

BAB I

PENDAHULUAN

1.1 Latar Belakang

Topik finansial menjadi perbincangan yang sangat menarik oleh masyarakat dunia terutama pada aset digital yaitu *cryptocurrency*. Hal tersebut ditandai dengan salah satu *cryptocurrency* yang bernama Bitcoin, mengalami kenaikan harga secara signifikan hingga lebih dari 2000% pada tahun 2017 dan menyentuh harga kisaran US\$20000 [1]. Kejadian tersebut membuat orang-orang ingin mencoba untuk menginvestasikan uang mereka ke dalam *cryptocurrency*, mengakibatkan terdapat 300000 aktivitas transaksi harian pada jaringan Bitcoin pada tahun 2017 [2]. Ketenaran akan aset digital ini disebabkan oleh kenaikan harga yang begitu signifikan dalam kurun waktu yang relatif singkat dibandingkan produk keuangan lainnya. Alasan lain dibalik ketenaran ini adalah adopsi masal oleh masyarakat karena inovasi teknologi yang ditawarkan *cryptocurrency* yang memungkinkan transaksi antara dua belah pihak tanpa melibatkan pihak ketiga dan memiliki tingkat keamanan yang tinggi karena memanfaatkan teknologi *blockchain* [3].

Banyak isu terkait *cryptocurrency* yang membuat nilainya tidak stabil, seperti tidak ramah lingkungan karena beberapa *cryptocurrency* menggunakan mekanisme *Proof of Work* (PoW) yang melibatkan aktivitas penambangan sehingga berakibat pada konsumsi energi yang masif [4]. Selain itu, *cryptocurrency* juga dikabarkan digunakan oleh para kriminal untuk transaksi gelap, pencucian uang, skema ponzi, dan terorisme [5]. Berdasarkan beberapa isu tersebut membuat pandangan masyarakat mengenai *cryptocurrency* sebagai suatu hal yang buruk, sejumlah negara juga membuat larangan untuk memperdagangkan *cryptocurrency*, dan mulai membuat regulasi seputar aset digital ini. Indonesia sendiri melarang penggunaan *cryptocurrency* sebagai alat tukar, namun memperbolehkan untuk diperdagangkan sebagai komoditas [6]. Harapannya dengan hadirnya regulasi terkait *cryptocurrency* dapat melindungi masyarakat dari beberapa masalah yang sudah disebutkan di atas. Meskipun masih banyak masyarakat yang skeptis terhadap

cryptocurrency, namun terdapat sekelompok orang yang mengupayakan dalam mempertahankan eksistensi *cryptocurrency* hingga saat ini salah satunya Negara El Salvador yang menjadikan *cryptocurrency* sebagai *legal tender* atau alat tukar yang diyakini dapat menekan inflasi [7]. Upaya lainnya dari *developer* yaitu menciptakan mekanisme-mekanisme baru untuk *cryptocurrency* seperti *Proof of Stake* (PoS) pada jaringan Ethereum, dimana mekanisme PoS tidak mengkonsumsi energi yang banyak seperti Bitcoin sehingga membuatnya lebih ramah lingkungan [8]. Serta berita yang baru-baru ini terjadi tepatnya pada 10 Januari 2024, akhirnya *U.S. Securities and Exchange Commission* (SEC) menyetujui *Bitcoin Exchange Traded Funds* (ETF) setelah beberapa kali melakukan penolakan [9]. Berbagai berita atau isu di atas menyebabkan harga dari suatu *cryptocurrency* memiliki volatilitas yang tinggi sehingga membuatnya dikategorikan sebagai salah satu instrumen investasi yang *high risk*, apalagi masih banyak orang yang skeptis terhadap aset digital ini karena umurnya relatif muda.

