Faculty of Natural and Mathematical Sciences
Department of Informatics

King's College London
Strand Campus, London,
United Kingdom

7CCSMPRJ

Individual Project Submission 2022/23

Name:	Yu-Hsun Wang
Student Number:	21150305
Email Address:	yu-hsun.wang@kcl.ac.uk
Degree Programme:	MSc in Computational Finance
Project Title:	Consumer Sentiment towards Various Apple Products in Different Regions and its Impact on Apple's Stock Price
Supervisor:	Peter McBurney
Word Count:	13647

RELEASE OF PROJECT

Following the submission of your project, the Department would like to make it publicly available via the library electronic resources. You will retain copyright of the project.

☑ I **agree** to the release of my project

☐ I **do not** agree to the release of my project

Signature: Yu-Hsun Wang **Date:** August 14, 2023

Department of Informatics
King's College London
United Kingdom

7CCSMPRJ Individual Project

Consumer Sentiment towards Various Apple Products in Different Regions and its Impact on Apple's Stock Price

Name: **Yu-Hsun Wang**
Student Number: 21150305
Course: MSc in Computational Finance

Supervisor: Peter McBurney

This dissertation is submitted for the degree of MSc in Computational Finance.

Acknowledgements

My sincere appreciation goes out to Professor Peter McBurney, my supervisor, for his unwavering support, astute observations, and proficient direction that proved essential in ensuring the success of this project. His expertise and persistence have been crucial in facilitating the achievement of my research objectives.

I extend my sincere appreciation to the creators and maintainers of both CardiffNLP and MacBERT, whose crucial involvement in the analysis of data for this project has been invaluable. Their noteworthy contributions to the field have facilitated the success of my work.

I would like to express my appreciation towards Reddit for furnishing a platform that was crucial in acquiring the necessary data for my research. Additionally, I sincerely thank Reddit's lucid and comprehensive regulations that ensured the preservation of ethical standards throughout this study.

Abstract

This research aims to scrutinise the multifaceted impact of consumer sentiment on the stock performance of Apple Inc. It harnesses the sentiment data mined from Reddit, a popular social media platform, and examines its correlation with the market performance of Apple. Furthermore, this investigation seeks to elucidate the relationships between the stock performance, revenue proportions of Apple's diverse product portfolio across different regions, and their associated sentiment scores. The primary objective of this study is to unravel the underlying influences of public sentiment and regional product performance on a globally recognised technology company's stock prices. It endeavours to aid in refining investment strategies by shedding light on these relationships.

The methodology employed in this study is two-fold. Firstly, it utilises Natural Language Processing (NLP) techniques to dissect a sizable dataset of Reddit posts centred around Apple products. Each post is assigned a sentiment score, essentially converting qualitative data into quantitative data that can be analysed statistically. Secondly, it incorporates an examination of Apple Inc.'s historical stock prices alongside data related to the revenue proportions generated by various Apple products across different regions.

In order to study the correlations and causal relationships between the gathered variables, a range of advanced analytical techniques are used, including but not limited to Granger causality tests. This approach allows for the investigation of potential causal relationships between the variables, contributing to a more robust understanding of the factors influencing Apple's stock prices.

Through this research, it is anticipated to make several critical contributions to the academic field. Firstly, it aims to provide an enriched understanding of how social media sentiment regarding specific Apple products and their associated regional revenue proportions affect Apple's stock price. Secondly, this research intends to contribute to the validation and further application of NLP techniques within sentiment analysis. This enhances the methodological repertoire available for sentiment analysis and underscores the value of incorporating social media data into financial market research. The findings generated from this study have the potential to offer a more nuanced comprehension of the intricate relationship between consumer sentiment and stock performance, particularly within the context of a multinational technology company like Apple. This, in turn, can guide future academic inquiries and offer valuable insights to stakeholders in the investment domain.

Contents

1 Introduction

Social media platforms like Reddit have become hotbeds for real-time consumer opinion sharing in our globalised world. These platforms significantly influence financial markets, especially for globally influential companies. This research focuses on Apple Inc., a dominant player whose stocks are consistently prominent in the U.S. stock market. The focal point of this research is to unearth any potential correlations between the sentiment expressed towards Apple's diverse product range on Reddit and the fluctuations observed in Apple's stock price. Not content with studying this dynamic in isolation, the research intends to weave in an additional layer of complexity by considering the revenue proportions generated by different Apple products in various geographical markets. This step aims to create a more nuanced picture of regional influences on the performance of Apple's stock.

The anticipated findings from this research are expected to give investors a more profound understanding of the influences and repercussions of social media trends on their investment strategies. Furthermore, by offering a fresh perspective on how regional revenue proportions and consumer sentiment on platforms like Reddit can sway the fortunes of companies like Apple, this study hopes to broaden the discourse around investment decision-making in the digital age. Ultimately, this project aims to untangle the intricate correlations between factors impacting Apple Inc.'s stock performance. The aspiration is that the knowledge acquired through this undertaking will equip investors with a comprehensive array of resources to effectively deal with the difficulties associated with making investment decisions about tech corporations of global significance.

1.1 Motivation

Understanding how consumer preferences and trends affect the financial market was the driving force behind this project. Social media platforms like Reddit have transformed how information is disseminated and become virtual channels for disseminating information and shaping public opinion. Analysing sentiments from such platforms can furnish critical insights into consumer preferences and inclinations. These insights are integral to understanding the dynamic and swift nature of the financial market, which often hinges on individual behavioural and psychological factors. Investors and traders in the financial market are essential participants in a vast game theory playground, where the opinions and emotions of others significantly sway their actions. During reinforcement learning combined with NLP that analyses environments can capture overall market sentiment, it may neglect specific product-related sentiments that can also shape stock prices.

This research focuses on Apple Inc., a global powerhouse whose stock remains among the most actively traded securities in the U.S. stock market. The company's vast array of products stirs conversations worldwide, offering a rich data source for sentiment analysis. Additionally, the regional revenue proportions generated by these different products serve as crucial economic indicators. With this in mind, this study aims to investigate the intricate web of factors influencing stock prices, unravelling the relationships between Apple's stock price fluctuations, public sentiment towards their products, and the revenue proportions these products generate in various regions. The ultimate objective is to provide a comprehensive understanding of the multiple facets influencing Apple's stock performance, offering investors a more holistic perspective for informed decision-making.

1.2 Scope

This dissertation aims to carry out a thorough investigation into the correlation between consumer opinions toward the products articulated on social networking sites and their influence on the stock price performance of Apple Inc. This study concentrates mainly on scrutinising the sentiments conveyed by English and Chinese communities using the widely used international social media platform known as Reddit. Through a detailed sentiment analysis of Apple's wide-ranging product line, classified as durable and non-durable goods, this research enables a comparative analysis of the impact of these sentiments on Apple's stock price.

The first stage of the analysis will involve the segregation of sentiments into non-durable and durable goods categories. The investigation will examine how sentiments related to these different product categories and the overall sentiment from all comments, respectively, impact Apple's stock price. This nuanced categorisation will facilitate a deeper understanding of distinct consumer behaviours and their potential influence on Apple's financial performance in the stock market.

In the second stage, comments will be further classified based on different Apple products. The analysis will consider Apple's revenue percentages for each product as an index, thereby ensuring a weighted examination of sentiments' impact on Apple's stock performance. This approach will yield insights into how sentiments associated with each Apple product can sway the company's stock price.

In the final stage, sentiments from Reddit's Chinese communities will be incorporated, with a particular emphasis on the Greater China region. The sentiment analysis will be indexed against Apple's regional revenue contribution from the Greater China area, ensuring a representative depiction of sentiments from this vital market. Concurrently, the sentiment analysis for the English comments will be adjusted based on the regional revenue contributions of North America and Europe to Apple Inc., reinforcing the accuracy of the study. This approach aspires to determine whether a more precise correlation can be identified between regional sentiment analysis and stock prices when the corresponding region's revenue contributions are taken into account.

By delving into these various aspects, this dissertation intends to offer a detailed analysis of the relationships between Reddit-based sentiments, product categorisations, and regional sentiments and their collective influence on Apple Inc.'s stock performance. The goal is to contribute to a holistic understanding of the elements affecting the stock prices of multinational corporations like Apple. These insights could prove to be a valuable resource for investors, assisting them in making well-informed decisions regarding their investments in Apple.

1.3 Aims

This research project embarks on a journey to investigate the connection between Reddit sentiment concerning Apple products, classified into durable and nondurable goods, and the implications it has on Apple Inc.'s stock price. By dissecting this relationship, the study aims to unravel how varying sentiments in different product categories can impact the company's stock market performance. Building on this sentiment analysis, the project will delve into a comparative analysis of how sentiments associated with different Apple products, and each accounted for with their respective revenue contributions, sway the company's stock prices. This novel approach aims to offer a unique perspective on how sentiment related to specific products can potentially influence stock prices.

In addition to product-specific sentiment analysis, the project extends its reach to understand the influence of regional sentiments on Apple Inc.'s stock price, particularly from the Greater China region. This will involve integrating sentiments from Reddit's Chinese communities, indexed against Apple's regional revenue contribution from the Greater China area. It also adjusts the sentiment analysis for the English comments based on the regional revenue contributions of North America and Europe to Apple Inc. The objective is to explore if a more precise correlation between regional sentiment analysis and stock prices can be established when considering regional revenue contributions.

Akin to how reinforcement learning has incorporated NLP methodologies for developing tactics at specific time intervals, this research project seeks to demonstrate that product-specific and region-specific sentiment analysis can also be considered in RL learning procedures. Ultimately, this project aims to contribute to the existing knowledge pool by offering a fresh perspective on stock market investments, emphasising the potential influence of product-specific and region-specific sentiments on stock performance. The ambition is to present a more comprehensive and insightful view of investment strategies, thereby aiding investors in making more informed decisions.

1.4 Objectives

The main objectives of the project's development are:

- Conduct research on available academic papers and journals regarding NLP in order to acquire an in-depth understanding of deriving sentiment scores from both English and Chinese textual data.

- Accumulate a comprehensive dataset of Reddit comments related to Apple products, with a focus on English and Chinese communities, and preprocess/classify these comments for sentiment analysis.
- Utilise NLP techniques to perform a sentiment analysis on the collected dataset. This analysis will involve evaluating the sentiment scores of Reddit comments and categorising them into durable and non-durable goods. The sentiments will also be classified based on different Apple products.
- Gather historical stock price data of Apple Inc. and regional revenue contribution data, particularly from the Greater China region, North America, and Europe. These data will be employed to adjust sentiment scores based on regional and product revenue contributions.
- Conduct a robust regression analysis to identify and compare correlations between the sentiment scores (overall, product-specific, and region-specific) and Apple's stock prices. This analysis will also investigate the impact of sentiments from different product categories (durable and non-durable goods) on Apple's stock performance.
- Aim to uncover a more accurate correlation between region-specific sentiment analysis and stock prices when regional revenue contributions are factored in. This objective extends to understanding the differentiated impacts of sentiments related to different Apple products, each considered with their respective revenue percentages.
- Through this thorough analysis, strive to provide valuable insights into investment strategies, aiding investors in making well-informed decisions regarding their investments in Apple.

1.5 Report Structure

The arrangement of this study's findings is presented through the subsequent sections:

- Chapter 2: Background – This chapter provides a firm grounding in the fundamentals of Natural Language Processing (NLP), emphasising sentiment analysis conducted on Reddit and its potential impact on stock market performance. It also explores the significance of different product categories, specifically durable and non-durable goods, and the role they play in shaping public sentiment. In addition, the chapter underlines Apple Inc.'s market positioning and the importance of regional contributions to its revenue, focusing on the Greater China region, North America, and Europe.

- Chapter 3: Related Work – This section reviews scholarly literature relevant to the project's scope. It investigates research on sentiment analysis applied to social media platforms and its effect on the stock market. The chapter also evaluates research studies investigating an appropriate model for analysing the current language used on social media platforms. Moreover, given the unique linguistic traits of the Chinese language, an appropriate sentiment analysis model will also be examined. Lastly, an investigation into fundamental techniques for uncovering correlations and creating a straightforward trading strategy based on the findings will be conducted.

- Chapter 4: Approach – This chapter elucidates the methodology employed in the study, commencing with data collection from Reddit and how the dataset is prepared and classified. The process of conducting a comprehensive sentiment analysis using advanced NLP techniques is detailed, with particular emphasis on classifying sentiments into product categories and regional sentiment analysis. Additionally, it elaborates on the process of acquiring and adjusting Apple's historical stock price data and regional revenue contribution data. The chapter also thoroughly explains the regression analysis techniques used to identify and compare correlations and the backtest method.

- Chapter 5: Results and Analysis – This section presents the findings of the study. It delves into the regression analysis results, exploring the correlation between overall and product-specific sentiment scores and Apple Inc.'s stock prices. The chapter also outlines the differential impacts of sentiments associated with various product categories on Apple's stock performance. Moreover, the section assesses the relationship between region-specific sentiment analysis and Apple's stock price when considering regional revenue contributions. These findings provide insights into the intertwined relationships between Reddit-based sentiments, product categorisations, regional sentiments, and their collective influence on Apple's stock performance.

- Chapter 6: Legal, Social, Ethical and Professional Issues – presents a rational discourse concerning ethical, legal, social, and professional concerns in relation to the project issue. The chapter additionally serves as evidence of my familiarity with the Code of Conduct & Code of Good Practice published by the British Computer Society (BSC).

