# Abstract

Final Report

A mobile application provide trading signal based on the cryptocurrency market tweets sentiments

**TEAM BEAR**

TEAM MEMBER:

HUNG WAI CHAK, Bear

TSOI SIU FUNG, Rex

LAM CHUN HIN, Clinton

CHENG YIU HANG, Toddy

SUPERVISOR:

DR. LEUNG MAN FAI HENRY

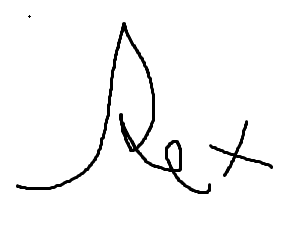
A trading signal is an act that recommends the investor to buy or sell a cryptocurrency generated by the analysis and gain insight into the current market before trading in this project. In this project, since there is a correlation between

# Declaration Page

We, Tsoi Siu Fung(12586593), Lam Chun Hin(12549757), Cheung Yiu Hang(12598033), Hung Wai Chak(12555360), certify that the work is original and we have utilized guidance of our supervisor in completing this project, and that the content which is not our own has been attributed and referenced properly. There should be no copyrighted content without permission to use. There should be no confidential data. We declare that the description and information outlined in the individual team member reports are true reflection of the project status to the best of our knowledge.

signatural

一張含有 箭 的圖片

自動產生的描述 ,, 一張含有 文字 的圖片

自動產生的描述and Date: 12/05/2022

# Acknowledgments

This work was supported in part by the Hong Kong Metropolitan University, School of Science and Technology. We wish to express our deepest gratitude to the supervisor Dr. Leung Man Fai Henry. His knowledge in machine learning, deep learning and data mining has been a valuable source of inspiration in the development of theories and application. He also is the ultimate source of support so that we have the opportunity to go this far in this study. His endless encouragement, patience and dedication has given us the confidence to overcome so many difficulties and challenges.

We also like to thank William and our friends who kindly helped to test and evaluate our mobile application.

**Content Page**

[Abstract 1](#_Toc103791506)

[Declaration Page 3](#_Toc103791507)

[Acknowledgments 4](#_Toc103791508)

[List of Tables 7](#_Toc103791509)

[List of Figures 7](#_Toc103791510)

[Chapter 1. Introduction 7](#_Toc103791511)

[A. Background 7](#_Toc103791512)

[B. Project Aim 8](#_Toc103791513)

[C. Project Objectives 8](#_Toc103791514)

[D. Value Propositions 8](#_Toc103791515)

[Chapter 2. Literature Review 9](#_Toc103791516)

[A. Problem Discussion and Analysis 9](#_Toc103791517)

[B. Review of Existing or Related Solutions for the Problem 10](#_Toc103791518)

[C. Related Supporting Technologies 12](#_Toc103791519)

[Chapter 3. Methodology and Solution 14](#_Toc103791520)

[A. Overview 14](#_Toc103791521)

[B. Requirements, Supporting Technologies, and Technical Gap 16](#_Toc103791522)

[C. Architecture or High-Level System Design 18](#_Toc103791523)

[Evaluation method and design 22](#_Toc103791524)

[Chapter 4. Prototype Result and Discussion 23](#_Toc103791525)

[A. Implementation results of the model 23](#_Toc103791526)

[Comparison 27](#_Toc103791527)

[Selecting regression model 28](#_Toc103791528)

[B. Implementation result of the general application 29](#_Toc103791529)

[Issues about implementation 29](#_Toc103791530)

[Description of the Prototype (organization, function list, screen-dump) 30](#_Toc103791531)

[Chapter 5. Evaluation and Discussion 34](#_Toc103791532)

[A. Approach Evaluation (Trading Signal) 34](#_Toc103791533)

[B. Application Performance 36](#_Toc103791534)

[Chapter 6. Conclusion 38](#_Toc103791535)

[References 39](#_Toc103791536)

[Appendix A Team Members’ Roles and Responsibility 41](#_Toc103791537)

[Appendix B1: Hung Wai Chak’s Final Report 43](#_Toc103791538)

[Appendix B2: Tsoi Siu Fung’s Final Report 50](#_Toc103791539)

[Appendix B3: Lam Chun Hin’s Final Report 57](#_Toc103791540)

[Appendix B4: Cheng Yiu Hang’s Final Report 63](#_Toc103791541)

[Appendix C Progress Report 70](#_Toc103791542)

[Appendix D Team Meeting Minutes 72](#_Toc103791543)

# List of Tables

[Table 1 A table for compare different feature between the chosen platforms 11](#_Toc103788777)

[Table 2 A table of the difference between User estimated price and the actual price 11](#_Toc103788778)

[Table 3 Available Datasets for machine models 23](#_Toc103788779)

[Table 4 TextBlob classification\_report 25](#_Toc103788780)

[Table 5 TextBlob + Clustering classification\_report 26](#_Toc103788781)

[Table 6 CNN-LSTM classification\_report 27](#_Toc103788782)

[Table 7 CNN-LSTM bitcoin tweets classification\_report 29](#_Toc103788783)

[Table 8 Trading signal accuracy 37](#_Toc103788784)

# List of Figures

[Figure 1 A schematic diagram of methodology overview . 16](#_Toc103791560)

[Figure 2 Use case diagram of application 19](#_Toc103791561)

[Figure 3 A component diagram of application 19](#_Toc103791562)

[Figure 4 A data-flow diagram of application 20](#_Toc103791563)

[Figure 5 A machine learning model flowchart diagram. 21](#_Toc103791564)

[Figure 6 Mar 2022 Bitcoin percentage change of price. 22](#_Toc103791565)

[Figure 7 confusion matrix TextBlob sentiment 26](#_Toc103791566)

[Figure 8 Jenks break clustering for sentiment 26](#_Toc103791567)

[Figure 9 Jenks break result for TextBlob 27](#_Toc103791568)

[Figure 10 confusion matrix TextBlob sentiment 28](#_Toc103791569)

[Figure 11 CNN-LSTM bitcoin tweets model train history 28](#_Toc103791570)

[Figure 12 confusion matrix CNN-LSTM bitcoin tweets model sentiment 29](#_Toc103791571)

[Figure 13 mobile application screen-dump of homepage 32](#_Toc103791572)

[Figure 14 mobile application screen-dump of watchlist 32](#_Toc103791573)

[Figure 15 mobile application screen-dump of detailpage 33](#_Toc103791574)

[Figure 16 mobile application screen-dump of news function 33](#_Toc103791575)

[Figure 17 mobile application screen-dump of prediction function 34](#_Toc103791576)

[Figure 18 Plot between Actual, Test and Predicted Trend in Bitcoin 35](#_Toc103791577)

[Figure 19 Plot between Test and Predicted Trend in Bitcoin 35](#_Toc103791578)

[Figure 20 Plot between Actual, Test and Predicted Trend in Dogecoin 36](#_Toc103791579)

[Figure 21 Plot between Test and Predicted Trend in Dogecoin 36](#_Toc103791580)

[Figure 22 Plot between Actual, Test and Predicted Trend in Ethereum 36](#_Toc103791581)

[Figure 23 Plot between Test and Predicted Trend in Ethereum 37](#_Toc103791582)

[Figure 24 Survey results on trading signal performance 38](#_Toc103791583)

[Figure 25 Survey results on trading signal satisfaction 38](#_Toc103791584)

[Figure 26 Users satisfaction on the application 39](#_Toc103791585)

# Chapter 1. Introduction

## Background

The Cryptocurrency market, a market for trading decentralized digital currency, has become a rising trend recently; even some novice investors want a piece. The cryptocurrency market is usually available to trade 24 hours a day, seven days a week, because there is no centralized market governance, central bank, or single administrator.[[1]](#footnote-1) The previous research from Ryan Farell has summarized the three leading indicators that show the cryptocurrency industry. Market capitalization, the estimated number of cryptocurrency users, and daily transaction volume.[1] The global crypto market cap of over 10,000 cryptocurrencies has reached a new all-time high of over $2.8 trillion on 2-11-2021[[2]](#footnote-2). These numbers will serve as indicators of the extent to which the public has already accepted cryptocurrency.

The cryptocurrency price changes very fiercely; investors may return a high-interest rate in the short term. Taking the 2017 Bitcoin price as an example: it increased 2000%, going from $863 on January 9, 2017[8], to a high of $17,550 on December 11, 2017. Meanwhile, it has a high-risk investment. The Mt. Gox failure in February 2014 showed that even if a significant exchange might suddenly exit the market[2], investors are exposed to huge risks. As a beginner or amateur investor, there are some human limitations, such as they cannot monitor the crypto market in 24 hours and make wrong trading decisions by their bias. As a result, some platforms for trading cryptocurrencies have emerged; those tools might help the investor analyze the past data of the crypto market.

However, having old data is not enough to help the investor predict the future trends or prices fluctuation of cryptocurrency. The research done by Connor Lamon, Eric Nielsen, Eric Redondo has shown that the prices are usually affected by the news and social media posts[3]. For example, in 2021, Elon Musk, the founder of Tesla, posted an announcement on Twitter, a famous social media platform, that their company will stop accepting Bitcoin as a payment method starting from May. After this post, the price of bitcoin decreased. Then, after one month, Elon Musk posted another announcement saying Tesla will accept bitcoin again. Immediately, the Bitcoin and the Ethereum price increased.

So there is a problem: can we accurately predict a cryptocurrency's future trends or prices fluctuation by analyzing the tweets sentiment and implementing this technology in a mobile application? We will solve this problem within this project.

## Project Aim

This project aims to develop a cryptocurrency analysis mobile application for beginners and amateur investors. The application provides a price tracker function and related information about the cryptocurrency. Moreover, to help investors predict a cryptocurrency's future trends or price fluctuation, it should give a trading signal base on analyzing the sentiment of related crypto tweets by implementing some machine learning technique. Also we need to design an approach to give the trading signal of the cryptocurrency to the users based on the analyzed data.

## Project Objectives

1. Collect and Organize social media post data from the selected platform.
2. Analyze Machine Learning models and determine the appropriate model for this application.
3. Design an approach to implement the analyzed results as a trading signal for the app
4. Develop a mobile application that gives users information about the crypto market
5. Evaluate the accuracy of the prediction and the usability of the mobile application.

## Value Propositions

For immediate benefit, we will provide an application that can perform a better approach to determine the trading decision for investors to solve the problem of human limitations. Furthermore, the application implements a mech machine learning technique that analyzes investors’ comments on tweets to find out some possible timing to buy or sell the cryptocurrency, letting users understand when they should make their trade.

For the ripple effect, the approach of analyzing related tweets sentiment as a trading signal can contribute the knowledge of this approach to another developer, which proposes a new method to give trading recommendations to investors.

For the long-term effect, the application facilitates an assistant investment tool for beginner investors in the cryptocurrency market in the next five years. The 24-hours update frequency of sentiment analysis can become a prime function of the new-come application of cryptocurrency market investment tools.

# Chapter 2. Literature Review

## Problem Discussion and Analysis

The problem of "the cryptocurrency's future trends prediction application" should be solved by developing a mobile application and implementing a sentiment analysis model into it.

First of all, to provide an application to give a signal of the potential market change for cryptocurrency investors, the application should contain a sentiment analysis model, which is deployed on the server. Besides, the application should also provide cryptocurrencies' news and the coins details, such as current price, the percentage of price change in twenty-four hours, and historical prices of different periods.

However, the fluctuation of cryptocurrency could be extremely fierce. This makes the cryptocurrency forecast very difficult [[4]](https://arxiv.org/pdf/1812.02987v2.pdf). A research done by Siddhi Velankar, Sakshi Valecha, Shreya Maji have mentioned the price of Bitcoin did not depend on the business events or intervening government unlike stock market [[5]](https://www.icact.org/upload/2018/0618/20180618_finalpaper.pdf) which mean some traditional factors that usually affect the stock price change are not suitable to use in predicting the cryptocurrency market trend. Besides, some research done in the past shows that if the prediction only used the historical price data of cryptocurrencies, the accuracy is low [[6]](http://www.diva-portal.org/smash/get/diva2:1110776/FULLTEXT01.pdf)[[7]](https://arxiv.org/abs/1612.01277v5). Although these data are the performance of the cryptocurrency in the past, history trends cannot accurately reflect future trends, especially in the cryptocurrency market. Therefore, to predict the cryptocurrency market trends correctly, there are some new factors needed to be used. As mentioned in chapter 1, some research and cases have proved a correlation between media sentiment and the trends of cryptocurrency [[3]](http://cs229.stanford.edu/proj2017/final-reports/5237280.pdf). According to the research done by Yu Wang and Runyu Chen, adding social media features can greatly increase the accuracy of price prediction [[8]](https://scholarspace.manoa.hawaii.edu/bitstream/10125/63875/1/0109.pdf). However, considering the time limit and the cost, analyzing all the social media posts in different platforms is unrealistic. In order to achieve the project goal within the time, this project will only focus on analyzing tweets as there are a lot of papers that have proved the efficiency of analyzing twitter posts to do cryptocurrency prediction[5][9][[10]](http://cs229.stanford.edu/proj2015/029_report.pdf)[[11]](https://ieeexplore.ieee.org/abstract/document/8530659). Twitter is the most popular social media source and it provides the earliest and fastest news updates in a concise format[[10]](http://cs229.stanford.edu/proj2015/029_report.pdf). As a result, a sentiment analysis model and related dataset are needed to provide an accurate signal of the cryptocurrency's future trends prediction.

Since the sentiment analysis model can analyze the words in the dataset and classify them into different emotions, such as positive, negative, and natural. Nonetheless, to ensure the sentiment analysis model can correctly classify the emotions expressed in the words, the model must be trained. In order to train the analysis sentiment model, a dataset consisting of a time series of the tweets sentiment of the cryptocurrency market is needed. This dataset will be created using a web scraper or related API.

As the project aims to develop a mobile application, the sentiment analysis model needs to be implemented as a trading signal function. We will deploy the model on the webserver and call it in the mobile application.

## Review of Existing or Related Solutions for the Problem

There are plenty of existing online platforms to help users trade or predict the cryptocurrency market, among them, each platform has their pros and cons and we have listed the difference between them in a table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Coinbase | Cryptocurrency forecast: AI Prediction | Cryptocurrency hero | CoinMarketCap |
| AI Prediction | ❌ | ✔ | ✔ | ❌ |
| News Feed | ✔ | ✔ | ❌ | ✔ |
| Live Trading | ✔ | ❌ | ✔ | ✔ |
| Live Cryptocurrency Prices | ✔ | ✔ | ✔ | ✔ |
| Consumption | ❌ | ✔ | ✔ ( subscribe to unlock more) | ❌ |
| Mobile App | ✔ | ✔ | ✔ | ✔ |

Table 1 A table for compare different feature between the chosen platforms

[Coinbase](https://www.coinbase.com/) is a liberate, release, and popular cryptocurrency exchange, users can purchase, sell and exchange cryptocurrency there. It also provides the latest news of cryptocurrency and supports mobile platforms. Although Coinbase has not provided an ai prediction feature, it has all the basic features for the cryptocurrency and the interface is concise for the cryptocurrency beginner, therefore when we are developing our cryptocurrency application we can refer to Coinbase.

[Cryptocurrency Forecast: AI Prediction](https://www.appannie.com/en/apps/ios/app/crypto-forecast-ai-prediction/) is a mobile application, it has used the Neural Network to train with and to provide state-of-the-art predictions of cryptocurrency prices. It provided hourly and daily price predictions of the cryptocurrencies and live market stats, news, sentiment analysis to let cryptocurrencies users refer to. Unfortunately, this mobile application requests users to pay money to join the Professional plan to get access to the AI Prediction, therefore the users of this application have limited access to it.

