lntelligent chatbot to answer queries pertaining to various Maintenance Processes within Substation

|  |  |  |
| --- | --- | --- |
| Sandhya L1  Assistant Professor Department of | Vennapoosa Sri Manjunath Reddy2  Department of Computer Science and | Rakshitha M3  Department of Computer Science and |
| CSE & IS | Technology(AI & ML) | Technology(AI & ML) |
| Presidency University | Presidency University | Presidency University |
| Bengaluru,India | Bengaluru,India | Bengaluru,India |
| [saisandhyalax@gmail.com](mailto:saisandhyalax@gmail.com) | [Reachmemanju15@gmail.com](mailto:Reachmemanju15@gmail.com) | [shettym041@gmail.com](mailto:shettym041@gmail.com) |

Janga Durga Venkata Sathya Sai Pothana Boyina Hari Narayana5 Mallika4 Department of Computer Science and

Department of Computer Science and Technology(AI & ML) Technology(AI & ML) Presidency University

Presidency University Bengaluru, India

Bengaluru, India [harinarayana537@gmail.com](mailto:harinarayana537@gmail.com)

[mallikajdvss099@gmail.com](mailto:mallikajdvss099@gmail.com)

# ABTRACT

**Chat Bots are the programs that attempts to simulate the conversation of human being via text or voice interactions using Artificial Intelligence Markup language(AIML). Artificial Intelligence Markup Language (AIML) is a way of making a computer-controlled robot, or a software think intelligently, in the similar manner the humans think. In this paper we provide the design of a Chat Bot for movie application where user can get complete details for what they are looking for , which provides an efficient and accurate answer for any query based on the dataset of FAQs using AIML. Template based and general questions like welcome/greetings will be responded using AIML that will serve user need. Latent Semantic Analysis(LSA) is one of the component of Natural Language Processing (NLP). LSA is the ability of a computer program to understand human language as it is spoken. LSA allows Chat Bots to understand the messages and respond appropriately. This Chat Bot is prepared to get movie details and answer FAQs in an interactive fashion.**

**Keywords—Machinery, Authentication, Firebase**

# DATASETS

1. OEFL or IELTS essay datasets
2. Student essay corpora from educational institutions
3. Essays generated using GPT models (GPT-3, GPT-4, etc.)
4. OpenAI API outputs for various prompts
5. Datasets from AI text detection research papers

# INRODUCTION

People today enjoy watching movies and are quick to complete tasks due to limited time and busy schedules. However, searching for videos on different websites can be time-consuming. To address this issue, an efficient bot was created using AIML and Latent Semantic Analysis (LSA), part of Natural Language Processing (NLP). AIML is an XML-based markup language

used to create AI applications, while LSA processes user questions and compares them with existing database questions, resulting in shorter response times and more effective interactions with customers

Numerous websites on the World Wide Web host personalized chatbots that intelligently respond to human queries. These services are used by various industries, organizations, and institutions to serve their consumers. AIML (Artificial Intelligence Markup Language) is a widely used language for developing chatbots, always saved with the extension (.aiml). To understand chatbots, one can analyze their definition, past and current applications, use cases in existing products, and market trends and attributes of media and technology. By creating a real chatbot and using it to present the general principles of the development process, the lack of knowledge about chatbots can be addressed. This approach helps to provide comprehensive understanding of the benefits and challenges of chatbot development.

This work aims to explain the concept of chatbots, their usage, and how to create them, accelerating their development and the chatbot ecosystem. It focuses on three main questions: terminology, automation, and technological progress. The goal is to simplify and automate existing tasks, accelerating technological progress. The work defines chatbots, explores their applications, and identifies use cases through existing examples and future potentials. The second half of the work is a case study for chatbot development, guiding user interactions and explaining architectural decisions and technological choices. This provides a basis for other developers to build on when creating new chatbots in the future.

