales.

Desde una perspectiva teórica, el estudio contribuye al campo del comercio internacional al aplicar técnicas de Machine Learning en el análisis de productos agrícolas no tradicionales, un ámbito donde la evidencia empírica aún es limitada. Desde el punto de vista metodológico, la investigación propone la implementación y comparación de modelos predictivos basados en Random Forest y XGBoost, incorporando variables macroeconómicas bilaterales y temporales que permitan capturar de manera más precisa las dinámicas del comercio internacional de flores.

En términos aplicados, los resultados podrán ofrecer insumos analíticos relevantes para entidades como PROCOLOMBIA y el Ministerio de Comercio, Industria y Turismo, facilitando la formulación de estrategias de promoción comercial, planificación logística y gestión del riesgo asociado a la volatilidad macroeconómica.

La literatura reciente (Yang et al., 2023; Sellami et al., 2024; Gulzar et al., 2024) reporta reducciones de hasta el 30% en los errores de predicción al usar modelos de Machine Learning frente a los enfoques tradicionales. Si eso se traslada al sector floricultor colombiano, podría traducirse en exportaciones más eficientes, mejores decisiones y mayor competitividad del país en los mercados internacionales.

1.4 Objetivo general

Implementar, evaluar y comparar modelos predictivos basados en Machine Learning como Random Forest y XGBoost, que integren variables macroeconómicas de Colombia y de sus 3 principales países importadores, con el fin de analizar su capacidad para mejorar la precisión de las proyecciones de exportaciones de flores frente a las estimaciones provistas por entidades del sector floricultor colombiano.

1.5 Objetivos específicos

- Determinar la influencia de las variables macroeconómicas (nacionales y del top 3 países importadores) en el comportamiento histórico de las exportaciones de flores colombianas.
- Desarrollar modelos predictivos basados en los algoritmos Random Forest y XGBoost que capturen las relaciones no lineales entre las variables macroeconómicas y el volumen de exportaciones.
- Cuantificar la precisión y el error de los pronósticos generados por los modelos de ML para establecer su confiabilidad en contextos de volatilidad comercial
- Validar estadísticamente a través del método Wilcoxon signed-rank, que modelo tiene mejor desempeño en la predicción del mercado de flores.

1.6 Alcance y delimitaciones del proyecto

El estudio se enfocará en la implementación y evaluación de modelos predictivos supervisados, particularmente Random Forest y XGBoost, aplicados al análisis de exportaciones florales colombianas. Los modelos serán entrenados y evaluados utilizando datos correspondientes al período 2015–2025, complementados con indicadores macroeconómicos nacionales e internacionales como PIB, inflación, tasa de cambio y variables de demanda de los principales países importadores.

El análisis se hará en Python, usando librerías de ciencia de datos como Scikit-learn, TensorFlow y XGBoost. El desempeño de los modelos se evaluará con MAE, RMSE, MAPE y R^2 , complementados con validación cruzada para dar más robustez a los resultados.

Quedan fuera del alcance técnico de este estudio los modelos basados en grafos y las arquitecturas de deep learning más especializadas (Monken et al., 2021; Sellami et al., 2024);

se mencionan más adelante como posibles líneas de trabajo futuro, pero no se implementan aquí.

La investigación también se limita a un enfoque cuantitativo y predictivo: no entra en dimensiones cualitativas como políticas públicas, estrategias empresariales o análisis institucional del sector.

Se espera llegar a un modelo predictivo con mayor precisión para estimar el comportamiento de las exportaciones florales colombianas, y que ese resultado sirva de base para herramientas analíticas útiles en la planificación comercial, la diversificación exportadora y la gestión del riesgo macroeconómico del sector floricultor.

2. Marco referencial

2.1 Marco conceptual

Este estudio se apoya en conceptos de tres campos: la analítica de datos, la inteligencia artificial y el comercio internacional, que en conjunto permiten comprender y modelar el comportamiento de las exportaciones florales colombianas.

En primer lugar, la analítica de datos se define como el conjunto de técnicas orientadas a la extracción de conocimiento útil a partir de datos, mediante procesos de limpieza, transformación, modelado y evaluación (Provost & Fawcett, 2013). En el contexto económico, la analítica permite identificar patrones, tendencias y relaciones entre variables macroeconómicas que inciden en el comercio internacional.

El Machine Learning (ML), por su parte, es la rama de la inteligencia artificial dedicada a algoritmos que aprenden patrones a partir de datos históricos y hacen predicciones sin necesidad de estar programados explícitamente para cada caso (T. Mitchell, 1997). A diferencia de los modelos econométricos tradicionales, no depende de supuestos estrictos de

linealidad, lo que lo vuelve especialmente útil para modelar sistemas complejos y dinámicos como el comercio internacional, donde las relaciones entre variables rara vez son lineales.

También son centrales las variables macroeconómicas (PIB, inflación, tipo de cambio, demanda internacional), que inciden directamente en los flujos comerciales entre países y que permiten mirar el comercio internacional como un sistema interdependiente, no como flujos aislados (Krugman, Obstfeld, & Melitz, 2018).

Por último, este trabajo se ubica dentro del análisis de sistemas complejos, en los que múltiples variables interactúan de forma no lineal y generan dinámicas que los enfoques tradicionales tienen dificultad para capturar (M. Mitchell, 2009). Ahí es donde los modelos de inteligencia artificial aportan herramientas más flexibles y robustas para el análisis predictivo.

2.2 Marco teórico

La predicción de flujos comerciales internacionales se ha abordado desde distintos enfoques, que aquí se agrupan en tres grandes categorías: los modelos econométricos tradicionales, los modelos de Machine Learning y los enfoques más avanzados basados en deep learning y grafos.

En primer lugar, los modelos econométricos tradicionales, como las regresiones lineales y los modelos gravitacionales del comercio, han sido ampliamente utilizados para explicar los flujos comerciales entre países. Estos modelos se basan en supuestos de linealidad y relaciones estáticas entre variables, lo cual limita su capacidad para capturar dinámicas complejas y no lineales presentes en el comercio internacional.

En segundo lugar, los modelos de Machine Learning han ganado relevancia debido a su capacidad para modelar relaciones no lineales y manejar grandes volúmenes de datos.

Entre los más utilizados se encuentran:

- Random Forest: modelo basado en ensambles de árboles de decisión que mejora la precisión mediante la agregación de múltiples predictores.
- XGBoost (Extreme Gradient Boosting): técnica de boosting que optimiza iterativamente el error de predicción, destacándose por su alto rendimiento y eficiencia computacional.
- Support Vector Machines (SVM) y redes neuronales artificiales, utilizadas en diversos estudios para predicción económica.

Diversos trabajos han demostrado que estos modelos superan a los enfoques tradicionales en términos de precisión predictiva. Por ejemplo, las redes neuronales han mostrado mejor desempeño que modelos logísticos y SVM en la predicción de exportaciones tecnológicas (Gulzar et al., 2024) .

En tercer lugar, se encuentran los enfoques avanzados basados en deep learning y grafos, como las redes neuronales profundas y las Graph Neural Networks (GNN). Estos modelos permiten capturar relaciones estructurales entre países y dependencias complejas en redes comerciales internacionales (Monken et al., 2021; Sellami et al., 2024). Asimismo, modelos como los autoencoders basados en grafos y enfoques híbridos han demostrado mejoras significativas en la predicción de flujos comerciales (Minakawa et al., 2024; Pan et al., 2024).

Como ya se mencionó, estos enfoques más avanzados no se implementan en este estudio, aunque sí quedan como una base importante para investigaciones futuras.

2.3 Estado del arte

Varias investigaciones recientes han abordado la predicción del comercio internacional con técnicas de analítica de datos e inteligencia artificial, y muestran avances significativos en el campo.

Yang et al. (2023) desarrollaron modelos de deep learning para predecir comercio global y lograron reducciones significativas en RMSE y MAPE frente a modelos tradicionales, lo que confirma la capacidad de los modelos profundos para capturar dinámicas complejas en los flujos comerciales internacionales.

Sellami et al. (2024), por su parte, exploran el uso de Graph Neural Networks (GNN) para modelar relaciones entre países y encuentran que incorporar estructuras de red mejora la precisión de la predicción de flujos comerciales. En una línea parecida, Monken et al. (2021) usan redes neuronales sobre grafos para analizar causalidad en el comercio internacional y resaltan la importancia de las interdependencias entre economías.

Dentro del Machine Learning aplicado a la economía, Gulzar et al. (2024) encuentran que las redes neuronales artificiales superan a los modelos tradicionales en la predicción de exportaciones tecnológicas, otro dato a favor de usar estos enfoques en contextos económicos complejos.

Jiang et al. (2023) y Karim et al. (2023), por otro lado, usan big data y análisis temporal para estudiar el comercio agrícola y los efectos del libre comercio, lo que confirma la importancia de cruzar varias fuentes de información y variables macroeconómicas.

Sin embargo, la mayoría de estos trabajos se concentran en sectores industriales o tecnológicos, mientras que el análisis de productos agrícolas, especialmente perecederos como las flores, ha recibido menor atención en la literatura; situación que compromete directamente el fin de esta investigación ya que el sector floricultor del país si bien está

tecnificado en torno a su producción, NO está tecnificado en modelos de predicción.

