

Cryptocurrency Prediction With Social Media

Darryl Patrick Matheuw Kurniawan^{a1}, Drs. I Wayan Santiyasa^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Bali, Indonesia

¹Darrylpatrck82@gmail.com

²Santiyasa@unud.ac.id

Abstract

Social Media adalah suatu wadah bagi kita manusia dalam berkomunikasi, tidak jarang juga kita banyak membagikan kehidupan kita didalam social media. Baik berupa hiburan, bisnis atau hal lainnya.maka dari itu penelitian ini akan menggunakan social media konten untuk memprediksi masa depan , contohnya adalah memprediksi harga dari cryptocurrency. Dengan menggunakan twitter sebagai media social, jurnal ini akan menunjukkan suatu model yang dapat kita gunakan untuk memprediksi harga cryptocurrency.

Keywords: Coin, Predict, Social Media

1. Introduction

Seiring dengan berkembangnya ekonomi dan teknologi, cryptocurrency atau coin adalah alternatif lain mata uang dimana coin ini dijadikan sebagai uang digital yang dapat diperjualbelikan dan digunakan [1]. coin digunakan untuk mendentralisasi nilai mata uang yang beredar di dunia agar lebih mudah dilakukan, coin pertama yang dikeluarkan adalah bitcoin dimana bitcoin ini dikeluarkan pada tahun 2009 oleh Satoshi Nakamoto sebagai software open-source[2]. kemudian coin ini menarik banyak perhatian orang-orang dan mengakibatkan munculnya coin alternatif sekitar 10.000 lebih coin hingga Desember 2021 terakhir[3]. Saat ini coin – coin yang menjadi coin utama adalah bitcoin (BTC), Ethereum(ETH), Litecoin(LTC) dan Cardano (ADA). Yang berhasil bertumbuh secara eksponensial hingga mencapai capitalization market lebih dari 2 triliun dollar dimana bitcoin memegang 40% total , ethereum 21%, Litecoin sekitar 0,5% dan cardano hampir 2%[4].

Dengan memperhatikan aktifitas – aktifitas yang dilakukan oleh masyarakat di media social penelitian ini akan memprediksi harga coin dengan mengandalkan aktifitas masyarakat. Media yang digunakan nantinya adalah twitter dimana twitter adalah salah satu platform yang sangat besar dengan pengguna sekitar 400 juta orang lebih dan 200 juta pengguna aktif setiap harinya[5]. Twitter digunakan karena jika dibandingkan dengan platform media social lainnya yang lebih banyak menggunakan gambar dan visual lainnya mengakibatkan sulitnya dilakukan analisis dan fakta bahwa twitter memiliki textual context terbesar sehingga kita dapat menganalisis lebih mudah dan akurat[6].

Metode yang akan digunakan nantinya adalah menggunakan tweet volumes dan sentiment analysis, sentiment analysis adalah opini mining atau emosi ai dimana sentiment analysis ini menggunakan Natural Language Processing (NLP) untuk mengesktrak data twitter yang berupa text nantinya[7]. Sementara tweet volumes adalah volume dari tweet yang ada pada saat waktu – waktu yang ditentukan[8].

2. Reseach Methods

2.1 Pengumpulan Data

Pengumpulan data dilakukan dengan menggunakan twitter API dan google trends untuk mendapatkan data. Antara lain Tweepy adalah suatu python library yang digunakan untuk mengakses twitter API dan mengumpulkan data. Library ini dapat membantu kita memfilter data berdasarkan hastags atau text yang ada pada twitter. Dalam penelitian ini menggunakan “#bitcoin” dan “#ethereum”. Selain itu google juga memberikan akses untuk mengakses data melalui google trends. Google adalah salah satu mesin pencari yang paling populer didunia dimana google digunakan lebih dari 74% dari setiap pencarian yang dilakukan pada web, jadi tidak diragukan lagi bahwa google adalah yang terbaik dalam memberikan informasi. Data yang diberikan oleh google adalah Search Volume Index (SVI) yang menunjukkan seberapa sering istilah penelusuran tertentu ditelusuri relatif terhadap total volume penelusuran di seluruh dunia, selama rentang tanggal tertentu yang dimasukkan pengguna.

