

**PENERAPAN MODEL XGBOOST UNTUK SISTEM  
REKOMENDASI SAHAM SEKTOR PROPERTI  
MENGUNAKAN ANALISIS TEKNIKAL**

**LAPORAN TUGAS AKHIR**

Laporan Ini Disusun untuk Memenuhi Salah Satu Syarat Memperoleh Gelar  
Sarjana Strata 1 (S1) pada Program Studi Teknik Informatika  
Fakultas Teknologi Industri Universitas Islam Sultan Agung Semarang



**DISUSUN OLEH:**

**MUHAMMAD IKO WARDHANA**

**32602100079**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNOLOGI INDUSTRI  
UNIVERSITAS ISLAM SULTAN AGUNG  
SEMARANG**

**2025**

***IMPLEMENTATION OF THE XGBOOST MODEL FOR A STOCK  
RECOMMENDATION SYSTEM IN THE PROPERTY SECTOR USING  
TECHNICAL ANALYSIS***

***FINAL PROJECT***

*Proposed to complete the requirement to obtain a bachelor's degree (S1) at  
Informatics Engineering Departement of Industrial Technology Faculty Sultan  
Agung Islamic University*



***ARRANGED BY:***

**MUHAMMAD IKO WARDHANA**

**32602100079**

***MAJORING OF INFORMATICS ENGINEERING  
INDUSTRIAL TECHNOLOGY FACULTY  
SULTAN AGUNG ISLAMIC UNIVERSITY  
SEMARANG  
2025***

**LEMBAR PENGESAHAN  
TUGAS AKHIR**

**PENERAPAN MODEL XGBOOST UNTUK SISTEM REKOMENDASI  
SAHAM SEKTOR PROPERTI MENGGUNAKAN ANALISIS TEKNIKAL**

**MUHAMMAD IKO WARDHANA  
NIM 32602100079**

Telah dipertahankan di depan tim pengujian ujian sarjana tugas akhir  
Program Studi Teknik Informatika  
Universitas Islam Sultan Agung  
Pada tanggal : 21 Agustus 2025

**TIM PENGUJI UJIAN SARJANA :**

**Andi Riansyah S.T., M.Kom**

NIK. 210616053

(Penguji 1)

**Ghufron S.T., M.Kom**

NIK. 210622056

(Penguji 2)

**Imam Much Ibnu Subroto**

**S.T., M.Sc., Ph.D**

NIK. 210600017

(Pembimbing)

26-08-2025

25-08-2025

26-08-2025

Semarang, 27 Agustus 2025

Mengetahui,

Kaprodi Teknik Informatika  
Universitas Islam Sultan Agung

**Moch. Taufik., ST, MIT**

NIK. 210604034

## **SURAT PERNYATAAN KEASLIAN TUGAS AKHIR**

Yang bertanda tangan dibawah ini :

Nama : Muhammad Iko Wardhana

NIM : 32602100079

Judul Tugas Akhir : PENERAPAN MODEL XGBOOST UNTUK SISTEM  
REKOMENDASI SAHAM SEKTOR PROPERTI  
MENGUNAKAN ANALISIS TEKNIKAL

Dengan bahwa ini saya menyatakan bahwa judul dan isi Tugas Akhir yang saya buat dalam rangka menyelesaikan Pendidikan Strata Satu (S1) Teknik Informatika tersebut adalah asli dan belum pernah diangkat, ditulis ataupun dipublikasikan oleh siapapun baik keseluruhan maupun sebagian, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka, dan apabila di kemudian hari ternyata terbukti bahwa judul Tugas Akhir tersebut pernah diangkat, ditulis ataupun dipublikasikan, maka saya bersedia dikenakan sanksi akademis. Demikian surat pernyataan ini saya buat dengan sadar dan penuh tanggung jawab.

Semarang, 21 Agustus 2025

Yang Menyatakan,



Muhammad Iko Wardhana

## PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Saya yang bertanda tangan dibawah ini :

Nama : Muhammad Iko Wardhana

NIM : 32602100079

Program Studi : Teknik Informatika

Fakultas : Teknologi Industri

Alamat Asal : Semarang

Dengan ini menyatakan Karya Ilmiah berupa Tugas akhir dengan Judul : Penerapan Model XGBoost Untuk Sistem Rekomendasi Saham Sektor Properti Menggunakan Analisis Teknikal.

Menyetujui menjadi hak milik Universitas Islam Sultan Agung serta memberikan Hak bebas Royalti Non-Eksklusif untuk disimpan, dialihmediakan, dikelola dan pangkalan data dan dipublikasikan diinternet dan media lain untuk kepentingan akademis selama tetap menyantumkan nama penulis sebagai pemilik hak cipta. Pernyataan ini saya buat dengan sungguh-sungguh. Apabila dikemudian hari terbukti ada pelanggaran Hak Cipta/Plagiarisme dalam karya ilmiah ini, maka segala bentuk tuntutan hukum yang timbul akan saya tanggung secara pribadi tanpa melibatkan Universitas Islam Sultan agung.

Semarang, 21 Agustus 2025

Yang menyatakan,



Muhammad Iko Wardhana

## KATA PENGANTAR

Puji syukur penulis panjatkan kepada ALLAH SWT, yang telah memberikan rahmat, taufik serta hidayah-Nya, sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul : “Penerapan Model XGBoost Untuk Sistem Rekomendasi Saham Sektor Properti Menggunakan Analisis Teknikal.” ini dengan baik. Dengan penuh rasa hormat, penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Rektor UNISSULA Bapak Prof. Dr. H. Gunarto, SH., M.Hum yang mengizinkan penulis menimba ilmu di kampus ini.
2. Dekan Fakultas Teknologi Industri Ibu Dr. Ir. Hj. Novi Marlyana, S.T., M.T
3. Dosen pembimbing penulis Bapak Imam Much Ibnu Subroto, S.T., M.Sc., Ph.D yang telah memberikan arahan, bimbingan, dan saran yang berarti dalam penyelesaian tugas akhir ini.
4. Seluruh dosen Program Studi Teknik Informatika, Fakultas Teknologi Industri UNISSULA yang telah memberikan ilmunya kepada penulis.
5. Orang tua penulis yang selalu memberikan segala doa, dukungan, biaya, dan motivasi dengan penuh kasih sayang sehingga penulis dapat menyelesaikan tugas akhir ini dengan baik.
6. Teman-teman seperjuangan atas kebersamaanya yang telah bekerja keras serta semangat dalam proses penyelesaian tugas akhir ini.

Semarang, 21 Agustus 2025



Muhammad Iko Wardhana

## DAFTAR ISI

|                                                     |           |
|-----------------------------------------------------|-----------|
| COVER .....                                         | i         |
| LEMBAR PENGESAHAN TUGAS AKHIR .....                 | ii        |
| SURAT PERNYATAAN KEASLIAN TUGAS AKHIR.....          | iii       |
| PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH ..... | iv        |
| KATA PENGANTAR.....                                 | v         |
| DAFTAR ISI.....                                     | vi        |
| DAFTAR TABEL .....                                  | viii      |
| DAFTAR GAMBAR .....                                 | ix        |
| ABSTRAK .....                                       | x         |
| <b>BAB I PENDAHULUAN.....</b>                       | <b>1</b>  |
| 1.1 Latar Belakang .....                            | 1         |
| 1.2 Rumusan Masalah .....                           | 3         |
| 1.3 Pembatasan Masalah .....                        | 3         |
| 1.4 Tujuan Tugas Akhir.....                         | 4         |
| 1.5 Manfaat .....                                   | 4         |
| 1.6 Sistematika Penulisan.....                      | 4         |
| <b>BAB II TINJAUAN PUSTAKA DAN DASAR TEORI.....</b> | <b>6</b>  |
| 2.1 Tinjauan Pustaka .....                          | 6         |
| 2.2 Dasar Teori .....                               | 8         |
| 2.2.1 Saham.....                                    | 8         |
| 2.2.2 Sistem Rekomendasi.....                       | 9         |
| 2.2.3 XGBOOST .....                                 | 10        |
| 2.2.4 Analisis Teknikal.....                        | 13        |
| <b>BAB III METODOLOGI PENELITIAN .....</b>          | <b>15</b> |
| 3.1 Metode Penelitian.....                          | 15        |
| 3.1.1 Studi Literatur .....                         | 15        |
| 3.1.2 Pengumpulan Data .....                        | 15        |
| 3.1.3 Perancangan Model.....                        | 17        |
| 3.1.4 Tahapan Perancangan Model .....               | 18        |
| 3.2 Analisis Kebutuhan .....                        | 24        |
| 3.3 Analisis Sistem.....                            | 27        |
| 3.4 Perancangan User.....                           | 29        |
| 3.4.1 Halaman Awal Sistem Rekomendasi.....          | 29        |
| 3.4.2 Halaman Hasil Sistem Rekomendasi .....        | 30        |
| <b>BAB IV HASIL ANALISIS PENELITIAN .....</b>       | <b>32</b> |
| 4.1 Hasil Penelitian .....                          | 32        |
| 4.1.1 Pengumpulan Data .....                        | 32        |
| 4.1.2 Tahapan Perancangan Model .....               | 35        |
| 4.2 Hasil Implementasi.....                         | 46        |
| 4.2.1 Halaman Awal Sistem Rekomendasi.....          | 47        |
| 4.2.2 Halaman Hasil Sistem Rekomendasi .....        | 48        |
| <b>BAB V KESIMPULAN DAN SARAN .....</b>             | <b>52</b> |
| 5.1 Kesimpulan .....                                | 52        |
| 5.2 Saran.....                                      | 53        |
| <b>DAFTAR PUSTAKA.....</b>                          | <b>54</b> |

|               |    |
|---------------|----|
| LAMPIRAN..... | 56 |
|---------------|----|



## DAFTAR TABEL

|                                                  |    |
|--------------------------------------------------|----|
| Tabel 3. 1 Tabel YahooFinance .....              | 15 |
| Tabel 3. 2 <i>Dataset</i> .....                  | 16 |
| Tabel 4. 1 Data Hasil <i>Scraping</i> .....      | 32 |
| Tabel 4. 2 <i>Dataset</i> .....                  | 33 |
| Tabel 4. 3 <i>Missing Value</i> .....            | 36 |
| Tabel 4. 4 Menampilkan <i>Split Data</i> .....   | 37 |
| Tabel 4. 5 Menampilkan <i>Oversampling</i> ..... | 38 |
| Tabel 4. 6 Konfigurasi Hiperparameter.....       | 40 |
| Tabel 4. 7 Hasil Konfigurasi Terbaik .....       | 41 |
| Tabel 4. 8 Evaluasi Model.....                   | 42 |
| Tabel 4. 9 Evaluasi Pengujian Model.....         | 42 |
| Tabel 4. 10 Hasil Rekomendasi.....               | 51 |



## DAFTAR GAMBAR

|                                                             |    |
|-------------------------------------------------------------|----|
| Gambar 2. 1 Arsitektur XGBoost.....                         | 10 |
| Gambar 3. 1 <i>Flowchart</i> Perancangan Model .....        | 18 |
| Gambar 3. 2 <i>Flowchart Preprocessing</i> .....            | 19 |
| Gambar 3. 3 <i>Flowchart</i> kerja model XGBoost .....      | 21 |
| Gambar 3. 4 <i>Flowchart</i> Alur Kerja Sistem .....        | 28 |
| Gambar 3. 5 Rancangan Halaman Awal.....                     | 30 |
| Gambar 3. 6 Rancangan Halaman Hasil .....                   | 31 |
| Gambar 4. 1 Korelasi Data.....                              | 39 |
| Gambar 4. 2 <i>Confusion Matrix</i> .....                   | 44 |
| Gambar 4. 3 Halaman Awal <i>Recommendation Result</i> ..... | 45 |
| Gambar 4. 4 Hasil <i>Recommendation Result</i> .....        | 46 |
| Gambar 4. 5 Tampilan Awal Halaman Rekomendasi.....          | 47 |
| Gambar 4. 6 Tampilan Konfigurasi Sistem.....                | 48 |
| Gambar 4. 7 Sinyal <i>Hold</i> .....                        | 48 |
| Gambar 4. 8 Sinyal <i>Buy</i> .....                         | 49 |
| Gambar 4. 9 Sinyal <i>Sell</i> .....                        | 50 |



## ABSTRAK

Peningkatan jumlah investor di pasar modal Indonesia mencerminkan meningkatnya ketertarikan terhadap investasi saham, khususnya pada sektor properti. Sektor ini memiliki prospek pertumbuhan jangka panjang yang menjanjikan, meskipun di sisi lain juga memiliki karakteristik volatilitas harga yang tinggi. Untuk menjawab tantangan tersebut, penelitian ini mengembangkan sebuah sistem rekomendasi saham berbasis analisis teknikal, dengan menggunakan algoritma *Extreme Gradient Boosting* (XGBoost). Data yang digunakan berupa data historis harga harian saham dari Bursa Efek Indonesia, diperoleh dari Yahoo Finance dengan cakupan periode Januari 2020 hingga Juli 2025. Proses rekayasa fitur menghasilkan 19 indikator teknikal, dan pelabelan data dilakukan ke dalam tiga kategori kelas (BELI, TAHAN, JUAL) berdasarkan wawancara dengan pakar pasar saham. Model dilatih menggunakan data yang telah diseimbangkan melalui metode SMOTE dan dioptimalkan dengan algoritma *Optuna* untuk memperoleh konfigurasi hiperparameter terbaik. Hasil evaluasi menunjukkan bahwa model XGBoost mampu mencapai tingkat akurasi sebesar 88,64% pada data pengujian, dengan performa presisi dan *f1-score* tinggi untuk kelas BELI dan JUAL. Namun demikian, performa pada kelas TAHAN masih menghadapi kendala akibat ambiguitas label dan keterbatasan data. Secara keseluruhan, sistem rekomendasi yang dikembangkan terbukti efektif dalam memberikan rekomendasi saham yang akurat dan objektif, khususnya untuk mendukung pengambilan keputusan berbasis teknikal di tengah kondisi pasar yang fluktuatif.

**Kata Kunci:** XGBoost, Sistem Rekomendasi, Analisis Teknikal, Sektor Properti

## ABSTRACT

*The increasing number of investors in the Indonesian capital market reflects a growing interest in stock-based investments, particularly in the property sector. This sector offers promising long-term growth prospects but is also characterized by high price volatility. To address these challenges, this study develops a stock recommendation system based on technical analysis, utilizing the Extreme Gradient Boosting (XGBoost) algorithm. The data used consists of historical daily stock prices from the Indonesia Stock Exchange, retrieved from Yahoo Finance, covering the period from January 2020 to July 2025. Feature engineering produced 19 technical indicators, and data labeling was conducted into three categories (BUY, HOLD, SELL) based on expert interviews. The model was trained using balanced data through the SMOTE method and optimized with the Optuna algorithm to determine the best hyperparameter configuration. Evaluation results show that the XGBoost model achieved an accuracy of 88.64% on the test data, with high precision and f1-score performance for the BUY and SELL classes. However, the HOLD class showed lower performance due to label ambiguity and limited data. Overall, the developed recommendation system proves effective in generating accurate and objective stock recommendations, particularly for supporting technical based decision making in a volatile market environment.*

**Keywords:** XGBoost, Recommendation System, Technical Analysis, Property Sector

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan aktivitas pasar saham di Indonesia menunjukkan tren yang semakin meningkat. Berdasarkan data Bursa Efek Indonesia (BEI), pada akhir tahun 2021 jumlah investor saham tercatat mencapai lebih dari 6 juta dengan *Single Investor Identification* (SID). Hingga Oktober 2024, jumlah tersebut telah melonjak drastis hingga melampaui 14 juta SID. Peningkatan investor pada investasi saham karena menawarkan keuntungan yang lebih besar dibandingkan investasi lainnya seperti deposito atau obligasi. Lonjakan ini menjadi indikasi meningkatnya literasi dan inklusi keuangan masyarakat Indonesia dalam sektor pasar modal. Namun, investasi saham memiliki kompleksitas tinggi yang memengaruhi proses pemilihan saham, seperti perkembangan perusahaan, kondisi industri, dan pengaruh investor asing. Kompleksitas tersebut menciptakan kebutuhan akan sistem rekomendasi investasi yang dapat membantu investor, terutama pemula, dalam mengambil keputusan.

Salah satu sektor yang sering menjadi pilihan investor adalah sektor properti. Sektor ini memiliki karakteristik yang relatif stabil dalam jangka panjang, menjadikannya sebagai instrumen investasi yang menarik. Menurut (Huda dan Satrio, 2020) dalam perusahaan di sektor Properti dan *Real Estate* dinilai peranan yang sangat penting dalam menyumbang perekonomian negara. Dalam sektor ini, diyakini bahwa hal tersebut dapat menjadi tolak ukur penting dalam pertumbuhan ekonomi saat ini. Namun, volatilitas harga saham di sektor properti tetap menjadi tantangan tersendiri bagi investor dalam menentukan waktu yang tepat untuk masuk dan keluar dari pasar. Oleh karena itu, diperlukan suatu sistem rekomendasi yang dapat membantu investor dalam pengambilan keputusan investasi berbasis data secara lebih akurat dan objektif.

Dalam praktik investasi saham, terdapat dua pendekatan utama yang digunakan untuk menganalisis kelayakan suatu emiten, yakni analisis *fundamental* dan analisis teknikal. Analisis *fundamental* berfokus pada evaluasi aspek makro dan mikro ekonomi seperti kondisi ekonomi nasional dan kinerja perusahaan. Sementara itu,

analisis teknikal memanfaatkan data historis harga dan volume saham yang divisualisasikan dalam grafik guna mengidentifikasi tren dan memprediksi pergerakan harga di masa depan (Mahendra dkk., 2022). Indikator teknikal merupakan hasil dari perhitungan matematis terhadap data harga saham dalam periode waktu tertentu, yang umumnya meliputi harga pembukaan (*open*), harga tertinggi (*high*), harga terendah (*low*), dan harga penutupan (*close*), dengan harga penutupan sering dijadikan sebagai komponen utama dalam perhitungan. Oleh karena itu, penggunaan berbagai indikator teknikal seperti *Moving Average* (MA), RSI, dan MACD dapat membantu menganalisis tren harga (Uhunmwangho, 2024).

Rekomendasi investasi saham merupakan suatu sistem yang bertujuan untuk memberikan saran atau penilaian terhadap saham tertentu guna membantu investor dalam mengambil keputusan yang lebih tepat, apakah akan membeli, menjual, atau mempertahankan saham tersebut dalam portofolio mereka. Sistem rekomendasi ini berfungsi sebagai alat bantu yang menyaring informasi dan menyederhanakan proses pengambilan keputusan dengan mempertimbangkan berbagai indikator relevan, sehingga dapat meminimalkan potensi kerugian finansial dan meningkatkan peluang keuntungan (Tanto dan Kurniawan, 2022). Dalam konteks pasar saham, kemampuan untuk memperkirakan arah pergerakan harga menjadi sangat krusial, meskipun proses tersebut tidaklah sederhana. Hal ini disebabkan oleh sifat pasar yang sangat fluktuatif dan dinamis, di mana harga saham dipengaruhi oleh banyak faktor seperti kondisi ekonomi, sentimen pasar, berita global, serta faktor teknikal dan *fundamental* perusahaan. Oleh karena itu, sistem rekomendasi yang berbasis data dan analisis yang akurat menjadi semakin penting dalam mendukung pengambilan keputusan investasi yang lebih terinformasi dan strategis.

Seiring dengan pesatnya kemajuan teknologi, penerapan kecerdasan buatan (AI), khususnya algoritma *Machine Learning*, telah menjadi solusi dalam pengembangan sistem prediksi dan rekomendasi saham. Teknik dalam pembelajaran mesin sendiri memungkinkan perilaku cerdas dengan mempelajari metode pembelajaran mendalam yang memanfaatkan jaringan saraf, seperti *Extreme Gradient Boosting* (XGBoost), telah menjadi populer sebagai algoritma

berbasis pohon keputusan yang dirancang khusus untuk mengolah data tabular non-sekuensial dengan tingkat efisiensi dan akurasi yang tinggi. XGBoost membangun model klasifikasi secara bertahap menggunakan pendekatan *ensemble*, di mana setiap pohon yang dibentuk berfungsi untuk mengoreksi kesalahan dari pohon sebelumnya. Algoritma XGBoost (*Extreme Gradient Boosting*) menunjukkan potensi kuat untuk mengintegrasikan data teknikal, guna menghasilkan sistem rekomendasi yang adaptif dan prediktif (Liwei dkk., 2021).

Dengan memanfaatkan algoritma XGBoost yang dikombinasikan dengan pendekatan analisis teknikal, sistem rekomendasi saham yang dibangun pada penelitian ini diharapkan mampu menghasilkan klasifikasi yang lebih komprehensif dan akurat. Sistem ini diharapkan dapat membantu investor, khususnya investor pemula, dalam membuat keputusan investasi di sektor properti dengan lebih percaya diri serta berbasis data yang objektif.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, penulis mengidentifikasi permasalahan yang dibahas yaitu, membangun model XGBoost menjadi sistem rekomendasi saham pada sektor properti dengan pendekatan analisis teknikal ?

## 1.3 Pembatasan Masalah

Batasan masalah dari penulisan proposal ini adalah sebagai berikut:

1. Penelitian ini terbatas pada saham-saham yang terdaftar di Bursa Efek Indonesia, dengan fokus pada sektor properti.
2. Penelitian ini hanya menghasilkan sinyal rekomendasi berupa *Buy*, *Sell*, dan *Hold*.
3. Analisis teknikal dilakukan berdasarkan data historis harga saham, yang mencakup variabel-variabel seperti tanggal (*Date*), harga pembukaan (*Open*), harga tertinggi (*High*), harga terendah (*Low*), harga penutupan (*Close*), volume perdagangan (*Volume*), kode emiten (*Ticker*), serta indikator teknikal seperti *Moving Average* (MA\_5, MA\_10, MA\_20), *RSI* (*Relative Strength Index*), *MACD* (*Moving Average Convergence Divergence*), *MACD\_Signal*,

MACD\_Hist, OBV (*On-Balance Volume*), BB\_Width (*Bollinger Band Width*), dan ATR\_14 (*Average True Range*).

#### 1.4 Tujuan Tugas Akhir

Tugas Akhir ini bertujuan untuk mengimplementasikan model XGBoost dalam perancangan sistem rekomendasi saham pada sektor properti, dengan memanfaatkan pendekatan analisis teknikal. Sistem yang dikembangkan diharapkan mampu menghasilkan rekomendasi saham yang lebih tepat.

#### 1.5 Manfaat

Manfaat Tugas Akhir ini dapat memberikan kontribusi dalam membantu investor pemula memperoleh wawasan tambahan dalam pengambilan keputusan investasi saham. Sistem rekomendasi yang dikembangkan memanfaatkan algoritma XGBoost, yang dikenal luas dalam tugas klasifikasi, serta mengintegrasikan pendekatan analisis teknikal untuk meningkatkan akurasi dan relevansi hasil rekomendasi.

#### 1.6 Sistematika Penulisan

Sistematika penulisan yang akan digunakan oleh penulis dalam sebuah pembuatan laporan tugas akhir sebagai berikut :

##### BAB I : PENDAHULUAN

Pada Bab 1, penulis menjelaskan urgensi dari penelitian yang dilakukan, mulai dari pembuatan latar belakang, membuat rumusan masalah dalam penelitian, membatasi permasalahan yang dibahas, tujuan dan manfaat yang diperoleh, dan sistematika penulisan.

##### BAB II : TINJAUAN PUSTAKA DAN DASAR TEORI

Pada bab 2, penulis membuat dasar teori yang digunakan, serta rujukan dari penelitian terdahulu yang akan digunakan dalam perancangna sistem, dan membantu penulis untuk memahami teori yang berhubungan dengan Saham, Analisis Teknikal, Sistem

Rekomendasi, *Extreme Gradient Boosting* (XGBoost) selama proses penelitian.

**BAB III : METODE PENELITIAN**

Pada bab 3, penulis menjelaskan proses dan langkah-langkah penelitian yang dimulai dari memperoleh dataset hingga proses pembuatan model topik data yang ada.

**BAB IV : HASIL DAN ANALISIS PENELITIAN**

Pada bab 4, penulis menjelaskan hasil penelitian yang telah dilakukan, dimulai dengan menjelaskan hasil akhir sistem, klasifikasi data uji, dan akurasi dari sistem.

**BAB V : KESIMPULAN DAN SARAN**

Pada bab 5, penulis memaparkan kesimpulan dari hasil proses penelitian mulai dari awal sampai akhir.



## BAB II

### TINJAUAN PUSTAKA DAN DASAR TEORI

#### 2.1 Tinjauan Pustaka

Sistem rekomendasi saham adalah sebuah aplikasi yang dirancang untuk memberikan saran terkait saham tertentu berdasarkan analisis pasar saham. Umumnya, kegiatan ini dilakukan oleh analis profesional dari suatu perusahaan sekuritas sebagai bentuk layanan bagi para investor. Dalam menjalankan tugasnya, analis akan mengamati dan mengevaluasi pergerakan saham dengan memanfaatkan berbagai indikator yang tersedia di pasar saham (Tanto dan Kurniawan, 2022). Dalam hal ini, pembuatan sistem rekomendasi saham dimaksudkan untuk mendapatkan sinyal dalam memilih sebuah saham yang sudah di analisis dengan indikator teknikal.

Model untuk memprediksi harga saham dibagi menjadi dua jenis utama, yaitu model linier dan model nonlinier. Model linier mencakup metode seperti *Autoregressive* (AR), *Moving Average* (MA), *Autoregressive Moving Average* (ARMA), serta *Autoregressive Integrated Moving Average* (ARIMA), yang telah banyak digunakan dalam analisis deret waktu (Oukhouya dkk., 2024). Di sisi lain, model nonlinier meliputi teknik seperti *Autoregressive Conditional Heteroskedasticity* (ARCH), *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH), serta algoritma *machine learning*. Kelemahan utama dari model linier adalah ketidakmampuannya dalam merepresentasikan dinamika kompleks pada data pasar saham, sehingga pendekatan nonlinier menjadi alternatif yang lebih adaptif dalam menangani variabilitas tersebut.

Algoritma XGBoost memiliki performa yang cukup baik dalam melakukan prediksi harga saham, khususnya pada saham Bank BCA. Penggunaan teknik rekayasa fitur berbasis indikator teknikal serta penyetelan parameter hiper (*hyperparameter tuning*) secara tepat terbukti dapat menurunkan tingkat kesalahan prediksi, ditunjukkan dengan nilai MAPE (*Mean Absolute Percentage Error*) sebesar 4,01% (Jange, 2022). Meskipun demikian, penambahan jumlah indikator teknikal dan parameter yang disesuaikan akan berdampak pada peningkatan waktu pemrosesan model. Hasil ini menunjukkan bahwa XGBoost mampu memberikan

hasil prediksi yang lebih akurat dibandingkan dengan metode Prophet, yang dalam penelitian terdahulu tidak mengintegrasikan indikator teknikal dalam rekayasa fitur.

Penelitian yang menggunakan algoritma XGBoost untuk melakukan klasifikasi terhadap kualitas air berdasarkan 2400 data. Dataset tersebut dibagi menjadi data latih dan data uji dengan rasio 80:20, di mana 80% digunakan untuk pelatihan model dan 20% sisanya untuk pengujian. Berdasarkan hasil evaluasi yang dilakukan menggunakan confusion matrix, diperoleh nilai akurasi sebesar 82,29%, precision 78,62%, recall 85,90%, dan f1-score sebesar 82,09%. Dengan hasil tersebut, dapat disimpulkan bahwa algoritma XGBoost yang dikombinasikan dengan teknik oversampling dan tuning parameter menunjukkan performa klasifikasi yang cukup baik untuk diterapkan pada kasus klasifikasi kelayakan air minum (Maulana dkk., 2023).

Penelitian lain yang mengukur performa XGBoost dan Random Forest dalam memprediksi nilai target pada data uji, yang terlebih dahulu harus diuji menggunakan *testing set* sebagai input. Hasil evaluasi menggunakan *confusion matrix* menunjukkan bahwa XGBoost memiliki performa yang lebih unggul dibanding Random Forest. Model XGBoost mampu menghasilkan prediksi yang konsisten dengan nilai *F1-score* sempurna (1.0) pada berbagai kondisi data, menunjukkan keandalannya dalam klasifikasi keputusan kredit. Sebaliknya, Random Forest menunjukkan kinerja yang kurang stabil terutama pada data kecil dan tidak seimbang, dengan *F1-score* hanya sebesar 0.700 pada dataset berjumlah 10.000, dan meningkat menjadi 0.995 pada data berjumlah 100.000. Meskipun begitu, Random Forest tetap layak digunakan dalam klasifikasi keputusan kredit, terutama jika didukung dengan strategi penyeimbangan data seperti *oversampling* atau *undersampling* (Dachi dan Sitompul, 2023). Oleh karena itu, penanganan ketidakseimbangan kelas dalam dataset menjadi penting dan perlu dieksplorasi lebih lanjut dalam proses *preprocessing* data.

Berdasarkan penelitian yang telah dilakukan sebelumnya, model XGBoost telah menunjukkan kemampuan yang baik dalam menentukan prediksi harga saham kedepannya serta memberikan dukungan dalam pengambilan keputusan investasi.

Sehingga pada penelitian ini bertujuan untuk merancang sistem rekomendasi saham menggunakan XGBoost berbasis analisis teknikal.

## 2.2 Dasar Teori

### 2.2.1 Saham

Saham merupakan salah satu instrumen investasi yang setiap kepemilikan saham berhak mendapatkan porsi kepemilikan di sebuah perusahaan berdasarkan jumlah lembar saham yang dimiliki, keuntungan saham bisa berupa dividen atau pembagian keuntungan yang diperoleh dari keuntungan perusahaan tersebut dan keuntungan lainnya yaitu *capital gain* atau selisih antara harga beli dan harga jual. Sebagai contoh seorang investor membeli saham perusahaan A dengan harga per sahamnya RP 2.000 kemudian menjualnya saat harganya naik menjadi Rp 4.000 per saham yang berarti investor tersebut memperoleh keuntungan sebanyak 50% dari saham tersebut.

Pasar saham saat ini menunjukkan harga yang fluktuasi karena dipengaruhi oleh dinamika permintaan dan penawaran di pasar modal. Oleh karena itu prediksi saham menjadi salah satu metode bagi semua investor, banyak analis mengembangkan strategi dan algoritma untuk memprediksi harga saham dan membuat keputusan dalam membeli saham (Mohan dkk., 2022). Beragam indikator, teknik pembelajaran mesin, dan metode *deep learning* seperti LSTM, XGBoost, dan GRU telah digunakan untuk memprediksi nilai dan harga saham maupun data keuangan. Jika dibandingkan dengan metode dan indikator konvensional, pendekatan jaringan saraf buatan modern terbukti memiliki performa yang lebih unggul.

Saham juga menjadi jalan keluar bagi perusahaan yang memerlukan modal untuk berbagai kepentingan, misalnya perluasan bisnis, menambah kapasitas kerja, melunasi utang, dan sebagainya (Janaina dan Yudiantoro, 2022). Di Indonesia, sejumlah perusahaan daerah telah menjual sahamnya kepada publik agar dapat dimiliki oleh masyarakat luas. Langkah ini tidak hanya bertujuan untuk memperluas kepemilikan, tetapi juga sebagai bentuk transparansi agar perusahaan daerah lebih terbuka dan mengurangi potensi praktik korupsi di dalamnya.

### 2.2.2 Sistem Rekomendasi

Sistem rekomendasi merupakan perangkat lunak atau metode yang menghasilkan keputusan berupa *item-item* spesifik yang mungkin menarik bagi suatu pengguna tertentu. Dalam sistem rekomendasi terbagi menjadi dua pendekatan dalam memberikan usulan, yaitu sistem rekomendasi *personalized* dan *non-personalized* (Arfisko dan Wibowo, 2022). Pendekatan *non-personalized* memberikan usulan dengan mengevaluasi keseluruhan data dalam sistem secara sekaligus tanpa mempertimbangkan preferensi pengguna, Sedangkan *personalized* menghasilkan usulan dengan mempertimbangkan preferensi setiap pengguna.

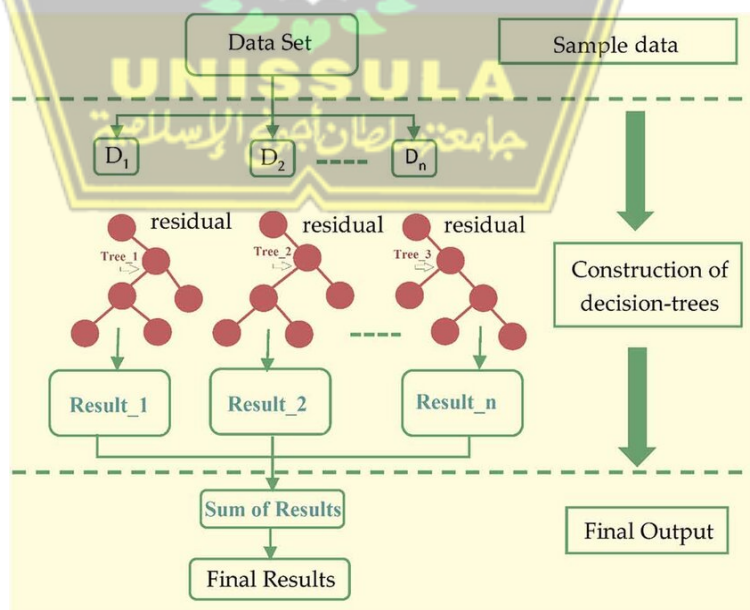
Pada sistem rekomendasi desain menjadi tahap perantara yang penting dalam menentukan kebutuhan pengguna dan spesifikasi, karena ini merupakan sistem rekomendasi yang datanya berasal dari penggunanya itu sendiri. Kesalahan dalam menginput data juga bisa menyebabkan hasil akurasi yang kurang baik dan akan memberikan pengalaman buruk ke pengguna, apalagi nantinya sistem ini akan digunakan oleh investor awam yang kebanyakan tidak memiliki ilmu investasi yang cukup baik. Berdasarkan hal tersebut, menjadi pertimbangan dalam memahami kebutuhan dan preferensi pengguna, yang kemudian digunakan menyusun sistem rekomendasi yang lebih relevan. Sistem ini nantinya akan memberikan penjelasan dalam memperoleh data dan perbersihan data sebelum di input pada aplikasi, sistem ini juga memanfaatkan riwayat pencarian, dan data terkait lainnya untuk menghasilkan rekomendasi yang sesuai.

Tujuan dari sistem rekomendasi bukan hanya memberikan rekomendasi atau memberikan peringkat pada suatu item. Pada intinya sistem rekomendasi menjadi alat untuk membantu pengguna dalam memperoleh keputusan yang lebih baik dari sebelumnya, karena tidak semua orang ahli dalam segala hal layaknya mahasiswa yang sedang di tahap belajar biasanya ingin mengetahui semua hal (Munawar dkk., 2021). Mempertimbangkan segala hal dengan penuh harapan, sistem rekomendasi dapat memberikan pelayanan untuk propaganda yang lebih besar dan lebih baik dengan demikian, hal yang sama dianggap kurang dieksplorasi, sementara itu sistem rekomendasi memiliki masa depan yang berkembang pesat dengan

menambahkan variabel variabel dalam membantu memberikan keputusan yang lebih optimal.

### 2.2.3 XGBOOST

XGBoost merupakan sebuah kerangka kerja dalam pembelajaran mesin yang memanfaatkan pemrosesan paralel guna mencapai efisiensi komputasi, fleksibilitas, serta portabilitas yang tinggi. Algoritma ini merupakan pengembangan lanjutan dari metode *gradient boosting decision tree* (GBDT) yang dioptimalkan untuk kecepatan dan kinerja tinggi. XGBoost dibangun dengan dasar algoritma *boosting*, yang bertujuan untuk meminimalkan fungsi objektif melalui teknik optimasi, seperti penurunan *gradient*, dan dilengkapi dengan komponen regularisasi guna mengurangi risiko *overfitting*. XGboost juga memiliki pengertian lain berupa implementasi dari metode *Gradient Boosting Machines* (GBM) yang dikenal memiliki performa tinggi dalam pembelajaran terawasi (*supervised learning*). Algoritma ini dapat diaplikasikan baik pada permasalahan regresi maupun klasifikasi. XGBoost menjadi pilihan utama di kalangan data *scientist* karena kemampuannya mengeksekusi proses pelatihan model dengan sangat cepat, bahkan di luar lingkungan komputasi inti (Ibrahim Ahmed Osman dkk., 2021). Pada gambar 2.1 menjelaskan arsitektur dari model XGBoost.



Gambar 2. 1 Arsitektur XGBoost

Gambar 2.1 menjelaskan bahwa model XGBoost dibangun secara bertahap (*sekuensial*), di mana setiap pohon keputusan (*Tree 1* hingga *Tree n*) dilatih untuk memprediksi *residual* atau kesalahan prediksi dari model kumulatif sebelumnya. Artinya, setiap pohon baru berfungsi memperbaiki kesalahan dari gabungan pohon-pohon sebelumnya. Proses ini berulang secara iteratif hingga jumlah pohon yang ditentukan tercapai. Hasil akhir diperoleh dengan menjumlahkan prediksi semua pohon (*Sum of Results*), sehingga menghasilkan model prediktif yang lebih akurat. Pendekatan ini merupakan inti dari metode *boosting*, yaitu meningkatkan performa model secara progresif melalui perbaikan bertahap.

Formulasi matematis XGBoost dijelaskan melalui komponen berikut :

$$obj^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) + constant \quad (1)$$

Keterangan :

$n$  : Jumlah total data latih (training samples)

$y_i$  : Label atau nilai aktual dari data ke- $i$

$\hat{y}_i^{(t-1)}$  : Hasil prediksi kumulatif hingga iterasi ke  $t - 1$

$f_t(x_i)$  : *Output* dari pohon ke- $t$  yang ditambahkan di iterasi ke -  $t$

$l(y_i, \hat{y}_i)$  : Fungsi kerugian (*loss*) antara nilai aktual dan prediksi

$\Omega(f_t)$  : Fungsi regularisasi yang menghitung kompleksitas pohon  $f_t$

$constant$  : nilai tetap yang tidak berpengaruh terhadap proses optimasi (biasanya diabaikan dalam turunan)

Fungsi objektif diatas, terdiri dari dua komponen utama, yaitu fungsi kerugian (*loss function*) dan fungsi regularisasi (*regularization term*). Fungsi kerugian berfungsi untuk mengukur perbedaan antara nilai aktual  $y_i$  dan hasil prediksi kumulatif model pada iterasi ke- $t$ , sedangkan fungsi regularisasi digunakan untuk mengontrol kompleksitas struktur pohon keputusan agar dapat menghindari *overfitting*. Selain itu, terdapat penambahan konstanta dalam formulasi fungsi objektif, yang bertujuan untuk menyederhanakan perhitungan ekspansi *Taylor* saat proses optimasi dilakukan. Nilai konstanta ini tidak memengaruhi hasil turunan pertama maupun kedua, sehingga tidak berdampak pada proses minimisasi fungsi objektif itu sendiri.

Untuk menilai sejauh mana model memberikan hasil yang akurat dan efektif, dilakukan evaluasi menggunakan sejumlah metrik kinerja, yaitu *Precision*, *Recall*, dan *F1-Score*. Ketiga metrik ini digunakan untuk mengukur keseimbangan antara prediksi yang benar dan kesalahan klasifikasi, terutama dalam konteks data yang tidak seimbang. *Precision* menilai ketepatan model dalam memprediksi kelas positif, *Recall* mengukur kemampuan model dalam menangkap seluruh instance dari kelas positif, sedangkan *F1-Score* merupakan harmonisasi antara keduanya. Selain itu, evaluasi dilakukan dengan membandingkan performa model pada data pelatihan dan data pengujian guna mengetahui apakah model mengalami *overfitting* atau *underfitting*. Analisis ini penting agar model yang dikembangkan tidak hanya unggul di data pelatihan, tetapi juga mampu melakukan generalisasi dengan baik terhadap data baru.

Salah satu pendekatan yang umum digunakan untuk mengukur performa model klasifikasi adalah *confusion matrix*. *Matrix* ini memungkinkan perhitungan berbagai indikator evaluasi yang mencerminkan kapabilitas model dalam membedakan setiap kelas target dengan benar.

Beberapa komponen penting dalam *confusion matrix* meliputi:

- *True Positive* (TP): Klasifikasi benar ketika model mengklasifikasikan data sebagai positif dan label aslinya juga positif.
- *True Negative* (TN): Klasifikasi benar ketika model mengklasifikasikan data sebagai negatif dan label aslinya juga negatif.
- *False Positive* (FP): Klasifikasi salah ketika model mengklasifikasikan data sebagai positif padahal label aslinya negatif.
- *False Negative* (FN): Klasifikasi salah ketika model mengklasifikasikan data sebagai negatif padahal label aslinya positif.

Dari keempat komponen tersebut, sejumlah metrik performa utama dapat dihitung sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1\ score = \frac{2 \times (precision \times recall)}{precision + recall} \quad (5)$$

#### 2.2.4 Analisis Teknikal

Analisis teknikal merupakan pendekatan analitis yang berfokus pada pengamatan pola pergerakan harga saham dalam rentang waktu tertentu. Metode ini memanfaatkan data historis pada sebuah saham, serta menggunakan grafik sebagai instrumen utama untuk mengidentifikasi tren dan pola pasar (Octasyilva, 2022). Tujuan utamanya adalah untuk memproyeksikan pergerakan harga di masa mendatang guna mendukung pengambilan keputusan investasi yang tepat. Analisis teknikal berasumsi bahwa semua informasi pasar tercermin dalam harga, dan bahwa harga bergerak dalam tren yang cenderung berulang akibat perilaku psikologis pelaku pasar.

Berdasarkan *The Dow Theory* yang di perkenalkan oleh Charles H. Dow pada tahun 1897, menjadi kerangka awal dalam memahami dinamika tren harga di pasar modal. Dow berasumsi bahwa harga saham sudah mencerminkan semua informasi yang tersedia, baik itu berita ekonomi, politik, atau faktor lain. Dow menjelaskan bahwa pasar memiliki tiga jenis tren, yaitu tren utama (jangka panjang), tren sekunder (jangka menengah), dan tren minor (jangka pendek). Selain itu, setiap tren utama terdiri atas tiga fase, yakni fase akumulasi, fase partisipasi publik, dan fase distribusi. Teori ini juga menyatakan bahwa tren harga akan terus berlangsung sampai terdapat sinyal pasti bahwa tren tersebut telah berbalik arah. Validitas tren dalam teori ini diperkuat dengan adanya konfirmasi dari *volume* perdagangan dan keselarasan pergerakan antara indeks-indeks pasar (Irahadi dkk., 2022).

Dalam analisis teknikal, seorang investor perlu memperhatikan pola pergerakan harga sebagai dasar dalam pengambilan keputusan investasi. Oleh

karena itu, data historis harga saham biasanya disajikan dalam bentuk grafik untuk mempermudah interpretasi visual. Beberapa jenis grafik yang umum digunakan antara lain *line chart* (grafik garis), *bar chart* (grafik batang), serta *candlestick chart* (grafik lilin), yang masing-masing menyajikan informasi pergerakan harga secara berbeda namun saling melengkapi (Irahadi dkk., 2022).

Salah satu indikator teknikal yang banyak digunakan adalah *Moving Average* (MA), yaitu metode perataan data harga saham dalam periode waktu tertentu yang hasilnya ditampilkan dalam bentuk garis pada grafik. Garis MA berfungsi untuk mengidentifikasi tren harga, memberikan sinyal awal terhadap munculnya tren baru, serta sebagai konfirmasi terhadap potensi pembalikan arah (*reversal*). Indikator lainnya adalah *Relative Strength Index* (RSI) berperan dalam mengatasi fluktuasi harga yang tidak menentu dengan mengukur kekuatan relatif pergerakan harga dalam suatu periode tertentu, sehingga dapat mengidentifikasi kondisi pasar yang berada pada posisi *overbought* atau *oversold*.

Sementara itu, indikator *Moving Average Convergence Divergence* (MACD) digunakan untuk mengevaluasi hubungan antara dua *Exponential Moving Average* (EMA) dengan periode yang berbeda. Perhitungan MACD umumnya dilakukan dengan mengurangkan EMA 26 hari dari EMA 12 hari. Indikator ini banyak digunakan untuk mendeteksi perubahan momentum harga dan memberikan sinyal beli atau jual berdasarkan pola perpotongan antar garis (Farild dkk., 2023).

## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Metode Penelitian

Penelitian ini menerapkan metode *Extreme Gradient Boosting* (XGBoost) yang dikombinasikan dengan pendekatan analisis teknikal saham. Algoritma tersebut diintegrasikan dengan indikator teknikal guna membangun sebuah sistem rekomendasi yang mampu menghasilkan sinyal investasi berupa *buy*, *sell*, dan *hold* untuk suatu saham. Adapun tahapan yang dilalui dalam penelitian ini mencakup langkah-langkah sebagai berikut:

##### 3.1.1 Studi Literatur

Penelitian ini akan dilakukan tinjauan terhadap beberapa makalah, jurnal, artikel, tesis, dan skripsi terdahulu akan diulas selain mengunjungi berbagai situs web seperti medium dan github. Tujuan dari tinjauan merupakan agar bisa mempelajari teori di balik konsep sistem rekomendasi saham dengan XGBoost menggunakan analisis teknikal.

##### 3.1.2 Pengumpulan Data

Tahap ini merupakan langkah awal dalam pengembangan model, di mana data dikumpulkan dan disusun menjadi sebuah dataset yang bersumber dari Yahoo Finance. Data yang dikumpulkan mencakup histori pergerakan saham harian. Pada saham dari sektor properti yang terdaftar di Bursa Efek Indonesia, data di ambil dari tahun 2020 hingga tahun 2025. Data yang sudah berhasil di ambil akan disimpan kedalam file dengan format CSV (*comma-separated values*). Pada tabel 3.1 menampilkan data yang tersedia pada halaman Yahoo Finance.

Tabel 3. 1 Tabel YahooFinance

| <i>Date</i>  | <i>Open</i> | <i>High</i> | <i>Low</i> | <i>Close</i> | <i>Adj Close</i> | <i>Volume</i> |
|--------------|-------------|-------------|------------|--------------|------------------|---------------|
| May 21, 2025 | 24.600      | 24.775      | 23.425     | 24.575       | 24.575           | 9,900         |
| May 20, 2025 | 24.000      | 24.725      | 23.475     | 24.400       | 24.400           | 8,600         |
| May 19, 2025 | 24.200      | 24.200      | 24.000     | 24.000       | 24.000           | 9,400         |
| May 16, 2025 | 24.250      | 24.250      | 24.250     | 24.250       | 24.250           | 7,600         |
| May 14, 2025 | 24.250      | 24.275      | 24.000     | 24.275       | 24.275           | 18,200        |
| May 9, 2025  | 24.225      | 24.275      | 24.000     | 24.275       | 24.275           | 7,900         |
| May 21, 2025 | 24.300      | 24.300      | 24.275     | 24.300       | 24.300           | 7,700         |

Tabel 3.1 menunjukkan data historis saham yang tersedia di situs Yahoo Finance. Dalam gambar tersebut terlihat bahwa data yang dapat diakses mencakup nilai *open*, *high*, *low*, *close*, dan *adj close*, yang nantinya akan digunakan sebagai data, untuk mengambil data tersebut melalui proses *scraping*. Pada tabel 3.2 merupakan *dataset* yang akan digunakan model.

Tabel 3. 2 *Dataset*

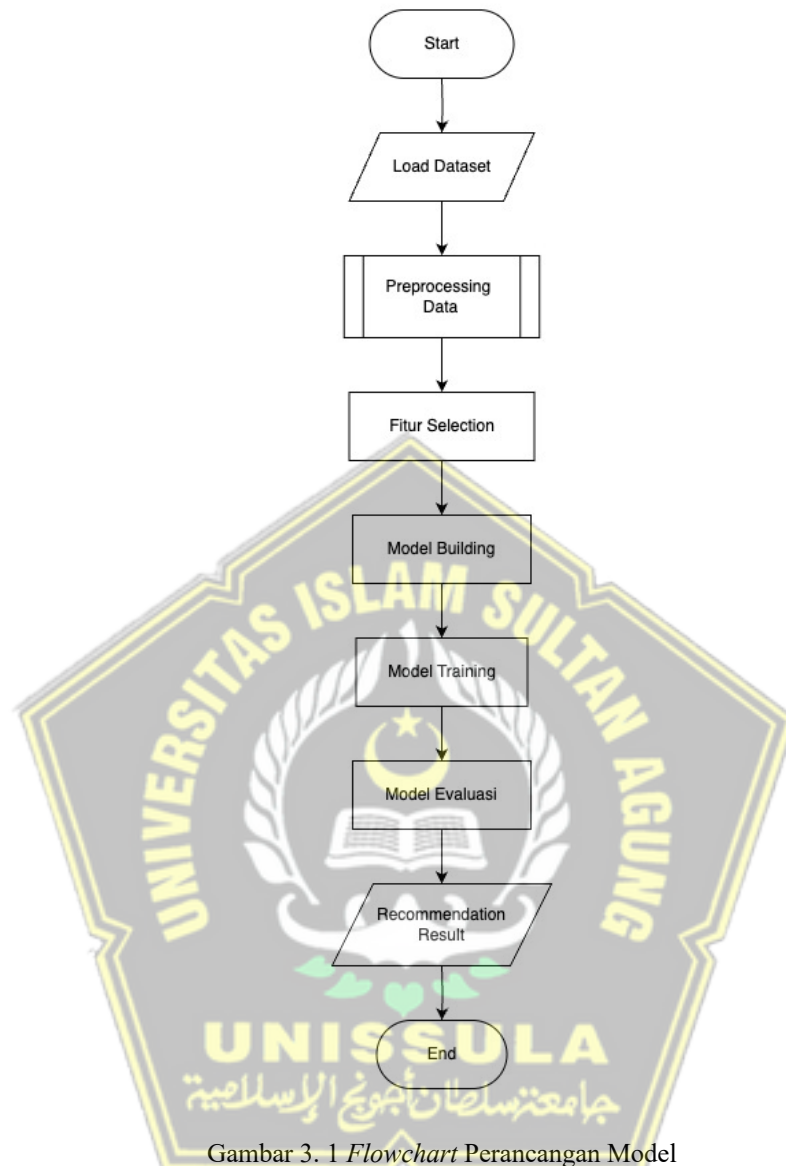
| <i>Date</i>  | <i>Ticker</i>   | <i>Open</i>    | <i>High</i> | <i>Low</i>    | <i>Close</i> | <i>Volume</i> | <i>Ma_5</i> | <i>Ma_10</i> |
|--------------|-----------------|----------------|-------------|---------------|--------------|---------------|-------------|--------------|
| 2025/07/04   | SMDM            | 995            | 995         | 915           | 935          | 247600        | 900         | 900          |
| 2025/07/07   | SMDM            | 935            | 940         | 900           | 915          | 76400         | 940         | 922          |
| 2025/07/08   | SMDM            | 915            | 940         | 875           | 895          | 263900        | 943         | 942          |
| 2025/07/09   | SMDM            | 895            | 920         | 890           | 890          | 145500        | 926         | 945          |
| 2025/07/10   | SMDM            | 890            | 900         | 860           | 885          | 153700        | 904         | 926          |
| 2025/07/11   | SMDM            | 885            | 920         | 865           | 885          | 119400        | 894         | 816          |
| 2025/07/14   | SMDM            | 885            | 885         | 840           | 850          | 397200        | 881         | 910          |
| 2025/07/15   | SMDM            | 885            | 895         | 830           | 875          | 177600        | 877         | 910          |
| 2025/07/16   | SMDM            | 875            | 880         | 825           | 825          | 353100        | 864         | 895          |
| <i>Ma_20</i> | <i>RSI</i>      | <i>MACD</i>    |             | <i>MACD_S</i> |              | <i>MACD_H</i> |             |              |
| 896          | 46.45496687     | -59.03215148   |             | -90.14922991  |              | 31.11707843   |             |              |
| 892          | 45.35048275     | -55.74915038   |             | -83.26921401  |              | 27.52006363   |             |              |
| 890          | 44.21830633     | -54.13712216   |             | -77.44279564  |              | 23.30567348   |             |              |
| 889          | 43.92308365     | -52.65605022   |             | -72.48544655  |              | 19.82939633   |             |              |
| 888          | 43.60952869     | -51.29445768   |             | -68.24724878  |              | 16.9527911    |             |              |
| 887          | 43.60952869     | -49.64313067   |             | -64.52642516  |              | 14.88329449   |             |              |
| 884          | 41.22060318     | -50.57564505   |             | -61.73626913  |              | 11.16062408   |             |              |
| 883          | 46.45496687     | -59.03215148   |             | -90.14922991  |              | 31.11707843   |             |              |
| 881          | 45.35048275     | -55.74915038   |             | -83.26921401  |              | 27.52006363   |             |              |
| <i>OBV</i>   | <i>BB_Width</i> | <i>ATRR_14</i> |             | <i>Signal</i> |              | <i>Label</i>  |             |              |
| 1544947500   | 387.286104      | 114.1110568    |             | BUY           |              | 0             |             |              |
| 1544871100   | 376.6598354     | 104.2295825    |             | BUY           |              | 0             |             |              |
| 1544607200   | 372.8129205     | 98.99897154    |             | BUY           |              | 0             |             |              |
| 1544461700   | 372.6943153     | 89.79910867    |             | BUY           |              | 0             |             |              |
| 1544308000   | 372.5799102     | 83.15922751    |             | SELL          |              | 2             |             |              |
| 1544308000   | 371.7653476     | 79.40466385    |             | SELL          |              | 2             |             |              |
| 1543910800   | 372.8072735     | 74.81737533    |             | SELL          |              | 2             |             |              |
| 1544088400   | 372.3835035     | 73.50839196    |             | SELL          |              | 2             |             |              |
| 1543735300   | 375.5066753     | 71.04060636    |             | SELL          |              | 2             |             |              |

Tabel 3.2 menampilkan data historis saham dengan *ticker* SMDM yang diperoleh dari situs Yahoo Finance. *Dataset* ini mencakup informasi harga (*Open*, *High*, *Low*, *Close*), volume transaksi, serta indikator teknikal yang telah dihitung melalui *feature engineering*, antara lain: *Moving Average* (MA) periode 5, 10, dan 20 hari, *Relative Strength Index* (RSI), *Moving Average Convergence Divergence* (MACD) beserta komponennya (*MACD Signal* dan *MACD Histogram*), *On Balance Volume* (OBV), *Bollinger Band Width* (BB\_Width), dan *Average True Range* periode 14 (ATRr\_14).

Proses pelabelan sinyal dilakukan berdasarkan wawancara dengan pakar saham, Arief Maulana, yang merupakan pendiri komunitas Green Stock Trader. Dalam contoh data terakhir yang ditampilkan, sinyal *HOLD* dan *SELL* dihasilkan sesuai interpretasi indikator teknikal tersebut.

### 3.1.3 Perancangan Model

Pada tahap ini peneliti merencanakan langkah-langkah proses perancangan model. Pada gambar 3.1 menggambarkan alur proses pengembangan sistem rekomendasi saham, yang diawali dengan *load dataset*, dilanjutkan dengan *preprocessing* untuk membersihkan dan menyiapkan data. Tahap berikutnya adalah *feature selection* guna mengidentifikasi dan memilih fitur-fitur yang relevan. Data yang telah disiapkan kemudian digunakan dalam tahap model *building*, model *training* dan model evaluasi. Sebagai hasil akhir dari proses ini, sistem akan menghasilkan rekomendasi keputusan beli, tahan, atau jual berdasarkan klasifikasi model.



Gambar 3. 1 *Flowchart* Perancangan Model

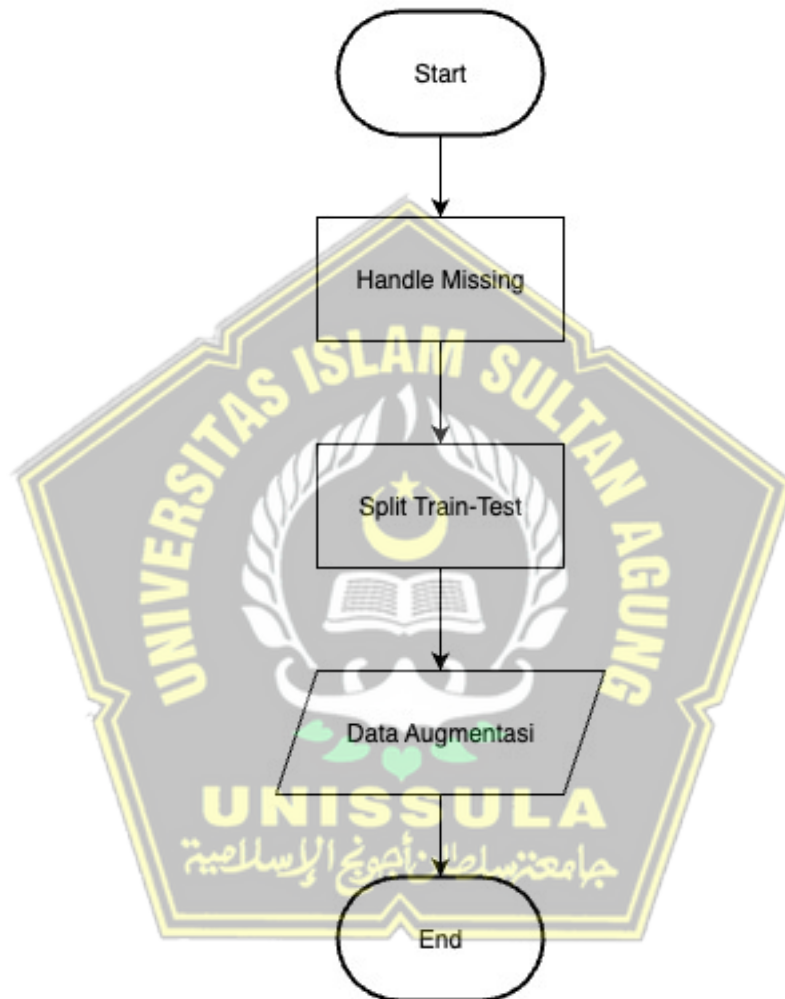
### 3.1.4 Tahapan Perancangan Model

#### 1) *Input Data*

Data yang digunakan sebagai input dalam sistem ini berupa data historis saham yang sudah diberi label oleh pakar ahli. Data tersebut mencakup informasi pergerakan harga saham harian serta indikator teknikal dan disimpan dalam format CSV (*Comma-Separated Values*) agar mudah diolah dalam proses pelatihan dan pengujian model. Format CSV dipilih karena kompatibel dengan berbagai *tools* analisis data dan memungkinkan pemrosesan yang efisien dalam membangun model klasifikasi berbasis *machine learning*.

## 2) *Data Preprocessing*

Pada tahap ini dijelaskan apa saja yang akan dilakukan dalam persiapan data yang akan digunakan model, pada gambar 3.2 menampilkan *flowchart preprocessing* pada *dataset*.



Gambar 3. 2 *Flowchart Preprocessing*

Gambar 3.2 merupakan tahap awal dalam perancangan model dimulai dengan proses *data preprocessing*, yaitu serangkaian langkah untuk mempersiapkan dan membersihkan data agar layak digunakan dalam pelatihan model. Adapun beberapa tahapan utama dalam *preprocessing* data adalah sebagai berikut:

### a) *Handle Missing*

Pada tahap ini, dilakukan proses identifikasi data yang hilang (*missing values*) pada fitur menggunakan fungsi `.isnull()`. Jika ditemukan nilai yang hilang, maka

akan ditangani dengan metode penghapusan baris terkait menggunakan *.dropna()*. Penanganan *missing values* ini penting dilakukan untuk mengurangi potensi bias dalam *dataset*, sehingga model yang dikembangkan dapat menghasilkan klasifikasi yang lebih akurat dan optimal.

b) Split Train - Test

Pada tahap pembagian data, digunakan metode *time series split* untuk memastikan bahwa data uji (*test*) berasal dari periode waktu setelah data latih (*train*), sehingga mencerminkan kondisi riil dalam klasifikasi saham dan mencegah kebocoran informasi dari masa depan ke dalam model. Pendekatan ini penting dalam analisis deret waktu seperti harga saham, di mana urutan kronologis harus dijaga. Berdasarkan hasil pembagian, total data berjumlah 12.575 sampel, dengan 80% *dataset* digunakan untuk pelatihan dan 20% *dataset* sisanya digunakan sebagai data pengujian.

c) Data Augmentation

Pada tahap *augmentasi* data, digunakan metode SMOTE (*Synthetic Minority Over-sampling Technique*) untuk menyeimbangkan distribusi kelas pada data pelatihan. Tidak seperti metode duplikasi data sederhana, SMOTE menghasilkan sampel sintetis dari kelas minoritas melalui pendekatan interpolasi. Proses ini dilakukan dengan memilih satu titik data minoritas, lalu mencari tetangga terdekatnya dalam ruang fitur, dan membentuk data baru di antara kedua titik tersebut. Teknik ini memungkinkan terciptanya variasi data yang lebih alami serta menghindari risiko *overfitting* akibat pengulangan data yang sama.

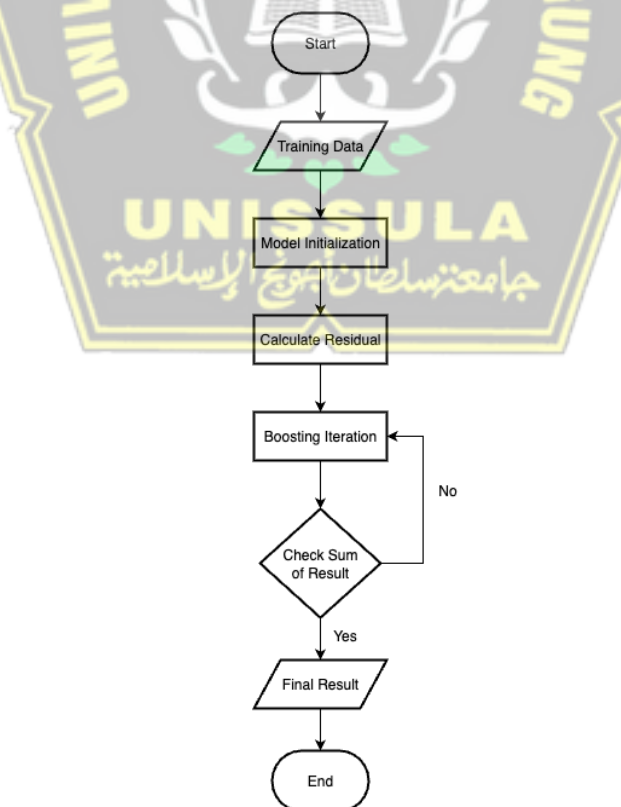
Dengan menyeimbangkan distribusi kelas secara lebih realistis, SMOTE berperan penting dalam mencegah bias model terhadap kelas mayoritas dan secara signifikan meningkatkan kemampuan generalisasi model. Penerapan SMOTE dibatasi hanya pada data pelatihan, sehingga evaluasi model tetap dilakukan terhadap data uji yang orisinal dan representatif. Secara keseluruhan, penerapan SMOTE pada tahap ini bertujuan untuk meningkatkan kinerja model klasifikasi dalam mengenali ketiga sinyal investasi (*BUY*, *HOLD*, dan *SELL*) secara adil dan seimbang.

### 3) *Feature Selection*

*Feature selection* atau seleksi fitur dilakukan dengan tujuan untuk meningkatkan performa dan akurasi model prediktif dengan hanya mempertahankan variabel-variabel yang memiliki pengaruh paling signifikan terhadap hasil klasifikasi. Dengan mengeliminasi fitur yang kurang relevan atau bersifat redundan, proses pelatihan model menjadi lebih efisien dan risiko *overfitting* dapat diminimalkan. Selain itu, seleksi fitur juga membantu meningkatkan interpretabilitas model, karena model menjadi lebih mudah dipahami dan dianalisis oleh peneliti atau pengguna akhir, terutama dalam konteks pengambilan keputusan berbasis data.

### 4) *Model Building*

Dalam proses pengembangan model, digunakan algoritma XGBoost untuk membangun sistem rekomendasi keputusan investasi yang diklasifikasikan ke dalam tiga kategori utama, yaitu beli (*buy*), jual (*sell*), dan tahan (*hold*). Pada gambar 3.3 menjelaskan langkah model yang outputnya berupa hasil klasifikasi.



Gambar 3. 3 *Flowchart* kerja model XGBoost

Langkah alur model XGBoost:

1. **Training Data:** Merupakan kumpulan data yang digunakan untuk melatih model. Data ini berisi fitur (variabel independen) dan label target (variabel dependen). Model akan mempelajari pola dari data ini untuk membuat klasifikasi yang akurat.
2. **Model Initialization:** Pada tahap ini, model awal dibuat. Biasanya model sederhana, seperti memprediksi nilai *log odds* (untuk klasifikasi). Model ini menjadi titik awal untuk proses boosting.
3. **Calculate Residual:** *Residual* dihitung sebagai selisih antara nilai aktual dengan klasifikasi model saat ini. *Residual* menunjukkan seberapa besar kesalahan yang masih perlu diperbaiki oleh model.
4. **Boosting Iteration:** Tahap ini melibatkan pelatihan model lemah baru (misalnya pohon keputusan sederhana) berdasarkan residual. Setiap iterasi bertujuan untuk mengurangi kesalahan model sebelumnya sehingga kinerja keseluruhan meningkat secara bertahap.
5. **Check Sum of Result:** Algoritma akan memeriksa apakah kondisi berhenti sudah terpenuhi. Kondisi ini bisa berupa jumlah iterasi tertentu, atau peningkatan performa yang sudah tidak signifikan.
  - Jika kondisi belum terpenuhi → kembali ke proses *iterasi boosting*.
  - Jika kondisi sudah terpenuhi → lanjut ke hasil akhir
6. **Final Result:** Model akhir diperoleh dengan menggabungkan seluruh model lemah yang dihasilkan pada iterasi boosting. Hasil gabungan ini memberikan kemampuan klasifikasi yang lebih kuat dibandingkan dengan satu model lemah saja,

Model ini dirancang untuk mengenali pola dari data historis saham yang telah melalui tahapan *preprocessing*, seperti pembersihan data, ekstraksi fitur teknikal, dan penyeimbangan distribusi kelas. Untuk mengoptimalkan kinerja model, diterapkan teknik *hyperparameter tuning* dengan bantuan *library Optuna*. *Optuna* secara otomatis mencari kombinasi parameter terbaik melalui proses optimasi berbasis trial, sehingga model yang dihasilkan memiliki performa klasifikasi yang

lebih akurat dan generalisasi yang baik terhadap data yang belum pernah dilihat sebelumnya.

### 5) **Model Training**

Pada tahap training, model XGBoost dioptimalkan menggunakan data yang telah melalui serangkaian proses *preprocessing*, termasuk ekstraksi indikator analisis teknikal seperti *Moving Average*, RSI, dan MACD. Tujuan utama dari tahap ini adalah agar model mampu mempelajari pola historis pergerakan harga saham dan menghasilkan klasifikasi atau rekomendasi berupa sinyal beli (*buy*), jual (*sell*), atau tahan (*hold*). Sebelumnya, proses tuning *hyperparameter* dilakukan menggunakan *library Optuna* untuk memperoleh konfigurasi model terbaik. Setelah parameter optimal diperoleh, model XGBoost kemudian dilatih secara menyeluruh pada data pelatihan untuk membangun kemampuan prediktif yang andal dan akurat. Tahapan ini menjadi inti dari sistem rekomendasi, karena kualitas model sangat bergantung pada proses pelatihan yang baik dan representatif terhadap kondisi pasar yang sebenarnya.

### 6) **Model Evaluation**

Setelah proses pelatihan selesai, tahap berikutnya adalah mengevaluasi kinerja model untuk memastikan hasil klasifikasi tidak menimbulkan kesalahan. Pada penelitian ini, evaluasi dilakukan menggunakan *confusion matrix* yang berfungsi untuk menilai sejauh mana model mampu melakukan klasifikasi dengan benar. Penjelasan lebih rinci mengenai metode evaluasi ini telah disampaikan pada Bab II dalam uraian mengenai algoritma XGBoost.

### 7) **Recommendation Results**

Hasil rekomendasi sistem menampilkan klasifikasi dari model XGBoost dalam bentuk sinyal investasi, yaitu *BUY*, *SELL*, atau *HOLD*. Untuk menjalankan dan memvisualisasikan hasil sistem rekomendasi saham ini, digunakan platform *Gradio* yang dijalankan di lingkungan Google Colab. Pemilihan *Gradio* didasarkan pada kemudahan penggunaan serta kemampuannya dalam membuat antarmuka interaktif secara cepat. Selain itu, hasil rekomendasi ini juga berperan sebagai tahap evaluasi model, guna memastikan performa dan akurasi klasifikasi sudah sesuai harapan

sebelum nantinya sistem diimplementasikan secara penuh pada platform website menggunakan *Streamlit*.

### 3.2 Analisis Kebutuhan

#### 1. *Python*

*Python* adalah bahasa pemrograman tingkat tinggi yang dikenal karena fleksibilitas, keterbacaan sintaksis, serta kemudahan dalam penggunaannya. Dikembangkan oleh Guido van Rossum dan dirilis pertama kali pada tahun 1991, *Python* telah berkembang menjadi salah satu bahasa pemrograman paling populer dan banyak digunakan oleh komunitas pengembang di berbagai bidang, termasuk ilmu data dan kecerdasan buatan. Dalam penelitian ini, digunakan *Python* versi 3.10.1 karena menyediakan berbagai pustaka (*library*) yang lengkap, mendukung analisis data dan pengembangan model *machine learning*, serta bersifat *open source* sehingga dapat digunakan secara bebas tanpa lisensi berbayar.

#### 2. Google Colaboratory

Google Colaboratory, atau yang biasa dikenal sebagai Google Colab, merupakan layanan berbasis cloud dari Google yang memungkinkan pengguna untuk menulis, mengembangkan, dan menjalankan kode program dalam lingkungan komputasi interaktif secara langsung melalui browser. Platform ini sangat mendukung aktivitas penelitian dan pengembangan karena menyediakan akses gratis ke GPU/TPU serta integrasi dengan Google Drive. Dalam penelitian ini, Google Colab digunakan sebagai media utama untuk menulis dan mengeksekusi kode program, sekaligus sebagai sarana pencatatan dan dokumentasi proses kerja secara efisien.

#### 3. *Library NumPy*

*NumPy* (*Numerical Python*) merupakan *library Python* yang dirancang khusus untuk mendukung komputasi numerik dan ilmiah. *Library* ini menyediakan struktur data array multidimensi yang efisien dan berbagai fungsi matematis yang cepat untuk operasi pada array tersebut. Selain itu, *NumPy* juga mencakup berbagai kemampuan dalam aljabar linier, transformasi Fourier, manipulasi matriks, serta

operasi statistik dasar, sehingga menjadikannya fondasi penting dalam pengolahan data dan analisis numerik di berbagai bidang penelitian dan pengembangan.

#### 4. *Library Scikit-Learn*

*Scikit-learn* merupakan *library Python* yang digunakan untuk implementasi berbagai algoritma pembelajaran mesin (*machine learning*) dan analisis data. *Library* ini menyediakan beragam fungsi untuk klasifikasi, regresi, klastering, pengurangan dimensi, pemilihan fitur, dan prapemrosesan data. *Scikit-learn* dibangun di atas *library* numerik seperti *NumPy*, *SciPy*, dan *matplotlib*, sehingga sangat cocok digunakan dalam pengembangan model prediktif dan eksplorasi data secara efisien dan terstruktur.

