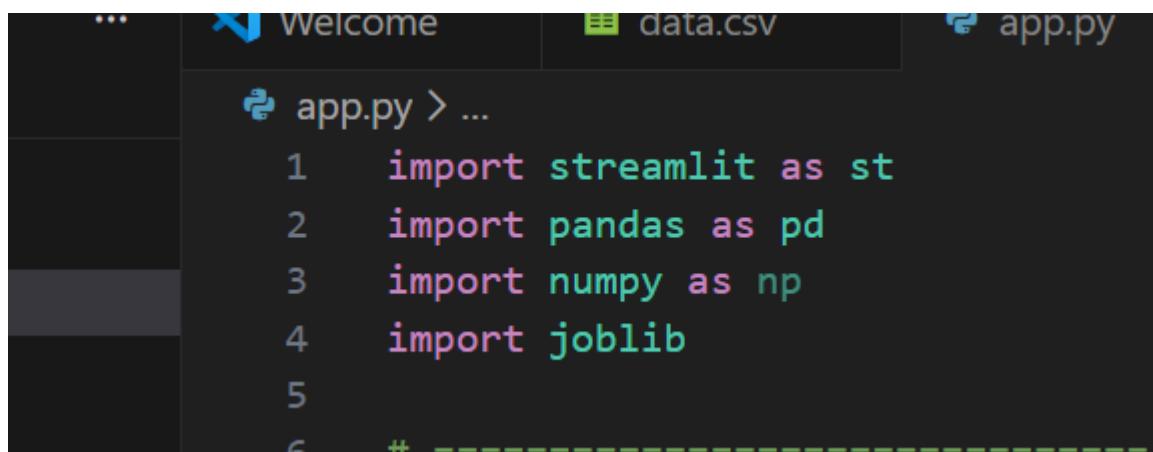


Deployment

1.Import Library

Pada bagian awal program, aplikasi memuat beberapa pustaka Python yang diperlukan untuk mendukung proses deployment. Pustaka Streamlit digunakan untuk membangun antarmuka web interaktif, sedangkan Pandas dan NumPy digunakan untuk pengolahan data dalam bentuk tabel dan numerik. Selain itu, pustaka Joblib digunakan untuk memuat model machine learning serta objek preprocessing yang sebelumnya telah dilatih dan disimpan pada tahap pemodelan.



```
... ➜ Welcome data.csv app.py
app.py > ...
1 import streamlit as st
2 import pandas as pd
3 import numpy as np
4 import joblib
5
6 # -----
```

2.Konfigurasi Halaman

Pada tahap ini dilakukan pengaturan konfigurasi halaman aplikasi menggunakan fungsi bawaan streamlit. Konfigurasi tersebut meliputi penentuan judul halaman, ikon yang ditampilkan pada tab browser, serta pengaturan tata letak halaman agar menggunakan mode tampilan lebar.

```
st.set_page_config(
    page_title="Prediksi Kebangkrutan Perusahaan",
    page_icon="📊",
    layout="wide"
)
```

3.Load Model & Dataset

Pada tahap ini, aplikasi memuat seluruh komponen utama yang diperlukan untuk menjalankan proses prediksi kebangkrutan perusahaan. Model Decision Tree yang telah dilatih pada tahap pemodelan sebelumnya dimuat kembali ke dalam sistem menggunakan pustaka Joblib. Model ini berperan sebagai inti dari sistem prediksi yang bertugas untuk mengklasifikasikan status kebangkrutan perusahaan berdasarkan indikator keuangan yang diberikan.

Selain model, aplikasi juga memuat objek StandardScaler yang digunakan untuk melakukan normalisasi data input. Proses normalisasi ini penting untuk memastikan bahwa skala data yang digunakan pada tahap prediksi sesuai dengan skala data yang digunakan pada saat pelatihan model, sehingga hasil prediksi yang dihasilkan tetap konsisten dan valid.

