This website discusses all issues related to text classification (stemming, stop words, etc)

<https://www.codeschool.com/blog/2016/03/25/machine-learning-working-with-stop-words-stemming-and-spam/>

Check Linkurious for visualization app.

Porter Stemmer algorithm is good to review, and code is available. Can be used in most text processing methods. IDEA: can it be used to improve string distances like Levenstein?

Download: <http://www.sourcecodeonline.com/details/porter_stemming_algorithm.html>

# Matlab text analytics toolbox

# normalizeWords

Remove inflections from words using the Porter stemmer

[collapse all in page](javascript:void(0);)

normalizeWords uses the Porter stemmer to group different forms of English words by reducing them to a common stem. This common stem is not necessarily a proper English word.

## Syntax

newDocuments = normalizeWords(documents)

newWords = normalizeWords(words)

===================================================================