Sentiment Analysis of Ethereum on X(Twitter): Insights from Social Media Trends

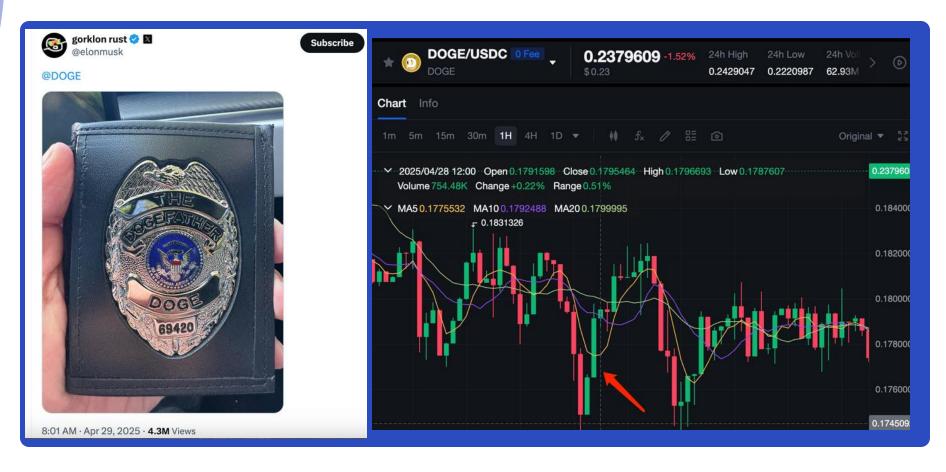






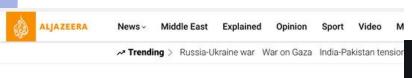
THE IMPACT OF SOCIAL MEDIA





THE IMPACT OF SOCIAL MEDIA





News | Crypto

Argentina's Javier Milei faces fraud allegations over cryptocurrency post

Critics have accused Milei of involvement in the rapid rise and fall of a cryptocurrency, costing investors thousands.



Argentina's President Javier Milei has positioned himself as a champion of the free market [Natacha Pisarenko/AP Photo]



