Enhancing Student Record Searches at Caraga State University using Mistral 7B Model

Jenelyn R. Asis, Edwin Manuel M. Jarlata, Hannah Faye A. Capoy
CCIS, Caraga State University, Butuan City, Philippines
jrasis@carsu.edu.ph

Abstract— This study evaluates the implementation of a generative AI system using the Large Language Model architecture with the Mistral 7B model to enhance student record searches at Caraga State University-Main Campus, particularly within the College of Computing and Information Sciences (CCIS). The current record retrieval process at Caraga State University, which relies on rigid system dropdown menus and predefined search conditions, is time-consuming, especially during student enrollment evaluations. Faculty members often need more time due to manual evaluations, specifically irregular student records, such as identifying subjects not yet enrolled and listing incomplete grades. This research assesses how generative AI with Mistral 7B can transform the process by allowing faculty to input natural language queries and receive accurate results without navigating multiple modules or systems. Data collection involved manual faculty evaluations and compiling of student profiles and grades from the university's MyWork website. Using the T-test statistical method, the AI-based system with Mistral 7B Model demonstrated a 97% reduction in time spent on record retrieval compared to traditional methods while maintaining 100% accuracy in generating relevant results. This significant improvement in operational efficiency alleviates the faculty workload and enables quicker decision-making during academic evaluations. The findings emphasized the broader applicability of generative AI in academic management, providing insights into how AI can improve the efficiency of administrative processes. This research contributes to the limited body of work on AI-LLM applications in student record management and highlights the potential of AI to streamline institutional operations in higher education. Future studies should explore system scalability and performance across academic programs and datasets.

Keywords – AI in Education, Automation, Generative AI, Large Language Model, Mistral 7B

1 Introduction

In the digital age, the efficient management and retrieval of student records are crucial for improving academic outcomes and supporting informed decision-making within educational institutions [1][2]. In higher education, student records primarily focus on semester grades, which is essential for monitoring academic progress and identifying areas requiring intervention. However, at Caraga State University-Main Campus, the current methods of accessing and analyzing these records, especially during enrollment evaluation, rely heavily on system rigid dropdown menus, predefined search conditions, and accessing multiple mywork website modules (e.g., student grades, evaluation, and Excel sheet curriculum guide), substantially leading faculty members in slowing down the enrollment process. Faculty often face the tedious task of manually evaluating especially irregular students, which involves tracing subjects not yet enrolled and listing INC grades. This process is time-consuming and prone to delays, limiting the faculty's ability to provide timely academic support [3]

Recent advancements in machine learning and deep learning architectures in almost every field [3]have also driven a surge of interest in Natural Language Processing techniques. Recently, the many approaches presented at SemEval 2024 [4] have further pushed the growing use of the Generative AI Large Language Model (LLM)-based architectures in academic research. Mistral 7B is one of the most discussed language models in the AI community because of its performance with just 7B parameters [4]. Mistral has been engineered for superior performance and efficiency, outperforming the best open 13B models across all evaluated benchmarks. Thus, this study explores the potential of generative artificial intelligence focusing on Mistral 7B to reduce the workload associated with student evaluations by transforming how records are accessed at CCIS-Caraga State University. Implementing a generative AI-based search system allows faculty to move beyond static queries and utilize natural language inputs to retrieve relevant student data more efficiently without relying on dropdown predefined search conditions and accessing multiple records.

This approach aims to streamline the evaluation process by automating complex queries and reducing the time faculty spend on manual record searches. This study aims to assess the effectiveness of generative AI using the LLM Mistral 7B Model in alleviating faculty workload during student evaluations, focusing on its ability to improve the accuracy of search results and reduce the Time spent on manual data retrieval. The findings from this study offer valuable contributions to academic record management and operational efficiency in higher education by demonstrating how generative AI can transform the retrieval and evaluation of student records.

