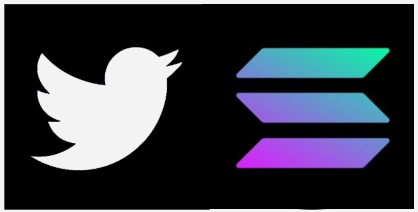
BOGDAN CONSTANTINESCU
KAI HAYDEN
ROB VOLGMAN

THE UNIVERSITY OF CHICAGO

Twitter Sentiment Analysis for Real-Time Solana Trading

Real Time Intelligent Systems Final Project



2022

THE UNIVERSITY OF CHICAGO

Contents

2021

BOGDAN CONSTANTINESCU
KAI HAYDEN
ROB VOLGMAN

THE UNIVERSITY OF CHICAGO

Project Overview

01

THE UNIVERSITY OF CHICAGO

Why This Project?

- Cryptocurrency trading volumes have skyrocketed to as high as $300B USD/day
- With no underlying asset, cryptocurrency valuations are often volatile and “hype-driven”
- Twitter activity can signal both positive and negative interest in a coin.
- Viral tweets can often have an near-immediate effect on the market
- The goal of this project is to create a profitable real time trading system that can generate signals in real time based on the sentiment of a tweet.

THE UNIVERSITY OF CHICAGO

Project Steps

Data Collection

- **Twitter** API for historical tweet data
- **CryptoCompare** API for historical Solana price data
- **Binance** API for live Solana price data

Data Engineering

- Label tweets with 24h trailing return
- Tokenize historical and real time tweets to standardize length

Model Training

- Train **LSTM** model on two weeks of historical data
- **Predict** 24-hour returns

Trading Strategy

- Utilize **sentiment analysis** to predict price movement
- Transform price predictions into **trading signals**

Real-Time Trading

- Stream Twitter data in **real-time**
- Generate trading signals and track PNL
- **Execute** long/buy and short/sell orders

BOGDAN CONSTANTINESCU
KAI HAYDEN
ROB VOLGMAN

THE UNIVERSITY OF CHICAGO

Data Collection

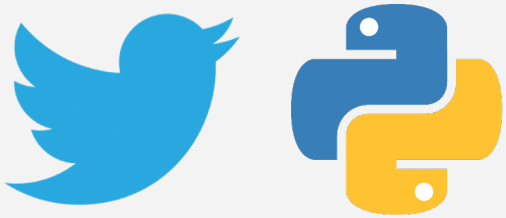
02

THE UNIVERSITY OF CHICAGO

Data Collection with APIs

Twitter Data

- Developer account with elevated access to Twitter API
- Tweepy library to access the Twitter API through Python
  - Search_recent_tweets function to collect historical tweets
  - Stream.filter function to stream tweets in real time

