Proceedings of the 3rd International Conference on Business and Policy Studies
DOI: 10.54254/2754-1169/81/20241843

Traditional CAPM Model to Sentiment Analysis

Ruohan Zheng[1,a,*,†], Yile Li[2,b,†], Yicheng Wang[3,c,†]

[1]arts and science, Ohio State University, Columbus, 43210, United States
[2]Physical Science, University of California, Irvine, 92612, United State
[3]arts and science, State University of New York at Stony Brook, Stony Brook, 11794, United States
a. Zheng.1827@osu.edu, b. 2000leahli@gmail.com, c. yicheng.wang@stonybrook.edu
**corresponding author*
[†]These authors contributed equally to this work and should be considered co-first authors.

Abstract: In this article, we will combine the novel applications of asset pricing, which is developed from a classic theory of this field. We will present this theory and then the latest found of the application. A large part will be the results based on the classic model CAPM. This part includes a review of Then present the other models or methods based on this result. Sentiment analysis is a vital component in stock price prediction, utilizing natural language processing and machine learning to extract and evaluate emotional information from textual data. This article reviews traditional lexicon-based methods and advanced sentiment analysis techniques, emphasizing the growing importance of sentiment and neural networks for a more nuanced understanding of sentiment's impact on financial markets. Integrating sentiment from social media, financial news, earnings reports, and analyst opinions provides a holistic view of market sentiment.

Keywords: CAPM model, finance model, asset pricing. stock prediction, sentiment analysis

1. Introduction

The classic model CAPM has been developed for many years, and there is a lot of research based on that. This research can be sorted by reviewing CAPM and modifying this model for different purposes. It can be used to price different objects or in different environments. However, it is not the only theory that can be used. There are other models as well. The most common application for that theory is stock price prediction; accurate information about price can benefit investing by getting the arbitrage space between the current price and real price. This article will present the conventional theories of asset pricing and focus on stock pricing methods, which are applications based on developed methods. The world of stock price prediction has evolved significantly in recent years, transcending traditional numerical models. In this rapidly changing landscape, understanding market sentiment, the collective emotional undercurrent that often dictates market behaviour, has become an essential element in forecasting stock prices. Sentiment analysis, a dynamic field at the intersection of natural language processing and machine learning, has emerged as a powerful instrument for decoding and quantifying the emotional aspects embedded within textual data. This review embarks on an exploration of the pivotal role of sentiment analysis in the realm of stock price prediction. It delves into the spectrum of techniques employed to extract sentiment from sources such as news articles, social media, financial reports, and other textual data, shedding light on using traditional

DOI: 10.54254/2754-1169/81/20241843

lexicon-based methods and advanced deep learning models. Moreover, we investigate how these sentiment-derived insights are seamlessly woven into predictive models, enriching the accuracy and comprehensiveness of stock price predictions. In an era of big data and technological advancement, sentiment analysis offers a new dimension to the age-old pursuit of predicting stock prices and, in doing so, provides investors and analysts with a valuable edge in the market.

2. Pricing based on classic theory

CAPM is a model that has bolstered the analysis of market shifts based on the prevailing market needs to reduce unforeseen risks [1]. By focusing on different aspects of this model, many models are produced for different applications. The following session will focus on different theories or models developed from this theory.

2.1. Review of CAPM

Based on the classic CAMP model, researchers verified whether those factors or calculating methods are still useful for today's market. If not, then they will provide one way to modify the model to allow this model to be updated. "The Capital Asset Pricing Model (CAPM) revolutionized modern finance [2]. This article illustrates the significance of the CAMP theory as the framework of finance pricing. Rossi, Matteo insights into the key ideas of the CAPM, the history of empirical work on the CAPM, and the implications of this work on the shortcomings of the CAPM. Ansari, Valeed A. in his paper verifies that if this model is still valid today. There are also many other papers focusing on this field [3].

