A Seminar Report

On

SEMANTIC AND SENTIMENTAL ANALYSIS USING NLP.

By

SHRIRANG RAJENDRA MHALGI.

T150394324

Under the guidance of

Prof. L. A. Bewoor.

DEPARTMENT OF COMPUTER ENGINEERING

VISHWKARMA INSTITUTE OF INFORMATION TECHNOLOGY, Pune

Affiliated to

Savitribai Phule Pune University

[2018-19]

Semantic and Sentimental Analysis using NLP.

Department of Computer Engineering

Vishwakarma Institute of Information Technology,

Pune

CERTIFICATE

This is to certify that **Shrirang Rajendra Mhalgi** from Third Year Computer Engineering has successfully completed his seminar work titled **"Semantic and Sentimental Analysis using NLP"** at Vishwakarma Institute of Information Technology, Pune in partial fulfilment of bachelor's degree in Engineering.

Prof. L. A. Bewoor.   Dr. Prof. S. R. Sakhare.   Dr. Prof. B. Karkare

Guide   Head of Dept.   Director.

Semantic and Sentimental Analysis using NLP.

Abstract :

With the increase in data, there is a need to understand the correct and relevant meaning of the data and understand the underlying hidden sentiments which come out of it. Semantic word spaces have been very useful and can express the meaning of longer phrases in a principled way. Further progress towards understanding compositionality in tasks such as sentiment detection requires richer supervised training and evaluation resources and more powerful models of composition. To capture the effects of negation and its scope at various tree levels for both positive and negative phrases is a hot and trending topic of the 21st century.

Natural Language Processing is a technique which can be used to do accurate analysis of data and its underlying hidden sentiments and is widely used. Although getting 100% accuracy is next to impossible, but scientists have made a progress of achieving a fair amount of high accuracy.

With the help of accurate data, it is much easier to do actual and accurate analysis of data which can be further used by the experts to derive much more concrete results.

Keywords :

Semantics, Meanings, Linguistic, Pragmatic, Sentiments, Emotions, Language Processing, Opinion Mining, Natural language processing.

Semantic and Sentimental Analysis using NLP.

ACKNOWLEDMENTS

Semantic and Sentimental Analysis using NLP.

TABLE OF CONTENTS.

Semantic and Sentimental Analysis using NLP.

LIST OF FIGURES

Semantic and Sentimental Analysis using NLP.

1. INTRODUCTION :

Human language is special for several purposes. It is specifically constructed to convey the speakers / writers meaning. The human brain is very difficult to understand. The capability of the languages that the brain can process is tremendous. In today's world, a normal person knows almost 3 to 4 languages which include English as the universal language, their regional native language and some extra languages. It is strange that not only that the person knows all the languages but also, he is able to understand the hidden sentiments behind it. (5.)According to Chris Manning (Machine Learning Professor at Stanford University), it is a discrete, symbolic, categorical signaling system. This means that we can convey the same meaning by using different ways, like signs, speech, gesture, etc. The encoding of these languages by our human brain is a continuous pattern of activation, where the symbols are transmitted via continuous signals of sound and vision.

Understanding human language is considered a difficult task due to its complexity. For example, there is an infinite number of different ways to arrange words in a sentence. Also, there are several meanings for words and contextual information is necessary to correctly interpret them. On top of it, every language is unique and ambiguous. There might be a case where a same word has different meaning in two different languages.

Just look at the following newspaper headline "The Pope's baby steps on gays". This sentence clearly has two very different interpretations and also two different sentiments, which is a pretty good example of the challenges in NLP.

A perfect understanding of language by a computer would result in an AI that can process the whole information that is available on the internet, which in turn would probably result in artificial general intelligence.

2. NATURAL LANGUAGE :

2 people interacting with each other.

Natural language is like an ordinary language that has evolved naturally in humans because of repetition without conscious planning or premeditation. Natural language means human language. For example, English, Hindi, France, German are natural languages while on the other hand, languages like C, C++, Java, Python are machine languages i.e. they can be understood and executed by the machine, while natural language cannot.

The most challenging problem in computer science is to develop computers that can understand natural languages. As of now, the complete solution of the problem has not been found, but a great progress in field has been achieved by the scientists. Fourth-generation languages which give the result of simple queries are the closest to natural languages. Fourth-generation languages mainly include different Database languages.

