

ECS 193AB Winter/Spring2017

Text Mining Web Portal

This project involves writing R scripts to perform various text mining and analysis function and a web-based interface for building analysis corpus, performing parametrized analysts (running the R scripts with tuning parameters) and displaying results tabular, visualized, etc. formats as appropriate. In the DSL we perform a variety of Natural Language Processing (NLP) and Machine Learning analysis. This project will build a web portal that allows users to select a corpus on which they want to work (a certain set of texts) and the kind of analysis they want to perform (hapax richness, topic modeling, etc.). Each of the various forms of analysis can typically be “tuned” with a variety of parameters that affect output. The system will allow the user to set parameters and then hit a button to run the analysis. The system will then return the results as a data file and/or visualization as appropriate.

Contact:

Carl G. Stahmer, PhD
Director of Digital Scholarship
Library of the University of California, Davis
Associate Director for Humanities
UC Davis Data Science Initiative
cstahmer@ucdavis.edu (<mailto:cstahmer@ucdavis.edu>)
[530-752-1138](tel:530-752-1138) (<tel:530-752-1138>)

[Create a free website or blog at WordPress.com. Do Not Sell My Personal Information](#)