

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

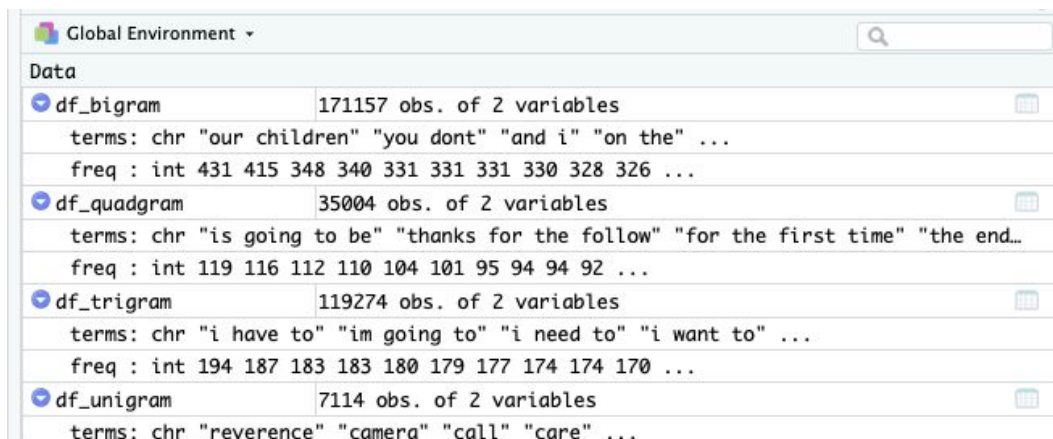
TO INSTALL THE USED PACKAGES IN THE FILE:

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.



The screenshot shows the 'Global Environment' window in R Studio. It lists four data objects under the 'Data' section:

Object Name	Observations	Variables	Sample Data
df_bigram	171157 obs.	2 variables	terms: chr "our children" "you dont" "and i" "on the" ... freq : int 431 415 348 340 331 331 331 330 328 326 ...
df_quadgram	35004 obs.	2 variables	terms: chr "is going to be" "thanks for the follow" "for the first time" "the end..." freq : int 119 116 112 110 104 101 95 94 94 92 ...
df_trigram	119274 obs.	2 variables	terms: chr "i have to" "im going to" "i need to" "i want to" ... freq : int 194 187 183 180 179 177 174 174 170 ...
df_unigram	7114 obs.	2 variables	terms: chr "reverence" "camera" "call" "care" ...


- **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



The screenshot shows a web browser window titled "Next Word Predictor". On the left, there is a text input field labeled "Enter your text here" containing the text "how are". On the right, there is a table with two columns: "Next_Word" and "Probability". The table lists 10 predicted words and their probabilities. Below the table, it says "Showing 1 to 10 of 50 entries" and has pagination controls: "Previous", "1", "2", "3", "4", "5", and "Next". At the bottom, it says "Developed By Abhilash, Lokesh, Vishal".

	Next_Word	Probability
1	you	55.81
2	u	7.66
3	things	5.89
4	we	4.25
5	ya	3.52
6	they	1.96
7	my	1.95
8	a	0.38
9	t	0.35
10	going	0.32

Showing 1 to 10 of 50 entries

Previous 1 2 3 4 5 Next

Developed By
Abhilash, Lokesh, Vishal

- 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"

Next Word Predictor

Enter your text here

jjj |

Error: Column `NextPrediction` is unknown

Developed By

Abhilash, Lokesh, Vishal