Text Suggestion Presentation

cwl

15/06/2021

Introduction

Objective

- Build a predictive model of English text
- Develop an interface for our model via a Shiny application.
- When a user type text, he can see predictions for his next word

Background

- Final task in the Data Science Capstone to create a usable data product

Data

- Provided by the Data Science Capstone
- en_US.blogs.txt, en_US.news.txt, en_US.twitter.txt

Model

- Apply N-gram to build our model
- (N-gram: a contiguous sequence of n items from a given sample of text)

Preprocessing

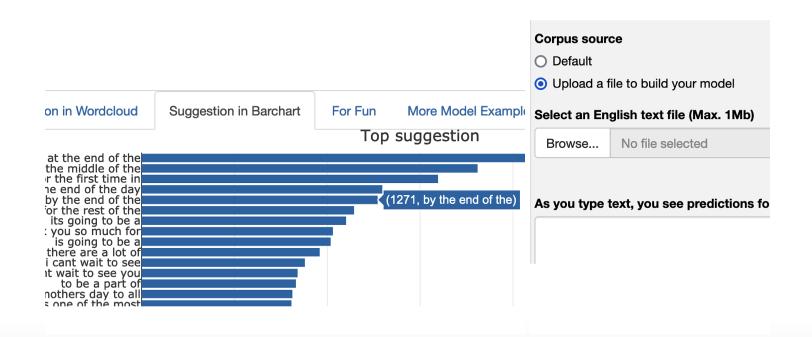- Remove URLs
- Remove punctuation
- Remove numbers
- Remove English stopwords
- Remove special characters

N-gram

- Train Bigram, Trigram, 5-grams and 7-gram

Application Features

- Automatic suggestion

- Build your own model

- Generate "composition"

- Interactive selection

- Suggestion visualization

Futher Information

- View Source
- Data Science Capstone
- Report [bug](bug)