

BAB II

LANDASAN TEORI

2.1 *Cryptocurrency*

Mata uang kripto atau *cryptocurrency* adalah aset digital yang memiliki suatu fungsi untuk digunakan sebagai media pertukaran dimana catatan kepemilikan koin individu disimpan dalam buku besar dan neraca keuangan yang ada dalam database yang terdiri dari serangkaian mekanisme kriptografi [2]. Mata uang kripto merupakan sistem keuangan elektronik *peer-to-peer* (P2P) yang memungkinkan penggunanya melakukan pembayaran secara online untuk langsung dikirim dari satu pihak ke pihak lain tanpa melalui lembaga keuangan seperti bank. Selain dapat digunakan sebagai alat pembayaran, mata uang kripto juga dapat digunakan sebagai alat investasi dan juga alat dalam melakukan perdagangan aset digital [11]. Hingga saat ini terdapat lebih dari 10 ribu mata uang kripto yang tersebar di dunia [4]. Berikut adalah mata uang kripto 10 besar di dunia dengan kapitalisasi pasar terbesar.

2.1.1 *Bitcoin*

Setelah krisis ekonomi yang terjadi pada tahun 2008, seorang tokoh yang disebut dengan nama Satoshi Nakamoto menciptakan sistem elektronik P2P yang disebut dengan *Bitcoin* [12]. *Bitcoin* adalah aset digital yang terdesentralisasi dan dikenalkan kepada dunia pada tahun 2008 dan diluncurkan pada awal tahun 2009. Menurut para ahli, terciptanya *Bitcoin* adalah dibutuhkan suatu sistem keuangan dimana seseorang dapat melakukan suatu transaksi yang aman dan cepat tanpa adanya pihak ketiga seperti bank, suatu lembaga institusi yang memonopoli uang rakyat [7].

Sejak diluncurkannya *Bitcoin*, terdapat lebih dari 11.000 mata uang kripto yang tersebar di dunia. Mata uang kripto kini dapat digunakan untuk membeli barang dan jasa contohnya seperti negara El Salvador yang telah mengadopsi bitcoin sebagai alat

tukar yang sah untuk membeli barang dan jasa [13]. *Bitcoin* dipercaya sebagai alternatif pada sistem keuangan dunia yang nilainya semakin menurun yang diakibatkan oleh inflasi, *Bitcoin* juga dapat digunakan sebagai alat untuk investasi yang akan berguna dimasa depan [14].

2.1.2 Ethereum

Ethereum adalah suatu blockchain yang dapat diprogram dengan koin aslinya sendiri yang disebut dengan *Ether* [15]. *Ethereum* adalah mata uang kripto yang terdesentralisasi dengan kode lambang *ETH*. *Ether* tidak dikendalikan oleh suatu lembaga pemerintah maupun suatu organisasi dan dapat digunakan untuk investasi dan melakukan suatu pembayaran. Ethereum adalah mata uang kripto terbesar kedua setelah *Bitcoin*. *Blockchain Ethereum* adalah platform komputasi terdistribusi *open source* yang menyoroti kegunaan *Smart Contract (scripting)*. Seseorang dapat dengan mudah menulis aplikasi terdesentralisasi pada tingkat yang signifikan dan keuntungan dari distribusi yang diperoleh dari teknologi *Blockchain* [16].

2.1.3 Binance Coin

Binance Coin atau BNB adalah suatu mata uang kripto yang diciptakan dan dimiliki oleh Binance yang merupakan pasar jual beli mata uang kripto terbesar didunia [17]. Binance Coin dapat digunakan untuk investasi dan dijadikan suatu alat pembayaran dan membeli mata uang kripto khusus pada platform Binance. Binance Coin pertama kali diluncurkan kepada masyarakat sebanyak 100 juta coin dan setiap 4 bulan akan dilakukan Coin Burn yaitu menghapus, membakar, dan menghancurkan sebanyak 20% koinnya setiap 4 bulannya [18]. Hal ini dilakukan untuk menjaga harga Binance Coin dan membuat Binance Coin semakin langka sehingga nilai dari mata uang kripto tersebut dapat terus naik.

