INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Recession Prediction using SVM and Linear Regression

Siddharth Nashikkar[1], Aryaa Hanamar[2], Nuren Pathan[3], Heena Mulla[4]

[1, 2, 3, 4]*Department of Computer Science and Engineering, D.Y. Patil College of Engineering and Technology, Kasaba Bawada, Kolhapur, 416006, Maharashtra, India.*

Abstract: In this study, we present a groundbreaking approach utilizing machine learning to predict economic recessions. By analyzing key indicators like real gross domestic product, income, employment, manu- facturing, and retail sales, we can forecast the likelihood of a recession occurring in a specific year using Python 3 and employing Linear Regression and Support Vector Regression algorithms. This user-friendly application not only provides advanced warning of recessions but also assists policymakers and businesses in making timely and informed decisions for economic stability and resilience.
Keywords: Recession, Support Vector Machine (SVM), Linear Regression, Tkinter

I. INTRODUCTION

A recession is when the economy experiences a marked slippage in economic activity. There's a drop in the following five economic indicators: real gross domestic product, income, employment, manufacturing, and retail sales. People often say a recession is when the GDP growth rate is negative for two consecutive quarters or more. These are the indicators to watch if you want to know when the economy is in a recession.

1) Real income measures personal income adjusted for inflation. Transfer payments, such as Social Security and welfare payments, are also removed. When real income declines, that reduces consumer purchases and demand.
2) Employment as measured by the monthly jobs report. Here's an analysis of the current jobs statistics.
3) The health of the manufacturing sector, as measured by the Industrial Production.
4) Manufacturing and wholesale retail sales adjusted for inflation.
5) The NBER also looks at monthly estimates of GDP provided by Macroeconomic Advisers.

Recessions are characterized by a rash of business failures and often bank failures, slow or negative growth in production, and elevated unemploy- ment. The economic pain caused by recessions, though temporary, can have major effects that alter an economy. Hence analysis is required to give advance warning of a recession and avoid ill effects on the economy.

II. OBJECTIVES

1) To predict the future recession that is it will tell the period of recession in future.
2) Period of the current recession.

III. REQUIREMENT SPECIFICATION

A recession is commonly understood to be a widespread and prolonged decline in economic activity which leaves its impact on that economy for a very long time. Predicting this recession before its occurrence can help in taking up preventive measures in time and avoid the loss to a great extent. The prime requirement of this project is to predict the probability of recession in the entered year on the basis of the data from previous years that will be provided to train the model. The project will be a window-based application where the user will be asked to enter the year for which prediction should be done and click on predict. Once the year is taken the model will predict the recession possibility in that year.