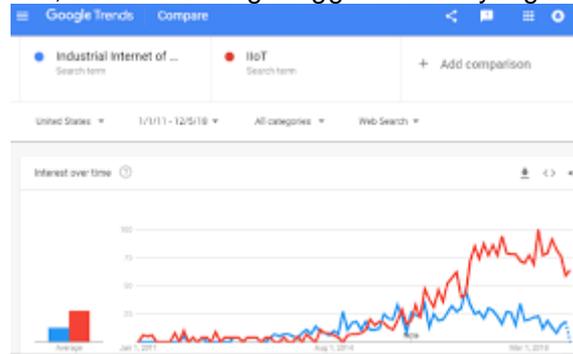


Figure 1. contoh penggunaan google trends

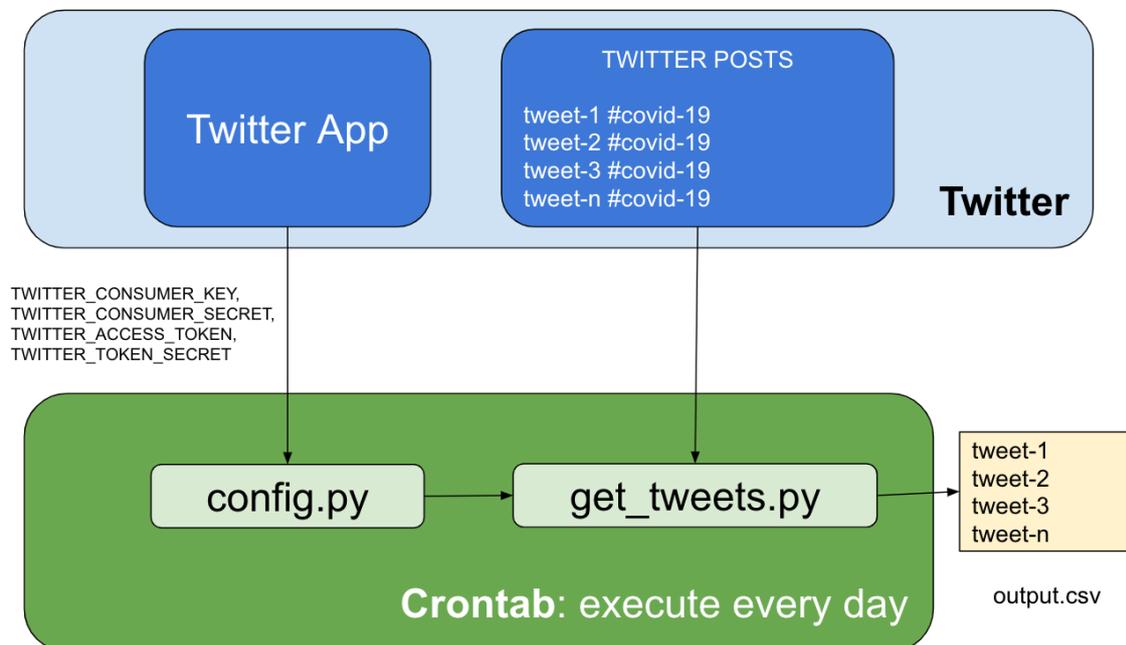


Figure 2. cara kerja tweepy

2.2 Cleaning data

Tahap setelah pengumpulan data adalah pembersihan data dimana pada proses ini data dibersihkan dengan menghilangkan hastags ,petik dan tanda – tanda lainnya sehingga data hanya tersisa text utuh agar lebih mudah diproses nantinya.

2.3 Sentiment Analysis

Pada tahap ini menggunakan VADER (Valence Aware Dictionary for Sentiment Reasoning) adalah suatu python library untuk memproses NLP dan subjektivitas dan polaritas.

