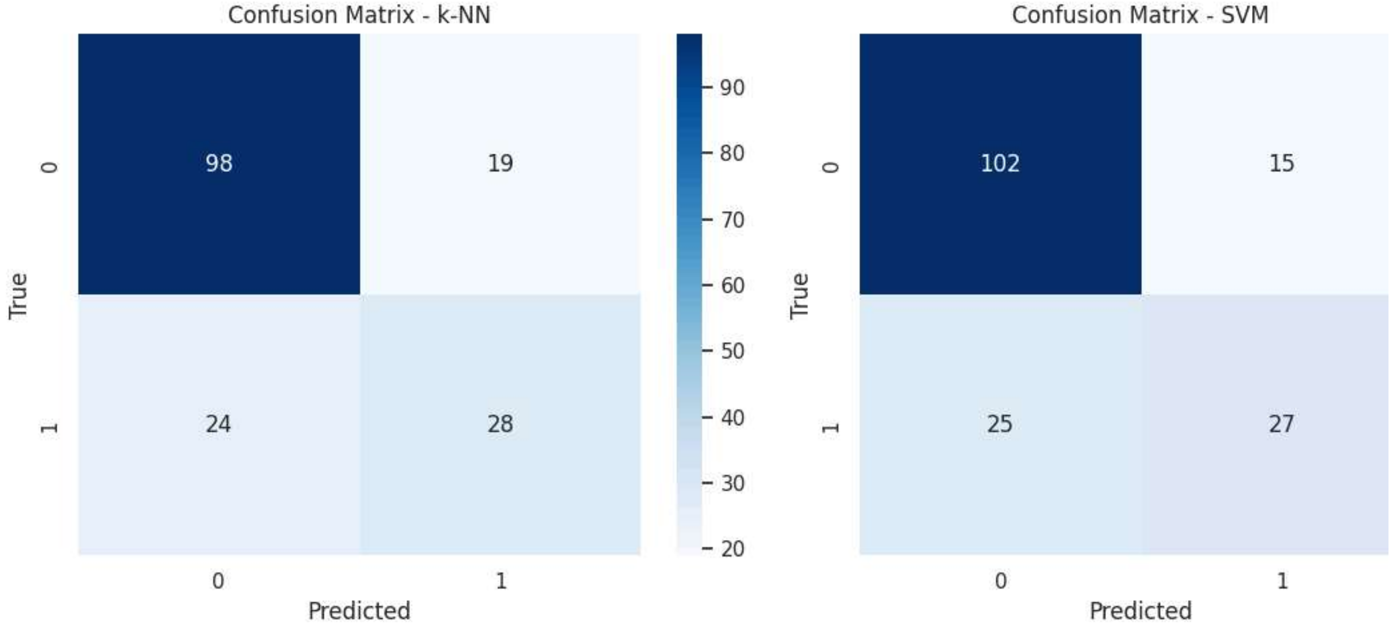
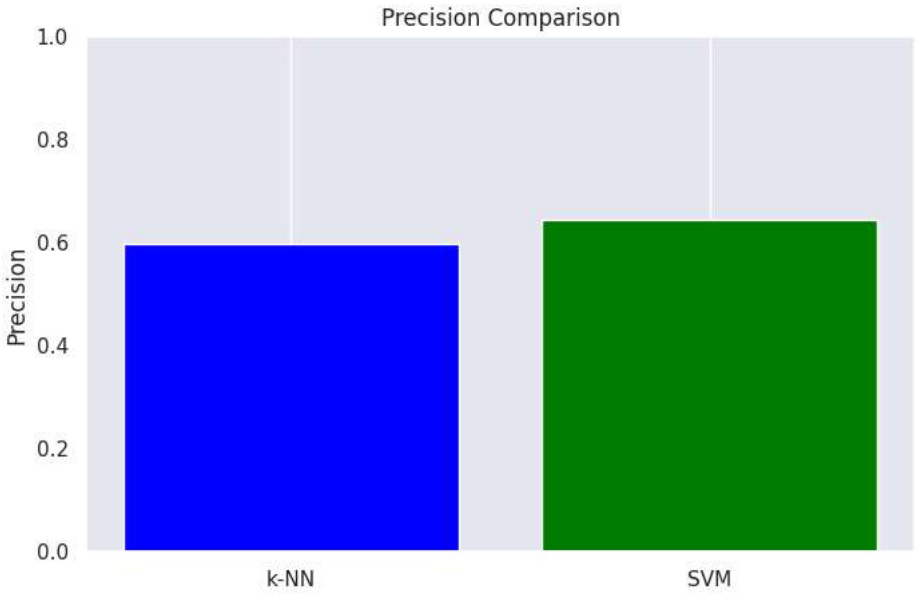
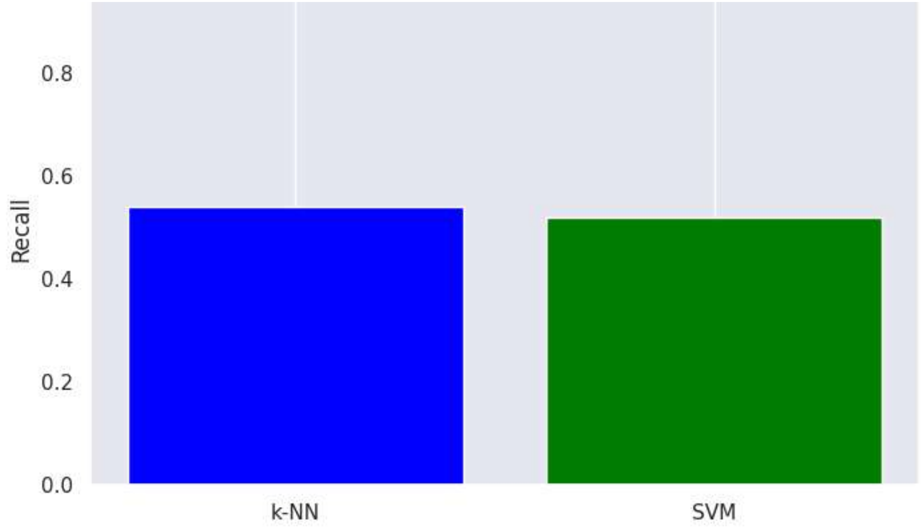
```
# Import necessary libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.preprocessing import StandardScaler
from sklearn.neighbors import KNeighborsClassifier
from sklearn.svm import SVC
from sklearn.metrics import classification_report, confusion_matrix, precision_score, recall_score, f1_score
import joblib

# Set style for plots
sns.set()

# Load the dataset
data = pd.read_csv('/content/diabetes prediction dataset.csv')

# Handle zero values (treat them as missing)
cols_with_zero = ['BMI', 'BloodPressure', 'Glucose', 'Insulin', 'SkinThickness']
for col in cols_with_zero:
    data[col] = data[col].replace(0, np.nan)
    data[col].fillna(data[col].mean(), inplace=True)

# Outlier removal
def remove_outliers(df, column, quantile):
    upper_limit = df[column].quantile(quantile)
    return df[df[column] < upper_limit]

data = remove_outliers(data, 'Pregnancies', 0.98)
data = remove_outliers(data, 'BMI', 0.99)
data = remove_outliers(data, 'SkinThickness', 0.99)
data = remove_outliers(data, 'Insulin', 0.95)
data = remove_outliers(data, 'DiabetesPedigreeFunction', 0.99)
data = remove_outliers(data, 'Age', 0.99)

# Features and labels
X = data.drop(columns=['Outcome'])
y = data['Outcome']

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25, random_state=0)

# Standardize features
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

# Initialize models
knn_model = KNeighborsClassifier()
svm_model = SVC(probability=True)

# Cross-validation
knn_cv = cross_val_score(knn_model, X_train, y_train, cv=5, scoring='accuracy')
svm_cv = cross_val_score(svm_model, X_train, y_train, cv=5, scoring='accuracy')

# Fit models
knn_model.fit(X_train, y_train)
svm_model.fit(X_train, y_train)

# Predictions
knn_pred = knn_model.predict(X_test)
svm_pred = svm_model.predict(X_test)

# Reports
print("k-NN Classification Report:\n", classification_report(y_test, knn_pred))
print("SVM Classification Report:\n", classification_report(y_test, svm_pred))

# Confusion matrices
fig, axes = plt.subplots(1, 2, figsize=(12, 5))
sns.heatmap(confusion_matrix(y_test, knn_pred), annot=True, fmt='g', cmap='Blues', ax=axes[0])
axes[0].set_title('Confusion Matrix - k-NN')
axes[0].set_xlabel('Predicted'); axes[0].set_ylabel('True')

sns.heatmap(confusion_matrix(y_test, svm_pred), annot=True, fmt='g', cmap='Blues', ax=axes[1])
axes[1].set_title('Confusion Matrix - SVM')
```

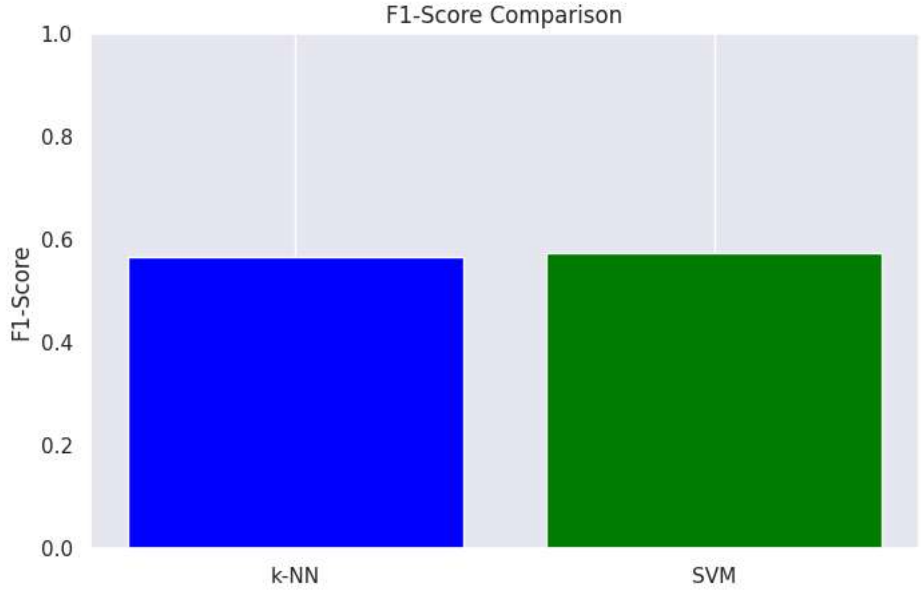
```
    axes[1].set_xlabel('Predicted'); axes[1].set_ylabel('True')

