Ethical Considerations and Best Practices for Using Large Language Models in Decision - Making

Ashish K. Saxena

Abstract: ***Large language models (LLMs) like GPT - 4 and BERT have revolutionized the way critical sectors make decisions. Independent of the ability to train such models with multiple data sources, state - of - the - art LLMs such as GPT - 3 have exhibited remarkable capacity for producing realistic textual output, and conversing with users in a coherent manner. Over time, as LLMs become more sophisticated and are used by a broad range of organizations, they will have the capacity to shape decision - making and in some cases become decision makers in various fields such as the healthcare, financial, governmental and other sectors. However, there is an issue of ethical concerns when relying on the imperfect or biased LLM in making crucial decisions. In this white paper we look into how LLMs are already being used to suggest or even dictate decisions that affect people's lives; review the dangers posed by bias, accountability, and explainability of LLMs; explain why protecting standards is important but difficult due to present technological restraints; and give considerations on best practices for proper usage of LLMs in essential applications. The critical success factors explained are the acquisition of training datasets that are diverse and of high quality, integration checks with human involved in the process, audits of the system on a continuous basis, the provision of solutions which are explicable and transparent and clarity on the standards of the performance of the system or measures for risk control.***

Keywords: LLMs, decision - making, ethical concerns, bias, best practices

1. Introduction

Large language models or LLMs are self - learning artificial intelligence systems designed to write text similar to human writing and are developed from large volumes of textual information. Attributing to the recent improvement in the natural language processing algorithms and the easy availability of large datasets, the state - of - art models such as GPT - 4 and BERT have been able to generate coherent long - form text at a very high level, in addition to answering questions, summarizing documents and even translating between languages (Devlin, Chang & Lee, 2019). These models, some of which have over 100 billion parameters, have shown seemingly unlimited knowledge about how language works and the content of the data fed into the algorithms. Having reached that point of development, LLMs are capable of influencing numerous and rather important decision - making processes in such spheres as healthcare, finance, law, and government policy (Brown et al., 2020). This white paper will look at the development of LLMs and consider their strengths and weaknesses and the potential ramifications as these models become more and more applied in application areas that are critical to society. Lastly, the risks, governance issues, and their management are also going to be covered.

Thesis Statement

It is also a concern when large language models (LLMs) such as GPT - 3/4 or ChatGPT and BERT are increasing in capability and can affect high - stakes decisions in fields like medicine, law, finance, and policy; therefore, it is essential to assess these systems' accuracies, biases, and the overall nature before deploying them for such critical tasks.

2. The Technology behind Large Language Models

According to Wolf et al., (2020) NLP is an artificial intelligence paradigm that used to facilitate the use of natural language by computers that means to translate human language to a computer understandable language. Natural Language Processing is much dependent on training machine learning models on raw textual data to perform language related operations such as translation, summarization and question answering (Wolf et al., 2020). As stated by Bender, Gebru, McMillan - Major, & Shmitchell, (2021) extremely big LLMs are subclasses of deep learning NLP models for which there has been significant advancements in terms of capability recently. They are trained on large volumes of text – sometimes in the hundreds of billions of words from the web and textbooks – to complete the next word in a string of words. This enables them to develop very detailed structures on how language is formed and what the meaning is (Bommasani et al., 2021).

The work of Rae et al., (2022) explores that majority of modern LLM models are developed using an architecture known as transformers, which was proposed in 2017. To address the issues of modulating long - distance dependencies and paying attention to the context, Transformers utilize an attention mechanism to model the relation between all words in a sentence. They rose to greater heights than prior recurrent neural network architectures by affording more parallelism during training. In the case of LLM, usually each succeeding generation entails greater model parameters, data set size, and computation used in training.

For instance, GPT - 3, the model offered by OpenAI in 2020 has 175 billion parameters. It was trained on 45 terabytes of internet text for 3640 GPU years, and the computational cost was valued at over $10 million. It has only 340 billion parameters, but it was trained for 1.5 years with a budget over $100 million (OpenAI, 2023).

As evidenced by Gupta, Raj, Puri, & Gangrade, (2024) due to the vast data and compute resources needed to train LLMs, work in this area has so far been mostly concentrated in large tech firms and research labs. But in the case of LLMs, the situation is a little different as the expansion seems to be almost exponential and has not witnessed any decline. As each round proceeds, capabilities expand to more types of language tasks for content creation, translation, classification, and speech recognition.

The investigation by Sajun, Zualkernan, & Sankalpa, (2024) explores that some of the most prominent LLMs at present are Google's LaMDA and PaLM, Microsoft's Turing NLG, Anthropic's Claude, and Meta's OPT. A lot is expected about these models in terms of the effects that they may bring to society – the positive ones, as well as the negative ones. It is crucial to gain some insight into the principles of operation of LLMs as well as the processes of creating such systems so that one can make a knowledgeable contribution to the ongoing debates on the further evolution of LLMs and on the range of possibilities for their utilization.

3. Critical Sectors Impacted by Large Language Models

3.1 Healthcare

The use of large language models can help short - sighted healthcare practitioners in making immediate and accurate decisions to enhance patient care. These models can help to scroll the patient records and medical databases and offer the diagnostic clues and individual treatment advisories. They may also use the information technology system to get the most current results of clinical trials to help physicians make informed decisions about implementing new forms of treatment. Furthermore, language models can be applied also for patient's education or even to make a patient feel that there is someone who understands his or her situation when they are under medical care. However, to prevent any influence or adverse effects of bias or negative suggestions by the algorithm, measures must be put in place. These systems need to undergo thorough evaluation in a controlled setting with strict norms to prove that such systems are effective and not a danger to the actual healthcare environment. Specific examples of how the application of language models has been beneficial and risky would be useful for determining whether or not language models are feasible for decision making support in this high risk industry (Devlin, Chang & Lee, 2019).

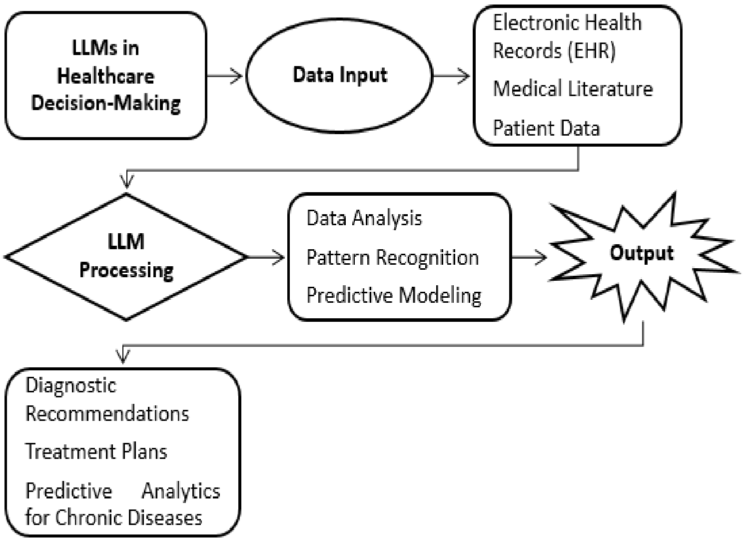

Figure 1: Integration of LLMs in Healthcare Decision – Making

3.2 Finance

New large language models have the potential to help in automating tasks, generating insights from big data, and in risk assessment in the context of deciding in the finance industry. But with the utilization of LLMs, they also come with new risks such as the concerns for biases, model interpretability, and model responsibility (Esteva et al., 2019).

Evaluations can be made through earnings calls, filings and news, or for generating trade ideas or tweaking risk models on the basis of alternative data. This can translate to more quantitative approach to the investment process as well as, higher chances of generating improved returns. However, if the models capture societal prejudices or are opaque, they force unfair results in most cases. Tight requirements control measures are required (Brown et al., 2020).

