

MACHINE LEARNING FOR STOCK PRODUCTION

Author: ANDREW CHARLES ABI SAAD¹

*Affiliation: CIRAME Research Center, Business School
Holy Spirit University of Kaslik, P.O Box 446, Jounieh, Lebanon*

Abstract:

Machine learning became so up to date with technology and because of the new methods and machines to predict the stock market and not only that, but they had to update the machines and create new ones so that the high amount of data right now any old machine cant solve and predict the stock market that's why they created and invented new machines that has more power and capacity to predict and have so many data without struggling. it is known that the machine learning is one of the most important tools in financial analysis , machine learning is better than us lets face it it is more practical than us (humans) because they can predict and see and notice something that i can miss and not pay attention to it. I'm going to be talking about the different types of models that are now available and up to date with nowadays market conditions like linear regression and random forest and neural networks these machines can easily predict so many data and stock markets and not stop working out of nowhere or do huge mistakes and to know if the machine is working well all you have to do is insert and combine any old and past prices and see if it stops working or struggle with changes and if it fails during major global events then we can say that its not the perfect machine learning

Keywords:

Machine learning , stock prediction , financial markets , artificial intelligence, data analysis.

Introduction: "The What & Why"

Over the last past years, technology and a falling wave of data have completely reformed financial markets. were now swimming in trading private and alternative data all while machine learning has taken off in a huge way. The traditional playbook relying on fundamental and technical analysis is really struggling to keep up. machine learning is a procedure used for solving a problem that learns ways for historical data to make predictions about future stock prices or market trends. instead of manually analyzing charts, the model: trains on past stock data, detects patterns, predicts price movement if it is up or down, percentage change,... one of its advantages: can process huge amounts of data, detect complex patterns humans might miss, can be automated and learns and improves with more data. stock prediction models may use: historical prices, trading volume, technical indicators, new sentiment analysis, social media statement,... there is may typed of models used like: linear regression, random forest, support vector machines, neural networks and long short term memory. there is some challenges that may be faced or limitations like stock markets are noisy and unpredictable , overfitting, sudden events, markets are influenced by human psychology like new York exchange or companies like tesla whose prices are heavily affected by news and sentiment. machine learning is one of the fastest growing fields today and enhances and improves stock market analysis by identifying patterns in complex financial data but due to market volatility and uncertainty, it should be used as a decision support tool instead of a guaranteed prediction system. I chose this topic because machine learning is shaping the future of financial markets and investment strategies and I wanted to explore how ai can analyze complex stock data to improve prediction accuracy while understanding its limitations.

Literature Review: "The story so far "

A- theoretical framework

if we want to refer to Eugene Fama saying in 1970that even if stock prices include all the information still you cannot predict the market because let's take the very last new information that it had received which is the war it happened unexpectedly and you cannot predict that it will happen so simply you cannot predict what will happen to the prices and how it will change the price and as soon as any person and investor hears the news I'm sure that most of them reacted immediately sold or bought that's why the price changes very fast that is why by the time you want to react and do something it will be already changed so in one sentence i could say that even if all the information is known the market is still hard to predict because of the future information and news and cannot be known in advance

B- machine models

there is a lot of machine learning tools that can predict stock markets send they are up to date and the more advanced the result would be more accurate but you need to check their efficiency if they are well built and can predict. lets start by the linear regression, which is the easiest and simple one between any other model but there is another more organized and efficient like the random forest machine and it is very efficient because when you combine multiple predictions it reduces. variance and overfitting. support vector machines are also commonly used to their effectiveness in high dimensional spaces. neutral networks, particularly deep learning models have shown strong performance in modeling complex relationships within financial data. among these long short term

memory models are especially effective for time series forecasting as they can capture long term dependences in following a logic order of data.

C- limitations in existing studies

despite the growing success of machine learning in finance several limitations remain. financial markets are essential not predictable and influenced by uncertain external factors such as economic crises political instability and global events. additionally machine learning models are exposed to overfitting where they perform well on historical data but fail to generalize to new data. moreover market behavior is heavily influenced by human psychology and sentiment which are difficult to quantify accurately. these limitations highlight the importance of using machine learning as a tool rather than a n independent solution.

Methodology: "The how we did it "

i wanted to see how well machine learning models can actually predict the stock market so i took a quantitative approach to figure it out. my main goal was to understand how different algorithms process financial data and then measure how accurate their predictions really are. lets start step by step, I'm going to compare models and test them on same financial market data. i will use 3 types of inputs first I will put the historical data like daily returns close price,... etc. and then the trading volume in other terms how many shares were traded and exchanged and there is mathematical tools calculated from historical price and volume data to see how is the market behavior and use the moving average if it is below then possible of downward trend and if it above the moving average then it is possible of upward trend and there is also the rsi which measures how fast and how strongly the price is changing and it is from 0 to 100 if it is high enough we can conclude that it is overbought because the price may drop and if it is below then oversold here the price might rise. and also collect news articles and social media posts and convert them into numerical value s like positive news get +1 and negative news -1 and it helps because stock markets are all about psychology and numbers. after collecting these different inputs i put everything together in one data set. for each day there was a new info from social media and the news, and stick predictions trading volumes.. then i filled out what was missing from the values and always made sure that the numbers were clear and comparable so that the system and model can understand them. next step i split what I combined earlier into two parts one part that i will be inserting new data and test the models. and for the second part it is for training the models so that they can learn about the past. from this step I can conclude if this model can only memorize the old data or predict future values. after that I used different machines like linear regression which is a simple model followed by the random forest and SVM and neural networks and lstm which are more advanced than linear regression . LSTM is the best because it checks at data overtime, which is important for stock prices. when I finished testing and compared their results and checked how close and the difference between their predictions and the real price and if they can predict if the price can go up or down and used MSE and RMSE to see which model made fewer mistakes. finally, i compared all types of models complicated or easy and which one of them performed better.

