

ISBN: 978-81-965283-7-9



EMERGING TRENDS IN INFORMATION TECHNOLOGY Vol - I

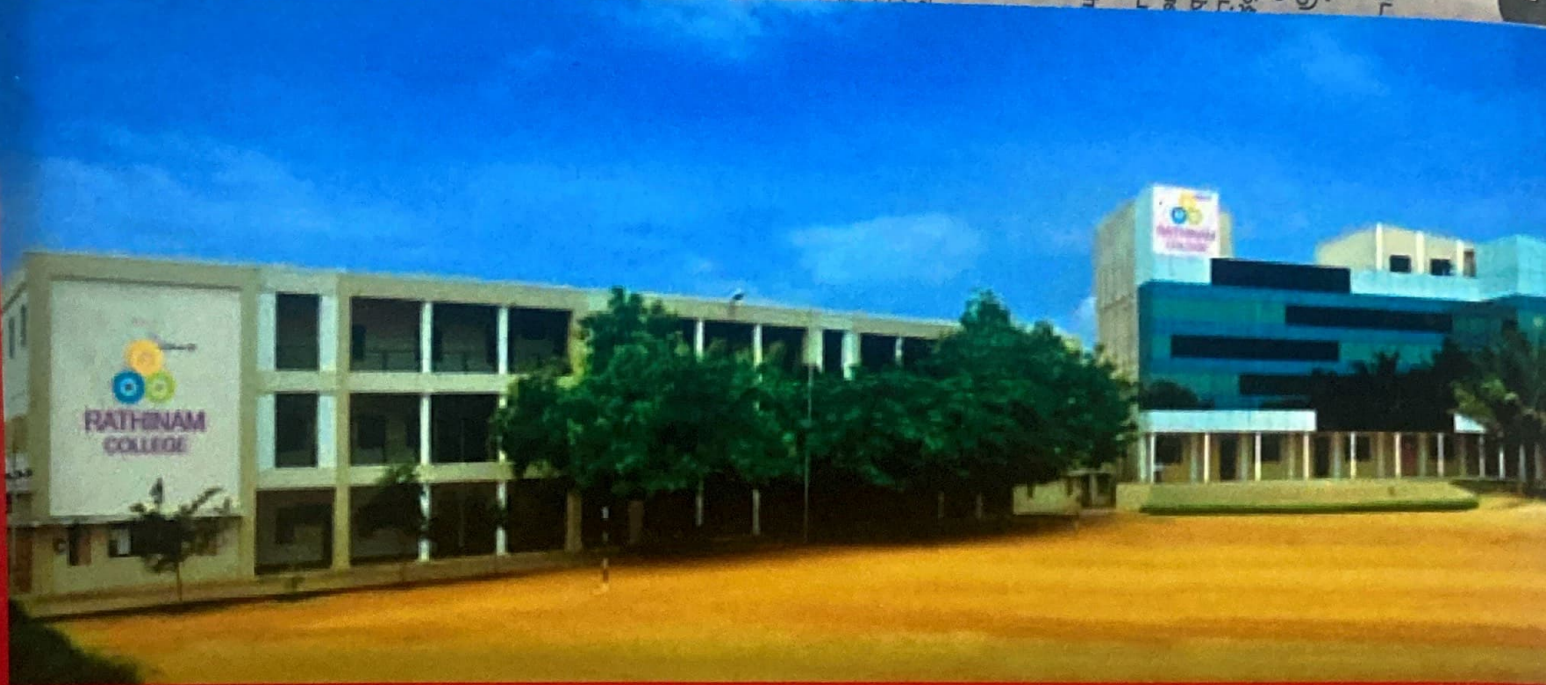
Editors

*Dr.T.Velumani | Dr.R.Lakshmi Devi
Dr.K.Prabavathy | Dr.V.Kalaimani*



RATHINAM
COLLEGE OF ARTS & SCIENCE





Dr. T. Velumani M.Sc(CS), M.Phil, MBA, B.Ed, M.Sc(Psychology), Ph.D is working as an Head & Assistant Professor in Department of Information Technology at Rathinam College of Arts and Science (Autonomous), Coimbatore 641 021 from 14-09-2020 to till date. He worked as Assistant Professor in Department of Computer Science at Kongu Arts and Science College (Autonomous), Erode from 11-06-2009 to 30-03-2020 TamilNadu, India. He has got more than 11 years of teaching experience. He has obtained his B.Sc(CT), M.Sc(CS), M.Phil and MBA Degrees from Periyar University Salem, Tamil Nadu, India. He has obtained his B.Ed Degree from Indira Gandhi National Open University (IGNOU) at Delhi, his M.Sc (Psychology) from Madras University Madras, TamilNadu, India. He has obtained his Ph.D in Computer Science from Manonmaniam Sundaranar University at Tirunelveli, TamilNadu, India. His area of interest is Image Processing research directions involved knowledge discovery, data mining, pattern recognition, knowledge-based neural networks, Software Engineering and IOT. He has Published More than 25 Articles in top most International journals like Scopus, Web of Science, UGC care Listed etc. He presented more than 20 papers in International conference. He is serving as an Editor and Reviewer in International journals. He Serving as a Research Supervisor in Computer Science under Bharathiar University. He acted as a Resource person and Guest lecturer in various topics in more than 20 Colleges. He acts as a BOS member in Many Colleges under Bharathiar University. He got 25 International Membership around world wide He filed 3 patents. He acted as a Chair session and key note speaker in World record Conference He Wrote 3 Books and Published with ISBN Number.



Dr. R. Lakshmi Devi, MCA., Ph.D, Assistant Professor, Women's Christian College, Chennai, Tamil Nadu, India. Dr. R. Lakshmi Devi is a dedicated and accomplished Assistant Professor with over a decade of teaching experience and a rich background in programming languages. Her expertise spans multiple domains, including Data Mining, Soft Computing, Image Processing, Android Programming, and IoT. She is a prolific researcher with a track record of publishing papers in esteemed international journals and conferences. Dr. Lakshmi Devi's commitment to excellence is underscored by her recent receipt of the Best Paper Award at the International Conference on Advancements in Intelligent Systems and Computing (ICAISC 2023). As an academic leader, she has served as a valued member of the Board of Studies, contributing to curriculum development and academic excellence. She