

BAB I

PENDAHULUAN

Bab ini memberikan gambaran umum tentang penelitian, termasuk latar belakang, rumusan masalah, tujuan, manfaat, dan batasan masalah. Pendahuluan ini disusun dengan tujuan untuk membantu memberikan gambaran dalam pemahaman terkait konteks penelitian sebelum masuk ke pembahasan lebih lanjut terkait proses dan tahapan pengolahan data hingga menghasilkan hasil akhir penelitian.

1.1 Latar Belakang

Indonesia merupakan salah satu negara di dunia yang memiliki kekayaan alam yang melimpah. Kekayaan tersebut mencakup sumber daya air, sumber daya lahan, sumber daya hutan, sumber daya laut, serta keanekaragaman hayati [1]. Potensi tentu dapat dimanfaatkan oleh Indonesia sebagai modal penting dalam menunjang pelaksanaan pembangunan ekonomi nasional, salah satunya melalui sektor pertanian.

Sektor pertanian memegang peranan penting dalam perekonomian Indonesia. Selain berkontribusi terhadap Produk Domestik Bruto (PDB) nasional, sektor ini juga menjadi penyedia lapangan pekerjaan bagi sebagian besar masyarakat Indonesia. Berdasarkan hasil Sensus Pertanian (ST) yang dilakukan oleh Badan Pusat Statistik (BPS) pada tahun 2023, tercatat sekitar 27,36 juta *unit* usaha pertanian di Indonesia [2], yang menunjukkan besarnya peranan sektor pertanian dalam mendukung ketahanan pangan dan stabilitas ekonomi nasional. Salah satu subsektor paling penting dalam sektor pertanian adalah hortikultura, yang mencakup komoditas bernilai tinggi seperti bawang merah.

Bawang merah (*Allium cepa* L. var. *aggregatum*) merupakan salah satu komoditas hortikultura unggulan yang memiliki nilai ekonomi tinggi di Indonesia [3]. Tanaman ini termasuk dalam kelompok sayuran semusim yang dibudidayakan untuk digunakan sebagai bumbu dasar dalam berbagai jenis masakan. Bawang merah umumnya dibudidayakan di daerah dataran rendah hingga dataran tinggi dengan suhu berkisar antara 25-32°C dan tingkat penyinaran matahari minimal 75% dengan masa budidayanya yang relatif singkat, yaitu sekitar 50-60 hari. Kemudahan

dalam proses budidaya serta cepatnya masa panen, menjadikan bawang merah memiliki potensi pasar yang luas dengan tingkat permintaan yang cukup tinggi dan stabil.

Berdasarkan laporan Statistik Holtikultura 2023 yang dipublikasikan oleh BPS, tingkat partisipasi konsumsi rumah tangga terhadap bawang merah mencapai 96,82% [4]. Angka tersebut menunjukkan bahwa hampir seluruh rumah tangga di Indonesia mengonsumsi bawang merah. Selain itu, laporan Buletin Konsumsi Pangan Semester I Tahun 2024 menunjukkan bahwa rata-rata konsumsi masyarakat Indonesia terhadap bawang merah mencapai 2,94 kilogram per kapita per tahun [5]. Data tersebut memperkuat posisi bawang merah sebagai salah satu komoditas yang memiliki tingkat permintaan yang cukup tinggi dan stabil sepanjang tahun.

Meskipun permintaan bawang merah relatif tinggi dan stabil, harga komoditas ini sering mengalami fluktuasi signifikan [6], di mana kondisi tersebut dapat mengganggu stabilitas ekonomi nasional [7]. Keberadaan gangguan stabilitas ekonomi nasional ini tentu akan berdampak pada efisiensi rantai pasok akibat adanya kerugian yang dialami oleh petani, pelaku usaha, konsumen, serta menjadi tantangan bagi pemerintah dalam menjaga keseimbangan antara tingkat pasok dan permintaan komoditas bawang merah [8].

Provinsi Jawa Timur merupakan salah satu daerah penghasil bawang merah terbesar di Indonesia dengan total produksi mencapai 4.766.656,52 kwintal pada tahun 2024 [9]. Berdasarkan data Badan Pusat Statistik (BPS) Jawa Timur yang dipublikasikan pada Juni 2025, Kota Probolinggo tercatat sebagai salah satu sentra utama produksi bawang merah di provinsi tersebut dengan jumlah produksi mencapai 53.386 kwintal pada tahun 2024. Capaian tersebut melampaui kontribusi dari beberapa daerah perkotaan lain, seperti Kota Batu, Kota Madiun, Kota Blitar, Kota Surabaya, dan Kota Kediri [10]. Selain itu, varietas bawang merah asal Probolinggo dikenal memiliki cita rasa yang kuat serta aroma khas sehingga memiliki daya saing tinggi di pasar [11].

Meskipun memiliki kontribusi yang cukup signifikan terhadap produksi bawang merah di tingkat provinsi, Kota Probolinggo menghadapi tantangan berupa tingkat inflasi daerah yang relatif tinggi. Pada Maret 2024, tingkat inflasi Kota Probolinggo tercatat sebesar 0,77 persen, menempatkan kota ini pada urutan

ketujuh dari sebelas kota yang mengalami inflasi di Jawa Timur [12]. Kondisi tersebut menunjukkan adanya tekanan harga yang cukup signifikan pada kelompok pangan, termasuk bawang merah.

Situasi tersebut menjadikan Kota Probolinggo sebagai wilayah yang penting untuk dikaji lebih mendalam, khususnya terkait dinamika fluktuasi harga bawang merah sebagai salah satu komoditas hortikultura strategis. Pemahaman yang lebih baik mengenai pola perubahan harga diperlukan untuk mendukung upaya pemerintah dalam menjaga stabilisasi harga, mengendalikan inflasi, serta mengambil keputusan yang tepat oleh pemerintah daerah, pelaku usaha, maupun masyarakat.

Fluktuasi harga bawang merah yang cukup tinggi juga menunjukkan kebutuhan akan adanya sebuah sistem prediksi yang akurat dan cukup adaptif mengikuti kondisi pasar yang dinamis. Namun, hingga saat ini penelitian yang secara spesifik membahas prediksi harga bawang merah di Kota Probolinggo masih sangat terbatas. Hal tersebut membuka peluang penelitian baru, terutama dengan memanfaatkan metode yang lebih modern dan mampu menangani karakter data yang kompleks. Salah satu pendekatan yang berkembang pesat adalah penggunaan teknologi kecerdasan buatan, khususnya *machine learning*, yang terbukti efektif dalam memodelkan data *non-linear* seperti harga komoditas pertanian.

Penelitian yang dilakukan oleh Rahmat Hidayat et al., pada tahun 2024 dengan judul “*Rice Price Prediction with Long Short-Term Memory (LSTM) Neural Network*” menunjukkan bahwa arsitektur *Long Short-Term Memory (LSTM)*, terbukti efektif dalam memprediksi harga beras. Hasil penelitian menunjukkan tingkat akurasi yang tinggi, ditunjukkan dengan nilai *Root Mean Square Error (RMSE)* sebesar 0,054, dengan mempertimbangkan variabilitas cuaca, hasil panen, serta luas lahan yang bersifat *time-series data* sebagai variabel *input* [13]. Namun demikian, model tersebut masih memiliki keterbatasan, terutama dalam menangkap pola musiman dan pengaruh variabel eksternal yang kompleks. Hal tersebut disebabkan karena LSTM lebih fokus pada pemrosesan data yang memiliki deret waktu atau historis yang kurang optimal dalam mengolah fitur kategorikal atau non-sekuensial.

Untuk mengatasi keterbatasan tersebut, salah satu pendekatan yang relevan adalah dengan menggabungkan LSTM dengan algoritma lain yang mampu menangani fitur non-sekuensial secara efisien. Salah satu algoritma yang memiliki keunggulan dalam pengolahan data kategorikal adalah *Category Boosting* (*CatBoost*). Penelitian yang dilakukan oleh Usman Syapotro et al. pada tahun 2024, dengan judul “*Prediction of Jakarta's Air Quality Using a Stacking Framework of CLSTM, CatBoost, SVR, and XGBoost*” menunjukkan bahwa *CatBoost* mampu memberikan hasil prediksi yang akurat pada data dengan karakteristik kompleks, yaitu kualitas udara di Jakarta. Pada penelitian tersebut, *CatBoost* menghasilkan nilai RMSE sebesar 13.4113, yang merupakan hasil terbaik diantara model individual lain yang digunakan dalam penelitian [14]. Tingginya tingkat akurasi yang didapat menunjukkan potensi algoritma *CatBoost* dalam mengolah data yang bersifat *non-linear* dan multivariabel.

Penelitian yang dilakukan pada tahun 2022 oleh Miaomiao Yang et al., yang berjudul “*Bearing Vibration Signal Fault Diagnosis Based on LSTM-Cascade CatBoost*” menunjukkan keberhasilan dalam pengembangan model diagnosis kerusakan pada bantalan mesin berbasis sinyal getaran menggunakan pendekatan yang menggabungkan model LSTM dan model *CatBoost* dalam struktur *cascade* yang berarti hasil dari satu tahap dapat mempengaruhi tahap berikutnya. Dengan adanya kombinasi ini, tahap *pre-processing* terhadap sinyal tidak diperlukan sehingga proses diagnosis menjadi lebih sederhana. Proses pengujian model dilakukan dengan menggunakan dua *dataset* besar, yang masing-masing menghasilkan nilai akurasi tinggi sebesar 99,72% dan 99,33%, lebih unggul dibandingkan penggunaan metode individual seperti LSTM, *gcForest*, maupun *CatBoost* saja [15]. Keberhasilan integrasi dalam penelitian ini menunjukkan bahwa kombinasi antara model *deep learning* dan *machine learning* berbasis pohon keputusan dalam struktur *cascade* mampu meningkatkan kemampuan klasifikasi secara signifikan, terutama pada data yang bersifat kompleks dan variatif.

Meskipun demikian, untuk menghasilkan performa model yang optimal, diperlukan proses *tuning hyperparameter* yang efektif. Salah satu pendekatan optimasi yang banyak digunakan saat ini adalah *Optuna*, sebuah *framework* berbasis *Bayesian Optimization* yang mengimplementasikan algoritma *Tree-*

structured Parzen Estimator (TPE). Penelitian yang dilakukan oleh Li Yang et al., yang berjudul “*On Hyperparameter Optimization of Machine Learning Algorithms: Theory and Practice*” menunjukkan bahwa pendekatan berbasis *Bayesian Optimization* (BO) seperti *Tree-structured Parzen Estimator* (TPE) yang diimplementasikan dalam *Optuna* dan *Hyperopt*, mampu menghasilkan akurasi prediksi yang tinggi dengan waktu komputasi yang jauh lebih efisien dibandingkan metode lain seperti *gridsearch*, *Genetic Algorithm* (GA), atau *Particle Swarm Optimization* (PSO) dengan hasil pada model *K-Nearest Neighbor* (KNN), pendekatan ini mencapai akurasi 96,83% hanya dalam waktu 2,33 detik [16].

Berdasarkan latar belakang tersebut, penelitian ini mengusulkan implementasi model LSTM-*CatBoost* yang dioptimasi menggunakan *Optuna* untuk memprediksi harga bawang merah di Kota Probolinggo. Kombinasi algoritma dua algoritma *machine learning* dan satu *framework* optimasi pendukung yang diterapkan ini diharapkan mampu mengatasi keterbatasan pada penelitian sebelumnya, dengan memanfaatkan kemampuan LSTM dalam menangkap pola temporal, *CatBoost* dalam mengelola data *non-linear* dan kategorikal, serta *Optuna* dalam menyusun konfigurasi model terbaik melalui *tuning hyperparameter* yang efisien sehingga model prediksi yang dihasilkan lebih akurat dan adaptif terhadap kondisi musiman maupun faktor eksternal lainnya yang memengaruhi harga bawang merah. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan model prediksi harga komoditas pertanian, khususnya bawang merah, dengan menyajikan informasi prediktif yang akurat mengenai fluktuasi harga. Model ini dapat digunakan sebagai alat bantu analisis untuk memantau tren harga ke depan, sehingga dapat memberikan gambaran awal bagi petani, pedagang, maupun instansi pemerintah daerah dalam mengambil keputusan yang lebih antisipatif.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, terdapat beberapa permasalahan yang perlu dikaji lebih lanjut dalam penelitian ini. Oleh karena itu, rumusan masalah dalam penelitian ini adalah sebagai berikut :

1. Bagaimana penerapan model LSTM-*CatBoost* dengan optimasi *Optuna* dalam memprediksi harga bawang merah di Kota Probolinggo?

2. Bagaimana pengaruh penerapan optimasi *Optuna* pada model LSTM-*CatBoost* terhadap peningkatan akurasi prediksi harga bawang merah di Kota Probolinggo?

1.3 Tujuan Penelitian

Penelitian ini dilakukan untuk mengatasi beberapa permasalahan yang ditemukan dalam proses prediksi harga bawang merah. Adapun tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut :

1. Menerapkan model LSTM-*CatBoost* dengan optimasi *Optuna* dalam proses prediksi harga bawang merah di Kota Probolinggo.
2. Menganalisis pengaruh penerapan optimasi *Optuna* pada model LSTM-*CatBoost* terhadap peningkatan akurasi hasil prediksi harga bawang merah di Kota Probolinggo.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi dalam berbagai aspek, baik secara akademis maupun praktis. Secara akademis, penelitian ini dapat menjadi referensi dalam pengembangan model LSTM-*CatBoost* dengan optimasi *Optuna* dalam memprediksi harga komoditas pertanian khususnya komoditas bawang merah, serta menambah literatur terkait penerapan *hyperparameter tuning* berbasis *Bayesian Optimization* dalam meningkatkan performa model prediksi. Secara praktis, penelitian ini ditujukan untuk menyediakan informasi prediktif mengenai dinamika harga bawang merah di masa mendatang yang dapat dimanfaatkan sebagai alat bantu analisis dalam pengambilan keputusan, khususnya oleh instansi pemerintah daerah dalam merumuskan strategi produksi, distribusi, serta kebijakan harga yang lebih tepat sasaran.

1.5 Batasan Masalah

Batasan ini ditetapkan untuk memastikan bahwa pembahasan tidak meluas ke luar cakupan yang telah ditentukan. Adapun batasan masalah yang telah ditetapkan dalam penelitian ini adalah sebagai berikut :

1. Penelitian ini berfokus pada penerapan model *LSTM-CatBoost* dengan optimasi *Optuna* dalam memprediksi harga bawang merah di Kota Probolinggo serta melakukan analisis terhadap nilai prediksi yang didapat.
2. Variabel yang digunakan dalam penelitian ini, meliputi waktu, harga (Rp/kg), jumlah produksi (kw), curah hujan, musim (hujan, panen / hujan, tanam / peralihan / kemarau, panen / kemarau, tanam), inflasi (%), serta IHK yang didapatkan melalui berbagai sumber, seperti DPKPP Kota Probolinggo, *website* SISKAPERBAPO, *website* BMKG, serta publikasi yang dikeluarkan oleh Badan Pusat Statistik Indonesia dengan periode data Januari 2020 hingga Desember 2024.
3. Evaluasi performa model dilakukan menggunakan metrik *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan koefisien determinasi (R^2), dengan membandingkan model LSTM, model *LSTM-CatBoost* serta model *LSTM-CatBoost* dengan optimasi *Optuna*.