

Utilizing a sophisticated language model for the coal mining sector.

Prof. Trupti Farande¹, Vishal Waghmare², Adesh Shinde³, Omkar Naikade⁴, Rushikesh Barkade⁵

¹Professor at Department of Artificial Intelligence & Data Science, Shree Ramchandra College of Engineering, Pune.

^{2,3,4,5} Student at Department of Artificial Intelligence & Data Science, Shree Ramchandra College of Engineering, Pune.

Abstract— The intricate legal landscape of the coal mining industry presents a constant challenge for stakeholders and customers seeking accurate and timely information. This research proposes a novel solution: a 24/7 chatbot powered by a large language model (LLM) to address legal queries pertaining to critical regulations and acts. Leveraging the LLM's capabilities for natural language processing, the chatbot offers users immediate access to a comprehensive knowledge base encompassing the Coal Mines Act, 1952, Indian Explosives Act, 1884, and other relevant regulations. Furthermore, it delves into land-related laws like Community Benefits Agreements and Land Acquisition, solidifying its position as a robust Management Information System tailored for the industry's specific needs. This paper explores the development and implementation of this innovative chatbot, highlighting its potential to revolutionize legal information access, promote compliance, and empower stakeholders within the coal mining industry.

Keyword: Large Language Model, Chatbot, Mining Industry, Natural Language Processing, Regulatory Compliance, AI Automation.

INTRODUCTION

Navigating the complex legal landscape of the mining industry poses a significant challenge for stakeholders, demanding constant awareness and interpretation of various acts and regulations. This paper proposes an innovative solution: a large language model (LLM) chatbot designed to streamline legal information access and promote compliance.

Empowered by natural language processing capabilities, the chatbot functions as a readily available virtual assistant, offering immediate responses to user queries regarding key mining regulations. Users can pose specific questions, such as "What does the Coal Mines Act, 1952, stipulate regarding safety inspections?" and receive clear, comprehensive answers in real-time, regardless of the hour.

Leveraging a vast knowledge base encompassing critical mining acts, land-related agreements, and relevant regulations, the LLM chatbot serves as a robust Management Information System tailored to the industry's specific needs. This readily accessible and user-friendly tool empowers stakeholders at all levels to readily comprehend and adhere to crucial regulations, fostering a culture of compliance and informed decision-making.

In essence, this paper explores the conceptualization and implementation of this AI-powered solution, highlighting its potential to revolutionize legal information access, optimize compliance adherence, and empower stakeholder engagement within the mining industry.

LARGE LANGUAGE MODEL

The cornerstone of the proposed chatbot's development lies in the integration of advanced Large Language Models (LLMs)^[1] meticulously tuned to comprehend and respond to intricate queries within the mining industry's expansive legal framework.

LLMs embody groundbreaking advancements in both Natural Language Processing (NLP)^[2] and Artificial Intelligence (AI). Defined by their sheer size and parameter count, these models demonstrate an unparalleled ability to grasp language structure, formation, and the underlying

concept of meaning. Their functionality hinges on processing massive datasets and leveraging substantial computing power to learn intricate patterns and word representations.

Our research employed Llama2, a cutting-edge LLM developed by Meta. This model undergoes pre-training on vast textual datasets before fine-tuning for specific tasks and domains. In our investigation, we utilized the pre-trained Llama2 model and further fine-tuned it using data tailored to the coal mining industry. This refined model forms the core of our proposed chatbot, enabling it to effectively navigate the complexities of mining-related legal information.

The development of this LLM-powered chatbot marks a significant step towards democratizing legal information access within the mining industry. It has the potential to revolutionize the way stakeholders interact with legal regulations, ultimately contributing to a safer, more efficient, and compliant mining sector.

METHODOLOGY

Following methods were used to develop the chatbot for coal mining industry and integrating LLM:

A. 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.

B. Training and Fine-tuning the LLM

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.

C. Implementing Chatbot and User Interface

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^[3], 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.

D. Evaluating Performance 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.

Result

The integration of Large Language Models (LLMs)^[4] into the development of the Mining Industry Regulatory Information Chatbot has yielded promising results, significantly enhancing the Chatbot's ability to comprehend and respond to complex text queries pertaining to the diverse Acts, Rules, and Regulations governing the mining sector. Leveraging the advanced natural language^[5] understanding capabilities of LLMs, the Chatbot demonstrated an unprecedented capacity to interpret intricate legal language, thereby ensuring the delivery of contextually relevant and precise information to stakeholders and customers.

A. Enhanced Query Comprehension

Through the incorporation of LLMs, the Chatbot showcased an enhanced ability to decipher user queries that involved nuanced legal language and intricate regulatory frameworks. The sophisticated language processing algorithms^[6] enabled the Chatbot to accurately interpret the intent behind the queries, providing comprehensive and informative responses in real time.

B. Contextually Relevant Responses

The chatbot is able to understand and process user queries. It is able to generate contextual responses with the training dataset^[7]. The user-friendly interface of the Chatbot, coupled with its seamless accessibility, contributed to a streamlined user experience, empowering stakeholders and customers to seek and obtain pertinent regulatory information^[8] effortlessly. The integration of LLMs played a pivotal role in enhancing the Chatbot's responsiveness and user engagement, thereby fostering a culture of transparency and accessibility within the mining industry.

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.

ACKNOWLEDGMENT

The research team acknowledges the invaluable support and help provided by Prof. Trupti Farande in carrying out the project.

REFERENCES

Reference Papers

- [1]. Ashish Vaswani, Noam Shazeer, "Attention Is All You Need.", Advances in Neural Information Processing Systems, 2017.
- [2]. ALAN F. SMEATON, "Using NLP or NLP Resources for Information Retrieval Tasks", Natural Language Information Retrieval, volume 7, pp 99–111.
- [3]. Norman, D. A., & Draper, S. W. "User Centered System Design: New Perspectives on Human-Computer Interaction", J. EDUCATIONAL COMPUTING RESEARCH, vol 3(1) 1987.
- [4]. Enkelejda Kasneci, E. Kasneci, "ChatGPT for good? On opportunities and challenges of large language models for education", Learning and Individual Differences, 2023, volume 103.
- [5]. Christopher Collins, "Artificial intelligence in information systems research: A systematic literature review and research agenda", International Journal of Information Management, Volume 60, 2021.

Books

- [6]. "Speech and Language Processing" book by Daniel Jurafsky and James H. Martin (2000).
- [7]. "Natural Language Processing with Transformer" by Lewis Tunstall, 2022.

Web Documents

- [8]. Ministry of Coal, Government of India. (2022). "The Coal Mines (Special Provisions) Act, 2015."