**Weekly Assignment – 5**

Saisrikar Paruchuri

44751575

INFS7410

**Weekly Assignment – 5**

**Query based Stemming**

Stemming is an application of natural language processing that is applied at indexing time to reduce morphological variants to a common root form to understand and cluster related words which are similar in meaning and retrieve relevant documents to improve efficiency of search engine.

**Advantages**

The advantage of query based stemming is it will collect all the related words that is locate, location etc into single common root to get the related documents even if there is a small difference in words it will eliminate enormous difference between query and retrieved document.

It will reduce the vocabulary space and improve the size of the index and dictionary based or rule based stemmers are very fast.

**Disadvantages**

The implementation for stemming are rule based also not all the words are stemmed to their right root word. For example, while stemming there is a lot of ambiguity which may cause several false positives, axes is both a plural form of axe and axis but you won’t be able to relate x axis in a plane with x-y axes. If the query is longer,

stemming process will take long time to process the query. The impact of response time will depend on the degree of query expansion.