#### 5. *Library XGBoost*

*XGBoost (Extreme Gradient Boosting)* adalah *library machine learning* yang dioptimalkan untuk kecepatan dan performa tinggi, dan banyak digunakan dalam kompetisi data *science* serta aplikasi industri. *Library* ini merupakan implementasi dari *algoritma boosting* berbasis pohon keputusan yang efisien dan scalable. *XGBoost* dirancang untuk menangani data dalam jumlah besar, menangani *missing value*, serta mendukung berbagai teknik regularisasi untuk menghindari *overfitting*, sehingga sangat cocok untuk tugas klasifikasi, regresi, dan *ranking*. Dalam penelitian ini, *XGBoost* digunakan sebagai model utama untuk membangun sistem rekomendasi saham karena kemampuannya dalam mengolah fitur kompleks dan memberikan hasil klasifikasi yang akurat.

#### 6. *Library yfinance*

*yfinance* adalah *library Python* yang digunakan untuk mengambil data keuangan historis secara langsung dari Yahoo Finance. *Library* ini menyediakan antarmuka sederhana untuk mengunduh data harga saham seperti harga pembukaan (*open*), penutupan (*close*), tertinggi (*high*), terendah (*low*), serta volume perdagangan dalam rentang waktu tertentu. Dalam konteks penelitian ini, *yfinance* dimanfaatkan untuk mengumpulkan data saham sektor properti yang terdaftar di Bursa Efek Indonesia sebagai basis input model. Kelebihan *yfinance* adalah gratis, mudah digunakan, dan mendukung pengambilan data secara otomatis tanpa perlu melakukan scraping manual dari situs keuangan.

### 7. *Library Gradio*

*Gradio* merupakan pustaka *Python* sumber terbuka yang memungkinkan pengembang untuk membangun antarmuka demo atau aplikasi web secara cepat, khususnya untuk model pembelajaran mesin, antarmuka pemrograman aplikasi (API), maupun fungsi *Python* lainnya. Salah satu keunggulan utama *Gradio* adalah kemampuannya dalam menghasilkan tautan berbagi secara otomatis, sehingga pengguna dapat membagikan hasil demo dalam waktu singkat tanpa perlu melakukan proses *deployment* secara manual. Selain itu, *Gradio* dapat dijalankan secara langsung pada Google Colaboratory (Colab), sehingga sangat efisien digunakan pada tahap awal pengembangan sistem. Kemampuan ini menjadikan *Gradio* sebagai platform prototipe yang efektif sebelum aplikasi dilanjutkan ke tahap implementasi lebih lanjut menggunakan *framework* seperti *Streamlit*.

### 8. *Library Optuna*

*Optuna* merupakan *library* yang dirancang untuk mempermudah dan mengotomatiskan proses pencarian hiperparameter terbaik. Dengan melakukan serangkaian uji coba secara otomatis, *Optuna* mampu menemukan kombinasi nilai hiperparameter yang menghasilkan performa model optimal. Saat ini, *Optuna* tersedia untuk digunakan dalam bahasa pemrograman *Python*.

Dalam mekanisme kerjanya, *Optuna* memanfaatkan data riwayat eksperimen sebelumnya untuk menentukan nilai hiperparameter berikutnya yang akan dieksplorasi. Berdasarkan informasi tersebut, sistem memperkirakan area parameter yang potensial untuk menghasilkan kinerja lebih baik dan melakukan pengujian di area tersebut. Setelah memperoleh hasil baru, *Optuna* kembali melakukan estimasi area yang semakin menjanjikan. Proses ini berulang dengan memanfaatkan rekaman uji coba yang telah dilakukan. Secara khusus, metode optimasi yang digunakan *Optuna* adalah algoritma *Bayesian optimization* dengan pendekatan *Tree-structured Parzen Estimator* (TPE).

### 9. Visual Studio Code

Visual Studio Code (VS Code) dipilih sebagai *text* editor dalam pengembangan aplikasi pada penelitian ini karena menawarkan berbagai keunggulan yang mendukung produktivitas pengembangan perangkat lunak. VS Code merupakan

editor kode sumber yang ringan, cepat, dan mendukung berbagai bahasa pemrograman serta framework populer. Selain itu, editor ini bersifat multiplatform (dapat dijalankan di Windows, macOS, dan Linux) serta menyediakan ekosistem ekstensi yang luas, memungkinkan pengguna menambahkan fitur-fitur seperti debugging, integrasi Git, linting, dan auto-completion. Kemudahan kustomisasi dan dukungan komunitas yang kuat menjadikan VS Code sebagai pilihan yang efisien dalam proses pengembangan.

#### 10. *Framework Streamlit*

*Streamlit* merupakan *framework open source* berbasis Python yang dirancang khusus untuk memudahkan pembuatan aplikasi web interaktif, terutama dalam bidang data *science* dan *machine learning*. Dalam penelitian ini, *Streamlit* digunakan sebagai media pengembangan antarmuka aplikasi web karena memiliki integrasi yang sangat baik dengan ekosistem *Python*. Mengingat inti (*core*) dari sistem rekomendasi saham yang dikembangkan ditulis dalam bahasa *Python*, penggunaan *Streamlit* memungkinkan visualisasi hasil rekomendasi serta interaksi pengguna dengan model secara efisien dan intuitif tanpa memerlukan keahlian dalam pengembangan web konvensional.

### 3.3 Analisis Sistem

Dalam tahap analisis sistem, dilakukan perancangan alur kerja menggunakan pendekatan *flowchart* untuk menggambarkan interaksi pengguna dengan sistem rekomendasi sinyal saham. Saat pengguna membuka aplikasi, antarmuka utama akan menampilkan panel kontrol di sisi kiri dan hasil rekomendasi di sisi kanan. Pengguna dapat memilih jenis grafik harga (*Candlestick* atau *Line*), menentukan periode tampilan grafik dalam satuan hari, memilih kode saham dari daftar yang tersedia, serta menentukan tanggal rekomendasi. Setelah konfigurasi selesai, pengguna dapat menekan tombol "Jalankan Rekomendasi", dan sistem akan menjalankan model XGBoost untuk memproses data harga saham historis. Hasil klasifikasi akan ditampilkan dalam bentuk rekomendasi strategi trading (*BUY*, *HOLD*, atau *SELL*), lengkap dengan *entry level*, *resistence*, dan *stoploss*, serta visualisasi grafik teknikal yang mencakup pergerakan harga saham, garis MA5,

MA10, MA20, dan titik rekomendasi. *Flowchart* kerja sistem akan di jelaskan pada gambar 3.4.



Gambar 3. 4 *Flowchart* Alur Kerja Sistem

Berdasarkan gambar 3.4, alur sistem rekomendasi saham terdiri dari beberapa tahapan yang dapat dijelaskan sebagai berikut:

- Tahap pertama pengguna mengakses tampilan awal sistem rekomendasi saham. Pada halaman ini, pengguna akan disuguhkan dengan dasbor rekomendasi, sementara di sisi kanan layar tersedia sejumlah konfigurasi yang perlu disesuaikan sebelum melakukan rekomendasi.
- Selanjutnya, pengguna dapat mulai mengatur konfigurasi sistem yang tersedia. Pengaturan ini mencakup: (1) pemilihan jenis tampilan grafik harga saham,

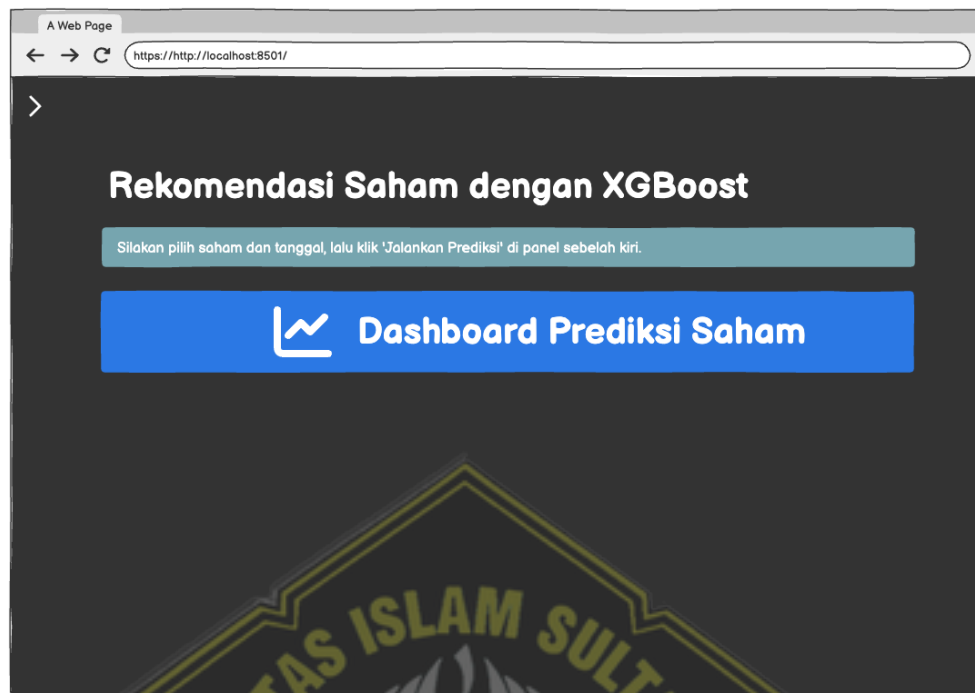
yaitu antara *candlestick* atau *line chart*; (2) penyesuaian rentang waktu data historis yang ingin ditampilkan, mulai dari 30 hingga 200 hari terakhir; (3) pemilihan kode saham yang tersedia dalam sistem; serta (4) penentuan tanggal rekomendasi yang diinginkan.

- c. Setelah seluruh konfigurasi selesai ditetapkan, sistem akan memproses data yang telah diinputkan. Hasil dari proses tersebut berupa sinyal rekomendasi investasi serta penetapan *level entry*, yang diambil dari harga penutupan satu hari sebelum tanggal rekomendasi yang dipilih. Selain itu, sistem juga secara otomatis menentukan *level resistance* atau *stop loss* dengan mengacu pada nilai *moving average* terdekat atau berdasarkan prinsip psikologis harga yang umum digunakan dalam analisis teknikal.
- d. Setelah proses pemodelan dan analisis selesai dilakukan, sistem rekomendasi akan menampilkan hasil akhir berupa sinyal investasi saham, yaitu *BUY*, *HOLD*, atau *SELL*. Rekomendasi ini dihasilkan dari model XGBoost yang telah dilatih menggunakan data historis dan indikator teknikal. Model akan memproses input yang telah ditentukan oleh pengguna pada tahap konfigurasi awal, seperti pemilihan saham, jangka waktu, dan tanggal rekomendasi. Dengan demikian, sistem dapat memberikan rekomendasi yang relevan dan berbasis data, guna membantu pengguna dalam mengambil keputusan investasi secara lebih objektif.

### 3.4 Perancangan *User*

#### 3.4.1 Halaman Awal Sistem Rekomendasi

Halaman awal pada sistem rekomendasi saham merupakan tampilan antarmuka yang muncul sebelum pengguna menjalankan proses rekomendasi. Tampilan ini menyajikan dasbor utama serta konfigurasi yang perlu diatur terlebih dahulu, seperti terlihat pada Gambar 3.5.

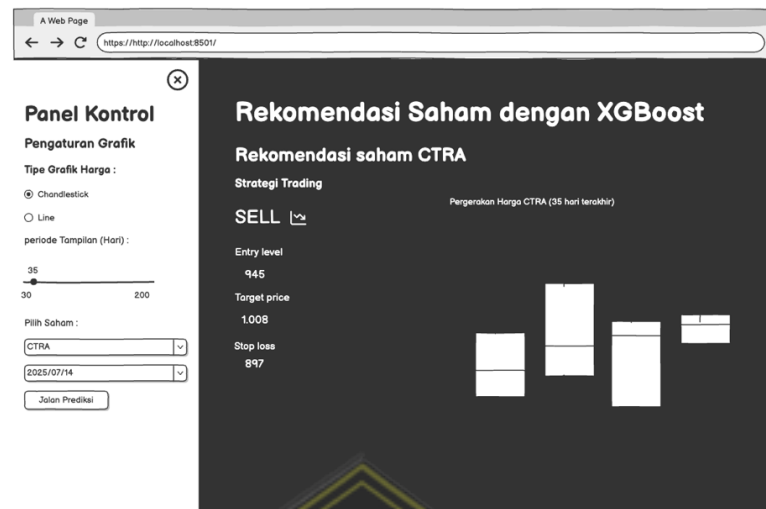


Gambar 3. 5 Rancangan Halaman Awal

Gambar 3.5 memperlihatkan tampilan awal dari sistem rekomendasi saham yang dapat diakses oleh pengguna. Pada tampilan tersebut terdapat ikon berbentuk tanda “>” yang berfungsi sebagai pemicu untuk menampilkan panel kontrol. Setelah ikon tersebut diklik, akan muncul panel konfigurasi yang memungkinkan pengguna untuk mengatur berbagai parameter, seperti pemilihan kode saham dan tanggal yang diinginkan untuk memperoleh rekomendasi. Fitur ini dirancang untuk memberikan fleksibilitas dan kemudahan dalam penggunaan sistem secara interaktif.

### 3.4.2 Halaman Hasil Sistem Rekomendasi

Halaman hasil menampilkan output rekomendasi saham berdasarkan konfigurasi yang telah diatur sebelumnya oleh pengguna. Pada halaman ini, sistem menampilkan sinyal investasi berupa *BUY*, *HOLD*, atau *SELL* sesuai dengan hasil analisis dari model. Visualisasi tampilan halaman ini dapat dilihat pada Gambar 3.6.



Gambar 3. 6 Rancangan Halaman Hasil

Gambar 3.6 menjelaskan hasil konfigurasi pada panel kontrol di sisi kiri, seperti memilih tipe grafik harga (*Candlestick* atau *Line*), menentukan periode tampilan data harga (dalam rentang 30 hingga 200 hari), memilih kode saham (dalam contoh ini: CTRA), dan memilih tanggal prediksi. Selanjutnya, pengguna dapat menekan tombol "Jalankan Prediksi" untuk memulai proses rekomendasi. Setelah itu, pada sisi kanan layar, sistem akan menampilkan hasil rekomendasi berdasarkan model XGBoost yang telah dilatih sebelumnya. Hasil prediksi mencakup:

- Sinyal Rekomendasi: Dalam gambar, sinyal yang diberikan adalah *SELL*, yang mengindikasikan bahwa model memprediksi potensi penurunan harga dalam waktu dekat.
- *Entry Level*: Harga masuk atau acuan posisi awal (dalam contoh ini adalah Rp945).
- *Target Price*: Harga target yang diantisipasi untuk dijadikan acuan jual (Rp1.008).
- *Stop Loss*: Batas harga terendah sebagai bentuk proteksi risiko (Rp947).

Selain informasi numerik, terdapat pula grafik harga saham CTRA dalam 35 hari terakhir dengan visualisasi *Candlestick*, yang memperlihatkan pergerakan harga untuk mendukung keputusan investor secara visual. Fitur ini memberikan kemudahan bagi pengguna, khususnya investor ritel atau pemula, untuk mengambil keputusan investasi secara lebih berbasis data dan visual.

## BAB IV

### HASIL ANALISIS PENELITIAN

#### 4.1 Hasil Penelitian

##### 4.1.1 Pengumpulan Data

Tahapan awal dalam penelitian ini dimulai dengan pengumpulan data historis yang berperan sebagai dasar untuk pelatihan dan pengujian model klasifikasi. Data diperoleh melalui pemanfaatan *library yfinance*, yang dipilih karena keandalannya dalam menyediakan data keuangan publik secara akurat, lengkap, dan mudah diakses. Ketersediaan data yang relevan dan berkualitas menjadi elemen krusial dalam membangun model *machine learning* yang mampu memberikan hasil klasifikasi yang optimal.

Rentang waktu pengumpulan data dimulai dari 1 Januari 2020 hingga 18 Juli 2025, yang mencakup berbagai fase pasar seperti kondisi tren naik (*bull market*) dan tren turun (*bear market*). Rentang waktu yang luas ini bertujuan untuk memberikan pemahaman menyeluruh terhadap dinamika pasar yang bervariasi. Total data yang berhasil dikumpulkan berjumlah 12.575 baris, yang berasal dari 10 saham sektor properti yang terdaftar di bursa efek Indonesia, dan selanjutnya akan dibagi menjadi data pelatihan (*training*) dan pengujian (*testing*) guna membangun serta mengevaluasi performa model klasifikasi. Pada tabel 4.1 menampilkan data yang sudah diambil dari web YahooFinance.

Tabel 4. 1 Data Hasil *Scraping*

| <i>Date</i> | <i>Close</i> | <i>High</i> | <i>Low</i> | <i>Open</i> | <i>Volume</i> | <i>Ticker</i> |
|-------------|--------------|-------------|------------|-------------|---------------|---------------|
| 2020/01/02  | 930.000061   | 952.682989  | 930.000061 | 943.609818  | 8099700       | CTRA          |
| 2020/01/03  | 952.682983   | 961.756155  | 930.000055 | 934.536641  | 17901300      | CTRA          |
| 2020/01/06  | 925.463440   | 952.682953  | 925.463440 | 952.682953  | 15719600      | CTRA          |
| 2020/01/07  | 925.463440   | 934.536611  | 911.853683 | 925.463440  | 20064800      | CTRA          |
| 2020/01/08  | 920.926880   | 925.463465  | 907.317123 | 916.390294  | 29455400      | CTRA          |
| 2025/07/11  | 885.0        | 920.0       | 865.0      | 885.0       | 119400        | SMDM          |
| 2025/07/14  | 850.0        | 885.0       | 840.0      | 885.0       | 397200        | SMDM          |
| 2025/07/15  | 875.0        | 895.0       | 830.0      | 855.0       | 177600        | SMDM          |
| 2025/07/16  | 825.0        | 880.0       | 825.0      | 875.0       | 353100        | SMDM          |
| 2025/07/17  | 830.0        | 835.0       | 815.0      | 825.0       | 116000        | SMDM          |

Tabel 4.1 menampilkan data mentah yang diperoleh melalui platform Yahoo Finance, dengan cakupan rentang waktu dari tahun 2020 hingga 2025. Setelah proses pengumpulan data selesai, tahap selanjutnya adalah melakukan pembuatan fitur teknikal serta pelabelan data. Proses pelabelan ini didasarkan pada hasil wawancara dengan pakar, yaitu Mas Arief selaku pendiri komunitas Green Stock Trader, untuk memastikan bahwa klasifikasi sinyal investasi sesuai dengan pengetahuan praktis di lapangan. Pada tabel 4.2 menampilkan dataset yang akan digunakan.

