

BAB I

PENDAHULUAN

1.1. Latar Belakang

Beras merupakan komoditas pangan utama yang memiliki peranan krusial bagi mayoritas penduduk Indonesia, khususnya di Provinsi Jawa Timur yang menjadi salah satu sentra produksi dan konsumsi terbesar [1]. Ketersediaan dan stabilitas harga beras tidak hanya memengaruhi aspek ketahanan pangan regional tetapi juga berkontribusi terhadap indikator ekonomi makro [2]. Oleh karena itu, pemahaman mendalam mengenai dinamika harga beras menjadi esensial bagi berbagai pemangku kepentingan.

Volatilitas harga beras di Jawa Timur merupakan sebuah keniscayaan yang dipicu oleh interaksi kompleks berbagai faktor. Fluktuasi ini secara langsung memengaruhi daya beli masyarakat dan dapat memicu tekanan inflasi, yang pada gilirannya berdampak pada stabilitas ekonomi secara lebih luas [3]. Implikasi yang lebih jauh mencakup potensi gangguan terhadap kesejahteraan petani dan konsumen, serta dapat mempengaruhi stabilitas sosial jika tidak dikelola dengan baik [4].

Selain faktor ekonomi, variabel lingkungan seperti suhu, curah hujan, kelembapan, dan lama penyinaran matahari juga berperan dalam memengaruhi dinamika harga beras, sehingga pemodelan prediksi perlu mempertimbangkan pendekatan multivariat [2], [5]. Faktor-faktor lain seperti luas panen dan tingkat konsumsi per kapita juga turut memberikan kontribusi signifikan [6]. Lebih lanjut, fenomena cuaca ekstrem yang semakin sering terjadi akibat perubahan iklim menambah lapisan kompleksitas dalam upaya memprediksi harga secara akurat dan tepat waktu [7], [8].

Dalam menghadapi tantangan prediksi harga komoditas yang memiliki karakteristik data deret waktu yang kompleks dan *non-linier*, metode-metode konvensional seringkali menunjukkan keterbatasan [11]. Keterbatasan ini mendorong eksplorasi pendekatan yang lebih canggih dan adaptif. Dari perspektif

ilmu data, pengembangan model prediktif yang mampu menangkap pola-pola tersembunyi dalam data historis harga menjadi sebuah kebutuhan mendesak.

Perkembangan pesat dalam ilmu data dan teknologi komputasi telah membuka jalan bagi implementasi algoritma *machine learning* dan *deep learning* untuk analisis data deret waktu. Pendekatan ini menawarkan kapabilitas yang superior dalam memodelkan hubungan yang rumit dan dinamika temporal yang seringkali sulit diidentifikasi oleh teknik statistik tradisional [9]. Keunggulan ini menjadikan *machine learning* sebagai alat yang menjanjikan untuk prediksi harga komoditas.

Salah satu arsitektur *deep learning* yang mendapatkan perhatian signifikan untuk pemodelan data sekuensial adalah *Recurrent Neural Network* (RNN). RNN dirancang khusus untuk memproses urutan data, di mana output dari langkah waktu sebelumnya dapat memengaruhi prediksi pada langkah waktu berikutnya. Varian RNN seperti *Long Short-Term Memory* (LSTM) [10] dan *Gated Recurrent Unit* (GRU) telah dikembangkan untuk mengatasi beberapa keterbatasan RNN standar, terutama terkait dengan masalah *vanishing gradient* pada dependensi jangka panjang.

Secara khusus, *Gated Recurrent Unit* (GRU) menawarkan arsitektur yang lebih sederhana dibandingkan LSTM namun tetap efektif dalam menangkap pola temporal [13]. GRU menggunakan mekanisme *reset gate* dan *update gate* untuk secara selektif memperbarui dan melupakan informasi dari urutan sebelumnya, memungkinkan model untuk mempelajari dependensi jangka panjang secara lebih efisien dan dengan komputasi yang relatif lebih ringan [14], [15]. Efisiensi ini menjadi pertimbangan penting dalam pengembangan model prediktif yang operasional.

Di sisi lain, *Support Vector Regression* (SVR), yang merupakan turunan dari *Support Vector Machine* (SVM) untuk tugas regresi, dikenal sebagai metode *supervised learning* yang tangguh dalam menangani hubungan non-linier antara variabel *input* dan *output* [17]. SVR bekerja dengan prinsip mencari *hyperplane* optimal yang dapat memetakan data *input* ke dalam ruang fitur berdimensi tinggi, dengan tujuan meminimalkan *margin* kesalahan prediksi [16]. Keunggulan SVR

terletak pada kemampuannya menghasilkan generalisasi yang baik bahkan dengan data pelatihan yang terbatas seperti SVR *grid search* dengan menggunakan *kernel Radial Basis Function (RBF)* yang terbukti optimal dalam memodelkan hubungan *nonlinier* pada data ekonomi yang memiliki pola fluktuatif dan tidak stabil [17].

Berbagai penelitian telah dilakukan untuk memprediksi harga komoditas menggunakan pendekatan *machine learning* dan *deep learning*, yang menunjukkan pergeseran dari metode statistik konvensional [21], [39]. Model berbasis *Recurrent Neural Network* (RNN), khususnya *Long Short-Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU), telah menunjukkan kinerja yang menjanjikan dalam menangani data deret waktu. Sanjaya dan Heksaputra [1] menerapkan LSTM untuk memprediksi rerata harga beras grosir di Indonesia. Secara lebih spesifik untuk wilayah Jawa Timur, penelitian oleh [4] menggunakan varian LSTM, yaitu *Stacked LSTM*, untuk memodelkan harga beras medium. Studi lain oleh [8] juga membandingkan performa LSTM dengan *Multilayer Perceptron* (MLP) untuk kasus prediksi harga beras.

Penerapan luas GRU, yang menawarkan arsitektur lebih sederhana dan efisiensi komputasi, juga terlihat dalam berbagai studi perbandingan dengan LSTM untuk prediksi data deret waktu seperti harga emas [13], [14] dan saham [35], di mana GRU seringkali menunjukkan hasil yang kompetitif. Di sisi lain, *Support Vector Regression* (SVR) juga menjadi metode yang populer karena kemampuannya dalam menangani hubungan *non-linier*. Beberapa penelitian membandingkan SVR dengan model *deep learning*, seperti yang dilakukan oleh [16] pada data harga saham, yang menunjukkan bahwa setiap model memiliki karakteristik keunggulannya masing-masing.