In fraud detection, the system could possibly aid in discovering new fraudulent patterns, which could be used to avoid future occurrences. Authoritative model behavior must therefore be audited to ensure the minimization of blind spots or false positives. Any alerts would need someone to review them given the nature and consequences involved (Brundage et al., 2020).

In sum, while the discussed LLMs offer numerous possibilities for improving decision making, it is important to approach their application in highly significant fields such as finance with due consideration of the existing ethical and governance concerns, as well as the responsible AI development. The technology is still relatively new and about best practice is still surfacing. Tight cooperation between academic disciplines will prove critical, especially for reaping value while shielding interested parties (Wolf et al., 2020).

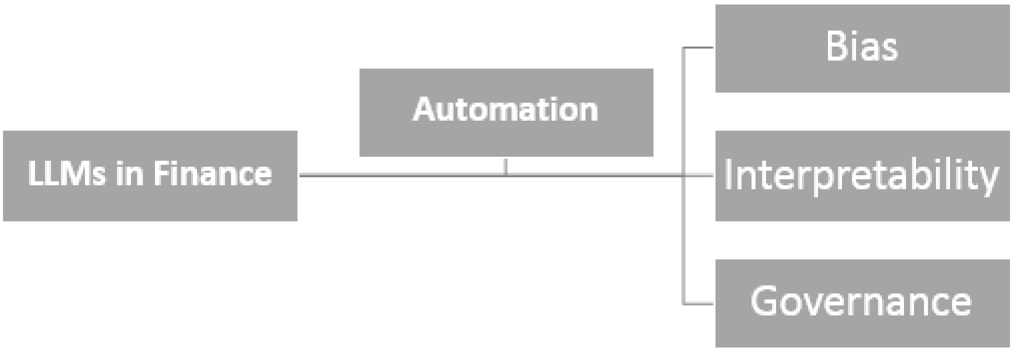

Figure 2: Impacts and Considerations of Large Language Models in Finance

3.3 Legal

The scale of these large language models hold the capability to revolutionize the way legal contracts are analyzed and reviewed by employing the latest in natural language processing. Bender, Gebru, McMillan - Major, & Shmitchell, (2021) study have found that by examining thousands of pages of legal documents, these systems can assist lawyers in performing fast and thorough contract due diligence by identifying the terms, risks, obligations and issues with specific contracts. Likewise, through their capability to perform protein structure prediction in a short while, large language models can help with database searches in case law research and legal precedents. They may also be able to look at the specific facts of the case when given and look at previous similar cases and previous pattern in relation to the case that would determine the strategy of the case (Bommasani et al., 2021).

Also, some the large language models are being trained to predict the legal consequences and probability of legal case based on the case characteristics and historical statistics. Though such predictive analytics tools remain in a fairly nascent stage currently, such forms could provide greater clarity as to the possible outcomes to the parties, thereby helping to inform the terms of the settlement. But there are also anxieties related to the possibility of the biases being encoded into the training data, which in turn would produce unfair outcomes (Rae et al., 2022).

As explained by Chowdhery et al., (2022) in general, as large language models become even more sophisticated in terms of their reasoning and language processing ability, lawyers could utilize these technologies in terms of contract analysis, legal research, and potentially even in the sphere of legal analytics of the cases. However, there is the need to put in place measures that could check the recommendations in line with legal and ethical perspectives. Thus, the large language models' ability to provide a solution to legal issues is promising as shown by case studies but also come with issues as illustrated.