Tabel 4. 2 *Dataset*

| <i>Date</i>     | <i>Close</i> | <i>High</i>    | <i>Low</i> | <i>Open</i>        | <i>Volume</i> | <i>Ticker</i>    | <i>MA_5</i> | <i>MA_10</i> | <i>MA_20</i> |
|-----------------|--------------|----------------|------------|--------------------|---------------|------------------|-------------|--------------|--------------|
| 2020/01/29      | 210          | 216            | 208        | 212                | 12363400      | ASRI             | 211.6       | 218.8        | 228.3        |
| 2020/01/30      | 206          | 214            | 204        | 212                | 6709600       | ASRI             | 210.4       | 215.8        | 226.7        |
| 2020/01/31      | 198          | 208            | 198        | 208                | 8545000       | ASRI             | 207.6       | 212.2        | 224.8        |
| 2020/02/03      | 186          | 199            | 180        | 198                | 15133600      | ASRI             | 202.4       | 207.8        | 222.4        |
| 2020/02/04      | 185          | 190            | 185        | 186                | 10564000      | ASRI             | 197         | 204.5        | 219.85       |
| <i>RSI</i>      |              | <i>MACD</i>    |            | <i>MACD_Signal</i> |               | <i>MACD_Hist</i> |             | <i>OBV</i>   |              |
| 20.18509371     |              | -7.145537557   |            | -4.63489157        |               | -2.510645987     |             | -51284800    |              |
| 17.72023865     |              | -7.70860306    |            | -5.249633868       |               | -2.458969192     |             | -57994400    |              |
| 14.03014454     |              | -8.700080556   |            | -5.939723205       |               | -2.76035735      |             | -66539400    |              |
| 10.49853599     |              | -10.3349979    |            | -6.818778145       |               | -3.516219759     |             | -81673000    |              |
| <i>BB_Width</i> |              | <i>ATRr_14</i> |            | <i>Signal</i>      |               | <i>Label</i>     |             |              |              |
| 50.28089519     |              | 7.245860564    |            | SELL               |               | 2                |             |              |              |
| 53.14686702     |              | 7.613079155    |            | SELL               |               | 2                |             |              |              |
| 58.1772247      |              | 7.931335268    |            | SELL               |               | 2                |             |              |              |
| 66.96299056     |              | 9.407157232    |            | SELL               |               | 2                |             |              |              |

Tabel 4.2 menyajikan dataset yang telah siap digunakan untuk proses pelatihan dan pengujian model klasifikasi saham. *Dataset* ini telah melalui tahapan pelabelan dari pakar saham, serta proses *feature engineering* untuk menambahkan fitur-fitur yang relevan dengan analisis teknikal. Beberapa fitur yang dimasukkan antara lain meliputi indikator-indikator teknikal yang umum digunakan dalam analisis pergerakan harga saham:

- *Date*: Menunjukkan tanggal pencatatan data.

- *Close*: Harga penutupan saham pada hari tersebut, yang juga menjadi target utama dalam proses klasifikasi.
- *High*: Harga tertinggi yang dicapai oleh saham dalam satu hari perdagangan.
- *Low*: Harga terendah yang dicapai dalam periode yang sama.
- *Open*: Harga pembukaan saham pada hari tersebut.
- *Volume*: Jumlah saham yang diperdagangkan pada hari tersebut.
- *Ticker*: Kode atau simbol saham yang merepresentasikan masing-masing perusahaan.
- *MA\_5, MA\_10, MA\_20*: Merupakan indikator *Moving Average* (MA) yang menggambarkan rata-rata harga penutupan dalam 5, 10, dan 20 hari terakhir secara berturut-turut, berguna untuk mengidentifikasi tren jangka pendek dan menengah.
- *RSI (Relative Strength Index)*: Indikator momentum yang digunakan untuk menilai kekuatan pergerakan harga dan mengidentifikasi kondisi *overbought* (jenuh beli) atau *oversold* (jenuh jual). Nilai RSI di atas 70 umumnya menandakan kondisi *overbought*, sedangkan nilai di bawah 30 menandakan *oversold*.
- *MACD (Moving Average Convergence Divergence)*: Indikator yang menunjukkan hubungan antara dua *moving average* (EMA\_12 dan EMA\_26) untuk mendeteksi arah tren dan momentum harga.
- *MACD\_Signal*: Garis sinyal (EMA-9 dari MACD) yang digunakan sebagai acuan untuk mengidentifikasi momen potensial beli atau jual. Sinyal beli muncul saat MACD memotong garis sinyal dari bawah, sedangkan sinyal jual muncul saat memotong dari atas.
- *MACD\_Hist*: Histogram yang merepresentasikan selisih antara garis MACD dan garis sinyal. Nilai positif menunjukkan tren naik yang menguat, sedangkan nilai negatif menunjukkan tren turun yang menguat.
- *OBV (On-Balance Volume)*: Indikator volume yang menggabungkan pergerakan harga dan volume perdagangan untuk mengukur tekanan beli dan jual. Nilai OBV meningkat ketika harga penutupan hari ini lebih tinggi dari hari sebelumnya (menandakan tekanan beli), dan menurun ketika harga

penutupan hari ini lebih rendah dari hari sebelumnya (menandakan tekanan jual).

- **BB\_Width (*Bollinger Band Width*)**: Merupakan indikator yang menghitung jarak antara pita atas dan pita bawah pada *Bollinger Bands*. Nilai ini memberikan gambaran mengenai tingkat volatilitas saat ini dengan mengacu pada penyebaran harga relatif terhadap rata-rata pergerakan. Semakin lebar pita, semakin tinggi volatilitas harga pada saat itu.
- **ATRR\_14 (*Average True Range*)**: Indikator volatilitas yang menghitung rata-rata rentang pergerakan harga harian (*high-low*) selama periode 14 hari terakhir. ATR lebih fokus pada seberapa besar harga berfluktuasi dari hari ke hari, tanpa memperhatikan arah tren.
- **Signal**: Label sinyal berdasarkan pendekatan ahli, seperti *BUY*, *HOLD*, atau *SELL*.
- **Label**: Representasi numerik dari sinyal untuk keperluan klasifikasi oleh algoritma machine learning.

Seluruh fitur yang disajikan dalam dataset ini dirancang untuk menyediakan informasi yang menyeluruh guna mendukung pengembangan sistem rekomendasi saham. Sistem ini tidak hanya memanfaatkan pendekatan analisis teknikal, tetapi juga mempertimbangkan masukan dari pakar saham, sehingga model yang dibangun menggunakan algoritma XGBoost, diharapkan mampu menghasilkan rekomendasi investasi yang lebih akurat dan relevan.

#### 4.1.2 Tahapan Perancangan Model

Pada tahap ini dijelaskan secara sistematis mengenai proses pengembangan model klasifikasi yang digunakan dalam penelitian. Proses perancangan model melibatkan beberapa tahapan penting, sebagai berikut :

##### 1. Load data

Pada tahap ini dilakukan proses pemuatan dataset yang digunakan dalam pembangunan model sistem rekomendasi. Dataset tersebut berisi 19 fitur dengan total 12.575 entri yang mencakup periode tahun 2020 hingga 2025. Data diklasifikasikan ke dalam tiga kelas, yaitu Hold sebanyak 1.918 entri, Sell sebanyak 4.913 entri, dan Buy sebanyak 5.744 entri.

## 2. *Preprocessing* data

Tahapan ini merupakan bagian krusial dalam pembangunan model, karena data mentah yang telah dimuat akan melalui serangkaian proses *preprocessing* guna memastikan data tersebut layak dan dapat digunakan oleh model dalam pengembangan sistem rekomendasi. Adapun tahapan *preprocessing* yang dilakukan adalah sebagai berikut:

### a. *Handle Missing*

Dalam pengelolaan *dataset*, sering kali ditemukan kesalahan yang disebabkan oleh faktor manusia maupun berasal dari sumber data itu sendiri. Untuk memastikan kinerja model yang optimal, dilakukan pemeriksaan *missing value* pada dataset menggunakan fungsi *.isnull()*. Apabila ditemukan nilai yang hilang, data tersebut akan segera dihapus guna menjaga kualitas dan konsistensi data yang akan digunakan dalam proses pelatihan model. Pada tabel 4.3 akan menampilkan hasil dari *missing value*.

Tabel 4. 3 *Missing Value*

| <i>Fitur</i>       | <i>Missing Value</i> |
|--------------------|----------------------|
| <i>Open</i>        | 0                    |
| <i>High</i>        | 0                    |
| <i>Low</i>         | 0                    |
| <i>Close</i>       | 0                    |
| <i>Volume</i>      | 0                    |
| <i>MA_5</i>        | 0                    |
| <i>MA_10</i>       | 0                    |
| <i>MA_20</i>       | 0                    |
| <i>RSI</i>         | 0                    |
| <i>MACD</i>        | 0                    |
| <i>MACD_Signal</i> | 0                    |
| <i>MACD_Hist</i>   | 0                    |
| <i>OBV</i>         | 0                    |
| <i>BB_Width</i>    | 0                    |
| <i>ATRr_14</i>     | 0                    |

Hasil pemeriksaan *missing value* yang ditampilkan pada Tabel 4.3 menunjukkan bahwa tidak terdapat data yang hilang atau bernilai *NaN* pada *dataset*. Dengan demikian, data dinyatakan layak untuk diproses lebih lanjut pada tahapan berikutnya.

b. Split Train-test

Pembagian data dilakukan dengan pendekatan *time series split* untuk menjaga kontinuitas dan ketepatan urutan deret waktu, yang sangat penting dalam pemodelan data runtun waktu (*time series*). Data dibagi dengan proporsi 80% untuk pelatihan (*training*) dan 20% untuk pengujian (*testing*). Pada tabel 4.4 akan ditampilkan hasil dari pembagian data.

Tabel 4. 4 Menampilkan *Split Data*

| Keterangan              | Hasil                     |
|-------------------------|---------------------------|
| Jumlah Data             | 12.575                    |
| Data Pelatihan          | 10.058                    |
| Data Pengujian          | 2.517                     |
| Rentang Waktu Training  | 2020-01-29 s.d 2024-06-06 |
| Rentang tanggal testing | 2024-06-07 s.d 2025-07-17 |

Tabel 4.4 menunjukkan hasil pembagian data tersebut, di mana total data awal sebanyak 12.575 baris telah terbagi secara proporsional menjadi dua bagian, sesuai dengan prinsip *time series split*. Teknik ini digunakan untuk menghindari data *leakage* serta memastikan bahwa model diuji pada data yang benar-benar belum pernah dilihat sebelumnya dalam urutan waktu yang realistis.

c. Augmentasi Data

Pada tahap ini, dilakukan proses penyeimbangan distribusi kelas (*class balancing*) karena ditemukan ketidakseimbangan jumlah sampel, khususnya pada kelas *HOLD* yang memiliki jumlah data lebih sedikit dibandingkan kelas lainnya. Untuk mengatasi permasalahan ini, digunakan teknik SMOTE (*Synthetic Minority Over-sampling Technique*) yang menghasilkan data sintetis guna menyeimbangkan jumlah sampel antar kelas.

Proses *oversampling* ini diterapkan hanya pada data pelatihan, dengan tujuan menjaga keaslian data pengujian. Hal ini penting dilakukan untuk mencegah

terjadinya data *leakage* dan menjaga integritas evaluasi model. Dengan menggunakan data pengujian yang tetap asli, proses evaluasi menjadi lebih objektif dan mencerminkan performa model dalam menghadapi data nyata yang belum pernah dilihat sebelumnya. Selain itu, hal ini juga meminimalkan risiko munculnya *noise* atau bias selama proses pengujian. Pada table 4.5 akan ditampilkan hasil dari proses penyeimbangan kelas.

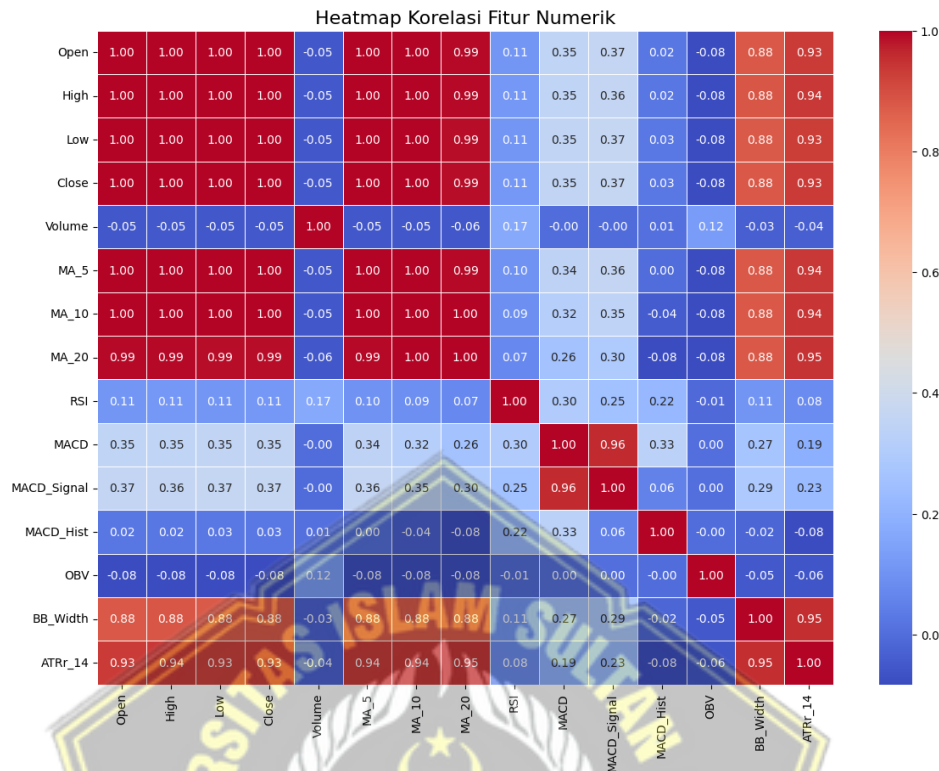
Tabel 4. 5 Menampilkan *Oversampling*

| Label         | Data Pengujian |         |         |
|---------------|----------------|---------|---------|
|               | <i>SMOTE</i>   | Sebelum | Sesudah |
| 0             | <i>BUY</i>     | 4.524   | 4.524   |
| 2             | <i>SELL</i>    | 3.939   | 4.524   |
| 1             | <i>HOLD</i>    | 1.595   | 4.524   |
| Data Training |                | 10.058  | 13.575  |

Tabel 4.5 menyajikan hasil dari penerapan metode SMOTE pada data pelatihan. Setelah dilakukan proses augmentasi, jumlah data meningkat dari semula 10.058 entri menjadi 13.572 entri. Peningkatan ini merupakan hasil dari penambahan sampel sintetis pada kelas minoritas, sehingga distribusi antar kelas menjadi lebih seimbang dan mendukung kinerja model klasifikasi yang lebih adil.

### 3. *Feature selection*

Pada tahap ini, dilakukan proses *feature selection* guna memastikan bahwa model hanya dilatih menggunakan fitur-fitur yang memiliki korelasi signifikan terhadap variabel target. Untuk itu, analisis korelasi antar fitur dilakukan untuk mengidentifikasi dan mengecualikan fitur yang tidak relevan atau bersifat redundan. Hasil dari pengecekan korelasi data ditampilkan pada gambar 4.1.



Gambar 4. 1 Korelasi Data

Gambar 4.1 menyajikan hasil analisis korelasi antar fitur. Nilai korelasi yang mendekati 1.00 mengindikasikan hubungan positif yang sangat kuat (korelasi positif sempurna), sebagaimana ditunjukkan oleh fitur-fitur harga seperti *Open*, *High*, *Low*, dan *Close*, yang saling berkorelasi tinggi karena merupakan komponen dari harga harian yang saling terkait. Hal serupa juga terlihat pada fitur *moving average* (*MA\_5*, *MA\_10*, dan *MA\_20*) yang menunjukkan korelasi positif yang kuat akibat basis perhitungannya yang menggunakan data historis harga dalam periode waktu yang berdekatan.

Sebaliknya, korelasi mendekati 0.00, seperti antara *Volume* dengan indikator teknikal seperti *RSI* dan *MACD*, menunjukkan hubungan linier yang sangat lemah atau tidak signifikan, menandakan bahwa fitur-fitur tersebut mungkin merepresentasikan dimensi informasi yang berbeda. Berdasarkan hasil analisis korelasi ini, peneliti memutuskan untuk mempertahankan seluruh fitur dalam proses pemodelan, karena meskipun beberapa fitur memiliki korelasi rendah, masing-masing diyakini memberikan kontribusi unik yang berpotensi meningkatkan performa model secara keseluruhan.

#### 4. Model Building

Dalam pengembangan model XGBoost, peneliti menerapkan *Optuna* sebagai metode optimasi *hyperparameter* untuk memperoleh konfigurasi parameter yang optimal, sehingga model dapat beroperasi secara lebih efisien dan akurat. Tahap pembangunan model merupakan bagian penting dalam penelitian ini, di mana algoritma XGBoost digunakan sebagai inti dari sistem rekomendasi saham berbasis klasifikasi sinyal.

