Q1.Use supervvised learning to build a model to dectect spam mail.

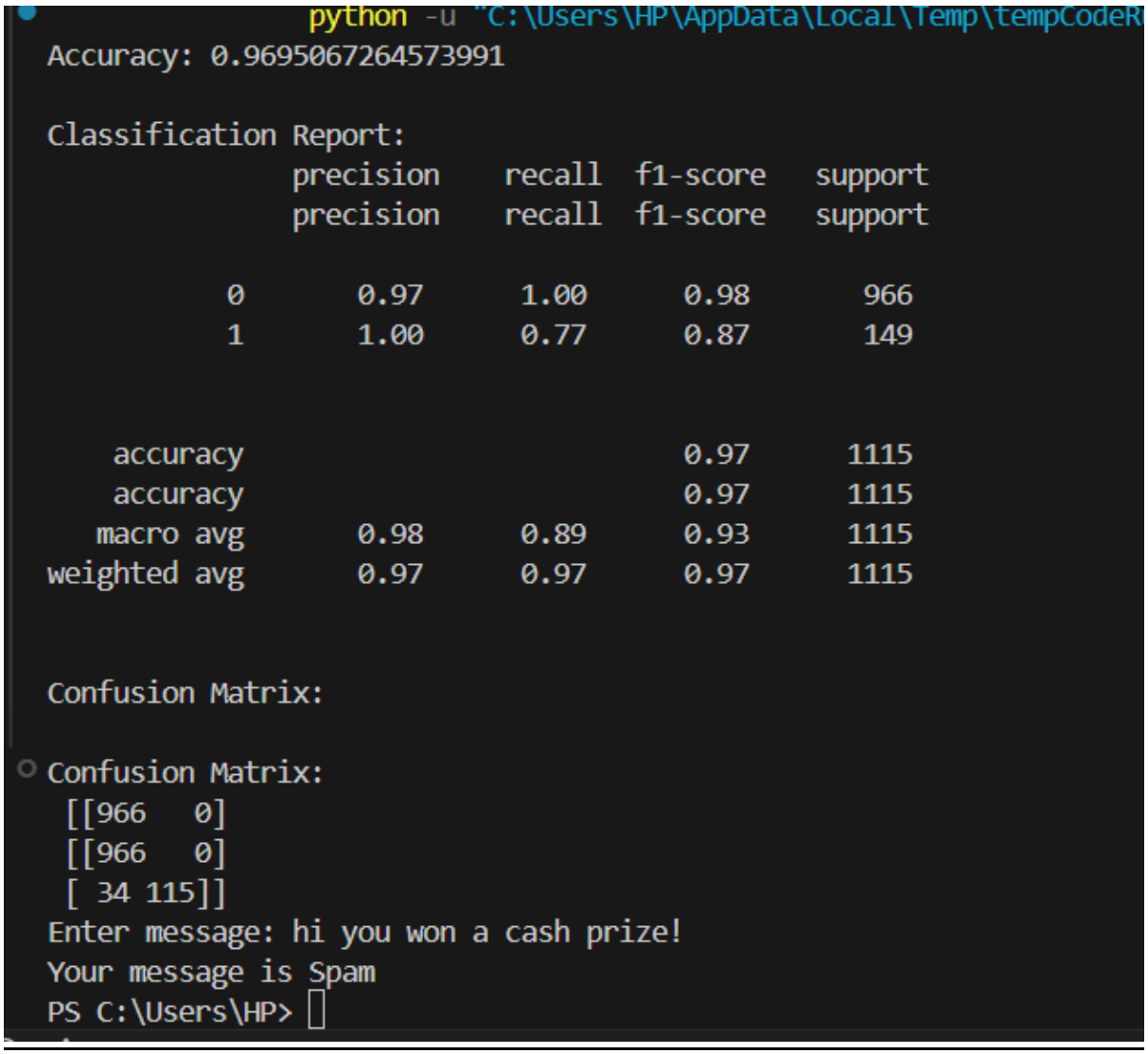
Code

```python
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score,classification_report,confusion_matrix
import pandas as pd
df = pd.read_csv("C:/Users/HP/Pictures/Downloads/SMSSpamCollection.csv",
                 sep='\t', header=None, names=['label', 'message'])
df['label_num'] = df.label.map({'ham': 0, 'spam': 1})
X_train,X_test,Y_train,Y_test=train_test_split(df['message'],df['label_num'],test_size=0.2,random_state=42)
vectorizer=TfidfVectorizer(stop_words='english')
X_train_vec=vectorizer.fit_transform(X_train)
x_test_vec=vectorizer.transform(X_test)
model=model = LogisticRegression(max_iter=1000)
model.fit(X_train_vec,Y_train)
Y_pred=model.predict(x_test_vec)
print("Accuracy:", accuracy_score(Y_test, Y_pred))
print("\nClassification Report:\n", classification_report(Y_test, Y_pred))
print("\nConfusion Matrix:\n", confusion_matrix(Y_test, Y_pred))
user_input=input("Enter message: ")
user_input_vec=vectorizer.transform([user_input])
prediction=model.predict(user_input_vec)[0]
if prediction == 1:
    label = "Spam"
else:
    label = "Ham"

print("Your message is", label)

```

Outout


python -u "C:\Users\HP\AppData\Local\Temp\tempCodeR
Accuracy: 0.9695067264573991
Classification Report:
precision recall f1-score support
precision recall f1-score support
0 0.97 1.00 0.98 966
1 1.00 0.77 0.87 149
accuracy 0.97 1115
accuracy 0.97 1115
macro avg 0.98 0.89 0.93 1115
weighted avg 0.97 0.97 0.97 1115
Confusion Matrix:
Confusion Matrix:
[[966 0]
[[966 0]
[34 115]]
Enter message: hi you won a cash prize!
Your message is Spam
PS C:\Users\HP>