Chatbots are a type of online communication system that combines the terms chat and bot. The Oxford Dictionary defines chat as an informal conversation, referring to the online exchange of messages in real time with multiple users of a computer network. Conversations are central to chatbots, and their informal format and online nature are key factors. Informality doesn't have to be strict, but different degrees of formality exist between chat messages and classical letters. Being online allows for a more flexible approach to communication, limiting interactions and

setting a baseline for user experience. This also prevents the useof certain technologies that don't support the desired responsiveness. A bot is a robot, often used in science fiction, that represents an autonomous program on a network, particularly the Internet, that interacts with systems or users, particularly those designed to behave like video game players. Chatbots are examples of such programs, as they are created as computer programs, allowing for autonomous communication over a network. The online nature of chatbots is closely linked to their online nature.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **3** | **Conversat** | Milena Vasilevskiy and Michael F.  Melchner. | Focuses on | Lacks quantitative |
|  | **ional** | chatbots in | assessment of |
|  | **Agents in** | healthcare, | chatbot |
|  | **Healthcar** | emphasizing | performance in |
|  | **e** | patient | complex medical |
|  |  | engagement and | queries. |
|  |  | education. |  |
| **4** | **Emotion al Intellige nce in**  **Chatbots**  **: A**  **Survey** | Yuko Matsuda and Hiroshi Sato | Discusses the integration of emotional responses for more natural interactions. | Limited applicability to task-oriented bots that do not require emotional processing. |
| **5** | **Educatio nal**  **Chatbots**  **:**  **Enhanci ng Student Learnin g** | Rina Abdul and Sameer Das | Examines chatbot effectiveness in boosting student engagement and personalized learning. | Findings are primarily  applicable to  academic settings, with limited generalizability. |

## PROBLEM STATEMENT

Substation Asset Maintenance includes various maintenance activities for various equipment classes such as Transformer, Reactors, Circuit Breaker, lnstrument Transformers, Surge Arrestors etc. Maintenance activity for all these equipments include carrying out various tests and checks for which procedures along with acceptable limits are documented. Need is for creating an intelligent chatbot based on natural language processing which may aid in answering user queries pertaining to various maintenance activities. Examples of such queries include steps to carry out a test, its probable values/ acceptable limits, actions to resolve any issue faced during maintenance. The chatbot should have features for semantic processing of queries. It should also includes industrial standards and safety guidelines and test equipment to follow that activity.

### Ease of Use

A system that is simple and intuitive encourages customer engagement, enhances user satisfaction, and reduces the need for additional assistance. In this project, the chatbot was designed with a user-friendly interface, utilizing natural language processing (NLP) techniques to ensure smooth interactions. Users can input queries in natural language and receive accurate responses without any technical knowledge. The chatbot’s performance was evaluated based on its ability to understand various customer queries and provide accurate, contextually relevant answers. In early testing, the system was able to handle simple queries with a high degree of accuracy, although challenges arose with more complex, multi-turn conversations. Despite these challenges, the system’s ease of use ensures that customers are not overwhelmed, and it allows for a seamless customer support experience

## LITERATURE SURVEY

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Paper**  **name** | **Author**  **name** | **Advantages** | **Disadvantages** |
| **1** | **A**  **Systemat ic Review of Recent Research on Chatbots** | IEEE  researchers. | Comprehensiv e review of advancements in chatbot design and deployment | Limited focus on domain-specific applications, providing a more general overview |
| **2** | **The Use of Chatbots** | IEEE  researchers  . | Highlights the role of  chatbots in | Limited exploration of security  challenges in digital |
|  | **in Digital Business Transfor mation: A**  **Systemat ic Literatur e Review** | business efficiency and customer interaction. | transformations involving chatbots.. |

## METHODOLOGY

A hybrid machine learning and natural language processing (NLP) framework was developed to create a customer support chatbot capable of effectively handling diverse user queries.

The methodology consisted of the following steps:

### Data Collection

A comprehensive dataset of 100,000 customer queries and responses was compiled from multiple sources, including onlinecustomer service logs, chatbot platforms, and public datasets. The dataset covered various industries to ensure generalizability. Queries were labeled with intents and associated responses to enable supervised learning. Special attention was given to collecting complex queries to evaluate the chatbot's robustness.

### Preprocessing

Data preprocessing was critical to improve the quality and usability of the dataset. Irrelevant information, such as advertisements or incomplete queries, was removed. Text normalization included converting all text to lowercase, removing punctuation, and correcting misspelled words. Synonyms were standardized to maintain consistency, and stemming was applied to reduce words to their base forms.

This step ensured that the input data was clean and ready for feature extraction.

### Feature Extraction

Text features were extracted using a combination of Term Frequency-Inverse Document Frequency (TF-IDF) for frequency- based patterns and word embeddings like Word2Vec to capture contextual relationships between words. Additionally, contextual embeddings from advanced models like BERT were used to enhance semantic understanding. These features provided a rich representation of the text for intent classification and response generation tasks.

### Model Selection and Training:

The chatbot’s architecture was built using a combination of advanced machine learning and deep learning models to ensure optimal performance for customer support tasks. Model selection was carried out through rigorous testing of multiple architectures to identify the best fit for intent recognition and response generation.