*Dataset* yang digunakan terlebih dahulu diseimbangkan menggunakan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) untuk mengatasi ketidakseimbangan distribusi kelas (*class imbalance*) yang dapat berdampak negatif terhadap performa model. Melalui konfigurasi *hyperparameter* yang telah dioptimalkan, model diharapkan mampu melakukan proses pembelajaran secara lebih efektif terhadap data historis yang tersedia. Objektif yang digunakan dalam optimasi adalah nilai *F1-score* makro pada data pengujian. Proses tuning dilakukan sebanyak 50 *trials*, dan parameter yang dioptimasi beserta ruang pencariannya. Rincian nilai *hyperparameter* optimasi yang digunakan dalam pelatihan model ditampilkan pada Tabel 4.6

Tabel 4. 6 Konfigurasi Hiperparameter

| Keterangan              | Rentang Pencarian |
|-------------------------|-------------------|
| <i>n_estimators</i>     | [200, 400]        |
| <i>max_depth</i>        | [3, 7]            |
| <i>learning_rate</i>    | [0.01, 0.3]       |
| <i>subsample</i>        | [0.7, 1.0]        |
| <i>colsample_bytree</i> | [0.6, 0.9]        |
| <i>reg_alpha</i>        | [0.1, 10.0]       |
| <i>reg_lambda</i>       | [0.1, 10.0]       |

#### 5. Model Training

Proses pelatihan model dilakukan dengan menggunakan konfigurasi *hyperparameter* terbaik yang telah diperoleh melalui proses optimasi menggunakan

*Optuna*. *Dataset* yang digunakan pada tahap ini sama dengan data yang digunakan dalam pembangunan model sebelumnya, yaitu data yang telah diseimbangkan melalui teknik *Synthetic Minority Over-sampling Technique* (SMOTE) untuk mengatasi ketidakseimbangan distribusi kelas. Dalam konfigurasi tersebut, parameter *n\_estimators* berperan menentukan jumlah pohon yang dibangun oleh model, sementara *max\_depth* mengatur kedalaman maksimum setiap pohon untuk menjaga keseimbangan antara *overfitting* dan *underfitting*.

Parameter *learning\_rate* mengatur kontribusi masing-masing pohon terhadap klasifikasi akhir, di mana nilai yang lebih rendah dapat menghasilkan proses pembelajaran yang lebih stabil namun memerlukan lebih banyak iterasi. Selain itu, *subsample* dan *colsample\_bytree* digunakan sebagai mekanisme regularisasi tambahan dengan mengambil sebagian data dan fitur secara acak dalam pembangunan tiap pohon, sehingga meningkatkan kemampuan generalisasi model. Terakhir, *reg\_alpha* dan *reg\_lambda* merupakan parameter regularisasi L1 dan L2 yang berfungsi untuk mengendalikan kompleksitas model dengan menekan bobot fitur yang tidak relevan, sehingga model dapat bekerja secara lebih efisien dan *robust* terhadap *overfitting*. Pada tabel 4.7 ditampilkan hasil dari model terbaik dari fungsi objektif *optuna*.

Tabel 4. 7 Hasil Konfigurasi Terbaik

| Keterangan              | Hasil |
|-------------------------|-------|
| <i>n_estimators</i>     | 226   |
| <i>max_depth</i>        | 4     |
| <i>learning_rate</i>    | 0.17  |
| <i>subsample</i>        | 0.74  |
| <i>colsample_bytree</i> | 0.70  |
| <i>reg_alpha</i>        | 7.44  |
| <i>reg_lambda</i>       | 6.27  |

Tabel 4.7 menampilkan hasil konfigurasi terbaik dari model XGBoost, di mana model tersebut berhasil mencapai nilai *F1-score* tertinggi sebesar 0,83. Nilai ini mencerminkan keseimbangan yang baik antara *presisi* dan *recall* dalam klasifikasi sinyal saham. Berdasarkan hasil tersebut, model kemudian diuji lebih lanjut pada

data uji guna mengevaluasi generalisasi performa serta untuk menyimpulkan efektivitas pendekatan yang telah diterapkan.

#### 6. Model Evaluasi

Pada tahap ini dilakukan evaluasi terhadap kinerja model XGBoost berdasarkan hasil pengujian pada data pelatihan dan data pengujian. Berdasarkan konfigurasi *hyperparameter* terbaik yang telah ditentukan dan disajikan pada Tabel 4.8, model berhasil mencapai akurasi sebesar 93,19% pada data pelatihan. Capaian ini menunjukkan bahwa model mampu mempelajari pola-pola penting dalam data secara efektif dan menunjukkan performa klasifikasi yang baik selama proses pelatihan. Nilai *precision*, *recall*, dan *F1-score* mengindikasikan bahwa model bekerja cukup seimbang dalam mengklasifikasi setiap kelas pada data pelatihan.

Tabel 4. 8 Evaluasi Model

| <b><i>Dataset</i></b> | <b><i>Precision</i></b> | <b><i>Recall</i></b> | <b><i>F1-score</i></b> |
|-----------------------|-------------------------|----------------------|------------------------|
| <i>Training</i>       | 0.90                    | 0.92                 | 0.91                   |
| <i>Testing</i>        | 0.82                    | 0.85                 | 0.83                   |

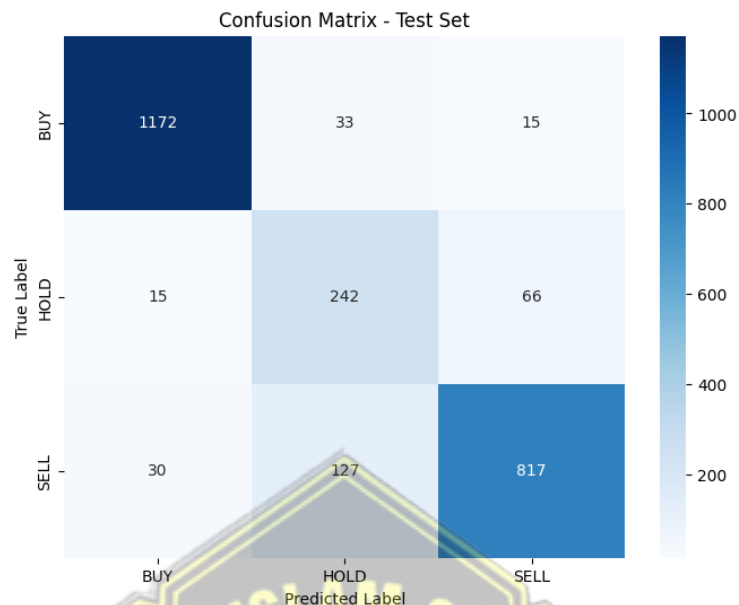
Namun demikian, hasil evaluasi pada data pengujian menunjukkan penurunan performa dengan akurasi sebesar 88,64%. Perbedaan performa ini tercermin dalam Tabel 4.8 yang memperlihatkan perbandingan metrik evaluasi antara data pelatihan dan data pengujian. Penurunan tersebut mengindikasikan bahwa meskipun model memiliki kemampuan belajar yang baik dari data pelatihan, kemampuan generalisasinya terhadap data baru masih perlu ditingkatkan. Oleh karena itu, hasil ini menunjukkan bahwa model masih memerlukan pengembangan lebih lanjut, baik dari sisi optimasi parameter maupun peningkatan kualitas dan keberagaman data pelatihan, guna mencapai performa klasifikasi yang lebih optimal pada data yang belum pernah dilihat sebelumnya. Pada tabel 4.9 menjelaskan hasil evaluasi dari ketiga kelas.

Tabel 4. 9 Evaluasi Pengujian Model

| <b>Kelas</b>       | <b><i>Precision</i></b>                                  | <b><i>Recal</i></b>                                      | <b><i>F1-Score</i></b>                                                       |
|--------------------|----------------------------------------------------------|----------------------------------------------------------|------------------------------------------------------------------------------|
| <b><i>BUY</i></b>  | $\frac{1172}{1172 + 15 + 30} = \frac{1172}{1217} = 0,96$ | $\frac{1172}{1172 + 33 + 15} = \frac{1172}{1220} = 0,96$ | $\frac{2 \times (0,96 \times 0,96)}{0,96 + 0,96} = \frac{1,85}{1,92} = 0,96$ |
| <b><i>HOLD</i></b> | $\frac{242}{242 + 33 + 127} = \frac{242}{402} = 0,60$    | $\frac{242}{242 + 15 + 66} = \frac{242}{323} = 0,75$     | $\frac{2 \times (0,60 \times 0,75)}{0,60 + 0,75} = \frac{0,9}{1,35} = 0,67$  |
| <b><i>SELL</i></b> | $\frac{817}{817 + 15 + 66} = \frac{817}{898} = 0,91$     | $\frac{817}{817 + 30 + 127} = \frac{817}{974} = 0,84$    | $\frac{2 \times (0,91 \times 0,84)}{0,91 + 0,84} = \frac{1,53}{1,75} = 0,87$ |

Tabel 4.9 menyajikan hasil evaluasi model XGBoost pada data pengujian berdasarkan masing-masing kelas sinyal: *BUY*, *HOLD*, dan *SELL*. Berdasarkan hasil evaluasi tersebut, kelas *BUY* menunjukkan performa tertinggi dengan nilai *precision*, *recall*, dan *F1-score* masing-masing sebesar 0,96. Hal ini menunjukkan bahwa model mampu mengenali sinyal beli dengan sangat baik dan memiliki tingkat kesalahan klasifikasi yang sangat rendah pada kelas ini. Sinyal *SELL* juga menunjukkan performa yang cukup baik, dengan *precision* sebesar 0,91 dan *F1-score* sebesar 0,87. Namun, nilai *recall*-nya yang berada di angka 0,84 mengindikasikan bahwa masih terdapat sejumlah sinyal jual yang gagal terdeteksi secara benar oleh model, meskipun secara umum performanya masih tergolong tinggi.

Di sisi lain, kelas *HOLD* menjadi perhatian utama dalam evaluasi karena menunjukkan performa yang paling rendah, dengan *precision* sebesar 0,60, *recall* sebesar 0,75, dan *F1-score* sebesar 0,67. Nilai-nilai tersebut mengindikasikan bahwa model masih mengalami kesulitan dalam membedakan sinyal *HOLD* dengan kelas lainnya. Meskipun demikian, kemampuan model untuk mencapai *precision* sebesar 60% pada kelas ini tetap menunjukkan adanya pembelajaran yang berlangsung, meskipun perlu ditingkatkan lebih lanjut. Secara keseluruhan, hasil ini menunjukkan bahwa model sudah sangat baik dalam mengenali sinyal beli dan jual, namun masih memerlukan peningkatan dalam menangani sinyal *HOLD* agar performa klasifikasi lebih seimbang di semua kelas. Pada gambar 4.2 menampilkan hasil *confusion matrix* dari model klasifikasi yang dibangun, hasil evaluasi ini menjadi validasi model.



Gambar 4. 2 *Confusion Matrix*

Gambar 4.2 menyajikan matriks konfusi yang merefleksikan kemampuan model dalam mengklasifikasikan data pada tahap pengujian. Berdasarkan hasil tersebut, model menunjukkan performa yang sangat baik dalam mengenali kelas *BUY*, dengan 1.172 klasifikasi yang tepat dan hanya sedikit kesalahan klasifikasi, yaitu 33 sampel diklasifikasikan sebagai *HOLD* dan 15 sebagai *SELL*. Kinerja untuk kelas *SELL* juga cukup memuaskan, di mana 817 instans berhasil diklasifikasikan dengan benar, meskipun terdapat 127 sampel yang salah diklasifikasikan ke dalam kelas *HOLD*, menunjukkan adanya ambiguitas antara dua kelas tersebut. Sebaliknya, kelas *HOLD* menjadi tantangan terbesar bagi model, dengan hanya 242 klasifikasi yang benar dari total sampel yang tersedia. Sebanyak 66 instans *HOLD* salah diklasifikasikan sebagai *SELL*, mengindikasikan keterbatasan model dalam membedakan kedua kelas tersebut. Performa yang lebih rendah ini diduga disebabkan oleh ketimpangan jumlah data antar kelas dalam dataset pengujian, di mana kelas *HOLD* memiliki representasi yang jauh lebih sedikit dibandingkan kelas lainnya.

#### 7. *Recomendation Result*

Hasil akhir dari pembangunan model XGBoost ditujukan untuk menghasilkan rekomendasi saham berdasarkan data terbaru. Model akan memberikan sinyal rekomendasi (*BUY*, *HOLD*, atau *SELL*) dengan menggunakan data terkini yang

diperoleh dari Yahoo Finance. Pemilihan tanggal analisis disesuaikan secara manual oleh user dengan ketersediaan data historis yang diperbarui oleh Yahoo Finance, sehingga sistem dapat memberikan rekomendasi yang relevan sesuai dengan kondisi pasar terkini. Untuk membuat *recommendation* result peneliti menggunakan *library gradio* agar tampilannya lebih menarik. Pada gambar 4.3 menampilkan tampilan awal dari sistem rekomendasi menggunakan *library gradio*.

Gambar 4. 3 Halaman Awal *Recommendation Result*

Gambar 4.3 menampilkan antarmuka sistem beserta konfigurasi awal yang harus diisi oleh pengguna sebelum mendapatkan hasil rekomendasi. Melalui tampilan ini, pengguna dapat memilih saham yang ingin dianalisis serta menentukan tanggal rekomendasi yang diinginkan. Setelah konfigurasi diatur, sistem akan menampilkan hasil klasifikasi dalam bentuk sinyal rekomendasi, yaitu *BUY*, *HOLD*, atau *SELL*. Sinyal ini diharapkan dapat menjadi panduan bagi pengguna, khususnya peneliti maupun investor, dalam mengambil keputusan transaksi saham secara lebih terinformasi berdasarkan analisis model yang telah dibangun.

Dengan adanya sistem ini, pengguna tidak perlu melakukan analisis teknikal secara manual. Hal ini sangat membantu dalam mempercepat proses pengambilan keputusan, khususnya bagi investor individu, peneliti, atau analis yang membutuhkan rekomendasi saham secara efisien dan berbasis data. Model ini

dirancang untuk bekerja secara adaptif terhadap data terbaru, sehingga mampu memberikan panduan yang relevan sesuai dengan dinamika pasar saham saat ini. Pada gambar 4.4 menampilkan hasil rekomendasi dari konfigurasi yang dimasukkan.

 **Sistem Rekomendasi Sinyal Saham**

Pilih saham dan tanggal untuk mendapatkan sinyal rekomendasi.

Pilih Saham

Pilih Tanggal

ASRI

2025-07-30

**Dapatkan Rekomendasi**

Informasi Data

Tanggal: 2025-07-30  
Harga Penutupan: Rp 162.00

Sinyal Rekomendasi

BUY

Tingkat Keyakinan

99.77%

Fitur yang Digunakan untuk Prediksi

| Open | Hi... | Low | Close | Volume | MA... | MA_10 | MA_20 | RSI          | MACD            | MACD_Signal     | MACD_Hist   |
|------|-------|-----|-------|--------|-------|-------|-------|--------------|-----------------|-----------------|-------------|
| 169  | 169   | 161 | 162   | 652015 | 157.  | 147.9 | 138.5 | 72.222222081 | 7.6546455007267 | 4.1152139790241 | 3.539431521 |

Gambar 4. 4 Hasil *Recommendation Result*

Pada Gambar 4.4 ditampilkan hasil rekomendasi dari sistem yang telah dibangun. Dalam contoh tersebut, peneliti memilih saham ASRI dengan tanggal rekomendasi 30 Juli 2025, dan sistem menghasilkan sinyal *BUY* dengan harga penutupan sebesar Rp162 per lembar saham. Selain sinyal rekomendasi, sistem juga menampilkan tingkat keyakinan (*confidence level*) dari hasil rekomendasi tersebut, yang memberikan indikasi seberapa kuat model merekomendasi sinyal tersebut. Ditampilkan pula beberapa fitur teknikal yang relevan sebagai informasi tambahan untuk membantu peneliti dalam mengevaluasi dan memvalidasi hasil rekomendasi yang diberikan. Hasil ini menunjukkan bahwa sistem dapat berjalan sesuai dengan tujuan yang telah dirancang sebelumnya, serta mampu memberikan rekomendasi yang informatif.

## 4.2 Hasil Implementasi

Setelah seluruh tahapan pemodelan selesai dilakukan, langkah selanjutnya adalah tahap implementasi sistem ke dalam sebuah platform berbasis web. Dalam

penelitian ini, implementasi dilakukan menggunakan *Streamlit*, sebuah *framework Python* yang memungkinkan pembuatan antarmuka pengguna secara interaktif dan cepat. Pengembangan antarmuka ini mengacu pada desain *wireframe* yang telah dirancang sebelumnya dan disajikan pada Bab III, sehingga seluruh komponen yang ditampilkan sesuai dengan rancangan sistem yang telah ditetapkan.

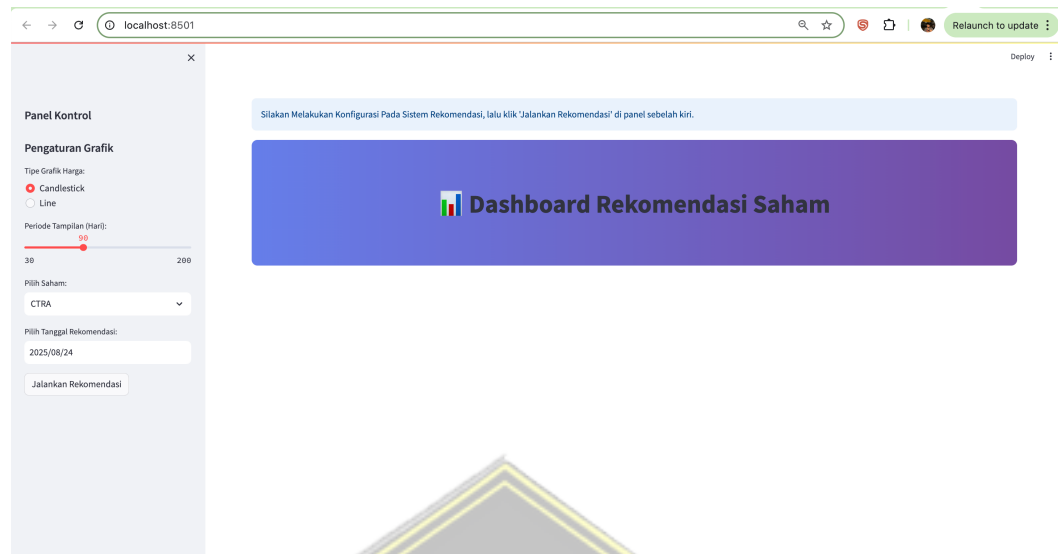
#### 4.2.1 Halaman Awal Sistem Rekomendasi

Pada tahap implementasi antarmuka halaman sistem rekomendasi saham, tampilan yang dihasilkan relatif sesuai dengan desain yang telah dirancang sebelumnya. Pada halaman awal, pengguna disajikan dengan judul sistem serta petunjuk singkat mengenai cara penggunaan sistem. Di sisi kiri halaman, tersedia panel kontrol yang memungkinkan pengguna untuk mengatur preferensi tampilan grafik, memilih periode historis data harga saham, serta menentukan saham yang akan dianalisis oleh model XGBoost. Panel kontrol dapat diakses melalui tombol ">" yang terletak di sisi kiri atas tampilan, dan akan membuka *sidebar* berisi berbagai pengaturan. Pada gambar 4.5 merupakan halaman awal dari hasil *deployment*.



