

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi *blockchain* telah menciptakan revolusi dalam sistem keuangan digital, dengan *cryptocurrency* menjadi aset investasi yang semakin diminati [1]. Bitcoin sebagai *cryptocurrency* pertama dan terbesar berdasarkan kapitalisasi pasar menunjukkan pertumbuhan yang signifikan namun disertai volatilitas harga yang ekstrem [2]. Tingginya volatilitas harga Bitcoin dipengaruhi oleh berbagai faktor seperti sentimen pasar, berita finansial, dan kondisi ekonomi global, sehingga menyulitkan proses prediksi harga secara akurat [3]. Beberapa kajian literatur menggarisbawahi bahwa volatilitas yang tinggi membuat prediksi harga jangka panjang menjadi tantangan tersendiri, sehingga menuntut pendekatan metodologis yang kuat untuk mendukung keputusan investasi berbasis data [4][5].

Model *deep learning* telah terbukti efektif dalam memprediksi *time series* finansial, dengan *Long Short-Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU) menunjukkan performa unggul dalam menangkap pola non-linear dan dependensi jangka panjang [6]. Studi terbaru menunjukkan bahwa dalam prediksi harga Bitcoin, model GRU mampu mencapai nilai *Mean Absolute Percentage Error* (MAPE) sebesar 0,38% dan *Root Mean Square Error* (RMSE) sebesar 343,02, sementara model LSTM mencatat MAPE sebesar 0,41% dan RMSE sebesar 356,01, yang menunjukkan bahwa kedua model dapat menghasilkan prediksi dengan tingkat akurasi yang kompetitif [7]. Namun, sebagian besar penelitian sebelumnya hanya mengandalkan data historis harga tanpa mempertimbangkan faktor eksternal yang signifikan mempengaruhi pasar *cryptocurrency* [8]. Berdasarkan analisis terhadap penelitian terdahulu, ditemukan gap di mana integrasi analisis sentimen berita finansial khususnya untuk prediksi jangka panjang masih terbatas [9]. Jaquart et al. (2021) misalnya, meskipun telah mengintegrasikan fitur sentimen ke dalam model LSTM dan GRU untuk prediksi harga Bitcoin, penelitian tersebut hanya mencakup horizon prediksi sangat pendek yaitu 1 hingga 60 menit dengan sumber sentimen terbatas pada data Twitter, sehingga kontribusi sentimen berita finansial pada cakupan waktu yang lebih panjang belum tereksplorasi [10]. Sejalan dengan itu, tinjauan sistematis oleh Otabek & Choi (2024) yang mengulas berbagai studi prediksi harga *cryptocurrency* berbasis sentimen menunjukkan

bahwa seluruh studi yang dibahas hanya mencakup integrasi sentimen dari platform sosial media seperti Twitter dan Reddit pada horizon prediksi jangka pendek, tanpa satu pun yang secara khusus mengkaji kontribusi sentimen berita finansial pada prediksi jangka panjang [11]. Padahal, sentimen pasar yang terefleksi melalui berita finansial memiliki korelasi kuat dengan pergerakan harga Bitcoin [3][9]. Hal ini menunjukkan bahwa integrasi analisis sentimen berpotensi meningkatkan akurasi prediksi harga Bitcoin.

