

AIFE: UNIT-3

MS. SHWETA TIWARI

Published: DECEMBER, 2021

“Artificial Intelligence for Engineering/Engineers (KMC-201)”_

UNIT-3: NATURAL LANGUAGE PROCESSING

FALL SEMESTER, YEAR (2ND, 1ST)

FALL SESSION (2021-22)

(AIFE)

MS. SHWETA TIWARI

Published: DECEMBER, 2021

Rajkiya Engineering College | Ambedkar Nagar, UP, India

**Faculty Name: Miss. Shweta Tiwari, Subject: Artificial Intelligence for Engineering (KMC101) Year- 1st Year, Semester- 2nd Sem,
Branch- CE, Session- Odd Semester (2021-22)**



PREPARED FOR

Engineering Students

All Engineering College

PREPARED BY

SHWETA TIWARI

Guest Faculty

UNIT-3

“Artificial Intelligence for Engineering/Engineers (KMC-201)”

By SHWETA TIWARI

Artificial Intelligence for Engineering/Engineers (KMC-201)”

UNIT-3: NATURAL LANGUAGE PROCESSING

December, 2021

Artificial Intelligence for Engineering

UNIT-III

UNIT 3: NATURAL LANGUAGE

PROCESSING

December, 2021

UNIT 3: Natural Language Processing

3.1 What is NLP?

Natural Language Processing (NLP) refers to the technology that is used by machines to understand, analyze, manipulate, and



The input and output of an NLP system can be-
Speech

Written Text

computerized data into natural language representation.

3.2 Steps in NLP

- 1. Lexical Analysis (Morphology)** – It involves identifying and analyzing the structure of words. It divides the whole chunk of text into paragraphs, sentences, and words.

- 2. Syntactic Analysis (Parsing)** – It

Components of NLP

There are two components of NLP:

- 1. Natural Language Understanding (NLU):** It helps the machine to understand and analyse human natural language.
- 2. Natural Language Generation (NLG):** It is a translator that converts the

December, 2021

involves analysis of words in the sentence for grammar and arranging words in a manner that shows the relationship among the words. The sentence such as “The school goes to boy” is rejected by English syntactic analyzer.

3. **Semantic Analysis** – It draws the exact meaning or the dictionary meaning from the text. The text is checked for meaningfulness. The semantic analyzer disregards sentence such as “hot ice cream”.
4. **Discourse Integration** – The meaning of any sentence depends upon the meaning of the sentence just before it. In addition, it also brings about the meaning of immediately succeeding sentence.
5. **Pragmatic Analysis** – During this, what was said is re- interpreted on what it actually meant. It involves deriving those aspects of language which require real world knowledge.

December, 2021

3.3 Applications of NLP

1. **Speech recognition:** This technology uses natural language processing to transform spoken language into a machine-readable format.
2. **Chatbot:** They are used for automatic question answering, designed to understand natural language and deliver an appropriate response through natural language generation.
3. **Virtual Assistants:** A voice assistant is a software that uses speech recognition, natural language understanding, and natural language processing to understand the verbal commands of a user and perform actions accordingly.
4. **Character Recognition:** Optical Character Recognition (OCR) is the process of converting images of handwritten, typed, or printed text into machine-encoded language. It is one of the commonly used approaches to digitize printed texts so it can be saved, edited, searched electronically.
5. **Text Extraction:** It automatically detects specific information in a text, such as names, companies, places, and more. This is also known as named entity recognition. You can also extract keywords within a text, as well as pre-defined features such as product serial numbers and models.
6. **Machine Translation:** It is an automatic translation from one language to another. The benefit of machine translation is that it is possible to translate large amount of text in a very short time. It is one of the first applications of NLP.
7. **Auto-Correct:** NLP plays a vital role in grammar checking software and auto-correct functions. Tools like Grammarly, for example, use NLP to help you improve your Writing, by detecting grammar, spelling, or sentence structure errors.
8. **Email Filtering:** The emails are filtered using text classification, which is a natural language processing technique. It classifies all the into the sections of primary, social, and promotions.

3.4 Speech Recognition

Speech recognition is the process that enables a computer to recognize and respond to spoken words and then converting them in a format that the machine understands.

Example: Voice Typing in GBoard and other transcription programs use speech recognition to convert your spoken words into text while digital assistants like Siri and Alexa respond in text format or voice.

Speech recognition focuses on the translation of speech from a verbal format to a text one whereas **voice recognition** just seeks to identify an individual's voice.

voice recognition. New users must first "train" the software by speaking to it, so the computer can analyze how the person talks.

2. **Speaker Independent:** software is designed to recognize anyone's voice, so no training is involved. This means it is

Why do we need Speech Recognition?

Most natural form of communication.

Differently abled people can use it with ease.

Helps people who can't read or write, Hence they can communicate with computers normally.

Increase adaption of technology by making it easier to use.

Types of Speech Recognition

There are three types of speech recognition.

1. **Speaker Dependent:** software works by learning the unique characteristics of a single person's voice, in a way similar to

December, 2021

the only real option for applications such as interactive voice response systems. Hence it's mostly voice recognition along with speech recognition.

3. **Speaker Adaptive:** They usually begins as a speaker independent model and slowly adapts and adjusts to the individual using the systems. Virtual Assistants like, Google, Alexa, Siri etc. uses similar methods.

Applications of Speech Recognition:

1. **Voice to text:** Speech recognition enables hands free computing. Users don't need to type emails, reports, and other documents.
2. **Voice commands to smart home devices:** Smart home applications are mostly designed to take a certain action after the user gives voice commands.
3. **Security:** As technology integrates into our daily lives, security protocols are an increasing priority. Voice-based authentication adds a viable level of security.