2.2. Modivied CAPM for different objects

For example, the CAPM can be used to analyze the consumption. The consumption base (CAPM) has different parameters that assess the connection between asset and investors consumption decision over a given period. Therefore, it tends to consider changes that affect investor's decision on asset pricing. Study conducted by Ghosh et al., asserts that the introduction of machine learning has enhanced the power to estimate framework of consumption-based model accurately [4]. The algorithms integrated into machine learning can process a large volumes of data based on input to reveal information about an individual's consumption behaviour, demographic data, and preference for accurate decision-making [5]. The ability of machine learning to process large volumes of personal data gives room for accurate modelling for potential investors. In addition, processing large data sets allows investors to make changes on challenging quantities [1]. Machine learning is fundamental in stock assessment and management because it can show how a shift in consumer preference impacts asset pricing, thus offering robust information on consumption-based asset pricing. The modified model for specific areas is like different industries, but the point to focus is different. For the U.S. construction industry, there is a pricing strategy." Construction Management & Economics. A change of bidding procedure is proposed so that all parties in construction can maximize the benefits of market-based pricing strategies [6]. For different models, this is important as well; in a two-period pricing model for a new fashion style launching strategy, the author constructs a model for the new fashion field [7].

2.3. Modeified CAPM for different environments

For different environments, researchers adjust their strategies. They do not necessarily use the classic model and are more likely to have some specific method or factor to fit the situation of industries. "Financial markets have become increasingly global in recent decades, yet the pricing of internationally traded assets depends strongly upon local risk factors, leading to several observations that are difficult to explain with standard frameworks. Equity returns depend upon both domestic and global risk factors [8]. This paper shows that in a global environment, asset pricing needs to be adjusted. Also, in the Capital Asset Pricing Model and the Liquidity Effect: A Theoretical Approach, the author discussed the cost of liquidity to fit the model to the environment [9]. In A Novel Pricing Method for European Options based on Fourier-cosine series expansions, the author discusses the model for European options. By improving the factors of CAMP, a more accurate model can be made to achieve different purposes. CAMP is also called the temporal capital asset pricing model, and it incorporates the traditional CAMP by considering the changes in economic parameters on asset returns. Machine learning has advanced the role of ICAPM by giving room for the integration of real-time data sets for accurate and adaptive modeling [10]. Machine learning algorithms could process social media information, news releases, and economic indicators in time to factor such information into investors' preferences [11]. The ability of machine learning to factor all these parameters in asset capital has improved intertemporal asset pricing and gives investors timely information for a faster decision-making process.

2.4. Other models

There are also some new models, usually including a different number of factors. Commonly, there are 2 to 5 factors. The number of factors depends on how detailed the model is. However, it does not mean that more factor means better. A smaller number of factors requires less data to use, and sometimes approximate prediction can meet the needs. It is good to collect every detail, but it is not always easy to achieve. The single-index model is a critical model with a simple framework in a single unit of systematic risk dictated by the prevailing market indexes. Anwar, Jasuni & Juniarso asserts that incorporation of machine learning in single index model allow the model to identify hidden variables that affects market demand [2]. Factors affecting market gaps includes economic and information related variables. By considering these parameters, machine learning has enhanced the single-index model to give accurate and reliable information on asset returns above the traditional market index frameworks. The multiple-factor model plays a critical role in extending the asset pricing above the limitation of a single-unit model such as CAPM. For instance, the Fama-French three factors model incorporates value and size factors besides the market parameters. A study conducted by Rehnby on the Swedish stock market asserts that machine learning has been essential in boosting the ability of multiple-factor models. Machine learning could uncover "alpha signals" like technical indicators and news sentiment [5,12]. In addition, machine learning parameters could factor complex and non-linear relationships between different factors to deduce accurate pricing.

3. Stock price prediction

One of the most common uses of asset pricing is measuring stock price. The stock price might be affected by information or different pricing. Thus, accurately getting the stock price is important when predicting the future price.

3.1. Sentiment Analysis in Stock Price Prediction

Sentiment analysis, or opinion mining, is a crucial component in stock price prediction. It involves using natural language processing and machine learning techniques to extract and analyze sentiment or emotional information from textual data, such as news articles, social media posts, financial reports, and more. This section provides an in-depth review of sentiment analysis in the context of stock price prediction. The integration of sentiment analysis into stock price prediction models has gained significant attention in recent years, as sentiment plays a vital role in influencing market dynamics. Sentiment analysis aims to determine whether textual content contains positive, negative, or neutral sentiment and how this sentiment can be leveraged to improve stock price forecasts [13].

