International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 08 | Aug - 2021 ISSN: 2582-3930

Use Cases of Artificial Intelligence in Information Extraction

Manisha Mali[1], Dr. Neeraj Sharma [2]
[1]Research Scholar, Dept. of CSE, SSSUTMS, Sehore, MP
[2]Associate Professor, Dept. of CSE, SSSUTMS, Sehore, MP

Abstract - Artificial Intelligence (AI) is the branch of computer sciences that emphasizes the development of intelligent machines, thinking, and working like humans. For example, speech recognition, problem-solving, learning, and planning. It is a way of providing a computer or a robot or simply some machine to think, process information, and act on their own or in simple words providing the machines with an ability to think like a human. No doubt that recreating a human brain seems like a dream and still, we haven't achieved it yet but that one thing that is impossible to recreate is human consciousness. The AI includes several other fields like neural networks, deep learning, statistics, machine learning which is proving to be successful in various domains like security, research, robotics, voice recognition, transportation, and many more.AI is proving to not only reducing the workload of humans but also opening new fields that once only dreamed by us. So in this paper, we are going to first discuss AI followed by various use cases in NLP as well as other domains.

Key Words : Information Extraction, Artificial Intelligence, NLP, Use Cases

1. INTRODUCTION

Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving. The ideal characteristic of artificial intelligence is its ability to rationalize and take actions that have the best chance of achieving a specific goal. A subset of artificial intelligence is machine learning, which refers to the concept that computer programs can automatically learn from and adapt to new data without being assisted by humans. Deep learning techniques enable this automatic learning through the absorption of huge amounts of unstructured data such as text, images, or video. Natural language processing (NLP) is a major area of artificial intelligence research, which in its turn serves as a field of application and interaction of several other traditional AI areas. Until recently, the focus in AI applications in NLP was on knowledge representation, logical reasoning, and constraint satisfaction - first applied to semantics and later to grammar. In the last decade, a dramatic shift in the NLP research has led to the prevalence of very large-scale applications of statistical methods, such as machine learning and data mining. Naturally, this also opened the way to the learning and optimization methods that constitute the core of modern AI, most notably genetic algorithms and neural networks. Information Extraction (IE) is one of the main tasks in NLP. In this paper, we give an overview of the possible use cases of AI techniques and their combination in NLP as well as in other domains.

Information Extraction (IE) refers to the automatic extraction of structured information such as entities, relationships between entities, and attributes describing entities from unstructured sources. This facilitates much richer forms of queries on the plentiful unstructured sources than possible with keyword searches alone. The extraction of structure from noisy, unstructured sources is a challenging task, that has involved a veritable community of researchers for over two decades now. With roots in the Natural Language Processing (NLP) community, the topic of structure extraction now engages many different communities covering artificial intelligence, machine learning, information retrieval, database, web, and document analysis. Early extraction tasks were concentrated around the identification of named entities, like people and company names and relationships among them from natural language text. IE techniques enable us to distill Knowledge from the abundantly available unstructured content. Some of the basic IE methods include the automatic extraction of relevant entities from the text (e.g. places, dates, people, ...), understanding relations among them, building semantic resources (dictionaries, ontologies) to inform the extraction tasks, connecting extraction results to standard classification resources.

The rest of the paper is organized as follows. Fundamental artificial intelligence is explained in section II. Various use cases of IE and AI are described in section III, Finally, the conclusion and possible directions for future research have discoursed in Section IV.

2. FUNDAMENTALS OF ARTIFICIAL INTELLIGENCE

AI is a branch of computer science by which we can create intelligent machines which can behave like a human, think like humans, and able to make decisions on their own. Artificial Intelligence is composed of two words artificial which define something that's made by humans and intelligence which refers to the ability to think on its own, and hence this makes artificial intelligence "thinking power made by human". This field was formed with an idea that one day

machines will be able to think or in other words, we can replicate the feature that makes us human, our intelligence along with our consciousness. This may sound like a sci-fi or a new age

concept but the fact is, there has been a reference to such things in mythologies and several other texts, scriptures, and artifacts as well.AI have been a turning point not only in the field of research but has also been playing a keen role in revolutionizing industries and work as we know it today.