4. **Voice Search:** This is the most common use of speech recognition. Users prefer to use voice searches Siri and Google voice search, over the traditional method of typing.
5. **Automotive:** In-car speech recognition systems, aim to remove the distraction of looking down at your mobile phone while you drive. Drivers can use simple voice commands to initiate phone calls, select radio stations or play music.

to translate large amount of text in a very short time.

- ✓ You can type text or upload whole documents and receive translations in dozens of languages using machine translation tools.

2. Automated Reasoning

3.5 Natural Language Understanding

- Natural language understanding (NLU) is a sub-topic of natural language processing, which involves breaking down the human language into a machine-readable format.
- Interesting applications include text categorization, machine translation, and question answering.
- NLU uses grammatical rules and common syntax to understand the overall context and meaning of “natural language,”.
- In shorts, Its goal is to understand written or spoken language the way a human would.

Examples of NLU:

1. Machine Translation (MT):

- ✓ MT is an automatic translation from one language to another. The advantage of machine translation is that it is possible

December, 2021

- ✓ Automated reasoning is a sub-field of cognitive science which is used to automatically prove mathematical theorems about a medical diagnosis.
- ✓ It gives machines a form of reasoning or logic, and allows them to infer new facts by deduction.

3. Automatic Ticket Routing

- ✓ With text analysis solutions, machines can understand the content of customer support tickets and route them to the correct departments without employees having to open every single ticket.
- ✓ This not only save customer support teams hundreds of hours, but it also helps them prioritize urgent tickets.

4. Question Answering

- ✓ Question answering uses NLU to help computers automatically understand natural language questions.

✓For example, If we ask Google Assistant:

"What is the weather like tomorrow?"

✓NLP tools can split this question into topic (weather) and date (tomorrow), understand it and gather the most appropriate answer from unstructured collections of "natural language documents": online news reports, collected web pages etc.

✓The goal of question answering is to give the user response in their natural language, rather than a list of text answer.

2. **Chatbots:** The best chatbots are the ones that give users the impression that they are chatting with a real human. These chatbots can be used for multiple purposes, such as complaint and query resolution and virtual assistance for online processes (e.g., form filling).

3.6 Natural Language Generation

Natural language generation (NLG) is a process of producing meaningful phrases and sentences in the form of natural language.

It automatically generates narratives that describe, summarize or explain input structured data in a human like manner at the speed of thousands of pages per second.

Applications of NLG

1. **Analytics Dashboards:** Analytics dashboards are among the earliest and the most obvious application areas for natural language generation. NLG-powered tools can generate information in an easily comprehensible format from the Analytics so that they can make quick and effective decisions.

December, 2021

3. **Content Creation:** AI research will truly achieve new heights when machines that are capable of creating content with the naturalness and quiriness of human writers are developed. Such as part and product descriptions, internal communications, agreements and contracts, and other similar forms of textual communications.

4. **Grammaticalization:** Grammaticalization stage makes sure that the whole report follows the correct grammatical form, spelling, and punctuation.

3.7 Chatbots

A chatbot is artificial intelligence (AI) software that can simulate a conversation (or a chat) with a user in natural language through messaging applications, websites, mobile apps or through the telephone.

A chatbot is often described as one of the most advanced and promising expressions of interaction between humans and machines.

- ✓ A contextualized piece of information based on data the user has provided.
- ✓ The result of an action that the chatbot performed by interacting with one or more backend application.
- ✓ A disambiguation question that helps the chatbot to correctly understand the user's request.

Working of a Chatbot

There are two different tasks at the core of a chatbot:

1. **User request analysis:** This is the first task that a chatbot performs. It analyzes the user's request to identify the user intent and to extract relevant entities.

It is the first condition and the most relevant step at the core of a chatbot: If you are not able to correctly understand the user's request, you would not be able to provide the correct answer.

2. **Returning the response:** Once the user's intent has been identified, the chatbot must provide the most appropriate response for the user's request. The answer maybe:

- ✓ A generic and predefined text.
- ✓ A text retrieved from a knowledge base that contains different answers.

December, 2021

Importance of chatbots

- ✓ Chatbot applications streamline interactions between people and services, enhancing customer experience.
- ✓ Offer companies new opportunities to improve the customer's engagement process and operational efficiency by reducing the typical cost of customer service.
- ✓ AI chatbots can make attending to customers quick and effortless. There is no off day for them and can engage with customers at all time.
- ✓ Using AI chatbots, don't need to invest in 24x7 customer service teams.
- ✓ AI chatbots interact with customers, collect valuable data, and analyze this data. We can systematically collect and use this data to understand your audience and market to them.

December, 2021

3.8 Machine Translation

- Machine translation (MT) is an automatic translation from one language to another.
- The benefit of machine translation is that it is possible to translate large swathes of text in a very short time.
- If you specially train the machine to your needs, machine translation provides the perfect combination of quick and cost-effective translations.
- With a specially trained machine, MT can capture the context of full sentences before translating them, which provides you with a high quality and human-sounding output.

Advantages of Machine Translation

- ✓ Fast and does not require vetting and managing of translators
- ✓ Cost-efficient for large volumes of translations
- ✓ Reduced time-to-market due to faster translation delivery
- ✓ Flexibility from a number of Machine Translation source engines
- ✓ Adaptable, programmable, and developer-friendly
- ✓ Ability to retrain MT into customized workflows and strings

December, 2021

Limitation of Machine Translation

- × Inability to account for certain local phrases due to lack of context
- × Possibility for diluted marketing and brand messages due to word-for-word translations
- × Difficulty to accurately translate nuances, slang, and other culturally relevant phrases
- × Possibility for brand damage due to lack of cultural awareness and cohesiveness

- × Difficulty translating complicated or industry-specific terms
- × Difficulty predicting and correcting specific grammatical and cultural errors