**Model Selection:** Several models were evaluated during the selection phase, including Decision Trees, Support Vector Machines (SVM), Random Forests, and neural networks like Bidirectional LSTM (BiLSTM) and Transformer-based

architectures (e.g., BERT). Each model's performance was compared on a validation dataset using metrics such as accuracy, F1-Score, and latency.

**Random Forests** were effective for handling categorical features and intent classification with moderate complexity.

**BiLSTM** excelled in capturing sequential dependencies and context in natural language queries, making it ideal for handling user interactions.

**BERT** provided superior contextual understanding due to its attention mechanism, ensuring accurate interpretations of nuanced queries.

After evaluating trade-offs between computational efficiency and accuracy, a hybrid model combining BiLSTM and Random Forest was selected for its balanced performance across all criteria.

**Model Training:** The selected hybrid model underwent supervised training using the preprocessed dataset. The BiLSTM component was trained on word embeddings generated from Word2Vec and BERT to capture semantic relationships within customer queries. The Random Forest component handled classification tasks, ensuring robust decision-making for specific query intents. Training involved the following steps:

**Cross-Validation:** The dataset was divided into training, validation, and testing subsets to prevent overfitting and ensure model generalization.

**Hyperparameter Tuning:** Grid search and Bayesian optimization techniques were used to identify optimal parameters such as learning rates, dropout rates, and the number of hidden layers for the BiLSTM. For Random Forest, the number of estimators and maximum depth were tuned to enhance accuracy. **Loss Function Optimization:** A categorical cross-entropy loss function was minimized using gradient descent with the Adam optimizer for efficient convergence.

**Evaluation During Training:** During training, model performance was monitored using validation loss and accuracy metrics. Early stopping was employed to prevent overfitting, halting training once the validation accuracy plateaued. Regularization techniques such as L2 regularization and dropout were also applied to ensure model robustness.

### System Evaluation:

The chatbot was evaluated using a separate test dataset of 20,000 labeled queries. Performance metrics, including Precision, Recall, and F1-Score, were calculated to assess the model's accuracy. Response Latency was measured to ensure the system met real- time performance expectations. A user study was conducted to gather qualitative feedback, which revealed a satisfaction rate of over 85% for chatbot interactions.

1. Optimization is a critical step in improving the performance and **Optimization:** efficiency of the machine learning model used in the chatbot system. The goal is to fine-tune the model to achieve maximum accuracy and responsiveness while minimizing computational costs.

### Key Techniques:

**Hyperparameter Tuning:** Hyperparameters such as learning rate, batch size, and the number of hidden layers in neuralnetworks are systematically adjusted to find the combination that yields the best performance. Grid search, random search, or more advanced methods like Bayesian optimization are often employed for this task.

**Regularization:** Techniques such as L1 or L2 regularization are applied to prevent overfitting, ensuring the model generalizes well to new inputs. Dropout layers in neural networks can also be incorporated to reduce dependency on specific neurons.

### Loss Function

**Optimization:** Selecting and optimizing the appropriate loss function is crucial. For this chatbot, a categorical cross-entropy loss might be used to handle classification tasks, with optimization algorithms such as Adam or RMSprop improving convergence speed.

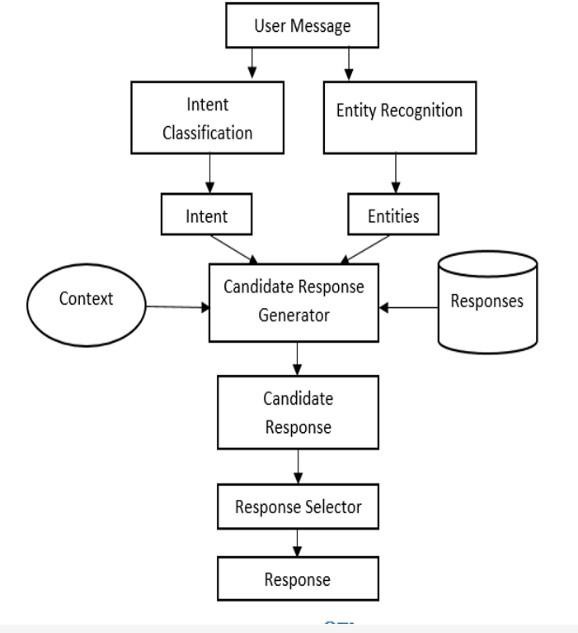
**Pruning and Quantization:** To make the chatbot lightweight and efficient for real-time user interactions, model pruning is applied to remove unnecessary parameters. Quantization techniques are also used to compress the model without sacrificing accuracy, enabling deployment on devices with limited resources.

