

Image illustration of Text using Natural Language Processing

Pirani Zainab

M.H.S.S.C.O.E SabooSiddik
Poly Road, Mumbai8

Amru Ritika

M.H.S.S.C.O.E SabooSiddik
Poly Road, Mumbai-8

Thakur Aman

M.H.S.S.C.O.E SabooSiddik
Poly Road, Mumbai8

Tapadia Sumit

M.H.S.S.C.O.E SabooSiddik Poly Road, Mumbai8

Wagh Maitreya

M.H.S.S.C.O.E SabooSiddik Poly Road, Mumbai8

ABSTRACT

There is a need in today's world to make learning and understanding various concepts easier. The written text is better understood pictorially. To understand a particular sentence it is necessary to know in what context and what emotion is being portrayed in it. There are various emotions, meaning, significance and content in a particular text. Thus we are going to display the meaning of the text in an image or stream of images form so as to comprehend the text in a better way. To convert the meaning of the text into its images we are making use of Natural Language Processing. An appropriate image related to that text will be feed through this process. This will make the common man understand the text in a more clear form. For this purpose we are developing a software called Image illustration of text using Natural Language Processing.

General Terms

Image Extraction

Keywords

Natural Language processing, Natural Language toolkit (nltk), image story, image extraction.

1. INTRODUCTION

Text could be easier to learn and comprehend if we would have a visual view of it. Concepts are much more likely to be remembered in practicality if they are presented as pictures rather than as words. There is a notion called as picture superiority effect. The primary meaning of this notion is that concepts that are learned through viewing pictures are more likely to be recalled than the concepts learnt through words in future.

Also it has become mandatory to express our thoughts in a practical manner rather than by verbally expressing them. A picture gives a better understanding and explanation of words. A visual thinking can make people comprehend the context in a better manner. People who have impaired hearing as they can't hear the voice expression will be able to better understand and imagine through visual thinking. This involves all the emotions related to a human being for example sadness, happiness, anger, aggression etc.

In today's world there are not too many softwares and tools available which display the images related to the text entered. But none are perfect. There are many flaws and limitations in them. The software project that we are developing displays the meaning of the text into its image equivalent in the form of a story. Basically it will show a particular text or a paragraph or a sentence in an image or flow of images form.

Let us explain the basic idea of our project with an example. There is a software called as Amazon studios online. What it

basically does is it takes a story as an input and displays a video of images as its output. The drawback of this software is that it has a database stored in its backend. So because of this only those stories which are present in its database can be processed. It is not dynamic. It has a limited usage.

The advantage of our software will be that anyone can access this software and can enter any kind of text or a sentence or a paragraph. The objective of our software is that it should give the most appropriate images to the text entered. It can be a word, phrase or paragraph and a stream of images are given as output. If more than one sentence has the same meaning then only one image will be displayed related to those sentences. It is going to be processed using Natural Language processing. The software is going to extract images from Google or Yahoo search engine. There is a necessity of an internet connection to use this software. The software will be simple and will have an attractive UI for users and to improve ease of use. The software is designed to be cost effective so that it can be affordable to everyone.

2. LITERATURE REVIEW

We have studied several papers regarding NLP, Emotion mining, software for minimizing the cost to as low as possible without compromising the efficiency of the system. However, we have short listed a set of pertinent papers which will be useful for us in our endeavor to develop our system.

2.1 Probabilistic part-of-Speech Tagging Using Decision Trees[1]

There is a wide range of technologies in Human Language Technology (HLT). Natural Language Processing (NLP) – Text is typically syntactically parsed using information form a formal grammar and a lexicon, the resulting information is then interpreted semantically and used to extract information.

2.2 Text Mining and Natural Language Processing-Introduction for the Special Issue[2]

Text Mining – is more recent, uses technologies primarily developed in the fields of information retrieval, statistics, and machine learning. Its aim is to extract patterns across a large number of documents. Text Mining consists of areas such as automatic text classification, text clustering, automatic summarization, extraction of topics from texts or groups of text and analysis of topic trends in text streams.

