

DOCUMENTACIÓN DEL PROYECTO DE MIGRACIONES

Flujos Migratorios en el Ámbito laboral

Agudelo, Julián - Machine Learning Engineer

Francisco, Sheila - Data Engineer

Muchino, Valentino - Data Architect

Trejo, Carlos - Data Analyst

Zerpa, Omarly - Data Analyst

Índice

[Planteamiento del problema](#)

[Objetivo general](#)

[Etapas Iniciales del Proyecto](#)

[Data Engineering: Pipeline de Datos en AWS](#)

[Data Analytics: PowerBI](#)

[Machine Learning](#)

[Conclusión](#)

[Referencias](#)

Planteamiento del problema

“Make America Great Again” fue el slogan utilizado en la política estadounidense popularizado por Donald Trump en su exitosa campaña presidencial de 2016. Varios periodistas analíticos, académicos y comentaristas la relacionan con el racismo en Estados Unidos, considerándola una política de silbato de perro y un uso del lenguaje codificado.

Estados Unidos es la potencia económica mundial, cuya estatua de la libertad declarada monumento nacional de los EEUU es un obsequio de Francia a EEUU en 1886...

EEUU es también territorio de emblemáticos puentes pero también de largos muros...

Sin embargo el racismo y la xenofobia no es un problema aislado de EEUU, es un problema altamente extendido a lo largo del mundo; movimientos en contra de migrantes y ataques a otras culturas suceden a diario, argumentando una invasión y pidiendo un alto a la migración.

¿Cómo puede existir fobia, rechazo e incluso atentados entre animales de una misma especie?

Los humanos somos por naturaleza seres curiosos y exploradores, desde nuestros orígenes no conocemos la palabra frontera y aunque antes nos detenían los océanos conquistamos un planeta entero y en este momento nuestra única frontera debería ser el universo o mejor dicho el descubrimiento mismo.

Después de todo cuanto todos comenzamos la vida y no tenemos ningún storytelling en la cabeza... somos tan iguales. Tenemos los mismos sueños, las mismas fantasías, compartimos juegos y formas de divertirnos y sobre todo tenemos el mismo lenguaje... sonreímos igual.

Somos un equipo de jóvenes entusiastas de los datos y amamos hacerlos estrujarlos hasta que cuenten su verdad.

Que en nuestro equipo seamos de diversas nacionalidades no es más que otro storytelling, después de todo todos habitamos el mismo planeta cuya naturaleza no sabe de nacionalidades ni de países, y creemos que si un storytelling es tan poderoso para mover sociedades enteras solo existe un recurso tan poderoso para contraatacar y es el poder de otro storytelling.

Para ello buscamos estudios sobre los efectos de migración en un país y aunque encontramos un tema bastante complejo y de muchas aristas, encontramos correlaciones positivas entre el desarrollo de los países y su nivel de migración, que aunque ligera y difícil de estudiar, existe.

Dentro de los estudios destacan :

- Vador Peter, 2021, Research: Why Immigrants are more likely to become entrepreneurs, Harvard Business Review.
- Anderson Stuart, 2018, Immigrants and Billion-Dollar Companies, National Foundation For American Policy.
- Jean-Christophe Dumont, et al., Is migration good for the economy?, OECD.

En dichos estudios destacan que la migración tiene una correlación positiva con el desarrollo económico de un país, sin embargo es ligera y multivariable.

Analizamos la correlación entre la Migración Neta y el crecimiento de PIB por año de EEUU y encontramos una correlación positiva, dando validez a la teoría de los estudios anteriores.

Por lo que dados los efectos positivos de la migración en este proyecto decidimos enfocarnos en analizar la correlación de distintas variables de las cuáles tenemos la hipótesis que influyen en la atracción de migración. Y como producto de este análisis construir dos modelos de Machine Learning, el primero capaz de clasificar países acorde a variables tanto socio-políticas y socio-económicas, mientras que el segundo predice el flujo migratorio futuro que tendrá dicho país en caso de continuar el escenario actual , dicha herramienta es útil para que un país pueda implementar medidas y estrategias a fin de lograr mejorar la captación de migrantes calificados y con ello impulsar su desarrollo.

Para que el país o interesados puedan tener métricas de referencia que sirvan como meta, implementamos KPI's basados en las variables correlacionadas tanto por análisis como por modelo de Machine Learning.

Objetivo general

- Obtener insights y KPI's que permitan a los países e interesados realizar análisis, construir políticas e incentivar sectores económicos específicos para captar migración con el objetivo de impulsar su desarrollo.
- Producir una Inteligencia Artificial que prediga el flujo migratorio futuro de un país bajo el contexto socio-político y socio-económico actual.

Etapas Iniciales del Proyecto

Se enfocó en la necesidad de brindarles a los países medidas que permitan construir políticas y motivar sectores económicos específicos para la captación de

inmigrantes para impulsar la economía del país destino/receptor e incrementar su fuerza laboral y así competir a nivel mundial.

Logramos impulsar la economía de un país a través de la captación de inmigrantes para incrementar su fuerza laboral y así competir a nivel mundial, a través de la fomentación a la integración socioeconómica y sociopolítica de los inmigrantes calificados en el país. Con la información brindada de las siguientes organizaciones: ILOSTAT (International Labour Organization), LinkedIn, The World Bank, Our World in Data y OCDE(Organización de cooperación y desarrollo económico), se obtuvieron los datasets para el estudio, los cuales son: Porcentaje de empleo, desempleo y participación de la población nativa e inmigrantes, porcentaje de participación de la población o índice de desempleo de inmigrantes, salarios mínimos por país, nivel de cumplimiento nacional de los derechos laborales, coeficiente gini por país, índice per cápita del PIB por país, migración neta y LinkedIn Talent Migration.

Data Engineering: Pipeline de Datos en AWS

En primer lugar, optamos por la arquitectura de Amazon Web Services debido a su muy **completa documentación** (a la que recurrimos muchas veces) y porque es la plataforma con la que estábamos más **familiarizados**.

Otro punto muy importante es la capa gratuita que ofrece AWS, lo que nos permitió diseñar una **arquitectura de servicios** sin incurrir en **ningún tipo de gasto**.

Descripción de servicios:

- **AWS S3:** Es un servicio de almacenamiento de objetos que ofrece escalabilidad, disponibilidad de datos y seguridad.