3. NATURAL LANGUAGE PROCESSING (NLP) :

3.1 Introduction to NLP :

Natural language processing (NLP) is an area of computer science and artificial intelligence that is concerned with the interaction between computers and humans in natural language. The goal of NLP is to enable computers to understand language as well as we do. It has a wide variety of applications which include chatbots, analysis, report generation, machine translation, and many more. The popular examples include, chatbots like Siri, Alexa, etc. speech recognition software's like google translator, sentimental analyzer like parallel-dots and much more.
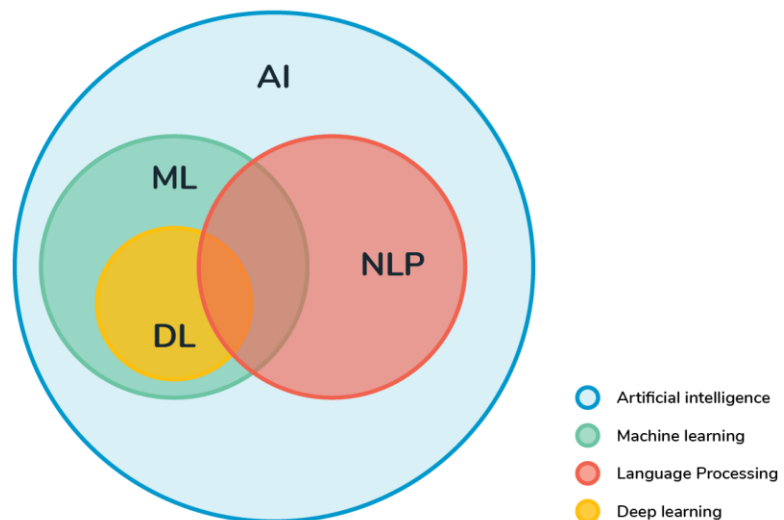
Interaction of NLP with various other domains.

Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) in which there is a little intersection of various AI domains namely Machine Learning (ML), and Deep Learning (DL) as well. On the other hand, it intersects with Linguistics, Pragmatics as well. Thus, it is a complex branch to understand and work upon. (4.)The focus and goal of natural language processing is to make the dumb machine interact with smart people. NLP is used to make the computer intelligent and do communication with humans. Applications of NLP techniques are voice assistants like Alexa and Siri, Machine Translation, sentimental and semantic analysis, POS tagging and text-filtering and many more.

3.2 Types of NLP :

NLP is the field which is benefitted the most as the basic language of medium of communication is language. NLP has benefited from the recent advances in Machine Learning and from Deep Learning techniques. This field is divided into number of parts which are uncountable. But let us try to list a few of them :

1. **Optical Character Recognition :** Converting written or printed text into data.

   E.g. Portable scanner.

2. **Speech Recognition :** Converting spoken words into data.

   E.g. Amazon Alexa, Apple Siri, Google Assistant.

3. **Machine Translation :** Translating text from one language to another.

   E.g. Google Translator, etc.

4. **Natural Language Generation :** Formatting information as a natural language.
   E.g. a weather system that generates a weather forecast in paragraph form.

5. **Sentimental Analysis :** Interpreting basic information from the language such as the topic being discussed and whether comments are positive or negative.

   E.g. Brand-engagement metrics might look at how many people are talking about a brand and how much of that talk is positive.

   Another example can be to analyze users' sentiments for a comment.

6. **Semantic Search :** Understanding and answering questions which are put forward in natural language.

   E.g. accurately processing a search query such as "give me the address of that bar

I went to last weekend."

The output of the above query will depend upon accurate analysis of Parts of Speech. If instead of bar the system returns the restaurant then the user would get annoyed.

7. **Machine Learning :** Using natural language to train artificial intelligence. This can be learning related to language itself or another topic such as economics.

   E.g. Self-driving car based on voice instructions.

8. **Natural Language Programming :** Tools that allow end-users to create or customize computer programs with natural language.

   E.g. "make me an application that suggests new careers paths to me based on people with similar skills and experiences who have successfully improved their salary and quality of my life with a new career."

   Analyzing the query in the first go is sometimes even difficult for humans. But when the machines are trained accurately, a greater amount of accuracy can be achieved.

9. **Affective computing :** Affective computing is the machine equivalent of emotional intelligence. Allows user interfaces to understand human ideas and emotional states and to take on human-like conversations.

   E.g. Crowd-Emotion software.

