Unit 5(Part1)

Introduction to chatbot application: Chatbots, also known as conversational agents, are designed with the help of AI (Artificial Intelligence) software. They simulate a conversation (or a chat) with users in a natural language via messaging applications, websites, mobile apps, or phone.

There are two primary ways chatbots are offered to visitors:

1. Web-based applications
2. Standalone applications

Chatbots represent a potential shift in how people interact with data and services online. While there is currently a surge of interest in chatbot design and development, we lack knowledge about why people use chatbots.

Here are specific steps to keep in mind for chatbot development.

1. Defined Objectives & Aspirations

Chatbots today mimic human conversations. Thanks to their learning ability and 24/7 presence. They can optimize communication and create real engagement. The client must build a creative and user-friendly interface for communication. Over-burdening your chatbot with traits and crafting it to ace all undertakings will probably set you up for disappointment.

The Approach

Instead of spreading the chatbot too thin over multiple functions, it can be crafted to focus entirely on one essential command. Always keep in mind; individuals need quality, not quantity.

2. Shorter Responses

In today's fast-paced world, with attention spans growing shorter every second, no one has the time to read out long conversations. Jutting in complicated languages and lengthy dialogues will make the chatbot seem tedious. Though your bot is capable of handling long messages and sending responses to the user, we need a mechanism to ensure that the bot is interactive and capable of responding to diverse and yet common queries that the user might have.

The Approach

When it comes to general usage, a bot should reply to the query in advance during interactions involving single-line or two-line messages. Be imaginative. Keep it basic; the bot must be clear about the next step. To achieve this, we need to train the bot to reply quicker for frequently asked messages.

3. Bot Humanization

There is a fine line between a decent bot and an incredible bot, and the latter is possible only if you give your bot a genuine identity (named as human).

It is not merely enough to pack a sequence of answers and algorithms with a human touch. Never neglect to humanize your bot, as it can leave your potential customer with mixed feelings. Users prefer to have a human conversation, irrespective of the knowledge that they are chatting with a chatbot.

The Approach

You can give a human personality to your bot with a cool title. Discover a particular and personalized name for your bot so that your users can find it easily. Additionally, educate your bot about its representation. Ensure your bot has specific information about its personification, especially when users attempt to get some info about your bot name, age, or its central goal. Ensure to keep your chatbot holistic in approach.

4. Design the conversation

Chatbot conversations are designed to attract customers. But when this is not executed correctly, it can be taxing on the customer experience. Most chatbots redirect to the Live agent quickly, wherever there are in-depth queries and conversations required. This can help retain the customers' interest.

The Approach

Conversational chatbots are now enabling you to comprehend your customer's demands better and collect more significant information to make the interaction between your bot & customer more open and easier. We need to monitor the use-cases in clients' current activities and save communication flows.

5. Poor escalation protocol

Undoubtedly, a chatbot can communicate with the clients and help them to settle on an active choice by demonstrating them the data from the database and contingent upon the information given by clients.

The bots can likewise get to the client's expressed objective, keeping in mind that the end goal is to allow adequate search and results. They can gather and present pre-chosen reviews to clients amid an interaction, given the keyword found in the data.

However, if any procedure is extremely customized and profitable, for example, marketing, then it may be helpful also to keep a human in the string. It is vital to comprehend how and where are you going to utilize your chatbot for effective business practices.

