

Large Language Model for Chatbot

Prof. Trupti Farande¹, Vishal Waghmare², Rushikesh Barkade³, Adesh Shinde⁴, Omkar Naikade⁵

Professor, Department of Artificial Intelligence and Data Science¹
Students, Department of Artificial Intelligence and Data Science^{2,3,4,5}
Shree Ramchandra College of Engineering, Pune, Maharashtra, India

Abstract: The integration of Artificial Intelligence (AI) and Natural Language Processing (NLP) has led to the development of sophisticated Chatbots capable of mimicking human conversation and providing automated responses. In the context of the mining industry, which operates under a complex framework of Acts, Rules, and Regulations, there is a growing need for a comprehensive and easily accessible information system. This research proposes the implementation of a 24/7 available Chatbot, equipped with the ability to address stakeholder and customer queries regarding various legal aspects, including the Coal Mines Act, 1952, Indian Explosives Act, 1884, Colliery Control Order, 2000, Colliery Control Rules, 2004, The Coal Mines Regulations, 2017, and The Payment of Wages (Mines) Rules, 1956. Furthermore, the Chatbot's scope will encompass land-related laws, such as Community Benefits Agreement (CBA), Land Acquisition (LA), and Resettlement and Rehabilitation (RandR), thereby establishing a robust Management Information System tailored to the specific needs of the mining industry.

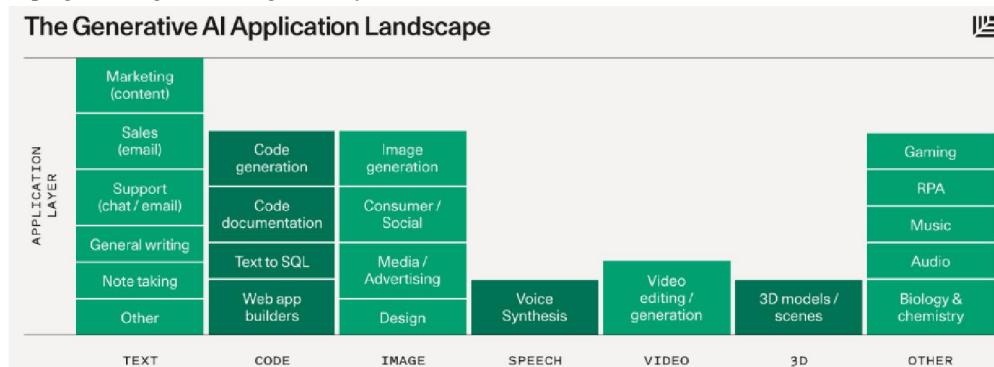
Keywords: Chatbot, Mining Industry, Natural Language Processing, Regulatory Compliance, AI Automation

I. INTRODUCTION

The mining industry involves adherence to numerous Acts, Rules, and Regulations, making it crucial to provide stakeholders and customers with an accessible platform to obtain relevant information. This paper proposes the development of an AI-driven Chatbot with Large Language Model capabilities that is powered by transformer technology^[1], catering to queries related to the aforementioned legal frameworks. The integration of land-related laws within the Chatbot's scope also aims to establish a comprehensive Management Information System^[2]. The proposed system enhances user experience^[3] by making the legal information available through a chat interface. It aims to utilize generative AI for education and learning^[4]. The core technologies powering this system are Natural Language Processing^[5], Large Language Model and Web technologies.

II. LARGE LANGUAGE MODEL

The development of the proposed Chatbot was primarily facilitated through the integration of advanced Large Language Models (LLMs), specifically tailored to comprehend and respond to complex queries related to the extensive legal landscape governing the mining industry^[6].



Source: Venture Beat^[9];

Large language models (LLMs) represent revolutionary advances in the fields of natural language processing (NLP) and artificial intelligence (AI). Characterized by their size and number of parameters, these models show an unprecedented capacity to understand language, formation, and the concept of meaning. LLMs work on the principle of using large amounts of data and computing power to learn complex patterns and representations of words.

Llama2 is a state-of-the-art large language model developed by Meta. It is pretrained on vast textual data and fine-tuned for specific tasks and domains. For our study, we used the pretrained Llama2 model and fine-tuned it using data from the coal mining industry.

III. MATERIALS AND METHODS

The following steps and methods were used to develop the system for chatbot and integrating it with LLM.

3.1. Data Collection and Preprocessing

A comprehensive dataset encompassing a diverse range of legal documents, including Acts, Rules, and Regulations pertinent to the mining industry, was curated and preprocessed for the training of the LLM. This dataset was carefully selected to ensure the inclusion of relevant information necessary for accurate responses to user queries.

3.2. LLM Training and Fine-Tuning

The preprocessed dataset was employed to train and fine-tune the Large Language Model (LLM), enabling the Chatbot to grasp and interpret the subtle language found in legal documents pertaining to the mining sector. The training of the LLM focused on building a strong grasp of the contextual complexities present in the Acts, Rules, and Regulations relevant to the mining industry.

3.3. Chatbot Implementation and Interface Design

During the implementation phase, we integrated the trained LLM into the Chatbot framework, enabling the system to handle user queries and offer contextually precise and reliable responses. The Chatbot's interface was carefully crafted for user-friendliness and accessibility^[7], ensuring that stakeholders and customers can effortlessly engage with the system and access timely information related to the legal and regulatory frameworks within the mining industry.

3.4. Performance Evaluation and Optimization

We conducted a thorough performance evaluation of the Chatbot, with a specific emphasis on its accuracy in understanding and responding to a wide range of queries pertaining to Acts, Rules, and Regulations in the mining sector. To improve the Chatbot's comprehension and its ability to provide precise and relevant information to users, we employed iterative optimization techniques.

IV. CONCLUSION

The deployment of the AI-driven Chatbot for Mining Industry Regulatory Information represents a significant leap forward in facilitating accessibility and ensuring compliance with the complex regulatory frameworks governing the mining sector. By harnessing the power of Large Language Models (LLMs), the Chatbot has successfully streamlined the process of disseminating pertinent information related to Acts, Rules, and Regulations, providing stakeholders and customers with a user-friendly interface for accessing real-time updates and insights. The Chatbot's capacity to deliver contextually relevant and precise responses underscores its pivotal role in fostering transparency, informed decision-making, and proactive adherence to regulatory standards within the dynamic landscape of the mining industry.

Acknowledgements: The research team acknowledges the invaluable support and guidance provided by Prof. Trupti Farande.

Declaration of competing interest. The authors declare no conflict of interest.

REFERENCES

- [1]. Vaswani, A., et al. (2017). "Attention Is All You Need." Advances in Neural Information Processing Systems.
- [2]. "Artificial intelligence in information systems research: A systematic literature review and research agenda" (2021) - International Journal of Information Management (IJIM)

- [3]. "The Design Everyday Things" book by Don Norman (1988).
- [4]. "ChatGPT for good? On opportunities and challenges of large language models for education" (2023) - Learning and Individual Differences
- [5]. "Speech and Language Processing" book by Daniel Jurafsky and James H. Martin (2000).
- [6]. Ministry of Coal, Government of India. (2022). "The Coal Mines (Special Provisions) Act, 2015."
- [7]. Norman, D. A., & Draper, S. W. (1986). "User Centered System Design: New Perspectives on Human-Computer Interaction."