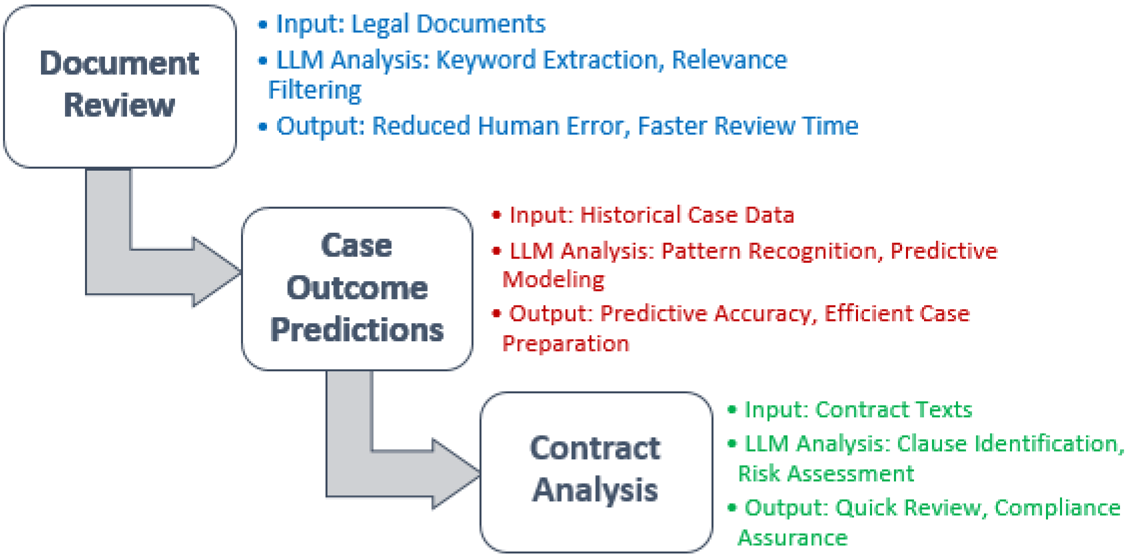

Figure 3: Efficiency Gains in Legal Decision - Making Using LLMs

3.4 Government and Public Policy

As reported by OpenAI, (2023) ChatGPT and other large language models may transform government decision - making and public policy since they are enormous systems capable of providing valuable information to governments. Policy makers may use these models to quickly scan through thousands of pieces of legislation, comments, and case laws and other policy documents and produce policy briefs, to help identify key actors and concerns, evaluate implications of proposed ideas, and propose research supported policy options (Bordt & von Luxburg, 2023). However, the use of such analysis also comes with the disadvantage of replicating and reinforcing pre - existing biases and trends in the training

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

sets. Policymakers should therefore employ large language models primarily as diagnostic aids rather than entrusting them with decision making and take measures to guarantee that the models were trained on high - quality and diverse data sets to reduce these risks (Gupta, Raj, Puri, & Gangrade, 2024). Extra monitoring might be needed to monitor for model biases; testing might be needed to audit them. When designed and executed appropriately, large language models have the potential to revolutionize many aspects of public policy by improving the synthesis of evidence, communicating with citizens, and considering secondary impacts, but these tools also require protective measures to ensure that they are used in a responsible manner to support sound, ethical policy decisions (Sajun, Zualkernan, & Sankalpa, 2024).

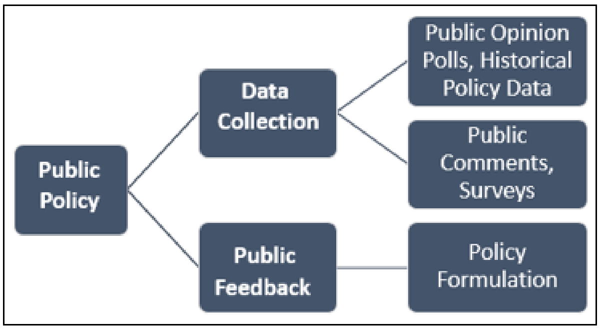

Figure 4: Policy Development Cycle with LLM Integration

3.5 Education

Self - learning through large language models such as ChatGPT has a high level of potential for revolutionizing education systems by making them more individual and dynamic. These models can offer personalized tutoring and feedback in terms of messages and lessons that may be provided as per the ability and intelligence level of the particular learner. They may also help teachers design more creative or appealing lessons and evaluate students' needs based on assessment results that may flag learning deficits. However, issues of the plagiarism, reduced reasoning capability, and dependency on the AI remain critical and warrant further study. Pilot implementations would be summarized by case studies, thereby highlighting the benefits and risks in contexts ranging from K - 12 education to university. In summary, it is highly possible for large language models to contribute to the enhancement of access, participation, as well as achievement, by being incorporated into educational contexts in a purposeful manner; however, it is important for this to be done intentionally with the guidance of educators and in a carefully planned and controlled manner (Devlin, Chang & Lee, 2019).

