**SYNOPSIS**

**Report On**

**VISUALISING AND FORECASTING**

**STOCKS USING PYTHON**

**Submitted By**

**MUDIT RASTOGI**

**(2000290140072 )**

**LOKESH GANGWAR**

**(2000290140062 )**

**PRABHAT CHAUDHARY**

**(2000290140087)**

**Submitted in partial fulfilment of the**

**Requirements for the Degree of**

**MASTER OF COMPUTER APPLICATIONS**

**Under the Supervision of**

**Dr. Amit Kumar Gupta**

**Associate Professor**



**Submitted to**

**Department Of Computer Applications**

**KIET Group of Institutions, Ghaziabad**

**Uttar Pradesh-201206**

**(Jan 2022)**

**ABSTRACT**

The report has been made in fulfilment of the requirement for the

subject: Pattern Recognition and Data Mining in March 2016 under the

supervision of Dr.Ming-Hwa Wang. For making this project we have

studied various concepts related to the stock market and how they can

be used. We also studied about various Machine Learning algorithms

and tools that can be used to solve the problem easily. The project aims

at applying two machine learning algorithms; Decision Trees and

Support Vector Machines and analyze how these algorithms performs at

predicting the stock market.

The prediction of a stock market direction may serve as an early

recommendation system for short-term investors and as an early

financial distress warning system for long-term shareholders.

Forecasting accuracy is the most important factor in selecting any

forecasting methods. Research efforts in improving the accuracy of

forecasting models are increasing since the last decade. The appropriate

stock selections those are suitable for investment is a very difficult task.

The key factor for each investor is to earn maximum profits on their

investments. In this paper Support Vector Machine Algorithm (SVM) is

used. SVM is a very specific type of learning algorithms characterized

by the capacity control of the decision function, the use of the kernel

functions and the scarcity of the solution. In this paper, we investigate

the predictability of financial movement with SVM. To evaluate the

forecasting ability of SVM, we compare its performance with Decision

trees. These methods are applied on 2 years of data retrieved from

Yahoo Finance. The results will be used to analyze the stock prices and

their prediction in depth in future research efforts.

**TABLE OF CONTENTS**

1. Introduction
2. Literature Review
3. Project / Research Objective
4. Research Methodology
5. Project / Research Outcome
6. Proposed Time Duration

**INTRODUCTION**

In the past decades, there is an increasing interest in predicting markets among

Economists, policymakers, academics and market makers. The objective of the

proposed work is to study and improve the supervised learning algorithms to

predict the stock price. Technical Objective The technical objectives will be

implemented in R. The system must be able to access a list of historical prices. It

must calculate the estimated price of stock based on the historical data. It must also

provide an instantaneous visualization of the market index. Experimental Objective

Two versions of prediction system will be implemented; one using Decision trees

and other using Support Vector Machines. The experimental objective will be to

compare the forecasting ability of SVM with Decision Trees. We will test and

evaluate both the systems with same test data to find their prediction accuracy.

1.2

WHAT IS THE PROBLEM? Investors are familiar with the saying, “buy low, sell

high” but this does not provide enough context to make proper investment

decisions. Before an investor invests in any stock, he needs to be aware how the

stock market behaves. Investing in a good stock but at a bad time can have

disastrous results, while investment in a mediocre stock at the right time can bear

profits. Financial investors of today are facing this problem of trading as they do

not properly understand as to which stocks to buy or which stocks to sell in order

to get optimum profits. Predicting long term value of the stock is relatively easy

than predicting on day-to-day basis as the stocks fluctuate rapidly every hour based

on world events.

**Literature Review**

DEFINITION OF THE PROBLEM

Stock market attracts thousands of investors’ hearts from all around the world. The

risk and profit of it has great charm and every investor wants to book profit from

that. People use various methods to predict market volatility, such as K-line

diagram analysis method, Point Data Diagram, Moving Average Convergence

Divergence, even coin tossing, fortune telling, and so on. Now, all the financial

data is stored digitally and is easily accessible. Availability of this huge amount of

financial data in digital media creates appropriate conditions for a data mining

research. The important problem in this area is to make effective use of the available data.

2.2

THEORETICAL BACKGROUND OF THE PROBLEM

Stock market is highly volatile. At the most fundamental level, it is said that supply

and demand in the market determines stock price. But, it does not follow any fixed

pattern and is also affected by a large number of highly varying factors The

investors on the Wall Street are split in two largest factions of adherents; those

who believe the market cannot be predicted and those who believe the market can

be beat.

