

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara yang kaya akan sumber daya alam, baik hayati maupun non-hayati. Sumber daya alam hayati yang terkandung di dalamnya tersebar luas di setiap pulau yang ada di negeri ini. Dengan adanya kekayaan alam tersebut, Indonesia memiliki modal untuk pelaksanaan pembangunan ekonomi dan penyokong ketahanan pangan. Berdasarkan arah kebijakan BPP untuk anggaran tahun 2026 yang disebutkan oleh pemerintah bahwa ketahanan pangan menjadi prioritas utama dalam anggaran tersebut [1]. Maka ketahanan pangan tidak dapat dilepaskan dari sektor pertanian yang berperan sebagai penghasil tanaman pangan.

Dalam kurun waktu satu dekade (2013–2023), sektor jasa pertanian di Jawa Timur mengalami penurunan signifikan sebesar 67,58% berdasarkan data BPS Provinsi Jawa Timur (2023) [2]. Jumlah Rumah Tangga Usaha Pertanian Perorangan (RTUP) di subsektor palawija juga turun drastis dari 2.529.602 unit pada tahun 2013 menjadi 1.886.425 unit pada tahun 2023. Sebaliknya, subsektor palawija pada Perusahaan Pertanian Berbadan Hukum (UPB) justru mengalami peningkatan sebanyak 31 unit, sementara Usaha Pertanian Lainnya (UTL) naik dari 39 unit menjadi 110 unit pada periode yang sama. Penurunan RTUP ini berimplikasi pada menyusutnya luas panen jagung di Jawa Timur, yaitu sebesar 2,49 juta hektare pada tahun 2023, turun dari 2,76 juta hektare pada tahun 2022 atau sekitar 10,03% [3].

Kondisi penurunan luas panen tersebut telah berkontribusi terhadap meningkatnya harga jagung di tingkat konsumen di Jawa Timur sebagai provinsi penghasil jagung terbesar [3]. Tanpa adanya model prediksi harga yang andal, fluktuasi harga yang dipicu penurunan produksi semacam ini berpotensi mengganggu ketahanan pangan, menekan daya beli masyarakat, serta menimbulkan ketidakstabilan pasar jagung baik di level provinsi maupun nasional. Hal ini sejalan dengan temuan penelitian tahun 2024 yang menekankan bahwa volatilitas harga pangan di daerah sentra produksi memiliki implikasi langsung terhadap keberlanjutan sistem pangan nasional dan memerlukan intervensi berbasis prediksi harga yang lebih akurat [4].

Pemilihan Kabupaten Sampang sebagai studi kasus didasarkan pada perannya di Pulau Madura sebagai salah satu kawasan sentra jagung di Jawa Timur, dengan indikator luas panen dan struktur usaha tani yang menonjol sebagaimana tercatat dalam publikasi resmi BPS serta Sensus Pertanian 2023. Hal ini menunjukkan bahwa komoditas jagung di Sampang memiliki posisi strategis dalam sistem pangan lokal dan sangat erat kaitannya dengan dinamika jagung Madura secara keseluruhan [5]. Secara historis, berbagai kajian akademik juga menegaskan kontribusi Madura termasuk Sampang, terhadap pasokan jagung Jawa Timur, serta memperlihatkan kesamaan karakter agronomi dan pola budidaya antar kabupaten di Madura. Dengan demikian, analisis terhadap harga jagung di Sampang berpotensi memiliki tingkat keterterapan yang tinggi pada konteks regional [6]. Lebih jauh, studi berbasis data resmi turut mendokumentasikan pangsa produksi jagung Madura terhadap Jawa Timur serta tantangan produktivitas khas lahan kering. Kondisi tersebut memperkuat urgensi pemodelan prediksi harga di wilayah ini, karena fluktuasi harga jagung di Sampang tidak hanya memengaruhi petani dan konsumen lokal, tetapi juga berpotensi memberikan dampak langsung pada stabilitas rantai pasok jagung di Madura dan Jawa Timur secara keseluruhan [7].

Kementerian Pertanian melaporkan bahwa pada tahun 2023 terjadi penurunan hasil panen jagung sekitar $\pm 3,17\%$, yang secara langsung turut mendorong kenaikan harga jagung di Kabupaten Sampang [8]. Situasi ini menuntut adanya langkah antisipatif melalui penerapan model prediksi berbasis regresi guna meminimalkan volatilitas harga pada periode berikutnya. Dalam ranah pembelajaran mesin, *Support Vector Machine* (SVM) telah banyak digunakan dan terbukti handal dalam meramalkan harga komoditas maupun variabel agrikultur lain, khususnya pada data dengan pola non-linear dan jumlah sampel yang terbatas. Namun demikian, berbagai penelitian juga menunjukkan bahwa metode regresi yang ada, termasuk SVM, masih menghadapi kendala dalam hal akurasi prediksi dan efisiensi waktu komputasi. Oleh sebab itu, dibutuhkan pendekatan alternatif berupa *Twin Support Vector Regression* (TSVR), yang menawarkan proses komputasi lebih cepat sekaligus potensi peningkatan ketepatan prediksi.

SVM adalah salah satu metode pembelajaran mesin yang dapat digunakan untuk tugas klasifikasi maupun regresi. Pada ranah regresi, teknik ini dikenal

sebagai *Support Vector Regression* (SVR). SVR dimanfaatkan untuk menganalisis data numerik dengan memproyeksikannya ke ruang fitur berdimensi tinggi melalui fungsi kernel [9]. Pendekatan ini memungkinkan terbentuknya model yang mampu memprediksi nilai kontinu berdasarkan variabel *input* sehingga menghasilkan bilangan riil. Selain itu, SVR memiliki kemampuan dalam mengurangi risiko *overfitting*, sehingga model yang dibangun tetap andal saat digunakan pada data berskala besar maupun kompleks [10].

Penelitian tahun 2021 menyoroti penggunaan metode *Support Vector Regression* untuk memperkirakan harga *Crude Palm Oil* di Indonesia serta nilai tukar EUR/USD, dengan memanfaatkan dataset FOREX bagi kurs mata uang dan data CPO untuk harga komoditas kelapa sawit [11]. Hasil penelitian ini didapatkan nilai akurasi MAPE sebanyak 99,97% dengan menggunakan *hyperparameter* $c=1e3$, $\gamma = 0,1$ dan $\text{Degree} = 3$. Sedangkan dengan menggunakan evaluasi MSE didapatkan akurasi sebanyak 98%.

Penelitian lain dilakukan dengan menggunakan SVR sebagai regresi pada tahun 2023 untuk menganalisa hasil prediksi harga komoditas nasional [12]. Berdasarkan penelitian tersebut hasil evaluasi model didapatkan melalui MSE, MAD, MAPE, dan RMSE. Nilai MSE adalah 6,062, nilai MAD adalah 6,6831, nilai MAPE adalah 0,0147, dan nilai untuk RMSE adalah 7,746.

Penerapan SVR yang sebelumnya dinilai memiliki performa yang baik untuk regresi untuk menormalisasi *error* pada saat penalti *outlier* dan menyelesaikan masalah lokal dari optimasi non-linier, akan tetapi waktu komputasi yang digunakan masih terbilang cukup besar [13]. Metode yang ditawarkan untuk menyelesaikan permasalahan ini adalah *Twin Support Vector Regression* (TSVR). Teknik ini merupakan pengembangan dari *Support Vector Regression* yang membangun dua buah *hyperplane* sehingga model lebih presisi dengan waktu komputasi yang lebih singkat. Pertama kali diperkenalkan oleh Peng Xinjun pada tahun 2010, TSVR memanfaatkan fungsi $\mathcal{E}$ -insensitive *up-bound* dan *down-bound* sehingga perhitungan dapat dilakukan lebih efisien.

Meskipun SVR telah banyak digunakan dalam prediksi harga komoditas, literatur lain seperti studi Prediksi Harga Emas [14] menunjukkan bahwa *Random Forest Regression* (RFR) juga memiliki keunggulan, khususnya dalam

menghasilkan kesalahan prediksi yang lebih rendah dibandingkan SVR dalam kasus prediksi harga emas. Namun, kajian yang secara khusus membandingkan TSVR dengan metode regresi berbasis ensemble seperti Random Forest masih sangat terbatas, sehingga membuka ruang penelitian baru untuk mengevaluasi sejauh mana TSVR mampu memberikan performa prediksi yang lebih baik.

Penelitian yang dilakukan pada tahun 2022 terkait Model *Twin Support Vector Regression* pada *Heterodastic Gaussian Noise* dan aplikasinya dinilai bahwa penggunaan TSVR memiliki waktu komputasi yang lebih cepat, kompleksitas rendah dan hasil regresi yang lebih baik dibandingkan SVR konvensional [15]. Hasil evaluasi menggunakan 5 metode evaluasi MAE, RMAE, SSE, SSE/SST dan SSR/SST. Dari ketiga dataset yang digunakan di penelitian ini, hasil TSVR-HGN mendapatkan hasil evaluasi yang lebih baik daripada metode pembanding lainnya.

Penelitian terkait TSVR digunakan untuk regresi karakterisasi ketidakpastian dalam rekonstruksi permukaan pada tahun 2024 dengan metode pembanding yang digunakan adalah CCSVIR [16]. Hasil yang didapatkan adalah UCTSVR lebih baik daripada metode pembanding lainnya dengan menggunakan evaluasi model $\sum_RMSD$ dan MRE. Selain itu TSVR juga dapat diterapkan pada estimasi respons frekuensi kanal dan memiliki hasil SSE yang baik dibandingkan metode pembandingnya [17].

Pengembangan model TSVR lainnya adalah SABO-ILSTSVR yang mengintegrasikan *Least Squares Twin Support Vector Regression* (LSTSVR) dengan regularisasi Lasso dan optimisasi berbasis *Subtraction Average Based Optimizer* (SABO) [18]. Model ini dirancang untuk mengatasi tantangan *overfitting* yang umum terjadi pada data genotipe dengan dimensi tinggi dan jumlah sampel terbatas. Dengan mengaplikasikan SABO-ILSTSVR pada empat dataset tanaman berbeda, hasilnya menunjukkan bahwa model ini memiliki kinerja prediksi yang setara atau lebih baik dibandingkan metode prediksi menggunakan *Support Vector Regression*. Hasil prediksi SABO-ILSTSVR meningkat 4% di *Pearson Correlation Coefficient* dan mengalami penurunan 2% dengan evaluasi model MSE.

Alur proses yang akan diterapkan pada penelitian ini berpedoman dari metode CRISP-DM untuk mendefinisikan proses *data mining*. Pemilihan CRISP-DM sebagai metode pada penelitian ini disebabkan oleh fleksibilitasnya terhadap

proses apabila menemui data atau evaluasi yang belum sesuai, sehingga dapat kembali ke proses sebelumnya. CRISP-DM sudah ada sejak tahun 1999 sebagai metode alur proses untuk pengolahan *data mining* [19]. Pada penelitian ini terdapat 5 proses evaluasi model yang akan diterapkan untuk menilai performa model. Proses evaluasi tersebut terdiri dari MAPE, SSE, MSE, *R-Squared* (R^2), dan *Error Variance*.

Evaluasi MAPE (*Mean Percentage Error*) digunakan untuk mengukur performansi model. MAPE mampu mengukur kesalahan relatif berdasarkan nilai absolut [20]. Keuntungan menggunakan MAPE adalah mampu menjaga kebenaran nilai dari nilai negative dan tidak bergantung pada skala variabel independen. Evaluasi SSE (*Sum Square Error*) adalah metode yang digunakan untuk mengukur seberapa baik model memprediksi nilai sebenarnya dengan menjumlahkan kuadrat selisih antara nilai aktual dan nilai prediksi [21]. Evaluasi MSE (*Mean Square Error*) digunakan untuk pengukuran kinerja sistem dengan menghitung hasil prediksi terhadap dataset asli [22]. Evaluasi *R-Squared* (R^2) digunakan untuk mengukur nilai variasi di dalam variabel independen pada model *forecasting* [23]. Dan evaluasi *Error Variance* digunakan untuk mengetahui interval dari nilai prediksi yang dihasilkan [24].

Pada penelitian ini diterapkan metode TSVR untuk melakukan regresi pada data pertanian dengan tujuan meningkatkan akurasi prediksi harga konsumen menjadi lebih baik. Tahapan untuk mengetahui efektifitas metode ini dilakukan dengan evaluasi melalui tiga tahapan berbeda dan membandingkan hasilnya pada setiap evaluasi tersebut. Penggunaan metode TSVR pada dataset realtime masih sangat sedikit digunakan dan fokus penerapan metode ini lebih ke arah pengembangan metode TSVR dengan menggabungkan metode lain untuk memperoleh akurasi dan efisiensi waktu yang lebih baik. Sehingga pada penelitian ini, metode tersebut digunakan pada dataset realtime yang diperoleh dari Dinas Pertanian dan Ketahanan Pangan Provinsi Jawa Timur pada tahun 2022 hingga 2024. Hasil akhir dari penelitian ini dibuat menjadi dashboard yang dapat dipahami dan memiliki fitur regresi di dalamnya.

Berdasarkan pemaparan terkait topik dan permasalahan di atas, peneliti bertujuan melakukan penelitian dengan judul “*Prediksi Harga Konsumen Jagung Menggunakan Twin Support Vector Regression*”.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Permasalahan dalam penelitian ini berawal dari penurunan jumlah Rumah Tangga Usaha Pertanian (RTUP) sektor palawija di Jawa Timur periode 2013–2023 yang berdampak pada berkurangnya luas panen jagung dan menyebabkan fluktuasi harga konsumen di Kabupaten Sampang. Kondisi tersebut menuntut model prediksi harga yang akurat dan efisien. Metode SVR yang umum digunakan masih memiliki keterbatasan pada efisiensi dan sensitivitas parameter, sedangkan TSVR menawarkan potensi peningkatan akurasi dan kecepatan komputasi namun belum banyak diterapkan pada prediksi harga jagung. Selain itu, belum tersedia dashboard perbandingan hasil penelitian yang menampilkan performa berbagai metode regresi secara visual untuk mendukung analisis akademik.

1.2.2 Metode Usulan

Solusi yang ditawarkan dalam penelitian ini adalah penggunaan metode TSVR untuk meningkatkan akurasi dan efektivitas prediksi harga jagung. Dibandingkan dengan SVR konvensional, metode TSVR telah terbukti memberikan hasil evaluasi model yang lebih baik serta efisiensi komputasi yang lebih tinggi.

1.2.3 Pertanyaan Penelitian

1. Berapa error dari penggunaan metode TSVR jika dibandingkan dengan SVR dan *Random Forest Regression* (RFR)?
2. Bagaimana dashboard perbandingan hasil penelitian dapat menyajikan visualisasi performa prediksi antar metode secara jelas dan informatif untuk mendukung analisis akademik?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan Penelitian

1. Untuk menganalisa dan membandingkan kesalahan (error) prediksi antara metode TSVR, SVR dan *Random Forest Regression* pada dataset pertanian harga konsumen.

2. Untuk mengembangkan dashboard perbandingan hasil penelitian sebagai media visualisasi untuk menunjukkan performa masing-masing metode secara jelas dan informatif.

1.3.2 Manfaat Penelitian

1. Memberikan kontribusi terhadap pengembangan metode regresi dalam *machine learning*, khususnya dengan menunjukkan keunggulan TSVR dibandingkan SVR dan *Random Forest Regression* pada prediksi harga komoditas pertanian.
2. Menyediakan *dashboard* perbandingan hasil penelitian yang dapat membantu peneliti dan akademisi dalam mengevaluasi performa berbagai metode regresi secara visual dan informatif.

1.4 Batasan Masalah

Berikut ini adalah batasan masalah dalam penelitian ini :

1. Informasi diperoleh dari portal SISKAPERBAPO (<https://siskaperbapo.jatimprov.go.id/>).
2. Dataset mencakup harga jagung di Kabupaten Sampang periode 2022–2024.
3. Data yang tersedia hanya dari dua pasar, yakni Pasar Rongtengah dan Pasar Srimangunan, kemudian dihitung nilai rata-ratanya.
4. Evaluasi model yang digunakan adalah MAPE, SSE, MSE, R^2 .
5. Evaluasi prediksi yang digunakan adalah dan MAPE, SSE, MSE, R^2 dan *Error Variance*.

1.5 Sistematika Penulisan Skripsi

BAB I Pendahuluan

Bab ini menampilkan uraian singkat mengenai isu utama penelitian, arah tujuan yang hendak diraih, manfaat yang diharapkan, serta batasan penelitian untuk memastikan kesesuaian dengan lingkup yang telah ditetapkan.

BAB II Tinjauan Pustaka

Bab II memuat kajian literatur sebagai landasan penelitian, sekaligus menampilkan studi terdahulu yang relevan untuk memberikan gambaran perkembangan riset di bidang serupa.

BAB III Metode Penelitian

Bab III memuat rancangan sistem yang dikembangkan, meliputi tahapan penelitian, data yang digunakan, serta proses penerapan metode regresi.

BAB IV Hasil dan Pembahasan

Pada bab IV berisi tentang hasil dari penerapan metode usulan yang digunakan dan pembahasan terkait hasilnya dengan perbandingan metode. Hasil yang digunakan adalah nilai evaluasi error yang dibandingkan antara metode TSVR, SVR, dan *Random Forest Regression*.

BAB V Penutup

Bab V memuat rangkuman hasil serta rekomendasi penelitian, meliputi metode yang diterapkan, data yang dianalisis, dan ukuran evaluasi yang digunakan.