- Chapter 7: Conclusion – This chapter provides a comprehensive study wrap-up, encapsulating the essential findings and presenting insightful recommendations for investors, analysts, and future research. It recaps the intricate relationships unveiled between Reddit-based sentiments, product categorisations, and regional sentiments and their collective influence on Apple Inc.'s stock performance. It proposes how these findings could guide investment strategies and decision-making processes. Additionally, the chapter acknowledges the limitations encountered in the study, offering suggestions for future research to tackle these challenges. It emphasises the necessity for continued exploration into the interplay between public sentiment, regional factors, product-specific sentiments, and stock prices in an increasingly globalised and digitally connected market landscape.

2 Background

This research aims to utilise NLP, precisely sentiment analysis, to interpret and categorise sentiments expressed in Reddit comments. Sentiment analysis, a subfield of NLP, uses algorithms to identify and classify opinions in text data, thereby ascertaining whether the writer's attitude towards a particular topic or product is positive, negative, or neutral. This process will involve assigning sentiment scores based on the content and subsequently examining the correlation between these sentiment scores and the stock performance of Apple Inc., a leading technology company with products distributed across various regions worldwide. Sentiment analysis provides insights into public opinion, customer sentiment, and market trends. For this research, the application of sentiment analysis would enable a more nuanced understanding of how public sentiment, as expressed on social media platforms like Reddit, could influence the stock market dynamics of a global company like Apple Inc. In doing so, it may also reveal the differential impacts of sentiments surrounding durable and non-durable goods on stock performance, providing potentially valuable insights for investment strategies and stock market forecasts.

In their survey, Medhat, Hassan, and Korashy (2014)[1] provide a comprehensive overview of sentiment analysis techniques, algorithms, and applications. They outline various sentiment analysis levels, including the document, sentence, and aspect levels. Similarly, Liu (2012)[2] describes the crucial step of text preprocessing in NLP. This process involves cleaning and formatting the raw data into an easily interpretable form for the algorithms. Preprocessing techniques often include tokenisation, removing stop words, and stemming.

Increasingly, researchers are leveraging the wealth of data available on social media to predict stock market trends, corroborating the predictive power of public sentiment. For instance, Nguyen, Shirai, and Velcin (2015)[3] applied sentiment analysis techniques to social media data in an attempt to forecast stock market movements. Their findings were illuminating - models incorporating mood information and human sentiment outperformed those solely based on price information in predicting various stock prices[3]. This suggests that social media sentiment can significantly enhance the predictive accuracy of stock market forecasts, lending credence to the fundamental premise of this research project. Their research[3], along with that of Bollen, Mao, and Zeng (2011)[4], reinforces the hypothesis that social media sentiment, specifically on platforms like Reddit, can be instrumental in predicting stock market performance.

These studies, therefore, highlight the significance and timely relevance of the proposed research project. While previous studies have demonstrated the predictive power of social media sentiment on stock market performance, this research aims to delve deeper by examining the potential influence of sentiments specifically tied to different product categories and services Apple Inc offers.

A vital facet of this research project is the exploration of sentiments associated with different categories of Apple Inc.'s products, which fall under durable and non-durable goods. Durable goods are those that have an extended lifespan, such as MacBooks, while non-durable goods are consumed over the short term, like music or movie purchases on iTunes. The importance of differentiating between durable and non-durable goods in sentiment analysis is underscored by Bing Liu's (2012) study [2]. Bing Liu[2] noted that the sentiment elicited by durable goods, such as excitement or speculation triggered by the release of a new iPhone model, could potentially influence the company's stock price more substantially than the sentiment around non-durable goods due to the potential for long-term effects on the company's revenue and market position. Additionally, as substantiated by Vlastakis and Markellos (2012) [5], the steady consumption and consistent revenue from non-durable goods, such as services, can also impact investor sentiment and, by extension, stock prices due to steady consumption and consistent revenue from these goods can signal stability in the economy, which can affect how investors perceive market risk. Therefore, segmenting sentiment analysis based on product categories can provide a more nuanced understanding of their respective influences on stock performance. This kind of detailed, product-category-specific sentiment analysis and its impact on stock performance has yet to be the focus of previous research. This research project aims to uncover potential differences in the impact of sentiment around durable and non-durable goods on Apple's stock performance, hoping to provide more precise guidance for investment strategies and stock market forecasts.

Furthermore, the research aims to examine the geographical origin of sentiments obtained from Reddit comments and correlate them with the regional revenue contribution of Apple Inc. This unique approach is designed to identify possible associations between region-specific and product-specific sentiments and the stock performance of Apple Inc. By considering sentiments alongside the proportion of regional revenue and different product indices, this research intends to provide a more comprehensive outlook. The goal is to integrate insights from product and region-specific sentiment analysis with broader trends in public opinion. This study aims to offer a more nuanced understanding of how public sentiment when segmented by product category and geographical location, can impact stock market dynamics. Essentially, this research project seeks to contribute valuable insights that could potentially refine and enhance investment strategies by integrating sentiment analysis into traditional stock market forecasts.

3 Related Work

3.1 Introduction to Sentiment Analysis

Sentiment Analysis, a subfield of Natural Language Processing (NLP), is a computational approach to determining the emotions conveyed within a piece of text. It involves identifying the sentiment polarity, which could be positive, negative, or neutral, towards the context or subject under discussion.

Medhat, Hassan, and Korashy (2014)[1] provide a comprehensive overview of sentiment analysis and its diverse applications. They explain that sentiment analysis can be conducted on different scales: document level, sentence level, and aspect level. For example, document-level sentiment analysis classifies an entire document into one of the sentiment categories, such as positive or negative. Sentence-level sentiment analysis does the same but uses individual sentences as the basic unit of information. Aspect-level sentiment analysis goes deeper by identifying sentiments towards specific aspects or features of a product or service. To illustrate this point further, consider aspect-level sentiment analysis in the hotel industry. A hotel may receive customer feedback about various aspects such as room cleanliness, staff friendliness, and food quality. Using aspect-level sentiment analysis techniques to analyse customer reviews for each aspect separately, hotels can identify areas where they need to improve their services. Moreover, Medhat et al. (2014)[1] highlight that there are many applications of sentiment analysis beyond product reviews and social media platforms. For instance, stock market analysts use it to predict market trends based on investors' sentiments towards particular stocks or companies.

Similarly, Liu (2012)[2] has delved into the subject of opinion mining, a closely related concept. In his work, Liu suggests that sentiment analysis can be seen as an instance of opinion mining. It leverages NLP and text analytics techniques to mine subjective information from a given text. Opinions, which express sentiments, evaluations, or emotions towards entities or events, can be extracted through sentiment analysis. This approach enables us to capture the emotional nuance behind words, which is highly valuable in applications such as brand monitoring, product analytics, and understanding market sentiment.

Therefore, sentiment analysis stands as a potent instrument that helps us decipher public opinion, track sentiment towards a brand or product, and comprehend consumer needs. With the surge of digital communication platforms such as social media and review websites, sentiment analysis has become even more critical. Its diverse applications continue to grow as businesses increasingly realise the value of understanding public sentiment for decision-making and strategy formulation.

3.2 CardiffNLP

The suite of tools provided by TweetNLP is specially designed for the tokenisation, part-of-speech tagging, chunking, and named entity recognition of English tweets[6]. One notable tool incorporated in this suite is the CardiffNLP / twitter-roberta-base-sentiment-latest model, which applies the RoBERTa architecture for conducting sentiment analysis on Twitter data[6][7]. This model has been trained on a broad dataset, encompassing approximately 124 million tweets collected over four years, from January 2018 to December 2021[6][7]. The size and diversity of this dataset enhance the model's capability to comprehend and predict sentiments in a myriad of contexts.

The CardiffNLP / twitter-roberta-base-sentiment-latest model classifies sentiment into three labels: Negative, Neutral, and Positive, providing a more nuanced understanding of sentiment as compared to binary classifications, thereby capturing a broader spectrum of human emotions[6][7]. The reliability and credibility of this model are strengthened through evaluation using the TweetEval benchmark, a platform for the comparison and assessment of different models for various tweet classification tasks[6][7].

The model's successful integration into the TweetNLP toolkit ensures that it can take advantage of other functionalities offered by the platform, enhancing its efficacy in processing and analysing Twitter data[6][7]. Owing to its up-to-date training dataset, the model is adept at accurately interpreting contemporary phrases and terminology, making it especially suited for analysing social media platforms, which continually evolve with current linguistic trends[7]. Preliminary testing has also revealed that this model can precisely analyse emojis, aligning with the prevalent trend of habitual emoji usage in social media interactions[7].

In summary, the CardiffNLP / twitter-roberta-base-sentiment-latest model is an efficient tool for sentiment analysis on social media data. Its integration into the TweetNLP toolkit and extensive training dataset ensure precise sentiment prediction across various contexts[6][7]. Although the style and norms of communication vary between Twitter and Reddit, the common linguistic characteristics these platforms share suggest that this model may also be suitable for analysing Reddit data.

3.3 HFL

Practical sentiment analysis of the Chinese language requires the creation of specialised models owing to the inherent linguistic and cultural intricacies involved. Chinese text exhibits marked dissimilarities from English in various aspects, such as grammatical structure, word segmentation, dialectal diversity, cultural context, and usage of emojis and stickers, which significantly affect sentiment analysis.

Tokenisation is a simple process in English, as words are separated by spaces. However, the Chinese language does not use spaces, thereby presenting a formidable obstacle in word segmentation, which is an essential initial phase in several NLP duties[8]. Moreover, the grammatical structures of the two languages differ significantly. While English employs inflexions to indicate various factors like tense, mood, and number and has a firm reliance on word order, Chinese being an isolating language, depends upon fixed word forms with grammatical connections often demonstrated by function words or word order[8].

The cultural context and the use of idioms, metaphors, or cultural references exclusive to the Chinese language can significantly impact sentiment analysis. It is often tricky for sentiment analysis tools to accurately interpret these aspects. Furthermore, the diversity in Chinese dialects adds another layer of complexity as each one possesses unique vocabulary, grammar, and pronunciation. Additionally, implied meanings and usage of emojis and stickers differ considerably across cultures and languages - which can also affect sentiment analysis outcomes.

Polysemy poses a significant obstacle in the analysis of sentiment as it refers to words that have multiple meanings depending on their contextual usage. While both English and Chinese languages exhibit this phenomenon, it is more noticeable in Chinese due to the abundance of homophonic words. Consequently, identifying the appropriate sentiment becomes more challenging[8].

In order to tackle these obstacles, particular models like MacBERT have been created. These models were trained using an extensive pre-training dataset consisting of 5.4 billion words from multiple sources, including encyclopedias, news articles, and question-answering websites[8]. Additionally, during training, the Language Technology Platform (LTP) was utilised in order to identify the boundaries of Chinese words accurately[8].

MacBERT utilises techniques from the "Pre-Training with Whole Word Masking for Chinese BERT" approach[9] in addition to its own strategies. This method is aimed at resolving the challenge of partial word masking in the Chinese language, which can be complicated due to the absence of distinct word boundaries and the multiple implications that a character can possess when it is part of a multi-character word as opposed to standing alone[9]. By treating a whole word as a single entity for masking purposes, this strategy helps enhance the model's ability to comprehend contextual meaning[9].

Moreover, MacBERT suggests alterations to BERT's MLM task, involving the adoption of N-gram masking techniques and comparable word substitution for masking. These adjustments, coupled with the implementation of the Sentence Order Prediction (SOP) task that includes negative samples created through reversing the original order, have significantly enhanced MacBERT's performance in Chinese sentiment analysis[8][9].

3.4 Impact of Social Media Platforms on Stock Price

The impact of social media networks on the stock market has become an increasingly popular topic in the realm of research. According to various studies, there is a correlation between the sentiment gathered from social media data and trends in the stock market. As an example, Bollen, Mao, and Zeng (2011)[4], utilised Twitter data to project changes in the Dow Jones Industrial Average(DJIA) with remarkable accuracy at a rate of 87.6%. The study[4] exemplified how valuable social media platforms can function as predictive tools within financial markets. It revealed their potential for evaluating future performances of stocks by analysing public sentiments conveyed through such platforms.

Pagolu, Reddy, Panda, and Majhi (2016)[10] undertook a thorough analysis of Twitter sentiment and its link to Microsoft's stock prices during a year-long period. In order to collect data, they[10] utilised the Twitter API to obtain 250,000 tweets that included keywords related to Microsoft, which were then filtered. Their subsequent investigations[10] discovered a compelling association between the sentiment expressed in tweets about Microsoft and the stock prices for the next day. These results offer concrete support for employing social media as an advantageous instrument when anticipating trends within the stock market.

As mentioned earlier, the studies emphasise the potential of sentiment analysis conducted on social media platforms for predicting fluctuations in stock prices. In order to contribute to this discourse, my investigation endeavours to explore whether analysing the sentiment expressed towards various products and regions, along with their corresponding revenue ratios, can result in even more accurate forecasts of future stock prices.