2.1.4 Dogecoin

Dogecoin atau yang lebih dikenal dengan *DOGE* adalah mata uang kripto yang sebenarnya diluncurkan untuk sebuah lelucon yang bereferensi pada sebuah meme [19]. Namun, disaat Elon Musk selaku CEO Tesla menyebutkan nama *DOGE* pada akun *Twitter*nya, harga *DOGE* ini langsung naik hingga ratusan persen.



Gambar 2.1 Cuitan Elon Musk mengenai mata uang kripto *DOGE* [20]

Mata uang kripto DOGE yang awalnya dibuat hanya untuk sebuah lelucon kini menjadi mata uang kripto dengan kapitalisasi pasar terbesar nomor 10 didunia berkat Elon Musk. DOGE coin ini sangat cocok untuk berinvestasi dikarenakan telah bekerja sama dengan Elon Musk dalam mengerjakan proyeknya yang berbasis Blockchain .

2.1.5 Ripple

Ripple adalah mata uang kripto yang menggunakan jaringan peer-to-peer terdesentralisasi yang menyediakan protokol pembayaran digital ke lembaga keuangan [21]. *Ripple* memungkinkan transfer uang digital tanpa batas, baik jika akan mengirimkan uang konvensional dalam dolar atau mata uang kripto seperti Bitcoin. Jaringan memiliki tokennya yang dilambangkan dengan *XRP*. Berdasarkan modal pasar, *XRP* saat ini merupakan mata uang kripto terbesar keenam [22].

2.2 Blockchain

Blockchain diartikan sebagai buku besar yang terdigitalisasi, terdesentralisasi dan tersebar yang dimana seluruh transaksi tercatat dan disimpan secara berurutan untuk membuat suatu daftar transaksi yang bersifat permanen dan tidak dapat diubah oleh siapapun. Teknologi *Blockchain* adalah teknologi yang dikelola oleh sebuah jaringan *peer to peer* dan diamankan oleh teknik kriptografi yang memungkinkan adanya transaksi yang bersifat transparan [23].

Teknologi *blockchain* telah mencuri banyak perhatian dunia, khususnya bank dan perusahaan yang khususnya beridri dalam bidang keuangan atau biasa disebut dengan *FinTech* untuk mengadopsi teknologi *blockchain* dalam menyalurkan jasa dalam bidang keuangan [24].

2.3 Twitter

Twitter adalah suatu aplikasi sosial media yang secara resmi diluncurkan kepada masyarakat pada tanggal 13 Juli 2006. Fungsi utama dari *Twitter* adalah membuat sebuah postingan atau cuitan yang pendek mengenai sebuah berita, ungkapan

perasaan dan suatu opini melalui website atau aplikasi pada mobile atau smartphone. *Twitter* memiliki beberapa poin yang unik yang membedakannya dari aplikasi sosial media lainnya, yaitu [8] :

1. *Twitter* memiliki batasan panjang karakter yang dituliskan yaitu 140 karakter.
2. Data *Twitter* dapat diakses dengan bebas oleh siapapun melalui *Twitter API* yang mempermudah untuk mengumpulkan data dari ribuan hingga jutaan cuitan yang ada di *Twitter*.
3. Pengguna *Twitter* dapat secara bebas mengirimkan pesan, topik atau berita dan yang berlaku ke seluruh pengguna *Twitter* didunia.

2.4 Sentiment Analysis

Sentimen analisis adalah suatu proses yang digunakan untuk mendeteksi dan mengklasifikasikan pendapat dan opini oleh seseorang dalam bentuk kalimat dan kata terhadap suatu hal yang sedang dan telah terjadi. Sentimen analisis dibagi menjadi klasifikasi 2 yaitu sentimen negatif dan positif. Kedua sentimen tersebut digunakan untuk memprediksi *sentiment polarity* berdasarkan data data yang telah didapatkan dari user atau pengguna sosial media. *Textual sentiment analysis* banyak digunakan dalam area ilmiah tetapi juga sering digunakan dalam area business dan marketing untuk mengetahui opini dan peluang yang akan muncul dari masyarakat melalui sosial media [8].