2 Related Literature

The increasing spread of hate speech on social media has led to significant advancements in detection models, particularly in multilingual contexts. Hate speech, defined as offensive language targeting individuals or groups, contributes to misinformation, psychological harm, and social division [21]. While considerable research has focused on English-language hate speech detection, low-resource languages such as Cebuano and Tagalog remain underrepresented, creating a gap in automated content moderation [22], [23]. Machine learning has become a key tool in automating hate speech detection, with traditional classifiers like Support Vector Machines (SVM), Naïve Bayes, and Random Forest demonstrating effectiveness in monolingual settings [24], [25]. However, these models struggle with multilingual text and the common phenomenon of code-switching in Philippine digital discourse [26].

Deep learning models, particularly transformer-based architectures such as BERT, mBERT, and XLM-Roberta, have demonstrated superior performance due to their ability to analyze contextual meaning across multiple languages [27], [28]. Despite their advantages, these models require large, high-quality training datasets, which remain scarce for Cebuano and Tagalog [29]. The lack of annotated data limits the generalization capability of these models, reducing their effectiveness in low-resource settings [30], [31]. Moreover, even state-of-the-art transformer models struggle with recall when trained on limited datasets, reinforcing the need for dataset augmentation and language-specific fine-tuning [32].

One strategy to address data scarcity is the use of secondary datasets, which provide pre-labeled corpora sourced from social media platforms and prior research [33]. While secondary datasets offer a cost-effective solution, they pose challenges such as annotation inconsistencies, dataset bias, and domain mismatches [34], [35]. Researchers emphasize the need for validation techniques, including manual annotation checks and inter-rater reliability assessments, to ensure data consistency and quality [36]. This study applies preprocessing techniques, including tokenization, stop word removal, and TF-IDF vectorization, to refine secondary datasets and improve model accuracy.

Another significant challenge in hate speech detection is dataset imbalance, where non-hate speech instances vastly outnumber hate speech examples, leading to biased model performance [6], [37]. Various resampling techniques have been developed to address this issue, with the Synthetic Minority Over-sampling Technique (SMOTE) proving to be one of the most effective in text-based classification [38]. SMOTE generates synthetic samples for the minority class, enhancing recall and model robustness across multiple languages. This study implements SMOTE to mitigate class imbalance and improve hate speech detection performance for Cebuano, Tagalog, and English.

Despite advances in natural language processing (NLP), deploying hate speech detection systems remains a challenge. Misclassification, whether false positives or false negatives, presents ethical and practical concerns, as automated systems must balance censorship risks with the protection of marginalized communities [39]. Bias in machine learning models, particularly against dialects and underrepresented languages, continues to be a pressing issue [40], [41]. Hybrid approaches integrating machine learning with rule-based filtering have shown promise in improving detection accuracy, with ensemble techniques proving particularly effective in multilingual environments [42].

This study contributes to the growing body of research on multilingual hate speech detection by addressing the limitations of dataset imbalance, code-switching complexities, and bias in machine learning models. By leveraging both traditional and transformer-based models, this research provides valuable insights into NLP applications for low-resource languages and enhances the effectiveness of AI-driven content moderation in diverse linguistic communities.

3 Methods

This study adopts a quantitative approach to evaluate the effectiveness of implementing generative AI in enhancing the process of student record searches at CCIS - Caraga State University-Main Campus. The research involves quantitative data collection, emphasizing system usability and the efficiency of the AI-based search system compared to the existing method.