Berita atau informasi mengenai *cryptocurrency* tersebar luas dan cepat ke seluruh dunia berkat media sosial. Informasi yang tersebar pada media sosial umumnya mengandung sentimen yang dapat mempengaruhi pandangan masyarakat atau pelaku pasar dalam memperdagangkan *cryptocurrency*. Kekuatan sentimen pada media sosial dapat menyebabkan harga suatu *cryptocurrency* mengalami peningkatan atau penurunan secara signifikan pada waktu tertentu [10]. Salah satu media sosial yang sering digunakan untuk menyebarkan informasi mengenai *cryptocurrency* adalah X yang awalnya bernama Twitter, Elon Musk yang merupakan salah satu orang terkaya di dunia pernah membagikan *tweets* seputar *cryptocurrency* menyebabkan harga Dogecoin naik lebih dari 100% dalam hitungan minggu [11]. Data pada *platform* X digunakan pada penelitian ini karena disebutkan bahwa sentimen pada X terbukti dapat mempengaruhi kenaikan atau penurunan pada pasar *cryptocurrency* [12]. Namun untuk membaca keseluruhan *tweets* atau opini masyarakat pada *platform* X akan sangat memakan waktu, apabila data yang dilihat hanya sedikit maka informasi yang didapatkan akan bias, hal tersebut akan membuat hasil evaluasi untuk mengetahui tren pasar *cryptocurrency* tidak sempurna. Salah satu pendekatan yang dapat dijadikan solusi atas

permasalahan di atas yaitu dengan menerapkan *sentiment analysis*. *Sentiment analysis* adalah proses menganalisis dan mengekstraksi pengetahuan dari suatu informasi [13]. Tujuan diterapkan *sentiment analysis* agar proses menganalisis opini atau informasi dapat dilakukan secara efektif dan efisien, mampu menganalisis sentimen dalam jumlah besar, serta dapat menghemat waktu dalam mengidentifikasi data yang dapat dikategorikan sebagai sentimen positif, negatif, atau netral pada suatu topik yang dibahas di media sosial dalam kurun waktu tertentu [14][15].

Data yang diambil dari media sosial umumnya memiliki karakteristik seperti berdimensi tinggi atau banyak data yang tidak relevan dan memiliki banyak *noise* [16]. Hal tersebut membuat sulit untuk mendapatkan hasil yang baik apabila masih menggunakan *sentiment analysis* konvensional. Terdapat 3 pendekatan yang biasanya digunakan penelitian lain dalam melakukan *sentiment analysis* antara lain pendekatan berbasis leksikon, pendekatan berbasis *machine learning*, dan *hybrid*. *Sentiment analysis* dengan pendekatan berbasis leksikon dapat dilakukan pada data *unsupervised* atau tidak memiliki label, sedangkan penerapan *sentiment analysis* berbasis *machine learning* memerlukan data yang sudah diberikan label kemudian membagi data menjadi *data training* dan *data testing* untuk mengklasifikasikan teks berdasarkan sentimennya [17]. Adapun pendekatan *hybrid* yang menggabungkan kedua pendekatan tersebut dengan menerapkan leksikon sebagai *feature selection* dan *machine learning* untuk klasifikasi. Berdasarkan beberapa penelitian sebelumnya menunjukkan bahwa penerapan *sentiment analysis* dengan metode *hybrid* dapat meningkatkan akurasi secara signifikan sekitar 10% dibandingkan *sentiment analysis* konvensional [18]. Contoh dari *sentiment analysis* konvensional adalah hanya menggunakan *machine learning* biasa, namun terdapat pengembangan dari *machine learning* yaitu *deep learning* yang mampu menganalisis data secara *unsupervised* dalam jumlah besar dan dapat mengekstraksi pola yang kompleks [19]. Berdasarkan beberapa penelitian terdahulu, algoritma *deep learning* memberikan kinerja yang lebih baik dibandingkan *machine learning* konvensional dilihat dari akurasi yang lebih tinggi dan efisiensi energi yang signifikan [20]. Alhasil banyak peneliti yang mencoba untuk melakukan diagnosis

penyakit menggunakan algoritma *deep learning* karena sejumlah kelebihan yang ditawarkan [21][22][23].

