predict_svm

November 9, 2019

In [17]:
```
# import necessary libraries
import pickle
import json
from sklearn.calibration import CalibratedClassifierCV
from sklearn.feature_extraction.text import TfidfVectorizer
```

In [3]:
```
# Load saved SVM model that performed best
filename = 'best_svm_model.sav'
loaded_model = pickle.load(open(filename, 'rb'))

# load calibrated SVM model to be able to predict probabilities
calibrated_svm_model = pickle.load(open('svm_model.sav', 'rb'))

# load saved tfidfvectorizer
tfidf_ngram_transformer = pickle.load(open("tfidf_ngram_transformer.pkl", 'rb'))
tfidf_to_predict = TfidfVectorizer(lowercase = True, analyzer='word', token_pattern=r'
                                   ngram_range=(1,3), max_features=500, vocabulary = tfidf_ngra
```

In [27]:
```
def svm_predict(document):
    doc_to_predict = tfidf_to_predict.fit_transform([str(document)])
    predicted_label = loaded_model.predict(doc_to_predict)
    predicted_prob = calibrated_svm_model.predict_proba(doc_to_predict)

    print("Predicted_label =%s" % (predicted_label))
    print("Predicted_probability (confidence of predicted label) =%s" % (predicted_pro

    with open('svm_predcition.json', 'w') as fp:
        json.dump(str(predicted_label[0]), fp)
    # print out saved dictionary
    print("")
    print("Saved as json")
    return predicted_label, predicted_prob
```

In [7]:
```
# Prediction 1 : Original Label 5
```

In [28]:
```
document = "Went here for a lonely sad man's beer and the young bartender who goes by
svm_predict(document)
```

```
Predicted_label =[5]
Predicted_probability (confidence of predicted label) =[[0.13253115 0.07092784 0.1025373  0.242

Saved as json


Out[28]: (array([5]),
          array([[0.13253115, 0.07092784, 0.1025373 , 0.24217359, 0.45183012]]))

In [ ]: # Prediction 2 : Original label 1

In [29]: document2 = "Horrible food and place. Hated it"
         svm_predict(document2)

Predicted_label =[1]
Predicted_probability (confidence of predicted label) =[[0.34197942 0.11040105 0.09919944 0.102

Saved as json


Out[29]: (array([1]),
          array([[0.34197942, 0.11040105, 0.09919944, 0.10271418, 0.34570592]]))

In [ ]: # Prediction 3 : Original label 2

In [30]: document3 = "Food was bland but ambience was okay. Average"
         svm_predict(document3)

Predicted_label =[1]
Predicted_probability (confidence of predicted label) =[[0.33005716 0.11136837 0.10912927 0.105

Saved as json


Out[30]: (array([1]),
          array([[0.33005716, 0.11136837, 0.10912927, 0.10511928, 0.34432593]]))

In [ ]:
```