Gambar 4. 5 Tampilan Awal Halaman Rekomendasi

Selanjutnya, pengguna dapat menentukan tanggal rekomendasi berdasarkan data harga terakhir yang dimiliki. Setelah seluruh parameter dikonfigurasi, pengguna dapat menekan tombol "Jalankan Rekomendasi" untuk memulai proses analisis, yang bisa dilihat pada gambar 4.6.



Gambar 4. 6 Tampilan Konfigurasi Sistem

Sebagaimana ditampilkan pada Gambar 4.6, pengguna memiliki fleksibilitas untuk melakukan konfigurasi sistem sesuai kebutuhan. Sistem akan memproses masukan yang diberikan dan menampilkan hasil rekomendasi berdasarkan pengolahan data tersebut.

#### 4.2.2 Halaman Hasil Sistem Rekomendasi

Pada tahap ini, sistem menampilkan hasil rekomendasi berdasarkan konfigurasi yang telah ditentukan oleh pengguna. Gambar 4.7 memperlihatkan hasil keluaran dari sistem rekomendasi saham yang telah diproses sesuai dengan parameter input yang diberikan oleh pengguna.

Gambar 4. 7 Sinyal *Hold*

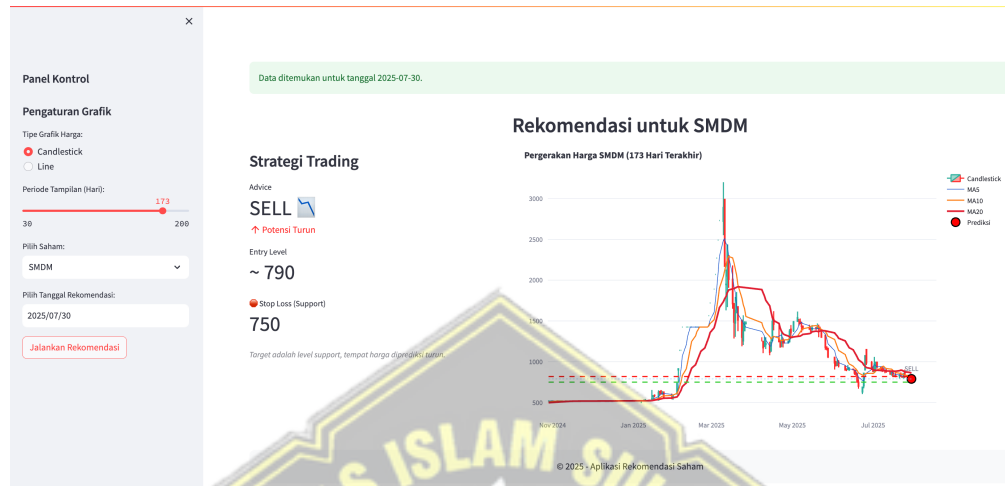
Gambar 4.7 menampilkan hasil konfigurasi ini, saham yang dianalisis dalam konfigurasi ini adalah saham dari Ciputra Group dengan kode emiten *CTRA*. Berdasarkan hasil analisis, sistem menghasilkan sinyal *Hold* yang menunjukkan bahwa kondisi pasar belum memenuhi kriteria teknikal untuk memberikan rekomendasi beli (*Buy*) maupun jual (*Sell*). Oleh karena itu, sistem menyarankan pengguna untuk bersikap netral dan menunggu perkembangan pasar selanjutnya sebelum melakukan transaksi. Selain menampilkan sinyal, sistem juga menyajikan grafik *candlestick* dari harga saham sebagai referensi visual tambahan. Grafik ini dilengkapi dengan tiga indikator *Moving Average*, yang merupakan komponen penting dalam analisis teknikal dan digunakan sebagai dasar dalam pengambilan keputusan oleh sistem rekomendasi. Pada gambar 4.8 menampilkan hasil lain dari rekomendasi.



Gambar 4. 8 Sinyal *Buy*

Gambar 4.8 menunjukkan hasil konfigurasi lain yang dilakukan oleh pengguna. Pada konfigurasi ini, pengguna memilih tampilan grafik dalam bentuk garis (*line chart*) dengan periode waktu historis yang lebih singkat, yaitu selama 73 hari. Saham yang dianalisis adalah saham PT Alam Sutera Realty Tbk dengan kode emiten *ASRI*, dan tanggal rekomendasi yang dipilih adalah 30 Juli 2025. Berdasarkan konfigurasi tersebut, sistem menghasilkan sinyal *Buy* karena parameter dan kondisi yang dibutuhkan untuk memberikan sinyal beli telah terpenuhi. Selain sinyal tersebut, sistem juga menampilkan *target price* yang diperoleh dari *level resistance* terdekat atau harga penutupan pada tanggal

rekomendasi. Sebagai bagian dari manajemen risiko, sistem turut memberikan informasi *stop loss* untuk membatasi potensi kerugian yang mungkin dialami oleh pengguna. Pada gambar 4.9 menampilkan hasil lain dari konfigurasi.



Gambar 4. 9 Sinyal *Sell*

Gambar 4.9 menyajikan konfigurasi sistem yang berbeda dari sebelumnya. Pada konfigurasi ini, pengguna memilih visualisasi harga dalam bentuk grafik *candlestick* serta menetapkan periode historis yang lebih panjang, yaitu selama 173 hari. Saham yang dianalisis merupakan saham dari perusahaan PT Suryamas Dutamakmur Tbk dengan kode emiten *SMDM*. Berdasarkan hasil pemrosesan data dengan konfigurasi tersebut, sistem menghasilkan sinyal *Sell* karena kondisi yang memenuhi syarat untuk memberikan rekomendasi jual telah terpenuhi. Sinyal ini menunjukkan bahwa sistem menyarankan pengguna untuk melakukan penjualan terhadap saham yang dimiliki sebagai bagian dari strategi pengelolaan investasi berdasarkan indikator teknikal yang terdeteksi.

Hasil rekomendasi yang dihasilkan oleh sistem menunjukkan variasi sinyal berdasarkan konfigurasi yang diberikan oleh pengguna. Pada konfigurasi awal, sistem memberikan sinyal *Hold* yang merefleksikan ketidakpastian pasar dan ketidakterpenuhan kondisi teknikal untuk aksi beli maupun jual. Pada konfigurasi berikutnya, dengan perubahan pada periode data dan jenis saham, sistem memberikan sinyal *Buy*, mengindikasikan adanya peluang kenaikan harga saham berdasarkan analisis teknikal. Sementara itu, konfigurasi terakhir menghasilkan

sinyal *Sell*, menandakan bahwa kondisi teknikal telah memenuhi kriteria untuk aksi jual.

Perbedaan sinyal ini menunjukkan bahwa parameter seperti periode data historis, jenis saham, dan tanggal analisis memiliki pengaruh signifikan terhadap keputusan yang dihasilkan oleh sistem. Hal ini menegaskan pentingnya konfigurasi yang sesuai dengan kebutuhan dan strategi pengguna dalam menggunakan sistem rekomendasi saham berbasis kecerdasan buatan. Tabel 4.10 berikut merangkum hasil sinyal rekomendasi dari beberapa konfigurasi yang diuji oleh pengguna. Setiap baris menampilkan kombinasi parameter seperti jenis grafik, periode historis, kode saham, tanggal analisis, serta sinyal yang dihasilkan beserta nilai *target price* dan *stop loss* jika tersedia.

Tabel 4. 10 Hasil Rekomendasi

| No | Grafik              | Periode<br>(Hari) | Saham | Tanggal      | Sinyal      | Target<br>Price | Stop<br>Loss |
|----|---------------------|-------------------|-------|--------------|-------------|-----------------|--------------|
| 1  | <i>Chandlestick</i> | 100               | CTRA  | 31 Juli 2025 | <i>HOLD</i> | -               | -            |
| 2  | <i>Line</i>         | 73                | ASRI  | 30 Juli 2025 | <i>BUY</i>  | 204             | 186          |
| 3  | <i>Chandlestick</i> | 173               | SMDM  | 30 Juli 2025 | <i>SELL</i> | 132             | 118          |

Tabel ini bertujuan untuk mendokumentasikan efektivitas sistem dalam menghasilkan sinyal sesuai dengan konteks data historis dan kondisi pasar. Informasi ini juga dapat digunakan sebagai dasar evaluasi terhadap performa model dan tingkat akurasi rekomendasi yang dihasilkan.

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil penelitian, penerapan metode XGBoost pada sistem rekomendasi saham sektor properti yang terdaftar di Bursa Efek Indonesia dengan pendekatan analisis teknikal menunjukkan hasil yang efektif. Sistem rekomendasi yang dikembangkan dapat digunakan sebagai sistem pembantu keputusan investasi, khususnya bagi investor pemula. Melalui rekomendasi berbasis data historis dan analisis teknikal yang objektif, sistem ini mampu menyederhanakan proses analisis, mengurangi potensi kerugian, serta meningkatkan rasa percaya diri investor dalam melakukan transaksi di pasar modal.

Pengembangan sistem dilakukan dengan memanfaatkan data historis harga saham dari Yahoo Finance, yang kemudian diperkaya melalui proses *feature engineering* dan pelabelan data berdasarkan pertimbangan pakar saham. Model XGBoost dibangun dengan dukungan algoritma *Optuna* untuk melakukan pencarian hiperparameter secara otomatis, sehingga diperoleh konfigurasi model yang optimal. Pendekatan ini terbukti lebih efisien dan akurat dibandingkan penyusunan model manual. Selain itu, permasalahan ketidakseimbangan kelas pada data pelatihan diatasi dengan menerapkan metode *Synthetic Minority Over-sampling Technique* (SMOTE).

Hasil evaluasi menunjukkan bahwa model yang dikembangkan memiliki kinerja klasifikasi yang baik, dengan akurasi pelatihan mencapai 93,19% dan akurasi pengujian sebesar 88,64%. Nilai *macro F1-score* juga tergolong tinggi, yaitu 0,91 pada data pelatihan dan 0,83 pada data pengujian. Capaian ini mengindikasikan bahwa model mampu melakukan generalisasi dengan cukup baik terhadap data baru, sehingga dapat diandalkan untuk mendukung pengambilan keputusan investasi pada saham sektor properti.

## 5.2 Saran

Berdasarkan hasil penelitian dan implementasi sistem rekomendasi saham berbasis model XGBoost dengan pendekatan analisis teknikal, terdapat beberapa saran yang dapat dijadikan acuan untuk pengembangan dan penelitian lebih lanjut:

1. Peningkatan Keseimbangan Kelas Data: Mengingat performa model pada kelas *Hold* masih relatif rendah, disarankan untuk memperbanyak jumlah data pada kelas tersebut. Penambahan data historis atau sintesis data berbasis kondisi pasar nyata dapat membantu meningkatkan representasi kelas *Hold* dalam pelatihan model.
2. Mitigasi *Overfitting* Ringan pada Kelas *Hold*: Terdapat indikasi *overfitting* ringan yang terlihat dari rendahnya performa klasifikasi pada kelas *Hold*. Hal ini kemungkinan dipengaruhi oleh jumlah data yang terbatas serta posisi kelas yang berada pada wilayah teknikal yang ambigu antara indikator MA5 dan MA20. Penelitian lanjutan disarankan untuk difokuskan pada peningkatan performa klasifikasi kelas *Hold* melalui pemilihan fitur yang lebih relevan atau penerapan model yang lebih adaptif terhadap ketidakseimbangan kelas.
3. Penerapan Model Alternatif dan *Ensemble*: Selain XGBoost, pemanfaatan model lain seperti LightGBM, CatBoost, ataupun pendekatan ensemble multi-model dapat dieksplorasi untuk melihat potensi peningkatan performa klasifikasi dan ketahanan terhadap ketidakseimbangan data.
4. Integrasi Data *Fundamental*: Sebagai pelengkap analisis teknikal, integrasi dengan data *fundamental* seperti laporan keuangan, rasio profitabilitas, dan sentimen pasar dapat dijadikan acuan untuk membangun sistem rekomendasi yang lebih holistik.

Dengan pengembangan-pengembangan tersebut, sistem rekomendasi saham yang telah dibangun diharapkan dapat memberikan hasil yang lebih akurat, adaptif, serta mampu diimplementasikan secara praktis di lingkungan pasar modal Indonesia.

## DAFTAR PUSTAKA

- Arfisko, H., & Wibowo, A. (2022). Sistem Rekomendasi Film Menggunakan Metode Hybrid Collaborative Filtering Dan Content-based Filtering. *e-Proceeding of Engineering*, 9(Vol. 9 No. 3 (2022): Juni 2022), 2149–2459.
- Dachi, J., & Sitompul, P. (2023). Analisis Perbandingan Algoritma XGBoost dan Algoritma Random Forest Ensemble Learning pada Klasifikasi Keputusan Kredit . *JURNAL RISET RUMPUN MATEMATIKA DAN ILMU PENGETAHUAN ALAM*, 2(2), 87–103. <https://doi.org/10.55606/jurrimipa.v2i2.1336>
- Farild, M., Izzulhaq Sawaji, M., & Poddala, P. (2023). *Analisis teknikal sebagai dasar pengambilan keputusan dalam transaksi saham* (Vol. 25).
- Huda, M., & Satrio, B. (2020). PENGARUH RASIO KEUANGAN TERHADAP HARGA SAHAM PROPERTI DAN REAL ESTATE DI BEI. *Jurnal Ilmu dan Riset Manajemen*, 09(Vol 9 No 7 (2020): Jurnal Ilmu dan Riset Manajemen), 1–20.
- Ibrahim Ahmed Osman, A., Najah Ahmed, A., Chow, M. F., Feng Huang, Y., & El-Shafie, A. (2021). Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Engineering Journal*, 12(2), 1545–1556. <https://doi.org/10.1016/j.asej.2020.11.011>
- Irahadi, D., Sianturi, M., & Kim, S. (2022). PENGGUNAAN INDIKATOR ANALISA TEKNIKAL PADA PASAR SAHAM DI INDONESIA. *JMBI UNSRAT*, 09(Vol. 9 No. 2 (2022): JMBI UNSRAT Volume 9 Nomor 2), 808–827.
- Janaina, N. N., & Yudiantoro, D. (2022). Pengaruh EPS, ROE Dan DER terhadap Harga Saham Properti dan Real Estate yang Terdaftar di JII70. *Al-Kharaj : Jurnal Ekonomi, Keuangan & Bisnis Syariah*, 5(2), 762–771. <https://doi.org/10.47467/alkharaj.v5i2.1314>
- Jange, B. (2022). Prediksi Harga Saham Bank BCA Menggunakan XGBoost. *ARBITRASE: Journal of Economics and Accounting*, 3(2), 231–237. <https://doi.org/10.47065/arbitrase.v3i2.495>
- Liwei, T., Li, F., Yu, S., & Yuankai, G. (2021). Forecast of LSTM-XGBoost in stock price based on Bayesian optimization. *Intelligent Automation and Soft Computing*, 29(3), 855–868. <https://doi.org/10.32604/iasc.2021.016805>
- Mahendra, K., Satyahadewi, N., & Perdana, H. (2022). ANALISIS TEKNIKAL SAHAM MENGGUNAKAN INDIKATOR MOVING AVERAGE CONVERGENCE DIVERGENCE (MACD). Dalam *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)* (Vol. 11, Nomor 1).
- Maulana, M., Hadiana, A., & Umbara, F. (2023). ALGORITMA XGBOOST UNTUK KLASIFIKASI KUALITAS AIR MINUM. *Jurnal Mahasiswa Teknik Informatika*, Vol. 7 No. 5 (Vol. 7 No. 5, Oktober 2023), 3251–3256.

- Mohan, A., Soni, T. K., Vamshisai, P., & Sateesh, B. (2022). deep Q network framework on stock price prediction. *International journal of health sciences*, 3050–3061. <https://doi.org/10.53730/ijhs.v6ns5.9318>
- Munawar, Z., Herdiana, Y., Indah Putri, N., Informatika, M., Informatika, T., Bandung, I., & Bale Bandung, U. (2021). *SISTEM REKOMENDASI HIBRID MENGGUNAKAN ALGORITMA APRIORI MINING ASOSIASI* (Vol. 8, Nomor 1).
- Octasylya, A. (2022). ANALISIS TEKNIKAL SAHAM KONTRUKSI. *Jurnal IPTEK*, 6(2), 23–32. <https://doi.org/10.31543/jii.v6i2.211>
- Oukhouya, H., Kadiri, H., El Himdi, K., & Guerbaz, R. (2024). Forecasting International Stock Market Trends: XGBoost, LSTM, LSTM-XGBoost, And Backtesting XGBoost Models. *Statistics, Optimization and Information Computing*, 12(1), 200–209. <https://doi.org/10.19139/soic-2310-5070-1822>
- Tanto, V., & Kurniawan, T. (2022). Pengembangan Sistem Rekomendasi Investasi Saham berbasis Web (Studi Kasus: Reliance Sekuritas Malang). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(Vol. 6, No. 6, Juni 2022), 2813–2821.
- Uhunmwangho, O. P. (2024). Comparing XGBoost and LSTM Models for Prediction of Microsoft Corp's Stock Price Direction. Dalam / *Mountain Top University Journal of Applied Science and Technology (MUJAST)* (Vol. 4, Nomor 2).