1) *Technologies used*
a) Python 3:version 3.6
b) Machine learning algorithms

2) Software Requirements

a) Microsoft Windows 8 or later

b) Pycharm

c) Microsoft Excel

3) Hardware Requirements

a) Intel Core i3 processor

b) RAM:4GB

c) Disk space: 2.5GB and another 1 GB for caches

d) Monitor Resolution: 1024*768

4) Libraries Used

a) Pandas & Matplotlib

b) Numpy

c) Sklearn

d) tkinter

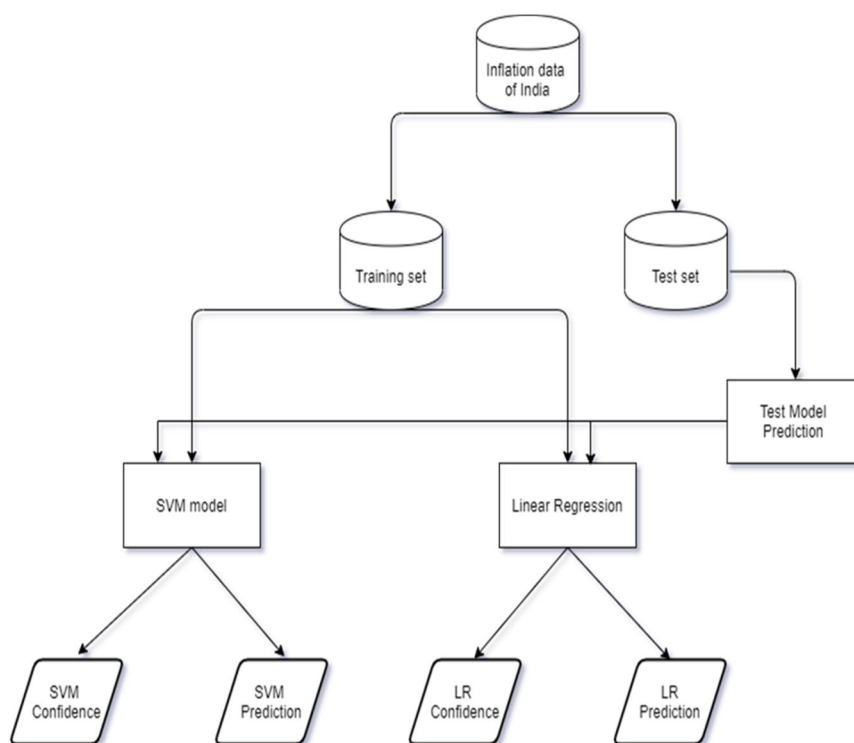## IV.   DESIGN

A. System Architecture

Fig. 1 System Architecture

The dataset used here consists of the data about the inflation rate in India from the year 1958 to 2018. After data cleaning it was split into train data and test data with the help of train test split function from sklearn. We have used two different algorithms for model training i.e. SVM and linear regression. On the basis of test model prediction and the specific model(SVM or LR), a prediction of the rate of inflation for the entered year is made and displayed along with the confidence score of that model so that we can consider the results of the model with more accuracy.

B. Flow chart
1) Start.
2) A window window with all the components.
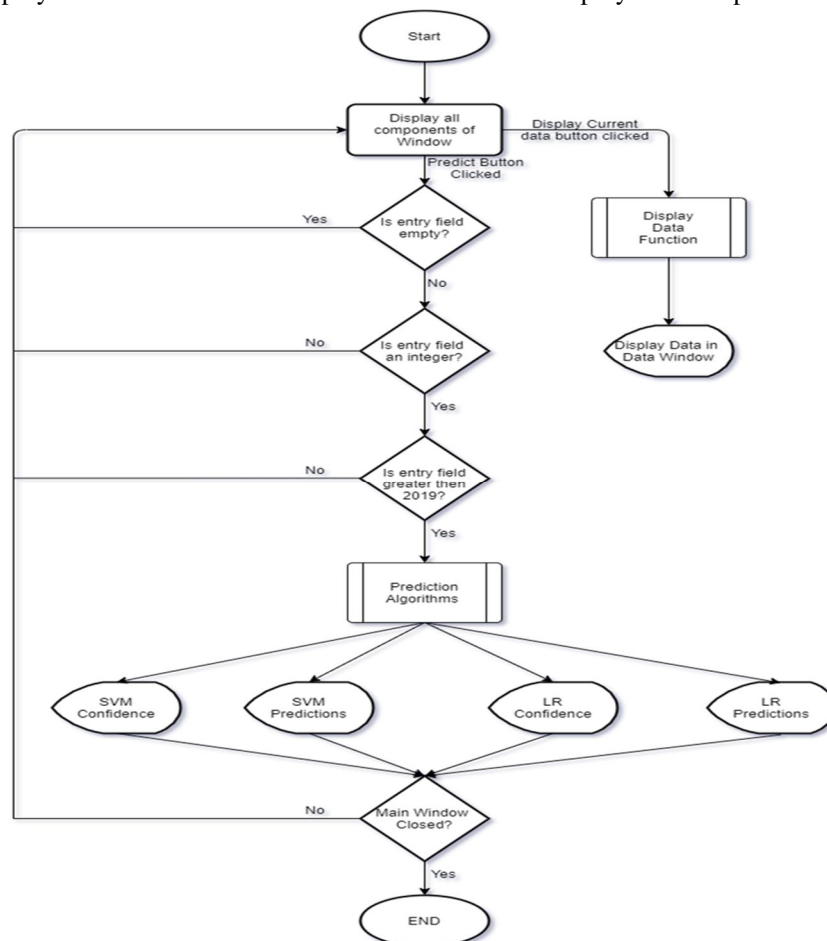3) On clicking the Display Current Dataset button the entire dataset is displayed in a separate window.

