[Day-38 2211cs020196]Write a Python program to classify text using a simple Decision Tree model from `sklearn. tree. DecisionTreeClassifier`. Convert text data into numerical format using `TfidfVectorizer`.Example input: ["Spam messages are annoying", "I won a lottery", "This is a normal message"]

In [1]:
```
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split
documents = ["Spam messages are annoying", "I won a lottery", "This is a normal message",
             "You have been selected for a prize", "This is not spam", "Urgent: Claim your reward now"]
labels = ["spam", "spam", "not_spam", "spam", "not_spam", "spam"]
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(documents)
X_train, X_test, y_train, y_test = train_test_split(X, labels, test_size=0.2, random_state=42)
classifier = DecisionTreeClassifier()
classifier.fit(X_train, y_train)
new_text = ["Congratulations! You have won a prize"]
X_new = vectorizer.transform(new_text)
predicted_category = classifier.predict(X_new)
print(f"The predicted category for the new text '{new_text[0]}' is: {predicted_category[0]}")

```

```
The predicted category for the new text 'Congratulations! You have won a prize' is: spam
```

In [ ]:
```

```