**Population Predictor**

**Submitted for**

**Statistical Machine Learning CSET211**

Submitted by:

**(E23CSEU1886) Shashank Pratyush**

**(E23CSEU1866) Akash Bisht**

Submitted to

Dr. Shakshi Sharma

**July-Dec 2024**

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**

A close-up of a logo

Description automatically generated

**INDEX**

|  |  |  |
| --- | --- | --- |
| Sr.No | Content | Page No |
| 1 | **Abstract** | 1 |
| 2 | **Introduction** | 2 |
| 3 | **Related Work (If Any)** | 3 |
| 4 | **Methodology** | 4 |
| 5 | **Hardware/Software Required** | 5 |
| 6 | **Experimental Results** | 6 |
| 7 | **Conclusions** | 7 |
| 8 | **Future Scope** | 8 |
| 9 | **GitHub Link of Your Complete Project** | 9 |

**ABSTRACT**

This project aims to forecast the future population and GDP of countries using ARIMA models and provide an interactive 3D visualization of these predictions.

The system includes a Flask-based API that allows users to select a country on a 3D model of Earth, where it then displays the population and GDP predictions for that country.

Using datasets from reliable sources, the population and GDP time series data for different countries are preprocessed and cleaned to handle missing values and outliers. The ARIMA (AutoRegressive Integrated Moving Average) model is applied to forecast the population and GDP for the next few years, providing insights into future economic trends.

The Flask API is designed to provide an easy-to-use interface for querying predictions, making it accessible to users without technical expertise. This system can be useful for policymakers, researchers, and businesses seeking to make data-driven decisions based on predicted demographic and economic changes

**Introduction**

The purpose of this project is to predict the future population and GDP of countries, utilizing ARIMA models for time series forecasting.

These predictions are displayed in an interactive 3D model of Earth, where users can select a country to view the respective forecasts. The significance of this project lies in its ability to make economic forecasting more accessible and engaging by leveraging a 3D visualization platform.

By integrating ARIMA with modern visualization techniques, users can interactively explore predictions for population and GDP trends, making complex data easier to understand.

The system provides both valuable data and an engaging user experience, ensuring that individuals without a technical background can still comprehend global economic changes.

**Related Work**

ARIMA models have been extensively used for time series forecasting of economic indicators such as population and GDP. Several studies have explored their application in different fields, from finance to health.

In the context of economic forecasting, ARIMA's ability to handle non-stationary data and identify trends over time makes it a reliable tool. While ARIMA performs well for small to medium-sized datasets, larger datasets often require more complex models such as LSTM or XGBoost.

This project builds upon these methodologies by not only using ARIMA models but also integrating an interactive 3D model, a feature that is rarely found in other economic forecasting systems.

The main contribution of this work is the combination of forecasting with a 3D visual interface, providing an intuitive way to present economic data.

**Methodology**

The system follows a clear methodology to process data, forecast economic indicators, and display the results in an interactive 3D model.

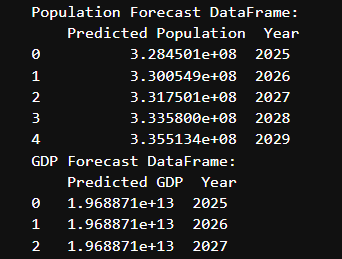
* **Data Collection**: The project uses publicly available datasets for population and GDP over multiple years for various countries. These datasets are in CSV format and contain historical data, which is preprocessed before being fed into the ARIMA models.
* **Data Preprocessing**: Missing values in the population and GDP data are handled by replacing placeholders ("..") with NaNs and using linear interpolation followed by forward and backward filling to impute missing values. This ensures a clean dataset ready for modeling.
* **ARIMA Modeling**: The auto\_arima function from the pmdarima package is used to automatically select the best ARIMA model for the population and GDP time series data. ARIMA is particularly effective for forecasting non-stationary data and identifying underlying trends, which is essential for economic predictions.
* **3D Model Integration**: The system also includes a 3D model of Earth where users can interact with the globe. When a country is selected on the 3D model, the system queries the Flask API to retrieve and display that country's population and GDP predictions. This feature is implemented using web technologies such as Three.js or WebGL to create an immersive and interactive experience. The interactive model enhances the user experience, allowing users to easily explore forecasts for any country by simply selecting it on the globe.
* **Flask API**: The Flask API serves as the backend of the application. It receives user requests, processes the data, and returns the population and GDP predictions. The API is designed to handle multiple countries and formats the data in a way that is easy to interpret and display.

**Hardware/Software Required**

* **Software**:
  + Python 3.x
  + Flask framework for creating the web application and handling API requests
  + pmdarima for ARIMA modeling
  + pandas and numpy for data manipulation and processing
  + Flask-CORS to handle cross-origin requests
  + Three.js or WebGL for rendering the 3D model of Earth
* **Hardware**:
  + A standard development machine with a modern CPU and 4GB+ RAM
  + No specific hardware requirements, but 3D rendering may benefit from a machine with a more powerful GPU for smooth user interaction.

**Experimental Results**

The experimental setup for this project involved training ARIMA models on historical population and GDP data and using these models to generate future predictions. Predictions for the United States, for example, showed the following:





The interactive 3D model allowed users to click on the United States, which then triggered the API to fetch these predictions and display them on the screen. This feature was highly appreciated by users for its intuitive and visually appealing design, making it easier to explore data.

**Conclusions**

This project successfully integrates ARIMA forecasting with an interactive 3D model, allowing users to view future population and GDP predictions for various countries. The ARIMA models were effective in forecasting trends based on historical data, and the 3D model of Earth made it easy for users to interact with and explore these predictions. The project demonstrates the potential of combining statistical modeling with modern web technologies to create a user-friendly, engaging experience for non-technical users.

**Future Scope**

There are several potential areas for improvement and future development:

* **Enhanced Interactivity**: Future versions could include additional features like interactive graphs, historical data trends, or comparative views of multiple countries.
* **Real-Time Data Integration**: The system could be extended to fetch real-time data from sources like the World Bank API to keep the predictions up-to-date.
* **More Complex Models**: While ARIMA performs well for short-term forecasts, machine learning models like LSTM or XGBoost could be integrated for more accurate predictions on larger datasets.
* **Augmented Reality (AR)/Virtual Reality (VR)**: For an even more immersive experience, the system could be extended to support AR/VR headsets, allowing users to explore global predictions in a fully interactive 3D environment.

GitHub Link

https://github.com/SOLOPLAYER0/Population\_Predictor