

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pada beberapa tahun terakhir *cryptocurrency* menjadi perbincangan hangat di kalangan masyarakat. *Cryptocurrency* merupakan mata uang digital yang bersifat virtual dan dapat dijadikan sebagai alat untuk bertransaksi di seluruh dunia. *Cryptocurrency* menjadi perbincangan di kalangan masyarakat dunia dikarenakan aset digital tersebut tidak terikat oleh otoritas ekonomi, seperti perbankan [1]. Pada tahun 2024, para pakar ekonomi dari seluruh dunia memprediksi bahwa tren *cryptocurrency* akan mengalami kenaikan dari tahun sebelumnya. Hal tersebut didukung dengan pernyataan dari Forbes majalah bisnis dan finansial asal Amerika Serikat yang memberitakan bahwa tahun 2024 akan menjadi tahun kebangkitan dari *cryptocurrency*. Salah satu alasan yang melandasi pernyataan tersebut adalah adanya penyetujuan kebijakan *Bitcoin Spot Exchange Trade Funds* (ETF) oleh Komisi Sekuritas dan Bursa Amerika Serikat. Penyetujuan kebijakan tersebut menyebabkan lonjakan kenaikan harga yang drastis terhadap *cryptocurrency*. Salah satu instrumen *cryptocurrency*, yakni Bitcoin mengalami peningkatan harga yang signifikan. [2]. Kenaikan harga Bitcoin dari Rp 2.080.001 menjadi Rp 6.114.877 pada awal tahun 2024 menunjukkan keyakinan dan minat yang tinggi dari masyarakat dalam berinvestasi pada pasar *cryptocurrency* [3]. Lonjakan tersebut juga membuat popularitas *cryptocurrency* menjadi terus menanjak dan menarik perhatian para investor dari seluruh dunia. Popularitas Bitcoin yang terus menanjak membuat tren harga *cryptocurrency* menjadi naik. Hal tersebut terlihat pada Gambar 1.1 yang menggambarkan tren harga bitcoin, saham dan properti.



Gambar 1. 1 Tren Harga *Cryptocurrency* (Bitcoin) Vs Instrumen Investasi Lainnya

Sumber: [4]

Pada grafik yang terlihat pada Gambar 1.1 menunjukkan bahwa tren *cryptocurrency* terutama pada bitcoin mengalami kenaikan signifikan dari tahun-tahun. Pada tahun 2023 menjadi tahun puncak dari perkembangan harga bitcoin, dimana menyentuh harga \$3.000.000. Walaupun terdapat penurunan harga pasar pada bitcoin pada tahun sebelumnya, namun bitcoin tetap mengungguli tren pasar dibandingkan dengan emas, minyak bumi, dan instrumen investasi lainnya [4].

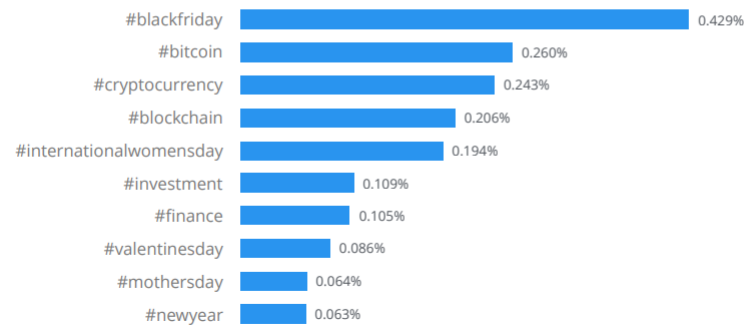
Namun, terdapat masyarakat yang skeptis atau memilih untuk menahan diri pada tren *cryptocurrency* tersebut. Hal tersebut dikarenakan tidak pernah ada pernyataan resmi dari lembaga manapun dan siapa yang mengatur alur perkembangan *cryptocurrency*. Oleh karena itu, banyak negara yang melarang perdagangan *cryptocurrency* [5]. Terdapat beberapa negara yang melarang *cryptocurrency* sebagai alat pembayaran, seperti China, Hong Kong, Iran, dan Maroko [6]. Selain itu, resiko kejahatan siber juga semakin meningkat dikarenakan banyak orang mulai membeli aset digital tersebut dalam jumlah yang sangat besar. Akibatnya, mulai banyak asumsi yang mengatakan jika aset digital tersebut merupakan tempat pencucian uang, dan bukan merupakan aset legal yang dapat bertahan dalam jangka yang panjang. Ditemukan juga kecurangan yang terjadi pada aset digital tersebut, seperti pada *cryptocurrency* FTX. Aset digital