[Cryptocurrency hero](https://cryptohero.ai/) is a website that allows consumers to trade cryptocurrency in an automated, AI-powered environment 24 hours a day, seven days a week. It also provides a backtest environment for users to leverage previous cryptocurrency hero market data to discover the best investing strategy for them, as well as insight about the strategy's performance, complexity, and profitability. However, cryptocurrency hero provided only one active bot and one linked exchange inside the liberate, release trial plan, so consumers must subscribe to a premium plan for assistance and to gain additional active bots and connected exchanges.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CoinMarketCap estimate  cryptocurrency(11/2021) | Number of user estimated cryptocurrency price | User estimated median price | Actual cryptocurrency price | Differences(In percentages) |
| Bitcoin | 116,025 | $57,000.00 | $65.560.52 | -13.06% |
| SHIBA INU | 82,086 | $0.00002758 | $0.00005259 | -47.56% |

Table 2 A table of the difference between User estimated price and the actual price

[CoinMarketCap](https://coinmarketcap.com/) compiles all projected data from user input, calculates the average estimated price, and compares it to the actual price. However, this method of predicting cryptocurrencies based on user voting is rather imprecise; even when there are over 10000 individuals to forecast, the prediction is still erroneous.

The above existing solution platform that consumers must subscribe to in order to have artificial intelligence prediction; otherwise, the consumer can only access limited features such as a news feed, live cryptocurrency, and live trading but no artificial intelligence prediction.

## Related Supporting Technologies

There is some research[[3]](http://cs229.stanford.edu/proj2017/final-reports/5237280.pdf)[[9]](https://scholar.smu.edu/cgi/viewcontent.cgi?article=1039&context=datasciencereview) that showed social media posts are related to the cryptocurrency price, and the research used the following technology (or technical) aspects to discover the relationship between social media data, and the cryptocurrency price change.

**Deep learning model**

Deep Neural networks(DNN)[[3]](#footnote-3) are inspired by the learning process occurring in human brains. They with more than two layers have reached a given level of complexity. To analyze input in complicated ways, deep neural networks need sophisticated mathematical models.

Multilayer perceptron(MLP)[[12]](https://www.sciencedirect.com/science/article/pii/S0957417411002740), a subset of DNN, is a class of feedforward artificial neural networks that create a number of outputs from a collection of input. [[4]](#footnote-4)This technique is frequently utilized in speech recognition, image or video recognition, and machine translation.

**Data collection**

Web scraping[[5]](#footnote-5) is the automatic extraction or mining of data from a structured web page. This approach is used to collect data for a variety of purposes, including price monitoring, news monitoring, and market research, among others.

Natural language processing (NLP)[[6]](#footnote-6) is a technology that allows artificial intelligence to interpret text and spoken words in the same manner that humans do. For instance, it allows artificial intelligence to do sentiment analysis[[13]](https://ieeexplore.ieee.org/iel7/6287639/8600701/08684825.pdf) to classify whether the data is positive, negative, or neutral. Natural Language Toolkit[[7]](#footnote-7)(NLTK) is one of the toolkits used a lot when training a machine learning model with the Python package.

Data preprocessing [[14]](https://link.springer.com/book/10.1007/978-3-319-10247-4) is a data mining approach for converting unstructured data into a useful and efficient format. Data cleansing, noisy data, data transformation, and data reduction are all components of data processing. In general, this approach is employed prior to machine learning to remove unnecessary data via data cleaning.

**Classification model**

Logistic Regression[[15]](https://www.tandfonline.com/doi/abs/10.1080/00220670209598786) is a machine learning classification technique that may be used to describe the probability of the likely outcomes of a particular experiment. It is straightforward and easy to comprehend since it can indicate the effect of multiple independent factors on a single result variable, allowing us to swiftly do error analysis to determine how to continue. However, this approach will only work if the anticipated variable is binary; otherwise, it will fail.

Naive Bayes algorithm[[8]](#footnote-8) is a classification algorithm that uses Bayes' theorem to classify an object. Its classifiers assume strong, or naive, independence between attributes of data points. Its algorithm was used in our daily lives, such as document classification and spam filtering. Unfortunately, the result of Naive Bayes is known as a bad estimator, because in real life, not all features are independent, and the calculation will be incorrect.

A Support Vector Machine (SVM)[[9]](#footnote-9) is a supervised machine learning algorithm that can be used for classification as well as regression. SVMs are more commonly used in classification problems such as text classification tasks such as spam detection and image recognition challenges, where they perform well in recognizing color-based and aspect-based features. However, it is very heavy, which means that it is not recommended to use this method in testing models because the training time can be very long with large datasets.

To summarize, when it comes to deep learning models, neural networks with MLP should be used to produce superior results. Each of the above classification models has advantages and disadvantages; for example, the Support Vector Machine performs well in classification but requires a considerable training period. As a result, while selecting a categorization model, we should examine which is most suited to our needs. In terms of data gathering, we may apply the same strategy as before because it is common and we have a superior source.

[**Trading signal**](https://www.investopedia.com/terms/t/trade-signal.asp)

A trade signal triggers a trading action - either buying or selling a security or other asset generated by analysis. The analysis is created with a suitable mathematical algorithm based on market action, or it can be human-generated using the technical indicator

**Data visualization of the cryptocurrency prices for the application**

The React Native Animated Charts[[10]](#footnote-10) provide different types of charts, log plots that contain more cryptocurrency prices data in a chart, depending on the date or different time slot.

**Connecting model and the mobile application**

A REST web service[[11]](#footnote-11) is a platform-independent application-to-application service that employs web technologies such as HTTP to facilitate interoperable interactions between heterogeneous systems.

1. **Conclusion**

By doing the literature review, there are some existing cryptocurrency analysis tools. Those tools might help the investor analyze the past data of the crypto market. However, users need to pay for the ai prediction function, unlike the fundamental parts like providing the live and historical prices and related news. Furthermore, having old data is not enough to help the investor predict future crypto trends. For example, a correlation between Tweets sentiment and the trends of cryptocurrency was found by research. Moreover, adding social media features can significantly increase the price prediction accuracy for the cryptocurrency.

# Chapter 3. Methodology and Solution

## Overview

We will build a mobile application that provides cryptocurrency information and prediction functions that gives a trading signal using React Native. After the app is developed, we will evaluate the accuracy of the prediction and the usability of the mobile application. Finally, we will invite 100 users to utilize our app and collect their feedback with a questionnaire.

We will provide a local server and database for the deep learning base model to run and implement two models:

a classification model that can analyze the sentiment on related selected tweets

a Regression Analysis Model that could accurately predict a cryptocurrency's future trends or prices fluctuation based on the result of the analyzed sentiment result.

Finally, we will develop an approach to change the given results to a trading signal for users.

The schematic diagram below shows the basic structure of the methodology

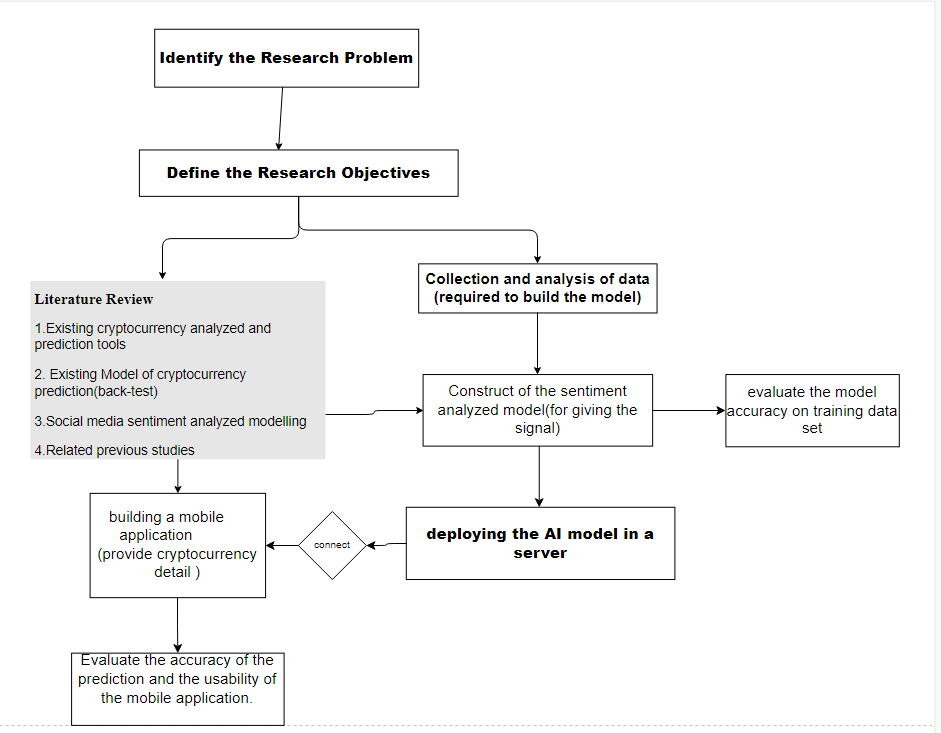


Figure 1 A schematic diagram of methodology overview

## Requirements, Supporting Technologies, and Technical Gap

To develop a cryptocurrency's future trends and price prediction mobile application,

We use React Native, a mobile development tool, to develop a mobile application. Since that React Native can perform cross-platform to let the user choose either Android or ios. Moreover, we can use the same code for deployment on iOS and Android to maximize code reuse instead of building iOS and Android native apps.

This mobile application aims to have these functional and non-functional requirements.

Our mobile application aims to help people gain insights; data is much more valuable when it is visualized. Therefore, Charts and graphs are needed and it will help people identify the chart pattern when people communicate data findings.

**The list of the technical requirements are as follow:**

1. A list will show the last price of different cryptocurrencies.
2. Provide related news of selected cryptocurrencies for the users
3. A prediction function that gives the user a buy or sells signal.
4. Detail of the cryptocurrency that the user selected.

**The list of the non-technical requirements are:**

1. Fast and stable response between our applications between the predictions and news server. We have no control over the API provider.
2. 75% of first-time users found the app is user-friendly.
3. The accuracy of the prediction can reach 80%.

**Some supporting technologies help us implement the above functions:**

As for the first requirement, many up-to-date price data are needed to provide a list that shows the latest price of different cryptocurrencies. By considering the update rate and convenience, we intend to use some existing APIs to give the newest price data immediately after sending a request, such as [coingecko](https://www.coingecko.com/en/api) ,CoinDesk, Cryptocompare API, etc.

For returning the detail for user-selected crypto, we need a massive number of historical data of different cryptocurrencies and plot a price chart with those data to give the user a better understanding of the price trend of the chosen crypto. As a result, [the react-native-animated-charts](https://github.com/rainbow-me/react-native-animated-charts) library is chosen for plotting aesthetic, animated linear charts based on a given input. Instead of hosting a database to store plenty of data, using APIs is more appropriate. C[oingecko](https://www.coingecko.com/en/api) [[12]](#footnote-12)API is chosen as it can also provide a timestamp of the price data.

For providing related news of selected cryptocurrencies for the users, [Cryptopanic](https://cryptopanic.com/developers/api/) API[[13]](#footnote-13) is used for attaching the pertaining news.

As for the mobile application prediction function, [Tweepy](https://www.tweepy.org/)[[14]](#footnote-14) Python library is used to access the Twitter API social platform to collect tweets related to cryptocurrency. [TextBlob](https://textblob.readthedocs.io/en/dev/index.html)[[15]](#footnote-15) Python libraries (use NLTK processing libraries) provide sentiment analysis, classification function. NLTK a Natural Language Toolkit that provides over 50 corpora and lexical resources, and it is used for data pre-processing to filter out useless data, such as stop words, wordnets, etc. CNN-LSTM sentiment analysis is a deep learning model for predicting the sentiment of the tweets. Decision tree and random forest regression machine model for predicting the future cryptocurrency price based on the last day sentiment result.

[Google Colab](https://colab.research.google.com/)[[16]](#footnote-16) and [Tensorflow](https://www.tensorflow.org/)[[17]](#footnote-17) are used for the machine learning process and observing the result.

[Snscrape](https://github.com/JustAnotherArchivist/snscrape)[[18]](#footnote-18)Python library is used to scrape the Twitter historical cryptocurrency related tweets.

[Pymongo](https://pymongo.readthedocs.io/en/stable/)[[19]](#footnote-19) Python library is used to work with MongoDB, and store all the cryptocurrency related tweets to databases.

[Rapminer](https://rapidminer.com/)[[20]](#footnote-20) a data analysis tool is used to train the regression model and test the model in a simple way.

The technical gap is that the returned result is a predicted price of the related cryptocurrency by combining the sentiment analyzing model and regression algorithm. Therefore, an approach or formula must be developed to change this result as a trading signal.

