

## 19CSE453 Natural Language Processing Assignment 3

### POS and NER

1. Read a sentence of your choice (eg: "At eight o'clock on Thursday morning  
Arthur didn't feel very good).

2. Tokenise the sentence.
3. Find the n-grams : bi-grams and tri-grams
4. Find the POS tag of each token.
5. Identify the named entities (use `maxent_ne_chunker`)
6. From treebank (treebank.tagged sentences) print a sentence and corresponding tags.
7. How many unique tags are present in the corpus?
8. Find the most commonly occurring tag in the corpus.
9. Which tag is most frequently assigned to the word “bank”?
10. Implement an HMM pos tagger in python.

1.