Casos de uso tales como lagos de datos, sitios web, aplicaciones móviles, copia de seguridad y restauración, archivado, aplicaciones empresariales, dispositivos IoT y análisis de big data. Amazon S3 proporciona funciones de gestión para que pueda optimizar, organizar y configurar el acceso a sus datos para satisfacer sus requisitos empresariales, organizativos y de conformidad específicos.

- **AWS Lambda:** Es un servicio informático que permite ejecutar código sin aprovisionar ni administrar servidores.

Lambda ejecuta el código en una infraestructura de computación de alta disponibilidad y realiza todas las tareas de administración de los recursos de computación.

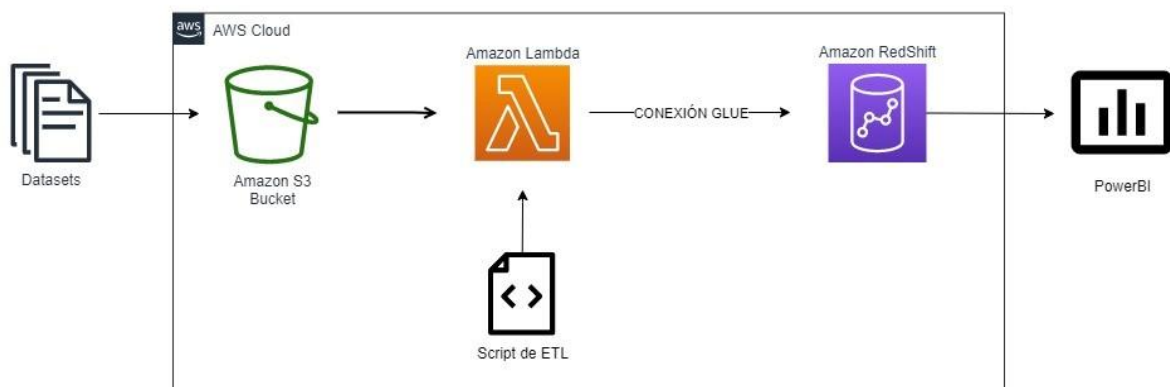
Con Lambda, puede ejecutar código para prácticamente cualquier tipo de aplicación o servicio de backend. Lo único que tiene que hacer es suministrar el código en uno de los lenguajes que admite Lambda (Python, JavaScript).

- **AWS Redshift:** Es un servicio de almacenamiento de datos completamente administrado que opera a escala de petabytes en la Nube de AWS.

Un almacenamiento de datos de Amazon Redshift es una colección de recursos informáticos denominados *nodos* que se organizan en un grupo llamado *clúster*. Cada clúster ejecuta un motor de Amazon Redshift y contiene una o más bases de datos.

Amazon Redshift es un sistema de administración y consulta de bases de datos relacionales de clase empresarial. Amazon Redshift admite las conexiones de clientes con muchos tipos de aplicaciones, incluidas las herramientas de análisis, datos, generación de informes e inteligencia empresarial (BI).

Pipeline de datos:



La primera parte de nuestro pipeline de datos en AWS consiste de una ingesta de datos, a un “bucket” de S3. Este bucket es un repositorio en el que se ubican todos los datasets que obtuvimos de diversas fuentes.

Una vez finalizada la ingesta, continuamos con la limpieza de los datos en Lambda.

Lambda nos permite automatizar todo el proceso de ETL, ya que extrae los datos directamente de S3, los limpia y los envía a nuestra base de datos en Redshift. Lo único que tenemos que hacer es escribir el código de ETL en Python y pasárselo a Lambda.

Otro punto importante es que AWS posee una librería llamada "Data Wrangler". Esta librería es compatible con Pandas, lo que disminuye drásticamente la complejidad de la limpieza y nos facilita el código.

Una vez finalizada la limpieza, Lambda carga los datos en nuestro DataWarehouse de Redshift.

En Redshift podemos realizar consultas SQL e incluso conectarnos a una herramienta de BI como PowerBI o Tableau para crear reportes y visualizaciones de los datos.

Validación de datos:

La validación de datos hace referencia a que tipo de datos presenta cada tabla de la base de datos. También define Primary Keys y Foreign Keys.

Así es la estructura de la base de datos:

Tabla: fact_empleo_condicion			
Objetivo: Almacena información sobre las condiciones de empleo dependiendo el lugar de nacimiento.			
Key	Columna	Tipo de dato	Descripción
FK	ID pais	Int	Hace referencia al ID en la tabla dim_pais
FK	ID anio	Int	Hace referencia al ID en la tabla dim_anio
FK	ID lugar_de_nacimiento	Int	Hace referencia al ID en la tabla dim_condicion
	empleados	Decimal	Porcentaje de población con empleo según el lugar de nacimiento.
	desempleados	Decimal	Porcentaje de población en desempleo según el lugar de nacimiento.
	participación	Decimal	Porcentaje de participación según el lugar de nacimiento.
	tasa_participacion	Decimal	Tasa de participación en la fuerza de trabajo según lugar de nacimiento.

Tabla: fact_ind_sociopoliticos			
Objetivo: Almacena información sobre indicadores sociopolíticos			
Key	Columna	Tipo de dato	Descripción
FK	ID país	Int	Hace referencia al ID en la tabla dim_pais
FK	ID año	Int	Hace referencia al ID en la tabla dim_año
	derechos_laborales	Decimal	Nivel de cumplimiento nacional de los derechos laborales (libertad sindical y negociación colectiva)
	derechos_humanos	Decimal	Indicador de derechos humanos
	coeficiente_gini	Decimal	Indicador de desigualdad salarial
	brecha_salarial_empleo	Decimal	Brecha salarial entre hombres y mujeres que se encuentran empleados por un tercero
	brecha_salarial_autoempleo	Decimal	Brecha salarial entre hombres y mujeres que están en condición de autoempleo

Tabla: fact_ind_economicos			
Objetivo: Almacena información sobre indicadores económicos			
Key	Columna	Tipo de dato	Descripción
FK	ID pais	Int	Hace referencia al ID en la tabla dim_pais
FK	ID anio	Int	Hace referencia al ID en la tabla dim_anio
	pib	Decimal	Crecimiento del PIB (Producto Interno Bruto) per cápita en porcentaje anual
	salario_minimo	Decimal	Salario mínimo nominal mensual bruto en USD
	salario_promedio	Decimal	Salario promedio en USD
	ingresos_altos	Decimal	Proporción de trabajadores que perciben más de 1.5 veces el ingreso promedio que hay en el país indicado
	ingresos_bajos	Decimal	Proporción de trabajadores que perciben menos de 2/3 del ingreso promedio que hay en el país indicado
	%empleo_agricultura	Decimal	Porcentaje del total de empleos, que se ubica en el sector agrícola
	%empleo_industria	Decimal	Porcentaje del total de empleos, que se ubica en el sector industrial
	%empleo_servicio	Decimal	Porcentaje del total de empleos, que se ubica en el sector de servicios

Carga incremental de datos:

La carga incremental hace referencia a cómo reacciona nuestro pipeline frente a nueva información.