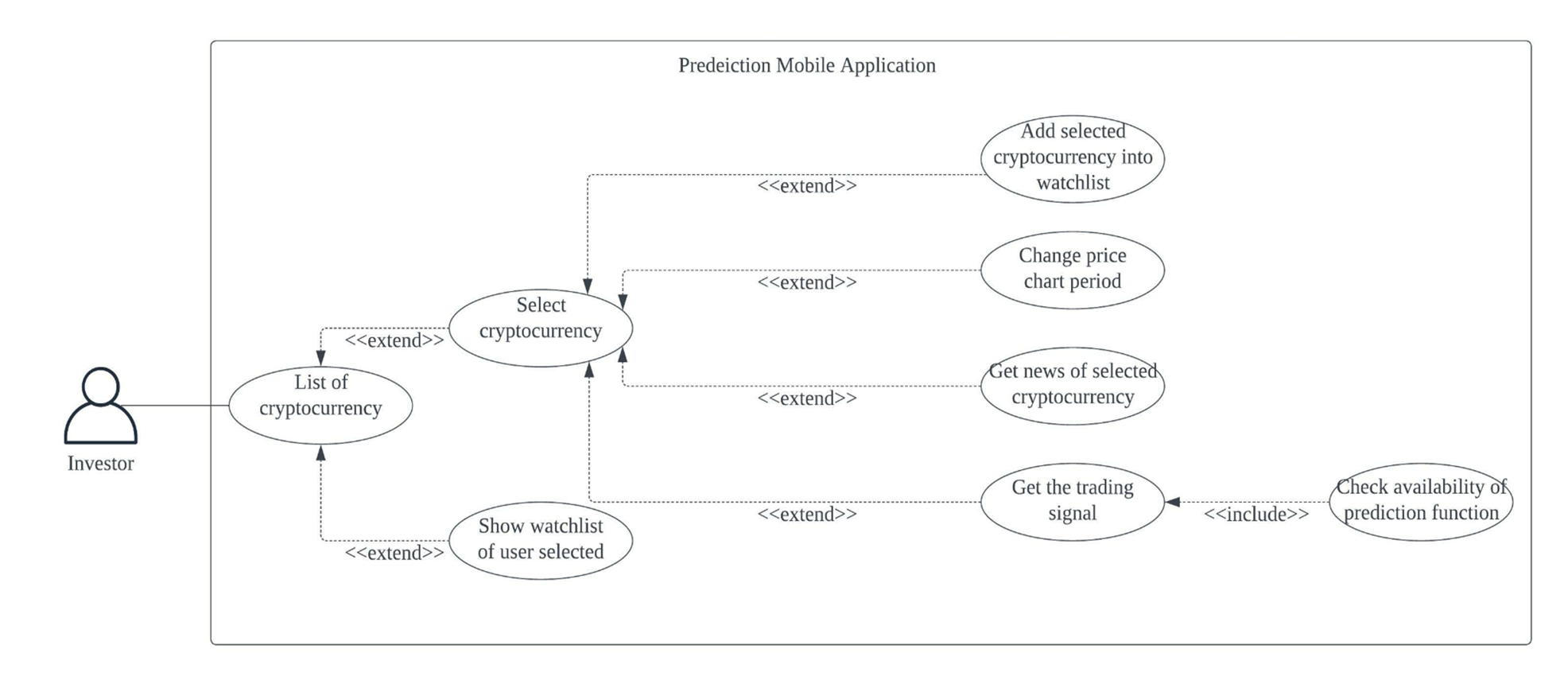


Figure 2 Use case diagram of application

## Architecture or High-Level System Design

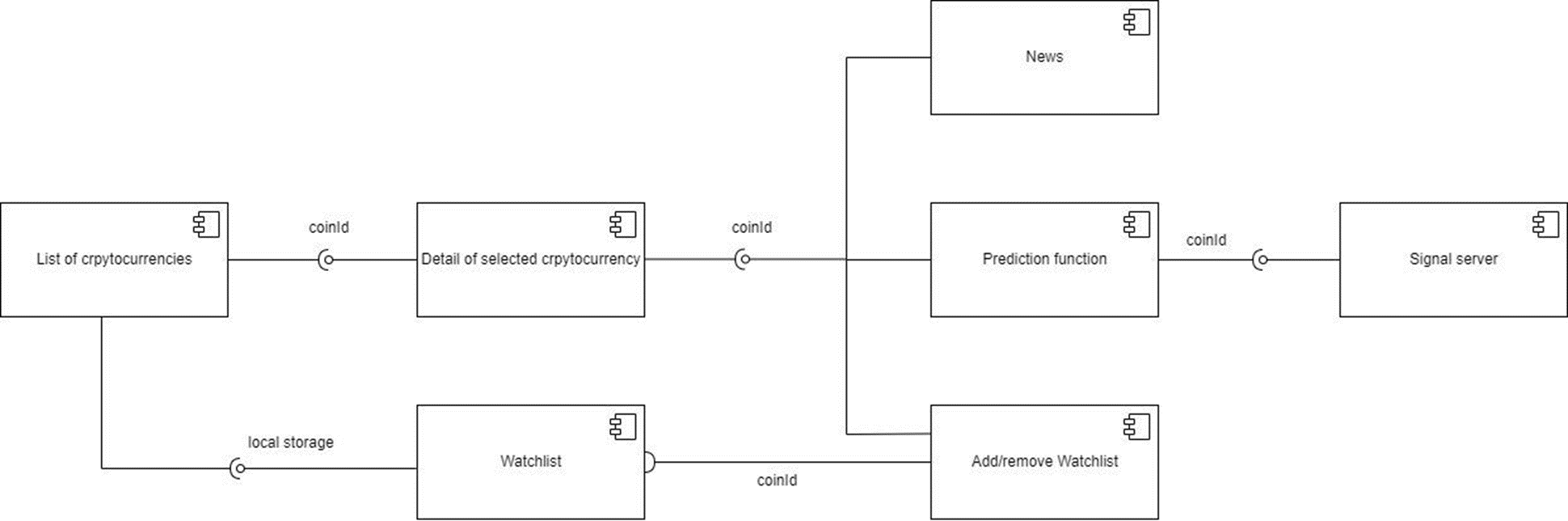


Figure 3 A component diagram of application

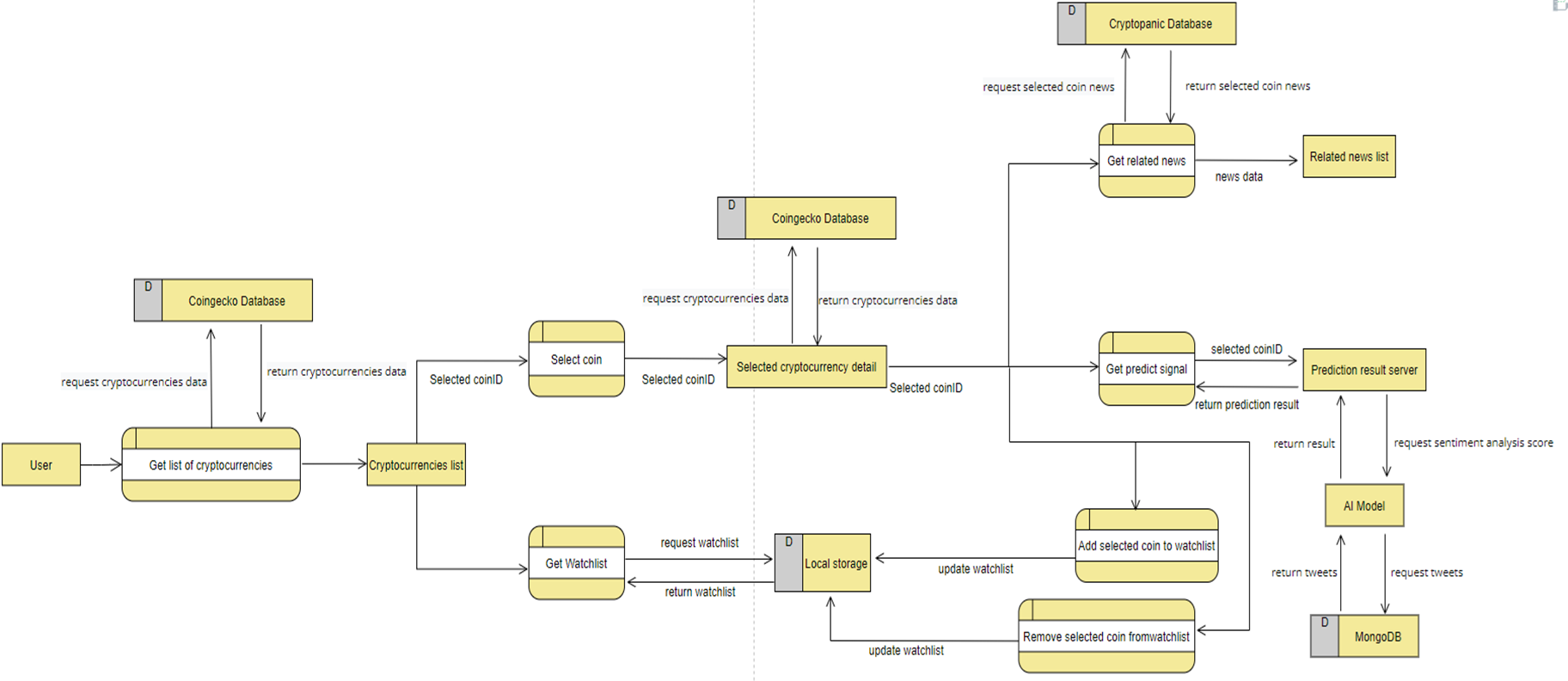


Figure 4 A data-flow diagram of application

Firstly, the user will access the homepage. Meanwhile, the application will request the cryptocurrency data by calling Coingecko API. Then, the user can select the cryptocurrency for more details or select the watchlist to see the cryptocurrency stored by the user. Moreover, the watchlist data will store in the local database with the async-storage library. Furthermore, after the user selects a cryptocurrency for detail, the user will access the detail page. The detail will request the detail of selected cryptocurrency data by calling Coingecko API. After the data has been returned, the user will see the detail of the cryptocurrency (i.e., real-time price, price chart in a different period, percentage of price change). Besides, the user can perform another action (i.e., view the related news, add or remove cryptocurrency from the watchlist or get the trading signal). For the watchlist function, the user can add or remove the selected cryptocurrency on the detail page. When the user performs the above action, the application will add or remove the cryptocurrency from the Async-Storage. In addition, the user can access the news page, and the application will request the related news data from Cryptopanic API. The news will be listed on the page after the application obtains the data from the third-party service. Lastly, for the prediction function, the prediction button will only be available in BTC, ETH, and DOGE. Otherwise, the prediction function is not available in the user-selected cryptocurrency, and the button will be disabled. Then, when the user executes the prediction function, the application will send an API request to the trading signal server. Then, the server will call the AI model, and the AI model will send an API request to the third-party service to obtain the related data and analysis. After that, the AI model will return the result to the client via the server.

**Machine Learning Model Flowchart**

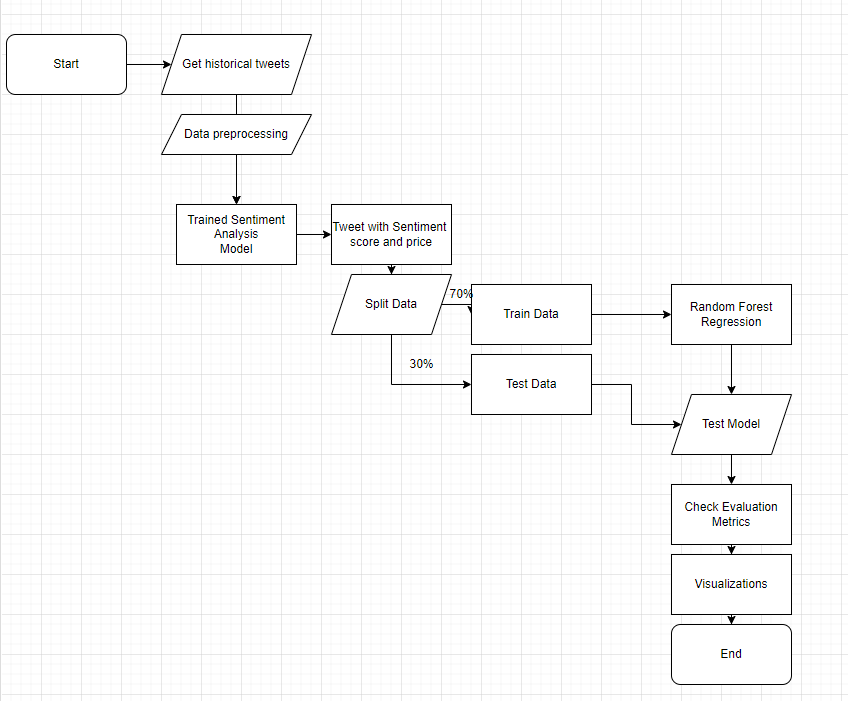


Figure 5 A machine learning model flowchart diagram.

Firstly, we need to prepare historical cryptocurrency-related tweets and preprocess them to make them suitable for sentiment model training, such as excluding tweets that obviously have nothing to do with the target. Then, we train the model with a preprocessed dataset and evaluate it. The best-performing model is selected as the cryptocurrency tweets sentiment model. We use sentiment models to predict the Jan 2022 Twitter cryptocurrency-related tweets sentiment score. Next, we split the Jan 2022 sentiment dataset with actual cryptocurrency price from [investing.com](https://www.investing.com/) into 70% training set and 30% testing set for train model and evaluate model performance. After the trained random forest regression model is finished, evaluation metrics are used to measure the model performance and plot a graph compared to the actual cryptocurrency price.

**How To Create The Signal**

To create the trading signal for cryptocurrency trading, we make use of the above model output and make use of our forecasting approach to forecast a trading signal.

For the trading signal, we tried to calculate the percentage change (pd)between every single predicted price(A) with the last predicted close price of that cryptocurrency (B) by using a percentage difference formula.

By looking at the (pd) results, we found the percentage difference in price between last day, and if the result is positive then the signal will return “Buy” and vice versa. Also, if the percentage change is equal or larger than 5 % then the signal will return “Strong Buy” because we observed that the maximum percentage change is around 5%.

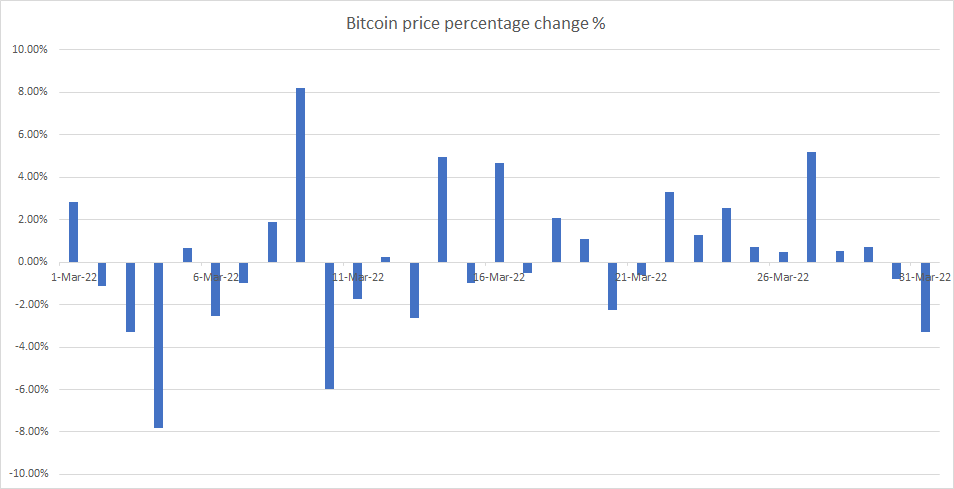


Figure 6 Mar 2022 Bitcoin percentage change of price.

All the above signal results after the calculation will be sent to the prediction API server, by this way the mobile application can get our predicted signal by posting a request.

## Evaluation method and design

We plan to use an accuracy formula to evaluate our trading signal. The prediction will use the Jan 2022 testing set for each cryptocurrency.

For the performance evaluation, we need to ensure that the fundamental component of the mobile application is well-functioning.

Furthermore, to evaluate the performance of the sentiment analysis model, we will split it into two ways. First, we will measure the accuracy of the sentiment analysis model by calculating the accuracy of the test result with the malignant and benign datasets as following:

[[21]](#footnote-21)

Moreover, we will calculate the confusion matrix value of the model.

After the comparison between three sentiment analysis model performance, the best performance of sentiment analysis model will be selected.

As for selection of the regression model, we will consider the different model pro and cons

As for evaluating the performance regression model, we will measure the mean absolute error (MAE) and the mean squared error (MSE) to evaluate the model.

𝑦𝑖 is predicted value, xi is actual value, n = the number of observations

is the ith observed value , is predicted value, n = the number of observations

and compare it with the market trend in the current date. At last, we will take thirty days as a period to evaluate the accuracy of the AI model when compared to the real market trend by selecting cryptocurrency.

For the user evaluation, we will focus on the interface design of our mobile application and the usability of the trading signal. We will investigate the aspects of user perception and feedback by following Shneiderman's eight golden rules.[[22]](#footnote-22) Therefore, we can evaluate whether our mobile application fulfilled those requirements of a better interface design and trading signal function.

For the analysis tool, Python is used for model training and visualization model performance. Rapidminer is used for training regression model random forest.

# Chapter 4. Prototype Result and Discussion

## Implementation results of the model

**Datasets for analysis**

|  |  |  |
| --- | --- | --- |
| Name | Source | Sentiment |
| Twitter Tweets Sentiment Dataset | https://www.kaggle.com/datasets/yasserh/twitter-tweets-sentiment-dataset | 27481 tweets, neutral 40%,positive  31%, negative 28% |
| Bitcoin Tweets | https://www.kaggle.com/datasets/skularat/bitcoin-tweets | 23965 tweets, positive 45%, neutral 43%, negative 12% |
| Bitcoin/Ethereum/Dogecoin historical prices (Jan 2022) | https://www.investing.com/crypto/cryptocurrencyname/historical-data | Not suitable |
| Historical Tweets (Jan 2022) | scraper | Not suitable |

Table 3 Available Datasets for machine models

**Pre-processing of the datasets**

A tweet contains lots of different content by different users, some of the tweets may be noisy. The raw data having polarity is susceptible to redundancy. Tweets preprocessing is needed, and the preprocessing of tweets included in following points.(Identifying (cryptocurrency-related) Twitter bot accounts[16])

* Remove all URLs (e.g. www.abc.com), targets (@username),hash tags (e.g.#topic)
* Replace all the emoticons and smileys
* Remove Stop Words and Punctuation
* Replace extra white spaces
* Remove duplicates tweet
* Remove Tweet contain “give away”, “pump”, “join”
* Remove Tweet with more than 14 hashtags
* Remove Tweet if the creator name contain “bot”

Experiments on sentiment analysis

Sentiment analysis model is for understanding the tweets’s piece of writing or emotion is positive, negative or neutral. In order to understand the emotion of a tweet and find the target sentiment analyzer, three different sentiment analysis models are evaluated with Twitter Tweets Sentiment Dataset from kaggle to train and test.

**TextBlob**

Textblob is the first sentiment analysis model we have tried. We tried to understand what sentiment analysis is and see the polarity relationship between tweets. Textblob returns two properties of the tweets, polarity and subjectivity. The [polarity score](https://textblob.readthedocs.io/en/dev/quickstart.html)[[23]](#footnote-23) is a float within the range [-1.0, 1.0]. The subjectivity is a float within the range [0.0, 1.0] where 0.0 is very objective and 1.0 is very subjective.

Also, Textblob is a pre-trained Lexicon-based sentiment analyzer, therefore we do not need to do anything to be able to use this model.

