International Journal of Computing and Artificial Intelligence

E-ISSN: 2707-658X
P-ISSN: 2707-6571
IJCAI 2021; 2(1): 45-52
Received: 09-02-2021
Accepted: 20-03-2021

Aman Sehgal
IIM Lucknow, BITS Pilani

Corresponding Author:
Aman Sehgal
IIM Lucknow, BITS Pilani

Forecasting stock price movements for intra-day trading using transformers and LSTM

Aman Sehgal

DOI: https://doi.org/10.33545/27076571.2021.v2.i1a.25

Abstract
For several years people have tried to find a scientific technique for time series forecasting in the world of stock market trading. A widely applicable model to forecast stock fluctuations can prove to be revolutionary in the worlds of both finance and data analysis. The advent of artificial intelligence and machine learning has sparked a new energy in the quest for algorithm designs. This quest is ignited further since the advent of neural networks and state-of-the-art transformer models. This research envisages taking this quest further ahead by developing a transformer-based model for intra-day stock forecasting. The study takes into account select Nifty 50 stocks and focuses primarily on the Indian stock market.

Keywords: Forecasting stock price, intra-day trading, transformers and LSTM

1. Introduction
Since 1980s there has been an attempt at solving the problem of time series forecasting in finance domain. The objective is to outperform the financial market and earn profits. Even till date, however, financial projection is one of the most difficult time series problems. Financial markets are affected by a large number of exogenous factors like economic, political and even psychological
Financial time series are inherently non-stationary and noisy. Problem of financial stock prediction can be divided into multiple sub problems. This research will primarily focus on intraday trading scenario, trying to identify top and bottom performing stocks from a basket.

2. Literature Review
Fischer and Krauss, 2018 used LSTM networks to predict directional movements for S&P 500 stocks. The study obtained daily returns of 0.46 percent and Sharpe ratio of 5.8 from 1992 to 2009, before accounting for transaction costs, evidencing LSTMs outperformed random forest algorithms. For data post 2010, higher returns seemed to have arbitraged away. Basis the work, they devised a strategy yielding a 0.23 percent return in short term before transaction costs.
(Ghosh *et al.*, 2020) [19], used LSTMs and random forest algorithms to predict intra-day movement of stocks. The study took multiple engineered features as inputs. Key engineered features include:
1. Opening price vis-a-vis previous day closing price
2. Intra-day returns

A combination of random forest and LSTM networks (CuDNNLSTM) was employed as a training methodology. The results show that a multi-feature setting provides a daily return of 0.54 percent for the random forest and 0.64 percent for LSTM based network.

3. Aim
Key objectives of the current research conducted:
- To develop a deep learning-based architecture for intra-day stock prediction
- To evaluate efficacy of proposed architecture vs. other relevant architectures

4. Dataset
Four stocks were taken from NSE, namely Asian Paints, Grasim, Cipla and Eicher Motors. The span of data for these 4 stocks ranged from 1st Jan 2000 to 31st Dec 2010.

on a daily basis. Textile (Grasim), Chemicals (Asian Paints), Pharmaceuticals (Cipla) and Automobile (Eicher Motors) industry have been chosen for analysis.

All the chosen sectors are among the leading contributors to India's GDP. India is the 4th biggest commercial vehicles market in the world and 2nd in the global two-wheeler market. Indian Pharmaceutical sector ranks 4th in the world pertaining to volume of sales. India is the 2nd largest manufacturer of cotton yarn in the world. Chemical industry accounts for approximately 2% of India's GDP.

One leading stock from each industry has been picked up to represent the chosen industry.

The time period from 2000 to 2010 is chosen to provide an adequate amount of time for analysis and also for easy validation of the results (compared to the actual data and vast amount of data on factors contributing to stock movement).

Opening and Closing prices of the above-mentioned stocks are recorded separately.

5. Methodology

Raw data is transformed to turn it into usable format.

Key steps included in the data transformation step include:

- Data pre-processing
- Target selection
- Exploratory data analysis (EDA)

Data pre-processing step focusses on attributes like dividing data set into study periods and engineering input features

Target selection step focuses on developing a problem statement that can be measured quantitively

EDA step focuses on finding key patterns in the input dataset

5.1. Data Pre-processing

The dataset has been divided into "study periods". Total data across 20 years is divided

using a four- year window and one-year stride. Each study period consists of roughly 3 years of training period and 1 year of testing period resulting in 17 non overlapping study periods.

