import pandas as pd

import numpy as np

from sklearn.model\_selection import train\_test\_split

from sklearn.preprocessing import LabelEncoder

from sklearn.ensemble import RandomForestClassifier

from sklearn.metrics import accuracy\_score

# Load dataset

data = pd.read\_csv('dataset.csv')

# Preprocess data

X = data.iloc[:, :-1].values

y = data.iloc[:, -1].values

# Encode target labels

encoder = LabelEncoder()

y = encoder.fit\_transform(y)

# Split data

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Train model

model = RandomForestClassifier()

model.fit(X\_train, y\_train)

# Make predictions

y\_pred = model.predict(X\_test)

# Evaluate model

accuracy = accuracy\_score(y\_test, y\_pred)

print(f'Accuracy: {accuracy \* 100:.2f}%')

# Function to predict disease based on symptoms

def predict\_disease(symptoms):

symptoms = symptoms.split(',')

input\_data = [0] \* len(X[0])

for symptom in symptoms:

index = data.columns.get\_loc(symptom.strip())

input\_data[index] = 1

input\_data = np.array(input\_data).reshape(1, -1)

prediction = model.predict(input\_data)

return encoder.inverse\_transform(prediction)[0]

# Example usage

print(predict\_disease("Itching, Skin Rash, Nodal Skin Eruptions"))

📊 Input and Output Example

Input: "Itching, Skin Rash, Nodal Skin Eruptions"

Output: "Fungal infection"