Machine Learning Project
GDP Prediction

Abderrazak Bouasker
ITBS Nabeul

Application's QR Code

Introduction

In the era of data-driven decision-making, the application of machine learning techniques to economic forecasting has emerged as a powerful tool. This project focuses on leveraging machine learning algorithms to predict the Gross Domestic Product for Tunisia.

Problem Statement

Tunisia have been in a bad situation economically in the last ten years with major shifts in its budget , debt, and raising interest rates there are a lot of uncertainties around the future of the country and its economy .

Proposed Solution

The solution that we decided to use was to create a machine learning model that will help us predict the Gross Domestic Product for the next year given a set of variables that are relevant .

Dataset

The Source for all the data in our Dataset is the Central Bank of Tunisia

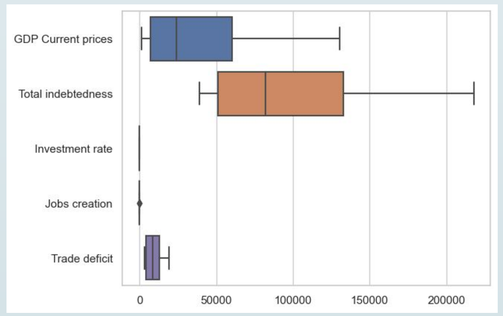
البنك المركزي التونسي
Banque Centrale de Tunisie — Central Bank of Tunisia

Dataset Description

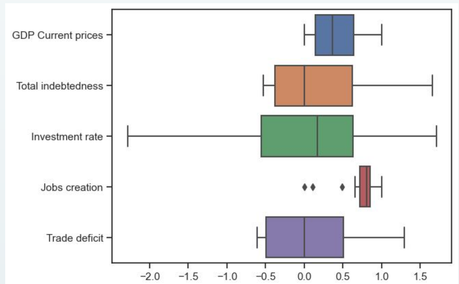
GDP :	Type : Float Unit : Million Dinar
Total Indebtedness :	Type : Float
Investment Rate :	Type : Float Unit : Percent
Jobs Created :	Type : Float Unit : Thousands
Trade Deficit :	Type : Float Unit : Million Dinar

Exploratory Data Analysis

For the Exploratory Data Analysis we went through multiple steps including :
Using Boxplot to study the data skewness and outliers.
Using Pairplot to study data density , distribution, and relations between two variables .
Using Heatmap to study the variable's correlation



After the initial data analysis we did some data cleaning :
Removing the empty data rows
Normalizing the data to the same scale



Model

The Machine Learning Model we used is Linear Regression which is a supervised algorithm that learns to model a dependent model y , as a function of some independent variables x .
The reason we chose to use linear regression is :

Interpolation : Linear regression can be effective for predicting values within the range of the observed data, making it useful for interpolation .

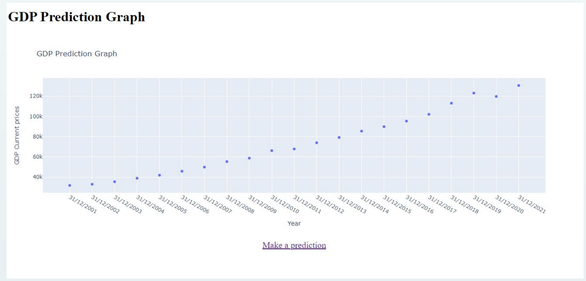
Simplicity and Interpretability :Linear regression models are simple to understand and interpret .

Training

For the training we split the data into two parts one part for training and another part for testing .
after training the model we did run multiple Cross Validation tests using loocv algorithm to determine what combination of features give us the best result .
By the end of the test we found that giving all the variables to the model to train on is the better option .
To determine the models quality we used Three Model Evaluation functions : Mean Absolute Error
Mean Squared Error
R-Squared

User Interface

To make the usage of the model easier we made a User Interface using flask for the backend and html , css for the frontend including a graph that easy to understand and interpret .



Hosting

To make the application available to everyone to use and test it we had to host it to a cloud server , for that we have used a cloud service called Render that gave us a public url that we shortened to the QR Code above for fast access .

Conclusion

Machine Learning is a very powerful tool to achieve multiple purposes one of which is prediction , while training a model isn't hard the work needed to make sure the model is of value and gives good results are hard , time consuming and requires a certain degree of competence and understanding of the whole process .