Key terms used in the study include: Total number of days considered for analysis, total number of stocks considered for analysis, number of stocks in study period having complete available history at time j, Closing price for stock at time t, Opening price for stock at time t For prediction time = t, The following inputs and engineered features have been used for the model:

- Historical opening prices O_t for all the stocks at time = 1,2. t-1,t
- Historical closing prices Ct for all the stocks at time = {1,2…t-1}
- Intraday returns for any stock s: $i_{t,m} = \frac{c_{t-m}}{o_{t-m}}$ -1

Return vis-à-vis opening price for stock s: $cr_{t,m} = \frac{c_{t-1}}{c_{t-m-1}} - 1$

Returns vis-à-vis last closing price for stock s: $or_{t,m} = \frac{o_t}{c_{t-m}} - 1$

5.2. Target Selection

A classification problem is designed, where each stock has been divided into two classes based on the intra-day return.

- If the intra-day return of a stock at time t is lower than the median intra-day return, it is put under class 1
- If the return of the stock at time t is greater (>) than median intra-day return it is put under class 2

5.3. Exploratory Data Analysis (EDA)

A detailed EDA of the stocks and relevant sectors has been performed. Key features covered include:

1. Plotting the time series trends
2. Rolling Average time series trend for monthly, quarterly and annual values
3. Auto Correlation Function (ACF) Plot
4. Partial Auto Correlation Function (PACF) Plot
5. Trend, Seasonality and Residue analysis EDA figures are posted in the appendix section.

6. Layout

LSTM and Transformer models are trained on the above-mentioned training data. LSTM and transformer specifications are as follows:

Table 1: *LSTM Specifications*

Dropout	0.1
Activation Function	SoftMax
Batch Size	512
Epochs	1000

Table 2: *Transformer Specifications*

Timestamp	5
Epochs	300
Learning rate	0.001
Timestamp	5

7. Results

This section discusses the results obtained from the LSTM and Transformer models.

7.1. LSTM Prediction Results

As indicated from the table 3, accuracy numbers are very low and the model is unable to explain most of the variance in the original Grasim forecasts

Table 3: *LSTM Prediction Results*

	Grasim Ind	Asian Paints	Cipla	Eicher Motors
Mean Accuracy	44.31%	89.5%	85.5%	28.22%
Std. Deviation	0.45	0.52	0.18	0.99
Coefficient of Variance	0.21	0.27	0.03	0.99
Correlation between actual and LSTM forecast	0.48	0.28	0.59	-0.19

Fig 1: *Grasim Stock Prediction*

For Asian Paints LSTM forecasts show good average values. However, the forecast values have low variance and hence do not capture the variation in the original stock data to a significant extent.

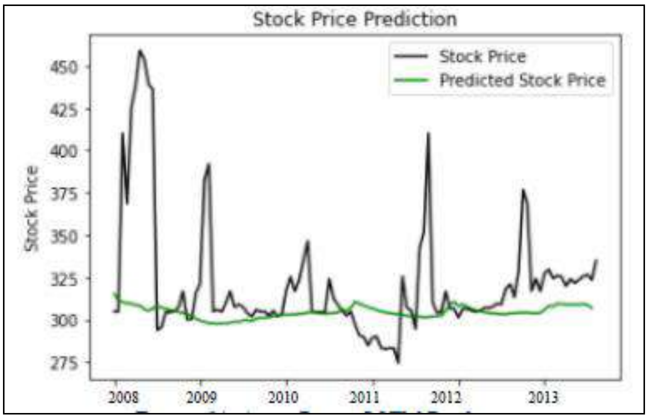

Fig 2: *Asian Paints Prediction*

Cipla forecasts are close to the moving average values of actual stock time series. However, the variance values are even lower than the values for other stocks. This further demonstrates the low extent of variance explained by the LSTM model.

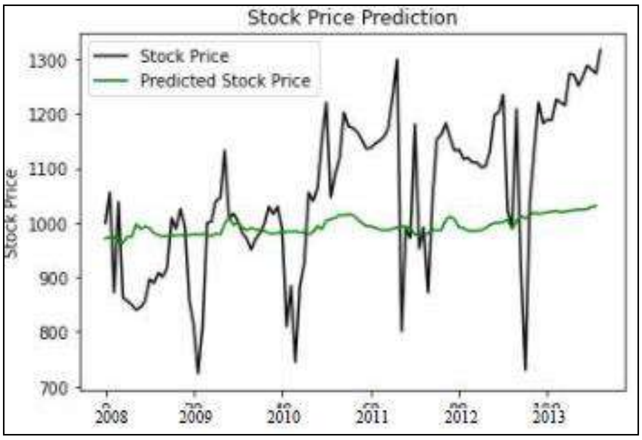

Fig 3: *Cipla Stock Prediction*

The mean accuracy value of Eicher motors forecasts is very low for the LSTM based model. Further, this model is not able to capture the variance in the original time series either