2.5 Rapid Miner

Rapid Miner adalah suatu aplikasi yang menggunakan bahasa pemrograman *Java* untuk menjalankan fungsinya. *Rapid Miner* adalah sebuah *software* atau perangkat lunak yang digunakan untuk mengolah suatu *data*, *text mining*, *data mining* dan *predictive analytics* [25]. *Rapid Miner* memberikan kemudahan pada penggunaanya untuk menghitung data dalam jumlah yang sangat besar dengan menggunakan operator

yang berfungsi untuk mengubah suatu data tertentu. *Rapid miner* memiliki sekitar 500 operator yang memiliki fungsi utama yang digunakan untuk menganalisa suatu data seperti permutan dan transformasi data, data visualisasi, pemodelan data dan evaluasi model data.

2.6 *Support Vector Machine*

Support Vector Machine adalah suatu algoritma yang biasa dan sangat sering digunakan untuk melakukan klasifikasi dan regresi. Algoritma *Support Vector Machine* memiliki suatu prinsip dasar dalam klasifikasi linier, yaitu kasus klasifikasi dapat dipisahkan secara linier dan nonlinier dengan cara menambahkan konsep kernel [26]. *Hyperplane* digunakan untuk membuat maksimal margin dan jarak antarkelas data yang berbeda. Data kemudian akan dijadikan 2 klasifikasi yaitu 1 untuk positif dan -1 untuk negatif.

$$Y_i (w \cdot X_i + b) \geq 1 - \xi_i, \xi_i \geq 0$$

Rumus 2.1 *Support Vector Machine*

Keterangan:

X_i = data ke -i

W = nilai bobot vektor

B = nilai bias

Y_i = target data ke -i

ξ_i = data variable slack untuk ke -i

2.7 *Naïve Bayes*

Naive Bayes adalah suatu algoritma yang biasa digunakan untuk melakukan suatu klasifikasi secara statistik yang dapat memprediksi suatu probabilitas dari kelas tertentu [27].

Rumus pada algoritma *Naive Bayes* adalah:

$$P(C_i|X) = \frac{p(X|C_i)p(C_i)}{P(X)}$$

Rumus 2.2 Rumus Naïve Bayes

Keterangan:

X = Data dengan kelas yang belum diketahui

C_i = Hipotesis data X Merupakan suatu class spesifik

$p(C_i|X)$ = probabilitas hipotesis berdasarkan kondisi (posteriori probability)

$p(C_i)$ = probabilitas hipotesis (prior probability)

$p(X|C_i)$ = probabilitas berdasarkan kondisi pada hipotesis

$p(X)$ = probabilitas dari C_i

2.8 CRISP - DM

Cross - Industry Standard process atau *CRISP - DM* adalah standarisasi yang biasa digunakan dalam melakukan suatu proses data mining yang digunakan untuk memecahkan suatu masalah secara umum yang dimulai dari suatu bisnis tertentu sampai suatu penelitian mengenai suatu masalah. *CRISP - DM* memiliki 6 tahapan kerja yaitu [28]:

1. Business Understanding phase

Tahapan pertama atau *Business Understanding Phase* adalah fase yang digunakan untuk memahami inti dari suatu kegiatan melalui sudut pandang dalam melakukan suatu bisnis. Dalam fase ini yang akan dilakukan adalah menentukan tujuan, strategi dan kebutuhan secara keseluruhan dalam mencapai tujuan yang akan dituju [29].

2. *Data Understanding phase*

Tahapan kedua atau *Data Understanding Phase* adalah fase untuk mengumpulkan berbagai data yang dibutuhkan, mempelajari data untuk dipahami yang nantiinya akan dilakukan kepada suatu penelitian [29]

3. *Data preparation phase*

Fase ketiga atau *Data Preparation Phase* adalah fase yang digunakan untuk mempersiapkan data dan memilih variable yang akan digunakan pada penelitian yang akan digunakan [29].

