

Version

1

ZHOUNING MA

Text Summarization Tool

User Manual

Table of Contents

Features	1
System Requirement.....	1
Download the code.....	2
Run the software	3
How to use	5
Question	6

Features

Today's world is information explosion world and people receive hug amount of news, documents and text context. The amount of text requires to be effectively summarized for usage. It is impossible to do that all by human while automatic text summarization is an important technology to solve this problem.

This software 'Text Summarization tool' are able to do the text summarization. The current version is using extractive summarization method. The tool supports the following algorithms:

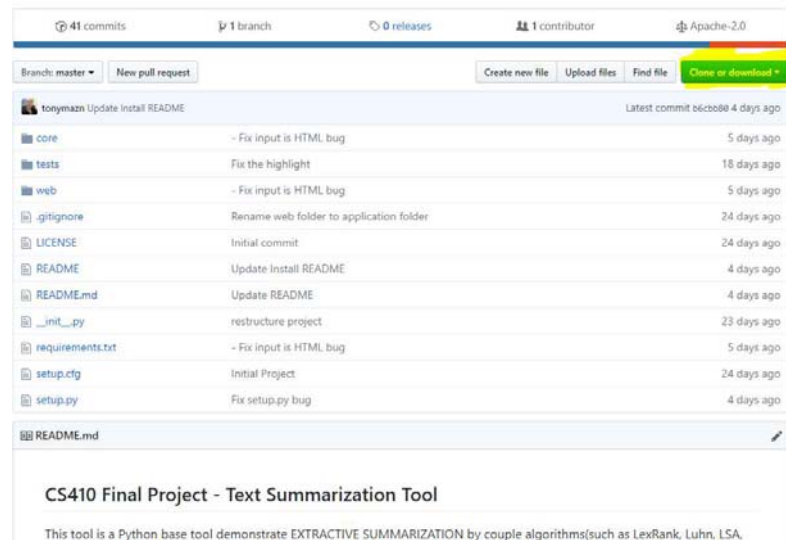
- Luhnⁱ - is a heuristic method
- Text Rankⁱⁱ - exploits the structure of the text and determine key words "central"
- LSA - Use the latent semantic analysis technique to identify semantically import sentences.
- Lex Rankⁱⁱⁱ – Unsupervised approach inspired by PageRank

System Requirement

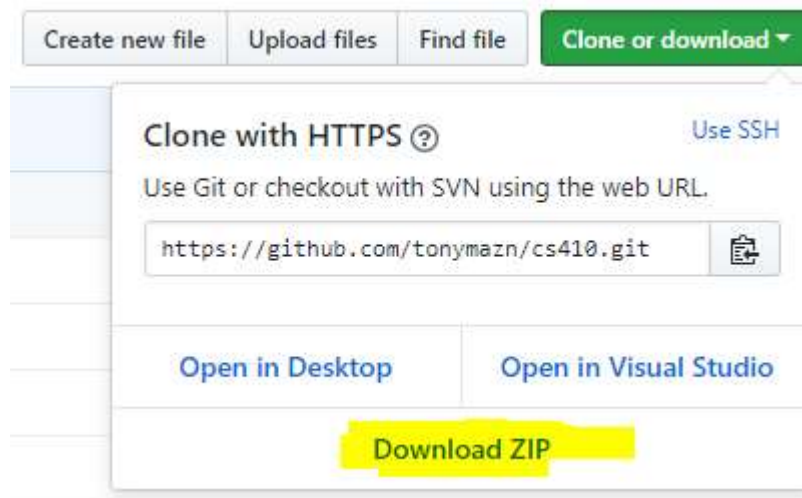
- Windows, Linux, Mac
- Python 3.0 or above with pip

Download the code

1. Go to <https://github.com/tonymazn/cs410>



2. Click “Clone or download” button



3. Selected “Download ZIP” option, will download a zip file, Unzip the file to a folder.
4. Go to that folder, you will see the following screen shot

cs410-master				
Name		Date modified	Type	Size
core		2018-12-10 6:56 AM	File folder	
tests		2018-12-10 6:56 AM	File folder	
web		2018-12-10 6:56 AM	File folder	
.gitignore		2018-12-10 6:56 AM	GITIGNORE File	2 KB
__init__		2018-12-10 6:56 AM	Python File	1 KB
LICENSE		2018-12-10 6:56 AM	File	12 KB
README.md		2018-12-10 6:56 AM	MD File	4 KB
requirements		2018-12-10 6:56 AM	Text Document	1 KB
setup.cfg		2018-12-10 6:56 AM	CFG File	1 KB
setup		2018-12-10 6:56 AM	Python File	1 KB