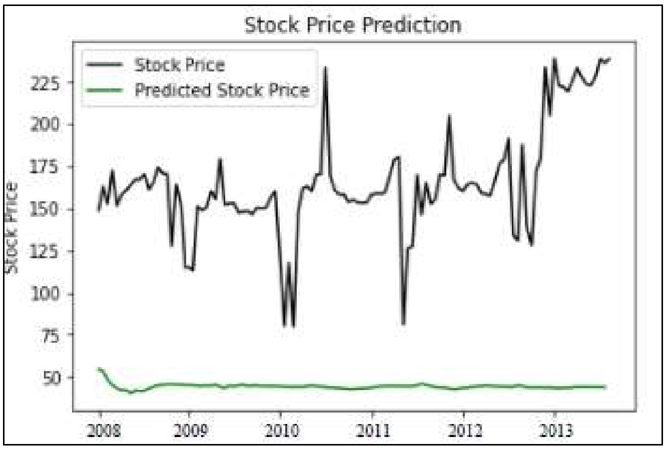

Fig 4: *Eicher Motors Prediction*

Above results show that LSTM based model results are not ideal for forecasting stock price data in the Indian equity market. The mean accuracy hugely varies between stocks and none of the forecasts have a high variance as demonstrated by the original stock price time series.

7.2. LSTM Simulation Results

Both the long and short positions vary above and below the zero value.

Table 4: *LSTM based trading results*

Year	Long Positions Average Daily Returns	Short Position Average Daily Returns
2003	0.3%	0.0094%
2004	-0.07%	0.34%
2005	-0.086%	0.24%
2006	0.084%	0.24%
2007	-0.027%	0.42%
2008	-0.10%	0.33%
2009	0.19%	-0.15%
2010	9.06 e-5	0.15%
2011	-0.032%	0.15%
2012	0.14%	0.09%

The average return of long position from 2003 to 2012 is 0.4% and average return of short position from 2003 to 2012 is 1.8%. This shows that across the years, short position has been more beneficial.

The net results after aggregating long and short position for each of the year. The average return value from 2003 to 2012 is 2.22%

7.3. Transformer Prediction Results

This section shows the results of the transformer-based model developed. Multiple transformer run (10 to be precise) were performed for each of the Stock to check which run gives the best result.

(accuracy is measured as 1 – root mean square error for each of the iteration in percentage separately).

Table 5: Transformation Prediction Result

Iteration Number	Asian Paints Mean Accuracy (in%)	Cipla Mean Accuracy (in%)	Grasim Mean Accuracy(in%)	Eicher Mean Accuracy (in%)
1	85.94%	75.55	70.97	75.18
2	68.21%	74.46	72.9	76.8
3	83.12%	76.03	72.92	68.58
4	78.82%	75.22	72.66	61.41
5	75.93%	76.46	69.3	59.37
6	74.1%	77	71.53	72.15
7	84.3%	76.96	71.67	76.7
8	77.2%	74.73	71.86	76.24
9	81.93%	75.29	70.62	67.14
10	69.6%	75.19	66.75	65.92

For the time period from 10-08-2003 to 31-12-2003, an average accuracy of 76.84% has been observed.

