NEXT WORD PREDICTION APP

PLATFORM USED: R Studio

FILES NEEDED TO RUN THE CODE:

- Clean_TextString.R
- NGram_generator.R
- server.R
- ui.R
- Word_prediction.R
- /final (Folder)
    1. en_US/en_US.blogs.txt
    2. en_US/en_US.news.txt
    3. en_US/en_US.twitter.txt
- /RData (Folder)
    1. DF_unigram.RData
    2. DF_bigram.RData
    3. DF_trigram.RData
    4. DF_quadgram.RData

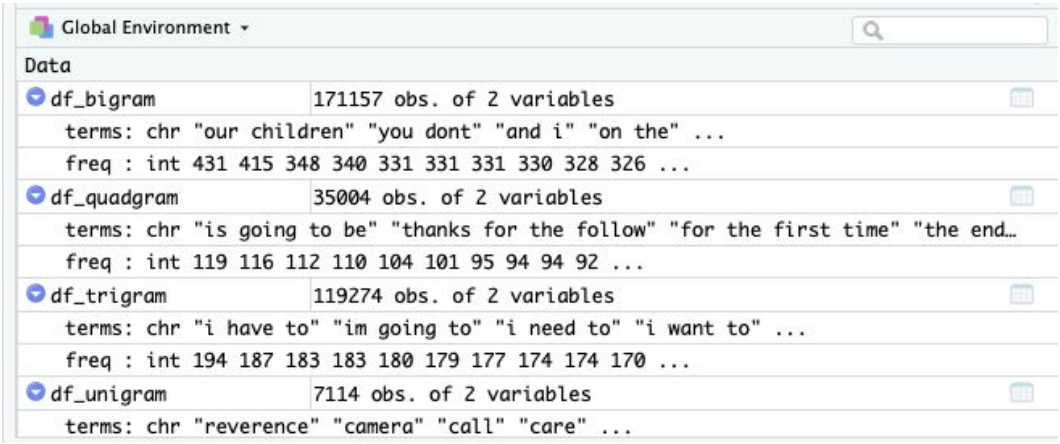
TO INSTALL THE USED PACKAGES IN THE FILE:

Eg. Type install.packages("some_package_name") in the R studio terminal

```
> install.packages("tm")
```

FILES DESCRIPTION:

- NGram_generator.R:
  Creates unigram, bigram, trigram, quadgram files. For the sake of easily running the code, already generated Ngram files can be used which are there in RData folder as mentioned above.



- Clean_TextString.R:
  It contains a function which takes the input String and cleans it. This file is imported inside server.R

- Word_prediction.R:
  A function which takes in the input String, and based on the length after the split, it checks the Ngram RData file and calculates the conditional probability of a word and sorts in the decreasing order of probability and sends it to the data table

- server.R:
  Contains shinyServer function which calls the word prediction function and generates output and sends it to ui.R for rendering on the web app

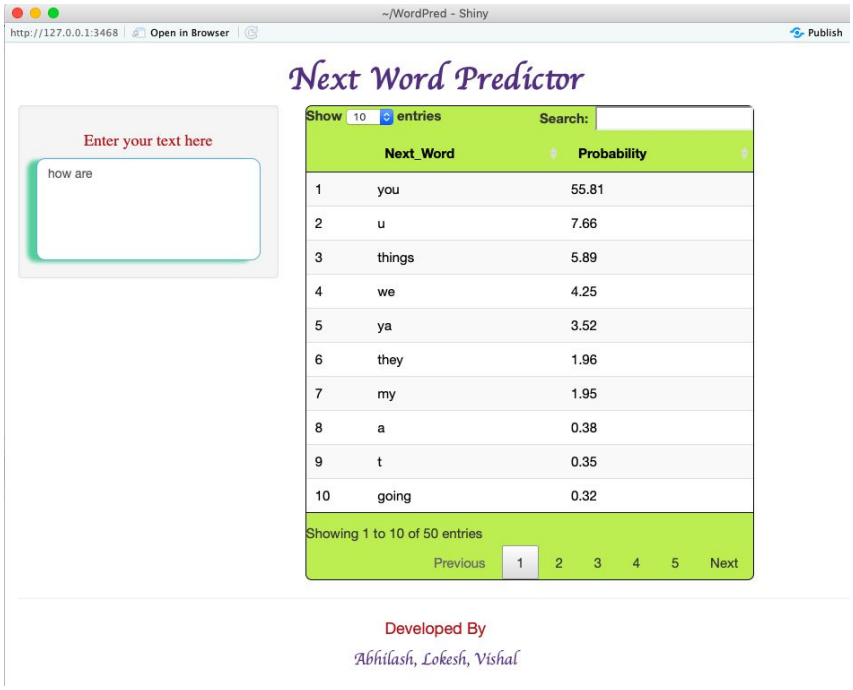
- ui.R:
  Contains shinyUI function which contains the code

RUNNING THE APP:

- Create a new project in R Studio and put all the files in the project directory.
- Go to server.R and press Run App button. It would open a local browser with some assigned port for the web app.

```
> runApp()

Listening on http://127.0.0.1:3468
```



- As you can see the image above, we can enter the text in the side text box. As soon as you enter a word, it starts predicting the next word and gives its probability percentage in the data table which is on the right side.

COMMENTS:

- This live working code has been presented during the poster presentation.
- There is no separate MAKEFILE for this.
- Doesn't work for proper nouns, non-meaningful words like eg. "jjj"