Selanjutnya, dataset laporan keuangan perusahaan dimuat ke dalam aplikasi untuk mendukung proses pengolahan data. Dataset ini berisi sejumlah indikator keuangan serta satu atribut target yang menunjukkan status kebangkrutan perusahaan. Setelah dataset berhasil dimuat, atribut target dipisahkan dari fitur-fitur keuangan agar fitur tersebut dapat digunakan sebagai input pada proses prediksi tanpa memengaruhi label klasifikasi. Dataset juga dimanfaatkan untuk memperoleh nilai rata-rata setiap indikator keuangan serta untuk mengambil contoh data perusahaan bangkrut dan tidak bangkrut yang digunakan pada skenario demonstrasi aplikasi.

Dengan memuat model, scaler, dan dataset secara bersamaan, aplikasi memastikan bahwa seluruh komponen yang digunakan pada tahap deployment memiliki keterkaitan langsung dengan proses pemodelan yang telah dilakukan sebelumnya.

```
# =====
model = joblib.load("decision_tree_bankruptcy.joblib")
scaler = joblib.load("scaler.joblib")
```

```
df = pd.read_csv("data.csv")
X = df.drop("Bankrupt?", axis=1)
```

```
feature_names = X.columns.tolist()
mean_values = X.mean()
```

5.Contoh data dari dataset

Pada tahap pengambilan contoh data, aplikasi memilih beberapa data perusahaan secara langsung dari dataset laporan keuangan yang digunakan pada proses pelatihan model. Pengambilan contoh data ini dilakukan dengan memanfaatkan atribut target yang menunjukkan status kebangkrutan perusahaan, sehingga sistem dapat membedakan antara perusahaan yang berisiko bangkrut dan perusahaan yang tidak bangkrut.

Aplikasi mengambil satu contoh perusahaan yang memiliki label bangkrut serta satu contoh perusahaan yang memiliki label tidak bangkrut. Kedua contoh data tersebut digunakan sebagai skenario pengujian pada tahap demonstrasi aplikasi. Dengan menggunakan data historis yang berasal langsung dari dataset, contoh data yang ditampilkan bersifat valid dan representatif terhadap kondisi keuangan perusahaan yang sebenarnya.

Penggunaan contoh data ini bertujuan untuk mempermudah proses demonstrasi dan pengujian sistem, terutama pada saat presentasi atau evaluasi aplikasi. Selain itu, penampilan nilai indikator keuangan dari contoh perusahaan dalam bentuk tabel memberikan transparansi kepada pengguna mengenai data yang digunakan dalam proses prediksi. Pendekatan ini memungkinkan pengguna untuk memahami perbedaan karakteristik keuangan antara perusahaan yang berisiko bangkrut dan perusahaan yang tidak bangkrut secara lebih jelas.

Dengan demikian, pengambilan contoh data dari dataset tidak hanya berfungsi sebagai alat bantu demonstrasi, tetapi juga sebagai sarana pembelajaran bagi pengguna dalam memahami pola kebangkrutan perusahaan berdasarkan indikator keuangan.

```
# Contoh data dari dataset
contoh_bangkrut = X[df["Bankrupt?"] == 1].iloc[0]
contoh_tidak_bangkrut = X[df["Bankrupt?"] == 0].iloc[0]
```

6.Pilih Mode Input Data

Memberi 3 pilihan cara input data

Memudahkan demo tanpa input manual

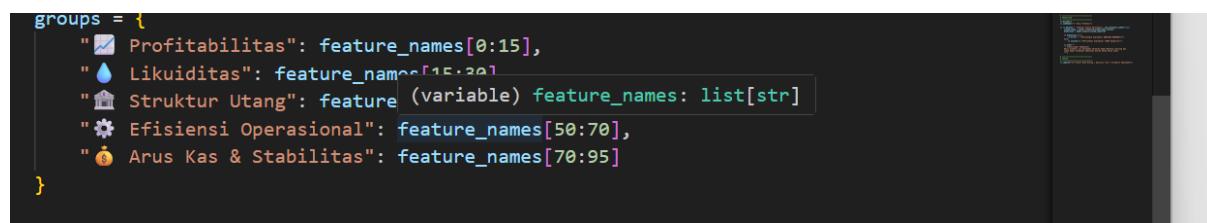
```
mode = st.radio(
    "Pilih skenario pengujian:",
    (
        "⚡ Contoh Perusahaan Berisiko Bangkrut",
        "📈 Contoh Perusahaan Tidak Bangkrut",
        "✍️ Input Manual (Lengkap 95 Indikator)"
    )
)
```