In the first try, we assume polarity score values below [0] are considered as negative sentiment data, values equal to [0] are neutral sentiment, values above [0] are positive sentiment.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-Score | Support |
| 0 | 0.66 | 0.47 | 0.55 | 1556 |
| 1 | 0.59 | 0.50 | 0.54 | 2224 |
| 2 | 0.54 | 0.80 | 0.65 | 1717 |
| accuracy |  |  | 0.58 | 5497 |
| macro avg | 0.6 | 0.59 | 0.58 | 5497 |
| weighted avg | 0.59 | 0.58 | 0.57 | 5497 |

Table 4 TextBlob classification\_report

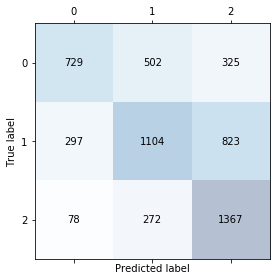


Figure 7 confusion matrix TextBlob sentiment

**TextBlob + Clustering**

Textblob is the first sentiment analyzer and its accuracy is 58%. Since we were blindly clustering the polarity score and classifying it as positive, negative, and negative, it may lead to inaccurate results. In order to achieve more consistent clustering, [Jenks natural breaks](https://www.ehdp.com/methods/jenks-natural-breaks-explain.htm)[[24]](#footnote-24) optimization is implemented. Jenks natural groupings of items that are near together and maximizing the spaces between the other groupings. In this case, It clusters the one-dimensional dataset with polarity values into different actual classes of data.

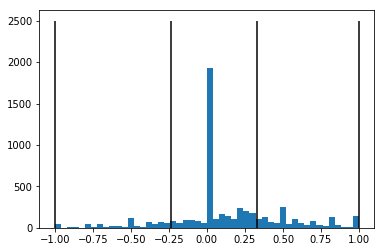


Figure 8 Jenks break clustering for sentiment



Figure 9 Jenks break result for TextBlob

Jenks break clustered polarity scores between [-1.0, -0.236] are negative sentiment data, values between [-0.236, 0.328] are neutral, values between [0.328, 1.0] are positive sentiment. And after the Jenks break clustering, the new model accuracy is 59%.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-Score | Support |
| 0 | 0.79 | 0.32 | 0.45 | 1556 |
| 1 | 0.51 | 0.82 | 0.63 | 2224 |
| 2 | 0.69 | 0.53 | 0.60 | 1717 |
| accuracy |  |  | 0.59 | 5497 |
| macro avg | 0.67 | 0.56 | 0.56 | 5497 |
| weighted avg | 0.65 | 0.59 | 0.57 | 5497 |

Table 5 TextBlob + Clustering classification\_report

**CNN-LSTM (Long Short-Term Memory Network)**

CNN-LSTM is a class of deep learning models that uses two models, CNN to extract the features of the tokenized tweets and uses LSTM to predict the sentiment of the tweet. CNN-LSTM is also handling Natural Language Processing well, such as text classification.

Unlike Textblob without any training to be able to use, CNN-LSTM requires time for training and pre-processing. Therefore it requires more workload to be able to use a CNN-LSTM model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-Score | Support |
| 0 | 0.69 | 0.71 | 0.70 | 1556 |
| 1 | 0.65 | 0.72 | 0.68 | 2224 |
| 2 | 0.82 | 0.69 | 0.75 | 1717 |
| accuracy |  |  | 0.70 | 5497 |
| macro avg | 0.72 | 0.70 | 0.71 | 5497 |
| weighted avg | 0.71 | 0.70 | 0.71 | 5497 |

Table 6 CNN-LSTM classification\_report



Figure 10 confusion matrix TextBlob sentiment

### Comparison

By taking three different model performance results, we can find that CNN-LSTM model has the highest accuracy and all the advantages. But the downside is that the CNN-LSTM model requires a training process and it needs to learn lots more words related to cryptocurrency in order to understand tweets about cryptocurrency.

**CNN-LSTM (Long Short-Term Memory Network) for cryptocurrency tweet**

As above mentioned, we have trained the [CNN-LSTM model](https://www.kaggle.com/code/dundee2002/bitcoin-tweets-sentiment-analysis-glove-cnn-lstm)[[25]](#footnote-25) with a dataset Bitcoin Tweets from Kaggle. In this way, this text analyzer is closer to our project aim and better understands cryptocurrency tweet sentiment.

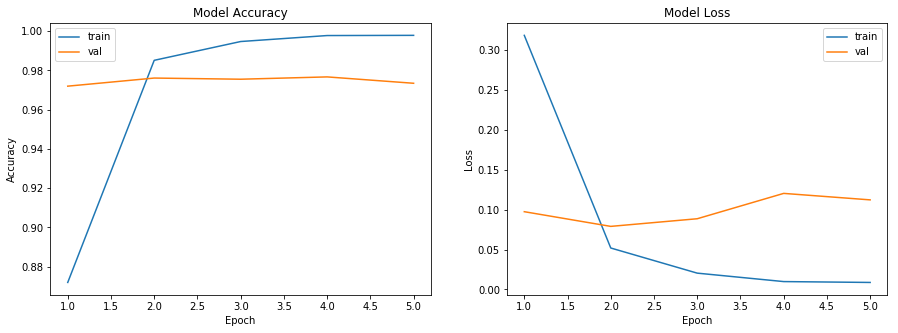


Figure 11 CNN-LSTM bitcoin tweets model train history

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-Score | Support |
| 0 | 0.95 | 0.95 | 0.95 | 1197 |
| 1 | 0.98 | 0.97 | 0.98 | 4388 |
| 2 | 0.97 | 0.98 | 0.98 | 4578 |
| accuracy |  |  | 0.97 | 10172 |
| macro avg | 0.97 | 0.97 | 0.97 | 10172 |
| weighted avg | 0.97 | 0.97 | 0.97 | 10172 |

Table 7 CNN-LSTM bitcoin tweets classification\_report

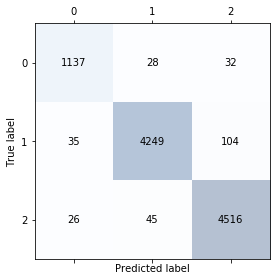


Figure 12 confusion matrix CNN-LSTM bitcoin tweets model sentiment

### Selecting regression model

As we know sentiment has an effect when forecasting cryptocurrency prices, therefore we need a regression machine-learning algorithm to predict future cryptocurrency prices with the sentiment. Decision Tree is a supervised classification machine learning algorithm. It was chosen for forecasting future prices in the beginning, as it is a simpler AI model. Decision tree is much more explainable and people can understand how the decision is made easier. However, Decision tree is prone to overfitting and the results are not accurate.

On other hand, Random Forest Regression is robust to outliers and more accurate compared to Decision trees. Random Forest algorithm consists of a large number of decision trees. These randomly selected decision tree results inside Random Forest will be combined and make the final decision.

By using the Rapidminer tool to train machine learning model, three Random Forest Regressor training with a combination of cryptocurrency Jan 2022 prices and Jan 2022 cryptocurrency sentiment, each random forest regressor model correspond to their cryptocurrency, such as bitcoin random forest regressor learn from tweets related to bitcoin and the bitcoin historical price.

We use mean absolute error (MAE) and mean squared error (MSE) to evaluate the model. Result of the BTC - random forest model MAE is 2788 and MSE is 10540621. As for the ETH - random forest model MAE is 319 and MSE is 134573. For the Doge - random forest model MAE is 0.0134 while the MSE is 0.000219.

**MACHINE LEARNING SETUP:**

Issues about implementation

1. Languages: Python

2. Platforms: Window

3. Libraries used: @gensim, @pymongo, @pandas, @datetime, @tweepy

4. Other specific implementation issues that other developers should know if they want to repeat your implementation

**Setup:** In order to make Twitter's tweets streaming API work, a local MongoDB server is needed for storing three different cryptocurrency tweets collection and sentiment scores. RapidMiner Studio is also needed to install and put the trained model into a custom process. After that, you need to change the MongoDB URL to your local MongoDB URL. Next, install all dependencies and Python 3.6 equal or above, the lower version of Python may not work for this project. Once all the above set-up is done, you need to start the MongoDB with a command prompt, enter command “mongod --dbpath \yourdatabase\location\” and start streaming tweets using StreamTweetMongoDB files. However, a developer API key is needed in order to start getting data from tweets otherwise Twitter will not provide this service. API keys need to be written to config.py and the streaming tweet services should work. Once all the above steps are done, you can start the query\_mongodb\_twttier.py, it will start returning new predictions every day.

## Implementation result of the general application

### Issues about implementation

1. Languages: Javascript; Framework: ReactNative

2. Platforms: ios and Android

3. Libraries used: @expo-web-browser, @react-navigation/native, @rainbow-me/animated-charts, @react-native-async-storage/async-storage, @react-navigation/native-stack

4. Other specific implementation issues that other developers should know if they want to repeat your implementation

When developers design the system of the mobile application, they should consider more about either using a database to store the data they need to provide in the application or using an API service to get data each time they need it. In this application, the API service has been chosen to implement the cryptocurrency data since it contains a vast amount of data and avoids some data loss of the historical data. Using API services can reduce the cost of using database services and avoid the risk of SQL injection by the user. However, suppose the other developers want to repeat our implementation. In that case, they should build a database for storing all cryptocurrency data instead of using API services if they can afford the cost of database services and if the application needs to provide high availability. Therefore, they can avoid the limitation of this application that if the API service is in downtime, this application can not provide any information to users.

### Description of the Prototype (organization, function list, screen-dump)

Setup: move the directory into the project in the command prompt, enter command “npm install -f” to install the dependencies of the project needed. Then, enter the command “npm start” to start the mobile application.

Key functions for each user type as a list or hierarchy: Prediction, related news, view detail of cryptocurrency, watchlist

**User interfaces:**

The users will receive the list of cryptocurrencies first and it will show the overview of each cryptocurrency. Then, user can choose to select the watchlist page and the mobile app will show all the cryptocurrencies in watchlist which added by the user. Meanwhile, user can have an option that to select a cryptocurrency and the app will show the detail of the selected cryptocurrency like the price-char and latest price. Moreover, after the access the detail page. User can perform four actions. Such as add selected cryptocurrency into watchlist, remove selected cryptocurrency from watchlist, execute the prediction function and show related news of selected cryptocurrency.

