



Semester : \_\_\_\_\_

Subject : \_\_\_\_\_

Academic Year: 20 - 20

➤ Challenges in NLP:-

### 1. Language differences

In the United States, most people speak English, but if you're thinking of reaching an international and/or multicultural audience, you'll need to provide support for multiple languages.

Different languages have not only vastly different sets of vocabulary, but also different types of phrasing, different modes of inflection, and different cultural expectations. You can resolve this issue with the help of "universal" models that can transfer at least some learning to other languages. However, you'll still need to spend time retraining your NLP system for each language.

### 2. Training data

At its core, NLP is all about analyzing language to better understand it. A human being must be immersed in a language constantly for a period of years to become fluent in it; even the best AI must also spend a significant amount of time reading, listening to, and utilizing a language. The abilities of an NLP system depend on the training data provided to it. If you feed the system bad or questionable data, it's going to learn the wrong things, or learn in an inefficient way.

### 3. Development time

Along similar lines, you also need to think about the development time for an NLP system.

To be sufficiently trained, an AI must typically review millions of data points. Processing all those data can take lifetimes if you're using an insufficiently powered PC. However, with a distributed deep learning model and multiple GPUs working in coordination, you can trim down that training time to just a few hours. Of course, you'll also need to factor in time to develop the product from scratch—unless you're using NLP tools that already exist.

### 4. Phrasing ambiguities

Sometimes it's hard even for another human being to parse out what someone means when they say something ambiguous. There may not be a clear concise meaning to be found in a strict analysis of their words. In order to resolve this, an NLP system must be able to seek context to help it understand the phrasing. It may also need to ask the user for clarity.



## 5. Misspellings

Misspellings are a simple problem for human beings. We can easily associate a misspelled word with its properly spelled counterpart, and seamlessly understand the rest of the sentence in which it's used. But for a machine, misspellings can be harder to identify. You'll need to use an NLP tool with capabilities to recognize common misspellings of words, and move beyond them.

## 6. Innate biases

In some cases, NLP tools can carry the biases of their programmers, as well as biases within the data sets used to train them. Depending on the application, an NLP could exploit and/or reinforce certain societal biases, or may provide a better experience to certain types of users over others. It's challenging to make a system that works equally well in all situations, with all people.

## 7. Words with multiple meanings

No language is perfect, and most languages have words that have multiple meanings. For example, a user who asks, "how are you" has a totally different goal than a user who asks something like "how do I add a new credit card?" Good NLP tools should be able to differentiate between these phrases with the help of context.

## 8. Phrases with multiple intentions

Some phrases and questions actually have multiple intentions, so your NLP system can't oversimplify the situation by interpreting only one of those intentions. For example, a user may prompt your chatbot with something like, "I need to cancel my previous order and update my card on file." Your AI needs to be able to distinguish these intentions separately.

## 9. False positives and uncertainty

A false positive occurs when an NLP notices a phrase that should be understandable and/or addressable, but cannot be sufficiently answered. The solution here is to develop an NLP system that can recognize its own limitations, and use questions or prompts to clear up the ambiguity.

## 10. Keeping a conversation moving

Many modern NLP applications are built on dialogue between a human and a machine. Accordingly, your NLP AI needs to be able to keep the conversation moving, providing additional questions to collect more information and always pointing toward a solution.



➤ Difference between lexical ambiguity and syntactic ambiguity.

Ambiguity is the quality of having more than one interpretation. A word, phrase, or a sentence becomes ambiguous if it can be interpreted with more than one meaning. Ambiguity can be classified into two different categories named lexical and structural ambiguity. **Lexical ambiguity occurs when a word has more than one possible meanings. Structural ambiguity is a situation where one sentence has more than one meaning due to its sentence structure.** This is the **key difference** between lexical and structural ambiguity.

**What is Lexical Ambiguity?**

Lexical ambiguity, also known as **semantic ambiguity**, occurs when a sentence has an ambiguous word or phrase (which has more than one possible meaning). This phenomenon is a result of polysemy. Lexical ambiguity is sometimes used deliberately to create pun and other wordplays. Given below are some examples of lexical ambiguity.

1. We saw her duck.

We saw her pet animal.

We saw her bend to avoid something. (verb duck)

2. The minister married her sister.

Her sister married a minister.

The minister performed the wedding ceremony.

3. Harriet cannot bear children.

Harriet cannot give birth to children.

Harriet cannot tolerate children.

4. The fisherman went to the bank.

The fisherman went to the river bank.

The fisherman went to a financial institution.

Although lexical ambiguity can cause problems in meanings, it is not difficult to understand the intended meaning of the writer by looking at the context. For example,



“We saw her duck when we visited her last month. She has made a special pond in the garden to keep it.” – Duck here refers to an animal.