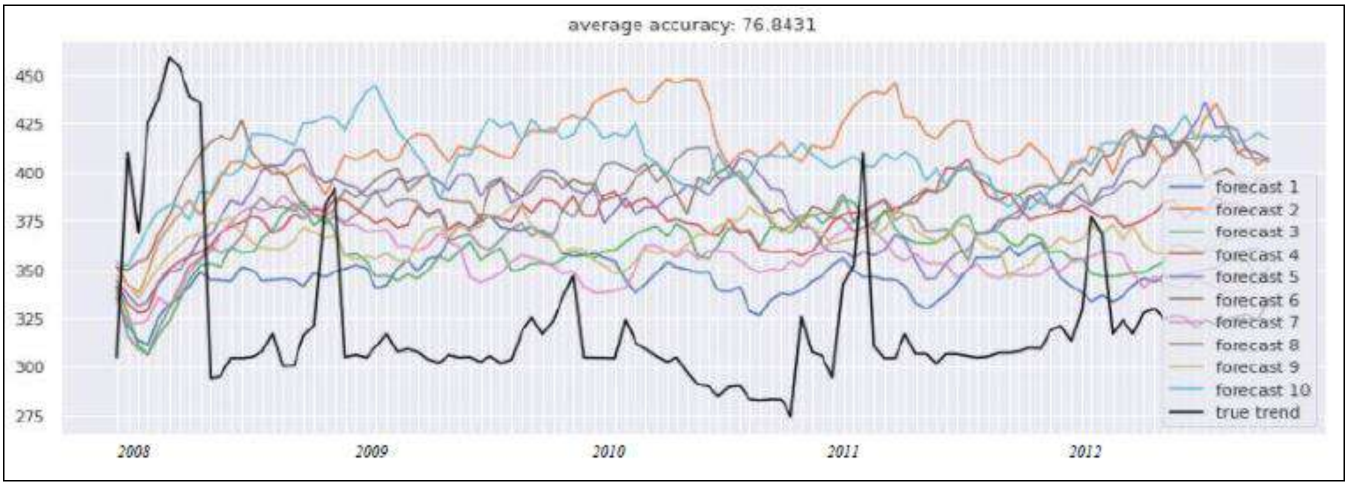

Fig 5: *Asian Paints Transformer Prediction*

Accuracy numbers vary across iterations, with the first iteration giving the best accuracy and the final iteration giving the worst results.

For Cipla, for the duration of 10-08-2003 to 31-12-2003, an average accuracy of 75.66% is observed. However, results are not able to capture the large fluctuations in the actual time series.

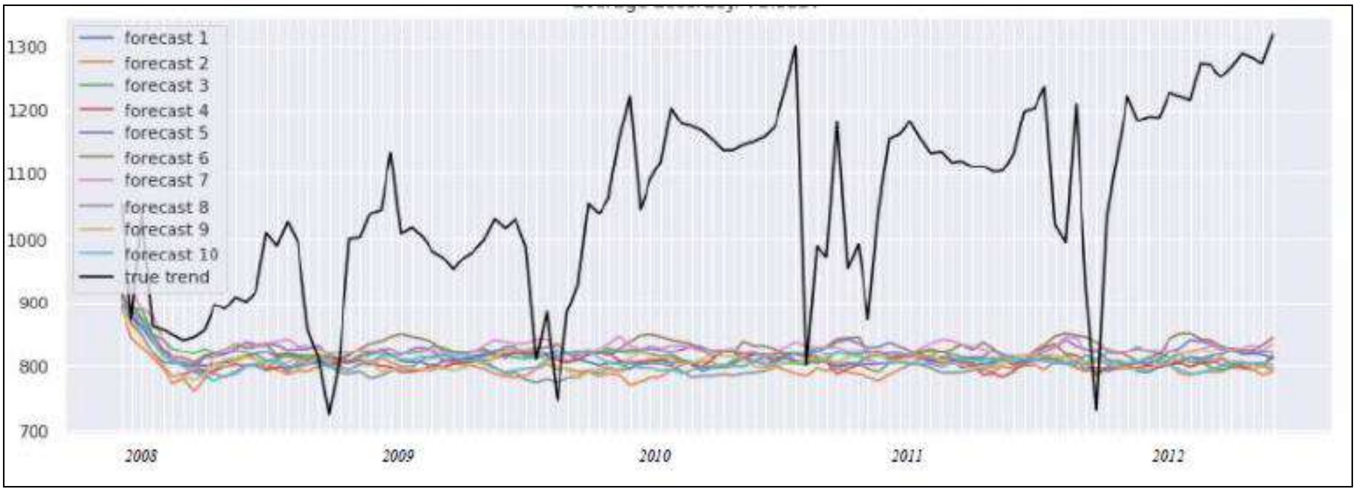

Fig 6: Cipla Transformer Prediction

Across iterations, the mean accuracy doesn't fluctuate much, highest accuracy of 77% is seen in iteration 6 and lowest at iteration 10. Within the year answer, standard deviation of each accuracy measure is not high and ranges between 1.09 to 1.34 across iterations.

For Grasim, for the duration of 10-08-2003 to 31-12 2003, an average accuracy of 70.8% is observed.