tersebut terbukti melakukan *rug pull* atau pengambilan dana besar terhadap para masyarakat yang membeli aset digital tersebut [7]. Kemudian, banyaknya opini negatif terkait investasi pada aset digital *cryptocurrency* membuat tren harga pasar aset digital tersebut menjadi mengalami penurunan. Pada tahun 2021 adalah 859,4 triliun rupiah. Kemudian angka tersebut turun menjadi 94,41 triliun rupiah pada tahun 2023. Penurunan sebesar 69% tersebut merupakan akibat dari banyaknya opini-opini negatif yang bertebaran di kalangan masyarakat, khususnya melalui media sosial [8]. Terdapat beberapa *platform* yang sering digunakan untuk membahas mengenai *cryptocurrency* seperti Yahoo Finance dan X. Yahoo Finance merupakan media digital yang menyediakan layanan informasi keuangan *online* untuk data keuangan dan analisis pasar, seperti data *cryptocurrency* [9]. X merupakan *platform* media sosial yang digunakan untuk berkomunikasi dengan sesama secara digital [10]. Kedua *platform* tersebut merupakan sumber informasi yang sangat baik untuk mencari informasi terkait *cryptocurrency*. Meskipun Yahoo Finance merupakan sumber informasi yang sangat baik untuk data keuangan dan analisis pasar, *platform* tersebut tidak dirancang untuk diskusi sosial dan opini pengguna secara langsung seperti media sosial X [9].

Informasi yang tersebar pada media sosial menjadi perdebatan tersendiri bagi masyarakat yang ingin membeli aset digital tersebut. Akibat dari perdebatan tersebut, harga *cryptocurrency* cenderung mengalami tingkat kenaikan dan penurunan yang cukup tinggi. Salah satu contoh kasusnya adalah Elon Musk, penemu Tesla dan pemilik media sosial X pernah memberikan *tweet* terkait salah satu jenis *cryptocurrency*, yaitu Dogecoin. Hal tersebut membuat terjadinya kenaikan harga Dogecoin sampai dengan 100% hanya dalam hitungan minggu [11]. Kasus tersebut membuktikan bahwa media sosial X berpengaruh terhadap kenaikan ataupun penurunan harga *cryptocurrency*. Adapun grafik yang menggambarkan jenis *hashtag* yang cukup populer di kalangan masyarakat media sosial X. Berikut adalah grafik yang menunjukkan bahwa topik *cryptocurrency* merupakan topik yang hangat dibicarakan di media sosial X atau Twitter.

FINANCIAL SERVICES  
Twitter hashtags

Top hashtags by engagement rate (used by more than 15 companies in this study)



Gambar 1. 2 Tren Tagar Bitcoin pada Twitter (X) Periode 2022

Sumber: [12]

Gambar 1.3 memvisualisasikan tren-tren yang terjadi selama periode 2022 pada media sosial X. Tagar bitcoin menjadi *trending topic* selama penghujung tahun 2022 tepat setelah tagar *blackfriday* [12]. Komentar-komentar yang diberikan pun beragam, mulai dari positif, negatif, hingga netral. Keragaman komentar tersebut dijadikan acuan sebagai indikator dalam mengetahui tren harga *cryptocurrency* selanjutnya. Namun pada praktiknya, akan membutuhkan waktu yang sangat lama untuk membaca semua komentar-komentar yang dilemparkan pada media sosial X. Sebaliknya, bila data yang digunakan juga terlalu sedikit maka akan menyebabkan hasil prediksi menjadi tidak tepat atau *bias*. Oleh karena itu, dibutuhkan pendekatan yang tepat untuk dapat memberikan informasi yang akurat mengenai sentimen yang terdapat pada data-data tersebut. Pendekatan tersebut ialah *sentiment analysis*.

