**Trimester: FIFTH TRIMESTER**

**Name(s): ANAND MOHAN  
SRN: R19MBA53**

|  |  |
| --- | --- |
| Project Name: | Directional Analytics forDay Trading in Stock Market |
| Background Information | Algorithmic Trading gives benefits like reduced expenses, reduced latency, and no dependence on sentiments. However, it brings up challenges for retail investors as they do not have the desired technology to create such systems. With new algorithms continuing to flood the markets every day, comparison of the effectiveness and accuracy of these algorithms pose nonetheless an added challenge.  Any one or two associated formulas or techniques may go fine on back testing in controlled environments, but the main challenge is live testing, as a result of many things like price variations, quiet news, and existing noise.  An oversized inventory of stock prediction techniques has been developed over the years, though the consistency of the particular prediction performance of most of those techniques remains debatable.  The requirement is to beat the deficiencies of Fundamental and technical analysis, and also the evident advancement within the modeling techniques has driven numerous researchers to review new strategies for stock value prediction. |
| Statement of the Problem | Investors are looking at algorithmic trading as an option to reduce volatility. Fundamental analysis is being used for evaluating a share's intrinsic value for long-term investment opportunities. Technical analysis on the other hand assists the traders to evaluate trends in the stock's price, momentum, and volume from a statistical perspective.  However, the consistency of the prediction performance of most of these techniques remains debatable and the volatility of the market is still unpredictable.  Therefore, it is the constant endeavor of investors to find better, easy, and simple Modelling techniques for forecasting any share’s price for day trading in the stock market.  Such a process should also evaluate the degree of risks concerned and minimize the chances of loss with the highest possible accuracy. |
| Proposed Solution | The objective of the project is to build various classification models to predict the direction of the closing price for the stock under consideration. The direction of the closing price is to be detected to determine the extent of accuracy of Modelling predictions.  6-day consecutive closing price for the stock under consideration is being taken. These 6 days consecutive closing prices will be tabulated week on week for the entire dataset and will be utilized as 6 different feature variables for building the classification Model.  Computation is being done to evaluate whether it is positive change, negative change or no change between 7th and 8th day closing price. The rule is being set to determine as to what has to be seen as direction change.0.7% change,1% change and 1.5% change -these are different classes of direction for which rule is being set which  is to be followed for computing the direction change as either positive change, negative change or no change.  Therefore, given 6-day data it will be predicted whether on 8th data the closing price of stock under consideration is going to increase or decrease or remain the same. Based on the close price, the direction of the next day closing price is to be predicted as to whether it is going to increase or decrease. A number of target variables can be created based on whether the change is on 0.7%,1% or 1.5% and then it is to be determined for each of these target variables what would be the prediction accuracy.  once it is determined say for example 0.7% change has the best prediction accuracy among all different classes of direction namely 0.7% change,1% change and 1.5% change then the range of consecutive days to be utilized as feature variable is increased to 10 days and 14 days consecutively. These 10 days and 14 days consecutive closing prices will be tabulated week on week for the entire dataset and will be utilized as different feature variables for building the classification Model.  Similarly, all technical indicators can be utilized in Technical Analysis to build another sets of classification Models. All different types of technical indicators namely momentum indicators, trend indicators, volatility indicators, volume indicators can be utilized as feature variables based on the input dataset and different classification models can be built to determine their prediction accuracy. Generally Open price, High price, low price, close price and volume for the stock under consideration will be utilized to derive feature variables from technical indicators. These derived feature variables will then be used as the feature variables to predict the direction of the close price.  If for example say 10000 is invested in HDFC stock, and say it is predicted as positive change for the next day. The same prediction process is repeated for say 100 times and evaluated how much is the net gain and loss based on that. |
