Investment Sentiment Analysis Capstone Project Report

Ramón.Torres

Northwest Missouri State University, Maryville MO 64468, USA
S556497@nwmissouri.edu

Abstract. The abstract should briefly summarize the contents of the paper in 150–250 words. The abstract should briefly summarize the contents of the paper in 150–250 words. The abstract should briefly summarize the contents of the paper in 150–250 words. The abstract should briefly summarize the contents of the paper in 150–250 words.

Keywords: natural language processing · financial analysis· sentiment analysis · capital preservation

1 Introduction

The planned domain is Natural language processing and visualization to demonstrate trends on investment options. The goal is to analyze sentiment analysis of experts, investors and institutional investment firms. Based on current world events and the generalization of the cryptocurrencies, particularly the bitcoin, I would like to analyze investment trends to preserve capital. I will use Kaggle data sources to obtained data sets to analyze financial trends and compare to financial reports and determine the sentiment or opinions of experts on this field . The process will include the following steps:

1. Explore for possible investment analysis scenarios 2. Search for data sources and save it to computer in CSV and MS Excel formats 3. Use python to organize and clean 4. Select appropriate data sources and columns 5. Design and run analysis model 6. Use Tableau to present visualization of relevant results 7. Present project report

Subsequent paragraphs, however, are indented.

1.1 Goals of this Research

Specify exactly your aims of this paper. Also, write a sentence how you will address for each individual goal.

1.2 Literature Review

Write something related to your paper.

2 Methodology

3 Data Collection and processing

Data source: Cryptory library for Python and https://www.kaggle.com/datasets/adilbhatti/bitcoin-and-fear-and-greed. Data sources in CSV and MS Excel formats. Data scraping techniques: Use python to organize and clean. Select appropriate data sources and columns. Data attributes: Cryptograms bitcoin: seven rows, transaction date, open market value, low value, high value, total volume, closing values, market capitalization

Other specific data extraction details related to your project: Based on current world events, clean up items for dates before year 2019 and analyze effects after the COVID 19 Pandemic.

4 Data Cleaning

Using Python to upload the appropriate data sets. Review and check for null values, values before 2019 and possible duplicates. Made backups of original file and working file. Tools and techniques are used in the process: Pandas, Numpy Checking for missing values or deleting roes with missing results to avoid false or misrepresentation of results by improving the quality of the data. Seven columns and 2205 records. The definitions of important data attributes. date = date of transaction

Open = value of bitcoin when market open

Cloe = value of bitcoin when market close

High = highest value of bitcoin on a trading day

Low = lowest value of bitcoin on a trading day

Volume = total shares traded during a trading day

MarketCap = total value in dollars per trading day.

For analysis and forecasting: the independent variables volume and marke-CAp by month. The dependent variable will be next day open.

5 Exploratory Data Analysis

Exploratory data analysis is a stage where data is organized and process to determine the main attributes and characteristics of the data set. Generally, data is organized using a visualization methods, tools, or techniques, like graphs, charts, and tables. It is important in any given data analytics projects to identify trends, outliers, tendencies and prepare the scenario for in depth analysis and discussion of findings. It helps us avoid making any assumptions that might overstated, understated or plain wrong. This will avoid adding additional errors to the forecast or predictive models, using Python or R for example. In research methodology it set the stage for an appropriate statistical summary to perform hypothesis testing.

Some of the exploratory data analysis techniques are univariable non graphical and graphical analysis, multivariable nongraphical analysis and graphical analysis. My project will include historical analysis of market values of the cryptocurrencies to consider the trends, frequencies, moving average and seasonality analysis. It may include time series analysis to determine variability rates and univariable correlation analysis among market values and prices. These tools are still under consideration.

The techniques included analysis of market values of the cryptocurrencies to consider the trends, frequencies, moving average and seasonality analysis. The process is as follows:

Phase 1: Historical cryptocurrency market value analysis for the period of 2015 to 2019. This technique will provide market value trends.

Phase 2: Historical cryptocurrency trading volume analysis for the period of 2015 to 2019. This technique will provide volume trends.

Phase 3: Moving averages or mean values of cryptocurrencies main attributes, price and volume for the period of 2015 to 2019. This technique will facilitate the visualization of upward and downward trends.

Phase 4: Seasonality analysis on cryptocurrency market value for the period of 2015 to 2019. This technique will decompose the data by trends, seasonality, and noise. Therefore, providing the levels of variability on market value and volume.

Intial findings includes that the prices of cryptocurrencies showed a tendency of having market value under pressure for months of September and October of every year. This showed a downward trend on the values of the digital asset. Then after these periods of overselling, upward trends appear to present until the end of each year. By December of every year, a buying rally reached a peak, as well as the value of the asset.

6 Discussion and findings

7 Conclusions

8 Ethical considerations and limitations

9 Recommendations

References

1. Sheynin, N.: 3 ways to apply natural language processing (nlp) in financial research. AlphaSene (2022)
2. Vimalachandran, A., Jayachandran, T.: The financial crash of 2020 and the retail trader's boon: a correlation between sentiment and technical analysis. SN Business Economics