El sector que se desea abordar trabaja bajo modelos muy básicos, por lo que, se pretende que la investigación se realice con modelos básicos como (Random Forest y XG Boost), para generar una adherencia rápida del sector.

2.4 Identificación de vacíos o limitaciones en la literatura

A pesar de los avances en la aplicación de técnicas de Machine Learning y deep learning en el análisis del comercio internacional, la literatura presenta varias limitaciones relevantes que justifican el desarrollo del presente estudio.

La primera limitación es temática: la literatura se concentra en sectores manufactureros y tecnológicos, y deja relativamente de lado los productos agrícolas perecederos como las flores. Este tipo de producto tiene características propias (alta estacionalidad, dependencia logística, sensibilidad a factores externos) que piden un enfoque metodológico distinto.

En segundo lugar, buena parte de los estudios recurre a modelos muy complejos (redes profundas, modelos basados en grafos) que, aunque precisos, son difíciles de interpretar y de aplicar en la práctica, sobre todo para quienes toman decisiones en contextos empresariales o institucionales.

La tercera limitación tiene que ver con las variables macroeconómicas bilaterales, que los modelos predictivos integran poco, restando capacidad para capturar las dinámicas propias entre países exportadores e importadores.

Además, varios trabajos se centran solo en mejorar el desempeño predictivo y dejan de lado la validación estadística de los resultados o la comparación sistemática entre modelos,

algo fundamental para que las conclusiones sean robustas.

Finalmente, existe una escasez de estudios aplicados al contexto colombiano, particularmente en el sector floricultor, a pesar de su relevancia económica y estratégica. Esta brecha limita la disponibilidad de herramientas analíticas adaptadas a las necesidades del país.

En este contexto, el presente proyecto busca contribuir a la literatura mediante la implementación y comparación de modelos de Machine Learning como Random Forest y XGBoost, incorporando variables macroeconómicas relevantes y enfocándose en un sector poco explorado, con el fin de mejorar la precisión predictiva y aportar evidencia empírica para la toma de decisiones.

3. Metodología

3.1 Tipo y enfoque de la investigación

La presente investigación se enmarca dentro de un enfoque cuantitativo, de tipo aplicado y alcance predictivo-comparativo. Es aplicada porque busca resolver una necesidad concreta identificada en el sector floricultor colombiano la ausencia de herramientas de proyección basadas en técnicas avanzadas de analítica de datos mediante la implementación de modelos ya validados en la literatura (Random Forest y XGBoost), en lugar del desarrollo de nuevos algoritmos. Su alcance es predictivo, en tanto el objetivo central es estimar el comportamiento de las exportaciones florícolas colombianas a partir de variables macroeconómicas históricas, y comparativo, dado que contrasta el desempeño de dos algoritmos de Machine Learning entre sí y frente a las estimaciones actualmente utilizadas por entidades del sector floricultor.

Desde el punto de vista metodológico, el estudio adopta un diseño no experimental, longitudinal retrospectivo, pues se apoya en series históricas de datos (2015-2025) sin manipulación de variables por parte del investigador, y correlacional-explicativo, en la medida en que busca establecer el grado de influencia de las variables macroeconómicas nacionales e internacionales sobre el volumen de exportaciones de flores. El enfoque es cuantitativo, ya que todo el proceso de análisis, modelado y validación se sustenta en datos numéricos y en técnicas estadísticas y de aprendizaje automático susceptibles de medición objetiva, tales como el MAE, el RMSE, el MAPE, el R^2 y la prueba de Wilcoxon signed-rank (HernándezSampieri & Mendoza, 2018).

3.2 Descripción de las fases del proyecto

El desarrollo del proyecto se estructura en seis fases secuenciales, cada una con entregables específicos que alimentan la fase siguiente, siguiendo una adaptación del proceso estándar para proyectos de minería de datos y aprendizaje automático conocido como CRISP-DM (Chapman et al., 2000):

- *Fase 1. Definición y planteamiento.* Delimitación del problema, formulación de objetivos e hipótesis de trabajo, y revisión del estado del arte (desarrollada en los Capítulos 1 y 2).
- *Fase 2. Recolección y consolidación de datos.* Identificación y descarga de series históricas de exportaciones florícolas colombianas (2015-2025) y de variables macroeconómicas nacionales y de los principales países importadores, provenientes de fuentes oficiales.
- *Fase 3. Preprocesamiento y construcción de variables.* Limpieza de datos, tratamiento de valores atípicos y faltantes, homologación de periodicidad y unidades, y construcción de variables derivadas (rezagos temporales, variables estacionales y tasas de crecimiento) que alimentarán los modelos.
- *Fase 4. Modelado predictivo.* Entrenamiento de los modelos Random Forest y XGBoost, incluyendo la partición de datos en conjuntos de entrenamiento y

prueba, y la optimización de hiperparámetros.

- *Fase 5. Evaluación y validación estadística.* Cálculo de métricas de error (MAE, RMSE, MAPE, R^2) mediante validación cruzada, y comparación formal del desempeño de los dos modelos a través de la prueba de Wilcoxon signed-rank.
- *Fase 6. Análisis de resultados y documentación.* Interpretación de los hallazgos frente a los objetivos planteados, elaboración de conclusiones y recomendaciones, y consolidación del informe final.

3.3 Fuentes de datos y criterios de selección

El proyecto integra dos tipos de información: (i) datos de comercio exterior correspondientes a las exportaciones florícolas colombianas y (ii) variables macroeconómicas de Colombia y de sus principales países importadores. Ambos conjuntos de datos se construyen a partir de fuentes oficiales y de organismos multilaterales, con el fin de garantizar la trazabilidad y confiabilidad de la información.

Datos de exportaciones: series mensuales/anuales de volumen y valor de exportaciones de flores colombianas, obtenidas de la base de datos para comercio exterior Legis Comex (Legis, s.f.), provista por el CRAI de la Universidad Santo Tomás.

Variables macroeconómicas: PIB, inflación y tasa de cambio de Colombia, obtenidas del Banco de la República (s.f.) y el DANE (s.f.); e indicadores equivalentes de los tres principales países importadores de flor colombiana, obtenidos de fuentes como el Banco

Mundial (s.f.), el Fondo Monetario Internacional (2025) y la Organización para la Cooperación y Desarrollo Económico (2024). La identificación definitiva de dichos países se realizará a partir de su participación histórica dentro de las exportaciones colombianas, verificada en la fase de recolección de datos.

Para la selección de las fuentes se aplican los siguientes criterios: (1) carácter oficial o institucional de la fuente; (2) cobertura temporal continua para el periodo 2015–2025; (3) nivel de completitud aceptable, admitiendo un porcentaje de datos faltantes inferior al 5% por serie, subsanable mediante técnicas de imputación; y (4) consistencia en la unidad de medida y la periodicidad entre las distintas series, de manera que puedan integrarse en un único conjunto de datos homologado.

3.4 Herramientas, técnicas y tecnologías empleadas

Todo el desarrollo técnico del proyecto se hizo en Python, por su ecosistema maduro de librerías para análisis de datos y aprendizaje automático y por ser el lenguaje más usado en la literatura revisada en el Capítulo 2. Como entorno de ejecución se usó Google Colab: permite correr notebooks de Python sin instalar nada localmente, da acceso a recursos de cómputo adicionales (GPU/TPU) si se necesitan, y facilita documentar el proceso de forma reproducible y trabajar de forma colaborativa.

Para la manipulación y preparación de los datos se emplean las librerías pandas y NumPy. Los modelos predictivos se implementan mediante Scikitlearn (Pedregosa et al., 2011), que provee la implementación de Random Forest y las utilidades de validación cruzada y partición de datos, y XGBoost (Chen & Guestrin, 2016), para el algoritmo homónimo de gradient boosting. La visualización de resultados se apoya en Matplotlib y Seaborn, mientras que la validación estadística de la comparación entre modelos se realiza

con el módulo `scipy.stats`, específicamente la prueba de Wilcoxon signed-rank (Wilcoxon, 1945). El control de versiones del código se lleva a cabo mediante Git y GitHub.

3.5 Métodos de análisis y criterios de evaluación

El proceso de modelado se apoya en un esquema de validación cruzada (kfold, adaptado a la naturaleza temporal de los datos mediante particiones que respetan el orden cronológico) para reducir el riesgo de sobreajuste y obtener estimaciones robustas del desempeño de cada modelo. La búsqueda de hiperparámetros óptimos para Random Forest (número de árboles, profundidad máxima, número mínimo de muestras por hoja) (Breiman, 2001) y XGBoost (tasa de aprendizaje, número de estimadores, profundidad máxima, regularización) (Chen & Guestrin, 2016) se realiza mediante búsqueda en cuadrícula (grid search) o búsqueda aleatoria (randomized search) sobre el conjunto de entrenamiento.

El desempeño de los modelos se evalúa mediante cuatro métricas complementarias: el Error Absoluto Medio (MAE), la Raíz del Error Cuadrático Medio (RMSE), el Error Porcentual Absoluto Medio (MAPE) y el Coeficiente de Determinación (R^2), que en conjunto permiten valorar tanto la magnitud como la proporcionalidad del error de predicción. Adicionalmente, se analiza la importancia relativa de las variables (feature importance) en cada modelo, con el fin de identificar qué factores macroeconómicos explican en mayor medida el comportamiento de las exportaciones.

