

## Semantic Analysis

### # What is Semantic Analysis?

=> As a human being making sense of text is very simple. We recognize every word and the context in which it is used.

- For example if you read this sentence:  
"Your teaching is very good, I scored full marks in my viva."

- So here you understand that student is very happy with the teaching skills of the teacher because he scored full marks in his viva.

- However machines first need to be trained to make sense of the human language and understand the context in which words are used. Otherwise they might misinterpret the word "good" as negative.

- With the power of Machine Learning algorithms and Natural Language Processing, Semantic analysis systems can understand the context of natural language, detect emotions and sarcasm and extract valuable information from unstructured data, achieving human-level accuracy.

- In a very simple terms we can say Semantic Analysis is the process of finding out meaning from the text.
- The word Semantics means the study of meaning. Semantic Analysis allows computers to understand and interpret sentences, paragraphs or whole documents, by analyzing their grammatical structure, and by identifying relationships between individual words in a particular context.
- In other or technical terms we can say that Semantic analysis is the process whereby meaning representations are composed and assigned to linguistic inputs.



# # Applications of Semantic analysis

## 1. Information Extraction

Information Extraction is a precise process of extracting information from unstructured textual sources to enable finding entities as well as classifying and storing them in the database.

## 2. Text Summarization

Text Summarization can automatically shorten longer texts and extract summaries of sections of text without losing the message.

## 3. Information Retrieval System

Information Retrieval System is the process of tracing and recovering specific information from stored data.

## 4. Machine Translation

Machine Translation is the process of translating one source language or text into another language and it is one of the most important applications.

## 5. Expert system

Expert System is a system which gives you expert advice.  
For example: Grammarly

## # Wordnet

- WordNet is a huge lexical database of English.
- Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms, each expressing a distinct concept.
- These Synonyms are called as Synsets and they are interlinked by means of conceptual-semantic and lexical relations.
- The resulting network of meaningfully related words and concepts can be navigated with the browsers.
- WordNet is freely available on Internet, it is public and anyone can download it and make use of it.
- The structure of WordNet make it a useful tool for computational linguistics and natural language processing (NLP).



## # Structure of wordNet

- The main relation among words in WordNet is Synonymy, as between the words shut and close or car and automobile.
- Synonyms - words that denote the same concept and are interchangeable in many contexts - are grouped into unordered sets (Synsets).
- There are ~~various~~ many Unique forms of different categories in WordNet.

Category	Unique Forms
Noun	117,097
Verb	11,488
Adjective	22,141
Adverb	4,609

- Each of wordnets Synsets is linked to other Synsets by means of a small number of "Conceptual Relations". Additionally, a Synset contains a brief definition ("gloss") and, in most cases, one or more short sentences illustrating the use of the synset members.

- Words forms with several distinct meanings are represented in as many distinct Synsets.

- Thus each form-meaning pair in WordNet is Unique.

## # Applications of WordNet

- WordNet has lot of Applications in IR and NLP (IR - Information Retrieval)  
Some of applications are mentioned below

### 1. Automatic Query Expansion:

WordNet Semantic Relations can be used to expand queries so that the search for a document is not confined to the pattern-matching of query terms, but also covers synonyms.

### 2. Document Summarization: WordNet has found useful application in text summarization. Few approaches utilize information from WordNet to compute lexical chains.

### 3. Concept Identification in Natural Language: WordNet can be used to identify concepts pertaining to a term, to suit them to the full semantic richness and complexity of a given information need.

### 4. Word Sense Disambiguation wordnet.com combines features of a number of the other resources commonly used in disambiguation work. It offers sense definitions of words, identifies synsets of synonyms, defines a number of semantic relations and is freely available. This makes it the best known and most



Used used resource for word sense disambiguation



## # Word Sense Disambiguation

=> Many of times a single word might have multiple meanings.

- Word Sense Disambiguation in Natural Language Processing is defined as the ability to determine which meaning of word is activated by the use of the word in a particular context.

- For example there is a ~~word to~~ sentence

She killed him with a baseball bat.

In this sentence one meaning of ~~bat~~ the word "bat" is it is a mammal and the other meaning of the word "bat" is it is a wooden object.

- Now if we consider this statement the word bat has multiple meaning.

- So the process of identifying the correct meaning of such a word according to the statement in which the word is, is called word sense disambiguation.