3.3 Semantic Analysis :

3.3.1 Definition :

       Semantics is all about meaning. It cannot be just defined for language but also can be defined for many different things such as arithmetic expressions.

3.3.2 Semantic Analysis :

Consider the following expression.

(5 + 2) * 7

= 7 * 7

= 49

The parse tree for the above expression can be drawn as follows :
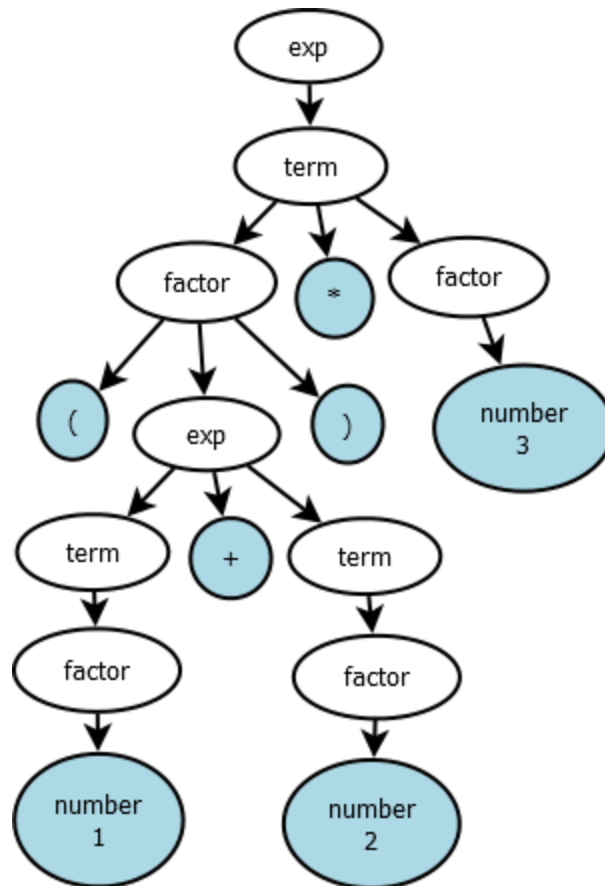
Parse tree for the expression (5 + 2) * 7.

The meaning that can be derived from the above expression is as follows :

The expression is nothing but a multiplication of 2 numbers in which the first number consists of addition of 2 numbers.

We can write it as follows :

temp = add(4, 3)

ans = mult(temp, 7) . . . . (1)

which is equivalent to

ans = mult (add (4, 3), 7) . . . .(2)

The above two equations will produce the same result.

Now let us study semantic analysis in case of words and sentences :
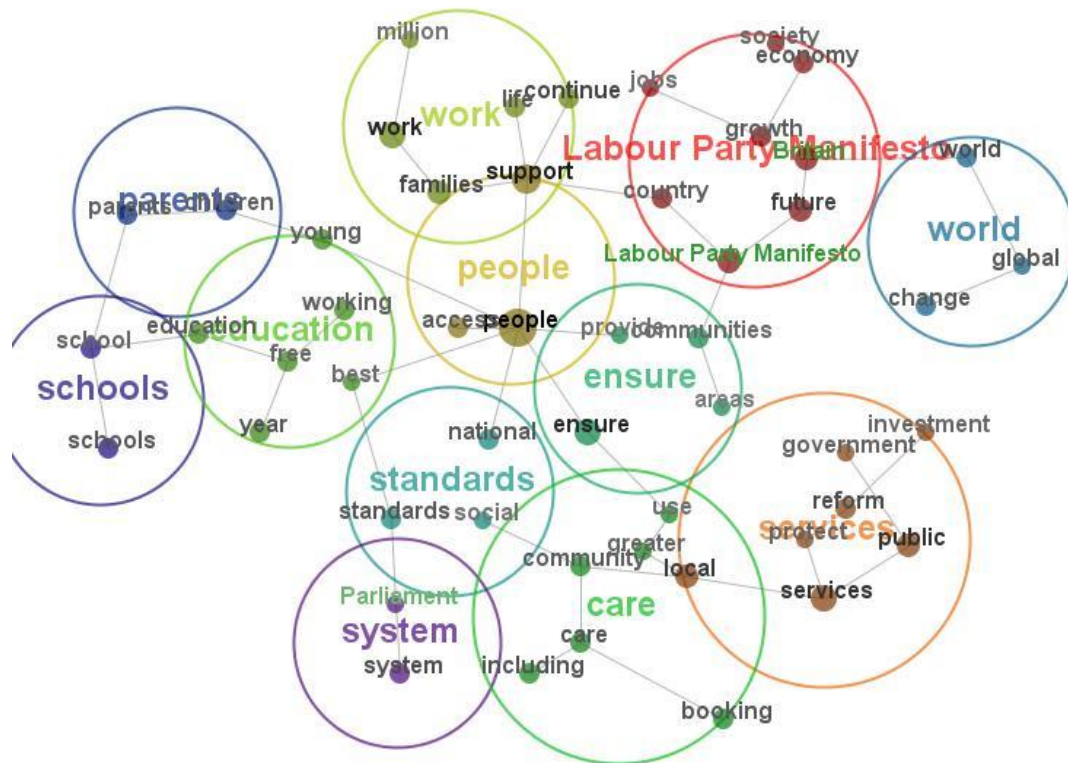
Clusters of words which have similar meanings.