also passionately mentors successful Hackathon projects, nurturing students' practical skills and innovative thinking. Dr. Lakshmi Devi is an inventor with a published patent titled "A System for Automatic Pallet Changer in Warehouse by Using Machine Learning," which was published 05-11-2021. She has authored a highly regarded book titled "Artificial Intelligence – A Modern Approach," which further establishes her expertise in the field of Artificial Intelligence. Her combination of teaching, research, and leadership accomplishments reflects her dedication to excellence in academia and technology.



Dr. K. Prabavathy is an accomplished academic professional with an impressive array of qualifications, boasting a M.Sc., MCA, M.Phil, and a Ph.D. in Computer Science. With over a decade of dedicated experience in research, teaching, and academic leadership, she currently serves as an Assistant Professor in the Department of Data Science at Sree Saraswathi Thyagaraja College (Autonomous), Pollachi, a position she has held since June 9, 2023. Dr. K. Prabavathy's illustrious career has seen her contribute her expertise as an Assistant Professor at prestigious institutions such as Rathinam College of Arts and Science, Coimbatore, and PSG College of Arts and Science. She has also served as a Lecturer at Sri Subramanya College of Engineering and Technology, Palani. She has supervised and produced 2 M.Phil and 2 Ph.D. scholars. Her scholarly endeavors are reflected in her extensive publication record, with over 20 articles featured in renowned UGC, Scopus Indexed, and Web of Science Journals. In addition to her written contributions, Dr. K. Prabavathy has presented her research findings at various conferences, both on the national and international stages, including overseas events. Her academic journey culminated in the successful completion of her Ph.D. at Manonmaniam Sundaranar University, Tirunelveli and completed M.Phil in PSG College of Arts and Science and MCA in Bharathiyar University. Dr. K. Prabavathy's commitment to the academic community extends beyond her own research, as she actively contributes as a reviewer for esteemed journals and serves as a valuable resource person for multiple conferences. In addition to her academic roles, Dr. Prabavathy has also excelled as a Software Trainer at Inova Infotech, Coimbatore, showcasing her versatility and expertise in the field of computer science. Her dedication to academia and her numerous accomplishments makes her a respected figure in her field. She is Proficient in Python, R, Java, specializing in data science, data mining, and data analytics, and adept at utilizing machine learning algorithms to address real-world challenges. She actively engages in community outreach and education initiatives, inspiring the next generation of computer scientists and researchers.