    plt.tight_layout()
    plt.show()

    # Evaluation metrics for plotting
    models = {'k-NN': knn_pred, 'SVM': svm_pred}
    precision_scores = {}
    recall_scores = {}
    f1_scores = {}

    for name, pred in models.items():
        precision_scores[name] = precision_score(y_test, pred)
        recall_scores[name] = recall_score(y_test, pred)
        f1_scores[name] = f1_score(y_test, pred)

    # Plotting metrics
    def plot_metric(metric_dict, metric_name):
        plt.figure(figsize=(8, 5))
        plt.bar(metric_dict.keys(), metric_dict.values(), color=['blue', 'green'])
        plt.title(f'{metric_name} Comparison')
        plt.ylabel(metric_name)
        plt.ylim(0, 1)
        plt.grid(axis='y')
        plt.show()

    plot_metric(precision_scores, "Precision")
    plot_metric(recall_scores, "Recall")
    plot_metric(f1_scores, "F1-Score")

    # Save models as .pkl
    joblib.dump(knn_model, "knn_model.pkl")
    joblib.dump(svm_model, "svm_model.pkl")
    print("✅ Models saved as knn_model.pkl and svm_model.pkl")
```

```
<ipython-input-1-ec6d4ea4c23d>:23: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignm
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting valu

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].me


  data[col].fillna(data[col].mean(), inplace=True)
k-NN Classification Report:
              precision    recall  f1-score   support

           0       0.80      0.84      0.82       117
           1       0.60      0.54      0.57        52

    accuracy                           0.75       169
   macro avg       0.70      0.69      0.69       169
weighted avg       0.74      0.75      0.74       169

SVM Classification Report:
              precision    recall  f1-score   support

           0       0.80      0.87      0.84       117
           1       0.64      0.52      0.57        52

    accuracy                           0.76       169
   macro avg       0.72      0.70      0.71       169
weighted avg       0.75      0.76      0.76       169
```








F1-Score Comparison

```
✅ Models saved as knn_model.pkl and svm_model.pkl
```

```
!pip install gradio
```

```
Collecting starlette<1.0,>=0.40.0 (from gradio)
  Downloading starlette-0.47.0-py3-none-any.whl.metadata (6.2 kB)
Collecting tomlkit<0.14.0,>=0.12.0 (from gradio)
  Downloading tomlkit-0.13.2-py3-none-any.whl.metadata (2.7 kB)
Requirement already satisfied: typer<1.0,>=0.12 in /usr/local/lib/python3.11/dist-packages (from gradio) (0.15.3)
Requirement already satisfied: typing-extensions~=4.0 in /usr/local/lib/python3.11/dist-packages (from gradio) (4.13.2)
