Assignment 3: Write a program to recognize a document is positive or negative based on polarity words using suitable classification method.

Code =>

```
import nltk

from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.naive_bayes import MultinomialNB

# Ensure necessary downloads
nltk.download('punkt')
nltk.download('stopwords')

# Sample dataset
documents = [
    ("I love this product, it is fantastic and great!", "positive"),
    ("This is a terrible experience, I hate it!", "negative"),
    ("I am so happy with the service, excellent support!", "positive"),
    ("The movie was horrible, I dislike it!", "negative"),
    ("It was a good day, everything went well!", "positive"),
    ("Awful customer service, very bad experience!", "negative")
]

# Preprocessing function
def preprocess(text):
    tokens = word_tokenize(text.lower())
    filtered_tokens = [word for word in tokens if word.isalpha() and word not
in stopwords.words("english")]
    return " ".join(filtered_tokens)

# Prepare dataset
texts, labels = zip(*documents)
texts = [preprocess(text) for text in texts]

# Convert text to feature vectors
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(texts)
y = [1 if label == "positive" else 0 for label in labels]

# Train Naïve Bayes classifier
classifier = MultinomialNB()
classifier.fit(X, y)

# Function to predict sentiment
```

```python
def predict_sentiment(text):
    processed_text = preprocess(text)
    text_vector = vectorizer.transform([processed_text])
    prediction = classifier.predict(text_vector)
    return "positive" if prediction[0] == 1 else "negative"

# Debugging: Check processing steps
print("Processed texts:", texts)
print("Vocabulary:", vectorizer.get_feature_names_out())

# Test the function
sample_text = "The product is great and I enjoy using it!"
print(f"Sentiment: {predict_sentiment(sample_text)}")
```

Output =>

```
[Running] python -u "c:\Users\Shreyash
Musmade\Desktop\Practical\MIDS\MIDS_Prac-3\Practical.py"

[nltk_data] Downloading package punkt to C:\Users\Shreyash
[nltk_data]     Musmade\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to C:\Users\Shreyash
[nltk_data]     Musmade\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
Processed texts: ['love product fantastic great', 'terrible experience hate',
'happy service excellent support', 'movie horrible dislike', 'good day
everything went well', 'awful customer service bad experience']
Vocabulary: ['awful' 'bad' 'customer' 'day' 'dislike' 'everything' 'excellent'
 'experience' 'fantastic' 'good' 'great' 'happy' 'hate' 'horrible' 'love'
 'movie' 'product' 'service' 'support' 'terrible' 'well' 'went']
Sentiment: positive

[Done] exited with code=0 in 4.165 seconds
```