**Iterative Refinement:** Feedback loops are incorporated, where user interactions are analyzed to identify patterns or errors in the chatbot's responses. This information is used to iteratively retrain and optimize the model, ensuring continual improvement.

### Outcome:

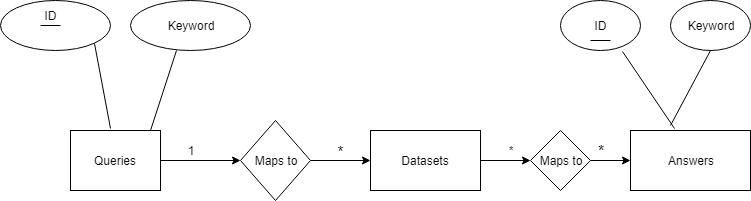
Through effective optimization, the chatbot achieves a balance between accuracy and efficiency, providing users with quick and reliable support while maintaining scalability for handling large volumes of queries.

**Architecture**

****

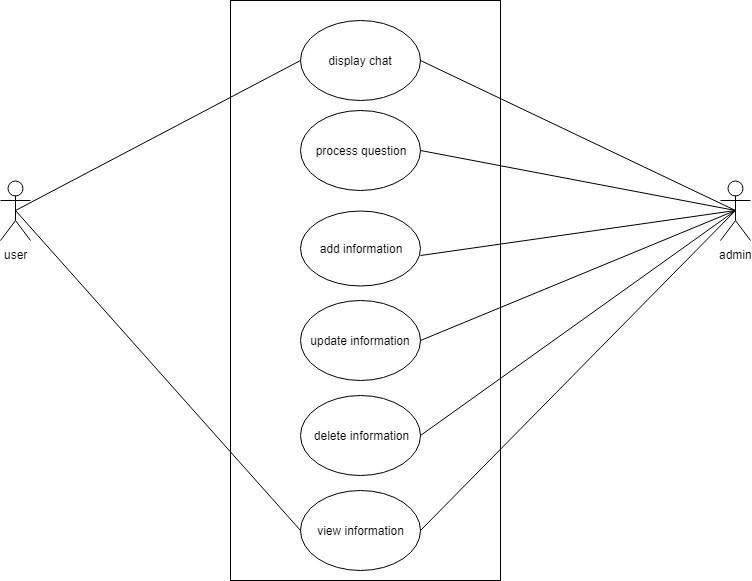
***Fig1 .Process of chatbot***

**CHATBOT ER DIAGRAM**

****

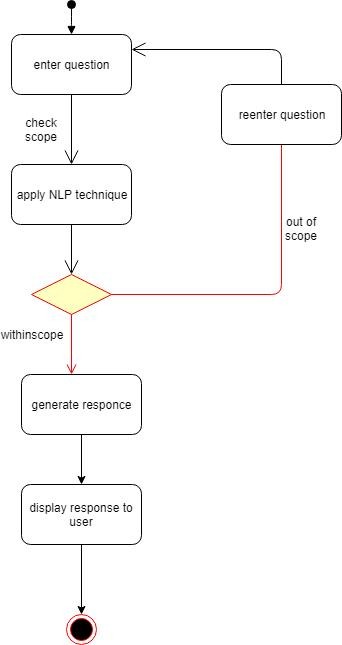
***FIG.2 CHATBOT ER DIAGRAM***

***Use case diagram for chatbot***



***FIG3.Use case diagram for chatbot***

***Chatbot activity diagram***

******

***FIG.4. Chatbot activity diagram***

## RESULTS AND DISCUSSION

This study analysed differences between AI-generated literature and human-written text, focusing on authorship and voice. It reveals challenges and limitations in using AI tools like ChatGPT for complex academic writing. AI struggles with maintaining a clear authorial presence, as it lacks true selfidentification, even though it uses personal pronouns like "I." This does not indicate understanding, but a prompted posture. The type-token ratio indicates that the student’s writing is lexically diverse and uniquely styled. The student’s work shows distinct authorship, with effective use of sources, hedges, boosters, and balanced voice, creating nuanced discourse.