### What is Structural Ambiguity?

Structural ambiguity, also known as **syntactic ambiguity**, occurs when a phrase or sentence has more than one underlying structure. Such a sentence can be interpreted in more than one way. Given below are some examples of structural ambiguity.

1. Miriam hit the boy with a book.

Miriam used the book to hit the boy.

Miriam hit the boy who was holding a book.

2. The teacher said on Friday he would give a test.

On Friday, the teacher said he would give a test.

The test will be on Friday.

3. Visiting relatives can be boring.

It is boring to visit relatives.

Relatives who are visiting are boring.

4. They are cooking apples.

A group of people are cooking apples.

They are apples that can be cooked.

5. Peter saw his neighbour with binoculars.

Peter has a telescope, and he saw his neighbour while using the binoculars.

Peter saw the neighbour who has binoculars.

What is the difference between Lexical and Structural Ambiguity?

Cause:



Lexical Ambiguity: Lexical Ambiguity occurs due to polysemy – words having more than one meaning.

Structural Ambiguity: Structural Ambiguity occurs due to the structure of the sentence.

The Intended Meaning:

Lexical Ambiguity: The intended meaning can be understood by the context.

Structural Ambiguity: The intended meaning can be understood by the prosodic features such as stress, intonation, etc.

### ➤ Applications of NLP

#### 1. Sentiment Analysis

Understanding natural language can be challenging for machines, especially when dealing with human opinions, often expressed through sarcasm and irony. However, sentiment analysis has the capability to discern subtle emotional nuances and opinions, determining their positivity or negativity. Real-time sentiment analysis enables the monitoring of social media mentions, allowing for the proactive management of negative comments before they escalate. It also provides insights into customer reactions to ongoing marketing campaigns or recent product launches, offering an overall understanding of the public sentiment toward your company.

Periodic sentiment analysis is one of the applications of NLP in artificial intelligence that allows for a deeper comprehension of customer preferences and concerns related to specific aspects of your business. For instance, it can reveal that customers appreciate a new feature but are dissatisfied with customer service. These insights serve as valuable inputs for making informed decisions and identifying areas for improvement in your business.

#### 2. Text Classification

Text classification, a facet of text analysis that encompasses sentiment analysis, entails the automated comprehension, processing, and categorization of unstructured text.

Imagine having to analyze numerous open-ended responses from your recent NPS survey manually – a time-consuming and potentially costly endeavor. However, what if you could train a natural language



processing model to swiftly categorize your data within seconds, using predefined categories and applying your own criteria?

For instance, you could employ a topic classifier tailored for NPS survey responses. This classifier would automatically tag your data based on topics such as Customer Support, Features, Ease of Use, and Pricing. This streamlined approach enhances efficiency and facilitates the extraction of valuable insights from your data.

### 3. Chatbots & Virtual Assistants

Chatbots and virtual assistants are some of the applications of NLP in real life that serve the purpose of automated question answering, designed to comprehend natural language and provide relevant responses through natural language generation.

While traditional question-answering systems adhere to predefined rules, AI-powered chatbots and virtual assistants possess the ability to learn from each interaction, adapting their responses accordingly. Notably, these intelligent systems continuously improve over time through learning from their interactions.

Increasingly deployed at the forefront of customer support, these machines can efficiently handle up to 80% of routine queries, directing more complex issues to human agents. Operating around the clock, chatbots and virtual assistants contribute to faster response times and alleviate human agents from dealing with repetitive and time-consuming inquiries.

### 4. Text Extraction

Text extraction, also referred to as information extraction, involves the automatic identification of specific details within a text, encompassing names, companies, locations, and more, commonly known as named entity recognition. Additionally, it includes extracting keywords and predefined features like product serial numbers and models.

One practical application of text extraction is in managing incoming support tickets, where specific data such as company names, order numbers, and email addresses can be identified without the need to manually open and read each ticket. Another use case for text extraction is in data entry. For example, individuals and organizations can use NLP-based data extraction tools such as [Imagetotext.io](https://www.imagetotext.io) to quickly extract data from images, documents, invoices, bank statements, etc. Extracted information can be seamlessly integrated into a database through automated triggers, streamlining the process.

The application of NLP in machine learning text extraction becomes evident. Combining keyword extraction with sentiment analysis offers an enhanced understanding of customer sentiments. This synergy





allows you to identify the most frequently used words by customers to express negativity toward your product or service, providing valuable insights for machine learning applications.

