AIDI 11840 Knowledge and Expert Systems

Best Stock to Invest In

Project Report

Dheeraj Reddy – 100811270

Pooja Vasan – 100799175

Purnima Yadav – 100800499

Vasundara – 100800734

7th April 2021

Uzair Ahmad

Table of Contents

Abstract

Prediction of stock market aims to determine the future movement of the stock value of a financial exchange. Predicting how the stock market will move is one of the most challenging issues due to many factors that involved in the stock prediction, such as interest rates, politics, and economic growth that make the stock market volatile and very hard to predict accurately.

Nowadays, more and more valuable market information like financial news, earning reports are publicly available online. This brings the importance of text mining strategies to extract significant information to analyze market behavior.

In this project, we have used technologies like NLP, we extracted meaningful information from unstructured texts such as financial news, developed an expert system that uses its knowledge base and inferencing techniques and indicators such as MACD and Stochastic to predict the best time to invest in Stocks.

Introduction

The Covid-19 pandemic has been one of the most burdensome times for all of us. That could be in terms of jobs, health, losing loved ones, studies, and it plays a major role in 'money'. There are cases where people are not able to get their daily resources. This is an opportunity to invest in some stocks where you get promising returns.

Prediction of stocks is a risky task. Why do we even need a prediction here? Of course, to buy or to sell. There are chances of huge losses in the worst case. When it is done by a human without having knowledge on the previous data, there is at least 99 percent certainty of having loss. This is when an automated system comes to the rescue. Scrape the latest news, load the data and just relax! Yes! An instance where you can absolutely trust the automated system to buy or sell your stock according to the updated market without losing a single dollar.

The next question that comes to mind is that- how an automated system can perform this task? The answer is- by using an Expert System, which as patterns and a set of rules that could predict future stock price movements. Defined patterns are classified into five groups with respect to their meanings: falling, rising, neutral, trend-continuation and trend-reversal patterns. There is a research which proves that the prediction of stock market using Expert System – fuzzy logic, yields an accuracy of *98.3%* for the Istanbul Stock Exchange.

Why is ES needed to predict/ buy the best available stocks?

Expert systems can explain market status, interprets patterns and can be helpful to investors' decision making. It helps to distribute the expertise of a human. One expert system may contain knowledge from more than one human expert thus making the solutions more efficient.

There are many firms in the market which uses automatic stock control and ordering techniques as a major part of their inventory control systems. Generally, there are huge chances of human making mistakes in prediction or control stock. Whereas, using an expert system, it makes much easier to diagnose anomalies in the data provided.

There are certain basic questions that could be asked by the expert system, like:

- Does the investor mind invest in risky companies?
- Does the investor have specific field/ area of interest to invest in?
- Are there any companies that the investor wants to avoid?

Expert System Indicators – Trading Strategy
Moving Average Convergence Divergence (MACD) indicator

It is one of the simplest and most effective indicators. As its name indicates, it is all about convergence and divergence of the two moving averages. As MACD is unbounded, traders can look for signal line crossovers, centerline crossovers and divergences to generate signals.

Though values can be substituted depending upon trading style, 12,26 and 9 are the typical setting used with MACD.

1. *Step 1:* Calculate the MACD line:
   a. Calculate the 26-day exponentially weighted moving average of the price. This is the long-term line.
   b. Calculate the 12-day exponentially weighted moving average of the price. This is the short-term line.
   c. Calculate the difference between the 26-day EMA and 12-day EMA lines. This is the MACD line.
2. *Step 2:* Calculate the Signal line from the MACD line:
   a. Calculate the 9 days exponentially weighted moving average of the MACD line. This is known as the signal lines.
3. *Step 3:* Compute the histogram: Distance between MACD and the Signal.
   a. We can then calculate the difference between the MACD and the Signal line and then plot it as a histogram. The histogram can help us find when the cross-over is about to happen.

Fig. 1 MACD positive line

In the above picture, the MACD has a positive value (blue line) whenever the 12-period EMA (red line) is above the 26-period EMA and a negative value when the 12-period EMA is below the 26-period EMA. The more distant the MACD is above or below its baseline indicates that the distance between the two EMAs is growing.

EMA refers to exponential moving average

The formula for MACD is:

MACD = 12-Period EMA - 26-Period EMA

Stochastic Indicator

This indicator measures the relationship between an issue's closing price and its price range over a predetermined period. It determines whether a stock has moved into overbought or oversold position.

The main parts of **Stochastic Indicator:**

1. **Price Action-** it indicates the range of closing according to the trends. In an upward trending market, its closing price tends to be at the high end of the range of price list at which the stock traded over that day. In down trending markets, it tends to be at or near the session's low.
2. **Relative Strength Index:** it is a stochastic oscillator that has a predetermined boundary on high and low trends. The range in RSI includes 0 to 100, 20 to 80 and 30 to 70.

Formula

Stochastic Indicator is measured with the K line and a D line. We follow D line closely because it determines the major signals in the chart.

$$\%K = 100 \times \frac{CP - L14}{H14 - L14}$$

To calculate the slow stochastic, replace "n" with the range you are monitoring which is number of past data days. The slow stochastic can be calculated on any time frame, although the default value is 14, as mentioned above.

Once using the formula above, the trader should then take the three period Simple Moving Average (SMA) of this value in order to calculate the %K result for the Stochastic Slow Strategy.

%D: Slow %D is equal to the three period SMA of Slow %K that the trader calculates in Step 2.

Reading the Chart

The K line is faster than the D line; the D line is the slower of the two. The investor needs to watch as the D line and the price of the issue begin to change and move into either the overbought (over the 80 line) or the oversold (under the 20 line) positions. The investor needs to consider selling the stock when the indicator moves above the 80 levels.

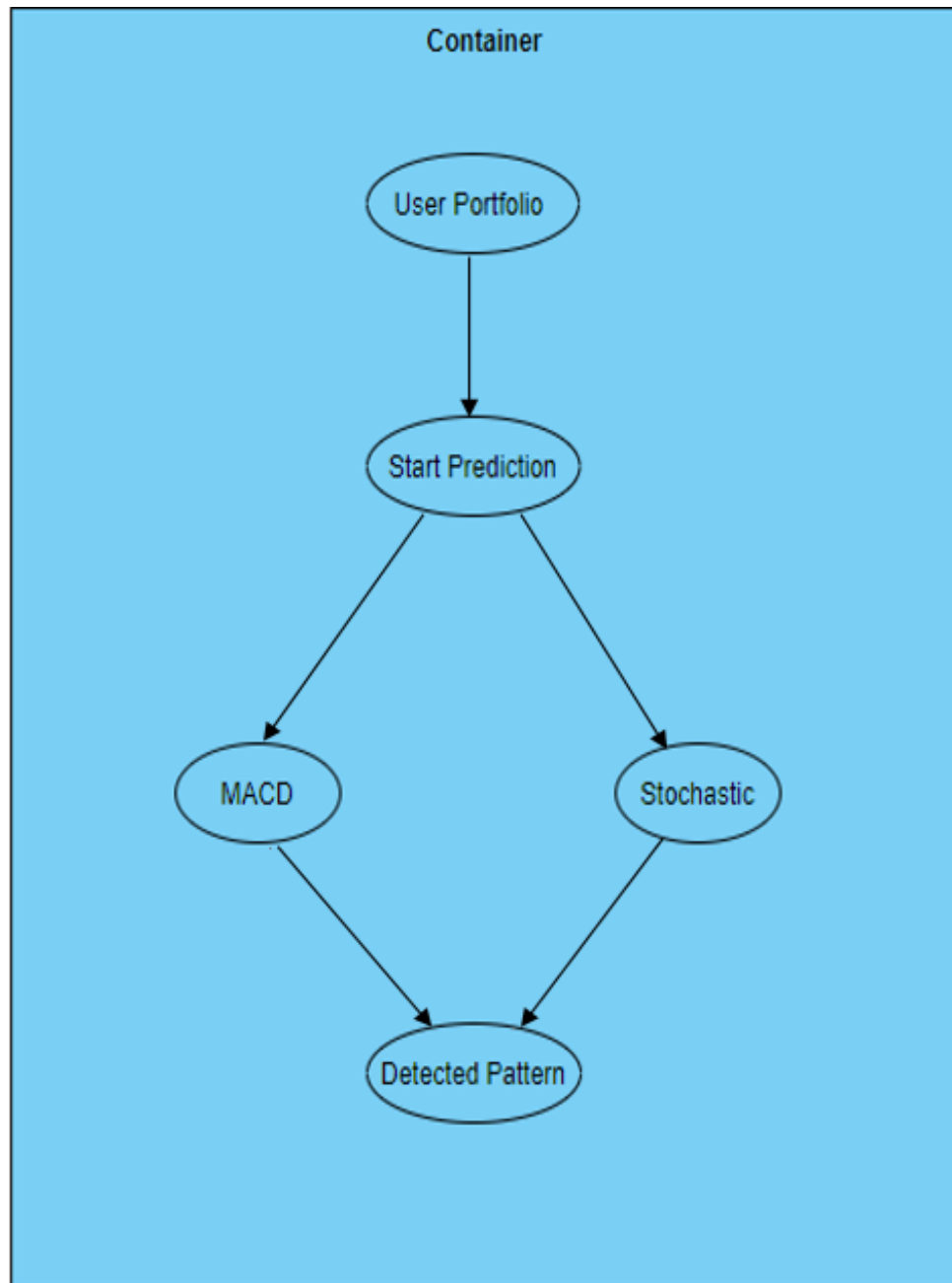*Fig. 2 Reading K and D lines of Stochastic Indicator*

Tools and techniques

- Python- We used Python3 for this project. Some libraries that were used are: Pandas and Numpy.
- Data:
    - Since it is the stock data that we are working on, it must be live which keeps updating every 5minutes at least. We came across several stock market websites like The Motley Fool, Yahoo! Finance, MetaStock, Morningstar, Bloomberg.
    - We realized that scraping data from *Yahoo! Finance* is going to be easier because of the information that the website provides. They provide straight financial news which is easier to scrape using Python and Beautiful Soup.
    - Though we say easier, we came across several challenges while scraping data from web. The web usually consists of different technologies, styles and personalities and it keeps growing every day. One of the main challenges was the *variety,* every other website is different in its own way. When it comes to Yahoo!, we found too many partitions and was difficult to fetch the exact information that we wanted to fetch. Second challenge is *durability*. Websites keep changing. There is a high probability that the code would work perfectly fine to fetch information today and might fail tomorrow due to changes that happens constantly. So, there is a need to update our scraper.
- Web Scraping:
    - Some of the websites do give us an option to download CSV data format or could be accessible via API but we scrape data when these convenient options are not available.
    - As the team was familiar with BeautifulSoup rather than Selenium, we went with the same to scrape and parse data from the website. BeautifulSoup is a Python library for parsing structured data. It allows us to interact with HTML in a similar way to how we would interact with a web page using developer tools.
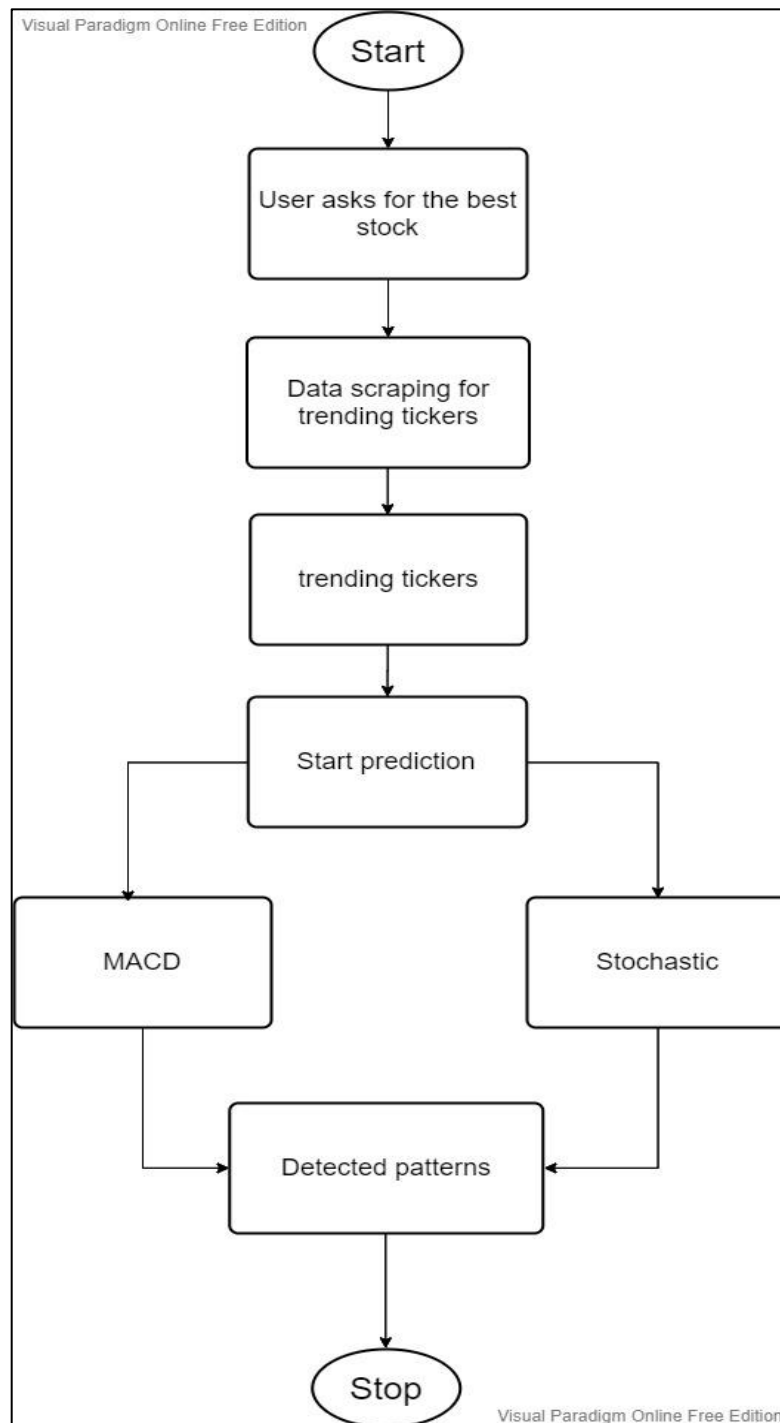
Methodology

Method-1

Method-2

```
              ┌───────────┐
              │   Start   │
              └─────┬─────┘
                    │
                    ▼
        ┌───────────────────────┐
        │ User asks for the best│
        │        stock          │
        └───────────┬───────────┘
                    │
                    ▼
        ┌───────────────────────┐
        │   Data scraping for   │
        │   trending tickers    │
        └───────────┬───────────┘
                    │
                    ▼
        ┌───────────────────────┐
        │   trending tickers    │
        └───────────┬───────────┘
                    │
                    ▼
        ┌───────────────────────┐
        │   Start prediction    │
        └───────────────────────┘
          │                   │
          ▼                   ▼
    ┌──────────┐        ┌──────────────┐
    │   MACD   │        │  Stochastic  │
    └────┬─────┘        └──────┬───────┘
         │                     │
         ▼                     ▼
        ┌───────────────────────┐
        │   Detected patterns   │
        └───────────┬───────────┘
                    │
                    ▼
              ┌───────────┐
              │   Stop    │
              └───────────┘
```

Features of the Model

Method-1:

1. User will be asked to input the shares he wants to watch i.e. the personal portfolio.
2. The expert system will grab the data for each stock in the portfolio using yahoo finance API
3. Then MACD and Stochastic signal indicators will be calculated.
4. User will be presented with the different indicator suggestions and finally can decide on the portfolio management.
5. User has the option to view the graph for each ticker separately if he wants to do further research.

Method-2:

1. User asks the system to suggest the stocks to look for today
2. The expert systems scrape the data from PR news website to find the trending tickers
3. For each ticker both MACD and Stochastic indicator will be calculated.
4. User will be presented with the trending tickers and suggestions based on both the experts.
5. Finally, user has the option to view the graph for each ticker separately if he wants to do further research.

Summary and Conclusion

Predicting the stock market has been a long-standing problem and there have been many attempts to predict the future price of a share based on different techniques. With the growth of technology and infrastructure machine learning based stock predictions have also become popular in the recent past. However, still there are some traditional calculations-based signal predictions to buy or sell a stock. These methods have been proven to be working for majority of the popular stocks when the volatility is low.

So, a rule-based stock buy or sell indicator is designed using two main criteria, Moving average Convergence Divergence (MACD) and a Stochastic Oscillator. In this developed system we have Two options,

1.   Suggestions on the existing user portfolio

2.   Suggesting trending stocks and the expert suggestions on them.

PR news website is used as a main source for data scraping as it is considered to be accurate and keeps track of all the press releases on different stocks listed in NYSE and NASDAQ.

This is our attempt to suggest the user about best stocks based on the MACD and Stochastic Oscillator experts.

References

- Using an Expert System to Monitor an Automatic Stock Control System – J.C. Thorpe

  https://link.springer.com/chapter/10.1007/978-1-349-12362-9_10

  Accessed Apr 6, 2021.

- Investment Analysis https://ahadexsys.weebly.com/investment-analysis.html#:~:text=A%20number%20of%20major%20investors,in%20the%20knowledge%20base%20suggest

- MACD MACD (Moving Average Convergence/Divergence Oscillator) [ChartSchool] (stockcharts.com)

- Top 8 best stock market sites to check investments: RapidAPI staff

  https://rapidapi.com/blog/best-stock-market-websites/

  Published: Mar 16, 2021. Accessed: Apr 6, 2021.

- Beautiful Soup: Build a web scraper with python: Martin Breuss

  https://realpython.com/beautiful-soup-web-scraper-python/

- Expert System for predicting stock market: K.H.Lee

  https://www.sciencedirect.com/science/article/abs/pii/S0957417499000111#:~:text=The%20expert%20system%20has%20patterns,continuation%20and%20trend%2Dreversal%20patterns.

- Sailing through the Covid-19 crisis by using AI for financial market predictions: Manjit Kaur

  https://www.hindawi.com/journals/mpe/2020/1479507/

  Published Dec 30, 2020. Accessed Apr 6, 2021.

- Flowchart and use case diagrams are made using https://online.visual-paradigm.com/

- Advances in Stock Market https://www.intechopen.com/online-first/recent-advances-in-stock-market-prediction-using-text-mining-a-survey