Let us try to understand with a help of an example.

- ***It is 8pm.***

  The above sentence can be plainly interpreted as :

  The time is 8pm.

  This is called as the linguistic meaning of the sentence.

  Linguistic meaning refers to the actual meaning of the sentence and it does not convey any emotions to the audience.

  On the other hand, consider the following sentence :

- ***It is the time to leave***.
  The above sentence does not determine the actual meaning of the sentence but tries to convey hidden emotions behind it.
  This sentence cannot stand by itself but would need help of some underlying sentence.
 If we combine the above example, then we can derive a clear hidden meaning of the sentence along with its emotions
  The time is 8pm. It is the time to leave.
  The above sentence tries to depict the meaning as well as emotions.
  We can derive the following meaning :
  Maybe the speaker in the sentence is going somewhere and the time is past 8. That is the reason why he is saying in anxiety that 'The time is 8pm, It is the time to leave.'
  This is known as the pragmatic meaning of the sentence.
  The word Pragmatic is an adjective which has the meaning as follows :
  Dealing with things sensibly and realistically in a way that is based on practical rather than theoretical considerations.
  Pragmatic qualities build up with experience whereas linguistic qualities can be made concrete by understanding the language.

3.3.3 Types of Semantic Analysis :

(S. Miller et al. 2000)In semantic analysis we assign each word some sort of meaning and combine the meanings to form a phrase, sentence, paragraphs, stories, narration, etc.

The types of semantic analysis are as follows :

i. Entailment :

In entailment one fact follows from the other.

Consider the following example :

"All cats have whiskers."

"Tom is a cat."

The above two sentences derive the conclusion that

"Tom has whiskers."

This can be depicted in terms of mathematical model as follows :

If A is related to B and B is related to C

Then A is related to C as well.

ii. Presupposition :

We may draw some conclusions based on our observations :

Maybe the book is wrapped in the paper.

Here we are not sure whether the book is wrapped in the paper or not. But we are assuming it to be true based on some observations which can be the size, shape, appearance, etc.

iii. Understanding meaning :

If the agent hears the sentence and acts accordingly then it is said that he has understood the meaning. Consider the following example.

Leave the book on the table

The common questions which will arise in everyone's mind are

a. Which book?

b. Which table?

If you know the answer of both then you will act accordingly.