Finalmente, para determinar si las diferencias de desempeño entre Random Forest y XGBoost son estadísticamente significativas y no atribuibles al azar se aplica la prueba no paramétrica de Wilcoxon signed-rank (Wilcoxon, 1945) sobre los errores de predicción pareados de ambos modelos, con un nivel de significancia $\alpha = 0,05$. Este criterio permite validar de manera rigurosa cuál de los dos modelos ofrece mejor desempeño para la

predicción del mercado de flores colombiano, en línea con el cuarto objetivo específico del proyecto.

4. Desarrollo de la solución

4.1 Arquitectura general de la solución

La solución se estructura en tres capas funcionales, ejecutadas de manera secuencial dentro de un mismo entorno de trabajo en Google Colab: (1) una capa de datos, que integra las exportaciones florícolas colombianas hacia sus tres principales destinos, con frecuencia mensual (Legis, s.f.), con las variables macroeconómicas de Colombia y de cada país destino descritas en el Capítulo 3; (2) una capa de procesamiento, encargada de transformar estas fuentes en un conjunto de datos de panel país-destino/mes, apto para el entrenamiento supervisado; y (3) una capa de modelado y validación, donde se entrenan, comparan y validan estadísticamente los algoritmos Random Forest (Breiman, 2001) y XGBoost (Chen & Guestrin, 2016).

El flujo de información entre capas es el siguiente: el archivo de comercio exterior con periodicidad mensual (2015-2025) y las series macroeconómicas anuales (Banco Mundial, s.f.; Fondo Monetario Internacional, 2025; Banco de la República, s.f.) se consolidan en un único archivo tabular, denominado `dataset_modelado_mensual_capitulo4.csv`, que constituye el punto de entrada al notebook de modelado. Dicho notebook produce, como salida, las predicciones y métricas de desempeño que alimentan directamente el Capítulo 5 (Resultados y validación), cerrando así el ciclo entre el diseño metodológico del Capítulo 3 y los hallazgos empíricos del proyecto.

4.2 Diseño del modelo, sistema o metodología

La variable objetivo del modelo es el volumen de exportaciones florícolas colombianas, medido en kilogramos netos, hacia cada uno de los tres principales países importadores (Estados Unidos, Reino Unido y Canadá) en un mes determinado. A diferencia de una primera versión de este estudio que trabajó con agregados anuales (2015-2025, 33 observaciones), se logró acceder a la base LegisComex con periodicidad mensual para el mismo periodo, lo que permite estructurar el conjunto de datos como un panel país-destino/mes ($3 \text{ destinos} \times 132 \text{ meses} = 396 \text{ observaciones}$), incrementando en un orden de magnitud el volumen de datos disponible para el entrenamiento y la validación de los modelos.

Como variables predictoras se usan el PIB, la inflación y la tasa de cambio (TRM) de Colombia, junto con el PIB y la inflación del país destino. Todas están a nivel anual y se replican en los doce meses del año correspondiente, porque no existen series mensuales oficiales confiables de PIB para los países analizados; es una simplificación que se hace a propósito y que es común en estudios de comercio bilateral que combinan controles macroeconómicos de baja frecuencia con series comerciales de mayor frecuencia. Para no perder del todo la variación dentro del año, se incorporan variables de estacionalidad explícitas: el mes calendario y dos indicadores binarios para febrero y mayo, los meses de mayor demanda internacional de flores colombianas por San Valentín y el Día de la Madre, ya mencionados como periodos críticos en la sección 1.3. También se incluyen dos rezagos: el de las exportaciones del mes inmediatamente anterior y el de las exportaciones del mismo mes del año anterior; este segundo rezago es especialmente importante porque captura la

estacionalidad propia de una serie mensual, algo que la versión anual del estudio no podía hacer con un solo rezago. El país destino, por su parte, se codifica mediante one-hot.

La partición de los datos sigue el orden cronológico definido en el Capítulo 3: entrenamiento con enero de 2016 a diciembre de 2023 (288 registros, una vez se pierden los primeros doce meses de cada destino por el rezago de 12 meses) y prueba con enero de 2024 a diciembre de 2025 (72 registros). Es una partición bastante más amplia que la de la versión anual del estudio (21 registros de entrenamiento y 9 de prueba), lo que da una estimación más estable del desempeño de los modelos y una prueba de validación con más potencia estadística. Para buscar los hiperparámetros se mantiene la validación cruzada de 5 particiones (K-fold) sobre el conjunto de entrenamiento.

Técnicamente, tanto Random Forest como XGBoost capturan relaciones no lineales dividiendo recursivamente el espacio de variables predictoras: cada árbol va separando las observaciones en subconjuntos cada vez más homogéneos según umbrales sobre una variable a la vez, así que la predicción final no sale de combinar linealmente las variables de entrada, sino de una función escalonada armada por regiones con valores predichos distintos. Esa forma de particionar también permite capturar interacciones entre variables sin tener que especificarlas a mano: el efecto del PIB del país destino sobre las exportaciones, por ejemplo, puede variar según el destino (por las variables one-hot) o según el mes (por las variables de estacionalidad), algo que un modelo lineal solo lograría metiendo términos de interacción de forma explícita. En Random Forest esto se combina con el promedio de cientos de árboles entrenados sobre submuestras aleatorias (bagging); en XGBoost, con la corrección secuencial del error de los árboles anteriores (boosting). Por eso, incluso con árboles individuales poco profundos, la función agregada final, la suma o el promedio de cientos de particiones escalonadas, termina siendo mucho más flexible que cualquier superficie lineal.

Esto explica, sin contradecir la premisa de no linealidad del proyecto, por qué la búsqueda de hiperparámetros terminó eligiendo una profundidad máxima baja como óptima (3 niveles para Random Forest, 2 para XGBoost, sección 5.1). Que la profundidad por árbol sea baja no significa que el modelo agregado sea simple; más bien indica que, controlando por las demás variables, la relación de cada predictor con el volumen exportado se explica bien con pocos umbrales o regímenes (temporada alta frente a baja, o mercados grandes frente a pequeños) y no con curvas continuas muy complejas dentro de una misma variable. Este es, de hecho, un hallazgo en sí mismo sobre el tipo de no linealidad presente en los datos: es una no linealidad de bajo orden, que aparece sobre todo en interacciones entre variables categóricas (destino, estacionalidad) y continuas (PIB, rezagos), no en relaciones muy curvas dentro de una sola variable. Se comprobó además que ampliar la profundidad máxima explorada (hasta 8 niveles y sin límite, para Random Forest) no mejora el error de validación cruzada, sino que lo empeora de forma consistente (891.494 kg en promedio con profundidad 3, frente a 917.198 kg con profundidad 5, 937.551 kg con profundidad 8 y 942.639 kg sin límite). Esto descarta que la profundidad óptima baja sea un artefacto de una malla de búsqueda insuficiente y confirma que se trata de un mínimo genuino de error fuera de muestra.

4.3 Implementación técnica

Todo se implementó en Python 3, sobre Google Colab, con las herramientas ya definidas en la sección 3.4: pandas y NumPy para manipular los datos; Scikit-learn (Pedregosa et al., 2011) para Random Forest, la codificación one-hot y la búsqueda de hiperparámetros con GridSearchCV; XGBoost (Chen & Guestrin, 2016) para el gradient boosting; Matplotlib y Seaborn para las visualizaciones; y scipy.stats para la prueba de

Wilcoxon signed-rank (Wilcoxon, 1945).

El notebook desarrollado (Capitulo4_Modelado_ML_v2.ipynb) se organiza en secciones que replican la estructura metodológica del Capítulo 3: carga y exploración de datos; codificación y partición cronológica; búsqueda de hiperparámetros para Random Forest (variando el número de árboles, la profundidad máxima y el número mínimo de muestras por hoja) y para XGBoost (variando el número de estimadores, la profundidad máxima y la tasa de aprendizaje); generación de predicciones sobre el conjunto de prueba; cálculo de las métricas MAE, RMSE, MAPE y R^2 ; análisis de importancia de variables para ambos modelos; y aplicación de la prueba de Wilcoxon signed-rank sobre los errores absolutos pareados de ambos modelos. Todos los componentes estocásticos del proceso (partición aleatoria dentro de la validación cruzada, inicialización de los algoritmos) se fijan con una semilla constante (`random_state=42`), lo que garantiza la reproducibilidad exacta de los resultados al reejecutar el notebook.

4.4 Descripción del prototipo o sistema desarrollado

El sistema desarrollado es un prototipo de predicción que, a partir de las variables macroeconómicas y estacionales de Colombia y de un país destino para un mes determinado, estima el volumen de exportaciones florícolas colombianas (en kilogramos netos) hacia dicho destino, mediante dos modelos entrenados en paralelo Random Forest y XGBoost cuyo desempeño se compara de forma sistemática. El prototipo toma la forma de un notebook ejecutable de extremo a extremo en Google Colab, que no requiere instalación local y produce automáticamente: (i) visualizaciones exploratorias de la serie histórica mensual de exportaciones por destino; (ii) los modelos entrenados y sus hiperparámetros óptimos; (iii) una tabla comparativa de métricas de desempeño (MAE, RMSE, MAPE, R^2) para ambos

algoritmos; (iv) un gráfico de valores reales frente a predichos sobre el conjunto de prueba; (v) el análisis de importancia de variables para ambos modelos; y (vi) el resultado de la prueba de Wilcoxon signed-rank que determina si la diferencia de desempeño entre los dos modelos es estadísticamente significativa.