- Now word sense disambiguation has different approaches

1. Knowledge based approach or Dictionary approach
2. Supervised and Unsupervised approach
3. Hybrid Approach.

## # Knowledge based Approach

• Requirements:

1. Raw corpora
2. Machine Understandable / readable dictionary eg: Indoword.

## # How Knowledge based Approach works

There is a creation of two bags: Sense bag and a context bag

consider this sentence and the <sup>below</sup> meaning of the word "bank" in the dictionary wordnet <sub>below</sub>

Sentence: The bank can guarantee deposits will eventually cover future tuition costs because it invests in adjustable rate mortgage.



The meaning / senses of the word bank in the wordnet dictionary are

Bank<sup>1</sup> sense: A financial institution that accepts deposits and channels the money into lending activities.

Example: "That bank holds the mortgage on my home."

Bank<sup>2</sup> sense: Sloping land (especially beside a body of water)

Example: "They pulled the canoe up on the bank."

- In Knowledge based approach there are two bags Sense bag and context bag
- Sense bag consists of all the senses of the word bank (because bank is the word with multiple meanings).
- And context bag consists remaining content other than the word "bank" from the sentence.
- Now context bag is compared with each of the sense of the word bank.

- The Sense from which maximum words are matched with the sentence is chosen as the meaning of the word "bank".
- In this case Sense 2 doesn't have any word matching with any word from context bag.
- But in the case of Sense 1 there are two words (Deposit and Mortgage) which matches with words from context bag.
- So ~~the~~ Sense 1 is the meaning of the word bank.



## # Lesk Algorithm

- Lesk Algorithm is a very classical algorithm for word sense disambiguation.  
~~Lesk A~~
- Lesk Algorithm is a knowledge based approach.
- Lesk Algorithm is an classical algorithm that is used for word Sense Disambiguation.

## # Homonyms

- Homonyms are those words that have same spelling and pronunciation but when it comes to their meaning, their meaning is different.

for example:

i) bat (wooden object) .  
vs bat (flying weird mammal)

ii) bank (river bank or riverside)  
vs bank (financial institution)



## # Polysemes

- Polysemes are words with different but related meanings
- A word will become polysemous if it expresses different meanings
- Here the difference between meanings can be subtle ~~and~~ or obvious
- If we examine the origins of words it can help to decide whether a word is polysemous or homonymous. This might need to be done because it is sometimes difficult to ~~decide~~ make out whether a word is polysemous or not because the connection between words can be unclear.

example:

1. ~~He drank a glass of chocolate~~
1. The angry artist sued the newspaper
2. She read <sup>the</sup> newspaper.

## # Synonymy

- ~~to~~ Synonyms are the words with similar meanings.
- Some synonyms donot have exactly the same meaning - there might be a tiny difference.

Example :

1. Beautiful - Gorgeous
2. Mistake - Error
3. Rich - wealthy.

## # Antonymy

- Words with opposite or contrasting meanings ~~are~~ share the relationship called as Antonymy
- There are three types of Antonyms ~~are~~ which are complementary, gradable and relational antonyms.
- Complementary antonyms are pairs of words which have opposite meanings that donot lie on a continuous spectrum  
Example : interior - exterior  
right - wrong  
Exhale - inhale etc.



- Gradable antonym are pairs of words with opposite meanings that lie on a continuous spectrum.

Example if we take age as a continuous spectrum, young and old are two ends of the spectrum.

- Relational antonyms are the pairs of words that refer to a relationship from opposite points of view.

- Example: Patient: doctor  
brother: sister etc.



## # Hypernymy and Hyponymy

### → Hyponymy

It is a sense which is a subclass of another sense

- i) Car is a hyponym of vehicle
- ii) dog is a hyponym of animal
- iii) Strawberry is a hyponym of fruit

### → Hypernym is a sense which is a superclass

- i) animal is a hypernym of dog
- ii) fruit is a hypernym of strawberry
- iii) vehical is a hypernym of car

### → Hyponymy is a transitive relation, if a is hyponym of b, b is hyponym of c then a is hyponym of c

- for example if violet is hyponym of purple and purple is hyponym of color then violet is hyponym of color.

- A word can be both a hypernym and hyponym: for example purple is a hyponym of color but itself is a hypernym of its shades.

## # Meronymy

- A meronym is a word which represents a constituent part of something
- if  $x$  is a part of  $y$  then  $x$  is meronym of  $y$
- for example coconut is a meronym of coconut tree.