Penelitian ini menggunakan algoritma *deep learning Long Short-Term Memory Networks* (LSTM) karena berdasarkan beberapa penelitian terdahulu hasil dari algoritma LSTM memberikan hasil terbaik dibandingkan model yang lain [10][15][24]. Alasan menggunakan algoritma LSTM untuk *sentiment analysis* didasari oleh kemampuan model dalam menangkap informasi berurutan dalam data teks, memungkinkan model untuk memahami konteks dan urutan kata yang penting dalam *sentiment analysis* [25][26]. Algoritma LSTM juga dirancang khusus untuk mengatasi masalah *vanishing gradient* yang sering terjadi pada *Recurrent Neural Networks* (RNN) tradisional, sehingga mampu mengingat informasi penting dalam jangka waktu yang lebih lama [27][28]. Algoritma ini sangat baik bila digunakan untuk *sentiment analysis* pada data yang berasal dari media sosial seperti ulasan komentar di *platform X*, karena kemampuannya untuk menangkap konteks dari teks yang tidak terstruktur [29][30]. Algoritma LSTM seringkali menunjukkan performa yang lebih baik dalam hal akurasi dan kemampuan generalisasi dibandingkan dengan model tradisional lainnya seperti *Support Vector Machine* (SVM) dan *Convolution Neural Network* (CNN), khususnya dalam tugas *sentiment analysis* [31]. Meskipun terdapat banyak kelebihan yang ditawarkan algoritma LSTM, algoritma ini mempunyai beberapa permasalahan dalam melakukan *sentiment analysis*. Algoritma LSTM memiliki kecenderungan untuk *overfitting*, terutama ketika dilatih pada *dataset* kecil. *Overfitting* terjadi ketika model terlalu menyesuaikan dengan *data training* sehingga kehilangan kemampuan untuk menggeneralisasi pada data baru [32]. Algoritma *deep learning* seperti LSTM juga terkenal dari kebutuhannya akan sumber daya komputasi yang besar dan memakan waktu, terutama untuk *dataset* yang besar dan kompleks dikarenakan sifat LSTM yang rekursif [33]. Adapun pemilihan *hyper-parameter* yang optimal seperti jumlah *neuron*, *learning rate*, dan jumlah *layer* juga mempengaruhi kinerja dari algoritma LSTM. Maka dari itu diperlukan beberapa hal untuk optimasi dalam meningkatkan kinerja dari algoritma LSTM, salah satunya dapat ditingkatkan melalui proses *pre-*

processing menggunakan leksikon apabila dibandingkan dengan hanya mengandalkan kemampuan dari algoritma *deep learning* itu sendiri [24].

