

Camille Esteves
IDC6940
9/6/2025

Article Summaries

The first article, **Financial Sentiment Analysis: Techniques and Applications (Du, 2024)**, is a broad introduction to how sentiment analysis is used across the financial sector. It explains the two major approaches: lexicon-based methods, which rely on predefined dictionaries of positive and negative words, and machine learning methods that learn patterns from labeled data. The article emphasizes that financial text (like earnings reports, analyst notes, or social media) is very noisy and context-specific, which means preprocessing steps like tokenization, stopword removal, and handling domain-specific jargon are critical. This is directly relevant to our project because we'll be cleaning and preparing X posts, which are also messy and filled with slang, abbreviations, and sometimes sarcasm. Another key point is that the article highlights how the effectiveness of models depends heavily on feature representation—like using TF-IDF or word embeddings—which aligns with our plan to transform X posts into numeric input for SVM.

The second article, **Sentiment Analysis Stock Market: Sources and Challenges (Al Multiple)**, focuses on where sentiment data usually comes from and what makes it hard to use effectively in finance. It points out that news headlines, social media platforms like X, and RSS feeds are the most common sources, but each comes with its own problems. For example, short posts on X may lack context, and sarcasm or spam can mislead sentiment models. This is important for us since our dataset will likely face these exact issues, so we'll need to account for noise when preprocessing. The article also outlines the general workflow for sentiment analysis: data collection, preprocessing, feature extraction, model training, and evaluation. That workflow is almost identical to the one we'll follow in our capstone, which reassures us that our plan is on the right track. Finally, it stresses that combining sentiment with structured financial data (like stock prices) is where the real predictive power comes in, which is exactly what we're trying to test in linking X sentiment to stock movements.