Laporan Tugas Data Mining

Deteksi Anomali pada Dataset Credit Card Fraud

Pendekatan Statistical, Proximity, dan Density

Identitas Kelompok 12

NRP	Nama
5054241029	Wayan Raditya Putra
5054241026	Arya Pratama Rhama Putra
5054231020	Dean Arkeanu Andharu

BAB 1 — Pendahuluan

Deteksi anomali (anomaly detection) adalah proses mengidentifikasi objek atau titik data yang menunjukkan perilaku menyimpang dari pola umum mayoritas. Teknik ini memiliki peran penting dalam berbagai sistem yang membutuhkan deteksi dini terhadap aktivitas tidak wajar, seperti keamanan jaringan, monitoring kesehatan sistem, dan deteksi transaksi finansial yang berpotensi fraud.

Pada tugas ini, pendekatan anomaly detection diterapkan pada *Credit Card Fraud Detection Dataset* dari Kaggle. Dataset tersebut sering digunakan sebagai studi kasus karena karakteristiknya yang menantang dan relevan dengan skenario nyata, di mana transaksi fraud hanya merupakan sebagian kecil dari keseluruhan data, sehingga tidak ideal ditangani dengan supervised learning secara langsung.

Laporan ini bertujuan untuk:

1. Mempelajari konsep anomaly dan outlier sesuai Bab 10 dari buku *Data Mining: Concepts and Techniques* oleh Pang-Ning Tan.
2. Mengimplementasikan tiga pendekatan anomaly detection, yaitu:
   - Statistical Approach
   - Proximity Approach
   - Density Approach
3. Menerapkan ketiga pendekatan tersebut pada dataset Credit Card Fraud Detection.
4. Mengevaluasi performa setiap pendekatan menggunakan metrik akurasi, precision, recall, dan waktu komputasi.
5. Memberikan analisis perbedaan performa dan menarik kesimpulan mengenai metode yang paling sesuai untuk mendeteksi anomali pada skenario fraud detection.

Melalui laporan ini, diharapkan pembaca dapat memahami perbedaan karakteristik setiap pendekatan anomaly detection serta bagaimana penerapannya pada data dunia nyata.

BAB 2 — Dasar Teori

2.1 Anomaly dan Outlier

Anomaly merupakan objek atau titik data yang memiliki karakteristik berbeda secara signifikan dibandingkan sebagian besar data lainnya. Dalam banyak aplikasi, anomali dapat merepresentasikan kejadian penting, seperti kesalahan sistem, aktivitas mencurigakan, atau potensi fraud.

Outlier adalah bentuk anomaly yang muncul sebagai titik penyimpang dalam distribusi data. Keberadaan outlier dapat disebabkan oleh noise, kesalahan pencatatan, atau indikasi kejadian nyata yang tidak mengikuti pola umum.

Deteksi anomaly bertujuan untuk mengidentifikasi titik-titik tersebut secara otomatis dengan berbagai pendekatan, baik yang berbasis statistik, jarak, maupun kepadatan data.

2.2 Statistical Approach

Pendekatan statistik didasarkan pada asumsi bahwa data normal mengikuti pola distribusi tertentu, umumnya distribusi Gaussian. Data yang berada jauh dari pusat distribusi atau memiliki probabilitas sangat kecil di bawah model probabilistik dianggap sebagai anomaly.

Metode populer dalam pendekatan ini adalah *Elliptic Envelope*, yang memodelkan data normal dalam bentuk kurva elips berdasarkan mean dan covariance. Titik data yang berada di luar batas elips tersebut diklasifikasikan sebagai outlier.

Pendekatan ini efektif jika distribusi data relatif teratur, namun sensitif terhadap data yang tidak memenuhi asumsi distribusi tertentu.

2.3 Proximity Approach

Pendekatan proximity mendeteksi anomaly berdasarkan jarak antar titik data. Ide dasarnya adalah bahwa titik yang jauh dari sebagian besar titik lain, atau memiliki sedikit tetangga dekat, cenderung merupakan outlier.

Contoh metode proximity adalah *k-Nearest Neighbors (k-NN) Outlier Detection*, di mana jarak ke tetangga terdekat digunakan sebagai skor anomaly. Semakin besar jaraknya, semakin besar kemungkinan titik tersebut merupakan anomaly.

