



**TRABAJO DE GRADO**  
**Opción Seminario-Diplomado.**

**ALGORITMO COMPUTACIONAL PARA EL ANÁLISIS Y TOMA DE DECISIONES EN  
DATOS DE LA MIGRACION DE VENEZOLANOS EN PAMPLONA, UTILIZANDO  
ESTRATEGIAS DE MACHINE LEARNING**

Corporación Universitaria Remington.  
Faculta de ingeniería  
Ingeniería de Sistemas

**SERGIO BAYONA**  
Tutor: Juan Carlos Briñez de León  
Opción de Trabajo de grado Seminario-Diplomado.  
2024.

## **Agradecimientos**

Agradezco a Dios por darme la sabiduría durante este proceso, así mismo agradezco a mi familia por brindar su apoyo incondicional en cada una de las etapas, pero en especial agradezco a mi esposa la cual fue un pilar fundamental durante toda mi formación académica y a quien debo la fortaleza en los momentos más difíciles porque sin su soporte este trabajo no sería posible.

**Tabla de Contenidos**

Resumen.....	6
Palabras clave.....	6
Marco Conceptual Y Contextual .....	7
Marco Conceptual.....	7
Marco Contextual.....	7
Pregunta problema .....	13
Acercamiento a los datos: .....	13
Descripción de variables. ....	13
Posibles aplicaciones. ....	14
Aproximaciones con gráficos. ....	14
Objetivos: .....	17
Objetivo general.....	17
Objetivos específicos. ....	17
Desarrollo e implementación del aprendizaje.....	18
Designar los datos .....	18
Modelado de datos .....	18
Mapeo de datos según el modelo .....	21
Implementación en contextos reales .....	24
Conclusiones .....	25

	4
Referencias.....	26

## Tabla de figuras

<b>Figura 1</b> Migrantes desde Venezuela en Colombia (2014-2019) .....	10
<b>Figura 2</b> Distribución de inmigrantes y no inmigrantes desde Venezuela por sector de actividad económica .....	11
<b>Figura 3</b> Diagrama de tora genero .....	15
<b>Figura 4</b> Grafico de dispersión grupo étnico .....	15
<b>Figura 5</b> Grafico de edad .....	16
<b>Figura 6</b> Datos iniciales .....	19
<b>Figura 7</b> Dataset sin columnas innecesarias .....	20
<b>Figura 8</b> Dataset con todos los datos numéricos .....	20
<b>Figura 9</b> Algoritmo de clúster .....	21
<b>Figura 10</b> Dataset con los clusters .....	22
Figura 11 Centroides de los clústers .....	23
Figura 12 Comparación cluster por edad y etnia .....	23
<b>Figura 13</b> Nuevos usuarios a la base de datos .....	24

## **Resumen**

El actual proyecto implementa el machine Learning sobre una base de datos de migrantes venezolanos en el municipio de Pamplona, con el propósito de mejorar el proceso de toma de decisiones de una ONG.

Este entrenamiento se lleva a cabo gracias al modelado de los datos para ser de fácil comprensión por el modelo, ayudándole a identificar tendencias, con un menor consumo de recursos, para su posterior análisis mediante el método de aprendizaje automático clustering, el cual es especial para estos problemas donde se necesita la agrupación de datos.

Todo esto se desarrolla en empatía con el pueblo venezolano quien enfrenta una situación complicada y Norte de Santander se ha visto afectada por este contexto, pero en especial Pamplona al ser una ciudad de paso obligado para los migrantes pero que por sus condiciones demográficas complican la situación para ambas partes.

Por lo cual con este proyecto se contribuye a la toma de decisiones efectivas en ámbitos de elaboración de planes y focalización de recursos de la ONG involucrada.

## **Palabras clave**

Migración Venezolana, Dataset, Análisis de datos, Machine learning, Clasificación, Clustering.

## Marco Conceptual Y Contextual

### Marco Conceptual

**Machine learning:** “Rama de la inteligencia artificial que utiliza la computación científica, las matemáticas y la estadística a través de técnicas automatizadas para resolver problemas basados en clasificación, regresión y clustering” (Mamani Rodríguez, 2023, p. 275)

**Predicción:** “Posibilidad de anticipar algo que va a suceder”(Pérez Bedia & Rojas Segovia, 2020, p.5)

**Python:** Lenguaje de programación de alto nivel interpretado, multiparadigma utilizado para el análisis de datos científicos (Pérez Bedia & Rojas Segovia, 2020, pág.7)

**Algoritmo:** en sistemas es una (De Carvalho Aquino, 2020) forma organizada de solucionar una problemática mediante la computación (p.45)

**Inteligencia Artificial:** (Estupiñán Ricardo et al., 2021) “Nombre que se le asigna a una serie de tecnologías con características o capacidades que antes eran exclusivas del intelecto humano” (p.363).

