Deep Science Publishing, 2025
https://doi.org/10.70593/978-93-7185-365-1_3

Chapter 3: Challenges and opportunities in modern artificial intelligence systems: A focus on natural language processing

Priyambada Swain

Infosys Limited

1. Introduction to Modern AI Systems

Recent artificial intelligence (AI) breakthroughs have led to systems capable of performing myriad tasks in industrial applications. Natural language processing (NLP) is a field of AI responsible for analysing and understanding natural language text and speech [1-2]. Several fundamental NLP problems exist, including text classification and sentiment analysis. These problems benefit from the latest advances in AI and deep learning, such as the revolutionary large language models trained on vast datasets [3-5]. The field of AI at the edge has also emerged, bringing new opportunities and challenges. NLP techniques have been applied in several industry use cases, such as healthcare, finance, and customer services. These advancements have made AI models particularly attractive for deployment on constrained devices at the edge. Nonetheless, a wide range of critical challenges must be addressed before these technologies can be widely adopted on a large scale.

Two significant challenges introduced by the development of NLP technologies are data privacy and ethical considerations. Data privacy and protection are directly related to GDPR, affecting the development and operation of AI at the edge. Recent research has attempted to address privacy issues related to AI at the edge. Ethical implications arise because several laws require transparency in AI decision-making processes [6-8]. Another important consideration is the risk of model bias, which also impacts model responses on the edge. Currently, there is no unified solution to these challenges.

2. Overview of Natural Language Processing

Natural language processing (NLP) is a crucial capability for machines to understand and interact with people. It deals with natural language inputs of various types and enables tasks such as speech recognition, dialogue and conversation systems, contextual translation, and contextual write support. NLP has experienced widespread success through the adoption of Transformer models on large amounts of data. Text classification is a broad NLP task that organizes or categorizes text. It has seen many important advances in recent years and is applied in a wide range of industries [7,9-10].

Sentiment analysis is an impactful NLP task that identifies and extracts the sentiment or polarity of a piece of text. The techniques used for sentiment analysis can be roughly divided into lexicon-based and machine-learning-based methods. Lexicon-based techniques rely on predefined sentiment lexicons such as AFINN, SentiWordNet, Opinion Lexicon, and SentiWords. Machine-learning techniques can be supervised-learning models such as Naïve Bayes and support vector machines and deep-learning models such as convolutional neural networks (CNNs) and long short- term memory (LSTM) networks. AI at the edge refers to the deployment of AI/ML on devices at the edge of the network.

3. Large Language Models: An In-Depth Analysis

Large language models (LLMs) are language models characterized by their extensive number of parameters and the availability of large training datasets. Owing to their substantial data requirements, their training often involves vast text corpora such as the Common Crawl, along with sources like books and Wikipedia. Although LLMs have been developed for diverse languages or even designed as multilingual systems, English commonly exhibits superior performance in evaluations [1,11-14]. Additionally, multilingual models have found applications beyond language tasks, including protein sequence annotation and forecasting epidemic trends. Despite the dominance of transformer-based architectures among open models, some large models employ alternative designs, such as CNN-based structures.

The evolution of LLMs has spurred growing interest in training methodologies and the characterization of their learning capabilities. The data and training expense associated with commercial LLMs has prompted investigations into knowledge transfer from large teacher models to smaller student models, as a means of enhancing infrastructure efficiency and reducing environmental impact. Moreover, the emerging application of LLMs at the edge, incorporating strategies like parameter-efficient fine-tuning, presents a distinct area of concern, characterized by performance trade-offs and resource constraints [13,15-17]. Their integration can be framed as a substitute or complement to traditional models in NLP applications such as text classification and sentiment analysis.

3.1. Architecture of Large Language Models

Large language models (LLMs) are neural network models with hundreds of billions of parameters specifically designed to process and generate text. They are currently at the forefront of natural language processing research due to their outstanding capabilities in a wide range of language-related tasks. Since the emergence of AlexNet in 2012 as the first deep learning model that significantly improved standard benchmarks, the field of natural language processing has witnessed rapid progress in the design of new architectures that can effectively generate human-like language or perform various natural language processing tasks. In fact, many texts mining algorithms, including those used for text classification and sentiment analysis, heavily rely on techniques designed for the training of LLMs. The size of these models, combined with the extremely large training datasets, enables them to learn highly effective representations of language, which can then be employed to achieve state-of-the-art results in the applications.

In recent years, transformer architectures have largely replaced traditional long short-term memory (LSTM) models due to the superior performance of the former in dedicated natural language processing tasks. More recently, GPT-2, GPT-3, and GPT-4 models have demonstrated convincing performance in text generation, as well as any model that is based on the transformer encoder-decoder architecture. Understanding the architecture of these large language models is crucial for grasping the foundation of modern natural language processing systems [18-20].

4. Advancements in Text Classification

Recent developments and breakthroughs in artificial intelligence (AI) have led to the creation of modern AI systems. These systems are implemented in a variety of novel real-world applications tackling deprivation of high-quality healthcare, poverty, mitigation of climate change, and many other problems. The field of a natural language processing has seen a constant stream of improvements and dedicated research [19,21-22]. The essence of natural language processing (NLP) is seen in the way it enables communication between humans and devices, including considerations of ethics and limitations.

Text classification is a task of assigning pre-defined class labels to documents. Classic algorithms like Support Vector Machines (SVM), k-Nearest Neighbours (KNN) and Naive Bayes (NB) combined with simple Bag-of-Words text representation have been studied for decades. Modern classification models like Convolutional Neural Networks (CNN), Long-Short Term Memory (LSTM) and others produce state-of-the-art results on many benchmarks. The creation of text classification systems for industrial use has also become relatively straightforward. Large corporations extensively apply these techniques for sorting documents or routing customer requests.

4.1. Techniques and Algorithms

Text classification is a transformative task within natural language processing that assigns predefined categories to textual documents. It involves mapping an input text to other values, such as the sentiment of a positive or negative movie review or the topic covered by a news article. Modern text classification algorithms support a wide variety of applications, including topic categorization, sentiment analysis, spam detection, suggestion mining, and opinion mining. This broad applicability underlines the significant influence of text classification techniques on diverse industries.

Contributions to text classification are widespread in the literature. An exemplary work proposes practical brute force solutions to train various deep learning models, including proposed Long Short-term Memory (LSTM) models, for large-scale industrial text classification problems. Long-distance dependencies in natural languages can be captured by the LSTM models and make use of the data balance to ensure robustness.

4.2. Applications in Industry

The efficacy of text classification algorithms is not confined to academic research but extends into various industrial applications. The importance of text classification and the effectiveness of different algorithms have been comprehensively demonstrated in earlier works [11,23-25]. There is a consensus that the advancement of artificial intelligence has triggered a new industrial revolution. Incorporating AI into enterprise information systems enables automation, rapid response to market changes, and cost reduction.

The significance of text classification creates demand for improved classifiers. Deep Convolutional Neural Networks (CNN) have been widely applied to text classification, proving their effectiveness under broad circumstances. Convolutional operations capture local features, and hierarchical operations extract high-level features from texts. Employment of CNN-based methods has enhanced the capability of text classifiers and strengthened the infrastructure for building AI companies.

5. Sentiment Analysis: Techniques and Tools

Sentiment analysis is a typical use in the NLP space and represents a process through which one can analyze text and detect the tone or polarity expressed in it. There is a wide range of algorithms, mainly falling into two categories: lexicon-based methods and machine learning-based methods. Both categories have been extensively researched, with many novel implementations applied to real-world scenarios.

Lexicon-based sentiment classification works rely on sentiment dictionaries. Obtaining sentiment categories and scores for text expressions allows for the evaluation of the sentiment orientation of the expression. Obtaining the final sentiment orientation of a sentence can be done by integrating the scores of all the component words [26-28]. Therefore, the key to sentiment classification for lexicon-based methods is how to effectively construct the sentiment dictionary. Machine learning-based classification methods regard sentiment classification as a text classification problem and apply classification models to the task. The main difference between the two methods lies in the sentiment feature information used in text representation. Typically, supervised machine learning classification methods use machine learning algorithms, such as Support Vector Machine (SVM) and Naive Bayes, through manually labeled sentiment texts, feature representation, and label information, and build classifiers to mine sentiment features.

5.1. Lexicon-Based Approaches

Sentiment analysis represents a subfield of NLP aiming to discern the emotional orientation—positive, negative, or neutral—in textual content. This orientation indicates the emotional response a text is likely to elicit in its readers. Commonly known as opinion mining, surveilling the tone of social media is a frequently highlighted application.

Lexicon-based techniques utilize dictionaries that conjoin every term or phrase with a specific emotional value. They operate on the principle that sentiments express themselves through influentials words, with successive terms and expressions bestowing unique strengths on the ensuing sentiment. An early contribution to this domain employed Pointwise Mutual Information (PMI) to score phrase-junctive sentiment.

5.2. Machine Learning Approaches

With the continuous growth of online resources and text information, the task of automatic sentiment polarity detection has acquired immense social and economic importance [29-32]. Broadly, approaches for sentiment classification can be categorized into lexicon-based and machine learning methods. Several recent surveys have addressed these methodologies in detail, encompassing numerous algorithms primarily designed to tackle binary two-class classification problems. Recognizing the limitations of manual data labeling, particularly its time consumption and high cost, extensive research efforts have focused on employing semi-supervised and unsupervised methods—such as semi-supervised learning, weakly supervised learning, self-supervised learning, active learning, reinforcement learning, and distant supervision—for polarity detection. Within the machine learning paradigm, classification methods span from decision trees and probabilistic algorithms to k-nearest neighbor (k-NN), support vector machines (SVMs), and artificial neural networks, each applied to the sentiment classification of tweets. Contemporary approaches often merge support vector machines with features like unigrams, bigrams, and lexical attributes, while also exploring the classification of Twitter data into three or more distinct categories [31,33-35]. In contrast, lexicon-based strategies define sentiment orientation through a sentiment lexicon composed of manually labeled opinion words. A hybrid approach that integrates both supervised and lexicon-based features has also been investigated for polarity classification. Recent explorations include leveraging convolutional neural networks (CNNs) and support vector regression for twitter sentiment analysis and predicting Bitcoin cryptocurrency price movements.

6. Challenges in Natural Language Processing

Artificial intelligence (AI) encompasses a broad spectrum of techniques and approaches, applying to neural networks, machine learning, DeepMind's reinforcement learning methodologies, convolutional neural networks, and more. Modern AI systems build upon significant innovations achieved in recent years across several approaches, leading to applications in robotics, computer vision, and autonomous agents, among others. Natural Language Processing (NLP), an important AI subfield, focuses intensively on the interaction between computers and human languages. It involves processing and analyzing large quantities of natural language data, enabling machines to comprehend and understand natural language at foundational levels [36-38]. Common NLP tasks include text classification, sentiment analysis, machine translation, and text summarization.

A particularly significant breakthrough for NLP applications, including chatbot agents, is large language models (LLMs) such as Generative Pretrained Transformer (GPT). Vast data and computational resources are utilized to train these models for natural language understanding and generation. LLMs underpin chatbots like ChatGPT and are readily accessible to users worldwide. Many companies are preparing to integrate LLMs into their businesses and products. Nonetheless, the large-scale training of language models raises concerns regarding immense energy consumption and the necessity for safeguards to prevent the misuse of these models for malicious actions.

6.1. Data Privacy and Ethical Considerations

Expanding the field of NLP to new application areas requires collecting and manipulating a huge amount of data, assuming that the obtained data are, all together, able to provide the best answer to a specific problem. For this reason, one of the major underground problems that a new AI solution must consider is the issue of privacy [1,39-41]. The sensitive subfields of data and artificial intelligence relate not only to data privacy and the ethical implications of this technology but also rise concerns about dangerous biases in AI. Collecting data is crucial when designing a product, but many of these data include sensitive information that should be protected by applying best practices for data engineering and data governance [42-44]. Data protection is vital, especially when facing analysis scenarios associated with laws such as the Regulation on European data protection (GDPR). These principles guide the way companies handle information. Despite these regulations, many challenges remain regarding how to safely explore, and use collected data. Practical yet powerful techniques have been developed to tackle privacy concerns, such as differential privacy, k-anonymity, and multiparty homomorphic encryption.

6.2. Bias and Fairness in AI Models

Human-generated data often enable success in AI and machine learning. Unsupervised learning eliminates this dependency—for example, by pre-training large language models. Despite this, biases in such data limit models' potential and raise ethical concerns [45-46]. These biases meaningfully impair model performance for some groups and can replicate and amplify social biases, leading to adverse outcomes in real-world applications when model decisions appear unethical or discriminatory.

Therefore, training data quality is crucial, as models generally mirror the patterns present in them. Acknowledge and measure potential risks of bias in the training corpus and ensure model sensitivity to bias. Long-standing debate exists about bias in NLP models and methods to mitigate and reduce it. Mitigation techniques strive to alter internal representations to produce lower bias scores, frequently listing the downstream task by which bias was assessed. However, bias is highly task-specific; for instance, a model trained to predict income level might be considered less biased under the definition that varies group outcomes, since job type—correlated with income—is also likely correlated with identity.

7. AI at the Edge: Opportunities and Limitations

Deploying powerful natural language processing algorithms on edge devices is an emerging approach intended to realise AI applications at the edge. Its goal is to address the critical challenges of data privacy and transmission latency associated with cloud-based AI processing [18,47-49]. The aims are: (1) to protect user's privacy by keeping local data on a deployed edge device; and (2) to reduce the latency of sending requests to a cloud and waiting for inference results. The deployment of NLP algorithms on edge devices promises the development of a fully smart ecosystem, but it also highlights the limitations of such devices, with their few resources and limited computational power. NLP algorithms have large input demands, and the data sets required for their creation are abundant in features that provide additional meaning during the implementation of machine learning algorithms [50-52].

AI at the edge is a prominent research topic in AI development. The growing technological interest in AI leads to corresponding growth in the computational times that need to be considered when executing an application. Data requirements for AI implementation in NLP remain large, and the requisite data sets exhibit a richness of features that demands more from the deployed algorithms. Finally, resources available on edge devices remain limited, with constrained memory and Computational Power. To address these limitations, careful database design and selection of appropriate NLP algorithms become necessary.

7.1. Deployment Strategies

The ever-growing demand for Natural Language Processing (NLP) services is converging with the ongoing migration of these services toward the network edge. The combination of these factors should be inherently synergistic, as the network edge offers the opportunity to achieve the low latency, enhance the privacy, and reduce the bandwidth requirements that must be supported by many NLP applications [53,54]. It is similarly anticipated that the broadening of NLP applications will enable new domains, with a corresponding increase in demand for NLP services becoming evident at the user level.

The outcome of these factors is that the deployment of NLP at the edge is emerging as a key component of the future application ecosystem. In this context, research is beginning to focus on how the deployment of NLP at the edge can be optimized to maximize service provisioning in support of live user demands. Two key questions that emerge in this context relate to the selection of the most appropriate algorithm for deployment, as well as the identification of the most appropriate device for hosting the chosen algorithm. Various candidate device types exist within the network edge, ranging from low-complexity, low-cost devices found in the sensory and user equipment to the relatively high-capacity smartphones and desktops located in the hands and offices of the end user. Key opportunities and challenges arise due to the variety of devices capable of hosting NLP services [55-57].

7.2. Performance Metrics and Evaluation

Performance metrics and evaluation methods play a pivotal role in the development of advanced text classification methodologies. The silhouette index has been applied to assess the appropriateness of clusters determined by the number of classes. Support vector machines equipped with a radial basis function kernel have produced high F-measure values, surpassing traditional classifiers such as random forest and k-nearest neighbor. The silhouette index is computed as follows:

$$\text{silhouette index} = \frac{b - a}{\max(a, b)}$$

where a is the average dissimilarity among all samples within a cluster and b is the average dissimilarity of all samples to their closest neighboring cluster. Support vector machines utilize the Gaussian Radial Basis Function (RBF) kernel given by:

$$K(x_i, x_j) = \exp\left(-\gamma \| x_i - x_j \|^2 \right)$$

Here, $\| x_i - x_j \|^2$ represents the squared Euclidean distance between two data points x_i and x_j, with γ operating as a regularization parameter.

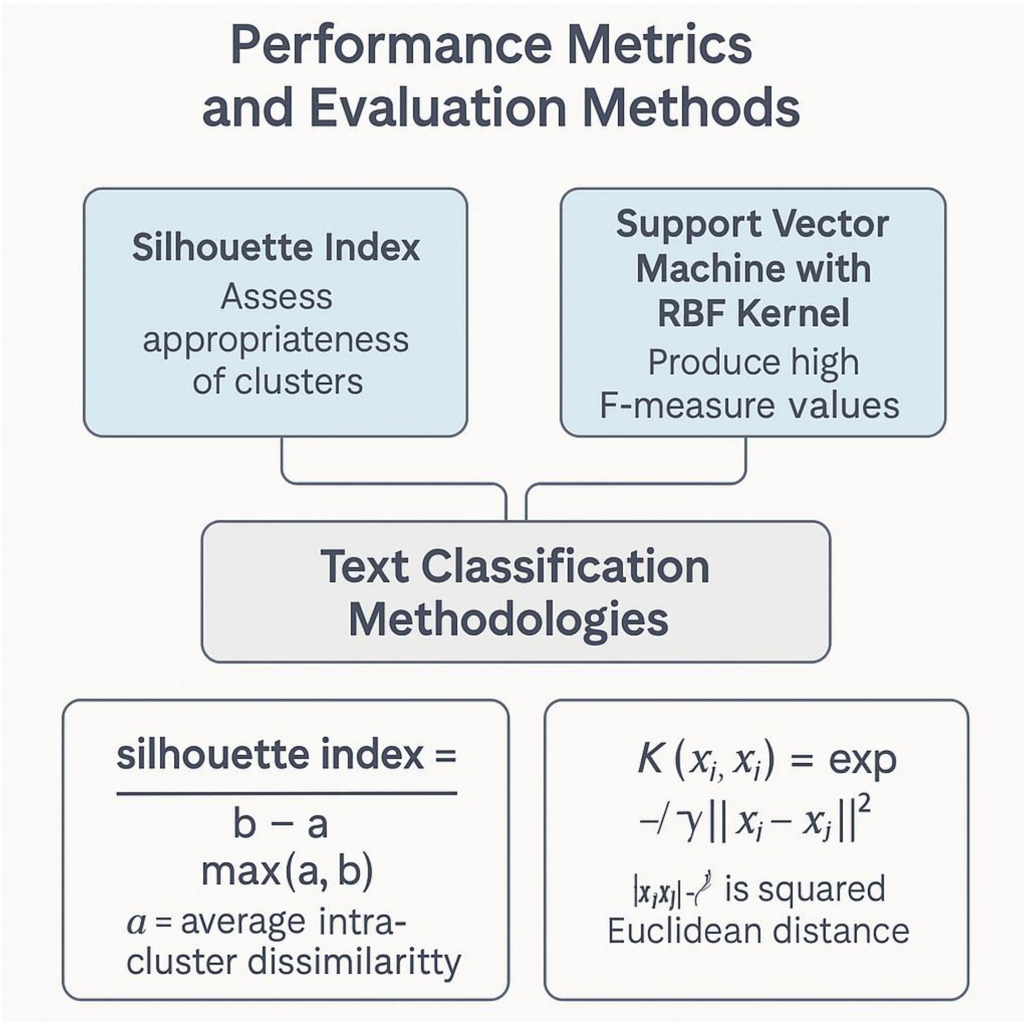

Fig1. Performance metrics and evaluation

8. Future Directions in Natural Language Processing

The field of Natural Language Processing (NLP) is presently undergoing a surge of novel research directions. Diversity in methodology and area of application has become the de facto standard, yet two approaches are particularly noteworthy. Transformer architectures are used in state-of-the-art models in NLP, leveraging the self-attention mechanism to produce strong representations of input tokens. Combining several of these large models results in mixtures of experts, where each model plays the role of an individual expert on a specific task. Additionally, the integration of multimodal data and applications inspired by ecology are gaining traction.

Nature presents a broad range of species, where each performs a unique ecological role—some produce energy, others process nutrients, and some act as top predators. This ecological viewpoint can serve as an inspiration for new NLP applications. Unlike current NLP approaches that largely treat the human language capability as focused solely on producing and consuming speech or text, ecological approaches aim for a more comprehensive representation of human behavior and cognition. Incorporating multimodal data and modeling the complexity of human interactions aligns language models more closely with neuroscience-based frameworks that describe human cognition and behavior.

8.1. Emerging Technologies

Modern artificial intelligent systems are recognized for their revolutionary breakthroughs stemming from recent innovations in the field of natural language processing [58,59]. Contrary to classical NLP techniques, which concentrate on the design and implementation of specific algorithms for each subtask, modern NLP systems focus on the development of a large foundation model for the training of massive amounts of data in a self-supervised fashion. The resulting representations from these large models can then be transferred and fine-tuned for use in particular applications, such as text classification and sentiment analysis.

A large language model (LLM) is a language model characterized by a remarkably high number of parameters and a substantial volume of training data. These transformers are designed to excel at self-supervised training and are optimized for natural language generation tasks. Consequently, they are endowed with an impressive capacity to generate novel content that is both contextually relevant and capable of mimicking human writing styles across diverse data modalities, including text, audio, and images. Beyond textual content creation, LLMs are considered generalist AI models due to their adaptability to virtually any textual task, encompassing translation, summarization, classification, and named entity recognition.

8.2. Interdisciplinary Applications

Natural language processing (NLP) applications involving human behavior and communication remain fertile ground for interdisciplinary research. The complex language-processing capabilities of humanoid robots require advances in NLP as well as in the human–robot interface, including speech, spatial awareness, and physical interaction. Investigations into trust between different groups of people may employ NLP to analyze surveys or social media posts, and NLP models can be used to predict users' political ideology, likely conventions of interaction when conversing with others, or moral foundations. These applications, exemplified by the Exploratory Interdisciplinary Research Portfolio of the Alberta Machine Intelligence Institute (Amii), highlight how NLP research integrates with diverse fields.

More sophisticated joint modelling techniques have also enabled improvements for analogous objectives in computer vision. Alongside these developments are growing expectations about the diverse capabilities of large, pre-trained language models at the core of many current NLP applications. These expectations can create challenging circumstances for developers and users alike.

9. Case Studies in AI Applications

This section presents case studies that illustrate the impact of innovative AI technologies in healthcare, finance and customer service.

9.1 AI for Patient Care in the Covid-19 Pandemic The COVID-19 pandemic rapidly spread worldwide from late 2019, presenting an urgent global health crisis. Healthcare services were challenged to identify the safest and most effective treatments and to rapidly diagnose patients in need. AI and Life Sciences Artificial Intelligence can support healthcare services through an increased understanding of the virus and through its mediation to patients [3,60-62]. The integration of technology and AI in purpose-built applications can assist both patients and healthcare workers, effectively bridging the gap between the two. To date, the minutely detailed structure of the SARS-CoV-2 proteome remains challenging despite recent advancements in structural biology. The necessity to better understand the virus has led the research community to explore alternative approaches such as Deep Learning and Natural Language Processing to leverage the information contained in the massive amount of scientific literature available since the outbreak.

9.2 AGGI Fin App: A Financial Management Assistant Full financial control is a difficult goal to achieve due to the many positive and negative events that can impact our finances during the year. Income and expenses are influenced not only by external events like a job loss or an increase in prices but also by our habits, both positive and negative. Moreover, each monetary decision that we make will affect our finances and, indirectly, our future. AGGI Fin App helps users in managing their own financial situation. In particular, the tool allows users to make financial decisions with greater foresight and control by proposing an analysis of all the income and expense items entered. It also presents projections and predictive models in order to support the planning of future activities and to evaluate the potential risks of leaving certain expenses unaddressed.

9.3 Are You Satisfied? Let the Machine Understand The satisfaction of a consumer is the fundamental point of any customer's activity and is capable of directly affecting their behaviour; but how do we know if a customer is not satisfied with a specific product or service? An approach often adopted to detect the satisfaction of consumers or users is through surveys or questionnaires; unfortunately, these approaches are not always used, especially since they require an important commitment of the consumers. Since the customer is the heart of any company, consumer satisfaction should always be obtained, and the decision taken by the consumers regarding a company, place, product, service and so on should also be monitored. If consumer dissatisfaction is left uncovered for a long period of time, the company may lose many potential customers, leading to significant financial damages. A possible solution is to easily access and collect customers' opinions.

9.1. Healthcare

Artificial intelligence (AI) is often heralded as a transformative technological innovation in healthcare [3,5]. The corresponding hype is fueled mainly by recent advances in natural language processing (NLP) and vision transformer models, which enabled remarkable enhancements and innovations in language- and vision-based tasks. These models have been applied to healthcare-related tasks such as medical text classification and report generation, and hold great promise as probable catalysts for the next wave of innovation and scientific discovery in medicine.

However, recent developments have also sparked an important conversation about and widespread recognition of AI's limitations and potential harms. Many question whether and how patient data should be used for training AI models given the privacy, confidentiality, security, and ownership concerns. The conversation on AI ethics has also been expanded to healthcare, focusing on preventing harm caused by AI models. In particular, bias mitigation and fairness promotion in AI have become active research areas, ensuring more equitable performance across demographic segments when AI is deployed in healthcare. The opportunity and challenges of deploying AI at the edge have attracted increased attention, especially during the past two years. Although the edge computing paradigm is still in its infancy, it is expected to emerge as a dominant service architecture for industrial online applications that require real-time responses, privacy preservation, robustness, and stable user experience.

9.2. Finance

Natural language processing is a widely studied application within the fields of fintech and financial risk management, primarily due to the potential value of predictive models but also the current regulatory mandates around risk identification. The prevailing incorporation of natural language processing techniques reveals a common reliance on classical methods such as term frequency-inverse document frequency (TF-IDF) weighting, n-gram features, pointwise mutual information, and target-encoded susceptible features, which are subsequently applied with general supervised classification models. Deep learning models for text classification—especially those utilizing architectures like convolutional neural networks to capture word, phrase, and sentence level features—and, more recently, transformer-based methods have begun to further advance the state of the art. These techniques aid in identifying the pertinence and nuances of financial disclosures.

Sentiment analysis is a standard application of interest within commercial and academic contexts; for example, consumer opinion analysis can provide key strategic insights. Approaches in this area typically divide between those based on lexica and those founded upon machine learning techniques. Lexicon-based methods leverage word- or phrase-level sentiment annotations that may span class labels or continuous polarity scores. Analysis can proceed via semantic orientation—comparing the presence of words from sets of positive and negative seed terms—or by summing oriented scores. In contrast, machine learning methods rely on annotated data to train a model to classify whole samples in terms of polarity.

9.3. Customer Service

There is a large number of customer service applications such as chat bots for websites or consumer-facing interactions. Although Natural Language Processing has made some significant advances in the past few years, these applications remain relevant and also present crucial challenges. The Covid-19 pandemic has generated a surge in the usage of contact centre conversational agents that respond to patients for common queries such as location and other details, thus reducing the need for human agent interaction in public services. This has led to a significant reduction in the need for human agents. General conversational systems for daily use can also trigger shopping online and assist customers in completing their journey in a variety of industries.

Neural Network language models decompose the probability of a sentence into products of conditional probabilities of the next word given the previous sequence of words. The recent pioneering models that set the new standards of natural language processing generate all these conditional probabilities through transformer architectures with self-attention. Large Language Models have millions or billions of parameters. They require a huge amount of text data to train, together with substantial computational resources and expert knowledge. Large Language Models have shown competitive performance in many important natural language processing tasks such as text classification, sentiment analysis, and question answering. Furthermore, it is demonstrated that with the technique of fine-tuning, Large Language Models can effectively adapt to specific tasks.

10. Ethical Implications of AI in Society

The ethics of artificial intelligence (AI) encompasses a broad range of issues, from the governance of lethal autonomous weapons and operation of driverless cars to human use of AI in everyday tasks and connections to AI-empowered robots and cyborgs. Emerging concerns for groups involved in optimization and training of generative AI models relate to data privacy, data governance, and the ownership of copyrighted content. Other predominant concerns include fairness, bias, transparency, and accountability. The importance of these concerns depends on the role of the AI system and the stakeholders affected. An AI system used for exploratory data analysis in medical research becomes increasingly important as human decisions are gradually more influenced by its hints. A deployed facial recognition system can have a large impact on individual lives and thus requires a much higher level of scrutiny. Recognizing the nuanced role AI plays in various aspects of life helps us better address the ethical questions arising along such a spectrum. Rawlsian social contract theory has been used to propose that AI should be accorded special consideration, or rights, within the broader context of AI society.

From a normative perspective, ethical inquiries explore how things should or ought to be, what they should or ought to be like, and which things are good or bad. Ways of acting ethically or morally emerge from such inquiries. Social effects of AI are indeed an area of strong importance and growing concern that has been the subject of substantial discussion. In research, there have been demonstrated instances of deep learning systems reproducing racist or sexist biases found in their training data, and adversarial attacks can deceive models into generating incorrect or sexist responses. Because text generated by transformers resembles human speech patterns, it is difficult to detect when these models generate harmful or biased content. A solution to this ongoing problem has been the development of reinforcement learning-based models that avoid bias by conversationally interacting with humans, using feedback obtained through such interactions as part of the optimization process.

11. Regulatory Frameworks and AI Governance

Regulatory frameworks and AI governance represent an essential topic during the development and deployment of Artificial Intelligence and Natural Language Processing techniques in everyday applications. Recent breakthroughs in Artificial Intelligence, particularly in Deep Learning, have significantly enhanced the abilities of Natural Language Processing. For instance, text classification can now be performed using very large language models, such as Google's BERT, OpenAI's GPT-x, BLOOM, among others, achieving very high accuracy and robustness. Nevertheless, the recent progress in AI has brought about substantial challenges, such as the demand for large amounts of training data and its associated privacy requirements.

Data protection imposes strict constraints regarding the collection, storage, and processing of personal data, leading to various ethical concerns. These concerns become even more critical when working with large language models, as their training necessitates vast quantities of varied data. The protected nature of personal data limits the availability of such information, which may contain sensitive or identifiable details. An identified individual is considered a natural person for whom, directly or indirectly, identification is feasible. Therefore, requirements related to data protection that must be met include: the implementation of a valid legal basis; adherence to the guiding principles of personal data processing; recognition of the data subjects' rights; and the adoption of appropriate technical and organizational security measures.

12. Collaborative Efforts in AI Research

A large language model (LLM) is an artificial intelligence (AI) model with many parameters (typically hundreds of millions to hundreds of billions) trained with extremely vast quantities of text data. LLMs have demonstrated strong natural language processing (NLP) capabilities and are a recent break-through in the field of NLP. LLMs are the most notable category of foundation model.

Text classification is an active research area in Industry 4.0 and is one of the most frequently adopted Artificial Intelligence techniques. Email, web page, and job title categorization are some common text classification applications. The review paper presents the advanced techniques used for various text classification tasks. Sentiment analysis of texts is one of the upcoming fields in the context of social network service data. Both lexicon-based and machine learning techniques are used for sentiment analysis. The support vector machine is widely used for sentiment classification.

13. Conclusion

Artificial Intelligence (AI) encompasses a broad range of branches that have experienced immense progress in recent years. Modern AI systems rely on new and innovative paradigms that allow them to learn, understand, and perform more complex and difficult tasks in many areas of human activity. Among the branches of AI, Natural Language Processing (NLP) has seen rapid development due to the generation of models with increasingly superior results, capable of solving various tasks in different fields of knowledge. The creation of models with billions of parameters allows for the extraction of deep contextualized knowledge; this includes the domain of text classification, which can be adapted to several NLP tasks, as for example Aspect-Based Sentiment Analysis. Large language models also play an important role in sentence-level sentiment tasks, yielding robust solutions.

The rapid development of AI has propagated to a wide range of areas. As a result of this advance, important topics that need to be addressed have arisen. In the area of text classification, topics such as data privacy and the execution of NLP tasks in edge environments are highlighted. While the former is a critical aspect in model training, the latter is fundamental for achieving good levels of accuracy in inference operations, since the more parameters an AI model has, the longer it takes to provide a result. Regarding the latter, the implications on AI of text classification and sentiment analysis in the areas of ethics and bias/fairness are also prominent. Despite its limitations, with careful handling and complementary measures, AI at the Edge is a highly attractive alternative to overcome problems related to speed, privacy, connectivity, and more, making it an essential analytical tool for future NLP applications.

References

[1] Soori M, Arezoo B, Dastres R. Machine learning and artificial intelligence in CNC machine tools, a review. Sustainable Manufacturing and Service Economics. 2023 Apr 1;2:100009.

[2] Arel I, Rose DC, Karnowski TP. Deep machine learning-a new frontier in artificial intelligence research [research frontier]. IEEE computational intelligence magazine. 2010 Oct 18;5(4):13-8.

[3] Nichols JA, Herbert Chan HW, Baker MA. Machine learning: applications of artificial intelligence to imaging and diagnosis. Biophysical reviews. 2019 Feb 7;11(1):111-8.

[4] Jones LD, Golan D, Hanna SA, Ramachandran M. Artificial intelligence, machine learning and the evolution of healthcare: A bright future or cause for concern?. Bone & joint research. 2018 Mar 1;7(3):223-5.

[5] Nuka ST. Next-Frontier Medical Devices and Embedded Systems: Harnessing Biomedical Engineering, Artificial Intelligence, and Cloud-Powered Big Data Analytics for Smarter Healthcare Solutions. Deep Science Publishing; 2025 Jun 6.

[6] Hyder Z, Siau K, Nah F. Artificial intelligence, machine learning, and autonomous technologies in mining industry. InResearch Anthology on Cross-Disciplinary Designs and Applications of Automation 2022 (pp. 478-492). IGI Global Scientific Publishing.

[7] Maguluri KK. Machine learning algorithms in personalized treatment planning. How Artificial Intelligence is Transforming Healthcare IT: Applications in Diagnostics, Treatment Planning, and Patient Monitoring. 2025 Jan 10:33.

[8] Mijwil MM, Aggarwal K, Sonia S, Al-Mistarehi AH, Alomari S, Gök M, Zein Alaabdin AM, Abdulrhman SH. Has the Future Started? The Current Growth of ArtificialIntelligence, Machine Learning, and Deep Learning. Iraqi Journal for Computer Science and Mathematics. 2022;3(1):13.

[9] Panda S. Observability in DevOps: Integrating AWS X-Ray, CloudWatch, and Open Telemetry. International Journal of Computer Application. 2025 Jan 1.

[10] Dunjko V, Briegel HJ. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. Reports on Progress in Physics. 2018 Jun 19;81(7):074001.

[11] Cioffi R, Travaglioni M, Piscitelli G, Petrillo A, De Felice F. Artificial intelligence and machine learning applications in smart production: Progress, trends, and directions. Sustainability. 2020 Jan 8;12(2):492.

[12] Tapeh AT, Naser MZ. Artificial intelligence, machine learning, and deep learning in structural engineering: a scientometrics review of trends and best practices. Archives of Computational Methods in Engineering. 2023 Jan;30(1):115-59.

[13] Panda SP. Artificial Intelligence Across Borders: Transforming Industries Through Intelligent Innovation. Deep Science Publishing; 2025 Jun 6.

[14] Theodosiou AA, Read RC. Artificial intelligence, machine learning and deep learning: Potential resources for the infection clinician. Journal of Infection. 2023 Oct 1;87(4):287-94.

[15] Kersting K. Machine learning and artificial intelligence: two fellow travelers on the quest for intelligent behavior in machines. Frontiers in big Data. 2018 Nov 19;1:6.

[16] Panda SP, Muppala M, Koneti SB. The Contribution of AI in Climate Modeling and Sustainable Decision-Making. Available at SSRN 5283619. 2025 Jun 1.

[17] Goldenberg SL, Nir G, Salcudean SE. A new era: artificial intelligence and machine learning in prostate cancer. Nature Reviews Urology. 2019 Jul;16(7):391-403.

[18] Shivadekar S. Artificial Intelligence for Cognitive Systems: Deep Learning, Neuro-symbolic Integration, and Human-Centric Intelligence. Deep Science Publishing; 2025 Jun 30.

[19] Sultan AS, Elgharib MA, Tavares T, Jessri M, Basile JR. The use of artificial intelligence, machine learning and deep learning in oncologic histopathology. Journal of Oral Pathology & Medicine. 2020 Oct;49(9):849-56.

[20] Guo K, Yang Z, Yu CH, Buehler MJ. Artificial intelligence and machine learning in design of mechanical materials. Materials Horizons. 2021;8(4):1153-72.

[21] Panda SP. Securing 5G Critical Interfaces: A Zero Trust Approach for Next-Generation Network Resilience. In2025 12th International Conference on Information Technology (ICIT) 2025 May 27 (pp. 141-146). IEEE.

[22] Rane J, Chaudhari RA, Rane NL. Data Privacy and Information Security in Deep Learning Applications: Risk Assessment and Patient Safety Protocols for Big Data Analytics. Ethical Considerations and Bias Detection in Artificial Intelligence/Machine Learning Applications. 2025 Jul 10:54.

[23] Mohapatra PS. Artificial Intelligence-Driven Test Case Generation in Software Development. Intelligent Assurance: Artificial Intelligence-Powered Software Testing in the Modern Development Lifecycle. 2025 Jul 27:38.

[24] Mich L. Artificial intelligence and machine learning. InHandbook of e-Tourism 2020 Sep 1 (pp. 1-21). Cham: Springer International Publishing.

[25] Mohapatra PS. Artificial Intelligence-Powered Software Testing: Challenges, Ethics, and Future Directions. Intelligent Assurance: Artificial Intelligence-Powered Software Testing in the Modern Development Lifecycle. 2025 Jul 27:163.

[26] Suura SR. Integrating Artificial Intelligence, Machine Learning, and Big Data with Genetic Testing and Genomic Medicine to Enable Earlier, Personalized Health Interventions. Deep Science Publishing; 2025 Apr 13.

[27] Panda SP. Relational, NoSQL, and Artificial Intelligence-Integrated Database Architectures: Foundations, Cloud Platforms, and Regulatory-Compliant Systems. Deep Science Publishing; 2025 Jun 22.
[28] Zhu S, Gilbert M, Chetty I, Siddiqui F. The 2021 landscape of FDA-approved artificial intelligence/machine learning-enabled medical devices: an analysis of the characteristics and intended use. International journal of medical informatics. 2022 Sep 1;165:104828.
[29] Mukhamediev RI, Popova Y, Kuchin Y, Zaitseva E, Kalimoldayev A, Symagulov A, Levashenko V, Abdoldina F, Gopejenko V, Yakunin K, Muhamedijeva E. Review of artificial intelligence and machine learning technologies: classification, restrictions, opportunities and challenges. Mathematics. 2022 Jul 22;10(15):2552.
[30] Gerke S, Babic B, Evgeniou T, Cohen IG. The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. NPJ digital medicine. 2020 Apr 7;3(1):53.
[31] Paramesha M, Rane N, Rane J. Big data analytics, artificial intelligence, machine learning, internet of things, and blockchain for enhanced business intelligence. Artificial Intelligence, Machine Learning, Internet of Things, and Blockchain for Enhanced Business Intelligence (June 6, 2024). 2024 Jun 6.
[32] Arel I, Rose DC, Karnowski TP. Deep machine learning-a new frontier in artificial intelligence research [research frontier]. IEEE computational intelligence magazine. 2010 Oct 18;5(4):13-8.
[33] Paramesha M, Rane N, Rane J. Artificial intelligence, machine learning, deep learning, and blockchain in financial and banking services: A comprehensive review. Machine Learning, Deep Learning, and Blockchain in Financial and Banking Services: A Comprehensive Review (June 6, 2024). 2024 Jun 6.
[34] Ongsulee P. Artificial intelligence, machine learning and deep learning. In2017 15th international conference on ICT and knowledge engineering (ICT&KE) 2017 Nov 22 (pp. 1-6). IEEE.
[35] Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. Journal of global health. 2018 Oct 21;8(2):020303.
[36] Das S, Dey A, Pal A, Roy N. Applications of artificial intelligence in machine learning: review and prospect. International Journal of Computer Applications. 2015 Jan 1;115(9).
[37] Balyen L, Peto T. Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology. The Asia-Pacific Journal of Ophthalmology. 2019 May 1;8(3):264-72.
[38] Michalski RS, Carbonell JG, Mitchell TM, editors. Machine learning: An artificial intelligence approach. Springer Science & Business Media; 2013 Apr 17.
[39] Soori M, Arezoo B, Dastres R. Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cognitive Robotics. 2023 Jan 1;3:54-70.
[40] Ghahramani Z. Probabilistic machine learning and artificial intelligence. Nature. 2015 May 28;521(7553):452-9.
[41] Jakhar D, Kaur I. Artificial intelligence, machine learning and deep learning: definitions and differences. Clinical and experimental dermatology. 2020 Jan 1;45(1):131-2.
[42] Siau K, Wang W. Building trust in artificial intelligence, machine learning, and robotics. Cutter business technology journal. 2018;31(2):47.

[43] Panda SP. The Evolution and Defense Against Social Engineering and Phishing Attacks. International Journal of Science and Research (IJSR). 2025 Jan 1.
[44] Kühl N, Schemmer M, Goutier M, Satzger G. Artificial intelligence and machine learning. Electronic Markets. 2022 Dec;32(4):2235-44.
[45] Karthikeyan A, Priyakumar UD. Artificial intelligence: machine learning for chemical sciences. Journal of Chemical Sciences. 2022 Mar;134(1):2.
[46] Helm JM, Swiergosz AM, Haeberle HS, Karnuta JM, Schaffer JL, Krebs VE, Spitzer AI, Ramkumar PN. Machine learning and artificial intelligence: definitions, applications, and future directions. Current reviews in musculoskeletal medicine. 2020 Feb;13(1):69-76.
[47] Shivadekar S, Halem M, Yeah Y, Vibhute S. Edge AI cosmos blockchain distributed network for precise ablh detection. Multimedia tools and applications. 2024 Aug;83(27):69083-109.
[48] Helm JM, Swiergosz AM, Haeberle HS, Karnuta JM, Schaffer JL, Krebs VE, Spitzer AI, Ramkumar PN. Machine learning and artificial intelligence: definitions, applications, and future directions. Current reviews in musculoskeletal medicine. 2020 Feb;13(1):69-76.
[49] Panda SP. Augmented and Virtual Reality in Intelligent Systems. Available at SSRN. 2021 Apr 16.
[50] Bhat M, Rabindranath M, Chara BS, Simonetto DA. Artificial intelligence, machine learning, and deep learning in liver transplantation. Journal of hepatology. 2023 Jun 1;78(6):1216-33.
[51] Tyagi AK, Chahal P. Artificial intelligence and machine learning algorithms. InChallenges and applications for implementing machine learning in computer vision 2020 (pp. 188-219). IGI Global Scientific Publishing.
[52] Shivadekar S, Kataria DB, Hundekar S, Wanjale K, Balpande VP, Suryawanshi R. Deep learning based image classification of lungs radiography for detecting covid-19 using a deep cnn and resnet 50. International Journal of Intelligent Systems and Applications in Engineering. 2023;11:241-50.
[53] Kühl N, Goutier M, Hirt R, Satzger G. Machine learning in artificial intelligence: Towards a common understanding. arXiv preprint arXiv:2004.04686. 2020 Mar 27.
[54] Rane J, Chaudhari RA, Rane NL. Data Analysis and Information Processing Frameworks for Ethical Artificial Intelligence Implementation: Machine-Learning Algorithm Validation in Clinical Research Settings. Ethical Considerations and Bias Detection in Artificial Intelligence/Machine Learning Applications. 2025 Jul 10:192.
[55] Neapolitan RE, Jiang X. Artificial intelligence: With an introduction to machine learning. CRC press; 2018 Mar 12.
[56] Rane NL, Paramesha M, Choudhary SP, Rane J. Artificial intelligence, machine learning, and deep learning for advanced business strategies: a review. Partners Universal International Innovation Journal. 2024 Jun 25;2(3):147-71.
[57] Rubinger L, Gazendam A, Ekhtiari S, Bhandari M. Machine learning and artificial intelligence in research and healthcare. Injury. 2023 May 1;54:S69-73.
[58] Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. Molecular diversity. 2021 Aug;25(3):1315-60.

[59] Entezari A, Aslani A, Zahedi R, Noorollahi Y. Artificial intelligence and machine learning in energy systems: A bibliographic perspective. Energy Strategy Reviews. 2023 Jan 1;45:101017.
[60] Mohapatra PS. Artificial Intelligence and Machine Learning for Test Engineers: Concepts in Software Quality Assurance. Intelligent Assurance: Artificial Intelligence-Powered Software Testing in the Modern Development Lifecycle. 2025 Jul 27:17.
[61] Bini SA. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care?. The Journal of arthroplasty. 2018 Aug 1;33(8):2358-61.
[62] Wang W, Siau K. Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: A review and research agenda. Journal of Database Management (JDM). 2019 Jan 1;30(1):61-79.