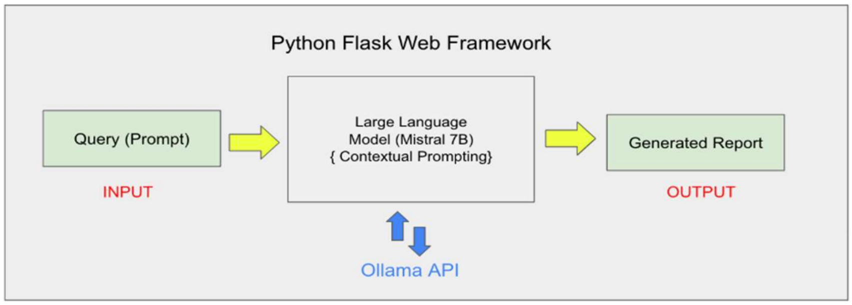

Fig. 1. Conceptual Framework of the Study

Figure 1 illustrates the process of generating a report from a user query using a Python Flask web framework and a Mistral 7B Model with the help of the Ollama API. The flow starts with the user (e.g., a faculty or administrator) submitting a query or request, such as "Show subjects with incomplete (INC) grades that have extended beyond two semesters.". The Mistral 7B Model interprets the prompt using natural language processing and retrieves relevant data from the context (student records), including courses, curriculum, subjects, and grades. The Mistral 7B Model generates a structured report based on the input query. The generated report is displayed on the user interface, providing actionable insights based on the original prompt

3.1 System Development

A generative AI-based search system is developed to address the current student record retrieval system's limitations. The system utilized the English Large Language Model - Mistral with 7B Parameters, integrated with the existing student database at Caraga State University, which focuses primarily on student profiles and grades per Semester. The study uses a Contextual-Prompting method for fine-tuning the model, which feeds the model-specific information in context. The AI system interprets natural language queries and returns relevant results based on student records.

Table 1. System Implementation

Key Steps	**Description**
Data Integration	The student profile and grade records will be extracted and formatted into a structure compatible with the AI-LLM system.
Mistral 7B Fine-Tuning	The Large Language Model, specifically Mistral with 7B Parameters, utilized a contextual prompting method for student record queries to accurately interpret and respond to user inputs based on the provided context.
User Interface Design	A simple and intuitive interface will be designed for users to input natural language queries without needing technical knowledge or predefined dropdowns.

3.2 Data Integration

The study used 100 sample student records from Caraga State University-Main campus specific to CCIS Third-Year IT students, representing the 2020 curriculum only. Each record includes the student's course, curriculum, enrolled subjects, and grades.

3.3 Formatting Data into Models

Data must be structured appropriately to use generative AI for effective student record searches. This enables the model to understand relationships and generate relevant results. Using contextual prompting, the model responds to queries based on the provided context. The following section outlines how data is formatted, from context definition to academic details.

Table 2. Data Format & Structure

Process	**Format**	**Actual Implementation**
Outlining Context	Student name || ID number || Enrolled Program	Arjane Claire Sanchez is a student with ID number 081-00113 enrolled in the Bachelor of Science in Information Technology program (BSIT).
Listing Curriculum and Subjects	Course Year Curriculum Year - Semester Subjects	The BSIT 2020 Curriculum has the following subjects per Year and Semester. These subjects are for First Year - 1st Semester ITE 10 - Introduction to Computing ITE 11 - Mathematical Applications for ITE ITE 15 - Social Issues and Professional Issues CSC 102 - Discrete Structures 1 PC - Purposive Communication STS - Science Technology Society PE - Physical Fitness NSTP - CWTST/LTS/ROTC 1 SE 101 - Orientation to University Life
Compiling Grades	Course Year Curriculum ID number || Grades || Subjects	BSIT 2020 Curriculum 221-01231 has a grade of 2.75 in TCW 221-01231 has a grade of 2.25 in ITE 16 221-01231 currently enrolled in CSC 106

3.4 Fine-Tuning Model

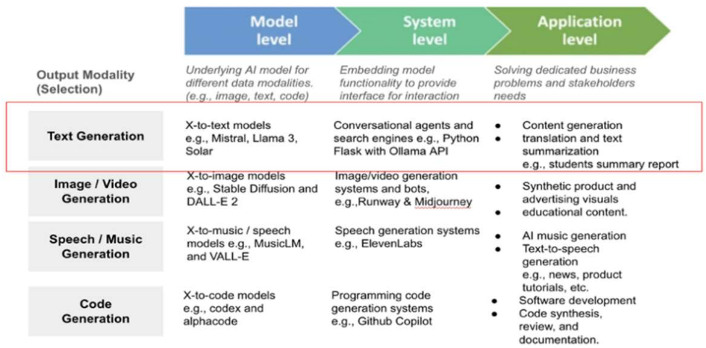

Fig. 2. Generative AI View Levels [6]

The figure shows generative AI's output modalities across model, system, and application levels, highlighting areas like text, image, speech, and code generation. This study focuses on **text generation**, using models like **Mistral with 7B parameters** for content creation, translation, and summarization, as well as applications like Python Flask Framework for generating student reports.