1st statement :

Overall sentiment dictionary is : {'neg': 0.165, 'neu': 0.588, 'pos': 0.247, 'compound': 0.5267}
sentence was rated as 16.5 % Negative
sentence was rated as 58.8 % Neutral
sentence was rated as 24.7 % Positive
Sentence Overall Rated As Positive

2nd Statement :

Overall sentiment dictionary is : {'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}
sentence was rated as 0.0 % Negative
sentence was rated as 100.0 % Neutral
sentence was rated as 0.0 % Positive
Sentence Overall Rated As Neutral

3rd Statement :

Overall sentiment dictionary is : {'neg': 0.508, 'neu': 0.492, 'pos': 0.0, 'compound': -0.4767}
sentence was rated as 50.8 % Negative
sentence was rated as 49.2 % Neutral
sentence was rated as 0.0 % Positive
Sentence Overall Rated As Negative

Figure 3. contoh penggunaan sentiment analysis

3. Result and Discussion

Berikut adalah hasil dari Vader figure 1 bitcoin dan figure 2 ethereum

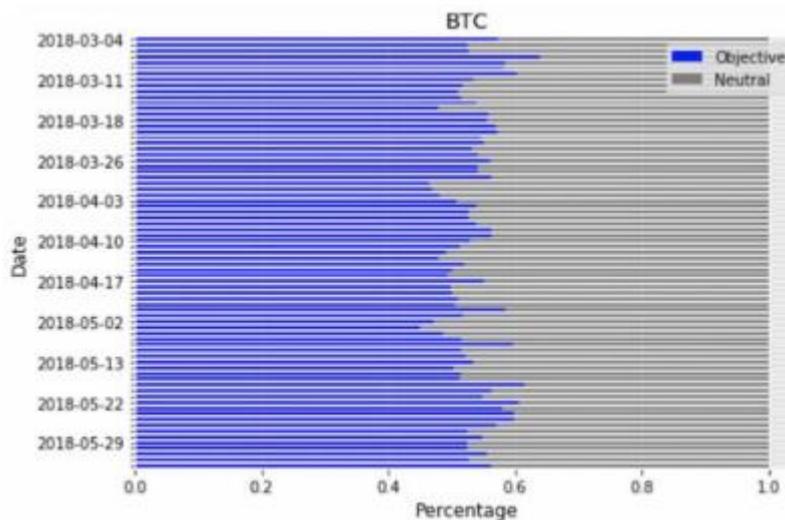


Figure 1. persentase objektif bitcoin (memiliki nilai positif atau negatif) dengan neutral

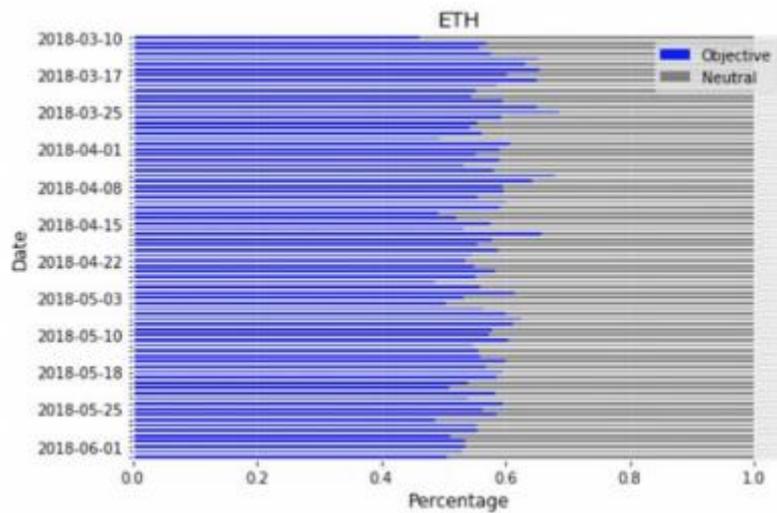


Figure 2. persentase objektif ethereum (memiliki nilai positif atau negatif) dengan neutral