| Detailed Scope of Work: | Daily Trading Data of HDFC company from the year 2000 to 2022 is being used for this study. This study uses NSE Data.  6-day consecutive closing price for the stock under consideration is being taken. These 6 days consecutive closing prices will be tabulated week on week for the entire dataset and will be utilized as 6 different feature variables for building the classification Model.  The data is being prepared week on week to determine how exactly computation is being done for what is up, what is down and what is neutral. Say for example, anything more than 0.7% change can be positive up, anything less than -0.7% change can be positive down, anything between 0.7% and -0.7% change can be taken as Neutral.  The difference between 7th and 8th day Closing price is determined. If the 8th day closing price is seen an increase from the 7th day by 0.7% or more, the direction of the closing price can be made as positive.  If the 8th day closing price is seen a decrease from the 7th day by -0.7% or less, the direction of the closing price can be made as negative. Between -0.7% and 0.7% that the direction of the closing price for the stock under consideration can be treated as sideways.  For data within the 0.7% and -0.7% band, usually the advice to the investor will be to hold on to existing portfolios and wait for the direction of the closing price to show as either negative or positive change. If there is a negative change, usually the advice to the investor will be to not to invest in such a circumstance. If there is a positive change the investor will be suggested to invest.  It is to be determined how many times the positive changes are identified by predicting and how many times positive changes are there in the actual data. This will be utilized to evaluate how many times true positives were detected and how many times the false positives were predicted in the prediction. Similar process to be followed for detecting true negatives and false negatives. Similar process to be followed for detecting true neutrals and false neutrals. Based on prediction accuracy, it can be suggested whether to invest or not to invest to the prospective investor.  Computation is being done to evaluate whether it is positive change, negative change or no change between 7th and 8th day closing price. The rule is being set to determine as to what has to be seen as direction change.0.7% change,1% change and 1.5% change -these are different classes of direction for which rule is being set which is to be followed for computing the direction change as either positive change, negative change or no change.  Therefore, given 6-day data it will be predicted whether on 8th data the closing price of stock under consideration is going to increase or decrease or remain the same. Based on the close price, the direction of the next day closing price is to be predicted as to whether it is going to increase or decrease. A number of target variables can be created based on whether the change is on 0.7%,1% or 1.5% and then it is to be determined for each of these target variables what would be the prediction accuracy.  It will be identified regarding the extent of accuracy by which positive, negative or neutral changes can be predicted based on 0.2 of the existing test data. Based on whatever is the prediction, the prediction accuracy is determined.  once it is determined say for example 0.7% change has the best prediction accuracy among all different classes of direction namely 0.7% change,1% change and 1.5% change then the range of consecutive days to be utilized as feature variable is increased to 10 days. Therefore,10-day consecutive closing price for the stock under consideration is being taken. These 10 days consecutive closing prices will be tabulated week on week for the entire dataset and will be utilized as different feature variables for building the classification Model.  The difference between 11th and 12th day Closing price is determined. If the 12th day closing price is seen an increase from the 11th day by 0.7% or more, the direction of the closing price can be made as positive. If the 12th day closing price is seen a decrease from the 11th day by -0.7% or less, the direction of the closing price can be made as negative. Between -0.7% and 0.7%, the direction of the closing price for the stock under consideration can be treated as sideways. The prediction accuracy is determined to confirm that say 0.7% change has the best prediction accuracy among all different classes of direction even when range of consecutive days to be utilized as feature variable is increased to 10 days.  Similar process is again repeated for range of consecutive days to be utilized as feature variable increased to 14 days. The prediction accuracy is determined to confirm that say 0.7% change has the best prediction accuracy among all different classes of direction even when range of consecutive days to be utilized as feature variable is increased to 14 days.  Similarly, all technical indicators can be utilized in Technical Analysis to build another sets of classification Models. All different types of technical indicators namely momentum indicators, trend indicators, volatility indicators, volume indicators can be utilized as feature variables based on the input dataset and different classification models can be built to determine their prediction accuracy. Generally Open price, High price, low price, close price and volume for the stock under consideration will be utilized to derive feature variables from technical indicators. These derived feature variables will then be used as the feature variables to predict the direction of the close price. Four different Classification models based on four different types of technical indicators are being built.  For momentum indicators, Awesome Oscillator Indicator, KAMA Indicator, Percentage Price Oscillator, Percentage Volume Oscillator, ROC Indicator, RSI Indicator, Stochastic Oscillator, TSI Indicator, Ultimate Oscillator, WilliamsR Indicator are being utilized as the feature variables to predict the direction of the closing price and determine the prediction accuracy.  