Dr. V. Kalaimani received the following degrees: MCA., Bharathiar University ME., Anna University, M.Phil., Bharathiar University, Ph.D., and Bharathiar University under the main field of Data Mining. Her main research fields are Bio-informatics and machine learning. Her teaching and research assignments are at various prominent institutions in and around Coimbatore like PSGR Krishnammal College for Women. She has published more than 10 research papers in Journals and Conferences. Her additional contributions are books, book chapters, special issues, monographs, and edited proceedings. She has five years of industrial experience and nineteen years of academic experience. Her memberships are as CSI and as a Life Member of ISCA. Her assignments included teaching and research supervising UG and PG Graduates and Research Scholars. She is a reviewer of Research Books, Journal, and Conference papers. She is a member of various committees like NAAC, NIRF, Organizing Committees of International Conferences and Academic In-charge, Assistant Controller of Examinations, and extension activities. She has offered several professional lecture programmes, to site a few are advanced techniques in computer networks and recent trends in cyber security tools and techniques.



Published by
SK Research Group of Companies
The International Journal & Book - SKRGC Publication



142, Periyar Nagar, Madakulam,
Madurai 625003, Tamil Nadu, India
skrgc.publisher@gmail.com / +91 9790120237
<https://skrgcpublishing.org/isbn/>

ISBN 978-81-965283-7-9



9 788196 528379

CONTENT

S. No.	Particulars	Page No.
1	AN INTERACTIVE GENETIC ALGORITHM FOR FRACTAL DESIGN SYSTEM <i>Dr.R.MANICKAM</i>	1
2	CONVOLUTIONAL NEURAL NETWORK BASED ALZHEIMER'S DISEASE DETECTION <i>Dr.T.VELUMANI</i>	8
3	AN APPLICATION FRAMEWORK OF AUGMENTED REALISTIC DISPLAY: A DYNAMIC VISUALIZATION PERSPECTIVE <i>Dr.VIJAYA SAMUNDEESWARI.V, Dr.R.LAKSHMI DEVI</i>	17
4	COMPARATIVE STUDY OF MACHINE LEARNING TECHNIQUES FOR MALWARE CLASSIFICATION: A COMPREHENSIVE ANALYSIS AND PERFORMANCE EVALUATION <i>Dr.K.PRABAVATHY</i>	25
5	CYBER SECURITY THREATS IN CLOUD: LITERATURE REVIEW <i>Mrs.J.GOKULAPRIYA, Mr.R.CHENNAPPAN</i>	30
6	PRIVACY AND CRYPTOLOGY IN PROTECTION OF SENSITIVE INFORMATION <i>Ms.R.PARAMESWARI, Ms.J.GOKULAPRIYA</i>	39
7	MACHINE LEARNING: A COOPERATIVE VIEW OF ALGORITHMS AND ITS APPLICATIONS <i>Dr.P.ARUNA DEVI, Dr.K.SASIREKHA</i>	44
8	IDENTIFICATION OF BRAIN TUMOR USING PROJECTION PURSUIT BIVARIATE MULTILAYER PERCEPTRED CLASSIFICATION <i>Ms.P.S.RENJENI, Dr.L.JAYA SINGH DHAS</i>	51
9	AN EXAMINATION ON ADAPTIVE ALGORITHMS <i>Ms.R.DEEPA, Ms.K.SAROJA</i>	61
10	FUZZIFIED CRISP RELATIVE CURE HIERARCHICAL CLUSTERING FOR TEMPORAL RELATIONAL DATA ANALYSIS <i>Dr.L.JAYA SINGH DHAS, Mrs.P.S.RENJENI</i>	65
11	IN-CITY ARTIFICIAL INTELLIGENCE, ROBOTICS, AND AUTOMATION <i>Dr.K.SANTHOSH KUMAR</i>	75
12	PERFORMANCE MEASURE OF SOCIAL MEDIA & CHANGING COMMUNICATION PATTERNS IN DIGITAL WORLD <i>Ms.P.THENMOZHI, Mr.P.MOHANRAJ</i>	81
13	BIG DATA ANALYTICS FOR INTERNET OF THINGS <i>Mr.R.MURUGESAN, Dr.M.DEEPA</i>	85
14	APPLICATIONS OF MACHINE LEARNING AND DEEP LEARNING <i>Mr.E.TAMIZHAN, Mr.M.P.VENKAT VIJAY</i>	91
15	PRIVACY AND CRYPTOLOGY <i>Ms.V.SOWMIYA, Ms.S.SRI RANJANI</i>	96
16	BIG DATA AND BIG DATA ANALYTICS IN HEALTHCARE <i>Ms.M.SOUNDARYA, Ms.V.PRABHA</i>	101
17	AI IN ROBOTICS <i>Mr.S.KALAICHELVAN, Ms.S.MADHUSRI</i>	108