Collecting uvicorn>=0.14.0 (from gradio)
  Downloading uvicorn-0.34.2-py3-none-any.whl.metadata (6.5 kB)
Requirement already satisfied: fsspec in /usr/local/lib/python3.11/dist-packages (from gradio-client==1.10.1->gradio) (2025.3.2)
Requirement already satisfied: websockets<16.0,>=10.0 in /usr/local/lib/python3.11/dist-packages (from gradio-client==1.10.1->gradio)
Requirement already satisfied: idna>=2.8 in /usr/local/lib/python3.11/dist-packages (from anyio<5.0,>=3.0->gradio) (3.10)
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.11/dist-packages (from anyio<5.0,>=3.0->gradio) (1.3.1)
Collecting starlette<1.0,>=0.40.0 (from gradio)
  Downloading starlette-0.46.2-py3-none-any.whl.metadata (6.2 kB)
Requirement already satisfied: certifi in /usr/local/lib/python3.11/dist-packages (from httpx>=0.24.1->gradio) (2025.4.26)
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.11/dist-packages (from httpx>=0.24.1->gradio) (1.0.9)
Requirement already satisfied: h11>=0.16 in /usr/local/lib/python3.11/dist-packages (from httpcore==1.*->httpx>=0.24.1->gradio) (0.16
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.28.1->gradio) (3.18.0)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.28.1->gradio) (2.32.3)
Requirement already satisfied: tqdm>=4.42.1 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.28.1->gradio) (4.67.1
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/dist-packages (from pandas<3.0,>=1.0->gradio) (2.9
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas<3.0,>=1.0->gradio) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas<3.0,>=1.0->gradio) (2025.2)
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/dist-packages (from pydantic<2.12,>=2.0->gradio) (
Requirement already satisfied: pydantic-core==2.33.2 in /usr/local/lib/python3.11/dist-packages (from pydantic<2.12,>=2.0->gradio) (2
Requirement already satisfied: typing-inspection>=0.4.0 in /usr/local/lib/python3.11/dist-packages (from pydantic<2.12,>=2.0->gradio)
Requirement already satisfied: click>=8.0.0 in /usr/local/lib/python3.11/dist-packages (from typer<1.0,>=0.12->gradio) (8.2.1)
Requirement already satisfied: shellingham>=1.3.0 in /usr/local/lib/python3.11/dist-packages (from typer<1.0,>=0.12->gradio) (1.5.4)
Requirement already satisfied: rich>=10.11.0 in /usr/local/lib/python3.11/dist-packages (from typer<1.0,>=0.12->gradio) (13.9.4)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.8.2->pandas<3.0,>=1.0->gr
Requirement already satisfied: markdown-it-py>=2.2.0 in /usr/local/lib/python3.11/dist-packages (from rich>=10.11.0->typer<1.0,>=0.12