Pendekatan ini intuitif dan tidak membutuhkan asumsi distribusi, namun performanya dapat terpengaruh oleh dimensi data yang tinggi.

2.4 Density Approach

Pendekatan density mendeteksi anomaly berdasarkan tingkat kepadatan lokal suatu titik. Jika suatu titik memiliki kepadatan yang jauh lebih rendah dibandingkan tetangga sekitarnya, maka titik tersebut dianggap sebagai outlier.

Metode yang umum digunakan adalah *Local Outlier Factor (LOF)*, yang menghitung perbandingan local reachability density antara sebuah titik dengan tetangganya. Nilai LOF lebih besar dari 1 menunjukkan bahwa titik tersebut memiliki kepadatan lebih rendah dan berpotensi menjadi anomaly.

Pendekatan ini fleksibel untuk data yang memiliki variasi kepadatan, namun lebih mahal secara komputasi dibandingkan pendekatan statistik sederhana.

Dengan memahami ketiga pendekatan tersebut, langkah selanjutnya adalah menerapkannya pada dataset Credit Card Fraud Detection serta membandingkan performanya dalam mendeteksi transaksi yang tidak normal.

BAB 3 — Exploratory Data Analysis (EDA) dan Preprocessing

Pada bagian ini dilakukan eksplorasi awal terhadap dataset untuk memahami struktur data, mengecek nilai kosong, serta melihat distribusi label. Tahapan EDA dilakukan secara sederhana karena fokus utama tugas adalah implementasi tiga pendekatan anomaly detection.

In [34]:
```
import pandas as pd

df = pd.read_csv("data/creditcard.csv")

df.head()
```

Out[34]:

	Time	V1	V2	V3	V4	V5	V6	V7	
0	0.0	-1.359807	-0.072781	2.536347	1.378155	-0.338321	0.462388	0.239599	0.0
1	0.0	1.191857	0.266151	0.166480	0.448154	0.060018	-0.082361	-0.078803	0.0
2	1.0	-1.358354	-1.340163	1.773209	0.379780	-0.503198	1.800499	0.791461	0.2
3	1.0	-0.966272	-0.185226	1.792993	-0.863291	-0.010309	1.247203	0.237609	0.3
4	2.0	-1.158233	0.877737	1.548718	0.403034	-0.407193	0.095921	0.592941	-0.2

5 rows × 31 columns

In [35]:
```
df.isnull().sum()
```

Out[35]:
```
Time      0
V1        0
V2        0
V3        0
V4        0
V5        0
V6        0
V7        0
V8        0
V9        0
V10       0
V11       0
V12       0
V13       0
V14       0
V15       0
V16       0
V17       0
V18       0
V19       0
V20       0
V21       0
V22       0
V23       0
V24       0
V25       0
V26       0
V27       0
V28       0
Amount    0
Class     0
dtype: int64
```

In [36]:
```
df['Class'].value_counts(normalize=True)
```

Out[36]:
```
Class
0    0.998273
1    0.001727
Name: proportion, dtype: float64
```

4.1 Statistical Approach

Metode yang digunakan adalah Elliptic Envelope, yang memodelkan distribusi data normal berdasarkan mean dan covariance. Titik yang berada di luar batas elips diprediksi sebagai outlier (fraud).

Pada tahap implementasi:

- Data dinormalisasi menggunakan StandardScaler.
- Model dilatih menggunakan seluruh fitur kecuali label.
- Proporsi fraud pada data training digunakan sebagai parameter contamination.

Hasil prediksi diubah dari label model: -1 → 1 (fraud)
1 → 0 (normal)

Metrik evaluasi yang dihitung meliputi akurasi, precision, recall, serta waktu komputasi. Interpretasi disampaikan pada bagian analisis hasil.