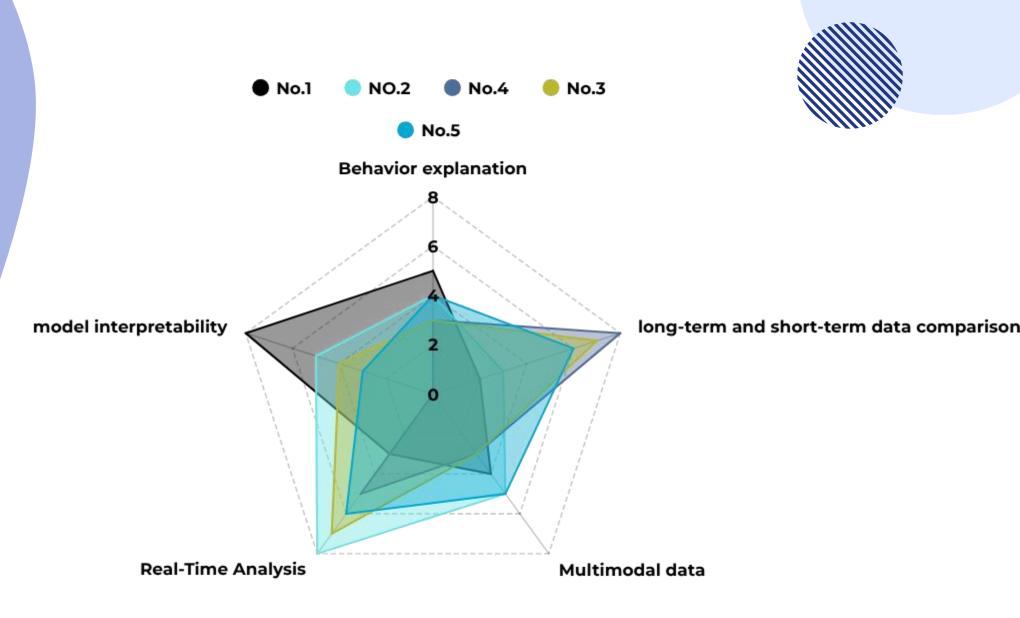
CONCEPTS IN ETHEREUM SENTIMENT ANALYSIS



COMPARISON OF PREVIOUS RESEARCH

No.	Title & Authors	Year	Methodology	Key Findings
1	Twitter Sentiment Analysis on the Cryptocurrency Market Frank van Engelen & Levente Kulcsár	2023	VADER sentiment analysis on 5M+ tweets, Pearson correlation, Cointegration, Granger causality tests	Long-term sentiment-price cointegration found for ETH. Only BTC showed short-terr Granger causality from returns to sentimen
2	Cryptocurrency Price Prediction Based on Social Network Sentiment Analysis Using LSTM-GRU and FinBERT Abba Suganda Girsang & Stanley	2023	FinBERT sentiment scoring, hybrid LSTM- GRU deep learning model	FinBERT-based sentiment significantly improved ETH and Solana price prediction accuracy.
3	Pump It: Twitter Sentiment Analysis for Cryptocurrency Price Prediction Vladyslav Koltun & Ivan P. Yamshchikov	2023	VADER sentiment, LSTM, NHITS forecasting on 567K+ tweets	Sentiment data enhanced model performance across all market phases; dail Twitter sentiment closely aligned with cryp price movements.
4	FinBERT-BiLSTM: A Deep Learning Model for Predicting Volatile Cryptocurrency Market Prices M.F. Hossain et al.	2024	FinBERT for sentiment, BiLSTM deep neural networks	Integrating financial sentiment with BiLSTN yielded high accuracy in predicting volatile cryptocurrencies like ETH.
5	Transformer-Based Approach for Ethereum Price Prediction Using Cross- Currency Correlation and Sentiment Analysis Shubham Singh & Mayur Bhat	2024	Transformer model combining sentiment and inter-crypto correlations	Outperformed ANN and MLP in ETH price forecasting. Transformer architecture effectively learned sentiment-driven marke trends.

IDENTIFIED RESEARCH GAPS





RESEARCH POSITIONING

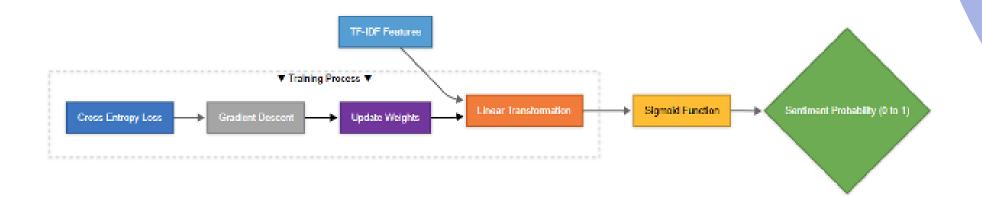
Enhances understanding of why sentiment affects ETH prices, grounded in behavioral finance Captures richer
Captures richer
emotional cues beyond text, improving sentiment accuracy
Increases transparency and trust; helps identify which features truly drive price movements

DATA PREPROCESSING

```
# 移除非 ASCII 字符
def remove non ascii(text):
   return ''.join(c for c in text if 0 < ord(c) < 127)
# 清洗文本 + 保留英文
def clean_and_filter_english(text):
   if not isinstance(text, str):
       return None
   # 1. 去除非 ASCII 字符
   text = remove_non_ascii(text)
   # 2. 转小写
   text = text.lower()
   # 3. 珍除链接、0提及、#
   text = re.sub(r"http\S+", "", text)
   text = re.sub; #", "", text)
   # 4. 移除特殊符号、数字, 仅保留字母和空格
   text = re.sub(r''[^{sa-zA-Z}]'', '''', text)
   # 5. 去掉多余空格
   text = re.sub(r"\s+", " ", text).strip()
   # 6. 语言检测 (只保留英文)
   try:
       if detect(text) == 'en':
           return text
```