3.5 Model Architecture

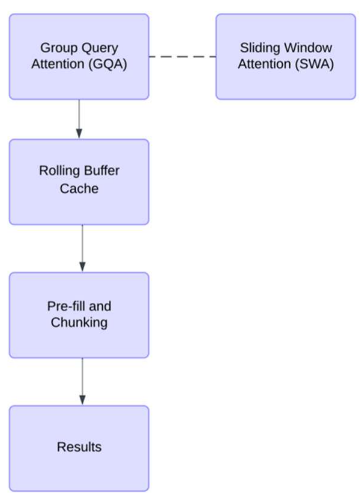

Fig. 3. Mistral 7B Architecture

The figure shows an optimized approach of Mistral 7B to handle long input sequences through advanced attention mechanisms. Group Query Attention (GQA) and

Sliding Window Attention (SWA) work together to manage large data efficiently. GQA reduces computational load by grouping similar queries, while SWA keeps track of context using a sliding window. A Rolling Buffer Cache stores previously computed information, preventing redundant processing, and a Pre-fill and Chunking step breaks down data into manageable pieces for efficient sequential or parallel processing. This setup enables Mistral 7B to produce accurate results on lengthy inputs, ideal for tasks like text summarization or multi-turn dialogues.

Table 3. Pseudocode of Mistral 7B using Contextual Prompting

```
Step 1. Libraries Initialization
Step 2. Dataset Preprocessing
        Create data contexts
        Tokenize dataset using Mistral tokenizer
Step 3. Initialize Mistral 7B
Step 4. For each batch of preprocessed data
        Input contextual prompt into the model
        Calculate loss using model predictions and true labels.
        Backpropagate loss to update model weights using the optimizer
Step 5. For each batch of validation data
        Input contextual prompt into the model
        Predict output using the model
        Calculate the evaluation metric
        Collect results for analysis
```

The process starts with Libraries Initialization and Dataset Preprocessing, where data is prepared, segmented, and tokenized. After initializing the Mistral 7B model, the Training phase involves feeding each data batch into the model, calculating the loss, and updating weights through backpropagation. During Validation, the model generates predictions for validation data, computes metrics, and collects results for further analysis, providing insights into model performance.

3.6 User Interface

Fig. 4. Sample prompt of subjects not yet taken

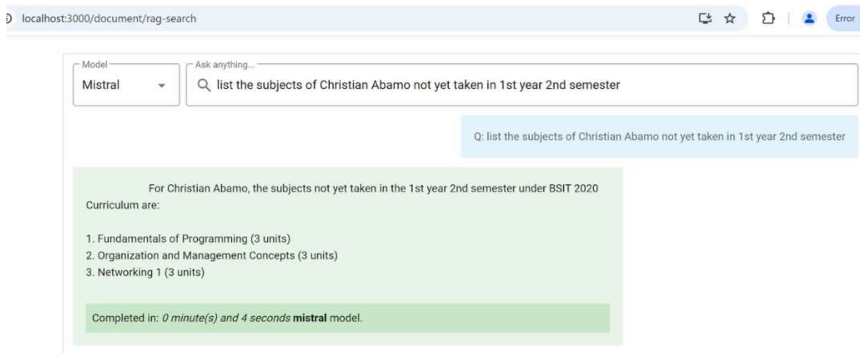

Fig. 5. Result for 1st Year-2nd semester subjects not yet taken by a student

3.7 Participants

The study involved faculty members from the CCIS college at Caraga State University-Main Campus. These participants are selected based on their involvement in managing and reviewing student records during enrollment. At least ten faculty are recruited for the system usability testing phase.