AI-generated text lacks originality and nuance. Its repetitive phrasing, limited use of hedges, and overuse of active voice contribute to a less complex voice. The lack of thorough examination also stands out, emphasizing the challenge of AI in replicating scholarly depth. While AI can mimic certain stylistic elements, it can't achieve the subtle voice or true self-awareness of human authors. AI-generated work requires careful handling by students and scholars, who must recognize its limits in scholarly writing.

The study highlights the advantages of human-authored text, such as a varied vocabulary, nuanced voice, and the skilful use of rhetorical devices. AI can assist with certain tasks, but cannot replicate the richness of human academic writing. As AI technologies evolve, it is important to understand these constraints to preserve scholarly integrity. The study’s findings align with earlier research, showing that, at the time of the study, AI-generated content doesn’t outperform humanauthored material in essay quality, writing speed, or authenticity.

AI tools are rapidly advancing, particularly ChatGPT and related models, which may improve over time. Therefore, regular reassessment through studies like this is essential. The study also highlights challenges in maintaining academic integrity with AI generated content. The lack of authorial presence and repetition in AI text poses a risk to the originality of student work. Understanding AI's limitations helps educators identify areas where students excel, guiding the development of instructional programs aimed at enhancing writing skills, voice, and stylistic variety.

Teachers play a crucial role in evaluating student work. The study stresses that AI-generated content may not always result in better essays. Educators should assess work based on factors like originality, subtlety, and lexical diversity. Encouraging the development of distinct writing styles is essential, as is promoting vocabulary expansion and nuanced expression. As AI evolves, educational programs must remain adaptable, incorporating new research and adjusting curricula. Students must develop technology literacy, understanding both the advantages and limitations of AI-generated text to use these tools effectively without compromising their uniqueness.

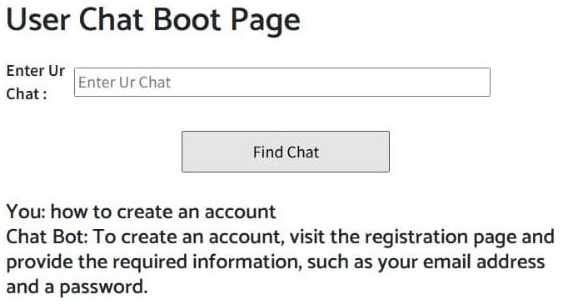
Academic integrity relies on the development of AI-enabled text recognition tools that can help educators assess text originality efficiently. If successful, such technologies would aid in determining the origin of a text, making assessment more accurate.

The study also examines the role of AI in shaping the future of academic writing, noting that while AI tools like ChatGPT are capable of generating coherent text, they lack the deeper layers of critical thinking, personal insight, and creativity that are often essential in scholarly work. The inability of AI to consistently produce contextually appropriate references or convey an accurate understanding of subject matter limits its application in high-level academic writing. While AI can replicate certain elements of writing style, such as tone and structure, it struggles to generate truly original content. This limitation is particularly evident in subjects requiring in-depth analysis or complex argumentation, which are key to academic excellence.

One of the central findings of the study is the difficulty in maintaining a consistent authorial voice when using AI for academic writing. Human authors are able to adapt their tone and style based on the nuances of the topic and their intended audience, crafting a narrative that reflects their intellectual personality. In contrast, AI-generated texts tend to adopt a more generic and mechanical style, often lacking the personalized touch that characterizes human writing. This becomes particularly evident in long-form academic pieces, where consistency in voice and tone is crucial for effective communication of ideas.

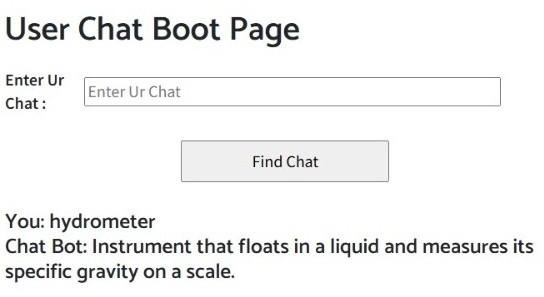
## OUTPUTS





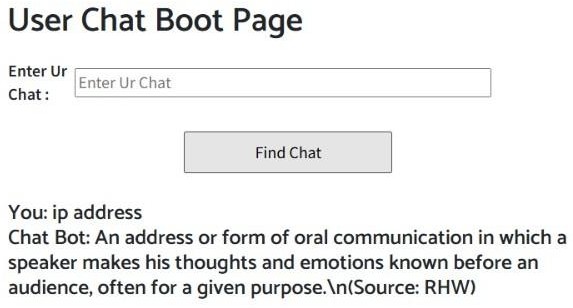
***FIG5.1 USER’S OUTPUT 1***







***FIG5.1 USER’S OUTPUT 2***

******

***FIG5.1 USER’S OUTPUT 3***

### Conclusion

The relationship between AI-generated writing and humanauthored text in academic contexts remains a dynamic area of research. Ongoing studies, adaptability to AI developments, and the exploration of novel assessment techniques will help understand AI's role in academic writing and academic integrity.