3.2. Lexicon-Sentiment Analysis:

Lexicon-based sentiment Analysis (LBSA) stands as a cornerstone technique in natural language processing, playing an integral role in forecasting stock movements by quantifying the sentiment embedded within textual data. LBSA hinges on using sentiment lexicons or dictionaries, which comprise predefined words or phrases, each tagged with sentiment labels, such as positive, negative, or neutral. These lexicons act as a compass for evaluating the sentiment associated with words or phrases found in the text. For example, words like "profit," "growth," or "opportunity" are typically linked to positive sentiment scores, while terms like "loss," "decline," or "uncertainty" tend to carry negative scores. The sentiment of a given text is determined by aggregating the sentiment scores of its constituent words or phrases, culminating in a quantitative measure of its overall sentiment. Within the context of stock price prediction, this technique is applied to diverse textual sources, including financial news articles and social media content, to assess market sentiment. The resultant sentiment scores are subsequently incorporated into predictive models to account for sentiment-driven variables influencing stock prices [13,14]. Based on the findings of the review study, there has been extensive prior research focused on sentiment analysis for predicting stock market movements. This research has involved the analysis of data gathered from diverse sources, including web scraping data from platforms like Twitter, Facebook and news headlines [15].

To illustrate, let's consider a financial news headline: "Tech Stocks Soar as Innovation Thrives." Lexicon-based sentiment Analysis dissects this headline into its individual components, such as "Tech," "Stocks," "Soar," "Innovation," and "Thrives." Each of these elements is cross-referenced with entries in the sentiment lexicon, resulting in the retrieval of associated sentiment scores. In this scenario, words like "Soar" and "Thrives" will likely yield positive sentiment scores. By summing or averaging these scores, the analysis produces an overall sentiment score for the headline, which is probably positive, echoing the optimistic sentiment conveyed in the text. This quantified sentiment can serve as an additional input in stock price prediction models, capturing the prevalent market sentiment and its potential sway over stock prices, thus heightening the models' predictive precision and comprehensiveness [14].

In practical application, the sentiment lexicons employed in LBSA are often more nuanced, assigning sentiment scores on a graduated scale to accommodate varying degrees of positivity or negativity. Furthermore, domain-specific sentiment lexicons, customized for the financial sector, are frequently brought into play to ensure the relevance and precision of sentiment analysis within financial contexts. Researchers and analysts rely on this methodology to process textual data from diverse sources, including financial news headlines, earnings reports, and social media posts, thereby unearthing sentiment-related insights that complement the conventional numerical data in stock price prediction models. This strategic amalgamation results in a more holistic comprehension of the emotional factors that sway market dynamics and stock price fluctuations [13,14]. In summation, Lexicon-Based Sentiment Analysis is a fundamental tool for quantifying sentiment within textual data, making it an indispensable asset for predicting stock movements. It leans on predetermined sentiment lexicons to allot sentiment scores to words or phrases in the text, empowering analysts and researchers to gauge market sentiment within financial news and other text-based sources. The quantified sentiment is subsequently woven into predictive models, offering an all-encompassing perspective on sentiment-induced elements impacting stock prices and contributing to more accurate and well-informed stock price projections.

DOI: 10.54254/2754-1169/81/20241843

3.3. Advanced Sentiment Analysis Techniques

Advanced sentiment analysis techniques have revolutionized the field of stock price prediction by providing a deeper and more nuanced understanding of the influence of sentiment on financial markets. These techniques leverage Natural Language Processing (NLP) and machine learning models, such as recurrent neural networks (RNNs) and deep learning, to extract sentiment from complex textual data. While lexicon-based analysis offers a valuable foundation, the integration of advanced sentiment analysis methods has gained prominence due to their ability to capture the subtleties of market sentiment. For instance, in the article by Wu et al., advanced sentiment analysis is employed as a feature extraction method using a combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory networks (LSTMs) to predict stock prices [13]. This approach enables the model to not only recognize the presence of sentiment but also discern the intricate relationships between sentiment-bearing words within the text. By incorporating CNNs, which are adept at capturing local patterns, and LSTMs, which excel in modeling sequential dependencies, the model can extract and analyze sentiment in a more context-aware and dynamic manner. This allows it to capture evolving sentiment trends and their impact on stock price movements.