Menyadari keterbatasan model tunggal, beberapa peneliti mulai mengeksplorasi pendekatan *hibrida* atau *ensemble*. Penelitian oleh [15] membandingkan kinerja GRU dan SVR secara terpisah untuk prediksi data penjualan, yang secara tidak langsung mengindikasikan potensi untuk menggabungkan kekuatan keduanya. Secara signifikan, penelitian oleh [18] telah mengusulkan model hibrida GRU-SVR dan menunjukkan kinerjanya yang baik. Namun, penelitian mereka diterapkan untuk tujuan yang berbeda, yaitu *monitoring*

Indeks Luas Daun (*Leaf Area Index* - LAI) pada tanaman padi menggunakan data penginderaan jauh, bukan untuk prediksi harga pasar secara langsung.

Selain dari tinjauan pustaka tersebut, teridentifikasi beberapa celah penelitian (*research gap*) yang signifikan. Meskipun model *hibrida* GRU-SVR telah terbukti potensial [18], Berdasarkan literatur yang dirujuk dalam penelitian ini, hanya dua penelitian yang secara spesifik mengkaji prediksi harga beras di wilayah Provinsi Jawa Timur menggunakan pendekatan berbasis deep learning, sementara penelitian lainnya masih berfokus pada cakupan nasional atau wilayah yang berbeda. [1], [4]. Sebagian besar penelitian sebelumnya lebih berfokus pada pendekatan model tunggal atau menggunakan data harga historis saja (pendekatan univariat), sehingga belum mengintegrasikan informasi eksternal seperti variabel cuaca yang berpotensi memengaruhi dinamika harga [7], [41]. Penelitian ini akan mengisi celah tersebut dengan mengintegrasikan *variabel eksogen* yang relevan yaitu data cuaca historis seperti suhu, curah hujan, kelembapan, dan lama penyinaran matahari ke dalam kerangka kerja model untuk meningkatkan konteks dan potensi akurasi prediksi. Terakhir, terdapat kesenjangan yang lebih luas antara hasil riset akademis dan pemanfaatan praktisnya. Mayoritas penelitian cenderung berhenti pada validasi metrik akurasi [4], tanpa menyediakan jembatan menuju implementasi nyata. Untuk mengatasi hal ini, penelitian ini tidak hanya berfokus pada pengembangan model, tetapi juga bertujuan menghasilkan luaran berupa Antarmuka Pengguna Grafis (GUI) interaktif menggunakan *Streamlit* [39], yang dirancang untuk menjembatani kesenjangan antara analisis teknis dan pengambilan keputusan oleh para pemangku kepentingan.

Berdasarkan celah penelitian tersebut, pengembangan pendekatan *ensemble* yang mengombinasikan keunggulan beberapa model menjadi relevan untuk dilakukan penelitian. Penelitian ini mengusulkan sebuah model *ensemble* yang mengintegrasikan GRU dan SVR. Ide dasar dari pendekatan *ensemble* GRU-SVR adalah untuk memanfaatkan kemampuan GRU dalam mengekstraksi fitur-fitur *temporal* dan dependensi jangka panjang dari data deret waktu harga beras. *Output* atau representasi fitur yang dihasilkan oleh GRU kemudian akan dijadikan sebagai *input* bagi model SVR untuk melakukan prediksi harga akhir, sehingga SVR dapat

fokus pada pemodelan hubungan non-linier dari fitur-fitur temporal yang telah diproses [18].

Pendekatan *ensemble* GRU-SVR telah menunjukkan kinerja yang menjanjikan dalam berbagai domain aplikasi, yang mengindikasikan potensinya untuk menghasilkan akurasi prediksi yang lebih superior dibandingkan penggunaan model GRU atau SVR secara individual [18]. Dalam konteks prediksi harga beras, model *ensemble* ini diharapkan mampu memberikan estimasi yang lebih akurat dengan mempertimbangkan karakteristik data yang kompleks dan berbagai faktor eksternal seperti variabel cuaca (suhu rata-rata, curah hujan, kelembapan, dan lama penyinaran matahari) sebagai fitur prediktor tambahan.

Secara spesifik, penelitian ini akan berfokus pada implementasi model *ensemble* GRU-SVR untuk memprediksi harga beras kualitas medium di Provinsi Jawa Timur. Data historis harga harian dari periode 2021 hingga 2025 akan diperoleh dari Sistem Informasi Ketersediaan dan Perkembangan Harga Bahan Pokok (SISKAPERBAPO), sementara data cuaca akan bersumber dari lembaga klimatologi jawa timur resmi data *online* BMKG. Data klimatologi yang digunakan dalam penelitian ini merupakan data hasil pengamatan meteorologi yang diperoleh dari jaringan stasiun cuaca Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). Secara ilmiah, klimatologi didefinisikan sebagai nilai statistik jangka panjang, khususnya nilai rata-rata, yang dihitung dari hasil observasi meteorologi pada berbagai stasiun cuaca [19]. Sebagaimana didefinisikan dalam pedoman *World Meteorological Organization* (WMO), bahwa klimatologi merupakan statistik jangka panjang dari pengamatan meteorologi yang dikumpulkan dari berbagai stasiun cuaca [20]. Oleh karena itu, data klimatologi yang digunakan bersifat representatif secara regional dan mampu menggambarkan kondisi iklim Provinsi Jawa Timur secara menyeluruh. Aspek penting dari penelitian ini adalah penekanan pada validasi dan evaluasi model dari sudut pandang ilmu data untuk memastikan robustisitas dan reliabilitas prediksi.

Untuk meningkatkan aksesibilitas dan interpretasi hasil model bagi pengguna non-teknis, penelitian ini juga akan mencakup pengembangan Antarmuka Pengguna Grafis (*Graphical User Interface* - GUI) menggunakan platform

Streamlit. Diharapkan, penelitian ini tidak hanya memberikan kontribusi metodologis dalam penerapan teknik *ensemble learning* untuk prediksi deret waktu di bidang ilmu data, tetapi juga menghasilkan luaran praktis yang dapat mendukung upaya stabilisasi harga pangan dan pengambilan keputusan berbasis data oleh pemerintah daerah serta pemangku kepentingan terkait lainnya.

1.2. Rumusan Masalah

Merujuk pada latar belakang yang telah dijelaskan sebelumnya, maka rumusan masalah dalam penelitian ini dapat disusun sebagai berikut:

- 1) Bagaimana karakteristik data harga beras medium dan variabel cuaca harian di Provinsi Jawa Timur pada periode 2021–2025 yang digunakan dalam penelitian ini?
- 2) Bagaimana proses pra-pemrosesan data harga beras dan variabel cuaca, termasuk normalisasi dan pembentukan data deret waktu, sebelum digunakan dalam pemodelan?
- 3) Bagaimana penerapan model *Gated Recurrent Unit* (GRU) sebagai pengekstrak fitur *temporal* dalam memodelkan data harga beras di Provinsi Jawa Timur?
- 4) Bagaimana penerapan *Enhanced Feature Engineering* (EFE) dalam menghasilkan fitur statistik tambahan untuk mendukung proses prediksi harga beras?
- 5) Bagaimana proses optimasi model *Support Vector Regression* (SVR) menggunakan *grid search* dalam memprediksi harga beras berdasarkan fitur hasil ekstraksi GRU dan EFE?
- 6) Bagaimana performa model *ensemble* GRU–SVR dalam memprediksi harga beras di Provinsi Jawa Timur berdasarkan nilai *Mean Absolute Percentage Error* (MAPE)?
- 7) Bagaimana hasil *forecasting* harga beras medium di Provinsi Jawa Timur untuk periode tiga puluh hari ke depan berdasarkan model terbaik yang diperoleh?