Figure 3 dan 4 menunjukkan sentiment tweet lebih neutral daripada objektif

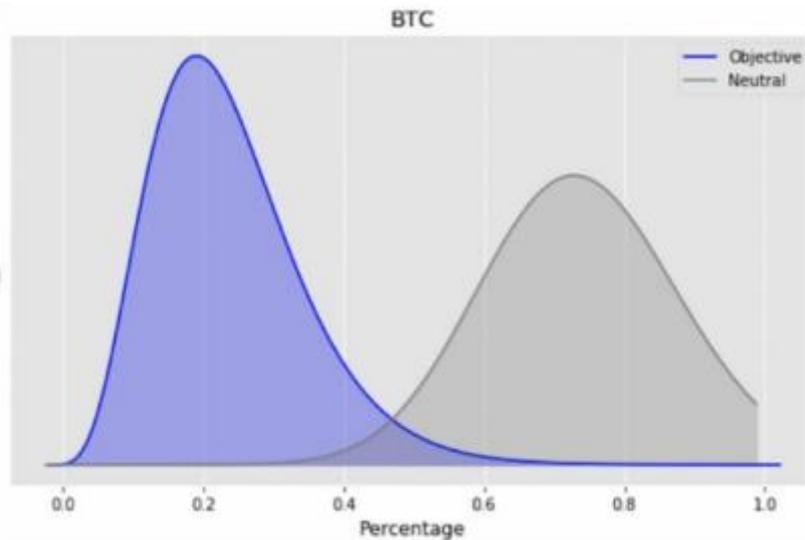


Figure 3. distribusi objektif bitcoin

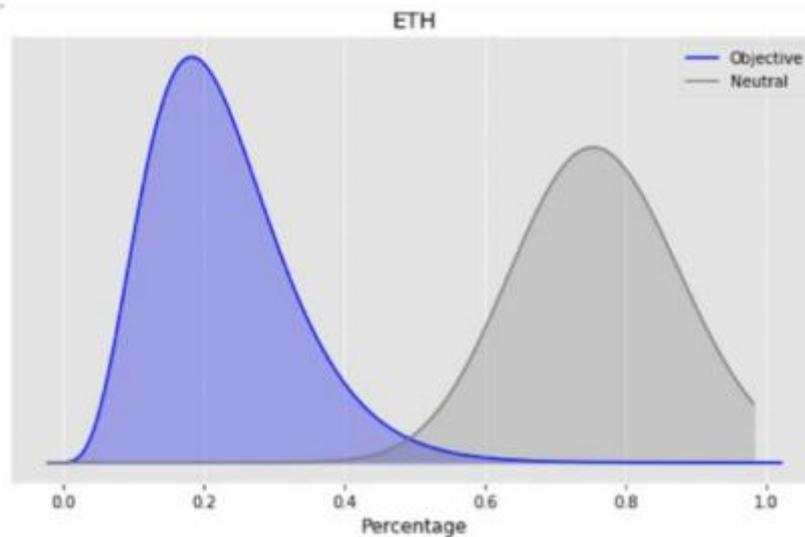


Figure 4. distribusi objektif Ethereum

Fakta bahwa setengah dari tweets memiliki sentiment netral, hal ini memungkinkan sentiment positif atau negative dapat memberikan insight bahwa model ini berhubungan perubahan harga. Hal tersebut dapat kita lihat pada fluktuasi harga, berikut figure 5 dan 6 adalah perubahan harga dan polarity pada tweet.