## 5. Machine Translation

Machine translation (MT) stands out as one of the earliest applications of natural language processing. Despite the declaration of Facebook's translations as superhuman, the challenge for machine translation lies in grasping context.

Nevertheless, for those who have been consistent users of Google Translate over the years, the significant progress it has made is evident, thanks to substantial advancements in neural networks and the increased availability of extensive datasets.

The application of natural language processing, especially in the context of automated translation, proves invaluable in business settings. It streamlines communication, enables companies to connect with broader audiences, and swiftly and cost-effectively comprehends foreign documentation.

## 6. Market Intelligence

Leveraging Natural Language Processing (NLP) can be a game-changer for marketers seeking deeper insights into their customers, ultimately enhancing their ability to craft more impactful strategies.

By delving into unstructured data and analyzing elements such as topics, sentiment, keywords, and intent, marketers can significantly augment their market research efforts. This approach not only illuminates current trends but also unveils potential business opportunities. Moreover, the ability to discern customer pain points and monitor competitor activities through data analysis adds another layer of strategic advantage.

The incorporation of example of NLP in daily life in market intelligence not only refines the understanding of customer preferences but also empowers marketers to stay agile in responding to dynamic market conditions, thus contributing to more informed and effective decision-making processes.

## 7. Speech Recognition

Harnessing the power of natural language processing, speech recognition technology converts spoken language into a machine-readable format.

Speech recognition systems are integral to virtual assistants like Siri, Alexa, and Google Assistant. Beyond consumer applications, the business landscape is witnessing an increasing adoption of speech recognition, highlighting diverse applications and the broader context of the application of AI in natural language



processing. For instance, integrating speech-to-text capabilities into business software enables companies to automate call transcriptions, streamline email communications, and even facilitate language translation.

The synergy between artificial intelligence (AI) and natural language processing in speech recognition applications underscores the evolution of technology, enabling more seamless and efficient communication. This not only enhances the user experience in virtual assistants but also presents businesses with innovative ways to boost productivity and streamline communication workflows in the broader framework of the application of AI in natural language processing.

## 8. Hiring and Recruitment

The Human Resource department holds a crucial role in every organization, tasked with the pivotal responsibility of selecting the most suitable employees. In today's highly competitive environment, recruiters often face the daunting task of reviewing numerous resumes, sometimes numbering in the hundreds or thousands, for a single position. This process of filtering resumes and shortlisting candidates can be time-consuming. However, there's a solution – automation through natural language processing (NLP).

By leveraging NLP, recruiters can streamline the process of identifying the right candidates. It is one of the important applications of NLP in AI. This means they no longer have to manually sift through each resume to filter potential candidates. Techniques like information extraction, coupled with named entity recognition, enable the extraction of essential details such as skills, names, locations, and education from resumes. These extracted features can then be used to represent candidates in a feature space, allowing for classification into categories such as fit or not fit for a specific role. Alternatively, candidates could be recommended for different roles based on the content of their resumes.

This automated approach ensures unbiased resume filtering, facilitating the selection of the most qualified candidates for a vacant position, all without extensive manual effort. Many companies employ Application Tracking Systems to efficiently screen resumes and enhance the overall recruitment process.

## 9. Email Filters

Among the various types of email filters, the spam filter stands out as a widely recognized example. Spam accounts for approximately 85% of global email traffic, underscoring the vital role that email filters play. How can the applications of natural language processing (NLP) contribute to the effectiveness of these filters? Over the years, you've likely observed the evolution of email filters aimed at helping users maintain an organized inbox. A prime illustration is Gmail, which provides distinct categories like primary, promotions, updates, and social for email organization.





At the heart of these email filters lies natural language processing, serving as a fundamental element. As emails land in your inbox, NLP techniques, including keyword extraction and text classification, play a crucial role in automatically scanning and categorizing the emails. This showcases the impactful example of NLP in daily life applications in enhancing the efficiency and organization of email filters.

#### 10. Customer Support

Have you ever found yourself needing to alter your flight plans or cancel a credit card? In such instances, you've likely encountered an automated answering system on the other end. While it may seem somewhat routine, these programmed systems serve a crucial purpose for corporations. They efficiently sift through customer information, ensuring a swift connection to the appropriate representative.

Moreover, these automated systems contribute valuable data for future interactions and enhancements. Don't be astonished if, in the near future, these answering machines start responding with a more human-like voice, offering a personalized touch to address your inquiries. The evolution of customer support mechanisms continues to embrace technological advancements, aiming to provide seamless and enhanced assistance to customers.



PARSHWANATH CHARITABLE TRUST'S  
**A.P. SHAH INSTITUTE OF TECHNOLOGY**  
Department of Computer Science and Engineering  
Data Science