2.3 Emotion Mining Techniques in Social Networking Sites[3]

From this paper we learn about the study of extracting the emotion of a text, i.e. Studying the feeling and emotion of the text which the author wants to convey through his text.

Emotions are a mandatory part of human nature that can be considered as hereditary.

2.4 Query Difficulty Prediction for Web Image Search[4]

In this we consider in detail predicting words associated with images. In this paper it describes the various models that link images and texts in various ways. It tries to predict annotations of the entire image using all information present.

2.5 Matching Words and Pictures[5]

This paper aims to investigate the QDP problem in Web image search. A novel method is proposed to automatically predict the quality of image search results for an arbitrary query. This model is built based on a set of valuable features that are designed by exploring the visual characteristic of images in the search results.

3. PROPOSED SYSTEM

The software that we are developing will take the text as an input and will display appropriate images or stream of images as an output. The images are going to be extracted using Natural Language Processing from Google or Yahoo search engine.

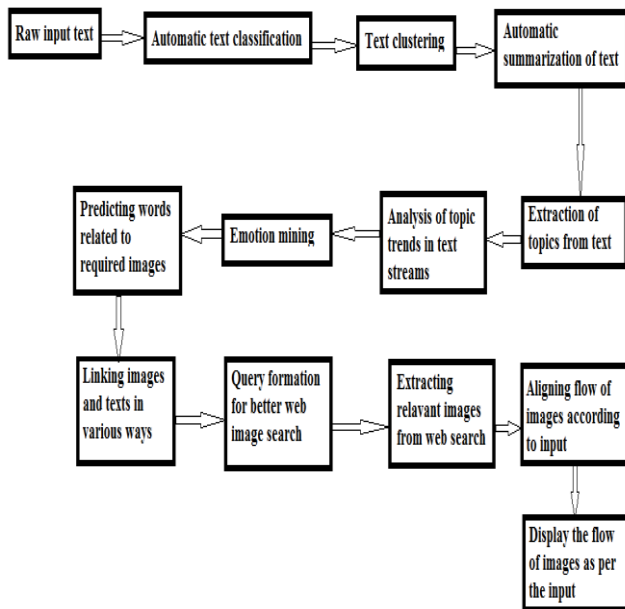


Fig 1: Block diagram of the software

4. WORKING OF THE SYSTEM

The three major phases in the application will be the

- 4.1 Tokenizing and tagging phase
- 4.2 Language processing phase
- 4.3 Image Extraction phase

4.1 Tokenizing and Tagging Phase[1]

In the proposed system we enable users to input a text or stream of text. It then tokenizes the input text; first we start with sentence tokenization and then word tokenization. This step breaks the text into sentences, and then those sentences into words. This step will make it easy for us to operate with the text and for ease of operation. Let's take for example; we take text as, "hello people. I am learning nltk." Will become ['hello peeps.', 'I am learning nltk.']. After sentence tokenization

and ['hello', 'peeps', '.', 'I', 'am', 'learning', 'nltk', '.'] after word tokenization.



Fig 2: Sentence Tokenization



Fig 3: Word Tokenization

Once the text is tokenized, it is then tagged for its part-of-speech. Part-of speech tagging helps us to learn the grammar of the sentence and the words in the text. For example, ['hello', 'peeps', '.', 'I', 'am', 'learning', 'nltk', '.'], the sentence after part-of-speech tagging will become [('hello', 'NN'), ('peeps', 'NNS'), ('.', '.'), ('I', 'PRP'), ('am', 'VBP'), ('learning', 'VBG'), ('nltk', 'NN'), ('.', '.')]]