In [37]:
```
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import numpy as np
X = df.drop(columns=["Class"])
y = df["Class"]

X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42, stratify=y
)
X['Amount'] = np.log1p(X['Amount'])
scaler = StandardScaler()
X_train_s = scaler.fit_transform(X_train)
X_test_s = scaler.transform(X_test)
```

In [38]:
```
import time
import numpy as np
from sklearn.covariance import EllipticEnvelope
from sklearn.metrics import accuracy_score, precision_score, recall_score

# contamination berdasarkan proporsi fraud di training
contamination = y_train.mean()
contamination
```

Out[38]:
```
np.float64(0.0017254870488152324)
```

In [39]:
```
stat_model = EllipticEnvelope(
    contamination=contamination,
    random_state=42
)
```

In [40]:
```
t0 = time.time()
stat_model.fit(X_train_s)
fit_time_stat = time.time() - t0
fit_time_stat
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-277.170309786192490 > -278.57329941
0424056). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-273.403645766972204 > -277.62015973
1646538). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-185.237800087885802 > -185.92887089
5933301). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-215.952886770687456 > -240.49210628
5528621). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-253.357737589942019 > -274.49468490
4542737). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-243.233725165078965 > -243.91402527
5566814). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-215.726411152438118 > -242.07795447
8330469). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-214.724192415204925 > -243.16442555
4403408). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-276.767421698827320 > -280.06615772
8433268). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-274.340067492347657 > -276.14792421
8463572). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-249.088711247365637 > -251.26099753
5749595). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-274.973398573185477 > -275.83823592
7799190). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-214.759888385087152 > -240.46231488
4177772). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-244.208972988928650 > -272.72495027
4833986). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-247.649460929705640 > -248.33670985
2136835). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-239.523831372832092 > -240.87926838
5737419). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-211.804003583931063 > -238.71875225
3840648). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-239.993563491622069 > -265.64749528
4177694). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-160.734375903709292 > -328.20341868
0908385). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-244.931303672190722 > -246.73626832
4316740). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-250.610151551411548 > -251.81613204
8219060). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-250.746292323232240 > -276.83470904
4663271). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-217.130614110964132 > -219.78522998
0037883). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-246.115905221702178 > -247.26797407
3445828). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-260.106285499507123 > -264.20035903
7293254). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-248.333231271623532 > -251.03509504
6458139). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-220.792250385837633 > -221.39226606
1356793). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-246.353289926390460 > -247.05872774
2999224). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-212.873268846207907 > -239.90464532
9834182). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-238.590611041011499 > -265.94365910
4722769). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-213.644361193303553 > -240.35619092
4585718). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-249.459923855009976 > -250.03977402
2422307). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-243.198702790028960 > -243.60990233
6301673). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-219.348822776553419 > -223.62613154
4812864). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-233.227779975236842 > -278.55531533
5733269). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-158.862818921235970 > -159.15237405
0802536). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-188.161926598045369 > -190.28764206
3216254). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-270.938610121122906 > -273.51411740
8939740). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-249.466587845323545 > -250.09766302
6435953). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-250.000673634461549 > -276.72916769
8836420). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-206.268441149885859 > -268.91939359
8405634). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-237.945442377854761 > -266.50907742
3397002). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-214.038100549353715 > -242.34306304
6849437). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-190.600013659233383 > -191.58819634
9412272). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-219.148714557008589 > -219.28347636
5356620). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-215.076957655413992 > -242.63709168
7320890). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-185.356511855645863 > -186.83394704
7764127). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-160.754426103133852 > -161.71707377
6602462). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-272.425304739262003 > -274.57506168
2701914). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-243.140936955908131 > -245.07727671
9200029). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-288.072958320805128 > -291.13753610
2508932). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-273.568058375854889 > -276.46961415
4455428). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-245.824479895688938 > -274.84863080
7935763). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-266.872070679470880 > -268.45061940
4110626). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-244.867969407121080 > -245.50255767