2.3

RELATED RESEARCH TO SOLVE THE PROBLEM Recently, a lot

of interesting work has been done in the area of applying Machine Learning

Algorithms for analyzing price patterns and predicting stock price. Most stock

traders nowadays depend on Intelligent Trading Systems which help them in

predicting prices based on various situations and conditions. Recent researches

uses input data from various sources and multiple forms. Some systems use

historical stock data, some use financial news articles, some use expert reviews

while some use a hybrid system which takes multiple inputs to predict the market.

Also, a wide range of machine learning algorithms are available that can be used to

design the system. These systems have different approaches to solve the problem.

Some systems perform mathematical analysis on historic data for prediction while

some perform sentiment analysis on financial news articles and expert reviews for

prediction. However, because of the volatility of the stock market, no system has a

perfect or accurate prediction

**Project / Research Objective**

RELATED RESEARCH TO SOLVE THE PROBLEM

Recently, a lot of interesting work has been done in the area of applying Machine

Learning Algorithms for analyzing price patterns and predicting stock price. Most

stock traders nowadays depend on Intelligent Trading Systems which help them in

predicting prices based on various situations and conditions. Recent researches

uses input data from various sources and multiple forms. Some systems use

historical stock data, some use financial news articles, some use expert reviews

while some use a hybrid system which takes multiple inputs to predict the market.

Also, a wide range of machine learning algorithms are available that can be used to

design the system. These systems have different approaches to solve the problem.

Some systems perform mathematical analysis on historic data for prediction while

some perform sentiment analysis on financial news articles and expert reviews for

prediction. However, because of the volatility of the stock market, no system has a

perfect or accurate prediction.

OUR SOLUTION TO SOLVE THIS PROBLEM

We will implement the system using two different machine learning techniques.

One using Support Vector Machines and the second implementation using

Decision Trees. We will train both the systems using 75% of 2 years of historic

data and then test our model to check which systems yields better output using

the remaining 25% of historic data.

**Research Methodology**

HOW TO COLLECT INPUT DATA?

Input data is taken from Yahoo Finance using following steps: 1. For our project,

we are considering S&P 500 Companies. The list of companies in S&P 500 can be

obtained from Wikipedia [3]. 2. Use stock’s ticker symbol from step a to get data

from Yahoo Finance. 3. System will take last 2 years’ stock data of the company

using quantmod package in R. 4. Further we divide the data into two parts,

training data and testing data, where 75% of the data will be used for training and

25% of the data will be used for testing.

HOW TO SOLVE THE PROBLEM?

We will solve the problem using below supervised learning techniques to build our

model - • Decision Trees with Technical Indicators. • Support Vector Machine

with Technical Indicators. To solve the problem, we will follow below steps - 1.

Fetch the data of a stock from Yahoo Finance of last 2 years. 2. Calculate the

values of technical indicators RSI, EMA, MACD, SMI, etc. 3. Train the model

using these indicators and training data. 4. Test the model using testing data. 5.

Evaluate our system using various evaluation techniques

**Project / Research Outcome**

In this paper, we study the use of decision trees and support vector machines to

predict financial movement direction. Of both these algorithms, we saw that

Support Vector Machine gave us better results. SVM is a promising type of tool for

financial forecasting. SVM is superior to the other individual classification

methods in forecasting daily movement direction. This is a clear message for

financial forecasters and traders, which can lead to a capital gain. However, each

method has its own strengths and weaknesses. In this model, the principal

components identified by the SVM are used along with internal and external

financial factors in SVM for forecasting. We also observed that the choice of the

indicator function can dramatically improve/reduce the accuracy of the prediction

system. Also a particular Machine Learning Algorithm might be better suited to a

particular type of stock, say Technology Stocks, whereas the same algorithm might

give lower accuracies while predicting some other types of Stocks, say Energy

Stocks.

7.2

RECOMMENDATIONS FOR FUTURE STUDIES

Two algorithms,

decision trees and SVM were used in this study and only one dataset from Yahoo

Finance was applied to train and test the models. The system can only predict the

direction(up/down) for the next trading day. In the future, SVM regression model

will be used in order to predict the price moment for the future and the results .