

BAB V KESIMPULAN

A. Kesimpulan

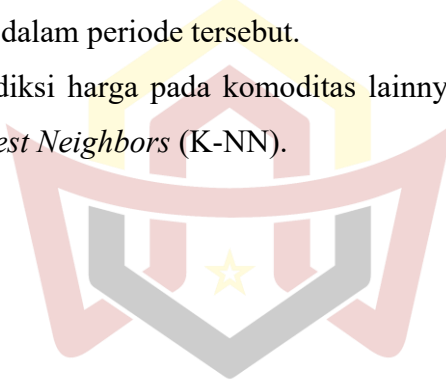
Berdasarkan hasil penelitian yang dilakukan, dapat disimpulkan bahwa penerapan algoritma *K-Nearest Neighbors* (K-NN) dalam memprediksi harga komoditas bawang merah di Kota Padang menunjukkan hasil yang cukup baik. Penggunaan pendekatan CRISP-DM sebagai metodologi dalam proses *data mining* membantu proses pemodelan dilakukan secara sistematis, mulai dari pemahaman bisnis, persiapan data, pemodelan, hingga evaluasi dan penyebaran hasil. Evaluasi performa model dengan metode *K-Fold Cross Validation* menghasilkan nilai *Koefisien Determinasi* (R^2) terbaik sebesar 0.8582, yang berarti bahwa model mampu menjelaskan 85,82% variasi harga bawang merah berdasarkan variabel-variabel yang digunakan.

Nilai *Mean Absolute Error* (MAE) terbaik sebesar 1363,76 menandakan bahwa rata-rata kesalahan prediksi harga berada pada kisaran Rp1.364, yang relatif kecil dan masih dalam batas toleransi. Sementara itu, nilai *Root Mean Squared Error* (RMSE) terbaik sebesar 1698.75 menunjukkan bahwa model memiliki tingkat kesalahan prediksi harga dengan rata-rata *kuadrat error* yang diakarkan sekitar Rp.1.698. Nilai ini juga termasuk kecil dan masih wajar sehingga dapat diterima dalam tujuan melakukan prediksi.

Proses normalisasi menggunakan metode *Min-Max Scaling* terbukti efektif dalam menyetarakan skala antar fitur sehingga meningkatkan akurasi prediksi. Selain itu, deteksi dan penanganan *outlier* menggunakan metode *Interquartile Range* (IQR) turut meningkatkan kualitas data dan hasil prediksi. Dengan demikian, model K-NN yang dibangun dapat dijadikan sebagai salah satu alternatif sistem prediksi harga yang dapat membantu pemerintah, petani, dan pelaku pasar dalam mengambil keputusan yang lebih tepat dalam mengantisipasi fluktuasi harga bawang merah di Kota Padang.

B. Saran

1. Penggunaan data eksternal tambahan, seperti inflasi, harga pupuk, dan variabel iklim lainnya, disarankan untuk meningkatkan akurasi prediksi harga bawang merah di masa depan.
2. Eksperimen dengan algoritma lain seperti *Random Forest*, *Gradient Boosting*, atau *XGBoost* dapat dilakukan untuk membandingkan performa prediksi dengan K-NN dan mungkin menghasilkan model yang lebih optimal.
3. Penggunaan data harga bawang merah dengan rentang minimal 1 tahun dapat menghasilkan prediksi yang lebih optimal, karena data tersebut mencakup pola harga yang dipengaruhi oleh hari-hari besar serta faktor historis lainnya dalam periode tersebut.
4. Melakukan prediksi harga pada komoditas lainnya dengan menggunakan metode *K-Nearest Neighbors* (K-NN).