7. Pegelompokan Fitur

Pada tahap persiapan fitur, aplikasi melakukan identifikasi terhadap seluruh indikator keuangan yang digunakan sebagai variabel input dalam proses prediksi kebangkrutan perusahaan. Indikator-indikator tersebut diperoleh dari dataset laporan keuangan perusahaan dan disusun dalam bentuk daftar fitur yang akan digunakan oleh sistem pada tahap deployment. Penyusunan daftar fitur ini bertujuan untuk memastikan bahwa urutan dan jumlah fitur yang digunakan pada tahap prediksi sesuai dengan fitur yang digunakan pada saat pelatihan model.

Selain itu, aplikasi menghitung nilai rata-rata dari setiap indikator keuangan yang terdapat dalam dataset. Nilai rata-rata ini digunakan sebagai nilai awal atau nilai default pada mode input manual. Penggunaan nilai rata-rata bertujuan untuk membantu pengguna dalam mengisi data input, terutama apabila pengguna tidak memiliki seluruh informasi keuangan perusahaan secara lengkap. Dengan demikian, pengguna tetap dapat melakukan pengujian sistem tanpa harus mengisi seluruh indikator secara manual.

Persiapan fitur dan nilai rata-rata ini juga berperan penting dalam menjaga konsistensi antara data pelatihan dan data prediksi. Dengan menggunakan struktur fitur dan nilai acuan yang berasal dari dataset yang sama, aplikasi dapat meminimalkan potensi kesalahan input serta memastikan bahwa data yang diproses oleh model memiliki karakteristik yang sesuai dengan data yang digunakan pada tahap pemodelan.



```
groups = [
    "📈 Profitabilitas": feature_names[0:15],
    "💧 Likuiditas": feature_names[15:30],
    "🏡 Struktur Utang": feature_names[(variable) feature_names: list[str]],
    "⚙️ Efisiensi Operasional": feature_names[50:70],
    "💰 Arus Kas & Stabilitas": feature_names[70:95]
]
```

A screenshot of a Jupyter Notebook cell showing Python code. The code defines a dictionary named 'groups' containing five entries. Each entry consists of a string icon followed by a colon, then the variable name 'feature_names' with a slice indicating the range of indices for that group. The first group is 'Profitabilitas' from index 0 to 15. The second is 'Likuiditas' from index 15 to 30. The third is 'Struktur Utang' with a complex key involving 'variable' and 'list[str]'. The fourth is 'Efisiensi Operasional' from index 50 to 70. The fifth is 'Arus Kas & Stabilitas' from index 70 to 95. The code is highlighted in blue and green, and the background of the code cell is dark grey.

8. Logika Input data

Logika input data pada aplikasi prediksi kebangkrutan perusahaan dirancang untuk memberikan kemudahan serta fleksibilitas kepada pengguna dalam melakukan pengujian sistem. Aplikasi menyediakan tiga skenario input data yang dapat dipilih sesuai kebutuhan pengguna, yaitu contoh perusahaan berisiko bangkrut, contoh perusahaan tidak bangkrut, dan input manual indikator keuangan.

Pada skenario contoh perusahaan berisiko bangkrut, sistem secara otomatis mengambil satu data perusahaan yang memiliki label kebangkrutan dari dataset asli. Data tersebut kemudian digunakan sebagai input prediksi tanpa dapat diubah oleh pengguna. Seluruh indikator keuangan pada skenario ini ditampilkan dalam bentuk tabel agar pengguna dapat melihat nilai-nilai rasio keuangan yang menyebabkan perusahaan tersebut diklasifikasikan sebagai berisiko bangkrut. Pendekatan ini bertujuan untuk mempercepat proses demonstrasi aplikasi serta meningkatkan transparansi sistem.

Skenario contoh perusahaan tidak bangkrut memiliki mekanisme yang serupa, di mana sistem mengambil data perusahaan dengan kondisi keuangan yang sehat dari dataset. Data ini juga ditampilkan dalam bentuk tabel indikator keuangan dan digunakan langsung sebagai input prediksi. Dengan adanya dua skenario contoh ini, pengguna dapat membandingkan karakteristik keuangan perusahaan yang berisiko bangkrut dan perusahaan yang tidak bangkrut secara langsung.