*Sentiment analysis* adalah sebuah teknik untuk menganalisis sebuah emosi dalam teks [13]. Teknik tersebut digunakan untuk menganalisis emosi positif dan negatif yang terdapat pada komentar media sosial X. Komentar-komentar tersebut akan diekstrak untuk dijadikan sebagai sumber data dalam penerapan analisis. Data-data yang diambil dari media sosial tersebut memiliki tingkat dimensi yang tinggi, oleh karena itu diperlukan *feature selection* untuk dapat meningkatkan hasil akurasi [14]. *Feature selection* digunakan untuk menseleksi fitur-fitur yang

terdapat pada data yang sudah dibersihkan [15]. Fitur-fitur yang diseleksi tersebut merupakan fitur-fitur yang terbaik pada data dan mereduksi fitur yang dirasa tidak memiliki kontribusi yang signifikan untuk analisis dan pemodelan data. *Swarm intelligence* merupakan salah bentuk dari *feature selection* yang memanfaatkan penggunaan *artificial intelligence* dengan mengadopsi perilaku sosial dari hewan yang hidup dalam kelompok [16]. Selain optimasi *feature selection*, juga dilakukan optimasi *hyperparameter* dengan menggunakan metode Grid Search. Hyperparameter akan dioptimasi pada model machine learning terlebih dahulu sebelum dilakukannya optimasi *feature selection*. Penelitian terdahulu, seperti [17] menggunakan grid search untuk mengoptimasi parameter yang digunakan pada model SVM dan mendapatkan hasil akurasi model yang lebih tinggi dan meningkat dari 80% menjadi 84%. Begitupun dengan penelitian [18] dan [19] yang melakukan optimasi *hyperparameter* pada algoritma Naïve Bayes dan KNN dengan menggunakan Grid Search. Terdapat peningkatan sebesar 1% untuk algoritma Naïve Bayes. Kemudian, pemodelan output yang digunakan akan menggunakan *machine learning* berbasis *supervised learning* yang meliputi algoritma SVM (*Support Vector Machine*), *naïve bayes*, dan KNN (*K-Nearest Neighbors*). Ketiga algoritma *machine learning* tersebut akan dilakukan perbandingan untuk mengetahui akurasi algoritma yang lebih tinggi dalam penerapan *sentiment analysis*. Terdapat penelitian terdahulu yang membandingkan algoritma SVM, KNN, dan *Naïve Bayes* untuk melakukan *sentiment analysis* pada *tweet* resmi di Indonesia. Penelitian tersebut juga menggunakan SMOTE untuk mengatasi permasalahan data tidak seimbang. Hasil penelitian tersebut menghasilkan akurasi tertinggi pada algoritma SVM-VADER-SMOTE yaitu 93% [20]. Penelitian sebelumnya melakukan komparasi *feature selection* dengan *supervised machine learning* atau algoritma klasifikasi untuk memprediksi pemilu di Indonesia periode 2019-2024 dengan menggunakan pendekatan *sentiment analysis* pada dataset Twitter. Peneliti menggunakan PSO (*Particle Swarm Optimization*) dan GA (*Generic Algorithms*) sebagai *feature selection* dan SVM sebagai algoritma klasifikasi. Hasilnya adalah model klasifikasi PSO-SVM memiliki nilai akurasi dan AUC (Area Under Curve) terbaik yakni secara berurutan 86,2% dan 93,4% dibandingkan dengan GA-SVM

