Digital Finance 3: Technology in Finance

Lesson 32: NLP in Finance

FHGR

January 3, 2026

NLP unlocks insights from the vast ocean of financial text data.

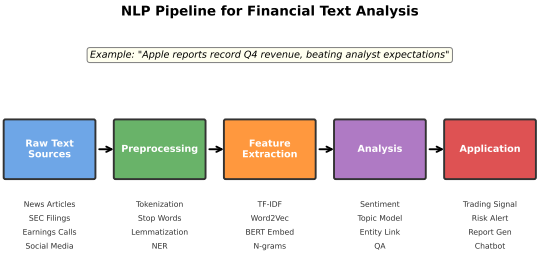
Learning Objectives

By the end of this lesson, you will be able to:

- Apply text preprocessing techniques to financial text
- Understand sentiment analysis for market prediction
- Explain word embeddings and transformer models
- Extract information from financial documents
- Evaluate NLP model performance in finance
- Recognize limitations of language models

Text preprocessing includes tokenization, stemming, and stop word removal.

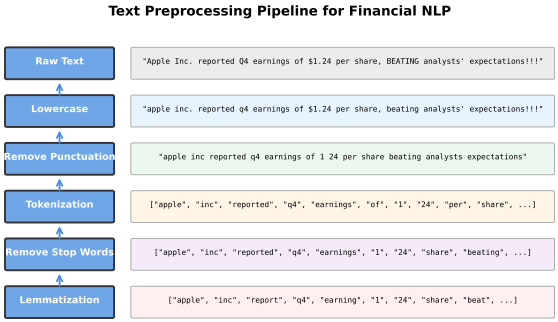
NLP Pipeline in Finance



Source: Jurafsky & Martin (SLP3), Loughran-McDonald, Kearney & Liu (2014)

Financial NLP pipelines transform raw text into structured features for downstream analysis.

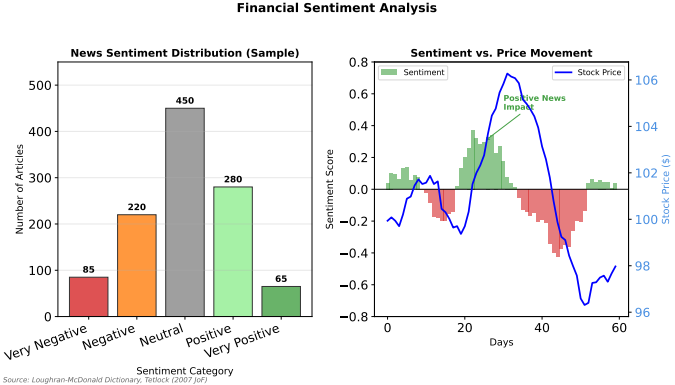
Text Preprocessing Steps



Source: nltk.org, Jurafsky & Martin (SLP3), spacy.io

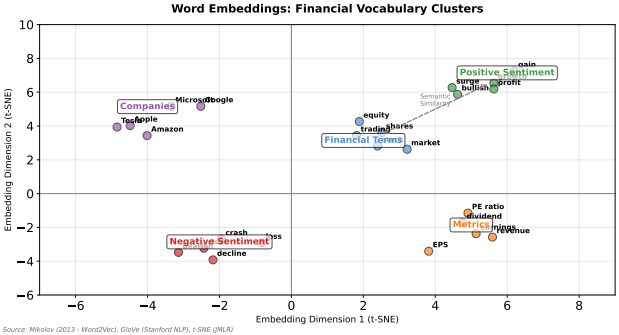
Preprocessing includes tokenization, stopword removal, stemming, and normalization of financial text.

Sentiment Analysis Methods



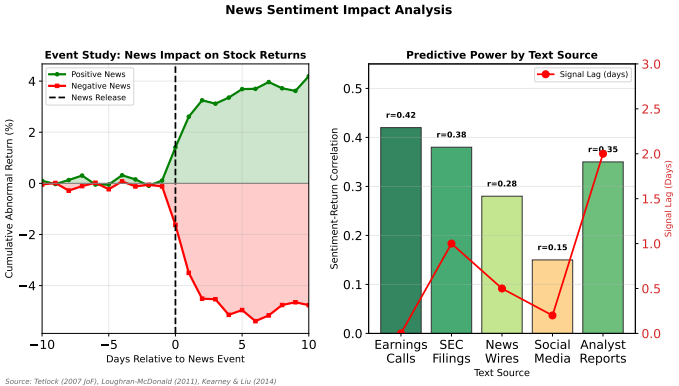
Sentiment analysis classifies financial text as positive, negative, or neutral using lexicons or ML models.

Word Embeddings Visualization



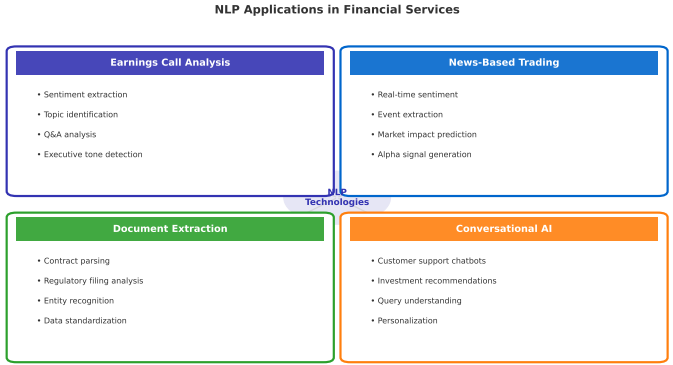
Word embeddings capture semantic relationships by representing words as dense vectors in continuous space.