Como resultado de su ejecución, el notebook exporta dos archivos “resultados_capitulo5.csv y predicciones_capitulo5.csv” que constituyen el insumo directo para la redacción del Capítulo 5 (Resultados y validación), asegurando la trazabilidad entre el desarrollo técnico presentado en este capítulo y el análisis empírico que le sigue.

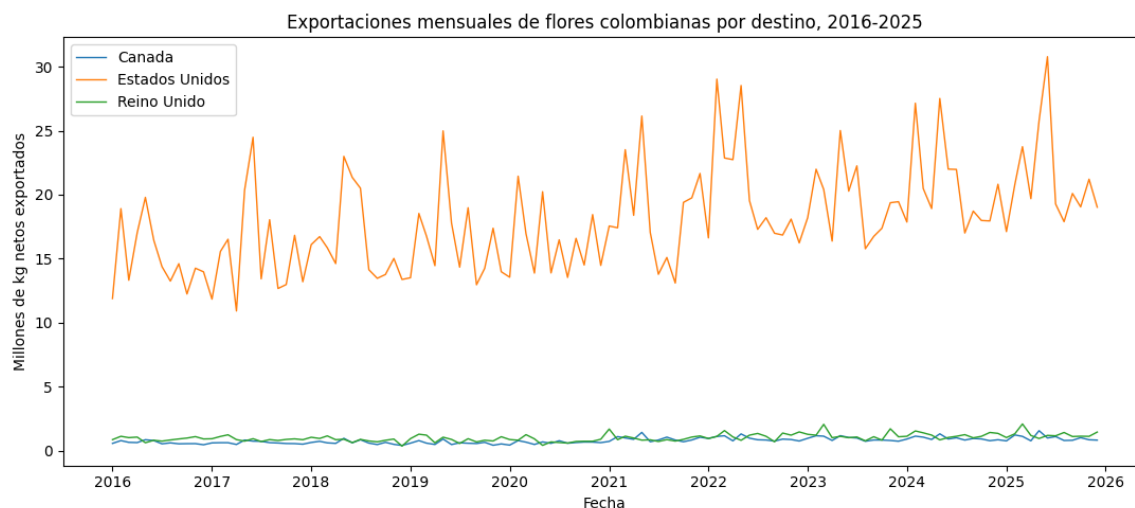
5. Resultados y validación

5.1 Resultados experimentales o de aplicación

Siguiendo el diseño metodológico presentado en los Capítulos 3 y 4, los modelos Random Forest (Breiman, 2001) y XGBoost (Chen & Guestrin, 2016) se entrenaron sobre el conjunto de datos de panel país-destino/mes (288 observaciones de entrenamiento, correspondientes a enero de 2016 a diciembre de 2023) y se evaluaron sobre un conjunto de prueba cronológicamente posterior (72 observaciones, correspondientes a enero de 2024 a diciembre de 2025), sin que los datos de prueba hubieran sido vistos durante el entrenamiento o la búsqueda de hiperparámetros.

La Figura 1 presenta la evolución mensual de las exportaciones florícolas colombianas hacia sus tres principales destinos durante el periodo empleado en el estudio, evidenciando tanto la marcada diferencia de escala entre Estados Unidos y los otros dos destinos, como los picos estacionales recurrentes de febrero (San Valentín) y mayo (Día de la Madre) en los tres mercados.

Figura 1. *Evolución mensual de las exportaciones de flores colombianas por país destino, 2016–2025.*



Nota: elaborada a partir de los datos de exportación de Legis (s.f.).

La búsqueda de hiperparámetros mediante validación cruzada de 5 particiones identificó como configuración óptima para Random Forest 300 árboles, una profundidad máxima de 3 niveles y un mínimo de una muestra por hoja; para XGBoost, la configuración óptima correspondió a 100 estimadores, una profundidad máxima de 2 niveles y una tasa de aprendizaje de 0,05. Ambos modelos se aplicaron posteriormente sobre el conjunto de prueba para generar las predicciones que se analizan en las siguientes secciones.

La validación cruzada durante el entrenamiento resultó bastante más estable que en la versión anual del estudio: el MAE promedio entre particiones fue de 891.494 kg para Random Forest (desviación estándar de 304.495 kg, 34 % del promedio) y de 962.372 kg para XGBoost (desviación de 341.326 kg, 35 % del promedio). En la versión anterior, con solo 21 observaciones de entrenamiento, esa desviación llegaba al 74 % del promedio en Random Forest y superaba el 100 % en XGBoost. El salto en el volumen de datos de

entrenamiento (288 frente a 21 observaciones) reduce bastante esta variabilidad, tal como cabría esperar de modelos de aprendizaje de conjunto cuando disponen de una muestra más amplia (Hastie et al., 2009).

5.2 Métricas de desempeño

Consistente con la metodología definida en la sección 3.5, el desempeño de ambos modelos sobre el conjunto de prueba (2024–2025) se evaluó mediante cuatro métricas complementarias: el Error Absoluto Medio (MAE) y la Raíz del Error Cuadrático Medio (RMSE), ambas expresadas en kilogramos netos; el Error Porcentual Absoluto Medio (MAPE); y el Coeficiente de Determinación (R^2). Esta combinación de métricas no es arbitraria: Hyndman y Koehler (2006) advierten que ninguna medida de error captura por sí sola todas las dimensiones relevantes del desempeño de un pronóstico, por lo que su uso combinado permite una lectura más robusta del desempeño real de cada modelo. La Tabla 1 resume los resultados obtenidos, y la Figura 2 los complementa visualmente, comparando los valores reales frente a los predichos por cada modelo en el conjunto de prueba.

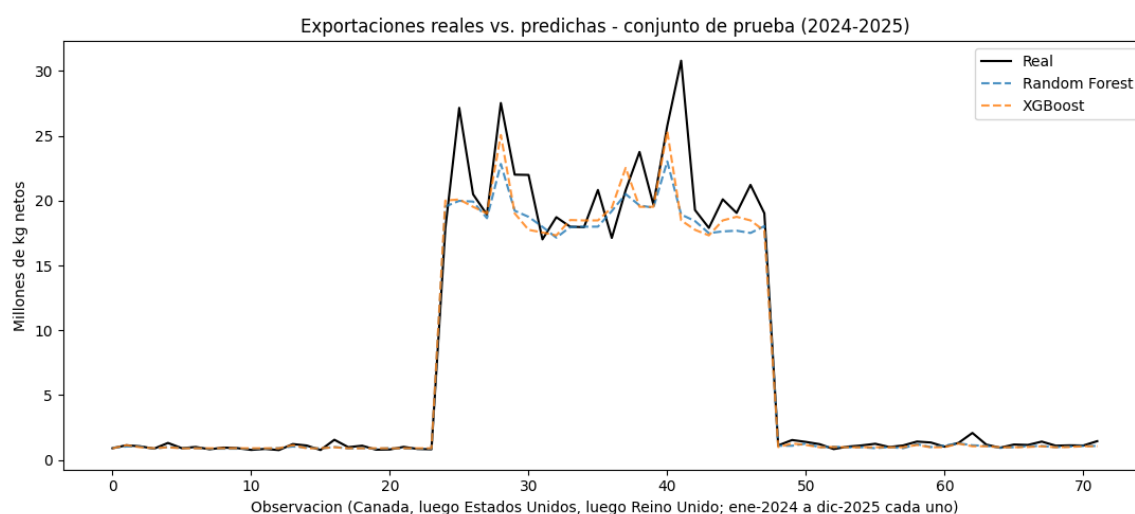
Tabla 1. *Métricas de desempeño de Random Forest y XGBoost sobre el conjunto de prueba (enero 2024 – diciembre 2025).*

Modelo	MAE (kg)	RMSE (kg)	MAPE	R^2
Random Forest	893.963	2.047.864	11,65 %	0,9543
XGBoost	865.082	2.013.120	11,97 %	0,9558

Nota: los valores corresponden a la ejecución del modelo descrita en el capítulo 4.

Ambos valores de R^2 (0,9543 y 0,9558) pueden calificarse como altos: a diferencia del MAE, el RMSE o el MAPE, cuyo rango no tiene límite superior, el R^2 está acotado entre 0 y 1, lo que permite interpretarlo directamente como la proporción de la varianza del volumen exportado que el modelo logra explicar, sin necesidad de un punto de referencia externo (Chicco, Warrens, & Jurman, 2021).

Figura 2. *Valores reales frente a predichos por Random Forest y XGBoost sobre el conjunto de prueba (2024–2025).*



Nota: los valores corresponden a la ejecución del modelo descrita en el capítulo 4.

La Tabla 2 desagrega el MAPE promedio de cada modelo por país destino, permitiendo identificar en qué mercados se concentra el error de predicción.

