Part 9: Step by Step Guide to Master NLP – Semantic Analysis

CHIRAG GOYAL — June 23, 2021

Advanced   NLP   Project   Python   Technique   Unstructured Data

This article was published as a part of the [Data Science Blogathon](#)

Introduction

This article is part of an ongoing blog series on Natural Language Processing (NLP). In the previous article, we discussed some important tasks of NLP. I hope after reading that article you can understand the power of NLP in Artificial Intelligence. So, in this part of this series, we will start our discussion on Semantic analysis, which is a level of the NLP tasks, and see all the important terminologies or concepts in this analysis.

This is part-9 of the blog series on the Step by Step Guide to Natural Language Processing.

Table of Contents

Semantic Analysis

Semantic analysis is the process of finding the meaning from text. This analysis gives the power to computers to understand and interpret sentences, paragraphs, or whole documents, by analyzing their grammatical structure, and identifying the relationships between individual words of the sentence in a particular context.

Therefore, the goal of semantic analysis is to draw exact meaning or dictionary meaning from the text. The work of a semantic analyzer is to check the text for meaningfulness.

As we have already discussed that lexical analysis deals with the meaning of the words, then a question comes to mind:

How is Semantic Analysis different from Lexical Analysis?

Lexical analysis is based on smaller tokens but on the contrary, the semantic analysis focuses on larger chunks.

Since semantic analysis focuses on larger chunks, therefore we can divide the semantic analysis into the following two parts:

Studying the meaning of the Individual Word

It is the first component of semantic analysis in which we study the meaning of individual words. This component is known as lexical semantics.

Studying the combination of Individual Words

In this component, we combined the individual words to provide meaning in sentences.

NOTE:

As we discussed, the most important task of semantic analysis is to find the proper meaning of the sentence.

For Example, consider the following sentence:

```
Sentence: Ram is great
```

In the above sentence, the speaker is talking either about Lord Ram or about a person whose name is Ram. That is why the task to get the proper meaning of the sentence is important.

Sentiment Analysis with Machine Learning

We can do semantic analysis automatically works with the help of machine learning algorithms by feeding semantically enhanced machine learning algorithms with samples of text data, we can train machines to make accurate predictions based on their past results.

While we implement a semantic-based approach for machine learning, there are various sub-tasks involved including

- Word sense disambiguation
- Relationship extraction.

let's discuss each of the above tasks one by one in detail.

Word Sense Disambiguation

As we have discussed that Natural language is ambiguous and polysemic; sometimes, the same word can have different meanings depending on its use in the sentence.

Therefore, in semantic analysis with machine learning, computers use Word Sense Disambiguation to determine which

For Example,

```
Consider the word: Orange
```

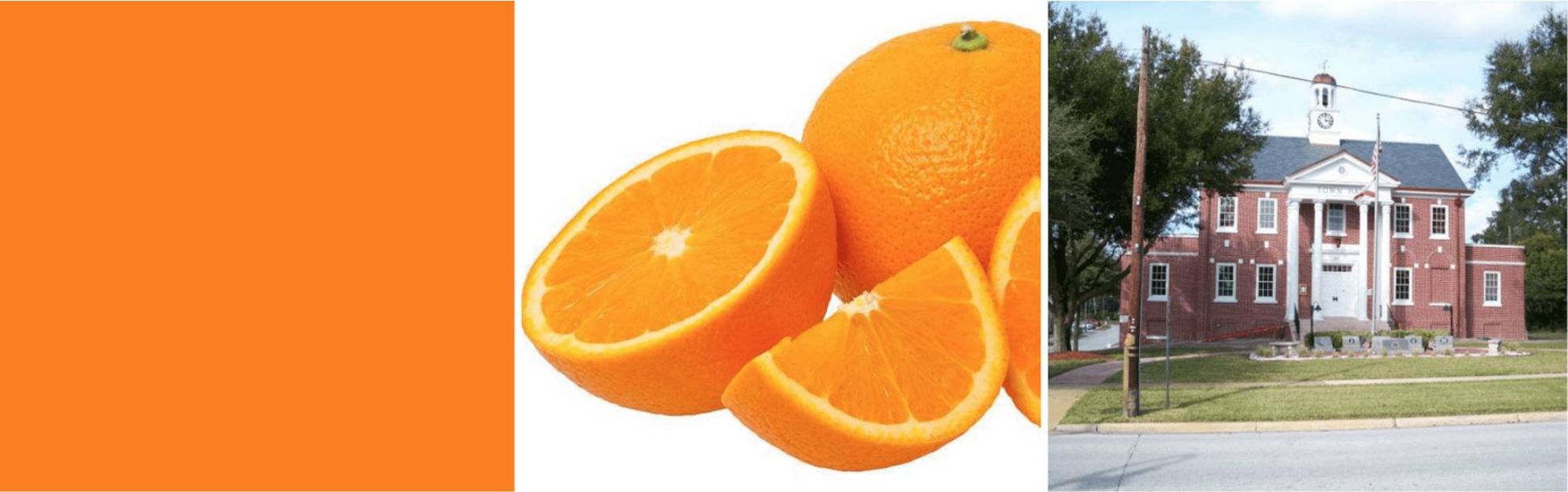The above word can refer to a color, a fruit, or even a city in Florida!

