Market Trade Forecast: A Probabilistic Approach to Price Reversion Analysis

Mabunda Hlulani
Department of Computer Systems Engineering
Tshwane University of Technology
Pretoria, South Africa
213067605@tut4life.ac.za

Dr. Oluwasogo Moses Olaifa (Supervisor)
Department of Engineering and the Built Environment
Tshwane University of Technology
Pretoria, South Africa
olaifamo@tut.ac.za

Dr. Chunling Du
Department of Computer Systems Engineering
Tshwane University of Technology
Pretoria, South Africa
duc@tut.ac.za

Dr. Oluwasogo Moses Olaifa
Department of Computer Systems Engineering
Tshwane University of Technology
Pretoria, South Africa
olaifamo@tut.ac.za

I. Introduction

Financial markets are a tricky space to navigate—stock prices move in ways that can feel almost chaotic, which is why figuring out how to predict them is such a big deal for researchers and traders alike. A lot of the models we use today are focused on one main question: are prices going up or down? They often lean on things like technical charts, company fundamentals, or even fancy algorithms to spot patterns. But while that's useful, there's something these tools tend to skip over—the possibility that prices might loop back to important levels they've already touched during a trading day.

This research takes a different angle, aiming to estimate the likelihood of a stock's price revisiting specific points, like:

1) The price it opened at today
2) The opening price from yesterday
3) Yesterday's closing price

I've spent some time looking at intraday price data from the past, and it's clear that stocks often circle back to these levels more than you might expect. Instead of just chasing trends like most models do, this approach zooms in on the odds of those returns happening at certain times—like early in the morning at 8:00 AM, around noon, or toward the end of the day at 4:00 PM. The idea is to give traders a better sense of when they might see an opportunity.

A. Aim and Objectives

This research intends to address gaps in conventional stock forecasting by creating a probability-driven model tailored for the requirements of intraday trading. The objective is to design a tool that will offer traders a competitive edge by enhancing precision, timing, and risk management. Here's what the project aims to accomplish:

- Analyze the frequency with which prices return to important reference points such as the day's open, prior day's open, or prior day's close which are commonly seen intraday trading.
- Determine the times of day when entering and exiting trades would be most optimal, something which models do not usually cater for.
- Forecast movements of price using the past intraday data, placing greater emphasis on probabilities of reversions rather than straightforwardly forecasting trends, which is what most machine learning algorithms focus on.
- Ensure the model's applicable scope is not limited to a specific asset class so that it can be used in different markets.
- Validate the model in a real trading environment for 30 days to assess performance in practical scenarios.

II. Problem Definition

One of the biggest issues with a lot of prediction models is that they assume stock prices follow a neat path—either steadily climbing or dropping without much back-and-forth. But if you've ever watched a trading session unfold, you know that's not how it works. Prices often bounce back to levels they've hit before, sometimes several times in a day. That can happen for all sorts of reasons—maybe there's a sudden shift in liquidity, or big institutional traders make a move, or the market's mood just changes [1]. When models don't account for this kind of behavior, they're missing a big part of what makes intraday trading so dynamic.

Then there's another problem: these models don't really line up with how a trading day flows. Traders often make decisions based on key moments—like the rush when the market opens, the quieter stretch in the middle of the day, or the flurry of activity right before the close. But most forecasting tools don't break things down that way. Even the really advanced machine learning techniques, which are great at picking up on bigger trends, usually don't focus on the odds of a price returning to a specific spot [2]. That leaves traders without the kind of detailed insights they need to make the most of their timing.

Right now, most stock prediction efforts are all about direction—will the price go up or down? [3]. That's a good starting point, but it doesn't tell you much about the likelihood of a price swinging back to a past level, which is something that matters a lot in short-term trading. This study is here to tackle that gap with an approach that's more focused on probabilities, especially for intraday scenarios.

Another challenge with conventional models is their lack of integration with specific intraday timeframes. Traders often make decisions based on critical market hours, such as the opening, midday, and closing periods. However, most forecasting approaches do not analyze price movements within these time-sensitive intervals, leading to suboptimal trading strategies. Additionally, current machine learning models primarily focus on predicting overall price trends rather than determining the probability of price returning to key levels [2]. This creates a gap in forecasting accuracy, as traders are unable to assess the likelihood of price reversion at specific points in time.