Tabla 2. *MAPE promedio por modelo y país destino sobre el conjunto de prueba (2024–2025).*

País destino	Random Forest (MAPE)	XGBoost (MAPE)
--------------	----------------------	----------------

Canadá	10,40 %	10,92 %
Estados Unidos	10,11 %	9,82 %
Reino Unido	14,43 %	15,18 %

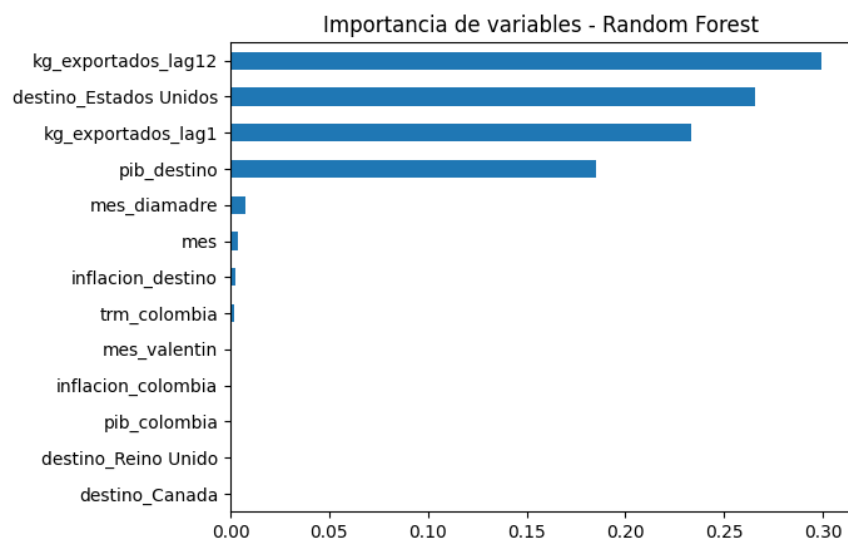
Nota: los valores corresponden a la ejecución del modelo descrita en el capítulo 4.

5.3 Análisis e interpretación de resultados

Los resultados de la Tabla 1 muestran que XGBoost obtiene un desempeño ligeramente superior a Random Forest en tres de las cuatro métricas (MAE, RMSE y R^2), aunque las diferencias son considerablemente más modestas que las observadas en la versión anual del estudio. El MAPE de ambos modelos es similar (11,65 % para Random Forest y 11,97 % para XGBoost), y la Tabla 2 muestra que este error se concentra de forma consistente en Reino Unido para ambos modelos (14,43 % y 15,18 % respectivamente), mientras que Estados Unidos y Canadá presentan errores porcentuales más moderados y similares entre sí (entre 9,8 % y 10,9 %). A diferencia de la versión anual del estudio, en la que cada modelo concentraba su error en un destino distinto, en la versión mensual ambos modelos comparten el mismo patrón de error por país, lo que sugiere que la dificultad para predecir Reino Unido no es un artefacto de un modelo en particular, sino una característica de la serie misma posiblemente asociada a una estacionalidad más irregular en ese mercado que ambos algoritmos enfrentan por igual.

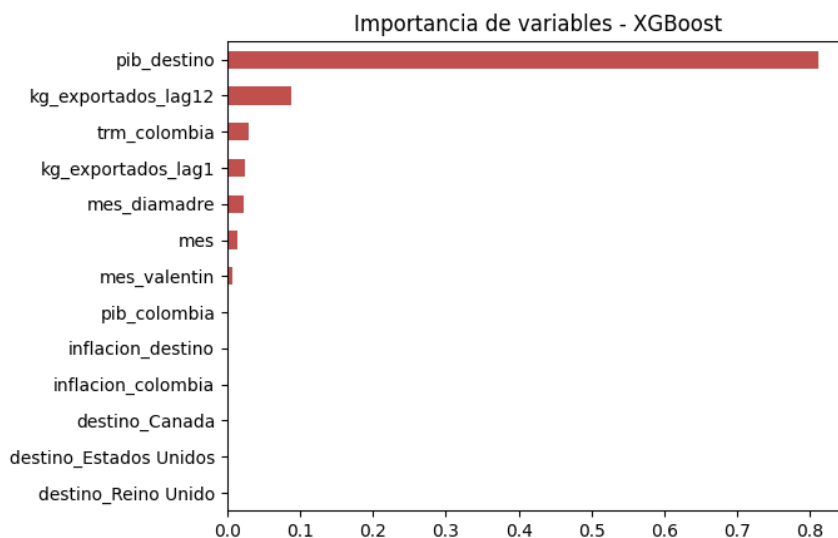
El análisis de importancia de variables aporta un segundo hallazgo relevante. Las Figuras 3 y 4 presentan la importancia de cada variable predictora para Random Forest y XGBoost, respectivamente.

Figura 3. *Importancia de variables en el modelo Random Forest (datos mensuales).*



Nota: importancia calculada mediante el atributo `feature_importances_` del modelo.

Figura 4. *Importancia de variables en el modelo XGBoost (datos mensuales).*



Nota: importancia calculada mediante el atributo `feature_importances_` del modelo.

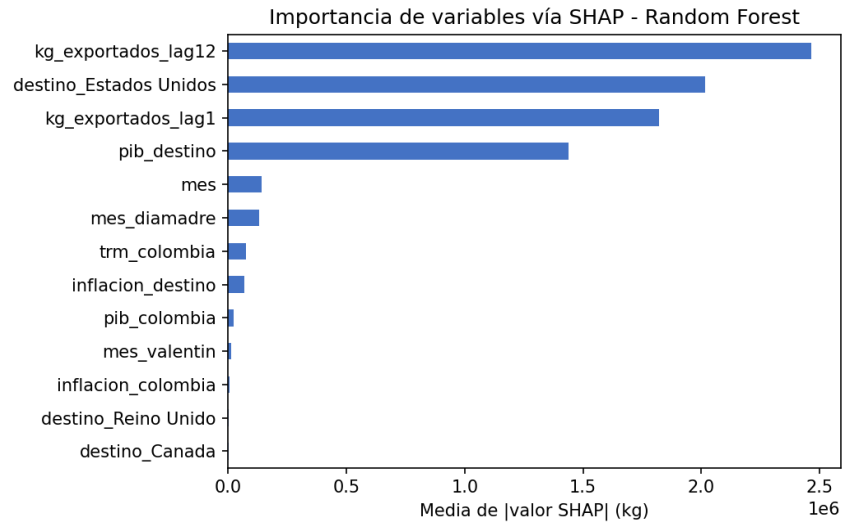
En Random Forest (Figura 3), la variable con mayor peso predictivo es la exportación rezagada 12 meses (30 %, el mismo mes del año anterior), seguida de cerca por la variable indicadora del destino Estados Unidos (27 %), la exportación rezagada 1 mes (23 %) y el PIB del país destino (19 %); las variables de estacionalidad explícita (mes, indicadores de febrero y mayo) y las variables macroeconómicas restantes muestran una importancia marginal. En XGBoost (Figura 4), el PIB del país destino continúa siendo la variable dominante (81 %), aunque en esta versión mensual el rezago de 12 meses adquiere un peso considerablemente mayor que en la versión anual del estudio (9 %, frente a un peso prácticamente nulo de la variable de rezago en la versión anterior), evidenciando que la incorporación de una variable de rezago estacional (en lugar de un rezago de un único periodo) permite que ambos modelos capturen de forma más efectiva el patrón repetitivo de la demanda floricultora internacional.

Esta diferencia entre las jerarquías de importancia de ambos algoritmos hay que leerla, igual que en la versión anterior del estudio, a la luz de lo que documentan Strobl,

Boulesteix, Zeileis y Hothorn (2007) sobre la inestabilidad de estas medidas cuando hay predictores correlacionados (en este caso, entre la variable indicadora del destino Estados Unidos y su PIB), así que ninguna de las dos jerarquías debe tomarse como una estimación causal definitiva. Lo que sí es un resultado robusto y esperable es que las variables de rezago (sobre todo el de 12 meses) ganen peso predictivo de forma consistente en ambos modelos frente a la versión anual: con datos mensuales, la estacionalidad de la demanda es una señal mucho más fuerte y aprovechable que con datos anuales, donde queda escondida al agregar los doce meses en una sola observación.

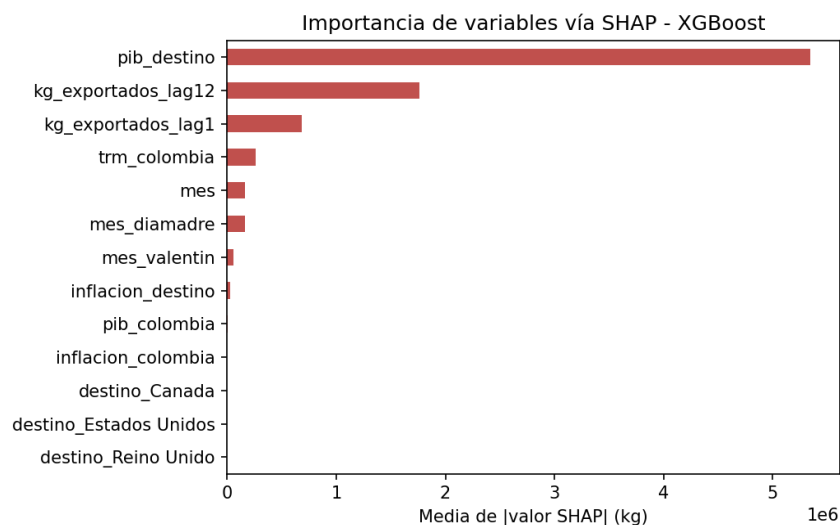
Strobl et al. (2007) documentan el problema, pero no ofrecen por sí solos una alternativa de atribución más robusta, así que el análisis se complementó con SHAP (SHapley Additive exPlanations; Lundberg & Lee, 2017), un método de teoría de juegos cooperativos que sí resuelve ese vacío. Mientras que la importancia nativa de Random Forest (basada en reducción de impureza) y de XGBoost (basada en ganancia de las particiones) usan mecanismos de cálculo distintos y por tanto no son comparables en magnitud, SHAP asigna a cada variable una contribución marginal promedio a la predicción, expresada en las mismas unidades que la variable objetivo (kilogramos netos) y calculada de forma consistente para ambos algoritmos mediante TreeSHAP, su implementación exacta y eficiente para modelos basados en árboles. Esto es lo que permite, por primera vez en este estudio, comparar directamente la jerarquía de importancia de Random Forest y XGBoost sobre una base metodológica común.