3.3.4 Flowchart :
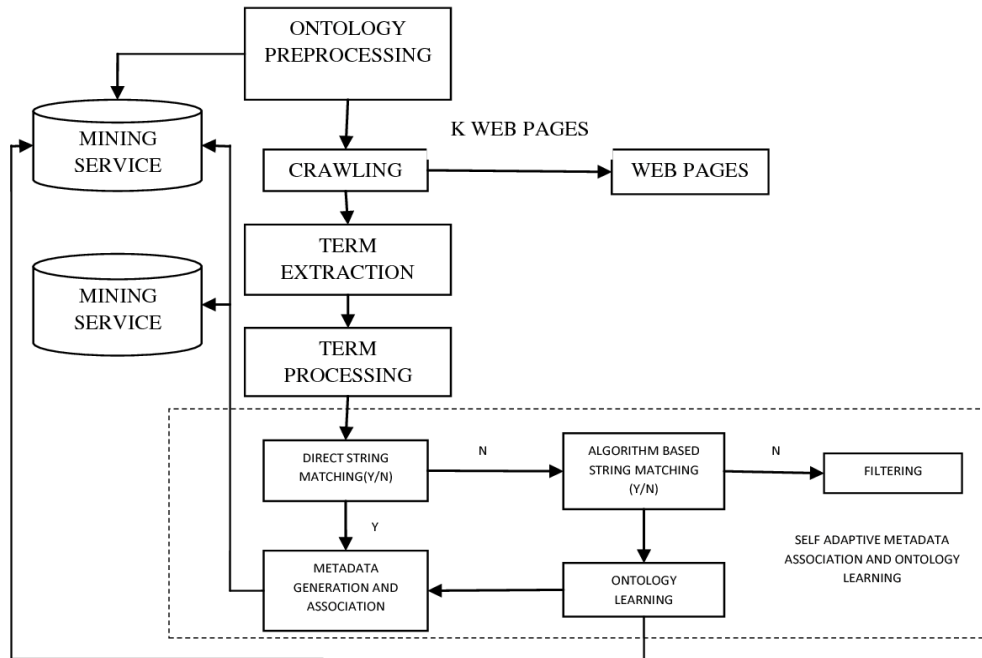
Flowchart for web based semantic analysis.

In the above figure, we can see the flowchart for web based semantic analysis. In this the starting point is ontology preprocessing.

Ontology refers to formal naming and definition of the types, properties, and interrelationships of the entities that really or fundamentally exist for a domain of discourse. It is thus a practical application of philosophical ontology, with a taxonomy. This preprocessing first refers to the database to check whether the word with the same meaning exists or not. If it exists, then it directly shows the output.

If it does not exist, then it makes a request to the web crawler to fetch the information from the www and then generate the links. After getting the required links, the terms are extracted, preprocessed and then semantic analysis is done.

In semantic analysis, direct string matching is performed, and the text is checked. If it the same, then result is shown else an algorithm is performed based on string matching and the relevant meaning is returned. If algorithm is not successful, then that text is filtered out.

3.3.5 Applications :

(1.)Social media, blog posts, comments in forums, documents, group chat applications or dialog with customer service chatbots: Text is at the heart of how we communicate with companies online.

Each type of communication, whether it's a tweet, a post on LinkedIn or a review in the comments section of a website, contains potentially relevant, even valuable information that must be captured and understood by companies who want to stay ahead. Capturing the information isn't the hard part. What's really difficult is understanding what is being said and doing it at scale.

For humans, the way we understand what's being said is almost an unconscious process. To understand what a text is talking about, we rely on what we already know about language itself and about the concepts present in a text. Machines can't rely on these same techniques.

Some technologies only make you think they understand text. An approach based on keywords or statistics, or even pure machine learning, may be using a matching or frequency technique for clues as to what a text is "about." These methods can only go so far because they are not looking at meaning.

Semantic analysis describes the process of understanding natural language–the way that humans communicate–based on meaning and context.

The semantic analysis of natural language content starts by reading all of the words in content to capture the real meaning of any text. It identifies the text elements and assigns them to their logical and grammatical role. It analyzes context in the surrounding text, and it analyzes the text structure to accurately disambiguate the proper meaning of words that have more than one definition.

(3.)Semantic technology processes the logical structure of sentences to identify the most relevant elements in text and understand the topic discussed. It also

understands the relationships between different concepts in the text. For example, it understands that a text is about "politics" and "economics" even if it doesn't contain the actual words but related concepts such as "election," "Democrat," "speaker of the house," or "budget," "tax" or "inflation."

Because semantic analysis and natural language processing can help machines automatically understand text, this supports the even larger goal of translating information–that potentially valuable piece of customer feedback or insight in a tweet or in a customer service log–into the realm of business intelligence for customer support, corporate intelligence or knowledge management.

3.4 Sentimental Analysis :

3.4.1 Definition :

Sentiment Analysis also known as *Opinion Mining* is a field within Natural Language Processing (NLP) that builds systems that try to identify and extract opinions within text.

3.4.2 Sentimental Analysis :

Sentiment Analysis is the most common text classification tool that analyses an incoming message and tells whether the underlying sentiment is positive, negative our neutral. Usually, besides identifying the opinion, these systems extract attributes of the expression e.g.:

- *Polarity*: if the speaker expresses a *positive* or *negative* opinion,
- *Subject*: the thing that is being talked about,
- *Opinion holder*: the person, or entity that expresses the opinion.

Currently, sentiment analysis is a topic of great interest and development since it has many practical applications. Since publicly and privately available information over Internet is constantly growing, a large number of texts expressing opinions are available in review sites, forums, blogs, and social media.