**Operation of the prototype (with screen dump) with respect to solving the problem:**

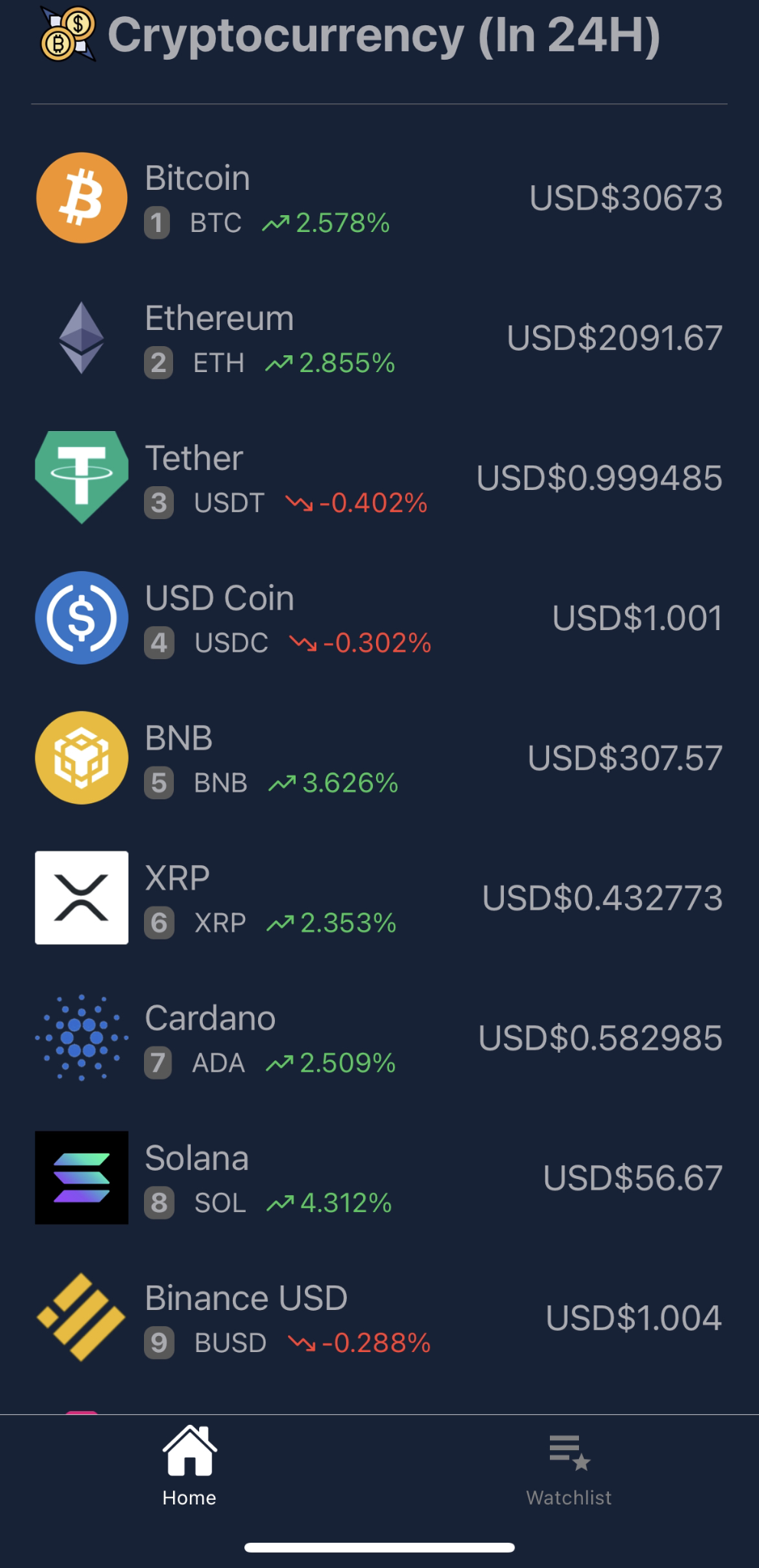


Figure 13 mobile application screen-dump of homepage



Figure 14 mobile application screen-dump of watchlist



Figure 15 mobile application screen-dump of detailpage

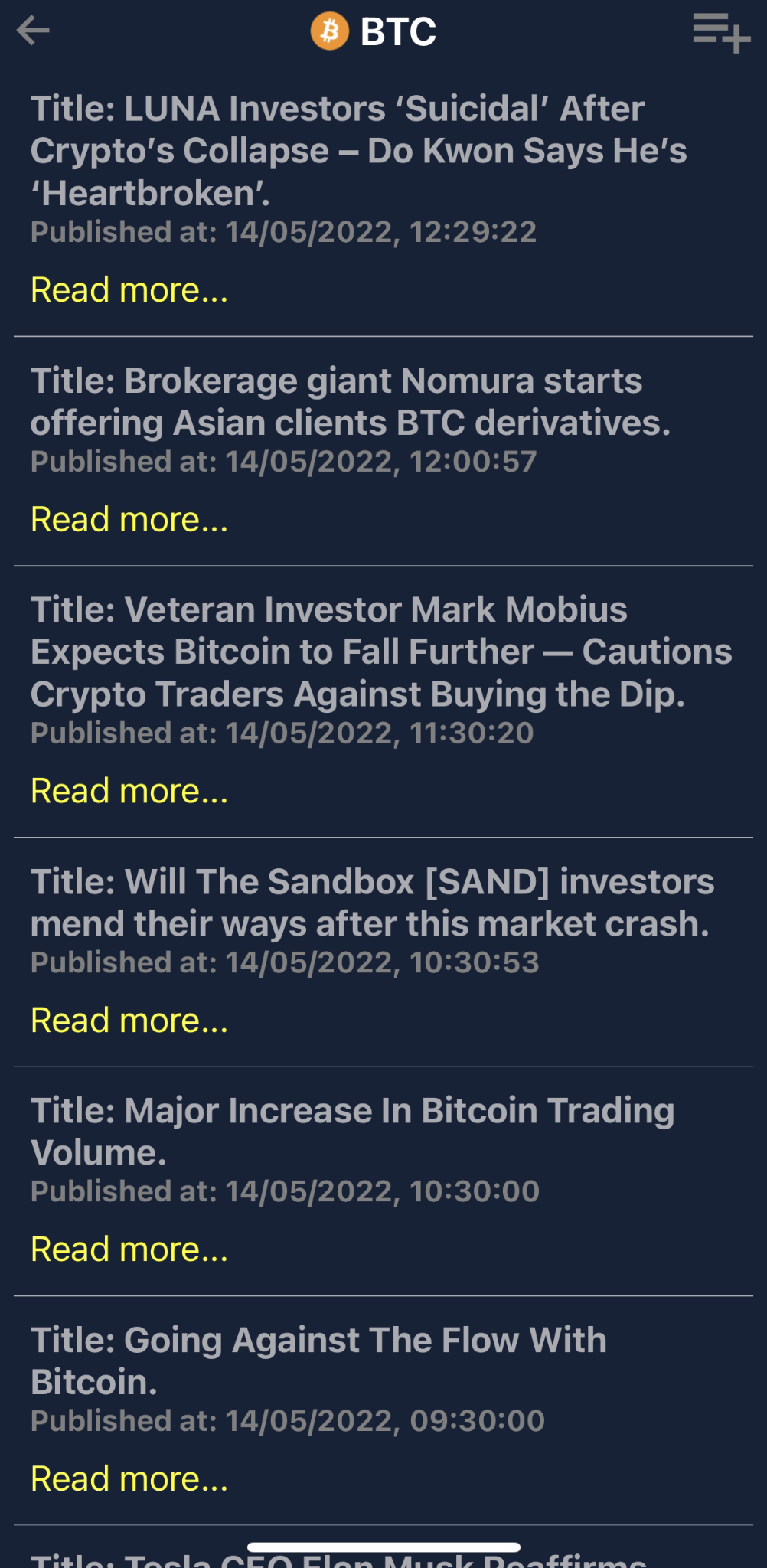


Figure 16 mobile application screen-dump of news function

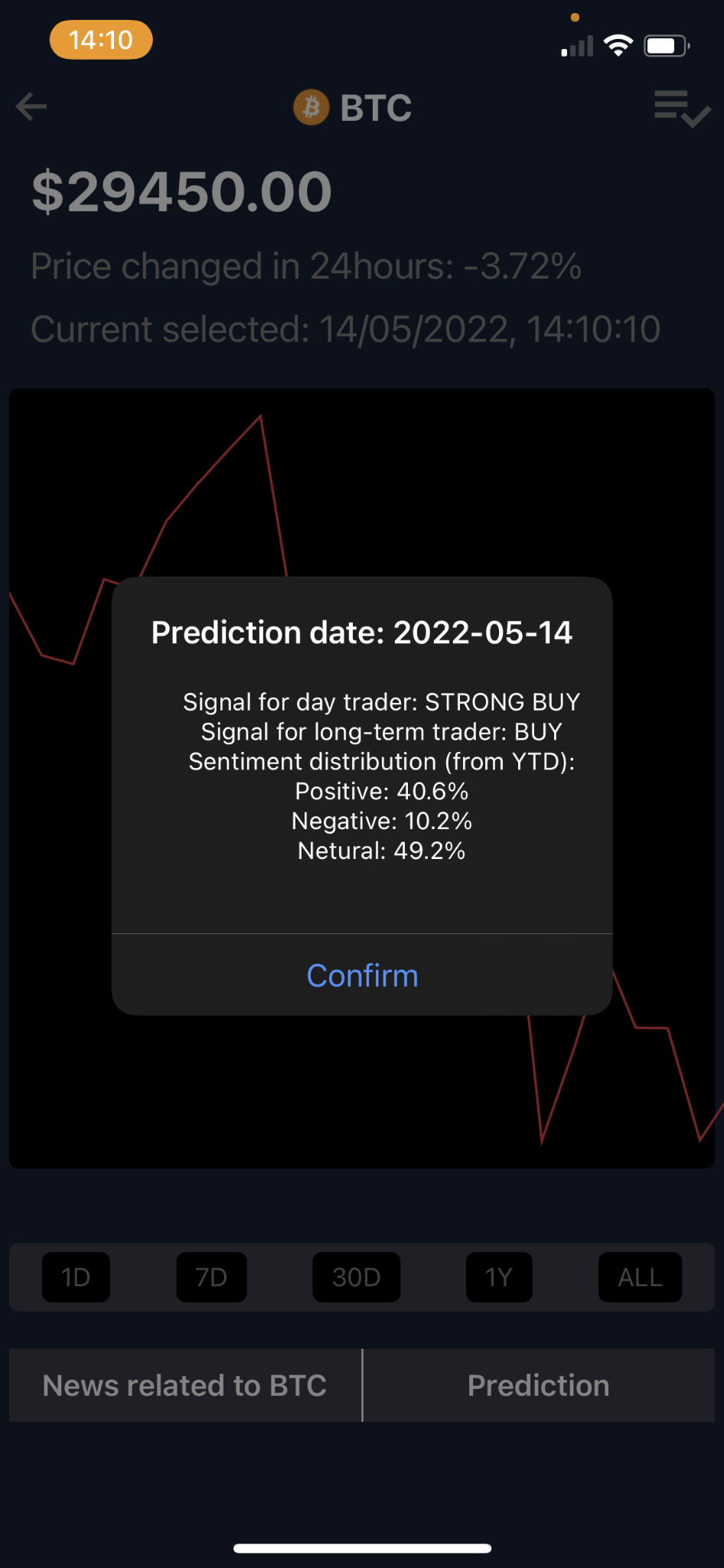


Figure 17 mobile application screen-dump of prediction function

From figure 13, we can see a list of cryptocurrencies with the price in 24 hours, which is data from the CoinGecko API. Then, a tab navigator at the bottom allows the user to choose whether homepage or watchlist. The user can check if there is any item on their watchlist by simply clicking on the watchlist tab, and it will show the result like figure 14. If users press on a cryptocurrency, they will move to the detail page.

In figure 15, we can see the current price, price change, current selected time, price chart, news button, and prediction button. Those data are obtained from the CoinGecko API, and if the user points to the price chart, the above value will keep changing to the pointing value. And if the user wants to see the other period of the price chart, they can press on the period bar to choose the period they want. Each time the user changes the period, the application will call the CoinGecko API again to obtain the data.

Furthermore, the user can press the "News related to BTC" button to execute the news function. In figure 16, the application obtained the news data by sending a request to the CryptoPanic API.

Lastly, the user can also press the "prediction " button to get the prediction result if the button is available. The result will show as a popup message like figure 17. And the result data is obtained from our trading signal server.

**Highlights of Key or Advanced Function(s):**

Prediction: When the user presses the prediction button, and if the function is available on the selected cryptocurrency, the mobile application will send a request to the trading signal server to get the prediction result. After the result has been returned, the result will show as an alert message in the mobile application.

Related news: When the user presses the news button, the mobile application will send a request to Cryptopanic API to get the news data. After the result has been returned, the mobile application will use the Flatlist component to list all related news.

# Chapter 5. Evaluation and Discussion

## Approach Evaluation (Trading Signal)

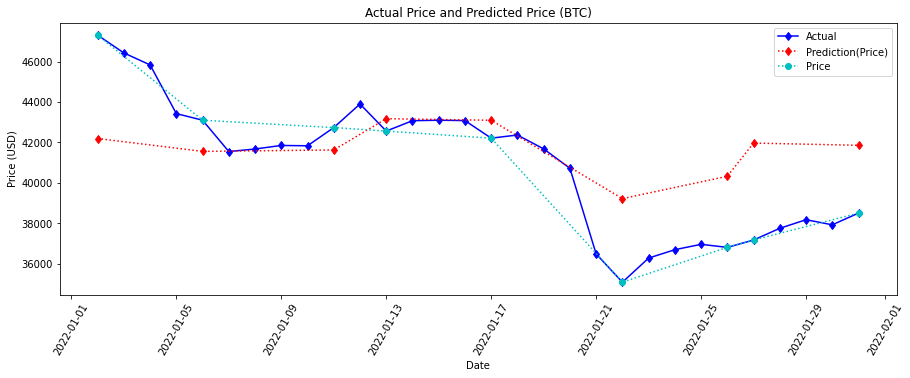


Figure 18 Plot between Actual, Test and Predicted Trend in Bitcoin

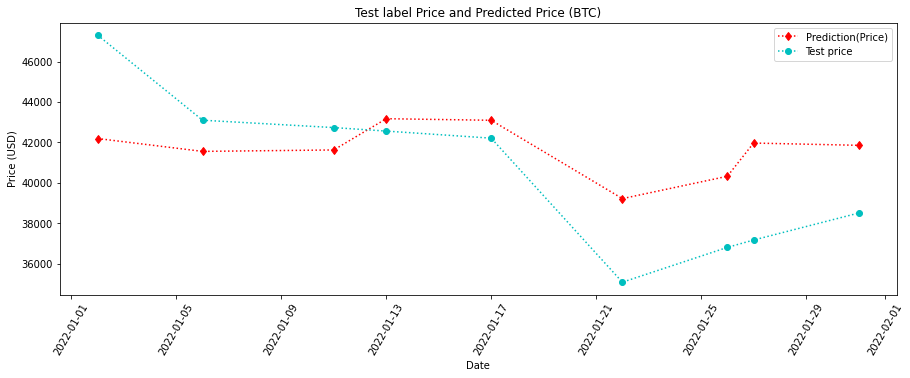


Figure 19 Plot between Test and Predicted Trend in Bitcoin

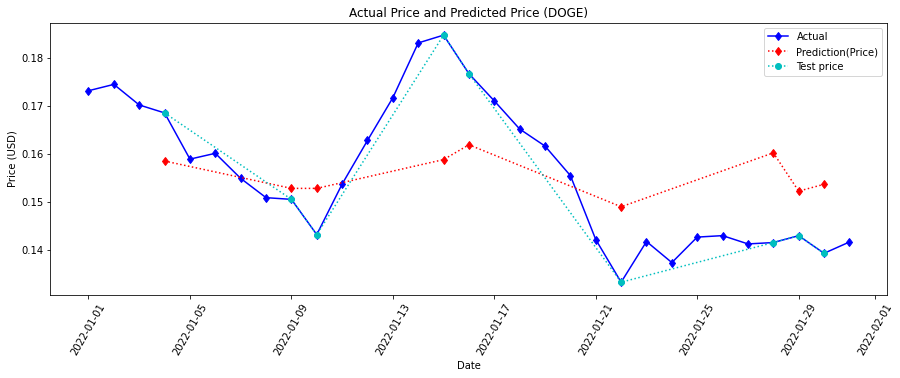


Figure 20 Plot between Actual, Test and Predicted Trend in Dogecoin

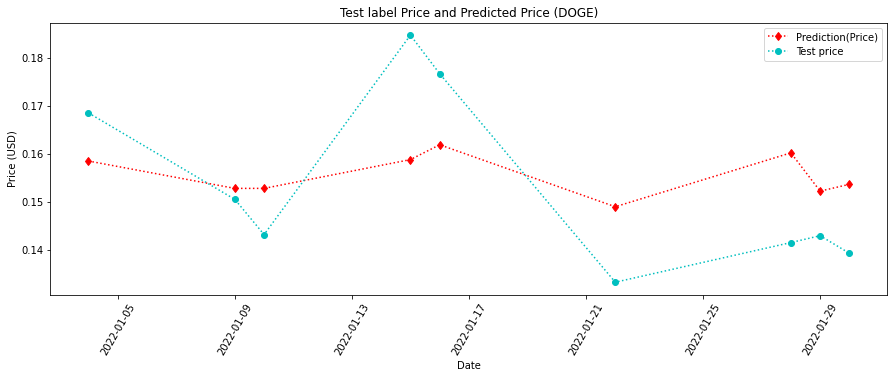


Figure 21 Plot between Test and Predicted Trend in Dogecoin

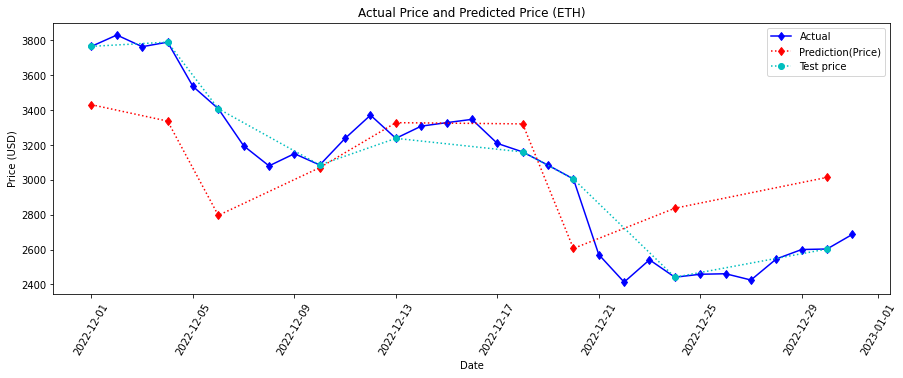


Figure 22 Plot between Actual, Test and Predicted Trend in Ethereum

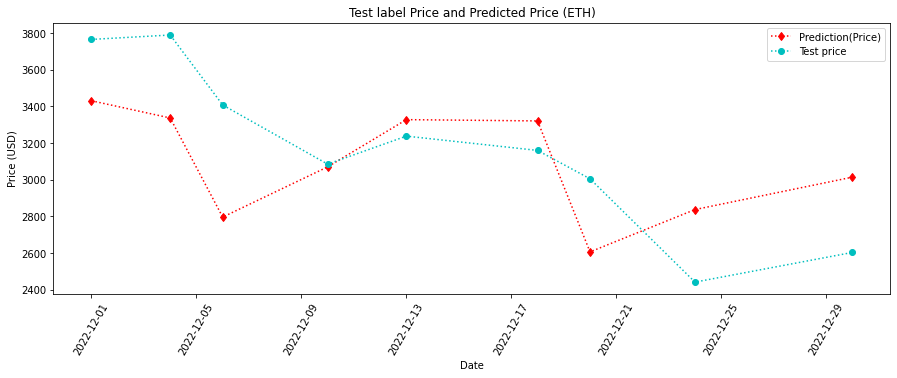


Figure 23 Plot between Test and Predicted Trend in Ethereum

|  |  |  |  |
| --- | --- | --- | --- |
| **Cryptocurrency** | **Test signal** | **Successful Signal** | **Accuracy** |
| Bitcoin | 8 | 5 | 62.5% |
| Dogecoin | 8 | 5 | 62.5% |
| Ethereum | 8 | 5 | 62.5% |

Table 8 Trading signal accuracy

The figures 18, 20, 22 are the results of the CNN-LSTM based model in 1 month of the Bitcoin, Dogecoin and Ethereum. The prediction of the trend is shown by the red line. The actual trend is shown by the blue line and the testing data which is randomly chosen from the actual trend is shown by the cyan line. Figure 19, 21, 23 has 8 periods and each edge between 2 points represents 1 period and clearly shows the comparison of the predicted trend and the testing trend. Table 8 listed the total signal and the successful signal in 1 month. The accuracy of the approach is 62.5% in predicting the selected cryptocurrencies. The results of the graphs and the tables proved that it is a stability approach and has the capability to predict the cryptocurrencies market trend and convert the results into a trading signal.

## Application Performance

The evaluation of the application performance is done by questionnaire. A 100 of survey is sent and received. All the interviewers are the users of the application.