The Approach

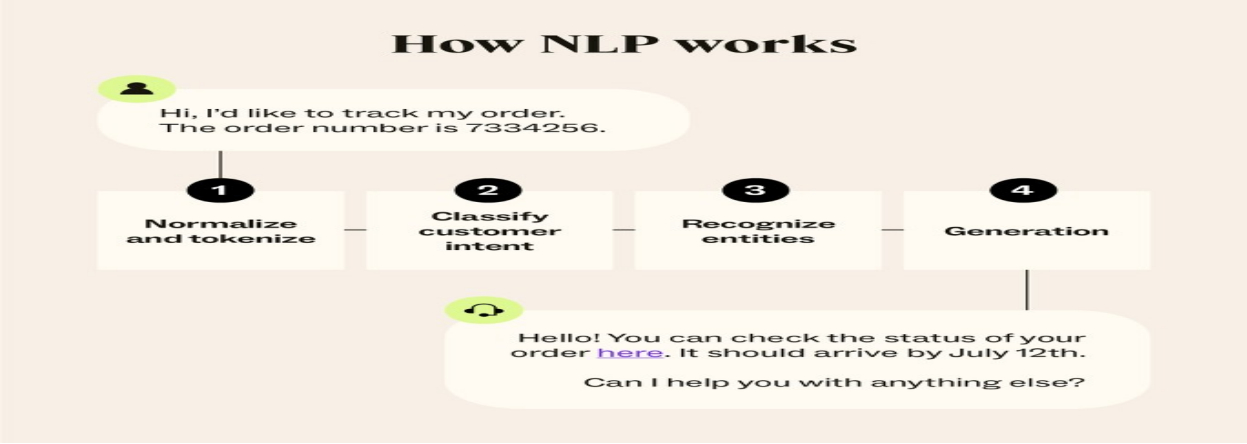
Chatbots can be utilized for various reasons. But one fundamental reason that organizations adapt to chatbots is the reliable customer service and customer engagement. Chatbots enhance user experiences and increase the value of the

organization. At the point that a customer has a positive involvement with an organization, they will, in the end, turn into a trustworthy customer.

Chatbots vs. AI chatbots vs. virtual agents vs. virtual assistants

	Chatbot	AI Chatbot	Virtual Agent
What is it?	Most inclusive. Any software that simulates human conversation	Chatbot that uses AI technology, from machine learning to NLP	AI chatbot that adds on robotic process automation to act upon user desires without prompting
Where are they often found?	Any communication channel, from social media to websites	On your phone or home	In specific customer service instances
Example	Customer service chatbot for health care company	Alexa and Siri	IBM watsonx Assistant

How a natural language processing chatbot works:



Bots using a conversational interface—and those powered by large language models (LLMs)—use major steps to understand, analyze, and respond to human language. For NLP chatbots, there's also an optional step of recognizing entities.

Let's take a closer look at how a natural language processing chatbot works:

1. Normalizing: Bots remove irrelevant details and convert words to a standardized version. For example, bots will lowercase language inputs.
2. Tokenizing: Chatbots chop the language input into pieces—or tokens—and remove punctuation.
3. Intent classification: With normalized and tokenized text, the bot uses AI to identify the issue or intent the customer is asking about.
4. Recognizing entities (optional): This optional step is where chatbots identify anything else referred to in a message, such as an order number, email address, or transaction ID.
5. Generation: For next-gen NLP AI agents, the AI model generates a number of responses and selects the most appropriate response to send to the user.

How to build a chatbot with NLP: There are several methods you can use to approach this. But, ultimately, your choices boil down to:

- Developing an NLP chatbot from scratch
- Using an existing chatbot NLP framework
- Designing bots with an NLP chatbot platform

Let's take a look at each of these methods of how to build a chatbot using NLP.

1. Creating an NLP chatbot in Python

If you decide to create your own NLP AI-powered chatbot from scratch, you'll need to have a strong understanding of coding both artificial intelligence and natural language processing. This option is recommended for experienced developers only.

The most common way to do this is by coding a chatbot in a programming language like Python and using NLP libraries such as Natural Language Toolkit (NLTK) or spaCy.

Building your own chatbot using NLP from scratch is the most complex and time-consuming method. So, unless you are a software developer specializing in chatbots and AI, you should consider one of the other methods listed below.

2. Use chatbot frameworks with NLP engines

Some of you probably don't want to reinvent the wheel and mostly just want something that works. In that case, you can use an existing chatbot framework. Thankfully, there are plenty of open-source NLP chatbot options available online.