With the help of sentiment analysis systems, this unstructured information could be automatically transformed into structured data of public opinions about products, services, brands, politics, or any topic that people can express opinions about. This data can be very useful for commercial applications like marketing analysis, public relations, product reviews, net promoter scoring, product feedback, and customer service.

3.4.3 Types of Sentimental Analysis :

(S. Miller et al., 2000)There are many types and flavors of sentiment analysis and SA tools range from systems that focus on polarity (positive, negative, neutral) to systems that detect feelings and emotions (*angry*, *happy*, *sad*, etc) or identify intentions (e.g. *interested* v. *not interested*). In the following section, we'll cover the most important ones.

1. **Fine-grained Sentiment Analysis :**

Sometimes you may be also interested in being more precise about the level of polarity of the opinion, so instead of just talking about *positive*, *neutral*, or *negative* opinions you could consider the following categories:

- Very positive
- Positive
- Neutral
- Negative
- Very negative

This is usually referred to as fine-grained sentiment analysis. This could be, for example, mapped onto a 5-star rating in a review, e.g.: Very Positive = 5 stars and Very Negative = 1 star.

Some systems also provide different flavors of polarity by identifying if the positive or negative sentiment is associated with a particular feeling, such as, anger, sadness,

or worries (i.e. negative feelings) or happiness, love, or enthusiasm (i.e. positive feelings).

2. Emotion detection :

Emotion detection aims at detecting emotions like, happiness, frustration, anger, sadness, and the like. Many emotion detection systems resort to lexicons (i.e. lists of words and the emotions they convey) or complex machine learning algorithms.

One of the downsides of resorting to lexicons is that the way people express their emotions varies a lot and so do the lexical items they use. Some words that would typically express anger like *shit* or *kill* (e.g. *in your product is a piece of shit* or *your customer support is killing me*) might also express happiness (e.g. in texts like *This is the shit* or *You are killing it*).

3. Aspect-based Sentiment Analysis :

Usually, when analyzing the sentiment in subjects, for example products, you might be interested in not only whether people are talking with a positive, neutral, or negative polarity about the product, but also which particular aspects or features of the product people talk about. That's what aspect-based sentiment analysis is about. In our previous example:
"The battery life of this camera is too short."
The sentence is expressing a negative opinion about the camera, but more precisely, about the battery life, which is a particular feature of the camera.

4. Intent Analysis :

Intent analysis basically detects what people want to do with a text rather than what people say with that text. Look at the following examples:

"Your customer support is a disaster. I've been on hold for 20 minutes".

"I would like to know how to replace the cartridge".

"Can you help me fill out this form?"

A human being has no problems detecting the complaint in the first text, the question in the second text, and the request in the third text. However, machines can have some problems to identify those. Sometimes, the intended action can be inferred from the text, but sometimes, inferring it requires some contextual knowledge.

5. Multilingual sentiment analysis :

Multilingual sentiment analysis can be a difficult task. Usually, a lot of preprocessing is needed, and that preprocessing makes use of a number of resources. Most of these resources are available online (e.g. sentiment lexicons), but many others must be created (e.g. translated corpora or noise detection algorithms). The use of the resources available requires a lot of coding experience and can take long to implement.

An alternative to that would be detecting language in texts automatically, then train a custom model for the language of your choice (if texts are not written in English), and finally, perform the analysis.

3.4.4 Flowchart :
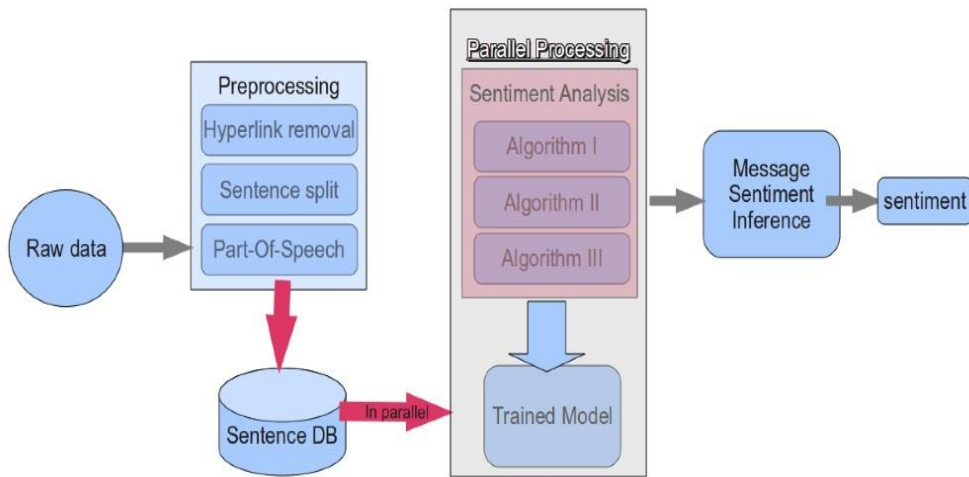
Flowchart for general sentimental analysis.

The above diagram shows the processing of raw data which gives the final output along with the sentiments. The raw data is passed into the preprocessing phase where it does sentence splitting, hypertext removal and parts of speech tagging. Then the sentences are stored into database. After that using the parallel processing algorithms sentimental analysis is done and the intent and emotion of the sentence is found and shown as the output

3.4.5 Applications :

Sentiment analysis has many applications and benefits to your business and organization. It can be used to give your business valuable insights into how people feel about your product brand or service. When applied to social media channels, it can be used to identify spikes in sentiment, thereby allowing you to identify potential product advocates or social media influencers.

The Role of Sentiment Analysis in Business :

The applications of sentiment analysis in business cannot be overlooked. Sentiment analysis in business can prove a breakthrough for the complete brand revitalization. The key to running a successful business with the sentiments data is the ability to exploit the unstructured data for actionable insights. Machine learning models, which largely depend on the manually created features before classification, have served this purpose fine for the past few years. However, deep learning is a better choice as it:

- Automatically extracts the relevant features.
- Helps to scrape off the redundant features.
- Rules out the efforts of manually crafting the features.

Sentiment Analysis in Business Intelligence Buildup:

(P. Koomen et al. 2005) Having insights-rich information eliminates the guesswork and execution of timely decisions. With the sentiment data about your established and the new products, it's easier to estimate your customer retention rate. Based on the reviews generated through sentiment analysis in business, you can always adjust to the present market situation and satisfy your customers in a better way. Overall, you can make immediate decisions with automated insights. Business intelligence is all about staying dynamic throughout. Having the sentiments data gives you that liberty. If you develop a big idea, you can test it before bringing life to it. This is known as concept testing. Whether it is a new

product, campaign or a new logo, just put it to concept testing and analyze the sentiments attached to it.

Sentiment Analysis in Business for Competitive Advantage:

If you are truly catching up with the applications of sentiment analysis in business, you should be open to experimenting with it tactfully. Like I mentioned before, sentiment analysis can be performed on any piece of text. So, why just settle for applying it to your brand? Getting x% negative or positive reviews on a certain product doesn't make much sense if you don't have a y% metric to compare it with. Knowing the sentiment data of your competitors gives you the opportunity as well as the incentive to perk up your performance. Sentiment analysis in businesses can be very helpful in predicting the customer trends. Once you get acquainted with the current customer trends, strategies can easily be developed to capitalize on them. And eventually, gain a leading edge in the competition.

Enhancing the Customer Experience through Sentiment Analysis in Business:

A business breathes on the gratification of its customers. The experience of the customers can either be positive, negative or neutral. Owing to the internet savvy era, this experience becomes the text of their social posting and online feedback. The tone and temperament of this data can be detected and then categorized according to the sentiments attached. This helps to know what is being properly implemented with regards to products, services and customer support and what needs improvement.

Getting a positive response to your product is not always enough. The customer support system of your company should always be impeccable no matter how phenomenal your services are.

Example :
48 hours locked out of @Snapchat and @snapchatsupport has still not responded. This is very disappointing. #customersupport #nosupport

Semantic and Sentimental Analysis using NLP.

— Chase Lepard (@chaselepard) December 10, 2016

Sentiment Analysis in Business for Brand Brisking:

A brand is not defined by the product it manufactures or the services it provides. The name and fame that build a brand majorly depend on their online marketing, social campaigning, content marketing and customer support services. Sentiment analysis in business helps in quantifying the perception of the present and the potential customers regarding all these factors. Keeping the negative sentiments in knowledge, you can develop more appealing branding techniques and marketing strategies to switch from torpid to terrific brand status. Sentiment analysis in business can majorly help you to make a quick transition.

The applications of sentiment analysis in business are plenty and overwhelming. Gaining a greater business value with sentiment analysis depends on what tool you use and how well you use it to your advantage.