For trend indicators, ADX Indicator, Aroon Indicator, CCI Indicator, Ichimoku Indicator, KST Indicator, MACD, PSAR Indicator, EMA Indicator, WMA Indicator, Vortex Indicator are being utilized as the feature variables to predict the direction of the closing price and determine the prediction accuracy.  for volatility indicators, Average True Range, Bollinger Bands, Donchian Channel, Keltner Channel, Ulcer Index are being used as feature variables. Lower and upper band of these volatility indicators are also utilized as feature variables and the direction of the closing price is predicted to determine what is the prediction accuracy.  for volume indicators, AccDistIndex Indicator, ChaikinMoneyFlow Indicator, EaseOfMovement Indicator, ForceIndex Indicator, MFI Indicator, OnBalanceVolume Indicator, VolumePriceTrend Indicator, VolumeWeightedAveragePrice, NegativeVolumeIndex Indicator, DailyLogReturn Indicator are used as feature variables and the direction of the closing price is predicted as to whether it is positive change, Negative change or Neutral to determine what is the prediction accuracy. **Go Long Direction Prediction using Technical Indicators:** The direction of the close price is estimated as percentage change of the close price between upper-band +0.5% and lower band -0.5%-if the percentage change of the closing price is more than 0.5%, the direction of the closing price is treated as positive and suitable for long Trading in stock market. Otherwise, the direction of the close price is treated as non-positive and not suitable for long Trading in stock market. **Go Short Direction Prediction using Technical Indicators:** The direction of the close price is estimated as percentage change of the close price between upper-band +0.5% and lower band -0.5%-if the percentage change of the closing price is less than -0.5%, the direction of the closing price is treated as Negative and suitable for Short Trading in stock market. Otherwise, the direction of the close price is treated as non-negative and not suitable for Short Trading in stock market.  When the majority of the 15 various models or all of them move in the same direction, a choice on whether to invest or not to invest on the stock under consideration must be made. What works in the Indian stock market must be proven with evidence.  The entire process is needed to be tried and tested for a different dataset altogether to ensure that Any stock on the stock market can utilise the same procedure to forecast whether to invest or not to invest, which is helpful.  Daily Trading Data of SBI and Kotak Mahindra company from the year 2000 to 2022 is being used to repeat the entire process which had been implemented for the HDFC dataset. |
| Support needed from Program office | Program Office |
| References | Aaron Patrick. (2020). *HDFC Bank Fundamental Analysis and Future Outlook*. https://billiondollarvaluation.com/hdfc-bank-fundamental-analysis-and-future-outlook/  Alhomadi, A. (2021). Forecasting stock market prices : A machine learning approach. *Digital Commons*, *11*(2), 16–36.  Biswas, M., Nova, A. J., Mahbub, M. K., Chaki, S., Ahmed, S., & Islam, M. A. (2021). Stock Market Prediction: A Survey and Evaluation. *2021 International Conference on Science and Contemporary Technologies, ICSCT 2021*, *December*. https://doi.org/10.1109/ICSCT53883.2021.9642681  Cornellius Yudha Wijaya. (2021). *CRISP-DM Methodology For Your First Data Science Project*. https://towardsdatascience.com/crisp-dm-methodology-for-your-first-data-science-project-769f35e0346c  Huang, Y., Capretz, L. F., & Ho, D. (2021). Machine Learning for Stock Prediction Based on Fundamental Analysis. *2021 IEEE Symposium Series on Computational Intelligence, SSCI 2021 - Proceedings*, *5*. https://doi.org/10.1109/SSCI50451.2021.9660134  Jierula, A., Wang, S., & Oh, T. (2021). *applied sciences Study on Accuracy Metrics for Evaluating the Predictions of Damage Locations in Deep Piles Using Artificial Neural Networks with Acoustic Emission Data*.  López del Val, J. A., & Alonso Pérez de Agreda, J. P. (1993). Principal components analysis. *Atencion Primaria / Sociedad Española de Medicina de Familia y Comunitaria*, *12*(6), 333–338. https://doi.org/10.5455/ijlr.20170415115235  moneycontrol. (n.d.). *HDFC Bank Ltd.TECHNICALS*. https://www.moneycontrol.com/technical-analysis/hdfcbank/HDF01/weekly  Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., & Kim, H. C. (2021). Stock market prediction using machine learning techniques: A decade survey on methodologies, recent developments, and future directions. *Electronics (Switzerland)*, *10*(21). https://doi.org/10.3390/electronics10212717  Series, I. (2021). Machine Learning Algorithms and Applications. In *Machine Learning Algorithms and Applications* (Vol. 7). https://doi.org/10.1002/9781119769262  Shah, D., Isah, H., & Zulkernine, F. (2019). Stock market analysis: A review and taxonomy of prediction techniques. *International Journal of Financial Studies*, *7*(3). https://doi.org/10.3390/ijfs7020026  Sonkiya, P., Bajpai, V., & Bansal, A. (2021). *Stock price prediction using BERT and GAN*. *6*. http://arxiv.org/abs/2107.09055  Vreeken, J., & Yamanishi, K. (2019). *Proceedings of the 25th {ACM} {SIGKDD} International Conference on Knowledge Discovery & Data Mining, {KDD} 2019, Anchorage, AK, USA, August 4-8, 2019*. 1946–1956. https://doi.org/10.1145/3292500  Сороко, Н. В. (2017). *Масові Відкриті Європейські Он-Лайн Курси Для Вчителів (2017 Р.). Інформаційний Бюлетень№ 1*. *801*, 1–23. |