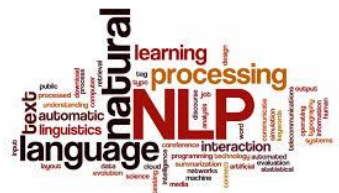


# Natural Language Processing

---

Text Summarization



# Agenda

---

- Text summarization
- Need for Text summarization
- Methods to extract Summary

# Text summarization

---

- Text Summarization is the process of creating a condensed form of text document which maintains significant information and general meaning of source text.
- Automatic text summarization becomes an important way of finding relevant information precisely in large text in a short time with little efforts.
- convey the key information in the input

# Automatic Text summarization - Need

---

- Summaries reduce reading time.
- When researching documents, summaries make the selection process easier.
- Automatic summarization improves the effectiveness of indexing.
- Automatic summarization algorithms are less biased than human summarizers.
- Personalized summaries are useful in question-answering systems as they provide personalized information.
- Using automatic or semi-automatic summarization systems enables commercial abstract services to increase the number of texts they are able to process.

# Examples of Text Summarization

---

- headlines (from around the world)
- outlines (notes for students)
- minutes (of a meeting)
- previews (of movies)
- synopses (soap opera listings)
- reviews (of a book, CD, movie, etc.)
- digests (TV guide)
- biography (resumes, obituaries)
- abridgments (Shakespeare for children)
- bulletins (weather forecasts/stock market reports)

# 8 Parts of Speech

---

- **Noun**
  - name of a person or an animal, a thing, a place or an idea
  - Nouns can be further classified as singular and plural Nouns.
- **Pronoun**
  - replaced by a noun. We , He, They
- **Verb**
  - Shows actions or states of being. Run , talk
- **Adjective**
  - describes or modifies a Noun or a Pronoun. A Red apple.
- **Adverb**
  - describes or modifies a verb , adjective. very
- **Conjunction**
  - Join . and
- **Interjection**
  - Excitement or emotion . Wow, oh
- **Preposition**
  - Relation between noun / pronoun and rest of sentence . In on

# Parts of Speech



## NOUN

*Name of a thing, a person, an animal, a place, or an idea.*

**Examples:** Daniel, London, table, hope  
- *Mary* uses a blue *pen* for her *letters*.

## PRONOUN

*A pronoun is used in place of a noun or noun phrase to avoid repetition.*

**Examples:** I, you, it, we, us, them, those  
- I want *her* to dance with *me*.

## ADJECTIVE

*Describes, modifies or gives more information about a noun or pronoun.*

**Examples:** cold, happy, young, two, fun  
- The *little* girl has a *pink* hat.

## VERB

*Shows an action or a state of being. It can show what someone is doing or did.*

**Examples:** go, speaking, lived, been, is  
- I *listen* to the word and then *repeat* it.

## ADVERB

*Modifies a verb, an adjective or another adverb. It tells how (often), where, when.*

**Examples:** slowly, very, always, well, too  
- *Yesterday*, I ate my lunch *quickly*.

## PREPOSITION

*Shows the relationship of a noun, noun phrase or pronoun to another word.*

**Examples:** at, on, in, from, with, about  
- I left my keys *on* the table *for* you.

## CONJUNCTION

*Joins two words, ideas, phrases together and shows how they are connected.*

**Examples:** and, or, but, because, until, if  
- I was hot *and* tired *but* I still finished it.

## INTERJECTION

*A word or phrase that expresses a strong emotion. It is a short exclamation.*

**Examples:** Ouch! Hey! Wow! Oh! Ugh!  
- *Wow!* I passed my English exam.

© Woodward English

# Features For Extractive Text Summarization

---

- Title word feature:
  - Sentences containing words that appear in the title are also indicative of the theme of the document. These sentences are having greater chances for including in summary.
- Sentence location feature:
  - Usually first and last sentence of first and last paragraph of a text document are more important and are having greater chances to be included in summary.
- Sentence Length feature:
  - Very large and very short sentences are usually not included in summary.



# Features For Extractive Text Summarization

---

- Proper Noun feature:
  - Proper noun is name of a person, place and concept etc. Sentences containing proper nouns are having greater chances for including in summary.
- Upper-case word feature:
  - Sentences containing acronyms or proper names are included.
- Cue-Phrase Feature:
  - Sentences containing any cue phrase (e.g. “in conclusion”, “this letter”, “this report”, “summary”, “argue”, “purpose”, “develop”, “attempt” etc.) are most likely to be in summaries.

# Features For Extractive Text Summarization

---

- **Biased Word Feature:**
  - If a word appearing in a sentence is from biased word list, then that sentence is important. Biased word list is previously defined and may contain domain specific words.
- **Font based feature:**
  - Sentences containing words appearing in upper case, bold, italics or Underlined fonts are usually more important.
- **Pronouns:**
  - Pronouns such as “she, they, it” cannot be included in summary unless they are expanded into corresponding nouns.