Optimal Buy Time Regression

Ritesh Patel, Ashay Pathak, Chatana Mandava

Simon Fraser University

INTRODUCTION:

Cryptocurrency is a digital asset designed to work as a medium of exchange that uses strong cryptography to secure financial transfer. Cryptocurrencies is turning heads in the financial services space because of the market place dynamics is similar to equities. And further more the biggest reason is financial returns are mind boggling

MOTIVATION:

The optimal buy time regression predicts the best time to buy an assert. This serves as a base to solve many other problems in finance industry.

This is challenging because its hard to find all the factors that can predict the market. When it comes to make a decision on buying or selling, few decisions are driven by individual emotions which makes this problem difficult.

DATA SETS:

All datasets were obtained from crypto compare ,yahoo finance and coinmarketcap. We have scraped data for the top 100 crypto currencies.

Dataset consists of the following columns:

High: The highest price of the day.
Open: The opening price is the starting price of the day.
Close: The closing price is the last price of the day.
Volume: Volume indicates the total number of coins that have been traded during the day.
Low: The lowest price of the coin during the day.
Date: Timestamp depending upon the type of data

TOOLS:

Pandas data frames, Flask, NumPy, Keras, Tensorflow , Sklearn, matplotlib, plotly, Dash

METHODOLOGY:

LSTM : In time series datasets there is a sequence of dependence among the input variables. Recurrent Neural Networks are very powerful in handling the dependency among the input variables. LSTM is a type of Recurrent Neural Network (RNN) that can hold and learn from long sequence of observations.

CNN (Multichannel Multistep forecasting) : A CNN works well for identifying simple patterns within your data which will then be used to form more complex patterns within higher layers. A 1D CNN is very effective when we expect to derive interesting features from shorter (fixed-length) segments of the overall data set.

Monte Carlo Simulation : This is a technique used to understand the impact of risk and uncertainty forecasting models.
We have used this simulation technique to have the idea about the market for past 30 days

DATA PIPELINE:

Data Preprocessing

- The original data is OHLV. For CNN, we normalized the data in windows and for every window we divided the data by the last closing price for that window. For LSTM we used minmax scaler for normalizing

Analytics

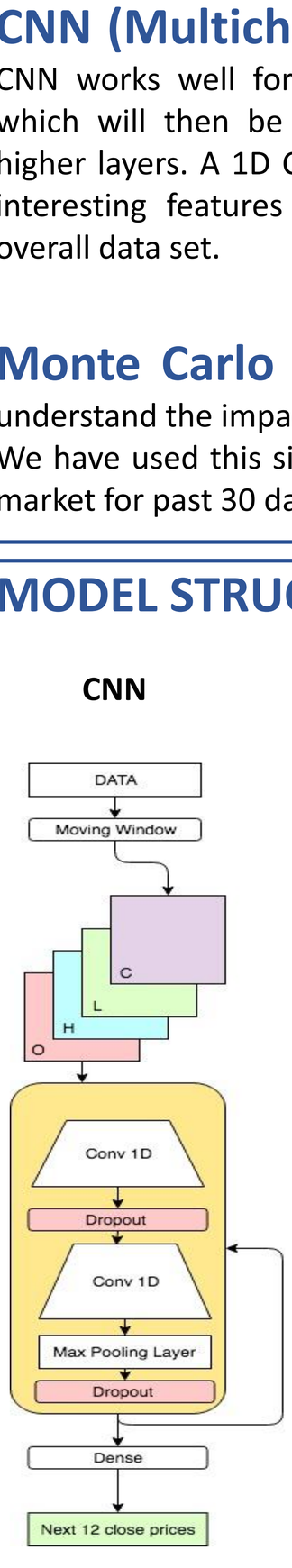
- Performed exploratory data analysis on the given data set. Found out the trend of different cryptocurrencies and also the correlation between them. Later we tried to find the factors that are influencing the variations of the price .

