IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207011 |

Large Language Models: From Research to Real-World Applications

Nazeer Shaik, Abdul Subhahan Shaik, Dr.C. Krishna Priya

Department of CSE, Srinivasa Ramanujan Institute of Technology (Autonomous), Anantapur, India

Department of CSE, Crimson Institute of Technology, Hyderabad, India

Department of Computer Science & IT, Central University of Andhra Pradesh, Anantapur, India

ABSTRACT: This paper explores the evolution, capabilities, and real-world applications of Large Language Models (LLMs), which have become central to the field of natural language processing. Beginning with the advent of the Transformer architecture and progressing through the development of models such as GPT-3 and BERT, LLMs have shown unprecedented proficiency in understanding and generating human language. Their applications span various domains, including healthcare, education, business, marketing, and creative industries, where they enhance productivity, improve decision-making, and foster innovation. Despite their transformative potential, LLMs pose significant challenges and ethical considerations. Issues such as bias in training data, data privacy concerns, potential misuse, environmental impact, and the need for transparency and accountability are critical areas that require attention. This paper discusses strategies to mitigate these challenges, including advancements in model architecture, the integration of LLMs with other AI technologies, and efforts to democratize access to these powerful tools. By examining both the opportunities and risks associated with LLMs, this paper underscores the importance of responsible and equitable use of these models. It highlights the exciting future directions for research and development in this field, aiming to harness the capabilities of LLMs to drive positive societal impact and innovation across multiple domains.

KEYWORDS: Large Language Models, Transformers, GPT-3, BERT, natural language processing, bias, data privacy, AI ethics, real-world applications, healthcare, education, business, creativity, sustainability.

I. INTRODUCTION

Language is a cornerstone of human interaction, facilitating communication, collaboration, and the exchange of ideas. The quest to enable machines to understand and generate human language has been a long-standing challenge in the field of artificial intelligence (AI). Over the years, Natural Language Processing (NLP) has evolved from simple rule-based systems to sophisticated models capable of performing complex linguistic tasks. Among these advancements, Large Language Models (LLMs) have emerged as groundbreaking tools, pushing the boundaries of what machines can achieve in terms of language understanding and generation [1].

Large Language Models, such as OpenAI's GPT series, have revolutionized the way we approach language-related tasks in AI. These models are characterized by their vast size, typically comprising billions of parameters, and their ability to learn from massive amounts of text data. By leveraging deep learning techniques and the Transformer architecture, LLMs have achieved remarkable performance across a variety of NLP tasks, including text generation, translation, summarization, and question-answering [2].

The impact of LLMs extends beyond academic research, finding practical applications in numerous fields. In healthcare, they assist in clinical decision-making and patient communication. In education, they provide personalized tutoring and generate educational content. Businesses use LLMs to enhance customer service, automate content creation, and gain insights from data. Even creative industries benefit from these models, as they aid in writing, composing music, and developing games.

However, the rise of LLMs also brings significant challenges and ethical considerations. Issues such as bias, data privacy, and the potential for misuse must be addressed to ensure that these powerful tools are used responsibly. As we continue to explore and expand the capabilities of LLMs, it is crucial to develop frameworks and guidelines that promote fairness, transparency, and ethical use.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207011 |

This paper aims to provide a comprehensive overview of Large Language Models, tracing their evolution from early NLP research to their current state and real-world applications. We will examine the technological advancements that have enabled their development, explore their diverse capabilities, and discuss the practical implications and challenges they present. By understanding the journey of LLMs and their impact on various sectors, we can better appreciate their transformative potential and navigate the ethical landscape surrounding their use.

II. EVOLUTION OF LARGE LANGUAGE MODELS

2.1. Early Developments in Natural Language Processing

The journey of natural language processing (NLP) began with rudimentary rule-based systems. These early systems, developed in the mid-20th century, relied heavily on handcrafted linguistic rules to process language. They were limited in scope and unable to handle the variability and complexity inherent in human language. The advent of statistical methods in the 1980s and 1990s marked a significant shift, allowing for more flexible and data-driven approaches. Techniques such as n-gram models and hidden Markov models (HMMs) laid the groundwork for future advancements, enabling better handling of sequential data and probabilistic language modeling [3].

2.2. The Rise of Machine Learning

The late 1990s and early 2000s saw the emergence of machine learning in NLP, particularly with the use of support vector machines (SVMs) and conditional random fields (CRFs). These models improved the ability to perform tasks like part-of-speech tagging, named entity recognition, and syntactic parsing by learning from annotated datasets. However, they were still constrained by the need for feature engineering and limited computational power.

2.3. The Breakthrough with Deep Learning

Deep learning, particularly the introduction of deep neural networks, revolutionized NLP by eliminating the need for extensive feature engineering and allowing models to learn directly from raw data. Recurrent Neural Networks (RNNs) and their variant, Long Short-Term Memory (LSTM) networks, were among the first to show promise in capturing temporal dependencies in sequential data. Despite their success, RNNs and LSTMs struggled with long-range dependencies due to issues like vanishing gradients [4].

2.4. Transformer Architecture

A breakthrough came with the introduction of the Transformer architecture by Vaswani et al. in 2017. The Transformer model leveraged self-attention mechanisms, enabling it to process input data in parallel rather than sequentially. This significantly improved the efficiency and scalability of training large models and better captured long-range dependencies. The self-attention mechanism allows the model to weigh the importance of different words in a sentence, making it highly effective for a wide range of NLP tasks.

2.5. Generative Pre-trained Transformers (GPT)

OpenAI's introduction of the Generative Pre-trained Transformer (GPT) series marked a significant milestone in the development of LLMs. The GPT models employed a two-step process: unsupervised pre-training on a massive corpus of text, followed by supervised fine-tuning on specific tasks.

- **GPT-1:** The first iteration, GPT-1, demonstrated the potential of transfer learning in NLP by achieving state-of-the-art results on various benchmarks through this pre-training and fine-tuning approach.
- **GPT-2:** GPT-2, with 1.5 billion parameters, showcased impressive text generation capabilities, raising awareness about the possibilities and risks of powerful language models. It highlighted the importance of balancing innovation with ethical considerations.
- **GPT-3:** GPT-3 further pushed the boundaries with 175 billion parameters, enabling it to perform a wide range of tasks with minimal fine-tuning. Its ability to generate human-like text, answer questions, and even write code demonstrated the vast potential of LLMs.
- **GPT-4:** Building on its predecessors, GPT-4 introduced improvements in understanding and generating nuanced and contextually relevant language. It demonstrated enhanced capabilities in areas such as multilingual translation, creative writing, and complex problem-solving.

III. CAPABILITIES OF LARGE LANGUAGE MODELS

Large Language Models (LLMs) have revolutionized the field of Natural Language Processing (NLP) by demonstrating a wide array of capabilities that were previously unattainable with traditional methods. These

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207011 |

capabilities stem from their vast size, extensive training data, and sophisticated architectures, particularly the Transformer model. This section explores the core competencies of LLMs and their practical applications [5].

Natural Language Understanding

One of the primary strengths of LLMs is their ability to understand and interpret human language. This includes:

1. **Sentiment Analysis**: LLMs can analyze text to determine the sentiment behind it, whether it is positive, negative, or neutral. This capability is widely used in customer feedback analysis, social media monitoring, and market research.
2. **Named Entity Recognition (NER)**: LLMs can identify and classify entities within a text, such as names of people, organizations, locations, dates, and more. This is crucial for tasks such as information extraction, document categorization, and data annotation.
3. **Text Classification**: LLMs can classify text into predefined categories. This includes spam detection, topic categorization, and intent recognition in conversational systems.
4. **Question Answering**: LLMs can understand and respond to questions posed in natural language. This capability is foundational for building intelligent assistants, customer support bots, and search engines.

Language Generation

LLMs are renowned for their ability to generate human-like text, making them invaluable for various content creation tasks. Key applications include:

1. **Content Creation**: LLMs can generate articles, blog posts, reports, and other written content. They assist writers by providing drafts, ideas, and stylistic improvements, thereby enhancing productivity and creativity.
2. **Creative Writing**: LLMs can produce poetry, stories, and other forms of creative writing. They serve as tools for authors, poets, and screenwriters, offering inspiration and diverse perspectives.
3. **Dialogue Systems**: LLMs power chatbots and conversational agents that can engage users in meaningful and coherent dialogues. They are used in customer service, virtual assistants, and interactive entertainment.
4. **Code Generation**: Advanced LLMs, such as OpenAI's Codex, can generate code snippets and assist in software development by understanding and producing programming language constructs.

Conversational Agents

The conversational capabilities of LLMs enable the development of sophisticated agents that can interact with users in natural language. Applications include:

1. **Customer Service Bots**: LLMs are used to automate customer support, providing instant responses to queries, resolving issues, and offering personalized recommendations.
2. **Virtual Assistants**: LLMs enhance virtual assistants like Siri, Alexa, and Google Assistant by improving their ability to understand user commands, provide accurate information, and perform tasks.
3. **Mental Health Support**: LLMs are employed in mental health applications to offer support and counseling through empathetic and context-aware conversations, providing users with timely assistance and resources.

Translation and Multilingual Applications

LLMs have made significant strides in the field of machine translation, delivering more accurate and contextually appropriate translations. Key aspects include:

1. **Machine Translation**: LLMs can translate text between multiple languages with high accuracy, supporting global communication and collaboration. They are used in translation services, international business, and language learning.
2. **Multilingual Understanding**: LLMs can process and understand text in various languages, making them suitable for applications that require language-agnostic capabilities, such as global customer support and international content curation.
3. **Cross-Language Information Retrieval**: LLMs facilitate the retrieval of information across different languages, enabling users to access and understand content irrespective of the language in which it is written[6].

IV. REAL-WORLD APPLICATIONS OF LARGE LANGUAGE MODELS

Large Language Models (LLMs) have transcended theoretical research and found applications in various real-world domains. Their ability to understand and generate human language with high accuracy and fluency makes them versatile tools in numerous fields. This section explores some of the key areas where LLMs are making a significant impact.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207011 |

4.1. Healthcare

In the healthcare sector, LLMs are transforming various aspects of medical practice and patient care:

1. **Clinical Documentation**: LLMs streamline clinical documentation by automatically generating and summarizing patient records, reducing the administrative burden on healthcare professionals. This improves the accuracy and completeness of medical records while allowing clinicians to focus more on patient care.
2. **Medical Research**: LLMs assist researchers by synthesizing vast amounts of medical literature, identifying relevant studies, and generating comprehensive summaries. This accelerates the pace of medical research and aids in the discovery of new treatments and therapies.
3. **Patient Interaction**: LLM-powered chatbots and virtual assistants enhance patient interaction by providing medical information, appointment scheduling, and preliminary diagnosis. They offer timely responses to patient inquiries, improving accessibility and patient satisfaction.
4. **Decision Support**: LLMs support clinical decision-making by analyzing patient data, medical histories, and current research to provide evidence-based recommendations. This aids healthcare providers in diagnosing conditions and selecting appropriate treatments.

4.2. Education

In the education sector, LLMs are revolutionizing the learning experience for both students and educators:

1. **Personalized Tutoring**: LLMs offer personalized tutoring by adapting to the learning style and pace of individual students. They provide explanations, practice problems, and feedback, catering to diverse learning needs and improving educational outcomes.
2. **Automated Grading**: LLMs assist educators by automating the grading of assignments and exams. This ensures consistency and frees up time for more personalized instruction and student engagement.
3. **Content Generation**: LLMs generate educational content, including textbooks, study guides, and quizzes. They create high-quality learning materials that are accessible to a wide range of learners, supporting education at all levels.
4. **Language Learning**: LLMs facilitate language learning by providing translations, grammar corrections, and conversational practice. They help learners improve their language skills through interactive and immersive experiences [7].

4.3. Business and Marketing

Businesses leverage LLMs to enhance efficiency, customer engagement, and data-driven decision-making:

1. **Customer Service Automation**: LLM-powered chatbots handle customer inquiries, process orders, and provide support, reducing response times and operational costs. They improve customer satisfaction by delivering accurate and timely assistance.
2. **Market Analysis**: LLMs analyze market trends, customer feedback, and competitive intelligence to provide actionable insights. This helps businesses make informed decisions, develop effective strategies, and stay ahead of market trends.
3. **Content Generation**: LLMs assist in creating marketing materials, social media posts, and product descriptions. They streamline content creation processes, ensuring consistency and enhancing brand communication.
4. **Sales and Lead Generation**: LLMs optimize sales processes by identifying potential leads, personalizing outreach, and automating follow-ups. This increases conversion rates and drives business growth.

4.4. Creative Industries

In creative industries, LLMs serve as tools for innovation and artistic expression:

1. **Writing Assistance**: LLMs help writers generate ideas, draft content, and refine their work. They support various forms of writing, from novels and screenplays to technical documentation and journalism.
2. **Music Composition**: LLMs aid musicians by generating melodies, harmonies, and lyrics. They foster creativity and experimentation in music production, enabling artists to explore new musical landscapes.
3. **Game Development**: LLMs contribute to game development by creating dialogue, storylines, and character interactions. They enhance the narrative and interactive elements of games, enriching the gaming experience.
4. **Visual Arts**: LLMs support visual artists by generating descriptions and concepts for artwork, facilitating creative brainstorming, and even assisting in the production of digital art [8].

4.5. Legal and Compliance

In the legal sector, LLMs improve efficiency and accuracy in various processes:

1. **Legal Research**: LLMs expedite legal research by summarizing case law, statutes, and regulations. They provide relevant legal information and precedents, aiding lawyers in case preparation.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207011 |

2. **Document Review**: LLMs assist in reviewing and drafting legal documents, contracts, and agreements. They ensure consistency, identify potential issues, and streamline document workflows.
3. **Compliance Monitoring**: LLMs help organizations monitor compliance with laws and regulations by analyzing documentation and communications for potential violations. They support risk management and regulatory adherence.

4.6. Finance

In the finance sector, LLMs enhance analysis, decision-making, and customer interactions:

1. **Financial Analysis**: LLMs analyze financial reports, market data, and economic indicators to provide insights and forecasts. They support investment strategies, risk management, and portfolio optimization.
2. **Customer Support**: LLM-powered chatbots handle customer inquiries related to banking, investments, and insurance. They provide timely and accurate responses, improving customer experience and satisfaction.
3. **Fraud Detection**: LLMs assist in detecting fraudulent activities by analyzing transaction patterns and identifying anomalies. They enhance security measures and protect against financial crimes.

V. CHALLENGES AND ETHICAL CONSIDERATIONS

The development and deployment of Large Language Models (LLMs) come with a host of challenges and ethical considerations that must be addressed to ensure their responsible use. As these models become increasingly integrated into various aspects of society, it is crucial to understand and mitigate the potential risks and negative impacts associated with their use.

5.1. Bias and Fairness

- **Bias in Training Data**: LLMs are trained on vast datasets that often contain biased information reflecting societal prejudices. As a result, these models can inadvertently learn and propagate these biases, leading to unfair and discriminatory outcomes. For example, biased language in training data can result in gender, racial, and cultural stereotypes being reinforced in the model's outputs [10,11].
- **Mitigating Bias**: Addressing bias requires careful curation and preprocessing of training data, as well as the development of algorithms and techniques to detect and reduce bias during training and inference. Researchers are exploring methods such as adversarial training, fairness constraints, and bias auditing to create more equitable models.
- **Fairness in Application**: Ensuring fairness also involves considering the broader context in which LLMs are used. This includes assessing the potential impact of LLMs on different demographic groups and implementing measures to prevent disparate treatment and outcomes.

5.2. Data Privacy

- **Sensitive Information**: The datasets used to train LLMs often contain sensitive and personal information, raising significant privacy concerns. There is a risk that LLMs might inadvertently generate outputs that reveal private data, leading to potential breaches of confidentiality.
- **Data Anonymization**: Techniques such as data anonymization and differential privacy can help protect individual identities and sensitive information during the training process. These methods aim to ensure that the outputs of LLMs do not compromise privacy while maintaining the utility of the model.
- **Regulatory Compliance**: Organizations using LLMs must comply with data protection regulations, such as the General Data Protection Regulation (GDPR) in the European Union, which mandates strict guidelines on data handling, consent, and privacy protection [9].

5.3. Misuse and Misinformation

- **Generation of Harmful Content**: The generative capabilities of LLMs can be misused to create harmful content, including hate speech, misinformation, and deepfakes. These malicious uses can have serious societal impacts, including the spread of false information, manipulation of public opinion, and harm to individuals and communities.
- **Content Moderation**: Effective content moderation strategies are essential to prevent the misuse of LLMs. This includes developing robust filters and monitoring systems to detect and mitigate harmful outputs. Collaboration with policymakers, industry stakeholders, and researchers is necessary to establish guidelines and best practices for content moderation.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207011 |

- **Ethical Guidelines**: The development of ethical guidelines and frameworks is critical to ensure the responsible use of LLMs. These guidelines should address issues such as transparency, accountability, and the ethical implications of deploying LLMs in various applications.

Environmental Impact

- **Energy Consumption**: Training and deploying LLMs require substantial computational resources, leading to significant energy consumption and carbon emissions. The environmental footprint of LLMs is a growing concern, especially as models continue to increase in size and complexity [12].
- **Sustainable AI Practices**: To mitigate the environmental impact, researchers and practitioners are exploring more efficient model architectures, optimization techniques, and the use of renewable energy sources. Efforts to improve the energy efficiency of data centers and promote sustainable AI practices are essential for reducing the ecological footprint of LLMs.

5.4. Accountability and Transparency

- **Opaque Decision-Making**: LLMs often operate as "black boxes," making it challenging to understand and interpret their decision-making processes. This lack of transparency can hinder accountability and trust in the technology.
- **Explainability**: Developing explainable AI techniques is crucial for making LLMs more transparent and interpretable. Explainability helps stakeholders understand how models arrive at their conclusions and provides insights into the factors influencing their outputs.
- **Ethical Auditing**: Regular ethical audits of LLMs can help identify potential risks, biases, and ethical issues. These audits should involve a diverse group of stakeholders, including ethicists, domain experts, and representatives from affected communities.

VI. FUTURE DIRECTIONS

6.1. Advancements in Model Architecture

- **Efficient Architectures**: Research continues to explore more efficient and effective model architectures, such as sparse transformers and reinforcement learning-based approaches, to enhance the performance and scalability of LLMs.
- **Specialized Models**: Developing specialized models tailored to specific tasks and domains can improve efficiency and reduce the need for large, generalized models. This approach can also help address some ethical concerns by limiting the scope of model applications.

6.2. Integration with Other AI Technologies

- **Multi-Modal AI**: Integrating LLMs with other AI technologies, such as computer vision and robotics, opens up new possibilities for multi-modal and embodied AI systems. These systems can perform more complex tasks and provide richer interactions [13].
- **Human-AI Collaboration**: Fostering collaboration between humans and AI systems can leverage the strengths of both, leading to more effective and ethical outcomes. Designing AI systems that augment human capabilities and involve human oversight is a key direction for future research.

6.3. Democratizing Access

- **Open-Source Initiatives**: Efforts to democratize access to LLMs through open-source initiatives and affordable cloud-based services aim to ensure that the benefits of these models are widely accessible. This includes providing resources for education, research, and innovation in diverse communities.
- **Inclusive Development**: Promoting inclusive development practices involves engaging a broad range of stakeholders in the design, deployment, and governance of LLMs. Ensuring that diverse perspectives are represented can help address ethical concerns and promote equitable outcomes [14,15].

VII. CONCLUSION

Large Language Models (LLMs) have ushered in a new era of artificial intelligence, demonstrating unparalleled capabilities in understanding and generating human language. From the evolution of the Transformer architecture to the advent of models like GPT-3, BERT, and their successors, LLMs have shown remarkable progress, making significant strides in both research and practical applications.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207011 |

LLMs are now at the forefront of numerous industries, revolutionizing healthcare with improved clinical documentation and patient interaction, enhancing education through personalized tutoring and automated grading, optimizing business and marketing strategies, and fostering creativity in writing, music, and game development. Their ability to generate coherent and contextually relevant text makes them invaluable tools across diverse fields, providing both efficiency and innovation.

However, the widespread deployment of LLMs also brings forth critical challenges and ethical considerations. Issues such as bias, data privacy, misuse, environmental impact, and the need for transparency and accountability must be addressed to ensure the responsible and equitable use of these powerful technologies. Bias in training data can lead to unfair outcomes, while the potential for generating harmful content raises concerns about misinformation and misuse. Moreover, the significant energy consumption associated with training LLMs calls for sustainable AI practices.

Future directions in the development of LLMs include advancements in model architectures to improve efficiency and scalability, integration with other AI technologies to create multi-modal systems, and efforts to democratize access through open-source initiatives and inclusive development practices. These endeavors aim to harness the full potential of LLMs while mitigating risks and ensuring their benefits are widely accessible.

In conclusion, Large Language Models represent a monumental leap in artificial intelligence, offering transformative potential across various sectors. The ongoing advancements in this field promise exciting possibilities, but it is imperative to address the ethical and practical challenges they present. By doing so, we can leverage the capabilities of LLMs responsibly and equitably, paving the way for a future where these models contribute positively to society and drive innovation across multiple domains.

REFERENCES

1. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. arXiv preprint arXiv:2005.14165.
2. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
3. Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 21(140), 1-67.
4. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. arXiv preprint arXiv:1907.11692.
5. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI Blog.
6. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems (pp. 5998-6008).
7. Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2019). ALBERT: A lite BERT for self-supervised learning of language representations. arXiv preprint arXiv:1909.11942.
8. Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). XLNet: Generalized autoregressive pretraining for language understanding. Advances in Neural Information Processing Systems, 32, 5753-5763.
9. Clark, K., Luong, M. T., Le, Q. V., & Manning, C. D. (2020). ELECTRA: Pre-training text encoders as discriminators rather than generators. arXiv preprint arXiv:2003.10555.
10. He, H., Liu, P., Wang, W., & McAuley, J. (2020). Learning to write with cooperative discriminators. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 2541-2553).
11. Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2020). BERTScore: Evaluating text generation with BERT. In International Conference on Learning Representations.
12. Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... & Zettlemoyer, L. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 7871-7880).
13. Touvron, H., Cord, M., Sablayrolles, A., Synnaeve, G., & Jégou, H. (2021). Training data-efficient image transformers & distillation through attention. In International Conference on Machine Learning (pp. 10347-10357). PMLR.
14. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.
15. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... & Liang, P. (2021). On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details