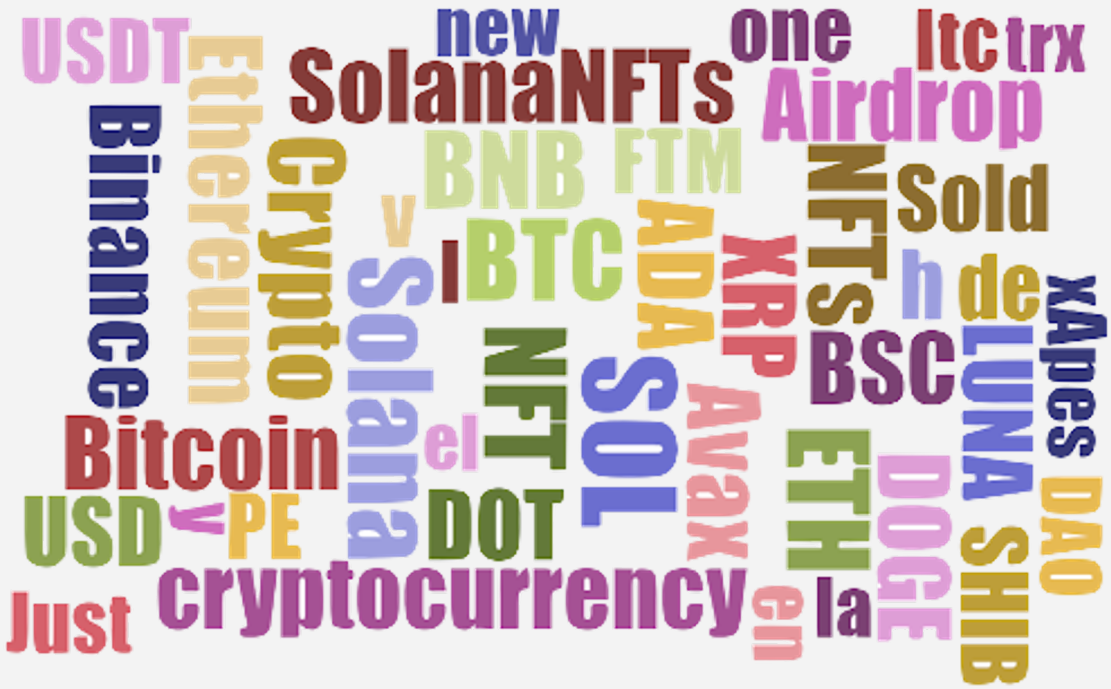


Price Data

- Cryptocompare web-socket API to collect historical Solana price data
  - From symbol - SOL
  - To symbol - USD
- Binance web-socket API to collect real time Solana price data
  - From symbol - SOL
  - To symbol - USDT

THE UNIVERSITY OF CHICAGO

Text Analysis



BOGDAN CONSTANTINESCU
KAI HAYDEN
ROB VOLGMAN

THE UNIVERSITY OF CHICAGO

Data Engineering

03

REAL TIME INTELLIGENT SYSTEMS
FINAL PROJECT

Data Engineering

- Associate tweet with 24-hour trailing returns minute-by-minute
- Remove punctuation and tokenize vectors on most common 1000 words to limit overfitting small samples
- Embedding layer to reduce input dimensionality
- Split train and test sets by time to avoid data leakage

Time	Followers	Text
2022-03-11 20:16:43	21880	Whats the highest price your favorite [...]
2022-03-11 20:16:44	17	This project is going to be great [...]

+

Time	Close	Move	Percent
2022-03-11 20:16:00	80.68	1.11	0.013758
2022-03-11 20:17:00	80.78	1.05	0.012998

↓

Time	Followers	Text	Returns
2022-03-11 20:16:43	21880	Whats the highest price your favorite [...]	0.012998
2022-03-11 20:16:44	17	This project is going to be great [...]	0.012998

BOGDAN CONSTANTINESCU
KAI HAYDEN
ROB VOLGMAN

THE UNIVERSITY OF CHICAGO

Model Training

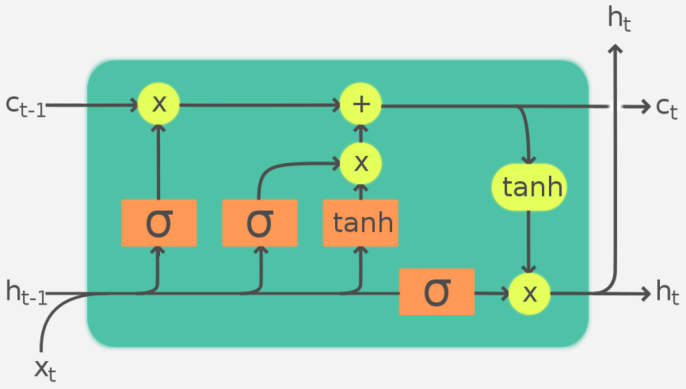
04

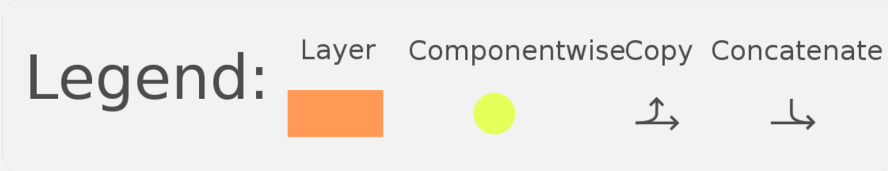
REAL TIME INTELLIGENT SYSTEMS
FINAL PROJECT

Model Selection

LSTM

- Allows us to associate words with price movement
- Long short-term memory (LSTM) models are a form of recurrent neural network with the ability to retain information over long sequences
- When applied to text recognition, allows for pattern recognition of phrases or word combinations separated by long stretches
- Implemented using the Keras library with MSE loss function, adam optimizer