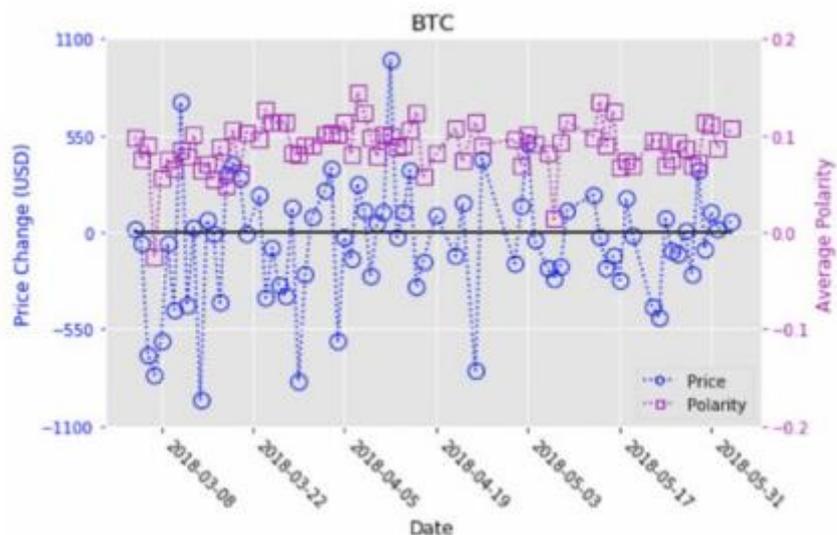


Figure 5. perubahan harga bitcoin dan polarity tweet

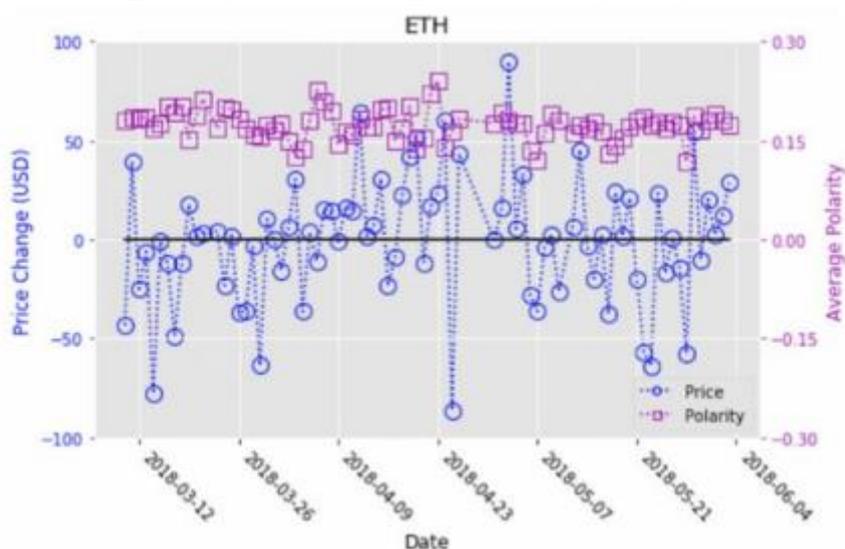


Figure 6. perubahan harga Ethereum dan polarity tweet

Dapat kita lihat polarity dan harganya cukup berbeda walaupun berhubungan tetapi Ketika harga mengalami penurunan polaritinya tetapi bagus, itu berarti orang-orang yang ada di twitter tidak membicarakan positif atau negative membelinya tetapi mereka suka dengan hal – hal lainnya seperti keamanan atau lain sebagainya. karena perbedaan yang drastic berarti sentiment analysis tidak dapat kita gunakan dan tweet volume dapat menjadi indicator yang lebih baik dibandingkan sentiment.

Karena sentiment sudah dikeluarkan maka tweet volume dan google trends menjadi indicator yang akan digunakan berikut SVI dan tweet volume sebagai indicator yang memiliki korelasi dengan naik turunnya harga crypto. Dengan menggunakan multiple linear regression berikut hasil prediksinya pada figure 7 dan 8.

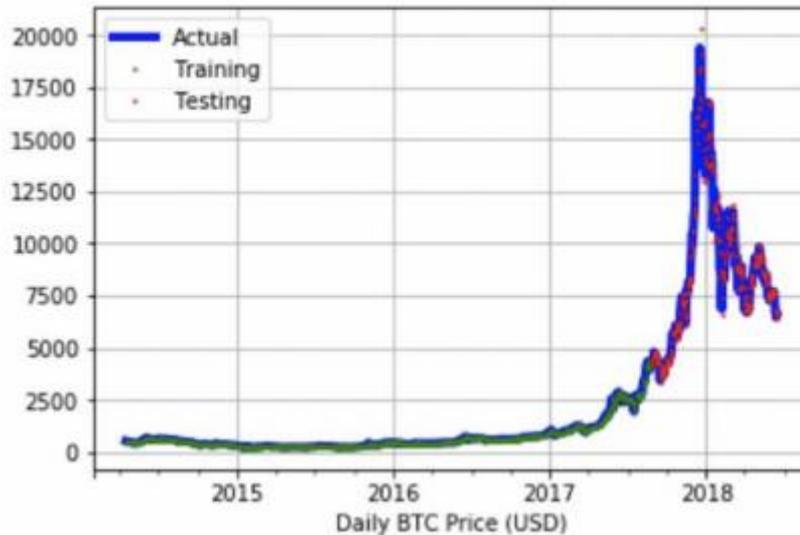


Figure 7. model menunjukkan kecocokan dengan harga asli bitcoin dengan warna hijau sebagai train data dan hasil tes adalah titik merah