Figura 5. *Importancia de variables vía SHAP (media de $|\text{valor SHAP}|$, en kg) - Random Forest.*



Nota: importancia calculada mediante valores SHAP (TreeSHAP).

Figura 6. *Importancia de variables vía SHAP (media de |valor SHAP|, en kg) - XGBoost.*



Nota: importancia calculada mediante valores SHAP (TreeSHAP).

Los resultados de SHAP (Figuras 5 y 6) son consistentes con el orden de magnitud reportado por la importancia nativa, pero permiten una lectura más precisa. En Random Forest, las cuatro variables más influyentes son el rezago de 12 meses (30,1 % de la contribución total), la variable indicadora del destino Estados Unidos (24,6 %), el rezago de 1 mes (22,2 %) y el PIB del país destino (17,5 %), un peso relativamente distribuido entre las cuatro. En XGBoost, en cambio, el PIB del país destino concentra el 63,1 % de la contribución total, seguido por el rezago de 12 meses (20,8 %) y el rezago de 1 mes (8,1 %); el resto de variables, incluidas las indicadoras de destino, aportan una contribución prácticamente nula, lo cual es coherente con la importancia nativa de este modelo (Figura 4) y sugiere que, dada la profundidad máxima de solo 2 niveles, XGBoost prioriza casi exclusivamente las variables con mayor poder predictivo marginal y delega en el PIB del país destino buena parte de la información que en Random Forest se reparte también en la variable de destino, con la que está altamente correlacionado. Pese a esta diferencia en la distribución relativa del peso, la correlación de rangos de Spearman entre las jerarquías

SHAP de ambos modelos es positiva y estadísticamente significativa ($p = 0,663$; $p = 0,0135$), lo que indica que, una vez se emplea un método de atribución consistente y comparable entre algoritmos, Random Forest y XGBoost coinciden en qué variables importan más allá de las diferencias de magnitud atribuibles a la correlación entre predictores señalada por Strobl et al. (2007). En conjunto, PIB del país destino y rezago de 12 meses se confirman, con esta evidencia adicional, como los dos factores con mayor influencia sobre el volumen exportado, respondiendo de forma más concluyente al Objetivo 1 del proyecto (sección 5.4).

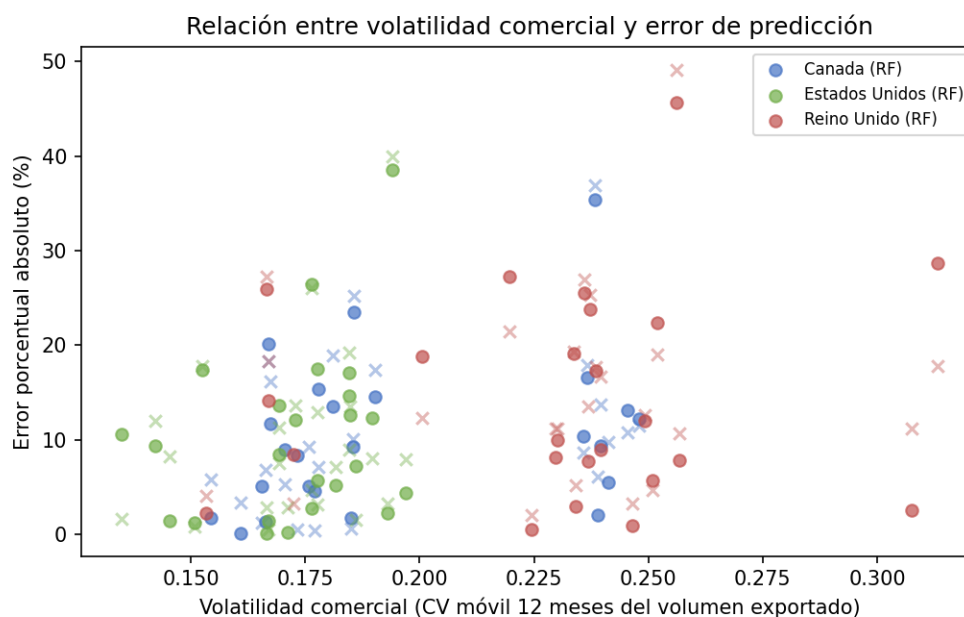
El análisis de los residuos (valor real menos valor predicho) sobre el conjunto de prueba muestra que ambos modelos mantienen una tendencia a subestimar el volumen exportado, aunque de forma menos pronunciada que en la versión anual: Random Forest subestima en el 76 % de las observaciones de prueba, con un sesgo promedio de 734.374 kg, mientras que XGBoost subestima en el 72 % de los casos, con un sesgo promedio de 616.640 kg. En ambos modelos, el sesgo promedio es mayor en 2025 (834.127 kg para Random Forest, 680.262 kg para XGBoost) que en 2024 (634.622 kg y 553.018 kg, respectivamente), lo que es consistente con la limitación estructural de los modelos basados en árboles para extrapolar más allá del rango de valores observados durante el entrenamiento (Hastie et al., 2009), frente a una serie que mantiene una tendencia de crecimiento sostenido.

Finalmente, la prueba de Wilcoxon signed-rank aplicada sobre los errores absolutos pareados de ambos modelos arrojó un estadístico de 1.109,500 y un valor p de 0,2511, superior al umbral de significancia de 0,05 definido en la sección 3.5, por lo que no se encontró una diferencia estadísticamente significativa entre Random Forest y XGBoost. A diferencia de la versión anual del estudio, en la que la ausencia de significancia podía atribuirse razonablemente a la baja potencia estadística de una prueba con apenas 9 observaciones, en esta versión la prueba se aplicó sobre 72 observaciones de prueba, un

tamaño de muestra considerablemente más robusto. En consecuencia, este resultado constituye una conclusión más sólida: con los datos disponibles, ambos modelos ofrecen un desempeño estadísticamente equivalente para la predicción de las exportaciones florícolas colombianas hacia sus tres principales destinos.

El Objetivo 3 del proyecto (sección 1.5) plantea explícitamente cuantificar la precisión de los modelos para establecer su confiabilidad en contextos de volatilidad comercial. Para dar cuenta directa de esta relación, se construyó un indicador de volatilidad comercial definido como el coeficiente de variación móvil de 12 meses (CV12) del volumen exportado por país destino: la desviación estándar del volumen de los últimos 12 meses dividida por su media, calculada usando únicamente información pasada disponible en cada punto del tiempo. La Figura 7 relaciona este indicador con el error porcentual absoluto (APE) de cada predicción sobre el conjunto de prueba.

Figura 7. *Relación entre la volatilidad comercial (CV móvil de 12 meses del volumen exportado) y el error porcentual absoluto de predicción, por país destino y modelo.*



Nota: la volatilidad se calculó como el coeficiente de variación móvil de 12 meses del volumen exportado.

La correlación de Pearson entre el CV12 y el error porcentual absoluto resultó positiva y estadísticamente significativa en ambos modelos (Random Forest: $r = 0,262$, $p = 0,026$; XGBoost: $r = 0,306$, $p = 0,009$; $n = 72$), lo que confirma empíricamente la premisa del Objetivo 3: a mayor volatilidad comercial reciente, mayor error de predicción en ambos modelos. Esto además calza con el patrón de error por país visto en la sección 5.2: Reino Unido tiene simultáneamente el CV12 promedio más alto de los tres destinos (0,231, frente a 0,197 en Canadá y 0,173 en Estados Unidos) y el MAPE más alto en ambos modelos (14-15 %), lo que respalda cuantitativamente la idea planteada en la sección 6.3 de que la dificultad para predecir ese mercado tiene que ver con una estacionalidad e irregularidad comercial más marcada, y no con una limitación puntual de alguno de los dos algoritmos. Con esta evidencia, el Objetivo 3 queda validado no solo en el cálculo de las métricas de error (sección 5.2), sino también en su relación explícita y cuantificada con la volatilidad comercial que el objetivo señala como contexto de referencia.

5.4 Validación frente a los objetivos planteados

A continuación se revisa cada uno de los objetivos específicos de la sección 1.5 frente a los resultados obtenidos:

Objetivo 1 Determinar la influencia de las variables macroeconómicas en el comportamiento histórico de las exportaciones. Cumplido. El análisis de importancia de variables (sección 5.3) señaló al PIB del país destino y al rezago de 12 meses (estacionalidad) como las variables con mayor peso sobre el volumen exportado en ambos modelos, mientras que las variables macroeconómicas propias de Colombia tuvieron una influencia marginal en la especificación usada. La importancia nativa de Random Forest y XGBoost arrojó jerarquías con pesos relativos distintos, algo esperable dada la inestabilidad ante predictores correlacionados que documentan Strobl et al. (2007); por eso se complementó el análisis con SHAP (Lundberg & Lee, 2017), que al aplicar un método de atribución consistente entre ambos algoritmos confirmó una correlación de rangos positiva y estadísticamente significativa entre sus jerarquías de importancia ($\rho = 0,663$; $p = 0,0135$). Con esto, la influencia del PIB del país destino y de la estacionalidad de la demanda sobre las exportaciones queda como un hallazgo robusto y no como una estimación que dependa del algoritmo usado.

