Pasin Chanaphan 6538121321
Pasut Aranchaiya 6538123621
Run tangkijngamwong 6538171721
Chanthep Sakulruangrak 6438049821

AI-Driven Stock Insights: Analyzing Financial News with LLM in BTC News

Introduction

Artificial Intelligence (AI) has revolutionized financial analysis by enabling the automation of data processing and interpretation. With advances in Natural Language Processing (NLP), Large Language Models (LLMs) can now analyze financial news, social media trends, and other textual sources to provide insights into stock and cryptocurrency movements. Bitcoin (BTC) is highly sensitive to market sentiment, regulatory developments, and macroeconomic factors. Analyzing financial news helps traders and investors anticipate price swings by understanding how sentiment influences BTC's market behavior. This project aims to leverage LLMs to analyze financial news related to BTC and determine sentiment trends that could impact BTC price movements. The expected outcomes include real-time sentiment classification of financial news, correlation analysis between sentiment and BTC price changes, and identification of significant events affecting BTC valuation.

Background & Literature Review

NLP techniques are widely used in financial markets to extract insights from news articles, earnings reports, and social media. LLMs like GPT, BERT, and Llama enable sophisticated sentiment analysis, allowing automated classification of news as positive, neutral, or negative. Numerous studies suggest that investor sentiment plays a crucial role in market movements. A sudden influx of negative news can lead to panic selling, while positive news often drives buying pressure. Prior research has explored using AI for stock predictions, demonstrating that sentiment analysis can enhance forecasting accuracy. Studies have also shown that BTC prices are influenced by Twitter sentiment, news headlines, and macroeconomic signals.

Methodology

The project gathers data from Twitter, financial news websites, and RSS feeds. Data preprocessing involves text cleaning, tokenization, and stopword removal to prepare data for model analysis. The implementation utilizes a BERT-based model for sequence classification. The model is pre-trained and used in a zero-shot setting for sentiment classification. Sentiment classification categorizes news as positive, neutral, or negative, and a correlation analysis is performed to understand sentiment-driven price trends.

Implementation & Code Explanation

Data Preparation

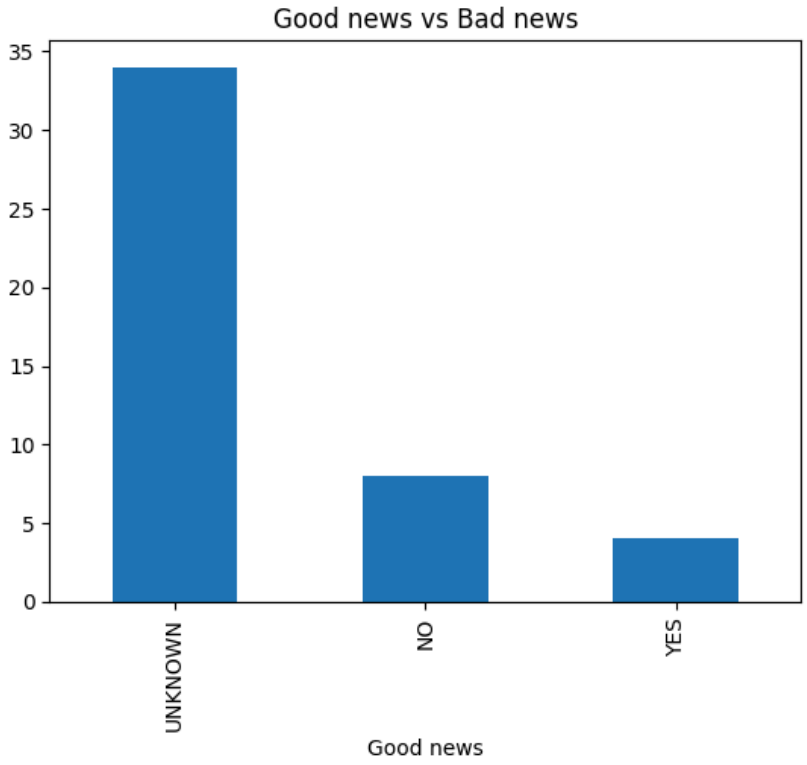
For the data preparation part, first, we filtered BTC_News_.csv and get data belonging to our desired start date and end date. Then, we drop unnecessary columns like 'Unnamed: 0', 'storyId', 'sourceCode'. Next, we create a date column from the versionCreated column where we only take the date and not the time. By dropping duplicates of date column, we are able to get the first news on each day. From this, the BTC data only gave us 46 rows of daily news which is very little because the BTC news data only started from 2025-02-01 to 2025-03-18. We then moved the versionCreated column as an index of the dataframe. Now, the data is in the same shape and form as the example given by P'Nuth, and is ready for Sentiment Analysis.

Sentiment Analysis

This sentiment analysis visualization categorizes financial news as positive ("YES"), negative ("NO"), or neutral/uncertain ("UNKNOWN"). The bar chart shows that most news articles fall under "UNKNOWN," indicating either neutral sentiment or classification uncertainty, while negative news ("NO") appears more frequently than positive news ("YES"). This distribution suggests that bearish sentiment is more prevalent in the dataset, which could influence BTC price movements. Improving the sentiment classification model could help better distinguish between positive, negative, and neutral news for more accurate market insights.

```
In [15]: filtered_news['Good news'].value_counts().plot.bar(title='Good news vs Bad news')

Out[15]: <Axes: title={'center': 'Good news vs Bad news'}, xlabel='Good news'>
```



Visualization

Visualization techniques are used to illustrate sentiment trends over time and their correlation with BTC price fluctuations. Tools such as Pandas and Plotly are utilized for data analysis and representation.

Future Work & Improvements

Future improvements include fine-tuning the model with a finance-specific dataset to improve sentiment classification precision. Expanding the analysis to non-English financial news could enhance global market insights. Integrating sentiment analysis into algorithmic trading strategies could provide real-time trading signals.

Conclusion

AI-driven sentiment analysis offers valuable insights into BTC price movements, with financial news sentiment showing a notable correlation with BTC price changes. Traders and investors can leverage AI-based sentiment analysis to make informed decisions. AI-driven financial analysis is rapidly evolving, with potential applications in real-time trading and risk assessment. Future advancements in LLMs will further enhance predictive capabilities in stock and crypto markets.