3.5 Granger Causality Test and Directional Change Point Analysis

3.5.1 Granger Causality Test

Determining the ideal lag in a Granger causality analysis constitutes a vital phase as it holds significant sway over the test's robustness. As per the work of Granger (1969)[11], the Granger Causality Test entails two fundamental stages. Firstly, a basic Autoregressive (AR) model anticipates Y's current value through its own preceding values. This stage strives to comprehend how much of Y's present value may be accounted for by its past values. Subsequently, an elaborate AR model is constructed to predict Y's present value using its own preceding values and X's lagged values. This step endeavours to determine whether incorporating past X values enhances projecting present Y values beyond what could be achieved by only considering Y's previous readings alone.

In the second model, if the coefficients of the previous X values are proven to be statistically significant, it indicates that past values of X hold valuable information in predicting

Y that is not present in its own past values. This ultimately leads to the conclusion that X is a causative factor for Y, which is commonly known as "Granger-causes". On the other hand, if incorporating lagged values of X does not enhance the prediction accuracy of Y in an acceptable statistical manner, it can be inferred that X does not serve as a "Granger cause" for Y.

The selection of lag length carries a significant weight in the framework of this investigation, as it has a direct impact on the speed at which stock prices respond to sentiments voiced on social media platforms. Although the most fitting duration for lag may vary depending on the specific company and period under scrutiny, selecting an appropriate length through rigorous empirical analysis and informed by theoretical principles can considerably amplify the effectiveness of the Granger causality test's predictive capacity.

3.5.2 Directional Change Point

In the study of financial time-series data, a concept of substantial interest is the notion of Directional Change (DC) points, which are seen as pivotal inflexion points within the data. Han Aon (2018)[12] provides an in-depth analysis of these DC points, noting that they are observed when the price shifts direction by a specific percentage from a previous extremum. As Han[12] describes, this' delta' or threshold denotes the minor relative change in value that must transpire to establish a new DC point.

Introducing these DC points into my research holds significant implications for the construction and backtesting of trading strategies. These strategies will leverage the estimated lag times determined through the Granger causality test. This project will examine different 'delta' levels to make them more intricate. Creating multiple sets of DC points based on varying delta levels will help identify turning points in the time series at different granularity levels. This multi-level approach offers a more nuanced understanding of the financial market's behaviour, enhancing the prediction accuracy of our trading strategies. It allows the exploration of how varying perceptions of market volatility, embodied in different delta levels, can inform the creation of more robust and adaptable strategies.

4 Approach

4.1 Research Questions

The primary aim of this research is to examine the complex mechanisms that influence the stock values of Apple Inc. (AAPL). The goal is to identify underlying factors that may provide predictive comprehension into AAPL's stock behaviour. The investigation has been limited to three principal domains:

1 The Sentiment-Stock Nexus: The emergence of social media has brought forth a new avenue for the dissemination of information, affecting numerous industries, such as financial markets. In light of this development, an investigation into the potential correlation between sentiments conveyed regarding Apple, whose products are sold worldwide, merchandise on Reddit - a well-liked social media platform - and AAPL's stock values is warranted. Furthermore, discerning between sentiments directed towards durable and non-durable goods and evaluating their corresponding influence on AAPL's stock performance will be pursued.

2 The Product-Specific Sentiment-Stock Nexus: Beyond general sentiment, this research aims to investigate the impact of product-specific sentiments on AAPL's stock prices. The analysis will classify sentiments based on different Apple products, each indexed by their respective revenue contributions. The aim is to determine whether product-specific sentiment provides a more robust lag day result and better backtest result.

3 The Region-Specific Sentiment-Stock Nexus: Given the global influence of Apple Inc., the impact of regional sentiments on its stock price cannot be undermined. This research seeks to scrutinise whether sentiments from specific regions, particularly the Greater China region, as expressed on Reddit, have implications for AAPL's stock prices. Additionally, the research will adjust the sentiment analysis based on the regional revenue contributions of North America, Europe, and the Greater China region to Apple Inc. The aim is to identify any potential time lags in the region-specific sentiment-stock relationship, thereby enabling us to anticipate future stock price movements.

This research aims to investigate the repercussions of opinions relating to specific products and regions, as conveyed through social media channels, on AAPL's stock prices. Moreover, this research anticipates exploring whether sentiments linked to distinct product categories and geographic regions have divergent impacts on AAPL's stocks' performance. The results from this research may offer significant breakthroughs in the field of predicting stock prices and furnish investors and market strategists with valuable insights. It could also serve as a new factor for reinforcement learning algorithms utilised for predicting future stock prices.

4.2 Methodology

The methodology section of this project presents an overview of the research approach utilised, elucidating the measures undertaken in initiating this investigative pursuit.

1 Data Acquisition through Python Reddit API Wrapper (PRAW): This research's initial phase used PRAW to gather data from Reddit systematically. Reddit is a prominent online community platform offering valuable consumer opinion insights. This study focused on acquiring data from the Apple subreddit, which is an active English forum where international users discuss Apple's product line. Commencing on March 28th, all submissions and related comments were collected. To ensure that this study effectively supports the examination of region-specific sentiment-stock association of a particular location, information was also gathered from a well-known Chinese subreddit: China_irl. Pertinent data involving various conversations and opinions regarding Apple's products were retrieved using keywords related to the company and its merchandise. The ultimate objective was to compile a complete dataset providing an extensive sentiment spectrum for analysis, considering the Chinese market's significant influence on Apple.

2 Data Preprocessing: To maintain the contextual integrity of sentiments and achieve more accurate sentiment scoring, a specific technique for data preprocessing was implemented. The approach involved adding the corresponding submission to each comment, thus preserving its original narrative context. This measure aimed at establishing a more dependable foundation for subsequent sentiment analysis, which would allow for a deeper comprehension of the relationship between sentiment and stock.

3 Sentiment Analysis: For the subsequent stage of this research, the CardiffNLP, a Twitter-roBERTa-base model,[6][7] was employed to conduct English sentiment analysis. This specific model was updated in 2022 and was trained on a comprehensive dataset comprising around 124 million tweets. Additionally, it was fine-tuned for sentiment analysis through the TweetEval benchmark. This model was selected with a purposeful intent due to its aptness for analysing sentiment within social media domains. One benefit of utilising the Twitter-roBERTa-base is its proficiency in comprehending up-to-date textual content and emoji nuances - a crucial component of online communication. Given that emojis are increasingly used to express tone and sentiment in internet discussions, their analysis contributes an additional layer of depth to our understanding of sentiment.

On the other hand, Chinese data required a different approach due to its unique linguistic and cultural complexities. Aspects such as word segmentation, grammatical structure, cultural context, dialectal diversity, and the usage of emojis and stickers significantly affect sentiment analysis in Chinese. To overcome these challenges, a specialised model named MacBERT[8][9] was employed. This model was trained using an extensive pre-training dataset, encompassing 5.4 billion words from diverse sources like encyclopedias, news articles, and question-answering websites [8][9]. The Language Technology Platform (LTP) was utilised during training to accurately demarcate the boundaries of Chinese words [8].

The utilisation of these two models - theCardiffNLP[6][7] for English and the MacBERT[8][9] for Chinese - was aimed at achieving more precise and refined sentiment evaluations, thereby enhancing the credibility of the subsequent analysis.

4 Classification of Durable and Non-Durable Goods: In order to undertake this study, it was imperative to categorise the products offered by Apple into durable and non-durable goods. The differentiation is crucial since these categories' purchasing habits and consumer preferences can differ significantly. To begin with, given the extensive range of Apple's non-durable products and services, this project chose to identify durable goods first. For this study, durable goods are defined as high-value items that are less likely to be replaced or renewed often due to their longevity. To achieve this classification, this project filtered out data that mentioned specific products like iPhone, Mac, iPad, AppleTV, Studio Display and Pro Display XDR - all of which represent long-lasting items with high-ticket values. Subsequently, after identifying the durable goods category from the dataset available at hand, all other remaining data without mentions of these selected items were classified as non-durable goods. Such goods could see more frequent replacement or upgrade cycles or services requiring regular subscription payments. This classification methodology provides a comprehensive understanding of how stock prices may be influenced by different types of products and their impact on consumer sentiment.

5 Sentiment Compound Score Calculation: In order to transform the sentiments that have been extracted into a measurable format suitable for the research, it will be necessary to compute the sentiment compound score. The CardiffNLP[6][7] using the Twitter-roBERTa-base model provides a three-part output representing the Negative, Neutral, and Positive sentiments, each ranging from 0 to 1, with their sum equating to 1. The MacBERT model[8][9], in contrast, only returns Negative and Positive scores, each ranging from 0 to 1, with their sum also equating to 1.

The challenge arises when neutral scores are present, as relying solely on negative or positive scores might not precisely represent the sentiment of the comment. Given that the MacBERT model only returns negative and positive scores, this research assumed that the neutral score had been averaged into these two scores. As such, the study introduced a compound score that encapsulates the overall sentiment leaning for each comment. This compound score was calculated as the difference between the positive and negative sentiment scores, with one subsequently added and the result divided by 2. This method ensures that the sentiment compound scores fall within a scale of 0 to 1, where 1 represents the most positive sentiment, and 0 signifies the most negative sentiment. This standardisation not only renders the scores interpretable but also facilitates their subsequent usage in the correlation analysis with stock prices. The analyses are effectively streamlined by representing sentiments on a precise, uniform scale. Furthermore, the sentiment score representing a specific day was calculated as the mean of all the compound scores for that day.

6 Product-Specific Data: Considering the diverse range of Apple's product offerings, it is highly probable that individual products elicit distinct consumer reactions. These varying responses possess the potential to affect stock prices in unique ways. To examine the impact of consumer sentiment towards different products on stock prices and whether this impact varies according to the product's share of Apple's total sales, a "Product-Specific Score" was derived through this study. This score involved assigning weightage to compound scores of all product-related comments based on their contribution to Apple's revenue during the preceding quarter. Such an approach facilitates greater comprehension of the intricate relationship between product-specific sentiment and fluctuations in stock prices.

7 Region-Specific Data: One of the primary objectives of this research is to examine the relationship between the sentiments specific to different regions and the volatility of stock prices. A segment called 'Region-Specific Data' has been incorporated into the study to accomplish this. Similar to computing 'Product-Specific Data', it considers how much each region contributes towards Apple's earnings. Ascertaining the geographical source of English data being uncertain, English compound scores are multiplied by an average proportion of revenues from North America and Europe. However, since Chinese data is definitively sourced from Greater China, it is multiplied by that region's revenue proportion. This approach facilitates incorporating regional sentiment variability and its potential influence on Apple's stock prices.

8 Employing Granger Causality Test for Lag Identification: The study proceeded by utilising the Granger Causality Test to identify any lag effects, specifically examining the relationship over time between the daily sentiment compound scores of durable and non-durable goods, Product-Specific Data, Region-Specific Data, and AAPL's adjusted closing prices. As the stock market operates on business days, there were instances where the market was closed during weekends and certain other days. Forward filling was adopted as a technique to address this issue in the time series data. This method involved replacing non-trading days with the adjusted closing price from the most recent preceding trading day, ensuring a continuous dataset and preventing distortions due to missing stock price values.

An evaluation was carried out using the Granger causality test to determine if a temporal relationship or potential lag existed between the sentiment scores and AAPL's stock prices. This statistical hypothesis test is commonly used in time series analysis to assess if one time series can be helpful in predicting another. By implementing the Granger causality test, the goal was to discern if changes in public sentiment about Apple products, as captured in the sentiment scores, preceded changes in AAPL's stock price. This process would indicate whether Reddit sentiments have potential predictive power over AAPL's stock performance.

9 Analysis of Directional Change Points and Local Extremes: If a causal relationship was established in the previous step, the investigation moved on to the analysis of directional change (DC) points and local extremes within the sentiment graph. Identifying these transition points — where the trend of emotions switches from a downward to an upward trajectory or vice versa — could reveal valuable insights about optimal trading times.

Every turn in sentiment, regardless of its magnitude, is captured by fine-tuning the detection of DC points by setting the delta (which measures the level of change needed to identify a DC point) to zero. Then, an intriguing trading strategy was hypothesised. If there is a genuine lagged relationship between sentiment and stock price, one could potentially gain positive returns by trading on the day of a sentiment local extreme plus the discovered lag amount. Specifically, one would purchase AAPL stock when the sentiment trend shifts from downward to upward (assuming it will lead to a rise in stock price after the identified lag period) and sell when the sentiment trend reverses from upward to downward.

In order to further investigate this theory, the research continues to conduct experiments by adjusting different delta values when identifying DC points. This is used to test how different levels of fluctuation tolerance would impact the trading outcome. The smaller the delta, the less tolerant one is to fluctuations. For instance, with a small delta, one would sell whenever the sentiment increases by the delta amount and buy when it drops by the same magnitude. This additional testing provided a more nuanced understanding of the potential trading strategies informed by sentiment analysis.

5 Results and Analysis

5.1 The Sentiment-Stock Nexus

5.1.1 Results

The sentiment analysis of Reddit discussions regarding Apple products has been presented in Figures 1 through 3. These figures showcase the sentiment towards durable and non-durable goods, with unique markers indicating local extremes and directional change points where the delta is zero. The black line graph represents the sentiment, while the green graph demonstrates all local extremes. Additionally, red spots denote DC points that appear on a subsequent day following a local extreme observation and fluctuation in sentiment on the following day. A noteworthy observation derived from these figures highlights that the non-durable goods' sentiment graph shows a stronger correlation with overall sentiments than its durable counterparts, implying that non-durable goods may hold greater significance in determining community perception of Apple products on Reddit.