4421951). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-154.641501700621376 > -267.19285524
6795546). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-214.600389086451827 > -243.67834284
4017408). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-213.934406209102804 > -240.55733421
9235230). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-220.571633142702439 > -221.22973205
7912031). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-208.646609802050136 > -263.03085532
0205831). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-237.795207714952596 > -238.52976832
7402280). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-161.038829892551462 > -161.21357768
1772705). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-271.452849591327379 > -274.09061357
4627469). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-270.122694437005009 > -272.32234606
2832537). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-273.492387010291907 > -273.51314221
1989020). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-258.033178377457830 > -259.55413167
2995709). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-259.201381727073908 > -261.19992885
7397026). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-211.578000885557913 > -241.04467295
1608618). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-241.035519038014456 > -241.17902625
2417316). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-275.927859731027638 > -277.54331255
1481279). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-214.159722153826351 > -216.60951011
2778452). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-150.498828345933504 > -151.05930844
1159061). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-181.779911142452107 > -183.26290815
8309870). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-271.643991911260628 > -273.89075949
4875851). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-217.099707787553939 > -241.13710818
0462803). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-182.184094888097746 > -267.93811312
5449263). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-245.252873170726559 > -245.92052063
7058843). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-272.692296081580764 > -276.06878131
2595206). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-218.197827603622358 > -243.01188863
4424309). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-165.832997166187170 > -297.30081173
2514489). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-250.953787503773526 > -252.93006771
3615756). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-217.470388665990839 > -217.70542410
7577812). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-250.746292323232240 > -276.83470904
4663271). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-215.915654478954195 > -242.77867639
8262974). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-219.907886756479513 > -221.14718648
5928728). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-213.098073313553357 > -242.98361766
7394014). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-212.752340091978226 > -241.29424801
2998025). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-215.186210856124546 > -241.47016740
5027354). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-247.643695700466367 > -251.17995003
0243134). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-244.179046381742126 > -245.66961495
6298318). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-239.708795467668409 > -267.25829035
4422741). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-238.642797661968018 > -269.22718539
8344432). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-241.062948787454644 > -265.69399911
6664486). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-188.444512431383515 > -188.66632921
2288701). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-274.627612971471592 > -275.26709698
6711806). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-248.508169647287076 > -248.64815991
1497032). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-215.301283334431304 > -241.88982254
7955589). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-247.930327788507043 > -273.73973559
5124216). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-244.563283915564966 > -246.85123533
3579808). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-213.821897832093896 > -241.24895911
5791365). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-238.007532493939834 > -238.04097632
8559708). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-182.419906001760438 > -184.23375207
7710676). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-215.739342424527564 > -241.28903418
7840628). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-216.773873276366601 > -217.34527349
0604313). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-246.885609474653251 > -250.13652841
0104972). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-248.374253184559024 > -249.50308008
0816801). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-176.514417638673279 > -238.43757096
0010021). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-216.489832168844771 > -216.71528358
5182533). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-258.101576246981665 > -276.10843379
7479904). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-233.476468159822076 > -276.72497973
5921750). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
```

```
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-218.821531762511768 > -219.37331636
8323856). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-244.537489765305537 > -245.99796214
7446771). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-261.604347025171819 > -262.88181178
7569632). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-274.103191979701876 > -275.46731246
3207463). You may want to try with a higher value of support_fraction (current va
lue: 0.501).
  warnings.warn(
/home/themechanic/envs/ML-project/lib/python3.12/site-packages/sklearn/covarianc
e/_robust_covariance.py:185: RuntimeWarning: Determinant has increased; this shou
ld not happen: log(det) > log(previous_det) (-118.090964829610485 > -269.21330617
5166394). You may want to try with a higher value of support_fraction (current va
lue: 0.500).
  warnings.warn(
```

