

2CSDE53 IRS List of Practicals (Tentative)

Sr No	Task	Link
1	Getting familiar with Kaggle platform and implementing MNIST digits classification.	
2	Text Preprocessing using NLTK. Visualization <ul style="list-style-type: none"> - Word Cloud - Histogram of top N frequent terms 	https://www.kaggle.com/sudalairajkumar/getting-started-with-text-preprocessing https://www.kaggle.com/anandhu/word-cloud-in-python-for-beginners https://www.kaggle.com/aashita/word-clouds-of-various-shapes Datasets which may be used : https://www.kaggle.com/c/nlp-getting-started/data?select=train.csv
3	Vector Space model (TFIDF) <ul style="list-style-type: none"> • Countvectorizer 	https://www.kaggle.com/adamschroeder/countvectorizer-tfidfvectorizer-predict-comments
4	Implementing similarity measures	https://www.kaggle.com/antriksh5235/cosine-similarity-using-tfidf-weighting/notebook
5	Performance Evaluation of IR systems	
6	Language models	http://nlp-iiith.vlabs.ac.in/
7	Text Classification	
8	Meta search	
9	Mini search engine using GUI	
10	Information Retrieval from multimedia data	