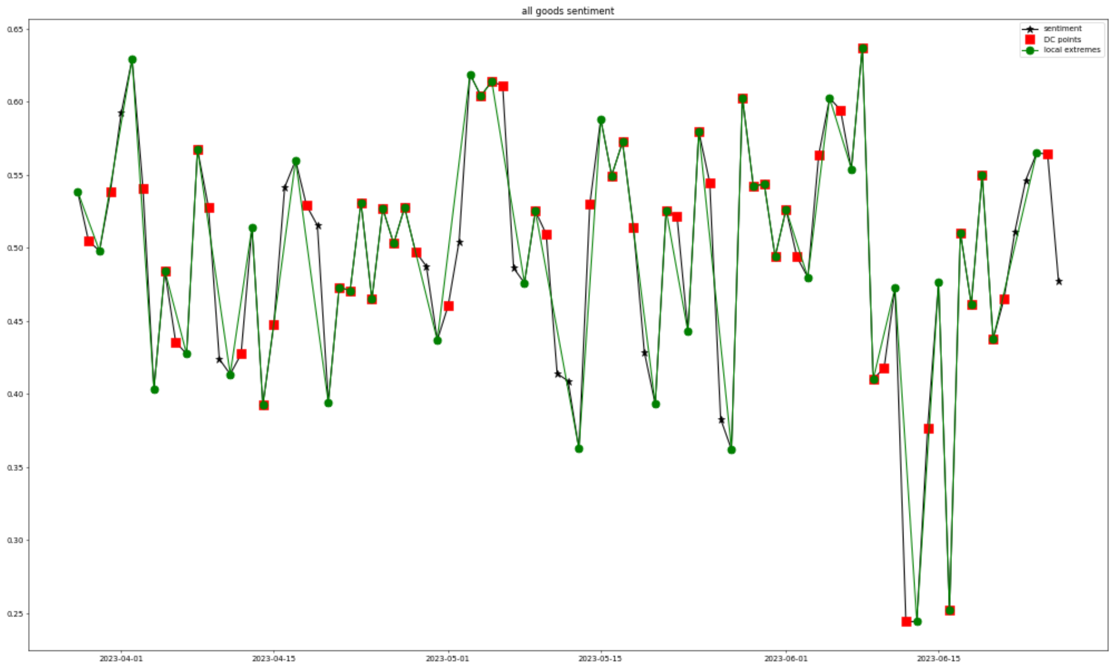

Figure 1: Sentiment towards all Apple goods.

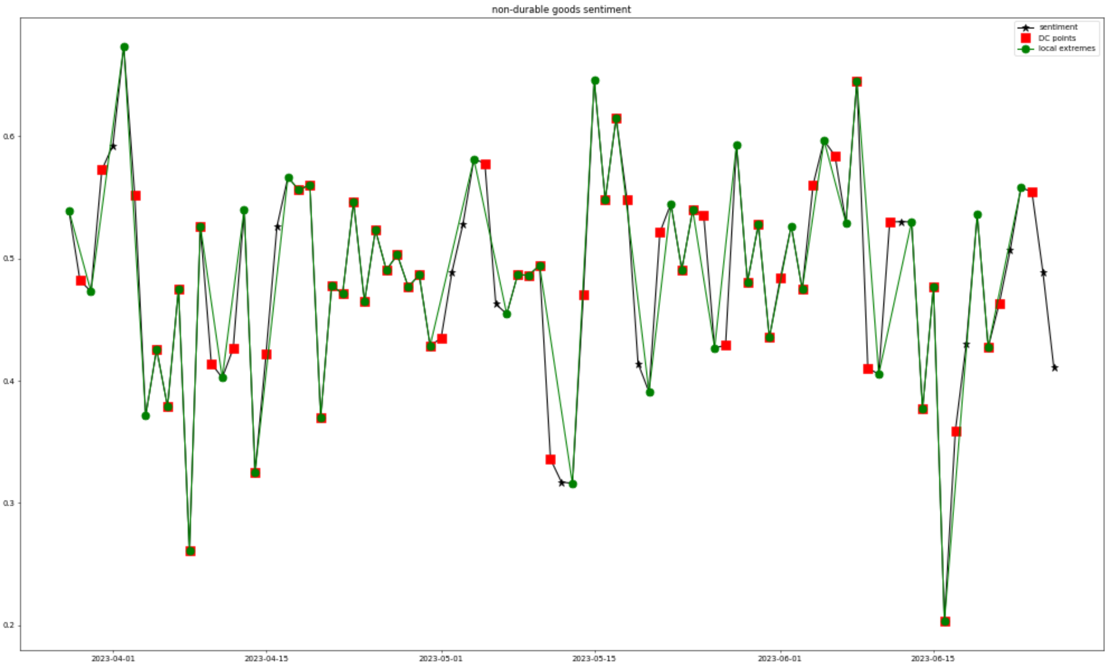

Figure 2: Sentiment towards non-durable goods.

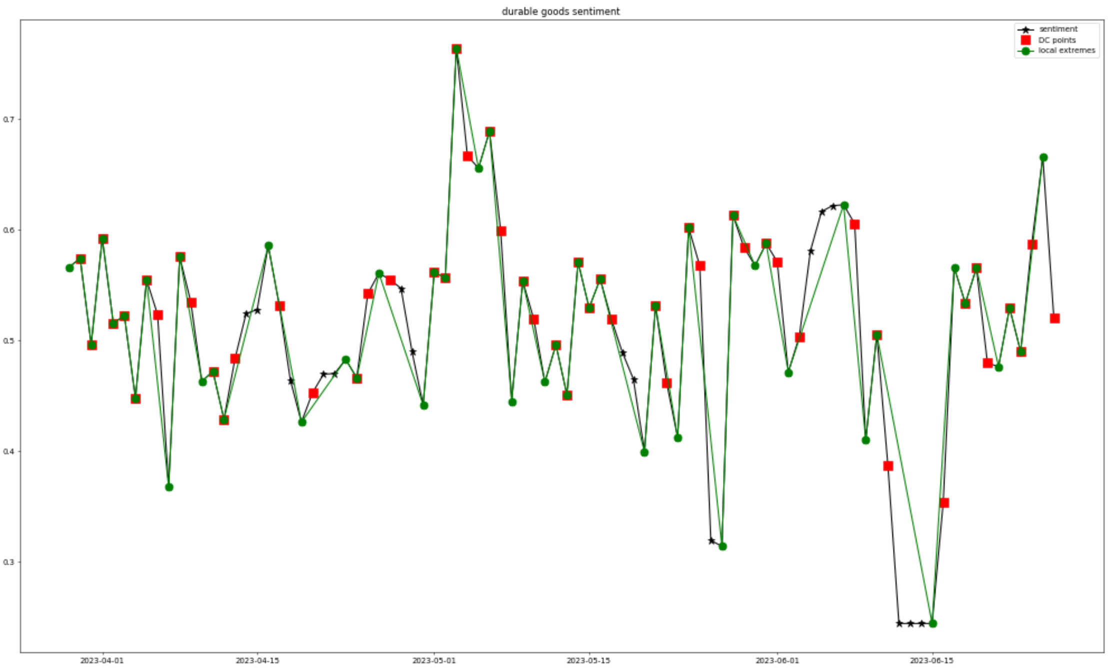

Figure 3: Sentiment towards durable goods.

- All Products and Services: The results of the Granger Causality test depicted in Figure 4 reveal a significant association between the public's sentiment towards Apple's products and services and the fluctuations observed in the company's stock prices. This correlation is predominantly apparent when there is a lag of two days. Various statistical tests were conducted, including ssr based F-test ($p = 0.0409$), ssr based chi-squared test ($p = 0.0296$), and a likelihood ratio test ($p = 0.0338$), all of which authenticate significant causal relationships between changes in sentiment and alterations in AAPL's stock price movements at this particular lag. These discoveries strongly suggest that any modifications in public sentiment regarding Apple products could accurately predict similar changes to AAPL's stock price almost two days later.

  It is of utmost importance to acknowledge that beyond a delay of two days, the relationship between sentiment and stock price appears less significant. For instance, the p-values for the ssr-based F-test, ssr-based chi-squared test, and likelihood ratio test at a lag of 5 days all exceed the commonly accepted significance level of 0.05, indicating insufficient evidence to support a causal connection between sentiment and stock price movement. This may imply that there is a delayed effect between changes in sentiment and fluctuations in stock prices, which diminishes over time.

  The results indicate a robust correlation between public perception of Apple products and the company's stock prices when there is a two-day time lag. Furthermore, this connection can somewhat anticipate the company's stock performance. However, it is essential to note that although the findings demonstrate a statistically significant association, their practical implications may differ depending on factors like market volatility, investor conduct, and other macroeconomic influences. Therefore, this forecasting capability should be viewed as just one piece of a comprehensive investment approach.

```
Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.8473  , p=0.3599  , df_denom=87, df_num=1
ssr based chi2 test:   chi2=0.8765  , p=0.3492  , df=1
likelihood ratio test: chi2=0.8723  , p=0.3503  , df=1
parameter F test:         F=0.8473  , p=0.3599  , df_denom=87, df_num=1

Granger Causality
number of lags (no zero) 2
ssr based F test:         F=3.3216  , p=0.0409  , df_denom=84, df_num=2
ssr based chi2 test:   chi2=7.0386  , p=0.0296  , df=2
likelihood ratio test: chi2=6.7741  , p=0.0338  , df=2
parameter F test:         F=3.3216  , p=0.0409  , df_denom=84, df_num=2

Granger Causality
number of lags (no zero) 3
ssr based F test:         F=2.1688  , p=0.0980  , df_denom=81, df_num=3
ssr based chi2 test:   chi2=7.0688  , p=0.0697  , df=3
likelihood ratio test: chi2=6.7992  , p=0.0786  , df=3
parameter F test:         F=2.1688  , p=0.0980  , df_denom=81, df_num=3

Granger Causality
number of lags (no zero) 4
ssr based F test:         F=1.7572  , p=0.1460  , df_denom=78, df_num=4
ssr based chi2 test:   chi2=7.8399  , p=0.0976  , df=4
likelihood ratio test: chi2=7.5066  , p=0.1114  , df=4
parameter F test:         F=1.7572  , p=0.1460  , df_denom=78, df_num=4

Granger Causality
number of lags (no zero) 5
ssr based F test:         F=1.4458  , p=0.2179  , df_denom=75, df_num=5
ssr based chi2 test:   chi2=8.2894  , p=0.1410  , df=5
likelihood ratio test: chi2=7.9138  , p=0.1610  , df=5
parameter F test:         F=1.4458  , p=0.2179  , df_denom=75, df_num=5
```

Figure 4: All goods Granger causality tests.

- Durable Goods: The findings of the Granger Causality examination, which examined the connection between durable goods sentiment and AAPL stock prices, diverged from those obtained in the previous analysis that evaluated overall sentiment scores. As shown in Figure 5, all p-values from each lag level test exceeded the widely recognised significance threshold of 0.05. This indicates a lack of a statistically significant Granger causality link between durable goods sentiment and subsequent fluctuations in AAPL's stock price.

  For instance, consider the lag of two days wherein the sentiment concerning all products was observed to have a correlation with AAPL's stock price. However, the results varied when the focus shifted solely towards durable goods sentiment. The F-test based on SSR showed a p-value of 0.3397, while the chi-squared test based on SSR yielded a p-value of 0.3136. In addition, the likelihood ratio test indicated a p-value of 0.3183. All these p-values significantly exceeded the significance threshold, suggesting that there is no prolonged temporal correlation when looking only at durable goods sentiment.

The absence of a statistically significant correlation suggests that the attitude towards durable goods by itself may not have enough influence to cause a noticeable change in AAPL's share value during the specific time frames analysed. However, it is important to interpret these results concerning the broader context of this study. The lack of a significant correlation does not negate the possible impact of durable goods sentiment on stock prices; instead, it underscores the complex nature of stock market behaviour.

```
Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.3499  , p=0.5557  , df_denom=86, df_num=1
ssr based chi2 test:   chi2=0.3621  , p=0.5474  , df=1
likelihood ratio test: chi2=0.3613  , p=0.5478  , df=1
parameter F test:         F=0.3499  , p=0.5557  , df_denom=86, df_num=1

Granger Causality
number of lags (no zero) 2
ssr based F test:         F=1.0938  , p=0.3397  , df_denom=83, df_num=2
ssr based chi2 test:   chi2=2.3194  , p=0.3136  , df=2
likelihood ratio test: chi2=2.2893  , p=0.3183  , df=2
parameter F test:         F=1.0938  , p=0.3397  , df_denom=83, df_num=2

Granger Causality
number of lags (no zero) 3
ssr based F test:         F=0.8897  , p=0.4502  , df_denom=80, df_num=3
ssr based chi2 test:   chi2=2.9026  , p=0.4069  , df=3
likelihood ratio test: chi2=2.8552  , p=0.4145  , df=3
parameter F test:         F=0.8897  , p=0.4502  , df_denom=80, df_num=3

Granger Causality
number of lags (no zero) 4
ssr based F test:         F=1.1456  , p=0.3416  , df_denom=77, df_num=4
ssr based chi2 test:   chi2=5.1179  , p=0.2754  , df=4
likelihood ratio test: chi2=4.9714  , p=0.2902  , df=4
parameter F test:         F=1.1456  , p=0.3416  , df_denom=77, df_num=4

Granger Causality
number of lags (no zero) 5
ssr based F test:         F=1.1390  , p=0.3475  , df_denom=74, df_num=5
ssr based chi2 test:   chi2=6.5417  , p=0.2570  , df=5
likelihood ratio test: chi2=6.3022  , p=0.2779  , df=5
parameter F test:         F=1.1390  , p=0.3475  , df_denom=74, df_num=5
```

Figure 5: Durable goods Granger causality tests.

