labMT-simple Documentation

Release 2.3.1.5

Andy Reagan

July 20, 2015

CONTENTS

Contents:

CHAPTER

ONE

GETTING STARTED

TL;DR a simple labMT usage script

1.1 Usage

This script uses the language assessment by Mechanical Turk (labMT) word list to score the happiness of a corpus. The labMT word list was created by combining the 5000 words most frequently appearing in four sources: Twitter, the New York Times, Google Books, and music lyrics, and then scoring the words for sentiment on Amazon's Mechanical Turk. The list is described in detail in the publication Dodds' et al. 2011, PLOS ONE, "Temporal Patterns of Happiness and Information in a Global-Scale Social Network: Hedonometrics and Twitter."

Given two corpora, the script "storylab.py" creates a word-shift graph illustrating the words most responsible for the difference in happiness between the two corpora. The corpora should be large (e.g. at least 10,000 words) in order for the difference to be meaningful, as this is a bag-of-words approach. As an example, a random collection of English tweets from both Saturday January 18 2014 and Tuesday January 21 2014 are included in the "example" directory. They can be compared by moving to the test directory, using the command

```
python example.py example-shift.html
```

and opening the file `example-shift.html` in a web browser. For an explanation of the resulting plot, please visit

http://www.hedonometer.org/shifts.html

1.2 Installation

Cloning the github directly is recommended, i.e.

```
git clone https://github.com/andyreagan/labMT-simple.git
```

and then installing locally using

```
python setup.py install
```

This repository can also be installed using pip

```
pip install labMTsimple
```

in which case you can download the tests from github and run them, if desired.

1.3 Running tests

Tests are based on nose2, `pip install nose2`, and can be run by executing

```
nose2
```

in the root directory of this repository.

This will compare the two days in test/data and print test.html which shifts them, allowing for a changable lens.

1.4 Building these docs

Go into the docs directory (activate local `virtualenv` first), and do the following:

```
sphinx-apidoc -o . ../labMTsimple
make html
make latexpdf
erm
git add *
git commit -am ``new docs, probably should just add a pre-commit hook''```
```

CHAPTER

TWO

DETAILED EXAMPLES

2.1 Preparing texts

2.2 Loading dictionaries

CHAPTER

THREE

MAKING WORDSHIFTS

3.1 More Installation

3.2 Example

CHAPTER

FOUR

ADVANCED USAGE

4.1 About Tries

4.2 Advanced Parsing

CHAPTER

FIVE

LABMTSIMPLE

5.1 labMTsimple package

5.1.1 Submodules

5.1.2 labMTsimple.speedy module

class labMTsimple.speedy.**sentiDict**(*corpus*, *datastructure='dict'*, *bootstrap=False*, *stopVal=0.0*, *bananas=False*, *loadFromFile=False*)

bootstrapify()

folders = ('labMT', 'ANEW', 'LIWC', 'MPQA-lexicon', 'liu-lexicon', 'Warriner')

loadDict(*bananas*)

makeDaTrie()

makeListsFromDict(*userdict*)

makeMarisaTrie()

matcherTrieBool(*word*)
matcherTrieBool(word) just checks if a word is in the list

matcherTrieDa(*word*, *wordVec*, *count*)

matcherTrieDict(*word*, *wordVec*, *count*)

matcherTrieMarisa(*word*, *wordVec*, *count*)

openWithPath(*filename*, *mode*)

scoreTrieDa(*wordDict*)

scoreTrieDict(*wordDict*)

scoreTrieMarisa(*wordDict*)

titles = ['LabMT', 'ANEW', 'LIWC', 'MPQA', 'Liu', 'Warriner']

wordVecifyTrieDa(*wordDict*)

wordVecifyTrieDict(*wordDict*)

wordVecifyTrieMarisa(*wordDict*)

labMTsimple.speedy.**u**(*x*)

5.1.3 labMTsimple.storyLab module

labMTsimple.storyLab.**allEmotions**(*tmpStr*, **allDicts*)

labMTsimple.storyLab.**emotion**(*tmpStr*, *someDict*, *scoreIndex=1*, *shift=False*, *happsList=[]*)

labMTsimple.storyLab.**emotionFileReader**(*stopval=1.0*, *lang='english'*, *min=1.0*, *max=9.0*, *returnVector=False*)

labMTsimple.storyLab.**emotionFileReaderRaw**(*stopval=1.0*, *fileName=u'labMT1raw.txt'*, *min=1.0*, *max=9.0*, *returnVector=False*)

labMTsimple.storyLab.**emotionV**(*frequencyVec*, *scoreVec*)

labMTsimple.storyLab.**plothapps**(*happsTimeSeries*, *picname*)

labMTsimple.storyLab.**shift**(*refFreq*, *compFreq*, *lens*, *words*, *sort=True*)

labMTsimple.storyLab.**shiftHtml**(*scoreList*, *wordList*, *refFreq*, *compFreq*, *outFile*)

labMTsimple.storyLab.**shiftHtmlSelf**(*scoreList*, *wordList*, *compFreq*, *outFile*)

labMTsimple.storyLab.**stopper**(*tmpVec*, *labMTvector*, *labMTwords*, *stopVal=1.0*, *ignore=[]*)

5.1.4 Module contents

CHAPTER

SIX

INDICES AND TABLES

- genindex
- modindex
- search

l