PROYECTO

ESTUDIANTES:

CAMILO URREGO BELTRAN

LEON FAEL ROCHA PEREZ

DOCENTE

SEBASTIAN PERDOMO LEIVA

ESCUELA TECNOLÓGICA INSTITUTO TÉCNICO CENTRAL

TECNOLOGIA EN DESARROLLO DE SOFTWARE

CIENCIA DE DATOS Y BIG DATA

BOGOTA D.C

2024

**Problema de Optimización Elegido**

El problema de optimización que hemos decidido abordar es la optimización de la tasa de reciclaje de residuos plásticos a nivel global. La cantidad de residuos plásticos generados anualmente es un desafío creciente para el medio ambiente y la salud pública. Los residuos mal gestionados pueden acabar en los océanos o contaminar el suelo y el aire, y el reciclaje ineficaz exacerba este problema.

El objetivo de este proyecto es mejorar la tasa de reciclaje global, utilizando un modelo de optimización basado en factores socioeconómicos y de políticas que puedan influir en la efectividad del reciclaje. Nos centraremos en identificar los factores que impactan la tasa de reciclaje para entender cómo modificar políticas o inversiones que mejoren este indicador en diferentes países o regiones.

**Metodología Utilizada**

La metodología hasta el momento se ha dividido en tres etapas principales: exploración de datos, modelado de regresión y optimización con un algoritmo genético.

**Exploración de Datos**

La primera etapa fue una exploración exhaustiva del dataset *Global Plastic Waste 2023*, que contiene información sobre la generación y gestión de residuos plásticos en diferentes países y regiones. Las columnas principales incluyen:

* **Cantidad total de residuos plásticos generados**.
* **Tasa de reciclaje** (porcentaje de los residuos plásticos que son reciclados).
* **PIB per cápita** y **nivel de industrialización**.
* **Población total** y otras variables socioeconómicas.

La exploración inicial se llevó a cabo mediante el análisis de estadísticas descriptivas y visualización de datos. Utilizamos gráficos de dispersión, histogramas, y mapas de calor para entender las correlaciones entre variables. Esto ayudó a identificar que factores como el PIB per cápita y la inversión en infraestructura tienen una correlación positiva con la tasa de reciclaje.

**Modelo de Regresión con Gradiente Descendente**

Para predecir la tasa de reciclaje, implementamos un modelo de **regresión lineal** optimizado mediante el algoritmo de **gradiente descendente**. En este caso:

* **Variable dependiente**: Tasa de reciclaje.
* **Variables independientes**: PIB per cápita, población, nivel de industrialización y políticas ambientales.

Usamos el algoritmo de gradiente descendente para ajustar los coeficientes de nuestro modelo de regresión lineal. Este modelo proporciona una primera aproximación para comprender la relación entre la tasa de reciclaje y las variables independientes, aunque es una solución limitada, ya que la tasa de reciclaje es un fenómeno complejo influido por muchos factores no lineales.

**Optimización con Algoritmo Genético**

Debido a la complejidad de mejorar la tasa de reciclaje, decidimos emplear un **algoritmo genético**. Los algoritmos genéticos son ideales para problemas de optimización en los que la función objetivo no es lineal y se pueden presentar restricciones o múltiples variables de influencia.

1. **Inicialización de Población**: Creamos una población inicial de posibles combinaciones de parámetros de políticas, que incluyen niveles de inversión en reciclaje, incentivos económicos, y mejoras en la infraestructura de gestión de residuos.
2. **Función de Fitness**: La función de fitness evalúa cada combinación de políticas, considerando factores como la tasa de reciclaje obtenida, los costos de implementación y el impacto ambiental. Las soluciones que logran una tasa de reciclaje más alta con un costo menor obtienen una mejor evaluación.
3. **Selección, Cruce y Mutación**: A través de varias generaciones, las soluciones mejor calificadas se seleccionan y combinan (cruce) para crear nuevas soluciones. Además, aplicamos mutaciones aleatorias para diversificar la búsqueda y evitar estancamientos en soluciones locales.
4. **Convergencia**: La optimización continúa hasta que la población converge hacia una combinación de parámetros óptima o hasta que se alcanza un número máximo de generaciones.

**Resultados Preliminares**

Hasta el momento, hemos obtenido los siguientes resultados preliminares:

1. Modelo de Regresión: El modelo de regresión lineal ajustado con gradiente descendente ha mostrado una correlación positiva entre el PIB per cápita y la tasa de reciclaje, aunque la relación no es totalmente lineal. Esto confirma que los países con un mayor PIB per cápita tienden a tener tasas de reciclaje más altas debido a mejores infraestructuras y políticas de gestión de residuos.
2. Optimización con Algoritmo Genético:
   * Las primeras iteraciones del algoritmo genético sugieren que combinaciones de alta inversión en infraestructura y políticas de incentivos económicos tienen un impacto positivo en la tasa de reciclaje.
   * La tasa de reciclaje óptima encontrada hasta ahora muestra una mejora del 15% en comparación con el promedio actual del dataset, aunque este resultado aún requiere validación y ajuste.
3. Visualización de Resultados: Los gráficos de correlación y mapas de calor han sido útiles para entender la relación entre variables, y los gráficos de dispersión han visualizado el efecto de los diferentes niveles de PIB en la tasa de reciclaje. Estos resultados sugieren que no solo los factores económicos, sino también los sociales y gubernamentales, son claves para mejorar la gestión de residuos plásticos.

**Conclusiones y Siguientes Pasos**

Conclusiones preliminares: Los resultados hasta ahora indican que existe una fuerte relación entre el nivel de desarrollo económico y la tasa de reciclaje. Sin embargo, el análisis inicial también muestra que la implementación de políticas públicas juega un rol esencial. Las combinaciones de políticas óptimas que estamos obteniendo podrían servir de guía para los gobiernos que buscan mejorar sus tasas de reciclaje.

**Siguientes pasos**:

1. **Ajustar el Algoritmo Genético**: Continuaremos refinando los parámetros de cruce y mutación para mejorar la optimización de la tasa de reciclaje.
2. **Incorporar Nuevas Variables**: Evaluaremos la posibilidad de incluir factores sociales y culturales, como la educación ambiental, para aumentar la precisión del modelo.
3. **Validación del Modelo**: Validaremos los resultados obtenidos en el algoritmo genético con datos de estudios de caso o países con políticas exitosas en reciclaje.