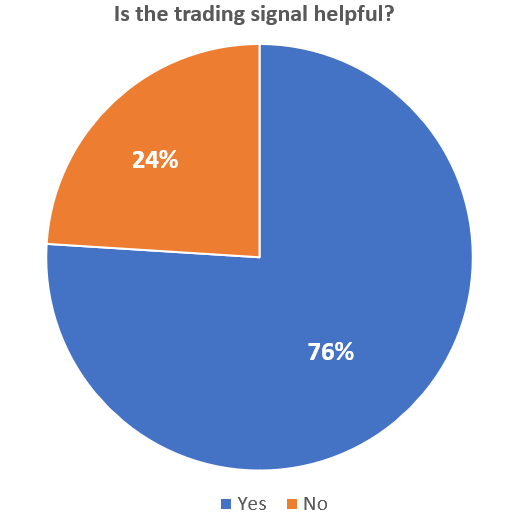


Figure 24 Survey results on trading signal performance

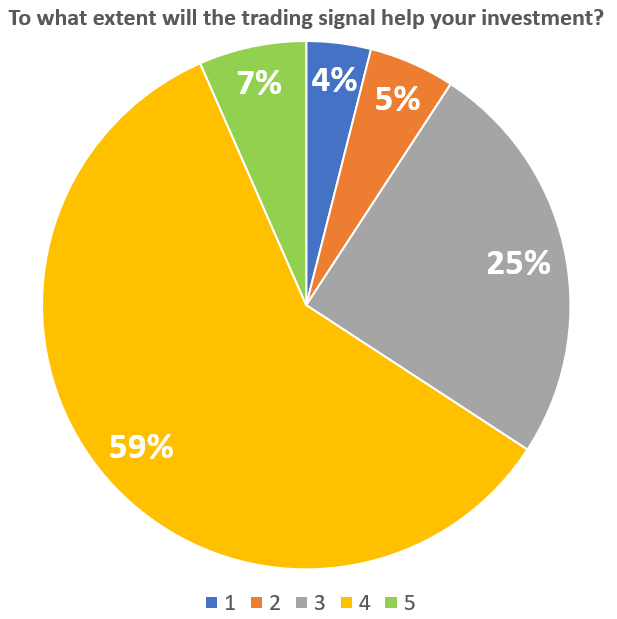


Figure 25 Survey results on trading signal satisfaction

The above pie chart figure 24 represents the user's reflection on the trading signal. There are 76% of the users think the trading signal is valuable. Figure 25 is the further question for the users that selected “Yes” in figure 24. The Results shows that 59% of the users scored the trading signal 4 out of 5 and 7% of the users scored 5. The two figures tells, the trading signal has successfully fulfilled the user’s expectation and requirement. However, there are still 24% of the users who think the trading signal can be improved.

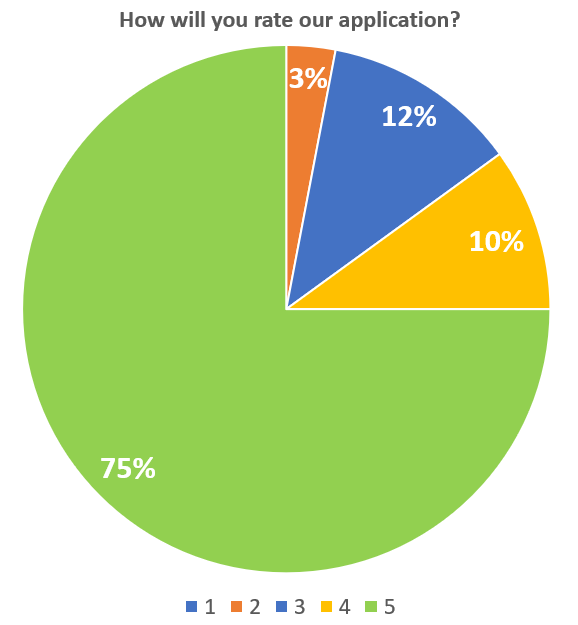


Figure 26 Users satisfaction on the application

Figure 26 is the overall performance of the application. There are 85% of the users give the application 4 stars or above. The results showed that a majority of the application users think their user experience in the application is good.

# Chapter 6. Conclusion

This project successfully develops a cryptocurrency price tracker mobile application and provides a trading signal based on the cryptocurrency market trends.

Different from the others projects, we focus on analyzing the sentiment of the 3 popular cryptocurrencies. To this end, we gathered the historic price data and scraped the tweets data of the user's reaction to the 3 popular cryptocurrencies. Then we can separate these data into price data and sentiment data. In order to analyze the sentiment data, we decided to use CNN-LSTM based model since it has the best accuracy compared to the other models. To improve the accuracy of the sentiment analysis, we create some train datasets which contain sentiment labels, “Positive”, “Negative” and “Neutral". After that, we use these train datasets to train the model and do the data preprocessing.

Moreover, we employ a regression model to predict the cryptocurrency price trend. The regression model is cooperating with the CNN-LSTM based model as the prediction result is based on the analyzed results. Then we import the prediction result into our approach. After that, the approach will generate the trading signal and display it in the mobile application. Besides, the results of the evaluation show that the users are satisfied when they are using the mobile application which means to predict cryptocurrency trends using machine learning and deep learning can help the investor to make decisions.

However, there are some limitations in our solution. First of all, the accuracy of the trading signals are not good enough. It can only reach 62.5% but our expectation is over 70%. The main factors affecting the prediction are not accurate because the data type is not enough. There are more factors that will also affect the cryptocurrency market trend, such as market related news, the status of world affairs and stock market trend. Apart from that, the amount of the used data in this project is small and there are some tweets posted by bot, therefore some error or problem may exist while doing the prediction. Besides, we found that the approach formula can be improved. As the trading signal is provided by the approach, we believe there are some defects in the formula.

In addition, there is a limitation in our mobile application. As the cryptocurrencies details in the application are obtained from API. If the API service is in downtime, the functions in the mobile application will stop working.

To solve the above problem, we planned to add changes in the future. First of all, more data is required. To improve the accuracy, we should use not only the tweets sentiment but also the cryptocurrency market news, the status of world affairs and stock market trends, so that the sentiment analysis model can more comprehensively analyze the cryptocurrency market trend. Also, the data preprocessing needs to be improved so that the bot posting can be reduced. And the formula should also be redesigned so that it can provide a more accurate trading signal.

Furthermore, an embedded databases can reduce the probability of the application crashing. As the application relies on the API to get the cryptocurrencies details, it is dangerous to only have one method to get the data. Holding an embedded database and storing the cryptocurrencies data in it is more safe.

To conclude, the application in this project is acceptable but can be improved. As it is created for the beginner and amateur cryptocurrency investor, it is an adequate application. In the future, the trading signal will be more accurate to help more investors make trading decisions.

# References

[1] Farell, R. (2015). An analysis of the cryptocurrency industry.

[2] Gandal, N., & Halaburda, H. (2014). Competition in the cryptocurrency market.

[3] Lamon, C., Nielsen, E., & Redondo, E. (2017). Cryptocurrency price prediction using news and social media sentiment. SMU Data Sci. Rev, 1(3), 1-22.

[4] Kim, K., Kim, J., & Rinaldo, A. (2018). Time series featurization via topological data analysis. arXiv preprint arXiv:1812.02987.

[5] Velankar, S., Valecha, S., & Maji, S. (2018, February). Bitcoin price prediction using machine learning. In *2018 20th International Conference on Advanced Communication Technology (ICACT)* (pp. 144-147). IEEE.

[6] Stenqvist, E., & Lönnö, J. (2017). Predicting Bitcoin price fluctuation with Twitter sentiment analysis.

[7] Jiang, Z., & Liang, J. (2017, September). Cryptocurrency portfolio management with deep reinforcement learning. In 2017 Intelligent Systems Conference (IntelliSys) (pp. 905-913). IEEE.

[8] Wang, Y., & Chen, R. (2020, January). Cryptocurrency price prediction based on multiple market sentiment. In Proceedings of the 53rd Hawaii International Conference on System Sciences.

[9] Abraham, J., Higdon, D., Nelson, J., & Ibarra, J. (2018). Cryptocurrency price prediction using tweet volumes and sentiment analysis. SMU Data Science Review, 1(3), 1.

[10] Colianni, S., Rosales, S., & Signorotti, M. (2015). Algorithmic trading of cryptocurrency based on Twitter sentiment analysis. *CS229 Project*, *1*(5).

[11] Jain, A., Tripathi, S., Dwivedi, H. D., & Saxena, P. (2018, August). Forecasting price of cryptocurrencies using tweets sentiment analysis. In *2018 eleventh international conference on contemporary computing (IC3)* (pp. 1-7). IEEE.

[12] Guresen, E., Kayakutlu, G., Daim, T.U.: Using artificial neural network models in stock market index prediction. Expert Systems with Applications 38(8) (2011) 10389 – 10397

[13] Xu, G., Meng, Y., Qiu, X., Yu, Z., & Wu, X. (2019). Sentiment analysis of comment texts based on BiLSTM. Ieee Access, 7, 51522-51532.

[14] García, S., Luengo, J., & Herrera, F. (2015). Data preprocessing in data mining (Vol. 72). Cham, Switzerland: Springer International Publishing.

[15] Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. The journal of educational research, 96(1), 3-14.

[16] Kraaijeveld, O & De Smedt, J 2020. The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. Identifying (cryptocurrency-related) Twitter bot accounts, 16(2.5).

# Appendix A Team Members’ Roles and Responsibility

**Roles table:**

|  |  |
| --- | --- |
| **Roles** | **Member(s)** |
| Team Coordinator | Bear |
| System Analyst and Designer | Bear, Rex, Toddy, Clinton |
| Programmer | Bear, Rex, Toddy, Clinton |
| Tester | Bear, Rex, Toddy, Clinton |
| UI Designer | Bear |
| Experiment in AI model | Clinton, Bear, Rex, Toddy |
| Report Writer | Bear, Rex, Toddy, Clinton |

**Responsibilities table:**

|  |  |  |
| --- | --- | --- |
| **Tasks** | **Member(s)** | **Target date** |
| Hold meeting periodically | Bear | N/A |
| Technology Test: AI model | Toddy | Nov 2021 |
| Find and analyze a suitable model | Rex, Toddy | Dec 2021 |
| Mobile Application: basic UI | Bear | Dec 2021 |
| Find corresponding API | Bear | Nov 2021 |
| Find related dataset | Bear, Clinton | Nov 2021 |
| Mobile Application: basic component | Clinton | Jan 2022 |
| Mobile Application: AI connect | Bear, Rex, Toddy, Clinton | Feb 2022 |
| Train AI: feed tweets data | Rex, Toddy | Feb 2022 |
| Server setup for the application | Clinton | Mar 2022 |
| Server setup for model and trading signal | Toddy | Mar 2022 |
| Application testing | Clinton | Apr 2022 |
| Mobile Application: UI improvement | Bear, Clinton | Mar 2022 |
| Mobile Application: functional improvement | Clinton | Mar 2022 |
| Model and application evaluation | Rex | Apr 2022 |
| Server Maintain | Clinton, Toddy, Bear, Rex | May 2022 |

# Appendix B1: Hung Wai Chak’s Final Report

1. **Information Box**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Information Box (Name and OUID, Team Name, Project Title, Supervisor’s Name)** | | | | |
| **Name** | **OUID** | **Team Name** | **Project Title** | **Supervisor’s Name** |
| Hung Wai Chak | 12555360 | Bear | A mobile application provide trading signal based on the cryptocurrency market tweets sentiments | DR. LEUNG MAN FAI HENRY |

1. **Declaration Statement**

I, Hung Wai Chak (SID:12555360), certify that the description and information included in this team member's report is true to the best of my knowledge.

Signature:

一張含有 箭 的圖片

自動產生的描述 and 12/05/2022

1. **Tasks Assigned to the Author and their Status**

|  |  |  |  |
| --- | --- | --- | --- |
| The following table outlines the task assigned to me in this project. | | | |
| Tasks | Responsible Member(s) | Target Date | Completion |
| Mobile Application: basic UI | Bear | Dec 2021 | 100% |
| Mobile Application: AI connect | Bear, Rex, Toddy, Clinton | Feb 2022 | 100% |
| Mobile Application: UI improvement | Bear, Clinton | May 2022 | 100% |
| Server Maintain | Clinton, Toddy, Bear, Rex | May 2022 | 100% |
| Find and test corresponding API | Bear | Nov 2021 | 100% |

1. **Review and Appraisal of Code Section**

**section 1:**

**一張含有 文字 的圖片

自動產生的描述**

appendix code screenshot 1

**一張含有 文字 的圖片

自動產生的描述**

appendix code screenshot 2

Section 2: Explain the design and operation of the code section.

The above screenshots shows the code section related to the news function for our moblie application, and the screenshot 1 shows the formation of a news item, as line 11, the item should contain a pair of JSON object which is pass from detail page and the “symbol” refer to the cryptocurrency symbol sunch as “btc”,”eth” , the “uri” refer to the relate icon , and the “id” refer to the item id.The function “fetchNews” in line 15 is use to get news data base on the “symbol”.the line 32-47 is use to provide a header view of the user-choesn crypto.The line 39-48 is use to return a view of the news for the user using a flatlist to loop though ListNewsItem, the in line 41 “id” refer to index key of JSON which get from Cryptopsnic API.

The screenshot 2 is a compoent of NewItem that make the code are reusable. The NewItem will call ListNewItem to list out each new data with the flatlist looping. Moreover, the NewItem also contain the function of open in-app browser with the library “expo-web-browser”, the user can press on the button in line 21 to open the read more of NewItem.

1. **Short Essay on Solving a Problem**

As the team leader and team Coordinator of this project, my primary duties are to ensure the process operates smoothly. Therefore, I need to assign different tasks to team members based on their ability and hold a meeting periodically to update the progress or bring up difficulties. For more efficient to manage, I divided the project into two-part: the general cryptocurrency analysis mobile app and the other is the formation of the trading signal, which involves implementing some machine learning techniques.

In the early of our report, we intend to collect the news data related to cryptocurrency for training a sentiment analysis model base on the method provided by previous research. But unfortunately, the design of the Coindesk news website has been changed. And the number of news headlines that can be web scraped is limited to a small among, which is not enough to train a good model, while there is no suitable dataset with labeled data for required data. Therefore, I decided to focus on another aspect, analyzing the sentiment in social media posts. From the literature review, we choose Twitter as the platform to collect related tweets. I found and tested some helpful APIs, Twitter API, which can retrieve 500,000 Tweets per month, and Tweepy is a Python library for accessing the Twitter API.

At the beginning of the application development, I had the task of designing a basic UI, and the initial design of the trading signal is as left of the following table:

|  |  |
| --- | --- |
| 一張含有 文字 的圖片  自動產生的描述 | 一張含有 文字 的圖片  自動產生的描述 |
| Before | After |

as there are half-page blank, only the upper part of the interface show some related detail; this design will affect the user experience. As shown on the right side of the table above, I change the function from a page to a pop-up screen using the alert library to provide a clean and user-friendly interface to improve this issue.

The other problem that I face is that we need to provide related data for different functions, such as the price data or the related news for the cryptocurrency we selected. Due to a large amount of data and high-frequency update requirements, we can't only use web-scraping to collect those data and store them in a database. Therefore, finding a suitable API is the best solution for the time cost and capacity issues.

After comparison, The Coingecko API is chosen to give price data, and Cryptopanic API is chosen to provide related news.