With an ultimate goal to create consciousness, AI goes through several stages of planning, reasoning, analyzing data, prediction of outcomes, and acting accordingly. AI also involves the use of statistics and probability and various other mathematical tools (neural networks and machine learning is mostly based on these).

2.A Elements Of AI

The three basic elements of AI are

1. The Human Element : This refers to a link between humans the machine. The machine or simply an algorithm understands a language of 0's and 1's which is hard to understand and process by humans. Thus comes in the role of an interactive interface that can take direction from the user, process, and then provide accurate results.

2. Knowledge Base : The AI works on analyzing data present in it. The more data is fed, the more efficient result will provide. The knowledge base also includes previous results to review and look for some sort of patterns. The AI analyzes the data, processes and compares large chunks of information to provide optimum results. It uses data algorithms and various other logic stored to find a solution to the problem present.

3. Algorithm set : The AI interface even though fed with a large chunk of data, it requires a certain set of instruction or algorithm to process and perform an operation on that data. Such algorithms are provided by programmers and data scientists working together using various mathematical tools such as statistics, probability, calculus, and algebra.

2.B. Components Of AI

Mimicking human intelligence or creating artificial intelligence requires several qualities to be focused on. These qualities rather than acting as a separate entity work as an array of components comprising the following items-

1. Learning : Learning can be categorized into various forms. One of the simplest of them all is the hit-and-try method. For example, while creating a drug to cure a disease various combinations of drugs are tested until a cure is found. The program analyzes and remembers each combination along with its effects. Thus when such a situation arises again the program can produce an answer immediately or as fast as it could produce. The process of memorizing simple items like phrases of words or solutions to a problem etc is known as Rote learning. Now suppose a program is made to learn tenses. So to make a word like "create" into the past tense the program has to encounter the word "created" at least once. A better way of dealing with this situation is by teaching it the "-ed rule" so that it can change the tense without encountering the word even once. This technique is known as Generalization.

2. Problem solving : Problems solving methods are divided into two categories special- purposes and general-purpose. Special-Purpose methods are specifically designed to solve specific problems like controlling the coolant release in a nuclear power plant. General-purpose, on the other hand, applies to a wide variety of problems like instructing a robot to move right, left, forward, backward. Further, the AI algorithm can be divided into two parts classifiers and controllers. Classifiers are pattern finders that find the closest pattern match to a problem. They act as a backbone for AI systems. Classifiers are fed with data and observation which gets stored in sets of predefined classes according to some pattern similarities among them. Thus more the data it's fed with, the more accurate result we will get. Controllers are instruction givers and takers who work on the results provided by the classifiers. For example, the purity of a mineral is tested and the classifiers which are already equipped with the data classify them, and then with the help of the classification, the controller separates the impure ones form those that are pure.

3. Logic and Reasoning : Reasoning refers to creating conclusions that are relevant to the task or situation at hand. There are two types of reasoning inductive and deductive reasoning. In inductive reasoning, general conclusions are drawn from specific instances like 'All men are strong'. In deductive reasoning, specifics are taken into consideration to draw a general conclusion. For example 'All cars have an engine. I have a car. So my car has an engine. Logic is used for information representation and problem-solving.

There are different types of logic used in AI algorithms.

• Propositional logic : Also known as sentential logic is a set of statements that can either be true or false.

• First-order logic : This type of logic is used to show facts about objects, their characteristics, and their relationship with each other.

• Fuzzy logic : It is a type of logic that allows the result(truth) to a statement between 0(false) and 1(true). This type of logic is used when the results are uncertain. These logics are mostly used in quantum computers.

• Subjective logic : This logic focuses on uncertain results more explicitly.

4. Perception : Perception means to scan and analyze objects or the environment along with their features and relationship through any sort of artificial organ or device at various depths and angles. Today artificial perception has advanced itself to enable cars on open roads or robots working as waiters at restaurants.

5. Language-understanding : Language is not only the words, we speak in certain forms as in English, Hindi, or any region, it also comprises our expression and gestures. The way we move our eyebrows or change our facial expressions according to our mood all comes under the domain of language. There are some AI algorithms these days that can read facial gestures and depict the mood of the user.