Analisis sentimen dilakukan menggunakan model FinBERT yang dirancang khusus untuk domain keuangan sehingga mampu menangkap persepsi pasar dari berita *cryptocurrency* secara lebih akurat dibandingkan model NLP umum [12]. Dalam penelitian ini, prediksi jangka panjang didefinisikan sebagai prediksi pada horizon waktu beberapa tahun ke depan, yang dibangun dari model yang dilatih menggunakan data historis harian selama kurang lebih sembilan tahun (Januari 2017 hingga Desember 2025), sehingga model mampu menangkap dependensi dan tren pergerakan harga Bitcoin secara menyeluruh [13]. Namun, penelitian yang mengintegrasikan analisis sentimen untuk prediksi harga Bitcoin jangka panjang masih terbatas, sehingga membuka peluang untuk pengembangan pendekatan yang lebih komprehensif. Oleh karena itu, penelitian ini mengusulkan metode prediksi harga Bitcoin jangka panjang dengan mengintegrasikan data historis harga dan analisis sentimen berita finansial menggunakan model *deep learning* berbasis LSTM dan GRU serta pemanfaatan FinBERT untuk ekstraksi sentimen. Pendekatan ini diharapkan mampu meningkatkan akurasi prediksi dengan mempertimbangkan pengaruh informasi eksternal terhadap pergerakan harga Bitcoin, sehingga penelitian ini berjudul **“Prediksi Harga Bitcoin Jangka Panjang Menggunakan LSTM dan GRU Berbasis Analisis Sentimen FinBERT”**.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana efektivitas model *deep learning* berbasis LSTM dan GRU dalam memprediksi harga Bitcoin pada horizon waktu jangka panjang?
2. Bagaimana pengaruh integrasi skor sentimen berita finansial yang diekstraksi menggunakan FinBERT terhadap akurasi prediksi harga Bitcoin jangka panjang dibandingkan dengan model yang hanya mengandalkan data historis harga?

1.3 Tujuan

Berdasarkan rumusan masalah di atas, tujuan penelitian ini adalah:

1. Mengevaluasi efektivitas model deep learning berbasis LSTM dan GRU dalam memprediksi harga Bitcoin pada horizon waktu jangka panjang.
2. Mengukur peningkatan akurasi prediksi harga Bitcoin jangka panjang yang dihasilkan dari integrasi skor sentimen berita finansial berbasis FinBERT dibandingkan dengan model yang hanya mengandalkan data historis harga, berdasarkan metrik R^2 , RMSE, MAPE, dan Directional Accuracy.

1.4 Manfaat

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Bagi pengembangan ilmu pengetahuan, penelitian ini menyediakan bukti empiris mengenai efektivitas integrasi analisis sentimen berita finansial berbasis FinBERT ke dalam model LSTM dan GRU untuk prediksi harga Bitcoin jangka panjang, termasuk perbandingan kuantitatif kinerja kedua belas skenario model berdasarkan metrik R^2 , RMSE, MAPE, dan *Directional Accuracy*.
2. Bagi institusi pendidikan dan peneliti selanjutnya, penelitian ini dapat menjadi rujukan konkret dalam pengembangan model prediksi *time series* finansial berbasis *deep learning*, khususnya sebagai acuan dalam pemilihan arsitektur model, strategi integrasi fitur sentimen, dan penggunaan Optuna sebagai metode optimasi *hyperparameter* pada domain *cryptocurrency*.
3. Bagi investor dan pelaku pasar kripto, penelitian ini memberikan informasi spesifik mengenai sejauh mana sentimen berita finansial berkontribusi terhadap akurasi prediksi harga Bitcoin jangka panjang, sehingga dapat dijadikan pertimbangan dalam menyusun strategi investasi berbasis data pada horizon waktu beberapa tahun ke depan.
4. Bagi pengembang sistem dan praktisi *data science*, penelitian ini memberikan panduan teknis mengenai *pipeline* lengkap integrasi FinBERT dengan model LSTM dan GRU, mulai dari ekstraksi sentimen, normalisasi fitur, pembentukan sekuens, hingga evaluasi model, yang dapat diadaptasi untuk pengembangan sistem prediksi harga aset digital lainnya.