1. **Self-Appraisal of Contributions**

|  |  |
| --- | --- |
| **Tasks** | **Member(s)** |
| System design (Mobile application) | 0% |
| System design (Model) | 0% |
| System implementation and testing  (Mobile application) | 30% |
| System implementation and testing  (Model) | 0% |
| Literature research and content analysis | 25% |
| Problem analysis and formulation | 40% |
| Development of evaluation plan | 0% |
| Development of integration plan | 0% |
| Project management | 70% |
| Project administrative work | 100% |
| Presentation preparation and delivery. | 40% |
| Liaison with external partners | 0% |

# Appendix B2: Tsoi Siu Fung’s Final Report

1. **Cover Page**

Final Report

A mobile application provide trading signal based on the cryptocurrency market tweets sentiments

**TEAM BEAR**

MEMBER NAME:

TSOI SIU FUNG, Rex (12586593)

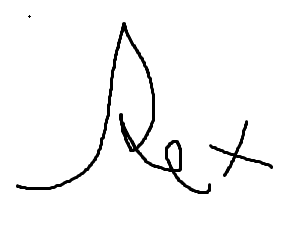
SUPERVISOR:

DR. LEUNG MAN FAI HENRY

1. **Declaration Statement**

I, (Tsoi Siu Fung, 12586593), certify that the description and information included in this team member's report is true to the best of my knowledge.

signatural

and 12/05/2022

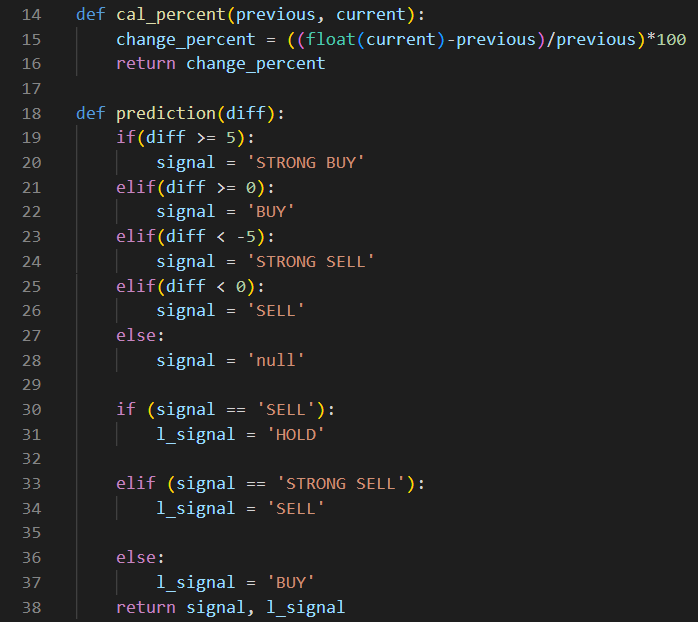
1. **Tasks Assigned to the Author and their Status**

**Responsibilities table:**

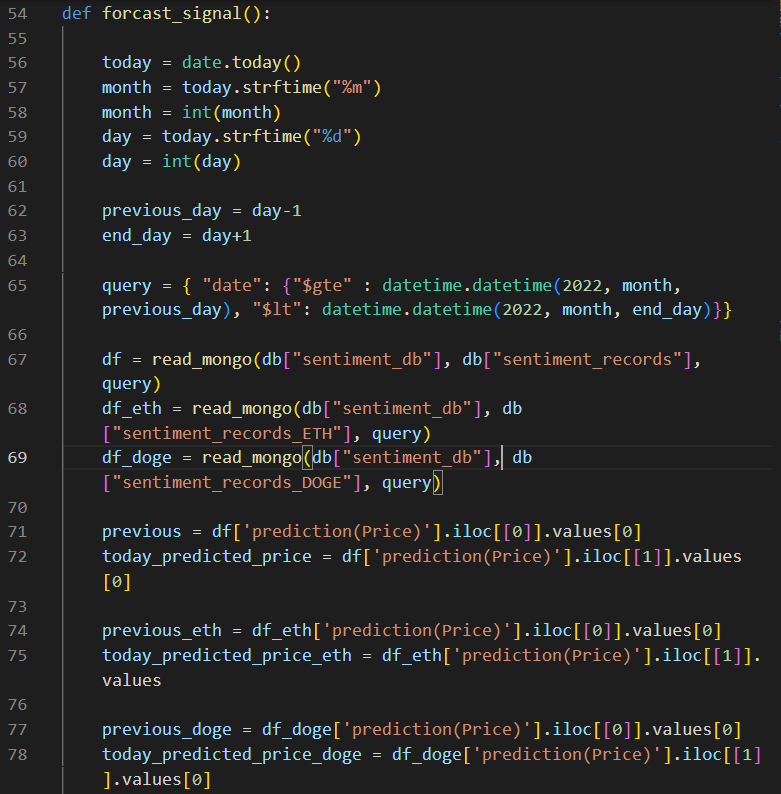
|  |  |  |  |
| --- | --- | --- | --- |
| The following table outlines the task assigned to me in this project. | | | |
| **Tasks** | **Member(s)** | **Target date** | **Completion** |
| Find and analyze a suitable model | Rex, Toddy | Dec 2021 | 100% |
| Mobile Application: AI connect | Bear, Rex, Toddy, Clinton | Feb 2022 | 100% |
| Train AI: feed tweets data | Rex, Toddy | Feb 2022 | 100% |
| Model and application evaluation | Rex | Apr 2022 | 100% |
| Server Maintain | Clinton, Toddy, Bear, Rex | May 2022 | 100% |

1. **Review and Appraisal of Code Section**

**Section 1**

****

Picture 1

****

Picture 2

****

Picture 3

****

Picture 4

****

Picture 5

**Section 2**

The picture 1 has 2 functions. The aim of the cal\_percent function is to calculate the price change percentage. Another function named prediction is to provide trading signals based on the condition.

Picture 2 shows the first part of the forcast\_signal function, the main purpose of this part is to get the data from the mongodb using read\_mongo() and import cryptocurrency's previous price into the data frame.

Picture 3 is the second part of the forcast\_signal function, line 86 to 88 is to calculate the difference between predicted price and the previous price. Line 90 to 92 is to provide the trading signal based on the price difference. Line 95 to 104 is to insert the Bitcoin price data into the table column. Line 107 to 110 is to convert the table into a json object. Line 113 to 148 is inserting another 2 price data into the table column and converting it into json.

Picture 5 line 151 to 156 is for testing if the program is running or not every 1 hour. Line 159 is to do the prediction everyday at 12am. Line 162 to 164 is the schedule library default code.

1. **Short Essay on Solving a Problem**

There are 5 tasks assigned to me. As the project aim is to provide a market trading signal based on the analyzed tweets sentiment. Therefore, finding a suitable sentiment analysis model is linked with the success of the project. However, there are a lot of sentiment analysis models on the internet. To choose a fitting model, me and my teammate use the analyze method mentioned in chapter 3. As a result, the CNN-LSTM model has the best performance, so CNN-LSTM model is choosed.

After selecting the CNN-LSTM model, I am able to solve the project problem which is to predict the cryptocurrency’s future trends by analyzing the tweets sentiment and implementing the model into the mobile application. Implementing the approach into the application is not a difficult job, I just need to deploy the model on the flask server. Once the application sends an api request to the server, the application will receive a json that includes the distribution of the sentiment analysis results and the trading signal. The biggest challenge that I faced is to find the tweets dataset and use it to train the sentiment model. Since there are only a small amount of the cryptocurrencies dataset on the internet. At first, I tried to build a tweets scraper to create the dataset for model training. However, it will take a lot of time to label each tweet. Fortunately, I found a website called Kaggle that has the suitable datasets for me to train the model.

Another important task assigned to me is model and application evaluation. As it is the part to see whether the model and application performance did reach our expectation, I considered sending questionnaires to evaluate the application performance and use python libraries to help me to check the signal accuracy and the sentiment model performance by creating a comparison graph and table shown in chapter 5.

1. **Self-Appraisal of Contributions**

|  |  |
| --- | --- |
| **Tasks** | **Self-Contribution** |
| System design (Mobile application) | 0% |
| System design (Model) | 0% |
| System implementation and testing  (Mobile application) | 0% |
| System implementation and testing  (Model) | 30% |
| Literature research and content analysis | 25% |
| Problem analysis and formulation | 30% |
| Development of evaluation plan | 100% |
| Development of integration plan | 0% |
| Project management | 10% |
| Project administrative work | 0% |
| Presentation preparation and delivery. | 20% |
| Liaison with external partners | 100% |

# Appendix B3: Lam Chun Hin’s Final Report

1. Cover Page

Final Report

A mobile application provide trading signal based on the cryptocurrency market tweets sentiments

**TEAM BEAR**

MEMBER NAME:

LAM CHUN HIN, CLINTON (12549757)

SUPERVISOR:

DR. LEUNG MAN FAI HENRY

1. Declaration Statement

I, Lam Chun Hin (12549757), certify that the description and information included in this team member's report is true to the best of my knowledge.

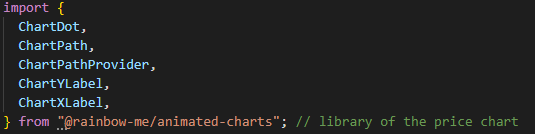
and 12/05/2022

1. **Tasks Assigned to the Author and their Status**

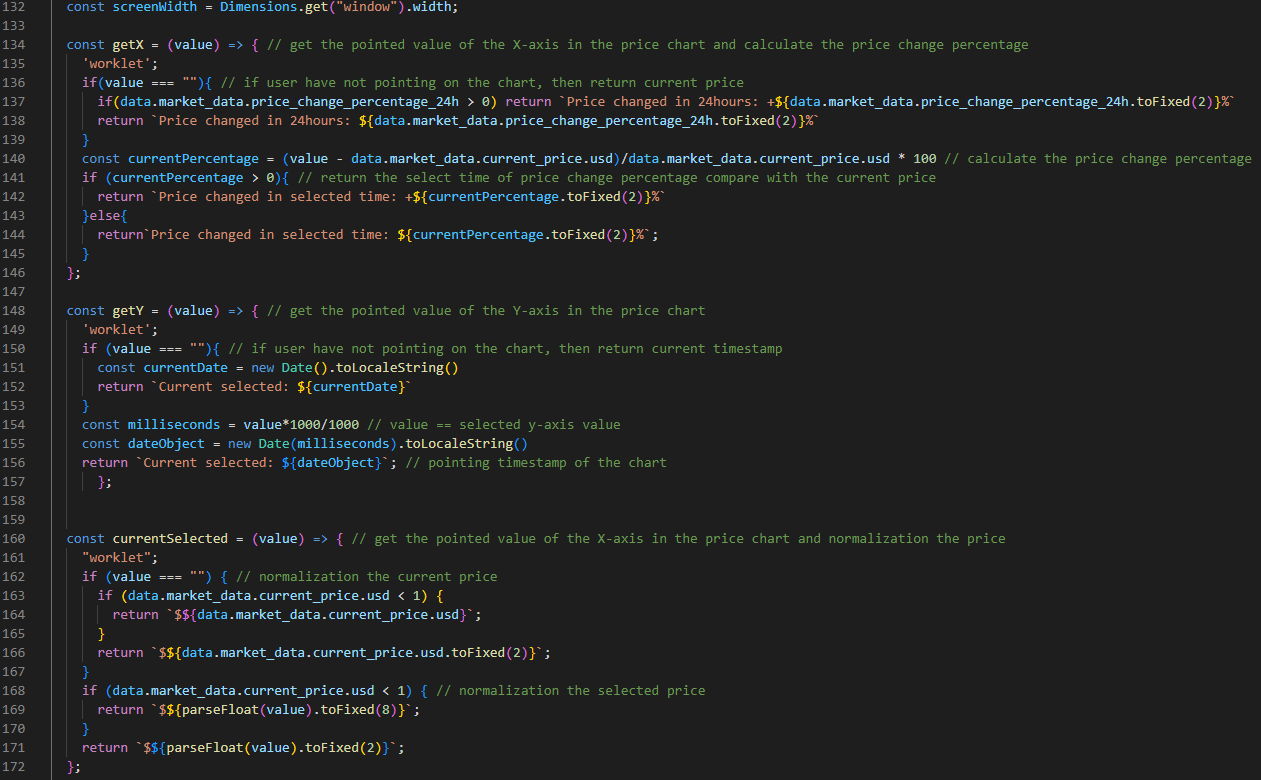
|  |  |  |
| --- | --- | --- |
| **Tasks** | **Member(s)** | **Completion** |
| Mobile Application: basic component | Clinton | 100% |
| Mobile Application: AI connect | Bear, Rex, Toddy, Clinton | 100% |
| Server setup for the application | Clinton | 100% |
| Application testing | Clinton | 100% |
| Mobile Application: UI improvement | Bear, Clinton | 100% |
| Mobile Application: functional improvement | Clinton | 100% |
| Server Maintain | Clinton, Toddy, Bear, Rex | 100% |

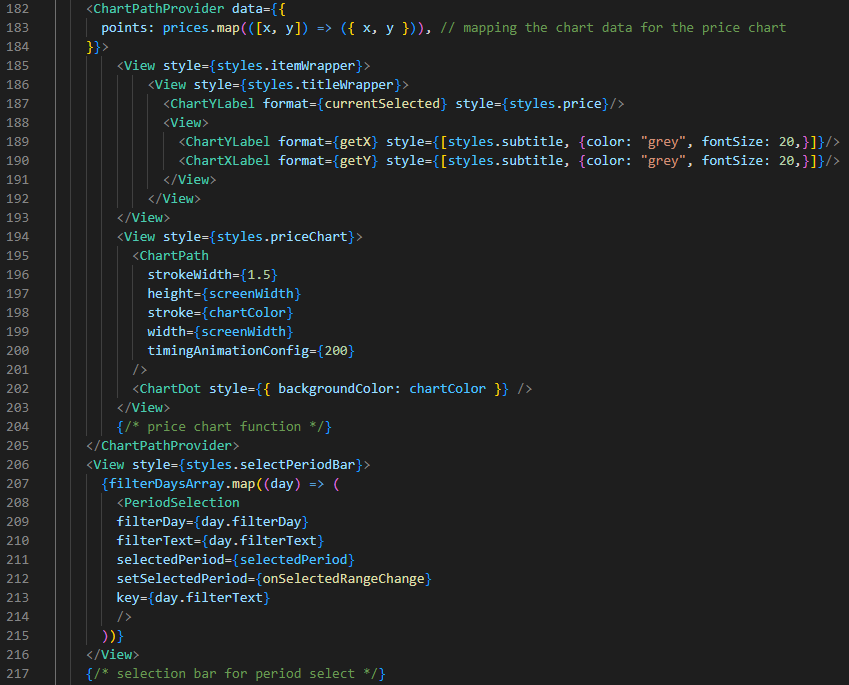
1. **Review of Code Selection**

Section 1: Display a selected code section



appendix code screenshot 1

appendix code screenshot 2



appendix code screenshot 3

Section 2: Explain the design and operation of the code section.

To display the price chart and show the price and timestamp of the selected point in the price chart. The "@rainbow-me/animated-charts" library has been selected.

Furthermore, there are three methods to get and normalize different values of the price chart. They will keep changing the return value when the user points at the price chart. Otherwise, they will show the original data from CryptoGecko API.

Firstly, from lines 134 to 146, the "getX" function is used to get the x-axis value of the selected point in the price chart. From lines 148 to 157, the "getY" function is used to get the y-axis value of the selected point in the price chart. From lines 160 to 172, the "currentSelected" function is used to get the x-axis value of the selected point in the price chart, and it is used to normalize the price. Then, lines 182 to 184 are used to assign data to the price chart and display the price chart in the application. After that, lines 187 to 190 are used to call the above functions and show the value of the returned value. Lastly, lines 195 to 202 define the price chart's screen size and the polyline color that is determined by the percentage of price change.

1. **Short Essay on Problem Solving**

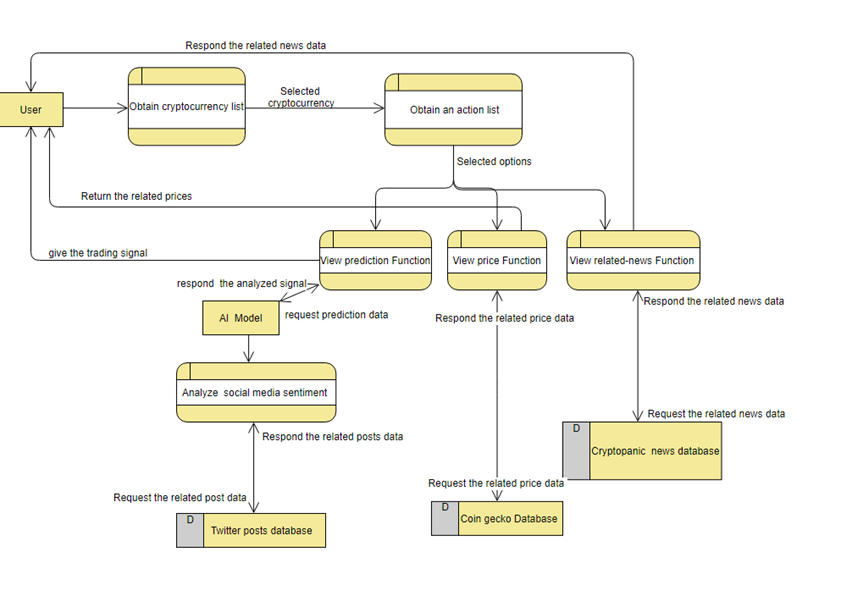
****

Figure x. An earlier version of data-flow diagram for mobile application \*\*未mark number

As a system designer and programmer of this project, my task is mainly to build the mobile application. Meanwhile, in the earlier section, my tasks are to collect a suitable dataset and find a suitable model. And I faced different problems when finding an appropriate dataset. Our first project aims to use sentiment analysis to analyze the cryptocurrency news title to determine future trends. Nevertheless, most of the dataset I found is unlabeled, but the size of the dataset is too large for us to label it manually. Therefore, I tried to use some classification models to label the dataset and decided to use the labeled dataset to train our model. However, before we took the labeled dataset as the training data, I found that the labeled result was uncertain and not our expectation. Therefore, I tried web scraping from Coindesk, but the number of historical news we can scrape is limited due to the website design change. Lastly, our solution uses Tweepy to collect the training datasets instead of web scraping, and the target changed to tweets sentiment analysis instead of news title.

