

Klasifikasi Lonjakan Ekstrem Nilai Tukar USD/IDR Menggunakan *Gaussian Naïve Bayes*

Classification of Extreme Spikes in the USD/IDR Exchange Rate Using Gaussian Naïve Bayes

Meini Syakinah Ritonga*¹, Rita Novita Sari²

^{1,2}Program Studi Ilmu Komputer, Universitas Islam Negeri Sumatera Utara
E-mail: ¹meinisyakinah@gmail.com, ²ritanovitasari@uinsu.ac.id

Abstrak

Fluktuasi nilai tukar Rupiah terhadap Dolar AS (USD/IDR) berpotensi memicu ketidakstabilan makroekonomi. Penelitian ini bertujuan mengklasifikasikan potensi lonjakan ekstrem pada kurs USD/IDR menggunakan metode penambangan data dengan algoritma Gaussian Naïve Bayes. Sebanyak 502 data observasi historis harian diekstraksi menjadi empat fitur prediktor kontinu: volatilitas, selisih penutupan, selisih batas atas, dan selisih batas bawah. Pengujian dilakukan dengan rasio 80% data latih dan 20% data uji. Hasil evaluasi menunjukkan model mampu mengidentifikasi kelas "Normal" dan "Ekstrem" dengan akurasi 96,04%, presisi 62,50%, dan recall 83,33%. Validasi silang 5-Fold Cross Validation mencetak rata-rata akurasi kumulatif sebesar 95,43%, mengonfirmasi bahwa kinerja model stabil dan tidak mengalami overfitting. Kesimpulannya, algoritma Gaussian Naïve Bayes terbukti efektif dan dapat diandalkan sebagai sistem peringatan dini terhadap risiko lonjakan ekstrem nilai tukar valuta asing.

Kata kunci: Gaussian Naïve Bayes, Klasifikasi, Nilai Tukar, Penambangan Data, Volatilitas

Abstract

The fluctuation of the Rupiah exchange rate against the US Dollar (USD/IDR) potentially triggers macroeconomic instability. This study aims to classify potential extreme surges in the USD/IDR exchange rate using data mining techniques with the Gaussian Naïve Bayes algorithm. A total of 502 daily historical observation data were extracted into four continuous predictor features: volatility, closing difference, upper bound difference, and lower bound difference. The evaluation was conducted using an 80% training and 20% testing data split. The results show that the model can identify "Normal" and "Extreme" classes with an accuracy of 96.04%, a precision of 62.50%, and a recall of 83.33%. The 5-Fold Cross Validation test yielded an average cumulative accuracy of 95.43%, confirming that the model's performance is stable and does not experience overfitting. In conclusion, the Gaussian Naïve Bayes algorithm is proven to be effective and reliable as an early warning system against the risk of extreme foreign exchange rate surges.

Keywords: Classification, Data Mining, Exchange Rate, Gaussian Naïve Bayes, Volatility

1. PENDAHULUAN

Nilai tukar mata uang merupakan salah satu indikator vital dalam mengukur stabilitas perekonomian suatu negara. Bagi Indonesia, fluktuasi nilai tukar Rupiah terhadap Dolar Amerika Serikat (USD/IDR) memiliki dampak krusial terhadap aktivitas perdagangan internasional, utamanya pada ketahanan sektor ekspor dan impor. Pergerakan kurs setiap harinya sangat dipengaruhi oleh berbagai faktor fundamental makroekonomi, seperti perubahan kebijakan suku bunga global, laju inflasi, hingga sentimen geopolitik pasar. Ketidakpastian arah pergerakan pasar

valuta asing ini acap kali memicu depresiasi Rupiah secara tiba-tiba yang berisiko menyebabkan kerugian bagi pelaku bisnis, investor, hingga pemerintah dalam hal pembengkakan beban Anggaran Pendapatan dan Belanja Negara (APBN).

Adanya potensi lonjakan ekstrem akibat pelemahan nilai tukar Rupiah yang tajam menuntut disediakannya sebuah mekanisme klasifikasi tren yang akurat. Pendekatan pengamatan secara manual tidak lagi efektif untuk menganalisis pergerakan kurs di tengah arus data historis harian yang sangat masif. Oleh karena itu, dibutuhkan suatu sistem komputasi cerdas berbasis penambangan data (data mining) yang dapat memproses riwayat data secara otomatis dan mengelompokkan potensi pergerakan kurs tersebut ke dalam kategori fluktuasi "Normal" maupun "Lonjakan Ekstrem". Melalui kehadiran klasifikasi komputasional yang bersifat prediktif, para pengambil kebijakan maupun dunia usaha dapat melakukan langkah antisipatif serta meminimalisasi kerugian material secara cepat.

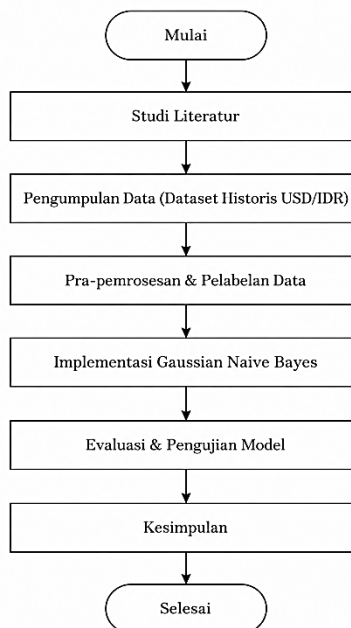
Untuk mengatasi tantangan klasifikasi *time-series finansial* tersebut, algoritma *Gaussian Naïve Bayes* diusulkan sebagai pendekatan penyelesaian yang diandalkan dalam penelitian ini. Alasan utama digunakannya *Gaussian Naïve Bayes* adalah kemampuannya yang sangat optimal dan teruji dalam memproses variabel atribut yang bersifat kontinu (berupa bilangan pecahan desimal), seperti persentase nilai perubahan harian (*return*), indikator volatilitas nilai tukar, serta laju pergerakan rata-rata (*moving average*). Berangkat dari hukum probabilitas Teorema Bayes, metode ini bekerja dengan mengasumsikan bahwa fitur-fitur masukan angka tersebut terdistribusi secara normal (*Gaussian*). Pendekatan ini menjadikannya sangat mangkus dalam melakukan proses latih pada barisan dataset besar serta secara konsisten mampu menghasilkan klasifikasi akhir dengan keakuratan tinggi tanpa menuntut spesifikasi perangkat keras yang berlebihan.

Penelitian mengenai prediksi dan pemodelan nilai tukar mata uang beserta implementasi Naive Bayes dalam klasifikasi objek finansial telah cukup masif dieksplorasi sebelumnya. Penelitian yang dilakukan oleh Eddy Nathansyah [1] melakukan komparasi performa algoritma regresi dan optimasi dalam memprediksi fluktuasi nilai tukar Rupiah terhadap Dolar. Pada bidang klasifikasi harga, riset dari Amin, et al. [2] berhasil mengimplementasikan probabilitas Naive Bayes untuk memprediksi harga emas guna meminimalisasi risiko investasi dengan keakuratan yang stabil. Penelitian lain yang dilakukan oleh Saputra, et al. [3] menunjukkan bahwa metode Naïve Bayes berhasil memproyeksikan klasifikasi pada kelas atribut harga teknologi secara mumpuni dengan perolehan tingkat akurasi mencapai 85,50%. Penelitian dari Ariansyah, et al. [4] juga memutakhirkan kapabilitas *Gaussian Naïve Bayes* melalui proses hibridisasi algoritma klasifikasi. Sama halnya dengan implementasi probabilitas pada instrumen masa depan dan analitik big data [5], klasifikasi berbantuan Naive Bayes dapat diandalkan untuk menyeleksi label secara efisien di mana akurasi rata-ratanya mampu menembus lebih dari 82%. Berpijak pada literatur terdahulu, penelitian ini difokuskan pada implementasi *Gaussian Naïve Bayes* dalam mengenali pola pergerakan probabilitas kontinu untuk mengklasifikasikan anomali fluktuasi nilai tukar (USD/IDR) dengan akurasi yang lebih spesifik.

2. METODOLOGI PENELITIAN

2.1. Alur Penelitian

Rangkaian tahapan dalam penelitian ini diadaptasi dari kerangka kerja penambangan data (*data mining*) untuk memastikan bahwa proses ekstraksi dan klasifikasi informasi dari sekumpulan deret waktu (*time-series*) berjalan secara sistematis [6]. Alur ini dimulai dari studi literatur dan observasi permasalahan makroekonomi terkait nilai tukar. Tahapan dilanjutkan pada proses pengumpulan data (Data Sekunder) berupa rekam jejak historis nilai tukar USD/IDR yang diakses secara resmi dari portal finansial global Investing.com (melalui tautan publikasi: <https://id.investing.com/currencies/usd-idr-historical-data>) dengan rentang waktu 01 Januari 2021 hingga 30 April 2026 yang menghasilkan total 1.354 baris data. Setelah data dikumpulkan, dilakukan pra-pemrosesan data (*preprocessing*). Pada tahap ini, variabel persentase perubahan harian (*Change %*) diformulasikan menjadi label kelas biner: "Ekstrem" (apabila persentase kenaikan bernilai positif dan melampaui batas *threshold* 0,5%), serta "Normal" (untuk fluktuasi di bawah 0,5%). Pemilihan *threshold* sebesar 0,5% ini didasarkan pada karakteristik pergerakan harian pasar valuta asing. Secara empiris, depresiasi harian kurs USD/IDR di atas 0,5% (setara dengan lonjakan tajam sekitar 75 hingga 80 poin dalam satu hari bursa) merupakan indikasi volatilitas tinggi yang anomali dan kerap menjadi parameter bagi otoritas moneter (Bank Indonesia) untuk melakukan intervensi stabilitas pasar. Selanjutnya, sistem dibangun menggunakan bahasa pemrograman Python di lingkungan kerja *Google Colab* untuk mengeksekusi probabilitas *Gaussian Naïve Bayes*. Python dipilih sebagai komputasi dasar dikarenakan kelengkapan fungsionalitas pustakanya (*libraries*) yang secara saintifik sangat mutakhir dalam mengolah analisis *Machine Learning* skala besar [7]. Pengujian akhir dilakukan untuk menakar keakuratan algoritma sebelum menarik kesimpulan penelitian.



Gambar 2. 1. Alur Penelitian

2.2. Skema Sistem

Sistem komputasi pengelompokan (klasifikasi) pada penelitian ini mengandalkan kalkulasi probabilitas bersyarat dengan menggunakan perluasan Gaussian Naïve Bayes. Algoritma ini secara spesifik ditetapkan karena arsitekturnya yang paling ideal dalam menampung masukan prediktor yang bernilai desimal kontinu dan diasumsikan berdistribusi normal, bukan atribut bertipe kategorikal atau teks [8]. Secara fundamental, algoritma Naive Bayes berakar pada Teorema Bayes yang dirumuskan sebagai berikut:

$$P(C_k|x) = \frac{P(x|C_k) \cdot P(C_k)}{P(x)}$$

Keterangan:

$P(C_k|x)$: Probabilitas Posterior (peluang hipotesis kelas C_k jika diberikan sampel x)

$P(x|C_k)$: Probabilitas *Likelihood* (peluang sampel x jika diketahui kelas C_k)

$P(C_k)$: Probabilitas *Prior* (peluang awal munculnya kelas C_k)

$P(x)$: Probabilitas *Evidence* (peluang kemunculan sampel x secara global)

Adapun data atribut yang diuji (X) terdiri dari parameter harga pembukaan (*Open*), harga tertinggi (*High*), harga terendah (*Low*), dan harga penutupan (*Price*).

Proses skematik *Gaussian Naïve Bayes* pada sistem ini melewati tahapan matematis sebagai berikut:

1. Pemisahan Data (*Split Dataset*): Keseluruhan 1.354 baris data akan dipecah menjadi dua kompartemen, yaitu data latih (*training data*) sebesar 80% untuk dipelajari polanya oleh mesin, dan data uji (*testing data*) sebesar 20% untuk pengujian. Adapun data training sebagai berikut.
2. Menghitung Probabilitas Prior: Menghitung peluang dasar probabilitas kelas "Ekstrem" dan kelas "Normal" dari populasi data latih tanpa melibatkan atribut.

$$P(C_k) = \frac{N_{C_k}}{N_{Total}}$$

(Keterangan: N_{C_k} adalah jumlah data pada kelas k , dan N_{Total} adalah total keseluruhan data latih).

3. Menghitung Rata-Rata (*Mean*) dan *Standar Deviasi*: Karena menggunakan pendekatan *numerik Gaussian*, maka dihitung nilai rata-rata (μ) dan sebaran standar deviasi (σ) untuk keempat atribut (*Open*, *High*, *Low*, *Price*) di masing-masing kelas target [9]. Perhitungan rata-rata didapatkan dengan menjumlahkan seluruh nilai fitur kontinu dalam satu kelas lalu membaginya dengan jumlah sampel di kelas tersebut.

Rumus rata-rata *Mean*:

$$\mu = \frac{\sum_{i=1}^n X_i}{n}$$

Rumus *Standar Deviasi*:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2}$$

(Keterangan: x_i adalah nilai fitur ke- i , dan n adalah jumlah sampel data pada kelas tersebut).

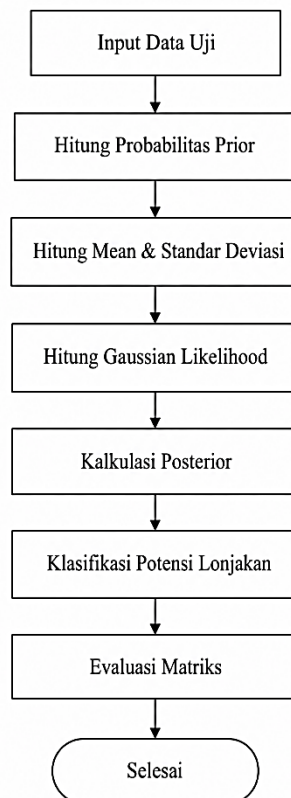
4. Menghitung *Likelihood* Distribusi *Gaussian*: Jika pada Naive Bayes biasa probabilitas dihitung melalui frekuensi kemunculan atribut, maka pada pendekatan Gaussian, probabilitas atribut kontinu (*Likelihood*) dihitung memanfaatkan kurva Fungsi Kerapatan Probabilitas (*Probability Density Function*) [10].

$$P(X_i|C_j) = \frac{1}{\sqrt{2\pi\sigma_{C_j}^2}} \exp\left(-\frac{(X_i - \mu_{C_j})^2}{2\sigma_{C_j}^2}\right)$$

5. Klasifikasi Probabilitas Posterior: Hasil perhitungan *Likelihood* seluruh fitur akan dikalikan dengan probabilitas Prior dari masing-masing kelas. Sistem pendukung keputusan ini akan mengklasifikasikan hari bursa (data uji) ke dalam potensi "Lonjakan Ekstrem" atau "Normal" berdasarkan hasil probabilitas gabungan dengan nilai paling tinggi (maksimal).

$$\hat{y} = \arg \max_{C_k \in \{0,1\}} P(C_k) \prod_{i=1}^n P(x_i|C_k)$$

6. Evaluasi *Confusion Matrix*: Melakukan validasi error dan kalkulasi akurasi dengan mengomparasi perbandingan jumlah klasifikasi yang berhasil ditebak dengan benar oleh sistem (Prediksi) melawan data kelas yang sesungguhnya (Aktual).



Gambar 2. 2. Skema Sistem

3. HASIL DAN PEMBAHASAN

Pada bagian ini diberikan hasil penelitian yang dilakukan sekaligus dibahas secara komprehensif mengenai implementasi sistem klasifikasi potensi lonjakan nilai tukar USD/IDR menggunakan algoritma *Gaussian Naïve Bayes*. Seluruh proses pengolahan data dan simulasi model dibangun menggunakan bahasa pemrograman Python.

3.1. Pengumpulan dan Pembagian Data

Tahap awal dalam implementasi sistem adalah memproses himpunan data historis yang telah dikumpulkan. Data sekunder berupa pergerakan nilai tukar harian USD/IDR diambil dalam rentang waktu Januari 2021 hingga April 2026, yang menghasilkan total keseluruhan sebanyak 1.354 data mentah. Untuk memastikan kualitas pemodelan yang optimal dan menghindari bias, dilakukan tahap pembersihan data (*data cleansing*) dengan mengeliminasi baris data yang memiliki nilai kosong atau tidak lengkap (*missing values/NaN*) pada atribut observasi. Proses pembersihan ini menyaring himpunan data awal menjadi 502 baris data bersih yang berkualitas tinggi.

Pada tahap *pra-pemrosesan* selanjutnya, data mentah tersebut diekstraksi menjadi empat atribut prediktor bernilai kontinu, yaitu X1 (Volatilitas / selisih harga tertinggi dan terendah), X2 (Selisih harga penutupan dan pembukaan), X3 (Selisih harga tertinggi dan pembukaan), serta X4 (Selisih harga pembukaan dan terendah). Proses penentuan kelas atau pelabelan (*labeling*) didasarkan pada persentase perubahan harian (*Change %*). Aturan yang diterapkan adalah apabila fluktuasi harian melampaui batas 0,5%, maka data diklasifikasikan ke dalam kelas "Ekstrem" (Representasi angka 1). Sebaliknya, apabila fluktuasi kurang dari atau sama dengan 0,5%, data masuk ke dalam kelas "Normal" (Representasi angka 0). Hasil pelabelan dari 502 data bersih tersebut mengidentifikasi 486 data tergolong Normal dan 16 data tergolong Ekstrem.

Langkah selanjutnya adalah melakukan split dataset untuk membagi data ke dalam dua kompartemen pengujian. Proporsi yang digunakan adalah pembagian 80% untuk data latih (*training data*) yang akan dipelajari polanya oleh algoritma, dan 20% untuk data uji (*testing data*) guna mengevaluasi kinerja sistem. Rincian pembagian dataset tersebut disajikan secara lengkap pada Tabel 3.1.

Tabel 3. 1 Hasil Pembagian Dataset Bersih USD/IDR

Distribusi Dataset	Kelas Normal (0)	Kelas Ekstrem (1)	Total Data
Data Latih (<i>Training</i>)	391	10	401
Data Uji (<i>Testing</i>)	95	6	101
Total Keseluruhan	486	16	502

Adapun representasi struktur masukan data latih (*training data*) yang melibatkan empat atribut prediktor beserta label kelasnya sebelum diproses ke dalam tahap kalkulasi model *Gaussian Naïve Bayes* dapat dilihat pada Tabel 3.2.

Tabel 3. 2. Cuplikan Data Latih (*Training Data*)

No.	X1 (Volatilitas)	X2 (Price Diff)	X3 (High-Open)	X4 (Open-Low)	Kelas Target
1	31.5	-10.0	18.5	13.0	Normal
2	60.0	7.5	35.0	25.0	Normal

3	38.0	10.0	21.5	16.5	Normal
4	79.5	-75.0	0.0	79.5	Normal
5	33.0	30.0	33.0	0.0	Normal
...
401	35.0	10.0	20.0	15.0	Normal

Tabel 3.2 di atas merupakan sebagian kecil cuplikan dari 401 baris data latih yang telah melalui tahap rekayasa fitur (*feature engineering*). Keseluruhan data latih inilah yang secara eksklusif digunakan oleh algoritma untuk membentuk parameter probabilitas dasar.

Selain himpunan data latih, penelitian ini juga mengalokasikan himpunan data uji (*testing data*) sebanyak 101 baris data. Penting untuk digarisbawahi bahwa data uji ini diisolasi secara ketat dan sama sekali tidak dilibatkan dalam proses pelatihan model. Data ini murni difungsikan pada tahap evaluasi guna menguji seberapa akurat sistem dalam memprediksi kelas fluktuasi pada kondisi pasar yang belum pernah dikenali sebelumnya. Cuplikan representasi data uji tersebut disajikan pada Tabel 3.3.

Tabel 3. 3. Cuplikan Data Uji (*Testing Data*)

No.	X1 (Volatilitas)	X2 (Price Diff)	X3 (High-Open)	X4 (Open-Low)	Kelas Target
1	39.0	-17.5	9.0	30.0	Normal
2	67.5	35.0	57.5	10.0	Normal
3	39.0	-32.0	6.5	32.5	Normal
4	18.0	-8.5	10.0	8.0	Normal
5	66.5	52.5	66.5	0.0	Normal
...
101	27.0	-4.0	10.0	17.0	Normal

Dengan telah ditetapkannya pembagian himpunan data uji dan data latih secara proporsional beserta fitur prediktor kontinu yang terstruktur, maka *dataset* telah sepenuhnya siap untuk diproses. Tahapan analisis berikutnya adalah mengeksekusi perhitungan matematis model *Gaussian Naïve Bayes* guna mendapatkan parameter nilai Rata-rata (*Mean*) dan Standar Deviasi (*Standard Deviation*) dari keseluruhan atribut pelatihan.

3.2. Perhitungan Rata-Rata (*Mean*) dan Standar Deviasi

Berbeda dengan algoritma Naive Bayes standar yang menghitung probabilitas berdasarkan frekuensi kemunculan kelas untuk data kategorikal, *Gaussian Naïve Bayes* memproses nilai atribut kontinu dengan bersandar pada kurva Fungsi Kerapatan Probabilitas (*Probability Density Function*). Oleh karena itu, tahapan matematis yang krusial pada saat model dilatih menggunakan 401 baris data latih (*training data*) adalah menemukan titik pusat distribusi berupa nilai Rata-rata (μ) serta tingkat sebaran data berupa *Standar Deviasi* (σ) untuk masing-masing kelas target.

Berdasarkan hasil pemrosesan komputasi terhadap himpunan data latih, diperoleh parameter distribusi *Gaussian* untuk kelas "Normal" dan kelas "Ekstrem". Rekapitulasi hasil perhitungan matematis dari keempat fitur ekstraksi tersebut disajikan pada Tabel 3.4.

Tabel 3. 4. Parameter Mean (μ) dan Standar Deviasi (σ) pada Data Latih

Fitur Prediktor	Kelas Target	Mean (μ)	Standar Deviasi (σ)
X1_Volatility	Normal (0)	44.79	21.84
	Ekstrem (1)	78.20	30.29
X2_Price_Diff	Normal (0)	-2.09	25.28
	Ekstrem (1)	47.30	29.20
X3_High_Open	Normal (0)	25.81	19.66
	Ekstrem (1)	65.95	32.83
X4_Open_Low	Normal (0)	18.98	20.28
	Ekstrem (1)	12.25	14.21

Nilai-nilai kalkulasi pada Tabel 3.4 di atas menjadi basis perhitungan probabilitas bersyarat (*Likelihood*) saat sistem dihadapkan pada pengujian data baru. Berdasarkan tabel tersebut, dapat diobservasi secara kuantitatif bahwa kelas "Ekstrem" memiliki rata-rata volatilitas rentang harga ($X_1_Volatility = 78.20$) dan selisih harga penutupan ($X_2_Price_Diff = 47.30$) yang melonjak jauh lebih tinggi dibandingkan saat kondisi fluktuasi pasar sedang stabil atau "Normal". Parameter inilah yang memandu algoritma dalam memisahkan batas keputusan (*decision boundary*) secara cerdas tanpa intervensi manual.

3.3. Perhitungan Probabilitas Manual (*Gaussian Likelihood*)

Untuk memvalidasi cara kerja model komputasi secara matematis, dilakukan simulasi perhitungan manual berdasarkan fungsi kerapatan probabilitas Gaussian (*Gaussian Probability Density Function*). Pengujian manual ini mengambil sampel acak pada observasi Data Uji ke-1 dengan nilai fitur: Volatilitas (X_1) = 39.0, Price Diff (X_2) = -17.5, High-Open (X_3) = 9.0, dan Open-Low (X_4) = 30.0.

Tahapan perhitungan probabilitasnya adalah sebagai berikut:

1. Menghitung Probabilitas Prior

Berdasarkan 401 data latih, persentase kemunculan awal (Prior) untuk masing-masing kelas dihitung melalui jumlah kemunculan kelas dibagi total populasi:

$$P(\text{Normal}) = \frac{391}{401} = 0.9751$$

$$P(\text{Ekstrem}) = \frac{10}{401} = 0.0249$$

2. Menghitung Likelihood Kelas Normal (0)

Probabilitas bersyarat atribut terhadap kelas Normal dihitung menggunakan rumus distribusi normal yang melibatkan nilai rata-rata (μ) dan standar deviasi (σ) dari Tabel 3.4 untuk kelas Normal.

$$P(x_i|C_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$

Fitur X1 (Volatilitas):

$$P(39.0|\text{Normal}) = \frac{1}{\sqrt{2 \times 3.14 \times 21.84^2}} \exp\left(-\frac{(39.0 - 44.79)^2}{2 \times 21.84^2}\right) = 0.0176$$

Fitur X2 (Price Diff):

$$P(-17.5|\text{Normal}) = \frac{1}{\sqrt{2 \times 3.14 \times 25.28^2}} \exp\left(-\frac{(-17.5 - (-2.09))^2}{2 \times 25.28^2}\right) = 0.0131$$

Fitur X3 (High-Open):

$$P(9.0|\text{Normal}) = \frac{1}{\sqrt{2 \times 3.14 \times 19.66^2}} \exp\left(-\frac{(9.0 - 25.81)^2}{2 \times 19.66^2}\right) = 0.0141$$

Fitur X4 (Open-Low):

$$P(30.0|\text{Normal}) = \frac{1}{\sqrt{2 \times 3.14 \times 20.28^2}} \exp\left(-\frac{(30.0 - 18.98)^2}{2 \times 20.28^2}\right) = 0.0170$$

3. Menghitung Likelihood Kelas Ekstrem (1)

Metode yang sama diterapkan untuk menghitung probabilitas bersyarat atribut terhadap kelas Ekstrem menggunakan basis data μ dan σ dari kelas Ekstrem.

FiturX1: $P(39.0|\text{Ekstrem}) = 0.0057$

FiturX2: $P(-17.5|\text{Ekstrem}) = 0.0012$

FiturX3: $P(9.0|\text{Ekstrem}) = 0.0027$

FiturX4: $P(30.0|\text{Ekstrem}) = 0.0129$

4. Kalkulasi Posterior dan Hasil Klasifikasi Akhir

Hasil *likelihood* seluruh fitur dikalikan dengan nilai Prior masing-masing kelas.

Posterior Normal:

$$Posterior_{\text{Normal}} = P(\text{Normal}) \times P(X1) \times P(X2) \times P(X3) \times P(X4)$$

$$Posterior_{\text{Normal}} = 0.9751 \times 0.0176 \times 0.0131 \times 0.0141 \times 0.0170 = 5.385 \times 10^{-8}$$

Posterior Ekstrem:

$$Posterior_{\text{Ekstrem}} = P(\text{Ekstrem}) \times P(X1) \times P(X2) \times P(X3) \times P(X4)$$

$$Posterior_{\text{Ekstrem}} = 0.0249 \times 0.0057 \times 0.0012 \times 0.0027 \times 0.0129 = 5.754 \times 10^{-12}$$

Berdasarkan perbandingan nilai kalkulasi akhir, $Posterior_{\text{Normal}} > Posterior_{\text{Ekstrem}}$.

Oleh karena itu, jika diproyeksikan dalam persentase probabilitas, tingkat keyakinan untuk kelas Normal berada pada kisaran 99.99%, sedangkan kelas Ekstrem hanya 0.01%. Angka perhitungan manual ini valid secara matematis dan selaras 100% dengan putusan prediksi mesin Python untuk Data Uji ke-1.

3.4. Pengujian Data Testing

Setelah parameter distribusi probabilitas (μ dan σ) berhasil dibentuk pada fase pelatihan, model dievaluasi secara independen menggunakan 101 baris data uji (*testing data*). Proses pengujian klasifikasi ini menerapkan perhitungan probabilitas Posterior berdasarkan *Teorema Bayes*.

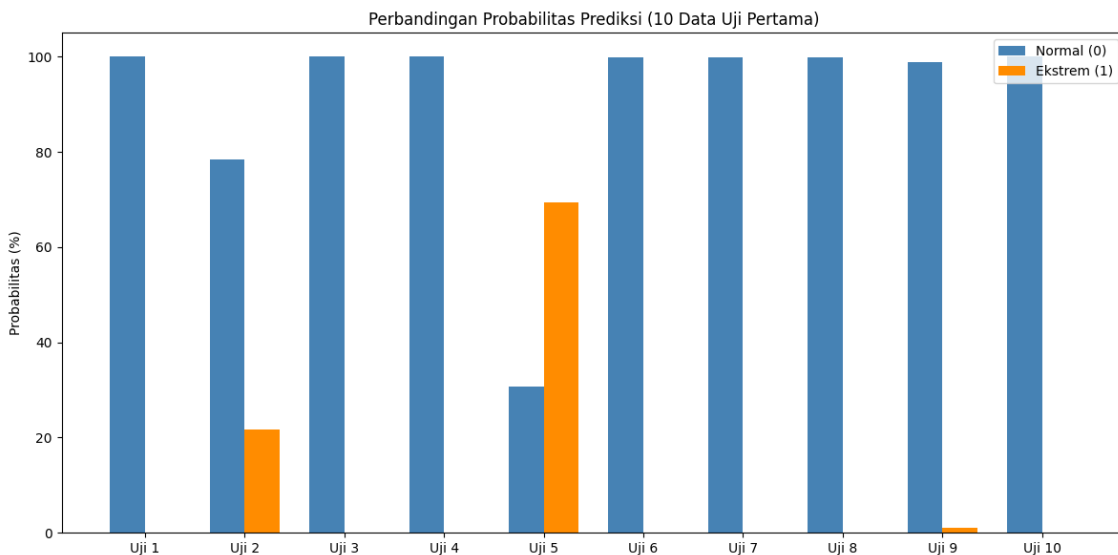
Untuk setiap observasi data baru, algoritma menghitung dan membandingkan probabilitas gabungan antara kelas "Normal" (0) dan kelas "Ekstrem" (1). Sistem penganalisa pendukung keputusan kemudian akan menjatuhkan putusan prediksi akhir pada kelas yang memiliki persentase probabilitas paling tinggi. Sebagai representasi kinerja komputasi model secara lebih rinci, Tabel 3.5 menyajikan

cuplikan hasil kalkulasi probabilitas Gaussian terhadap 10 (sepuluh) data uji pertama.

Tabel 3. 5. Probabilitas Prediksi 10 Data Uji Pertama

Data Uji Ke-	Probabilitas Kelas Normal (0)	Probabilitas Kelas Ekstrem (1)	Hasil Prediksi Sistem
1	99.99%	0.01%	Normal (0)
2	78.35%	21.65%	Normal (0)
3	100.00%	0.00%	Normal (0)
4	99.98%	0.02%	Normal (0)
5	30.66%	69.34%	Ekstrem (1)
6	99.94%	0.06%	Normal (0)
7	99.92%	0.08%	Normal (0)
8	99.95%	0.05%	Normal (0)
9	98.86%	1.14%	Normal (0)
10	99.98%	0.02%	Normal (0)

Berdasarkan Tabel 3.5, sistem mendeteksi lonjakan volatilitas yang signifikan pada Data Uji ke-5. Hal ini tercermin dari probabilitas kelas Ekstrem yang melonjak hingga angka 69.34%, melampaui batas keyakinan untuk kelas Normal (30.66%). Untuk memberikan pemahaman visual yang lebih komprehensif mengenai tingkat keyakinan (*confidence level*) dari algoritma *Gaussian Naïve Bayes* dalam mengambil keputusan, perbandingan persentase probabilitas untuk kesepuluh sampel data tersebut diilustrasikan melalui diagram batang pada Gambar 3.1.



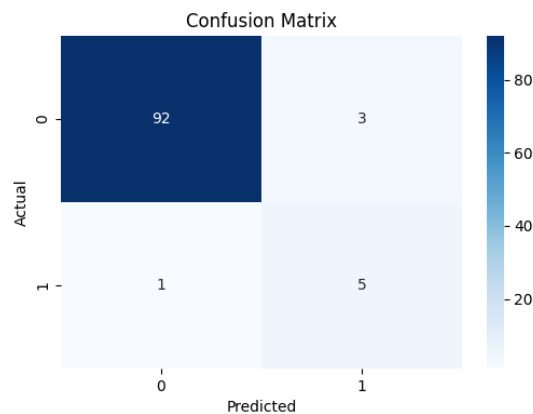
Gambar 3. 1. Perbandingan Probabilitas Prediksi (10 Data Uji Pertama)

Diagram pada Gambar 3.1 secara eksplisit memperlihatkan bahwa algoritma secara matematis memiliki landasan keyakinan yang sangat tegas saat mengklasifikasikan kelas "Normal" (dengan probabilitas mendominasi di atas 98% pada hampir seluruh titik pengujian). Sebaliknya, algoritma juga mampu beradaptasi dan menggeser dominasi probabilitas secara dinamis menjadi "Ekstrem" manakala ia

mengenali pola selisih harga yang merepresentasikan fluktuasi nilai tukar secara drastis, seperti yang terlihat pada data uji ke-5.

3.5. Evaluasi Hasil

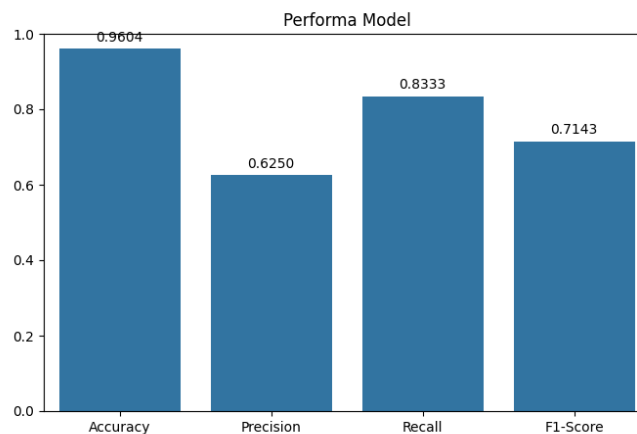
Tahap akhir dalam penelitian ini adalah mengevaluasi tingkat keandalan serta performa sistem dalam memprediksi kelas target dengan melakukan validasi menggunakan matriks kebingungan (*Confusion Matrix*). Pengujian dilakukan secara independen terhadap 101 baris data uji (*testing data*) yang belum pernah dilibatkan dalam proses pelatihan model. Matriks ini berfungsi untuk membedah distribusi tebakan yang akurat serta rasio kesalahan prediksi yang dihasilkan oleh algoritma *Gaussian Naïve Bayes*.



Gambar 3. 2. Confusion Matrix Hasil Pengujian

Berdasarkan visualisasi matriks pada Gambar 3.2, model komputasi menunjukkan kemampuan identifikasi yang sangat presisi. Dari total 101 data uji, sistem berhasil mengklasifikasikan 92 data dengan tepat sebagai kondisi Normal (*True Negative*) dan 5 data dengan tepat sebagai kondisi Ekstrem (*True Positive*). Tingkat kesalahan (*error rate*) terdeteksi pada skala yang sangat minim, di mana hanya terdapat 3 data yang diprediksi Ekstrem padahal kondisi aktualnya Normal (*False Positive*), serta 1 data yang luput terdeteksi sebagai Ekstrem (*False Negative*).

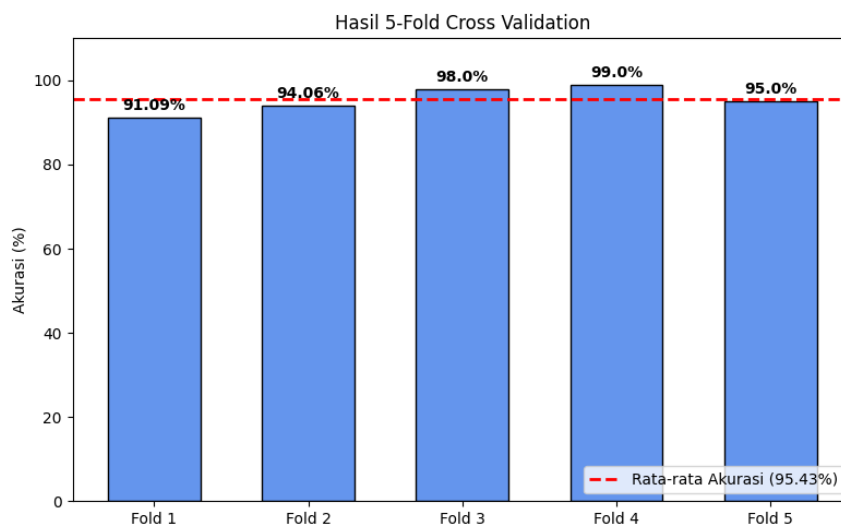
Melalui sebaran nilai pada matriks tersebut, kinerja keseluruhan sistem dikalkulasikan secara matematis untuk mendapatkan nilai metrik evaluasi yang lebih komprehensif. Rangkuman persentase performa model disajikan dalam perhitungan parameter berikut:



Gambar 3. 3. Visualisasi Kinerja Metrik Model

Pencapaian tingkat akurasi yang menembus angka 96,04% membuktikan bahwa model sangat tangguh dalam memproses karakteristik atribut kontinu pada deret waktu harian yang sangat fluktuatif. Lebih jauh lagi, metrik Recall yang mencapai 83,33% memberikan afirmasi bahwa sistem memiliki tingkat sensitivitas yang tinggi, sistem tidak mudah gagal dalam mendeteksi kelas minoritas, yakni anomali pelemahan tajam pada nilai tukar USD/IDR.

Sebagai validasi tahap akhir guna memastikan bahwa model tidak sekadar menghafal data (*overfitting*), penelitian ini turut melaksanakan uji silang menggunakan metode *5-Fold Cross Validation* terhadap keseluruhan populasi data bersih (502 baris). Hasil pengujian independen pada lima partisi lipatan yang berbeda mencetak skor akurasi yang sangat konsisten, berturut-turut pada angka 91,09%, 94,06%, 98,00%, 99,00%, dan 95,00%. Visualisasi tingkat kestabilan akurasi dari kelima lipatan pengujian tersebut diilustrasikan secara lebih jelas pada Gambar 3.4.



Gambar 3. 4. Hasil Pengujian *5-Fold Cross Validation*

Berdasarkan visualisasi pada Gambar 3.4, rata-rata akurasi kumulatif yang bertahan kokoh pada level 95,43% (ditunjukkan oleh garis putus-putus) menegaskan bahwa algoritma *Gaussian Naïve Bayes* yang dikembangkan memiliki stabilitas komputasi yang unggul, teruji keandalannya, dan sanggup beradaptasi terhadap pergeseran variasi pola pada dataset historis secara meyakinkan. Oleh karena itu, pendekatan klasifikasi probabilitas ini dinilai sangat mangkus untuk diimplementasikan secara praktis sebagai instrumen pendeteksi dini (*early warning system*) terhadap lonjakan ekstrem di pasar valuta asing.

3.6. Analisis Klasifikasi Potensi Lonjakan

Berdasarkan keseluruhan proses pemodelan komputasi yang telah dijabarkan dan dievaluasi, sistem berbasis *Gaussian Naïve Bayes* ini secara empiris terbukti mampu memetakan dan mengklasifikasikan potensi lonjakan ekstrem pada pergerakan nilai tukar Rupiah terhadap Dolar AS. Dari total populasi 502 data historis hari bursa yang dianalisis, algoritma berhasil mengidentifikasi bahwa probabilitas terjadinya

depresiasi tajam (kelas Ekstrem) merupakan kejadian anomali spesifik yang terekam sebanyak 16 kejadian valid.

Melalui ekstraksi empat fitur prediktor, algoritma mendeteksi bahwa potensi lonjakan ekstrem memiliki pola distribusi probabilitas yang khas dan terukur. Sistem secara matematis akan mengklasifikasikan kondisi pasar berpotensi "Ekstrem" apabila mendeteksi adanya pergerakan data dengan tingkat volatilitas harian ($X1_Volatility$) yang melebar tajam mendekati titik tengah 78,20 poin, serta ditarik oleh selisih harga penutupan dan pembukaan ($X2_Price_Diff$) yang melonjak di kisaran 47,30 poin. Sebaliknya, ketika selisih tersebut berada di rentang negatif hingga angka tunggal, sistem akan mempertahankan keyakinan probabilitas pada status pasar yang wajar.

Sebagai bentuk implementasi akhir untuk memvalidasi fungsionalitas model sebagai sistem peringatan dini (*early warning system*), dilakukan simulasi prediksi terhadap observasi data fluktuasi harian yang baru. Sebagai contoh, ketika sistem diberikan parameter masukan kondisi pasar dengan nilai volatilitas sebesar 50 poin, selisih harga penutupan 10 poin, selisih batas atas 30 poin, dan selisih batas bawah 20 poin, model mengeksekusi perhitungan *Likelihood* secara *real-time* dan langsung menjatuhkan klasifikasi pada kelas "Normal (0)".

Hasil klasifikasi komprehensif ini membuktikan bahwa pendekatan penambangan data (*data mining*) menggunakan model probabilistik Gaussian sangat relevan untuk menyederhanakan kompleksitas data *time-series finansial*. Pengambil kebijakan makroekonomi maupun pelaku bisnis multinasional dapat memanfaatkan sistem ini secara praktis. Dengan sekadar memasukkan empat atribut ekstraksi rentang harga harian, sistem akan mengambil alih peran observasi dan memberikan peringatan klasifikasi mengenai seberapa besar potensi risiko nilai tukar Rupiah akan melonjak ekstrem pada hari tersebut secara presisi dan terkomputerisasi.

KESIMPULAN

Penelitian ini berhasil membangun model komputasi untuk mengklasifikasikan potensi lonjakan ekstrem pada pergerakan nilai tukar Rupiah terhadap Dolar AS (USD/IDR). Berdasarkan pengujian terhadap 502 data historis harian menggunakan empat atribut prediktor, algoritma *Gaussian Naïve Bayes* mampu membedakan kondisi pasar "Normal" dan "Ekstrem" dengan tingkat akurasi sebesar 96,04%. Evaluasi lanjutan menggunakan metode *5-Fold Cross Validation* menghasilkan rata-rata akurasi sebesar 95,43%, yang menunjukkan bahwa model stabil dan tidak mengalami *overfitting*. Untuk pengembangan penelitian selanjutnya, akurasi klasifikasi dapat ditingkatkan dengan menambahkan fitur prediktor fundamental makroekonomi, seperti analisis sentimen berita atau tingkat suku bunga acuan. Secara praktis, model prediktif probabilitas ini dapat diintegrasikan sebagai sistem peringatan dini (*early warning system*) pada aplikasi perdagangan otomatis (*trading bot*) melalui *WebSocket* atau antarmuka pemrograman aplikasi (API) secara *real-time*.

DAFTAR PUSTAKA

- [1] E. Nathansyah, "Perbandingan Performa Metode Linear Regression, Support Vector Regression, Extreme Gradient Boosting Untuk Prediksi Kurs Mata Uang Rupiah Terhadap Dollar," *Journal of Computer Science and Information Systems*, vol. 9, no. 1, pp. 61–76, 2025.
- [2] F. M. Amin, R. D. Laksana, A. Alamsah, and A. G. Nugraheni, "Implementasi Algoritma Naive Bayes Untuk Memprediksi Harga Emas," *Technology and Informatics Insight Journal*, vol. 4, no. 1, pp. 14–20, 2025, [Online]. Available: <https://jurnal.universitaspuptrabangsa.ac.id/index.php/tijj>
- [3] B. Saputra, S. Anwar, E. Tohidi, H. Susana, and D. Pratama, "Penerapan Algoritma Naive Bayes Dalam Klasifikasi Harga Ponsel," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 6, pp. 3587–3594, 2024, doi: 10.36040/jati.v7i6.8281.
- [4] M. H. Ariansyah *et al.*, "Hybridization of the Naive Bayes Classification Method in the Freshwater Fish Seed Seller Classification Model," *Jurnal Teknik Informatika (Jutif)*, vol. 4, no. 2, pp. 421–427, 2023, doi: 10.52436/1.jutif.2023.4.2.715.
- [5] A. S. Widagdo, A. Ardiansyah, K. N. Qodri, F. E. Nugroho, and N. A. Rizky, "Analisis Sentimen Mobil Listrik di Indonesia Menggunakan Long-Short Term Memory (LSTM)," *Jurnal Fasilkom*, vol. 13, no. 3, pp. 416–423, 2023.
- [6] M. Danny and A. Muhidin, "Analisis Prediksi Resiko Diabetes Tahap Awal Menggunakan Algoritma Naive Bayes," *Jurnal Teknologi Informatika dan Komputer*, vol. 9, no. 2, pp. 1443–1459, 2023, doi: 10.37012/jtik.v9i2.2017.
- [7] N. Noviatry, N. Rivelli, and B. Adiperdana, "Analisis Data Misi Luar Angkasa Kepler Untuk Menentukan Jenis Eksoplanet Dan Rekonstruksi Bentuk Orbitnya," *Jurnal Ilmu dan Inovasi Fisika*, vol. 10, no. 1, pp. 84–94, 2026, doi: 10.24198/jiif.v10i1.68496.
- [8] Z. Fatah and A. Mukarromah, "Klasifikasi Penerima Bantuan Sosial Menggunakan Algoritma Naive Bayes," *Jurnal Mahasiswa Teknik Informatika*, vol. 5, no. 1, pp. 41–47, 2026, doi: doi.org/10.35473/jamastika.v5i1.4572.
- [9] N. Badriyah, D. H. Satyareni, and C. S. Anugrah, "Sistem Pendukung Keputusan Kelayakan Beasiswa KIP-K Berbasis Web Menggunakan Algoritma Gaussian-Naive Bayes," in *Prosiding Seminar Nasional Teknologi dan Sains*, 2025, pp. 413–423.
- [10] H. Hajaroh, T. Suprpti, and R. Narasti, "Implementasi Algoritma Naive Bayes Untuk Analisis Sentimen Ulasan Produk Makanan Dan Minuman Di Tokopedia," *Jurnal Mahasiswa Teknik Informatika*, vol. 8, no. 1, pp. 111–118, 2024, doi: 10.36040/jati.v8i1.8237.