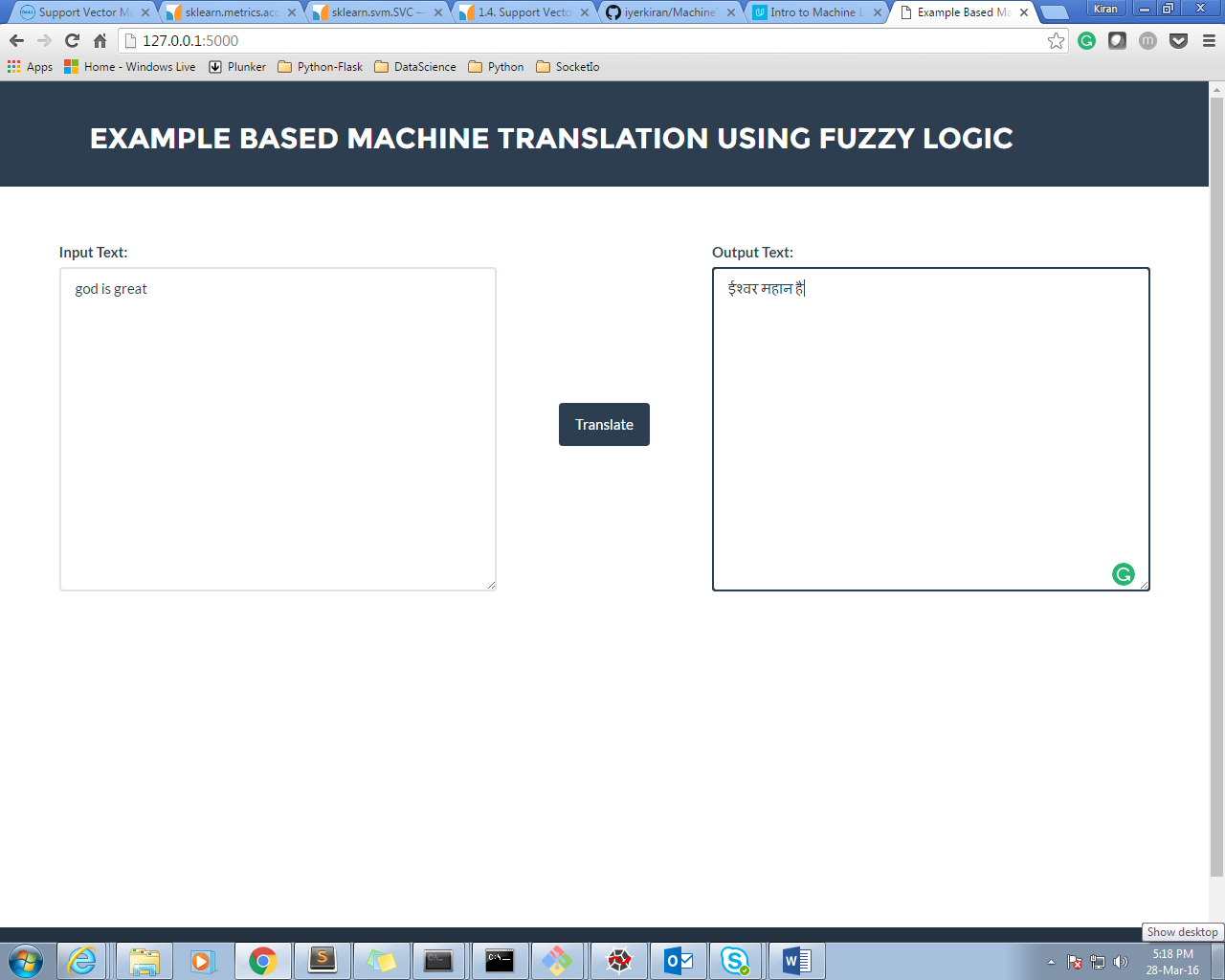
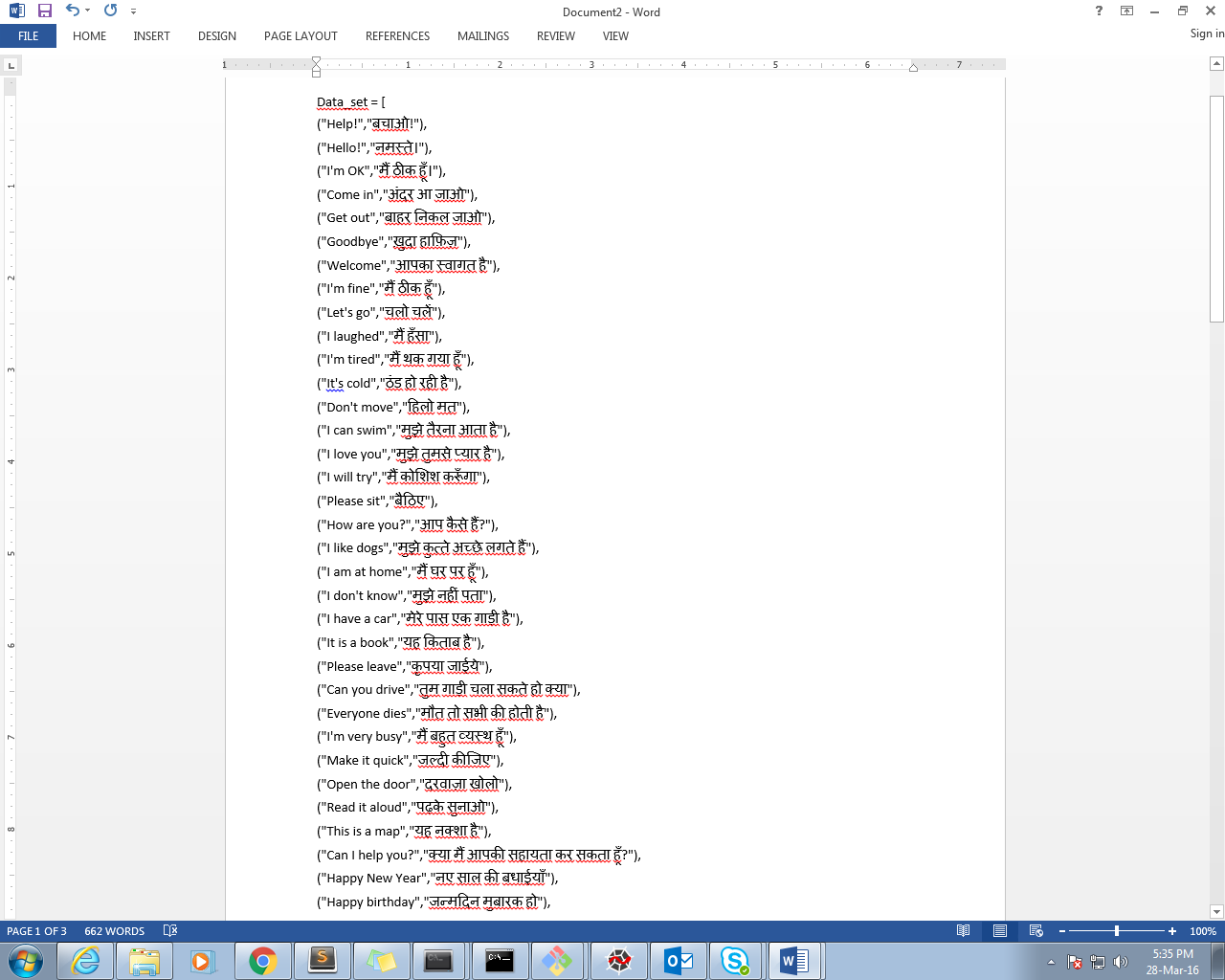
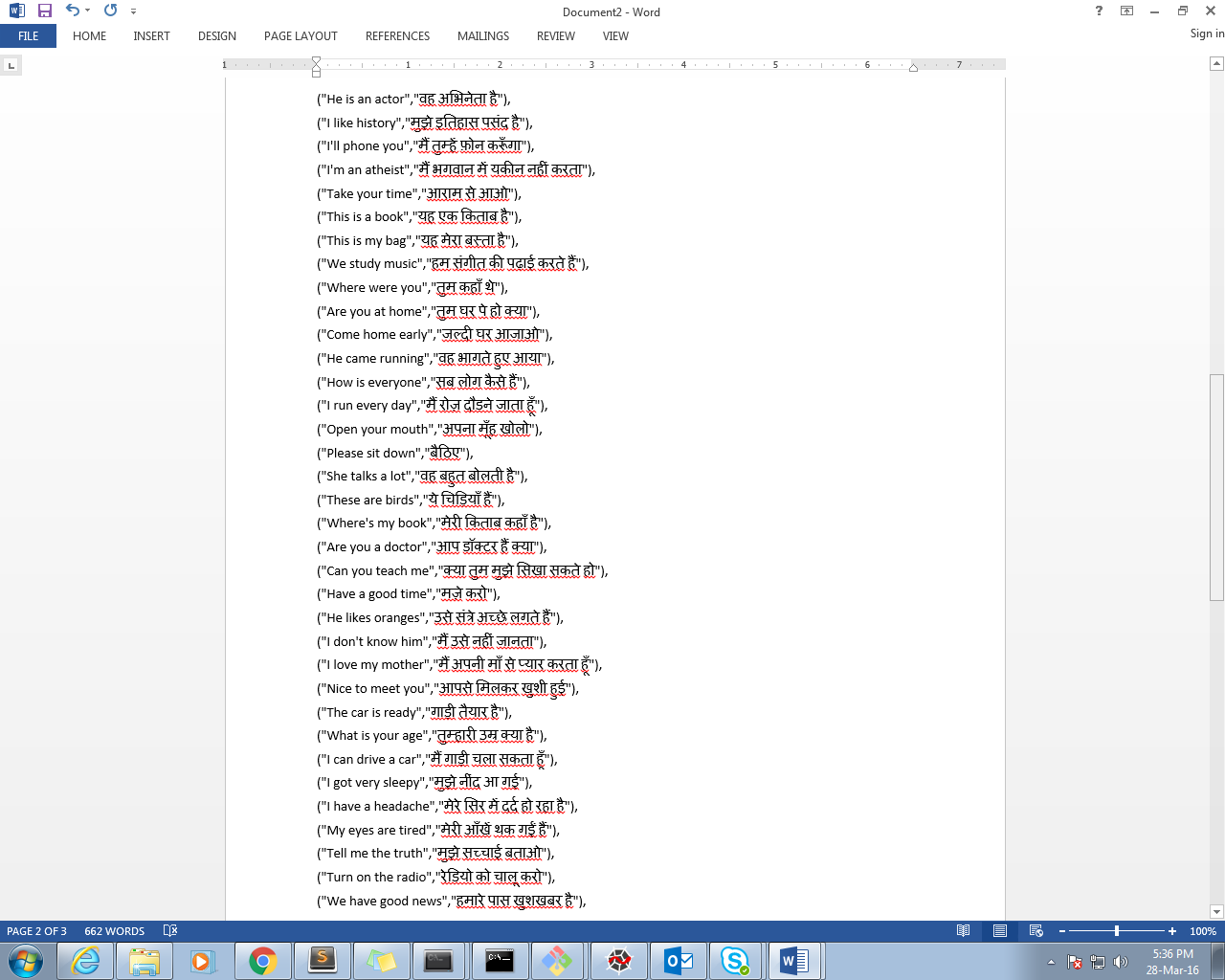
UI snapshot



Data\_set snapshot 1



Data\_set snapshot 2

**Referred following links for knowledge:**

http://www.nltk.org/

http://scikit-learn.org/stable/index.html

http://worldcomp-proceedings.com/proc/p2015/ICA2489.pdf

http://www.academia.edu/12878577/Example\_Based\_Machine\_Translation\_Using\_Various\_Soft-Computing\_Techniques\_ReviewExample\_Based\_Machine\_Translation\_Using\_Various\_Soft-Computing\_Techniques\_Review

http://pybrain.org/

**Python Packages used:**

1. fuzzy-wuzzy

2. nltk (Natural Language Text Processing Toolkit)

**Approach: How it is done technically**

Pre-processing:

A. Prepare a training dataset containing english equivalent hindi sentences.   
B. Initial run to train the application with the training set data.   
C. Minimum 1 million datasets required to achieve an efficiency of 80% and above. So we have used some manual dataset as well as from Microsoft's vast collection of corpora.

Working process:

1. Capture the input text from the user through front end.   
2. Now tokenize the sentence using Natural Language Processing  
3. Carry out POS Tagging and label them accordingly  
4. Replace all the nouns and adjectives to hindi directly  
5. Now generate sentences with possible placements of prepositions in hindi. Now these sentences will be compared for the grammatical precision with our datasets which contain predefined sentences. Now by matching the grammar patterns of our sentence and the dataset a probability ratio is generated. Out of all the sentences the one having highest ratio is selected and shown as output. 