

# Sectoral analysis of NIFTY50

*A project report submitted by*  
**Sachin Negi**

*in partial fulfillment of the requirements for the award of the degree of*  
**Master of Technology**



॥ त्वं ज्ञानमयो विज्ञानमयोऽसि ॥

**Indian Institute of Technology Jodhpur**  
**Department of Mathematics**  
*June 2021*

## **Declaration**

I hereby declare that the work presented in this project *Sectoral Analysis of NIFTY50* submitted to the Indian Institute of Technology Jodhpur in partial fulfillment of the requirements for the award of the degree of Master of Technology, is a bonafide record of the research work carried out under the supervision of Dr. Kirankumar Hiremath. The contents of this project in full or in parts, have not been submitted to, and will not be submitted by me to, any other Institute or University in India or abroad for the award of any degree or diploma.



*Sachin Negi*  
**MT19DCS005**

## **Certificate**

This is to certify that the project titled *Sectoral Analysis of NIFTY50*, submitted by *Sachin Negi (MT19DCS005)* to the Indian Institute of Technology Jodhpur for the award of the degree of *Master of Technology*, is a bonafide record of the research work done by him under my supervision. To the best of my knowledge, the contents of this project, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

*Dr. Kirankumar Hiremath*  
*M.Tech.Project Supervisor*

## **Abstract**

Nifty50 index consists of 50 companies. These constituent companies can be grouped into different sectors. Each sector shows a different trend of movement at a given time compared to that of the Nifty50 index. The main objective of this study is to understand this trend and apply technical indicators to predict the movements for portfolio optimization.

This report is a detailed account of all the work done from starting of the project. The study started with looking for different models to predict the stock price. Many models were tested including traditional models consisting of technical indicators and modern models consisting of machine learning and deep learning models. Further work was concentrated on Nifty50 and its components. The trend of the components is analyzed to see any potential trend with the Nifty50. Next, the study is done on looking for a potential relationship between the GDP of the country and Nifty50. At last the Nifty50 sectors analysis is done using clustering algorithms. It was found that the sector movement gives a better picture to predict the future.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Literature Review</b>	<b>2</b>
<b>3</b>	<b>Problem Statement</b>	<b>4</b>
3.1	Methodology . . . . .	4
<b>4</b>	<b>Experimental Findings</b>	<b>5</b>
4.1	Stock prediction using LSTM network data . . . . .	5
4.2	Analysis of Nifty index and its component . . . . .	7
4.3	Analysis of GDP and Nifty 50 . . . . .	8
<b>5</b>	<b>Sectoral Analysis</b>	<b>10</b>
5.1	Sectoral analysis of Nifty 50 . . . . .	10
<b>6</b>	<b>Clustering</b>	<b>16</b>
6.1	How Nifty 50 index is calculated . . . . .	16
6.2	Weightage of Nifty50 component . . . . .	16
6.3	k-mean clustering . . . . .	16
6.3.1	k-mean clustering algorithm . . . . .	17
6.4	Application . . . . .	17
6.4.1	Dataset . . . . .	17
6.4.2	Methodology . . . . .	17
6.5	Result . . . . .	18
6.6	Observation . . . . .	18
<b>7</b>	<b>P/E Ratio</b>	<b>21</b>
7.1	Definition . . . . .	21

7.2	Application . . . . .	21
7.2.1	Data collection . . . . .	21
7.2.2	Methodology . . . . .	22
7.3	Result . . . . .	22
7.3.1	Inferences . . . . .	23
<b>8</b>	<b>P/B Ratio</b>	<b>24</b>
8.1	Application . . . . .	24
8.1.1	Data collection . . . . .	24
8.1.2	Methodology . . . . .	25
8.1.3	Result . . . . .	25
<b>9</b>	<b>Relationship among Weights, P/E and P/B ratio Clustering</b>	<b>27</b>
9.1	Relationship among Weights, P/E and P/B ratio Clustering . . . . .	27

# Chapter 1

## Introduction

The purpose of the study is to understand market behavior. A better understanding of it will help in making a better decision about selling and buying stocks.

The stock prediction means forecasting the price of a given stock in the future. In the future what is meant can be the next day, month, year, or even second. In the study mainly the prediction is done on the closing price of a stock on a particular day. To break the ice, firstly a reinforcement learning technique is used to predict the stock. The particular reinforcement learning model used is Long Short Term Memory(LSTM) model. The main advantage of the LSTM Model over others is that it consists of Forget Gate. The Forget Gate eliminates outlier from the data. Thus gives a better result. To study Nifty 50 and its component daily Closing Price is observed. It is observed on how many days the Closing Price of the component and Nifty 50 are moving up or down or in opposite direction. The correlation coefficient is calculated to quantify the relationship. Most basic Technical Indicators like Simple Moving Average and Exponential Moving Average is applied to make the prediction. The GDP of a country is an important parameter to study the growth of a country. The study is done to establish a relationship between the GDP of the country and the Nifty 50 stock. The Nifty 50 consists of 50 companies. These 50 companies can be classified under various sectors. A study is done to understand which sector is doing well and which are not doing well at a given time. Further work is on finding these clusters.

# Chapter 2

## Literature Review

People make and lose huge money in the stock market. The idea of predicting the stock market is very attractive. People have been using numerous techniques to predict the market. A lot of research has been done in this field. There are mainly two categories of analysis of the stock market which are popular 1) Technical Analysis and 2) Fundamental Analysis. The technical analysis involves the use of various Technical Indicators like simple Moving Average, Relative Strength Index, Exponential Moving Average, etc. The Fundamental Analysis involves analyzing balance sheets, statement of cash flow, income statement, etc

Jae Won Lee et al [4] in their work introduced reinforcement learning. Numerous investigations for stock price prediction and portfolio management using machine learning have a limitation in that they are based on supervised learning which is not so adequate for learning problems with long-term goals. Reinforcement learning learns from experience. It is used to learn states where state corresponds to the stock price at a given time.

Kae Chin et al [2] in their work used the LSTM. The stock price of China stock market is transformed into 30-days-long sequences with 10 learning features and 3-day earning rate labeling. The model was fitted by training on 900000 sequences and tested using the other 311361 sequences. The improvement in accuracy observed was 14.3 %

Zhifeng Dai et al [5] in their work combined de-noising stock returns by wavelet transform with new proposed technical indicators which have to improve the accuracy of stock return forecasts, in which the new technical indicators can directly reflect the trend of stock return forecast.

Gourav Kumar et al [3] presented an ANN-based approach to forecast the Nifty 50 Index. A feed-forward neural network using a multiple backpropagation algorithm has been used to forecast the next day's OHLC data. This model has used the pre-processed dataset of Open price (O), High price (H),



Low price (L), Close price (C), Volume Traded (V) and Turnover (T) for 10 years from 03 April 2006 to 16 May 2016.

A. Sathish et al [1]. presented an analysis of the stock price volatility of the automobile sector. The outcome summary of the study revealed that there is a significant impact of automobile sector index price movements after the COVID – 19 in India.

# Chapter 3

## Problem Statement

The problem statement is to make a prediction about the stock in future. The work is concentrated on the Nifty 50 and its components. The Nifty 50 sector-wise analysis is done to understand how different sectors behave at a given time. The objective is to understand and improve the existing ways to make stock predictions by optimizing the portfolio and minimizing the risk. This is done by clustering the Nifty 50 components into different groups and further doing technical indicator analysis on it.

### 3.1 Methodology

The various methods used in stock prediction include models like LSTM (Long-Short-Term-Memory) model. The technical indicators used are Simple Moving Average and Exponential Moving Average. Clustering technique like K-mean clustering and Hierarchical clustering is used. P/E and P/B ratios are used to do the clustering. The various sources are used to obtain financial data like Yahoo finance, Investopedia.com, nseindia.com.

# Chapter 4

## Experimental Findings

### 4.1 Stock prediction using LSTM network data

The data taken is of APPLE stock. Two-year data is taken from May 2015 to May 2017. The closing price is taken as the parameter for stock prediction. The total observation is 1258.

#### Training set and Test set

The first 65 % or 817 observation is taken as a training set and the remaining 441 observation is taken as a Test set.

#### Model: LSTM Networks

Long Short Term Memory networks, usually just called LSTM, are a special kind of RNN, capable of learning long-term dependencies. They were introduced by Hockreiter & Schmidhuber(1997) and were refined and popularized by many people in the following work.

LSTMs are explicitly designed to avoid the long-term dependency problem. Remembering information for long periods is practically their default behavior, not something they struggle to learn.

All recurrent neural networks have the form of a chain of repeating modules of a neural network. In standard RNNs. This repeating module will have a very simple structure, such as a single tanh layer.

The LSTM does have the ability to remove or add information to the cell state, carefully regulated by structures called gates.

Gates are a way to optionally let information through. They are composed out of a sigmoid neural net layer and a pointwise multiplication operation.

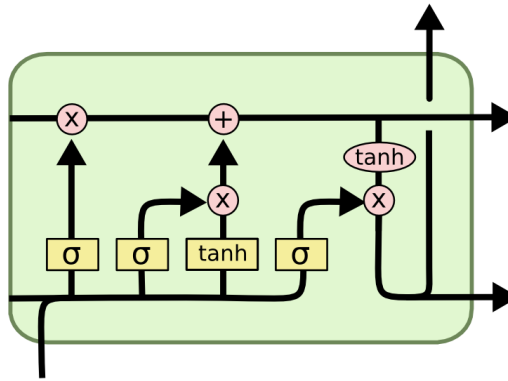


Figure 4.1: Long-Short-Term-Memory Model

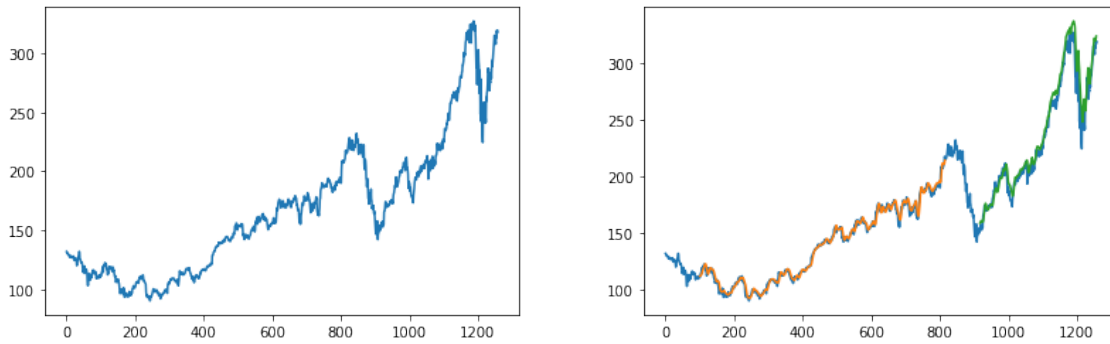


Figure 4.2: Stock Price prediction from LSTM

The sigmoid layer outputs numbers between zero and one, describing how much of each component should be let through. A value of zero means “let nothing through,” while a value of one means “let everything through!”. An LSTM has three of these gates, to protect and control the cell state.

### Data preparation

The window I have chosen is of 100 days. It means it keeps track of the previous 100 days to predict the next day’s stock price.

I have chosen the tensor flow inbuilt function to implement the LSTM model.

### Results and Observations

Diagram 4.2 a) is the actual data we have chosen of AAPL STOCK.

Diagram 4.2 b) diagram orange color curve is the training set and the green color curve is our prediction curve.

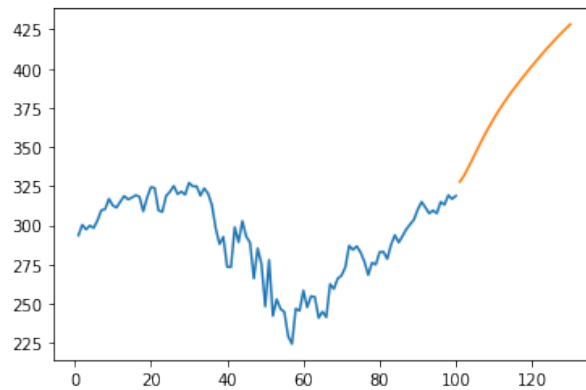


Figure 4.3: LSTM forecast

Diagram 4.3 is obtained when the blue curve is taken to predict the next day's prediction. The orange curve shows the predictions.

## 4.2 Analysis of Nifty index and its component

As an example, we study the temporal evolution of the Nifty index and one of its component stocks Coal India.

**Parameters calculated:** 1)Relation of Coal India with Nifty 50 2)Technical Indicators: i)SMA(Simple Moving Average) ii)RSA(Relative Strength Index) 3)Confusion matrix and Accuracy

Coal India weightage is 0.81% in Nifty 50.

The following things I observed.

1. The confusion matrix shows that nifty 50 and COAL INDIA simultaneously going up and going down 69.5%.
2. The Technical Indicator SMA(Simple Moving Average) predicted correctly 57.5%.
3. The Technical Indicator RSI(Relative Strength Index) predicted correctly 89.5%.

Other observation:

1. Although the Nifty50 and COAL INDIA have a negative correlation they are simultaneously going up and going down 69.5%. This value is high.
2. The RSI technical Indicator predictor can predict correctly 90%. This value is very high

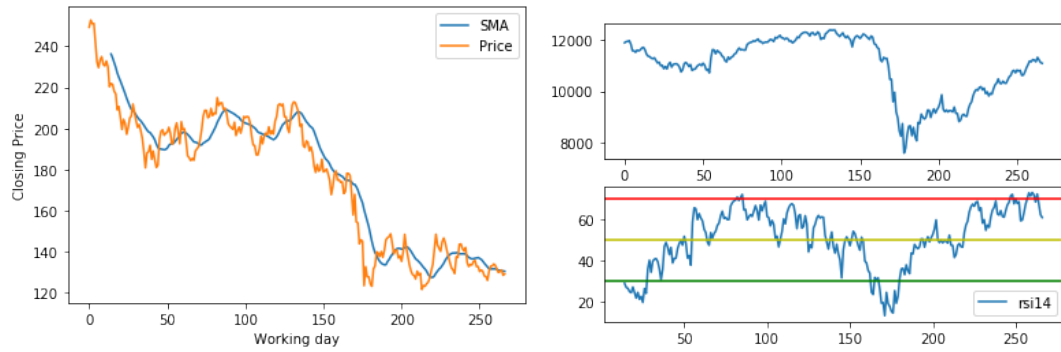


Figure 4.4: Technical Indicator Analysis: SMA and RSA

Diagram 4.4 a. Simple moving average on Coal India stock

Diagram 4.4 b. Relative Strength Index on Coal India Stock

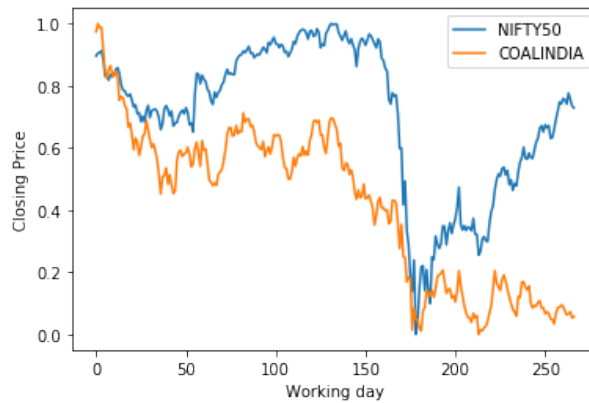


Figure 4.5: Relative movement of Nifty 50 and Coal India

### 4.3 Analysis of GDP and Nifty 50

In this study, the GDP data is collected from <http://mospi.nic.in/data>.

The GDP data is available quarterly. The data is taken in the period 2012-20.

The Nifty 50 closing price is available monthly. The average is taken of the three months to make it consistent with the GDP data.

The Nifty 50 and the GDP data are scaled to 0-1 value.

The graph is plotted between the Nifty 50 and GDP data.

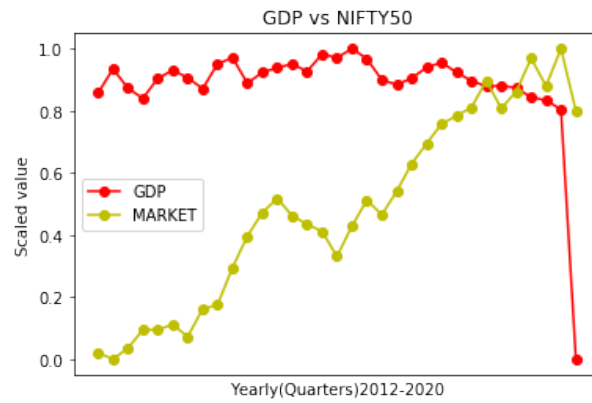


Figure 4.6: Relative movement of Nifty50 and GDP

## RESULT

There are a total of 34 observations. Out of these 17 times, the GDP and Nifty 50 stock go up and down simultaneously and 17 times in opposite directions.

There is no strong correlation that can be established based on the result.

# Chapter 5

## Sectoral Analysis

### 5.1 Sectoral analysis of Nifty 50

Nifty 50 consists of 50 companies. The 50 companies can be classified among different sectors. These sectors include Automobile, Finance, Cement, Oil, etc.

Each Nifty 50 components have weights. The weights of the individual components change with time. For the analysis purpose we have collected monthly data of weights of components of Nifty 50. The data is collected from <https://www.nseindia.com/>. The data is taken in the period from Jan 2020- Dec 2020.

Similarly, the data is collected of components of Nifty50 month-wise. Then components are divided among different sectors. In this study, we group them in 14 sectors.

In each sector, each component's closing price is multiplied by its corresponding weights. This is calculated for each month. All classes are plotted in a single diagram.



## SECTORS

Telecom	Shipping	Fertilizer	Engineering	Cigarettes
BHARTIARTL INFRA TEL	ADANI PORTS	UPL	LT	ITC

Financial Services	Automobile	Consumer Goods	Energy
AXIS BANK INFRA TEL BAJ FINANCE BAJAJ FINSV HDFC BANK HDFC ICICI BANK INDUS INDBK KOTAK BANK SBIN YES BANK	BAJAJ AUTO EICHER MOT HEROMOTOCO MM MARUTI TATAMOTORS	ASIAN PAINTL BRITANNIA HINDUNILVR NESTLE IND GAIL	BPCL GAIL IOC NTPC ONGC POWER GRID RELIANCE

Media	IT	Metal	Pharma	Cement
ZEEL	HCLTEC INFY TCS MM TECHM WIPRO	COALINDIA HINDALCO JSWSTEEL TATASTEEL VEDL	CIPLAL DRREDDY SUNPHARMA	GRASIM ULTRACEMCO

Each sector total value is calculated as:  $\sum w_i x_i$ , where  $w_i$  is the weight and  $x_i$  is the Closing Price

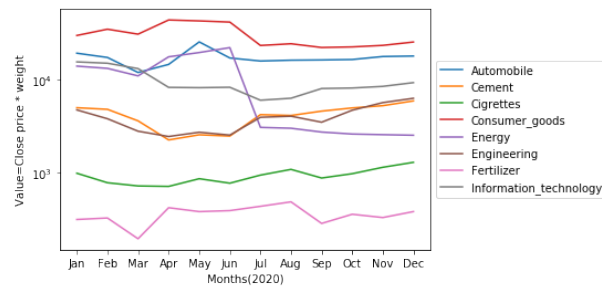


Figure 5.1: Nifty 50 components value

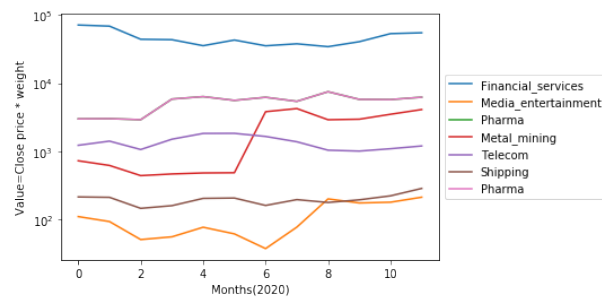


Figure 5.2: Nifty 50 components value

The relative behavior of each sector is calculated as:

The sector value calculated as before divide by the first value of each sector.

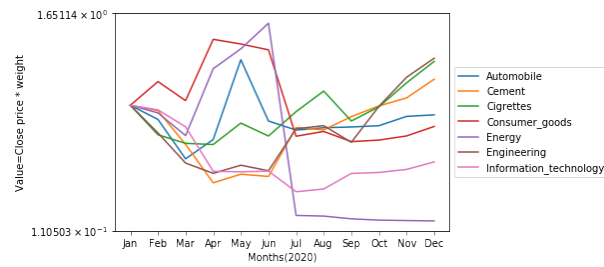


Figure 5.3: Nifty 50 components value

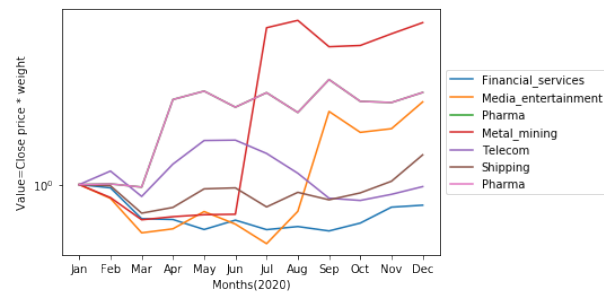


Figure 5.4: Nifty 50 components value

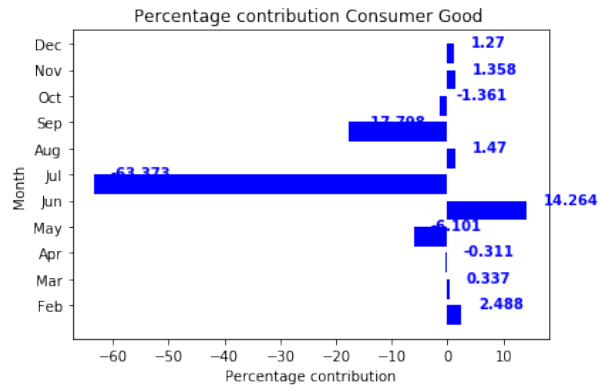


Figure 5.5: Contribution of consumer goods sector in NIFTY50

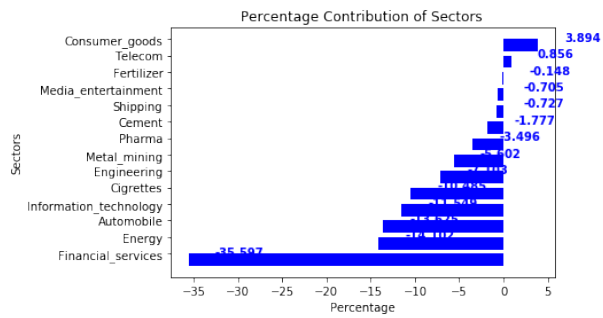


Figure 5.6: Contribution of different sector in Nifty 50(feb2020)

## RESULT

- All the 14 sector values are plotted on the same graph. The comparative study is done. It can be seen at a given time which sector is moving up and which sector is moving down as compare to other sectors.
- The percentage contribution of the different sectors at a given time can be ascertained.
- For a given sector its contribution in Nifty 50 is plotted. In the above diagram, Consumer Good contribution in Nifty 50 is calculated throughout the year. Similarly others sector contribution is analyzed. In a given month contribution of different sectors is analyzed.
- In the above diagram, the contribution of the different sectors in Feb is plotted. Similarly, for other months, the contribution of different sector is analyzed.

# Chapter 6

## Clustering

### 6.1 How Nifty 50 index is calculated

The index of Nifty50 is based on the market capitalisation method.

**Index Value = Index Market Capitalisation / Base Free Float Market Capitalisation of index \* Base Index Value**

where Index Market Capitalisation = Shares outstanding \* IWF \* Capping factor \* Price

IWF= Investible Weight Factor- It is the total number of floating stock which are available for trading.

These stocks are not held by someone who has a strategic interest in the company.

Capping Factor=The factor is used to reduce the Free Float Market Cap of equity in an index. It enables to set all index weight under a set capping threshold.

### 6.2 Weightage of Nifty50 component

The weightage of a component in Nifty50 is the percentage of contribution of individual stock in the overall Nifty50.

### 6.3 k-mean clustering

k-mean clustering is a type of clustering algorithm. It is an iterative algorithm. Its objective is to partition the data points in a predefined k distinct group. It tries to allocate groups to data points such that the sum of the squared distance of data points within a group from their centroid is minimum.

### 6.3.1 k-mean clustering algorithm

1. Assign the number of cluster k.
2. The dataset is shuffled and randomly k data point is chosen for the centroid.
3. Keep iterating till there is no change in datapoints centroid.
4. The sum of squared distance is calculated for each data point to all the centroid.
5. Each data point closest to a centroid is assigned that cluster.
6. The centroid is calculated by taking the average of all distance of data points in that cluster to its centroid.

The objective function for the k mean clustering is given as:

The diagram shows the objective function  $J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2$  with several annotations. An arrow points from the text 'objective function' to the variable  $J$ . An arrow points from 'number of clusters' to the summation index  $k$ . An arrow points from 'number of cases' to the summation index  $n$ . An arrow points from 'case i' to the term  $x_i^{(j)}$ . An arrow points from 'centroid for cluster j' to the term  $c_j$ . A bracket under the distance term  $\|x_i^{(j)} - c_j\|^2$  is labeled 'Distance function'.

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2$$

Figure 6.1

## 6.4 Application

### 6.4.1 Dataset

The Nifty50 dataset is chosen for the year 2020-21.

The dataset consists of the weightage of each stock in Nifty50 from Jan 2020 to Dec 2020. Month wise data is available.

The sources used is: <https://www.nseindia.com/>

### 6.4.2 Methodology

1. The percentage change in weight for each month is taken as a feature.
2. These 11 features are taken and k-mean clustering is applied.

3. The k is chosen as 10
4. The groups are arranged in the descending order of Euclidean distance.

$$dist(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (6.1)$$

$p, q$  = two points in Euclidean n-space

$q_i, p_i$  = Euclidean vectors, starting from the origin of the space (initial point)

$n$  = n-space

5. English alphabets A-H are used. A is given to the least distance and likewise, H is given to the greatest distance from the origin.

## 6.5 Result

Total 10 different groups were obtained:

GROUP A	GROUP B	GROUP C
INDUSINDBK	BAJFINANCE BAJAFINSV	AXISBANK

GROUP D	GROUP E	GROUP F	GROUP G
MM	ASIANPAINT	DRREDDY	ADANI PORTS
HINDALCO	NESTLEIND	RELIANCE	ONGC
TATMOTARS	POWERGRID	SUNPHARMA	NTPC
	HINDUNILVR	CIPLA	IOC
	ULTRACEMCO	BRITANNIA	ITC
	TCS	HCLTECH	BHARTIARTL
			COALINDIA

## 6.6 Observation

- Group A: Indusbank shows great variation over the period. Its value decrease from 1.44 in Feb to .59 in Mar. That may be the reason it is in a separate group.



GROUP H	GROUP I	GROUP J
INFY	UPL	ICICIBANK
TECHM	BAJAJ-AUTO	SBIN
TITAN	MARUTI	BPCL
WIPRO	JSWSTEEL	HDFCBANK
	HEROMOTOCO	GAIL
	GRASIM	KOTAKBANK
	EICHERMOT	TATASTEEL
		LT

- Group B: Bajaj Finance and Bajaj Finserv are in the same group. Both are financial companies and are owned by Bajaj.
- Group C: Axis bank shows large variation. Its value decrease from 3.28 in Jan to 2.17 in Sep.
- Group D: Mahindra and Mahindra and Tata Motors are automobile companies. Both are large vehicle manufacturers. Hindalco is a steel manufacturing firm.
- Group E: Hindustan Unilever and Nestle India are food-related companies. Ultratech cement and Asian Paint can be clubbed together. TCS is an IT firm.
- Group F: Dr Reddy, Sun Pharma, Cipla are pharmaceutical companies. Britannia is a food and beverage company while Hcltech is an IT firm.
- Group G: ONGC, NTPC, IOC, Coal India can be clubbed in the Energy sector. Bharti Airtel is a telecommunication firm. Adani Port is a multiport operator.
- Group H: Infosys, Tech Mahindra, Wipro are IT industry while Titan is a lifestyle company.
- Group I: Bajaj Auto, Maruti, Hero Moto are all automobile companies. JSW Steel is a steel manufacturer. Grasim is a manufacturing company.
- Group J: ICICI Bank, SBI, HDFC Bank, HDFC, Kotak Bank can be clubbed together. BPCL and GAIL are gas companies, Larson Turbo is a technology, engineering, construction, manufacturing, and financial services conglomerate. Tata Steel is a steel manufacturer.

## Outliers

There are two groups which have only one company in them.

- IndusInd Bank
- Axis Bank

The values of IndusInd Bank and Axis Bank shows large variation throughout the period.

The standard deviation observed was .3999 and 0.305 respectively. The value is quite as compared to others in the dataset

# Chapter 7

## P/E Ratio

### 7.1 Definition

It is the ratio of a company's share price to the company's earnings per share. The formula is given by:

$$PE = \frac{\text{Index Market Capitalisation}}{\text{Gross Earnings}}$$

where Index market capitalization of the Index constituents is the sum total of the outstanding equity shares or units considered for index computation multiplied by the close price of each index constituent adjusted for factors such as free-float, capping factor, etc. depending upon the index methodology; and

The earnings (including profits and losses) reported by each index constituent in the trailing 4 quarters (consolidated financials) are cumulated and adjusted for factors such as free-float, capping factor, etc. depending upon the index methodology to arrive at the gross earnings. In case, consolidated financials are not available, standalone financials for trailing 4 quarters will be considered.

### 7.2 Application

#### 7.2.1 Data collection

Nifty 50 component P/E ratio is taken.

The data is collected from the <https://www.valueresearchonline.com/> website.

Data taken is 1 Year median P/E ratio.

The data collected is from Jan 2020 to Dec 2020.

Each month-end day value is chosen (same as that of weightage).

## 7.2.2 Methodology

Each month P/E ratio is taken as separate features

The total features taken is 12.

K-mean clustering is applied by taking each month's values as features.

The total number of clusters chosen is 10.

The clusters obtained are sorted based on cluster center distance from the origin.

The least distance is given symbol 'A' and the largest distance as 'H'.

## 7.3 Result

CLUSTERS OBTAINED

GROUP A	GROUP B	GROUP C	GROUP D
COALINDIA	BAJAJFINSV	ICICI	CIPLA
GAIL	KOTAKBANK		DRREDDY
IOC	UPL		EICHERMOT
JSWSTEEL	ULTRACEMCO		HDFCBANK
NTPC			MARUTI
POWERGRID			TCS
TATASTEEL			

GROUP D	GROUP E	GROUP F
CIPLA	GRASIM	ADANI PORTS
DRREDDY	HDFC	BAJAJ-AUTO
EICHERMOT	ITC	BPCL
HDFCBANK	INDUSINDBK	HCLTECH
MARUTI	LT	HEROMOTOCO
RELIANCE	SBIN	INFY
SUNPHARMA		MM
TCS		TECHM
		WIPRO

GROUP G	GROUP H	GROUP I	GROUP J
HINDALCO	AXISBANK	ASIANPAINT	NESTLEIND
	KOTAKBANK	BRITANNIA	TITAN
	UPL	HINDUNILVR	
	ULTRACEMCO		

### 7.3.1 Inferences

#### Outliers

There are two groups which have only one company in them.

- ICICI
- HINDALCO

The fluctuation of values over the period is very high.

The standard deviation observed is 15.269 and 15.245 respectively which is very high

# Chapter 8

## P/B Ratio

subsectionDefinition Price to book value measures the enterprise value of the company. It is considered to be more stable than the P/E ratio in a volatile market.

$$PB = \frac{\text{Index market capitalization}}{\text{Gross book value or net-worth}}$$

where Index market capitalization of the Index constituents is the sum total of the outstanding equity shares or units considered for index computation multiplied by the last traded price of each index constituent adjusted for factors such as free-float, capping factor, etc. depending upon the index methodology; and

The equity capital and the reserves and surplus (networth) reported by each index constituent in the annual financial report (standalone financials) are cumulated and adjusted for factors such as free-float, capping factor, etc. depending upon the index methodology to arrive at the gross book value.

### 8.1 Application

#### 8.1.1 Data collection

Nifty 50 component P/B ratio is taken.

The data is collected from the <https://www.valueresearchonline.com/> website.

Data taken is 1 Year median P/B ratio.

The data collected is from Jan 2020 to Dec 2020.

Each month-end day value is chosen (same as that of weightage).

### 8.1.2 Methodology

Each month P/B ratio is taken as separate features

The total features taken is 12.

K-mean clustering is applied by taking each month's values as features.

The total number of clusters chosen is 10.

The clusters obtained are sorted based on cluster center distance from the origin.

The least distance is given symbol 'A' and the largest distance as 'H'.

### 8.1.3 Result

GROUP A	GROUP B	GROUP C	GROUP D
GAIL	AXISBANK	ADANI PORTS	BAJAJFINANCE
GRASIM	CIPLA	BAJAJ-AUTO	DRREDDY
HINDALCO	ICICIBANK	COALINDIA	EICHERMOT
IOC	INDUSINDBK	DRREDDY	HDFCBANK
JSWSTELL	LT	HCLTECH	MARUTI
NTPC	MM	HDFCBANK	ITC
ONGC	POWERGRID	HEROMOTOCO	
SBIN	RELIANCE	HDFC	
TATAMOTARS	SUNPHARMA	TECHM	
TATASTEEL	UPL	ULTRACEMCO	
	WIPRO		

GROUP D	GROUP E	GROUP F
BAJAJFINANCE	TCS	ASIANPAINT
EICHERMOT		BPCL
ITC		
INFY		
KOTAKBANK		
MARUTI		

GROUP G	GROUP H	GROUP I	GROUP J
BRITANNIA	BAJAJFINSV	NESTLEIND	HINDUNILVR
TITAN			

## Outliers

There are four groups which have only one company in them.

- TCS
- BAJAJFINSV
- NESTLEIND
- HINDUNILVR

The standard deviation observed of TCS, BAJAJFINSV, NESTLEIND, and HINDUNILVR is .291, 4.654, 17.248, 2.668 respectively.

Leaving first value the rest are the largest value in the dataset



## Chapter 9

# Relationship among Weights, P/E and P/B ratio Clustering

### 9.1 Relationship among Weights, P/E and P/B ratio Clustering

#### Weights and P/E ratio

- DR REDDY, RELIANCE, SUNPHARMA, CIPLA are in the same group. Moreover DR REDDY, SUNPHARMA, CIPLA belong to the same type of sector
- ONGC, NTPC, IOC are in the same group. Moreover, they belong to the same sector.
- TECHM, WIPRO are in the same group. They belong to the same sector.
- BAJAJ-AUTO and HEROMOTOCO are in the same group. Both belong to the same sector.

#### Weights and P/B ratio

- RELIANCE, SUNPHARMA, CIPLA are in the same group. Moreover SUNPHARMA, CIPLA belong to same type of sector

- ONGC, NTPC, IOC are in the same group. Moreover, they belong to the same sector.
- MARUTI and EICHERMOT are in the same group. Both belong to the same sector.
- BAJAJ-AUTO and HEROMOTOCO are in the same group. Both belong to same sector.
- HDFC and HDFCBANK are in the same group.

#### **Weights, P/E and P/B ratio**

- RELIANCE, SUNPHARMA, CIPLA are common in all. Moreover SUNPHARMA, CIPLA belong to the same type of sector
- ONGC, NTPC, IOC are common in all. Moreover, they belong to the same sector.

# Bibliography

- [1] Abdul Rahman Dr. S. Rajamohan A. Sathish. “Impact of COVID - 19 on stock price of NSE in automobile sector”. In: *International Journal of Advanced Multidisciplinary Research* (2020).
- [2] Fangyan Dai Kai Chen Yi Zhou. “A LSTM-based method for stock returns prediction: A case study of China stock market”. In: *IEEE International Conference on Big Data (Big Data)* (2015).
- [3] P Hemanth Kumar and S Basavaraj Patil. “Stock Market Index Forecasting of Nifty 50 Using Machine Learning Techniques with ANN Approach”. In: *International Journal of Modern Computer Science (IJMCS)* (2016).
- [4] Jae Won Lee. “STOCK PRICE PREDICTION USING REINFORCEMENT LEARNING”. In: *IEEE International Symposium on Industrial Electronics Proceedings* (2001).
- [5] Jie Kang Zhifeng Dai Huan Zhu. “New technical indicators and stock returns predictability”. In: *New technical indicators and stock returns predictability*. 2021, pp. 127–142.