

BAB I

PENDAHULUAN

1.1. Latar Belakang

Investasi adalah kegiatan menanamkan aset ke dalam suatu sektor atau area tertentu [1]. Ada banyak cara untuk melakukan investasi antara lain, memiliki aset tanah, emas, saham, deposito, obligasi, mata uang asing hingga mata uang kripto atau biasa dikenal dengan istilah *cryptocurrency* [2]. *Cryptocurrency* adalah mata uang digital yang semakin populer dimanfaatkan sebagai aset dan penanaman modal oleh penggunaanya [3]. Meski tidak memiliki bentuk fisik, mata uang ini dapat ditransaksikan melalui perdagangan *online*. Jumlah investor *cryptocurrency* atau mata uang kripto mengalami kenaikan setiap tahunnya sampai angka belasan juta pada tahun 2021 [4].

Salah satu jenis *cryptocurrency* yang telah meluas pemakaiannya yaitu *bitcoin*. *Bitcoin* pertama kali diperkenalkan tahun 2009 oleh seseorang (ataupun sekelompok) dengan nama samaran Satoshi Nakamoto [5]. Pengelola *bitcoin* bukan berasal dari bank sentral, melainkan dikelola langsung dari pengguna *bitcoin* yang identitasnya dirahasiakan. Transaksi *bitcoin* dicatat dalam blok-blok pada teknologi *blockchain* yang setiap bloknnya terhubung dan membentuk rantai yang susah untuk dimanipulasi [6]. Proses transaksinya yang aman, mudah, dan cepat serta memiliki potensi keuntungan yang signifikan di masa mendatang membuat *bitcoin* memiliki banyak peminat.

Walaupun demikian, perdagangan *bitcoin* merupakan aktivitas yang berisiko karena harga *cryptocurrency* yang terus berfluktuasi secara signifikan setiap harinya [7]. Berdasarkan data harga *bitcoin* selama 5 tahun kebelakang (2019-2023) pada website *Yahoo finance*, harga *bitcoin* pada Bulan Maret 2021 melambung tinggi mencapai 61243,69 *USD*. Namun, harga *bitcoin* mengalami penurunan yang cukup tajam pada Bulan Juli 2021 mencapai 31533,07 *USD*. Kemudian, harga *bitcoin* kembali meningkat secara signifikan dan mencapai puncak tertingginya pada Bulan November 2021, yaitu sebesar 64469,63 *USD*. Berdasarkan data tersebut, dapat disimpulkan bahwa salah satu harga

cryptocurrency, yaitu *bitcoin* sangat fluktuatif dari masa ke masa. Oleh karena itu, prediksi harga *bitcoin* sangat penting dilakukan untuk meminimalisir terjadinya kerugian dalam investasi *bitcoin*.

Pergerakan harga *bitcoin* yang sangat fluktuatif tidak hanya dipengaruhi oleh faktor internal. Namun, juga faktor eksternal seperti sentimen publik mengenai masalah politik maupun ekonomi global juga berpengaruh terhadap fluktuatifnya harga *bitcoin* [8]. Sentimen publik mencerminkan pandangan dan keyakinan masyarakat terhadap *bitcoin* yang secara langsung dapat mempengaruhi permintaan dan penawaran. Sebagai contoh pada Bulan Juni 2021, harga *bitcoin* mengalami penurunan karena Elon Musk mengungkapkan ketidaktertarikannya terhadap *cryptocurrency* dengan menambahkan tagar *bitcoin* di *Twitter* dan mengumumkan bahwa Tesla tidak akan menerima pembayaran menggunakan *bitcoin* lagi. Hal tersebut memicu dugaan masyarakat bahwa nilai *cryptocurrency*, khususnya sebagai alat pembayaran sangat fleksibel dan bergantung pada penggunaannya [9]. Keterkaitan antara sentimen publik dengan fluktuasi harga *bitcoin* juga dibuktikan oleh penelitian [10] yang dilakukan dengan menghitung korelasi antara harga *bitcoin* dengan sentimen data *Twitter*. Nilai koefisien ϕ yang dihasilkan dari perhitungan korelasi koefisien *pearson* sebesar 0.5 menunjukkan bahwa harga *bitcoin* dengan sentimen data *Twitter* yang membahas mengenai *bitcoin* memiliki hubungan positif kuat.

Selain sentimen publik, popularitas juga berpengaruh terhadap harga *bitcoin* yang dapat diakumulasikan melalui data pencarian *Google*. *Google* mendominasi ranah mesin pencarian dan menjadi tempat utama yang mencerminkan berbagai macam pandangan dan minat masyarakat [11]. Salah satu layanan *Google* yang dapat digunakan untuk menampilkan nilai indeks dari kata kunci pencarian oleh pengguna internet di seluruh dunia, yaitu *Google Trends* [12]. Nilai indeks pada *Google Trends* menunjukkan popularitas relatif yang berkisar antara 0 sampai 100 serta memiliki keterkaitan dengan harga *bitcoin* yang dibuktikan oleh penelitian [13], dimana hasilnya mengatakan bahwa nilai indeks *bitcoin* pada *Google Trends* memiliki hubungan yang positif dan signifikan terhadap harga *bitcoin*.

Adanya hubungan antara kedua faktor eksternal tersebut dengan harga *bitcoin* menjadi fokus pada penelitian ini untuk melihat sejauh mana akurasi terbaik dapat dihasilkan oleh model prediksi yang dibangun menggunakan teknik regresi. Teknik regresi memungkinkan suatu model dapat memproyeksikan kondisi di masa depan dengan cara membentuk pola hubungan antara satu atau lebih variabel independen dengan variabel dependen [14]. Pemilihan teknik regresi harus mempertimbangkan karakteristik data serta didasarkan pada hasil pengujian asumsi klasik untuk memastikan validitas model yang diterapkan terhadap data yang dianalisis [15]. Data historis harga *bitcoin* memiliki tipe data kontinu dan bersifat *time series*, sedangkan data sentimen dan *Google Trends* memiliki tipe data diskrit dan kategorikal. Sebaran data yang dihasilkan memiliki *outlier* karena harga *bitcoin* yang sangat fluktuatif. Selain itu, uji asumsi klasik yang diimplementasikan pada data mengindikasikan bahwa data tidak berdistribusi normal, terjadi multikolinearitas antar variabel bebas, tidak adanya linearitas, terdapat autokorelasi, dan terjadi gejala heteroskedastisitas pada data.

Berdasarkan karakteristik data pada penelitian ini, teknik regresi yang dapat digunakan yaitu regresi non-linear. Beberapa algoritma regresi non-linear yang biasa digunakan untuk memprediksi harga kripto antara lain yaitu, *XGBoost regression* dan *Long Short Term Memory (LSTM) for regression* [16], [17]. Algoritma *XGBoost* memiliki keunggulan dalam menangani data dengan fitur kategorikal dan *outlier* secara efektif, namun memiliki keterbatasan dalam menggeneralisasi pola data, terutama pada data yang memiliki hubungan berurutan kompleks, seperti data deret waktu. Hal ini disebabkan oleh ketidakmampuannya dalam menangani dinamika kompleks seperti tren, musiman, dan autokorelasi dalam data deret waktu, serta kesulitan mengatasi fluktuasi yang tidak tercakup secara memadai dalam kumpulan data pelatihan [18], [19]. Di sisi lain, algoritma *LSTM* unggul dalam memahami pola data berurutan jangka panjang karena kemampuannya mempertahankan informasi dalam memori, namun memiliki kelemahan dalam waktu pelatihan yang lebih lama dan kebutuhan komputasi yang lebih tinggi, terutama pada dataset berukuran besar [20], [21]. Kedua algoritma tersebut memiliki kelebihan dan kekurangan masing-

masing. Oleh karena itu, perbandingan dilakukan untuk menentukan algoritma yang paling optimal dalam memprediksi harga *bitcoin*, baik dari segi performa model maupun kemampuannya dalam menangani pola data yang kompleks [22].

Keunggulan *XGBoost regression* dibuktikan pada penelitian [16] yang menunjukkan bahwa model *Grid Search XGBoost* menghasilkan *MSE* dan *R2-Score* terendah dibandingkan *linear regression* dan *random forest regression* dalam memprediksi harga *bitcoin* berdasarkan harga harian *bitcoin* dan sentimen pasar. Selain itu, performa *XGBoost regression* juga cukup baik dalam memprediksi harga saham Perusahaan *Amazon.com, Inc*, dimana model berhasil mencapai nilai *RMSE* sebesar 0,009437 [23]. Penelitian lain juga dilakukan untuk memprediksi harga *bitcoin* menggunakan *LSTM*, dimana hasil penelitiannya mengatakan bahwa model memiliki akurasi yang cukup baik dengan nilai *RMSE* dan *MAPE* berturut-turut sebesar 2033,28 dan 3,53% [10]. Keandalan *LSTM* dalam kasus prediksi terutama pada data *time series* dibuktikan dengan hasil penelitian [24] yang mampu memprediksi harga *cryptocurrency*, terutama untuk koin *DOGE* dan *ADA* dengan nilai *RMSE* berturut-turut sebesar 0,0544 dan 0,1607.

Berdasarkan uraian yang telah dipaparkan, penelitian ini dilakukan untuk membandingkan algoritma regresi, yaitu *XGBoost regression* dan *LSTM for regression* untuk memprediksi harga penutupan *bitcoin* berdasarkan harga historis serta penambahan variabel eksternal, yaitu sentimen *Twitter* dan *Google Trends*. Adapun perbandingan performa model regresi diidentifikasi menggunakan metrik evaluasi, yaitu *Root Mean Squared Error (RMSE)* dan *R2-Score* untuk mengukur sejauh mana model mampu memprediksi harga dengan akurat serta seberapa kecil kesalahan yang dihasilkan dari prediksi berdasarkan variabel input yang telah ditetapkan. Penelitian ini tidak hanya berfokus pada pencapaian model prediksi yang akurat, tetapi juga bertujuan untuk mengidentifikasi variabel yang memiliki dampak signifikan terhadap prediksi harga *bitcoin*.

1.2. Rumusan Masalah

Harga *cryptocurrency*, khususnya *bitcoin* yang cenderung berfluktuasi tidak hanya dipengaruhi oleh faktor internal. Faktor eksternal, seperti sentimen publik dan nilai indeks pencarian pada *Google* (*Google Trends Index*) juga dapat mempengaruhi pergerakan harga *bitcoin*. Untuk meminimalisir kerugian akibat fluktuasi harga *bitcoin*, diperlukan model prediksi yang tepat. Model ini harus mampu memperhitungkan harga historis *bitcoin* serta variabel eksternal seperti sentimen publik dan *Google Trends Index*. Dalam hal ini, dibutuhkan perbandingan antara algoritma regresi yang berbeda, seperti *XGBoost regression* dan *LSTM for regression* untuk menghasilkan model prediksi yang lebih akurat dan dapat diandalkan dalam mengestimasi harga *bitcoin* di masa mendatang.

1.3. Pertanyaan Penelitian

Adapun pertanyaan pada penelitian ini adalah:

1. Bagaimana perbandingan performa algoritma *XGBoost regression* dan *LSTM for regression* dalam memprediksi harga *bitcoin*?
2. Apa saja variabel yang berpengaruh signifikan terhadap performa model terpilih dalam memprediksi harga penutupan *bitcoin*?
3. Bagaimana hasil prediksi harga *bitcoin* dalam 30 hari kedepan (Bulan Maret 2024)?

1.4. Batasan Masalah

Batasan masalah dari penelitian ini adalah:

1. Menggunakan dataset sentimen yang hanya diperoleh dari platform *Twitter* dan berbahasa inggris.
2. Menggunakan salah satu jenis *cryptocurrency*, yaitu *bitcoin*.
3. Variabel eksternal yang digunakan yaitu sentimen publik dan *Google Trends Index*.
4. Jumlah dataset yang digunakan sebanyak 366 baris dan 9 kolom dimulai sejak tanggal 1 Maret 2023 sampai 29 Februari 2024.

1.5. Tujuan Penelitian

Berikut merupakan tujuan dari penelitian ini, yaitu:

1. Mengevaluasi kinerja algoritma *XGBoost regression* dan *LSTM for regression* dalam memprediksi harga *bitcoin*.
2. Mengidentifikasi variabel-variabel yang secara signifikan dapat mempengaruhi performa model terpilih dalam memprediksi harga penutupan *bitcoin*.
3. Memprediksi harga *bitcoin* dalam 30 hari kedepan (Bulan Maret 2024).

1.6. Manfaat Penelitian

Manfaat penelitian ini adalah memberikan wawasan terkait perbandingan kinerja algoritma *XGBoost regression* dan *LSTM for regression* dalam memprediksi harga *bitcoin* dengan mempertimbangkan faktor eksternal seperti sentimen publik dan indeks pencarian *Google*. Penerapan *XGBoost* dan *LSTM* untuk prediksi ini berkontribusi pada penelitian terkait *Sustainable Development Goals (SDGs)* pilar 9 - *Industry, Innovation, and Infrastructure*, khususnya dalam menciptakan sistem keuangan yang lebih efisien dan transparan. Selain itu, penelitian ini bertujuan mengidentifikasi variabel yang secara signifikan mempengaruhi performa model prediksi, sehingga dapat menjadi landasan bagi investor *cryptocurrency* dalam mengambil keputusan yang lebih tepat dan berbasis data.