Writing Samples for CLMS

Youyun Zhang

1. Using Eigenvectors of Bigram Graph to Derive Parts of Speech for English and Mandarin

Filename	Size	
EigenV_POS_YZ.ipynb	279kb	The project file, runs in Jupyter notebook.
EigenV_POS_YZ.py	21kb	Same file
EigenV_POS_YZ.html	584kb	Same file
Austen5.txt	178kb	First 18 chapters of *Pride and Prejudice*, about 32624 words.
沉默之门 3.txt	78kb	5 chapters of *沉默之门*, a Chinese novel, about 25934 words.

- This is a project that I worked on in my free time, with some help from Professor Khalil Iskarous.
- The Chinese tokenization module I used: https://github.com/fxsjy/jieba#jieba-1. Automatic installation: *easy_install jieba* or *pip install jieba*.

2. Decision Tree Learning

Filename	Size	
DecisionTreeLearning_YZ.pdf	130kb	Brief writeup file
sample.h	474b	C++ file
decisiontree.h	282b	C++ file
makedecision.h	4k	C++ file
main.cpp	3k	C++ file
makefile	36b	C++ file
mushrooms.txt	260kb	Dataset file

- Coursework in CSCI 360 Artificial Intelligence.
- Runs in a C++ environment:

  make

  ./dt mushroom.txt

3. Model for epidemics

Files:

Filename	Size	
Epidemic_SIRmodeling_YZ.pdf	110kb	A SIR model for an epidemic

- Coursework in Math 466 Dynamic Modeling.

4. Effects of Age of Acquisition: Evidence from Grammaticality Judgement of Non-manual Markings in American Sign Language

Filename	Size	
EffectsOfAOA_YZ.pdf	350kb	

- A research hypothesis that I wrote in a linguistic course LING 410 Second language Acquisition.

Other projects I have worked on: https://github.com/youyunzh/RecentProjects