**Dataset:** Rodríguez-García et al., (2020) conjunto de datos utilizados en inteligencia artificial para entrenar modelos de machine learning (p.20)

### Marco Contextual

El desarrollo tecnológico ha llegado con innumerables herramientas que buscan la automatización y facilitación del trabajo, es precisamente en la automatización de las labores diarias donde ingresa la cuarta gran revolución tecnológica, como lo explica Mori Castro, (2021) “En esta cuarta RI se han desarrollado nuevas tecnologías dentro de las cuales podemos

destacar la inteligencia artificial” (p.37). Es la evolución del aprendizaje automático en la creación de modelos de inteligencia, la automatización de los procesos, donde grandes organizaciones buscan la forma de potencializar sus acciones mediante los sistemas de Machine learning obteniendo soporte en los procesos de toma de decisiones.

Es importante tener presente la conceptualización de machine learning dada anteriormente, pues es la construcción de estos modelos mediante la multiplicidad de datos orientados a un objetivo, lo que permite del aprendizaje automático la herramienta fiel para las organizaciones en las tomas de decisiones, tal como lo expone (Mamani Rodríguez, 2023) definiendo al machine learning como aquella herramienta de procesamiento de datos, mediante matemáticas orientadas a la programación y estadística con el fin de obtener la toma de decisiones y simulación de casos mediante programas informáticos. (p.276)

Si bien está clara la respuesta a ¿qué es machine learning y su relevancia? Otro eje central del presente proyecto es el asunto social de Venezuela y su afectación a Colombia, en específico a la ciudad de Pamplona. De este tema hace mención Paz Noguera et al., (2021) citando al Banco Mundial (2018) Donde recapitula como Venezuela fue un país al cual multitudes de personas migraban por las oportunidades que en este se ofrecían, sin embargo, estas oportunidades cambiaron con las problemáticas económicas y conflictos políticos que atravesó el país, desencadena una ola de migración sin precedentes en la región, de venezolanos en búsqueda de mejores oportunidades. (p.73).

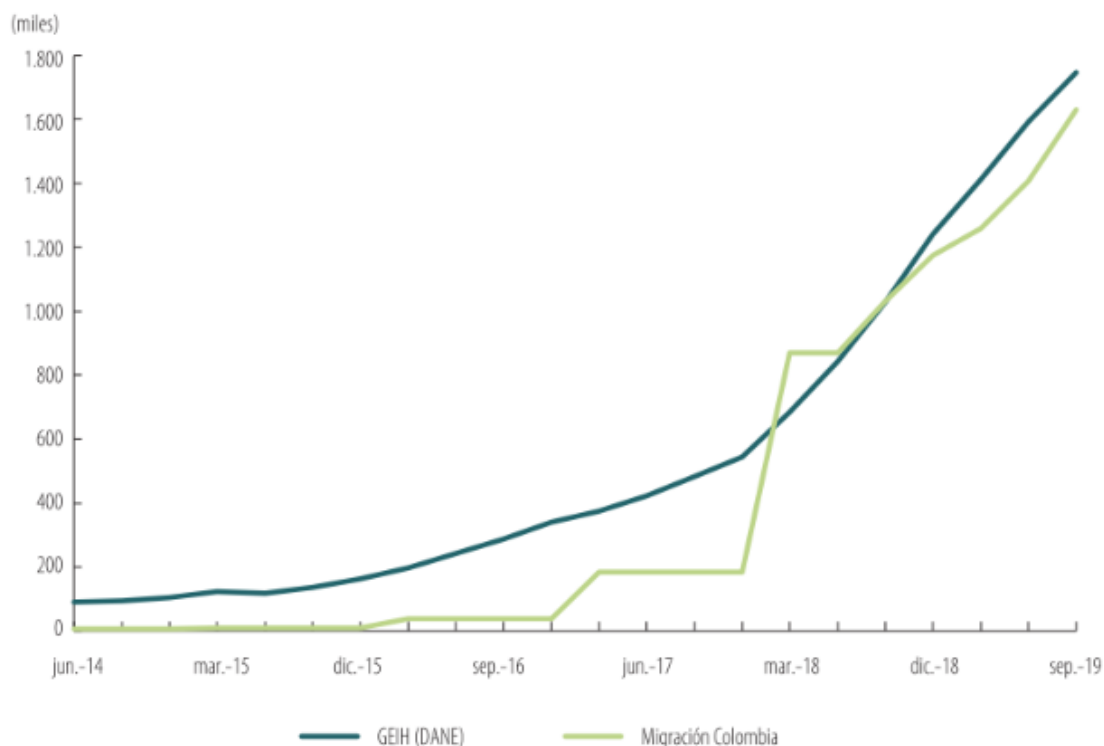
En su proceso de migración los venezolanos han decidió optar por innumerables destinos, pero por el contexto geográfico Colombia es un país de paso obligado en su deseo de buscar nuevas oportunidades, como lo explica (Zambrano Miranda et al., 2020) “La frontera entre Colombia y Venezuela comprende 2.219 kilómetros de proximidad geográfica y humana. El

territorio de Norte de Santander se destaca como la parte más dinámica y compleja.” (p. 261) por lo cual no se hace extraño que Arteada et al., (2017) comienzo su artículo con las palabras “El municipio de Pamplona Norte de Santander no es ajeno a esta problemática” (p.14) puesto que, por su ubicación geográfica alejada del epicentro de Norte de Santander, muchos llegarían a pensar que su afectación es menor en comparación con la capital de este departamento, pero la realidad es que por su demografía se han visto afectados por la diáspora de migrantes venezolanos, esto es claro con el comentario hecho por (Velandia Andrade, 2022) “Pamplona, Norte de Santander, no es ajeno a esta problemática por ser un paso obligado de la entrada de migrantes al país” (p.5).