5. Type command:

Python setup.py install

```

C:\Users\Tony\Downloads\cs410-master\cs410-master>python setup.py install

```

6. It will start to install system and all the dependences.
7. Done.

Run the software

1. Once the software dependences are installed, it is ready to run the software.
2. Go to the folder you unzip:

cs410-master				
Name		Date modified	Type	Size
core		2018-12-10 6:56 AM	File folder	
tests		2018-12-10 6:56 AM	File folder	
web		2018-12-10 6:56 AM	File folder	
.gitignore		2018-12-10 6:56 AM	GITIGNORE File	2 KB
__init__		2018-12-10 6:56 AM	Python File	1 KB
LICENSE		2018-12-10 6:56 AM	File	12 KB
README.md		2018-12-10 6:56 AM	MD File	4 KB
requirements		2018-12-10 6:56 AM	Text Document	1 KB
setup.cfg		2018-12-10 6:56 AM	CFG File	1 KB
setup		2018-12-10 6:56 AM	Python File	1 KB

- Run the following command,

```
python -m web.webserver
```

- The Flask web application 'webserver' will start

```
Administrator: Command Prompt - python -m web.webserver

C:\Users\tony.z.ma\Downloads\cs410-master\cs410-master>python -m web.webserver
* Serving Flask app "webserver" (lazy loading)
* Environment: production
  WARNING: Do not use the development server in a production environment.
  Use a production WSGI server instead.
* Debug mode: off
```

- Done.

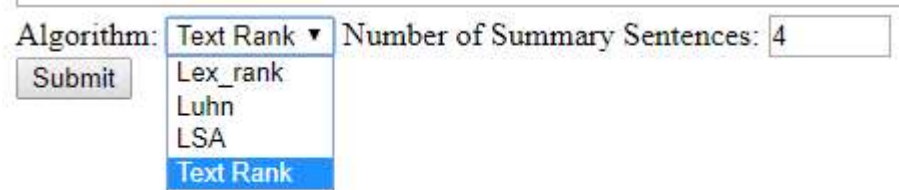
How to use

1. Open browser(IE ,Chrome or others)
2. Type the URL: <http://localhost:5000>, you will see the following screen.

Text Summarization Tool

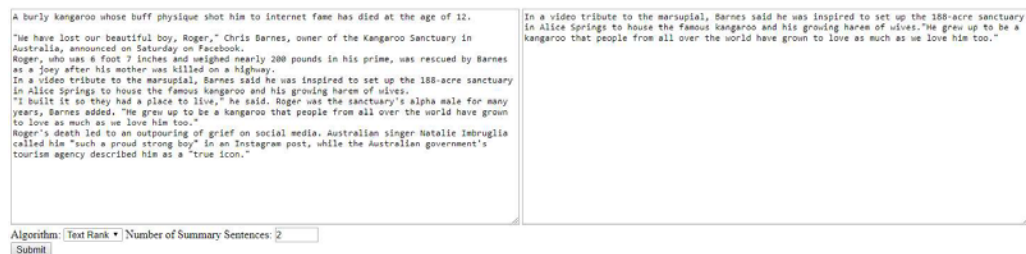


3. The left hand side Text Area with the “Enter plain text here ...” is the input, and right hand side Text Area with the “Summary here...” is the output.
4. The software supports 4 algorithms, such as Lex Rank, Luhn, LSA, and Text Rank.



5. You have option to change the number of summary sentences return, the default is 4 sentences.
6. Once you put the original text into the input, select the Algorithm and Number of Summary Sentences, click the Submit, the result will display on Summary text box.

Text Summarization Tool



7. Done.

Question

If you have any question or feedback, please send to zm11@illinois.edu

REFERENCES

ⁱ H.P.Luhn 'The Automatic Creation of Literature Abstracts, 1958

ⁱⁱ Mihalcea, R. and Tarau, TextRank: Bringing order into text, 2004

ⁱⁱⁱ Erkan, G. and Radev, D.R. Lexrank: Graph-based lexical centrality as salience in text summariazation, 2004