6. Neural network : The human brain is filled with hundreds of neurons that actively work every second in us and help us to analyze things, find a relation between them through electrical impulses. Collectively these neurons form the neural network of our brain. Similarly, AI has its neural network that is a series of algorithms that recognize the relationship between a set of data. These networks can adapt to changing input, thus the need to redesign the model each time a new set/type of data is entered is not needed. The neural network incorporates a set of layers of interconnected nodes. Each node is a perceptron (an algorithm for supervised learning of binary classifiers). In a multi-layered perceptron (MLP), perceptrons are interlinked in layers. The input layer receives input patterns while the output layer contains classification according to which input patterns may get arranged. There are hidden layers present that fine-tune the input sets until the neural network's error percentage reaches its lowest.

7. Machine learning : Machine learning is a part of artificial intelligence that allows the system to learn on its own. It is based on the observations, data, or instruction either with or without human intervention and observes patterns in previous decisions, and thus provides optimum results or approach towards a problem. For example, machine learning algorithms are used by various companies to study previous market trends to predict future sales and new innovative ways to raise their sales.

- Artificial intelligence refers to the simulation of human intelligence in machines.
- The goals of artificial intelligence include learning, reasoning, and perception.
- AI is being used across different industries including finance and healthcare.
- Weak AI tends to be simple and single-task oriented, while strong AI carries on tasks that are more complex and human-like.

2.C. Sectors Of Artificial Intelligence

1. Genetic programming : Just like humans evolved, the fittest genes mutated over multiple generations to evolve the human race. Similarly, inspired form the human evolution, genetic programming executes an algorithm that performs random mutation, crossovers, over multiple generations to find a solution for the user-defined task.
2. Data mining : Data mining is a process in which a large chunk of raw data is turned into a useful set of information with the help of AI algorithms which looks up patterns or trends in the data, filter it out, and brings in a clean set of useful information.
3. Pattern recognition : Pattern recognition is a process in which a machine-learning algorithm is used to classify and sort large databases based on information previously obtained or statistical records based on certain patterns present in the data.
4. Expert system : An expert system simulates the decision-making ability of a human expert. These systems are designed to take complex decisions, act the way that experts would have acted through complex reasoning algorithms, and using if-else structures.

2.D. Categorization of Artificial Intelligence

Artificial intelligence can be divided into two different categories: weak and strong.

1.Weak artificial intelligence embodies a system designed to carry out one particular job. Weak AI systems include video games such as the chess example from above and personal assistants such as Amazon's Alexa and Apple's Siri. You ask the assistant a question, it answers it for you.

2.Strong artificial intelligence systems are systems that carry on the tasks considered to be human-like. These tend to be more complex and complicated systems. They are programmed to handle situations in which they may be required to problem solve without having a person intervene. These kinds of systems can be found in applications like self-driving cars or hospital operating rooms.

3. USE CASES OF INFORMATION EXTRACTION AND ARTIFICIAL INTELLIGENCE

The applications for artificial intelligence and information extraction are endless. The technology can be applied to many different sectors and industries. Few sectors where Information Extraction can be used are as :

Business intelligence for enabling analysts to gather structured information from multiple sources;

Financial investigation for analysis and discovery of hidden relationships;

Scientific research for automated references discovery or relevant papers suggestions;
Media monitoring for mentions of companies, brands, people;
Healthcare records management for structuring and summarizing patients records;
Pharma research for drug discovery, adverse effects discovery, and clinical trials automated analysis.