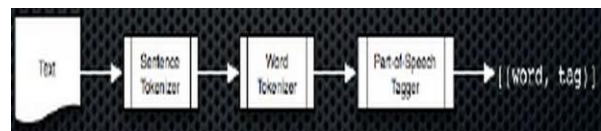


Fig 4: Part-Of-Speech Tagging

4.2 Language Processing Phase

In this phase we will perform language processing on the text which we have tagged and tokenized. In our system we are going to perform this in 6 steps excluding the input from the previous steps. Let's take for example a short story of a Moon.

4.2.1 Input

The Moon is a barren, rocky world without air and water. It has dark lava plain on its surface. The Moon is filled with craters. It gets its light from the Sun. The Moon keeps changing its shape as it moves round the Earth.

4.2.2 Text Classification[2]

In this step we study the text and classify it on the basis of the context of it. From the above text we learn that all the sentences are going in the context of what the surface of the moon is made up of and what are the properties of the moon.

So for example the above text will be classified as:-

1. What the moon is made up of- 1, 2, 3 sentences
2. Properties of the moon- 4, 5 sentences

4.2.3 Text Clustering[2]

In this step we form clusters of the words that can be ambiguous, i.e. Some words may have more than completely different meanings. So we form clusters of such words and try to remove the ambiguity by learning the message the author is trying to convey.

For example Text clustering is performed as follows:-

1. Barren – unproductive or infertile or waste
 - 1.1. Rocky – full of rocks or unsteady
2. Dark – with no light or black
3. Craters – hollow or cavity or hole

4.2.4 Summarization of text[2]

In this step we try and summarize the complete text so as to know in what context is the text, and what the base of the story is. This we do by picking the common traits in each sentence, and by picking who is it referring to, what are the purpose, and the method for the purpose. This gives us a clear idea about the text and will help for better image search.

For example, by summarizing the above text we get:-

The moon is dark and has no atmosphere and it gets light from the sun.

4.2.5 Extraction of topics

In this step we study each sentence, summarize it and try to predict the relevant topic of the sentence. This is done by studying the subject of the sentence.

For example, by extracting topics from each sentence, we say:-

1. The surface of moon
2. The surface of the moon
3. Craters on moon
4. Moon Light
5. Shape of Moon

4.2.6 Analysis of trends in text

Here we study the whole trend of the text, i.e. it describes what in general the whole text is about.

For example here the trend is:-

1. Characteristics

4.2.7 Emotion Mining[3]

In this step we study the emotion of each sentence. This we do by studying the words in a sentence and then looking up to their predefined emotion quotient.

For example, the emotions of the above sentences are as follows:-

1. Serenity
2. Pensive
3. Pensive
4. Helpless
5. Wonder

4.3 Image Extraction Phase

In this phase we have to extract and display the images from the web or our database so as to get our desired output of flow of images. For this we need to extract the keywords related to the images.

Then we need to form a query for better image search. Then all the relevant images are extracted randomly. We then align the flow of images according to the input and finally display them as a flow of images.

All these steps are described in detail below:-

4.3.1 Words related to images

In this step we study the keywords which are describing the noun in the sentences. We also add the emotion to it from the emotion mining step.

For example, by this step we get this output for each sentence:-

1. Moon without air and water
2. Dark lava on moon
3. Craters on moon
4. Moon light from sun
5. Shape of moon

4.3.2 Query formation for better Image search[4]

From the previous step we form queries out of those keywords, so that we get a better search output of the images and we extract the highest ranked image for the query from the search engine.

For example, we get the following queries for the above sentences:-

1. Moon without atmosphere
2. Dark lava on moon surface
3. Craters on the moon
4. Moon light from the sun
5. Orbit of moon around earth

4.3.3 Extracting relevant images from web

From the previous step we get queries to extract the images. From these queries we extract the highest ranked image from the search engine. The figures below show the extracted images from the web.