The most popular choices include Microsoft Bot Framework, Amazon Lex, IBM, and Google Dialogflow. These frameworks will give you the building blocks you need to develop a chatbot. However, you'll still have to put in a lot of work and know how to code to use these chatbot NLP architectures effectively.

3. Utilize NLP chatbot platforms

This is the easiest chatbot development method. So, if you want to avoid the hassle of developing and maintaining your own NLP conversational AI, you can use an NLP chatbot platform instead. These ready-to-use chatbot apps provide everything you need to create and deploy a chatbot, without any coding required.

All you have to do is set up separate bot workflows for different user intents based on common requests. From the user's perspective, you just need to type or say something, and the NLP support chatbot will know how to respond.

Retrieval-Based Chatbots: Retrieval-based chatbots are a type of conversational AI system that operate by matching user inputs to a predefined set of responses stored in a knowledge base. They use natural language processing techniques to understand user intents and queries, and then retrieve the most appropriate pre-written response from their database to provide back to the user.

How do they work?

Retrieval-based chatbots rely on a curated knowledge base containing conversational scripts, information, and responses customised to specific use cases and domains. When a user sends a message, the chatbot analyses the input using techniques like pattern matching, keyword identification, and intent recognition to map the user's query to the most relevant response in its database. The chatbot then returns this pre-written response to the user.

Examples of Retrieval-Based Chatbots:

- Mitsuku: This chatbot contains over 300,000 predefined response patterns and a knowledge base of over 3,000 objects. Mitsuku can construct songs and poems based on its knowledge base.
- A FAQs ChatBot for Babcock University, Nigeria: This chatbot answers questions that current and future students or the general public may have about the university.

Advantages of Retrieval-Based Chatbots:

- Predictable: They provide 100% consistent responses as they select from a predetermined set of data and answers.
- Less Risky: There is no likelihood of generating inappropriate or nonsensical answers.

- Integrated: They can connect to various systems, such as Student Information System(SIS), Customer Relationship Management (CRM), and Learning Management System (LMS), out-of-the-box, enabling them to handle a higher number of queries.
- Fixed Cost: There are no variable fees, like tokens, to consider.

Disadvantages of Retrieval-Based Chatbots:

- Limited Flexibility: They are restricted to their predefined responses and may appear less "intelligent" if they repeatedly provide the same responses to slightly varied questions.
- Longer Training Time: They require a significant amount of data, definitions, and intents to answer and provide data, links, or videos.

Challenges with Retrieval-Based Chatbots:
Retrieval-based chatbots rely on a pre-defined set of responses. They use techniques like keyword matching, machine learning, or deep learning to select the most appropriate response from their database. They are often used in closed-domain scenarios and are good at handling a large volume of requests. However, they may appear less flexible and may struggle with generating dynamic responses. when dealing with complex or ambiguous queries. They may struggle to understand the user's intent and provide generic or irrelevant responses. Additionally, they are limited by their pre-defined responses, making them less adaptable to changing needs.

Conversation based: Conversational artificial intelligence (AI) refers to technologies, such as chatbots or virtual agents, that users can talk to. They use large volumes of data, machine learning and natural language processing to help imitate human interactions, recognizing speech and text inputs and translating their meanings across various languages.
Conversational AI combines natural language processing (NLP) with machine learning. These NLP processes flow into a constant feedback loop with machine learning processes to continuously improve the AI algorithms.

Components of conversational AI:
Conversational AI has principle components that allow it to process, understand and generate response in a natural way.
Machine Learning (ML) is a sub-field of artificial intelligence, made up of a set of algorithms, features, and data sets that continuously improve themselves with

experience. As the input grows, the AI platform machine gets better at recognizing patterns and uses it to make predictions.

Natural language processing is the current method of analyzing language with the help of machine learning used in conversational AI. Before machine learning, the evolution of language processing methodologies went from linguistics to computational linguistics to statistical natural language processing. In the future, deep learning will advance the natural language processing capabilities of conversational AI even further.