Fig. 2 Enter Caption

4) Next, the user is asked to enter the required year and click on the predict button.
5) After entering the value is checked if the entry field is empty.
a) If yes then an exception occurs which has been handled in the code and goes to step 4.
b) If no then proceed to step 6.

6) Now we have to check if the entered value is an integer.
a) If yes proceed to step 7.
b) If no message "Enter valid number is dis- played" go to step 4

7) If the entered year is not greater than 2019 then the user is asked to re-enter the value which should be greater than 2019.
8) Later in the flow once the correct input from the user is obtained prediction for the entered year is done on the basis of our training algo- rithms.
9) The predicted result is then displayed on the window along with the confidence score.
10) The main window closed
a) If yes then END
b) If no then go to step 4.

V. ALGORITHMS USED

1) *Linear Regression:* Linear regression is a linear model, e.g. a model that assumes a linear rela- tionship between the input variables (x) which is the date of the year in our case and the single output variable (y) which is the inflation rate. More specifically, that y can be calculated from a linear combination of the input variables (x)

2) *Support Vector Regression (SVR):* Support Vec- tor Regression(SVR) is quite different than other Regression models. It uses the Support Vector Machine(SVM, a classification algo- rithm) algorithm to predict a continuous vari- able which is GDP in our case.

VI. IMPLEMENTATION DETAILS

A. *Function Used*

1) *def displayData():* This function is used for cre- ating buttons and labels. when we click on the button it will run the function with whom it is initialized and display the output.

2) *def displayPredict():* This function is used for predicting the inflation value after 2019. It takes valid input from users and gives the required results.

B. *Development Tools Used*

1) *PyCharm:* It is an integrated development environment (IDE) used in computer program- ming, specifically for the Python language. It is developed by the Czech company JetBrains. It provides code analysis, a graphical debug- ger, an integrated unit tester, integration with version control systems (VCSes), and supports web development with Django as well as Data Science with Anaconda. PyCharm is cross- platform, with Windows, macOS and Linux versions.

2) *Libraries*

a) *Pandas:* Pandas is a fast, powerful, flexible and easy-to-use open source data analy- sis and manipulation tool, built on top of the Python programming language. Data analysis was done using pandas.

b) *Numpy:* NumPy was used as an efficient multi-dimensional container of generic data. Arbitrary data types can be defined.

c) *Matplotlib:* Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. Data visualization was done using pyplot from matplotlib.

d) *Sklearn:* This module was used to implement the different machine learning algorithms used in the project.

e) *Tkinter:* Tkinter is the Python interface to the Tk GUI toolkit shipped with Python. The GUI of the projected id is created using the library.

3) *Microsoft Excel:* Used Microsoft's product Excel to cover a range of machine learning tasks such as data mining, data analytics, and smart visualization.

C. *Supporting Tools Used*

1) Train test split is a function in the Sklearn model selection for splitting data arrays into two subsets: for training data and for testing data. With this function, you don't need to divide the dataset manually.

2) matplotlib.pyplot is a collection of command style functions that make matplotlib work like MATLAB. Each pyplot function makes some change to a figure: e.g., creates a figure, creates a plotting area in a figure, plots some lines in a plotting area, decorates the plot with labels, etc.

