Investigating the Viability of AI-Driven Chatbots as Autonomous Agents for Predicting Stock Prices Using Historical Company Data: An Analysis of Models and User Interaction.

Contents

Abstract

This research explores the viability of AI-driven chatbots as autonomous agents for predicting stock prices using historical company data. By integrating advanced technologies such as Natural Language Processing (NLP), predictive modelling with Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN), and sentiment analysis, this study evaluates the effectiveness of chatbots in offering reliable financial advice. The project utilises real-time stock data from the yfinance API, sentiment analysis on financial news, and comparative testing between LSTM and CNN models to determine which offers superior predictive accuracy. Furthermore, the study incorporates user experience feedback through the User Experience Questionnaire (UEQ) to refine the chatbot's functionality and conversational interface. The findings aim to provide insights into the design and operational strategies of financial chatbots, contributing to the broader discourse on AI's role in enhancing financial decision-making tools.

Introduction

Background Research:

Overview of AI-driven chatbots and their growing role in finance

The rapid advancement of artificial intelligence (AI) has revolutionised various sectors, including finance. One notable innovation in the financial industry is the use of AI-driven chatbots, also known as conversational agents. These chatbots can engage in natural conversations with users, providing personalised financial advice, and performing complex financial tasks. In finance, they have emerged as Robo-Advisors or virtual financial assistants, capable of assisting with investment decisions, portfolio management, and stock price predictions. By leveraging machine learning algorithms and natural language processing (NLP), these agents aim to offer accurate, data-driven insights in real-time.

Inspired by companies such as Wealthfront (Stein, 2018). I became curious about pitting two models against each other and finding out which would perform better when it came to predicting stock and subsequently using the best one in an investment type chatbot.

I delved deeper into the research for my question eventually concluding that, stock price prediction is a significant challenge in finance due to the inherent volatility and complex market dynamics. Despite these challenges, predictive models like Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN) have shown promise in forecasting future stock prices by analysing historical data, and through some research came across the yfinance API which has stock prices dating back to the 1070s (Hofstra, 2023).

Through further research I discovered LSTM excels at learning and remembering patterns over time, making them suitable for sequential data like time series. Research shows that LSTM models outperform traditional machine learning models in predicting stock prices with higher accuracy due to their memory cells that manage past information effectively (networks, 2017); (Thomas Fischer, 2018). Whereas although CNNs are widely used for image processing they can also be adapted for stock price prediction. They can capture intricate patterns and features in stock price movements by treating time series data like a visual signal (Taewook Kim, 2019). CNNs are also capable of detecting local dependencies and patterns, which are crucial in identifying trends in stock market data. In particular, their ability to recognise trends and movements across various temporal scales makes them highly effective for financial predictions (Xiao Zhong, 2019). Due to this newfound information I decided to implement each model and see which of the two would give a more accurate result for the chatbot.

Throughout My research it became clear that implementing sentiment analysis of financial news articles would valuable tool in understanding market sentiment and investor behaviour. Integrating these predictive models and sentiment analysis into chatbots would offer users comprehensive, personalised investment advice.

As this application is primarily a chatbot therefore the importance of user feedback plays a critical role in the ongoing development. This direct input from users is indispensable for modifying the chatbot's knowledge base and optimising algorithms to improve response quality. Additionally, feedback sheds light on the user experience, revealing both strengths and weaknesses in the chatbot's interface and interactions. User Experience Questionnaire (UEQ) is a fast and reliable questionnaire to measure the User Experience of interactive products (Andreas Hinderks, 2008), and having used it previously on other projects I felt it was perfect for this one.

Please assess the product now by ticking one circle per line.

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

Purpose of the Research and Primary Question:

The primary purpose of this research is to investigate the viability of AI-driven chatbots as autonomous agents for predicting stock prices using historical company data. The research aims to assess the efficiency of different predictive models and evaluate how well the chatbot provides accurate investment advice while engaging in human-like conversations. The main research question guiding this investigation is:

Investigating the Viability of AI-Driven Chatbots as Autonomous Agents for Predicting Stock Prices Using Historical Company Data: An Analysis of Models and User Interaction.

Research Context

To address the primary research statement, the chatbot, StockBot, integrates several AI technologies to offer comprehensive investment advice. The architecture of StockBot includes:

- **Natural Language Processing (NLP):** To recognise user intents and engage in human-like conversations.
- **Predictive Modelling:** Utilises both LSTM and CNN models to predict future stock prices based on historical data.
- **Sentiment Analysis:** Analyses financial news articles to assess market sentiment and incorporate it into investment advice.

Main Research Questions and primary purpose of chatbot:

The primary purpose of the StockBot chatbot is to serve as an AI-driven financial advisor, providing real-time stock market predictions and financial guidance. It should utilise the machine learning models LSTM and CNN to forecast stock prices based on historical data from the yfinance API, and aiding users in making informed investment decisions.