[21]. Pada penelitian yang membahas mengenai komparasi model feature selection dengan model klasifikasi pada elektabilitas tokoh politik. Algoritma klasifikasi yang digunakan adalah SVM dan Naïve Bayes, *feature selection* yang digunakan adalah PSO. Hasilnya adalah terjadi peningkatan akurasi dan AUC pada model klasifikasi yang menggunakan *feature selection*. Model terbaik adalah SVM-PSO dengan akurasi dan AUC sebesar 78.4% dan 85% [22]. Penelitian sebelumnya membandingkan model klasifikasi KNN, *Information Gain* (IG) dan *Rough Set Attribute Reduction* (RSAR) dengan menggunakan *Ant Colony Optimization* (ACO) dan *Generic Algorithm* (GA) untuk *sentiment analysis* pada *customer review*. Hasil terbaiknya adalah ACO-KNN yang mampu untuk mendapatkan subset fitur yang paling optimal dan akurasi yang baik dengan rata-rata *f-score* 82.7%, mengungguli IG-RSAR dan IG-GA [23].

Kontribusi pada penelitian, diantaranya: 1) Menerapkan *sentiment analysis* terhadap tren *cryptocurrency*, 2) Menggunakan data komentar-komentar pada media sosial X terkait pembahasan *cryptocurrency* pada periode Desember 2023 – Januari 2024, 3) Menerapkan *feature selection* berbasis *swarm intelligence* pada pembahasan *cryptocurrency*, 4) Mengklasifikasikan mode *feature selection* dengan algoritma klasifikasi, seperti SVM, Naïve Bayes, dan KNN, 5) Membandingkan model PSO-SVM, ACO-SVM, CSO-SVM, PSO-Naïve Bayes, ACO- Naïve Bayes, CSO- Naïve Bayes, PSO-KNN, ACO- KNN, CSO- KNN untuk mengetahui *feature selection* dan algoritma klasifikasi terbaik dalam *sentiment analysis* pada *cryptocurrency*.

## 1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini, diantaranya:

1. Bagaimana proses pengambilan data pada media sosial X dengan menggunakan teknik *web scrapping data*?
2. Bagaimana proses pembersihan data yang dilakukan pada dataset yang digunakan dalam *sentiment analysis cryptocurrency*?
3. Bagaimana hasil perbandingan algoritma machine learning (SVM, Naïve Bayes, dan KNN) untuk *sentiment analysis* opini *cryptocurrency* pada X?

4. Bagaimana memodifikasi algoritma *machine learning* SVM, Naïve Bayes, dan KNN menggunakan algoritma optimasi berbasis *swarm intelligence* (PSO, ACO, dan CSO) sebagai *feature selection*?
5. Bagaimana hasil perbandingan dari algoritma SVM, Naïve Bayes, dan KNN sebelum dan setelah penerapan *feature selection* berbasis *swarm intelligence* (PSO, ACO, dan CSO) berdasarkan akurasi dan waktu pemrosesannya?

### 1.3 Batasan Masalah

Berdasarkan rumusan masalah yang dijabarkan sebelumnya. Batasan masalah diperlukan untuk mengarahkan topik ke inti pembahasan sehingga tujuan penelitian dapat tercapai. Berikut merupakan batasan masalah pada penelitian:

1. Objek yang diteliti adalah komentar-komentar mengenai *cryptocurrency* yang terdapat pada media sosial X
2. Periode waktu data yang diambil adalah Desember 2023 – Januari 2024
3. Data yang diambil dengan tagar #*cryptocurrency*, #*crypto*, #*bitcoin*, #*btc*, #*eth*, dan #*binance*
4. Bahasa pemrograman yang digunakan adalah Python
5. Algoritma *machine learning* yang digunakan adalah SVM, Naïve Bayes, dan KNN
6. Optimasi *feature selection* yang dilakukan adalah dengan menggunakan algoritma PSO, ACO, dan CSO.