En un análisis enfocado principalmente en el rol de pamplona y su relación con Venezuela, tratando de entender mejor las consecuencias del fenómeno migratorio a este municipio (Arteada et al., 2017) detalla, para pamplona su principal ingreso era el comercio de los productos cultivados y generados, productos que eran exportados a Venezuela para su comercialización, debido a la crisis muchos comercios cerraron, obligando a disminuir la demanda de mano de obra (p.23). Ahora es claro que Pamplona enfrentaba una crisis desde el cierre de la frontera en 2015, por lo cual, al momento del presente proyecto con la masiva llegada de venezolanos, la economía a declinado, creando una crisis económica, la cual se puede entender en la figura 1.

## Figura 1

*Migrantes desde Venezuela en Colombia (2014-2019)*

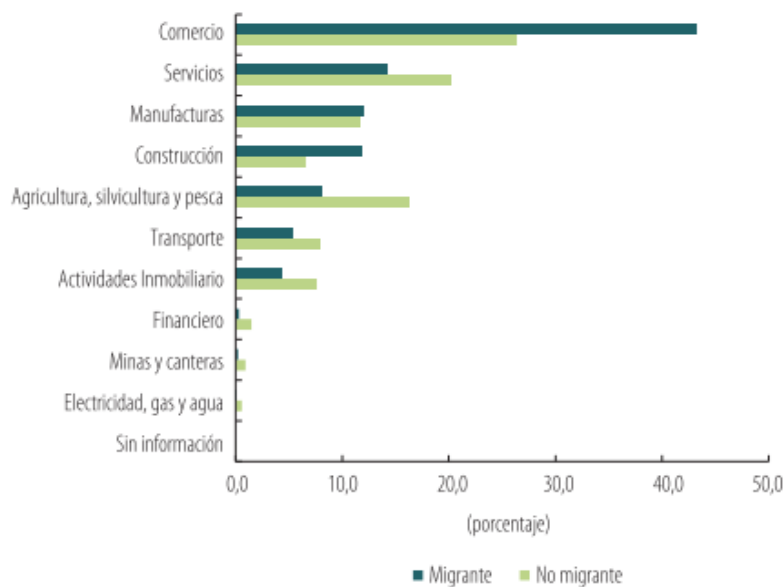


Nota: El aumento de venezolanos a Colombia ha ido creciendo de manera exponencial. Tomado de *Ensayos sobre política económica*. (p.7) Tribín Uribe et al., 2020, Banco de la república de Colombia.

A lo añadido anteriormente es importante entender el rol del migrante al llegar a Colombia, y de esto se encarga Paz Noguera et al., (2021) quien describe como el venezolano en su condición de migrante terminar ejerciendo labores “informales, principalmente, como vendedores ambulantes, lo que les permite el envío de remesas hacia Venezuela.” (p.72) véase figura 2.

**Figura 2**

*Distribución de inmigrantes y no inmigrantes desde Venezuela por sector de actividad económica*



Nota: La población migrante se concentra principalmente en los sectores comercio, manufactura y agricultura. Tomado de *Ensayos sobre política económica*. (p.14) Tribín Uribe et al., 2020, Banco de la república de Colombia.

En continuación con el actual proyecto donde se hace el planteamiento de incluir al machine learning en los procesos de toma de decisiones a nivel social, es imprescindible hacer una recopilación contextual del estado del arte con relación a machine en learning en ámbitos de toma de decisiones para contextos sociales, por lo cual un proyecto bastante interesante es el planteado por Arango Pastrana & Osorio Andrade, (2021) durante el cual utilizaron Machine learning para entrenar un algoritmo desarrollado en Python, el cual analiza los diferentes post de las redes sociales, y a partir de un algoritmo de lenguaje humano el software podía analizar el estado anímico de los usuarios, permitiendo entender como la política de aislamiento afectaba a

las personas y dando herramientas para que el gobierno de turno tomara decisiones sin afectar drásticamente el nivel emocional de las personas.

Lo anterior es una clara muestra de cómo las tecnologías, en especial el aprendizaje automático pueden ser un aliado en procesos delicados, no obstante, hay más ejemplos uno muy similar al anterior, pero en un contexto de migración es el desarrollado por González Rostani et al., (2023) quienes construyeron un algoritmo que analizaba los tweets para determinar el grado de xenofobia contra los migrantes venezolanos, este modelo a su vez fue adaptado y aplicado a otros contexto como refugiados por guerras, pero sin lugar a dudas es una herramienta de automatización útil a la hora de entender el comportamiento social de un país y su pensamiento con relación a una situación particular.

Otro ejemplo donde la inteligencia artificial cobra vida para facilitar el trabajo y contribuir en la toma de decisiones es el proyecto desarrollado por (Parada Parra, 2024) quienes apostaron por construir un modelo que diera soporte a la cadena de suministros usados para la agricultura colombiana, modelo que termino siendo un éxito, pues tras el ingreso de los datos relacionados con compra de suministros y el valor de estos en el tiempo, el modelo logro arrojar datos, los cuales permitieron a sus usuarios finales obtener más ganancias monetarias.

