

Stock Volatility & Sentiment

Does News Sentiment Improve Volatility Prediction?

Capstone Project Report

A model that classifies a stock's coming week as Low, Medium, or High volatility from price history and news-sentiment features — built primarily to answer one question through a controlled experiment: does the sentiment signal actually improve the prediction?

Arel · RLcapstone.ai

1. Summary

This project classifies a stock's coming week as **Low, Medium, or High volatility** — how much it moves, not its direction — from 12 features drawn from price history and daily news sentiment. It was built primarily to answer one question through a **controlled experiment**: does the sentiment signal actually improve the prediction?

The model reaches **57% accuracy** on a three-class problem where chance is 33%, because volatility clusters over time. The decisive result is the ablation: removing the sentiment features entirely yielded **57.5%** — marginally **higher** than the 57.1% obtained with sentiment. Sentiment added no predictive value.

Not financial advice. This is an educational machine-learning project. It predicts volatility, not price direction, and nothing here is a recommendation to buy, sell, or hold any security.

2. Motivation

Predicting whether a stock rises or falls tomorrow is close to a coin flip. Predicting how much it will move is more tractable, because volatility exhibits **autocorrelation** — calm periods tend to follow calm periods and turbulent ones follow turbulent ones. A project framed around news sentiment should also **test whether the sentiment signal earns its place** rather than assuming it does.

3. Data and features

Daily prices for 59 stocks were obtained from Yahoo Finance, and ticker-level news sentiment from a public 2023 dataset (Polygon). From these, **12 features** were engineered per stock per day — trailing returns, rolling volatility, volume, drawdown, and a daily sentiment score.

No look-ahead bias: every feature uses only past information, the target is the **next** week's realized volatility, and the train/test split is performed **by time**, so the test weeks all follow the training weeks.

4. Model and experiment

The classifier is a compact **multilayer perceptron** (3,397 parameters, 20 KB) that maps the 12 features to Low / Medium / High. To test the sentiment signal directly, an **identical model** was trained a second time on price features only — sentiment removed — and the two were compared.

5. Results

Model (on unseen weeks)	Accuracy
Price + sentiment	57.1%
Price only	57.5%

Chance baseline	33.3%
-----------------	-------

The model clears the chance baseline comfortably, but adding sentiment made it **0.4% worse** — no meaningful difference. This is a normal, honest outcome: a once-daily average of headline sentiment is a coarse, sparse signal, and the rolling-volatility features already capture most of what is predictable about next week's turbulence. **Reporting the negative result — rather than omitting the ablation and claiming sentiment 'works' — is what distinguishes an experiment from a demonstration.**

Error analysis

The errors fall where expected: the model reads calm weeks (Low, 75% recall) and turbulent weeks (High, 57%) reasonably well, but the Medium class is difficult (33% recall) because intermediate volatility sits on a blurred boundary between the other two. This is an interpretable failure mode rather than a mysterious one.

6. Deployment

The model was exported to ONNX for the browser demonstration and to TensorFlow Lite for the Android application (Risk Explorer). Both reproduce the original model, and the demonstration lets the user pick a real stock and week and compare the prediction against what actually happened.

7. Limitations

- Educational project only — not financial advice and not a recommendation.
- It predicts volatility, not direction, and is correct roughly half the time.
- A single year of data and a coarse daily sentiment signal.

8. Future work (v2)

For sentiment to earn its place it must be richer than a daily average: per-headline embeddings rather than a single number, a multi-year history spanning additional market regimes, and a test of whether sentiment contributes more to **direction** prediction than to volatility. Those results will be reported honestly — including if sentiment still does not help.