Fig 5(a): Image extracted from web

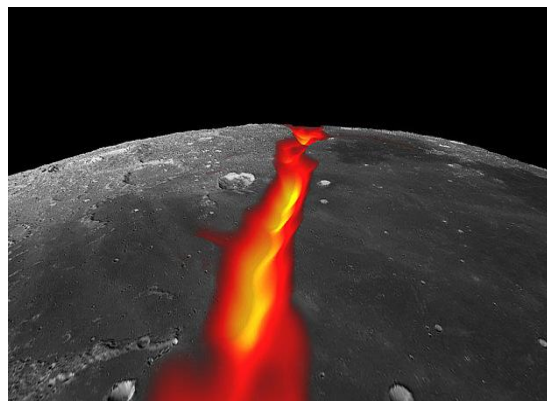


Fig 5(b): Image extracted from web

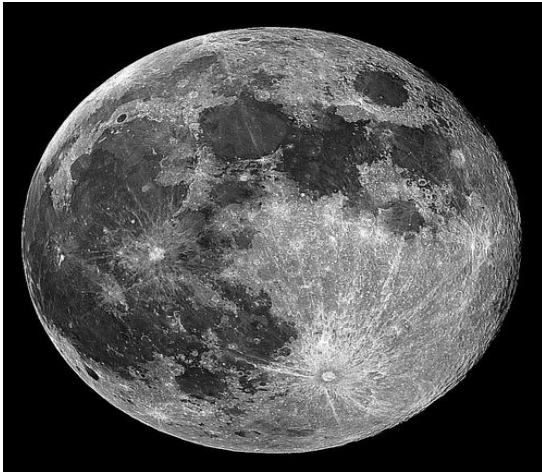


Fig 5(c): Image extracted from web

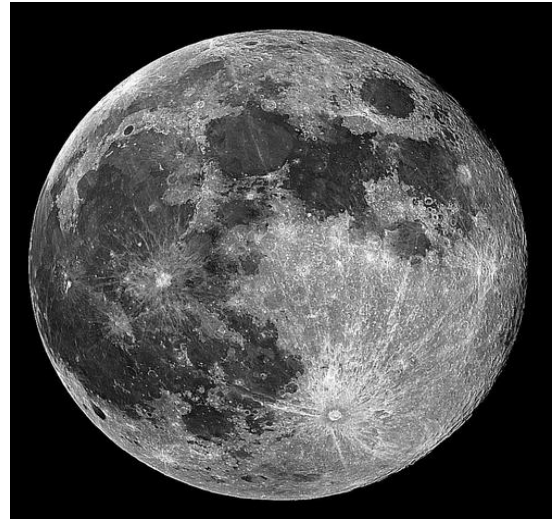


Fig 6(a): Image alignment as per input 1

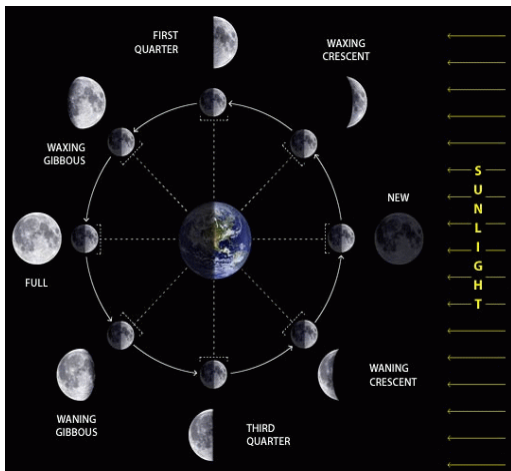


Fig 5(d): Image extracted from web

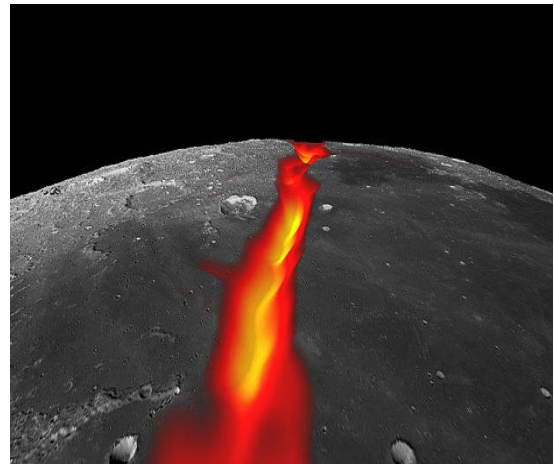


Fig 6(b): Image alignment as per input 2



Fig 5(e): Image extracted from web

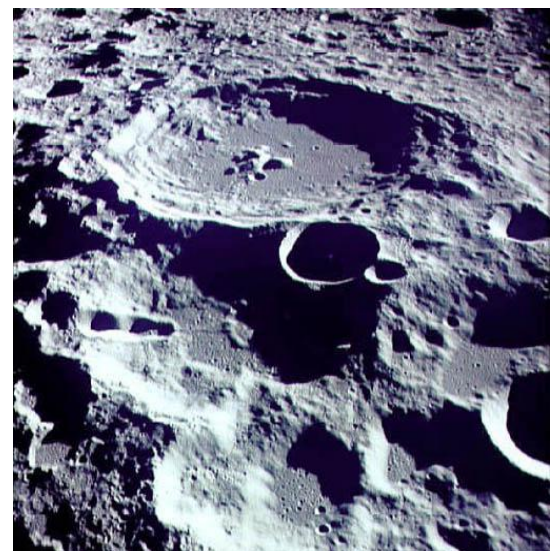


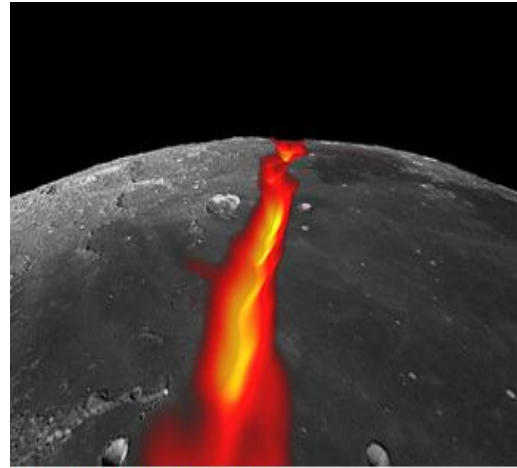
Fig 6(c): Image alignment as per input 3

4.3.4 Aligning flow of images According to input

In this step we align the randomly extracted images from the net according to the flow of sentences in our input text.