Extending the existing practices of information extraction, semantic information extraction enables new types of applications such as 1. highlighting, indexing, and retrieval; 2. categorization and generation of more advanced metadata; and 3. smooth traversal between unstructured text and available relevant knowledge

AI is being tested and used in the healthcare industry for dosing drugs and different treatments in patients, and for surgical procedures in the operating room. Today, Artificial Intelligence and information extraction are a very popular subject that is widely discussed in the technology and business circles. Many experts and industry analysts argue that AI or machine learning is the future – but if we look around, we are convinced that it's not the future – it is the present.
Other examples of machines with artificial intelligence and information extraction include computers that play chess and self-driving cars. Artificial intelligence and information extraction also have applications in the financial industry, where it is used to detect and flag activity in banking and finance such as unusual debit card usage and large account deposits—all of which help a bank's fraud department. Applications for AI are also being used to help streamline and make trading easier. This is done by making the supply, demand, and pricing of securities easier to estimate.
With the advancement in technology, we are already connected to AI in one way or the other – whether it is Siri, Watson, or Alexa. Yes, the technology is in its initial phase and more and more companies are investing resources in machine learning, indicating a robust growth in AI products and apps soon. Unlike general perception, artificial intelligence is not limited to just the IT or technology industry; instead, it is being extensively used in other areas such as medical, business, education, law, and manufacturing.
The statistics give us an idea of the growth of AI. In 2014, more than $300 million was invested in AI startups, showing an increase of 300%, compared to the previous year (Bloomberg). By 2018, 6 billion connected devices will proactively ask for support. (Gartner). By the end of 2018, "customer digital assistants" will recognize customers by face and voice across channels and partners (Gartner). Artificial intelligence will replace 16% of American jobs by the end of the decade (Forrester). 15% of Apple phone owners' users use Siri's voice recognition capabilities. (BGR)
Here, we are listing down some of the very intelligent AI use cases that we are using today,

1. Siri : Siri is one of the most popular personal assistants offered by Apple in iPhone and iPad. The friendly female voice-activated assistant interacts with the user on a daily routine. She assists us to find information, get directions, send messages, make voice calls, open applications and add events to the calendar. Siri uses machine-learning technology to get smarter and capable-to-understand natural language questions and requests. It is surely one of the most iconic examples of the machine learning abilities of gadgets.

2. Tesla : Not only smartphones but automobiles are also shifting towards Artificial Intelligence. Tesla is something you are missing if you are a car geek. This is one of the best automobiles available until now. The car has not only been able to achieve many accolades but also features like self-driving, predictive capabilities, and absolute technological innovation. If you are a technology geek and dreamt of owning a car like shown in Hollywood movies, Tesla is one you need in your garage. The car is getting smarter day by day through over-the-air updates.

3. Cogito : Cogito originally co-founded by Dr. Sandy and Joshua is one of the best examples of the behavioral version to improve the intelligence of customer support representatives, currently on the market. The company is a synthesis of machine learning and behavioral science to enhance customer collaboration for phone professionals. Cogito applies to millions of voice calls that take place daily. The AI solution analyses the human voice and provides real-time guidance to enhance behavior.

4. Netflix : Netflix needs no introduction – it is a widely popular content-on-demand service that uses predictive technology to offer recommendations based on consumers' reactions, interests, choices, and behavior. The technology examines from several records to recommend movies based on your previous liking and reactions. It is turning more intelligent with each passing year. The only drawback of this technology is that small movies go unnoticed while big films grow and propagate on the platform. But as I wrote earlier, it is still improving and learning to be more intelligent.

5. Pandora : Pandora is one of the most popular and highly demanded tech solutions that exist. It is also called the DNA of music. Depending on 400 musical characteristics, the team of expert musicians individually analyses the song. The system is also good at recommending the track record for recommending songs that would never get noticed, despite people's liking.

6. Nest (Google) : Nest was one of the most famous and successful artificial intelligence start-ups and it was acquired by Google in 2014 for $3.2 billion. The Nest Learning Thermostat uses behavioral algorithms to save energy based on your behavior and schedule. It employs a very intelligent

machine learning process that learns the temperature you like and programs itself in about a week. Moreover, it will automatically turn off to save energy, if nobody is at home. It is a combination of both – artificial intelligence as well as Bluetooth low-energy because some components of this solution will use BLE services and solutions.

7. Boxever: Boxever is a company that heavily relies on machine learning to enhance the customer experience in the travel industry and conveys micro-moments or experiences that can please the customers. Boxover significantly improves customer engagement through machine learning and Artificial Intelligence to rule the playing field, helping customers to find new ways and make memorable journeys.

