SENTIMENT INDICATORS IN FINANCIAL TIME SERIES ANALYSIS

Michael Langeder
Bernadett Neuhauser
Johannes Voraberger

https://github.com/miicchhii/fhtw-bdeng-project

AGENDA

01 Motivation

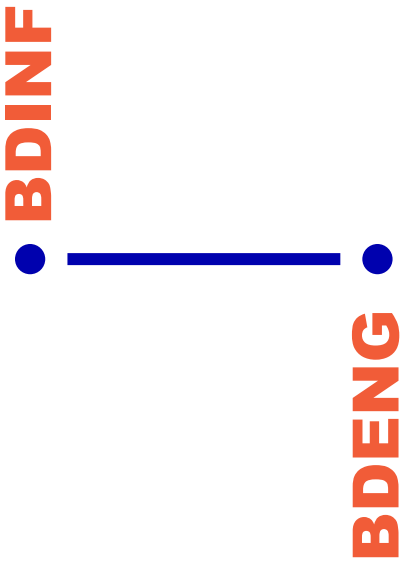
02 Dataflow

03 Data Transformation

04 Sentiment Analysis using FinBERT

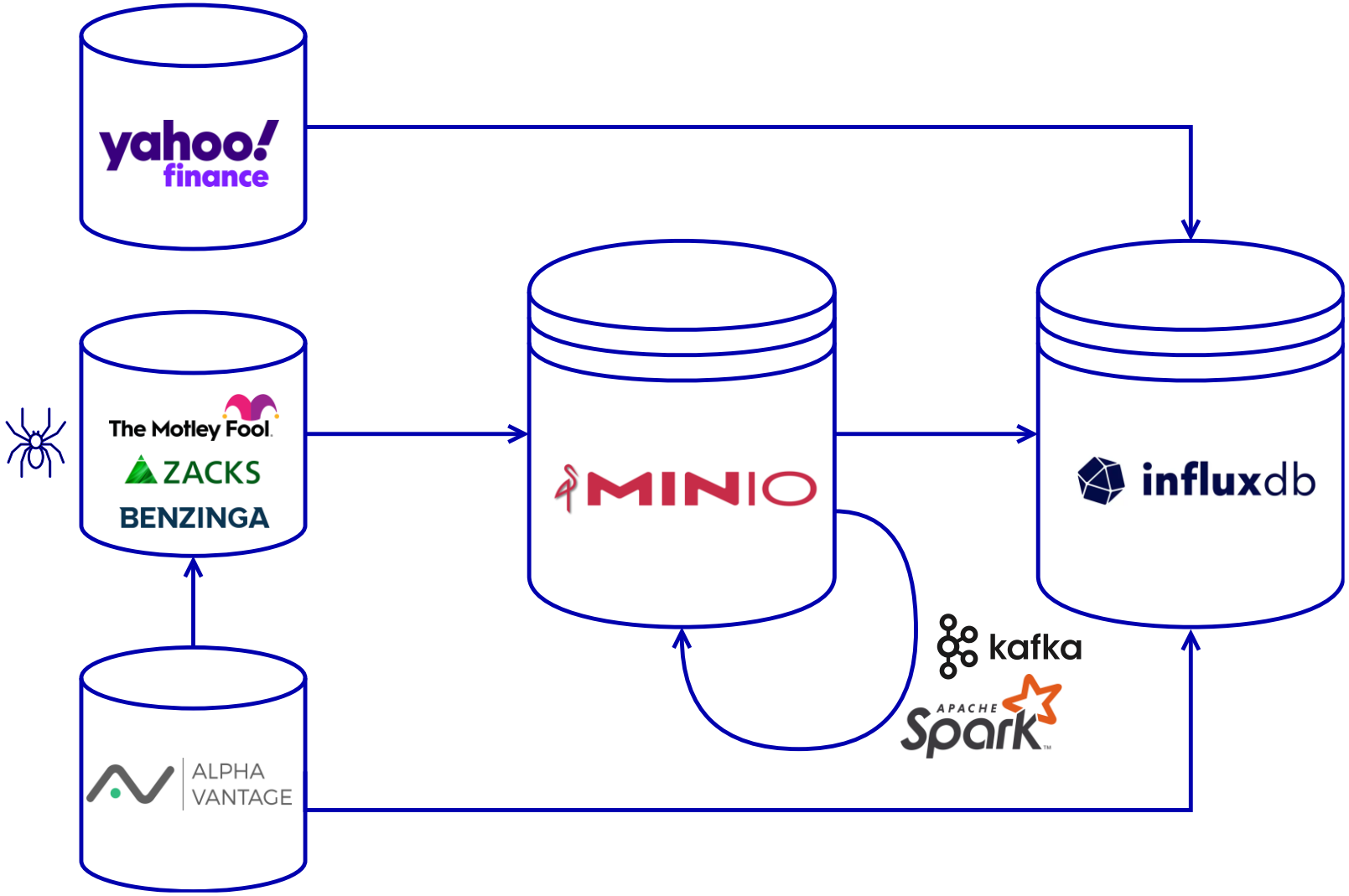
05 Results

MOTIVATION



- Lack of Transparency in Alpha Vantage‘s Sentiment scoring
- Desire for Full Control over data pipeline
- Reproducibility for Academic Use