UIN IMAM BONJOL
PADANG

DAFTAR PUSTAKA

- Aditya Pratama, M., Munawaroh, M., & Joko Pranoto, W. (2024). Perbandingan Performa Algoritma Linear Regresi dan Random Forest untuk Prediksi Harga Bawang Merah di Kota Samarinda. *Jurnal Ilmu Teknik*, 1(2), 172–182. <https://doi.org/10.62017/tektionik>
- Akbar, F. M. N. (2024). Metode K-NN (K-Nearest Neighbor) untuk Menentukan Kualitas Air. *Jurnal Teknokompak*, 18(1), 28–40.
- Allorerung, P. P., Erna, A., Bagussahrir, M., & Alam, S. (2024). Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit. *Jurnal Informatika Sunan Kalijaga*, 9(3), 178–191.
- Anugra, M., & Syukri, F. (2024). Pengaruh Harga Komoditas Pangan Terhadap Inflasi di Kota Pare Pare. *Cateris Paribus Journal*, 4(2).
- Apsari, R. M. (2024). Penerapan Metode Naïve bayes dalam Memprediksi Prestasi Siswa. *Jurnal Pustaka AI*, 4(2), 38–46. <https://doi.org/10.55382/jurnalpustakaai.v4i2.760>
- Astuti, L. T. W., Daryanto, A., Syaikat, Y., & Daryanto, H. K. (2019). Analisis Resiko Produksi Usahatani Bawang Merah pada Musim Kering dan Musim Hujan di Kabupaten Brebes. *Jurnal Ekonomi Pertanian dan Agribisnis*, 3(4), 840–852. <https://doi.org/10.21776/ub.jepa.2019.003.04.19>
- Balai Pengkajian Teknologi Pertanian Sumatera Barat. (2021). Teknologi Budidaya Bawang Merah Asal Biji. *Jurnal Sains Agro*, 6(1).
- Chaudhry, M., Shafi, I., Mahnoor, M., Vargas, D. L. R., Thompson, E. B., & Ashraf, I. (2023). A Systematic Literature Review on Identifying Patterns Using Unsupervised Clustering Algorithms: A Data Mining Perspective. Dalam *Symmetry* (Vol. 15, Nomor 9, hlm. 1–44). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/sym15091679>
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, 1–24. <https://doi.org/10.7717/PEERJ-CS.623>
- Marseva, A. D. (2023). Livelihood Strategy of Shallot Farmer Household in Dealing with Rainfall Variability. *Jurnal Ilmiah Inovasi*, 23(1), 17–21. <https://doi.org/10.25047/jii.v23i1.3818>
- Gorriz, M. C. S. R., Ortiz, & John Suckling. (2024). Is K-Fold Cross Validation The Best Model Selection Method For Machine Learning? *Arxiv*, 1–40. <https://doi.org/https://doi.org/10.48550/arXiv.2401.16407>

- Hayami, R., Soni, & Gunawan, I. (2022). Klasifikasi Jamur Menggunakan Algoritma Naïve Bayes. *Jurnal CoSciTech (Computer Science and Information Technology)*, 3(1), 28–33. <https://doi.org/10.37859/coscitech.v3i1.3685>
- Jamil, A., & Fatah, Z. (2024). Implementasi Data Mining Untuk Menganalisa Pola Penjualan Menggunakan Metode K-Means Pada Cv. Havas P2s2. *Riset Sistem Informasi, I*(4), 111–118. <https://doi.org/10.69714/rxnjxb68>
- Khoirani, N., Anees, R., Annisa, S., & Rusyda, H. (2022b). Peramalan Harga Bawang Merah di Pasar Tradisional Sulawesi Selatan dengan Metode ARIMA. *Jurnal Agribisnis, 24*(2), 274–287. <https://doi.org/10.31849/agr.v24i2.8670>
- Larissa, D., Fitri, F., & Fitria, D. (2024). Forecasting the Price of Shallots in Padang City Using the SARIMA Method. *Unp Journal Of Statistics And Data Science, 3*(1), 17–25. <https://doi.org/10.24036/ujsds/vol1-iss5/330>
- Larissa, D., Fitri, F., & Fitria, D. (2025). Forecasting the Price of Shallots in Padang City Using the SARIMA Method. *UNP Journal Of Statistics And Data Science, 3*(1), 17–25. <https://doi.org/10.24036/ujsds/vol1-iss5/330>
- Lumumba, V., Kiprotich, D., Mpaine, M., Makena, N., & Kavita, M. (2024). Comparative Analysis of Cross-Validation Techniques: LOOCV, K-folds Cross-Validation, and Repeated K-folds Cross-Validation in Machine Learning Models. *American Journal of Theoretical and Applied Statistics, 13*(5), 127–137. <https://doi.org/10.11648/j.ajtas.20241305.13>
- Maulana, A., & Ali, I. (2023). Prediksi Hasil Produksi Panen Bawang Merah Menggunakan Metode Regresi Linier Sederhana. *Jurnal Mahasiswa Teknik Informatika, 7*(4).
- Nasution, M. I., Syahni, R., & Yusmarni, D. (2020). Analisis Faktor-Faktor Yang Mempengaruhi Permintaan Bawang Merah Lokal Pada Konsumen Rumah Tangga Di Kota Padang. *Journal of Socio Economics on Tropical Agriculture, 2*(3), 258–267. <https://doi.org/10.25077/joseta>
- Nazar Yuniar, M. (2023). Klasifikasi Kualitas Air Bersih Menggunakan Metode Naïve bayes. *Jurnal Sains dan Teknologi, 5*(1), 243–246. <https://doi.org/10.55338/saintek.v5i1.1383>
- Nursyanti, R., Prakarsa, G., Rahmawati, S. N., & Andriana, A. (2025). Implementasi K-Nearest Neighbor untuk Prediksi Harga Bibit Kelapa Sawit dengan Metode CRISP-DM. *Jurnal Sistem Informasi dan Informatika, 7*, 124–134.