VII. TESTING AND RESULTS

Testing is the process of executing a program or application with the intent of finding software bugs. Software Testing is necessary because we all make mistakes. Some of those mistakes are unimportant, but some of them are expensive or dangerous.

We need to check everything and any- thing we produce because things can always go wrong – humans make mistakes all the time. So there are some test cases provided below:

A. Test Case 1

If we enter an invalid integer or string then it will give the error and ask us to enter the valid value.
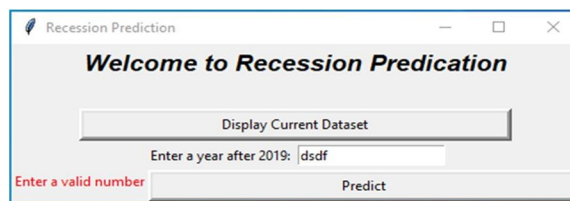

Fig. 3 test case 1

B. Test Case 2

If we don't give any value and push the predict button then it will give the error and ask to enter a value.
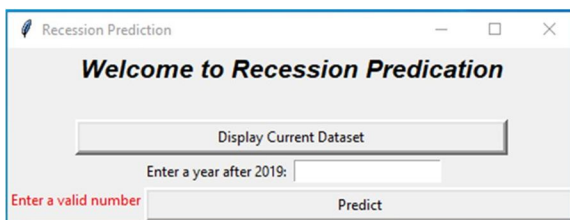

Fig. 4 test case 2

C. Test Case 3

If we give the previous year to predict then it also gives the error and asks for the years after 2019
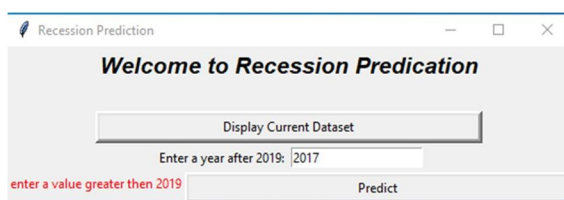

Fig. 5 test case 3

D. Test case 4

Input: Year after 2019 Output: SVM confidence: 0.1973 LR confidence: 0.3881 SVM prediction: 3.1503 LR prediction: 6.5915 As LR confidence is greater than SVM confidence so we will consider LR predication values.
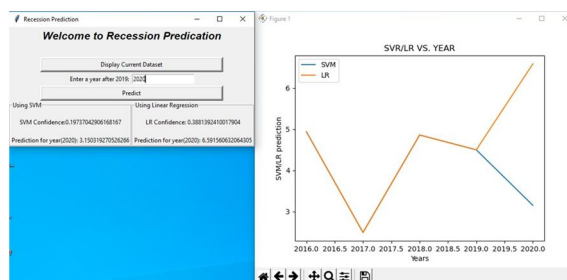

Fig. 6 test case 4

VIII. CONCLUSION

[H] Recession has been a major problem for any economy today. The department hit by this prob- lem needs a lot of time and effort to cope with the loss. The increased rates of unemployment, infla- tion and many other difficulties may also lead to other major problems which include suicide rates among youth due to unemployment and many more. As a preventive measure, we need to predict this situation beforehand to take the necessary steps. Hence we have created this model where we have made use of Supervised learning algorithms like Linear regression and support vector machine using regression(SVR).

We have successfully pre- dicted a recession on the basis of past data. The users of this application will be able to see the rate of inflation and accordingly understand the condi- tion of future recession which will be a great boon for any economy.

IX. STATEMENTS AND DECLARATIONS

A. Competing Interests

The authors declare that they have no conflict of interest. All the authors have agreed to the submission.

B. Code Base

https://github.com/sid732/RecessionPrediction

REFERENCES

[1] Dataset - https://machinelearningmastery.com/confidence-intervals-for-machine-learning/

[2] https://towardsdatascience.com/linear-regression-understanding-the-theory- 7e53ac2831b5

[3] https://medium.com/coinmonks/support-vector-regression-or-svr-8eb3acf6d0ff

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)