TASK 4: Disease Prediction from Medical Data

Objective: Predict the **possibility of diseases** based on patient data.
Approach: Apply **classification techniques** to structured medical datasets.
Key Features:

- Use features like symptoms, age, blood test results, etc.
- Algorithms: SVM, Logistic Regression, Random Forest, XGBoost.
- Datasets: Heart disease, Diabetes, Breast Cancer (UCI ML Repository).

4 task.py > ...

```python
# Required Libraries
import pandas as pd
import numpy as np
from sklearn.datasets import load_breast_cancer
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score, classification_report
from sklearn.svm import SVC
from sklearn.linear_model import LogisticRegression
from sklearn.ensemble import RandomForestClassifier
from xgboost import XGBClassifier

# Step 1: Load Dataset
data = load_breast_cancer()
X = pd.DataFrame(data.data, columns=data.feature_names)
y = pd.Series(data.target)

# Step 2: Train-Test Split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Step 3: Standardize Features
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
```

```python
# Step 3: Standardize Features
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

# Step 4: Initialize Models
models = {
    'Logistic Regression': LogisticRegression(max_iter=1000),
    'SVM': SVC(),
    'Random Forest': RandomForestClassifier(n_estimators=100),
    'XGBoost': XGBClassifier(use_label_encoder=False, eval_metric='logloss')
}

# Step 5: Train and Evaluate Models
for name, model in models.items():
    model.fit(X_train, y_train)
    predictions = model.predict(X_test)
    print(f"\nModel: {name}")
    print(f"Accuracy: {accuracy_score(y_test, predictions):.4f}")
    print("Classification Report:\n", classification_report(y_test, predictions))
```

Output:-

```
Model: Logistic Regression
Accuracy: 0.9737
Classification Report:
              precision    recall  f1-score   support

           0       0.98      0.95      0.96        43
           1       0.97      0.99      0.98        71

    accuracy                           0.97       114
   macro avg       0.97      0.97      0.97       114
weighted avg       0.97      0.97      0.97       114


Model: SVM
Accuracy: 0.9825
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.95      0.98        43
           1       0.97      1.00      0.99        71

    accuracy                           0.98       114
   macro avg       0.99      0.98      0.98       114
weighted avg       0.98      0.98      0.98       114
```

```
Model: SVM
Accuracy: 0.9825
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.95      0.98        43
           1       0.97      1.00      0.99        71

    accuracy                           0.98       114
   macro avg       0.99      0.98      0.98       114
weighted avg       0.98      0.98      0.98       114


Model: Random Forest
Accuracy: 0.9649
Classification Report:
              precision    recall  f1-score   support

           0       0.98      0.93      0.95        43
           1       0.96      0.99      0.97        71

    accuracy                           0.96       114
   macro avg       0.97      0.96      0.96       114
weighted avg       0.97      0.96      0.96       114

C:\Users\sushm\AppData\Local\Programs\Python\Python311\Lib\site-packages\xgboost\training.py:183: UserWarning: [11:46:24] WARNING: C:\act
t\xgboost\src\learner.cc:738:
Parameters: { "use_label_encoder" } are not used.
```

```
Model: XGBoost
Accuracy: 0.9561
Classification Report:
              precision    recall  f1-score   support

           0       0.95      0.93      0.94        43
           1       0.96      0.97      0.97        71
Model: XGBoost
Accuracy: 0.9561
Classification Report:
              precision    recall  f1-score   support
Model: XGBoost
Accuracy: 0.9561
Model: XGBoost
Model: XGBoost
Accuracy: 0.9561
Model: XGBoost
Model: XGBoost
Model: XGBoost
Accuracy: 0.9561
Classification Report:
Model: XGBoost
Accuracy: 0.9561
Classification Report:
```

```
Model: XGBoost
Model: XGBoost
Accuracy: 0.9561
Model: XGBoost
Model: XGBoost
Model: XGBoost
Accuracy: 0.9561
Classification Report:
Model: XGBoost
Accuracy: 0.9561
Classification Report:
              precision    recall  f1-score   support
Model: XGBoost
Accuracy: 0.9561
Model: XGBoost
Model: XGBoost
Accuracy: 0.9561
Classification Report:
              precision    recall  f1-score   support
Model: XGBoost
Accuracy: 0.9561
Classification Report:
Model: XGBoost
Model: XGBoost
Model: XGBoost
Accuracy: 0.9561
Classification Report:
              precision    recall  f1-score   support

           0       0.95      0.93      0.94        43
           1       0.96      0.97      0.97        71

    accuracy                           0.96       114
   macro avg       0.96      0.95      0.95       114
weighted avg       0.96      0.96      0.96       114

PS C:\Users\sushm>
```