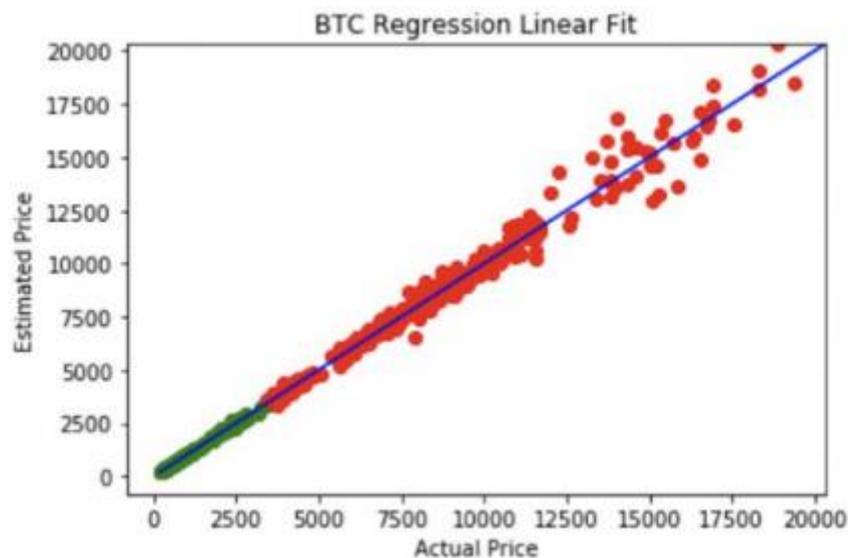


Figure 8. bitcoin regression menunjukkan kecocokan dengan perkiraan harga y-axis, dan x-axis adalah harga asli. titik hijau adalah train data dan titik merah adalah hasil

Figure 7 merupakan model machine learning yang digunakan yakni multiple linear regression dimana data dari Google Trends dan tweet volume dibagi dengan 80% dari total data adalah data training dan 20% adalah data test dan dapat kita lihat dari hasilnya model ini memiliki akurasi yang baik dimana data training dan test cocok dengan harga asli yang ada. Sementara figure 8 adalah tampilan dengan linear regression dimana graphic menunjukkan bahwa garis positif dan hubungan antara harga asli dan perkiraan harga mayoritas cocok, dimana titik hijau adalah data train dan titik merah adalah hasil test.

Hal ini dapat terjadi karena hubungan antara Google Trends data dengan perubahan harga crypto cukup baik. Dengan membandingkan berdasarkan R-pearson dan P-value metrik. R-pearsonnya adalah 0.817 yang berarti hubungan positif dengan Google Trend data, dan P-valuenya adalah 0.000, yang berarti hasilnya signifikan secara statistic.

4. Conclusion

Hasil prediksi pada model menunjukkan hasil yang baik hanya saja terdapat beberapa indicator yang tidak dapat digunakan dalam memprediksi harga dengan social media seperti sentiment analysis karena Ketika sentiment menunjukkan positif harga tidak lah turun mengakibatkan ketidak cocokan sentiment analysis sebagai indicator serta google trends dan volume dari tweet dapat menjadi indicator karena aktifitas yang masyarakat lakukan mempengaruhi pasar yang ada hal ini membuktikan bahwa kita dapat menggunakan media social untuk memprediksi masa depan.

References

- [1] Connor Lamon, Eric Nielsen ,Eric Redondo, Cryptocurrency Price Prediction Using News and Social Media Sentiment, 2017.
- [2] Michail Vlachos and Giovnopoulos, Forecasting Cryptocurrency Price Movements with Predictive Social Media Analytics,2022
- [3] Oikonomopoulos Sotirios, Cryptocurrency price prediction using social media sentiment analysis,2022.
- [4] Wolfgang Karl Härdle, Campbell R. Harvey and Raphael C. G. Reule, Understanding Cryptocurrencies,2018.
- [5] Sitaram Asur and Bernardo A. Huberman ,Predicting the Future With Social Media,2010

This page is intentionally left blank.