Throughout my background research for creating this application I felt the five questions below encompassed the main agenda of the goal of the project helping me stay on track throughout. The 5 questions are:

Question	**Explanation**
How does the chatbot recognise user intents and respond appropriately?	Assessing the effectiveness of NLP modules in understanding user queries and intents.
What predictive model is more effective for stock price prediction (LSTM vs. CNN)?	Analysis of LSTM and CNN models in predicting stock prices.
How does sentiment analysis of financial news articles contribute to investment advice given by the chatbot?	Evaluating the impact of sentiment analysis on the users.
How effectively can the chatbot simulate a human-like conversation while providing accurate financial advice?	Investigating how well the chatbot engages users in conversation and provides relevant financial advice.
How can the user experience of the chatbot be improved based on user feedback and UEQ results?	Using User Experience Questionnaire (UEQ) results to refine and enhance the chatbot's conversational design and user experience.

The outcomes of this research will provide insights into the viability of integrating predictive modelling, sentiment analysis, and NLP into autonomous financial agents. The findings will also help in identifying best practices and areas of improvement for developing intelligent financial chatbots.

Methodology

The two main things that are tested is the which of two models (LSTM and CNN) is best for predicting stock price, and also the user experience in using the chatbot.

<u>Chatbot Experience:</u>

For the user testing, a beta version of the chatbot will first be tested with the all the main functionalities, then after each user had finished answering the UEQ they will be asked: “what is the one thing you would include to improve the application?” I’ll then go ahead and attempt to improve the application based off the UEQ result and particularly this specific question.

I developed each model making sure they would both be given the same data when training, as well as keeping things such as the epochs and batch_size the same. In order to determine which model was best I a mix of stocks for a full working week, with my test strategy focusing on:

- **Volatility**: Chose stocks with different volatility levels to see how the chatbot's predictions handle sharp rises and falls in prices.
- **Sector Diversification**: Used stocks from various sectors to test the model's effectiveness across different economic conditions.
- **Market Cap Variation**: Included different market caps ensures the chatbot works well for companies of all sizes.

After exploring tradingview.com below are the 5 I felt met my test strategy criteria:

Company	Ticker	Sector	Market Cap	Volatility	Reason for Selection
Tesla, Inc.	TSLA	Automotive & Energy	Large Cap	High	Tesla is known for its high volatility and market impact due to technological innovations and headline-driven news. It will test the chatbot's ability to handle rapid and large fluctuations in stock prices.
Pfizer Inc.	PFE	Healthcare/Pharmaceuticals	Large Cap	Medium	Pfizer's involvement in high-demand products, including vaccines, can lead to volatility spikes based on clinical trial results and government agreements. This will test the prediction model's responsiveness to industry-specific news.
Square, Inc.	SQ	Financial Technology	Mid Cap	High	As a fintech company, Square experiences different market dynamics, including regulatory impacts and technology adoption rates. Its performance will check the chatbot's effectiveness in sectors at the intersection of technology and finance.
Etsy, Inc.	ETSY	E-commerce	Mid Cap	Moderate to High	Etsy will showcase how well the chatbot performs with consumer-driven stocks that are highly sensitive to market trends and economic conditions.
Chevron Corporation	CVX	Energy	Large Cap	Medium	As a major player in the energy sector, Chevron's stock movements are influenced by

					both global oil prices and geopolitical events. This choice will help evaluate the chatbot's capability in sectors affected by external global factors.

Metrics such as MAE (Mean Absolute Error), MSE (Mean Squared Error), and RMSE (Root Mean Squared Error) will also be used to compare the LSTM and CNN models' ability to forecast stock price using historical data. These calculations provide the numerical basis for comparing the performance of the LSTM and CNN models across different stocks. They will help me in assessing which model predicts stock prices more accurately by looking at how far off the predictions are, on average, from the actual prices. This is crucial for determining the model's effectiveness in real-world trading scenarios, particularly in a tool like a chatbot designed for providing investment advice based on predictive analytics.

The idea is to implement both models into the chatbot so the user can choose to get a prediction from both on the next day stock price if they wish. However, for deciding which model should be used to provide the verdict on if the user should invest in certain stock, will be determined by the model with the most accurate predictions of the stocks above over a full working week.

System Architecture and Design

NLP Architecture:

StockBot integrates a comprehensive NLP architecture and is designed to effectively process and understand user queries through several stages of text processing and response generation.

Text Preprocessing:

Tokenisation, Lemmatisation, and Stemming:

To enhance the understanding of user inputs, text preprocessing techniques such as tokenisation, lemmatisation, and stemming are implemented primarily within the **spell_check** module. These processes standardise text input by reducing words to their base or root forms, which helps in minimising the variability between words that have similar meanings.

- **Tokenisation** splits text into individual elements or tokens.
- **Lemmatisation** simplifies words to their lemma or dictionary form.
- **Stemming** reduces words to their stem or root form.