In the study by Jing et al., a hybrid model that integrates deep learning with investor sentiment analysis is employed for stock price prediction [16]. Deep learning techniques, including neural networks and attention mechanisms, are utilized to analyze and incorporate market sentiment data from social media and news sources. This sophisticated approach goes beyond traditional sentiment analysis by considering the broader context in which sentiment is expressed. It can identify sentiment nuances, such as sarcasm or mixed emotions, which are challenging for lexicon-based methods to capture. This contextual understanding enhances the accuracy of stock price predictions.

Moreover, in the research conducted by Jin et al., LSTM-based models are employed to predict stock closing prices by considering sentiment dynamics in financial texts [17]. These models excel at capturing long-range dependencies and sequential patterns in textual data. They enable the modeling of how sentiment evolves over time, allowing analysts to identify sentiment trends and their correlation with stock price movements. This level of temporal analysis provides valuable insights for investors seeking to time their trades based on sentiment fluctuations.

Additionally, Li et al. demonstrate the role of advanced sentiment analysis techniques in a multilingual context by using deep learning to extract text-extracted investor sentiment from Chinese financial news articles [18]. This showcases the versatility of advanced sentiment analysis across languages and its relevance in diverse global financial markets. Advanced sentiment analysis techniques, empowered by NLP and deep learning models, offer a comprehensive and dynamic approach to understanding sentiment's influence on stock prices. They enable models to extract, analyze, and contextualize sentiment from complex textual data, providing a deeper understanding of market sentiment dynamics. These techniques are pivotal in improving stock price prediction accuracy and guiding investment decisions in today's rapidly evolving financial landscape.

3.4. Sentiment from Multiple Data Sources:

Sentiment analysis has transcended its traditional boundaries and embraced many textual data sources, including financial news, social media, earnings reports, and analyst opinions. Wu et al. underscore the significance of incorporating sentiment from multiple data sources into stock price prediction models, recognizing that market sentiment results from a complex interplay of diverse textual narratives [13]. Social media platforms, such as Twitter, Stock Twits, and Reddit, have emerged as influential arenas for discussing stocks and financial markets. Investors and traders actively share their viewpoints, news, and even rumors on these platforms, collectively shaping market sentiment. For example, a surge in positive sentiment expressed on Twitter about a specific stock can signify a growing wave of optimism among retail investors, potentially impacting trading volumes and stock prices [16].

Financial news articles represent another invaluable source of sentiment. These articles frequently feature expert opinions, analyses, and insights into market trends. By aggregating sentiment data from financial news, models can capture the sentiment conveyed by financial experts and analysts. For instance, a positive analysis of a company's earnings report in a reputable financial news outlet can elevate investor sentiment, potentially resulting in an uptick in the company's stock price [17]. Earnings reports and analyst opinions also provide substantial cues for sentiment analysis. When companies release their quarterly or annual earnings reports, these documents contain both quantitative data and qualitative information. Analysts offer commentary and forecasts about a company's performance. Advanced sentiment analysis techniques can extract sentiment from these reports and opinions, assisting investors in assessing market sentiment regarding a specific company. Positive earnings reports and bullish analyst opinions can fortify investor confidence and contribute to upward movements in stock prices [18]. In conclusion, sentiment analysis from multiple data sources furnishes a multi-dimensional perspective on market sentiment. Social media platforms, financial news, earnings reports, and analyst opinions all contribute to the intricate fabric of sentiment that influences stock prices. When advanced sentiment analysis techniques are applied to aggregate and dissect data from these sources, they equip researchers and investors with deeper insights into the ever-evolving realm of market sentiment, thereby guiding their trading and investment decisions [13].