Model evaluation

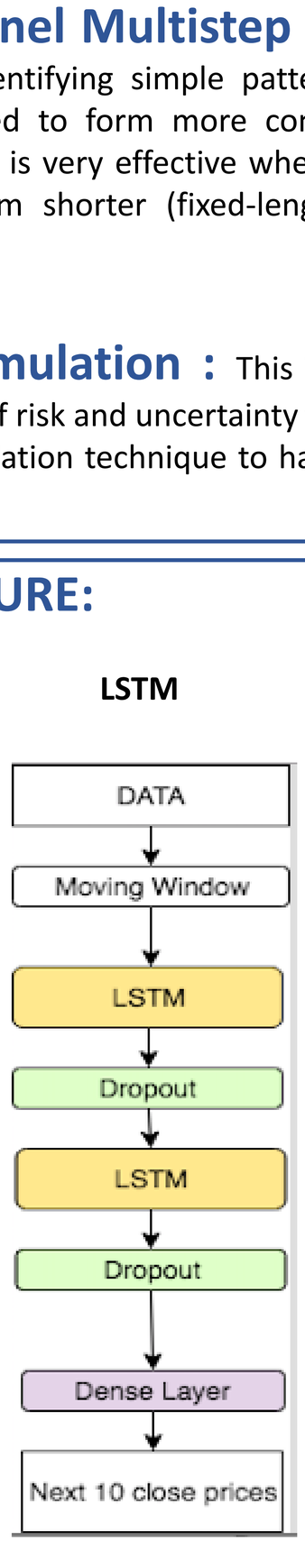
- For both the model **CNN** and **LSTM,** we chose MAE as a metric to calculate the absolute error.

MODEL STRUCTURE:

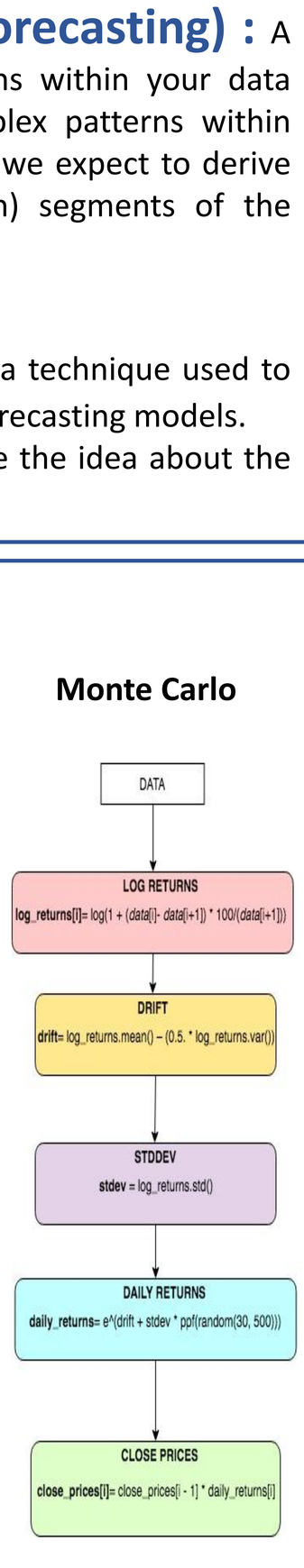
CNN



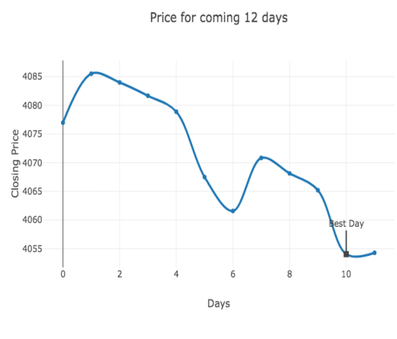
LSTM

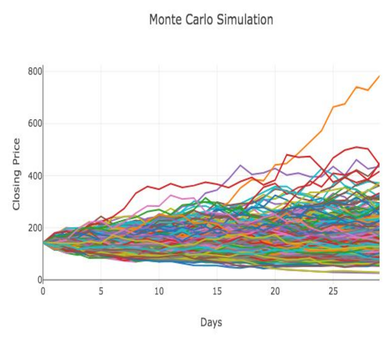


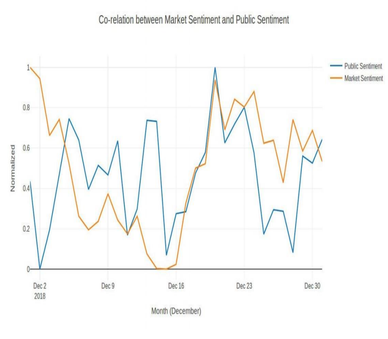
Monte Carlo



RESULTS:







FINDINGS :

- Window based Normalization improved CNN model accuracy.
- Adding News sentiment score, EMA, SMA as feature to the dataset degraded the model's performance.
- LSTM performs better with more data.
- Monte Carlo cannot be used to predict close prices. However, can be used to understand market trends over a period of time.

Model	Mean Absolute Error(days)	Mean Absolute Error(prices)
LSTM-CP	3.8	232.69741235351566
LSTM-CP-EMA	3.5	380.3250916503906
LSTM-CP-HOURLY	3.35	52.79510131835938
CNN	2.3	10.22332344244