These techniques are encapsulated in functions like **token_stemming**, **token_lemmatisation**, and **text_preprocessing**, enhancing the chatbot's ability to interpret and process user input accurately.

Text Representation:

TF-IDF Vectorisation:

Textual data within StockBot is transformed into numerical feature vectors using Term Frequency-Inverse Document Frequency (TF-IDF) vectorisation. This method is critical in the **name_management**, **QA**, and **small_talk** modules, helping the system to understand and categorise user queries based on the significance of each word within the dataset and queries.

- Functions like **name_response**, **answer_Q**, and **talk_response** utilise TF-IDF to convert text into a format that can be effectively analysed for similarity.

Similarity Calculation:

Cosine Similarity:

To determine the relevance of user queries to possible responses, StockBot employs cosine similarity. This metric measures the cosine of the angle between the TF-IDF vectors of the user's query and the possible responses stored in the system, facilitating accurate intent recognition.

- The **pairwise_distances** function is employed to compute these similarities, thereby identifying the best-matched responses to user queries.

Spell Correction:

Frequency-Based Correction:

Recognizing and correcting spelling errors is vital for effective communication. StockBot uses a frequency-based correction system that relies on a pre-trained model of word frequencies to suggest the most likely correction for misspelled words.

- Implemented through functions like **train** and **correct**, this system enhances the chatbot's ability to understand misspelled inputs by referencing a frequency-based dataset.

Conversational Design

StockBot is designed to engage users in a dynamic and context-aware conversation, adhering to established principles of conversational design:

Intent-Based Responses

The system is structured to recognise specific intents from user inputs, such as "checking balance," "buying stocks," or "predicting stock prices." This functionality is supported by the main chatbot interface (**chatbot_window**) and specialised functions like **get_investment_advice**, which directly respond to user queries based on identified intents.

User-Centric Design

StockBot emphasises personalised interaction, using functions like **name_change** and **name_response** to adjust responses based on the user's previous inputs and preferences, enhancing the personalised feel of the conversation.

Turn-Taking

The conversational UI is designed to mimic human-like interaction patterns, using the PySimpleGUI package to facilitate a clear and effective exchange of turns between the user and the chatbot.

Feedback Mechanism

Feedback to user actions is a critical component, provided through real-time updates on actions like wallet transactions or stock purchases, utilising functions such as **add_to_wallet** and **buy_stock**.

Implementation Details

Name Management Module

This module handles personalisation features such as recognising and responding to user name change requests. It leverages small datasets and TF-IDF vectorisation to detect and respond to these requests.

Question Answering (QA) Module
StockBot's QA module addresses predefined queries by matching user questions with stored answers, employing TF-IDF vectorization and cosine similarity to ensure relevant responses.

Small Talk Module
Handles informal interactions, enhancing user engagement and maintaining the flow of conversation through predefined responses.

Spell Check Module
This module ensures that user inputs are correctly interpreted by correcting spelling errors using a frequency-based approach.

Sentiment Analysis Module
By analysing the sentiment of financial news articles, this module informs the chatbot's financial advice, adding a layer of market sentiment analysis to its predictions and recommendations.

StockBot Main Code
Integrates all modules, providing a cohesive system that handles user interactions, processes queries, and delivers financial advice and stock market predictions through a user-friendly chat interface. This comprehensive integration ensures that StockBot operates efficiently as an autonomous agent capable of aiding users in navigating the complexities of stock market investments.

Results and Analysis
Link to LSTM & CNN test results

LSTM V CNN

Analysis Between LSTM and CNN Models
The dataset created provides a detailed comparison of predictive accuracies for LSTM and CNN models across a selection of stocks that span various sectors, including automotive, healthcare, financial technology, e-commerce, and energy. The summary of results highlighting key differences in how each model performs, reflecting their respective capabilities in interpreting and forecasting market data dynamics. Upon reviewing the results, it is evident that the CNN model consistently exhibits lower Mean Squared Error (MSE) values across the majority of the stocks when compared to the LSTM model. This suggests that CNN may be more adept at capturing and predicting non-linear patterns in stock price movements, potentially due to its inherent strengths in processing spatial and temporal data.

Model Performance Across Different Stocks and Sectors
- **Tesla, Inc. (TSLA)**: The CNN model outperforms LSTM with a significantly lower MSE, indicating better prediction capability in the volatile automotive and energy sector.
- **Pfizer Inc. (PFE)**: LSTM shows a slightly better performance in terms of MSE, suggesting it handles less volatile stocks like pharmaceuticals more effectively than CNN.
- **Square, Inc. (SQ)** and **Etsy, Inc. (ETSY)**: In the financial technology and e-commerce sectors, which are known for their high volatility, CNN again shows superior performance, handling rapid market changes better.

- **Chevron Corp. (CVX)**: In the energy sector, both models perform similarly, but CNN has a slight edge.