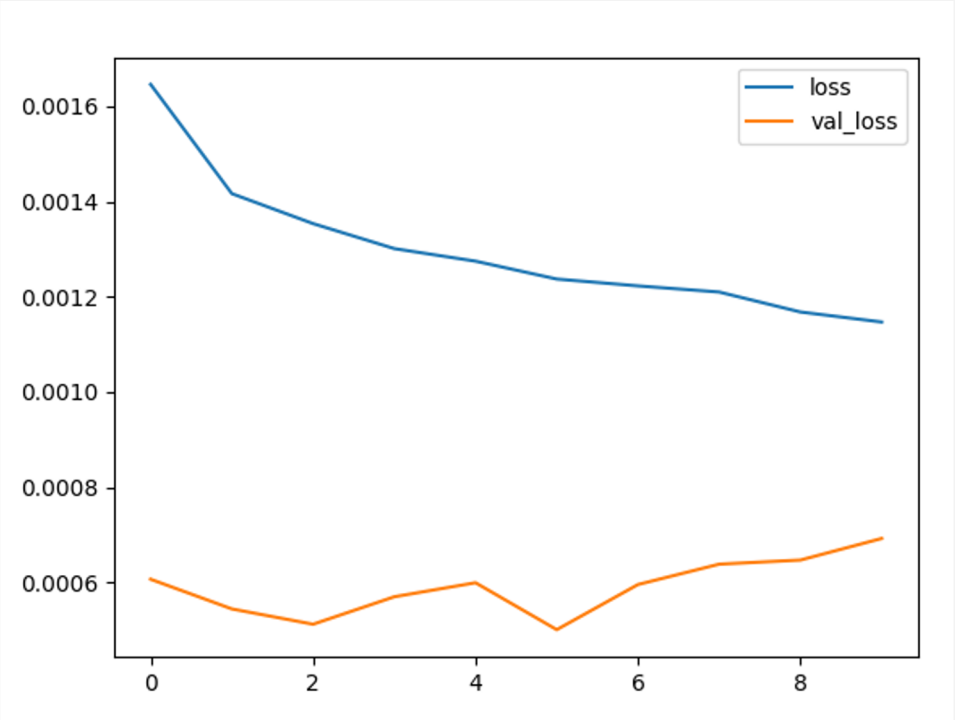




THE UNIVERSITY OF CHICAGO

Model Training

- Training set of 13,000 Tweets collected over 2 weeks
- Validation set time frame had significantly lower variance than training set
- Training beyond 5 epochs did not improve validation loss
    - 5 epochs used for final model



