Camille Esteves
IDC6940
9/15/2025

Summaries

Research Paper 1: Stock Market Forecasting Based on Text Mining Technology: A Support Vector Machine Method
Authors: Yancong Xie, Hongxun Jiang
Year: 2019

Goal: To forecast stock market tendencies (price movement/trend) for Chinese stocks by combining sentiment analysis of financial news with SVM-based models.

Why it is important: Because financial news has a significant influence on investor behavior and market movements, yet many forecasting models rely mainly on technical or historical data; adding sentiment aims to improve predictions. Also, there is scarce large-scale work on combining Chinese news sentiment and SVM in this way.

How it's solved (methods): They collected 2.3 million Chinese financial news items (from 2008-2015). They built a domain-specific stop-word dictionary and a sentiment dictionary. They then use text mining/sentiment extraction to produce features from news, and combine those with historical stock data. They try both *Support Vector Regression* (SVR) and *Support Vector Classification* (SVC) with parameter-optimization for SVM hyper-parameters (C, gamma etc.).

Results: The model shows that news content has a "significant influence" on stock market predictions. SVR (regression) does well in fitting stock price fluctuations, though with some time-lag. For classification (predicting upward/downward trend), SVC works well but the effective features and optimal hyperparameters differ depending on stock. Also, including days without news using expanded input vectors improves regression performance.

Limitations: There is a lag effect (i.e. news impact is not instantaneous). The classification model's performance depends heavily on how much relevant news there is. Also, the best hyperparameters vary per stock (so less generalizable). And because it's Chinese markets and news sources and language, some results may not directly transfer to US/X/Twitter settings.

Research Paper 2: Using News to Predict Investor Sentiment: Based on SVM Model
Authors: D. Wang et al.
Year: 2020

Goal: To build an SVM-based model that uses news text to predict investor sentiment and then use that sentiment as an input (or indicator) for financial predictions.

Why it is important: Investor sentiment is a latent force that often moves markets in ways not captured purely by past prices or technical indicators. If you can reliably estimate sentiment from news, it could improve prediction accuracy or give early signals. This is especially relevant in volatile or news-driven market periods.

How it's solved (methods): They collect news texts, preprocess them (tokenization, cleaning etc.), extract features (likely TF-IDF or similar text features), then use SVM to classify sentiment. Then they evaluate how well predicted sentiment aligns with actual investor behavior/market moves. (They measure accuracy and possibly other metrics.)

Results: The SVM model achieves "high accuracy" in classifying investor sentiment via news. This suggests that news text is a strong feature. The sentiment predictions can be used to predict or correlate with market changes.

Limitations: The granularity of sentiment (how many classes, neutral vs ambiguous cases) is usually coarse; news may have delays or noise. Also, it's not always clear how large the lead/lag effect is. They may also have used aggregated sentiment, which can wash out signals. And generalization to other markets/languages/news sources might be limited.