Mohamed Elenany

MAIS202 Project Deliverable 1

1. The Project and the Dataset:
   Harvard Business School has created a dataset that is named “Global Crisis by Country”, and it basically consists of data from 1800 to 2016 about most of the countries in the world. Data includes systematic crisis, inflation rate, USD exchange rate, etc. The dataset can be found here. I will be using data from this dataset to predict the USD exchange rate of any country.

2. Methodology:
   a. Data Preprocessing:
      The dataset comes in an xlsx and will be loaded into python by pandas. It has a lot of unnecessary columns that will not be used in the model, so those columns will be dropped. Also, the dataset has a lot of missing values which need to be handled. Other than that, the data is reliable and true, as all the data reflect true numbers that have been recorded by the countries in the data.
   b. Machine Learning Model:
      The objective of this project is to use the data from the dataset to predict the currency exchange price of USD for each country. I am proposing the use of three different machine learning models. The first is a random forest regression which I want to use due to the complexity of the data and due to the nonlinear property of data, however, it may not perform well due to not being able to extrapolate and in a lot of instances, it overfits the training data. I am also thinking of using linear regression as it is able to predict and extrapolate data and give weights to more important features, however, it may not perform well as the data is too complex and the linearity of data assumption it makes may not be applicable. The final model I am proposing is the LSTMs as it is known to perform well on time series data and can store past information and use it effectively, however, it is highly prone to overfitting and is very computationally expensive.
   c. Evaluation metric:
      I will be using mean squared error as an evaluation metric to see how accurate/close the USD exchange rate predicted was to the actual exchange rate.
3. Final Conceptualization/ Application:
   The project will be featured in a webapp where the user can choose a country and input the required data related to that country, and the webapp will show the user the predicted result.