Selain itu, aplikasi juga menyediakan skenario input manual indikator keuangan yang memungkinkan pengguna untuk memasukkan nilai rasio keuangan secara mandiri. Pada skenario ini, sistem menginisialisasi setiap indikator keuangan dengan nilai rata-rata dataset sebagai nilai awal. Hal ini dilakukan untuk menghindari kesalahan input serta membantu pengguna yang tidak memiliki seluruh data keuangan perusahaan. Pengguna dapat mengubah nilai indikator sesuai kebutuhan, sedangkan indikator yang tidak diubah akan tetap menggunakan nilai default.

Dengan adanya ketiga skenario input data tersebut, aplikasi mampu mengakomodasi berbagai kebutuhan pengguna, baik untuk keperluan demonstrasi, pembelajaran, maupun pengujian sistem secara lebih mendalam. Pendekatan ini juga memastikan konsistensi antara data yang digunakan dalam pelatihan model dan data yang digunakan pada tahap prediksi.

```
88  """
89  if mode == "🔴 Contoh Perusahaan Berisiko Bangkrut":
90      input_data = contoh_bangkrut.copy()
91
92      st.warning("Menggunakan contoh perusahaan dengan kondisi keuangan BURUK dari dataset.")
93
94      for group_name, features in groups.items():
95          with st.expander(group_name):
96              st.dataframe(
97                  input_data[features].to_frame(name="Nilai Indikator"),
98                  use_container_width=True
99              )
100
101 elif mode == "🟢 Contoh Perusahaan Tidak Bangkrut":
102     input_data = contoh_tidak_bangkrut.copy()
103
104     st.success("Menggunakan contoh perusahaan dengan kondisi keuangan SEHAT dari dataset.")
105
106     for group_name, features in groups.items():
107         with st.expander(group_name):
108             st.dataframe(
109                 input_data[features].to_frame(name="Nilai Indikator"),
110                 use_container_width=True
111             )
```

```

else:
    input_data = mean_values.copy()

    st.info("Silakan masukkan indikator keuangan atau gunakan nilai default.")
        (variable) groups: dict[str, list[str]]]

    for group_name, features in groups.items():
        with st.expander(group_name):
            cols = st.columns(3)
            for i, feature in enumerate(features):
                with cols[i % 3]:
                    input_data[feature] = st.number_input(
                        label=feature,
                        value=float(mean_values[feature]),
                        help="Rasio keuangan berdasarkan laporan perusahaan"
                    )

```

9. Proses Normalisasi dan Prediksi

Setelah data input ditentukan, sistem melakukan proses normalisasi menggunakan StandardScaler untuk menghasilkan kesesuaian skala data. Data yang telah dinormalisasikan kemudian diproses oleh model Decision Tree untuk menghasilkan prediksi status kebangkrutan perusahaan

```

if st.button("⚡️ Prediksi Status Perusahaan", use_container_width=True):
    scaled_input = scaler.transform([input_data.values])
    prediction = model.predict(scaled_input)[0]

```

10. Tampilan Hasil Prediksi

Hasil prediksi ditampilkan secara visual pada antarmuka web. Jika perusahaan diprediksi berisiko bangkrut, sistem akan menampilkan pesan peringatan, sedangkan jika perusahaan diprediksi tidak bangkrut, sistem akan menampilkan pesan keberhasilan. Ini bertujuan untuk memudahkan pengguna dalam memahami hasil prediksi.

```

if prediction == 1:
    st.error("⚠️ **Perusahaan diprediksi BERISIKO BANGKRUT**")
else:
    st.success("✅ **Perusahaan diprediksi TIDAK bangkrut**")

st.info("""
⚠️ **Disclaimer Akademik**
Hasil prediksi ini merupakan keluaran model Machine Learning dan
tidak dapat dijadikan keputusan mutlak dalam dunia nyata.
""")
```

11. Output

Prediksi Kebangkrutian Perusahaan

localhost:8501

Deploy

Informasi Aplikasi

- Model : Decision Tree
- Fitur : 95 indikator keuangan
- Dataset : Laporan keuangan perusahaan
- Tujuan : Prediksi risiko kebangkrutian



PREDIKSI KEBANGKRUTIAN PERUSAHAAN

Aplikasi ini memprediksi potensi kebangkrutian perusahaan berdasarkan 95 indikator keuangan menggunakan Decision Tree.