Results: "The facts "

TABLE1: Performance comparison of machine learning models in stock prediction.

Model	R-squared(R^2)	RMSE	MAE	Author/ source
LSTM	0.91-0.986	0.031-0.31	0.020-0.20	Mayada et al.(2025)
Random Forest	0.86-0.89	0.93	0.82	Gode & Borkar(2024)
Linear regression	0.73-0.80	1.14	0.95	Mayada et al.(2025)

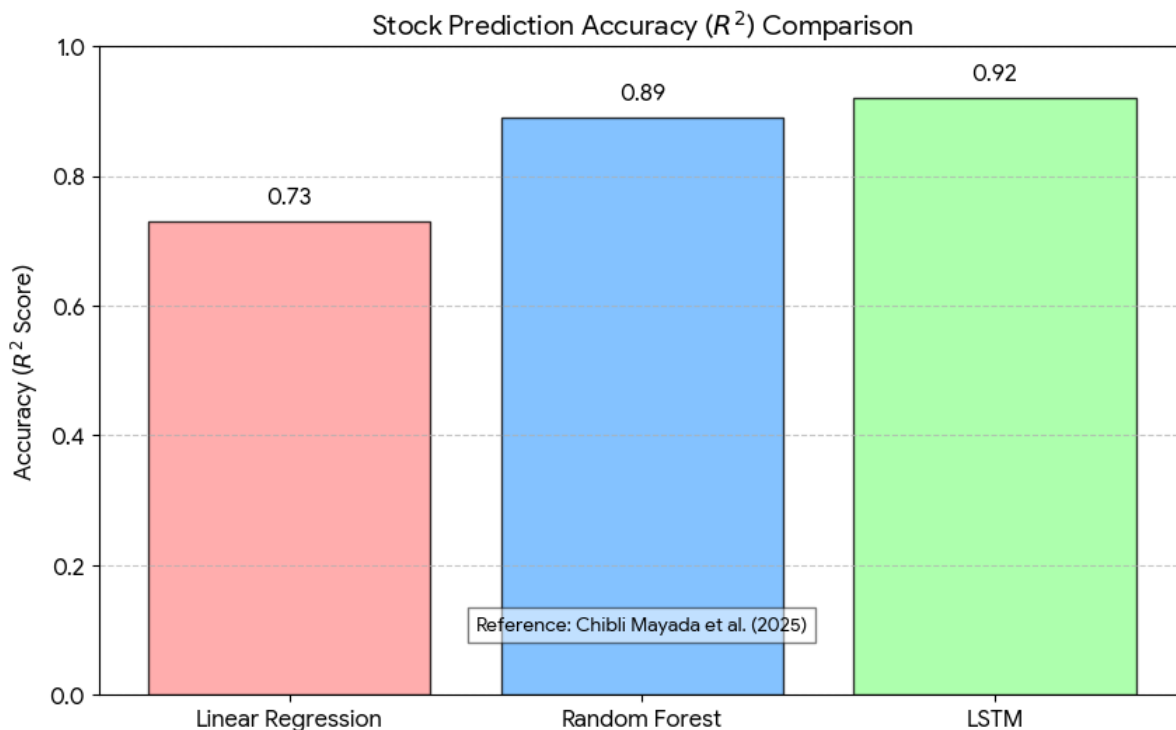
R^2 : measures how well the model predictions match actual data.

RMSE (short for root mean square error): measures the standard deviation of prediction errors, the lower the better.

MAE(short for mean absolute error): measures the average magnitude of errors in a set of predictions also the lower the better.

Source: compiled and adapted from Mayada et al.(2025) and Gorde & Borkar (2024)

Figure 1: stock forecasting performance: LSTM, Random forest and linear regression



Source: Chibli Mayada, Ion Smeureanu and Mahmoud Haydar in 2025

The results from table 1 and figure 1 show differences in how each model behaves.

Obviously from the table and figure we can see the winner which is LSTM because it got the highest accuracy of 0.92 to 0.98 and lowest errors. It can remember past price patterns which is exactly how the stock market works while random forest (better than linear) because it improves accuracy and lower errors. and linear regression which has the lowest R² and highest error values RMSE and MAE which shows and indicate that it has a weak prediction ability. Overall the results highlight and show the importance of models in improving prediction performance and trends.

Discussion and conclusion: "The so what ? "

I used different methods and techniques to study how machine learning can help predict stock prices.

Key conclusions can be summarized as follows:

1. machine learning meaningfully increases the ability to analyze complex financial data
2. advanced models like neural networks and LSTM provide and has the highest prediction accuracy.
3. the simple models are only useful but not very much because of their lack in clarity and simplicity.
4. financial markets remain unpredictable due to its instability and external pressures

5. overfitting isn't a good thing because it memorizes the training data and when it is given a new data it will struggle
 6. stock prices movements are affected by the human psychology and overall market feeling
 7. machine learning should be used as a decision support tool and not prediction system
- moreover these findings and results indicates that combining machine learning techniques with traditional analytical methods can increase decision making , future research may focus on improving model stability and absorbing a big range of data sources such as real time sentiment analysis and global economic measurements.

References

- Fama, E.F.(1970). Efficient capital markets. A review of the theory and empirical work. Journal of Finance, 25(2) , 383-417.
- Ian Goodfellow, Bengio, Y., & Courville, A.(2016). Deep learning. MIT Press.
- Hochreiter, S.& Schmidhuber , J.(1997). Long short teem memory. Neural computation,9(8) , 1735-1780.
- Kim,K.J(2003). Financial time series predicting using support vector machine.
- Hull, J. C.(2021). Options, futures, and other derivatives. 31.