**Weekly Assignment - 4**

**Name**  : Saisrikar Paruchuri

**Student Id** : 44751575

**Course Id**  : INFS7410

# **Weekly Assignment-4**

**Information to be Indexed other than Frequency and Word Location to Improve the Efficiency of Search Result**

In the process of indexing and query processing tasks, it is important to concentrate on providing efficient search that is document groups that is clusters or categories. In complicated situations cluster based retrieval first clusters documents that are more likely the given query are determined and then the documents from these clusters are selected to form the final rank outputs. For efficient cluster based retrieval we first identify some alternative query processing strategies then we introduce a new index organization called cluster-skipping inverted index structure (CS-IIS). It shows that typical cluster based retrieval with cluster skipping inverted index structure outperforms previous cluster based retrieval strategies with an ordinary index for the datasets and under varying search parameters. In the enhanced version of cluster skipping inverted index structure in which all the information to compute query cluster similarities during query evaluation is stored. An incremented cluster based retrieval strategy that operates on top of this latter index structure and demonstrate its search efficiency for different scenarios.

We explore query views that are obtained from the search engine query logs to tailor more effective static pruning techniques also related to the indexing task involved in search engine. Query view approach is incorporated into a set of existing pruning strategies. The query view based strategies significantly outperform the existing approaches in terms of query output quality, for both disjunctive and conjunctive evaluation of queries.