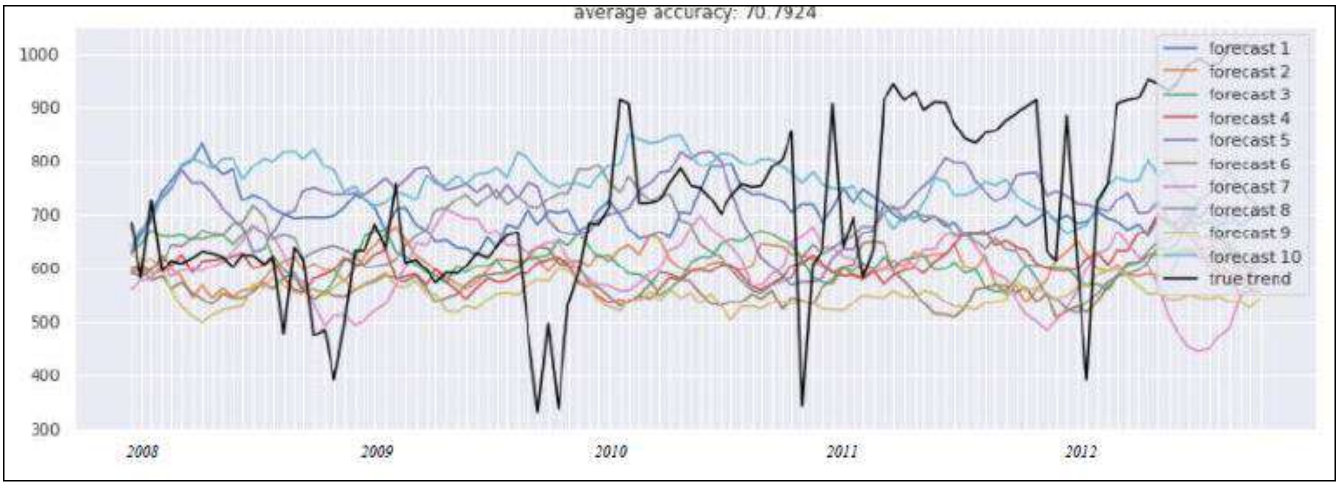

Fig 7: *Grasim Transformer Prediction*

Accuracy numbers varied with each iteration.
however, the variation is not very high across iterations. Mean Accuracy is highest for the first 3 iterations and is lowest for the final iteration.

For Eicher Motors, for the duration of 10-08-2003 to 31-12-2003 an average accuracy of 70.24% is observed.

Fig 8: *Eicher Motors Transformer Prediction*

7.4. Transformer Grain Level Results

Overall, a 49.18% average grain level match across the four stocks is observed.

Table 6: Transformer Grain Level Results

Year	Stock	Grain Match
2003	Asian Paints	47.3%
	Grasim Ind. Ltd	17.8%
	Cipla	32.14%
	Eicher	27.6%
2004	Asian Paints	43.75%
	Grasim Ind. Ltd	50%
	Cipla	47.32%
	Eicher	48.21%
2005	Asian Paints	55.35%
	Grasim Ind. Ltd	52.67%
	Cipla	57.14%
	Eicher	56.25%
2006	Asian Paints	52.67%
	Grasim Ind. Ltd	50.89%
	Cipla	52.67%
	Eicher	50.89%
2007	Asian Paints	51.78%
	Grasim Ind. Ltd	42.85%
	Cipla	46.42%
	Eicher	55.35%
2008	Asian Paints	55.35%
	Grasim Ind. Ltd	47.32%
	Cipla	51.78%
	Eicher	56.25%
2009	Asian Paints	48.21%
	Grasim Ind. Ltd	49.10%
	Cipla	55.35%
	Eicher	52.67%
2010	Asian Paints	43.75%
	Grasim Ind. Ltd	49.10%
	Cipla	49.10%
	Eicher	45.53%
2011	Asian Paints	53.57%
	Grasim Ind. Ltd	53.57%
	Cipla	54.46%
	Eicher	50.89%
2012	Asian Paints	54.46%
	Grasim Ind. Ltd	49.10%
	Cipla	54.46%
	Eicher	54.46%

Stock wise average grain level match from 2009 to 2012 numbers are as follows:

Table 7: Average Returns

Stock	Average Grain Level Match
Asian Paints	50.62%
Grasim	46.24%
Cipla	50.08%

Above numbers show that transformer-based model gives consistent results for the ten-year duration across the four stocks.

7.5. Transformer Forecast Summary

Table 8 shows that Transformer forecasts consistently give profits for the long positions. However, the results for the short position are not as promising. Further, long position results being consistently better than the LSTM based returns.