Moreover, when I finished the prototype of our mobile application with the earlier system design. My system design is rudimentary compared to existing applications like Binance and Cryptohero. The functions of the prototype provided are minimal, as shown in figure 1. Therefore, I added some advanced functions to make the mobile application more functional and contain more information for the user, such as showing detail of a selected cryptocurrency and the watchlist function. Besides, when we were testing the first version of the mobile application prototype, we also found a problem with the data-flow diagram of the application. As Figure 1 shows, after the obtain cryptocurrency list process, the application does not have an entity containing and showing the list data. Also, the process has not connected to the Coingecko API and request for the data. Therefore, I improved the data-flow diagram and rebuilt the mobile application to make it more user-friendly.

1. **Self-Appraisal of Contributions**

|  |  |
| --- | --- |
| **Tasks** | **Member(s)** |
| System design (Mobile application) | 100% |
| System design (Model) | 0% |
| System implementation and testing  (Mobile application) | 70% |
| System implementation and testing  (Model) | 0% |
| Literature research and content analysis | 25% |
| Problem analysis and formulation | 0% |
| Development of evaluation plan | 0% |
| Development of integration plan | 100% |
| Project management | 10% |
| Project administrative work | 0% |
| Presentation preparation and delivery. | 20% |
| Liaison with external partners | 0% |

# Appendix B4: Cheng Yiu Hang’s Final Report

1. **Cover Page**

Final Report

A mobile application provide trading signal based on the cryptocurrency market tweets sentiments

**TEAM BEAR**

MEMBER NAME:

Cheng Yiu Hang, TODDY (12598033)

SUPERVISOR:

DR. LEUNG MAN FAI HENRY

1. **Declaration Statement**

I, Cheng Yiu Hang (SID:12598033, certify that the description and information included in this team member's report is true to the best of my knowledge.

Signature:

 date:16/5/2022

1. **Tasks Assigned to the Author and their Status**

**Responsibilities table:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Tasks** | **Member(s)** | **Target date** | **Completion** |
| Technology Test: AI model | Toddy | Nov 2021 | 100% |
| Find and analyze a suitable model | Rex, Toddy | Dec 2021 | 100% |
| Mobile Application: AI connect | Bear, Rex, Toddy, Clinton | Feb 2022 | 100% |
| Train AI: feed tweets data | Rex, Toddy | Feb 2022 | 100% |
| Server setup for model and trading signal | Toddy | Mar 2022 | 100% |
| Server Maintain | Clinton, Toddy, Bear, Rex | May 2022 | 100% |

1. **Review and Appraisal of Code Section**

**Section 1**

|  |
| --- |
|  |
| Picture 1 |

|  |
| --- |
|  |
| Picture 2 |

|  |
| --- |
|  |
| Picture 3 |

**Section 2**

The code section is from the daily analysis area. The focus of the code is to handle daily cryptocurrency tweets and sentiment analysis. The prediction is based on the previous day’s sentiment and does the new prediction for today.

I have chosen bitcoin (BTC) for the demonstration, the other cryptocurrency code is removed for clean understanding. The year variable is set to 2022 (current) year for the prototype.

The first picture 1 has the function analysis(month, day), it shows the steps of bitcoin prediction. (month, day) is generated by the datetime.now() from the datetime module, which means the program is always running the current day prediction. The analysis function is combined with a few different python functions, by this way it is much easier to maintain.

Picture 2 is the first function called from analysis, the main purpose of this part is to preprocess all the previous day tweets, e.g. remove noise data, remove URL, remove tweet related to giveaway, and tweets created by bots…. After the preprocessing, now the bitcoin tweets are suitable for an AI model to analyze the sentiment of the tweets. And the result is saved to a CSV format, and the file name has been renamed to ready, and let others know this tweets data set is already preprocessed and prepared for prediction.

Picture 3 is the btc\_sentiment\_analysis.predict(month, day), line 56 to 64 is tokenization for the text from the preprocessed tweets, and the tweets will be split into words and more basic words. After that, Line 66 to 68 is using the trained BTC CNN-LSTM sentiment analysis model to analyze the previous day's tweets sentiment. Line 70 to Line 78 is writing the sentiment result of that day in percentage (Positive:40%/Neutral:50%, Negative:10%). Last, the analyzed tweet sentiment result will save to CSV format and rename to pre, and let others know this dataset is ready, and the next step is to add the cryptocurrency price to the dataset for the next regression model prediction.

1. **Short Essay on Solving a Problem**

There are 6 tasks assigned to me, the prediction pipeline from collecting cryptocurrency tweets to output a daily prediction. It is a challenge to me to make all the processes update daily.

The first challenge I have faced is preprocessing the cryptocurrency tweets. For example, I must think outside of the box for tweet preprocessing. Twitter “bot” and “give away account” are commonly spreading misinformation and noisy information. We must decide whether we let our sentiment model analyze this sentiment. Our final decision removes that noisy tweet, but the effects of removing noisy tweets are not evaluated, and we believe it is still a big area for people to research more.

To analyze a suitable model, we must find a suitable tweets-related dataset in order to evaluate the sentiment analysis model more appropriately. For instance, people who choose to post on Instagram are completely different to Twitter, people tend to post more carefully and formally. After choosing the model test for evaluation, my teammate and I use the analysis method mentioned in chapter 3, the accuracy and confusion matrix. And we found that CNN-LSTM neural network model performs the best for sentiment analysis.

After selecting the sentiment analysis model, the next step is to train a regression model to evaluate the relationship between cryptocurrency sentiment and cryptocurrency price. Like, will the sentiment score affect cryptocurrency price? To do this, it requires many Twitter (Bitcoin, Ethereum, Dogecoin) tweets to make a large dataset. But due to the limitation of the Twitter API policy, it only allows finding historical tweets within 14 days. Therefore, we have tried to use a scraper to find the historical tweets, but the scraping speed is also limited by Twitter, as a result, our training regression model data set only contains Jan 2022 with three different coins, and at least one million tweets are analyzed.

There is an alternative way, using a dataset collected 2018-2019 tweets posted on Twitter, but we believe the past historical tweets and the words are different from now, so we tend to use the tweets dataset we have collected although it may have a small sample size problem. As we believe, in future research collecting at least one year of data will improve the credibility of the prediction.

Finally is to put our whole prediction pipeline to cloud services. Containing the result from the model and predict signal approach. One limitation of choosing cloud services is that most of the platforms charge money for deployment. As a free platform for deployment with some limits, like no internet access. Finally, we chose PythonAnywhere and paid for the service to deploy our Flask application online. Now our mobile application can request cryptocurrency signal results from Flask anytime.

Lastly, after this project, I have completed lots of challenges and learned a lot about the importance of teamwork. If without teamwork then the project won't be successful.

1. **Self-Appraisal of Contributions**

|  |  |
| --- | --- |
| **Tasks** | **Self-Contribution** |
| System design (Mobile application) | 0% |
| System design (Model) | 100% |
| System implementation and testing  (Mobile application) | 0% |
| System implementation and testing  (Model) | 70% |
| Literature research and content analysis | 25% |
| Problem analysis and formulation | 30% |
| Development of evaluation plan | 0% |
| Development of integration plan | 0% |
| Project management | 10% |
| Project administrative work | 0% |
| Presentation preparation and delivery. | 20% |
| Liaison with external partners | 0% |

# 

# 

# Appendix C Progress Report

**Progress Report:**

The progress of this project is satisfactory. Two of the five major objectives have been completed and three of them are doing.

Tasks completed and tasks ongoing:

**The following lists the tasks completed:**

* Collect and Organize news or social media posts data from selected cryptocurrency news websites. (Objective 1)
* Develop an AI model to analyze the collected data. (Objective 2)

**The following lists the ongoing tasks:**

* Design and develop an approach for predicting cryptocurrency's future trends or prices fluctuation with given data. (Objective 3)
* Develop a mobile application for trading recommendations with an AI model. (Objective 4)
* Evaluate the accuracy of the prediction and the usability of the mobile application. (Objective 5)

**Gnatt chart:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Week** | **1 - 4** | **5 - 12** | **13 - 16** | **17 - 22** |
| Topic background research |  |  |  |  |
| Design the solution |  |  |  |  |
| Design application UI |  |  |  |  |
| Develop application component |  |  | |  |
| Find AI model |  | |  |  |
| Train the AI model |  |  | |  |
| Implement the AI model to the server |  |  | |  |
| Connect the AI to the application |  |  |  |  |
| Testing the app |  |  |  |  |

# Appendix D Team Meeting Minutes

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 29/9 | 2pm - 4pm | F2F |
| Member present: | Rex, Bear, Clinton, Toddy | | |
| Description: | * Discuss the selection of the topic * Do the topic research (all topic) | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 7/10 | 2pm - 4pm | F2F |
| Member present: | Rex, Bear, Clinton, Toddy | | |
| Description: | * Do the topic research (the selected topic) | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 13/10, 16/10, 17/10, 20/10 | 2pm - 6pm | Discord(online) |
| Member present: | Rex, Bear, Clinton, Toddy | | |
| Description: | * Discuss the proposal flow * Define the problem * Find the solution * Discuss the project aim * Find reference * Discuss the presentation flow | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 28/10 | 2pm - 3pm | F2F |
| Member present: | Rex, Bear, Clinton, Toddy | | |
| Description: | * Initial Report discussion | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 11/11, 13/11, 14/11 | 11am - 2pm | Discord(online) |
| Member present: | Rex, Bear, Clinton | | |
| Description: | * Seperate the work * Do the research of the topic deeply * Set the deadline of each part | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 11/11, 13/11, 14/11 | 3pm - 8pm | Discord(online) |
| Member present: | Rex, Bear, Clinton, Toddy | | |
| Description: | * Discuss what can be improved in Session 1 and 2 | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 15/11 - 17/11 | 11am - 2pm | Discord(online) |
| Member present: | Rex, Bear, Clinton | | |
| Description: | * Discuss what can be change in Session 2 * Find the reference that are useful in Session 2 * Discuss Session 3 | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 15/11 - 17/11 | 4pm - 9pm | Discord(online) |
| Member present: | Rex, Bear, Clinton, Toddy | | |
| Description: | * Discuss what can be change in Session 2 * Discuss Session 3 | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 25/11 | 7pm - 8pm | Discord(online) |
| Member present: | Rex, Bear, Clinton, Toddy | | |
| Description: | * Change the responsibilities | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 30/11 | 2pm - 4pm | Discord(online) |
| Member present: | Rex, Toddy | | |
| Description: | * Testing model and API | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 31/11 | 3pm - 5pm | Discord(online) |
| Member present: | Bear, Clinton | | |
| Description: | * Design application feature and UI * find API | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 20/1 | 2pm - 3pm | Discord(online) |
| Member present: | Rex, Toddy, Bear, Clinton | | |
| Description: | * Review the grade of initial report * Plan to build the application and design the system | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  | Date | Time | Place |
| 15/1 - 30/1 | 2pm - 6pm | Discord(online) |
| Member present: | Rex, Toddy, Bear, Clinton | | |
| Description: | * Discuss the interim Report * Find suitable model * Find API * Complete the application features * Improve the UI of the application | | |
|

1. https://www.ig.com/en/cryptocurrency-trading/benefits-of-cryptocurrency-trading [↑](#footnote-ref-1)
2. https://www.financialexpress.com/market/crypto-rally-total-market-cap-hits-new-all-time-high-of-2-8t-1t-added-in-just-over-a-months-time/2362504/ [↑](#footnote-ref-2)
3. Education, I. C. (2021, August 3). Neural networks. IBM Cloud Learn Hub. https://www.ibm.com/cloud/learn/neural-networks [↑](#footnote-ref-3)
4. What are the differences between a deep neural network and multilayer perceptron? | Data Science and Machine Learning | Kaggle. (n.d.). Kaggle: Your Machine Learning and Data Science Community. Retrieved May 18, 2022, from https://www.kaggle.com/getting-started/46727 [↑](#footnote-ref-4)
5. P. (2019, September 27). What is Web Scraping and What is it Used For? | Definition and Examples EXPLAINED[Video].YouTube. https://www.youtube.com/watch?v=Ct8Gxo8StBU&feature=youtu.be [↑](#footnote-ref-5)
6. Oliver Knocklein (2019, Jun 6) Classification Using Neural Networks. https://towardsdatascience.com/classification-using-neural-networks-b8e98f3a904f [↑](#footnote-ref-6)
7. NLTK :: Natural Language Toolkit. (n.d.). NLTK :: Natural Language Toolkit. Retrieved May 18, 2022, from https://www.nltk.org/ [↑](#footnote-ref-7)
8. Techopedia. (2019, February 5). Naive Bayes. Techopedia.Com. https://www.techopedia.com/definition/32335/naive-bayes [↑](#footnote-ref-8)
9. Pupale, R. (2019, February 11). Support Vector Machines(SVM) — An Overview - Towards Data Science. Medium. https://towardsdatascience.com/https-medium-com-pupalerushikesh-svm-f4b42800e989 [↑](#footnote-ref-9)
10. npm: @rainbow-me/animated-charts. (2021, September 20). Npm. https://www.npmjs.com/package/@rainbow-me/animated-charts [↑](#footnote-ref-10)
11. https://cs.calvin.edu/courses/cs/262/kvlinden/references/rodriguez-restfulWS.pdf [↑](#footnote-ref-11)
12. https://www.coingecko.com/en/api [↑](#footnote-ref-12)
13. https://cryptopanic.com/developers/api/ [↑](#footnote-ref-13)
14. Tweepy. (n.d.). Tweepy. Retrieved May 18, 2022, from https://www.tweepy.org/ [↑](#footnote-ref-14)
15. TextBlob: Simplified Text Processing — TextBlob 0.16.0 documentation. (n.d.). TextBlob: Simplified Text Processing — TextBlob 0.16.0 Documentation. Retrieved May 18, 2022, from https://textblob.readthedocs.io/en/dev/index.html [↑](#footnote-ref-15)
16. Google Colab. (n.d.). Google Colab. Retrieved May 18, 2022, from https://colab.research.google.com/ [↑](#footnote-ref-16)
17. TensorFlow. (n.d.). TensorFlow. Retrieved May 18, 2022, from https://www.tensorflow.org/ [↑](#footnote-ref-17)
18. JustAnotherArchivist. (n.d.). GitHub - JustAnotherArchivist/snscrape: A social networking service scraper in Python. GitHub. Retrieved May 18, 2022, from https://github.com/JustAnotherArchivist/snscrape [↑](#footnote-ref-18)
19. PyMongo 4.1.1 Documentation — PyMongo 4.1.1 documentation. (n.d.). PyMongo 4.1.1 Documentation — PyMongo 4.1.1 Documentation. Retrieved May 18, 2022, from https://pymongo.readthedocs.io/en/stable/ [↑](#footnote-ref-19)
20. RapidMiner | Amplify the Impact of Your People, Expertise & Data. (n.d.). RapidMiner. Retrieved May 18, 2022, from https://rapidminer.com/ [↑](#footnote-ref-20)
21. https://developers.google.com/machine-learning/crash-course/classification/accuracy [↑](#footnote-ref-21)
22. https://www.interaction-design.org/literature/article/shneiderman-s-eight-golden-rules-will-help-you-design-better-interfaces [↑](#footnote-ref-22)
23. Tutorial: Quickstart — TextBlob 0.16.0 documentation. (n.d.). TextBlob: Simplified Text Processing — TextBlob 0.16.0 Documentation. Retrieved May 18, 2022, from https://textblob.readthedocs.io/en/dev/quickstart.html [↑](#footnote-ref-23)
24. Jenks Natural Breaks Explained. (n.d.). Vitalnet Data Analysis Software. Retrieved May 18, 2022, from https://www.ehdp.com/methods/jenks-natural-breaks-explain.htm [↑](#footnote-ref-24)
25. dundee2002. (2018, March 26). Bitcoin Tweets Sentiment Analysis: cnn-lstm | Kaggle. Kaggle: Your Machine Learning and Data Science Community; Kaggle. https://www.kaggle.com/code/dundee2002/bitcoin-tweets-sentiment-analysis-glove-cnn-lstm [↑](#footnote-ref-25)