NLP consists of four steps: Input generation, input analysis, output generation, and reinforcement learning. Unstructured data transformed into a format that can be read by a computer, which is then analyzed to generate an appropriate response. Underlying ML algorithms improve response quality over time as it learns. These four NLP steps can be broken down further below:

1. Input generation: Users provide input through a website or an app; the format of the input can either be voice or text.
2. Input analysis: If the input is text-based, the conversational AI solution app will use natural language understanding (NLU) to decipher the meaning of the input and derive its intention. However, if the input is speech-based, it'll leverage a combination of automatic speech recognition (ASR) and NLU to analyze the data.
3. Dialogue management: During this stage, Natural Language Generation (NLG), a component of NLP, formulates a response.
4. Reinforcement learning: Finally, machine learning algorithms refine responses over time to ensure accuracy.

<u>Conversational AI use cases:</u>

When people think of conversational artificial intelligence, online chatbots and voice assistants frequently come to mind for their customer support services and omni-channel deployment. Most conversational AI apps have extensive analytics built into the backend program, helping ensure human-like conversational experiences.

Experts consider conversational AI's current applications weak AI, as they are focused on performing a very narrow field of tasks. Strong AI, which is still a theoretical concept, focuses on a human-like consciousness that can solve various tasks and solve a broad range of problems.

Despite its narrow focus, conversation AI is an extremely lucrative technology for enterprises, helping businesses more profitable. While an AI chatbot is the most popular form of conversational AI, there are still many other use cases across the enterprise. Some examples include:

- Online customer support: Online chatbots are replacing human agents along the customer journey. They answer frequently asked questions (FAQs) around topics, like shipping, or provide personalized advice, cross-selling products or suggesting sizes for users, changing the way we think about customer engagement across websites and social media platforms. Examples include messaging bots on e-commerce sites with virtual agents, messaging apps, such as Slack and Facebook Messenger, and tasks usually done by virtual assistants and voice assistants.
- Accessibility: Companies can become more accessible by reducing entry barriers, particularly for users who use assistive technologies. Commonly used features of Conversation AI for these groups are text-to-speech dictation and language translation.
- HR processes: Many human resources processes can be optimized by using conversational AI, such as employee training, onboarding processes, and updating employee information.
- Health care: Conversational AI can make health care services more accessible and affordable for patients, while also improving operational efficiency and the administrative process, such as claim processing, more streamlined.
- Internet of things (IoT) devices: Most households now have at least IoT device, from Alexa speakers to smart watches to their cell phones. These devices use automated speech recognition to interact with end users. Popular applications include Amazon Alexa, Apple Siri and Google Home.
- Computer software: Many tasks in an office environment are simplified by conversational AI, such as search autocomplete when you search something on Google and spell check.

While most AI chatbots and apps currently have rudimentary problem-solving skills, they can reduce time and improve cost efficiency on repetitive customer support interactions, freeing up personnel resources to focus on more involved customer interactions. Overall, conversational AI apps have been able to replicate human conversational experiences well, leading to higher rates of customer satisfaction.

Information extraction: Information extraction (IE) is the automated process of extracting structured information from semi-structured or unstructured text data, transforming human language text sources such as PDFs into a format that's organized, searchable and machine-readable. Natural language processing (NLP) relies on information extraction to identify important data within input text.

Information extraction algorithms can identify entities, including names, relationships, events, sentiment and more, then classify and store them in a database for further use. The resulting structured information has a standardized format and is typically stored in rows and columns that identify its attributes. The standardized storage is the primary differentiator between structured data and unstructured data.
All the data values within the same database adhere to the same structured format with the same defined attributes. Relational attributes are also highlighted to connect databases together based on shared attributes.

