

# MODELING AND FORECASTING GROSS DOMESTIC PRODUCT (GDP) USING LINEAR REGRESSION

<sup>1</sup>Aparna V.Mote, <sup>2</sup> Snehal Chauvan, <sup>3</sup>Kirtikummar Waykos

<sup>1</sup>Assistant Professor, <sup>2</sup>Student, <sup>3</sup>Student

<sup>1</sup>Computer Engineering Department,  
<sup>1</sup>ZCOER, Pune, India

**Abstract:** Gross Domestic Product (GDP) is mostly known as the main measure of economic performance. It is an economic survey of market values of all merchandise and services produced over a period of time, usually annually or quarterly. GDP is Nominal when calculated for current prices. When calculated for constant prices, it is Real GDP. The increase in real GDP indicates increase in the national output GDP per capita. GDP per capita at purchasing power parity (PPP) is possibly more useful while comparing the living standards between nations. This paper highlights the major components involved in GDP calculation. This paper reviews various indicators used for measurement of GDP and how they affect GDP. This paper used linear regression for GDP prediction.

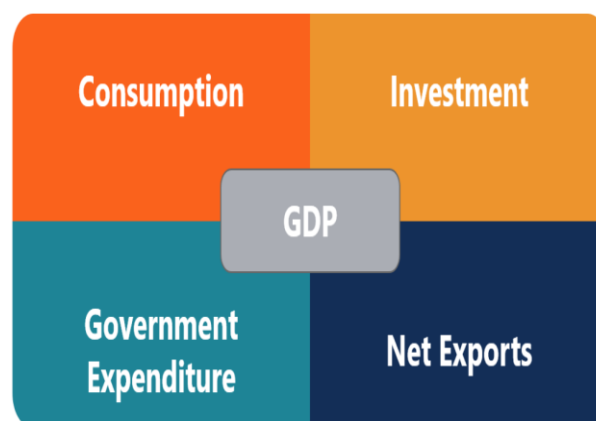
**Index Terms** – GDP, Machine Learning, Data Mining, Linear Regression

## I. INTRODUCTION

Gross Domestic Product (GDP) is an economic measure which records the quantity of goods as well as services produced within nation's economic period. It is an economic survey of market values of all merchandise and services produced over a period of time, usually annually or quarterly. The main aim is to record the level of production within an economy. This is considered as main factor for calculating the economic performance. GDP value gives the total size of an economy. Overall health of the economy is generally measured by considering the changes in GDP usually referred as real growth in GDP.

GDP combines all private as well as public consumptions, government expenditures, various investments, difference of exports and imports. Usually GDP is calculated with following formula:

Following figure shows the four important components of GDP:



**Fig 1: Components of GDP**

- Consumption (C):

This component calculates money value of consumer goods as well as services purchased by households or profitless institutions during the year. These are categorized into consumer durables, semi-durables, non-durables and services. These categories mainly consider length of time within which goods are used.

- Investment (I):

Investment is nothing but the addition to the physical stock during specific time period. Gross Private Domestic Investment gives the collaborative value. This considers manufacturing of housing construction's machinery, factories construction. Goods used in the process of manufacturing of other goods called as intermediate goods. These intermediate goods get partially consumed in the production of other goods as well as services. This is known as Depreciation of fixed capital goods. Depreciation is nothing but decrease in the value of current capital stock which is used in output production.

Investment can be either gross or net. Gross investment considers depreciation value and when depreciation is deduced from gross investment, it is net investment.

- Government Expenditure (G):

Government expenditures mainly focus on spending of government on goods as well as on services. This mainly includes:

- i. Intermediate goods purchase
- ii. Wages as well as salaries given

All these are substitute measure for government output. These government purchases are considered as final product. The payment transfers made by government to firms, households are not considered. This is done for avoiding the double counting as they are counted in consumption and in investment.

- Net Exports (X-M):

This component shows the difference between imports and exports that is difference between Spending on foreign goods and foreign spending on domestic goods that is Difference of Exports (X) and Imports (M) gives net Exports (X-M).

Thus Gross Domestic Product (GDP) is addition of all these components that is consumption, Investment, Government Expenditure and net exports (imports minus exports)

### 1.1 GDP Growth Rate

GDP growth rate is much useful while measuring the improvement in nation's GDP. This compares the current level of GDP with previous year. Percentage difference is considered as the growth rate.

Usually, Goal of Australia for Sustainable Economic Growth is between 3-4% which is set by considering inflationary pressures and unemployment and other factors responsible for recession.

### 1.2 GDP Per Capita:

GDP is an indicator to find the level of living standards within the economy of a nation. Whereas GDP per Capita metric is detailed indicator of living standards. This divides the overall level of GDP by population of that nation. This is very helpful for indicating the 'economic pie'. This is technique of calculating the living standard of a country.

Gross National Income per capita is equal to Gross National Income divided by Total Population.

For example: GDP Per Capita of India in 2017 was \$1963.55.

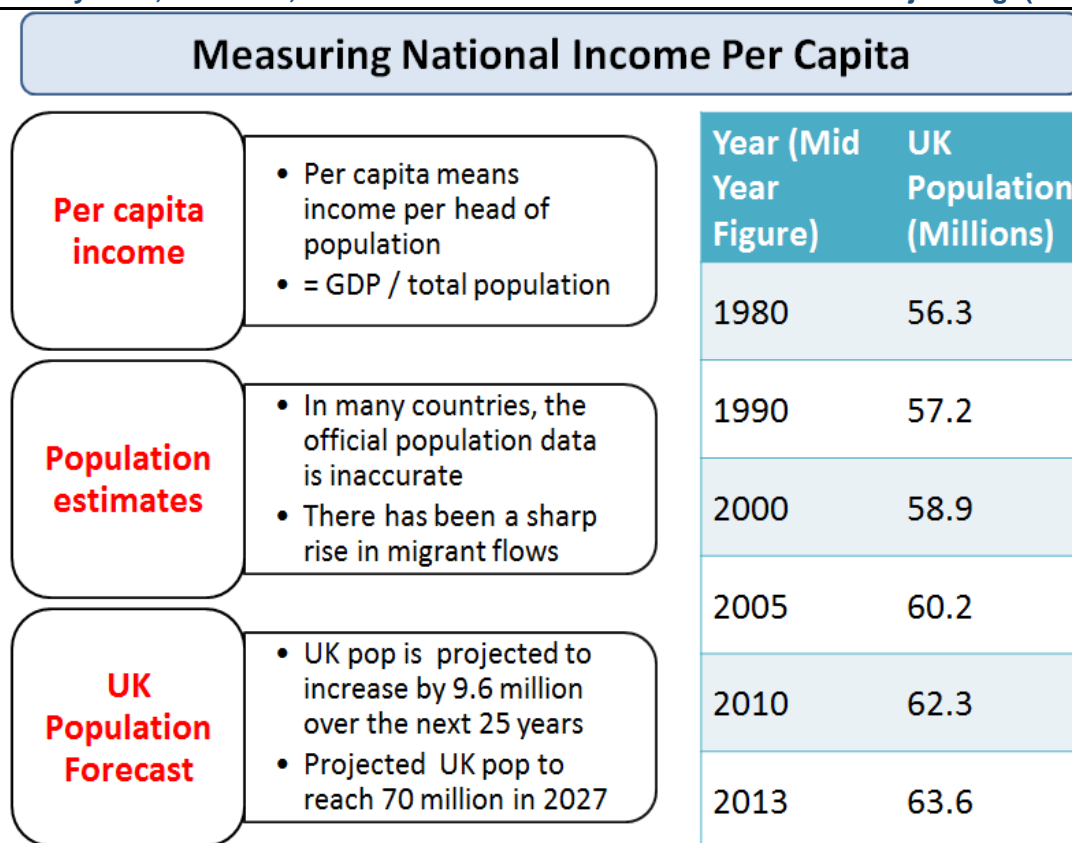


Fig. 2: Per Capita National Income

### 1.3 Real vs Nominal GDP:

Real GDP is a measure of how much is actually produced. Real GDP measures aggregate output using constant prices, thus removing the effect of changes in the overall price level. For example, in 2015 the value of Canada's output expressed in constant 2010 prices was CAN \$1,857 billion.

Whereas Nominal GDP is a measure of how much is spent on output. For example in 2015, total CAN \$1994.9 billion were spent on goods and services produced in Canada. Nominal GDP measures aggregate output that is the value of all the final goods and services produced using current prices.

Terms used for Real and Nominal GDP:

Term	Definition
GDP Deflator	A price index used for adjusting nominal GDP and real GDP. The GDP deflator measures average prices of all finished goods and services produced within a year.
Base year	The year used for comparison in the determination of price changes using the GDP Deflator price index. The deflator in base year is always 100.
Current Prices	The prices at which goods are sold in particular year
Constant Prices	The prices from a base year those are used for calculating real GDP in other years. This gives more accurate results to check how a country's actual output changes over time.

The two approaches used to adjust nominal GDP to get real GDP are:

- 1) Using the same prices every year
- 2) Using the GDP deflator.

- $\text{Real GDP} = \text{Nominal GDP} / \text{GDP Deflator in hundredth}$
- $\text{Nominal GDP} = \text{Real GDP} * \text{GDP deflator (in hundredth)}$

Comparison of Real and Nominal GDP:

Parameter	Nominal GDP	Real GDP
Definition	Sum of total of the economic output produced in a year evaluated with current market price.	Real GDP is the sum-total of the economic output produced in a year evaluated at a pre-determined base market price
Considers	Current price	Market price of base year
Effect of inflation	Does not take inflation into account	Takes inflation into account known as Inflation adjusted GDP.
GDP Value	Much higher as current prices are considered.	Much lower as base year's market price of is taken into consideration.
Popularity	Less popular	More popular
Complexity	Easy to compute	Bit complex
Growth of the economy	Economic growth can't be analyzed easily.	Economic growth can be analyzed easily.

#### 1.4 GDP Indicators:

Two types of indicators used in GDP are:

1. Leading indicators usually change to big economic adjustments and these can be used for predicting future trends.
2. Lagging indicators, reviews the economy's historical performance.

##### 1.4.1 Leading Indicators

As the leading indicators play important role in predicting where an economy is leading. With the help of these indicators, budgetary policymakers as well as government implement programs to prevent nation from recession or other negative economic things. Following are some leading indicators:

##### 1. Stock Market:

Stock market is liked by most of the people even though it is not the most important indicator. This is because most of the companies' earnings are based on stock market. This indicates the direction of economy if accurate estimates of earnings are calculated.

For example, strong market tells that estimates of earnings are up and overall economy is prospering. Whereas down market indicates that earnings of company are decreasing and economy may lead to recession.

However there are some defects if we consider on stock market it as leading indicator. First important defect is the estimation of earnings can be wrong. Another is it is susceptible to manipulation.

Finally, stock market may give false positive for market's direction as it is vulnerable to the creation of bubbles. Bubbles occurs when the investors ignores the fundamental economic indicators which leads to increase in unsupported price levels. This can create "perfect storm" for market correction. Example of same is market crash in 2008 due to overvalued subprime loans and credit default swaps.

##### 2. Manufacturing Activity:

Manufacturing activity is another leading indicator. This strongly influences GDP. Increase in this suggests demand for consumer goods indicating healthy economy. As this requires manpower for manufacturing, boosts employment and salary as well.

However, increase in this can be misleading. For example, goods produced sometimes do not pass it to the end consumers. They may be there at wholesale or retailers inventory for some time. This increases cost of asset holding. Thus while considering manufacturing data, retail sales data should also be considered. If both are on demand then it indicates increased demand for consumer goods. Inventory levels should also be considered.

### 3. Inventory Levels:

High inventory levels reflect two things:

- i. Increase in Inventory demand
- ii. Decrease in Inventory demand

In first case, businesses intentionally round up inventory to show increase in consumption. If customer demand is increased as per the expectations, business profit will be increased also there will be increase in inventory demand. Both are useful things for the economy.

In another case, sometimes high inventories tell that company provides exceed demand. This costs companies money as well as indicates that retail sales are down. This reflects that tough time is going to come.

### 4. Retail Sales:

These are as important as Inventory levels and Manufacturing activity are. When sales are increased, companies hire more employees which in turn put money back in customer's pockets. Increase in Retail sales indicates improvement in economy.

### 5. Building Permits:

These indicators give future views for levels of real estate supply. High volume indicates activeness of the construction industry which forecasts increase in jobs which automatically increase the GDP. Whereas the reverse effect will be dangerous.

### 6. Housing Market:

Decrease in housing prices indicates that there is excess supply than demand and existing prices are not affordable. These should be corrected to avoid housing bubbles. Decrease in housing price shows negative impact on economy. This may impact on:

- i. Decrease the wealth of homeowner
- ii. Decrease the number of jobs in construction
- iii. Decrease the property taxes

When housing data is considered, two things should be taken care of changes in housing values and sales. When sales decline, drop in values is observed. For example decrease in housing bubble in 2007 had worst effect on the economy.

## 1.4.2 Lagging Indicators:

Lagging indicators move after change in the economy. Though they do not inform us economy, they tell how economy changes after certain period of time which is helpful to find long term trends. Following indicators are considered as lagging Indicators:

### 1. Changes in the Gross Domestic Product (GDP):

GDP is important measure of the economy's current health. When there is increase in GDP, it indicates that economy is strong and so businesses will plan the production based on the output of GDP. GDP has some flaws. Like stock market, GDP can also be misleading.

This simply tells us what has happened and not what is going to happen in future.

### 2. Income and Wages:

When economy is working properly, earning should also increase regularly. When income decreases, it indicates that either employers will cut the pay rates or work hours will be reduced. Wages change based on different types of demographics used for incomes like age, level of education.

### 3. Unemployment Rate:

This is important measure to calculate the number of people looking for job. In a healthy economy, the unemployment rate will be between 3% and 5%. When unemployment rates are increased, it affects the GDP. This can also be misleading like other indicators.

### 4. Consumer Price Index (Inflation):

This indicator tells increased living cost or inflation. This is calculated by measuring the cost of necessary goods and services which includes vehicles, shelter electronics and so on. Inflation is calculated with average increased cost of total basket of goods over specific period. High inflation rate decrease the value of dollar. Inflation may also decrease the employment rate and also GDP. However it is not always bad. It has some benefits at intermediate levels of inflation like:

- a) Motivates spending and investing, which helps economy to grow.

- b) Keeps interest levels at intermediate level, which motivates people for investments
- c) It is not deflation which may cause economic depression

Deflation is something where living cost decreases. Though it sounds good, it actually indicates economy's poor shape. This occurs because of reduction in money supply. This forces retailers to reduce their prices, so it affects the profits.

Currency strengths, Interest rates, corporate profits, balance of trade these are also considered as lagging indicators.

## II LITERATURE REVIEW

### 2.1 Gross Domestic Product prediction using Autoregressive Models:

John Roush et al. [1], proposed autoregressive model for prediction of GDP. This model changes constantly based on new figures on productivity, consumption and investments. This can be used for any country in GDP prediction. But, this does not count unpaid volunteer work and quality of work. Also disaster can affect GDP.

### 2.2 Forecasting and simulation model:

George I. Treyz et al. [3] proposed simulation model for GDP prediction. This model is able to estimate economic and demographic effects of economic development which improved research based decision making. This updates dynamically. But, it is limited to regional forecasting and applied for larger platforms like at country level.

### 2.3 provincial panel Data:

X Yang et al. [5] studied the relationship between financial development and economic growth of China during period 1978 to 2007. A framework of data is used for the specification and considered as long term relationship for financial development and growth. Other factors are also analyzed. The results shown that there is unidirectional causality from financial development and economic growth in China

### 2.4 Bridge Models:

Alberto Baffigi et al. [6], proposed Bridge model for forecasting European GDP. They examined the forecast ability of bridge models for growth of GDP in Euro area. Bridge models are used to bridge the gap between timely updated indicator's information and delayed national account. The models are estimated for aggregate GDP. This paper shown that national Bridge models are better than benchmark models. GDP of Euro area and its components predicted results very precisely with the help of aggregation of national forecasts.

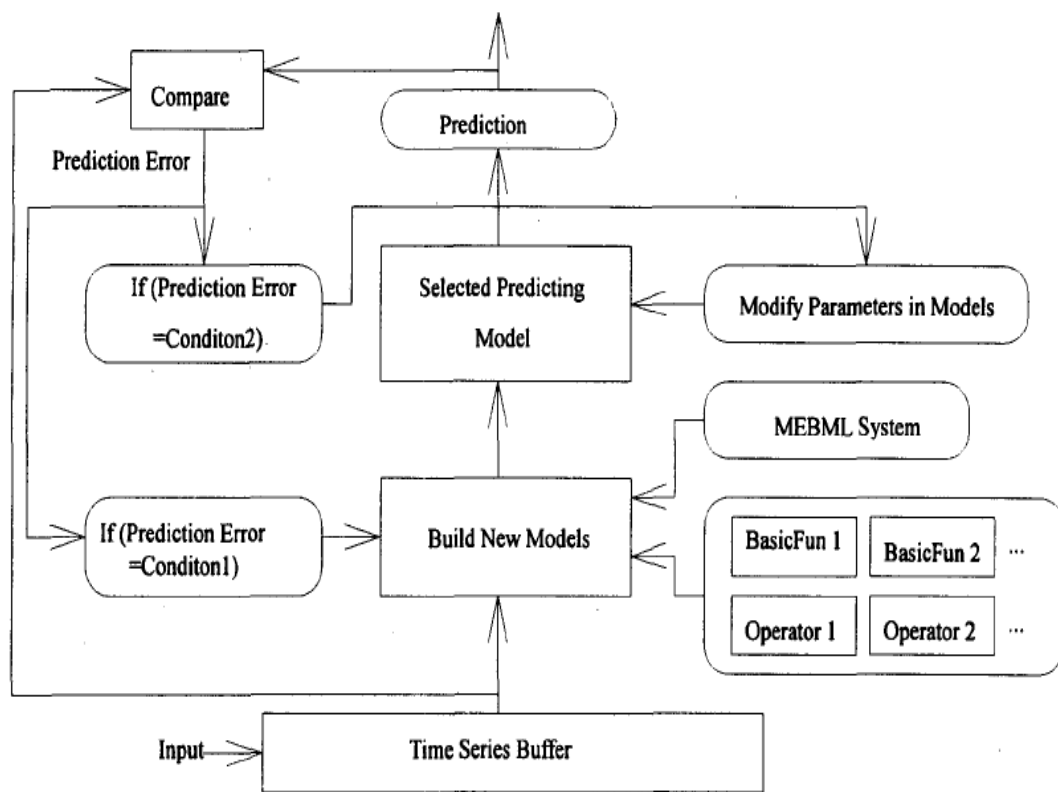
In [9], Karim Barhoumi et al. proposed a model to predict the GDP for France. This is designed to work for monthly economic data with the help of bridge models for supply as well as demand. For every component of GDP, bridge equations are used with the help of general to specific approach. This approach selects variables which are self explanatory among the dataset. The results shown that changing the set of equations over period of time i.e. quarterly gives good results. The GDP growth rate can also be predicted.

### 2.5 Dynamic Factor Model:

S Guichard et al. [7], explained monthly indicators useful for forecasting the world trade. They compared various types of forecasting models with the help of these indicators. This paper proposed a dynamic Factor Model which is capable of handling large datasets of information. Dynamic factor models are more capable than bridge models [6]. This allows adding many monthly indicators like financial indicators, supply and order indicators [8] etc. when compared with traditional bridge models and with autoregressive models, proposed model performed better especially when large number of indicators are used. But marginal gains affect the performance.

**III PROPOSED METHODOLOGY:****A. Proposed System:**

The basic flow of the system depends upon the scenarios present and items which work accordingly. The section gets divided to various modules. Following diagram shows the architecture of proposed system:



**Fig 3: System Architecture**

In this proposed system we are using linear regression algorithm. Linear regression plays an important role in the field of artificial intelligence such as machine learning. The linear regression algorithm is one of the fundamental supervised machine-learning algorithms due to its relative simplicity and well-known properties. Regression is a method of modeling a target value based on independent predictors. This method is mostly used for forecasting and finding out cause and effect relationship between variables. Regression techniques mostly differ based on the number of independent variables and the type of relationship between the independent and dependent variables.

**B. Mathematical Model: Set Theory:**

- Let S be closed system defined as,  
 $S = \{Ip, Op, Ss, Su, Fi, A\}$   
 $Ip = \{Username, Password\}$   
 To select the training dataset for performing various actions from the set of actions A so that Su state can be obtained.
- Set of Actions  $A = \{F1, F2, F3, F4\}$   
 Where F1= provide dataset as input  
 F2= Perform Prediction  
 F3= Compare prediction Model  
 F4= Build Model  
 S= set of users
- $Ss = \{rest\ state, login\ state, selection\ of\ training\ dataset, learning\ process, selection\ of\ basic\ model, predicting\ the\ result\ and\ display\}$
- $Su = success\ state\ is\ successful\ analysis$
- $Fi = Failure\ State$

**C. Linear Regression Algorithm:**

- Step 1: Start with a training set with  $x_1, x_2, x_3, \dots$  and  $y$ .
- Step 2: start with parameters  $c_0, c_1, c_3$  with random values.
- Step 3: with learning rate  $\alpha$

Step 4: Then repeat the following update

$$C0 = c0 - \alpha * h(x) - y$$

$$C1 = c1 - \alpha * (h(x) - y) * x$$

Step 5: Repeat step 4 till it converges.

#### D. Advantages of linear regression algorithm:

Linear Regression is a good supervised learning algorithm which is used to predict values. It finds the target variable by finding a best suitable fit line between the independent and dependent variables.

Its main advantage is, the best fit line is the line with minimum error from all the points. It has high efficiency but sometimes this high efficiency created disadvantage which is prone to over fitting of the data (i.e. some noisy data also considered as useful data), and also it can't be used when the relation between dependent and independent variable is not linear.

#### IV Strengths of the GDP:

Following are the strengths of GDP:

- This is very important indicator of economic output and growth.
- Real GDP considers inflation into account, and allows for comparisons against other historical time periods
- Helps policy makers and analysts for implementing economic policy.
- Provides an analysis measure of economic activities.

#### V The limitations of GDP:

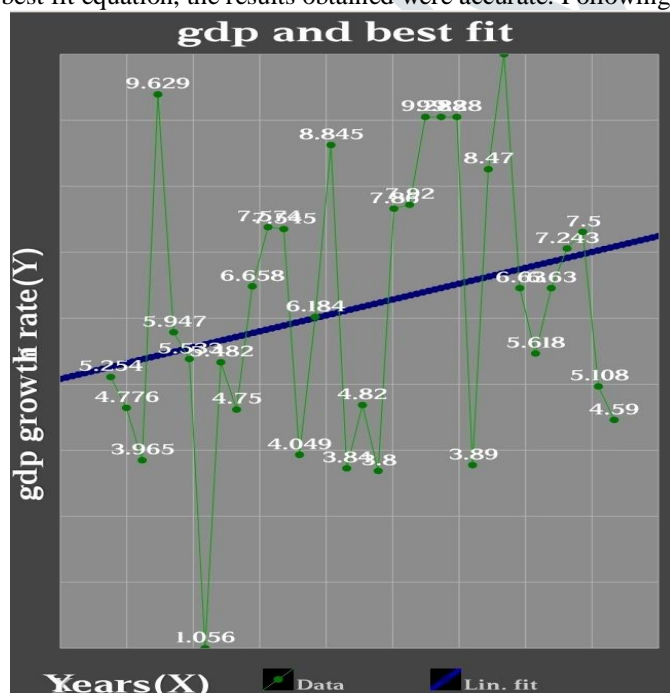
Gross Domestic Product is a useful indicator of a nation's economic performance, and it is the most commonly used measure of well-being. However, it has some important limitations, including:

- The exclusion of non-market transactions
- The failure to account for or represent the degree of income inequality in society
- The failure to indicate whether the nation's rate of growth is sustainable or not
- The failure to account for the costs imposed on human health and the environment of negative externalities arising from the production or consumption of the nation's output
- Treating the replacement of depreciated capital the same as the creation of new capital

#### VI RESULTS:

Linear Regression is a good supervised learning algorithm which is used to predict values. It finds the target variable by finding a best suitable fit line between the independent and dependent variables.

Its main advantage is, the best fit line is the line with minimum error from all the points. It has high efficiency. With the help of best fit equation, the results obtained were accurate. Following figure shows the best fit line for the dataset considered.





**VI. Conclusion:**

Gross Domestic Product is a useful indicator of a nation's economic performance. It is an economic survey of market values of all merchandise and services produced over a period of time, usually annually or quarterly. This paper explains various indicators used in GDP measurement. Overviews various techniques used for GDP measurements also the problem of GDP forecasting is analyzed. The usability of Learning Algorithms and problem formulation for these methods has been discussed. The importance of GDP forecasting has been rightly emphasized with the help of Linear Regression technique.

**References:**

- [1] John Roush, Keith Siopes, Gongzhu Hu, 'Predicting Gross Domestic Product Using Autoregressive Models', IEEE Computer Society, June, 2017.
- [2] David A. Dickey, Wayne A. Fuller, 'Distribution of the Estimators for Autoregressive Time series with a Unit Root', Journal of The American Statistical Association, 2002.
- [3] George I. Treyz, Dan S. Rickman, Gang Shao, 'The REMI Economic-Demographic forecasting and simulation model', International Regional Science Review, Vol. 14, pp. 221-253, 1992.
- [4] Alberto Baffigi, Roberto Golinelli and Giuseppe Parigi, 'Real-Time Gdp Forecasting In The Euro Area', 2002.
- [5] X Yang, W Yuan, 'Financial development and economic growth: An analysis based on provincial panel data', IEEE International Conference on Information and Financial Engineering, 2010.
- [6] Alberto Baffigi, Roberto Golinelli and Giuseppe Parigi, 'Bridge models to forecast the euro area GDP', International Journal of Forecasting, Vol. 20, Issue 3, pp 447-460, September 2004.
- [7] S Guichard, E Rusticelli, "A Dynamic Model For World Trade Growth", ECO/WKP, 2011
- [7] Gross Domestic Prediction. [https://en.wikipedia.org/wiki/Gross\\_domestic\\_product](https://en.wikipedia.org/wiki/Gross_domestic_product).
- [8] Aparna V. Mote, Pratima Patil, Tejaswini Mane, 'Information Retrieval in CBR using vertical association Knowledge and Shannon Information Gain', IEEE conf. on Information, Communication, Engineering and Technology, Aug. 2018.
- [9] Karim Barhoumi, Olivier Darné, Laurent Ferrara, Bertrand Pluyaud, 'Monthly GDP Forecasting Using Bridge Models: Application For The French Economy', Bulletin of Economic Research, 2011.
- [10] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung, 'Time series analysis: forecasting and control', John Wiley & Sons, 2015.
- [11] C. Brooks, "Introductory Econometrics for Finance", Cambridge University Press, Cambridge, 2008.
- [12] Helmut Lutkepohl. Vector autoregressive models. Springer, 2011.
- [13] John C Robertson and Ellis W Tallman. Vector autoregressions: forecasting and reality. Economic Review-Federal Reserve Bank of Atlanta, 84(1):4, 1999.