

NAME: J JANASREE

COLLAGE: PANIMALAR INSTITUTE OF TECHNOLOGY-2115

NAAN MUDHALVAN ID : au211521104056

DEPARTMENT: COMPUTER SCIENCE AND ENGINEERING
3rd YEAR 6th SEMESTER

DOMAIN NAME: GENERATRED AI

PROJECT TITLE: LYRICS GENERATED [DEEP LEARNING]

ALGORITHM : RNN-RECURRENT NEWRAL NETWORKS

LYRICS GENERATED IN DEEP LEARNING AGORITHM USING IN RNN-RECURRENT NEWRAL NETWORKS

INPUT:

#importing the libraries

import pandas *as* pd

import numpy *as* np

import seaborn *as* sns

import matplotlib.pyplot *as* plt

import string, os

import nltk

import re

import keras

import random

import io

from keras.utils *import* np_utils

from keras.models *import* Sequential

from keras.layers *import* LSTM, Dense, Dropout

from keras.optimizers *import* Adamax

import sys

from PIL *import* Image

from wordcloud *import* WordCloud, STOPWORDS, ImageColorGenerator

from PIL *import* Image, ImageDraw, ImageFont

import warnings

warnings.filterwarnings("ignore")

data = pd.read_csv("../input/lyrics/Songs.csv")

data.head()

print("Artists in the data:\n",data.Artist.value_counts())

print("Size of Dataset:",data.shape)

data["No_of_Characters"] = data["Lyrics"].apply(len)

```

data["No_of_Words"]=data.apply(lambda row: nltk.word_tokenize(row["Lyrics"]), axis=1).apply(len)
data["No_of_Lines"] = data["Lyrics"].str.split('\n').apply(len)
data.describe()
plt.figure(figsize=(15,15))
ax = sns.pairplot(data, hue="Artist", palette="plasma")
stopwords = set(STOPWORDS)
wordcloud = WordCloud(stopwords=stopwords, background_color="#444160", color map="Purples", max_words=800).generate(" ".join(data["Lyrics"]))
plt.figure(figsize=(12,12))
plt.imshow(wordcloud, interpolation="bilinear")
plt.show()
def My_song(song):
    img = Image.open("../input/image-for-notebook/Pink and White metric Marketing Presentation (1).png")
    Text_on_image = ImageDraw.Draw(img)
    myFont = ImageFont.truetype("../input/font-style/DancingScript-VariableFont_wght.ttf", 45)
    Text_on_image.text((620,90), song, font=myFont, fill=(255, 255, 255))
    return img
My_song(data.Lyrics[42][:500])

Corpus = ""
for listitem in data.Lyrics:
    Corpus += listitem

Corpus = Corpus.lower() #converting all alphabets to lowercase
print("Number of unique characters:", len(set(Corpus)))
print("The unique characters:", sorted(set(Corpus)))
to_remove = ['{', '}', '~', '©', 'à', 'á', 'â', 'ä', 'ç', 'è', 'é', 'ê', 'ë', 'í', 'ñ', 'ó', 'ö', 'ü', 'ö', 'e', 'í', 'س', 'ل', 'م', 'و', '\u2005', '\u200a', '\u200b', '-', '—', '‘', '’', '!', '“', '”', '...', '\u205f', '\uffff', '!', '&', '(', ')', '*', '-', ' ', '/']
for symbol in to_remove:
    Corpus = Corpus.replace(symbol, " ")
print("The unique characters:", sorted(set(Corpus)))
symb = sorted(list(set(Corpus)))

L_corpus = len(Corpus) #length of corpus
L_symb = len(symb) #length of total unique characters
mapping = dict((c, i) for i, c in enumerate(symb))
reverse_mapping = dict((i, c) for i, c in enumerate(symb))

print("Total number of characters:", L_corpus)
print("Number of unique characters:", L_symb)
length = 40
features = []
targets = []

```

```

for i in range(0, L_corpus - length, 1):
    feature = Corpus[i:i + length]
    target = Corpus[i + length]
    features.append([mapping[j] for j in feature])
    targets.append(mapping[target])

L_datapoints = len(targets)
print("Total number of sequences in the Corpus:", L_datapoints)
X = (np.reshape(features, (L_datapoints, length, 1))) / float(L_sy
    mb)
y = np_utils.to_categorical(targets)
model = Sequential()
model.add(LSTM(256, input_shape=(X.shape[1], X.shape[2])))
model.add(Dense(y.shape[1], activation='softmax'))
opt = Adamax(learning_rate=0.01)
model.compile(loss='categorical_crossentropy', optimizer=opt)
model.summary()
history = model.fit(X, y, batch_size=128, epochs=100)
history_df = pd.DataFrame(history.history)
fig = plt.figure(figsize=(15,4), facecolor="#B291B6")
fig.suptitle("Learning Plot of Model for Loss")
pl=sns.lineplot(data=history_df["loss"],color="#444160")
pl.set(ylabel="Training Loss")
pl.set(xlabel="Epochs")
def Lyrics_Generator(starter,Ch_count): #,temperature=1.0):
    generated= ""
    starter = starter
    seed=[mapping[char] for char in starter]
    generated += starter
    for i in range(Ch_count):
        seed=[mapping[char] for char in starter]
        x_pred = np.reshape(seed, (1, len(seed), 1))
        x_pred = x_pred / float(L_symb)
        prediction = model.predict(x_pred, verbose=0)[0]
        prediction = np.asarray(prediction).astype('float64')
        prediction = np.log(prediction) / 1.0
        exp_preds = np.exp(prediction)
        prediction = exp_preds / np.sum(exp_preds)
        probas = np.random.multinomial(1, prediction, 1)
        index = np.argmax(prediction)
        next_char = reverse_mapping[index]
        generated += next_char
        starter = starter[1:] + next_char
    return generated
song_1 = Lyrics_Generator("the shoe shrunk, and the school belt g
    ot ridiculously petit", 400)
My_song(song_1)
song_2 = Lyrics_Generator("i'm a sunflower, a little funny", 400)
My_song(song_2)

```

OUTPUT:

| | Artist | Title | Lyrics |
|---|--------------|--------------------------|---|
| 0 | Taylor Swift | cardigan | Vintage tee, brand new phone\nHigh heels on co... |
| 1 | Taylor Swift | exile | I can see you standing, honey\nWith his arms a... |
| 2 | Taylor Swift | Lover | We could leave the Christmas lights up 'til Ja... |
| 3 | Taylor Swift | the 1 | I'm doing good, I'm on some new shit\nBeen say... |
| 4 | Taylor Swift | Look What You Made Me Do | I don't like your little games\nDon't like you... |

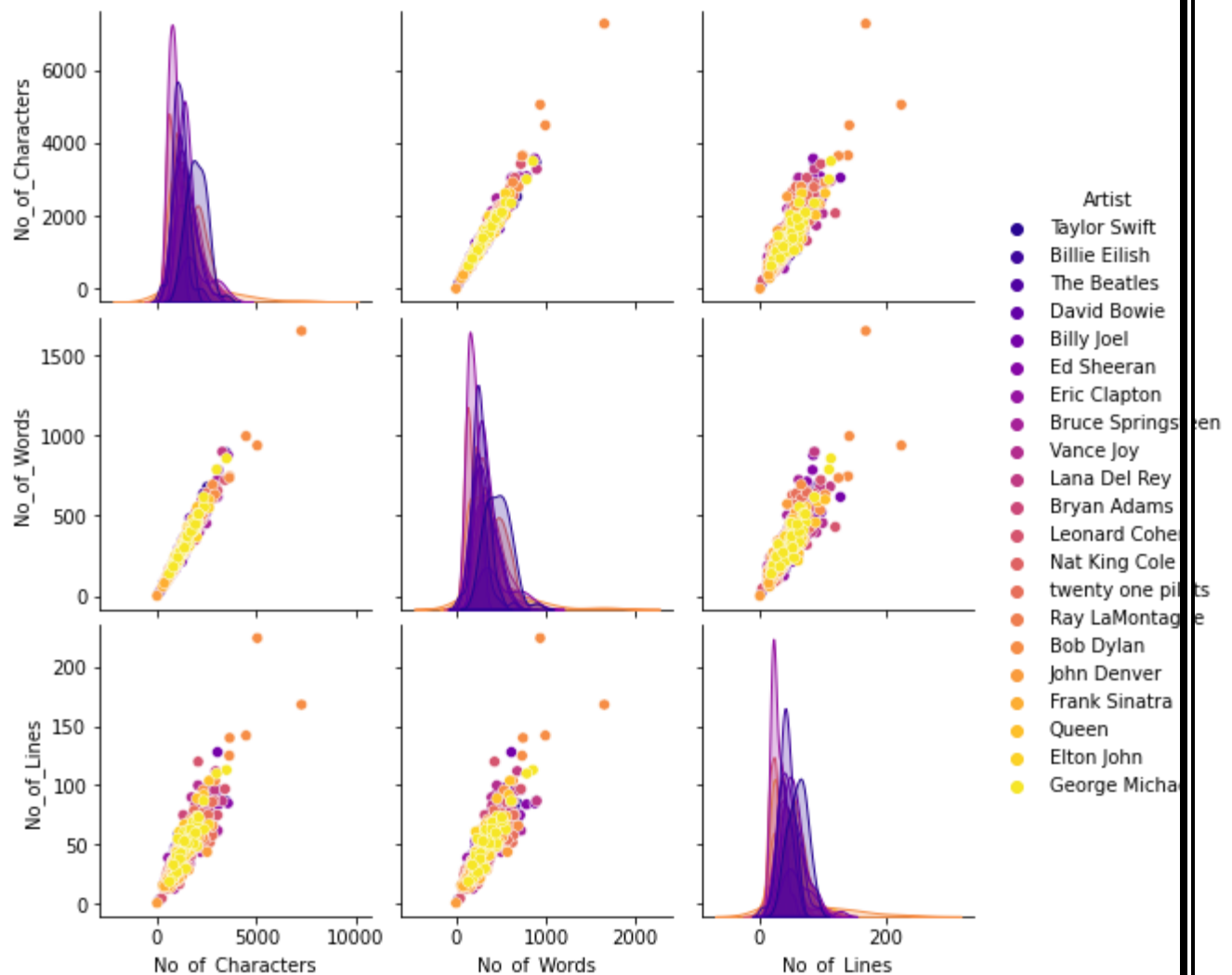
Artists in the data:

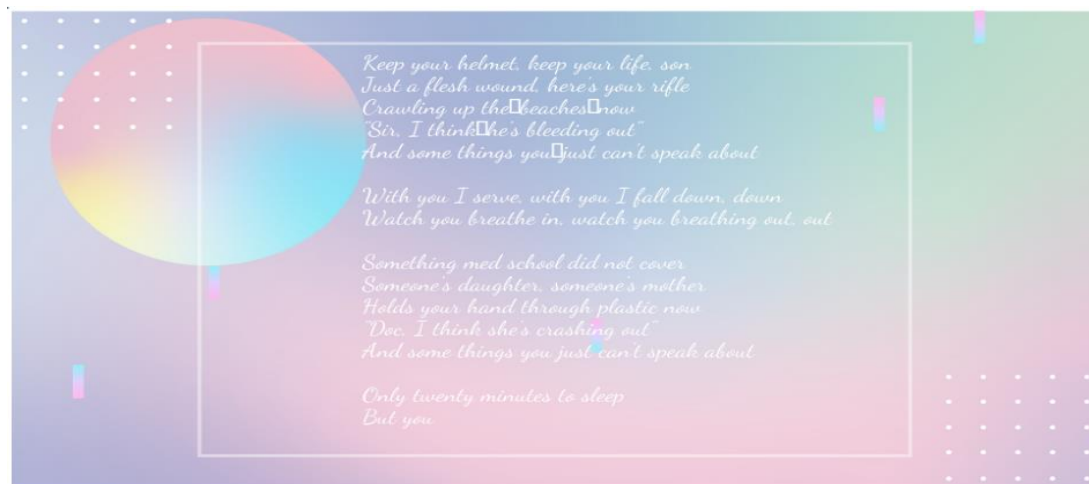
David Bowie 50
Billy Joel 50
Taylor Swift 50
Billie Eilish 50
Eric Clapton 50
Leonard Cohen 50
Bruce Springsteen 40
The Beatles 35
George Michael 30
Vance Joy 30
Frank Sinatra 30
Elton John 30
twenty one pilots 30
John Denver 30
Bryan Adams 30
Nat King Cole 30
Ray LaMontagne 30
Queen 30
Lana Del Rey 30
Ed Sheeran 20
Bob Dylan 20
Name: Artist, dtype: int64

Size of Dataset: (745, 3)

| | No_of_Characters | No_of_Words | No_of_Lines |
|-------|------------------|-------------|-------------|
| count | 745.000000 | 745.000000 | 745.000000 |
| mean | 1403.347651 | 319.338255 | 46.277852 |
| std | 666.721467 | 156.067038 | 21.180531 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 946.000000 | 215.000000 | 33.000000 |
| 50% | 1289.000000 | 291.000000 | 44.000000 |
| 75% | 1714.000000 | 389.000000 | 56.000000 |
| max | 7267.000000 | 1652.000000 | 224.000000 |

Figure size 1080x1080 with 0 Axes>





The unique characters: ['\n', ',', '!', '""', '&', '""', '(', ')', '*', ',', '-', '!', '/', '0', '1', '2', '3', '4', '5', '6', '7', '8', '9', ':', ';', '>', '?', '^', 'a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j', 'k', 'l', 'm', 'n', 'o', 'p', 'q', 'r', 's', 't', 'u', 'v', 'w', 'x', 'y', 'z', '{', '}', '~', '©', 'à', 'á', 'ä', 'ä', 'ç', 'è', 'é', 'ê', 'ë', 'í', 'ñ', 'ó', 'ö', 'ü', 'õ', 'e', 'و', 'م', 'ل', 'س', 'ل', '\u2005', '\u200a', '\u200b', '—', '—', '‘', '’', '‘', '’', '…', '\u205f', '\ufeff']

linkcode

Model: "sequential"

[illegible]

| | | |
|-------------|-------------|--------|
| lstm (LSTM) | (None, 256) | 264192 |
|-------------|-------------|--------|

| | | |
|-------------------|-------------|---|
| dropout (Dropout) | (None, 256) | 0 |
|-------------------|-------------|---|

| | | |
|---------------|------------|-------|
| dense (Dense) | (None, 47) | 12079 |
|---------------|------------|-------|

=====

=====

Total params: 276,271

Trainable params: 276,271

Non-trainable params: 0

[Text(0.5, 0, 'Epochs')]