Types of information extraction:
Information extraction tasks are categorized based on the type of information being identified and labeled. IE systems can handle tasks including:

1. Named entity recognition (NER)
2. Relation extraction
3. Event extraction
4. Sentiment analysis

1. Named entity recognition (NER) : Named entity recognition is the IE task of identifying named entities in unstructured data. Named entities are real-world objects that can be uniquely identified. Essentially, they are the proper nouns of data. Named entities include people, dates, corporations, places and products and can be both physical or abstract. In the sentence "As of January 2025, Arvind Krishna is the CEO of IBM," the named entities include January 2025, Arvind Krishna, CEO and IBM." Entity linking: Entity linking is the process of figuring out whether multiple entities refer to the same real-world object. When conducting IE on an article mentioning "Arvind Krishna," "Krishna" and "IBM's CEO," an entity linking subtask would identify all 3 as references to the same person. Entity linking is also referred to as coreference resolution.
2. Relation extraction (RE): Relation extraction is the information extraction task of identifying and categorizing the relationships between entities in a data source. Uncovering relationships between entities can open the door to insights that might otherwise go unnoticed. In our example sentence from the beginning of this section, the RE process would draw a "works at" connection between "Arvind Krishna" and "IBM" with the title of "CEO."Relation extraction versus relationship extraction: The terms relation extraction and relationship extraction are often used interchangeably, but some data scientists argue for a subtle distinction. While relationship extraction covers any attempt to discern the relationships between entities,

relation extraction is most often used regarding the application of machine learning models to accomplish this task.

3. Event extraction: Event extraction is how IE systems recognize discrete events in a body of input text. Words, such as "appointment" or "meeting," can trigger an event extraction sequence, as can dates. Event extraction covers the event itself, the time and date at which it occurred and any mentioned participants. In the sample sentence, "Arvind Krishna attended the conference in January 2025," an event extraction algorithm would identify that a conference took place in January 2025 and that 1 of the attendees was IBM CEO Arvind Krishna.
4. Sentiment analysis: Sentiment analysis determines the feeling communicated by a piece of text. Sentiment analysis is a valuable tool for conducting market research and understanding customer behavior. If given a dataset consisting of user reviews, an IE algorithm can provide semantic insights that reveal the percentages of consumers that feel positively, negatively or neutrally about a product. Product managers could then take those insights and tweak the product to make it more appealing to a greater portion of their current and potential users.

How does information extraction work?

Information extraction works by parsing unstructured data sources with machine learning algorithms to identify meaningful data. IE systems label the discovered data entities and store them in an organized, queryable database for efficient retrieval.

Information extraction techniques include:

1. Rule-based
2. Classification (machine learning)
3. Sequence labeling

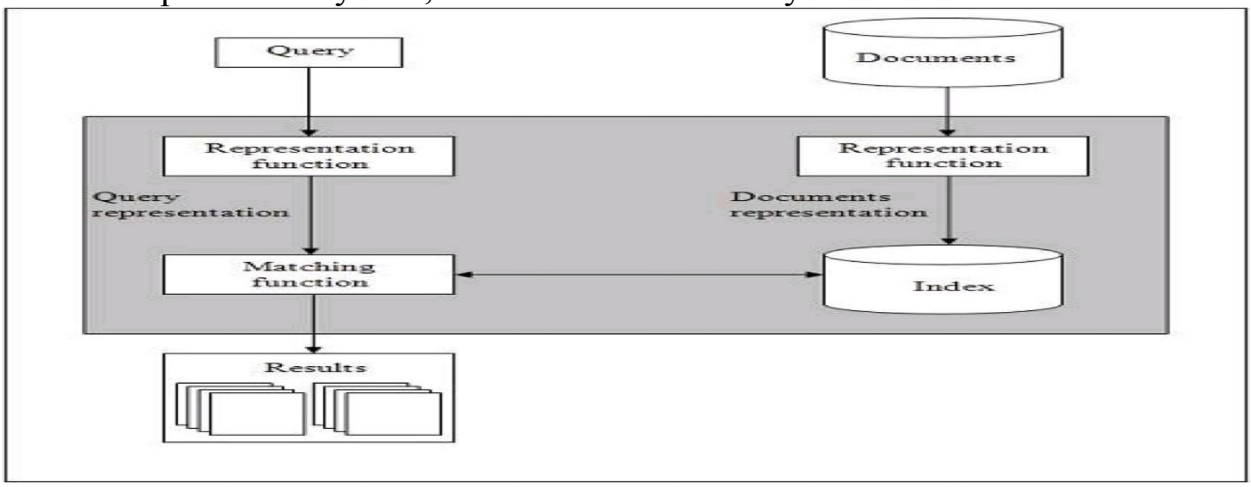
These methods are not mutually exclusive—advancements in IE have led to hybrid models that combine methods for improved results.