8. Flying Drones : The flying drones are already shipping products to customer's homes – though on a test mode. They indicate a powerful machine learning system that can translate the environment into a 3D model through sensors and video cameras. The sensors and cameras can notice the position of the drones in the room by attaching them to the ceiling. Trajectory generation algorithm guides the drone on how and where to move. Using a Wi-Fi system, we can control the drones and use them for specific purposes – product delivery, video-making, or news reporting.

9. Echo : Echo was launched by Amazon, which is getting smarter and adding new features. It is a revolutionary product that can help you to search the web for information, schedule appointments, shop, control lights, switches, thermostats, answers questions, reads audiobooks, reports traffic and weather, gives info on local businesses, provides sports scores and schedules, and more using the Alexa Voice Service.

10. Medical reports which include pathology laboratory and radiology reports are some of the most important medical documents that a diagnostician looks into, especially in the emergency context. They provide the emergency physicians with critical information regarding the condition of the patient and help the physicians take immediate action on urgent conditions. However, the reports are in the form of unstructured text, which makes them time-consuming for humans to interpret. In this Information Extraction and Artificial Intelligence play a very vital role. Applications are developed which can efficiently extract the clinically significant parts and their level of importance in medical reports, and classify the overall report into critical or non-critical categories which help doctors to identify potential high priority reports**.**

4. CONCLUSION

This paper focuses on various use cases of Information Extraction and Artificial Intelligence along with a brief introduction to Artificial Intelligence. AI is gaining popularity at a quicker pace; influencing the way we live, interact, and improve customer experience. There is much more to come in the coming years with more improvements, development, and governance. Though various use cases have been discussed in this paper, use cases are not limited to these. We can find use cases of AI in all domains like image processing, networking, IoT, civil engineering, Natural Language Processing, Text Mining, Health Care, and many more. There is a huge scope of research in extracting information from medical historical records and clinical data

REFERENCES

[1] Clarissa Castellã Xavier, Marlo Souza, “Extraction and Classification of Semantic Data from Twitter”, ACM ISBN 978-1-4503-5867-5/18/10.

[2] Deepti Ameta, Pokhar Mal Jat, “Information Extraction From Wikipedia Articles Using DeepDive”, 2018 International Conference on Communication, Information Computing Technology (ICCICT), Feb. 2-3, Mumbai, India

[3] Fuchun Peng, Andrew McCallum, “Accurate Information Extraction from Research Papers using Conditional Random Fields”

[4] Georg Fette, Maximilian Ertl, Anja Wörner, Peter Kluegl, Stefan Störk, Frank Puppe, “Information Extraction from Unstructured Electronic Health Records and Integration into a Data Warehouse”

[5] Hongmei Yang, Lin Li, Ridong Yang, Yi Zhou, “Towards Automated Knowledge Discovery of Hepatocellular Carcinoma: Extract Patient Information from Chinese Clinical Reports”, ICMHI 2018, June 8–10, 2018, Tsukuba, Japan

[6] Muawia Abdelmagid, Ali Ahmed and Mubarak Himmat, “Information Extraction Methods And Extraction Techniques In The Chemical Document's Contents: Survey”, ARPN Journal of Engineering and Applied Sciences, VOL. 10, NO. 3, FEBRUARY 2015

[7] Qi Zhu, Xiang Ren, Jingbo Shang, Yu Zhang, Frank F. Xu, Jiawei Han, “Open Information Extraction with Global Structure Constraints”, WWW 2018, April 23-27, 2018, Lyon, France

[8] Raymond J. Mooney and Un Yong Nahm, “Text Mining with Information Extraction”, Multilingualism and Electronic Language Management: Proceedings of the 4th International MIDP Colloquium, September 2003, Bloemfontein, South Africa, Daelemans, W., du Plessis, T., Snyman, C. and Teck, L. (Eds.), pp.141-160, Van Schaik Pub., South Africa, 2005

[9] Raymond J. Mooney and Razvan Bunescu, “Mining Knowledge from Text Using Information

Extraction", SIGKDD Explorations, Volume 7, Issue 1 - Page 3

[10] Varsha Pande, Dr. A. S. Khandelwal, "Information Extraction Technique: A Review", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727 PP 16-20