Table 8: *Transformer Forecast Results*

Year	Long Positions Average Daily Returns	Short Position Average Daily Returns
2003	1.29%	-3.78%
2004	0.26%	-8.77%
2005	0.94%	-0.43%
2006	1.54%	-1.95%
2007	1.50%	-0.77%
2008	0.67%	-2.59%
2009	3.86%	-2.19%
2010	1.82%	-0.9%
2011	0.75%	-0.15%
2012	1.48%	-0.53%

7.6. LSTM vs Transformer Summary

Following figures summarize the returns obtained from the LSTM and Transformer based strategies.

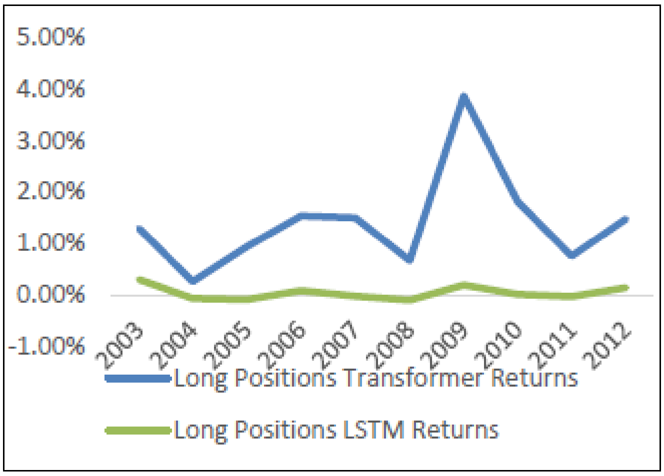

Fig 9: Long Position Returns

Figure 9. shows that the transformer based long position gives significantly better returns as compared to the transformer-based approach.

Fig 10: *Short Position Returns*

However, as depicted by Figure 10. further work needs to be done to improve the short position results.

8. Conclusion

Final results clearly showcase following characteristics of the Transformer based model vis-à-vis the LSTM based model:

- Transformer based prediction of top 2 and bottom 2 stocks, consistently match the actual top 2 and bottom 2 stocks in 49.18% of cases
- LSTM based forecasts show poor correlation with the actual stock time series and the mean accuracy levels vary significantly (from 20% to 80%). Further, LSTM based forecasts show low standard deviation and coefficient of variance. This signifies that the forecasted values don't fluctuate as much as the actual time series forecasts.
- Transformer based stock forecasts show mean accuracy in the range of 70% to 80% for the four stocks analysed. Further, the standard deviation and coefficient of variance for the transformer-based forecast are significantly higher than the LSTM based forecast's values. This shows that the forecasts are better able to capture the variance in the original time series trends.
- Long position returns of transformer-based model are consistently and significantly better than the LSTM based model. Short position returns of the LSTM based model are consistently and significantly better than the Transformer based model
- Above mentioned points clearly signify that a transformer-based architecture is a significantly better algorithm for intra-day long position than existing models
- The EDA demonstrates that all major stocks have auto correlation till lag 2 and have a trend component but lack a consistent seasonality. Further, returns over the 20-year period follow a bell- shaped curve with a few outliers around 2008 and 2020.
- The proposed model of selecting top 10 stocks for long position is ideal for stock selection.

9. Future Work

This research opens up potential for further future work. Key additional research areas that can improve the results include:

- Stock or Industry level analysis can be performed to identify best forecasting approach for individual stock / industry.
- Factors like industry growth prospects, economic prospects, fundamental parameters of the stock, etc. can be included to perform a multivariate analysis on the stock.
- Sentiment analysis can be performed on the stock market news to add additional features to the given analysis to improve the results for the grins where deep-learning based algorithms are lacking.

10. Appendix

Figures listed below show that PACF curve is significant for Asian paints till lag 2 and opening price is the most important feature in forecasting closing price. Also, stock prices across most points in 2020 have been marked as outliers.