18	MACHINE LEARNING AND DEEP LEARNING APPLICATIONS <i>Dr.K.KAVITHA, Dr.T.SARATHAMANI</i>	112
19	AI TRANSPORT MANAGEMENT SYSTEM <i>Ms. V.M.SAHIRA, Ms.K.P.SREELAKSHMI</i>	120
20	PARKINSONS DISEASE DETECTION USING CLASSIFICATION ALGORITHMS <i>Dr.R.SABIN BEGUM, Dr.A.HAJA ALAUDEEN</i>	122
21	SOCIAL MEDIA DATA MINING - NEED, TECHNIQUES AND CHALLENGES <i>Dr.T.SARATHAMANI, Dr.K.KAVITHA</i>	127
22	REMOTE SURVEILLANCE VIDEO ANALYTICS MODELS TO IMPROVE OBJECT ACCURACY <i>Dr.S.SELVI, Dr.D.MOHANAPRIYA</i>	133
23	COMPARATIVE STUDY OF ETHEREUM AND BITCOIN USING BLOCKCHAIN TECHNOLOGY <i>Mr.K.RAJESHKUMAR, Mr.R.RAMESH</i>	139
24	UNVEILING NEW HORIZONS IN RESEARCH: BIG DATA AND ITS ANALYTICAL ODYSSEY <i>Mr.M.VENKATESH SARAVANAKUMAR, Dr.M.SABIBULLAH</i>	143
25	SURVEY ON DATA MINING TECHNIQUES AND ITS APPLICATIONS <i>Ms.K.JAYADEVI, Ms.M.SWETHA</i>	152
26	HYDRO CHEMICAL PREDICTION USING CLASSIFICATION ALGORITHMS <i>Ms.U.SARALA DEVI, Ms.G.JOTHIPRIYA</i>	159
27	PROGNOSTICATION OF COVID-19 CASES UTILIZING AN ENHANCED K-NEAREST NEIGHBOR (KNN) APPROACH BY SMOTE MACHINE LEARNING <i>Ms.K.SINDHU, Ms.S.VANISRI</i>	166
28	BIG DATA ANALYTICS IN HEALTH CARE <i>Dr.V.JANE VARAMANI SULEKHA</i>	171
29	BLOCKCHAIN TECHNOLOGY AND CRYPTOCURRENCIES <i>Mr.S.ANANTHARAJ, Mr.T.KARTHIKEYAN</i>	176
30	AN OVERVIEW OF GENE EXPRESSION USING MICROARRAY <i>P.NANCY VINCENTINA MARY Dr.R.NAGARAJAN</i>	181
31	COMPARATIVE STUDY OF COMPRESSION ALGORITHM BASED ON BLOCK TRUNCATION CODING <i>Dr.T.ILAM PARITHI, Ms.K.ANTONY SUDHA</i>	188
32	PROFICIENT AUDIT SERVICE OUTSOURCING FOR DATA RELIABILITY IN CLOUDS <i>Mr.P.DHINESH KUMAR, Mrs.B.KAVIPRIYA</i>	191
33	MOBILE AD HOC NETWORK ROUTING STABLE PATH SELECTION USING ENRDP PROTOCOL <i>Mr.P.BALAMUTHUKUMAR, Dr.V.KATHIRESAN</i>	195
34	MULTI-AGENT SYSTEMS <i>Ms.M.ARCHANA, Ms.C.SRIMATHI</i>	201
35	ARTIFICIAL NEURAL NETWORK TECHNIQUES IN EDUCATION <i>Ms.N.VAISHNAVI, Dr.A.DEVI</i>	209

