AI-Powered Disease Prediction Based on Patient Data

```
1. Import Required Libraries
----------------------------
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
import matplotlib.pyplot as plt
import seaborn as sns

2. Load and Prepare the Data
----------------------------
data = pd.read_csv('patient_data.csv')  # Columns: age, sex, bp, cholesterol, glucose, smoking, disease

data['sex'] = data['sex'].map({'male': 0, 'female': 1})
data['smoking'] = data['smoking'].map({'no': 0, 'yes': 1})
data['disease'] = data['disease'].map({'no': 0, 'yes': 1})

X = data.drop('disease', axis=1)
y = data['disease']

3. Train/Test Split
-------------------
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

4. Train the Model
------------------
model = RandomForestClassifier(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

5. Make Predictions and Evaluate
--------------------------------
y_pred = model.predict(X_test)

print("Accuracy:", accuracy_score(y_test, y_pred))
print("\nClassification Report:\n", classification_report(y_test, y_pred))
sns.heatmap(confusion_matrix(y_test, y_pred), annot=True, fmt='d')
plt.title('Confusion Matrix')
plt.show()

6. Predict on New Data
----------------------
new_patient = pd.DataFrame([{
    'age': 45,
    'sex': 0,
    'bp': 130,
    'cholesterol': 200,
```

```
    'glucose': 120,
    'smoking': 1
}])

prediction = model.predict(new_patient)
print("Disease Prediction (1: Yes, 0: No):", prediction[0])
```