DOI: http://dx.doi.org/10.26483/ijarcs.v15i4.7111

Volume 15, No. 4, July-August 2024

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

A Literature Review on Emotional Intelligence of Large Language Models (LLMs)

Pinaki Raj
Aligarh, India

Abstract: Large Language Models(LLMs) are artificial intelligence models that use deep neural networks to perform Natural Language Processing (NLP) tasks. These tasks include interaction between humans and computers, enabling computers to interpret and generate human languages in a meaningful manner. Large Language models are called "large" because of the architecture's size and the huge sets of training text data. With the emergence of transformer-based LLMs, the game of NLPs has reached another level. This is due to their ability to handle long-range text dependencies in parallel. The growing prevalence of transformer-based LLMs in human lives has necessitated evaluating the scope of the Emotional Intelligence(EI) of LLMs. This paper will discuss the need for emotional intelligence in transformer-based LLMs and the various existing studies that have evaluated this aspect. The potential challenges of the LLMs along with the future directions for research in this field will also be discussed.

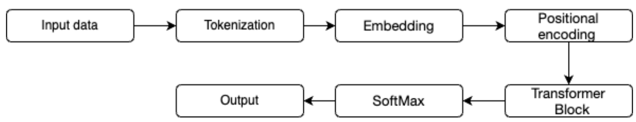
Keywords: Large Language Models (LLMs), Emotional Intelligence (EI), Natural Language Processing (NLP), Language Model, Machine Learning, Artificial Intelligence, Transformer Model

I. INTRODUCTION

A language model is the probability distribution of the sequence of words, using machine learning methods to predict the next word based on the previous words in sequence [25]. N-gram models were the first promising models used in the development of Large Language Models (LLMs) [1][25]. However, these models faced limitations due to the inability to handle long-range interdependencies between words, which led to difficulty in comprehending the contextual meaning of text [1][25]. Recurrent Neural Networks (RNNs) were introduced to rectify the issues present in n-gram models. These neural architectures can process any length of the input, allowing the modeling of sequential data [2]. Even with these prominent improvements, these models also faced limitations because of the vanishing gradient problems, exploding gradient problems and the availability of limited memory which can be troublesome for longer sequences as the network may forget the important information when dealing with these sequences [2][26]. The introduction of transformer architecture marked a significant advancement in the field of LLMs [3]. These models offer notable advantages over previous models as they allow parallelization of training data as compared to RNN where training was done in sequential manner. With the help of Graphical Processing Units(GPUs), it can further enhance the concurrent training of models enabling the construction of much larger language models [3]. With the emergence of transformer architecture, several LLMs like Generative Pre-trained Transformer(GPT), Bidirectional Encoder Representation from Transformer (BERT) have become hot topics in the market. Transformer-based LLMs models have become prevalent, finding applications in various fields such as healthcare, education, social media, and chatbots. The widespread use of these LLMs by humans underscores the need to determine whether these LLMs can understand humans' emotions. Also, it is important to evaluate the scope of incorporating EI in these transformer-based LLMs.

II. TRANSFORMER MODEL

Transformer model architecture uses the self-attention mechanism to compute the input and output sentences without using convolution [3]. It performs sequence transduction and pays attention to the important words in a sequence [3]. The transformer architecture maintains the encoder-decoder model [3]. The encoder takes an input and outputs a matrix representation of that input, while the decoder takes that encoded input and generates an output [3][4]. Fig.1 shows the main parts of the transformer model architecture and this section will be discussing these parts in brief.



1. The novel architecture of transformer model

A. ***Tokenization:*** The input is fed to the machine learning model for tokenization where the words of the input sequence are assigned as tokens [3].

B. ***Embedding:*** The tokens are then assigned unique numerical vector identifiers. This process of converting tokens in numerical format is known as embedding. The numerical vectors are then passed through an embedding layer [3].

C. ***Positional encoding:*** The positional encoding is used to encode the position of each token as an integer. It provides each token with a unique position vector by using the combination of various sine and cosine functions, thereby encoding sequences of any length [3].

D. ***Transformer block*:** A transformer block comprises of two components: the multi-head self-attention mechanism and the feed-forward neural network [3]. There can be multiple sets of self attention and feed forward network in the transformer block. The self-attention mechanism is the key component of the transformer block, as it adds context and meaning to tokens in sequences. It weights the tokens according to their contextual relevance and shifts them closer to the related word embeddings [3]. The feed-forward neural network is another important component of the transformer model that enhances the model's ability to understand complex patterns [3]. This network processes the input from the self-attention mechanism, passing it through a series of input layers and hidden states, and then transforming it into the model's output [3].

E. ***Softmax***: The output of the transformer block is then converted into scores. The softmax function transforms these scores into probabilities and the token with the highest probability for each position is considered the appropriate output [3].

III. NEED OF EMOTIONAL INTELLIGENCE IN LARGE LANGUAGE MODELS

Emotional Intelligence(EI) is the ability to understand our feelings and of others. It helps regulate emotions, foster empathy, imbibe motivation, and inculcate social awareness. Thus, learning emotional intelligence can help people enhance effective communication, critical thinking, and decision-making skills. Various studies indicate that people with higher Emotional Intelligence (EI) tend to be more successful than those with a higher intelligence quotient (IQ). Since EI is so important in effective human communication, it is also necessary to integrate EI into transformer-based LLMs. It can help in augmenting user satisfaction by comprehending their emotions in a better way and also responding to them appropriately. Reference[5][6] discussed the importance of EI in LLMs as chatbots or diagnostic tools are interacting directly with patients. It can help in understanding the emotional states of patients, thus enhancing patient experiences. Reference[7] conducted a study to compare the responses of ChatGPT and physicians to some randomly drawn patients from social media forums. The results show that responses given by ChatGPT were more empathetic as compared to physicians. Reference[8] also shows that chatbots can engage in emotional conversations with humans. These studies emphasize that user satisfaction can be augmented if EI is to be incorporated into LLMs. It can also help in providing inclusivity and acceptance in society [9]. Reference[9] discusses the challenges faced by queer community and they may seek online support and information on issues related to identity, acceptance, and mental health [9]. An empathetic LLM can provide affirmations to their feelings, help them navigate issues they are not comfortable discussing with others, and help other people understand and respect diverse communities, thus breaking down stigma [9]. From a theoretical perspective, [10][6] discussed five reasons for the necessity of taking human emotions into account when communicating with AI: The first was the human reaction to AI actions in typical situations of social interaction [10]. The second was the various emotional reactions of people during the AI implementation typically human areas [10]. The third reason was the "uncanny valley" phenomenon [10]. The fourth was provoking moral emotions during interaction with AI [10]. And the last reason was the specificity of some AI instruments designed to work with people's emotions [10]. These reasons support the integration of human emotions into LLMs. It also encourages us to understand how well the LLMs can learn the aspects of EI from the vast amount of training data.

IV. LITERATURE REVIEW

With the widespread use of LLMs across various fields, their compatibility with human emotions has become a point of discussion. Various studies have been conducted to evaluate the Emotional Intelligence (EI) of large language models (LLMs). To compare the EI, each study has used a different set of parameters to judge the LLMs, making it difficult to compare these studies. The comparisons done in these studies ultimately are judged against the human Emotional Intelligence as a benchmark.

Emotional Understanding (EU), a fundamental aspect of EI, assesses the ability to recognize and interpret emotions in social situations. The two studies that are discussed below will assess the EU of LLMs.

A test known as the Situational Evaluation of Complex Emotional Understanding (SECEU) has been used to evaluate the EU of both LLMs and humans [11]. SECEU is an objective, text-based assessment that evaluates complex emotions in realistic situations [11]. The test was conducted on 500 adults and various LLMs, including the OpenAI GPT series, Claude, LLaMA-based models, FastChat-T5, Pythia-based models GLM-based models and RWKV (Recurrent Weighted Key-Value) [11]. The results revealed that GPT-4 performed exceptionally well, with performance comparable to that of humans [11]. Since emotional intelligence is mainly dependent on situational contexts, achieving performance in language models that is equivalent to human levels is a good sign. This indicates that these models can understand and respond to complex emotional cues and scenarios, thus showing a deeper level of artificial intelligence.

Along the same line of comparing EU in different situations, EQ-Bench was designed to assess the ability of LLMs to understand complex emotions and interactions [15]. A specific question format was designed to rate the emotional intensity of LLMs in a dialogue of conflict and tension scenarios [15]. It provided a comprehensive and challenging assessment of LLMs without needing the assessor's interpretation [15]. The results showed that LLMs could effectively differentiate dialogue scenarios, even with narrowly focussed dialogue scenarios [15]. Also, it should be noted that the OpenAI's GPT-4 model has the highest score in the EQ-bench, thus emphasizing its paramountcy over the available models [15]. Both these studies were conclusive enough that LLMs indeed are able to understand the emotional context of the sentences.

Another set of studies used different parameters to compare the models. They either provided some sort of input stimuli or interacted with an end user to understand the response of the trained model.

Reference[12] proposed a Six Emotional Dimensions (6DE) model to evaluate human emotions in AI and LLMs. This model comprises arousal, valence, dominance, agency, fidelity, and novelty [12]. These dimensions are incorporated with sub-dimensions within each dimension, thus offering a comprehensive framework for understanding and quantifying emotions [12]. It uses prompts and instructions to analyze the emotions of LLMs, thereby enabling customized emotional responses which can help in having meaningful conversations and enhancing user satisfaction [12]. The author concludes that incorporating the 6DE model in AI can help in providing better emotional responses along with offering a common language for discussing emotions [12]. This can help in enhancing the relatability of human-AI relationships [12].

In another study, an Emotion Prompt was designed, using 11 sentences as emotional stimuli influenced by self-monitoring, social cognitive theory, and cognitive emotion regulation theory [14]. The emotional stimuli were added to the initial prompts to evaluate the EI of LLMs [14]. The LLMs used in this study included Flan-T5-Large, Vicuna, Llama2, BLOOM, ChatGPT, and GPT-4 [14]. The author concluded that all LLMs not only showed positive results, indicating that LLMs can understand and respond to emotional stimuli but can also be enhanced by emotional stimuli [14]. The study showed that LLMs enhanced by emotional intelligence can also achieve better performance, truthfulness and responsibility [14].

Reference[13] evaluated the impact of ChatGPT on user experiences using two questionnaires and the YOLOv5 emotion-detecting model. One questionnaire was formulated by humans, while the other was created by rephrasing human-generated questions using ChatGPT [13]. The study, conducted on 14 participants, analyzed data using Analysis of Variance (ANOVA) [13]. The results concluded that the questionnaire formulated with both human and ChatGPT inputs provided superior user satisfaction compared to those formulated solely by humans or ChatGPT [13]. The questionnaire generated by both humans and ChatGPT enhances participants' happiness levels and reduces their sadness levels. This emphasizes that AI-powered chatbots like ChatGPT can enhance user satisfaction when manipulated by humans [13].

In another study, a set of survey questions were asked to GPT-4 and each question was asked as a separate prompt to avoid the ability of GPT-4 to hold the context [6]. To ensure the reliability of responses, each question was asked three times and the responses were recorded. The Russian version of the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT V.2.0) was used in this research [6]. The author pointed out that the vast internet text data set is sufficient for assimilating emotional domain patterns [6]. GPT-4 is capable of identifying emotions in text and describing techniques for managing them, despite these abilities were not programmed into the model [6]. Also, certain intricate emotions such as motivational aspects might not be handled by LLMs [6]. The study also mentioned that understanding the motivation aspect of emotions will be difficult to learn from a large scattered set of data. It might be possible to learn emotions from a concentrated set of training data and will require more testing to investigate this further [6].

The other part of understanding the EI is how well it can be applied to society. To do so it needs to assess human nature. The Myers–Briggs Type Indicator (MBTI) tests were developed to evaluate the ability of LLMs to assess human personalities. The framework comprised of three components- unbiased prompts, subject-replaced query and correctness-evaluated instructions [16]. These components allow to reformulate the MBTI questions in a flexible manner which allows them to assess human personalities [16]. Also, the three quantitative evaluation metrics were used to measure the consistency of LLMs ’assessments. These were robustness, fairness, and consistency. The LLMs used for this study are ChatGPT, GPT-4 and InstructGPT. The results revealed that all the LLMs in the study can assess human personalities. However, ChatGPT and GPT-4 achieved more consistent and fair assessments of human personalities, with less gender bias than InstructGPT [16].

Reference[17] highlighted that LLM agents could mirror certain social behaviors typical of human collaboration, emphasizing the potential of human-AI interaction. The research framework consisted of a test that integrated multi-agent societies, thinking patterns and collaborative strategies [17]. The paper highlighted that the collaboration capabilities of LLM agents have the potential for human-AI interaction. The author also emphasized that collaborative LLM multi-agents can be used to make social-aware AI [17].

Overall, the ability of LLMs to understand and learn the emotions from the training data is quite promising, as evident from the studies. However, it does face a few challenges as we move forward on this path.

V. CHALLENGES AND FUTURE DIRECTIONS

The rapid development and extensive usage of large language models (LLMs) have significantly impacted society. While LLMs have undoubtedly assisted us in navigating various tasks, they also present challenges and shortcomings that need to be addressed. This section will discuss some of these challenges and future directions for overcoming them.

A. Challenges

The NLP tasks performed by LLMs have raised concerns about their ethical nature. As LLMs like ChatGPT are used for various tasks, it is important to question their ethical behavior. One major concern is the bias in LLMs, as they are trained on data available on the internet, which might contain biases [18]. This can lead to the generation of biased data, resulting in discrimination and prejudice against certain sections of society [18],[19]. Also, the data entered into ChatGPT is used to train these models, so it is important to be cautious with personal data. Without robust anonymization or redaction measures, personal data can become part of the model's training dataset and be used to generate responses for others, compromising privacy and data security [18],[19]. Another challenge is the tendency of LLMs to spread misinformation and disinformation [19],[20]. Since LLMs can generate human-like text, they can convincingly present fraudulent information, which can be used for political or financial gain, bolstering ideologies, or trolling, ultimately affecting public opinion and social harmony [18]. Hallucinations, where LLMs generate incorrect or irrelevant information due to limitations in understanding context or reliance on training data are also problematic, especially in fields requiring precise information such as healthcare, law, and engineering [18],[19]. Thus, relying entirely on LLMs in such situations can be detrimental. Transparency is another issue, as the black-box nature of LLMs makes it difficult to build trust [18]. This lack of transparency can be attributed to complex and uncertain model capabilities, massive and opaque architectures, and proprietary technology [18],[19]. The multimodal LLMs face inefficiencies, such as the challenge of coherently aligning data from different modalities like text and images, handling cross-modal understanding, and managing uncertainty and informed decision-making [21],[22]. Finally, the environmental impact of developing LLMs is a concern, as a significant amount of energy is required for training these large models [23],[24].

B. *Future Directions*

The future directions should aim to bridge the gap and overcome the challenge pertaining to LLMs. The ethical concerns require a multi-dimensional approach to resolve the challenges [18],[19]. Transparent and accountable LLMs should be developed. Privacy breaches can be minimized by using anonymized data in training that can help in mitigating the re-surfacing of data[18],[19]. Also, the decentralized data should be applied by federated learning and secure multiparty computation techniques [18]. Regular auditing of training data, scrutinizing of training biases and monitoring of privacy breaches should be done to build user's trust[18],[19]. For improving multimodal LLMs, designing architectures that effectively handle different modalities can be beneficial [21],[22]. Also, the energy used in the development of LLMs should be optimized by increasing the reuse rate and using sustainable high-performance computing technology[23],[24]. By addressing these aspects, future LLMs can become more ethical, efficient, and trustworthy.

VI. Conclusion

This literature review highlights the importance of Emotional Intelligence (EI) in transformer-based Large Language Models (LLMs). The studies examined in this paper demonstrate the ability of these models to understand and respond to human emotions when provided with appropriate instructions. The review suggests that incorporating EI into LLMs can enhance user interaction positively. In some cases, LLMs have even outperformed humans in providing better user experiences and responding empathetically, illustrating their capacity to comprehend and address emotions. However, this review has certain limitations. The studies discussed are based on subjective assessments and focus directly on EI. To conclude, as Emotional Intelligence is a fundamental aspect of human communication, LLMs can incorporate EI to bridge the gap between machines and humans. Nonetheless, further research is needed to refine EI in LLMs, thus expanding the potential applications of EI in these models.

VII. References

1. Liu, Jiacheng, et al. “Infini-gram: Scaling unbounded n-gram language models to a trillion tokens.” arXiv preprint arXiv:2401.17377 (2024).
2. Kombrink, Stefan, Tomas Mikolov, Martin Karafiát, and Lukás Burget. “Recurrent Neural Network Based Language Modeling in Meeting Recognition.” In Interspeech, vol. 11, pp. 2877-2880. 2011.
3. Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. “Attention is all you need.” Advances in neural information processing systems 30 (2017).
4. Ansar, Wazib, Saptarsi Goswami, Amlan Chakrabarti, and Basabi Chakraborty. “A novel selective learning based transformer encoder architecture with enhanced word representation.” Applied Intelligence 53, no. 8 (2023): 9424-9443.
5. Kerasidou, Angeliki. "Artificial intelligence and the ongoing need for empathy, compassion and trust in healthcare." Bulletin of the World Health Organization 98, no. 4 (2020): 245.
6. Vzorin, Gleb, Alexey Bukinich, Anna Sedykh, Irina Vetrova, and Elena Sergienko. "Emotional Intelligence of GPT-4 Large Language Model." (2023).
7. Meng, Jingbo, and Yue Dai. "Emotional support from AI chatbots: Should a supportive partner self-disclose or not?." Journal of Computer-Mediated Communication 26, no. 4 (2021): 207-222.
8. Uludag, Kadir. "The use of AI-supported Chatbot in Psychology." Available at SSRN 4331367 (2023).
9. Lissak, S., Calderon, N., Shenkman, G., Ophir, Y., Fruchter, E., Klomek, A.B. and Reichart, R., 2024. “The Colorful Future of LLMs: Evaluating and Improving LLMs as Emotional Supporters for Queer Youth.”arXiv preprint arXiv:2402.11886.
10. Shank, Daniel B., Christopher Graves, Alexander Gott, Patrick Gamez, and Sophia Rodriguez. “Feeling our way to machine minds: People's emotions when perceiving mind in artificial intelligence.” Computers in Human Behavior 98 (2019): 256-266.
11. Xuena Wang, Xueting Li, Zi Yin, Yue Wu and Jia Liu,“Emotional intelligence of Large Language Models,”(Article in Journal of Pacific Rim Psychology)
12. Jeremiah Ratican, James Hutson, ‘‘The Six Emotional Dimension (6DE) Model: A Multidimensional Approach to Analyzing Human Emotions and Unlocking the Potential of Emotionally Intelligent Artificial Intelligence (AI) via Large Language Models (LLM) ,” (Article in DS Journal of Artificial Intelligence and Robotics)
13. Zou, Zhao, et al. “A pilot study of measuring emotional response and perception of LLM-generated questionnaire and human-generated questionnaires.” Scientific reports 14.1 (2024): 2781.
14. Li, C., Wang, J., Zhang, Y., Zhu, K., Hou, W., Lian, J., ... & Xie, X. (2023). “Large language models understand and can be enhanced by emotional stimuli.” arXiv preprint arXiv:2307.11760.
15. Sabour, Sahand, Siyang Liu, Zheyuan Zhang, June M. Liu, Jinfeng Zhou, Alvionna S. Sunaryo, Juanzi Li, Tatia Lee, Rada Mihalcea, and Minlie Huang. “EmoBench: Evaluating the Emotional Intelligence of Large Language Models.” arXiv preprint arXiv:2402.12071 (2024).
16. Rao, H., Leung, C. and Miao, C., 2023. “Can chatgpt assess human personalities? a general evaluation framework.” arXiv preprint arXiv:2303.01248.
17. Zhang J, Xu X, Deng S. “Exploring collaboration mechanisms for llm agents: A social psychology view.”
18. Jiao, Junfeng, Saleh Afroogh, Yiming Xu, and Connor Phillips. “Navigating LLM Ethics: Advancements, Challenges, and Future Directions.” arXiv preprint arXiv:2406.18841 (2024).
19. Yan, Biwei, Kun Li, Minghui Xu, Yueyan Dong, Yue Zhang, Zhaochun Ren, and Xiuzheng Cheng. “On protecting the data privacy of large language models (llms): A survey.” arXiv preprint arXiv:2403.05156 (2024).
20. Barman, Dipto, Ziyi Guo, and Owen Conlan. “The dark side of language models: Exploring the potential of llms in multimedia disinformation generation and dissemination.” Machine Learning with Applications (2024): 100545.
21. Tong, Shengbang, et al. “Eyes wide shut? exploring the visual shortcomings of multimodal llms.” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.
22. Sulaiman, N., & Hamzah, F. “Multimodal Integration in Large Language Models: A Case Study with Mistral LLM.”

23. Lin, Hsiao-Ying, and Jeffrey Voas. "Lower Energy Large Language Models (LLMs)." Computer 56, no. 10 (2023): 14-16.
24. Stojkovic J, Choukse E, Zhang C, Goiri I, Torrellas J. "Towards Greener LLMs: Bringing Energy-Efficiency to the Forefront of LLM Inference." arXiv preprint arXiv:2403.20306. 2024 Mar 29.
25. Bonlarron, Alexandre, and Jean-Charles Régin. "Markov Constraint as Large Language Model Surrogate." arXiv preprint arXiv:2406.10269 (2024)
26. Hochreiter, Sepp. "The vanishing gradient problem during learning recurrent neural nets and problem solutions." International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 6, no. 02 (1998): 107-116.