4. *Modeling phase*

Fase keempat atau *Modeling Phase* adalah fase yang digunakan memilih dan menentukan teknik dari model berdasarkan tujuan yang sebelumnya sudah ditentukan [29] .

5. *Evaluation phase*

Fase Kelima atau *Evaluation Phase* adalah fase yang digunakan untuk melakukan evaluasi terhadap model yang akan digunakan pada penelitian untuk memastikan bahwa model yang dipilih akan memiliki tingkat kualitas dan efesiensi yang di harapkan [29].

6. *Deployment phase*

Fase keenam atau *Deployment Phase* adalah fase yang digunakan untuk memberikan suatu gambaran atau visualisasi terhadap laporan penelitian yang akan dibuat [29].

2.9 *Text Mining*

Text mining adalah suatu proses yang digunakan untuk menambang suatu data yang berbetuk *text* pada suatu dokumen atau media. Tujuan dari *text mining* adalah

mengambil suatu informasi yang masih berantakan dan tidak terstruktur dalam bentuk text yang nantinya akan digunakan pada suatu penelitian yang bersangkutan dengan text mining itu sendiri [30].

Dalam proses text mining memiliki beberapa proses yang disebut dengan tahap *pre-processing*. *Pre-processing* juga memiliki 2 sub proses yaitu:

1) *Data Cleansing*

Data Cleansing adalah proses yang digunakan untuk membersihkan suatu kata atau kalimat pada suatu data yang berbentuk *text*. *Data Cleansing* dapat digunakan dengan cara mengganti, memodifikasi, dan menghapus suatu kata namun tetap tidak merubah inti dari kalimat tersebut [29].

2) *Data Labeling*

Data Labeling adalah suatu proses yang digunakan untuk menklasifikasikan suatu data dan mengelompokkan beberapa data kedalam satu grup yang dilabel [29].

2.10 *TF – IDF*

TF – IDF adalah kepanjangan dari *Term Frequency Inverse Document Frequency*. *TF – IDF* adalah suatu metode pembobotan yang sangat umum digunakan untuk penelitian pembobotan kata. Pembobotan kata adalah suatu proses yang memberikan bobot kepada setiap kata didalam suatu dokumen. *TF- IDF* digunakan untuk menghitung seberapa sering suatu kata muncul dalam suatu dokumen [31].

2.11 *Confusion Matrix*

Confusion Matrix adalah suatu metode yang umumnya digunakan dalam perhitungan akurasi dalam data mining dan analisa sentimen. *Confusion Matrix* memiliki 3 macam parameter yaitu *accuracy*, *precision*, dan *recall* dimana semakin tinggi nilai dari ketiga parameter tersebut maka akan semakin baik [32]. Fungsi dari ketiga parameter tersebut adalah:

1) *Accuracy*

Accuracy adalah jumlah dan persentase keakuratan dari jenis model yang digunakan dalam penelitian. Rumus dari *Accuracy* adalah

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}$$

Rumus 2.3 Accuracy Confusion Matrix

2) *Precision*

Precision adalah tingkat keakuratan dari tiap dokumen *text* yang dianggap relevan dari seluruh dokumen *text* yang dipilih. Rumus dari *Precision* adalah

$$\text{Precision} = \frac{TP}{TP+FP}$$

Rumus 2.4 Precision Confusion Matrix

3) *Recall*

Recall adalah keakuratan dari tiap dokumen *text* terhadap seluruh dokumen *text* yang telah didapatkan. Rumus dari *Recall* adalah

$$\text{Recall} = \frac{TP}{TP+FN}$$

Rumus 2.5 Recall Confusion Matrix

2.12 *Vader*

Vader adalah kepanjangan dari *Valence Aware Dictionary for Social Reasoning* yang digunakan sebagai model dalam menganalisa suatu sentimen dan dapat menentukan perbedaan data melalui tingkat intensitas perasaan dari suatu data. *Vader* terdiri dari 4 bagian yaitu *pos*, *neg*, *neu*, dan *compound* [33].