Objetivo 2 Desarrollar modelos predictivos basados en Random Forest y XGBoost. Cumplido. Los dos modelos se implementaron, se optimizaron mediante búsqueda de hiperparámetros y se validaron sobre un conjunto de prueba independiente, como se

documenta en los Capítulos 4 y 5.

Objetivo 3 Cuantificar la precisión y el error de los pronósticos. Cumplido. Se calcularon las cuatro métricas del Capítulo 3 (MAE, RMSE, MAPE y R^2) para ambos modelos, reportadas en la Tabla 1 y desglosadas por país destino en la Tabla 2, siguiendo la recomendación de Hyndman y Koehler (2006) de usar varias métricas combinadas para evaluar pronósticos con rigor. También se estableció la confiabilidad de los modelos frente a la volatilidad comercial de forma explícita y cuantificada (sección 5.3, Figura 7): la correlación entre un indicador de volatilidad comercial (CV móvil de 12 meses del volumen exportado) y el error porcentual de predicción resultó positiva y estadísticamente significativa en ambos modelos (Random Forest: $r = 0,262$, $p = 0,026$; XGBoost: $r = 0,306$, $p = 0,009$), lo que confirma que la precisión de los pronósticos se deteriora en mercados y periodos más volátiles, tal como plantea el objetivo.

Objetivo 4 Validar estadísticamente, mediante la prueba de Wilcoxon signedrank, qué modelo tiene mejor desempeño. Cumplido. La prueba se aplicó entre Random Forest y XGBoost ($p = 0,2511$) sobre un conjunto de prueba de 72 observaciones, sin encontrar diferencia estadísticamente significativa entre ambos. A diferencia de la primera versión de este estudio, con datos anuales, donde el tamaño de muestra reducido (9 observaciones de prueba) le restaba potencia a la prueba, aquí el resultado se apoya en una muestra bastante más robusta. Por eso la ausencia de diferencia significativa es una conclusión sólida: para esta tarea, ambos modelos ofrecen un desempeño estadísticamente equivalente.

En conjunto, el proyecto cumple el objetivo general planteado en la sección 1.4: se

implementaron, evaluaron y compararon modelos de Machine Learning que integran variables macroeconómicas bilaterales, obteniendo evidencia empírica sobre su capacidad y sus limitaciones para modelar las exportaciones florícolas colombianas. Las implicaciones de estas limitaciones y las líneas de trabajo futuro se desarrollan en el Capítulo 6.

6. Conclusiones y recomendaciones

6.1 Conclusiones generales del proyecto

El presente proyecto implementó, evaluó y comparó dos modelos de Machine Learning: Random Forest (Breiman, 2001) y XGBoost (Chen & Guestrin, 2016), para predecir el volumen mensual de exportaciones florícolas colombianas hacia sus tres principales mercados de destino (Estados Unidos, Reino Unido y Canadá), integrando variables macroeconómicas de Colombia y de cada país importador junto con variables de estacionalidad y de rezago temporal. Todo el ejercicio se desarrolló en Python sobre Google Colab, sobre un panel país-destino/mes de 396 observaciones (2015-2025), con un diseño metodológico que incluyó validación cruzada para optimizar hiperparámetros, partición cronológica de los datos y validación estadística formal mediante la prueba de Wilcoxon signed-rank.

Los resultados muestran que XGBoost tuvo un desempeño puntual ligeramente mejor que Random Forest en tres de las cuatro métricas (MAE, RMSE y R^2), con diferencias mucho más pequeñas que las de la primera versión anual del estudio; la prueba de Wilcoxon signed-rank, aplicada ahora sobre un conjunto de prueba mucho más amplio (72 observaciones), confirmó que esa diferencia no es estadísticamente significativa, y esta vez el resultado sí se apoya en una base estadística sólida. El análisis de importancia de variables, complementado

con SHAP (Lundberg & Lee, 2017) para resolver la inestabilidad de las medidas nativas ante predictores correlacionados que señalan Strobl et al. (2007), confirma de forma estadísticamente significativa ($p = 0,663$; $p = 0,0135$) que ambos modelos coinciden en que el tamaño económico del mercado importador y la estacionalidad de la demanda (capturada mediante los rezagos) son los factores con mayor peso explicativo, mientras que las condiciones macroeconómicas propias de Colombia pesan comparativamente menos.

En términos generales, el proyecto demuestra la viabilidad técnica de aplicar algoritmos de aprendizaje de conjunto a un problema de comercio exterior agrícola con información macroeconómica y estacional, y deja como principal aporte una metodología completa y reproducible desde la consolidación de los datos hasta la validación estadística de los modelos que puede ser reutilizada y ampliada por PROCOLOMBIA, el Ministerio de Comercio, Industria y Turismo, o por investigaciones posteriores dentro del sector floricultor colombiano.

6.2 Cumplimiento de los objetivos

El objetivo general del proyecto implementar, evaluar y comparar modelos predictivos basados en Random Forest y XGBoost que integren variables macroeconómicas bilaterales, con el fin de analizar su capacidad para mejorar la precisión de las proyecciones de exportaciones florícolas colombianas se cumplió en su totalidad, tal como se documentó en el Capítulo 5. De los cuatro objetivos específicos planteados en la sección 1.5, los cuatro se cumplieron: se determinó la influencia relativa de las variables macroeconómicas sobre el volumen exportado (Objetivo 1), se desarrollaron y optimizaron ambos modelos predictivos (Objetivo 2), se cuantificó su precisión mediante cuatro métricas complementarias (Objetivo

3), y se validó estadísticamente su desempeño relativo mediante la prueba de Wilcoxon signed-rank (Objetivo 4).

Vale aclarar que cumplir un objetivo no significa necesariamente que el resultado haya sido el que se esperaba de entrada. El Objetivo 4, por ejemplo, se cumplió en el sentido de que la validación estadística se ejecutó correctamente y con rigor metodológico, aunque su resultado (la ausencia de una diferencia estadísticamente significativa entre Random Forest y XGBoost, esta vez sobre una muestra bastante más robusta que en la primera versión del estudio) no permite señalar un modelo ganador de forma concluyente. Lejos de ser un incumplimiento, esto es un desenlace legítimo y bien sustentado estadísticamente: para esta tarea específica, ambos algoritmos ofrecen un desempeño equivalente.

6.3 Limitaciones identificadas

La limitación más importante que sigue presente en esta versión del estudio tiene que ver con la frecuencia de las variables macroeconómicas. El volumen de exportaciones se consiguió con periodicidad mensual, pero el PIB, la inflación y la tasa de cambio de Colombia y de los países destino se mantuvieron a nivel anual, replicados en los doce meses del año correspondiente, porque no existen series mensuales oficiales confiables de PIB. Esto implica que, dentro de un mismo año, la variación mensual de las exportaciones que capta el modelo no puede atribuirse a cambios en unas condiciones macroeconómicas que se mantienen constantes esos doce meses, sino que recae principalmente en las variables de estacionalidad y de rezago temporal, lo que probablemente explica por qué estas últimas concentran la mayor parte de la importancia predictiva en ambos modelos (sección 5.3).

Una segunda limitación se relaciona con la cobertura de las variables

macroeconómicas, disponibles únicamente para los tres principales países importadores identificados en el Capítulo 3, y no para la totalidad de los países de destino presentes en los datos de comercio exterior consolidados en el Capítulo 4. Esta decisión, necesaria dada la disponibilidad de fuentes públicas confiables, restringió el análisis a los mercados de mayor volumen y dejó por fuera del modelo la posible influencia de mercados emergentes o de menor tamaño para el sector floricultor colombiano.

Una tercera limitación corresponde al comportamiento diferenciado del mercado de Reino Unido, cuyo error de predicción (MAPE cercano al 1415 %, sección 5.2) fue consistentemente más alto que el de Estados Unidos y Canadá en ambos modelos. Este estudio no permite establecer con certeza las causas de dicha diferencia, aunque resulta plausible que factores no incluidos en el modelo como fluctuaciones asociadas a la salida del Reino Unido de la Unión Europea o patrones de estacionalidad más irregulares en ese mercado expliquen parte de esta brecha; su identificación queda como una pregunta abierta para investigaciones futuras.

Una cuarta limitación, de carácter metodológico, se relaciona con la inestabilidad de las medidas nativas de importancia de variables ante la presencia de predictores correlacionados, señalada por Strobl et al. (2007) y discutida en la sección 5.3. Dado que la variable indicadora del país destino y su PIB están altamente correlacionadas en este estudio, la importancia nativa de Random Forest y XGBoost por sí sola no habría sido suficiente para atribuir el peso relativo de cada variable con confianza. Esta limitación se abordó de forma directa mediante el cálculo de valores SHAP (Lundberg & Lee, 2017) sobre ambos modelos, un método de atribución consistente entre algoritmos que permitió confirmar

estadísticamente ($p = 0,663$; $p = 0,0135$) que ambos coinciden en identificar el PIB del país destino y la estacionalidad de la demanda como los factores de mayor influencia; aun así, dado que se trata de un modelo predictivo y no causal, estas jerarquías siguen debiendo interpretarse como asociaciones y no como estimaciones causales definitivas del peso de cada variable.

