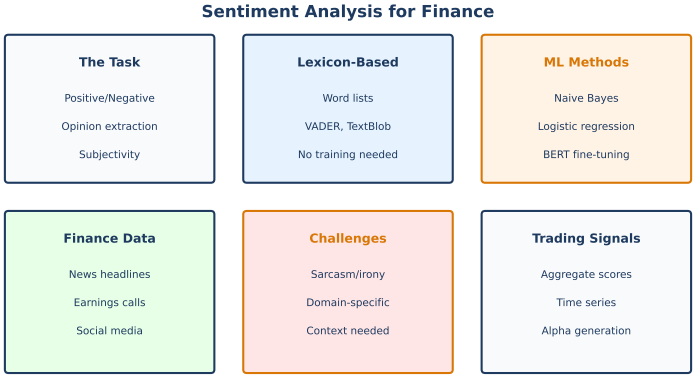
Lesson 40 Summary: Sentiment Analysis

Data Science with Python – Key Concepts

Data Science Program

Sentiment Analysis Overview



Extract opinions and emotions from text

The Task

Classify text polarity:

- **Binary:** Positive vs negative
- **Fine-grained:** 1-5 stars
- **Aspect-based:** Sentiment per topic

Key application for financial NLP

Lexicon-Based Methods

Use word lists:

- **VADER:** Social media optimized
- **TextBlob:** Simple polarity score
- **Advantage:** No training required

Quick baseline, works out-of-box

ML-Based Methods

Learn from labeled data:

- **Classic:** Naive Bayes, SVM, LogReg
- **Deep:** LSTM, BERT, transformers
- **FinBERT:** Finance-specific BERT

ML methods outperform lexicons with training data

Finance Applications

Trading signals from text:

- **News:** Headlines, articles
- **Earnings:** Call transcripts
- **Social:** Twitter, Reddit

Aggregate sentiment predicts returns

Quick Reference

Essential Commands:

Task	Code
VADER	`SentimentIntensityAnalyzer().polarity_scores()`
TextBlob	`TextBlob(text).sentiment.polarity`
ML pipeline	`TfidfVectorizer() + LogisticRegression()`
Hugging Face	`pipeline('sentiment-analysis')`

Sentiment analysis is foundational for finance NLP