Nilai contoh diambil langsung dari dataset untuk mempermudah demonstrasi.

Mode Input Data

Pilih skenario pengujian:

- Contoh Perusahaan Berisiko Bangkrut
- Contoh Perusahaan Tidak Bangkrut
- Input Manual (Lengkap 95 Indikator)

Menggunakan contoh perusahaan dengan kondisi keuangan BURUK dari dataset.

Profitabilitas

Nilai Indikator	
ROA(C) before interest and depreciation before interest	0.3706
ROA(A) before interest and % after tax	0.4244
ROA(B) before interest and depreciation after tax	0.4057
Operating Gross Margin	0.6015
Realized Sales Gross Margin	0.6015
Operating Profit Rate	0.999
Pre-tax net Interest Rate	0.7369
After-tax net Interest Rate	0.8088
Non-industry income and expenditure/revenue	0.3026
Continuous interest rate (after tax)	0.781

Likuiditas

Informasi Aplikasi

- Model : Decision Tree
- Fitur : 95 indikator keuangan
- Dataset : Laporan keuangan perusahaan
- Tujuan : Prediksi risiko kebangkrutan

Likuiditas

	Nilai Indikator
Net Value Per Share (B)	0.1479
Net Value Per Share (A)	0.1479
Net Value Per Share (C)	0.1479
Persistent EPS in the Last Four Seasons	0.1691
Cash Flow Per Share	0.3117
Revenue Per Share (Yuan RMB)	0.0176
Operating Profit Per Share (Yuan RMB)	0.0959
Per Share Net profit before tax (Yuan RMB)	0.1387
Realized Sales Gross Profit Growth Rate	0.0221
Operating Profit Growth Rate	0.8482

Struktur Utang

	Nilai Indikator
Total Asset Return Growth Rate Ratio	0.2631
Cash Reinvestment %	0.3637
Current Ratio	0.0023
Quick Ratio	0.0012
Interest Expense Ratio	0.63
Total debt/Total net worth	0.0213
Debt ratio %	0.2076
Net worth/Assets	0.7924
Long-term fund suitability ratio (%)	0.005
Borrowing dependency	0.3903

Efisiensi Operasional

	Nilai Indikator
Operating Funds to Liability	0.334
Inventory/Working Capital	0.2769
Inventory/Current Liability	0.001
Current Liabilities/Ability	0.6763
Working Capital/Equity	0.7213
Current Liabilities/Equity	0.3381
Long-term Liability to Current Assets	0.0256
Retained Earnings to Total Assets	0.9532
Total Income/Total expense	0.002
Total expense/Assets	0.0649

Arus Kas & Stabilitas

	Nilai Indikator
Net Income to Total Assets	0.7168
Total assets to GNP price	0.0092
No-credit Interval	0.6229
Gross Profit to Sales	0.6015
Net Income to Stockholder's Equity	0.8279
Liability to Equity	0.2902
Degree of Financial Leverage (DFL)	0.0266
Interest Coverage Ratio (Interest)	0.5641
Net Income Flag	1
Equity to Liability	0.0165

Informasi Aplikasi

- Model : Decision Tree
- Fitur : 95 indikator keuangan
- Dataset : Laporan keuangan perusahaan
- Tujuan : Prediksi risiko kebangkrutan

Arus Kas & Stabilitas

	Nilai Indikator
Net Income to Total Assets	0.7168
Total assets to GNP price	0.0092
No-credit Interval	0.6229
Gross Profit to Sales	0.6015
Net Income to Stockholder's Equity	0.8279
Liability to Equity	0.2902
Degree of Financial Leverage (DFL)	0.0266
Interest Coverage Ratio (Interest)	0.5641
Net Income Flag	1
Equity to Liability	0.0165

Prediksi Kebangkrut Perusahaan

localhost:8501

Informasi Aplikasi

- Model : Decision Tree
- Fitur: 95 indikator keuangan
- Dataset : Laporan keuangan perusahaan
- Tujuan : Prediksi risiko kebangkrutan

	Nilai Indikator
Liability to Equity	0.2902
Degree of Financial Leverage (DF)	0.0266
Interest Coverage Ratio (Interest)	0.5641
Net Income Flag	1
Equity to Liability	0.0165

Hasil Prediksi

Prediksi Status Perusahaan

Perusahaan diprediksi BERISIKO BANGKRUT

Disclaimer Akademik

Hasil prediksi ini merupakan keluaran model Machine Learning dan tidak dapat dijadikan keputusan mutlak dalam dunia nyata.