- Nurwahidah, Lokapitasari, P. L., & Herdianti. (2023). Implementasi Algoritma Neural Network untuk Memprediksi Harga Bawang Merah di Kabupaten Bima. *Buletin Sistem Informasi dan Teknologi Islam*, 4(2), 112–119.
- Octiva, C. S., Fajri, T. I., Sulistiarini, E. B., Suharjo, S., & Nuryanto, U. W. (2024). Penggunaan Teknik Data Mining untuk Analisis Perilaku Pengguna pada Media Sosial. *Jurnal Minfo Polgan*, 13(1), 1074–1078. <https://doi.org/10.33395/jmp.v13i1.13936>
- Omoseebi, A., Ola, G., & Tyler, J. (2024). Data Preparation and Feature Engineering. *ResearchGate*, 1–41.
- Pangestu, A., & Ridwan, D. T. (2022). Penerapan Data Mining Menggunakan Algoritma K Means Pengelompokan Pelanggan Berdasarkan Kubikasi Air Terjual Menggunakan Weka. *Just IT: Jurnal Sistem Informasi, Teknologi Informasi, dan Komputer*, 12(3), 67–71. <https://doi.org/https://doi.org/10.24853/justit.12.3.67-71>
- Putrama, I. M., & Martinek, P. (2024). Heterogeneous data integration: Challenges and opportunities. *Data in Brief*, 56, 1–23. <https://doi.org/10.1016/j.dib.2024.110853>
- Rahmadini, Enjel Erika LorencisLubis, Aji Priansyah, Yolanda R.W.N, & Tuti Meutia. (2023). Penerapan Data Mining Untuk Memprediksi Harga Bahan Pangan Di Indonesia Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Mahasiswa Akuntansi Samudra*, 4(4), 223–235. <https://doi.org/https://doi.org/10.33059/jmas.v4i4.7074>
- Rahmayuni, I., Sonatha, Y., Alanda, A., Erianda, A., Informasi, S., Komoditas, H., Untuk Pasar-Pasar, P., & Padang, K. (2020). Sistem Informasi Harga Komoditas Pangan untuk Pasar-Pasar di Kota Padang. *Jurnal Ilmiah Teknologi Sistem Informasi*, 1(1), 25–31. <https://doi.org/10.62527/jitsi.1.1.5>
- Rehman, A. ur, & Belhaouari, S. B. (2021). Unsupervised outlier detection in multidimensional data. *Journal of Big Data*, 8(1). <https://doi.org/10.1186/s40537-021-00469-z>
- Rianti, A., Nuur Wachid Abdul Majid, & Ahmad Fauzi. (2023). CRISP-DM: Metodologi Proyek Data Science. *Senatib*, 107–114. <https://www.researchgate.net/publication/377662458>
- Rofiq, N., & Salim, A. (2023). Prediksi Harga Bawang Merah menggunakan Algoritma Fuzzy Inference System (FIS). *Rekayasa Teknik Informatika dan Informasi*, 3(4), 154–162. <https://djournals.com/resolusi>

- Romli, I. (2021). Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Klasifikasi Penyakit Ispa. *Indonesian Journal of Business Intelligence (IJUBI)*, 4(1), 10–15. <https://doi.org/10.21927/ijubi.v4i1.1727>
- Sabda, M. A., & Suhardi, S. (2023). Implementasi Data Mining Dalam Memprediksi Penjualan Parfum Terlaris Menggunakan Metode K-Nearest Neighbor. *Jurnal Sistem Komputer dan Informatika (JSON)*, 5(2), 415. <https://doi.org/10.30865/json.v5i2.7194>
- Samruddhi, K., & Ashok Kumar, R. (2020). Used Car Price Prediction using K-Nearest Neighbor Based Model. *International Journal of Innovative Research in Applied Sciences and Engineering*, 4(3), 686–689. <https://doi.org/10.29027/ijirase.v4.i3.2020.686-689>
- Sastya, N. C., & Isna Nugraha. (2023). Penerapan Metode CRISP-DM dalam Menganalisis Data untuk Menentukan Customer Behavior di Meat Solution. *Jurnal Pendidikan dan Aplikasi Industri*, 10(2), 103–115. <https://doi.org/https://doi.org/10.33592/unistek.v10i2.3079>
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, 181, 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>
- Sharifnia, A. M., Kpormegbey, D. E., Thapa, D. K., & Cleary, M. (2025). A Primer of Data Cleaning in Quantitative Research: Handling Missing Values and Outliers. *Journal of Advanced Nursing*, 1–6. <https://doi.org/10.1111/jan.16908>
- Sholikin, A. R., & Haryono, D. (2019). Studi Perubahan Curah Hujan terhadap Produktivitas Tanaman Bawang Merah (*Allium ascalonicum* L.) di Beberapa Sentra Produksi. *Jurnal Produksi Tanaman*, 7(9), 1587–1594.
- Sunariadi, N. M., Keumala Intan, P., Candra, D., Novitasari, R., & Hariningsih, Y. (2022). Prediksi Produksi Bawang Merah Di Kabupaten Nganjuk Dengan Metode Seasonal Arima (Sarima). *Jurnal Pendidikan Matematika dan Matematika*, 6(1), 49–60. <https://doi.org/10.36526/tr.v%vi%i.1672>
- Wijaya, S. T., Hartami Santi, I., & Wulansari, Z. (2023). Penerapan Metode K-Nearest Neighbor Untuk Prediksi Harga Jagung Dengan Pengujian RMSE. *Jurnal Mahasiswa Teknik Informatika*, 7(2), 1255–1260.