- Non-Durable Goods: The outcomes of the Granger Causality examinations, which were administered to investigate the association between the sentiment relating to non-permanent goods and AAPL's stock price, as displayed in Figure 6, present a more refined perspective compared to previous analyses. Although the majority of p-values over various lag periods surpass the established significance level of 0.05, indicating an absence of a statistically significant Granger causal connection, there are noteworthy exclusions that warrant further discussion.

The findings of a time delay of 4 days are of particular interest. Notably, the ssr-based F-test shows a p-value of 0.0311, the ssr-based chi-squared test displays a p-value of 0.0138, and the likelihood ratio test exposes a p-value of 0.0197. All these numerical values fall below the widely accepted significance threshold value of 0.05, which strongly suggests that there exists a statistically significant Granger causal relationship between people's sentiment towards non-durable goods and subsequent stock price variations in AAPL over a period of four days.

Nevertheless, the pattern, as mentioned earlier, does not continue when considering a delay of 5 days. In this instance, the F-test based on SSR (p=0.0639) and likelihood ratio test (p=0.0386) outcomes tends to be around the significance threshold, implying a tenuous association. These discoveries indicate that the impact of non-durable goods' sentiments on AAPL's stock value appears to differ in terms of importance depending on the time lag. More precisely, the correlation seems most notable at a 4-day interval but diminishes slightly at a 5-day interval.

```
Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.2644  , p=0.6084  , df_denom=87, df_num=1
ssr based chi2 test:   chi2=0.2736  , p=0.6009  , df=1
likelihood ratio test: chi2=0.2732  , p=0.6012  , df=1
parameter F test:         F=0.2644  , p=0.6084  , df_denom=87, df_num=1

Granger Causality
number of lags (no zero) 2
ssr based F test:         F=3.0161  , p=0.0543  , df_denom=84, df_num=2
ssr based chi2 test:   chi2=6.3913  , p=0.0409  , df=2
likelihood ratio test: chi2=6.1723  , p=0.0457  , df=2
parameter F test:         F=3.0161  , p=0.0543  , df_denom=84, df_num=2

Granger Causality
number of lags (no zero) 3
ssr based F test:         F=1.9035  , p=0.1356  , df_denom=81, df_num=3
ssr based chi2 test:   chi2=6.2042  , p=0.1021  , df=3
likelihood ratio test: chi2=5.9952  , p=0.1118  , df=3
parameter F test:         F=1.9035  , p=0.1356  , df_denom=81, df_num=3

Granger Causality
number of lags (no zero) 4
ssr based F test:         F=2.8080  , p=0.0311  , df_denom=78, df_num=4
ssr based chi2 test:   chi2=12.5281 , p=0.0138  , df=4
likelihood ratio test: chi2=11.7043 , p=0.0197  , df=4
parameter F test:         F=2.8080  , p=0.0311  , df_denom=78, df_num=4

Granger Causality
number of lags (no zero) 5
ssr based F test:         F=2.1925  , p=0.0639  , df_denom=75, df_num=5
ssr based chi2 test:   chi2=12.5702 , p=0.0278  , df=5
likelihood ratio test: chi2=11.7323 , p=0.0386  , df=5
parameter F test:         F=2.1925  , p=0.0639  , df_denom=75, df_num=5
```

Figure 6: Non-durable goods Granger causality tests.

In brief, result 1 of this study demonstrates a multifaceted and complex correlation between Reddit sentiment and fluctuations in AAPL's stock price. Notably, the impact of sentiment appears to differ based on the categories of products being discussed and the duration of time under consideration. Upon analysing overall sentiment, a significant Granger causal relationship was found with AAPL's stock price when delayed by two days. However, upon examining sentiment linked solely to durable goods, no statistically significant Granger causality was observed during any of the examined time delays. Conversely, sentimental expressions associated with non-durable goods showed a statistically significant Granger causal relationship with AAPL's stock price at a lag time of 4 days.

The varied findings highlighted the intricate nature of the relationship between social media sentiment and stock market performance. The research shows that the impact of sentiment on stock prices can vary significantly depending on the type of goods and time lags under consideration. The outcome of this study confirms that Reddit's sentiment can have some effect on AAPL's stock price. However, with such diverse results, there is a pressing need for a more nuanced comprehension of these dynamics.

5.1.2 Backtest: Delta=0 (Consider Every Local Extreme)

The investigation successfully demonstrated a notable cause-and-effect association between Reddit's sentiment on Apple's various products and services and the fluctuation of AAPL stock prices. The research observed a time lag of two days for all Apple products and services and four days for non-durable ones. Upon establishing this connection, the study proceeded to implement a trading strategy that involved trading decisions centred around local extremes in sentiment, specifically at points where sentiment trends shifted from downward to upward or vice versa.

The trading strategy simulation assumed that the investor did not possess any AAPL shares initially but had adequate funds to purchase a single share at any point in time. Only sales were allowed when the investor had one or more shares, and each buy or sell transaction involved only one share. Hence, the investor could only hold either one AAPL stock or none. In addition, the initial investment return was set at zero, while gains or losses from trading activities were adjusted accordingly. Therefore, the investor's return is susceptible to negative values, indicating losses. Moreover, the implemented trading tactic required retaining ownership of the stock if the sentiment trend persisted on an upward trajectory, signifying that the last buying point was lower than the current market rates. Thus, if a share remained at the end of the plan's duration, it would be sold on that day to factor in any last potential returns and provide a more comprehensive and precise evaluation of its overall performance.

The trading strategy simulation outcomes were found to be favourable upon conducting a backtest. Specifically, conducting trade based on the sentiment of all Apple products and services with a two-day lag resulted in a positive return on investment, producing a gain of roughly 0.8731049999999811 USD. Similarly, adopting the same technique for non-durable Apple products and services generated a positive return on investment, yielding approximately 15.304993000000024 USD. These discoveries imply that using social media sentiment analysis as a basis for informed trading decisions is worth further investigation.

5.1.3 Backtest: Distinct Delta Values

After verifying the trading strategy regarding different lag days, this study investigates how distinct delta values influence the returns. Similar to the previous strategy, different deltas, ranging from 0% to 5% with an increment of 0.5%, were introduced in this phase of the analysis. Minor inflexion points in sentiment trends can be overlooked by increasing the delta value.

This approach aims to reduce noise in the sentiment curve, focusing instead on more significant sentiment trends. When changes in sentiment following an inflexion point are less than the specified delta value, they are dismissed as insignificant. In Figures 7 and 8, we can observe the effects of different delta values. When the delta value is set at 5% (as depicted in Figure 8), the number of local extreme points is noticeably fewer than in the situation where the delta equals 0 (as depicted in Figure 7), which considers all sentiment trend turns. Therefore, through this approach, we can more effectively isolate and focus on the major sentiment trends that potentially have a more substantial impact on returns.

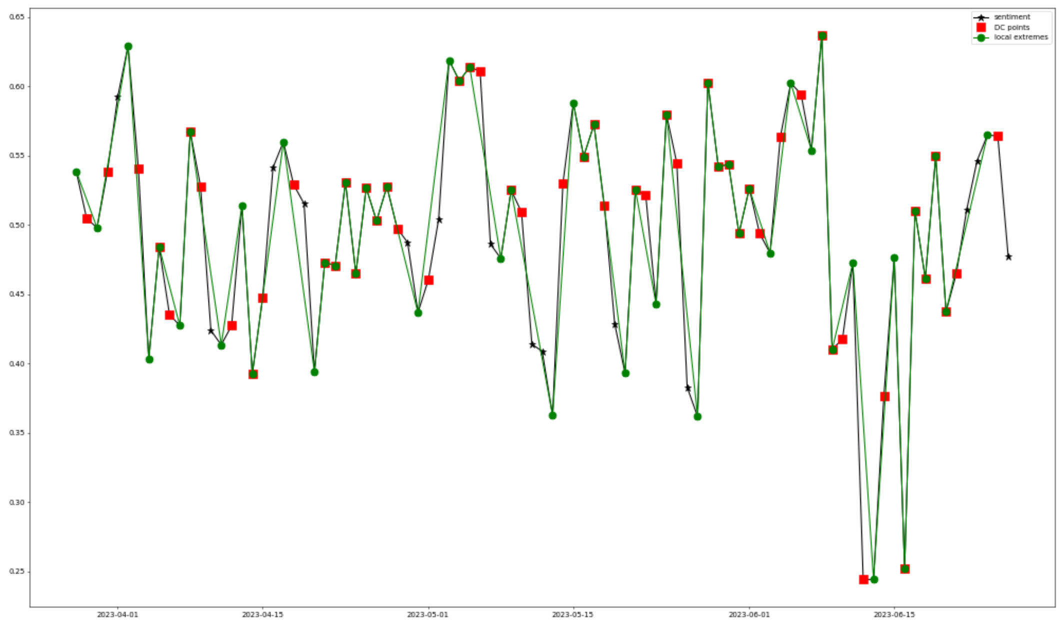

Figure 7: Zero delta in customer sentiment towards all Apple products and services.

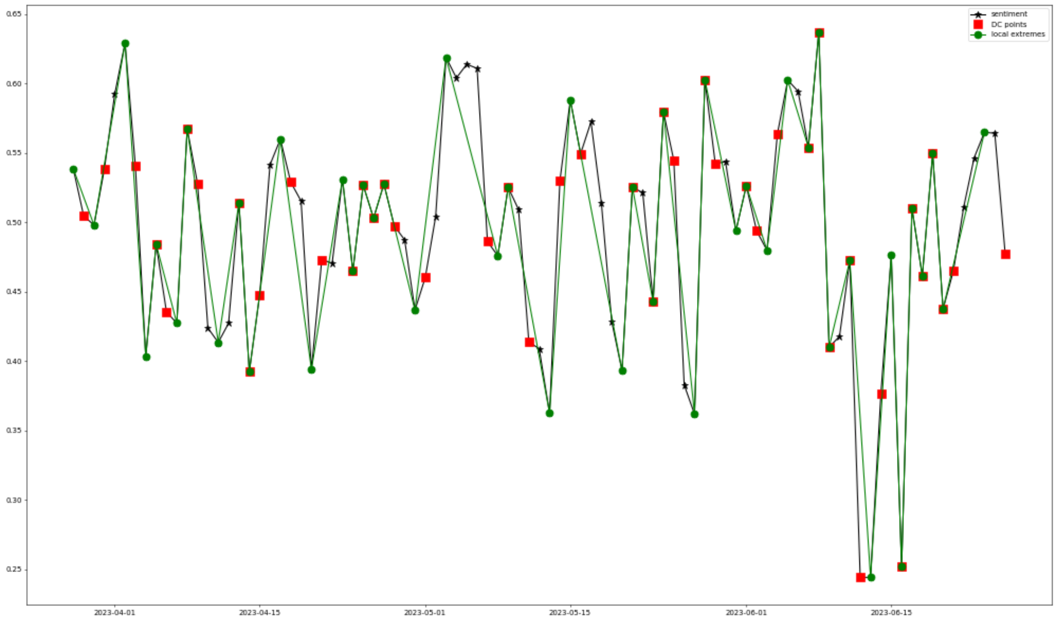

Figure 8: 5% delta in customer sentiment towards all Apple products and services.

The outcomes of trading methods based on different delta values of consumer sentiment towards all Apple and non-durable products are presented in Tables 1 and 2. It is worth mentioning that these approaches yield favourable returns for almost all delta values with one exception (non-durable products at delta = 1%).

Considering the entirety of the tactics, a customer sentiment-based strategy encompassing all Apple products generates an average return of around 16.29 USD. In contrast, such a strategy directed solely at non-durable Apple products yields an average return of approximately 14.5 USD. These outcomes serve to confirm the initial premise that consumer sentiment expressed on Reddit concerning Apple's products and services can predict its stock price to some degree. Nonetheless, additional investigation is required to ascertain whether non-durable items hold more significant sway over the stock price than other categories or not.

It has been noted that the incorporation of a rising delta value, which eliminates insignificant shifts in sentiment and reduces noise, can lead to higher returns. Nonetheless, this trend is not consistent throughout. For instance, in cases where the strategy concentrates on customer sentiment regarding all Apple products and raises the delta value to 4%, subsequent increases of 0.05% curiously lead to a decrease in returns. Despite this peculiarity, returns are still greater at these deltas than at lower ones, with the exception of the highest returns.

In summary, although amplifying the delta value seems to augment profits overall, it does not assure a surge. Based on this research, a greater delta might enhance returns.

However, the precise association and ideal delta value require additional scrutiny.

