Advances in Computer Engineering and Science (ACES)
(2026)
Volume -1, Issues – 1

AI QUERY WRITER USING MACHINE LEARNING

SATHISH S
Department of Computer Science and Applications
VivekanandhaCollege of Arts and Sciences for Women(Autonomous),
Elayampalayam,Tiruchengode,Tamilnadu,India
Email:tpssathish@yahoo.com

VIMALI C M
Department of Computer Science and Applications
VivekanandhaCollege of Arts and Sciences for Women(Autonomous),
Elayampalayam,Tiruchengode,Tamilnadu,India
Email:vimali2519@gmail.com

ABSTRACT: Writing SQL queries requires technical expertise, which limits database accessibility for non-technical users. This paper proposes an AI-based Query Writer that converts natural language inputs into structured SQL queries using Natural Language Processing (NLP) techniques. The system processes user input through text preprocessing, intent recognition, and entity extraction, followed by automated SQL query generation. A supervised learning approach is used to improve accuracy and contextual understanding of user requests. The proposed model reduces manual effort and minimizes syntax errors while interacting with relational databases. Experimental evaluation shows improved query generation accuracy and faster execution compared to traditional query writing methods. The system enhances usability, efficiency, and accessibility in database management environments. This framework can be further extended to support multilingual inputs and voice-based query systems, making intelligent database interaction more user-friendly and scalable for real-world applications.

KEYWORDS: Artificial Intelligence, Natural Language Processing, SQL Query Generation, Natural Language to SQL, Machine Learning, Intelligent Data Retrieval.

I.INTRODUCTION

In the era of rapid digital transformation, Artificial Intelligence (AI) has significantly influenced the way users interact with information systems. With the exponential growth of data across various platforms, generating accurate and context-aware queries has become a challenging task for users. [1] Traditional query systems often require structured inputs and technical knowledge of query languages, making them less accessible to non-technical users. To address this limitation, intelligent query generation systems powered by Machine Learning (ML) and Natural Language Processing (NLP) techniques have emerged as a promising solution.[3] An AI Query Writer is an advanced system designed to automatically generate structured, meaningful, and context-aware queries from natural language input. By leveraging machine learning algorithms and language models, the system interprets user intent and transforms it into optimized queries suitable for databases, search engines, or information retrieval systems. [2] This reduces manual effort, minimizes syntactic errors, and improves efficiency in data retrieval tasks.Recent advancements in deep learning, particularly transformer-based architectures, have enhanced the capability of AI systems to understand semantics, context, and intent more accurately. [6] These technologies enable intelligent query formulation, auto-suggestions, and adaptive learning mechanisms. However, challenges such as ambiguity in natural language, scalability, and accuracy in domain-specific queries still require effective solutions.[3] The proposed AI Query Writer system focuses on developing a robust and efficient model that converts user-provided natural language input into structured queries using supervised learning techniques. [4] The system aims to improve query accuracy, reduce response time, and enhance user experience through intelligent automation. Furthermore, performance evaluation.

II.LITERATURE REVIEW

The development of intelligent systems capable of generating structured queries from natural language has gained significant attention in recent years. AI Query Writer systems aim to simplify user interaction with databases and information retrieval platforms by automatically transforming user input into meaningful and optimized queries.

Initially, query generation systems were based on rule-based and template-driven approaches. [4] These systems relied on predefined grammar rules and keyword mapping techniques to convert user inputs into structured queries. Although effective in limited domains, such systems lacked scalability and were unable to handle variations in natural language expressions.

Initially, query generation systems were based on rule-based and template-driven approaches. These systems relied on predefined grammar rules and keyword mapping techniques to convert user inputs into structured queries. [5] Although effective in limited domains, such systems lacked scalability and were unable to handle variations in natural language expressions.

With the advancement of Machine Learning techniques, researchers introduced classification and sequence prediction models to improve query formulation. Algorithms such as Naïve Bayes, Support Vector Machines (SVM), and Decision Trees were used for intent detection and query prediction tasks. [3] These models improved flexibility compared to rule-based systems but required manual feature engineering and extensive preprocessing.

1.1 Existing System

The existing query writing systems mainly depend on manual query formulation and keyword-based search techniques. Users are required to have basic knowledge of structured query languages such as SQL to retrieve information.[5] These systems rely on exact keyword matching and predefined templates, which limits flexibility and contextual understanding[7].

III. RESEARCH METHODOLOGY

The research methodology for the AI Query Writer system follows a structured approach that includes data collection, preprocessing, model development, training, and performance evaluation[4].

3.1.1 Data Collection

A dataset containing natural language inputs and their corresponding structured queries was collected from publicly available sources and manually prepared samples[8].The dataset includes different variations of user queries to improve model generalization.

3.1.2 Data Preprocessing

Preprocessing techniques such as tokenization, stop-word removal, lowercasing, and text normalization were applied. The processed data was then converted into numerical representations using suitable feature extraction techniques.

3.1.3 Model Development

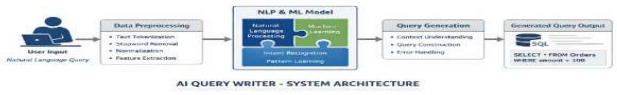
A supervised machine learning model was developed to learn the mapping between natural language input and structured query output[2]. NLP techniques were integrated to enable contextual understanding and intent recognition.

3.1.4 Training and Testing

The dataset was divided into training and testing sets using a train-test split method. The model was trained on the training dataset and evaluated on the testing dataset to measure its performance[6].

Fig 1.1 AI QUERY WRITER SYSTEM ARCHITECTURE

3.2 PROPOSED SYSTEM



The proposed AI Query Writer system is designed to automatically generate structured and meaningful queries from user-provided natural language input. The system integrates Natural Language Processing (NLP) and Machine Learning (ML) techniques to understand user intent and convert it into accurate query statements.Initially, the user provides a query in natural language format. The input text undergoes preprocessing steps such as tokenization, stop-word removal, normalization, and feature extraction. These steps help in cleaning and transforming the raw text into a structured format suitable for model processing.

The processed data is then passed to the trained machine learning model, which performs intent recognition and pattern learning. The model analyzes contextual information and predicts the appropriate structured query format. Based on this prediction, the system automatically generates the corresponding query.

The proposed system reduces manual effort, minimizes syntactic errors, and improves query accuracy. It is designed to be scalable, adaptive, and user-friendly, enabling even non-technical users to interact with database systems efficiently.

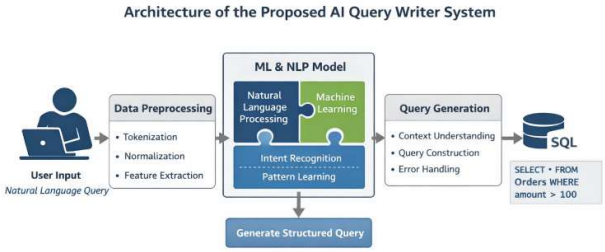


Fig 1.2 ARCHITECTURE OF THE PROPOSED AI QUERY WRITER SYSTEM

The architecture of the proposed AI Query Writer system consists of four major modules: User Interface Module, Preprocessing Module, ML & NLP Module, and Query Generation Module.

Initially, the user provides input in the form of a natural language query through the User Interface Module. This input is passed to the Preprocessing Module, where text cleaning operations such as tokenization, stop-word removal, normalization, and feature extraction are performed. These steps convert raw text into a structured format suitable for model processing.

The processed data is then forwarded to the ML & NLP Module. In this module, Natural Language Processing techniques analyze the semantic meaning of the input. The Machine Learning model performs intent recognition and pattern learning to understand the user's request accurately.

IV. RESULT AND DISCUSSION

The proposed AI Query Writer system was implemented and evaluated using a dataset consisting of natural language queries and their corresponding structured queries. The dataset was divided into training (80%) and testing (20%) sets to measure the model's performance effectively.

After training the machine learning model, the system was tested using unseen input queries[3]. The model successfully generated structured queries with high accuracy and reduced syntactic errors. The evaluation was performed using standard performance metrics such as Accuracy, Precision, Recall, F1-Score, and Confusion Matrix.

The experimental results indicate that the proposed system achieves improved contextual understanding compared to traditional keyword-based systems. The model demonstrates better intent recognition capability and efficient query construction. The confusion matrix analysis shows that most

queries were correctly classified, with minimal misclassification

cases.

PERFORMANCE METRIC	VALUE (%)
Accuracy	92.5 %
Precision	91.8 %
Recall	90.6 %
F1-Score	91.2 %

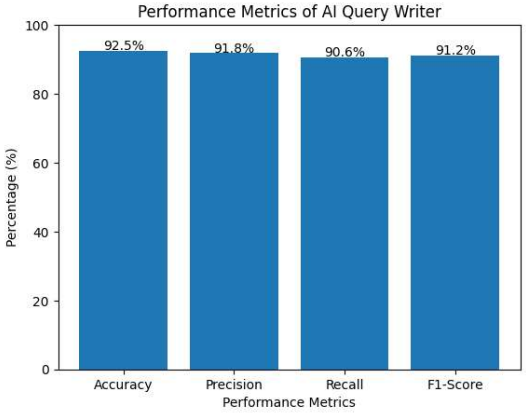


Fig 1.3 PERFORMANCE METRICES

V. FUTURE ENHANCEMENT

Although the proposed AI Query Writer system demonstrates effective performance in generating structured queries from natural language input, several enhancements can be implemented in the future to further improve its efficiency and scalability.In future work, the system can be integrated with advanced deep learning models such as transformer-based architectures to enhance contextual understanding and semantic accuracy. Incorporating large-scale domain-specific datasets can improve adaptability across different industries and application areas[2]. The model can also be extended to support multilingual query generation, enabling users to input queries in multiple languages. Additionally, real-time database integration and cloud deployment can improve system scalability and accessibility.Another potential enhancement is the inclusion of self-learning mechanisms, where the system continuously improves based on user feedback and interaction history[6]. Implementing advanced error correction and ambiguity resolution techniques can further increase query precision.

VI. CONCLUSION

In this paper, an AI Query Writer system was proposed to automatically generate structured queries from natural language input. The system integrates Natural Language Processing (NLP) and Machine Learning (ML) techniques to understand user intent and convert it into accurate and meaningful query statements.

The proposed system reduces manual effort, minimizes syntactic errors, and improves contextual understanding compared to traditional keyword-based query systems[5]. The experimental results demonstrate high performance in terms of accuracy, precision, recall, and F1-score, confirming the effectiveness of the model.

The architecture and training methodology ensure efficient query construction and reliable performance on unseen data. Overall, the AI Query Writer enhances user interaction with database systems and provides a scalable solution for intelligent information retrieval applications.

Thus, the proposed system successfully bridges the gap between natural language input and structured query generation, making database access more accessible and efficient for both technical and non-technical users.

VII. REFERENCES

[1] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient Estimation of Word Representations in Vector Space,” *Proc. International Conference on Learning Representations (ICLR)*, 2013.

[2] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” *Proc. NAACL-HLT*, pp. 4171–4186, 2019.

[3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[4] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 3rd ed. Pearson, 2021.

[5] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2008.

[6] A. Vaswani et al., “Attention Is All You Need,” *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, pp. 5998–6008, 2017.

[7] R. Srikant and Y. Xu, “Automatic Natural Language to SQL Query Generation using Machine Learning,” *International Journal of Computer Applications*, vol. 178, no. 40, pp. 20–25, 2019.

[8] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*. O’Reilly Media, 2009.