Most existing stock price prediction models focus solely on directional movements, aiming to predict whether a stock's price will increase or decrease [3]. While these models are widely used, they lack the ability to provide probabilistic insights into price reversion, which is a crucial aspect of short-term market behavior.

III. Literature Review

A. Machine Learning in Stock Price Forecasting

Machine learning has become a favorite tool for predicting stock prices, and it's easy to see why—it's really good at sorting through complicated financial patterns. Zhang et al. [1] took a deep dive into models like Support Vector Machines, Random Forests, and LSTM networks to see how well they could predict price directions. They found that these methods definitely improve accuracy, but they also noticed a catch: the models struggle when the market gets messy, and they don't really deal with probabilities. I've taken a lot of inspiration from their work, but I'm trying to go a step further by adding a probabilistic angle to look at price reversion, while still using machine learning to make the results more dependable.

B. Deep Learning for Temporal Dependencies in Stock Markets

Deep learning, especially with LSTM networks, has been a total game-changer when it comes to spotting patterns that play out over time in stock data. Kumar and Lee [2] did some interesting work comparing LSTM models to Random Forests for predicting intraday movements in the S&P 500. They found that LSTMs did better, pulling in a daily return of 0.64% compared to 0.54% for Random Forests, which makes sense since LSTMs are so good at handling sequences. But their study didn't look at the odds of prices swinging back to earlier levels, which is where I saw an opportunity. I'm taking their approach as a starting point and mixing in some probability calculations to estimate the chances of those reversions happening, which I think could be really useful for traders.

C. Probabilistic Models in Financial Forecasting

Some recent work has started to lean more toward probabilistic approaches instead of just making flat-out guesses. Qian et al. [3] came up with a Transformer-based network to figure out the probability distribution of stock returns, and they showed it's more reliable than older methods like regression. Wang et al. [4] also jumped in, using Hidden Markov Models to study how markets shift between different states, which really drives home how useful probability can be. I'm drawing on their ideas here, but applying them specifically to intraday price reversion, which hasn't been explored as much as it deserves.

D. Influence of Trading Volume on Intraday Stock Movements

Trading volume has a huge impact on how prices move during the day—it's a big factor in liquidity and volatility. Nguyen et al. [**?**] looked into how volume connects to price reversions, putting together a model to predict volume chunks throughout the day to help with forecasting. They noticed that when volume spikes, prices often end up hitting key levels, but they didn't take the next step into a full probabilistic setup. I'm building on that by weaving volume into my probability model, so traders can get a better sense of when liquidity might trigger a price swing.

E. Feature Engineering in Stock Price Forecasting

Getting the inputs right is so important for making prediction models work well. Li and Chen [**?**] dug into using custom features—like moving averages and momentum indicators—in deep learning models to spot when trends might flip. They found that choosing the right features makes a big difference, even though they were mostly focused on direction. I'm taking a similar approach, but tweaking it to focus on intraday reversion odds, using things like how often prices have reversed in the past and what the volume looks like at those turning points.

F. AI-Powered Stock Forecasting and Trading Systems

Getting the inputs right is so important for making prediction models work well. Li and Chen [**?**] dug into using custom features—like moving averages and momentum indicators—in deep learning models to spot when trends might flip. They found that choosing the right features makes a big difference, even though they were mostly focused on direction. I'm taking a similar approach, but tweaking it to focus on intraday reversion odds, using things like how often prices have reversed

in the past and what the volume looks like at those turning points.

IV. Methodology

A. Research Design

This study is guided by a detailed plan that systematically addresses machine learning with probabilistic techniques to address the challenges of intraday forecasting. Each step has been thought through carefully to make sure the approach is both thorough and practical, especially for a research project at this level

1) Analysis Tools, Software, and Data Sources

This section breaks down the tools, software, and data sources that will be used to make this research happen. A lot of thought went into picking these to keep things manageable for a master's project while still being up to the task of handling intraday forecasting. Here's the rundown:

- **Data Sources**: Pulling historical and real-time data feeds from Yahoo Finance (always free intraday goodies), Alpha Vantage (over 20 years of history), and perhaps Bloomberg for pro-grade checks.
- **Programming Environment**: Python has been my favorite for quite sometimes now: Pandas and NumPy for the time series heavy lifting, and Scikit-learn for my initial models and metrics.
- **Machine Learning Libraries**: For LSTM/Transformer builds, TensorFlow or PyTorch will do just fine. With GPU juice for bigger data pulls.
- **Probabilistic Modelling**: HMMs (hmmlearn) and Bayesian Networks (Pgmpy) cover the odds features, and ever were Monte Carlo simulations for the "what ifs".
- **Visualising and Dashboarding**: Matplotlib and Seaborn will do the charts while Plotly (Dash) provides for the elegant real time UI
- **Deployment framework**:Serve it as an API with Flask or Django, running on cloud or local rig for fast iterations.

