Challenge - Apna Time Aaega!

Apna Time Aayega - Lyrics Generation!

Use Markov Chains to create a Predictive Model for Text

In this fun challenge, you will generate song lyrics for 'Apna Time Aaega' using Machine Learning for the movie Gully Boy (2019 Indian Hindi-language musical drama film). You are given a training set which contains lyrics created by Ranveer Singh, your task is to train a model using Markov Chains to generate lyrics which look similar to actual lyrics.

Dataset

- Dataset contains the text file containing actual lyrics of the song. Since the data-set is scrapped from the internet you need to remove the starting and ending

  tags and clean it before feeding to model.

Submission Format

- Submit a '.txt' file containing generated lyrics upto 2000 characters. Your model should be able to generate new line characters as well.
- Your lyrics must start with the word 'apna' and use numpy random seed of 11 (for consistent result and avoid any randomization).
- You result will be evaluated word by word with the expected output given by Markov Chain.
- You can assume prediction of current character depends only on last 4 characters (Use K=4 in Markov Chain Model)

Scoring

Your score will depend upon number of words matched with the expected output. Take care of white spaces including new line characters.