El aprendizaje automático o Machine Learning es una de las herramientas más importantes con las cuales se cuenta hoy en día, por esto se deben aprovechar para explotar al máximo su bondades, no importa el contexto en el cual se aplique, siempre será un soporte en la toma de decisiones, de modo que en coherencia con las investigaciones planteadas en párrafos anteriores se implementa un modelo de machine learning a un aspecto de índole socio económica, como es la migración de venezolanos sobre el municipio de Pamplona.

## **Pregunta problema**

¿Cómo un algoritmo de Machine Learning, aplicado a una base de datos sobre la migración venezolana a Pamplona en 2023, puede ayudar a la toma de decisiones para la integración de la población migrante en 2024?

### **Acercamiento a los datos:**

Los datos son proporcionados por la ONG Care Colombia, donde buscar mitigar las consecuencias de los diferentes conflictos sociales que Colombia enfrenta, entre los cuales se destaca la diáspora de venezolanos migrando a Colombia en búsqueda de mejores oportunidades.

Estos datos son de uso interno de la empresa y son provistos por empleados con la intención de lograr crear un modelo que facilite la creación programas de ayuda social con un mayor éxito en los necesitados. Por consiguiente, estos datos son provenientes de una base de datos mucho mayor, pero por políticas de tratamiento solo se brindó la actual base de datos.

### **Descripción de variables.**

Dentro de la base de datos proporcionada se encuentran dados de diferentes valores como son cualitativos y cuantitativos, la idea es poder contar con una base de datos con variables fijas y determinantes, pero con varios datos dentro de estas variables, todo esto con la finalidad de lograr un procesamiento certero por parte del modelo.

Dentro de las variables se encuentra el departamento y municipio donde residen lo migrantes encuestado, seguido por el sexo y su fecha de nacimiento, así como la edad, estos datos dan una contextualización de la persona.

Posterior se cuentan con datos para conocer la razón de su migración, si es perteneciente algún grupo étnico y que tipo de beneficiario es, este último teniendo en cuenta que la base de datos pertenece a una ONG

### **Posibles aplicaciones.**

Entendiendo el origen de los datos, la idea es poder consolidar la ruta de acción para enlazar el modelo desarrollado a la base de datos grande, la cual está en constante evolución y cuenta con datos en streaming que permitirían un resultado mas eficaz en los datos salientes.

Así mismo este proyecto junto al modelo podría implementarse fuera de Google Colab, así de este una API consumible para la futura creación de una aplicación que se amigable con los empleados de la ONG garantizando su uso no solo en la creación de programas sociales de ayuda si no en el resto de las políticas internas de la empresa, como serian la administración de recursos tanto humanos, materiales y económicos.

### **Aproximaciones con gráficos.**

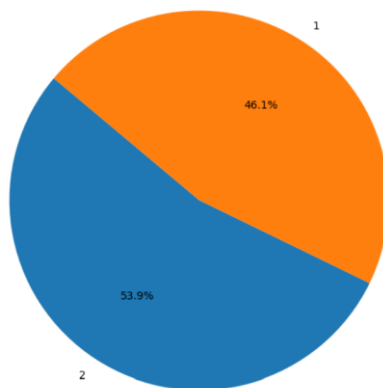
Para lograr un mejor acercamiento a los datos, se desarrolla en Python una descripción grafica de los datos seleccionados, los cuales se pueden observar en las graficas siguientes, donde se analizan datos como son las edades, la etnia y el género de los beneficiarios de la ONG.

En estos datos se resaltan las edades las cuales van desde los 21 años hasta los 94 años, pero teniendo una media de 40 años, de igual forma se denota también como la mayor parte de inmigrantes son de géneros masculino, lo cual tiene sentido si se entiende los medios por los cuales realizan la migración, siendo una de las mas comunes la caminata.

Por último, se analiza también la etnia entre la cual la menos señalada fue la mestiza y se conto con un grupo mayoritario de personas de etnia afrodescendiente e indígena.

**Figura 3**

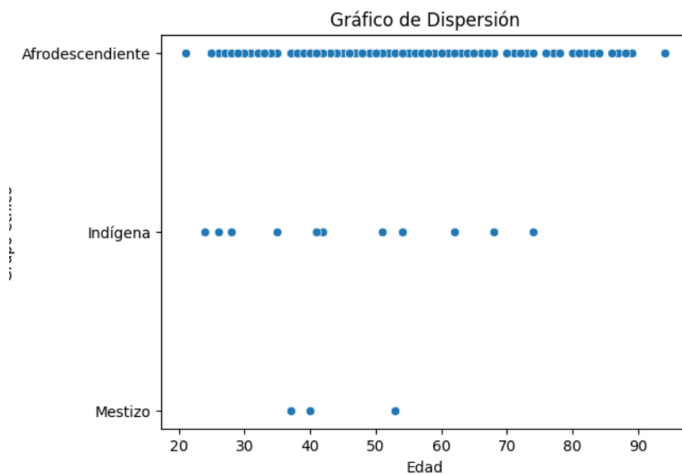
*Diagrama de tora genero*



Nota: se visualiza mayor cantidad de hombres que de mujeres Fuente: elaboración del autor