Salah satu cara untuk meningkatkan performa pada *sentiment analysis* yaitu dengan menerapkan proses *feature selection* berbasis leksikon [34]. Tujuan diterapkan *feature selection* yaitu mengurangi dimensi data dan meningkatkan kinerja model dengan menghilangkan fitur yang tidak relevan. Permasalahan yang dialami ketika melakukan *sentiment analysis* dengan menggunakan data dari media sosial terletak pada jumlah data yang terus berkembang dan semakin kompleks dari waktu ke waktu, membuat *feature selection* konvensional tidak cukup untuk mengatasi permasalahan tersebut. Oleh sebab itu, diperlukan *feature selection* yang canggih, cepat, dan dapat diandalkan untuk mengatasi permasalahan yang kompleks yaitu dengan *feature selection* berbasis *swarm intelligence*. *Swarm intelligence* adalah bagian dari *Artificial Intelligence* (AI) untuk menyelesaikan masalah optimasi yang kompleks, dengan menyediakan metode yang cepat dan andal [35]. Penggunaan *swarm intelligence* sebagai *feature selection* diharapkan dapat meningkatkan hasil optimasi saat melakukan *sentiment analysis* khususnya pada bahasan *cryptocurrency*. Algoritma *swarm intelligence* yang digunakan pada penelitian ini antara lain *Particle Swarm Optimization* (PSO), *Ant Colony Optimization* (ACO), dan *Cat Swarm Optimization* (CSO). Algoritma PSO digunakan sebagai *feature selection* pada *sentiment analysis* pada penelitian ini dikarenakan dapat meningkatkan akurasi pada model *machine learning* seperti SVM dari 78.70% menjadi 86.20% [36][37]. Adapun penelitian serupa yang membahas *sentiment analysis* menggunakan klasifikasi LSTM dengan *feature selection Adaptive Particle Swarm Optimization* (APSO) yang meningkatkan akurasi dari 95.1% menjadi 97.8% [38]. Algoritma ACO sebagai *feature selection* pada penelitian *sentiment analysis* juga memberikan hasil yang paling optimal dengan rata-rata *f-score* 82.7% pada model klasifikasi menggunakan *K-Nearest Neighbour* (KNN) [39]. Adapun penelitian lain yang membahas *sentiment analysis* dengan membandingkan algoritma PSO dan CSO pada model LSTM, dimana algoritma CSO mampu mengungguli PSO dalam menentukan sorotan fitur lebih dalam, lebih cepat, dan mendapatkan akurasi yang lebih baik yaitu 92.56% [15].

Oleh sebab itu dengan menggabungkan antara algoritma LSTM dengan algoritma swarm intelligence seperti PSO, ACO, dan CSO dalam *sentiment analysis cryptocurrency* dapat digunakan untuk mencari kombinasi *hyper-parameter* terbaik seperti LSTM *Unit* yang kemudian digunakan untuk melatih model LSTM.

Sudah banyak penelitian yang melakukan *sentiment analysis cryptocurrency* oleh karena itu berdasarkan penelitian sebelumnya, kontribusi yang diberikan pada penelitian ini antara lain: 1) menerapkan algoritma *deep learning* yaitu LSTM untuk *sentiment analysis cryptocurrency* berdasarkan opini di *platform X*, 2) menerapkan algoritma optimasi yaitu *feature selection* berbasis *swarm intelligence* PSO, ACO, dan CSO untuk meningkatkan kinerja algoritma LSTM berdasarkan *accuracy*, *loss*, dan *execution time* pada *sentiment analysis cryptocurrency* di *platform X*, 3) mengevaluasi hasil perbandingan kinerja pada algoritma LSTM sebelum dan sesudah diterapkan *feature selection* berbasis *swarm intelligence* PSO, ACO, dan CSO berdasarkan *accuracy*, *loss*, dan *execution time*.

1.2 Rumusan Masalah

Rumusan masalah dari karya ilmiah ini antara lain:

- 1) Bagaimana menerapkan algoritma *deep learning* yaitu LSTM untuk *sentiment analysis cryptocurrency* berdasarkan opini di *platform X*?
- 2) Bagaimana menerapkan algoritma optimasi berbasis *swarm intelligence* dengan PSO, ACO, dan CSO masing-masing pada algoritma LSTM untuk *sentiment analysis cryptocurrency* berdasarkan opini di *platform X*?
- 3) Bagaimana hasil perbandingan kinerja algoritma LSTM sebelum dan setelah diterapkan *feature selection* berbasis *swarm intelligence* dengan PSO, ACO, dan CSO masing-masing berdasarkan *accuracy*, *loss*, dan *execution time*?

1.3 Batasan Masalah

Batasan-batasan pada penyusunan karya ilmiah ini antara lain:

- 1) Data yang digunakan untuk melakukan *sentiment analysis* berupa komentar atau opini tentang *cryptocurrency* yang berasal dari *platform X*.