DATAFLOW


yahoo! finance
The Motley Fool
ZACKS
BENZINGA
ALPHA VANTAGE
MINIO
influxdb
kafka
APACHE Spark

TRANSFORMING SCRAPED ARTICLES

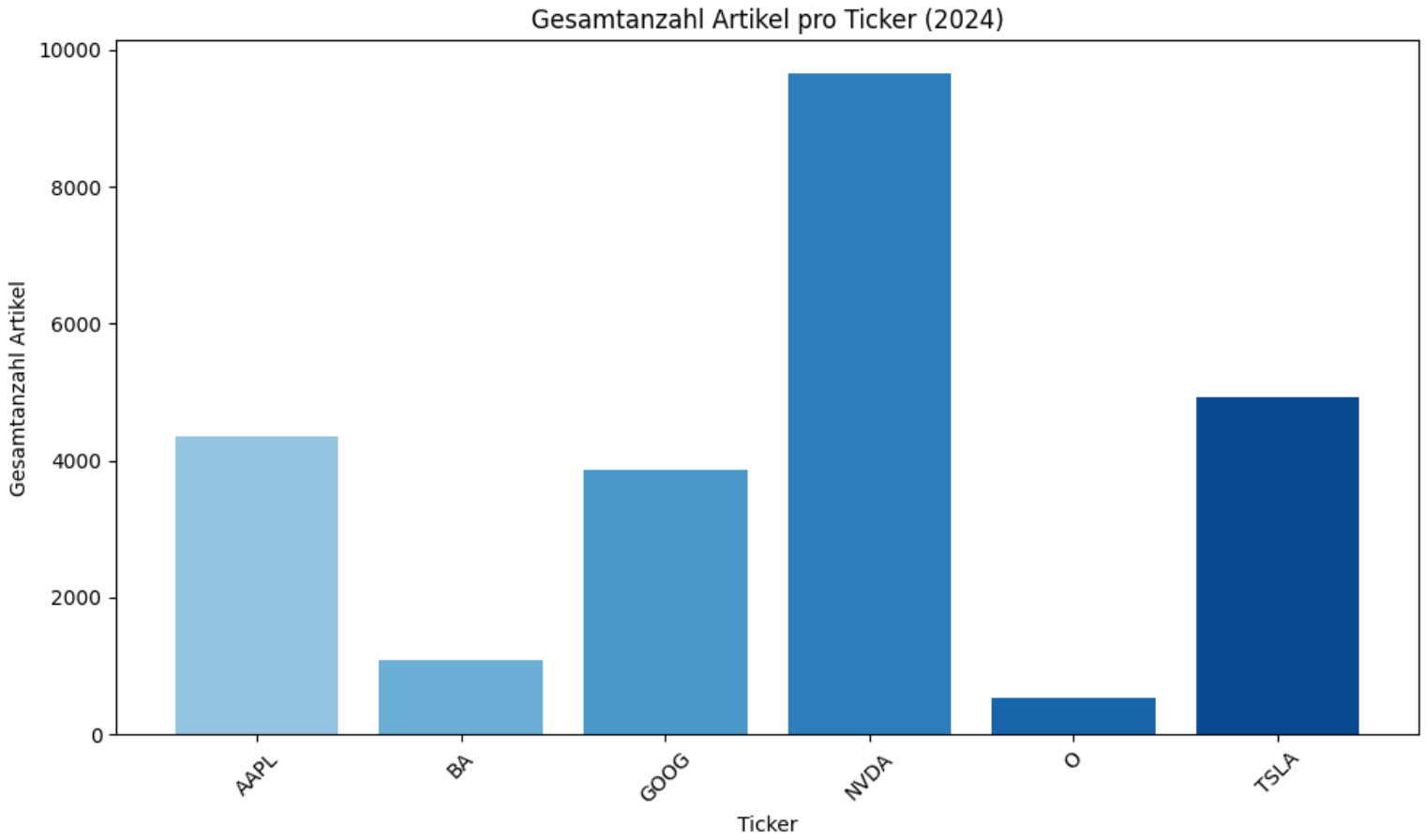
retrieve metadata	scraping	parsing	sentiment analysis	sentiment aggregation
use Alpha Vantage API to retrieve article metadata.	Extract URLs pointing to Zacks, Benzinga, and Motley Fool. scrape full HTML content of each article for analysis Save results to **MinIO**	Extract publishing date and article body from HTML process data using **Spark**. Save results to **MinIO**	a **Kafka** Producer pushes article data to the topic. a **FinBERT** Client listens to the topic and runs sentiment analysis per article	Group articles by publishing date and calculate average sentiment score per day using **Spark**. Save results to **InfluxDB**