Sector-Based and Volatility-Based Analysis

- **High Volatility Stocks**: In high volatility environments (e.g., Tesla and Square), CNN seems to provide more accurate forecasts, suggesting its better suitability for stocks that experience large price fluctuations.
- **Moderate to Low Volatility Stocks**: LSTM performs comparably well with less volatile stocks (e.g., Pfizer), indicating its efficiency in more stable environments.

From the analysis, CNN generally exhibits superior predictive performance across most stocks, particularly in sectors characterised by high volatility. The ability of CNN to minimise prediction errors across diverse market conditions underscores its robustness as a tool for financial forecasting within AI-driven chatbots. However, LSTM still shows competitive performance, especially in sectors with lower volatility, highlighting its utility in stable market predictions. Due to the outcome of the results CNN is the Model used in the chatbot for giving user advice.

Actual price	LSTM	CNN	Date test Ran	Company	LSTM Line picture	LSTM box plot picture	CNN line Graph pictu	CNN box Plot picture	LSTM MAE	LSTM MSE	LSTM RMSE	CNN MAE	CNN MSE	CNN RMSE
142.05	150.44	156.64	23/04/2024	Tesla, Inc. (TSLA) - Automotive & Energy					10.65	228.84	15.13	13.46	325.46	18.04
26.26	26.44	26.27	23/04/2024	Pfizer Inc. (PFE) - Healthcare/Pharmaceuticals					1.03	4.36	2.09	1.03	4.19	2.05
71.6	75.81	73.82	23/04/2024	Square, Inc. (SQ) - Financial Technology					5.82	80.43	8.97	5.82	87.38	9.35
66.45	66.97	66.79	23/04/2024	Etsy, Inc. (ETSY) - E-commerce					5.67	117.08	10.82	6.31	130.11	11.41
161.92	158.94	160.14	23/04/2024	Chevron Corporation (CVX) - Energy					3.48	23.78	4.88	3.31	20.88	4.57
144.68	140.39	151.99	24/04/2024	Tesla, Inc. (TSLA) - Automotive & Energy					17.18	493.13	22.21	10.87	242.02	15.56
26.32	26.11	26.52	24/04/2024	Pfizer Inc. (PFE) - Healthcare/Pharmaceuticals					0.92	3.89	1.97	1.08	4.58	2.14
75.21	70.59	71.17	24/04/2024	Square, Inc. (SQ) - Financial Technology					5.34	75.69	8.7	4.79	66.08	8.13
68.35	67.85	63.29	24/04/2024	Etsy, Inc. (ETSY) - E-commerce					5.92	116.61	10.8	6.85	130.24	11.41
162.92	158.63	161.56	24/04/2024	Chevron Corporation (CVX) - Energy					3.8	25.41	5.04	3.13	19.07	4.37
162.13	157.64	168.07	25/04/2024	Tesla, Inc. (TSLA) - Automotive & Energy					11.55	246.19	15.69	11.85	281.14	16.77
26.27	25.63	27.3	25/04/2024	Pfizer Inc. (PFE) - Healthcare/Pharmaceuticals					0.91	4.04	2.01	1.35	4.72	2.17
72.03	74.43	73.9	25/04/2024	Square, Inc. (SQ) - Financial Technology					5.8	86.79	9.32	5.36	75.88	8.71
66.87	65.29	65.13	25/04/2024	Etsy, Inc. (ETSY) - E-commerce					6.05	124.26	11.15	5.36	105.61	10.28
162.85	163.56	164.41	25/04/2024	Chevron Corporation (CVX) - Energy					3.51	24.36	4.94	3.79	26.56	2.15
170.18	150.14	173.2	26/04/2024	Tesla, Inc. (TSLA) - Automotive & Energy					13.68	347.31	18.64	11.61	270.79	16.46
25.26	26.17	26.54	26/04/2024	Pfizer Inc. (PFE) - Healthcare/Pharmaceuticals					1.02	4.33	2.08	1.01	4.09	2.02
74.48	72.88	72.32	26/04/2024	Square, Inc. (SQ) - Financial Technology					5.05	72.53	8.28	4.81	68.52	8.28
67.2	67.7	66.03	26/04/2024	Etsy, Inc. (ETSY) - E-commerce					6.05	120.11	10.96	6.26	139.91	11.83
165.28	173.59	162.79	26/04/2024	Chevron Corporation (CVX) - Energy					3.46	22.96	4.79	3.32	20.66	4.55
194.05	167.42	197.51	30/04/2024	Tesla, Inc. (TSLA) - Automotive & Energy					15.08	406.48	20.16	11.56	269.45	16.41
25.64	25.52	26.35	30/04/2024	Pfizer Inc. (PFE) - Healthcare/Pharmaceuticals					0.92	4.09	2.02	1.34	4.79	2.19
75.31	70.62	74.07	30/04/2024	Square, Inc. (SQ) - Financial Technology					5.47	77.63	8.21	4.79	67.36	8.21
68.88	65.83	59.17	30/04/2024	Etsy, Inc. (ETSY) - E-commerce					5.81	118.18	10.87	10.63	207.33	14.4
166.33	164.27	169.57	30/04/2024	Chevron Corporation (CVX) - Energy					3.57	22.78	4.77	6.66	62.84	7.93

UEQ Analysis

11 people participated in the test taking the test twice. The first time they tested the beta version, filling out the UEQ form and commenting on the one thing they would change. Once this was filled I attempted to improve the chatbot based on their suggestions.

The User Experience Questionnaire (UEQ) measures: Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation and Novelty.

1st UEQ test

[Link to 1st UEQ test excel file]

Attractiveness

Initial scores for attractiveness were moderate, suggesting that while users found the chatbot somewhat appealing, there were areas for improvement in visual design or initial user engagement. Comments indicated that users liked the concept but felt the interface could be more engaging or visually appealing.

Perspicuity

Lower scores in this dimension indicated that users found the chatbot somewhat difficult to understand or navigate initially. This could be due to complex instructions, a steep learning curve, or unclear functionalities that hindered immediate ease of use.

Efficiency

The mixed reviews on efficiency highlighted a split in user experiences, where some found the chatbot responsive and quick, while others noted delays or cumbersome interactions. This variability suggested the need for optimisation to ensure consistent performance across all user interactions.

Dependability

Average dependability scores indicated a moderate level of trust and reliability in the chatbot's responses and functionalities. Users felt that while the chatbot generally performed well, there were occasional inconsistencies or errors that affected their trust.

Stimulation

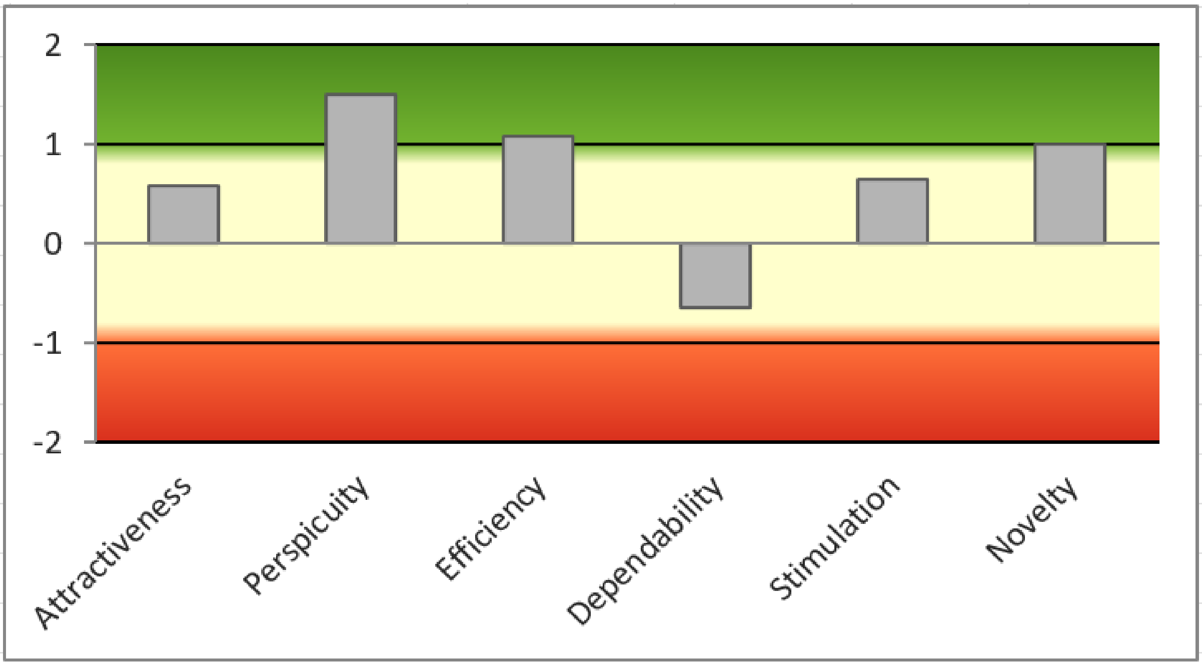
Initially, the chatbot was found to be somewhat stimulating, but it didn't particularly excite or engage users beyond basic interactions. This suggested a need for more dynamic features or content that could capture and hold user interest.

Novelty

The moderate novelty scores indicated that while the chatbot introduced some new features or approaches, it wasn't seen as highly innovative by users. This may have been due to users having prior experiences with similar technologies that set a baseline expectation.

These results showed that although there are clear signs of improvement I was overall heading in the right direction.

Items																									
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
3	5	3	2	4	4	6	5	2	3	4	3	5	5	4	4	6	4	5	4	3	5	4	3	2	6
4	5	3	2	4	5	6	5	1	2	5	4	5	5	6	5	6	4	6	5	2	5	3	3	1	5
4	5	3	2	4	5	6	5	2	3	4	4	5	5	5	5	6	4	6	4	3	5	3	4	3	6
4	6	3	2	3	4	6	5	1	3	4	3	5	5	5	4	6	4	6	4	3	5	3	3	3	4
5	5	3	2	5	4	6	5	1	2	4	4	5	4	5	5	6	4	5	4	3	4	3	3	3	4
5	5	3	2	3	5	5	5	2	2	4	4	5	4	5	5	7	3	5	4	2	4	4	4	3	5
3	5	4	2	4	4	6	5	1	2	4	4	5	4	5	5	7	3	5	5	2	6	3	4	1	5
4	5	4	2	4	4	6	6	1	2	4	3	5	5	5	4	7	3	5	4	2	4	4	3	4	5
3	6	3	1	5	5	6	5	2	3	4	3	5	5	4	4	7	4	6	4	3	6	3	3	1	5
4	6	4	1	5	5	6	6	1	3	5	3	5	5	4	4	7	4	6	5	2	4	4	4	6	5
4	6	4	1	4	5	6	6	1	3	4	3	5	5	6	5	6	3	6	5	2	6	4	4	6	4



Feedback Analysis and Implemented Changes

I made user feedback a primary focus as I feel user feedback is not just beneficial but crucial for aligning chatbot technology with user needs and expectations, driving continuous improvement. Below are the users answers to the question.

Question: what is the one thing you would include to improve the application?

Answer
Maybe a function where users can add money to a balance
Make a user interface
Make a GUI
Would it be possible for the user to get investment advice and buy stocks
Make it easier to read
Make the help stuff more informative
Improve the interface
Make it clearer what the chatbot is for
Make a GUI
Have the user able to buy stocks and see how much stock they own
Make it easier to read

Question feedback breakdown::

1. **Real-time Stock Price Updates:** Users expressed a need for more immediate and reliable stock price information.
2. **Enhanced Financial Transactions:** Feedback indicated that the process for adding money to wallets and buying stocks needed simplification and error reduction.
3. **Improved Prediction Accuracy:** Users sought more accurate and understandable predictions for stock prices.
4. **Better Interaction and Responsiveness:** There was a demand for a more engaging and responsive conversational interface.

Changes Made:

1. **Updated Stock Price Retrieval:** The **get_current_stock_price** function was optimized to ensure more reliable and timely retrieval of stock prices using direct calls to Yahoo Finance APIs, which provide up-to-date market data.
2. **Streamlined Financial Transactions:** The functions **add_to_wallet** and **buy_stock** was refined to handle user inputs more robustly, including better error handling and user input validation. This prevents common mistakes such as entering non-numeric values and provides clearer feedback to the user about the transaction status.
3. **Enhanced Prediction Models:** Based on feedback, the prediction models were refined. The **predict_stock_price** function was updated to include both LSTM and CNN models, allowing users to choose the model they prefer or compare results between the two. This was aimed at improving prediction accuracy and user trust in the system's outputs.
4. **Conversational Interface Improvements:** The conversational logic in **chatbot_window** was enhanced to provide more dynamic and context-aware responses. Natural Language Processing (NLP) capabilities were also expanded to better recognise and respond to a wider range of user intents, improving the overall conversational flow.

2^{nd} UEQ test

Link to 2^{nd} UEQ test excel sheet.

I believe the changes implemented in the chatbot aimed to address specific user concerns and improve overall functionality, likely contributing to the enhanced UEQ scores observed in the second test:

Impact on Efficiency and Dependability

By optimising the retrieval of stock prices through Yahoo Finance APIs, the chatbot became more efficient in delivering real-time data. This likely contributed to higher scores in efficiency as users experienced quicker responses and more up-to-date information. Improved reliability in data retrieval also enhanced the dependability score, as users could trust the chatbot to provide accurate market data consistently.

Impact on Efficiency and Perspicuity

Refining the financial transaction functions to handle inputs more robustly reduced the likelihood of errors during interactions such as adding funds to wallets or purchasing stocks. Better error handling and clear feedback mechanisms likely improved the perspicuity score, as users found the system easier to understand and interact with. Additionally, these changes enhanced efficiency by minimising delays and frustrations associated with transaction errors.

Impact on Dependability and Stimulation

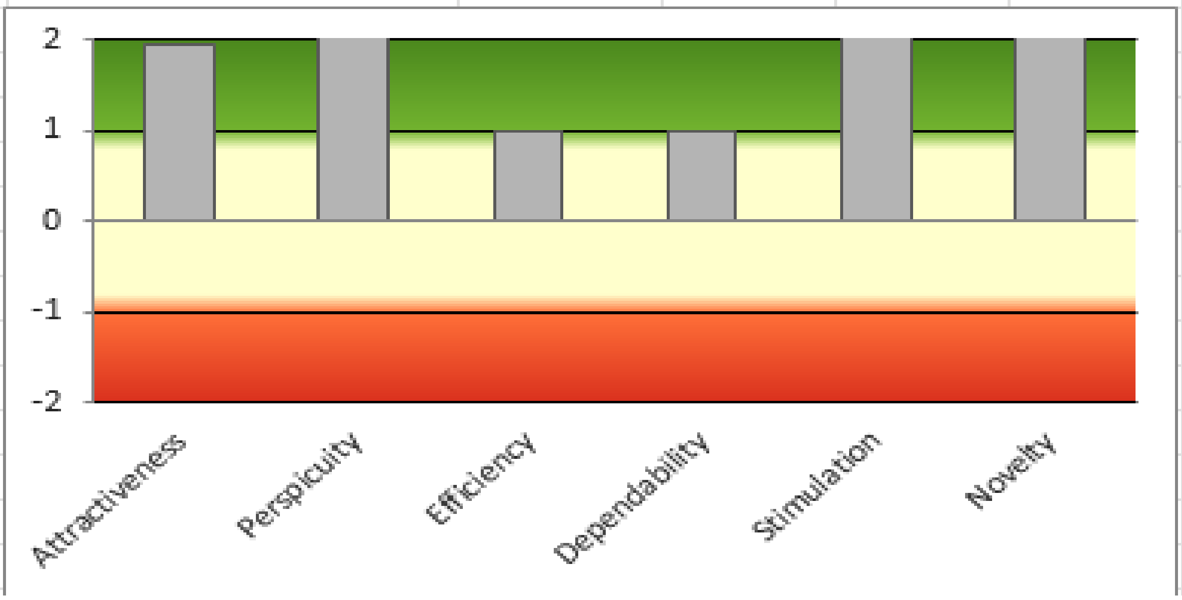
Allowing users to choose between LSTM and CNN models or to compare their outputs directly empowered users and may have increased engagement and interest (stimulation score), as they could actively participate in the analysis process. Improvements in the prediction models likely bolstered the dependability score as well, by providing more accurate and trustworthy forecasts that users could rely on for making investment decisions.

Impact on Perspicuity and Novelty

Enhancing the NLP capabilities and refining the conversational logic ensured that the chatbot could understand and respond to a broader range of user intents more accurately. This improvement in understanding user queries without misinterpretation boosted the perspicuity score, as interactions became more intuitive and less prone to errors. Moreover, advanced NLP features likely contributed to a higher novelty score, making the chatbot appear more innovative and engaging to users.

Overall, these targeted enhancements directly addressed the initial shortcomings identified in the first UEQ test, leading to improvements across all UEQ dimensions in the second test. By focusing on areas crucial for user satisfaction and operational efficiency, the revised chatbot offered a more refined, user-friendly, and reliable tool for financial analysis and decision-making.

Items																									
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
5	6	2	1	3	6	6	6	2	2	6	2	6	7	5	5	6	3	1	5	2	6	6	2	2	6
6	6	1	2	2	6	6	7	2	2	6	1	6	7	6	5	6	2	2	5	2	6	6	3	2	6
6	7	2	2	3	6	6	6	2	2	5	2	6	6	6	6	6	2	2	6	2	6	6	2	2	5
5	6	1	2	3	6	6	6	1	1	6	2	6	6	5	6	6	3	2	5	2	6	5	2	2	5
6	6	1	2	3	6	6	6	1	2	6	2	6	6	5	6	6	2	2	5	2	6	6	2	2	6
6	6	2	1	2	6	5	6	2	2	6	2	7	6	6	6	7	2	1	6	2	6	6	3	2	6
6	7	1	1	2	6	6	7	1	2	7	1	6	6	6	5	7	2	2	5	2	6	6	2	2	6
5	7	2	2	3	7	6	7	1	1	6	2	6	7	5	5	7	2	2	6	2	6	6	2	3	6
5	6	1	1	2	76	6	6	2	1	6	2	7	7	6	6	7	2	2	6	2	6	6	3	1	5
6	6	1	1	3	6	6	6	2	2	5	1	6	7	6	6	7	2	1	5	2	6	5	2	2	5
6	6	2	1	2	6	6	6	1	1	6	2	6	6	6	6	6	3	2	6	2	6	6	2	3	6



Investment Advice Analysis:

[Click to see video of chatbot in progress]

The quality and relevance of the investment advice provided by the chatbot were evaluated. This involved examining how the integration of sentiment analysis from financial news impacted the advice given, such as "buy," "sell," or "hold" recommendations.

Relevance and Quality of Advice Provided by the Chatbot

The chatbot, designed to assist users in making informed investment decisions, leverages advanced predictive analytics and natural language processing to provide relevant and actionable financial advice. The quality of the investment advice hinges on its accuracy, timeliness, and relevance to the user's financial goals and market conditions. To ensure this, the chatbot incorporates real-time data fetching and a sophisticated analysis mechanism that evaluates current market trends and historical data patterns.

Integration of Sentiment Analysis into Investment Recommendations

The integration of sentiment analysis into the chatbot's investment advice framework is a critical component that enhances its ability to deliver nuanced advice. Sentiment analysis is applied to financial news articles, providing the chatbot with a layer of qualitative data that complements its quantitative algorithmic predictions. This allows the chatbot to gauge market sentiment, which can be pivotal during periods of high volatility or in situations where market sentiment is significantly influenced by news events or public opinion.

Use of TextBlob for Sentiment Analysis

TextBlob, a popular Python library for processing textual data, is employed for its straightforward and effective sentiment analysis capabilities. TextBlob offers out-of-the-box support for sentiment analysis by providing a polarity score that ranges from -1 (very negative) to 1 (very positive), which helps in assessing the general sentiment of news articles concerning stocks or the financial market. This sentiment score is instrumental in refining the chatbot's financial advice, especially in discerning whether a stock is likely to experience positive or negative movement based on public sentiment.

Utilisation of News API for Fetching Relevant Articles

The News API is a vital tool used by the chatbot to fetch relevant news articles about specific companies or the broader financial market. This API allows the chatbot to access a wide array of news sources in real-time, ensuring that the sentiment analysis is based on the most current and relevant information available. By integrating the News API, the chatbot can systematically retrieve, parse, and analyse news content that could impact stock prices, providing users with insights that are not only based on historical data but also on current market events and trends.

Analysis Function: analyze_sentiment(headlines)

The **analyze_sentiment** function processes the headlines retrieved from the News API through TextBlob to assign a sentiment score to each article. Here's a breakdown of how this function contributes to investment advice:

By combining sentiment analysis with traditional financial metrics and predictive analytics, the chatbot provides a well-rounded investment advice service. This approach ensures that the advice is not only grounded in solid data but also context-aware, considering the mood and trends that can significantly influence market movements.

Conclusion

This study on AI-driven chatbots for stock price prediction demonstrates their potential in utilising machine learning and NLP to enhance financial decision-making. The comparative analysis showed CNN models generally outperformed LSTM in accuracy, particularly in volatile markets. Integrating sentiment analysis using TextBlob and the News API added depth to the investment advice by considering market sentiment, enhancing the chatbot's relevance and responsiveness to current events.

User feedback from the UEQ was instrumental in refining the chatbot's functionalities, leading to significant improvements in user interaction and satisfaction in subsequent tests. These enhancements confirmed the chatbot's capability to provide timely and relevant financial advice, thereby improving its utility as a financial advisory tool.

Moving forward, the project could expand by integrating more diverse data sources and advanced analytics to enhance predictive accuracy further. Exploring additional machine learning models and deep learning techniques could offer insights into complex pattern recognition and forecasting. I would also extend user study and get users opinion again on what they would like to improve. These steps would ensure continuous improvement and adaptation of the chatbot to meet evolving user needs and market conditions effectively.

References

Andreas Hinderks, M. S. J. T., 2008. *User Experience Questionnaire.* [Online]
Available at: https://www.ueq-online.org/
[Accessed 2024].

Benson, A., 2024. *Betterment Review 2024: Pros, Cons and How It Compares.* [Online]
Available at: https://www.nerdwallet.com/blog/investing/betterment-review/
[Accessed 2024].

Hofstra, 2023. *Where can I find historical stock prices for stock tickers no longer traded?.* [Online]
Available at:
https://libanswers.hofstra.edu/faq/128721#:~:text=Additionally%2C%20a%20web%20resource%20for,%2C%20weekly%2C%20or%20monthly%20data.
[Accessed 2024].

networks, S. m. p. m. p. w. L. n., 2017. *David M. Q. Nelson, Adriano C. M. Pereira, Renato A. de Oliveira.* [Online]
Available at: https://ieeexplore.ieee.org/document/7966019
[Accessed 2024].

Schrepp, D. M., n.d. *User Experience Questionnaire.* [Online]
Available at: https://www.ueq-online.org/Material/Handbook.pdf
[Accessed 15 03 2024].

Stein, 2018. *Is Wealthfront Worth It? Review From A Finance Nerd....* [Online]
Available at: https://www.theadvisorcoach.com/wealthfront-review.html
[Accessed 2024].

Taewook Kim, H. Y. K., 2019. *Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data.* [Online]
Available at: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0212320#sec018
[Accessed 2024].

Thomas Fischer, C. K., 2018. *Deep learning with long short-term memory networks for financial market predictions.* [Online]
Available at: https://doi.org/10.1016/j.ejor.2017.11.054
[Accessed 2024].

Xiao Zhong, D. E., 2019. *Forecasting daily stock market return using dimensionality reduction.* [Online]
Available at:
https://www.sciencedirect.com/science/article/abs/pii/S0957417416305115?via%3Dihub
[Accessed 2024].