

Determinantes económicos, sociales y demográficos de  
la migración en Colombia: Un enfoque probabilístico  
y machine learning (2023-2025)

Alan Steven Ríos Munar  
Juan Esteban Samaniego Poveda  
Sebastian Casas Poloche

Universidad Externado de Colombia  
IA con aplicaciones en economía  
Lina María Castro

### Objetivo:

El objetivo principal del proyecto es estimar la probabilidad de que un ciudadano colombiano migre dentro del territorio nacional y determinar hacia qué región se dirige esa migración, ya sea entre departamentos o municipios. En particular, se busca identificar qué factores económicos, demográficos, sociales y asociados a la violencia influyen en la decisión de migrar, cuantificando la probabilidad individual de desplazamiento y analizando los patrones de destino. De esta manera, el estudio no solo pretende explicar las causas de la migración interna, sino medir de forma precisa la probabilidad de migrar y establecer cuáles son las regiones con mayores flujos de salida y de entrada de población.

### Descripción:

Para cumplir este objetivo, se desarrollará un modelo econométrico y de machine learning que permita caracterizar las variables que inciden en la migración y estimar tanto la probabilidad de que una persona migre como la probabilidad del destino elegido. Se compararán distintos modelos probabilísticos para determinar cuál ofrece mayor capacidad predictiva, y se implementará un algoritmo de K-means para agrupar individuos con características similares e identificar perfiles con mayor propensión a migrar. Además, se construirán mapas de calor a nivel departamental y municipal que permitan visualizar los flujos de migración y clasificar los territorios según su nivel de movilidad y probabilidad estimada de migración, utilizando la integración de la GEIH 2023–2025 como base principal de análisis.

### Desafíos

Unir en una sola base de datos la GEIH (gran encuesta integrada de hogares) realizada por el DANE de los años 2023, 2024, 2025, para poder tener la mayor cantidad de datos y lograr una mayor tasa de efectividad en la predicción de la probabilidad de migración y sus causas. Teniendo en cuenta que la información dentro de cada encuesta está dada por meses y que cada una de ellas representa un año específico.

### Entrega de valor

Este proyecto aporta una innovación metodológica y aplicada al estudio de la migración interna en Colombia al integrar modelos probabilísticos, uso de machine learning y análisis espacial con base en la unificación de la GEIH del año 2023, 2024, 2025. A diferencia de los estudios de causalidad, la investigación estima la probabilidad individual de migración interna, identifica perfiles migratorios mediante algoritmos de clasificación como K-means, compara distintos modelos probabilísticos para determinar cuál ofrece mayor capacidad predictiva e incorpora análisis territorial a través de mapas de calor y flujos interregionales.

## Stakeholders

- Ministerio de Trabajo
- Planeación Nacional
- Gobernaciones
- Alcaldías
- Programas de desarrollo regional
- Planeación urbana

## Técnicas que se utilizaran

- Análisis exploratorio de datos
- Clustering
- Machine learning
- Modelos probabilísticos
- Análisis espacial
- Integración y transformación de datos

## Fuentes de datos

Gran Encuesta Integrada de hogares (GEIH) 2023-2025

Banco de la República

- Indicadores regionales de actividad económica.
- Desempleo regional.
- Informes económicos departamentales.

Departamento Administrativo Nacional de Estadística

- PIB departamental.
- Tasa de desempleo por ciudad.
- Índice de pobreza monetaria.
- Índice de pobreza multidimensional.

## Variables

Demográficas:

- Edad
- Sexo
- Tamaño del hogar
- Estado civil

Socioeconómicas:

- Ingresos
- Condición laboral
- Tipo de ocupación
- Formalidad/informalidad
- Nivel educativo

Territoriales:

- Cambio de municipio de residencia.
- Municipio de residencia anterior

Migración:

- Migración reciente (últimos 12 meses)
- Motivo de migración

Violencia:

- Desplazamiento por violencia asociada al conflicto armado.
- Desplazamiento por violencia no asociada al conflicto armado.
- Motivo de migración por inseguridad o amenazas.