Software: Python, Pandas, Numpy, Scikit-learn, Tensorflow/Pytorch, Flask/Django, Matplotlib, seaborn and Plotly/Dash.
AI/ML Tools: HMMs, Bayesian Networks, LSTMs, Transformers, Monte Carlo sims.
Testing Tools: Backtesting setups, MAPE, precision and recall, 30 days out live trading sandbox.

V. Expected Output

1) Benefits / Contributions

This research aims to deliver some practical benefits that could make a real difference for traders. Here's what's expected from the model once it's up and running:

1) **Better Forecasting**:This model steps up intraday predictions with probability vibes—not just "here's the price," but "here's how likely it is," giving traders a range to play with and confidence to time things right.
2) **Smarter Decisions**: With odds tied to key moves, traders can pick their spots—like shorting if today's open looks likely to pull back—balancing risks and wins better.
3) **Real-Time Flexibility**: The system will be designed to keep up with live data, shifting gears if the market flips, and ping traders with fresh insights fast.
4) **Scalability**: The model will be built to flex—stocks, forex, crypto, whatever—and plug into trading platforms, perfect for solo traders or big firms alike.

2) Limitations

No model is perfect, and this one will have its challenges to watch out for. Here are some potential limitations that might come up:

- **Market Chaos**: Sudden events can really shake things up and mess with historical patterns—like major updates from the Fed or unexpected market crashes—which might cause the model to stumble when the market gets wild and tough to predict.
- **Tech Load**: Crunching minute-by-minute data takes serious horsepower; lag could be a pain in live trading.
- **Data Risks**: The saying "garbage in, garbage out" applies here—if the data is messy or has gaps, the model's performance could take a hit. Clean, reliable feeds will be essential for it to work well.
- **Market Shifts**: What worked in the past might not hold up as markets change over time, so the model will need regular tweaks to stay relevant and effective.

3) Ethical, Health, and Safety Issues

This project also takes into account some ethical and safety concerns to ensure the model is used in a responsible way. Here's what's on the radar:

- **Privacy**: Working with live market data means following strict rules to keep everything secure and above board, ensuring no sensitive information gets mishandled.
- **Bias**: If the historical data has any unnoticed imbalances, the model might end up favoring certain outcomes in ways that aren't quite fair, so keeping things even and balanced will be a big priority.
- **Over-Reliance**: Traders shouldn't rely solely on the model—outside factors like breaking news or gut instincts still matter a great deal, and ignoring them could lead to risks building up over time.
- **Security**: Since the system will be running live, strong protections will be needed to guard against hacks or breaches, which could be a big problem if not managed well.
- **Market Ripple**: If a large number of traders start using the model, it might unintentionally affect prices in strange ways, so that's something to keep an eye on.

VI. CONCLUSION

This research offers a fresh perspective on stock price forecasting by focusing on a probability-based approach that centers on price reversion to key reference points. By blending machine learning with probabilistic models, a tool has been developed that traders can rely on to improve decision-making, manage risks more effectively, and execute trades with greater precision in the fast-paced world of intraday trading.

ACKNOWLEDGMENT

The author expresses gratitude to Dr. Oluwasogo Moses Olaifa and Prof. Chunling Du for their guidance and support. Thanks also to colleagues, family, and the institution for their encouragement throughout this research.

REFERENCES

[1] Zhang et al., ”Comparison of Machine Learning Models for Stock Price Direction Prediction,” 2023.

[2] Kumar and Lee, ”Effectiveness of LSTM and Random Forests in Forecasting Intraday Stock Price Movements,” 2024.

[3] Qian et al., ”Transformer-based Neural Network for Probability Distribution of Stock Returns,” 2023.

[4] Wang et al., ”Hidden Markov Models for State Transitions in Financial Markets,” 2024.