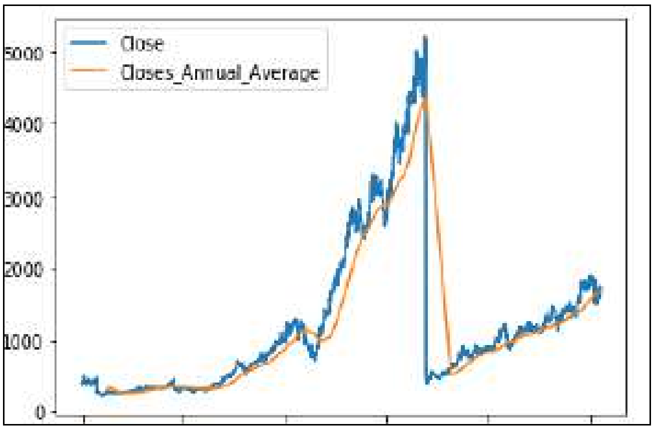

Fig 11: *Asian Paints Time Series*

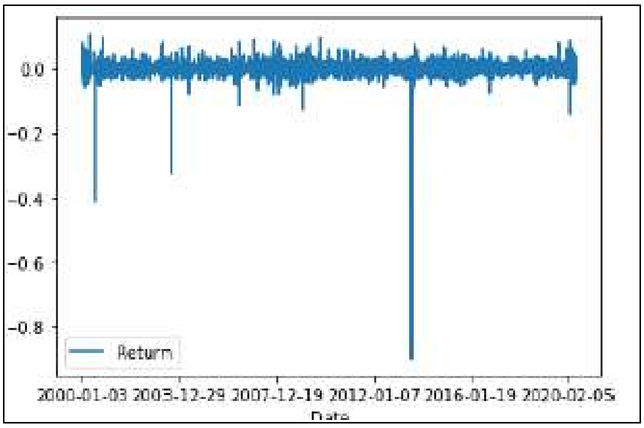

Fig 12: *Asian Paints Returns*

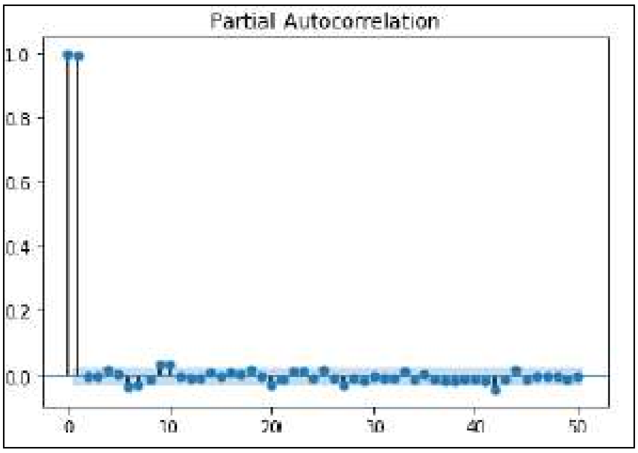

Fig 13: *Asian Paints PACF*

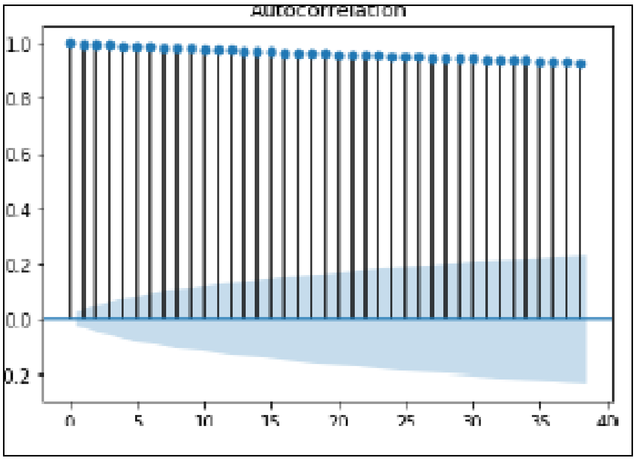

Fig 14: *Asian Paints ACF*

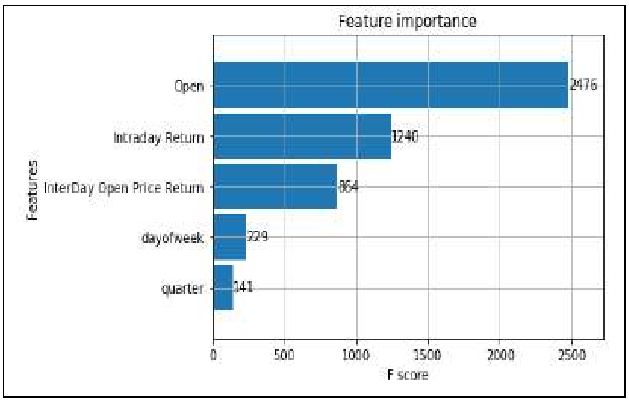

Fig 15: *Asian Paints Key Factors*

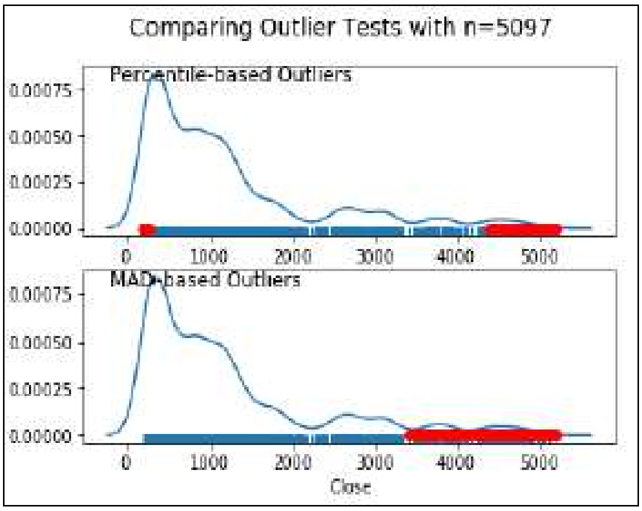

Fig 16: *Asian Paints Outlier Checks*

References

1. Abraham A, Nath B, Mahanti PK. Hybrid intelligent systems for stock market analysis. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 2001;2074:337-345.
2. Abraham A, Philip NS, Saratchandran P. (n.d.) Abraham Philip Saratchandran 2003.pdf.
3. Andrada Félix J, Fernández Rodríguez F, García Artiles MDD, Sosvilla Rivero SJ. An Empirical Evaluation of Non-Linear Trading Rules. SSRN Electronic Journal 2005.
4. Andreou AS, Neocleous CC, Schizas CN, Toumpouris C. Testing the predictability of the Cyprus Stock Exchange: The case of an emerging market. Proceedings of the International Joint Conference on Neural Networks 2000;6:360-365.
5. Anon. American Finance Association Predicting Stock Returns in an Efficient Market Author (s): Ronald J. Balvers, Thomas F. Cosimano and Bill McDonald Published by : Wiley for the American Finance Association Stable URL : http://www.jstor.org/stable/232871. 2016;454:1109-1128.
6. Atsalakis GS, Valavanis KP. Surveying stock market forecasting techniques - Part II: Soft computing methods. *Expert Systems with Applications*, [online] 363 PART 2009b;2:5932-5941. Available at: http://dx.doi.org/10.1016/j.eswa.2008.07.006.
7. Augusto Casas C. Tactical asset allocation: An artificial neural network based model. *Proceedings of the International Joint Conference on Neural Networks*, 2001;3:1811-1816.
8. Avellaneda M, Lee JH. Statistical arbitrage in the US equities market. Quantitative Finance 2010;107:761-782.
9. Baba N, Kozaki M. An intelligent forecasting system of stock price using neural networks 2003;1:371-377.