Fig 6(d): Image alignment as per input 4



It has dark lava plain on its surface.

Fig 7(b): Image related to the second sentence

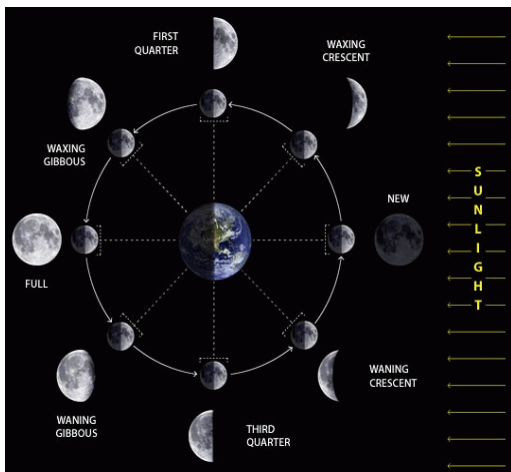


Fig 6(e): Image alignment as per input 5



The Moon is filled with craters.

Fig 7(c): Image related to the third sentence

4.3.5 Display image Story[5]

After aligning the images according to the sentences, we form a flow of these images by an arrow or as a motion picture, and we will assign the relevant text to those images for better understanding by looking at them and reading the text at the same time. We might also read out the text while the images are being displayed.