Risiko Bangkrut

localhost:8501

Mode Input Data

Pilih skenario pengujian:

- Contoh Perusahaan Berisiko Bangkrut
- Contoh Perusahaan Tidak Bangkrut
- Input Manual (Lengkap 95 Indikator)

Menggunakan contoh perusahaan dengan kondisi keuangan SEHAT dari dataset.

Profitabilitas

	Nilai Indikator
Operating Profit Rate	0.999
Pre-tax net Interest Rate	0.797
After-tax net Interest Rate	0.809
Non-industry income and expenditure/revenue	0.3028
Continuous interest rate (after tax)	0.7812
Operating Expense Rate	0.0004
Research and development expense rate	730000000

Likuiditas

	Nilai Indikator
Revenue Per Share (Yuan ¥)	0.0309
Operating Profit Per Share (Yuan ¥)	0.0999
Per Share Net profit before tax (Yuan ¥)	0.1289
Realized Sales Gross Profit Growth Rate	0.0221
Operating Profit Growth Rate	0.8483
After-tax Net Profit Growth Rate	0.6886
Regular Net Profit Growth Rate	0.6886
Continuous Net Profit Growth Rate	0.2176
Total Asset Growth Rate	5720000000
Net Value Growth Rate	0.0004

Struktur Utang

	Nilai Indikator
Total Asset Return Growth Rate	0.263
Cash Reinvestment %	0.3753
Current Ratio	0.0047
Quick Ratio	0.0029
Interest Expense Ratio	0.63
Total debt/Total net worth	0.0244
Debt ratio %	0.2161
Net worth/Assets	0.7839
Long-term fund suitability ratio (A)	0.005
Borrowing dependency	0.385

Informasi Aplikasi

Model : Decision Tree
Fitur : 95 indikator keuangan
Dataset : Laporan keuangan perusahaan
Tujuan : Prediksi risiko kebangkrutan

Efisiensi Operasional	
	Nilai Indikator
Operating Funds to Liability	0.3401
Inventory/Working Capital	0.2766
Inventory/Current Liability	0.0022
Current Liabilities/Liability	0.5081
Working Capital/Equity	0.7287
Current Liabilities/Equity	0.3374
Long-term Liability to Current Assets	0.0093
Retained Earnings to Total Asset	0.9114
Total Income/Total expense	0.0021
Total expense/Assets	0.0928

Arus Kas & Stabilitas

Arus Kas & Stabilitas	
	Nilai Indikator
Current Asset Turnover Rate	0.0001
Quick Asset Turnover Rate	9560000000
Working capital Turnover Rate	0.5939
Cash Turnover Rate	2390000000
Cash Flow to Sales	0.6716
Fixed Assets to Assets	0.3559
Current Liability to Liability	0.5081
Current Liability to Equity	0.3374
Equity to Long term Liability	0.1206
Cash Flow to Total Assets	0.6304

localhost:8501

Informasi Aplikasi

- Model : Decision Tree
- Fitur : 95 indikator keuangan
- Dataset : Laporan keuangan perusahaan
- Tujuan : Prediksi risiko kebangkrutan

Hasil Prediksi

Prediksi Status Perusahaan

Perusahaan diprediksi TIDAK bangkrut

Disclaimer Akademik
Hasil prediksi ini merupakan keluaran model Machine Learning dan tidak dapat dijadikan keputusan mutlak dalam dunia nyata.

Project Data Mining | Decision Tree + Streamlit Deployment

Aplikasi

Decision Tree
indikator keuangan
Laporan keuangan
Prediksi risiko

Profitabilitas

Silakan masukkan indikator keuangan atau gunakan nilai default.

