N-Gram Word Predictor

James Wang for the John Hopkins Coursera Data Science Capstone

Enter a sentence, press the 'Predict next' button, and see what the algorithm thinks will come next!

The underlying model is an n-gram model, with quadgrams from a dataset of Twitter, news articles, and blog posts. It uses 'Stupid Backoff' (Brants et al 2007) to deal with unseen n-grams.

1) Start by entering text here

3) See most likely & top 5 predictions

Input

> happy birthday to

Type in a sentence above, press the button below, and the results will display to the right.

[Predict next]

2) Click on "Predict next" (or just press enter)

happy birthday to

my

Top 5 Possibilities:

	Next Possible Word	Score (Adjusted Freq)
1	my	11
2	the	6
3	me	4
4	one	4
5	anyone	1