



SKRIPSI

**PREDIKSI JUMLAH PEMBELIAN DI KEDAI KOPI
AREA KOTA SURABAYA DENGAN METODE
CATBOOST BERBASIS GUI *REACTJS***

CALVIEN DANNY NARIYANA
NPM 21083010040

DOSEN PEMBIMBING
Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.
Trimono, S.Si., M.Si.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**



SKRIPSI

PREDIKSI JUMLAH PEMBELIAN DI KEDAI KOPI AREA KOTA SURABAYA DENGAN METODE CATBOOST BERBASIS GUI *REACTJS*

CALVIEN DANNY NARIYANA
NPM 21083010040

DOSEN PEMBIMBING
Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.
Trimono, S.Si., M.Si

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**



SKRIPSI

PREDIKSI JUMLAH PEMBELIAN DI KEDAI KOPI AREA KOTA SURABAYA DENGAN METODE CATBOOST BERBASIS GUI *REACTJS*

CALVIEN DANNY NARIYANA
NPM 21083010040

DOSEN PEMBIMBING
Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.
Trimono, S.Si., M.Si

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**

LEMBAR PENGESAHAN

PREDIKSI JUMLAH PEMBELIAN DI KEDAI KOPI AREA KOTA SURABAYA DENGAN METODE *CATBOOST* BERBASIS GUI *REACTJS*

Oleh:
CALVIEN DANNY NARIYANA
NPM. 21083010040

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Sidang Skripsi Program Studi
Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur
Pada tanggal 15 Mei 2025:

Menyetujui,

Dr. Ir. Mohammad Idhom., S.P., S.Kom., M.T.
NIP. 1983031020212 1 1006

(Pembimbing I)

Trimono, S.Si., M.Si.
NIP. 19950908 202203 1 003

(Pembimbing II)

Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT.,
IPU., Asean, Eng
NIP. 1980120520050 1 1002

(Ketua Penguji)

Shindi Shella May Wara, M.Stat.
NIP. 19960518 202406 2 003

(Penguji I)

Mengetahui,
Dekan Fakultas Ilmu Komputer

 Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

**PREDIKSI JUMLAH PEMBELIAN DI KEDAI KOPI AREA KOTA
SURABAYA DENGAN METODE *CATBOOST* BERBASIS GUI *REACTJS***

Oleh :
CALVIEN DANNY NARIYANA
NPM. 21083010040

Telah menyelesaikan ujian skripsi dan revisinya

Menyetujui,

**Koordinator Program Studi Sains Data
Fakultas Ilmu Komputer**

Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.
NIP. 19801205 200501 1 002

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini:

Nama	Calvien Danny Nariyana
NPM	21083010040
Program	Sarjana (S1)
Program Studi	Sains Data
Fakultas	Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila di kemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 9 Mei 2025

Yang Membuat Pernyataan,



CALVIEN DANNY NARIYANA
NPM. 21083010040

ABSTRAK

Nama Mahasiswa / NPM : Calvin Danny Nariyana / 21083010040
Judul Skripsi : Prediksi Jumlah Pembelian Di Kedai Kopi Area Kota Surabaya Dengan Metode *Catboost* Berbasis Gui *Reactjs*
Dosen Pembimbing : 1. Dr. Ir. Mohammad Idhom., S.P., S.Kom., M.T.
2. Trimono, S.Si., M.Si.

Konsumsi kopi di Indonesia menunjukkan tren peningkatan yang stabil dalam lima tahun terakhir. Pada tahun 2015, konsumsi kopi mencapai 265 ribu ton, dan pada tahun 2020 meningkat menjadi 294 ribu ton, dengan rata-rata kenaikan tahunan sebesar 2% dan pada proyeksi konsumsi dalam negeri pada tahun 2024 mencapai 368.000 ton. Penelitian ini bertujuan untuk membuat website prediksi jumlah pembelian kedai kopi Area kota Surabaya dengan metode *Catboost* berbasis GUI *Reactjs* dengan perbandingan metode *LightGBM*. Metode penelitian yang digunakan melibatkan enam tahapan utama yaitu pengumpulan data dari wawancara dan kuisisioner pemilik, karyawan dan pembeli kedai kopi *third new era* dan data laporan keuangan beberapa kedai kopi terkait, pembacaan dataset, eksplorasi data sederhana, preprocessing data, pembangunan model prediksi, serta evaluasi model menggunakan data pengujian dan hasil akhir berupa pembangunan website prediksi. Dataset jumlah pembelian kedai kopi diambil dari wawancara dan kuisisioner pemilik, karyawan dan pembeli kedai kopi *third new era* serta data laporan keuangan beberapa kedai kopi terkait dan kemudian melalui tahap preprocessing untuk membersihkan serta menyiapkan data untuk analisis lebih lanjut. Model utama yaitu *Catboost* dan *LightGBM* sebagai perbandingan dibangun menggunakan data yang telah diproses dan hasil dari evaluasi model menunjukkan *Catboost* sebagai model terbaik untuk memprediksi pembelian di kedai kopi Surabaya, dengan MAE 0,91, RMSE 0.91, MSE 2.31 dan MAPE 15%, melampaui *LightGBM* (MAE 1,13, RMSE 1,1, MSE 2.9, dan MAPE 18%). *Catboost* dioptimalkan menggunakan Ordered Target Statistic Encoding dan Leave One Out Cross Validation. Penelitian ini juga menyertakan antarmuka ramah pengguna, memudahkan kedai kopi dalam mengelola stok, mengidentifikasi pasar sasaran, berinovasi sesuai preferensi pelanggan, dan menyusun strategi bisnis yang akurat menggunakan *Reactjs* sebagai tampilan website.

Kata kunci : *Catboost*, Kedai Kopi, *LightGBM*, Prediksi, *Third New Era*

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Calvin Danny Nariyana / 21083010040
Thesis Title : *Predicting Coffee Shop Purchase Quantities In The Surabaya City Area Using The Catboost Method In A Reactjs Based Gui*
Advisor : 1. Dr. Ir. Mohammad Idhom., S.P., S.Kom., M.T.
2. Trimono, S.Si., M.Si.

Coffee consumption in Indonesia has shown a stable upward trend over the past five years. In 2015, coffee consumption reached 265 thousand tons, increasing to 294 thousand tons in 2020, with an average annual growth rate of 2%. Domestic consumption is projected to reach 368 thousand tons by 2024. This study aims to develop a website for predicting coffee shop purchases in Surabaya City using the Catboost method, featuring a Reactjs-based GUI, compared with the LightGBM method. The research methodology involves six main stages: data collection through interviews and questionnaires with owners, employees, and customers of third new era coffee shops, as well as financial Reports from several related coffee shops; dataset reading; simple data exploration; data preprocessing; predictive model development; and model evaluation using test data, culminating in the development of the prediction website. The dataset on coffee shop purchases was obtained through interviews and questionnaires with owners, employees, and customers of third new era coffee shops, as well as financial Reports from several related coffee shops. The data underwent preprocessing to clean and prepare it for further analysis. The main models, Catboost and LightGBM for comparison, were developed using the processed data. Model evaluation results indicated that Catboost outperformed, achieving MAE 0.91, RMSE 0.91, MSE 2.31, and MAPE 15%, surpassing LightGBM (MAE 1.13, RMSE 1.1, MSE 2.9, and MAPE 18%). Catboost was optimized using Ordered Target Statistic Encoding and Leave One Out Cross Validation. This study also includes a user-friendly interface, simplifying stock management, target market identification, customer preference-based innovation, and business strategy formulation for coffee shops, utilizing Reactjs as the website interface.

Keywords: *Catboost, Coffee Shop, LightGBM, Prediction , Third New Era*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Prediksi Jumlah Pembelian di Kedai Kopi Area Kota Surabaya dengan Metode Catboost Berbasis gui Reactjs”** dapat terselesaikan dengan baik.

Penulis mengucapkan terima kasih kepada Bapak Dr. Ir. M. Idhom., S.P., S.Kom., M.T. selaku Dosen Pembimbing utama yang telah meluangkan waktunya untuk memberikan bimbingan, nasehat serta motivasi kepada penulis. Dan penulis juga banyak menerima bantuan dari berbagai pihak, baik itu berupa moril, spiritual maupun materiil. Untuk itu penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT., IPU, selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur
3. Bapak Dr. Eng. Ir. Dwi Arman Prasetya., ST., MT., IPU., Asean. Eng., selaku Koordinator Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Bapak Trimono, S.Si., M.Si. selaku dosen pembimbing skripsi kedua yang selalu memberikan bimbingan, arahan, dan motivasi kepada penulis.
5. Bapak dan Ibu dosen jurusan Sains Data Universitas Pembangunan Nasional “Veteran” Jawa Timur
6. Orang tua dan keluarga yang setia menemani dan mendukung kelancaran pengerjaan proposal ini.
7. Seluruh teman – teman Sains Data Angkatan 2021 yang selalu memberikan semangat dan dukungan.

Penulis menyadari bahwa di dalam penyusunan skripsi ini banyak terdapat kekurangan. Untuk itu kritik dan saran yang membangun dari semua pihak sangat diharapkan demi kesempurnaan penulisan skripsi ini. Akhirnya, dengan segala keterbatasan yang penyusun miliki semoga laporan ini dapat bermanfaat bagi semua pihak.

Halaman ini sengaja dikosongkan

DAFTAR ISI

ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xv
DAFTAR LAMPIRAN	xvii
DAFTAR NOTASI.....	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian	6
1.5 Manfaat Penelitian	7
BAB II TINJAUAN PUSTAKA.....	9
2.1 Penelitian Terdahulu	9
2.2 Dasar Teori.....	13
2.2.1 Kedai Kopi	14
2.2.2 Minuman Kopi	15
2.2.3 Machine learning	23
2.2.4 Algoritma <i>Categorical Boosting (Catboost)</i>	24
2.2.5 Algoritma <i>Light Gradient Boosting Machine (LightGBM)</i>	31
2.2.6 Model Evaluasi.....	33
2.2.7 Data Mining dan Analisis Regresi	36
2.2.8 Reactjs	37
2.2.9 Python	39
2.2.10 Visual Studio Code	40
BAB III DESAIN DAN IMPLEMENTASI SISTEM	9
3.1 Langkah Analisis.....	9

3.2	Desain Sistem.....	23
BAB IV HASIL DAN PEMBAHASAN.....		26
4.1	Informasi Dataset	26
4.2	Inisialisasi Data	27
4.3	Analisis Deskriptif Data	30
4.4	Pra Pemrosesan Data.....	32
4.5	<i>Modelling Data</i>	36
4.6	Evaluasi Model.....	43
4.7	Model Prediksi	51
4.8	<i>API Building</i>	54
4.9	<i>Website Building</i>	56
4.10	Hasil Tampilan Website	68
BAB V PENUTUP		82
5.1	Kesimpulan	82
5.2	Saran Pengembangan	83
DAFTAR PUSTAKA		84
LAMPIRAN.....		92

DAFTAR GAMBAR

Gambar 2. 1 Kedai Kopi	14
Gambar 2. 2 Kopi Hitam.....	16
Gambar 2. 3 Kopi Susu	18
Gambar 2. 4 Kopi <i>Espresso</i>	18
Gambar 2. 5 Kopi <i>Americano</i>	19
Gambar 2. 6 Kopi <i>Cappucino</i>	20
Gambar 2. 7 Kopi <i>Latte</i>	20
Gambar 2. 8 Kopi <i>Mocha</i>	22
Gambar 2. 9 Diagram Alir <i>Catboost</i>	26
Gambar 2. 10 Algoritma dalam <i>LightGBM</i>	32
Gambar 2. 11 Proses <i>Data mining</i>	36
Gambar 2. 12 Logo <i>Reactjs</i>	38
Gambar 2. 13 Logo <i>python</i>	39
Gambar 2. 14 Logo <i>Visual Studio Code</i>	40
Gambar 3. 1 Diagram Alir Penelitian	9
Gambar 3. 2 Contoh Visualisasi.....	17
Gambar 3. 3 <i>Usecase</i> Desain Sistem.....	24
Gambar 3. 4 Halaman <i>Dashboard Website</i>	25
Gambar 4. 1 Output Inisialisasi Data	28
Gambar 4. 2 Statistik Deskriptif Data	29
Gambar 4. 3 Output Visualisasi Data.....	31
Gambar 4. 4 Top 3 Jenis Kopi Yang banyak Dibeli	32
Gambar 4. 5 Jumlah Data Yang Hilang	33
Gambar 4. 6 <i>Output Data Preprocessing</i>	35
Gambar 4. 7 Ilustrasi Pohon <i>Catboost</i>	36
Gambar 4. 8 Ilustrasi Pohon <i>LightGBM</i>	36
Gambar 4. 9 Distribusi Beberapa Variabel Kategori Pada <i>Catboost</i>	38
Gambar 4. 10 <i>Leave-One-Out Cross Validation</i>	42
Gambar 4. 11 Output Proses Model.....	45
Gambar 4. 12 Gambar Perbandingan Data Aktual <i>Catboost</i>	48

Gambar 4. 13 Perbandingan Hasil Prediksi dan Aktual Model	49
Gambar 4. 14 <i>Feature Importance Model</i>	50
Gambar 4. 15 <i>API Building</i>	55
Gambar 4. 16 Contoh <i>Bug Error</i>	61
Gambar 4. 17 Homepage Layanan JagoHosting.....	62
Gambar 4. 18 Tampilan Cpanel	63
Gambar 4. 19 <i>Setting Domain</i>	64
Gambar 4. 20 Setting Setup <i>Python</i>	64
Gambar 4. 21 Detail Setting Setup <i>Python Application</i>	65
Gambar 4. 22 <i>Setup File Manager Server</i>	65
Gambar 4. 23 Setup File Manager Prediksi	66
Gambar 4. 24 <i>Setup Cron Jobs</i>	67
Gambar 4. 25 File Output.log <i>Nohup</i>	68
Gambar 4. 26 Halaman <i>Login</i>	69
Gambar 4. 27 Halaman <i>Register</i>	70
Gambar 4. 28 Halaman <i>Email Login</i>	70
Gambar 4. 29 Halaman <i>Landing</i>	71
Gambar 4. 30 Halaman Rekomendasi.....	72
Gambar 4. 31 Halaman Kalender.....	73
Gambar 4. 32 Halaman Input Prediksi	74
Gambar 4. 33 Halaman Hasil Prediksi.....	75
Gambar 4. 34 Halaman Visi Misi	76
Gambar 4. 35 Halaman <i>Reach Us</i>	77
Gambar 4. 36 Halaman <i>Reports</i> Jenis Kopi.....	77
Gambar 4. 37 Halaman <i>Reports</i> Waktu Dan Preferensi Konsumen	78
Gambar 4. 38 Halaman <i>Reports</i> Gender Dan Umur	78
Gambar 4. 39 Halaman <i>Settings</i>	79
Gambar 4. 40 Halaman <i>Profile</i>	80

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	9
Tabel 2. 2 Tingkatan Akurasi MAPE.....	35
Tabel 3. 1 Variabel penelitian	12
Tabel 3. 2 Pembagian kecamatan Kota Surabaya menurut wilayah	13
Tabel 4. 1 Informasi Dataset	26
Tabel 4. 2 Hasil Evaluasi Model	46
Tabel 4. 3 Contoh Hasil Prediksi	54

Halaman ini sengaja dikosongkan

DAFTAR LAMPIRAN

Lampiran 1. <i>Frontend Source Code</i>	92
Lampiran 2 Link Dokumentasi Survei Penelitian	111
Lampiran 3. <i>Cross Validation Evaluation Model</i>	112

Halaman ini sengaja dikosongkan

DAFTAR NOTASI

$countInClass$	:	Jumlah nilai target dalam suatu kelas kategori tertentu.
$Prior$	:	Rata-rata global target dari seluruh data
$avg\ target$	:	Hasil encoding fitur kategori
$totalCount$	:	jumlah data dalam kategori tersebut.
y_i	:	Bobot pada objek ke-i tetap diterapkan dalam <i>Catboost</i> , yang biasanya mencerminkan seberapa penting kontribusi dari setiap sampel dalam data
$\hat{y}_i$	:	prediksi yang dibuat oleh model untuk sampel ke-i
g_i	:	Gradien atau kesalahan prediksi pada objek ke-i. Dalam regresi <i>Catboost</i> , ini adalah turunan dari fungsi kerugian, misalnya mean squared error (MSE), yang mengukur selisih antara prediksi dan nilai aktual. Gradien ini menentukan seberapa besar model perlu dikoreksi.
r_{m-1}	:	residual dari model sebelumnya
h_m	:	model residual ke m.
$\hat{y}_i^{(t-1)}$	:	prediksi pada iterasi sebelumnya,
$f_t(x_i)$	:	fungsi keputusan (biasanya pohon keputusan) yang dipelajari pada iterasi ke-t.
α_t	:	faktor pembobot untuk iterasi ke-t
$y_i - F(X_i)$	:	Selisih nilai asli dengan prediksi
$F(X_i)$	:	nilai prediksi dari model untuk observasi ke-i
y_i	:	nilai actual
$\frac{1}{N}$	:	rata-rata dari selisih nilai asli
$\frac{y_i - \bar{y}}{\bar{y}_i}$	:	Persentasi kesalahan nilai asli dengan prediksi jangka waktu tertentu
t	:	penjumlahan dimulai pada periode waktu tertentu dan berakhir pada periode waktu tertentu.
$\sum =$	:	menunjukkan penjumlahan dari semua selisih absolut
MSE	:	<i>Mean Squared Error</i>

<i>MAE</i>	:	<i>Mean Absolute Error</i>
<i>RMSE</i>	:	<i>Root Mean Squared Error</i>
<i>MAPE</i>	:	<i>Mean Absolute Percentage Error</i>