Pos, *neg*, dan *neu* adalah nilai rasio dari *text* yang dimana *pos* adalah positif, *neg* adalah negatif, dan *neu* adalah natural. Sedangkan nilai *compound* adalah skor valensi yang menilai suatu kalimat dan kata dengan mempertimbangkan pengukuran kata, bentuk kata, tanda baca, sensitivitas konjungsi. Nilai *compound* ini digunakan sebagai nilai satuan standar yang dimana jika suatu kalimat bersifat positif maka score

akan lebih besar dari 0, jika kalimat bersifat negatif maka akan diberi score lebih kecil dari 0, dan jika kalimat tersebut bersifat netral maka akan diberi score sama dengan 0.

2.13 Penelitian Terdahulu

Berikut adalah tabel penelitian terdahulu mengenai penelitian yang serupa yang dapat dijadikan sebagai referensi.

Tabel 2.1 Tabel Penelitian Terdahulu

No	Jurnal	Keterangan
1.	Penulis	Adam Prasetya, Ferdiansyah, Yesi Novaria Kunang, Edi Surya Negara, Winoto Chandra
	Nama Jurnal	<i>Journal of Information Systems and Informatics Vol. 3, No. 2, June 2021</i>
	Judul Jurnal	Sentimen Analisis Terhadap <i>Cryptocurrency</i> Berdasarkan <i>Comment</i> Dan <i>Reply</i> Pada Platform <i>Twitter</i>
	Metode	<i>Naïve Bayes</i> dan <i>Support Vector Machine</i>
	Hasil dan Kesimpulan	• Hasil akurasi SVM adalah <i>Bitcoin</i> 71.30% <i>Ethereum</i> 75.05 % dan <i>Ripple</i> 71.66 %. • Hasil akurasi <i>Naïve Bayes</i> adalah <i>Bitcoin</i> 65.65 % <i>Ethereum</i> 72.86 % dan <i>Ripple</i> 65.58 %.
	Adopsi	Penerapan algoritma SVM dan <i>Naïve Bayes</i>
2.	Penulis	Rein Rachman Putra, Monika Evelin Johan, Emil Robert Kaburuan
	Nama Jurnal	<i>International Journal of Advanced Trends in Computer Science and Engineering</i> , 8(5), September - October 2019, 1856 - 1860
	Judul Jurnal	A <i>Naïve Bayes</i> Sentiment Analysis for Fintech Mobile Application User Review in Indonesia

No	Jurnal	Keterangan
	Metode	<i>Naïve Bayes</i>
	Hasil dan Kesimpulan	Algoritma Naïve Bayes mampu menghasilkan akurasi yang memuaskan yakni 78% dan algoritma <i>Confusion Matrix</i> pada perhitungan <i>accuracy</i> , <i>precision</i> , dan <i>recall</i> dapat digunakan pada algoritma <i>Naïve Bayes</i> .
	Adopsi	Metode dalam penggunaan algoritma Naïve bayes dan penggunaan <i>Confusion Matrix</i>
3.	Penulis	Rizky Parlika , Sunu Ilham Pradika , Amir Muhammad Hakim , Kholilul Rachman N.M
	Nama Jurnal	JIFTI - Jurnal Ilmiah Teknologi Informasi dan Robotika Volume 2 Nomor 2 Bulan Desember 2020
	Judul Jurnal	Analisis Sentimen <i>Twitter</i> Terhadap <i>Bitcoin</i> dan <i>Cryptocurrency</i> Berbasis <i>Python TextBlob</i>
	Metode	<i>Python TextBlob</i>
	Hasil dan Kesimpulan	Seluruh data <i>tweets</i> yang berjumlah 3433 <i>data</i> telah diklasifikasikan menggunakan <i>library python textblob</i> dan didapatkan hasil dimana 41,3% <i>tweets</i> adalah positif, 44,9% adalah netral, dan 13,7% adalah negatif.
	Adopsi	Metode dalam menggunakan <i>API Twitter</i> dalam melakukan request untuk <i>text mining</i>
4.	Penulis	Raihan Astrada Fathurrahman
	Nama Jurnal	Makalah IF2211 Strategi Algoritma, Semester II Tahun 2020/2021
	Judul Jurnal	Analisis Sentimen Pasar Saham dari Berita Keuangan menggunakan Algoritma Pencocokan String
	Metode	<i>Vader</i>