36	MITIGATING CONGESTION BASED DOS ATTACKS WITH AN ENHANCED AQM TECHNIQUE <i>Dr.S.THILAGAVATHI, Dr.S.NIVAS</i>	212
37	SURVEY ON MACHINE LEARNING AND DEEP LEARNIG APPLICATIONS <i>Mr.A.MOHAMMED ARSHATH, Ms.R.DEEPA</i>	216
38	SURVEY ON DATA MINING APPROACH FOR ANALYSIS AND PREDICTION OF STUDENT PERFORMANCE <i>Dr.P.LOGESWARI, Ms.J.GOKULAPRIYA</i>	223
39	MACHINE LEARNING AND DEEP LEARNING <i>Mr.S.SACHINDEVARAJAN, Ms.T.DEVAKI</i>	228
40	USING THE RADON TRANSFORM, IDENTIFY THE TAMIL CHARACTERS IN INSCRIPTION <i>Mr.C.BALAJI, Dr.P.LALITHA</i>	234
41	LIGHTWEIGHT CRYPTOGRAPHY FOR RESOURCE-CONSTRAINED DEVICES: A BALANCING ACT BETWEEN SECURITY AND EFFICIENCY <i>Ms.B.USHA</i>	239
42	UNDERSTANDING NATURAL LANGUAGE PROCESSING FOR TEXT AND SPEECH <i>Dr.A.INDUMATHI, Dr.R.TAMILSELVI</i>	242
43	MULTIPLE CHAOTIC 3D KEYS TO PROTECT YOUR DIGITAL IMAGES <i>Mr.S.RAJA, Mr.C.ASHOK KUMAR</i>	247
44	SWOT ANALYSIS ON ROLE OF WOMEN ENTREPRENEURS IN THE DIGITAL ECONOMY - SPECIAL REFERENCE TO COIMBATORE CITY <i>Dr.V.GOKULANACHIYAR, Dr.R.PUSHPA LATHA</i>	251
45	UNLOCKING LANGUAGE'S DIGITAL FRONTIER: A COMPREHENSIVE REVIEW OF NATURE LANGUAGE PROCESSING <i>Dr.S.SHANTHINI, Dr.R.DIVYA</i>	255
46	DNA CRYPTOGRAPHY FOR DATA SECURITY <i>Mrs.V.MAGESHWARI</i>	263
47	EXPERT SYSTEM AND BUSINESS INTELLIGENCE <i>Dr.N.SARANYA, Ms.P.K.PAVATHARINI</i>	268
48	BLOCK CHAIN <i>Ms.C.BHAVADHARINI, Ms.N.SINDHU</i>	272
49	A NOVEL STUDY OF DATA SECURITY IN REAL WORLD ENVIRONMENT <i>Dr.P.PRASANYA DEVI, Dr.M.RAVINDRAN</i>	276
50	BIG DATA AND BIG DATA ANALYTICS: CONCEPTS, TOOLS AND APPLICATIONS <i>Dr.C.NAGARANI, Mrs.R.SATHYA</i>	284

ID: 42

UNDERSTANDING NATURAL LANGUAGE PROCESSING FOR TEXT AND SPEECH

Dr.A.INDUMATHI

*Associate Professor and Head,
Department of Computer Technology,
Kongunadu Arts and Science College,
Coimbatore, Tamil Nadu, India.*

Dr.R.TAMILSELVI

*Associate Professor and Head,
Department of Artificial intelligence and Data Science,
VET Institute of Arts and Science College,
Erode, Tamil Nadu, India.*

ABSTRACT

Natural language processing (NLP) is the set of methods for making human language accessible to computers. In the past decade, natural language processing has become embedded in our daily lives: automatic machine translation is ubiquitous on the web and in social media; text classification keeps our email in boxes from collapsing under a deluge of spam; search engines have moved beyond string matching and network analysis to a high degree of linguistic sophistication; dialog systems provide an increasingly common and effective way to get and share information. These diverse applications are based on a common set of ideas, drawing on algorithms, linguistics, logic, statistics, and more. NLP is a type of Artificial Intelligence.

KEYWORDS: *Natural language processing (NLP), Artificial Intelligence, Text Classification*

I. INTRODUCTION

Natural Language Processing (NLP) – the application of software systems to examining, interpreting and accurately responding to speech is viewed as the next big leap in user interface technology. However, human speech is far more complex than most people realize. There are rules, such as spelling and grammar. In human language, the words can say one thing, but the context and the tone make those words mean something else. It takes humans a half a lifetime to learn the subtle nuances of language. Even then, there will be words and phrases that some of us don't understand. Then, there are further complications in understanding language, such as dialects

and colloquialisms. So, how can a computer that “thinks” in binary be programmed, line by line, to become fluent in any language? But with the help of Artificial Intelligence (AI), a computer can now learn how to understand a language.

What is Natural Language Processing?

Natural language processing, shortened to NLP, is the science of building machines that can read and write similar to how humans do. The goal of NLP is to organize massive amounts of text data (ie) millions of documents into information so humans can use it to solve problems. Since machines can perform tasks much faster than humans, we can work with machines to read massive amounts of text in a fraction of the time. While NLP is just *one* type of artificial intelligence that is becoming important for analysts to understand, it's a critical piece to solving lots of problems.

II. RELATIONSHIP BETWEEN AI, ML, DL AND NLP

Today, Technology has been revolutionized in our lives and its constantly changing and progressing. The most flourishing technologies include Artificial Intelligence, Machine Learning, Natural Language Processing, and Deep Learning. These terms are generally used together in some contexts but do not mean the same and are related to each other in some or the other way. ML is one of the leading areas of AI which allows computers to learn by themselves and NLP is a branch of AI. The below Figure illustrates how AI, ML, DL and NLP overlap with each other.

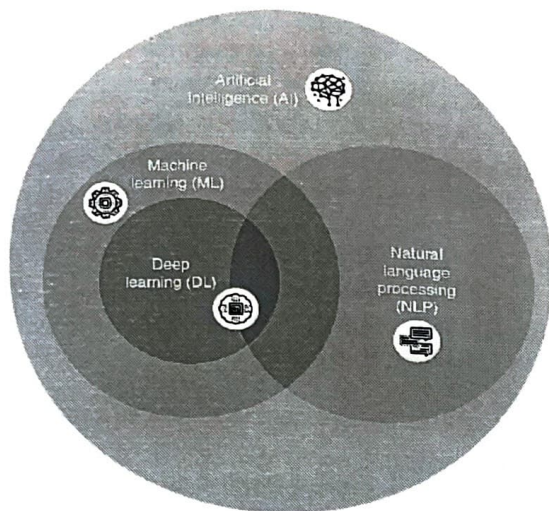


Fig.1 Relationship between AI, ML, DL and NLP

Artificial Intelligence (AI) is a broad umbrella field that is concerned with achieving human like intelligence using machines. It encompasses a wide range of subfields, including Machine Learning, Natural Language Processing, Computer Vision and Speech Recognition. The field also includes subfields such as reasoning, planning and search, which do not fall under either Machine Learning or Natural Language Processing.

Machine Learning (ML) is usually considered a subfield of Artificial Intelligence that is about improving computer algorithms through experience and data. This includes learning a general function that maps inputs to outputs based on past experience (Supervised Learning), drawing hidden patterns and structures from data (Unsupervised learning), and learning how to act in a dynamic environment based on indirect rewards (Reinforcement Learning). Supervised Machine Learning is the main paradigm for training NLP Models.

Deep Learning (DL) is a subfield of machine learning that usually uses deep Neural Networks. These Neural Network models are called "deep" because they consist of a number of layers. A Layer is a substructure of Neural Networks. With Many stacked Layers, deep Neural Networks can learn complex representations of data and can capture highly complicated relationships between the input and the output.

When the amount of available data and computational resources increases, NLP makes a heavier and heavier use of Machine Learning and Deep Learning. NLP applications and tasks are usually built on

top of Machine Learning Pipelines and trained from data. But NLP does not overlap with Machine Learning.

Computational Linguistics (CL) and Text Mining are other fields that are related to NLP. Computational Linguistics is a subfield of linguistics that uses computational approaches to study human language. CL is concerned with engineering approaches for making computers perform something useful related to language. Text Mining is a type of data mining targeted at textual data. Its focus is on drawing useful insights from unstructured textual data, which is a type of text data that is not formatted in a form that is easily interpretable by computers. Such data is usually collected from various sources like Web and Social media.

Why is NLP important?

In the past, computers could only work with structured languages. The language had to be precise and unambiguous. To program a computer to perform any task, you had to give it clear instructions. You could only use the limited number of commands that the computer understood. The syntax had to be perfect as well.

Even an end-user of a computer program needs to give the computer precise commands. Those who are old enough will remember that to use a PC you once had to know the common MS-DOS commands. That barrier was overcome, to a degree, with graphical user interfaces, such as Windows. Now, we can point to a file with a mouse, instead of having to know the name of the file. NLP promises to remove the need for being so precise. Instead of having to learn the computer's language, the computer will learn how to understand ours. A very basic application of NLP will be how we interface with computers.

How does Natural Language Processing work ?

NLP enables computers to understand natural language as humans do. Whether the language is spoken or written, natural language processing uses artificial intelligence to take real-world input, process it, and make sense of it in a way a computer can understand. Just as humans have different sensors such as ears to hear and eyes to see computers have programs to read and microphones to collect audio. And just as humans have a brain to process that input, computers have a program to process

their respective inputs. At some point in processing, the input is converted to code that the computer can understand. There are two main phases to natural language processing: data preprocessing and algorithm development.

Data preprocessing involves preparing and "cleaning" text data for machines to be able to analyze it. Preprocessing puts data in workable form and highlights features in the text that an algorithm can work with. There are several ways this can be done, including:

- **Tokenization.** This is when text is broken down into smaller units to work with.
- **Stop word removal.** This is when common words are removed from text so unique words that offer the most information about the text remain.
- **Lemmatization and stemming.** This is when words are reduced to their root forms to process.
- **Part-of-speech tagging.** This is when words are marked based on the part-of-speech they are such as nouns, verbs and adjectives.

Once the data has been pre-processed, an algorithm is developed to process it. There are many different natural language processing algorithms, but two main types are commonly used:

- **Rules-based system.** This system uses carefully designed linguistic rules. This approach was used early on in the development of natural language processing, and is still used.
- **Machine learning-based system.** Machine learning algorithms use statistical methods. They learn to perform tasks based on training data they are fed, and adjust their methods as more data is processed. Using a combination of machine learning, deep learning and neural networks, natural language processing algorithms hone their own rules through repeated processing and learning.

Techniques and Methods of Natural Language Processing

Syntax and semantic analysis are two main techniques used with natural language processing.

Syntax is the arrangement of words in a sentence to make grammatical sense. NLP uses syntax to assess meaning from a language based on grammatical rules. Syntax techniques include:

- **Parsing.** This is the grammatical analysis of a sentence. *Example:* A natural language processing algorithm is fed the sentence, "The dog barked." Parsing involves breaking this sentence into parts of speech i.e., dog = noun, barked = verb. This is useful for more complex downstream processing tasks.
- **Word segmentation.** This is the act of taking a string of text and deriving word forms from it. *Example:* A person scans a handwritten document into a computer. The algorithm would be able to analyze the page and recognize that the words are divided by white spaces.
- **Sentence breaking.** This places sentence boundaries in large texts. *Example:* A natural language processing algorithm is fed the text, "The dog barked. I woke up." The algorithm can recognize the period that splits up the sentences using sentence breaking.
- **Morphological segmentation.** This divides words into smaller parts called morphemes. *Example:* The word untestably would be broken into [[un][test]able][ly], where the algorithm recognizes "un," "test," "able" and "ly" as morphemes. This is especially useful in machine translation and speech recognition.
- **Stemming.** This divides words with inflection in them to root forms. *Example:* In the sentence, "The dog barked," the algorithm would be able to recognize the root of the word "barked" is "bark." This would be useful if a user was analyzing a text for all instances of the word bark, as well as all of its conjugations. The algorithm can see that they are essentially the same word even though the letters are different.

Semantics involves the use of and meaning behind words. Natural language processing applies algorithms to understand the meaning and structure of sentences. Semantics techniques include:

- **Word sense disambiguation.** This derives the meaning of a word based on context. *Example:* Consider the sentence, "The pig is in the pen." The word pen has different meanings. An algorithm using this method can understand that the use of the

word *pen* here refers to a fenced-in area, not a writing implement.

- **Named entity recognition.** This determines words that can be categorized into groups. *Example:* An algorithm using this method could analyze a news article and identify all mentions of a certain company or product. Using the semantics of the text, it would be able to differentiate between entities that are visually the same. For instance, in the sentence, "Daniel McDonald's son went to McDonald's and ordered a Happy Meal," the algorithm could recognize the two instances of "McDonald's" as two separate entities one a restaurant and one a person.
- **Natural language generation.** This uses a database to determine semantics behind words and generate new text. *Example:* An algorithm could automatically write a summary of findings from a business intelligence platform, mapping certain words and phrases to features of the data in the BI platform. Another example would be automatically generating news articles or tweets based on a certain body of text used for training.

Current approaches to natural language processing are based on deep learning, a type of AI that examines and uses patterns in data to improve a program's understanding. Deep learning models require massive amounts of labeled data for the natural language processing algorithm to train on and identify relevant correlations, and assembling this kind of big data set is one of the main hurdles to natural language processing.

Earlier approaches to natural language processing involved a more rules-based approach, where simpler machine learning algorithms were told what words and phrases to look for in text and given specific responses when those phrases appeared. But deep learning is a more flexible, intuitive approach in which algorithms learn to identify speakers' intent from many examples almost like how a child would learn human language.

Challenges of Natural Languages Processing

There are a number of challenges of natural language processing and most of them boil down to the fact that natural

language is ever evolving and always somewhat ambiguous. They include:

- **Precision.** Computers traditionally require humans to "speak" to them in a programming language that is precise, unambiguous and highly structured or through a limited number of clearly enunciated voice commands. Human speech, however, is not always precise; it is often ambiguous and the linguistic structure can depend on many complex variables, including slang, regional dialects and social context.
- **Tone of voice and inflection.** Natural language processing has not yet been perfected. For example, semantic analysis can still be a challenge. Other difficulties include the fact that the abstract use of language is typically tricky for programs to understand. For instance, natural language processing does not pick up sarcasm easily. These topics usually require understanding the words being used and their context in a conversation. As another example, a sentence can change meaning depending on which word or syllable the speaker puts stress on. NLP algorithms may miss the subtle, but important, tone changes in a person's voice when performing speech recognition. The tone and inflection of speech may also vary between different accents, which can be challenging for an algorithm to parse.
- **Evolving use of language.** Natural language processing is also challenged by the fact that language and the way people use it is continually changing. Although there are rules to language, none are written in stone, and they are subject to change over time. Hard computational rules that work now may become obsolete as the characteristics of real-world language change over time.

CONCLUSION

Natural language processing is a major leap forward in AI technology. It removes the communication barrier that has always existed between machines and humans. The potential for the application of NLP in business is immense. A computer could now answer customer queries and take orders. Even if the customer uses obscure language. NLP is likely to remove the need for input devices, such as the keyboard and mouse as well. NLP matters, because it is about to revolutionize the way that we communicate

with machines, and how they communicate with us.

REFERENCE

1. Liddy, E. D. (2001). Natural language processing.
2. Feldman, S. (1999). NLP Meets the Jabberwocky: Natural Language Processing in Information Retrieval. *ONLINE-WESTON THEN WILTON*-, 23, 62-73.
3. "Natural Language Processing." Natural Language Processing RSS. N.p., n.d. Web. 25 Mar. 2017
4. Hutchins, W. J. (1986). Machine translation: past, present, future (p. 66). Chichester: Ellis Horwood.
5. Hutchins, W. J. (Ed.). (2000). Early years in machine translation: memoirs and biographies of pioneers (Vol. 97). John Benjamins Publishing.
6. Green Jr, B. F., Wolf, A. K., Chomsky, C., & Laughery, K. (1961, May). Baseball: an automatic question-answerer. In Papers presented at the May 9-11, 1961, western joint IRE-AIEE-ACM computer conference (pp. 219-224). ACM.