1. Rule-based information extraction: Rule-based information extraction parses documents to identify entities based on established "rules"—predefined patterns and definitions that are known about the entities in the text. Rule-based IE is most often applied to semi-structured data sources—data that isn't fully structured but still has some identifying features such as tags or metadata. Top-down rule-based IE works by progressing from general cases to specific cases, while the bottom-up method does the opposite.
2. Classification-based information extraction: Classification-based IE is a 2-step process that approaches information extraction as a supervised learning classification task. First, machine learning models are trained on labeled

datasets to learn the connections between entities and their corresponding attributes. The models then predict labels for the entities they identify in new unstructured data.

3. Sequence labelling:Sequence labeling is the cornerstone of NLP and uses deep learning models to identify and label the components of an input sequence—for example, the words in a chatbot prompt. Sequence labeling is a critical NLP preprocessing step, helping ensure that neural networks know exactly how to interpret the input data. In addition to identifying entities in data, sequence labeling also captures dependencies between parts of an input sequence. Dependencies are a special type of relationship in which 1 part of an input sequence relies on another part to be correctly interpreted. Transformer models such as general-purpose technologies (GPTs) excel at capturing dependencies, which is why they can maintain contextual understanding across lengthy input sequences.

Information retrieval: Information retrieval is defined as the process of accessing and retrieving the most appropriate information from text based on a particular query given by the user, with the help of context-based indexing or metadata.
Google Search is the most famous example of information retrieval.
An information retrieval system searches a collection of natural language documents with the goal of retrieving exactly the set of documents that matches a user's question. They have their origin in library systems.
These systems assist users in finding the information they require but it does not attempt to deduce or generate answers. It tells about the existence and location of documents that might consist of the required information that is given to the user. The documents that satisfy the user's requirement are called relevant documents. If we have a perfect IR system, then it will retrieve only relevant documents.



From the above diagram, it is clear that a user who needs information will have to formulate a request in the form of a query in natural language. After that, the IR system will return output by retrieving the relevant output, in the form of documents, about the required information.

The step by step procedure of these systems are as follows:

- Indexing the collection of documents.
- Transforming the query in the same way as the document content is represented.
- Comparing the description of each document with that of the query.
- Listing the results in order of relevancy.

Retrieval Systems consist of mainly two processes:

1. Indexing
2. Matching

Indexing: It is the process of selecting terms to represent a text.Indexing involves:

1. Tokenization of string
2. Removing frequent words
3. Stemming

Two common Indexing Techniques:

- Boolean Model
- Vector space model

Matching :It is the process of finding a measure of similarity between two text representations.The relevance of a document is computed based on the following parameters:

1. TF: It stands for Term Frequency which is simply the number of times a given term appears in that document.

TF (i, j) = (count of ith term in jth document)/(total terms in jth document)

2. IDF: It stands for Inverse Document Frequency which is a measure of the general importance of the term.

IDF (i) = (total no. of documents)/(no. of documents containing ith term)

3. TF-IDF Score (i, j) = TF * IDF

Information Retrieval Models: Information retrieval models predict and explain what a user will find in relevance to the given query. These are basically a pattern that defines the above-mentioned aspects of retrieval procedure that we discussed in ad-hoc retrieval and consists of the following:

- A model for documents.
- A model for queries.
- A matching function that compares queries to documents.

Mathematically, a retrieval model consists of the following components:

- D: Representation for documents.
- R: Representation for queries.
- F: The modeling framework for D, Q along with the relationship between them.
- R (q, di): A ranking or similarity function that orders the documents with respect to the query.