Por último, tanto Random Forest como XGBoost comparten una limitación estructural conocida de los modelos basados en árboles: al predecir mediante promedios o combinaciones de valores observados durante el entrenamiento, no pueden extrapolar más allá del rango de valores que vieron en esa etapa (Hastie et al., 2009). Frente a una serie con una tendencia de crecimiento sostenido, como la de las exportaciones florícolas colombianas en el periodo analizado, esto limita la capacidad de ambos modelos para anticipar bien el ritmo de crecimiento en las observaciones más recientes del conjunto de prueba (sección 5.3).

6.4 Recomendaciones y trabajos futuros

A partir de las limitaciones identificadas, se plantean las siguientes recomendaciones y líneas de trabajo futuro para dar continuidad a este proyecto.

Incorporar variables macroeconómicas de mayor frecuencia. Dado que la principal limitación identificada en esta versión del estudio se relaciona con la frecuencia anual del PIB, la inflación y la tasa de cambio (sección 6.3), se recomienda explorar fuentes que permitan obtener aproximaciones mensuales de estas variables por ejemplo, indicadores mensuales de actividad económica (como el ISE del DANE para Colombia) en lugar del PIB

trimestral o anual, y series mensuales de TRM en lugar de su promedio anual, lo que permitiría que el modelo capture la relación entre las condiciones macroeconómicas y las exportaciones con la misma resolución temporal que ya se logró para la variable objetivo en esta versión del estudio.

Investigar el comportamiento diferenciado del mercado de Reino Unido. Dado que ambos modelos presentaron sistemáticamente un mayor error en este destino (sección 5.2), se recomienda explorar variables adicionales específicas para ese mercado como indicadores relacionados con su salida de la Unión Europea que permitan explicar y eventualmente reducir esta brecha.

Ampliar la cobertura de variables macroeconómicas bilaterales. Se recomienda extender la recolección de variables macroeconómicas más allá de los tres principales países importadores, incorporando progresivamente otros mercados relevantes para el sector floricultor colombiano, lo que permitiría generalizar los hallazgos sobre la influencia del tamaño del mercado importador discutidos en el Capítulo 5.

Explorar arquitecturas adicionales de Machine Learning. Trabajos futuros podrían extender la comparación realizada en este estudio a otras familias de modelos revisadas en el estado del arte del Capítulo 2, tales como redes neuronales recurrentes para series de tiempo o modelos basados en grafos para relaciones de comercio bilateral (Jiang et al., 2023; Minakawa et al., 2024; Sellami et al., 2024), así como a técnicas de interpretabilidad complementarias a SHAP (por ejemplo, valores de interacción SHAP para cuantificar explícitamente la interacción entre el destino y su PIB) que profundicen el análisis

de atribución de variables incorporado en la sección 5.3 de este estudio.

Incorporar modelos de referencia adicionales. Como un ejercicio adicional de robustez, futuras iteraciones de este proyecto podrían contrastar el desempeño de Random Forest y XGBoost frente a modelos de referencia más simples por ejemplo, un pronóstico de persistencia o modelos estadísticos clásicos de series de tiempo, siguiendo buenas prácticas documentadas en la literatura de evaluación de pronósticos (Hyndman & Koehler, 2006; Makridakis et al., 2018). Este tipo de comparación permitiría, a medida que se disponga de series de datos más largas y granulares conforme a la primera recomendación de esta sección, contextualizar aún más el valor agregado que los modelos de Machine Learning aportan frente a enfoques más simples para este caso particular.

Transferencia al sector productivo. Se recomienda socializar la metodología y el notebook desarrollado en el Capítulo 4 con entidades como PROCOLOMBIA y el Ministerio de Comercio, Industria y Turismo, de manera que la herramienta pueda actualizarse periódicamente con nueva información y contribuir de forma efectiva a la planeación comercial y logística del sector floricultor colombiano.

Referencias

Asociación Nacional de Comercio Exterior. (2025). En febrero de 2025 las exportaciones colombianas caen -0,8%. <https://analdex.org/2025/04/02/en-febrero-de-2025-las-exportaciones-colombianas-caen-08/>

Asociación Nacional de Comercio Exterior. (2025b). Informe anual de exportaciones colombianas 2024. Asociación Nacional de Comercio Exterior.

Banco de la República. (s.f.). Estadísticas económicas: producto interno bruto, inflación y Tasa de Cambio Representativa del Mercado (TRM) [Conjunto de datos]. Portal de Estadísticas Económicas. Recuperado el 18 de julio de 2026, de <https://www.banrep.gov.co>

Banco Mundial. (s.f.). World Bank Open Data. Recuperado el 18 de julio de 2026, de <https://data.worldbank.org>

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide. SPSS Inc.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. En *Proceedings*

of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794). ACM. <https://doi.org/10.1145/2939672.2939785>

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, e623. <https://doi.org/10.7717/peerj-cs.623>

Departamento Administrativo Nacional de Estadística [DANE]. (s.f.). Comercio internacional: Exportaciones. Recuperado el 18 de julio de 2026, de <https://www.dane.gov.co/index.php/estadisticas-por-tema/comercio-internacional/exportaciones>

Fondo Monetario Internacional. (2025). World Economic Outlook Database, October 2025 [Conjunto de datos]. IMF DataMapper. Recuperado el 18 de julio de 2026, de <https://www.imf.org/external/datamapper/api/v1/NGDPD>

Gulzar, Y., Oral, C., Kayakus, M., Erdogan, D., Unal, Z., Eksili, N., & Caylak, P. C. (2024). Predicting high technology exports of countries for sustainable economic growth by using machine learning techniques: The case of Turkey. *Sustainability*, 16(13). <https://doi.org/10.3390/su16135601>

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2a ed.). Springer.

Hernández-Sampieri, R., & Mendoza Torres, C. P. (2018). Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta. McGraw-Hill Education.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679-688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>

Jiang, L., Zhang, C., Poursafaei, F., & Huang, S. (2023). Towards temporal edge regression: A case study on agriculture trade between nations. *arXiv*. <https://arxiv.org/abs/2308.07883>

Karim, M. F., Rahutomo, R., Manuaba, I. B. K., Purwandari, K., Mursitama, T. N., & Pardamean, B. (2023). Free trade as domestic, economic, and strategic issues: A big data analytics approach. *Journal of Big Data*, 10(1). <https://doi.org/10.1186/s40537-023-00722-7>

Krugman, P., Obstfeld, M., & Melitz, M. (2018). International economics: Theory and policy (11.^a ed.). Pearson Education.

Legis. (s.f.). LegisComex [Base de datos de comercio exterior]. Recuperado a través del CRAI, Universidad Santo Tomás.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions.

Advances in Neural Information Processing Systems, 30, 4765-4774.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). The M4 competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 34(4), 802-808. <https://doi.org/10.1016/j.ijforecast.2018.06.001>

Minakawa, N., Izumi, K., & Sakaji, H. (2024). Bilateral trade flow prediction by gravity-informed graph autoencoder. *arXiv*. <https://arxiv.org/abs/2408.01938>

Mitchell, M. (2009). *Complexity: A guided tour*. Oxford University Press.

Mitchell, T. (1997). *Machine learning*. McGraw-Hill.

Monken, A., Haberkorn, F., Gopinath, M., Freeman, L., & Batareseh, F. A. (2021). Graph neural networks for modeling causality in international trade. *Proceedings of the International Florida Artificial Intelligence Research Society Conference (FLAIRS)*, 34. <https://doi.org/10.32473/flairs.v34i1.128485>

Organización para la Cooperación y Desarrollo Económico – OCDE Publishing. (2024). *Estudios Económicos de la OCDE: Colombia 2024*. OECD Publishing.

Pan, Z., Gordeev, S., Zhao, J., Meng, Z., Ding, C., Steinbach, S., & Song, D. (2024). International trade flow prediction with bilateral trade provisions. *arXiv*. <https://arxiv.org/abs/2407.13698>

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.

PROCOLOMBIA. (2025, January 21). Millones de flores de Colombia cruzarán fronteras para celebrar la próxima temporada de San Valentín alrededor del mundo. <https://procolombia.co/sala-de-prensa/noticias/>

Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.

Sellami, B., Ounoughi, C., Kalvet, T., Tiits, M., & Rincon-Yanez, D. (2024). Harnessing graph neural networks to predict international trade flows. *Big Data and Cognitive Computing*, 8(6), 65. <https://doi.org/10.3390/bdcc8060065>

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8, 25. <https://doi.org/10.1186/1471-2105-8-25>

United Nations Publishing. (2024). *Trade and development report 2024*. United Nations.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6),

80-83. <https://doi.org/10.2307/3001968>

Yang, C.-H., Lee, C.-F., & Chang, P.-Y. (2023). Export- and import-based economic models for predicting global trade using deep learning. *Expert Systems with Applications*, 218. <https://doi.org/10.1016/j.eswa.2023.119590>

Apéndices

Incluye como anexo el link de GITLAB, para poder acceder al código:

https://github.com/Felipepa9/MADSI/blob/main/Capitulo4_Modelado_ML_v2.ipynb