News Sentiment Impact on Prices



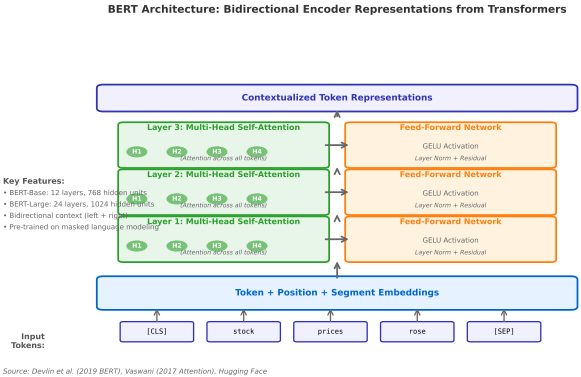
News sentiment correlates with short-term price movements but predictive power decays rapidly.

NLP Use Cases in Finance



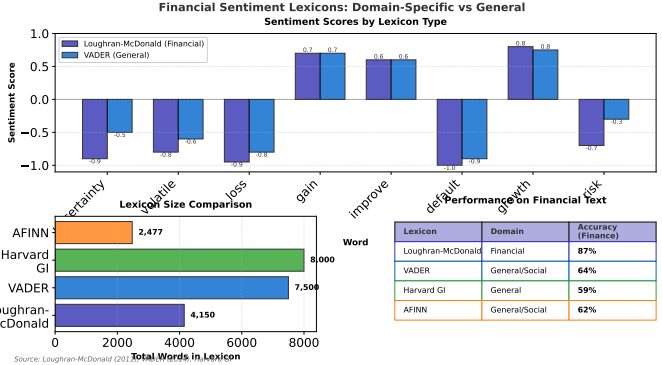
NLP applications span from sentiment analysis to automated report generation.

BERT Architecture



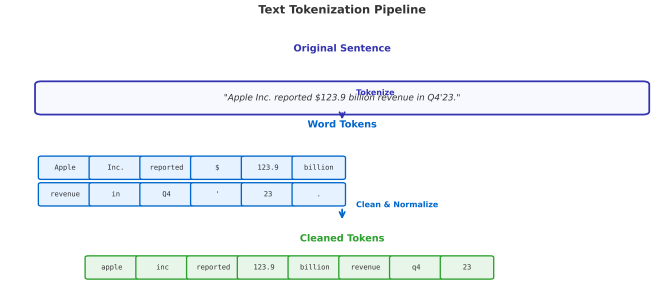
BERT revolutionized NLP by training bidirectional context representations.

Sentiment Lexicons for Finance



Domain-specific lexicons capture financial sentiment more accurately than general dictionaries.

Text Tokenization Example



Tokenization breaks text into meaningful units for NLP processing.

Nlp In Finance: Key Takeaways

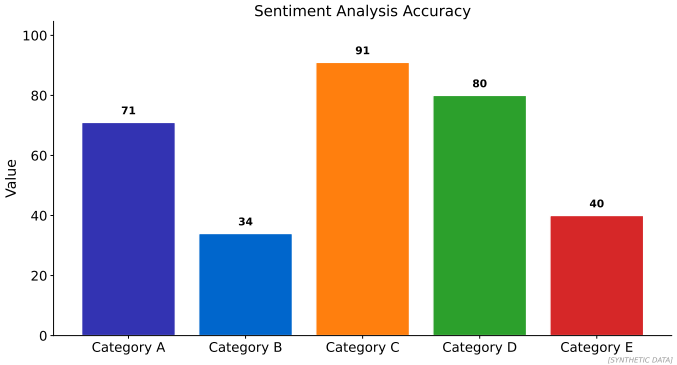
Key Takeaways:

- NLP extracts signals from unstructured financial text
- Sentiment analysis provides market mood indicators
- Word embeddings and transformers revolutionized finance NLP
- Applications: earnings call analysis, news trading, document extraction
- Challenges: domain adaptation, sarcasm, context-dependency
- LLMs (GPT, BERT) now dominate financial NLP tasks

Next Lesson: Robo-Advisors ML Aspects

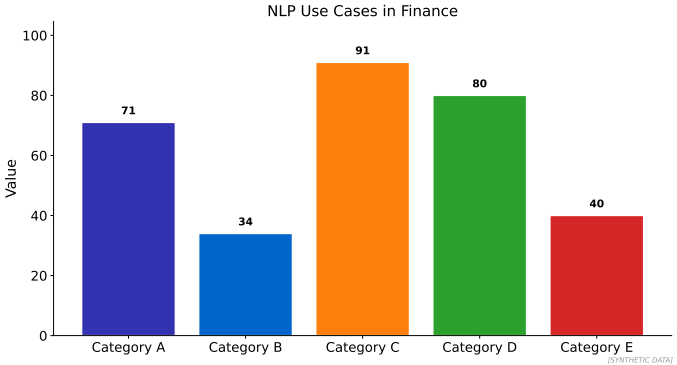
NLP pipelines transform unstructured text into structured features for analysis.

Sentiment Analysis Accuracy



Domain-specific models outperform general approaches.

NLP Applications in Finance



Document processing and sentiment lead adoption.