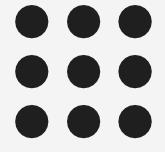# SNS COLLEGE OF ENGINEERING

Department of Artificial Intelligence and Data Science

Course Name – 19AD601 – Natural Language Processing

III Year / VI Semester

Unit 2 – WORD LEVEL ANALYSIS

Topic 8- Wordclass Pos Tagging

Word Class Pos Tagging

Parts of speech fall into two broad categories:
– closed class and
– open class.

Closed classes are those with relatively fixed membership, such as prepositions, new prepositions are rarely coined. By contrast, nouns and verbs are open classes.

Closed class words are generally function words like of, it, and, or you, which tend to be very short, occur frequently, and often have structuring uses in grammar.

Four major open classes occur in the languages of the world: nouns (including proper nouns), verbs, adjectives, and adverbs, as well as the smaller open class of interjections.

Nouns
Nouns are words for people, places, or things, but include others as well.
Nouns can be divided as:

☐   proper nouns -- names for specific entities such as Ankara, John, Ali. Proper nouns do not take an article but common nouns may take.

☐   common nouns – include concrete terms like cat and mango, abstractions like algorithm and beauty, and verb-like terms like pacing.

Verb
Verbs refer to actions and processes, including main verbs like draw, provide, and go.

English verbs have inflections (non-third-person-singular (eat), third-person-singular (eats), progressive (eating), past participle (eaten)).

Adjective
Adjectives often describe properties or qualities of nouns, like color (white, black), age (old, young), and value (good, bad), but there are languages without adjectives.

Word Class Pos Tagging

Interjection

Interjections (oh, hey, alas, uh, um) are a smaller open class that also includes greetings (hello, goodbye) and question responses (yes, no, uh-huh).

Preposition

English ad-positions occur before nouns, hence are called prepositions. They can indicate spatial or temporal relations, whether literal (on it, before then, by the house) or metaphorical (on time, with gusto, beside herself), and relations like marking the agent in Hamlet was written by Shakespeare.

Conjunctions

Conjunctions join two phrases, clauses, or sentences. Coordinating conjunctions like and, or, and but join two elements of equal status. Subordinating conjunctions are used when one of the elements has some embedded status.

Pos Tagging

Part of Speech

The part of speech explains how a word is used in a sentence. There are eight main parts of speech - nouns, pronouns, adjectives, verbs, adverbs, prepositions, conjunctions and interjections.

Parts of speech (also known as POS) and named entities are useful clues to sentence structure and meaning. Knowing whether a word is a noun or a verb tells us about likely neighboring words and syntactic structure making part-of-speech tagging a key aspect of parsing.

Part-of-Speech Tagsets

There are various tag sets to choose. The choice of the tag set depends on the nature of the application. We may use small tag set (more general tags) or large tag set (finer tags).

Some of widely used part-of-speech tag sets:

- ⬚   Penn Treebank has 48 tags
- ⬚   Brown Corpus has 87 tags
- ⬚   C7 tag set has 146 tags

Pos Tagging

Part of Speech

The part of speech explains how a word is used in a sentence. There are eight main parts of speech - nouns, pronouns, adjectives, verbs, adverbs, prepositions, conjunctions and interjections.

Parts of speech (also known as POS) and named entities are useful clues to sentence structure and meaning. Knowing whether a word is a noun or a verb tells us about likely neighboring words and syntactic structure making part-of-speech tagging a key aspect of parsing.

Part-of-Speech Tagsets

There are various tag sets to choose. The choice of the tag set depends on the nature of the application. We may use small tag set (more general tags) or large tag set (finer tags).

Some of widely used part-of-speech tag sets:

- Penn Treebank has 48 tags
- Brown Corpus has 87 tags
- C7 tag set has 146 tags

THANK YOU