SENTIMENT SCORING USING FINBERT

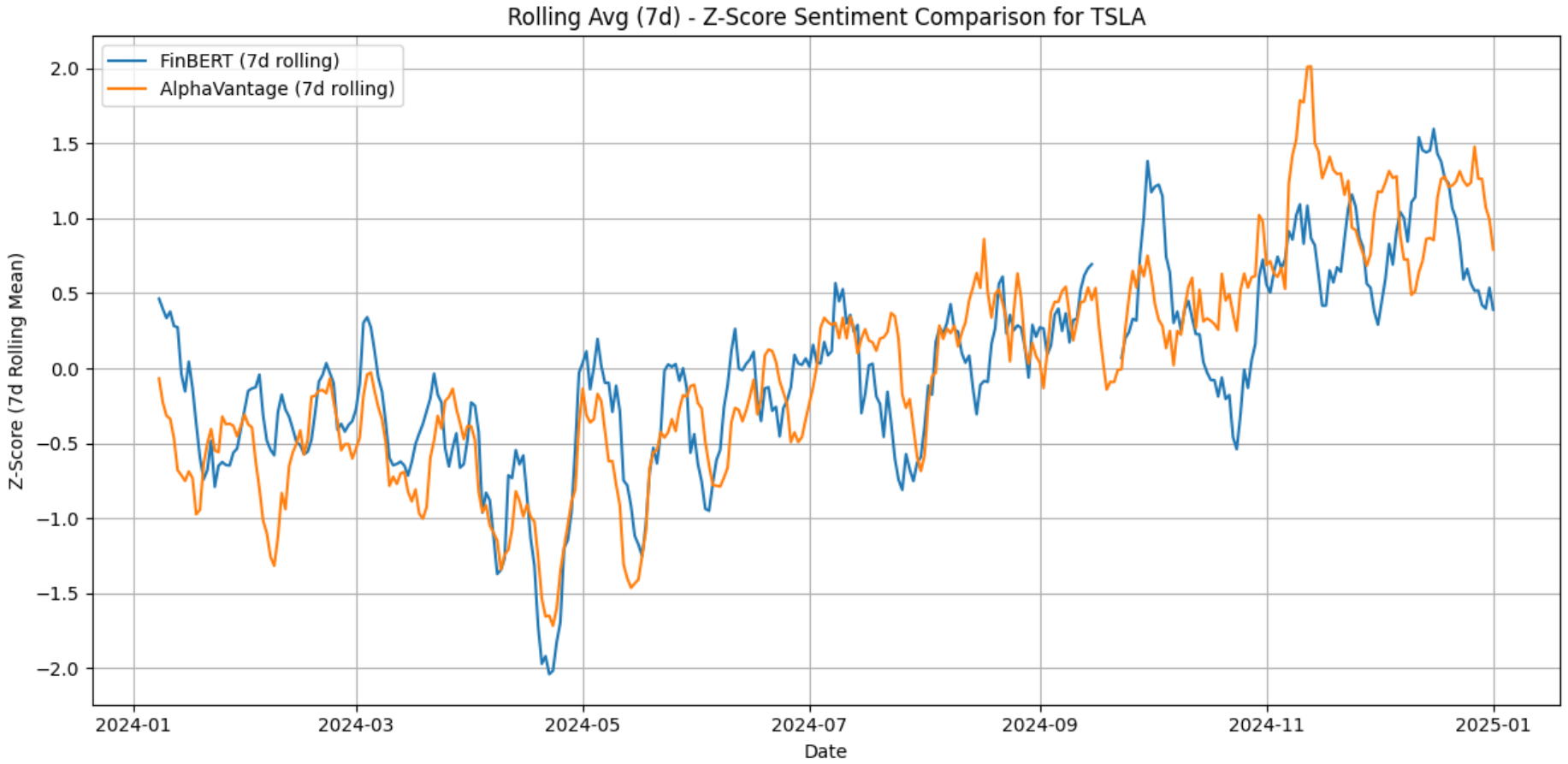
Hugging Face

- pre-trained NLP model specialized in financial sentiment
- Built on BERT, fine-tuned with Financial PhraseBank
- Classifies sentiment as:
  - Positive (1)
  - Neutral (0)
  - Negative (-1)

MEDIA COVERAGE



COMPARISON OF SENTIMENTSCORES



SENTIMENT INDICATORS IN FINANCIAL TIME SERIES ANALYSIS