BACKGROUND

As the economic and social impacts of cryptocurrencies continue to grow rapidly, so does the prevalence of related news articles and social media posts. Similar to traditional financial markets, it appears likely that there is a relationship between media sentiment and the prices of cryptocurrency coins.

Our goal was to explore whether sentiment analysis on news headlines, tweets, and Reddit posts can inform predictions on future price changes of the cryptocurrency, Bitcoin.

METHOD

DATA

- **News Data**: 4,000,000 news headlines scraped from various sources.
- **Tweets**: 100,000 tweets scraped from Twitter.
- **Reddit Posts**: 100,000 Reddit posts.

LABELING

- In the final implementation, data points are labeled with the percent change in coin price 1hr, 2hr, 6hr, 12hr, and 24hr in the future, as defined by the model.

MODEL

- A separate model was constructed for each media source; all use the Keras [8] framework, with Tensorflow [9] backend, and the same general design. First, input text passes through an LSTM layer. Second, the encoded text passes through a bi-directional LSTM layer. Finally, a linear activation is applied to perform the regression task in generating price change predictions. Models are trained using mini-batches (stochastic gradient descent).

RESULTS

- **T-SNE Visualization of Custom Crypto Word Vectors**:
  - Query Top N Similar Words (missing):
    - **query**: dog
    - **similar words**: 'doge', 'coin', 'coinbase'
- **Closest Vector to ICO**:
  - **result**: tokensale
  - **cos distance**: 0.663

AGGREGATED PRICE CHANGE PREDICTIONS

- **News Headlines**:
  - Accuracy: 94.31%, 45.87%
  - Mean: 46.25%, 51.47%
  - Predicted price change: 1hr:
    - Overall: +5.31%, 51.47%
    - Specific investment: 2hr, 6hr, 12hr, 24hr:
      - +1hr: 49.25%, 48.28%
      - +2hr: 46.36%, 44.68%
      - +6hr: 53.13%, 44.28%
      - +12hr: 54.04%, 46.68%
- **Tweets**:
  - Accuracy: 95.31%, 44.35%
  - Mean: 49.31%, 46.68%
  - Predicted price change: 1hr:
    - Overall: +5.31%, 46.68%
    - Specific investment: 2hr, 6hr, 12hr, 24hr:
      - +1hr: 49.41%, 47.97%
      - +2hr: 46.36%, 46.68%
      - +6hr: 50.00%, 50.00%
      - +12hr: 54.04%, 46.36%
- **Reddit**:
  - Accuracy: 93.15%, 45.31%
  - Mean: 49.31%, 46.68%
  - Predicted price change: 1hr:
    - Overall: +2.40%, 44.35%
    - Specific investment: 2hr, 6hr, 12hr, 24hr:
      - +1hr: 45.43%, 46.68%
      - +2hr: 49.41%, 46.68%
      - +6hr: 60.95%, 46.68%
      - +12hr: 59.15%, 46.68%

REFERENCES