In the name of Allah

Natural Language Processing   Instructor: Dr. Rasekh

Assignment	**Due Date: 1402/10/30**

1. Find the web address, a *Wikipedia* site and download the page

2. Extract the content of the page and copy it in a file called "*TextMining*"

3. Launch preprocessing (such as removing stop words, tokenize, stemming, parse tree, POS tagging, NER and etc.)

- You are free to use the *NLP Stanford* or *knime* tools

4. Use WordNet

   a. Finding the synonym of each word.

   b. Finding hypernym and hyponym relations in word sequences.

   c. Finding the semantic distance between two words.

5. Compute similarity between two words

Note that you are asked to

➢ Prepare a report in PDF format about your program.

➢ Submit your answer before the deadline (1402/10/30).

➢ Zip all files and submit as TM-HW-(Your Student Number).zip

Hint

If you have any question about the assignment, please email to rasekh@zand.ac.ir.