Machine Learning Internship – Population Growth Rate Prediction

Team Name: Rebels

Team Member 1: Aakesh Raj R (<u>aakeshrajr.22@gmail.com</u>)

Team Member 2: Deepak P (deepakprakasam11@gmail.com)

Team Member 3: Tharun S (tharunselladurai123@gmail.com)

Preference: Simple Machine Learning Project

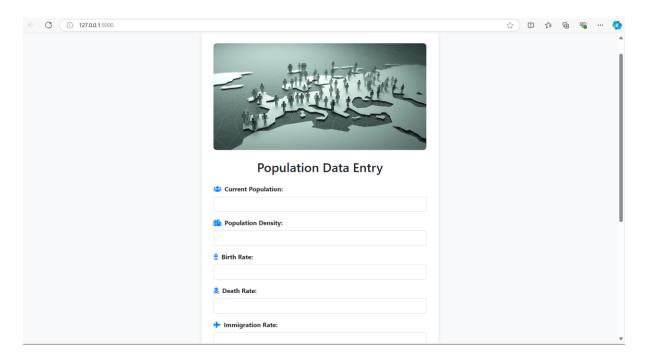
- Training/Test data split 80%/20%
- Machine Learning Algorithm Used:
 - GradientBoostingRegressor
 - Gradient Boosting, including variants like
 GradientBoostingRegressor and generally provides high prediction
 accuracy. It is known for producing models that achieve state-of-the-art performance on many tasks, especially when fine-tuned properly.
 - Gradient Boosting supports both regression and classification tasks, making it versatile for a wide range of use cases.
 - The algorithm allows for tuning various hyperparameters like learning rate, tree depth, and the number of trees, providing the flexibility to optimize performance based on the dataset.
 - Decision trees, the building blocks of gradient boosting models, are generally robust to outliers because they partition the feature space into regions rather than fitting a single line through the data.
 - Unlike deep learning models, which often require large datasets and substantial computational power, gradient boosting can

achieve **high performance** on smaller datasets, making it suitable for scenarios where data may be limited.

- Outlier Removal method used: Interquartile Range (IQR) method
 - It's a straightforward way to identify extreme values that deviate significantly from the majority of the data, making it easy to implement and understand.
 - Unlike the Z-score method, which relies on the mean and standard deviation, the IQR method is less influenced by extreme outliers since it uses medians (quartiles) instead of means.

Output:

Home Page- For Data Entry



o Result Page:

