

Prophet-LightGBM Hybrid Model Implementation in Cafe Menu Sales Prediction

Implementasi Model *Hybrid* Prophet-LightGBM dalam Prediksi Penjualan Menu Kafe

Erik Evranata Pardede, Fetty Tri Anggraeny, Achmad Junaidi

**Prodi Informatika, Fakultas Ilmu Komputer
Universitas Pembangunan Nasional Veteran Jawa Timur**

E-mail: 121081010114@student.upnjatim.ac.id

Abstract - This study aims to improve the accuracy of sales forecasting for cafe menu items through the development of a hybrid model that combines the Facebook Prophet and LightGBM algorithms. This hybrid model is designed to leverage the strengths of Prophet in detecting seasonal patterns and trends, as well as the ability of LightGBM to learn from residuals that are not captured by Prophet. The dataset used is sourced from Kaggle, containing cafe menu sales data, which includes information about the menu items, the quantity sold, and the transaction dates. Model evaluation was conducted using MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), and RMSE (Root Mean Squared Error) metrics. According to the results, the hybrid model shows significant improvement in forecasting accuracy, with MAPE of 5.83% for one menu item (cake), MAE of 0.84, and RMSE of 0.99, indicating better accuracy compared to the single models. This study is expected to provide valuable contributions to more efficient stock management and the development of more targeted marketing strategies for the cafe industry.

Keywords: Hybrid Model, Prophet, LightGBM, Sales Forecasting, Cafe Management, Time Series

Abstrak – Kajian ini bertujuan untuk meningkatkan akurasi peramalan penjualan menu kafe melalui pengembangan model hibrida yang menggabungkan algoritma Facebook Prophet dan LightGBM. Model hibrida ini dirancang untuk memanfaatkan keunggulan Prophet dalam mendeteksi pola musiman dan tren, serta kemampuan LightGBM dalam mempelajari residual yang tidak tertangkap oleh Prophet. Dataset yang digunakan berasal dari platform Kaggle, mencakup data penjualan menu kafe yang terdiri dari informasi tentang menu, jumlah produk yang terjual, serta tanggal transaksi. Evaluasi model dilakukan menggunakan metrik MAE (*Mean Absolute Error*), MAPE (*Mean Absolute Percentage Error*), dan RMSE (*Root Mean Squared Error*). Berdasarkan hasil kajian, model hibrida menunjukkan peningkatan signifikan dalam akurasi peramalan, dengan nilai MAPE salah satu menu yakni cake sebesar 5.83%, MAE sebesar 0,84, dan RMSE sebesar 0,99, yang mengindikasikan akurasi yang lebih baik dibandingkan dengan model tunggal. Kajian ini diharapkan dapat memberikan kontribusi penting dalam pengelolaan stok yang lebih efisien dan pengembangan strategi pemasaran yang lebih tepat sasaran bagi industri kafe.

Kata Kunci: Model Hibrida, Prophet, LightGBM, Peramalan Penjualan, Manajemen Kafe, Time Series

1. PENDAHULUAN

Industri kafe saat ini mengalami pertumbuhan yang pesat, yang disertai dengan semakin ketatnya persaingan di pasar [1]. Pemahaman terhadap tren penjualan sangat penting, sebab fluktuasi penjualan dapat berdampak pada performa bisnis. Penurunan penjualan menimbulkan hambatan dalam operasional dan pencapaian laba, sedangkan lonjakan penjualan yang tidak diprediksi dapat menyebabkan stok tidak mencukupi [2]. Kondisi ini seringkali menyebabkan masalah seperti kekurangan stok (*under stock*) atau kelebihan stok (*overstock*). Kekurangan stok dapat berujung pada ketidakpuasan pelanggan, karena produk yang diinginkan tidak

tersedia, sedangkan kelebihan stok dapat menyebabkan biaya penyimpanan yang tinggi dan risiko pemborosan produk. Oleh karena itu, prediksi penjualan yang akurat sangat penting untuk mendukung pengelolaan stok yang efisien dan pengambilan keputusan strategi pemasaran yang tepat.

Peramalan adalah teknik yang memanfaatkan data masa lalu untuk memperkirakan nilai yang mungkin terjadi di masa depan [3]. Tujuan dari prediksi ini adalah untuk mengidentifikasi pola dan tren yang dapat dijadikan acuan dalam mengambil keputusan strategis di masa mendatang [4]. Dalam hal ini, prediksi penjualan berbasis deret waktu (*time series*) sangat relevan, mengingat pola penjualan yang dipengaruhi oleh

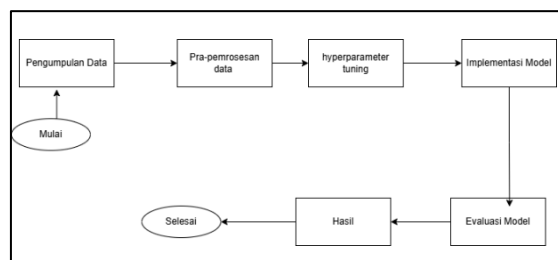
tren jangka panjang, musiman, dan faktor eksternal lainnya [5]. Menurut penelitian yang dilakukan oleh Winarno et al. (2022), penggunaan model berbasis deret waktu seperti Prophet efektif dalam memprediksi permintaan produk dengan akurasi yang baik, terbukti dengan nilai MAPE (*Mean Absolute Percentage Error*) yang rendah, yang menunjukkan kemampuan model dalam menangkap tren musiman dan pola permintaan [5].

Metode yang digunakan dalam kajian ini adalah kombinasi dari dua algoritma, yaitu Facebook Prophet dan LightGBM, yang keduanya terbukti efektif dalam memodelkan data deret waktu. Prophet sangat efektif dalam menangkap tren dan pola musiman, sementara LightGBM digunakan untuk menangani residual atau pola-pola yang tidak dapat dijelaskan oleh Prophet. Penelitian oleh Febrianto [6] juga menunjukkan bahwa LightGBM memiliki kemampuan untuk memproses data dalam jumlah besar dengan efisiensi tinggi dan memberikan hasil prediksi yang lebih baik dalam berbagai aplikasi, termasuk prediksi penjualan. Dengan pendekatan model hibrida yang menggabungkan keunggulan Prophet dalam mendeteksi pola musiman dan tren, serta kemampuan LightGBM dalam memperbaiki prediksi melalui pembelajaran residual, diharapkan dapat memberikan hasil yang lebih akurat [7]. Penerapan model hibrida ini dapat memberikan dampak signifikan dalam pengelolaan stok, mengurangi risiko kekurangan atau kelebihan stok.

Dengan menggabungkan kedua metode ini, kajian ini bertujuan untuk meningkatkan akurasi peramalan penjualan di industri kafe, agar bisa memberikan kontribusi dalam menciptakan sistem prediksi yang lebih handal, yang pada akhirnya akan membantu pengelola kafe dalam menjalankan operasi yang lebih efisien dan meningkatkan daya saing di pasar yang kompetitif.

2. METODE PELAKSANAAN

Kerangka kajian ditunjukkan pada Gambar 1. Pada **tahap pengambilan data**, dataset (Gambar 2) yang digunakan diperoleh dari platform Kaggle, yang menyediakan data penjualan menu cafe dengan informasi mengenai menu, jumlah produk yang terjual, dan tanggal transaksi. Data ini dimuat ke dalam lingkungan analisis menggunakan Python dengan bantuan library *pandas*. Kolom *Transaction Date* yang semula berbentuk *string*, kemudian diubah menjadi format *datetime* untuk memungkinkan analisis *time series* yang sesuai dengan kebutuhan model Prophet dan LightGBM.



Gambar 1. Kerangka Kajian

	Transaction ID	Item	Quantity	Price Per Unit	Total Spent	Payment Method	Location	Transaction Date
0	TXN_4292280	Cake	4	3.0	12.0	Cash	0	2023-01-01
1	TXN_2104473	Cake	3	3.0	9.0	Digital Wallet	0	2023-01-01
2	TXN_2104473	Cake	9	3.0	9.0	Cash	0	2023-01-01
3	TXN_5358805	Coffee	5	2.0	10.0	Digital Wallet	0	2023-01-01
4	TXN_3093284	Coffee	4	2.0	8.0	Digital Wallet	1	2023-01-01

Gambar 1. Dataset

Pra-pemrosesan data adalah tahap krusial sebelum pemodelan peramalan untuk memastikan data bersih, konsisten, dan sesuai format yang dibutuhkan. Proses ini mencakup konversi kolom *Tanggal* menjadi format *datetime* untuk analisis tren waktu pada model Prophet, agregasi data penjualan harian untuk stabilitas prediksi, serta penambahan fitur turunan seperti hari, bulan, dan hari dalam minggu untuk memperkaya informasi model dan membantu LightGBM menangkap pola residual. Selain itu, normalisasi fitur dilakukan untuk menyesuaikan skala data dalam rentang [0,1] dan data dibagi menjadi tiga bagian utama: data latih, validasi, dan uji, untuk pelatihan dan evaluasi model yang optimal.

Data cleaning adalah tahapan krusial untuk memastikan konsistensi dan kualitas data, karena kesalahan data dapat berdampak signifikan pada akurasi model *machine learning* dalam peramalan penjualan menu kafe [8]. Data mentah sering kali mengandung anomali, seperti nilai penjualan negatif, duplikasi transaksi, atau format tanggal yang tidak sesuai. Selain itu, nilai yang hilang (*missing values*) juga sering ditemukan. Untuk mengatasi masalah ini, dilakukan pemeriksaan dan penanganan terhadap nilai yang hilang dengan metode imputasi atau menghapus baris yang tidak relevan, pengecekan terhadap duplikasi data untuk menghindari bias, serta validasi konsistensi data seperti memastikan jumlah penjualan tidak negatif dan tanggal transaksi valid.

Tahap agregasi data dilakukan karena data transaksi biasanya masih berbentuk mentah (per item atau per transaksi). Agar sesuai dengan kebutuhan peramalan, data perlu diagregasi berdasarkan periode waktu tertentu, misalnya per hari, per minggu, atau per bulan. Sebagai

contoh, jika dalam satu hari terdapat ratusan transaksi, maka seluruh transaksi tersebut perlu dijumlahkan sehingga menghasilkan total penjualan harian. Proses agregasi ini memastikan pola tren dan musiman dapat lebih mudah dianalisis oleh model.

Pada **tahap transformasi data**, konversi format data, kolom *Transaction Date* yang awalnya berupa *string* diubah menjadi format *datetime* dan kolom *location* diubah rentangnya menjadi 0 dan 1 menggunakan teknik *min-max scaling*. Langkah ini penting karena model deret waktu seperti Prophet hanya dapat memproses data dengan struktur waktu yang valid. Dengan konversi ini, data transaksi dapat dipahami sebagai urutan kronologis, memungkinkan model untuk menganalisis tren, pola musiman, dan fluktuasi penjualan dengan lebih akurat. Jika tetap menggunakan format *string*, data tidak akan dapat diproses lebih lanjut untuk analisis *time series*.

Setelah proses pengolahan fitur dan target selesai, tahap selanjutnya adalah membagi dataset menjadi beberapa bagian, yaitu data latih (*training set*), data validasi (*validation set*), dan data uji (*test set*). Pembagian ini bertujuan agar model dapat dilatih menggunakan sebagian besar data, lalu diuji dan divalidasi pada data yang belum pernah digunakan dalam proses pelatihan, sehingga performa model dapat dievaluasi secara lebih objektif. Dalam kajian ini, dataset dibagi dengan rasio 80:10:10, di mana 80% digunakan untuk pelatihan, 10% untuk validasi saat tuning model, dan 10% untuk data uji akhir. Data latih berfungsi untuk mengajarkan model pola hubungan antara fitur dan target, sementara data validasi digunakan untuk *tuning* parameter dan menghindari *overfitting*. Data uji digunakan untuk menilai performa model pada data yang belum pernah dilihat sebelumnya, memastikan kemampuan model dalam melakukan generalisasi pada data baru. Dengan pembagian data yang tepat, model diharapkan memiliki akurasi tinggi tidak hanya pada data latih, tetapi juga pada data nyata di luar sampel pelatihan.

Proses hyperparameter tuning bertujuan untuk memperoleh konfigurasi parameter terbaik pada algoritma LightGBM (LGBM Regressor), yang mempengaruhi langsung kinerja prediksi. Parameter seperti *learning rate*, *max depth*, *num leaves*, dan *n_estimators* sangat berpengaruh terhadap kemampuan model dalam mempelajari pola data. Tahap pertama adalah mendefinisikan ruang pencarian parameter (*parameter grid*) dengan variasi nilai untuk parameter-parameter tersebut, diikuti dengan inisialisasi LightGBM menggunakan *random*

state=42 untuk memastikan konsistensi. *GridSearchCV* kemudian digunakan untuk mengevaluasi semua kombinasi parameter dengan *cross-validation* 3 lipatan dan metrik *neg_mean_absolute_error*. Parameter *verbose=1* ditambahkan untuk memberikan informasi rinci selama proses *tuning*, yang bertujuan untuk menemukan kombinasi parameter optimal guna meningkatkan akurasi model. Tujuannya adalah untuk menemukan kombinasi nilai *hyperparameter* yang menghasilkan nilai *error* terkecil berdasarkan metrik evaluasi MAE, RMSE, dan MAPE. Adapun parameter-parameter yang dioptimalkan pada model LightGBM dapat dilihat pada Tabel 1.

Tabel 1. *Tuning Gridsearch*

Nama	Rentang
<i>learning_rate</i>	[0.01, 0.05, 0.1]
<i>max_depth</i>	[3, 5, 7, 9]
<i>num_leaves</i>	[20, 31, 50, 70, 100]
<i>n_estimators</i>	[200, 400, 600, 800]

Learning_rate mengatur besar kontri-busi setiap pohon keputusan terhadap prediksi akhir. Nilai yang lebih kecil akan membuat model lebih konservatif dalam pembelajaran, namun bisa meningkatkan akurasi jika dilatih lebih banyak iterasi. *Max_depth* mengatur kedalaman maksimum pohon keputusan dalam model. Kedalaman yang lebih tinggi memungkinkan model untuk mempelajari interaksi yang lebih kompleks antara fitur. *Num_leaves* menentukan jumlah maksimum daun dalam setiap pohon keputusan. Semakin banyak daun yang digunakan, semakin banyak kompleksitas yang dapat dipelajari oleh model. *N_estimators* merujuk pada jumlah pohon keputusan yang digunakan dalam model. Semakin banyak pohon yang digunakan, semakin baik model dalam menangkap pola yang lebih kompleks, tetapi juga berisiko menyebabkan *overfitting* jika terlalu banyak.

Dengan menguji berbagai kombinasi parameter tersebut pada data validasi, model dievaluasi untuk mengukur performanya. Hasil dari proses *tuning* akan menunjukkan pengaturan parameter yang optimal berhasil meningkatkan akurasi prediksi penjualan secara signifikan dibandingkan dengan pengaturan *default*. Pendekatan *tuning* ini terbukti penting dalam mengoptimalkan kinerja model yang digunakan.

Prophet merupakan model peramalan deret waktu yang dikembangkan oleh Facebook untuk menangani data dengan tren jangka panjang dan pola yang tidak stasioner. Prophet mengandalkan model deret waktu yang terbagi menjadi dua komponen utama, yaitu tren (*trend*)

dan musiman (*seasonality*) [9]. Dalam kajian ini, Prophet digunakan untuk membentuk *baseline* prediksi dan mengekstrak komponen tren dari data penjualan harian. Kedua komponen ini digabungkan dalam persamaan:

$$y = g(t) + s(t) \quad 1)$$

Y merupakan nilai hasil prediksi pada waktu ke- t . Nilai ini dibentuk dari dua komponen utama, yaitu $g(t)$ dan $s(t)$. Komponen $g(t)$ adalah fungsi tren yang bertugas menangkap pola pertumbuhan umum dari waktu ke waktu, baik itu berupa kenaikan atau penurunan secara konsisten. Sementara itu, komponen $s(t)$ adalah fungsi musiman yang digunakan untuk mendeteksi pola-pola berulang dalam data.

LightGBM (*Light Gradient Boosting Machine*) adalah algoritma pembelajaran mesin berbasis pohon keputusan yang menggunakan teknik *Gradient Boosting* untuk membangun model prediksi yang akurat. LightGBM dirancang untuk menangani dataset besar dengan efisien dan cepat, serta memberikan hasil yang sangat baik dalam berbagai tugas pembelajaran terawasi. Model LightGBM diperoleh dengan meminimalkan fungsi *boosting loss* berdasarkan algoritma penurunan gradien. Setiap model baru ditambahkan, dan *loss function continues* untuk mendapatkan gradien variabel dengan informasi tertinggi [10]. Langkah awal dalam *boosting* adalah menghitung residual dari prediksi sebelumnya. Pada regresi dengan *squared error loss*, nilai gradien dan Hessian dihitung sebagai berikut:

$$g_i = y^{\wedge}(t) - y_i, \quad h_i = 1 \quad 2)$$

Pada regresi dengan *squared error loss*, gradien (g_i) dan Hessian (h_i) dihitung untuk mengoptimalkan model. Gradien mengukur perbedaan antara prediksi model ($y^{\wedge}(t)$) dan nilai aktual (y_i) pada titik data ke- i , menunjukkan besar perubahan fungsi *loss* terhadap perubahan parameter. Hessian (h_i), yang dalam kasus ini bernilai 1, mengukur kelengkungan fungsi *loss*, dan memberikan informasi tentang perubahan gradien. Dalam regresi dengan *squared error loss*, Hessian tetap konstan, yaitu 1, yang mempermudah perhitungan dalam optimasi. Gradien dan Hessian ini digunakan dalam algoritma optimasi untuk memperbarui parameter model dan meminimalkan kesalahan prediksi. Selanjutnya dilakukan *split* data menggunakan formula Gain dengan rumus:

$$Gain = \frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} + \frac{(G_L + G_R)^2}{(H_L + H_R) + \lambda} \quad 3)$$

Nilai G mempresentasikan total gradien dari fungsi *loss* terhadap hasil prediksi. H merupakan total Hessian atau turunan kedua dari fungsi *loss*. Nilai Hessian untuk setiap data adalah 1, sehingga total Hessian sama dengan jumlah data (n). λ merupakan parameter regularisasi yang digunakan untuk menghindari *overfitting* dengan menekan kompleksitas model melalui pengendalian terhadap nilai bobot pada setiap *leaf* pohon.

$$Output\ Leaf = \frac{G}{H + \lambda} \quad 4)$$

LightGBM membangun model berbasis pohon dengan memperbarui bobot model berdasarkan gradien kesalahan residual. Dalam kajian ini, LightGBM digunakan untuk memodelkan sisa *error* (residual) dari hasil model Prophet. Artinya:

$$residual = Y_{aktual} - (g(t) + s(t)) \quad 5)$$

Residual ini kemudian dijadikan target pelatihan LightGBM untuk menangkap pola-pola non-linier yang tidak dapat ditangkap Prophet, sehingga prediksi akhir:

$$y(t) = g(t) + r(t) \quad 6)$$

dimana $y(t)$ merupakan hasil prediksi pada waktu t yang ingin diprediksi oleh model. $g(t)$ merupakan komponen model atau prediksi yang dihasilkan oleh sistem pada waktu tertentu, yang bisa berupa tren atau pola yang dipelajari oleh model. $r(t)$ adalah komponen residual atau kesalahan, yaitu selisih antara nilai yang diprediksi oleh model dan nilai aktual yang diamati, yang menunjukkan kemampuan model dapat menangkap variabilitas data yang tidak dapat dijelaskan oleh komponen model.

3. HASIL DAN PEMBAHASAN

Dengan melakukan tahapan-tahapan pada metode kajian, dihasilkan matrik evaluasi (Tabel 3), yang meliputi MAE, RMSE, dan MAPE, sesuai untuk regresi dan memberikan gambaran kesalahan dalam satuan absolut maupun persentase. Hasil metrik evaluasi menunjukkan bahwa model *hybrid* prophet dan LightGBM menghasilkan performa terbaik dengan MAE sebesar 0.84, RMSE 0.99, dan MAPE 5.83%. Ini membuktikan bahwa integrasi tren dari Prophet dengan kemampuan LightGBM dalam menangani fitur kompleks memberikan prediksi yang jauh

lebih akurat. Setelah melakukan *Grid Search*, kombinasi parameter terbaik yang ditemukan adalah *learning_rate* = 0.1, *max_depth* = 3, *num_leaves* = 20, dan *n_estimators* = 400. Hasil ini menunjukkan bahwa kombinasi parameter tersebut memberikan kinerja terbaik dalam menghasilkan prediksi yang akurat.

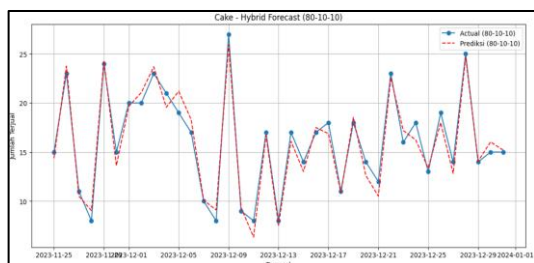
Tabel 3. Hasil Matriks Evaluasi

Model	MAE	RMSE	MAPE
Prophet + LightGBM	0.84	0.99	5.83%

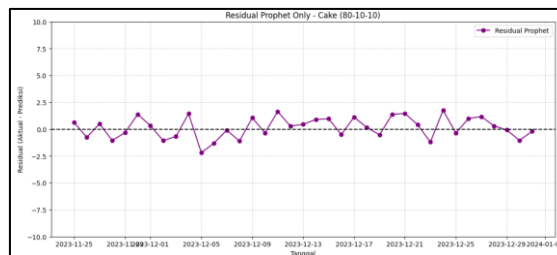
Hasil evaluasi (Gambar 3) menunjukkan bahwa model hibrida Prophet-LightGBM memberikan hasil prediksi yang lebih stabil dan akurat dibandingkan dengan model Prophet yang digunakan secara tunggal. Dari Gambar 3 terlihat bahwa pola prediksi yang dihasilkan oleh model *Hybrid* mengikuti tren umum dari data aktual dengan cukup baik. Sebagian besar lonjakan dan penurunan dalam grafik prediksi berhasil mereplikasi pergerakan yang terjadi pada data aktual, yang menandakan bahwa model mampu menangkap dinamika dan variasi dalam penjualan harian.

Grafik residual menunjukkan bahwa model hibrida mampu mengurangi kesalahan prediksi yang besar dan menghasilkan prediksi yang lebih baik pada periode-periode dengan lonjakan atau penurunan tajam. Sebagian besar titik residual terletak dekat dengan nol dapat dilihat pada Gambar 4, yang menandakan bahwa model hibrida berhasil memperbaiki kesalahan yang ada pada model Prophet.

Kajian ini berhasil mengembangkan model peramalan penjualan menu kafe berbasis hibrida Prophet-LightGBM yang terbukti lebih akurat dibandingkan penggunaan Prophet secara tunggal, khususnya pada menu cake dengan penurunan signifikan nilai MAE (0.84 vs 4.10), RMSE (0.99 vs 5.18), dan MAPE (5.83% vs 29.86%).



Gambar 3. Grafik Evaluasi *Hybrid*



Gambar 4. Grafik *Residual Hybrid*

Secara keseluruhan, hasil ini menegaskan bahwa model ini berhasil menggabungkan keunggulan komponen waktu dan fitur non-linear dengan efektif. Model tidak hanya unggul dalam segi akurasi numerik, tetapi juga memberikan stabilitas dan kemampuan adaptasi terhadap data baru. Hal ini sangat relevan untuk implementasi sistem prediksi penjualan dalam dunia nyata, terutama di sektor digital yang memiliki dinamika transaksi yang tinggi.

4. PENUTUP

Hasil kajian ini menunjukkan bahwa penggabungan Prophet dan LightGBM mampu mengatasi kelemahan Prophet dalam menangkap residual kompleks, sehingga menghasilkan prediksi yang lebih akurat, bahkan pada periode dengan fluktuasi tinggi. Pendekatan hibrida ini memberikan manfaat praktis bagi industri kafe, seperti efisiensi pengelolaan stok, perencanaan promosi yang lebih tepat, dan pengambilan keputusan bisnis yang lebih baik, serta berpotensi diterapkan di sektor lain dengan tantangan serupa.

Hasil kajian dapat ditingkatkan melalui *tuning hyperparameter* yang lebih optimal dengan *Random Search* atau *Bayesian Optimization* serta penambahan fitur eksternal, seperti cuaca, hari libur, atau kampanye promosi, untuk meningkatkan akurasi model secara lebih komprehensif.

DAFTAR PUSTAKA

- [1]. NA., "Banyak Usaha Kafe yang Bangkrut, Kenali Penyebabnya! – KABAR TODAY." Accessed: Aug. 25, 2025. [Online]. Available: https://kabartoday.co.id/banyak-usaha-kafe-yang-bangkrut-kenali-penyebabnya/?utm_source=chatgpt.com
- [2]. Ardiana Deka Maharani, A. Puspita Sari, C. Aji Putra, F., and J. Timur, P. Prediksi Penjualan pada Optimasi Stok dan Manajemen Persediaan Menggunakan Holt-Winters. "CICES (Cyberpreneurship Innovative and Creative Exact and Social Science)", doi: 10.33050/cices.v11i1.3499.

- [3]. Dimas Satria Prayoga, Puspita Sari Anggraini, and Achmad Junaidi, "Bitcoin Mining Hardware Profitability Prediction Using Categorical Boosting and Extreme Gradient Boosting Algorithms".
- [4]. L.L.D.I. Tasya Kurnia Sabila, "Sistem Prediksi Penjualan Di Toko Dasni Menggunakan Metode Double Exponential Smoothing," vol. 15, pp. 305–312, 2022, [Online]. Available: <http://journal.stekom.ac.id/index.php/pixel/page305>
- [5]. B. Heri Winarno, D. Kusumawati, and H. Agus Triyanto, "Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2024 Yogyakarta," 2024.
- [6]. E. Febriantoro, E. Setyati, and J. Santoso, "Pemodelan Prediksi Kuantitas Penjualan Mainan Menggunakan LightGBM," *SMARTICS Journal*, vol. 9, no. 1, pp. 7–13, Apr. 2023, doi: 10.21067/smartics.v9i1.8279.
- [7]. A. Shakeel, D. Chong, and J. Wang, "District heating load forecasting with a *hybrid* model based on LightGBM and FB-prophet," *J Clean Prod*, vol. 409, Jul. 2023, doi: 10.1016/j.jclepro.2023.137130.
- [8]. G. Y. Lee, L. Alzamil, B. Doskenov, and A. Termehchy, "A Survey on Data Cleaning Methods for Improved Machine Learning Model Performance," Sep. 2021, [Online]. Available: <http://arxiv.org/abs/2109.07127>
- [9]. D. Riau, "Prediksi Indeks Harga Saham Gabungan (IHSG) Menggunakan Prophet," 2022.
- [10]. K. Nisa, "Klasifikasi Penyakit Gangguan Mental dengan Algoritma LightGBM," *Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK)*, vol. 9, no. 2, pp. 1086–1094, 2024, [Online]. Available: <https://tunasbangsa.ac.id/ejurnal/index.php/jurasik>