

Universidad Francisco Marroquín
Facultad de Ciencias Económicas
Data Science

Modelo de Clasificación para Identificación de Mercados Atractivos para Guatemala

Enfoque AGEXPORT - Inteligencia de Mercados

Andrea Soto 20210349

Guatemala, Mayo 2025

[Repositorio en GitHub](#)

En un contexto globalizado, Guatemala busca fortalecer su presencia en mercados internacionales, especialmente para productos clave como café, banano y cardamomo. Sin embargo, la falta de datos específicos sobre la demanda de estos productos en otros países y las barreras arancelarias dificultan la identificación de mercados prioritarios. Este proyecto propone un **modelo de clasificación binaria** que, utilizando el dataset [World Export & Import Dataset \(1989-2023\)](#) y técnicas de feature engineering, que prediga qué países son más atractivos para las exportaciones guatemaltecas. A través de un enfoque basado en datos —como el volumen de importaciones, aranceles y crecimiento económico—, el modelo priorizará mercados con alta demanda, bajas barreras comerciales y potencial de crecimiento, alineándose con la misión de AGEXPORT de impulsar la inteligencia de mercados. Los resultados permitirán tomar decisiones estratégicas con menor incertidumbre, incluso ante limitaciones de datos específicos por producto.

Caso de Uso (Enfoque AGEXPORT)

AGEXPORT necesita priorizar mercados para productos guatemaltecos (café, banano, cardamomo) basándose en:

- **Demanda real:** Volumen de importaciones de productos *similares* (no hay datos directos).
- **Acceso al mercado:** Aranceles y políticas comerciales.
- **Crecimiento económico:** Países con mayor potencial futuro.

Limitaciones del Dataset:

- No tiene datos por producto específico (café, banano).
- Variables como Import Product Share (%) están incompletas o son genéricas.

Solución Propuesta:

1. **Feature Engineering estratégico** para aproximar la demanda de productos guatemaltecos.
2. **Target binario** con criterios de AGEXPORT y estándares internacionales.

Dataset y Feature Engineering

- **Origen:** [World Export & Import Dataset \(1989-2023\)" \(Kaggle\).](#)
- **Variables clave:**
 - Partner Name (países).
 - Import (US\$ Thousand) (demanda total).
 - AHS MaxRate (%) (barreras arancelarias).
 - Country Growth (%) (crecimiento económico).

Feature Engineering para Productos Guatemaltecos

Como no hay datos directos de café/banano, **se crearán variables** :

Nueva Variable	Fórmula/Justificación	Objetivo
Demanda_Productos_GT	$\text{Import (US\$ Thousand)} * 0.3$ (supuesto: 30% corresponde a productos similares a los de GT).	Estimar demanda para café, banano, etc.
Arancel_Especifico_GT	$\text{AHS MaxRate (\%)} * 1.2$ (supuesto: productos GT enfrentan aranceles 20% más altos).	Simular barreras para productos GT.
TLC_Con_GT	1 si el país tiene TLC con Guatemala (ej: EE.UU., UE), 0 si no.	Priorizar mercados con acceso preferencial.

Target: Mercado Atractivo

Un país es "atractivo" si cumple:

- Demanda alta:** $\text{Demanda_Productos_GT} > \text{percentil } 75$.
- Barreras bajas:** $\text{AHS MaxRate (\%)} < 10\%$ o $\text{TLC_Con_GT} = 1$.
- Crecimiento estable:** $\text{Country Growth (\%)} > 3\%$ (media global).

Justificación:

- Transparencia:** Los criterios se basan en estándares de la OMC y reportes de AGEXPORT.
- Accionabilidad:** Un mercado "atractivo" es aquel donde Guatemala puede competir.

Validación Cruzada:

- Comparar predicciones con mercados prioritarios conocidos (ej: EE.UU., México).
- Métricas: **Precision** (evitar falsos positivos) y **Recall** (captar todos los mercados relevantes).