Table 1: Distinct Delta Values in Customer Sentiment Towards All Apple Products and Services

Delta (%)	**Return (USD, 2 decimal places)**
0	0.87
0.5	9.83
1	11.27
1.5	11.27
2	23.47
2.5	23.47
3	23.47
3.5	23.47
4	23.47
4.5	15.49
5	13.09

Table 2: Distinct Delta Values in Customer Sentiment Towards Non-durable Products and Services

Delta (%)	**Return (USD, 2 decimal places)**
0	15.30
0.5	19.10
1	-2.98
1.5	17.23
2	17.23
2.5	1.79
3	18.37
3.5	18.37
4	18.37
4.5	18.37
5	18.37

5.2 The Product-Specific Sentiment-Stock Nexus

5.2.1 Results

In this phase, the investigation integrates the revenue percentages of individual Apple commodities for Q1 2023 to determine product-specific scores. Apple's items and facilities are classified into five wide-ranging categories: iPhone sales contribute 54.13% of revenue, iPad sales account for 7.03%, Mac sales comprise 7.56%, Services make up 22.05%, and Wearables and Home Accessories constitute 9.23%.

	title	time	pos	neu	neg	per	compound	iphone	ipad	services	mac	wearables & home accessories
0	PSA: iOS 16.4 reintroduced the 'Curl' page tur...	2023-03-28 18:02:19	0.305941	0.686402	0.007657	0.0923	0.059916	0.0000	0.0	0.0	0.0000	0.0923
1	I am so frustrated with Apple - Does anyone el...	2023-03-28 19:22:53	0.010620	0.073149	0.916231	0.0923	0.004356	0.0000	0.0	0.0	0.0000	0.0923
2	Story of Classical on Apple Music Classical	2023-03-28 21:07:09	0.391858	0.601303	0.006839	0.0923	0.063919	0.0000	0.0	0.0	0.0000	0.0923
3	Apple Contribution to 'Made in India' Smartpho...	2023-03-28 23:33:55	0.535911	0.459227	0.004862	0.0923	0.070658	0.0000	0.0	0.0	0.0000	0.0923
4	Apple still hasn't fixed macOS Ventura's netwo...	2023-03-29 00:17:26	0.016878	0.201729	0.781392	0.0756	0.008901	0.0000	0.0	0.0	0.0756	0.0000
...	...	...	...	...	...	...	...	...	...	...	...	...

Figure 9: The revenue percentage associated with each content.

As shown in Figure 9, when a subject matter pertains to a particular product, the corresponding column is allocated the revenue proportion of that merchandise type, while unspecified categories persist at zero. It is noteworthy that this study permits one remark to refer to various product types; therefore, the overall revenue proportions of all specified merchandise categories are tallied in the end. The compound score gets standardised after it has been multiplied by an overarching percentage and then generates a daily product-specific score by averaging it out.

```
Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.3696  , p=0.5448  , df_denom=87, df_num=1
ssr based chi2 test:   chi2=0.3823  , p=0.5364  , df=1
likelihood ratio test: chi2=0.3815  , p=0.5368  , df=1
parameter F test:         F=0.3696  , p=0.5448  , df_denom=87, df_num=1

Granger Causality
number of lags (no zero) 2
ssr based F test:         F=1.5792  , p=0.2122  , df_denom=84, df_num=2
ssr based chi2 test:   chi2=3.3464  , p=0.1876  , df=2
likelihood ratio test: chi2=3.2850  , p=0.1935  , df=2
parameter F test:         F=1.5792  , p=0.2122  , df_denom=84, df_num=2

Granger Causality
number of lags (no zero) 3
ssr based F test:         F=1.0405  , p=0.3792  , df_denom=81, df_num=3
ssr based chi2 test:   chi2=3.3912  , p=0.3352  , df=3
likelihood ratio test: chi2=3.3274  , p=0.3438  , df=3
parameter F test:         F=1.0405  , p=0.3792  , df_denom=81, df_num=3

Granger Causality
number of lags (no zero) 4
ssr based F test:         F=1.0650  , p=0.3796  , df_denom=78, df_num=4
ssr based chi2 test:   chi2=4.7515  , p=0.3138  , df=4
likelihood ratio test: chi2=4.6263  , p=0.3278  , df=4
parameter F test:         F=1.0650  , p=0.3796  , df_denom=78, df_num=4

Granger Causality
number of lags (no zero) 5
ssr based F test:         F=0.9962  , p=0.4260  , df_denom=75, df_num=5
ssr based chi2 test:   chi2=5.7115  , p=0.3353  , df=5
likelihood ratio test: chi2=5.5298  , p=0.3547  , df=5
parameter F test:         F=0.9962  , p=0.4260  , df_denom=75, df_num=5
```

Figure 10: Product-Specific Granger causality tests.

However, as indicated in Figure 10, the Granger causality tests did not establish any delayed relationship between Apple stock prices or product-specific scores. Nonetheless, the study persists in examining whether there is a more robust association between the score specific to Apple's products and its stock price during the annual Worldwide Developers Conference (WWDC) - a technology conference conducted by Apple Inc. This investigation extends from May 5th, approximately one month before the conference, until June 26th, when the research obtained the most recent available data.

```
Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.0569  , p=0.8125  , df_denom=49, df_num=1
ssr based chi2 test:   chi2=0.0604  , p=0.8059  , df=1
likelihood ratio test: chi2=0.0603  , p=0.8060  , df=1
parameter F test:         F=0.0569  , p=0.8125  , df_denom=49, df_num=1

Granger Causality
number of lags (no zero) 2
ssr based F test:         F=6.1470  , p=0.0043  , df_denom=46, df_num=2
ssr based chi2 test:   chi2=13.6304 , p=0.0011  , df=2
likelihood ratio test: chi2=12.0798 , p=0.0024  , df=2
parameter F test:         F=6.1470  , p=0.0043  , df_denom=46, df_num=2

Granger Causality
number of lags (no zero) 3
ssr based F test:         F=4.0711  , p=0.0124  , df_denom=43, df_num=3
ssr based chi2 test:   chi2=14.2016 , p=0.0026  , df=3
likelihood ratio test: chi2=12.5003 , p=0.0059  , df=3
parameter F test:         F=4.0711  , p=0.0124  , df_denom=43, df_num=3

Granger Causality
number of lags (no zero) 4
ssr based F test:         F=4.9799  , p=0.0024  , df_denom=40, df_num=4
ssr based chi2 test:   chi2=24.4013 , p=0.0001  , df=4
likelihood ratio test: chi2=19.8019 , p=0.0005  , df=4
parameter F test:         F=4.9799  , p=0.0024  , df_denom=40, df_num=4

Granger Causality
number of lags (no zero) 5
ssr based F test:         F=3.5885  , p=0.0096  , df_denom=37, df_num=5
ssr based chi2 test:   chi2=23.2770 , p=0.0003  , df=5
likelihood ratio test: chi2=18.9779 , p=0.0019  , df=5
parameter F test:         F=3.5885  , p=0.0096  , df_denom=37, df_num=5
```

Figure 11: Product-Specific Granger causality tests WWDC period.

As demonstrated in Figure 11, over this period of almost one month surrounding the conference, it becomes evident how the connection between the product-specific score and Apple's stock price unfolds. For a 1-day lag, the Granger Causality Test does not reject the null hypothesis, suggesting that the product-specific sentiment fails to Granger cause the Apple stock prices (F=0.0569, p=0.8125). This outcome infers that sentiment data from the previous day do not significantly influence the immediately following day's stock price. Nevertheless, this premise is compellingly challenged when the lag is extended to 2 days. The null hypothesis is decisively rejected at this juncture (F=6.1470, p=0.0043), suggesting a predictive power of the product-specific sentiment from two days prior on the present day's Apple stock prices. A similar trend can be noted while assessing the lags of 3, 4, and 5 days, with respective F-values of 4.0711, 4.9799, and 3.5885, along with corresponding p-values of 0.0124, 0.0024, and 0.0096. The associated p-values are all below the conventional 0.05 significance level, thereby confirming that the product-specific sentiment does indeed Granger cause Apple's stock price. These findings suggest that there is an increased relationship between the product-specific score and Apple's stock prices during this time; however, further exploration is necessary to gain a more detailed understanding of this correlation's nature and duration.

5.2.2 Backtest: WWDC With Varying Lag Days

Table 3 presents the returns obtained from product-specific sentiment during the WWDC period with typical lag days. The use of a 2-day lag in strategy resulted in a return of USD 3.44, while a 3-day lag caused a loss of USD 2.13, thereby highlighting that this strategy involves inherent risks. Conversely, applying a 4-day lag produced returns of USD 1.57 and using a lag of five days yielded the highest return of USD 8.15, which is noteworthy. These results confirm that there is a temporal correlation between product-specific sentiment and stock price during the WWDC period; however, the uneven performance across different lag periods emphasises the necessity for an attentive and dynamic approach when executing this strategy.

Table 3: Distinct Lag Days in Product-Specific Sentiment

Lag (day(s))	**Return (USD, 2 decimal places)**
2	3.44
3	-2.13
4	1.57
5	8.15

5.3 The Region-Specific Sentiment-Stock Nexus

5.3.1 Results

The outcomes concerning the Chinese Reddit sentiment on AAPL's stock price are presented in Figure 12. Upon scrutinising the Granger causality tests, it is evident that a correlation exists between Chinese Reddit sentiment and AAPL's stock price, with a delay ranging from 2 to 5 days. This is demonstrated by the SSR-based F test p-values beneath the 0.05 significance threshold. Nevertheless, identifying the most suitable lag range for predicting the stock price continues to be challenging.

```
Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.4438  , p=0.5071  , df_denom=87, df_num=1
ssr based chi2 test:   chi2=0.4591  , p=0.4981  , df=1
likelihood ratio test: chi2=0.4579  , p=0.4986  , df=1
parameter F test:         F=0.4438  , p=0.5071  , df_denom=87, df_num=1

Granger Causality
number of lags (no zero) 2
ssr based F test:         F=5.1498  , p=0.0078  , df_denom=84, df_num=2
ssr based chi2 test:   chi2=10.9126 , p=0.0043  , df=2
likelihood ratio test: chi2=10.2937 , p=0.0058  , df=2
parameter F test:         F=5.1498  , p=0.0078  , df_denom=84, df_num=2

Granger Causality
number of lags (no zero) 3
ssr based F test:         F=3.4977  , p=0.0192  , df_denom=81, df_num=3
ssr based chi2 test:   chi2=11.3998 , p=0.0097  , df=3
likelihood ratio test: chi2=10.7196 , p=0.0133  , df=3
parameter F test:         F=3.4977  , p=0.0192  , df_denom=81, df_num=3

Granger Causality
number of lags (no zero) 4
ssr based F test:         F=2.9746  , p=0.0243  , df_denom=78, df_num=4
ssr based chi2 test:   chi2=13.2713 , p=0.0100  , df=4
likelihood ratio test: chi2=12.3515 , p=0.0149  , df=4
parameter F test:         F=2.9746  , p=0.0243  , df_denom=78, df_num=4

Granger Causality
number of lags (no zero) 5
ssr based F test:         F=2.7509  , p=0.0246  , df_denom=75, df_num=5
ssr based chi2 test:   chi2=15.7718 , p=0.0075  , df=5
likelihood ratio test: chi2=14.4812 , p=0.0128  , df=5
parameter F test:         F=2.7509  , p=0.0246  , df_denom=75, df_num=5
```

Figure 12: Chinese Sentiment Granger causality tests.

The research combines the Chinese and English compound sentiments by utilising Apple's regional revenue proportions index. It has been previously observed that correctly classifying the region of each English Reddit post poses a challenge. Therefore, this study assumes that these posts originate from either North America or Europe, with an equal weightage of 50% assigned to each region. This implies that the specific sentiments related to the region are formulated by multiplying the English compound sentiment with an average proportion of North American and European revenues, which are 39.84% and 25.25%. Similarly, it multiplies Chinese compound sentiment with Greater China's revenue proportion (18.78%). These values are then standardised based on daily data quantity before being added together. The findings of the Granger causality test, which examined the correlation between region-specific sentiment and AAPL stock price, have been visually presented in Figure 13. The study has conclusively established a significant correlation between these two variables, with the only noteworthy lag being for two days. Notably, these results closely resemble those discussed in "The Sentiment-Stock Nexus" section pertaining to the correlation between sentiment towards all Apple products and fluctuations in AAPL stock price.

```
Granger Causality
number of lags (no zero) 1
ssr based F test:         F=1.5650  , p=0.2143  , df_denom=87, df_num=1
ssr based chi2 test:   chi2=1.6190  , p=0.2032  , df=1
likelihood ratio test: chi2=1.6046  , p=0.2053  , df=1
parameter F test:         F=1.5650  , p=0.2143  , df_denom=87, df_num=1

Granger Causality
number of lags (no zero) 2
ssr based F test:         F=3.1632  , p=0.0474  , df_denom=84, df_num=2
ssr based chi2 test:   chi2=6.7030  , p=0.0350  , df=2
likelihood ratio test: chi2=6.4626  , p=0.0395  , df=2
parameter F test:         F=3.1632  , p=0.0474  , df_denom=84, df_num=2

Granger Causality
number of lags (no zero) 3
ssr based F test:         F=2.0697  , p=0.1107  , df_denom=81, df_num=3
ssr based chi2 test:   chi2=6.7457  , p=0.0805  , df=3
likelihood ratio test: chi2=6.4996  , p=0.0897  , df=3
parameter F test:         F=2.0697  , p=0.1107  , df_denom=81, df_num=3

Granger Causality
number of lags (no zero) 4
ssr based F test:         F=1.7631  , p=0.1448  , df_denom=78, df_num=4
ssr based chi2 test:   chi2=7.8661  , p=0.0966  , df=4
likelihood ratio test: chi2=7.5306  , p=0.1104  , df=4
parameter F test:         F=1.7631  , p=0.1448  , df_denom=78, df_num=4

Granger Causality
number of lags (no zero) 5
ssr based F test:         F=1.5072  , p=0.1978  , df_denom=75, df_num=5
ssr based chi2 test:   chi2=8.6412  , p=0.1243  , df=5
likelihood ratio test: chi2=8.2341  , p=0.1438  , df=5
parameter F test:         F=1.5072  , p=0.1978  , df_denom=75, df_num=5
```

Figure 13: Region-Specific Granger causality tests.

The similarity observed in the study may be attributed to the data asymmetry between the Chinese and English content acquired from Reddit. The data volume of Chinese content is significantly lower than that of English, which minimises the impact of Chinese sentiment after standardisation, as illustrated in Figure 14. Hence, it can only be inferred that English and Chinese sentiments both have some capability to predict Apple's stock price, respectively. Furthermore, merging Chinese and English sentiment data under conditions of unequal data volumes does not change the relationship with a 2-day time gap. However, further comprehensive research is needed to better understand the interrelationships among these variables, requiring larger datasets for testing purposes.

	index	Date	English_compound_scores	Count_E	Chinese_compound_scores	Chinese_counts
0	0	2023-06-26	0.477594	1018	0.458017	4.0
1	1	2023-06-25	0.564166	461	0.395134	3.0
2	2	2023-06-24	0.564887	625	0.397084	22.0
3	3	2023-06-23	0.546381	1314	0.371331	5.0
4	4	2023-06-22	0.511117	1100	0.418958	49.0
...	...	...	...	...	...	...

Figure 14: Data Volume.

5.3.2 Backtest

Based on the Granger causality test, a trading strategy based on Chinese Reddit sentiment with a delay of 2 to 5 days is recommended to be employed as it is expected to generate favourable outcomes. Figure 15 shows a line chart delineating Chinese Reddit sentiment towards Apple items, displaying local extremes (green) and DC points (red) at zero delta. It is worth noting that, at specific points, the sentiment remains consistent for several days. This may be attributed to the fact that the amount of Chinese data analysed is relatively small compared to English data. As a result, this study assumes that the sentiment does not vary on such days and holds at the previous day's level. This aspect did not appear in the analysis of English sentiment for all Apple products outlined in "The Sentiment-Stock Nexus" section.

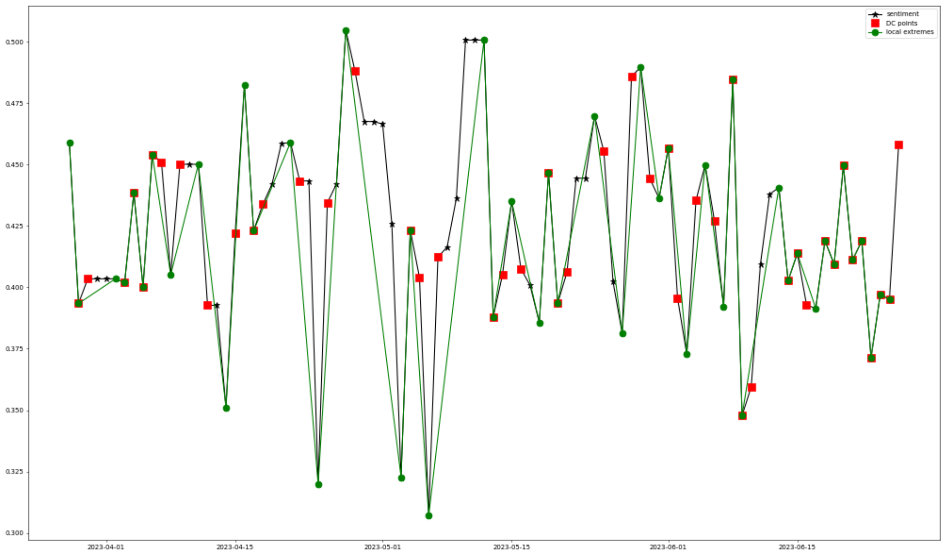

Figure 15: Chinese sentiment towards all Apple goods.

The outcomes for a delay ranging from 2 to 5 days, which resulted from the backtesting process, have been demonstrated in Table 4. As anticipated earlier, all four returns show a positive trend. Even though the outcome with a lag of two days appears optimal, as evidenced by significantly lower p values from Granger causality tests compared to other lags, it is imperative to acknowledge that the absence of Chinese data may have affected the results. Therefore, some predictability in stock prices could exist based on Chinese Reddit sentiment. However, it is crucial to conduct extensive testing and further research for a more comprehensive understanding of this relationship.

Table 4: Distinct Lag Days in Chinese Sentiment

Lag (day(s))	**Return (USD, 2 decimal places)**
2	16.20
3	6.52
4	4.41
5	9.56

Subsequently, an evaluation was conducted on the strategy backtest of region-specific sentiment. A prior assessment through the Granger causality test revealed that the correlation between region-specific sentiment and AAPL stock price materialises after two days. Figure 16 portrays a line chart representing region-specific sentiment. It is noteworthy that this particular chart bears dissimilar to those of pure English and pure Chinese sentiment (as depicted in Figures 1 and 15 correspondingly). However, the general pattern is similar to that of pure English, as demonstrated in Figure 1. This further clarifies that the impact of Chinese sentiment on region-specific outcomes might be restricted owing to a significant disparity in data volume.

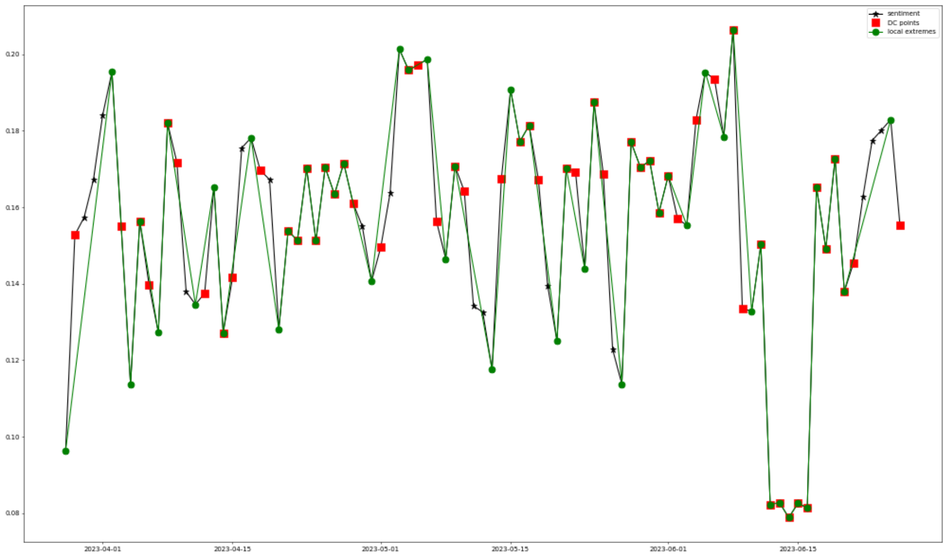

Figure 16: Region-Specific sentiment towards all Apple products and services.

A delay of two days and a delta value equal to zero were used to perform a backtest on a trading strategy. The result produced a positive value of 4.370010000000036, which appears to outperform the pure English Reddit value of 0.87 towards Apple products that were presented in the Sentiment-Stock Nexus backtest section. This test indicates that implementing a two-day lag trading strategy can be effective and supports the idea that region-specific sentiment is an essential predictor of AAPL stock prices. However, due to limitations in data sample size, further research and testing are necessary to fully understand this correlation and determine whether considering sentiment from various regions along with the products' regional revenue percentage could enhance the prediction of stock prices.

5.3.3 Backtest: Distinct Delta Values

This study conducted Granger causality tests to identify the regional sentiment as a factor of Apple's stock price, which was found to be a two-day lag. Subsequently, the study delved deeper into the impact of distinct delta values on the returns. Table 5 shows that the findings were consistent with a previous study that analysed pure English sentiment towards all Apple products and services - a general trend emerged where higher delta values corresponded to greater returns. However, exceptions were observed in the rates of 1.5%, 2.5%, and 4.5%, where a decrease in returns was noted. Despite the decline in returns at the above rates, the resulting return still surpassed the highest observed return among rates ranging from 0% to 1.5% by approximately 80%.

According to the Granger causality tests, a two-day lag was observed to have a substantial correlation with Apple's stock price in both pure English sentiment and region-specific sentiment experiments. Table 5 illustrates the contrast between returns obtained through pure English sentiment and region-specific sentiment at varied delta values over a two-day lag. The outcomes indicate that the inclusion of Chinese Reddit comments and consideration of regional revenue percentages for Apple products only yielded higher returns when the delta value equalled zero. The average return from pure English sentiment stood at roughly 16.29 USD, which outperformed the average return of nearly 13.45 USD derived from region-specific sentiment at different delta values. Nonetheless, overall region-specific sentiment produces the most noteworthy return when the delta equals 2%.

Conclusively, the present investigation established that the integration of Chinese Reddit data and consideration of Apple's regional earnings ratios did not enhance returns significantly when trading based on relevant lag days and delta values. However, it is imperative to acknowledge that returns were observed to increase when the delta equalled 0, 2 and 5, signifying a need for further inquiry into this aspect. Asymmetry in volume between English and Chinese data, along with testing methodologies, could have contributed to the less distinctive findings in this study. Overall region-specific sentiment produces the most noteworthy return when the delta equals two, highlighting potential

Table 5: Distinct Delta Values in English Sentiment and Region-Specific Sentiment

Delta (%)	English (USD)	Region-Specific (USD)
0	0.87	4.67
0.5	9.83	4.67
1	11.27	8.45
1.5	11.27	0.46
2	23.47	25.72
2.5	23.47	18.32
3	23.47	18.32
3.5	23.47	18.32
4	23.47	18.32
4.5	15.49	15.34
5	13.09	15.34

pathways for future research. This indicates that exploring additional variables and refining testing strategies could produce more substantial outcomes.

5.3.4 Limitations of Chinese Reddit Data Acquisition

The current investigation made use of the Python Reddit API Wrapper (PRAW) to retrieve data from Reddit that pertains to Chinese keywords related to Apple products. Nevertheless, it has been observed that this technique of gathering data has certain drawbacks because Reddit is primarily an English-language-based platform. The accuracy of PRAW in extracting Chinese keywords was discovered to be somewhat inadequate as it included irrelevant information, and there is no assurance that all pertinent data has been acquired with this approach. As mentioned above, the limitation has considerable implications on the dependability of the analyses carried out in "The Region-Specific Sentiment-Stock Nexus" section. As mentioned earlier, the disparity in data between English and Chinese Reddit content is possibly worsened by the imprecise approach to extracting Chinese keywords. This can lead to a likely underrepresentation of Chinese sentiment in the analysis of regional sentiment and subsequent backtesting of trading strategies.

It is noteworthy that although this constraint may have an effect on the results, the study carried out in these areas still offers valuable perspectives into the potential influence of region-specific Reddit sentiment on Apple's stock price. Nevertheless, this problem underscores the necessity for enhanced techniques or instruments to extract non-English material from other online platforms such as Weibo.

Conclusively, the observations made in the section of "The Region-Specific Sentiment-Stock Nexus" still hold value but should be considered with the awareness of these limitations. This could also open up future research, particularly in improving data collection methods for non-English sentiment analysis, thus leading to more precise and meaningful outcomes.

6 Legal, Social, Ethical and Professional Issues

6.1 Synopsis

1 Public Interest: Throughout this research endeavour, I have ensured that the public's interest is prioritised. It is worth mentioning that while handling data obtained from publicly accessible domains of Reddit, I have been highly cautious and mindful of preserving the privacy and confidentiality of users. This commitment aligns with Reddit's developer regulations that necessitate anonymising any published outcomes derived from their data. Additionally, the study aims to ensure no partiality or bias in the analysis or interpretation of findings.

2 Professional Competence and Integrity: Throughout the duration of the research project, I conscientiously followed the principles of professional competence and integrity, remaining cognizant of Reddit's code of conduct regarding the use of its developer tools. Given the rate limits imposed by Reddit's APIs and their possible impact on data collection, this research methodology was based on the need for ongoing professional development and awareness of relevant technological advancements. The sound execution of this study was crucial to ensure the accuracy and consistency of the research findings. Therefore, I approached the design of this research methodology with a resolute commitment to ethical standards, being extremely cautious not to cause harm or infringe upon Reddit's users or reputation. This approach reflects my firm dedication to upholding the values of integrity and respect in academic research.

3 Duty to Relevant Authority: My commitment to Reddit's policies for academic purposes has been unwavering in connection with my professional judgement and responsibilities, specifically regarding the Duty to Relevant Authority. The credibility of the research is supervised by the institution I am affiliated with, as well as my supervisor, who serves as relevant authority. Following Reddit's developer regulations, which prohibit the distribution of data or derivative products, this research has maintained the confidentiality of Reddit's data. Additionally, this research strictly adheres to the requirement to not use the content on Reddit for model training or commercial purposes without explicit consent from Reddit.

4 Duty to the Profession: As a prospective member of the academic community, I acknowledge my responsibility to maintain and enhance the standing and criteria of our profession. This encompasses acknowledging Reddit for granting access to their platform and data as per their policies. Furthermore, I am dedicated to building constructive professional associations with all members of the academic community in order to advance our field. Additionally, I am resolute in promoting ethical conduct and professional integrity while conducting and distributing my research.

6.2 Deliberation regarding the BCS Code of Conduct

1 Public Interest

a. Have due regard for public health, privacy, security and wellbeing of others and the environment.

– The research was conducted ethically with due regard for public health, privacy, and security of Reddit users' data.

b. Have due regard for the legitimate rights of Third Parties.

– The research anonymised any published outcomes derived from Reddit data as required by their policies to respect users' rights.

c. Conduct your professional activities without discrimination on the grounds of sex, sexual orientation, marital status, nationality, colour, race, ethnic origin, religion, age or disability, or of any other condition or requirement.

– The research was designed and conducted research without discrimination on any grounds.

d. Promote equal access to the benefits of IT and seek to promote the inclusion of all sectors in society wherever opportunities arise.

– The research aimed to contribute to a greater understanding of using sentiment analysis to promote equal access to technology benefits.

2 Professional Competence and Integrity

a. Only undertake to do work or provide a service that is within your professional competence.

– Only undertook sentiment analysis using NLP techniques within my area of competence.

b. NOT claim any level of competence that you do not possess.

– Did not claim any unwarranted level of competence in my abilities.

c. Develop your professional knowledge, skills and competence on a continuing basis, maintaining awareness of technological developments, procedures, and standards that are relevant to your field.

– Developed skills and knowledge in sentiment analysis and NLP through conducting this research project.

d. Ensure that you have the knowledge and understanding of Legislation* and that you comply with such Legislation, in carrying out your professional responsibilities.

– Knew of and complied with applicable legislation, including Reddit's terms of use and privacy policies.

e. Respect and value alternative viewpoints and, seek, accept and offer honest criticisms of work.
- Knew of and complied with applicable legislation, including Reddit's terms of use and privacy policies.

f. Avoid injuring others, their property, reputation, or employment by false or malicious or negligent action or inaction.
- The research avoided any false or negligent actions that could harm others.

g. Reject and will not make any offer of bribery or unethical inducement.
- Rejected any unethical inducements and conducted research with integrity.

3 Duty to Relevant Authority

a. Carry out your professional responsibilities with due care and diligence in accordance with the Relevant Authority's requirements whilst exercising your professional judgement at all times.
- Not Relevant.

b. Seek to avoid any situation that may give rise to a conflict of interest between you and your Relevant Authority.
- Not Relevant.

c. Accept professional responsibility for your work and for the work of colleagues who are defined in a given context as working under your supervision.
- Not Relevant.

d. NOT disclose or authorise to be disclosed, or use for personal gain, or to benefit a third party, confidential information except with the permission of your Relevant Authority, or as required by Legislation.
- The research did not disclose or authorise the disclosure of any confidential Reddit data, as per Reddit's developer policies prohibiting the distribution of data or derivative products.

e. NOT misrepresent or withhold information on the performance of products, systems or services (unless lawfully bound by a duty of confidentiality not to disclose such information), or take advantage of the lack of relevant knowledge or inexperience of others.
- Not Relevant.

4 Duty to the Profession

a. Accept your personal duty to uphold the reputation of the profession and not take any action which could bring the profession into disrepute.
- Conducted research ethically to uphold the reputation of the field.

b. Seek to improve professional standards through participation in their development, use and enforcement.

– The research contributed to a greater understanding of sentiment analysis applications.

c. Uphold the reputation and good standing of BCS, the Chartered Institute for IT.

– Not Relevant.

d. Act with integrity and respect in your professional relationships with all members of BCS and with members of other professions with whom you work in a professional capacity.

– Not Relevant.

e. Encourage and support fellow members in their professional development.

– Not Relevant.

7 Conclusion

7.1 Conclusion

To summarise, this investigation has revealed a strong correlation between the sentiments expressed on Reddit regarding Apple products and the fluctuation of Apple's stock price. This establishes the possibility of leveraging such data for devising strategies in stock trading. Furthermore, adjusting sentiment and the revenue percentage of different products makes it possible to predict stock prices during launches. Additionally, when increasing the delta, large-scale sentiment trends generally provide more lucrative returns since minor fluctuations can be overlooked. Although research into regional earnings and sentiment is ongoing, these findings make it clear that social media sentiment towards multinational corporations' products can offer significant value when processed via deep reinforcement learning as factors for trading. It is also worth paying attention to the ratio between revenue and sentiment surrounding individual products during launch events.

7.1.1 The Sentiment-Stock Nexus

The research was conducted to analyse the correlation between sentiment and stock, which revealed that for all Apple products and services, AAPL stock prices experience a delay of two days to English Reddit sentiment. This implies that the sentiment tone conveyed through social media requires a span of two days to develop and be manifested in the stock market. It is possible that this arises from the swiftness with which information and emotions are disseminated on social media in contrast to conventional news sources; however, owing to its lack of authentication, investors may need time to scrutinise and confirm information along with underlying sentiment on social media before arriving at trading decisions.

A lack of correlation was detected between the sentiment of durable goods and their corresponding stock prices compared to non-durable goods. One plausible reason for this could be that investors tend to regard the level of satisfaction with durable goods as having a relatively minor influence on a company's earnings change. This can be attributed to the fact that generally, durable goods are priced higher per unit and have a comparatively fixed usage period. When purchasing such goods, most consumers consider their needs carefully and do not necessarily cause an abrupt surge in a company's profits solely based on positive product reviews.

The distinction between durable and non-durable goods has become indistinct when referring to Apple products. Products traditionally considered durable are now being consumed like non-durable goods, which may be attributed to individuals actively seeking out the latest trends and colours. This trend could be responsible for a lag observed in analysing the connection between non-durable goods and Apple's stock price but challenging to determine why this lag changes from two days when all products and services are considered to four days for non-durable goods.

7.1.2 The Product-Specific Sentiment-Stock Nexus

The introduction of Apple products' revenue percentage as an index in the experimental segment resulted in the disappearance of the previously observed two-day lag relationship between English Reddit sentiment and stock prices. However, by narrowing down the timeline to around a month before and after WWDC, it was found that a correlation between Product-Specific Sentiment and Apple's stock prices could be established. This indicates that investors generally prioritise overall sentiment towards Apple products, but during product launches, sentiment and sales proportions of individual products become significant factors impacting stock prices. It is evident that investors primarily focus on the overall performance of Apple products and customer sentiment towards them. Nonetheless, individual product sentiments, particularly those with high sales ratios, play a vital role in affecting stock prices during launch periods, whereas products with lower sales ratios have a comparatively lesser impact.

In addition, it was observed that the property of emotions being a predictor of stock prices continues to hold accurate starting from a delay of two days. This occurrence may be because individuals' anticipations for the product prior to its release will exhibit a comparable pattern except in cases where new information is leaked. Similarly, subsequent to the launch of the product, public sentiment towards it is expected to conform to this same pattern until they have had personal experience with it. Consequently, the day of product release is likely to be the only probable instance when there could be a noteworthy change in sentiment outlook over the entire course of events.

7.1.3 The Region-Specific Sentiment-Stock Nexus

The study analysed comments on Apple products collected from the Chinese subreddit China IRL. Additionally, a regional revenue percentage of Apple was introduced as an index. As a result of this analysis, it was found that Chinese sentiment on Reddit regarding Apple products could predict their stock prices. Furthermore, the combined Region-Specific Sentiment showed a two-day lag correlation with the stock price. However, no marked improvement was observed during the strategy backtest stage. This could be attributed to the general emotional trends towards Apple products being similar on Reddit regardless of language (English or Chinese) or the constraints and inaccuracies associated with extracting Chinese keywords using PRAW. Moreover, given the limited data obtained and the disproportionate ratio of English to Chinese comments, it is possible that the impact of Chinese Reddit sentiment may not be as significant and accurate as anticipated.

It is of note that languages other than English and Chinese were not included in this study, nor were social platforms beyond Reddit taken into consideration, such as Weibo, which is frequently utilised in the Greater China region. As a result, comprehending the predictive efficacy of regional revenue and sentiment on stock prices necessitates a more thorough investigation. Additionally, besides the limitations of Chinese Reddit acquisition, as mentioned in 5.3.4, another significant challenge identified by this study pertains to defining the geographic location of each social media post. The online world operates without borders, and categorising posts exclusively based on language may lead to combined comments from diverse geographical regions. This may also contribute to similarities observed within the general trends of collective sentiments.

7.2 Future Work

Future research could explore various potential areas, each offering distinct opportunities to enhance and build upon the current study's findings. One such area is expanding the range of data sources used in the analysis, which can be approached from two angles: temporal expansion and cross-platform expansion. Temporal expansion entails analysing data over a more extended period to identify any emerging trends or patterns, while cross-platform expansion involves examining data from additional platforms or sources to achieve a broader perspective on the subject matter being studied. Both methods present unique opportunities to deepen our understanding of the research area and lead to further discoveries in future investigations.

Expanding the dataset over an extended time would be highly advantageous in attaining a more profound comprehension of the correlation between consumer sentiment on social media and a company's stock performance. The present research offers valuable insights into this connection but is constrained to an isolated time. By prolonging the time-frame, analysts could reveal patterns and fluctuations that are discernible solely across extended periods. This would enable them to scrutinise how variations in consumer sentiment induce changes in stock prices spanning weeks, months or even years. Furthermore, by examining this data over an extensive duration, specialists can perceive trends and make more precise predictions about future market conduct.

It is also necessary to incorporate a broader range of data sources from various social media platforms such as Twitter, Weibo, and 'Threads' introduced by Meta. These platforms possess unique language usage and presentation patterns, requiring diverse Natural Language Processing (NLP) techniques for sentiment analysis. Incorporating different social media platforms into sentiment-stock research can provide a much more comprehensive understanding of the correlation between public sentiment and stock performance while considering the nuanced differences among platform communities. Exploring region-specific social media platforms commonly used in different regions can further provide insight into identifying how regional sentiments could potentially impact stock performance. By integrating multiple social media platforms, researchers can obtain extensive data covering various languages, cultures, and regions, leading to more accurate results in their study on stocks' relationship with public sentiments expressed across these networks.

Broadening the research scope to encompass additional multinational enterprises would be a beneficial measure towards substantiating the identified correlations across diverse industries, products, and markets. Although analysing Apple provides valuable insights into the sentiment-stock relationship, it is imperative to acknowledge that this association may vary among various business types, consumer bases, and cultural contexts. Therefore, expanding this study to include a more varied range of corporations could allow scholars to validate their findings in a broader context and investigate how these correlations demonstrate differently across different situations. It would be worthwhile to explore the integration of sentiment-stock correlation into a deep reinforcement learning framework. This technique may entail using correlation as a fundamental input attribute to improve decision-making. The inclusion of sentiment-stock correlation in a deep reinforcement learning framework shows promising potential for improving trading strategies in several ways. Researchers should further examine it in this field.

As the discussion draws to a close, there are numerous potential directions for future research in this domain. These domains present considerable potential and could substantially enhance our understanding of the intricate interplay between social sentiment and the stock market's performance. While this study has illuminated certain aspects of this relationship, many more opportunities remain untapped by researchers eager to advance the frontiers of knowledge in this exciting field.

References

[1] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams engineering journal*, vol. 5, no. 4, pp. 1093–1113, 2014.

[2] B. Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1–167, 2012.

[3] T. H. Nguyen, K. Shirai, and J. Velcin, "Sentiment analysis on social media for stock movement prediction," *Expert Systems with Applications*, vol. 42, no. 24, pp. 9603–9611, 2015.

[4] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of computational science*, vol. 2, no. 1, pp. 1–8, 2011.

[5] N. Vlastakis and R. N. Markellos, "Information demand and stock market volatility," *Journal of Banking & Finance*, vol. 36, no. 6, pp. 1808–1821, 2012.

[6] J. Camacho-collados, K. Rezaee, T. Riahi, A. Ushio, D. Loureiro, D. Antypas, J. Boisson, L. Espinosa Anke, F. Liu, E. Martínez Cámara, *et al.*, "TweetNLP: Cutting-edge natural language processing for social media," in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, (Abu Dhabi, UAE), pp. 38–49, Association for Computational Linguistics, Dec. 2022.

[7] D. Loureiro, F. Barbieri, L. Neves, L. Espinosa Anke, and J. Camacho-collados, "TimeLMs: Diachronic language models from Twitter," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, (Dublin, Ireland), pp. 251–260, Association for Computational Linguistics, May 2022.

[8] Y. Cui, W. Che, T. Liu, B. Qin, S. Wang, and G. Hu, "Revisiting pre-trained models for Chinese natural language processing," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings*, (Online), pp. 657–668, Association for Computational Linguistics, Nov. 2020.

[9] Y. Cui, W. Che, T. Liu, B. Qin, Z. Yang, S. Wang, and G. Hu, "Pre-training with whole word masking for chinese bert," *arXiv preprint arXiv:1906.08101*, 2019.

[10] V. S. Pagolu, K. N. Reddy, G. Panda, and B. Majhi, "Sentiment analysis of twitter data for predicting stock market movements," in *2016 international conference on signal processing, communication, power and embedded system (SCOPES)*, pp. 1345–1350, IEEE, 2016.

[11] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: journal of the Econometric Society*, pp. 424–438, 1969.

[12] H. Ao, *A Directional Changes based study on stock market.* PhD thesis, University of Essex, 2018.