4. Conclusion

In conclusion, the CAPM, as the keystone of asset pricing, has developed and been applied in many fields. Stock prediction is just a common application. There are many other applications; this article focuses on the latest techniques and models of stock prediction and the foundation of that. In the future, this theory will keep developing and improving with better data collection techniques and matured analysis methods, yet the foundation will still be significant. As information becomes more and more clear, the investment market will be more equal to investors and allow investors to compete in a clean and fair environment. Sentiment analysis has become a cornerstone in the realm of stock price prediction. It offers valuable insights into market sentiment, extracting emotional nuances from textual data that may not be evident through traditional financial analysis. This review has showcased a spectrum of techniques, from lexicon-based methods to advanced deep learning models, illustrating how they've enhanced the accuracy and depth of stock price forecasts. Incorporating sentiment from various data sources, such as social media, financial news, earnings reports, and analyst opinions, has provided a holistic view of market sentiment. As technology and data analysis methods continue to advance, sentiment analysis will remain a crucial asset for investors and researchers, offering a deeper understanding of the emotional factors that drive financial markets. It is not just a complementary tool but an integral part of financial decision-making, helping stakeholders navigate the complexities of the global financial landscape and make more informed and strategic investment choices.

Acknowledgement

Ruohan Zheng, Yile Li, and Yicheng Wang contributed equally to this work and should be considered co-first authors.

References

[1] Arthur, W. B., Holland, J. H., LeBaron, B., Palmer, R., & Tayler, P. (2018). Asset pricing under endogenous expectations in an artificial stock market. InThe economy as an evolving complex system II (pp. 15-44). CRC Press.
[2] Perold, André F. "The capital asset pricing model." Journal of economic perspectives 18.3 (2004): 3-24
[3] Ansari, Valeed A. "Capital asset pricing model: should we stop using it?." Vikalpa 25.1 (2000): 55-64.

DOI: 10.54254/2754-1169/81/20241843

[4] Ghosh, A., Julliard, C., & Taylor, A. P. (2017). What is the consumption-CAPM missing? An information-theoretic framework for the analysis of asset pricing models.The review of financial studies, 30(2), 442-504.
[5] Giglio, S., Kelly, B., & Xiu, D. (2022). Factor models, machine learning, and asset pricing.Annual Review of Financial Economics, 14, 337-368.
[6] Mochtar, Krishna, and David Arditi. "Pricing strategy in the US construction industry." Construction Management & Economics 19.4 (2001): 405-415.
[7] Zhou, Erfeng, et al. "A two period pricing model for new fashion style launching strategy." International Journal of Production Economics 160 (2015): 144-156.
[8] Lewis, Karen K. "Global asset pricing." Annu. Rev. Financ. Econ. 3.1 (2011): 435-466.
[9] Jacoby, Gady, David J. Fowler, and Aron A. Gottesman. "The capital asset pricing model and the liquidity effect: A theoretical approach." Journal of Financial Markets 3.1 (2000): 69-81.
[10] Sakemoto, R. (2022). Multi-scale inter-temporal capital asset pricing model.International Journal of Finance & Economics, 27(4), 4298-4317.
[11] Bagnara, M. (2022). Asset Pricing and Machine Learning: A critical review.Journal of Economic Surveys.
[12] Rehnby, N. (2016). Does the Fama-French three-factor model and Carhart four-factor model explain portfolio returns better than CAPM.A study performed on the Swedish stock market.
[13] Wu, S., Liu, Y., Zou, Z., & Weng, T. H. (2022). S_I_LSTM: stock price prediction based on multiple data sources and sentiment analysis. Connection Science, 34(1), 44-62.
[14] Turner, Z., Labille, K., & Gauch, S. (2020). Lexicon-based sentiment analysis for stock movement prediction. International Journal of Mechanical and Industrial Engineering, 14(5), 220-226.
[15] Das, N., Sadhukhan, B., Chatterjee, T., & Chakrabarti, S. (2022). Effect of public sentiment on stock market movement prediction during the COVID-19 outbreak. Social network analysis and mining, 12(1), 92.
[16] Jing, N., Wu, Z., & Wang, H. (2021). A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. Expert Systems with Applications, 178, 115019.
[17] Jin, Z., Yang, Y., & Liu, Y. (2020). Stock closing price prediction based on sentiment analysis and LSTM. Neural Computing and Applications, 32, 9713-9729.
[18] Li, Y., Bu, H., Li, J., & Wu, J. (2020). The role of text-extracted investor sentiment in Chinese stock price prediction with the enhancement of deep learning. International Journal of Forecasting, 36(4), 1541-1562.