Out[40]: 22.15370488166809

In [41]:
```
t0 = time.time()
pred_stat = stat_model.predict(X_test_s)
pred_time_stat = time.time() - t0

# konversi:
# -1 (outlier) -> 1 (fraud)
#  1 (inlier)  -> 0 (normal)
pred_stat_bin = np.where(pred_stat == -1, 1, 0)
```

In [42]:
```
acc_stat = accuracy_score(y_test, pred_stat_bin)
prec_stat = precision_score(y_test, pred_stat_bin, zero_division=0)
rec_stat = recall_score(y_test, pred_stat_bin, zero_division=0)

acc_stat, prec_stat, rec_stat, fit_time_stat, pred_time_stat
```

Out[42]:
```
(0.9970389616469459,
 0.1568627450980392,
 0.16216216216216217,
 22.15370488166809,
 0.05080914497375488)
```

4.1 Proximity Approach (kNN-based Outlier Detection)

Tahap Implementasi:

1. Menentukan K= 5
2. Menghitung jarak ke 5 tetangga
3. Merata-ratakan jarak ke tetangganya
4. memasang threshold sebesar kuartil ke (1- nilai kontaminasi)
5. Jika jarak lebih besar dari threshold maka dia fraud dan jika tidak maka dia normal

In [43]:
```
from sklearn.neighbors import NearestNeighbors
```

In [44]:
```
k = 5
t0 = time.time()
knn_model = NearestNeighbors(n_neighbors=k)
knn_model.fit(X_train_s)
fit_time_knn = time.time() - t0
distances_train, _ = knn_model.kneighbors(X_train_s)
avg_dist_train = distances_train.mean(axis=1)
threshold = np.quantile(avg_dist_train, 1 - contamination)
distances_test, _ = knn_model.kneighbors(X_test_s)
avg_dist_test = distances_test.mean(axis=1)
t0 = time.time()
pred_knn_bin = (avg_dist_test > threshold).astype(int)
pred_time_knn = time.time() - t0
```

In [45]:
```
acc_knn = accuracy_score(y_test, pred_knn_bin)
prec_knn = precision_score(y_test, pred_knn_bin, zero_division=0)
rec_knn = recall_score(y_test, pred_knn_bin, zero_division=0)

print(f"k-NN Accuracy  : {acc_knn:.4f}")
print(f"k-NN Precision : {prec_knn:.4f}")
print(f"k-NN Recall    : {rec_knn:.4f}")
```

```
k-NN Accuracy  : 0.9947
k-NN Precision : 0.0744
k-NN Recall    : 0.1824
```

4.2 Density Approach (Local Outlier Factor)

Tahap Implementasi:

1. membuat sebuah variabel untuk menyimpan model LocalOutlierFactor dengan tetangga 20 novelty true untuk mendeteksi data baru.

cara kerja model ini simple. mereka menghitung kepadatan lokalnya terlebih dahulu. jarak 10 akan dianggap normal jika berada di kepadatan lokal yang rendah dibandingkan kepadatan lokal yang rapat

In [46]:
```
from sklearn.neighbors import LocalOutlierFactor
```

In [47]:
```
t0 = time.time()
lof_model = LocalOutlierFactor(
    n_neighbors=20,
```

```
    contamination=contamination,
    novelty=True
)

lof_model.fit(X_train_s)
fit_time_lof = time.time() - t0
```

In [48]:
```
t0 = time.time()
pred_lof = lof_model.predict(X_test_s)
pred_time_lof = time.time() - t0
pred_lof_bin = np.where(pred_lof == -1, 1, 0)
```

In [49]:
```
acc_lof = accuracy_score(y_test, pred_lof_bin)
prec_lof = precision_score(y_test, pred_lof_bin, zero_division=0)
rec_lof = recall_score(y_test, pred_lof_bin, zero_division=0)

print(f"LOF Accuracy  : {acc_lof:.4f}")
print(f"LOF Precision : {prec_lof:.4f}")
print(f"LOF Recall    : {rec_lof:.4f}")
```

```
LOF Accuracy  : 0.9967
LOF Precision : 0.0000
LOF Recall    : 0.0000
```

Perbandingan ketiga approach

In [50]:
```
results = pd.DataFrame({
    'Method': ['Statistical (Elliptic)', 'Density (LOF)', 'Proximity (k-NN)'],
    'Accuracy': [acc_stat, acc_lof, acc_knn],
    'Precision': [prec_stat, prec_lof, prec_knn],
    'Recall': [rec_stat, rec_lof, rec_knn],
    'Fit Time (s)': [fit_time_stat, fit_time_lof, fit_time_knn],
    'Pred Time (s)': [pred_time_stat, pred_time_lof, pred_time_knn]
})

# Tampilkan tabel
results.sort_values(by='Recall', ascending=False)
```

Out[50]:

	Method	Accuracy	Precision	Recall	Fit Time (s)	Pred Time (s)
2	Proximity (k-NN)	0.994651	0.074380	0.182432	0.014985	0.000353
0	Statistical (Elliptic)	0.997039	0.156863	0.162162	22.153705	0.050809
1	Density (LOF)	0.996664	0.000000	0.000000	41.395913	14.237450

BAB 5 — Analisis Hasil

Berdasarkan hasil evaluasi tiga pendekatan anomaly detection pada dataset Credit Card Fraud Detection, terlihat bahwa setiap metode menunjukkan karakteristik performa yang berbeda, terutama pada recall, precision, dan waktu komputasi.