- 8) Bagaimana pengembangan *Graphical User Interface* (GUI) berbasis *Streamlit* untuk mempermudah visualisasi, interaksi, dan interpretasi hasil prediksi harga beras oleh pengguna?

1.3. Batasan Masalah

Pembatasan masalah dalam penelitian ini dimaksudkan untuk menetapkan cakupan pembahasan, yaitu sebagai berikut:

- 1) Penelitian ini difokuskan pada objek harga beras medium di Provinsi Jawa Timur dan menggunakan data deret waktu harian yang mencakup periode dari tahun 2021 hingga 2025.
- 2) Penelitian ini hanya menggunakan data historis harga dan kondisi cuaca tanpa mempertimbangkan faktor eksternal lain seperti perubahan kebijakan atau kondisi ekonomi yang mungkin memengaruhi harga pangan.
- 3) Pengukuran evaluasi model hanya menggunakan metrik MAPE (*Mean Absolute Percentage Error*).
- 4) Optimasi parameter *Support Vector Regression* (SVR) dalam penelitian ini dibatasi kombinasi rentang nilai parameter $C = 100 - 10000$, $\gamma = 0,0001 - 0,1$, dan $\epsilon = 0,01 - 0,1$ yang telah ditentukan melalui metode *grid search*, untuk menjaga efisiensi komputasi, menghindari *overfitting* dan stabilitas performa model.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dirumuskan, penelitian ini bertujuan untuk membangun dan mengevaluasi model *ensemble* GRU – SVR dalam memprediksi harga beras medium di Provinsi Jawa Timur, serta mengimplementasikan antarmuka pengguna grafis untuk memudahkan visualisasi, interpretasi, dan pemanfaatan hasil prediksi oleh pengguna.

1.5. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat bagi pihak-pihak yang membutuhkan baik teoritis maupun praktis, diantaranya:

1. Manfaat teoritis

Penelitian Penelitian ini diharapkan dapat memberikan solusi inovatif dalam pengolahan dan analisis prediksi harga beras. Selain itu, hasil dari penelitian ini diharapkan dapat berfungsi sebagai dasar bagi penelitian selanjutnya yang berkaitan dengan peramalan harga beras di sektor ekonomi dengan menggunakan metode *ensemble* GRU - SVR.

2. Manfaat praktis

- a. Bagi Masyarakat Umum: Sistem yang dikembangkan diharapkan dapat berkontribusi dalam meningkatkan literasi masyarakat mengenai sektor ekonomi khususnya mengenai pasar, karena jika diaplikasikan secara luas masyarakat dapat berperan dalam mendukung pertumbuhan ekonomi yang stabil dan berkelanjutan dengan kemampuan membuat keputusan yang lebih bijak sehingga mendukung stabilitas ekonomi dalam jangka panjang.
- b. Bagi Pemangku Kebijakan: Sistem yang dikembangkan diharapkan dapat membantu dalam meramalkan harga beras dan menganalisis manajemen ekonomi dalam mengambil kebijakan khususnya yang berkaitan dengan pasar.
- c. Bagi Akademisi dan Peneliti: Sistem yang dikembangkan diharapkan dapat menjadi acuan atau alat bantu penelitian dan pembelajaran yang berguna untuk memahami lebih mendalam mengenai meramalkan harga beras. Peneliti berikutnya diharapkan dapat menembangkan model yang lebih kompleks.

BAB II

TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Dalam memahami konteks dan relevansi penelitian ini, penting untuk meninjau beberapa penelitian terdahulu yang relevan dengan penelitian penulis. Penelitian-penelitian ini memberikan wawasan berharga mengenai metodologi, temuan, dan model yang digunakan dalam prediksi harga beras. Berikut beberapa temuan penelitian terdahulu.

Tabel 2. 1 Penelitian Terdahulu

No	Profil Pustaka	Metode dan Temuan
1.	Judul: Prediksi Harga Komoditas Pangan Menggunakan Algoritma <i>Long Short-Term Memory</i> (LSTM) Penulis: 1. Rizki Mugi Setya Adi 2. Aulia cahya reni 3. Khofifah Amalia 4. Dwi Rolliawati Identitas artikel: <i>Building of Informatics, Technology and Science</i> (BITS), Vol. 4, No. 2, September 2022, Hal. 1137–1145.	Metode: Penelitian ini menggunakan model Algoritma <i>Long Short-Term Memory</i> (LSTM) untuk memprediksi harga komoditas pangan. Temuan: Penelitian ini menemukan bahwa hasil yang didapat, sanggup melakukan prediksi harga komoditas pangan. Sementara itu, model prediksi dengan <i>epoch</i> 50 menunjukkan (RMSE) terendah dengan nilai 79.19%.
2	Judul: <i>Prediction of Vehicle Driver's Facial Air Temperature With SVR, ANN, and GRU</i> Penulis:	Metode: Penelitian ini menggunakan metode (<i>SVR</i>), (<i>ANN</i>), dan (<i>GRU</i>) metode ini dipilih karena kemampuannya dalam memprediksi Suhu udara di wajah yang memiliki dampak

No	Profil Pustaka	Metode dan Temuan
	<p>1. Xiaohan Zhang 2. Yichun Wang 3. Xinglei He 4. Hongzeng Ji 5. Yawen Li 6. Xiuhui Duan 7. Fen Gui</p> <p>Identitas artikel: <i>IEEE ACCESS Digital Object Identifier</i>, Vol. 10, 1109, Hal. 2021 - 2022.</p>	<p>signifikan terhadap kenyamanan termal pengemudi. Model <i>Machine Learning</i> terbukti sangat efektif dalam memprediksi suhu. Dalam penelitian ini, tiga model digunakan untuk memprediksi suhu wajah pengemudi pada serangkaian kendaraan tertentu.</p> <p>Temuan: Penelitian tersebut menemukan bahwa evaluasi kinerja model dilakukan menggunakan RMSE, MAE, dan koefisien determinasi (R^2). MAE untuk ketiga model menunjukkan SVR: 0.8096, ANN: 0.4984, GRU: 0.7289 dalam kondisi terlatih, dan SVR: 1.0946, ANN: 0.7878, GRU: 0.7837 dalam kondisi tidak terlatih. Hasil menunjukkan bahwa ANN memiliki kinerja terbaik di antara ketiga model. Selain itu, akurasi model menurun saat <i>dataset</i> diuji dalam kondisi baru. Oleh karena itu, ANN mungkin menjadi metode yang lebih disukai untuk prediksi suhu udara wajah pengemudi.</p>
3	<p>Judul: <i>Remote Sensing Monitoring of Rice LAI Based on GRU-SVR.</i></p> <p>Penulis: 1. Chuhan Qi 2. Xiaofei Wang</p> <p>Identitas artikel:</p>	<p>Metode: Penelitian ini menggunakan dua metode yaitu GRU dan SVR untuk Pengukuran indeks luas daun (LAI) padi secara cepat dan akurat sangat penting untuk mengevaluasi pertumbuhan padi dan meningkatkan efisiensi manajemen lahan. Berdasarkan model PROSAIL, makalah ini</p>

No	Profil Pustaka	Metode dan Temuan
	<p><i>Journal of Network Intelligence. Taiwan Ubiquitous Information</i>, Vol. 7, No. 4, November 2022.</p>	<p>mengusulkan metode untuk pertama-tama mengekstraksi pita karakteristik menggunakan Spektrum Korelasi Dua-Dimensi (2DCOS), dan kemudian membangun jaringan GRU-SVR untuk membalikkan LAI padi.</p> <p>Temuan:</p> <p>Penelitian ini menemukan bahwa, Hasil eksperimen menunjukkan bahwa metode yang diusulkan memiliki akurasi inversi yang lebih tinggi ($R^2 = 0,92648$, RMSE = 0,63732), dan lebih cocok untuk tugas inversi LAI padi yang sebenarnya. Selain itu, peta distribusi pengelompokan LAI padi yang dihasilkan di area studi memungkinkan penguasaan situasi pertumbuhan padi secara makro, yang dapat memberikan panduan dan perspektif yang lebih luas untuk pengelolaan lahan selanjutnya.</p>
4	<p>Judul: Penerapan Regresi Linear Untuk Prediksi Harga Beras Di Indonesia.</p> <p>Penulis:</p> <ol style="list-style-type: none"> 1. Veri Arinal 2. Muhammad Azhari <p>Identitas artikel:</p>	<p>Metode:</p> <p>Penelitian ini menerapkan <i>Data Mining</i> dengan pendekatan SEMMA (<i>Sample, Explore, Modify, Model, Assess</i>). Tujuan utama dari penelitian ini adalah untuk menerapkan metode Regresi Linear guna memprediksi harga beras serta menghitung nilai RMSE. Data yang digunakan dalam</p>

No	Profil Pustaka	Metode dan Temuan
	<p>Jurnal Sains dan Teknologi – Vol. 5, No. 1, September 2023 – Hal. 341 - 346.</p>	<p>studi ini mencakup harga beras dari tahun 2021 hingga 2023.</p> <p>Temuan:</p> <p>Hasil prediksi harga beras yang dihasilkan dari penerapan Regresi Linier menunjukkan bahwa nilai prediksi cukup dekat dengan harga aktual. Penelitian ini juga mengukur performa model regresi linear dengan membagi data menjadi 70% untuk pelatihan dan 30% untuk pengujian. Dari pembagian data tersebut, diperoleh nilai RMSE (<i>Root Mean Squared Error</i>) sebesar 337.996 ± 0.000. Pengujian ini menggunakan perhitungan RMSE untuk mengevaluasi akurasi prediksi harga beras.</p>
5	<p>Judul: <i>Comparison Four Kernels of SVR to Predict Consumer Price Index</i></p> <p>Penulis:</p> <ol style="list-style-type: none"> 1. M F Rohmah 2. I K G D Putra 3. R S Hartati 4. L Ardiantoro <p>Identitas artikel: <i>Journal of Physics: Conference Series, Vol. 1737, No. 1, 2021.</i></p>	<p>Metode:</p> <p>Dalam penelitian ini, <i>Support Vector Regression</i> (SVR) digunakan untuk melakukan peramalan data deret waktu dengan membandingkan empat jenis <i>kernel</i>, yaitu <i>linear</i>, <i>polynomial</i>, <i>sigmoid</i>, dan <i>Radial Basis Function (RBF)</i>. Tujuan utama penelitian adalah mengidentifikasi <i>kernel</i> SVR yang paling optimal dalam memodelkan hubungan <i>nonlinier</i> pada data ekonomi yang memiliki pola fluktuatif dan tidak stabil. Proses pemodelan dilakukan dengan mengatur parameter SVR dan mengevaluasi kinerja masing-masing</p>

No	Profil Pustaka	Metode dan Temuan
		<p><i>kernel</i> menggunakan metrik kesalahan prediksi.</p> <p>Temuan:</p> <p>Hasil penelitian menunjukkan bahwa <i>kernel RBF</i> menghasilkan kinerja terbaik dibandingkan <i>kernel</i> lainnya. <i>Kernel RBF</i> memperoleh nilai RMSE sebesar 0.147 dan MAPE sebesar 1.52%, yang lebih rendah dibandingkan <i>kernel linear, polynomial, dan sigmoid</i>. Nilai kesalahan yang lebih kecil ini menunjukkan bahwa <i>kernel RBF</i> memiliki kemampuan yang lebih baik dalam menangkap pola <i>nonlinier</i> pada data deret waktu. Temuan tersebut menegaskan bahwa SVR dengan <i>kernel RBF</i> efektif digunakan untuk peramalan data ekonomi yang bersifat fluktuatif.</p>

Berdasarkan dari profil pustaka yang disajikan, beberapa penelitian menunjukkan kemajuan dalam penggunaan berbagai model pembelajaran mesin untuk prediksi harga komoditas, suhu pengemudi, dan luas daun padi, serta analisis jaringan sensor nirkabel. Namun, terdapat beberapa *research gap* yang belum sepenuhnya terjawab oleh penelitian-penelitian ini.

Pertama, meskipun model LSTM telah digunakan secara efektif dalam memprediksi harga komoditas pangan yang ditulis oleh [38] ANN juga terbukti menjadi metode yang lebih baik dalam memprediksi suhu pengemudi yang ditulis oleh [35], penelitian mengenai model yang menggabungkan kemampuan temporal LSTM atau GRU dengan akurasi tinggi dari model seperti SVR atau ANN belum sepenuhnya dieksplorasi. Misalnya, penelitian [18] telah memperkenalkan kombinasi GRU dan SVR untuk memprediksi indeks luas daun padi, yang

menunjukkan performa yang baik dalam akurasi. Namun, potensi model ini untuk aplikasi prediksi harga komoditas beras belum banyak dijadikan objek kajian. Kedua, penelitian sebelumnya yang menggunakan Regresi Linear untuk prediksi harga beras di Indonesia yang ditulis oleh [39] menunjukkan akurasi yang cukup baik. Namun, model yang lebih canggih seperti GRU-SVR, yang memiliki keunggulan dalam menangkap pola *non-linier* dan *tren* musiman, belum digunakan secara komprehensif untuk prediksi komoditas pangan di Indonesia.

Berdasarkan penelitian-penelitian terdahulu tersebut, dapat disimpulkan bahwa metode *deep learning* seperti LSTM dan GRU efektif dalam memodelkan data deret waktu, sementara SVR menunjukkan performa yang baik dalam regresi non-linear. Namun, penerapan model *ensemble* yang mengombinasikan GRU sebagai pengekstrak fitur temporal dan SVR sebagai *regresor* akhir pada prediksi harga beras, khususnya dengan melibatkan variabel cuaca dan implementasi sistem berbasis GUI, masih terbatas. Hal ini menjadi dasar bagi penelitian ini untuk mengembangkan pendekatan ensemble GRU-SVR yang lebih komprehensif.

2.2. Kerangka Teori

Pada penelitian ini, terdapat dasar teori yang menjadi landasan dalam melakukan analisis dan pemodelan data *time series*. Berikut adalah penjelasan singkat dari teori-teori yang relevan.

2.2.1. Harga beras

Indonesia merupakan negara agraris yang sangat bergantung pada sektor pertanian untuk memenuhi kebutuhan pokok masyarakatnya. Salah satu tantangan utama yang dihadapi saat ini adalah ketidakstabilan harga bahan pangan, terutama beras, yang dapat memengaruhi kesejahteraan masyarakat secara luas [1]. Terpenuhinya kebutuhan pangan menjadi indikator penting dalam menjaga stabilitas ekonomi dan sosial. Namun, fluktuasi harga yang tajam, perubahan kebijakan pemerintah, serta faktor cuaca dan iklim menjadi penyebab utama tidak stabilnya harga pangan [2], [3]. Fluktuasi harga yang tidak terkendali dapat memicu inflasi, menurunkan daya beli masyarakat, dan memengaruhi stabilitas perekonomian nasional secara keseluruhan [4].

Sebagian besar penduduk Indonesia bermata pencaharian sebagai petani, dan kesuburan lahan di Indonesia yang tinggi disebabkan oleh lokasinya yang berada di iklim tropis [5]. Namun, meskipun kondisi geografis menguntungkan, harga beras masih menunjukkan ketidakstabilan yang dipengaruhi oleh berbagai faktor seperti jumlah penduduk, inflasi, luas panen, konsumsi, serta ketidakpastian cuaca [6],[7]. Dalam konteks ini, pengembangan teknologi berbasis *machine learning* telah memberikan solusi inovatif, khususnya dalam analisis data deret waktu (time series), yang menjadi pendekatan utama untuk memahami pola historis harga beras[8].

Beras merupakan komoditas utama dan makanan pokok di Indonesia. Oleh karena itu, prediksi harga beras menjadi penting untuk mendukung pengambilan keputusan yang lebih baik bagi petani, konsumen, dan pemerintah [9]. Permintaan beras sangat sensitif terhadap perubahan harga dan pengeluaran masyarakat, sebagaimana dibuktikan dalam studi yang menunjukkan pengaruh signifikan antara harga beras dan pengeluaran rata-rata terhadap permintaan nasional [10].

2.2.2. Analisis Deret Waktu

Menurut [21] data deret waktu merupakan rangkaian pengamatan yang berurutan dalam waktu. Sedangkan, analisis deret waktu berguna untuk melakukan analisis data dengan mempertimbangkan pengaruh waktu. Pengumpulan data dilakukan secara berkala berdasarkan urutan waktu, yang dapat dinyatakan dengan jam, hari, minggu, bulan, triwulan, dan tahun. Pola data bisa dibedakan menjadi 4 jenis, yaitu musiman, siklus, tren, dan tidak teratur [22], [42].

2.2.3. Analisis Data Eksploratif

Menurut [23] Analisis Data Eksploratif merupakan pendekatan pembelajaran *machine learning* yang digunakan untuk memahami data sebelum pemodelan. EDA dapat membantu mengidentifikasi pola, menemukan anomali, menguji hipotesis, dan memverifikasi hipotesis menggunakan statistik dan visualisasi grafis. Proses ini penting untuk mendapatkan pengetahuan mendalam tentang data dan memastikan bahwa pemodelan *machine learning* yang dilakukan

pada data dapat mudah dipahami dengan baik. [24] menyebutkan bahwa banyak pustaka yang terdapat di Python untuk EDA salah satunya adalah *library* pandas digunakan untuk membaca data dari berbagai format, membersihkan dan merapihkan. Pada penelitian ini terdapat beberapa tahapan Analisis Data Eksploratif:

a. Pengumpulan Data

Langkah penting dalam proses penelitian merupakan pengumpulan data, yang mencakup pengambilan teknik dan sumber daya yang baik untuk mengumpulkan data yang diperlukan guna menjawab pertanyaan penelitian dan menguji hipotesis. Ada banyak cara lain untuk mengumpulkan data, seperti melalui survei, wawancara, observasi, dan analisis dokumen. Tujuan utamanya adalah untuk memastikan data akurat, relevan, dan dapat diandalkan untuk analisis lebih lanjut [25].

b. Pembersihan Data

Pembersihan data merupakan proses penting dalam analisis data yang bertujuan untuk mengidentifikasi dan memperbaiki kesalahan atau ketidakkonsistenan pada kumpulan data. Fase ini mencakup berbagai aktivitas, seperti penanganan data yang hilang, penghapusan duplikat, koreksi kesalahan format data, dan memastikan bahwa data sesuai dengan standar yang diperlukan untuk analisis. Pembersihan data yang efektif meningkatkan kualitas dan keandalan hasil analisis, menjadikannya sebuah langkah penting dalam proses pengolahan data [26].

c. Transformasi Data

Proses mengubah data dari banyaknya sumber ke dalam format konsisten yang sesuai untuk analisis lebih mendalam dikenal sebagai transformasi data. Langkah ini mencakup prosedur seperti normalisasi data, agregasi, pemfilteran, dan penggabungan [46]. Tujuan transformasi data adalah untuk memastikan bahwa data tersebut dapat diandalkan, komprehensif, dan relevan untuk digunakan dalam analisis konteks atau integrasi dengan kumpulan data lainnya. Prosedur ini sangat penting untuk memperoleh hasil yang dapat diandalkan dari data yang dikumpulkan [27].

d. Pemilihan Fitur

Tujuan pemilihan fitur merupakan langkah penting dalam machine learning, untuk mengidentifikasi subkumpulan fitur yang paling relevan. Hal ini juga sangat memengaruhi pilihan fitur prediktif, meningkatkan performa model dengan menghapus informasi yang tidak perlu atau berlebihan. Selain itu, hal ini dapat mempersingkat waktu komputasi, meminimalkan overfitting, dan menyederhanakan model, yang semuanya untuk meningkatkan efektivitas dalam pemodelan [28].

e. Pembagian Data

Pembagian data merupakan tahapan penting dalam proses machine learning yang melibatkan pemisahan kumpulan data menjadi 2 bagian, seperti data training dan data testing. Hal ini untuk mengembangkan model yang dapat belajar dari data pelatihan dan kemudian mengujinya dengan data pengujian untuk menunjukkan kinerjanya merupakan tujuan dari pembagian data ini. Proses ini untuk mencegah overfitting model, teknik berbagi data yang optimal dapat membantu menjamin bahwa model memiliki generalisasi yang baik [29].

2.2.4. *Forecasting*

Menurut buku *Forecasting: Theory and Practice* [30], peramalan adalah proses meramal nilai masa depan berdasarkan pola dalam data sejarah. Peramalan sering digunakan dalam bisnis, ekonomi, teknik, dan ilmu sosial untuk memfasilitasi pengambilan keputusan berdasarkan prediksi di masa depan [45], [47]. Berikut merupakan metode yang dijelaskan buku ini yaitu metode runtun waktu (*Time Series*) adalah metode kuantitatif yang paling banyak digunakan dalam peramalan. Data deret waktu merupakan kumpulan observasi yang dilakukan pada interval waktu tertentu. Berikut merupakan persamaan dasar untuk model *Simple Moving Average*:

$$\hat{y}_{t+1} = \frac{1}{n} \sum_{i=0}^{n-1} y_{t-i} \quad (2.1)$$

Dimana:

\hat{y}_{t+1}	: nilai hasil peramalan pada periode ke- $t + 1$,
n	: jumlah periode pengamatan dalam perhitungan rata-rata,
y_{t-i}	: nilai aktual pada periode ke- $(t-i)$
Σ	: penjumlahan seluruh nilai dalam window

Persamaan (2.1) menunjukkan bahwa nilai peramalan pada periode berikutnya diperoleh dengan menghitung rata-rata dari n nilai aktual pada periode sebelumnya. Metode ini mengasumsikan bahwa pola data masa lalu bersifat relatif stabil, sehingga nilai rata-rata historis dapat digunakan sebagai estimasi nilai di masa depan. Meskipun metode *Simple Moving Average* mudah diimplementasikan dan efektif untuk mereduksi fluktuasi acak, metode ini memiliki keterbatasan dalam menangkap pola *non-linear* dan ketergantungan *temporal* yang kompleks pada data deret waktu.

2.2.5. *Machine Learning*

Menurut IBM, *machine learning* merupakan cabang dari kecerdasan buatan (AI) yang berfokus pada penggunaan algoritma untuk mengolah data dan menyelesaikan masalah. Dengan algoritma ini, komputer dapat meniru cara manusia mengambil keputusan dan terus belajar untuk meningkatkan akurasinya [31]. Tujuan utamanya adalah menciptakan sistem yang bisa belajar dan mengambil keputusan secara mandiri dari data tanpa perlu diprogram ulang, serta mampu beradaptasi dengan informasi baru. Sistem ini tidak hanya membuat keputusan terbaik, tetapi juga terus berkembang. Tujuan dari *machine learning* adalah menghasilkan prediksi yang mudah dipahami manusia melalui pengolahan data. Untuk bisa meniru cara berpikir manusia, sistem ini memerlukan data yang cukup agar dapat membuat keputusan akurat tanpa campur tangan manusia. Sebagai contoh, komputer dapat secara mandiri menganalisis dan memprediksi harga beras di Jawa Timur.

Berdasarkan metode belajarnya, *machine learning* terbagi menjadi empat teknik utama dalam poin-poin dibawah ini [31].

- Pertama, *Supervised Machine Learning* (Pembelajaran Terarah): Teknik ini memerlukan data yang sudah diberi "label" atau jawaban yang benar.

Tujuannya adalah agar algoritma belajar dari data berlabel tersebut untuk membuat prediksi yang akurat. Ketika data baru dimasukkan, hasilnya harus sesuai atau sangat mendekati data utama yang telah dipelajari.

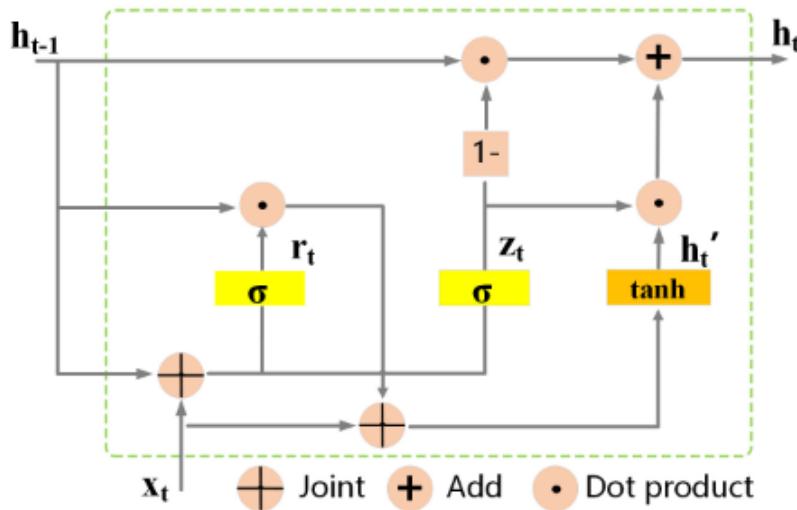
- Kedua, *Unsupervised Machine Learning* (Pembelajaran Tak Terarah): Berbeda dengan sebelumnya, teknik ini menggunakan data tanpa label. Tujuannya adalah untuk menemukan pola atau struktur tersembunyi di dalam data. Contoh paling umum adalah pengelompokan (*clustering*), di mana algoritma mengkategorikan data ke dalam kelompok-kelompok yang memiliki kemiripan.
- Ketiga, *Semi-supervised Machine Learning* (Pembelajaran Semi Terarah): Teknik ini adalah gabungan antara *supervised* dan *unsupervised*. Di sini, hanya sebagian kecil data yang memiliki label, sementara sebagian besar lainnya tidak. Prosesnya memanfaatkan data berlabel yang sedikit itu untuk membantu mengelompokkan dan memahami data tidak berlabel yang jumlahnya lebih banyak. Keuntungannya adalah efisiensi waktu karena tidak semua data perlu dilabeli secara manual.
- Keempat, *Reinforcement Learning* (Pembelajaran Penguatan): Teknik ini mengadopsi paradigma pembelajaran di mana sebuah agen (*agent*) belajar untuk membuat keputusan dengan berinteraksi dalam sebuah lingkungan. Agen tersebut akan menerima umpan balik berupa "ganjaran" (*reward*) atau "hukuman" (*punishment*) atas tindakannya, dengan tujuan untuk memaksimalkan total ganjaran dalam jangka panjang. Tujuan akhirnya adalah untuk mengembangkan model algoritma yang efektif dengan cara mengevaluasi kelebihan dan kekurangan dari setiap tindakan yang diambil.

2.2.6. Gate Recurrent Unit (GRU)

Gated Recurrent Unit (GRU), yang pertama kali diperkenalkan oleh Cho pada tahun 2014, adalah varian dari *Long Short-Term Memory* (LSTM) yang lebih sederhana dan efisien [32]. Sederhananya, GRU adalah varian LSTM yang mengurangi kompleksitas jaringan sambil tetap mempertahankan efektivitasnya dalam menangani data sekuensial dan deret waktu (*time series*). Karena lebih

sederhana, waktu komputasi dan pelatihan GRU menjadi lebih cepat. Hal ini membuatnya sangat cocok untuk lingkungan prediksi keuangan secara *real-time*.

Dalam jaringan saraf tiruan berulang (*recurrent neural networks*) seperti GRU, tujuan utamanya adalah meminimalkan kesalahan prediksi melalui proses optimisasi. Ini dicapai dengan memperbarui bobot (*weights*) model secara berulang menggunakan algoritma *backpropagation through time* [34].



Gambar 2. 1 GRU hidden unit structure diagram

Gambar 2.1 menunjukkan arsitektur GRU, yang memiliki dua gerbang (*gates*) yaitu *reset gate* dan *update gate* [34]. *Reset gate* mengontrol seberapa banyak informasi dari tahap sebelumnya yang akan diteruskan ke tahap berikutnya, sedangkan *update gate* menentukan seberapa banyak informasi yang akan diperbarui atau dipertahankan. Kombinasi kedua gerbang ini membuat GRU lebih efektif dalam menangkap informasi temporal dari data tanpa memerlukan struktur gerbang yang kompleks [34].

Arsitektur *Gated Recurrent Unit* (GRU) mencakup *update gate*, yang mengatur seberapa banyak informasi dari *hidden state* sebelumnya yang harus dibawa ke langkah waktu (*time step*) saat ini. Gerbang ini menerima masukan dari *hidden state* sebelumnya dan *input* saat ini, lalu menghasilkan nilai antara 0 dan 1 yang menunjukkan proporsi informasi yang akan dipertahankan. Mekanisme gerbang ini membantu mengatasi masalah *vanishing gradient* yang umum dihadapi

pada RNN tradisional, sehingga memungkinkan penanganan dependensi jangka panjang yang lebih baik [35]. Rumus untuk menghitung *Update Gate* dapat dinyatakan sebagai berikut.

$$z_t = \sigma (W_z \cdot [h_{t-1}, x_t]) \quad (2.2)$$

z_t merepresentasikan *Update Gate*. Nilainya dihitung menggunakan beberapa komponen: W_z adalah bobot (*weight*), σ adalah fungsi aktivasi Sigmoid, h_{t-1} adalah *hidden state* dari langkah waktu sebelumnya, dan x_t adalah *input* pada waktu t.

Reset Gate dalam GRU menentukan seberapa banyak informasi dari *hidden state* sebelumnya yang harus diabaikan atau dilupakan. Gerbang ini menghasilkan nilai antara 0 dan 1, di mana nilai yang mendekati 0 akan mengurangi pengaruh informasi masa lalu, sedangkan nilai yang mendekati 1 akan mempertahankannya. Dengan mengabaikan bagian dari *hidden state* secara selektif, *reset gate* memungkinkan jaringan untuk fokus pada data *input* yang lebih baru dan relevan, sambil menyaring informasi lama yang mungkin tidak lagi berguna [35]. Rumus untuk menghitung *Reset Gate* dapat dinyatakan sebagai berikut:

$$r_t = \sigma (W_r \cdot [h_{t-1}, x_t]) \quad (2.3)$$

r_t adalah *Reset Gate* dan W_r adalah bobot (*weight*) yang digunakan untuk perhitungannya, σ adalah fungsi aktivasi Sigmoid, h_{t-1} adalah *hidden state* dari langkah waktu sebelumnya, dan x_t adalah *input* pada waktu t.

Candidate State, seperti yang ditunjukkan pada persamaan (2.4), bertanggung jawab untuk menghitung *hidden state* pada langkah waktu saat ini. *Candidate state* dihasilkan menggunakan fungsi tangen hiperbolik (*tanh*) yang diterapkan pada kombinasi *input* saat ini dan *hidden state* sebelumnya yang telah dimodifikasi oleh *reset gate*. Penyesuaian ini memastikan bahwa hanya informasi masa lalu yang relevan yang memengaruhi kondisi saat ini. *Candidate state* merepresentasikan versi awal dari *hidden state* berikutnya sebelum diperbarui oleh

update gate [35]. Rumus untuk menghitung *candidate state* dapat dinyatakan sebagai berikut:

$$\check{h}_t = \tanh (W \cdot [r_t \cdot h_{t-1}, x_t]) \quad (2.4)$$

\check{h}_t adalah *Candidate State*, W adalah bobot (*weight*) untuk perhitungannya, dan \tanh adalah fungsi *aktivasi Tangen*.

Hidden State dalam GRU merepresentasikan *output* pada langkah waktu saat ini dan diteruskan ke unit GRU berikutnya. *Hidden state* dihitung dengan memadukan *candidate state* dan *hidden state* sebelumnya, dengan *update gate* yang mengontrol kontribusi keduanya. Keseimbangan ini memungkinkan *hidden state* untuk menyimpan informasi penting dari masa lalu dan masa kini, membuatnya lebih relevan untuk prediksi di masa depan. Dengan mengintegrasikan *update gate*, *reset gate*, dan *candidate state*, GRU dapat menangkap informasi jangka panjang secara lebih efektif daripada RNN tradisional [35]. Rumus untuk menghitung *hidden state* dapat dinyatakan sebagai berikut:

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \check{h}_t \quad (2.5)$$

h_t adalah *Hidden State* final yang nilainya merupakan hasil kombinasi dari *hidden state* sebelumnya (h_{t-1}) dan *candidate state* (\check{h}_t).

2.2.7. *Support Vector Regression (SVR)*

Support Vector Machine (SVM) adalah teknologi pembelajaran mesin yang populer, di mana perhitungannya bergantung pada fungsi *kernel*. *Support Vector Machine (SVM)* sendiri diterapkan pada dua masalah klasifikasi [32]. Dalam artikel ini, variasi penting dari *Support Vector Machine (SVM)* dalam regresi, yaitu *Nonlinear Support Vector Regression (SVR)*, digunakan. Dasar teori ini adalah menggunakan fungsi *kernel* untuk *mentransformasi* data ke tempat yang lebih tinggi, lalu menemukan fungsi terbaik $f(x)$ atau *hyperplane* sehingga *deviasi*

fungsi dari target tidak melebihi nilai ϵ (*epsilon*), dan pastikan fungsi tersebut sejajar mungkin. (2.6) [36].

$$f(x) = \sum_{n=1}^n (a_i^n - a_i) k(x_i, x) + b \quad (2.6)$$

Dalam persamaan ini, x_i adalah *vektor* dukungan dari nilai *input*, n adalah jumlah *vektor* dukungan, k adalah fungsi *kernel*, yang bisa berupa *Gaussian*, *linier*, atau *polinomial*, dan y_i adalah nilai target.

Dalam penelitian ini, *kernel Radial Basis Function* (RBF) dipilih karena fleksibilitas dan kemampuannya yang telah terbukti dalam menangani berbagai jenis data *non-linear* [31]. Fungsi *kernel* RBF didefinisikan sebagai.

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (2.7)$$

Dimana x_i dan x_j adalah dua sampel data, $\|x_i - x_j\|^2$ adalah kuadrat jarak *Euclidean* antara kedua sampel, dan γ *gamma* adalah parameter *kernel* yang mengatur lebar pengaruh dari fungsi *basis Gaussian*; pemilihan γ *gamma* yang tepat krusial untuk kinerja model [33].

Pada arsitektur model *ensemble* yang dikembangkan dalam penelitian ini, SVR bertindak sebagai model regresi utama yang menerima *input* berupa *vektor fitur* berdimensi 121. Vektor fitur ini merupakan hasil *ekstraksi* informasi *temporal* oleh lapisan *Gated Recurrent Unit* (GRU) sebelumnya. Kinerja SVR sangat dipengaruhi oleh pemilihan *hyperparameter*-nya. Oleh karena itu, dilakukan proses optimasi model menggunakan metode *grid search* untuk menentukan kombinasi parameter terbaik pada model *Support Vector Regression* (SVR) [33]. Parameter yang dioptimalkan meliputi parameter regularisasi C , yang berfungsi untuk mengontrol keseimbangan antara kompleksitas model dan toleransi terhadap kesalahan prediksi, nilai C yang diuji adalah 100 - 10000. Selain itu, parameter *kernel* γ untuk fungsi *kernel Radial Basis Function* (RBF) juga dioptimalkan, dengan nilai yang dieksplorasi adalah 0,0001 – 0,1.

2.2.8. Evaluasi Kinerja Model

Menilai kinerja model prediksi merupakan langkah penting untuk mengetahui sejauh mana model mampu memproyeksikan data yang belum diamati. Ukuran yang umum digunakan dalam evaluasi performa model prediktif adalah *Mean Absolute Percentage Error* (MAPE) [43]. MAPE merupakan metriks yang digunakan untuk menilai keakuratan metode prediksi dengan menghitung persentase kesalahan rata-rata antara nilai sebenarnya dan nilai prediksi. MAPE digunakan untuk mengukur seberapa besar perbedaan hasil prediksi metode dengan nilai sebenarnya dalam persentase [44]. Semakin kecil nilai MAPE maka semakin akurat model prediksinya. Berikut merupakan persamaan dari metriks MAPE.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \times 100 \quad (2.8)$$

Keterangan:

- A_t : nilai *actual* pada periode ke- t .
- F_t : nilai prediksi pada periode ke- t .
- n : banyak total periode.

Berikut merupakan tingkat signifikansi nilai MAPE yang digunakan untuk menilai tingkat akurasi model.

Tabel 2. 2 Tingkat Signifikansi Nilai MAPE

Nilai MAPE	Signifikasi
< 10%	Akurat
11% - 20%	Baik
21% - 50%	Cukup Baik
>51%	Tidak Akurat

2.2.9. *Framework Streamlit*



Gambar 2. 2 Logo *Streamlit*

Streamlit merupakan *framework* berbasis *Python* yang dirancang untuk membangun aplikasi *web interaktif* secara cepat dan efisien, terutama dalam konteks analisis data dan pembelajaran mesin. *Framework* ini memungkinkan pengembang untuk membuat antarmuka pengguna (GUI) dengan sedikit kode, karena setiap elemen antarmuka dapat dihasilkan langsung dari script *Python* tanpa memerlukan pengetahuan *HTML*, *CSS*, atau *JavaScript* [39]. Dalam konteks penelitian berbasis data *science*, *Streamlit* banyak digunakan untuk menyajikan hasil analisis secara *real-time* dan interaktif, sehingga mempermudah proses *eksplorasi* data dan komunikasi hasil prediksi kepada pengguna *non-teknis*. Keunggulan lainnya termasuk waktu pemuatan yang cepat, kemampuan untuk integrasi dengan model *machine learning* yang kompleks, serta kemudahan *deployment* ke cloud. Sebuah studi menunjukkan bahwa penggunaan *Streamlit* dalam pengembangan prototipe sistem prediksi berbasis pembelajaran mesin dapat meningkatkan efisiensi proses pengujian dan validasi model di lingkungan industri [39].

2.2.10. *Graphical User Interface (GUI)*

Antarmuka Pengguna Grafis (*Graphical User Interface/GUI*) merupakan representasi visual yang berfungsi untuk memfasilitasi interaksi antara pengguna dan sistem melalui elemen-elemen grafis, sehingga pengguna tidak perlu berinteraksi langsung dengan perintah berbasis teks. GUI dirancang dengan tujuan utama untuk meningkatkan pengalaman pengguna (*user experience*) melalui penyajian tampilan yang intuitif, konsisten, dan mudah dioperasikan. Sebuah GUI

yang efektif memungkinkan perangkat lunak dapat digunakan secara efisien dengan mengintegrasikan berbagai komponen interaktif seperti tombol, menu, kotak daftar, dan penggeser (*slider*), yang memudahkan navigasi serta pelaksanaan perintah tertentu [38].

Selain itu, GUI yang baik harus bersifat intuitif dan prediktif, artinya pengguna dapat memahami fungsi elemen-elemen visual tanpa kebingungan. Misalnya, saat pengguna menekan tombol dengan kursor, sistem secara otomatis menjalankan fungsi yang telah ditetapkan untuk tombol tersebut. Konsep ini dikenal sebagai *snap binding*, di mana label atau ikon membantu memperjelas fungsionalitasnya. Tujuan utama dari penerapan GUI adalah untuk menghubungkan kesenjangan antara bahasa perintah teks yang kompleks dan interaksi visual yang lebih sederhana, sehingga pengguna dapat mengoperasikan perangkat lunak tanpa membutuhkan keahlian teknis yang mendalam. Desain GUI yang konsisten, pelabelan yang informatif, serta respons yang cepat terhadap tindakan pengguna menjadi faktor penting dalam membangun kepercayaan dan kenyamanan pengguna saat berinteraksi dengan sistem [48].