THE UNIVERSITY OF CHICAGO

Model Predictions

Positive

- Memecoin
- Pengoicy
- Airdrop
- Project
- Ada
- Xrp
- Nftart

Negative

- Bnb
- Binance
- Avax
- Cro
- Tez
- Shill
- Chill
- Poly

BOGDAN CONSTANTINESCU
KAI HAYDEN
ROB VOLGMAN

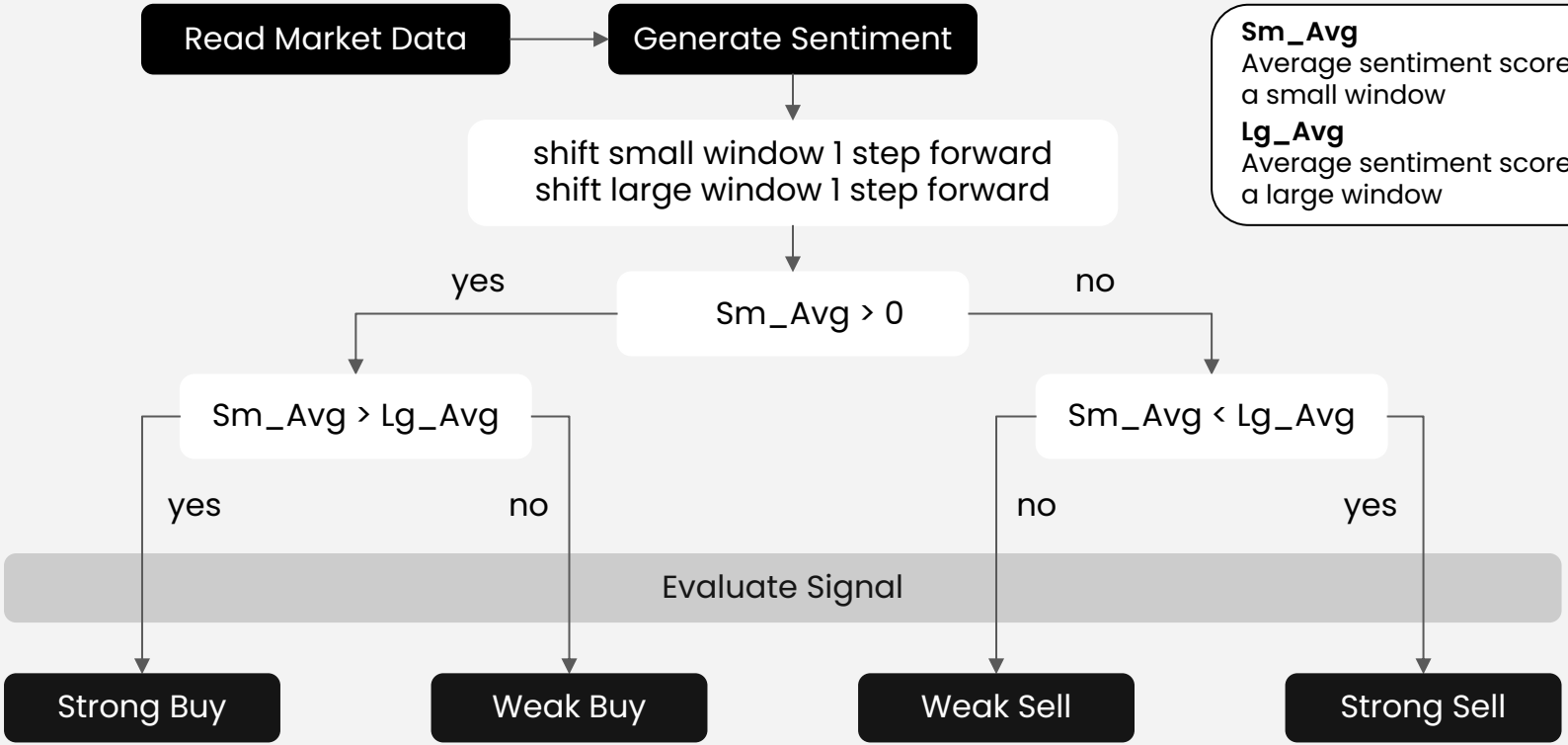
THE UNIVERSITY OF CHICAGO

Trading Strategy

05

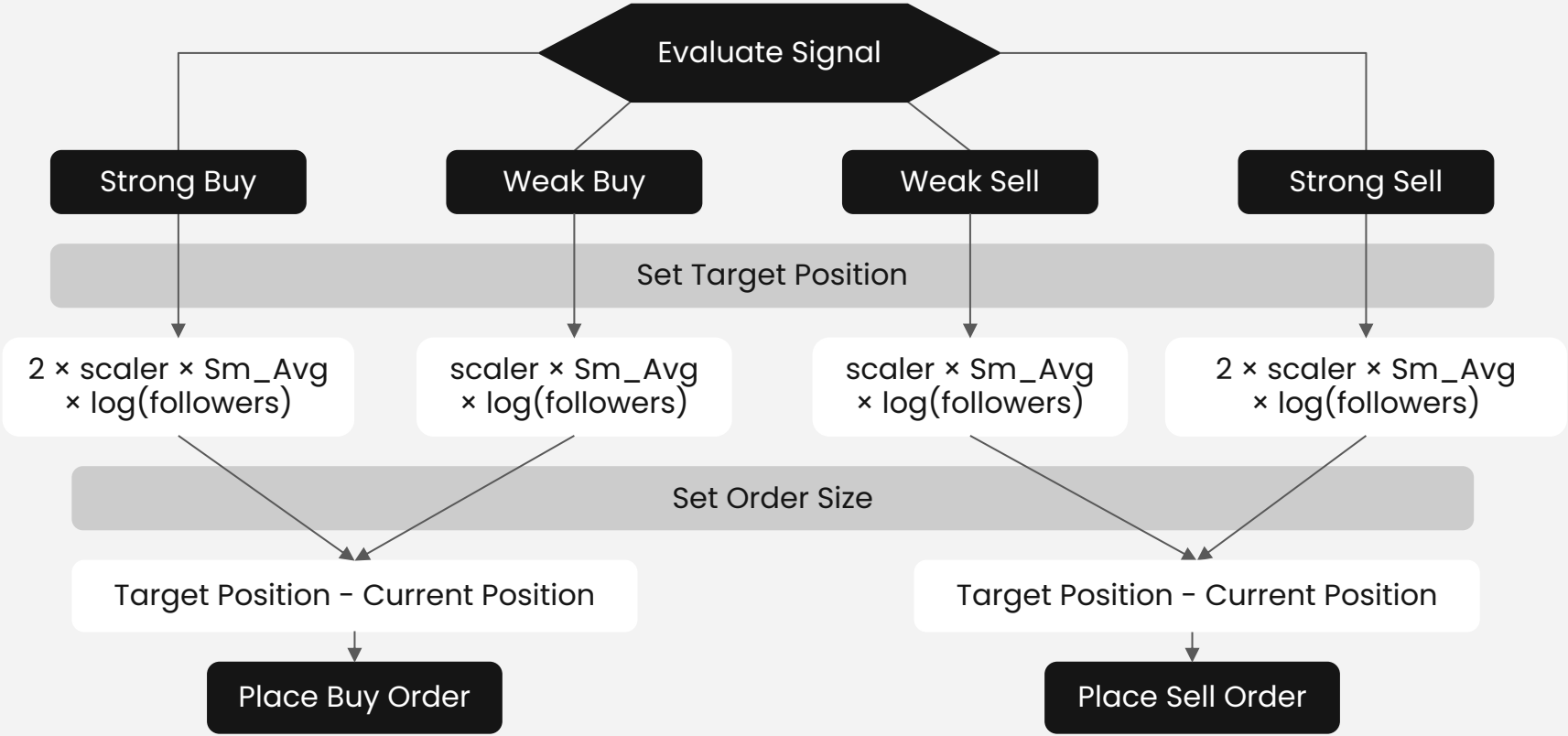
THE UNIVERSITY OF CHICAGO

Sentiment Interpretation



THE UNIVERSITY OF CHICAGO

Trading Algorithm



BOGDAN CONSTANTINESCU
KAI HAYDEN
ROB VOLGMAN

THE UNIVERSITY OF CHICAGO

Real-Time Trading

06

REAL TIME INTELLIGENT SYSTEMS
FINAL PROJECT

THE UNIVERSITY OF CHICAGO

Real-Time Trading

Trading Strategy Class

- Track positions, cash, holdings, profits and losses
- Define **trading strategy** and initialize parameters
- Instantiate queues to record tweet sentiment
- Create a **real time twitter stream** instance

↓

- **Trade** method to connect to twitter stream and begin trading
- **On_market_update** method to generate signal
- **Check_signal** method to execute orders

↓

- **Stop** method to disconnect from twitter stream and stop trading
- **Close** current position and liquidate all holdings

Twitter Stream Class

- Parent class inheriting **tweepy.Stream**
- Establish a **streaming connection with Twitter** API
- Start and stop streaming tweets based on its **trading strategy superclass** instance

↔

- Extract tweet fields in real time
- Perform **sentiment analysis** on extracted text
- Pass an **update** containing sentiment, price, and followers to the trading system

↔

- **Disconnect** from twitter stream
- Pass a “disconnect update” to instruct trading system to **close position** at current price

BOGDAN CONSTANTINESCU
KAI HAYDEN
ROB VOLGMAN

THE UNIVERSITY OF CHICAGO

Next Steps

07

Next Steps

- Gather tweets over a larger time frame and on multiple cryptocurrencies
- Gather data from other sources
  - Reddit
  - News
  - Discord
  - Telegram
- Explore whether phrase patterns and tweet activity translate across cryptocurrencies or are unique to each currency
- Utilize model predicted risk and return to refine sizing
- Alternative models (GRU, BERT)

BOGDAN CONSTANTINESCU
KAI HAYDEN
ROB VOLGMAN

THE UNIVERSITY OF CHICAGO

Real-Time Demo

REAL TIME INTELLIGENT SYSTEMS
FINAL PROJECT