Semantic and Sentimental Analysis using NLP.

4. Conclusion :

Thus, we have understood the meanings and applications of semantic analysis as well as sentimental analysis using the power of Natural Language Processing.

Semantic and Sentimental Analysis using NLP.

5. References

[1.] https://en.wikipedia.org/wiki/Natural_language

[2.] https://www.expertsystem.com/natural-language-process-semantic-analysis-definition/

[3.] https://simplicable.com/new/natural-language-processing

[4.] https://monkeylearn.com/sentiment-analysis/

[5.] https://www.paralleldots.com/sentiment-analysis

[6.] S. Miller, H. Fox, L. Ramshaw, and R. Weischedel. A novel use of statistical parsing to extract information from text. Applied Natural Language Processing Conference (ANLP), 2000.

[7.] S. Miller, J. Guinness, and A. Zamanian. Name tagging with word clusters and discriminative training. In Conference of the North American Chapter of the Association for Computational Linguistics & Human Language Technologies (NAACL-HLT), pages 337–342, 2004.

[8.] P. Koomen, V. Punyakanok, D. Roth, and W. Yih. Generalized inference with multiple semantic role labeling systems (shared task paper). In Conference on Computational Natural Language Learning (CoNLL), pages 181–184, 2005.

[9.] A Mnih and G. E. Hinton. Three new graphical models for statistical language modelling. In International Conference on Machine Learning (ICML), pages 641–648, 2007.

Semantic and Sentimental Analysis using NLP.

6. Plagiarism Scan Report

SmallSEOTools

PLAGIARISM SCAN REPORT

Words	420	Date	March 31,2019
Characters	2228	Exclude Url	

9% Plagiarism	91% Unique	2 Plagiarized Sentences	20 Unique Sentences

Content Checked For Plagiarism

Semantic Analysis : Definition : Semantics is all about meaning. It cannot be just defined for language but also can be defined for many different things such as arithmetic expressions. Semantic Analysis : Consider the following expression. (5 + 2) * 7 = 7 * 7 = 49 The parse tree for the above expression can be drawn as follows : The meaning that can be derived from the above expression is as follows : The expression is nothing but a multiplication of 2 numbers in which the first number consists of addition of 2 numbers. We can write it as follows : temp = add(4, 3) ans = mult(temp, 7) . . . . (1) which is equivalent to ans = mult (add (4, 3), 7) . . . .(2) The above two equations will produce the same result. Now let us study semantic analysis in case of words and sentences : Let us try to understand with a help of an example. • It is 8pm. The above sentence can be plainly interpreted as : The time is 8pm. This is called as the linguistic meaning of the sentence. Linguistic meaning refers to the actual meaning of the sentence and it does not convey any emotions to the audience. On the other hand, consider the following sentence : • It is the time to leave. The above sentence does not determine the actual meaning of the sentence but tries to convey hidden emotions behind it. This sentence cannot stand by itself but would need help of some underlying sentence. If we combine the above example, then we can derive a clear hidden meaning of the sentence along with its emotions The time is 8pm. It is the time to leave. The above sentence tries to depict the meaning as well as emotions. We can derive the following meaning : Maybe the speaker in the sentence is going somewhere and the time is past 8. That is the reason why he is saying in anxiety that 'The time is 8pm, It is the time to leave.' This is known as the pragmatic meaning of the sentence. The word Pragmatic is an adjective which has the meaning as follows : Dealing with things sensibly and realistically in a way that is based on practical rather than theoretical considerations. Pragmatic qualities build up with experience whereas linguistic qualities can be made concrete by understanding the language.

Sources	Similarity
1+1 (фильм) — ВикипедияCompare text «1+1» (фр. intouchables — «Неприкасаемые») — французская трагикомедия 2011 года, основанная на реальных событиях об успешном аристократе Филиппе, который в результате несчастного случая оказывается в инвалидном кресле и берёт себе в качестве помощника... https://ru.wikipedia.org/wiki/1+1_(фильм)	10%
Philosophy Flashcards	QuizletCompare text dealing with things sensibly and realistically attitude or policy. is decidedly humanistic emphasizes that multiple realities exist because of the mind's capacity to havedealing with things sensibly and realistically in a way that is based on practical rather than theoretical considerations. https://quizlet.com/97892423/philosophy-flash-cards/	5%

Plagiarism check report.