Prediction using the trained model

```
from google.colab import drive
drive.mount('/content/gdrive')
```

```
Mounted at /content/gdrive
```

Navigating to the location where model and parameter files are present

1. model.h5
2. parameters.sav

```
cd /content/gdrive/MyDrive/NLP/Language Identification
```

```
/content/gdrive/MyDrive/NLP/Language Identification
```

Function to predict the language from text

```
# Importing the libraries
from tensorflow import keras # keras
import pickle # pickle
import numpy as np #numpy
import re # regular expressions
import pandas as pd # pandas

try:
    model = keras.models.load_model('model.h5')
except:
    print("Error : Model not found!!!")
try:
    infile = open('parameters.sav','rb')
    train_max, train_min, vectorizer, feature_names, encoder = pickle.load(infile)
except:
    print("Error : Parameters not found!!!")
```

```
# Preprocessing each text line
def data_preprocess(text):
    # Removing numbers and symbols
    text = re.sub(r'[!@#$()-_,n"%^*?:;~`0-9]', ' ', text)
    text = re.sub(r'[[]]', ' ', text)
    # Lowercasing the text
    text = text.lower()
    return text

# Function to detect language
def detect_language(text):
    text = data_preprocess(text)
    X = vectorizer.fit_transform([text])
    X_feat = pd.DataFrame(data=X.toarray(),columns=feature_names)
    X_feat = (X_feat - train_min)/(train_max-train_min)
    predicted_my_val = model.predict(X_feat)
    val = np.where(predicted_my_val[0] == np.amax(predicted_my_val[0]))[0]
    print("Detected Language :", encoder.classes_[val[0]])
```

```
my_text = input("Enter the text : ")
my_val = detect_language(my_text)
```

```
Enter the text : 1846ء اور پنجاب میں سنہ 1849ء میں بطور دفتری زبان نافذ کیا۔ اس کے علاوہ خلیجی، یورپی، ایشیائی اور امریکی علاق
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:22: FutureWarning: Possible nested set at position 1
Detected Language : Urdu
```


0s completed at 15:13