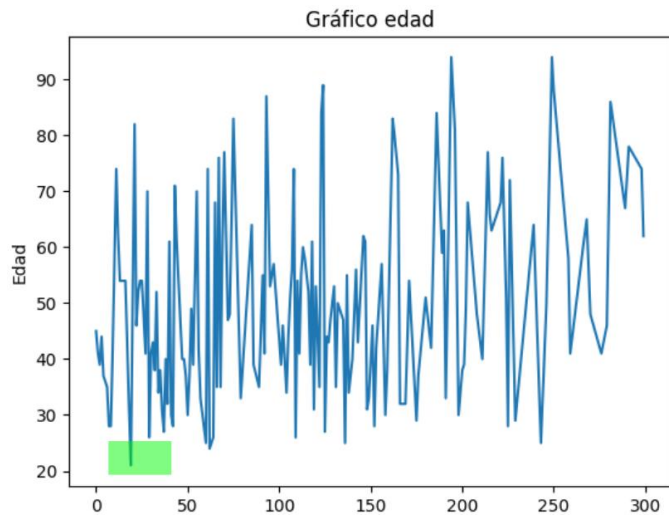
**Figura 4**

*Grafico de dispersión grupo étnico*



Nota: los grupos étnicos con mayor presencia en la base de datos son el indígena y el mestizo

Fuente: elaboración del autor.

**Figura 5***Gráfico de edad*

Nota: la edad menor es de 21 años y la mayor de 94, con un promedio entre 48 año. Fuente:

elaboración del autor

## **Objetivos:**

### **Objetivo general.**

Desarrollar e implementar un modelo de Machine Learning que, a partir de una base de datos sobre la migración venezolana a Pamplona en 2023, ayude a la toma de decisiones para la creación efectiva de programas de ayuda social.

### **Objetivos específicos.**

- Analizar y procesar la base de datos sobre la migración venezolana a pamplona en 2023 para obtener un conjunto de datos limpio y estructura adecuada para el modelado de Machine Learning.
- Seleccionar un modelo de Machine Learning apropiado que pueda procesar la base de datos sobre la migración venezolana a pamplona en 2023, con el objetivo de identificar patrones relevantes para la toma de decisiones
- Obtener conclusiones sobre la aplicación del modelo de Machine Learning, contribuyendo con los resultados a la toma de decisiones para la creación de los proyectos de a migrante durante el año 2024

## **Desarrollo e implementación del aprendizaje**

Para el desarrollo del proyecto y ejecución de los datos se seleccionó el método no supervisado de datos agrupados, dando la posibilidad de agrupar los datos en clusters creado automáticamente por Python mediante su librería sklearn, pero para lograr mejores resultados se tuvieron en cuenta los parámetros y pasos aconsejados por Galindo et al., (2022) quienes describen el proceso, y propone ciertos pasos a realizar con los datos cuando se desean obtener resultados asertivos y con mayor confiabilidad, 1. Designar los datos, 2 Modelado de los datos, 3 Mapear los datos y 4 Comprar los datos. (p.3)

### **Designar los datos**

El proceso de asignación de datos se centra esencialmente en la elección de la base de datos sobre la cual se entrenarán los modelos de predicción, es importante en este apartado entender, que los datos deben en su gran mayoría de tipo numérico para que el modelo pueda analizarlos de manera correcta.

Para el presente proyecto la base de datos ya ha sido provista por una ONG como se mencionó en apartados anteriores donde se relacionaron las variables que se tiene en cuenta.

### **Modelado de datos**

Para el modelado de datos fue necesaria la importación de la base de datos, para su manipulación, como se puede observar en la figura 6

## Figura 6

### Datos iniciales

```

import pandas as pd
from google.colab import files
uploaded = files.upload()
for filename in uploaded.keys():
    Dataset = pd.read_excel(filename)
Dataset.head(7)

```

Elegir archivos Base Migr...ion2 (1).xlsx

- Base Migracion2 (1).xlsx(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 22604 bytes, last modified: 28/3/2024 - 100% done
- Saving Base Migracion2 (1).xlsx to Base Migracion2 (1).xlsx

ID	Departamento	Municipio	Sexo	Fecha nacimiento	Edad	Perfil migratorio	Grupo étnico	Tipo de Beneficiario	
0	1	Norte de Santander	Pamplona	Hombre	1979-12-09	45	10-Desplazado	Afrodescendiente	Individuo
1	2	Norte de Santander	Pamplona	Hombre	1983-03-07	41	10-Desplazado	Afrodescendiente	Individuo
2	3	Norte de Santander	Pamplona	Hombre	1985-10-10	39	10-Desplazado	Afrodescendiente	Individuo
3	4	Norte de Santander	Pamplona	Mujer	1980-08-24	44	10-Desplazado	Afrodescendiente	Individuo
4	5	Norte de Santander	Pamplona	Mujer	1987-03-20	37	10-Desplazado	Afrodescendiente	Individuo
5	6	Norte de Santander	Pamplona	Mujer	1987-10-10	37	10-Desplazado	Afrodescendiente	Individuo
6	7	Norte de Santander	Pamplona	Hombre	1989-11-26	35	10-Desplazado	Afrodescendiente	Líder/lideresa Comunitaria

Fuente: elaboración del autor

Ahora si bien el modelado de la base de datos lleva varios pasos solo se resaltarán los más importantes, por ejemplo, en párrafos anteriores se mencionó la importancia de contar con datos numéricos, sin embargo, en la figura 6 se aprecian datos en cadenas de texto, además de datos innecesarios para la ejecución del modelo de clúster.