### 1.4 Tujuan dan Manfaat Penelitian

#### 1.4.1 Tujuan Penelitian

Tujuan penelitian ini, diantaranya:

1. Mengetahui teknik dan proses pengambilan data pada media sosial X dengan menggunakan teknik *web scrapping data*?
2. Mengetahui hasil dan proses pembersihan data yang dilakukan pada dataset yang digunakan dalam *sentiment analysis cryptocurrency*.
3. Membandingkan algoritma *machine learning* (SVM, Naïve Bayes, dan KNN) untuk *sentiment analysis* opini *cryptocurrency* pada X

4. Mengetahui hasil perbandingan berdasarkan akurasi dari algoritma SVM, Naïve Bayes, dan KNN yang sudah diterapkan *feature selection* berdasarkan *swarm intelligence* (PSO, ACO, dan CSO) untuk *sentiment analysis*
5. Memahami hasil perbandingan berdasarkan jumlah fitur yang berhasil direduksi oleh algoritma *feature selection* berbasis *swarm intelligence* (PSO, ACO, dan CSO)
6. Mengetahui hasil perbandingan berdasarkan waktu eksekusi dari algoritma SVM, Naïve Bayes, dan KNN sebelum dan setelah menerapkan *feature selection* berbasis *swarm intelligence* (PSO, ACO, dan CSO)

#### 1.4.2 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat memberikan manfaat praktis dan teoritis bagi para pembacanya. Berikut merupakan manfaat pada penelitian:

1. Teoritis

Manfaat teoritis pada penelitian, diantaranya:

- 1) Hasil penelitian lain dengan menggunakan algoritma SVM, Naïve Bayes, dan KNN
- 2) Hasil penelitian diharapkan dapat menjadi acuan dalam menggunakan algoritma SVM, Naïve Bayes, dan KNN yang sudah diterapkan *feature selection* berdasarkan *swarm intelligence* (PSP, ACO, dan CSO) untuk *sentiment analysis*

2. Praktis

Manfaat praktis pada penelitian ini, diantaranya:

- 1) Penelitian ini diharapkan dapat memanfaatkan hasil dari *sentiment analysis* pada komentar ataupun opini pada media sosial untuk pengambilan keputusan dalam berinvestasi kedepannya
- 2) Memperdalam ilmu dalam penggunaan *feature selection* berdasarkan *swarm intelligence* dalam menggunakan pendekatan *sentiment analysis*



- 3) Mampu memberikan informasi kepada para calon investor dalam pemilihan waktu terbaik untuk membeli *cryptocurrency* berdasarkan sentimen yang ada

## **1.5 Sistematika Penulisan**

### **BAB 1 PENDAHULUAN**

Pada Bab 1, diuraikan penjelasan mengenai konteks latar belakang penelitian melakukan penelitian, batasan masalah, rumusan masalah, tujuan dan manfaat, serta tata cara penyusunan skripsi.

### **BAB II LANDASAN TEORI**

Pada Bab 2 dipaparkan penjelasan mengenai landasan-landasan teori yang menjadi acuan dalam penulisan skripsi, penelitian sebelumnya, dan perbandingan penelitian sebelumnya dengan penelitian yang sedang dilakukan saat ini.

### **BAB III METODOLOGI PENELITIAN**

Pada Bab 3 dipaparkan penjelasan mengenai gambaran umum objek penelitian, yakni *cryptocurrency*, model penelitian, variabel penelitian, teknik pengumpulan data, teknik pengambilan sampel, dan teknik analisis data-data yang digunakan dalam penelitian.

### **BAB IV ANALISIS DAN HASIL PENELITIAN**

Pada Bab 4 dipaparkan penjelasan mengenai analisis masalah dan kebutuhan penelitian dan hasil analisisnya.

### **BAB V SIMPULAN DAN SARAN**

Pada Bab 5 dipaparkan penjelasan mengenai kesimpulan yang didapatkan dari hasil penelitian dan saran kepada peneliti selanjutnya.