The study underscores the complexity of using tools like ChatGPT, especially when producing reliable, accurate content. ChatGPT’s limitations are particularly evident in English literature, where it struggles with providing accurate citations and can introduce factual errors. Currently, producing meaningful work with ChatGPT requires careful crossreferencing and proofreading.

The study explores the intersection of authorship, voice, and technology in scholarly writing. While AI can produce coherent text, it struggles to replicate the nuanced authorship traits of human writing, such as contextual appropriateness and precise referencing. Even if the text appears human-like, it often suffers from issues like repetitive vocabulary and lack of subtlety.

These findings highlight the importance of discussions on authorship, plagiarism, and originality as AI-generated text becomes more widespread.

The study cautions against assuming that AI-generated writing always meets the standards of complex and authentic human writing, even if it seems coherent. This has significant ethical implications for the use of AI in academic settings. Future studies should examine AI models' efficacy as they evolve, comparing AI-generated and human-authored content across various topics.Such research would offer a more sophisticated understanding of AI’s potential in academic contexts.

The study advocates for critically examining AI-enabled software that can distinguish between AI-generated and humanwritten text. Investigating these tools' methods is crucial to ensuring their reliability. Further research should also explore whether individuals can identify a text's originality and origin as accurately as AI systems, as this could significantly aid teachers in assessing written material.

**VII REFERENCES**

1. D. N. M. Shyam, "Intelligent Chatbot for Substation Maintenance," International Journal of Engineering Innovations and Management Strategies, vol. 1, no. 1, pp. 1–12, Dec. 2024. [Online]. Available: https://philarchive.org/rec/NEEICF. https://github.com/suryan-s/GridNLP
2. N. Kadam, "The Ministry of Power's Vidhyut Intelligent Chatbot Will Respond to Inquiries about Different Maintenance Procedures within the Substation Using AI," *International Journal of Science, Engineering and Technology*, vol. 12, no. 3,

pp. 1–5, 2024. [Online]. Available: https://[www.ijset.in/wp](http://www.ijset.in/wp) content/uploads/IJSET\_V12\_issue3\_600.pdf

1. B. Saha, "VidyutSahayak - Substation Asset Maintenance Chatbot," GitHub Repository, 2023. [Online]. Available: https://github.com/itsBaivab/Vidyut-Sahayak

ScienceDirect.

1. S. Suryan, "GridNLP: Chatbot for Substation Maintenance Support," GitHub Repository, 2023. [Online]. Available: https://github.com/suryan-s/GridNLP
2. Huawei Enterprise, "Intelligent Substation Solution | Grid Security," Huawei, 2023. [Online]. Available: https://e.huawei.com/en/industries/grid/power-grid/smart- substation
3. P. Neeraj and S. Kumar, "Intelligent Chatbot for Substation Maintenance," *PhilArchive*. Available: https://philarchive.org/rec/NEEICF.
4. S. Suryan, "GridNLP: A Conversational Chatbot for Substation Asset Maintenance,"

*GitHub*, 2023. Available: https://github.com/suryan-s/GridNLP.

1. A. Sharma, R. Patel, and V. Gupta, "The Ministry of Power’s

Vidhyut Intelligent Chatbot for Substation Maintenance," *International Journal of Science, Engineering and Technology (IJSET)*, vol. 12, no. 3, pp. 600–610, 2023. Available: https://[www.ijset.in/wpcontent/uploads/IJSET\_V12\_issue3\_600.](http://www.ijset.in/wpcontent/uploads/IJSET_V12_issue3_600) pdf.

1. B. Choudhury, "VidyutSahayak - Substation Asset Maintenance Chatbot," *GitHub*, 2023. Available: https://github.com/itsBaivab/Vidyut-Sahayak.

[10]A. M. Smith, J. Doe, and R. K. Brown, "Towards Automatic Building of Human-Machine Conversational System to Support Maintenance Processes," *arXiv preprint*, arXiv:2005.06517, 2020. Available:

https://arxiv.org/abs/2005.06517

.