Table 1: Potential Benefits and Critical Issues of Using Large Language Models in Education

Aspect	Potential Benefits	Critical Issues	Recommendations
Personalized Tutoring	• Offers individualized lessons and feedback	• Plagiarism concerns	• Intentional incorporation with educator guidance
Creative Lesson Design	• Assists teachers in creating engaging lessons	• Reduced reasoning capability	• Controlled and purposeful implementation
Student Needs Evaluation	• Flags learning deficits based on assessments	• Dependency on AI	• Pilot implementations and case studies
Access and Participation	• Enhances educational access and participation		• Guided use in educational contexts
Achievement	• Contributes to improved student achievement		• Careful planning and control

Benefits of Large Language Models in Decision - Making

Promising technologies that reap scalable benefits include the large language models LLMs including GPT - 3, BERT among others in different key sectors. One of the positive effects of this is the increase in the accuracy and speed of the process. This makes the LLMs efficient in that, by analyzing large and complicated data sets, it is able to recognize patterns and linkages that would otherwise go unnoticed by human intelligence. This makes them to offer highly accurate predictions and recommendations that result in business advancements (Brown et al., 2020).

Furthermore, the insights from LLMs are evidence - based, which makes decisions more informed and less reliant on instinct or hunches. Where the application of general knowledge is not sufficient, such as in medicine and finance, the use of data is likely to produce better results. LLMs can take in patient data, financial statements and other types of information relevant in the field to showcase the best risk adverse decision (Bender, Gebru, McMillan - Major, & Shmitchell, 2021).

Thus, LLMs also reduce human error and bias that may arise when making decisions. As opposed to people, there are no issues with conceptual fatigue, or with the LLM getting exhausted or influenced by its emotions. This helps to minimize the likelihood that a decision will be reached based on cognitive biases, fatigue, and inattention or due to misunderstanding (Rae et al., 2022).

Last but not least, the availability and applicability of LLMs let the decision - making benefits of such systems be disseminated across the board. By using LLMs several user interfaces can be presented and there are no limitations in terms of volume of data within an LLM. This makes their insights always on the fingertips of decision makers in wholes organizations and industries (OpenAI, 2023).

Therefore, given that large language models bring accuracy, assist data - driven decision making, reduce the impact of human bias, and provide accessibility, these models can potentially enhance decision - making efforts in areas in which choices are critical. When they are applied and used appropriately and in an ethical manner, it will be effective in improving decisions and results (Gupta, Raj, Puri, & Gangrade, 2024).

4. Challenges and Risks

With a large and growing base of users and applications, there is rising concern about the reliability of the models underpinning artificial intelligence and their implications for society (Esteva et al., 2019).

One of the significant concerns is that any prejudice or mistakes may be inherited from the training data in the models and come up to discriminations or even harm. These are also black boxes with no readily discernable method to determine how they came to a particular decision, which makes it problematic to review them for flaws. It is contrary to individuals' expectation that they have to be able to understand why they receive certain recommendations or judgments, or are subjected to certain outcomes in their lives (Brundage et al., 2020).

Moreover, large language models rely on large datasets and there are privacy concerns as well as the matter of the actual consent regarding how people's personal sensitive data is employed. Breaches of data could also lead to leakage of highly sensitive data. However, as of now, there is little legal recommendation or model that can regulate the ethics of these models' formation and application (Bommasani et al., 2021).

If people take the advice from these bad models and do not question it, it will only extend the bias, much less breed public distrust. While language models become incorporated more into major institutions, individuals might rely heavily on computerized decisions with no information input or human intervention. This can potentially freeze undesirable trends, thus cementing problematic paradigms into place (Chowdhery et al., 2022).

According to Bordt & von Luxburg, (2023) continuing issues include that of seeing to it that models are socially responsible, and that Artificial Intelligence is created and deployed with the intention to be benign for man's kind. Solutions that build on these conditions require higher levels of transparency, auditing, the involvement of users and citizens in the design of ICTs, and policy - framework rules that ensure the integration of new ICTs with the public good (Sajun, Zualkernan, & Sankalpa, 2024).

5. Mitigation Strategies

In order to optimize the use of LLMs for making decisions and at the same time, manage corresponding risks, it is possible to apply the following strategies (Devlin, Chang & Lee, 2019)

Ensuring Ethical Use of LLMs

- **Bias Mitigation Techniques:** The main technique is to apply practices for dealing with bias found in training data and model outputs. In this, the use of varied data sets and the continued assessment of bias are also recommended (Brown et al., 2020).
- **Transparent Decision - Making Processes**: Another critical consideration is that it may be more comforting to LLMs if their decision - making processes are not opaque or black - boxed and if public reasons are given. These are all concerning justification about the process of arriving at a decision or the bodies that execute decisions (Wolf et al., 2020).

Enhancing Data Privacy and Security

- **Secure Data Handling Practices**: Measures that need to be put in place in order to control exposure of sensitive data is that high mature measures in handling and storing the sensitive data should be set. These are encryption, access control and security audit which are normally carried out at some interval time (Bommasani et al., 2021).
- **Compliance with Regulations**: Pursuant to these rules, the violation of user's rights to privacy and legal consequences of such laws as GDPR or HIPAA when using LLMs is excluded (Floridi et al., 2018).

Reducing Over - Dependence on LLMs

- **Human - in - the - Loop Approaches**: The involvement of the human factor in the decision - making process also has certain benefits in adding a human element to LLMs strengths. The strategy makes it possible to have an extensive study and validation of the most crucial decisions by human specialists (Rae et al., 2022).
- **Continuous Training and Education**: Ensuring that the LLMs themselves are employed by professionals with relevant continuing education maintains the effectiveness of the tools by preventing the professional from overestimating the LLM's capabilities, or underestimating them and failing to take full advantage of the LLM's potential.

Devlin, Chang & Lee, (2019) recommended with more evolution of LLMs in the future, they can potentially transform decision - making and support processes in various industries. However there are risks involved with the use of LLMs for critical decisions if the risks are not well controlled. It is recommended that sector - specific ethical requirements should be set and validation procedures should be put in place prior to the testing environment. There is the need for organizations to ensure proper training data for LLM, ensure that sample data is diversified and that there is careful observation of any biases (Brown et al., 2020). Outputs of LLMs should be reviewed by specialists in the subject area together with explainability techniques. In particular, LLMs should augment, not supersede human decisions in applications of higher risk and importance. New players will be expected to work together with the current ones while the latter will be forced to innovate and take risk in their operations but all under supervision of the regulators. In light of these issues, Bommasani et al., (2021) and (Rae et al., 2022) recommend that policy makers, engage technologists, legal scholars, and ethicists to provide understanding on the flexible governance systems that can fit an emerging technology. Concerted efforts by developers, users, and regulators, embracing principles of responsibility and openness, will ensure that LLMs can contributing to improving decision making in cases where this is possible and protecting against adverse effects. Further assessment as these models emerge will be crucial from time to time (OpenAI, 2023; Gupta, Raj, Puri, & Gangrade, 2024). By implementing these mitigation strategies, organizations can leverage the benefits of LLMs while addressing ethical, privacy, and dependence concerns. This balanced approach ensures responsible and effective integration of LLMs into decision - making frameworks (Sajun, Zualkernan, & Sankalpa, 2024).

6. Future Trends and Developments

LLMs are continuing to improve, and with each passing year, the capabilities and availability of models are growing. Since LLMs are increasing in hugeness and intricacy, they are supposed to be combined with other state - of - art methodologies like computer vision, robotics, and Multi - modal learning. It could also lead to even more potent implementation of applications in critical industries such as healthcare, education, finance and government (Esteva, C'esar, & Alfonso, 2019).

For example, LLMs can be combining with the computer vision to help identify the issues in the MRI scans and suggest the possible diagnosis to the doctors. In education, they can be behind intelligent tutoring systems that will improve learning to a certain student. With the assist of LLMs, quite a vast range of analysis functions in financial transactions can be implemented to improve the decision - making process for the various governmental organizations (Brundage et al., 2020).

This is why there is need to review the implications of the absorption of more LLMs on people=based organizations and the society in general. There are issues and questions that may be raised concerning bias, responsibility, and openness with LLMs that can be raised (Bommasani et al., 2021). Potential of the models in the future will be just as important as the one seen today to be kept safe and useful for the society; more research has to be done about AI alignment and control. He stresses that overall LLMs are a transformative technology and that managing their responsible development and deployment will be one of the key concerns of the following decades (Chowdhery et al., 2022).

7. Case Studies

In the field of healthcare and particularly in the medical profession, LLMs have proved useful in analyzing patients' data and literature in giving diagnosis and treatment advice. But more testing is currently required to be sure of the safety of the test and prevent the continuation of prejudices. There are various applications of LLMs in the legal sector for contract review and case advice (Rae et al., 2022). On the positive side, it has been noted that there has been an increased efficiency in the working of LLMs but the issue of the "black box" character of LLMs and its incapacity to expound on reasoning remains a concern. This kind of machine learning is still under research in the field of finance for portfolio analysis and credit risk determination (OpenAI, 2023). It is crucial to have these measures in place to avoid the emergence of more circumstances where market manipulations or crashes result from wrong predictions. To government, LLMs are potential areas to rationalize functions and offer services to citizens (Bordt & von Luxburg, 2023). Material and financial resources are not sufficient, yet accountability and transparency are nonexistent. In education, LLMs could positively contribute to availing quality and customized learning resources and materials, yet; the issues of protection of students' information and data, and equitable grading system cannot be turned a blind eye on (Gupta, Raj, Puri, & Gangrade, 2024). Indeed, in terms of scale and speed, LLMs have their advantage regardless of the industry; however, it is essential to put in place a mechanism of testing, auditing, and human intervention to set boundaries on LLMs. These are questions of fairness, accountability, and unforeseen side - effects that must not be left to emerge over time as the technology develops further (Sajun, Zualkernan, & Sankalpa, 2024).

8. Recommendations

Best Practices for Implementing LLMs in Decision - Making

- **Identifying Suitable Applications**: Organizations should carefully evaluate areas where LLMs can add the most value. This involves assessing the specific needs and potential benefits of integrating LLMs into decision - making processes (Devlin, Chang & Lee, 2019).
- **Ensuring Ethical Use**: The ethical standards and procedures have to be followed to eliminate bias to improve the efficiency of recruiting a diverse workforce. This comprises of deploying different training data and frequently checking model performances for any bias (Brown et al., 2020).
- **Building Robust Infrastructure**: It is therefore necessary that efficient technological frames work as the bedrock for the implementation and running of LLMs. This involves getting enough computational power, data space and security (Brundage et al., 2021).

Policy Recommendations for Regulators

- **Establishing Regulatory Frameworks**: Lack of guidelines and regulation seems to be an issue that regulators should take an interest in, with an emphasis on the ethical implications, data privacy, and security of LLMs (Rae et al., 2022).
- **Promoting Transparency and Accountability**: Using LLMs and policies should foster the disclosure of how the software operates and organizations being held accountable for its use. This also serves to sustain public confidence and address reasonable utilization (Brown et al., 2020).

