Review Task 2

Evelyn Su 110208040

Selected Method: SVM Classification

Intro

In this review, I'm examining the use of SVM Classification in stock price forecasting. In these three articles, SVM Classification is utilized to perform binary predictions focusing on whether stock prices will increase or decrease. The first article enhances SVM's accuracy by preprocessing market data with a wavelet de-noising model. The second uses an ensemble of SVMs to analyze sentiments from Stocktwits via the FinBERT model, aiming to accurately predict stock price movements. Lastly, the third integrates SVM with multi-view learning, using both market data and financial news to predict the direction of stock price changes.

Article 1

A Hybrid Forecasting Method for Anticipating Stock Market Trends via a Soft-Thresholding De-noise Model and Support Vector Machine (SVM),
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4408128

- The data analyzed using SVM in the study comes from the Shanghai Stock Exchange Composite Index. This dataset includes historical stock price movements, which are used to predict trends such as upward and downward movements in the stock market.
- In this article, the SVM is employed as a critical tool for operationalizing and measuring the DV, which in this case is the direction of stock market trends—upward or downward. The SVM model utilizes features extracted from the processed stock market data to infer these trends, effectively serving as a bridge between the raw financial data and actionable market insights. The article argues that SVM, enhanced by advanced parameter optimization, can lead to more accurate and reliable predictions of market movements.
- The SVM is not used standalone but is integrated with a wavelet soft-thresholding de-noising model for data preprocessing. This strategic integration is vital as the de-noising model preprocesses the stock market data by removing noise components from the stock price data, thereby enhancing the signal-to-noise ratio. Such preprocessing improves the quality of the data input into the SVM, allowing it to more effectively identify the correct market trends from the enhanced data features, which may include various statistical measures and technical indicators used in financial analysis. Additionally, the article describes

the use of the Harmony Search (HS) algorithm to further enhance the SVM's performance. This algorithm automates the optimization of SVM parameters, such as the kernel function and penalty factors. By integrating HS, the model tuning process becomes less reliant on manual parameter selection, which traditionally is labor-intensive and prone to errors. This integration not only streamlines the modeling process but also significantly improves the prediction accuracy by finding the optimal parameter settings faster and more efficiently than traditional methods. This combination of noise reduction, machine learning modeling, and automated parameter optimization forms a powerful approach to increasing the accuracy of financial predictions.

Article 2

Stock price movement prediction based on Stocktwits investor sentiment using FinBERT and ensemble SVM, https://peerj.com/articles/cs-1403/

- The data being analyzed using the SVM in the study consists of sentiment analysis results derived from investors' comments on Stocktwits, a social media platform for investors. These sentiment analyses were conducted using FinBERT, a language model specifically fine-tuned for financial contexts. The SVM uses these sentiment scores as input features to predict the future movement of stock prices, classifying them into binary categories—price increase or decrease.
- SVM plays a critical role in the article by operationalizing and quantifying the DV, which is the future movement of stock prices. The IV — investor sentiment scores — are derived from advanced natural language processing. By using SVM, the study is able to methodologically infer the DV from the IV, effectively linking qualitative sentiment data to quantifiable market outcomes. This not only supports but also strengthens the article's argument that investor sentiments, as expressed on social platforms, can be quantitatively linked to stock market dynamics.
- In this study, SVM is not used as a standalone method but is part of a more extensive ensemble learning strategy. This approach uses multiple SVM models to predict stock price movements, which are then aggregated to form a single, more accurate prediction. This technique, known as bagging or bootstrap aggregating, helps in addressing the individual limitations of each model, thereby reducing variance and avoiding overfitting. This ensemble method allows for a more robust and reliable prediction, which is critical in the volatile domain of stock market forecasting.

Article 3

A hybrid model for stock price prediction based on multi-view heterogeneous data, https://link.springer.com/article/10.1186/s40854-023-00519-w

- In the article, the SVM model analyzes both daily news text and historical stock price data to predict stock movements. The news text captures market sentiment and impacts from current events, while the stock data reflects historical trends. This data is sourced from financial news platforms and stock exchanges, providing a comprehensive base for the SVM to make informed predictions.
- The SVM is used for operationalizing and measuring the DV, which in this case is the future direction of stock prices (up or down). The IVs would include the features extracted from the news text and the historical stock prices. The SVM model infers the DV (future stock price movement) from these IVs by classifying the expected outcome based on the input data.
- In the article, the SVM model is not used by itself but as part of a hybrid approach that combines multi-view learning with it. This hybrid model is particularly designed to handle and integrate heterogeneous data from multiple views—specifically, financial market data and news text. This integration helps to address the challenges of combining data with completely different characteristics, such as structured numerical data and unstructured text data, which are typically hard to merge effectively using conventional methods.
- The hybrid model, by leveraging the strengths of both SVM and multi-view learning, allows for the simultaneous input of these diverse data types, reducing information loss compared to methods that might only handle a single type of data or need to pre-process one type into the format of another. The MVL-SVM approach thus offers a robust framework for capturing a more comprehensive set of features relevant to stock price movements, which significantly enhances prediction accuracy and reliability over traditional models that use either market data or news data alone.

Evaluation

The use of SVM (Support Vector Machine) Classification in stock price forecasting, as demonstrated by three scholarly articles, illustrates both significant contributions and notable challenges. SVM enhances predictive accuracy through advanced preprocessing techniques like wavelet soft-threshold de-noising, which refines input data quality. It also successfully integrates with sentiment analysis tools such as FinBERT to extract market insights from social media, adding depth to the predictive models. Additionally, SVM's application in multi-view learning frameworks showcases its adaptability in processing heterogeneous data sources, thereby capturing a wider array of market influences. However, challenges persist, including the dependency on high-quality data, the inherent complexity of financial markets which can confound predictive models, and the computational intensity required for processing large datasets or running ensemble methods. These factors necessitate ongoing improvements in model accuracy, efficiency, and scalability to maintain the efficacy of SVM in dynamic financial environments.