

Sudhanshu Kumar- 82

Bhargav Guirish- 81

Siddharth Malang- 69

Specification of title & objective of project

-> Financial market analysis using machine learning in python

Introduction of Project

-> Stock market prediction with data mining techniques is one of the most important issues to be investigated. We intend to present a system that predicts the changes of stock trend by analyzing the influence of news articles.

Literature Survey

-> Machine learning has had fruitful applications in finance well before the advent of mobile banking apps, proficient chat bots, or search engines. Given high volume, accurate historical records, and quantitative nature of the finance world, few industries are better suited for artificial intelligence. There are more uses cases of machine learning in finance than ever before, a trend perpetuated by more accessible computing power and more accessible machine learning tools (such as Google's Tensorflow). Today, machine learning has come to play an integral role in many phases of the financial ecosystem, from approving loans, to managing assets, to assessing risks. Yet, few technically-savvy professionals have an accurate view of just how many ways machine learning finds its way into their daily financial lives.

->When predicting the future prices of Stock Market securities, there are several theories available. The first is Efficient Market Hypothesis (EMH) (Fama 1964). In EMH, it is assumed that the price of a security reflects all of the information available and that everyone has some degree of access to the information. Fama's theory further breaks EMH into three forms: Weak, Semi-Strong, and Strong. In Weak EMH, only historical information is embedded in the current price. The Semi-Strong form goes a step further By incorporating all historical and currently public information into the price. The Strong form includes historical, public, and private information, such as insider information, in the share price. From the tenets of EMH, it is believed that the market reacts instantaneously to any given news and that it is impossible to consistently outperform the market. A different perspective on prediction comes from Random Walk Theory (Malkiel 1973). In this theory, Stock Market prediction is

believed to be impossible where prices are determined randomly and outperforming the market is infeasible. Random Walk Theory has similar theoretical underpinnings to Semi-Strong EMH where all public information is assumed to be available to everyone. However, Random Walk Theory declares that even with such information, future prediction is ineffective

It is from these theories that two distinct trading philosophies emerged; the fundamentalists and the technicians. In a fundamentalist trading philosophy, the price of a security can be determined through the nuts and bolts of financial numbers. These numbers are derived from the overall economy, the particular industry's sector, or most typically, from the comp

any itself. Figures such as inflation, joblessness, industry return on equity (ROE), debt levels, and individual Price to Earnings (PE) ratios can all play apart in determining the price of a stock. In contrast, technical analysis depends on historic

al and time-series data. These strategists believe that market timing is critical and opportunities can be found through the careful averaging of historical price and volume movements and comparing them against current prices. Technicians also believe that there

are certain high/low psychological price barriers such as support and resistance levels where opportunities may exist. They further reason that price movements are not totally random, however, technical analysis is considered to be more of an art form rather than a science and is subject to interpretation. Both fundamentalists and technicians have developed certain techniques to predict prices from financial news articles. In one model that tested the trading philosophies; LeBaron et. al. posited that much can be learned from a simulated stock market with simulated traders (LeBaron, Arthur et al. 1999). In their w

ork, simulated traders mimicked human trading activity.

Because of their artificial nature, the decisions made by these simulated traders can be dissected to identify key nuggets of information that would otherwise be difficult to obtain. The simulated traders were programmed to follow a rule hierarchy when responding to changes in the market; in this case it was the introduction of relevant news articles and/or numeric data updates. Each simulated trader was then varied on the timing between the point of receiving the information and reacting to it. The results were startling and found that the length of reaction time dictated a preference of trading philosophy. Simulated traders that acted quickly formed technical strategies, while traders that possessed a longer waiting period formed fundamental strategies (LeBaron, Arthur et al. 1999). It is believed that the technicians capitalized on the time lag by acting on information before the rest of the traders, which lent this research to support a weak ability to forecast the market for a brief period of time.

Dataset Choice

collection was one of the biggest tasks which we had to face, we first look out for BSE, but it gives only opening & closing price. We scanned through almost all the indices out there in the world to see if anyone has archive for the intra-day variations. These are all the indices which we tried out for <https://www.nseindia.com/products/content/equities/equities/> National Stock Exchange, Bombay stock exchanges.

References

<https://www.cse.iitb.ac.in/~saketh/research/ppBTP2010.pdf>

<https://www.techemergence.com/machine-learning-in-finance/>

<https://igniteoutsourcing.com/publications/machine-learning-in-finance/>

<https://thefinancialbrand.com/71350/ai-machine-learning-analytics-marketing-banking-trends/>

<https://www.quora.com/How-is-machine-learning-used-in-finance>