ROA(C) before interest and depreciation before interest	ROA(A) before interest and % after tax	ROA(B) before interest and depreciation after tax
0,51	0,56	0,55
Operating Gross Margin	Realized Sales Gross Margin	Operating Profit Rate
0,61	0,61	1,00
Pre-tax net Interest Rate	After-tax net Interest Rate	Non-industry income and expenditure/revenue
0,80	0,81	0,30
Continuous interest rate (after tax)	Operating Expense Rate	Research and development expense rate
0,78	1995347312,80	

```

import pandas as pd
import numpy as np

from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.preprocessing import StandardScaler
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score, confusion_matrix,
classification_report

import joblib

df = pd.read_csv("data.csv")
df.head()

Bankrupt?    ROA(C) before interest and depreciation before interest \
0           1                           0.370594
1           1                           0.464291
2           1                           0.426071
3           1                           0.399844
4           1                           0.465022

ROA(A) before interest and % after tax \
0                         0.424389
1                         0.538214
2                         0.499019
3                         0.451265
4                         0.538432

ROA(B) before interest and depreciation after tax \
0                         0.405750
1                         0.516730
2                         0.472295
3                         0.457733
4                         0.522298

Operating Gross Margin    Realized Sales Gross Margin \
0                     0.601457                      0.601457
1                     0.610235                      0.610235
2                     0.601450                      0.601364
3                     0.583541                      0.583541
4                     0.598783                      0.598783

Operating Profit Rate    Pre-tax net Interest Rate \
0                     0.998969                      0.796887
1                     0.998946                      0.797380

```

2	0.998857	0.796403
3	0.998700	0.796967
4	0.998973	0.797366
After-tax net Interest Rate \ Non-industry income and expenditure/revenue		
0	0.808809	
0.302646		
1	0.809301	
0.303556		
2	0.808388	
0.302035		
3	0.808966	
0.303350		
4	0.809304	
0.303475		
... Net Income to Total Assets Total assets to GNP price \		
0 ...	0.716845	0.009219
1 ...	0.795297	0.008323
2 ...	0.774670	0.040003
3 ...	0.739555	0.003252
4 ...	0.795016	0.003878
No-credit Interval Gross Profit to Sales \		
0	0.622879	0.601453
1	0.623652	0.610237
2	0.623841	0.601449
3	0.622929	0.583538
4	0.623521	0.598782
Net Income to Stockholder's Equity Liability to Equity \		
0	0.827890	0.290202
1	0.839969	0.283846
2	0.836774	0.290189
3	0.834697	0.281721
4	0.839973	0.278514
Degree of Financial Leverage (DFL) \		
0	0.026601	
1	0.264577	
2	0.026555	
3	0.026697	
4	0.024752	
Interest Coverage Ratio (Interest expense to EBIT) Net Income Flag \		
0		0.564050
1		
1		0.570175

```

1                               0.563706
2                               0.564663
1                               0.575617
1

    Equity to Liability
0      0.016469
1      0.020794
2      0.016474
3      0.023982
4      0.035490

[5 rows x 96 columns]

x = df.drop("Bankrupt?", axis=1)
y = df["Bankrupt?"]

scaler = StandardScaler()
x_scaled = scaler.fit_transform(x)

x_train, x_test, y_train, y_test = train_test_split(
    x_scaled, y, test_size=0.2, random_state=42
)

model = DecisionTreeClassifier(random_state=42)
model.fit(x_train, y_train)

DecisionTreeClassifier(random_state=42)

y_pred = model.predict(x_test)

print("Accuracy:", accuracy_score(y_test, y_pred))
print(confusion_matrix(y_test, y_pred))
print(classification_report(y_test, y_pred))

Accuracy: 0.9567448680351907
[[1282   31]
 [ 28   23]]
             precision    recall  f1-score   support
          0       0.98      0.98      0.98     1313
          1       0.43      0.45      0.44      51

   accuracy         0.96      0.96      0.96     1364
  macro avg        0.70      0.71      0.71     1364
weighted avg       0.96      0.96      0.96     1364

```

```
joblib.dump(model, "decision_tree_bankruptcy.joblib")
joblib.dump(scaler, "scaler.joblib")

['scaler.joblib']
```