- 2) Periode data *platform X* yang diambil yaitu dari tanggal 31 Desember 2023 – 31 Januari 2024 menggunakan *keywords* '(*cryptocurrency OR crypto OR bitcoin OR \$btc OR ethereum OR \$eth OR binance OR \$bnb OR solana OR \$sol*) lang:en'.
- 3) Pelabelan pada data hanya terdiri dari sentimen positif dan negatif untuk menghindari ambiguitas dalam analisis data.
- 4) *Feature selection* yang digunakan berbasis *swarm intelligence* pada domain *cryptocurrency*, dengan algoritma PSO, ACO, dan CSO.
- 5) Optimasi algoritma LSTM berfokus pada pengaturan parameter LSTM yang didapat dari *feature selection* berbasis *swarm intelligence*.
- 6) Metrik yang dibandingkan antar model klasifikasi sentimen adalah *accuracy*, *loss*, dan *execution time*.
- 7) Menggunakan bahasa pemrograman Python untuk melakukan penelitian.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, tujuan dari penelitian ini antara lain:

- 1) Menerapkan algoritma *deep learning* yaitu LSTM untuk *sentiment analysis cryptocurrency* berdasarkan opini di *platform X*.
- 2) Menerapkan algoritma optimasi yaitu *feature selection* berbasis *swarm intelligence* dengan PSO, ACO, dan CSO masing-masing untuk meningkatkan kinerja algoritma LSTM berdasarkan *accuracy*, *loss*, dan *execution time* pada *sentiment analysis cryptocurrency* di *platform X*.
- 3) Mengevaluasi hasil perbandingan kinerja pada algoritma LSTM sebelum dan sesudah diterapkan *feature selection* berbasis *swarm intelligence* dengan PSO, ACO, dan CSO masing-masing berdasarkan *accuracy*, *loss*, dan *execution time*.

1.4.2 Manfaat Penelitian

Manfaat penelitian yang diharapkan dari penelitian ini antara lain:

- 1) Manfaat Teoritis

Penerapan algoritma optimasi dengan *feature selection* berbasis *swarm intelligence* seperti PSO, ACO, dan CSO pada LSTM diharapkan dapat menjadi sumbangsih bagi penelitian serupa terkait *sentiment analysis* salah satunya pada bahasan *cryptocurrency*.

2) Manfaat Praksis

- a) Memberikan informasi kepada masyarakat atau pelaku pasar agar mampu memanfaatkan peluang atas sentimen baik berupa komentar atau opini pada media sosial dalam mengambil keputusan investasi pada aset *cryptocurrency*.
- b) Memberikan informasi atas pandangan masyarakat terhadap *cryptocurrency*.

1.5 Sistematika Penulisan

Penulisan karya ilmiah ini menggunakan *template* yang terstandarisasi untuk mahasiswa Sistem Informasi di Universitas Multimedia Nusantara. *Template* ini digunakan untuk mempermudah dalam melihat dan mengetahui pembahasan secara keseluruhan dari karya ilmiah ini. Karya ilmiah ini terbagi menjadi lima bab utama, berikut sistematika dari penulisannya:

BAB 1 PENDAHULUAN

Bagian ini berisi penjelasan mengenai alasan topik penelitian ini diangkat. Pendahuluan terdiri dari latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Bagian ini menjelaskan mengenai teori-teori yang berhubungan dengan topik penelitian yang dilakukan. Adapun penelitian terdahulu berisi *review* dari penelitian-penelitian sebelumnya dalam mendukung penelitian yang dilakukan.

BAB III METODOLOGI PENELITIAN

Bagian ini menjelaskan mengenai objek penelitian dan metode atau tahapan dalam melakukan penelitian meliputi teknik pengumpulan data, variabel penelitian, dan teknik analisis data.

BAB IV ANALISIS HASIL PENELITIAN

Bagian ini berisi proses yang dipaparkan pada bagian metodologi untuk mencapai tujuan penelitian.

BAB V SIMPULAN DAN SARAN

Pada bab ini menjelaskan mengenai kesimpulan dari keseluruhan penelitian yang dilakukan. Serta saran untuk penelitian di masa mendatang yang akan membahas topik serupa.



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA