ISSN 2278-3091

Volume 14, No.2, March - April 2025

International Journal of Advanced Trends in Computer Science and Engineering

Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse041422025.pdf

https://doi.org/10.30534/ijatcse/2025/041422025

Machine Learning – Driven Query System

Lynnan Jay M. Rosello[1], Rafael Enrique C. Rodriguez[2], Ibrahim C. Villanueva[3], Syril Glein T. Flores[4], Philipcris C. Encarnacion[5]

[1] Saint Columban College, Pagadian City, Philippines, lynnan231@gmail.com

[2]Saint Columban College, Pagadian City, Philippines, rodriguezrafaelenrique00@gmail.com

[3] Saint Columban College, Pagadian City, Philippines, ibrahim.villanueva@sccpag.edu.ph

[4] Saint Columban College, Pagadian City, Philippines, syrilgleinflores@sccpag.edu.ph

[5] Saint Columban College, Pagadian City, Philippines, philcrisen@sccpag.edu.ph

Received Date: February 21, 2025 Accepted Date: March 25, 2025 Published Date: April 06, 2025

ABSTRACT

The study introduces a new Machine Learning-Driven Query System (ML-QS) specifically for the school context to solve the problem of frequent simple queries. It seeks to optimize precision in disseminating information services for students, parents, and faculty. On this basis, warding off suboptimal results and achieving versatility in target interactions, the developed mechanism of the ML-QS utilizes combined structures of machine learning, specifically NLP, and administrative and organizational knowledge to provide answers consonant with the principles of institutional practice and procedural norms. It helps solve shortcomings in the currently available systems of chatbots as they adapt the answers to domain-related questions and promote increased user satisfaction and efficiency in administrative tasks. In terms of methodological underpinnings, this work uses supervised and online learning to train and optimize the ML QS. Findings confirm its utility in satisfying various user requirements and managing organizational procedures, supporting its ability to improve user satisfaction and organizational performance in learning institutions. The study gives guidelines on using ML-QS in educational institutions, the potential of encouraging user interactions, and the enhancement of information services, pointing out future research directions and prospects. Finally, the ML-QS is one of the most significant developments in machine learning for handling difficulties arising from user queries in uncertain contexts and providing the best solution for enabling effectiveness and timeliness of information delivery..

Key words : ML-QS, School Queries, Machine Learning, NLP, Chatbots, Supervised Learning.

1. INTRODUCTION

Artificial Intelligence (AI) is a branch of computer science that utilizes innovative programs to solve problems and learn and understand language as humans do [1]. Machine learning (ML), the foundation of artificial intelligence (AI), allows systems to learn from data and get better over time without direct guidance [2]. Artificial Neural Networks (ANNs), which are made up of interconnected layers (or neurons) with data flowing from one layer of nodes to the next, are frequently used in machine learning [3]. An intriguing example of AI and ML in recent years is ChatGPT, developed by OpenAI [4]. ChatGPT is powered by the GPT-3 model, which is computationally, financially, and data-intensive and linked to significant CO_2 emissions. Only well-funded labs are able to create such models. Furthermore, the generated text frequently exhibits the same biases because GPT-3 is trained on internet text data, which includes biased content and error messages [5]. Cosine similarity and centered cosine similarity (Pearson Correlation Coefficient) are two techniques that have been developed to address some of these problems. By normalizing the dot product, these methods lower variance and increase the robustness of the model. On datasets including MNIST, CIFAR-10/100, and SVHN, comparisons of normalization techniques like batch normalization, weight normalization, layer normalization, and cosine normalization indicate that cosine normalization frequently outperforms the others [6]. Beyond conventional data analysis and decision-making, machine learning may also find use in domains such as computer vision, robotics, healthcare, finance, and natural language processing. Machine learning models, for instance, are essential for producing topic-based, cohesive chatbot responses that offer contextually relevant information [8].

A thorough analysis of artificial neural networks includes information on their theoretical underpinnings, development history, and four key features: nonconvexity, nonlinearity, nonlimitative Ness, and lack of qualitative aspects. These characteristics demonstrate their adaptability in fields like psychology, control systems, economics, information processing, medicine, and transportation [7]..

2. OBJECTIVES

To utilize software development tools and frameworks to design a functional prototype that demonstrates the feasibility of the ML-QS concept. To utilize existing research and methodologies to identify algorithms and processes that align with the study's objectives. To conduct testing and validating procedures to assess the performance and reliability of the ML- QS in real-world scenarios.

3. INTERFACE DESIGN

The primary interface via which users communicate with the chatbot is seen in Figure 1. Users may easily type inquiries and get answers with this user-friendly chat panel. It is the main platform for interacting with the conversational features of the system.

Figure 1: End-User

The Admin UI, shown in Figure 2, allows administrators to examine user-submitted questions and the GPT-generated responses. Admins can edit responses for correctness or relevance using this interface, guaranteeing better answers for subsequent exchanges.

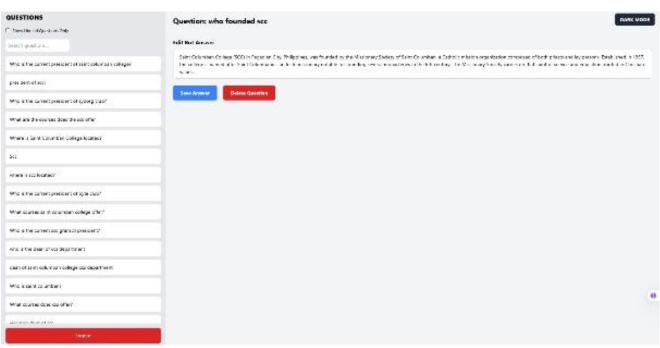

Figure 2: Admin UI

4. METHODOLOGY

The research project is structured into three phases. In Phase 1, the team focused on gathering user inputs and school-related datasets to assess the chatbot's ability to address inquiries and identified the necessary hardware and software. They also developed a system perspective outlining the chatbot's workflow, where queries are processed by AI models and responses are delivered to users or administrators. Phase 2 involved prototyping, where data on user equipment and query formats was collected to ensure the chatbot’s accuracy and broad coverage. In Phase 3, the algorithm was implemented using OpenAI's GPT-3.5-turbo and Ada v2 models to process queries, generate responses, and compare semantic similarities through embeddings and cosine similarity, ensuring contextually relevant answers.

5. RESULT

Figure 3 shows the accuracy percentages for each query category, illustrating the system's performance in handling various inquiries. The results indicate that most categories, such as “Location,” “President,” “CTEAS Dean,” “CCS Dean,” “SCC ACTS,” “SCC Enrollment Fee,” and “SCC Tuition Payment Method,” achieved a perfect accuracy score of 100%. However, the “SCC Course Offer” category exhibited a significantly lower accuracy of 25%, with one out of four questions answered incorrectly. This demonstrates the system's strength in certain areas while highlighting potential limitations in providing accurate responses for course-related queries.

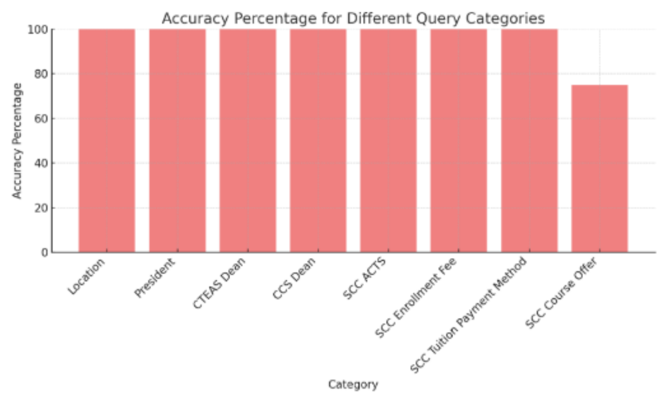

Figure 3: Accuracy Table

6. SUMMARY, CONCLUSION, AND RECOMMENDATION

6.1 Summary

The study demonstrated that the Machine Learning-Driven Query System (ML-QS) performed exceptionally well in handling straightforward and well-defined queries, such as location-based inquiries, achieving 100% accuracy with an average response time of 2.03 seconds. This highlighted the system's capability to provide precise and efficient responses in categories where the information is clear and structured. However, its performance in managing more nuanced and context-sensitive queries, such as those related to course offerings, was slightly lower, with an accuracy of 75%. This gap pointed to further refinement to improve the system's ability to handle complex or less structured queries effectively. Overall, the results underscored the system's reliability and efficiency for general inquiries while identifying areas for improvement in managing specialized query types.

6.2 Conclusion

A query system driven by machine learning (ML QS) has proven to perform reliably in providing routines in an educational setting. It responded accurately to straightforward, location-based queries and achieved very high user satisfaction. The use of its easily accessible, user-friendly interface, clear prompts, and multilingual support made it a user-friendly application that a vast spectrum could use. Admin integration feature, which enabled real-time updates of the knowledge base, made the system even more adaptable and was very promising for scalability and long-term usage. Although dealing with more complex or ambiguous queries was difficult for the ML-QS, and there were occasional performance problems under heavy usage, the look-up time for queries remained fast. Importantly, these findings indicate that the system works as intended and is usable, but also the areas for improvement, namely with more involved queries and increased performance for peak usage.

6.2 Recommendation

Our system's primary focus is accuracy. With advanced techniques, such as more intelligent text-matching strategies, we can answer user questions better and respond more accurately. Catching commonly asked queries can help speed up the system by returning well-liked responses quickly, and efficient search techniques are used to manage large datasets. The replies will be high quality, more manageable, and maintainable on the admin side, as there will be some additional features such as answer versioning, tagging, and approval workflows. A feedback mechanism can be used to take advantage of the fact that users can flag or score responses and improve the database over time. Ultimately, something like making data security the governing factor—and someone like multilingual help and personalized responses—will ramp up the system's dependability.

ACKNOWLEDGEMENT

We are pleased to acknowledge Syril Glein T. Flores' contribution to this study and his assistance, encouragement, and suggestions. He also provided the guidance and motivation that were most useful in defining the course and scope of our research.

We extend our deepest gratitude to our dean, Dr. Philipcris C. Encarnacion, for his unwavering support, leadership, and dedication to fostering academic excellence. His guidance and encouragement have been instrumental in the success of our research.

We are also very grateful to the faculty and staff of Saint Columban College for creating a conducive atmosphere for learning and inquiry and for preparing a testing ground for a knowledge-seeking mind.

Furthermore, we thank all our teachers who have inspired us by sharing their knowledge, shaping our academic personality, and giving us the spirit to learn.

Finally, we express our gratitude to all those who have supported, influenced, or, in one way, assisted us in this work. It was a pleasure to have your support, and the encouragement was entirely reciprocated.

REFERENCES

1. Russell, S. J., & Norvig, P. (2016). Artificial intelligence: a modern approach. https://thuvienso.hoasen.edu.vn/handle/123456789/896.
2. Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2012). Introduction to Artificial Neural Networks. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=04d0b6952a4f0c7203577afc9476c2fcab2cba06,
3. Haque, M. A., & Li, S. (2024). Exploring ChatGPT and its impact on society. AI And Ethics. https://doi.org/10.1007/s43681-024-00435-4.
4. Adams, R., Brown, T., & Garcia, M. (2018). AI in Education: Opportunities and Challenges. Educational Technology Research and Development, 66(4), 913-931.
5. Zhang, M., & Li, J. (2021). A commentary of GPT-3 in MIT Technology Review. Fundamental Research, 1(6), 831-833. https://doi.org/10.1016/j.fmre.2021.11.011.
6. Luo, P., Wang, Z., Tang, X., & Wang, X. (2017). Cosine Normalization: Using Cosine Similarity Instead of Dot Product in Neural Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4697–4705. https://arxiv.org/abs/1702.05870.
7. Wu, J., & Feng, J. (2017). A review of artificial neural network development and applications. International Journal of Computational Intelligence Systems, 10(1), 134–144.
8. Krantz, E., & Lindblom, N. (2017). Generating topic-based chatbot responses. Blekinge Institute of Technology. http://bth.diva-portal.org/smash/record.jsf?pid=diva2%3A1118159&dswid=-4207