No	Jurnal	Keterangan
	Hasil dan Kesimpulan	Algoritma <i>VADER</i> dapat digunakan untuk klasifikasi sentimen positif, negatif, dan netral dengan memberikan nilai mendekati 1 pada sentimen yang bersifat positif, nilai mendekati -1 pada sentimen negatif, dan 0 pada sentimen netral
	Adopsi	Metode dalam menggunakan algoritma <i>Vader</i> dalam melakukan data labeling
5.	Penulis	Christevan Destitus, Wella, Suryasari
	Nama Jurnal	ULTIMA InfoSys, Vol. XI, No. 2 Desember 2020
	Judul Jurnal	Support Vector Machine VS Information Gain: Analisis Sentimen Cyberbullying di Twitter Indonesia
	Metode	<i>SVM, Information Gain</i>
	Hasil dan Kesimpulan	<i>Support Vector Machine</i> memberikan <i>accuracy</i> 80%, <i>precision</i> 75,1%, dan <i>recall</i> 96%
	Adopsi	Metode dalam menggunakan algoritma <i>SVM</i> dan Confusion Matrix

Berdasarkan tabel 2.1 penelitian terdahulu diatas, terdapat 5 jurnal yang telah digunakan sebagai penelitian terdahulu dengan menggunakan berbagai macam cara yang berbeda dalam melakukan penelitian khususnya dalam hal sentimen analisis. Kelima jurnal yang sudah ditulis dan dimasukkan pada tabel 2.1 akan digunakan sebagai referensi dan petunjuk dalam melakukan penelitian ini. Berikut adalah cara dan algoritma yang digunakan kelima jurnal diatas dalam melakukan sentiment analisis:

1. Jurnal pertama ditulis oleh Adam Prasetya, Ferdiansyah, Yesi Novaria Kunang,Edi Surya Negara, dan Winoto Chandra menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* dalam melakukan sentimen analisis

2. Jurnal kedua ditulis oleh in Rachman Putra, Monika Evelin Johan, Emil Robert Kaburuan menggunakan algoritma *Naïve Bayes* dalam melakukan sentimen analisis dan *Confusion Matrix* dalam melakukan perhitungan akurasi
3. Jurnal ketiga ditulis oleh Rizky Parlika , Sunu Ilham Pradika , Amir Muhammad Hakim , dan Kholilul Rachman N.M menggunakan algoritma *Python TextBlob* dalam melakukan sentimen analisis.
4. Jurnal Keempat ditulis oleh Raihan Astrada Fathurrahman menggunakan algoritma Vader dalam melakukan data labeling pada proses sentimen analisa.
5. Jurnal kelima ditulis oleh Musthofa Galih Pradana ,Azriel Christian Nurcahyo ,Pujo Hari Saputro menggunakan algoritma *Support Vector Machine*, *Naïve Bayes*, dan *Liner Regression* dalam melakukan sentimen analisis.

Perbedaan dan keunikan dari penelitian terdahulu dengan penelitian yang akan dilakukan adalah penelitian yang dilakukan akan menggunakan 2 algortima yaitu Support Vector Machine dan Naïve bayes pada 5 mata uang kripto yaitu *Bitcoin*, *Ethereum*, *Binance Coin*, *DOGE*, dan *Ripple*. Kedua algoritma tersebut akan digunakan untuk mengklasifikasi sentiment positif dan sentiment negatif. Sumber data akan didapatkan dari sosial media *Twitter* berupa postingan dan isi komentar pada suatu postingan mengenai mata uang kripto atau *cryptocurrency*.