5.1 Analisis Performa

1. **Proximity Approach (k-NN)**
   Metode k-NN menghasilkan *recall* tertinggi (0.182432).
   Ini berarti lebih banyak transaksi fraud yang berhasil tertangkap dibandingkan dua metode lainnya.
   Hal ini selaras dengan sifat k-NN yang sensitif terhadap titik yang jauh dari tetangganya.
   Namun, precision-nya rendah (0.074380), menunjukkan banyak false positive.
   Dari sisi waktu, metode ini sangat efisien dengan waktu prediksi tercepat (0.000353 detik).

2. **Statistical Approach (Elliptic Envelope)**
   Metode ini menempati posisi tengah dengan recall sebesar 0.162162 dan precision lebih tinggi (0.156863) dibanding k-NN. Ini menunjukkan pendekatan statistik lebih selektif dalam menandai anomaly.
   Fit time cukup tinggi (22.153 seconds) karena perhitungan mean dan covariance multidimensi.

3. **Density Approach (LOF)**
   LOF memberikan performa paling rendah, dengan recall dan precision sama-sama 0.0.
   Penyebabnya adalah LOF mengandalkan variasi kepadatan lokal, sementara dataset Credit Card Fraud sudah melalui PCA sehingga distribusi fitur menjadi cukup homogen. LOF tidak menemukan perbedaan jelas antara titik fraud dan non-fraud.
   Selain itu, LOF juga memiliki waktu komputasi terlama.

5.2 Pola Perbedaan Metode

- **k-NN unggul recall**
  Cocok untuk tujuan "menangkap sebanyak mungkin fraud", meskipun banyak false alarm.

- **Elliptic Envelope lebih stabil**
  Precision lebih tinggi karena pendekatan statistik lebih konservatif dan tidak terlalu agresif.

- **LOF tidak cocok untuk dataset PCA**
  Density-based methods bekerja baik pada data dengan kepadatan yang bervariasi, bukan data high-dimensional yang sudah ditransformasi.

Secara keseluruhan, perilaku dataset menunjukkan bahwa metode proximity dan statistik lebih sesuai dibanding metode density untuk kasus anomaly detection pada fraud.

BAB 6 — Kesimpulan

Dari hasil eksperimen tiga pendekatan anomaly detection pada dataset Credit Card Fraud Detection, dapat disimpulkan hal-hal sebagai berikut:

1. **Proximity-based k-NN merupakan metode dengan recall terbaik**, sehingga paling efektif jika tujuan utama adalah menangkap sebanyak mungkin transaksi fraud.
2. **Pendekatan Statistical (Elliptic Envelope) memberikan keseimbangan precision dan recall yang lebih baik**, cocok jika ingin mengurangi false positive.
3. **Density-based LOF tidak cocok**, terutama karena fitur PCA menyebabkan kepadatan antar titik menjadi relatif seragam, sehingga LOF gagal membedakan fraud dan non-fraud.
4. Pemilihan metode anomaly detection harus mempertimbangkan bentuk distribusi data, transformasi fitur, serta tujuan analisis (mau memaksimalkan recall atau precision).

Dengan demikian, untuk kasus fraud detection pada dataset ini:

- **k-NN adalah pilihan terbaik untuk mendeteksi sebanyak mungkin fraud (high recall)**
- **Elliptic Envelope adalah pilihan terbaik untuk prediksi yang lebih stabil (higher precision)**

BAB 7 — Penutup

Tugas ini memberikan pemahaman mengenai konsep dasar anomaly detection serta perbedaan pendekatan yang digunakan untuk mendeteksi titik yang tidak mengikuti pola umum. Melalui implementasi tiga metode—statistical, proximity, dan density—terlihat bahwa karakteristik dataset sangat mempengaruhi performa model.

Dataset Credit Card Fraud Detection menunjukkan bahwa metode proximity dan statistical lebih efektif dibanding density-based, terutama karena transformasi PCA yang mengurangi variasi kepadatan lokal.

Diharapkan laporan ini dapat menjadi referensi dalam memahami anomaly detection dan bagaimana memilih metode yang sesuai berdasarkan struktur data dan tujuan analisis.