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/dist-packages (from rich>=10.11.0->typer<1.0,>=0.
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.28.1-
Requirement already satisfied: mdurl~=0.1 in /usr/local/lib/python3.11/dist-packages (from markdown-it-py>=2.2.0->rich>=10.11.0->type
Downloading gradio-5.31.0-py3-none-any.whl (54.2 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 54.2/54.2 MB 19.9 MB/s eta 0:00:00
Downloading gradio_client-1.10.1-py3-none-any.whl (323 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 323.1/323.1 kB 24.6 MB/s eta 0:00:00
Downloading aiofiles-24.1.0-py3-none-any.whl (15 kB)
Downloading fastapi-0.115.12-py3-none-any.whl (95 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 95.2/95.2 kB 8.5 MB/s eta 0:00:00
Downloading groovy-0.1.2-py3-none-any.whl (14 kB)
Downloading python_multipart-0.0.20-py3-none-any.whl (24 kB)
Downloading ruff-0.11.12-py3-none-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (11.5 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 11.5/11.5 MB 117.6 MB/s eta 0:00:00
Downloading safehttpx-0.1.6-py3-none-any.whl (8.7 kB)
Downloading semantic_version-2.10.0-py2.py3-none-any.whl (15 kB)
Downloading starlette-0.46.2-py3-none-any.whl (72 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 72.0/72.0 kB 6.1 MB/s eta 0:00:00
Downloading tomlkit-0.13.2-py3-none-any.whl (37 kB)
Downloading uvicorn-0.34.2-py3-none-any.whl (62 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 62.5/62.5 kB 6.0 MB/s eta 0:00:00
Downloading ffmpy-0.5.0-py3-none-any.whl (6.0 kB)
Downloading pydub-0.25.1-py2.py3-none-any.whl (32 kB)
Installing collected packages: pydub, uvicorn, tomlkit, semantic-version, ruff, python-multipart, groovy, ffmpy, aiofiles, starlette,
Successfully installed aiofiles-24.1.0 fastapi-0.115.12 ffmpy-0.5.0 gradio-5.31.0 gradio-client-1.10.1 groovy-0.1.2 pydub-0.25.1 pyth
```

```
%%writefile app.py

import gradio as gr
import numpy as np
import joblib

# Load models
knn_model = joblib.load("knn_model.pkl")
svm_model = joblib.load("svm_model.pkl")

def predict(pregnancies, glucose, bp, skin, insulin, bmi, dpf, age, model_choice):
    user_input = np.array([[pregnancies, glucose, bp, skin, insulin, bmi, dpf, age]])
    model = knn_model if model_choice == "k-NN" else svm_model
    pred = model.predict(user_input)[0]
    return "Diabetic" if pred == 1 else "Non-Diabetic"

inputs = [
    gr.Number(label="Pregnancies", value=0),
    gr.Number(label="Glucose Level", value=0),
    gr.Number(label="Blood Pressure", value=0),
```