Guidelines for Organizations

- **Training and Development**: By offering training and learning needs for the staff utilizing LLMs regularly and consistently, the limits of these models are made understood. It helps to create a work environment that encourages professionals to learn from experience and improve on that knowledge (OpenAI, 2023).
- **Ongoing Monitoring and Evaluation**: The performance of LLMs should be routinely assessed to establish any deviations, so the courses delivered meet the needs of the organization as well as the standards of ethical practice (Gupta, Raj & Gangrade, 2024).

These recommendations, therefore, help the organizations in implementing the LLMs so as to achieve the intended benefits with reduced risk.

9. Conclusion

Current developments of large language models demonstrate unprecedented capability to synthesize information and perform decisions on their own. However, their use in urgent necessary areas has come out to be issues to do with bias, fairness and information disclosure. It is because of this reason that the researchers, policymakers, and the companies must work hand in hand and design the ethical standards as well as the right evaluation frameworks as the LLMs are rapidly and continually emerging. Provided enough precaution measures are put in place, the potential benefits that come with the use of LLMs cannot be overstated as these technologies can improve the decision - making processes in areas of healthcare, finance, law, and many other sectors. In conclusion, there are many remaining questions, however, by taking the initiative and engaging in positive discourse, LLMs will be able to responsibly equip professionals for important positions.

References

[1] Bender, E. M., Gebru, T., McMillan - Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610 - 623. Retrieved from: https: //doi. org/10.1145/3442188.3445922

[2] 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*. Retrieved from: https: //doi. org/10.48550/arXiv.2108.07258

[3] Bordt, S., & von Luxburg, U. (2023). Chatgpt participates in a computer science exam. *arXiv preprint arXiv: 2303.09461*. Access from: https: //doi. org/10.48550/arXiv.2303.09461

[4] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert - Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J.,. . . Amodei, D. (2020). Language models are few - shot learners. *Advances in Neural Information Processing Systems, 33*, 1877 - 1901. Retrieved from: https: //doi. org/10.48550/arXiv.2005.14165

[5] Brundage, M., Avin, S., Wang, J., Belfield, H., Krueger, G., Hadfield, G.,. . . & Dafoe, A. (2020). Toward trustworthy AI development: Mechanisms for supporting verifiable claims. *arXiv preprint arXiv: 2004.07213*. https: //doi. org/10.48550/arXiv.2004.07213

[6] Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A.,. . . & Dean, J. (2022). PaLM: Scaling language modeling with Pathways. *arXiv preprint arXiv: 2204.02311*. Retrieved from: https: //doi. org/10.48550/arXiv.2204.02311

[7] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre - training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv: 1810.04805*. Retrieved from: https: //doi. org/10.48550/arXiv.1810.04805

[8] Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., Mottaghi, A., Liu, Y., Topol, E., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine, 25* (1), 24 - 29. Retrieved from: https: //doi. org/10.1038/s41591 - 018 - 0316 - z

[9] Gupta, A., Raj, A., Puri, M., & Gangrade, J. Ethical Considerations in the Deployment of AI. *Tuijin Jishu/Journal of Propulsion Technology, 45* (2), 2024. Retrieved from: https: //www.researchgate. net/profile/Aaryan - Gupta - 4/publication/380518796

[10] OpenAI. (2023). GPT - 4 technical report. *OpenAI*. Retrieved from: https: //www.openai. com/research/gpt - 4

[11] Rae, J. W., Borgeaud, S., Cai, T., Millican, K., Hoffmann, J., Song, F.,. . . & Irving, G. (2022). Scaling language models: Methods, analysis & insights from training Gopher. *arXiv preprint arXiv: 2112.11446*. Retrieved from: https: //doi. org/10.48550/arXiv.2112.11446

[12] Sajun, A. R., Zualkernan, I., & Sankalpa, D. (2024). A Historical Survey of Advances in Transformer Architectures. *Applied Sciences, 14* (10), 4316. Retrieved from: https: //github. com/huggingface/transformers.

[13] Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., & Rush, A. M. (2020, October). Transformers: State - of - the - art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations* (pp.38 - 45). Retrieved from: https: //github. com/huggingface/transformers.