NLP-Driven Sentiment Analysis in the Financial Market

1.Commonly Referred Sentiment Analysis Websites

A variety of online platforms and research initiatives offer sentiment analysis tools for the financial markets.

Some of the most frequently referenced sources:

- **Nasdaq FinSentS News Sentiment:**

https://data.nasdaq.com/data/NS1-finsents-web-news-sentiment/documentation

- **Assets Covered:** Over 23,000 global equities.
  - 5,000 North American stocks
  - 8,000 European stocks
  - 4,000 Japanese stocks
  - 14,000 other Asian stocks
  - 3,000 Australia/New Zealand stocks
  - 1,000 South American stocks
  - Tens of thousands of private companies, major indices, currencies, and commodities
- **Method:**
  - Utilizes proprietary machine- & deep-learning text analysis from news, blogs, filings, and social media
  - Generates metrics such as sentiment score, high/low intraday scores, news buzz, and volume indicators
  - However, the exact scoring algorithm and model details are not publicly disclosed

- **Investopedia Anxiety Index (IAI):**

https://www.investopedia.com/anxiety-index-explained/

- **Assets Covered**: Broad U.S. market sentiment; not specific to individual tickers.
- **Method**:

  - Tracks U.S. page views for anxiety-related topics (based on ~1 billion page views)
  - Normalized using historical baselines and compared versus defined article categories

- **IBD & Investors Intelligence:**

https://www.investorsintelligence.com/

https://www.investors.com/

- **Assets Covered**: U.S. stock market, particularly retail investor sentiment.
  - U.S. equities, with a focus on retail investor sentiment
  - Includes weekly readership surveys, newsletter writers, and advisor sentiment indexes
- **Method:** Uses surveys, margin debt, and ratios like put/call volume to estimate market sentiment. Aggregated by contrarian indicators.

  - Aggregates advisor newsletter sentiment (bullish vs bearish)

  - Uses contrarian measures: put/call ratios, margin debt, VIX, etc.

  - Numerical readings signal extremes

- **ISEE Sentiment Index:**

https://www.investopedia.com/terms/i/iseesentiment.asp?

- **Assets Covered:** U.S. options market (primarily equities options).
- **Method:** Ratio of opening long call to put positions among individual investors (excluding market maker data)

  - Ratio = (Number of retail long call buys) ÷ (Number of retail long put buys) × 100

  - Excludes institutional/market-maker trades, focusing only on new retail positions

- **State Street Investor Confidence Index:**

https://www.investopedia.com/terms/s/state-street-confidence-index.asp?

- **Assets Covered:** Global equities, segmented by region
  - Global institutional portfolio holdings across ~45 countries
  - Divided into regional sub-indices: North America, Europe, Asia-Pacific
- **Method:** Aggregates institutional trading behavior to measure changes in actual risk allocation (equity weighting changes)

  - Measures equity weighting shifts in real institutional portfolios

  - Index rebased in May 2009: 100 = neutral risk appetite; >100 = increased equity exposure

  - Data updated weekly, released last Wednesday monthly

- **FinBERT:** A transformer-based NLP model fine-tuned for financial texts, widely used in academic and industry research.

https://github.com/ProsusAI/finBERT?

- **Assets Covered:** Broad financial texts: news, disclosures, earnings calls, reports—for any asset (stocks, bonds, crypto, etc.); Not tied to a specific asset class

- **Method:** Transformer-based NLP model (BERT) fine-tuned on financial sentiment classification tasks using labeled datasets.

2. Accuracy and Proofs

While many sentiment indices are proprietary, some provide validation evidence:

- **Nasdaq FinSentS News Sentiment:**

  - No public backtest results, validation data, or statistical analysis are available

- As a commercial offering, detailed performance reporting remains proprietary

- **Investopedia Anxiety Index (IAI):**

  - Displays a strong correlation with VIX over nearly a decade

  - Detected elevated anxiety > 1 year before the 2008 financial crisis—an "early warning signal"

  - No detailed backtest results or academic statistical validation published

- **IBD & Investors Intelligence:**

  - Frequent anecdotal use and media coverage around market turning points

  - QuantifiedStrategies noted no published backtests, evidence is practitioner-based

- **ISEE Sentiment Index:**

  - Useful at bullish/bearish extremes—apanalysts note correlation with market turns

  - No formal published backtests found online; validation remains practitioner-focused