MODERL TRAINING

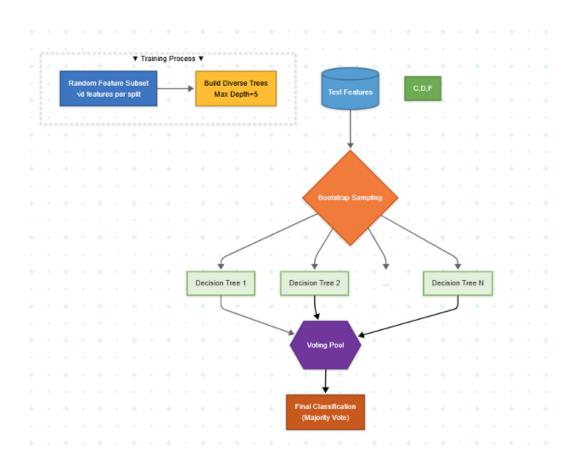
Logistic Regression





MODERL TRAINING

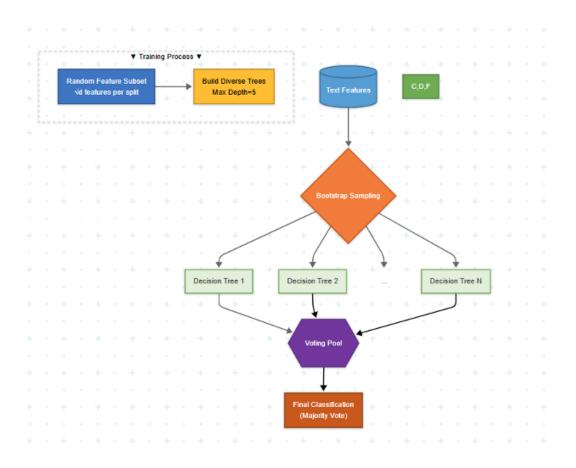
Random Forest





MODERL TRAINING

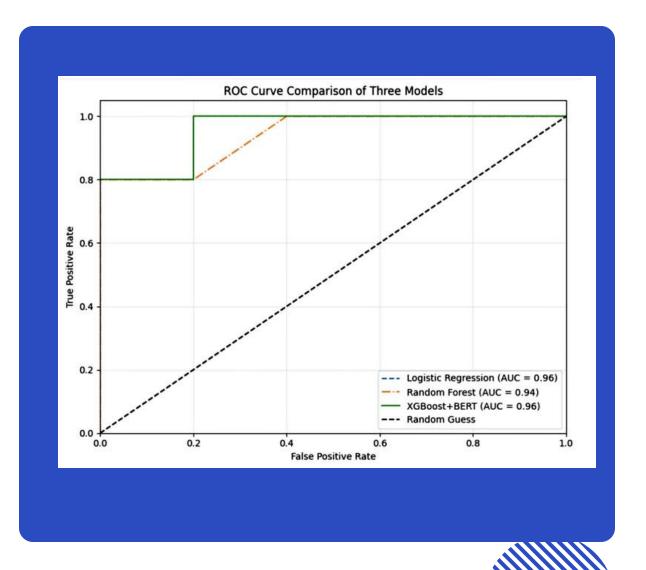
XGBoost+BERT







DATA EVALUATION





RESULTS AND FINDINGS

Top Performer: XGBoost + BERT

Achieved the highest overall accuracy.

 BERT embeddings captured nuanced emotions in text, greatly enhancing sentiment classification.

Baseline: Logistic Regression

- Provided consistent but limited performance.
- Lacked depth in capturing complex semantic patterns.

Alternative: Random Forest

- Strong at handling non-linear data.
- Stable but slightly underperformed in precision and recall compared to XGBoost + BERT.



FUTURE WORK

Automated trading systems

Combining sentiment analysis with technical indicators

Real-time sentiment processing

THANK YOU!