10. Baek J, Cho S. An up-trend detection using an auto-associative neural network: KOSPI 200 futures. *Lecture* Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2002;2412:359-365.
11. Desai VS, Bharati R. A comparison of linear regression and neural network methods for predicting excess

returns on large stocks. *Annals of Operations Research*, 1998;78:127-163.
12. Dickey DA, Fuller WA. Distribution of the Estimators for Autoregressive Time Series With a Unit Root. Journal of the American Statistical Association 1979;74366:427.
13. Dixon M, Klabjan D, Bang JH. Implementing deep neural networks for financial market prediction on the intel xeon phi. Proceedings of WHPCF 2015: 8th Workshop on High Performance Computational Finance - Held in conjunction with SC 2015: The International Conference for High Performance Computing, Networking, Storage and Analysis 2015.
14. Do B, Faff R, Hamza K. A new approach to modeling and estimation for pairs trading. Proceedings of 2006 Financial Management Association European Conference, [online] Monash Uni, 2006, 31. Available at:
15. Doeksen B, Abraham A, Thomas J, Paprzycki M. Real stock trading using soft computing models. International Conference on Information Technology: Coding and Computing, ITCC, 2005;2:162-167.
16. Dunis CL, Laws J, Evans B. Modelling and trading the gasoline crack spread: A non-linear story. Derivatives Use, Trading & Regulation 2006;121:126-145.
17. Engelberg JE, Reed AV, Ringgenberg MC. How are shorts informed?. Short sellers, news, and information processing. Journal of Financial Economics, [online] 2012;1052:260-278. Available at: http://dx.doi.org/10.1016/j.jfineco.2012.03.001.
18. Gençay R. Optimization of technical trading strategies and the profitability in security markets. *Economics Letters*, 1998;592:249-254.
19. Ghosh P, Neufeld A, Sahoo JK. Forecasting directional movements of stock prices for intraday trading using LSTM and random forests. [online] December 2020, 1-8. Available at: http://arxiv.org/abs/2004.10178.