- **State Street Investor Confidence Index:**

  - Not designed as a prediction tool—rather a risk-appetite gauge

  - Observations indicate confidence often leads price moves

- No formal backtesting or statistical performance data publicly available

- **FinBERT:**

  - Achieves 97% on full-agreement subset of Financial PhraseBank; 86% on noisy labels, which surpasses SOTA ~6–15% gains

  - ResearchGate papers confirm superior performance vs general BERT and classical ML methods

  - Recent studies show predictive power: e.g., Bayesian-enhanced FinBERT yields >70% F1 score and profitable backtest on SPY returns

3. Python Tools and Code for Sentiment Analysis

Several open-source Python projects and libraries support NLP-based sentiment analysis in finance:

- **FinBERT (HuggingFace/ProsusAI):** Finance-specific sentiment classifier with high accuracy. Open-source.
- **VADER:** Lightweight rule-based sentiment scorer. Works well for short financial text.
- **Backtrader:** Backtesting framework compatible with Pandas/YFinance. Integrates well with signals.
- **Fastquant:** Beginner-friendly strategy backtester; works with price/signal dataframes.
- **SAStocks / GenAI Algo (GitHub):** Integrated NLP + trading bots using FinBERT, OpenAI, VADER. Used for prototyping.

Summary

Among the tools evaluated, the Investopedia Anxiety Index and FinBERT offer the most validated and reproducible sentiment insights. FinBERT offers the strongest sentiment classification accuracy backed by financial-domain training and peer-reviewed benchmarks, while IAI provides empirical support for macro-level prediction. For trading system integration, using FinBERT or VADER with Backtrader or Fastquant provides a complete, reproducible framework for signal generation and evaluation.
Despite the lack of transparency from some commercial platforms, the open-source ecosystem provides strong alternatives for both sentiment generation and performance validation.

Reference Website Links and Citations

Araci, Duygu.
"FinBERT: Financial Sentiment Analysis with Pre-Trained Language Models." *arXiv preprint arXiv:2006.08097*, 2020.
https://arxiv.org/abs/2006.08097

Investopedia.
"Investopedia Anxiety Index (IAI)." *Investopedia*, Dotdash Meredith,
https://www.investopedia.com/investopedia-anxiety-index-5181205

Nasdaq.
"News Sentiment Analytics (FinSentS)." *Nasdaq Data Link*, Nasdaq,
https://www.nasdaq.com/solutions/news-sentiment-analytics

Investors Intelligence.
"Advisors Sentiment Survey." *Investors Intelligence*,
https://www.investorsintelligence.com

ISEE Index (ISE Sentiment Index).
"ISEE Sentiment Index." *Options by ISE*, Nasdaq,
https://www.nasdaq.com/solutions/isee-sentiment-index

State Street Global Markets.
"Investor Confidence Index (ICI)." *State Street Corporation*,
https://www.ssga.com/us/en/institutional/ic

ProsusAI.
"ProsusAI/finbert." *Hugging Face Model Hub*,
https://huggingface.co/ProsusAI/finbert

VADER Sentiment.
Hutto, C.J., and Eric Gilbert. "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text." *Proceedings of the Eighth International Conference on Weblogs and Social Media (ICWSM-14)*, 2014.
https://github.com/cjhutto/vaderSentiment

yfinance Developers.
"yfinance: Download Market Data from Yahoo! Finance." *GitHub*,
https://github.com/ranaroussi/yfinance

Backtrader.
Backtrader Documentation.
https://www.backtrader.com/docu/

Fastquant.
Ampil, Enzo. "FastQuant: Backtesting Made Simple." *GitHub*, https://github.com/enzoampil/fastquant

EthanFalcao.
"Sentiment Analysis in Financial Markets." *GitHub*, 2023. https://github.com/EthanFalcao/sentiment-analysis-in-financial-markets

Ja-Crispy.
"FinBERT Sentiment Analysis at Scale." *GitHub*, https://github.com/ja-crispy/finbert-sentiment-analysis

SaloniJhalani.
"Stock Market Live Sentiment Dashboard." *GitHub*, https://github.com/SaloniJhalani/Stock-Market-Live-Sentiment

NerdyBurner.
"SAStocks – News-Based Trading Bot Using VADER, GPT, and Polygon.io." *GitHub*, https://github.com/NerdyBurner/SAStocks