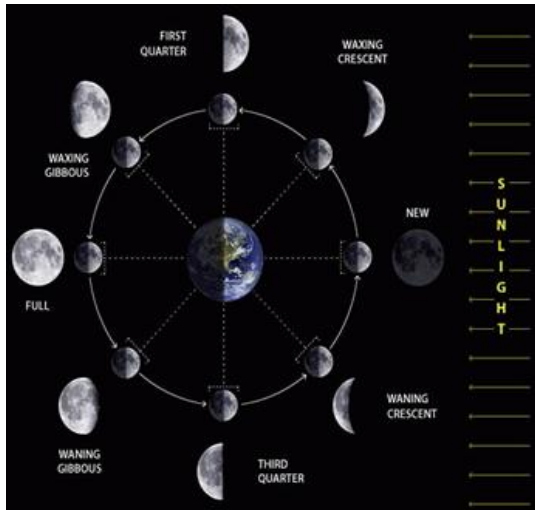
The Moon is a barren, rocky world without air and water.

Fig 7(a): Image related to the first sentence



It gets its light from the Sun.

Fig 7(d): Image related to the fourth sentence



The Moon keeps changing its shape as it moves round the Earth.

Fig 7(e):Image related to the fifth sentence

5. RESULT ANALYSIS

TABLE 1. Comparison of existing system and proposed system

System	Quality Attributes					
	Comprehensible	Functionality	Plainness	Applicability	Accessibility	Learning Independence
Amazon Studios[6]	✓	✓	✓	X	✓	X
Google image search [7]	✓	✓	✓	X	✓	✓
Proposed system	✓	✓	✓	✓	✓	✓

5.1 Comparison of the Parameters of our System and the Existing System

- **Comprehensible:** The proposed system will provide us appropriate pictures related to the text entered.
- **Functionality:** The proposed system will provide ease of use.

- **Plainness:** The proposed system will have a simple user interface so that the common man can use the software without any difficulties.
- **Applicability:** As the software takes any kind of input for example any text or a sentence or a paragraph it has a better applicability than the Google engine or Amazon studios.
- **Learning Independency:** The common man can understand and learn any kind of concepts on its own.

6. CONCLUSION

The software thus aims at understanding the text or a story better. It will develop interest in people to gain more information from images rather than a text. This software will be of great advantage to users conducting seminars and conferences. It will help reduce the reading time. The system will not only reduce physical stress for commuters but also psychological stress that results from frustration and anger when they don't understand the meaning of a particular word or sentence.

7. REFERENCES

- [1] Helmut Schmid, "Probabilistic part-of-Speech Tagging Using Decision Trees" presented at the International Conference on New Methods in Language Processing, 1994, Manchester, UK.
- [2] Anne Kao & Steve Poteet, "Text Mining and Natural Language Processing-Introduction for the Special Issue." presented at the SIGKDD Explorations 2004-05. Volume 7, Issue 1.
- [3] Sanjeev Dhavan, Kulvinder Singh & Deepika Sehrawat, "Emotion Mining Techniques in Social Networking Sites" presented at International Journal of Information & Computation Technology, ISSN 0974-2239, Volume 4.
- [4] Xinmei Tian, Yijuan Lu, Member; IEEE, and Linjuns Yang, Member; IEEE, "Query Difficulty Prediction for Web Image Search" presented at IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 14. AUGUST 2012.
- [5] Kobus Barnard, Pinar Duygulu, David Forsyth, Nando de Freitas, David M. Blei, Michael I. Jordan. "Matching Words and Pictures" presented at the Journal of Machine Learning Research 3 (2003).
- [6] Amazon studios- [http://www.electronista.com/articles/13/06/07/storyteller.scans.scripts.creates.storyboard.from.stock.assets/\[Application available online on Google chrome\]- 20th Oct 2014](http://www.electronista.com/articles/13/06/07/storyteller.scans.scripts.creates.storyboard.from.stock.assets/[Application available online on Google chrome]- 20th Oct 2014)
- [7] Google image search engine- [http://www.quora.com/How-does-google-image-search-engine-work \[Application available online on Google chrome\]- 20th Oct 2014](http://www.quora.com/How-does-google-image-search-engine-work [Application available online on Google chrome]- 20th Oct 2014)