Lo primero sería corroborar los datos verificar que a este dataset no le falten de datos y posterior eliminar, las filas que no son necesarias como son el ID, Departamento y municipio, de igual forma se eliminarán también la fecha de nacimiento, puesto que se cuenta con la edad de las personas, en este orden de ideas el perfil migratorio se hace innecesario puesto que aunque la ONG trabaja con varios perfiles el entrenamiento de este modelo se centra en inmigrantes de Venezuela, al hacer esta depuración se está consiguiendo centralizar la atención del modelo en los datos realmente importantes.

**Figura 7***Dataset sin columnas innecesarias*

	Sexo	Edad	Grupo étnico	Tipo de Beneficiario
0	Hombre	45	Afrodescendiente	Individuo
1	Hombre	41	Afrodescendiente	Individuo
2	Hombre	39	Afrodescendiente	Individuo
3	Mujer	44	Afrodescendiente	Individuo
4	Mujer	37	Afrodescendiente	Individuo
5	Mujer	37	Afrodescendiente	Individuo
6	Hombre	35	Afrodescendiente	Líder/lideresa Comunitaria
7	Mujer	28	Afrodescendiente	Individuo
8	Hombre	28	Indígena	Líder/lideresa Comunitaria
9	Hombre	41	Indígena	Líder/lideresa Comunitaria

Nota: en esta vista solo se observa un ejemplo con diez datos. Fuente: elaboración del autor

Finalmente, como se necesita contar con datos numéricos, se hará la siguiente conversión de datos a un identificador numérico (véase fig.8):

Sexo: 1 = hombres, 2 = mujeres

Grupo étnico: 1 = afrodescendiente, 2 = indígena, 3 = mestizo

Tipo beneficiario: 1 = individuo, 2 = líder/lideresa comunitaria

**Figura 8***Dataset con todos los datos numéricos*

	Sexo	Edad	Grupo étnico	Tipo de Beneficiario
0	1	45	1	1
1	1	41	1	1
2	1	39	1	1
3	2	44	1	1
4	2	37	1	1
5	2	37	1	1
6	1	35	1	2

Fuente: elaboración del autor

## Mapeo de datos según el modelo

El modelo seleccionado fue el clustering, por lo cual ha llegado el momento de pasarle los datos ya listo al modelo para que el cree los clústeres según los datos existentes, para el modelo presente y a petición de la ONG se elaboran tres grupos con el fin de crear tres campañas de operación diferentes.

### Figura 9

#### *Algoritmo de clúster*

```
# Importando las bibliotecas necesarias
import numpy as np
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

# Transformando el dataset a formato numérico
Dataset_numerico = np.array(Dataset)

# Elegir el número de grupos
k = 3

# Generando y entrenando el modelo
Modelo_Cluster = KMeans(n_clusters=k, n_init = 'auto', random_state=42)
Modelo_Cluster.fit(Dataset_numerico)

# Prediciendo etiquetas de grupos en el conjunto de datos
Predicciones = Modelo_Cluster.fit_predict(Dataset_numerico)

# Asignando etiqueta como columna nueva al dataset
Dataset_2 = Dataset.copy()
Dataset_2['Grupo'] = Predicciones
Dataset_2.head(20)
```

Fuente: elaboración del autor

Una vez ejecutado el código obtenemos los tres grupos como se puede observar en la figura 10

## Figura 10

### *Dataset con los clústers*

	Sexo	Edad	Grupo étnico	Tipo de Beneficiario	Grupo
0	1	45	1	1	2
1	1	41	1	1	2
2	1	39	1	1	2
3	2	44	1	1	2
4	2	37	1	1	2
5	2	37	1	1	2
6	1	35	1	2	2
7	2	28	1	1	2
8	1	28	2	2	2
9	1	41	2	2	2
10	1	41	2	2	2
11	1	74	2	2	1
12	1	62	2	2	0

Fuente: elaboración del autor

Al ejecutar los centroides para analizar los datos se observan (véase fig.11) algunas peculiaridades, como que el factor principal para hacer la distinción entre los grupos fue la edad y el género, obteniendo un grupo de mayores de 35 años, otros mayores de 54 y uno final de mayores de 77 años, otro dato interesante es que la mayor cantidad de mujeres quedaron en los grupos de 35 y 54 años.

Pero entre todo esto el grupo étnico también tuvo relevancia pues se contaba con tres tipos de etnias, entre la cuales el grupo cero quedo en su mayoría con personas de etnia afrodescendiente e indígena al igual que el grupo uno, donde se puede ver que hubo más personas de etnia mestiza fue en el clúster dos.