1.5 Ruang Lingkup

Penelitian ini akan dilaksanakan sepenuhnya pada platform Google Colab dengan ruang lingkup teknis berikut:

1. Platform dan tools yang digunakan dalam penelitian ini meliputi Google Colab dengan dukungan GPU runtime serta Google Drive sebagai media penyimpanan data. Proses pengolahan data dan pengembangan model dilakukan menggunakan bahasa pemrograman Python versi 3.8 atau yang lebih baru dengan beberapa pustaka pendukung seperti TensorFlow 2.8 atau yang lebih baru, Transformers 4.20 untuk implementasi FinBERT, Pandas dan NumPy untuk pengolahan data, Scikit-learn untuk proses preprocessing dan evaluasi model, Matplotlib dan Seaborn untuk visualisasi data, serta Optuna yang digunakan untuk proses optimasi hyperparameter berbasis Bayesian Optimization.
2. Data yang digunakan dalam penelitian ini terdiri dari dua jenis data, yaitu data harga historis Bitcoin dan data teks berita. Data harga Bitcoin berupa data harian dengan atribut Open, High, Low, Close, dan Volume (OHLCV) pada periode Januari 2017 hingga Desember 2025 dengan total sekitar 3.287 data observasi. Data tersebut diperoleh dari sumber data publik yang menyediakan informasi harga aset kripto secara terbuka. Periode ini dipilih karena merepresentasikan siklus pasar Bitcoin yang lengkap, mencakup *bull market* 2017, bear market 2018 - 2019, *halving* 2020, lonjakan harga 2021, koreksi besar 2022, serta pemulihan 2023 - 2025, sehingga model dapat belajar dari berbagai kondisi pasar yang beragam. Data sebelum tahun 2017 tidak digunakan karena *volume* perdagangan dan likuiditas pasar Bitcoin yang belum matang pada periode tersebut dapat menghasilkan pola yang tidak representatif untuk kondisi pasar saat ini. Sementara itu, data sentimen diperoleh dari keseluruhan 66.623 artikel berita CryptoVision Dataset yang membahas Bitcoin pada periode yang sama, dan seluruhnya digunakan dalam proses analisis sentimen [14].
3. Arsitektur model yang digunakan dalam penelitian ini terdiri dari 12 skenario model eksperimen yang mencakup: (1) LSTM, (2) GRU, (3) LSTM dengan Sentimen, (4) GRU dengan Sentimen, (5) Stacked LSTM, (6) Stacked GRU, (7) Stacked LSTM dengan sentimen, (8) Stacked GRU dengan sentimen, (9) Optimasi LSTM, (10) Optimasi GRU, (11) Optimasi LSTM dengan sentimen, serta (12) Optimasi GRU dengan sentimen.
4. Metode analisis yang digunakan meliputi tahap preprocessing data menggunakan teknik normalisasi MinMaxScaler, diikuti proses ekstraksi fitur sentimen menggunakan model FinBERT yang diakses melalui ProsusAI/finbert pada Hugging Face. Hasil ekstraksi sentimen kemudian digunakan sebagai fitur tambahan yang digabungkan dengan data harga Bitcoin dalam proses pelatihan model. Dataset dibagi menjadi data pelatihan dan data pengujian dengan rasio 80:20. Evaluasi kinerja model dilakukan menggunakan

beberapa metrik evaluasi yaitu R-squared (R^2), *Mean Absolute Percentage Error* (MAPE), *Root Mean Square Error* (RMSE), serta *Directional Accuracy* (DA). Selain itu, validasi model dilakukan menggunakan teknik time-series cross validation.

5. Output yang dihasilkan dari penelitian ini berupa notebook Google Colab yang berisi keseluruhan kode implementasi model, dataset yang telah melalui proses preprocessing, model hasil pelatihan dalam format file .h5, visualisasi hasil prediksi harga Bitcoin, serta analisis komparatif terhadap kinerja keduabelas skenario model beserta rekomendasi model dengan performa terbaik.
6. Batasan penelitian ini mencakup beberapa hal yang perlu diperhatikan. Pertama, penelitian ini tidak membangun sistem prediksi harga Bitcoin secara real-time seluruh proses pelatihan, pengujian, dan evaluasi dilakukan secara offline menggunakan data historis yang telah tersedia. Kedua, model yang dihasilkan hanya dievaluasi pada data historis dan tidak divalidasi pada kondisi pasar masa depan yang sesungguhnya. Ketiga, proses komputasi dilakukan sepenuhnya pada platform Google Colaboratory sehingga terikat pada keterbatasan sumber daya GPU yang tersedia.