Para demostrarlo, se realizó el siguiente ejemplo:

Primero creo un dataframe con nueva información. En este caso es un dataframe con valores de migración neta en distintos países para el año 2030.



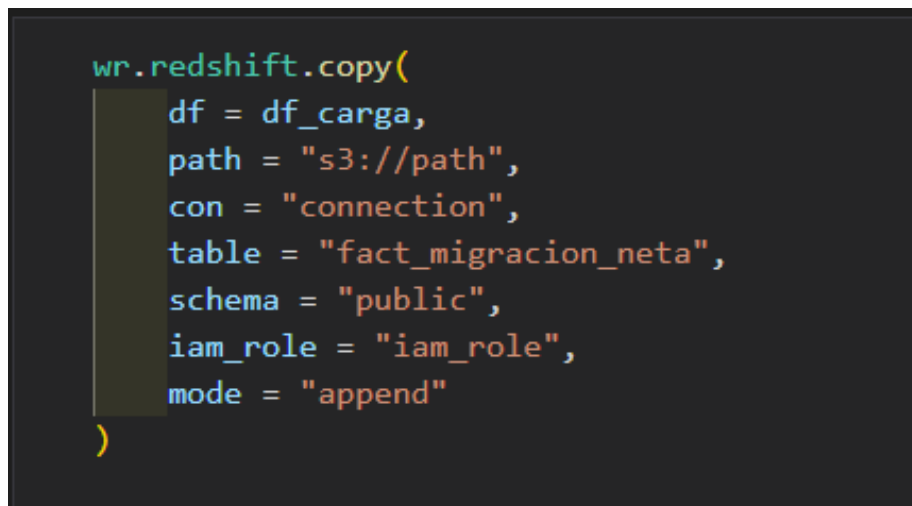
df_carga

✓ 0.0s

	codigo_pais	anio	migracion_neta
0	ARG	2030	10000
1	MEX	2030	48000
2	COL	2030	6000

Dataframe de Pandas

Luego, gracias a la librería de AWS Data Wrangler, con unas pocas líneas de código puedo subir el dataframe a Redshift.



```
wr.redshift.copy(  
    df = df_carga,  
    path = "s3://path",  
    con = "connection",  
    table = "fact_migracion_neta",  
    schema = "public",  
    iam_role = "iam_role",  
    mode = "append"  
)
```

Código AWS Wrangler

El argumento “append” al parámetro “mode” indica que los datos tienen que añadirse a la tabla ya existente. Por defecto este parámetro tiene el valor de “overwrite”, lo que sobrescribe los datos de la tabla.

Una vez ejecutado el código, los datos ya están disponibles en nuestra base de datos, como lo muestra la siguiente consulta SQL:

```

3 SELECT * FROM fact_migracion_neta
4 WHERE anio = 2030;

```

Result 1 (3)

<input type="checkbox"/> codigo_pais	anio	migracion_neta
<input type="checkbox"/> ARG	2030	10000
<input type="checkbox"/> COL	2030	6000
<input type="checkbox"/> MEX	2030	48000

Consulta SQL en Redshift

De esta manera finaliza el ejemplo de una carga incremental, el cual demuestra que el pipeline de datos que diseñamos es apto y funcional.

Data Analytics: PowerBI

El análisis se dividió en dos factores:

- Socio-políticos
- Socioeconómicos

Dentro de los análisis sociopolíticos encontramos:

- Derechos humanos
- Derechos Laborales

En los análisis socioeconómicos:

- Sectores económicos
- Tasa de empleo y desempleo
- Desigualdad (índice de Gini)

Y como índice compartido:

- Salario mínimo mensual bruto

Factores sociopolíticos

Derechos humanos

Los datos se extrajeron de “Our World in Data” cuya publicación está a cargo de la Universidad de Oxford (Inglaterra).

Dichos datos están basados en evaluaciones de expertos y el índice V-Dem (Varieties of Democracy).

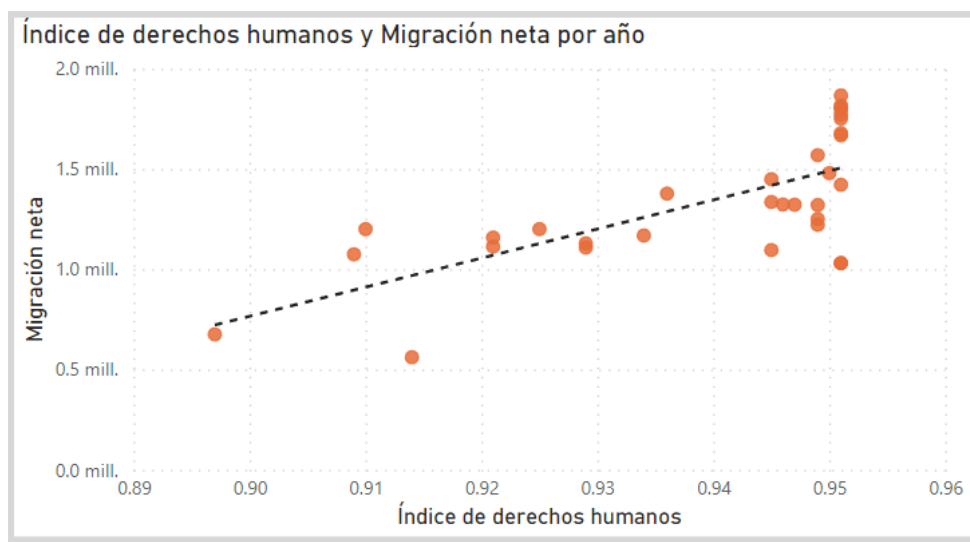
En estos datos se evalúan las siguientes libertades humanas:

- Tortura gubernamental
- Asesinatos políticos

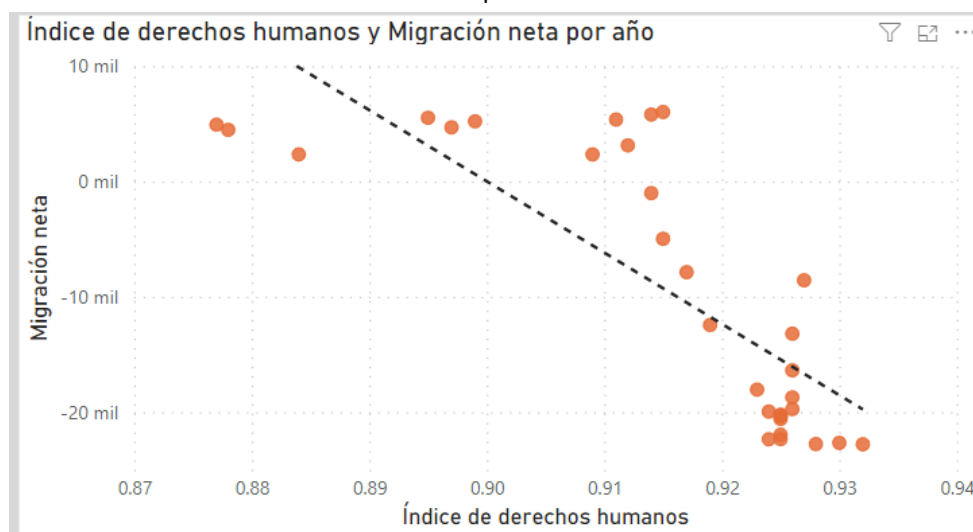
- Trabajos forzados
- Derechos de propiedad
- Libertad de circulación
- Libertad de asociación
- Libertad de expresión
- Libertad de religión

Al realizar la correlación se encontró una tendencia positiva en países desarrollados como EEUU, sin embargo se encontró correlación negativa en países pobres o en vías de desarrollo como Argentina, probablemente debido a que a pesar de que los derechos humanos sean aceptables, las condiciones económicas o políticas no son favorables y aún así tienen un flujo migratorio negativo.

Para la correlación se tomó como variable independiente los índices de derechos humanos (eje X) y como variable dependiente la migración neta (eje Y).



Correlación positiva: EEUU



Correlación negativa : Argentina

De este análisis tomamos un **KPI Derechos Humanos** el cual es la Media del índice de derechos humanos del país con respecto al valor máximo histórico alcanzado a nivel global, en este caso Suecia con un índice de 0.98.

Por lo que su cálculo se realiza con la siguiente fórmula:

Fórmula:

$$\frac{\text{Valor actual} + \text{Valor max global}}{2}$$

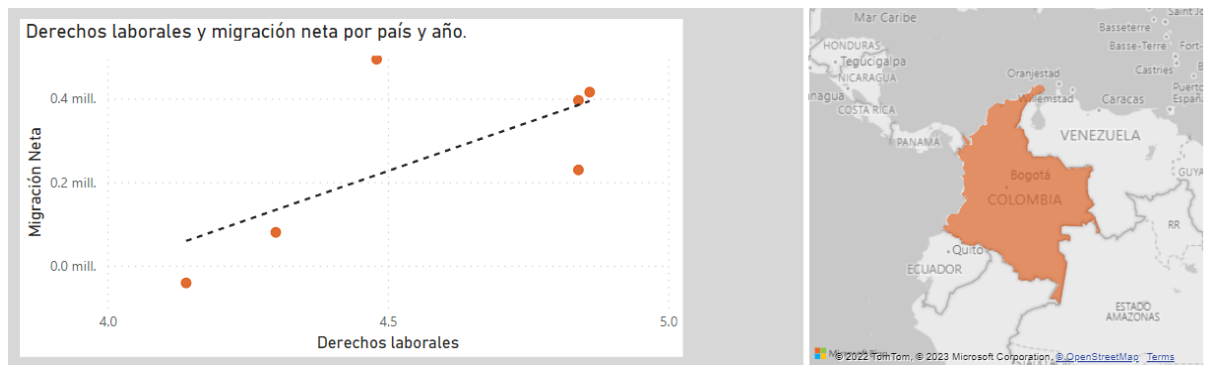
Derechos Laborales

Para este análisis se extrajeron datos de la OIT (Organización Internacional del Trabajo) disponibles en ILOSTAT.

El valor de estos datos está basado en el nivel de cumplimiento nacional de los derechos laborales (libertad sindical y negociación basado en fuentes textuales de la OIT (ILO) y la legislación nacional.

Con estos datos se realizaron correlaciones entre los derechos laborales y la migración neta.

En el análisis se detectó que no son variables con gran fuerza en la migración , únicamente resultaron en variables importantes cuando son países que por factores geopolíticos como el área Schengen tienen una correlación positiva, tal es el caso de Suiza , países bajos, Colombia y Canadá.



Salario mínimo

Para este análisis se extrajeron datos de ILOSTAT, cuyos valores corresponden al salario mínimo mensual bruto (Anual) y los valores se encuentran estandarizados a dólares estadounidenses (USD).

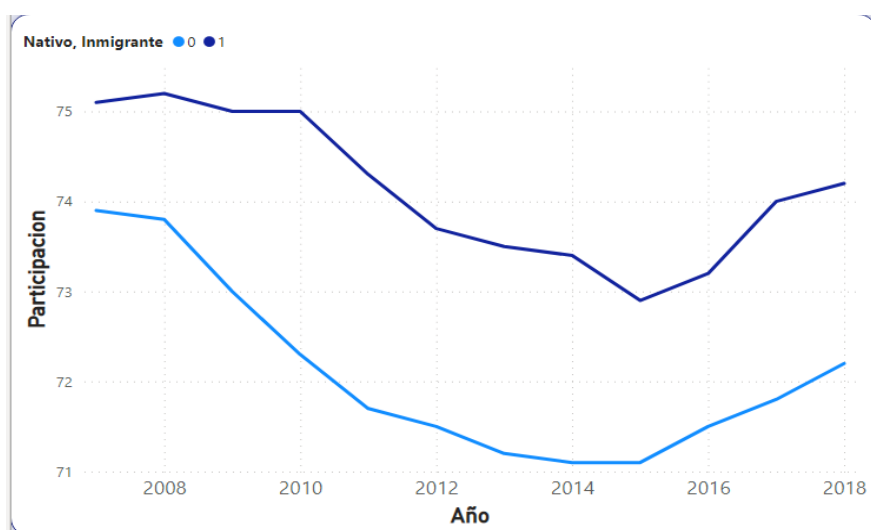
Factores Socioeconómicos

Mercado laboral

Los países modelos garantizan una participación laboral igual para la población nativa e inmigrantes y de esta manera elevan o mantienen su economía . Con los datos de empleo, desempleo y participación laboral de la población nativa e inmigrantes y participación laboral en general de la población del país, se establece

un objetivo de reflejo de participación ideal representativa u objetivo general para nativos e inmigrantes, de esta manera evaluar la participación de los inmigrantes de cada país o de países posibles clientes.

La ecuación utilizada para medir el **KPI de Índice de inmigrantes en el mercado laboral** se trabaja con el promedio de participación general más brecha promedio diferencial mínima entre inmigrantes y nativos.



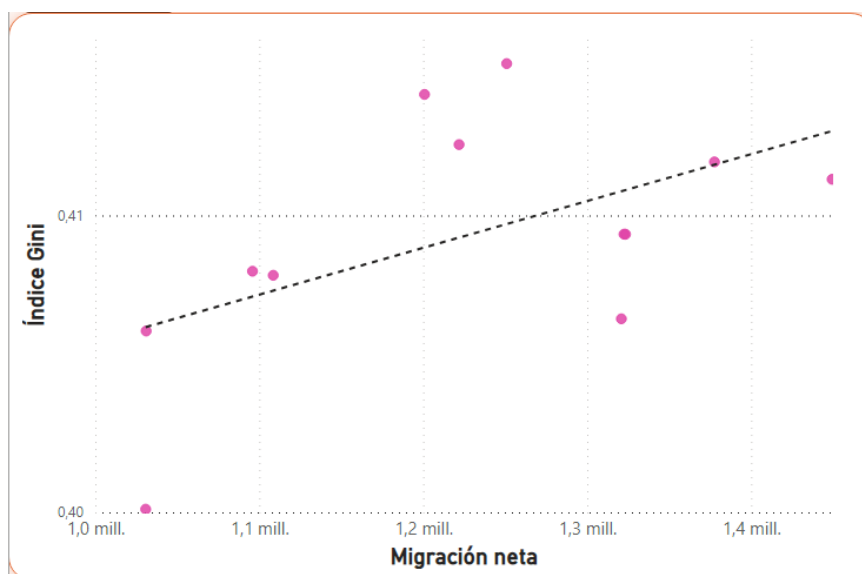
Dashboard. Porcentaje de participación de la población nativa vs inmigrantes por año, de los E.E.U.U para el periodo 2007-2018.

Desigualdad

La mayoría de las naciones junto con la organización **OCDE** (Organización de cooperación y desarrollo económico) empujan a la igualdad económica y bajar la desigualdad a un **30%** para que la población pobre crezca a una tasa superior que la media nacional, de esta manera potenciar la inclusión social y económica de todas las personas independientemente de su raza, etnia o situación económica, garantizando la igualdad de oportunidades. Analizando el gini y la migración neta se evidencia que ocurre tanto correlación negativa, en la que los valores de la variable migración neta tienden a incrementarse mientras que los valores de la variable de coeficiente de gini descenden, como correlación positiva, donde los valores de la variable migración neta tienden a incrementarse al igual que los valores de la variable de coeficiente de gini, esto dependiendo del año o periodo establecido; Además al analizar el coeficiente gini y el PIB per cápita de cada país se constata que tienen crecimientos y decrecimientos similares en los mismos años, es decir, que se relacionan mutuamente.

Utilizando el 30% de desigualdad objetivo propuesto por la **OCDE**, y usando la ecuación promedio de coeficiente gini multiplicado por 100%, se obtuvo el **KPI de Índice de desigualdad económica** o porcentaje de desigualdad de los países. En conclusión se

observó que los inmigrantes tienen que migrar donde la desigualdad esté por encima pero cercana a ese 30%.



Dashboard. Correlación de índice gini y migración neta para E.E.U.U para el periodo 2007-2018.

Profesiones y Sector Industrial

La colaboración de LinkedIn y del Banco Mundial brinda excelente información que sirve de ejemplo de cómo las empresas de tecnología pueden trabajar con instituciones de desarrollo para llevar nuevos datos e ideas a los países en desarrollo para abordar los desafíos de desarrollo apremiantes. Las oportunidades y los desafíos que presenta la economía global requieren que los sectores público y privado unan fuerzas, compartan información, compartan recursos y trabajen hacia una visión común para lograr un impacto significativo, positivo y escalable.

Las métricas y los datos son un producto de la asociación y cubren las métricas de empleo, habilidades y migración de la industria para más de 100 países. Específicamente, los datos cubren 4 métricas: 1) Cambios en el empleo de la industria.

2) Migración de talentos.

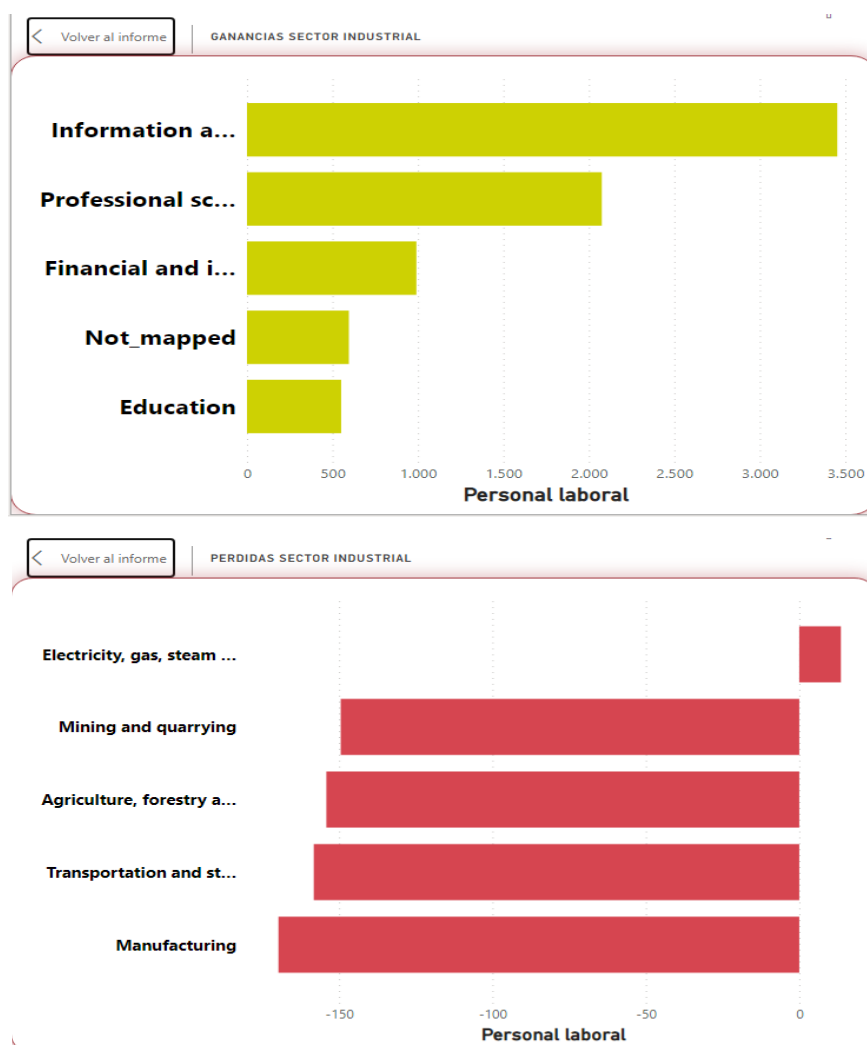
3) Necesidades de habilidades de la industria.

4) Penetración de habilidades.

LinkedIn y el Grupo del Banco Mundial actualizan estos datos anualmente.

Todas las métricas de Migración de Laboral se basan en la migración neta (entradas menos salidas). Estas cifras de migración neta están normalizadas por el número de miembros de LinkedIn en un país determinado para permitir comparaciones más justas entre muestras. Por ejemplo, considerando a Canadá como el país de interés,

los flujos netos de migrantes hacia/desde los EE. UU. en términos absolutos se normalizan por el recuento total de miembros de LinkedIn en Canadá. Se utiliza un método similar para calcular las ganancias/pérdidas de capacitación e industria asociadas con los flujos migratorios bilaterales. Este calcula todo sobre una base anual e informa un promedio de los últimos tres años. Para proteger la privacidad de los miembros, LinkedIn solo muestra datos agregados con al menos 50 observaciones. Los espacios en blanco anteriores generalmente se deben a que los flujos bilaterales no cruzan este umbral, o a que los **países solo tienen ganancias o pérdidas netas en LinkedIn**. Dado que los datos de LinkedIn tienen una mejor cobertura de los trabajadores administrativos en los sectores de servicios intensivos en conocimiento, es poco probable que estos flujos migratorios representen todo el panorama migratorio de un país, pero pueden **arrojar luz sobre algunos de los sectores más dinámicos e innovadores**.

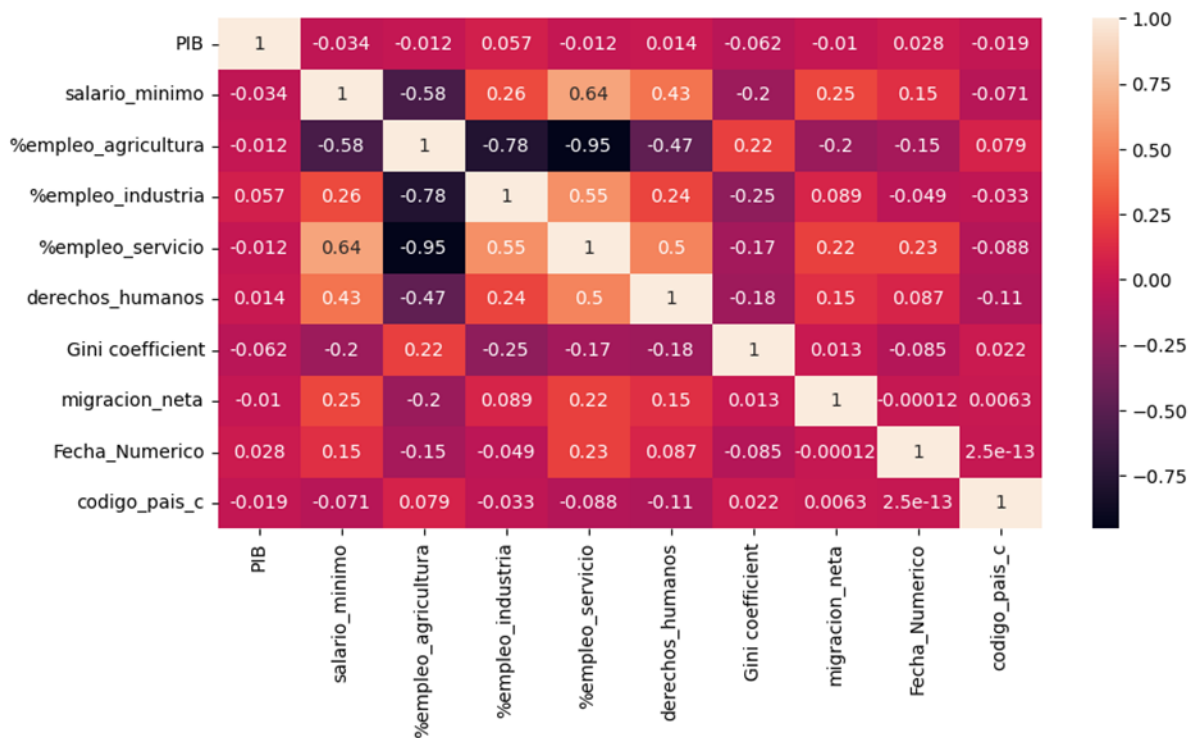


Dashboard. Flujos migratorios laborales en sector industrial ganancias y pérdidas de personal de E.E.U.U para el periodo 2015-2019.

Machine Learning

El objetivo del modelo de machine learning es predecir la migración neta a partir de ciertos indicadores o parámetros socioeconómicos.

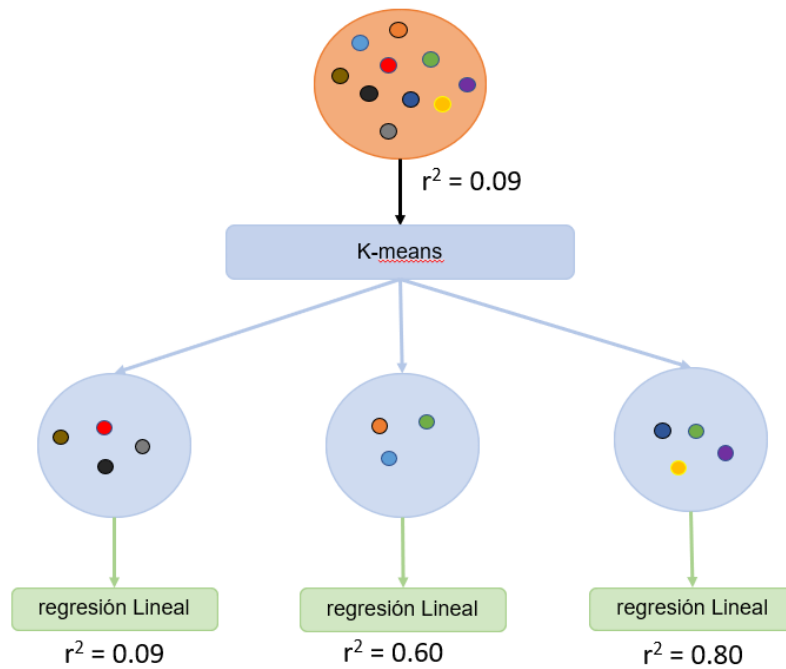
En el siguiente mapa de correlación se aprecian los diferentes indicadores y su grado de correlación con la migración neta.



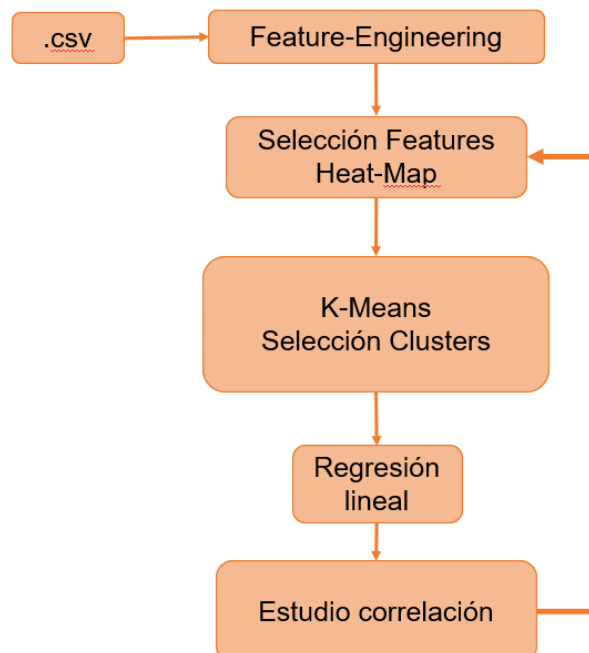
Al ejecutar una regresión lineal (forecast) sobre todo el conjunto de datos la correlación lineal arroja un valor muy bajo ($r^2 = 0,09$). Esto se debe a que las diferentes entidades (países) guardan entre sí diferencias en cuanto a tendencias y valores en sus indicadores socioeconómicos.

Para llegar a un modelo de predicción con mejores valores de correlación se propuso:

1. Ejecución previa de un modelo no supervisado sobre todo el conjunto de datos que haga un agrupamiento (clusters) de los diferentes países.
2. Ejecución de un modelo predictivo (forecast-regresión lineal) sobre cada clúster.



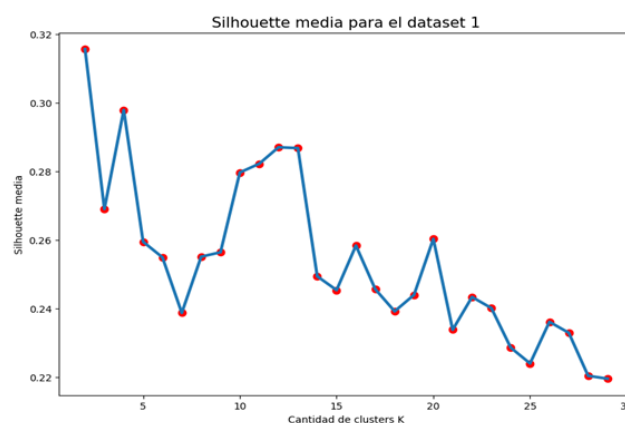
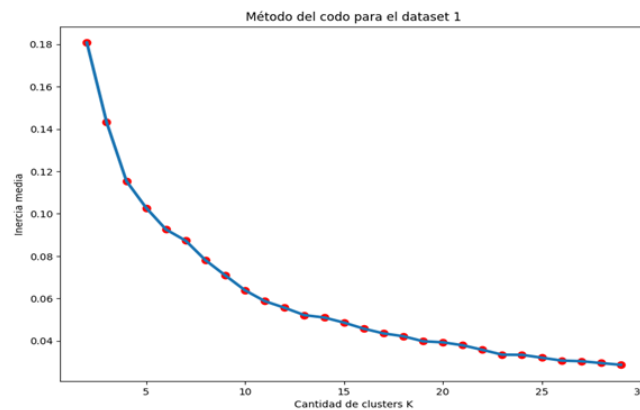
Para la ejecución del modelo de machine learning se siguió el siguiente proceso iterativo.



El objetivo final del procedimiento planteado es evaluar el modelo de machine learning en su conjunto, a partir de los coeficientes de correlación y el error en la

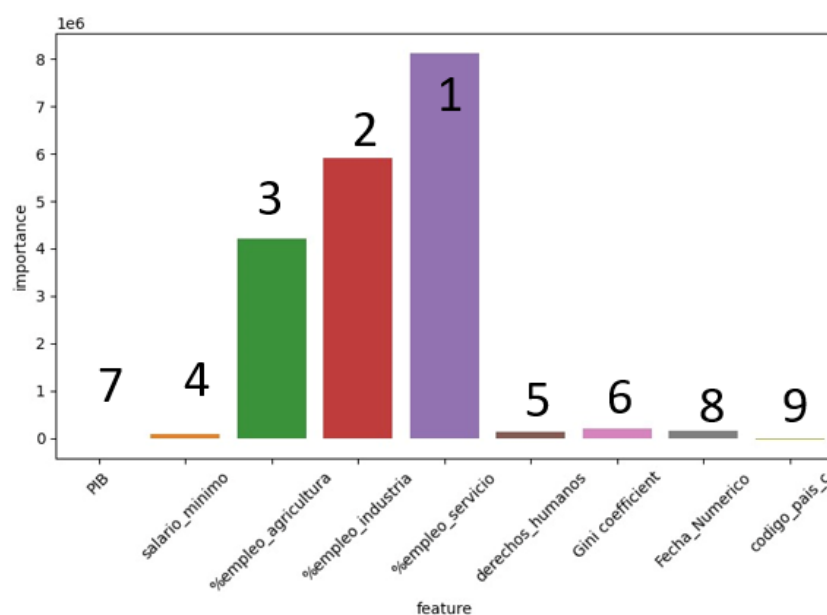
regresión lineal. Para ello también se estudió la influencia de los siguientes parámetros en el modelo:

- Selección de número de clusters: para la selección del número de clusters se hizo uso del método del codo y el valor de silhouette media. De acuerdo con los gráficos que se muestran a continuación se decidió elegir 13 como número óptimo de clusters.



- Selección de los features – reducción de parámetros: para la reducción de parámetros se decidió evaluar el error para la predicción (USA- año 2015) y la correlación mínima y máxima para todos los clusters. De acuerdo a los resultados obtenidos el número de features que garantizan una buena correlación lineal y bajos porcentajes de error, es un modelo de machine learning donde se utilicen 8 variables.

Features	4	5	6	7	8	9
USA (% error)	83,4	64,55	38,52	62,05	3,07	35,35
r^2_{min}	0,018	0,07	0,06	0,10	0,09	0,09
r^2_{max}	0,52	0,89	0,87	0,54	0,68	0,76



A continuación, se listan los valores elegidos para implementar el modelo de machine learning.

<u>Clusters</u>	13
Variables	8
<u>Años (Clustering)</u>	2015 - 2019
Años (Training)	2000 - 2014
Años (Test)	2015 - 2019

Con los valores elegidos se obtienen los siguientes resultados que representan el grado de correlación para cada cluster (13 clusters en total) y el error en la predicción para USA y Angola año 2015.

```
El coeficiente de determinación es: 0.29511001823449756
El coeficiente de determinación es: 0.385336659168369
El coeficiente de determinación es: 0.2307282809644241
El coeficiente de determinación es: 0.27144128955447644
El coeficiente de determinación es: 0.6441352143323044
El coeficiente de determinación es: 0.29887834131897884
El coeficiente de determinación es: 0.6795250822339014
El coeficiente de determinación es: 0.13273209832132826
El coeficiente de determinación es: 0.09067615802767626
El coeficiente de determinación es: 0.1997382255930722
El coeficiente de determinación es: 0.32025291471211503
El coeficiente de determinación es: 0.4900884959693138
El coeficiente de determinación es: 0.7608256996596261
La cantidad que se predice para migración es :1062324.93
La cantidad que se predice para migración es :28373.23
El error de la predicción para USA es:3.07
El error de la predicción para AGO es:68.48
```

Finalmente, para la puesta en marcha del modelo de machine learning se decidió crear una API utilizando el framework FastAPI y para el deployment se utilizó Deta.

Se seleccionó Deta ya que trabaja sin servidores, no necesita dockerización y requiere mínimos conocimientos de programación para tener un producto accesible a cualquier persona con acceso a internet.

Dentro de la API se definió una función con la cual obtener la predicción a futuro de la migración neta para los países evaluados, contando con predicciones hasta 2025.

La forma de acceder a dicha consulta en el navegador es la siguiente:

https://teamproject-1-p7026804.deta.app/prediccion/{codigo_pais}/{anio}

En donde {codigo_pais} corresponde al código del país del que se desee obtener la predicción y {anio} el año para el cual calcularla. Dichos parámetros deben ser sustituidos para poder visualizar la predicción.

En caso de no conocer el código del país de interés se cuenta con un diccionario de códigos, al cual se puede acceder desde la siguiente liga:

<https://teamproject-1-p7026804.deta.app/codigos>

Conclusión

El 70% de los aproximadamente 260 millones de migrantes que hay en el mundo son trabajadores migrantes, hombres y mujeres.

A partir de este dato, una de las conclusiones es que se deben construir políticas de migración laboral, esto es, dicho de otra forma, Mejorar la atractividad de un país a través de la libertad de facultad de derecho y el valor intrínseco de individuo, Crear hojas de ruta de empleo específicos para inmigrantes calificados, Gestionar la participación o servicios de oportunidad económica para inmigrantes, Ajustar y fijar el nivel de salario mínimo para captación de personal calificado de extranjero y Mejorar el impacto económico, social y político de país.

Ello supone ver al inmigrante y/o retornado también como una persona con potencialidades, experiencia, disponibilidad y/o competencias profesionales que pueden ser bien aprovechadas si las políticas públicas definen instrumentos y mecanismos adecuados. En este horizonte, los Ministerios de trabajo tienen que ser fortalecidos para poder asumir la cogestión de los flujos migratorios y aplicar una lógica de vinculación con el mercado del trabajo.

Entonces, en términos reales la migración laboral genera un efecto positivo neto al país de destino, y un efecto negativo neto al país de origen cuando no existen transferencias de los emigrantes a aquellos dejados atrás en su país de origen.

Siguientes pasos

Segmentar la migración por tipo (refugiados, personal calificado, reunificación familiar, etc.). Acorde a esta segmentación inferir nuevas medidas de integración a fin de obtener mayores beneficios de desarrollo económico y cultural del país.

Obstáculos

- Falta de democratización de datos de algunos países.
- Falta de datos sobre instrumentos de financiamiento a inmigrantes, así como de esfuerzos de inclusión social de los inmigrantes mediante cursos (idioma, legislación, historia, etc).
- Aunque se disponibilizó una API para consultar las predicciones sin necesidad de ejecutar código, y así, optimizar recursos informáticos, los modelos de ML hechos no tienen deployment directo debido al peso permitido.

Referencias

https://www.ilo.org/shinyapps/bulkexplorer21/?lang=es&id=MFL_FEMP_SEX_ECO_NB_A

<https://data.oecd.org/inequality/income-inequality.htm>

<https://datos.bancomundial.org/indicador/NY.GDP.PCAP.KD.ZG>

<https://datacatalog.worldbank.org/search/dataset/0038044/Talent-Migration--Linkerdln-Data->

<https://datos.bancomundial.org/indicador/SM.POP.NETM?end=2021&start=1962>

https://stats.oecd.org/Index.aspx?DataSetCode=DIOC_OCCUPATION