Figura 11

*Centroides de los clústers*

Centroides:

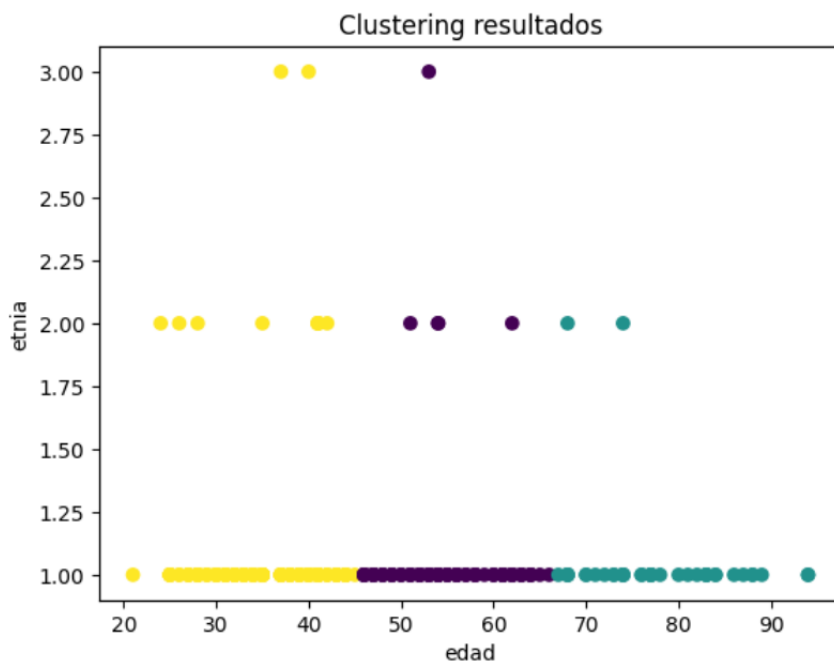
	Sexo	Edad	Grupo étnico	Tipo de Beneficiario
0	1.545455	54.872727	1.054545	1.018182
1	1.414634	77.804878	1.048780	1.024390
2	1.557047	35.402685	1.087248	1.033557

Fuente: elaboración del autor

Algún gráfico que los grupos permiten obtener son el de la edad con la etnia (véase fig.12) donde se observa que la mayoría de las personas son afrodescendientes y quienes son mestizos están en las edades de 40 a 55 años.

Figura 12

*Comparación clúster por edad y etnia*



Fuente: elaboración del autor

## Implementación en contextos reales

El modelo una vez entrenado comprobado con estos datos, se guardo el modelado y las reglas de los clústers para probar sus reacciones a nuevos usuarios y en que grupos les asignaban para lograr entender su funcionamiento y se desarrollan pruebas como se puede observar en la figura

### Figura 13

#### *Nuevos usuarios a la base de datos*

<pre># Ingresar datos del nuevo almacén (3 variables) Clientes_New=np.zeros((1,4)) Clientes_New[0,0]=2 #genero Clientes_New[0,1]=95 #edad Clientes_New[0,2]=3 #grupo etnico Clientes_New[0,3]=1 #tipo beneficiario  # Realizar la predicción del cluster del nuevo almacen Etiqueta_nueva = modelo_cargado.predict(Clientes_New) print('') print("El nuevo usuario pertenece al cluster:", Etiqueta_nueva)</pre> <p>El nuevo usuario pertenece al cluster: [1]</p>	<pre># Ingresar datos del nuevo almacén (3 variables) Clientes_New=np.zeros((1,4)) Clientes_New[0,0]=2 #genero Clientes_New[0,1]=22 #edad Clientes_New[0,2]=3 #grupo etnico Clientes_New[0,3]=2 #tipo beneficiario  # Realizar la predicción del cluster del nuevo almacen Etiqueta_nueva = modelo_cargado.predict(Clientes_New) print('') print("El nuevo usuario pertenece al cluster:", Etiqueta_nueva)</pre> <p>El nuevo usuario pertenece al cluster: [2]</p>
--	--

Fuente: elaboración del autor

Una vez corroborado su funcionamiento el modelo fue socializado con el fin de poderlo implementar, sin embargo, queda para futuros desarrollos el poder desarrollar un método que facilite su acceso y cargue de datos, por el momento este permitió crear tres campañas de acercamiento para personas mayores de 21 año, mayores de 54 y mayores de 77 entendiendo sus géneros y etnias

## Conclusiones

Las tecnologías de inteligencia artificial son una de las herramientas mas valiosas con las cuales las empresas y organización pueden contar, puesto que facilita la toma de daciones disminuyendo el riesgo. Sumado a esto su uso no está limitado para un cierto grupo de personas, al contrario, es una herramienta aplicable a cualquier modelo de negocio, sin importar su contexto, como se observa en el presente proyecto donde se implementa en un contexto social humanitario.

Aunque el contexto no importa, pues las tecnologías de machine learning se adaptan, es necesario determinar de manera correcta cual modelo de aprendizaje automatizado es el apropiado para el caso particular. Al seleccionar el clúster para la elaboración de grupos, el investigador entendió los datos y las necesidades expuestas por la empresa, consiguiendo los resultados esperados por los usuarios finales.

Venezuela se encuentra en un proceso difícil donde todos pueden aportar un grano de arena para mejorar la situación de quienes deben salir hoy de sus hogares con la esperanza de un mejor destino, por eso desde la parte de TI, el poder brindar herramientas de machine learning para ser mas eficaces en el tratamiento de este contexto es un aporte a la situación.