Image Source: Google Images

Relationship Extraction

In this task, we try to detect the semantic relationships present in a text. Usually, relationships involve two or more entities such as names of people, places, company names, etc.

These entities are joined through a semantic category, like "works at," "lives in," "is the CEO of," "headquartered at."

For Example, Consider the following phrase

```
Phrase: Steve Jobs is the founder of Apple, which is headquartered in California
```

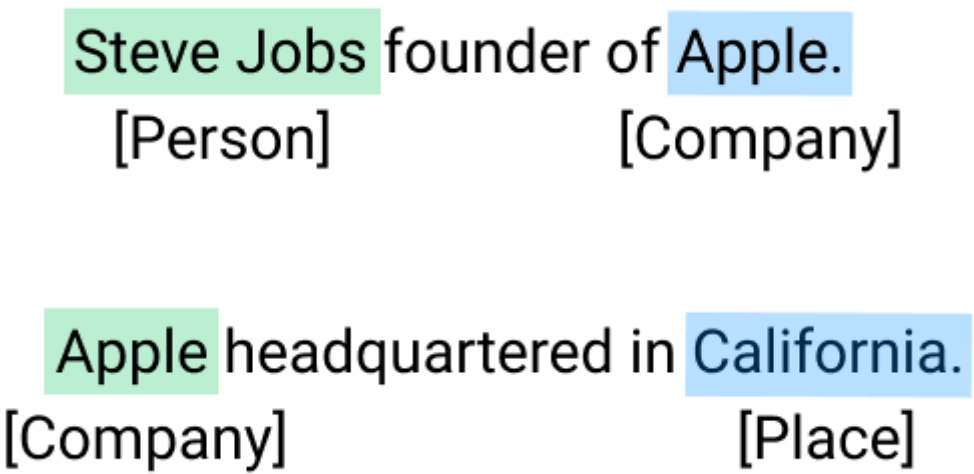The above phrase contains two different relationships:

Image Source: Google Images

Elements of Semantic Analysis

Some important elements of Semantic analysis are as follows:

Hyponymy

It represents the relationship between a generic term and instances of that generic term. Here the generic term is known as

> The word color is hypernym, and the colors blue, yellow, green, etc. are hyponyms.

Homonymy

It may be defined as the words having the same spelling or same form but having different and unrelated meanings.

For example,

> The word "Bat" is a homonymy word.

The above word is homonymy word because a bat can be an implement in two ways:

- To hit a ball,
- Bat is a nocturnal flying mammal also.

Polysemy

Polysemy is a Greek word, that means **"many signs"**. It is a word or phrase with a different but related sense. In other words, we can say that polysemy has the same spelling but different and related meanings.

For Example,

> The word "Bank" is a Polysemy word.

The above word is a polysemy word having the following meanings:

- A financial institution.
- The building in which such an institution is located.
- A synonym for "to rely on".

Difference between Polysemy and Homonymy

Both polysemy and homonymy words have the same syntax or spelling but the main difference between them is that in polysemy, the meanings of the words are related but in homonymy, the meanings of the words are not related.

For Example, if we talk about the same word "Bank" as discussed above, we can write the meaning as

- 'a financial institution' or
- 'a river bank'.

In that case, it becomes an example of a homonym, as the meanings are unrelated to each other.

Synonymy

It represents the relation between two lexical items of different forms but expressing the same or a close meaning.

For Example,

> 'author/writer', 'fate/destiny'

Antonymy

It is the relation between two lexical items having symmetry between their semantic components relative to an axis. The scope of antonymy is as follows –

Application of property or not:

 For Examples,

> 'life/death', 'certitude/incertitude'

For Examples,

```
'rich/poor', 'hot/cold'
```

Application of a usage:

For Examples,

```
'father/son', 'moon/sun'
```

Meronomy

It is defined as the logical arrangement of text and words that denotes a constituent part of or member of something.

For Example,

```
A segment of an orange
```

Meaning Representation

The semantic analysis creates a representation of the meaning of a sentence. But before deep dive into the concept and approaches related to meaning representation, firstly we have to understand the building blocks of the semantic system.

Building Blocks of Semantic System

While representing the meaning of the words, the following building blocks play an important role –

Entities

It represents the individual such as a particular organization, location, people's name, etc.

For Example,

```
Punjab, China, Chirag, Kshitiz all are entities.
```

Concepts

It represents the general category of the individuals such as a person, city, etc.

Relations

It represents the relationship between entities and concepts.

For Example,

```
Sentence: Ram is a person
```

Predicates

It represents the verb structures.

For Example,

```
Semantic roles and Case Grammar
```

Now, we have a brief idea of meaning representation that shows how to put together the building blocks of semantic systems. In other words, it shows how to put together entities, concepts, relations, and predicates to describe a situation. It also enables

The semantic analysis uses the following approaches for the representation of meaning –

- First-order predicate logic (FOPL)
- Semantic Nets
- Frames
- Conceptual dependency (CD)
- Rule-based architecture
- Case Grammar
- Conceptual Graphs

Need of Meaning Representations

The reasons behind the need for the meaning representation are as follows:

Linking of linguistic elements to non-linguistic elements

With the help of meaning representation, we can link linguistic elements to non-linguistic elements.

Representing variety at the lexical level

With the help of meaning representation, we can represent unambiguously, canonical forms at the lexical level.

Can be used for reasoning

The meaning representation can be used to reason for verifying what is correct in the world as well as to extract the knowledge with the help of semantic representation.

Lexical Semantics

It is the first part of semantic analysis, in which we study the meaning of individual words. It involves words, sub-words, affixes (sub-units), compound words, and phrases also. All the words, sub-words, etc. are collectively known as lexical items.

In simple words, we can say that lexical semantics represents the relationship between lexical items, the meaning of sentences, and the syntax of the sentence.

The steps which we have to follow while doing lexical semantics are as follows:

- Classification of lexical items.
- Decomposition of lexical items.
- Differences, as well as similarities between various lexical-semantic structures, are also analyzed.

Techniques of Semantic Analysis

We can any of the below two semantic analysis techniques depending on the type of information you would like to obtain from the given data.

- **text classification model**(which assigns predefined categories to text)
- **text extractor** (which pulls out particular information from the text).

Semantic Classification Models

Topic Classification

Based on the content, this model sorts the text into predefined categories. In a company, Customer service teams may want to classify support tickets as they drop into their help desk, and based on the category it will distribute the work.

Sentiment Analysis

In Sentiment analysis, our aim is to detect the emotions as positive, negative, or neutral in a text to denote urgency.

For Example, Tagging Twitter mentions by sentiment to get a sense of how customers feel about your product and can identify unhappy customers in real-time.

Intent Classification

We can classify the text based on the new user requirement.

You can these types of models to tag sales emails as either "Interested" or "Not Interested" to proactively reach out to those users who may want to try your product.

Semantic Extraction Models

Keyword Extraction

It is used to find relevant words and expressions from a text. This technique is used separately or can be used along with one of the above methods to gain more valuable insights.

For Example, you could analyze the keywords in a bunch of tweets that have been categorized as "negative" and detect which words or topics are mentioned most often.

Entity Extraction

The idea of entity extraction is to identify named entities in text, such as names of people, companies, places, etc.

This might be useful for a customer service team to automatically extract names of products, shipping numbers, emails, and any other relevant data from customer support tickets.

This ends our Part-9 of the Blog Series on Natural Language Processing!

Other Blog Posts by Me

You can also check my previous blog posts.

Previous Data Science Blog posts.

LinkedIn

Here is   my Linkedin profile   in case you want to connect with me. I'll be happy to be connected with you.

Email

For any queries, you can mail me on   Gmail   .

End Notes

Thanks for reading!

I hope that you have enjoyed the article. If you like it, share it with your friends also. Something not mentioned or want to share your thoughts? Feel free to comment below And I'll get back to you. 😉

The media shown in this article are not owned by Analytics Vidhya and are used at the Author's discretion.