En definitiva, el machine learning es una herramienta valiosa para cualquier empresa, corporación o pequeño negocio que quiera un soporte en la toma de decisiones, pero es importante contar con el personal cualificado para que los datos expuestos y procesados sean tomados de manera correcta, puesto que un sesgo en la cadena de confianza de los datos puede acarrear un mal entrenamiento del modelo, y en consecuencia resultados desfazados a la realidad.

## Referencias

- Arango Pastrana, C. A., & Osorio Andrade, C. F. (2021). Aislamiento social obligatorio: un análisis de sentimientos mediante machine learning. *Suma de Negocios*, 12(26), 1–13.  
<https://doi.org/10.14349/sumneg/2021.v12.n26.a1>
- Arteada, P. A., Carvajal Villamizar, A. F., & Guerrero Jaimes, R. M. (2017). La migración venezolana y su impacto socio económico en el municipio de Pamplona Norte de Santander. *Revista Científica Profundidad Construyendo Futuro*, 6(6), 13–25.  
<https://doi.org/10.22463/24221783.2292>
- DE CARVALHO AQUINO, E. L. (2020). *O algoritmo computacional como objeto sociotécnico: encontros da complexidade algorítmica*. Universidad de Federal de São.
- Estupiñán Ricardo, J., Leyva Vázquez, M. Y., Peñafiel Palacios, A. J., & El Assafiri Ojeda, Y. (2021). Inteligencia artificial y propiedad intelectual. *Revista Universidad y Sociedad*, 13(3), 362–368.
- Galindo, G., Navarro, J., Reales, J., Castro, J., Romero, D., Rodríguez, S., & Rivera-Royero, D. (2022). Immigrants resettlement in developing countries: A data-driven decision tool applied to the case of Venezuelan immigrants in Colombia. *PLoS ONE*, 17(1 1).  
<https://doi.org/10.1371/journal.pone.0262781>
- González Rostani, V., Incio, J., & Lezama, G. (2023). *Political Agenda Shaping and Differentiation in Response to Unprecedented Shocks: The Venezuelan Migration Crisis* \*.
- Mamani Rodríguez, Z. (2023). Proceso de machine learning para determinar la demanda social de puestos de empleo de profesionales de TI. *Industrial Data*, 25(2), 275–300.  
<https://doi.org/10.15381/idata.v25i2.21643>

- Mori Castro, A. K. (2021). Inteligencia Artificial y Sociedad ¿El fenómeno social tecnológico 4?0? *Futuro Hoy*, 2(1). <https://doi.org/10.52749/fh.v2i1.11>
- Parada Parra, N. (2024). *Demand driven reinforcement learning modeling for decision support on colombian agricultural supply chain*. Universidad de los Andes.  
<https://hdl.handle.net/1992/73948>
- Paz Noguera, B., Alpala Ramos, O., & Villota Vivas, E. (2021). Análisis de la migración venezolana en la ciudad de Pasto: características y percepciones de los migrantes. *Tendencias*, 22(1), 71–94. <https://doi.org/10.22267/rtend.202102.155>
- Perez Bedia, C. A., & Rojas Segovia, L. E. (2020). *Diseño de un sistema para predecir la deserción de los alumnos mediante machine learning en la Universidad Tecnológica del Perú*. Universidad Tecnológica del Perú.
- Rodríguez-García, J. D., Moreno León, J., Román González, M., & Robles, G. (2020). LearningML: A tool to foster computational thinking skills through practical artificial intelligence projects. *Revista de Educación a Distancia*, 20(63).  
<https://doi.org/10.6018/RED.410121>
- Tribín Uribe, A. M., Adhvaryu, A., Anzola Bravo, C., Ávila Montealegre, O., Bonilla Mejía, L., Castro Fernández, J. C., Flórez, L. A., Grajales Olarte, Á., Guarín López, A., Hamann Salcedo, F., Hermida-Giraldo, D., Khanna, G., Lasso Valderrama, F. J., Medina Durango, C., Melo Becerra, L., Méndez Vizcaino, J. C., Morales, L., Nyshadam, A., Ospina Tejeiro, J. J., ... Velásquez, S. (2020). Migración desde Venezuela en Colombia: Caracterización del fenómeno y análisis de los efectos macroeconómicos. *Ensayos Sobre política Económica*, 2020(97), 1–76. <https://doi.org/10.32468/espe.97>

Velandia Andrade, A. C. (2022). *Análisis de los Efectos de la Política Integral Migratoria en los*

*Ciudadanos que Utilizan a Pamplona, Norte De Santander, Como Corredor Migratorio*

*Desde el Año 2017- 2022.* Universidad Nacional Abierta y a Distancia UNAD.

Zambrano Miranda, M. de J., Galvis Rodríguez, D. C., Sierra De Rodríguez, O. M., & Toloza

Martínez, C. E. (2020). Movilidad humana: venezolanos en Cúcuta y Villa del Rosario, su

satisfacción en el territorio. *Semester Económico*, 23(55), 259–284.

<https://doi.org/10.22395/seec.v23n55a12>