3.8 Data Collection

The study utilizes a quantitative method to assess the effectiveness of a generative AI system for student record searches at CCIS Caraga State University-Main Campus. This approach evaluates the AI system's impact on enhancing time efficiency and accuracy of results in academic record management using the Mistral 7B Model. The research examines the system's influence on improving time efficiency and the accuracy of academic record management. Data collection includes manual assessments by faculty and information extracted from the MyWork website, specifically student profiles and grade records.

4 Results & Discussion

Implementing the generative AI-based search system using Mistral 7B at CCIS-Caraga State University aimed to streamline the evaluation process for faculty members. This section presents the testing results, focusing on its impact on faculty workload and time efficiency during enrollment student evaluations.

4.1 Time Spent on Record Retrieval

A comparison was conducted for student subject evaluation by measuring the time taken for manual record retrieval versus the time taken using the Mistral 7B method system of a third-year BSIT student, where the AI system had been preloaded with the 1st- and 2nd-year grades of 100 BSIT students.

Table 4. Comparison of Time Spent on Record Retrieval

Method	Attempts	Average Time Spent	Std Dev (seconds)	Percentage Reduction (%)
Manual Method	5x	7 minutes	32.58	----
Generative AI Mistral 7B Method	5x	10 seconds	2.19	97.62%

Using the T-test statistical method, this table presents the average time spent by faculty members on student subject evaluation using two different methods: the existing or manual method and the generative AI with Mistral 7B model. The table includes the percentage reduction in time achieved through implementing the generative AI system. Faculty members performed five (5) attempts that spent an average of 7 minutes retrieving student records using the existing system, which relies on manually listing the taken subjects to identify the unenrolled or behind ones. When using the generative AI system, the average time spent on retrieval decreased to 6 minutes and 50 seconds, making the process roughly 42x faster with a 97.62% reduction in time. The table also includes the standard deviation, indicating time variability. The existing method had a standard deviation of 32.58 seconds, while the generative AI with the Mistral 7B method had a standard deviation of just 2.19 seconds, reflecting more consistent performance.

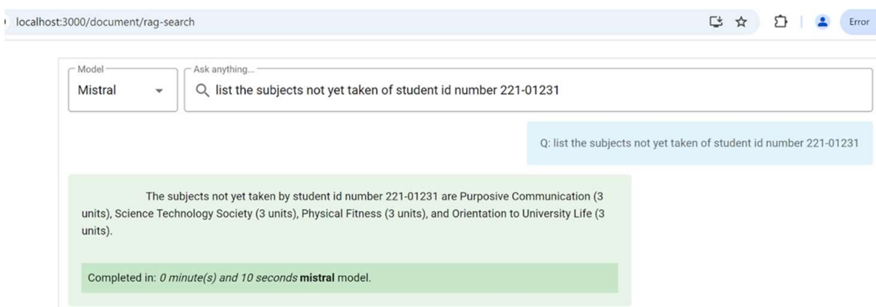

Fig. 6. Sample result of all subjects not yet taken from the previous semester

4.2 Comparative Analysis

All the computational operations were run on a laptop unit with 11th Gen Intel Core i5-11400H 2.70 GHz CPU processor, 32GB DDR4 RAM, 6GB NVIDIA RTX 3060 GPU, and Windows 11 Operating System with records of 100 BSIT 3rd-Year Student. Moreover, the following is the performance comparison between the three (3) Large Language Models (Mistral, Llama 3, and Solar) employed in Generative AI Systems.

Table 5. Comparative Analysis of Three (3) Large Language Models

Large Language Models	Average Time Spent	Result Accuracy (%)	Hardware Usage (%)		
			Memory	CPU	GPU
Mistral 7B	10 sec.	100	13.75	38	90
Llama 3	27 sec.	100	17.81	70	68
Solar	75 sec.	100	20.0	84	33

5 Conclusion

The comparison of Mistral with 7B Parameters, Llama 3, and Solar reveals notable differences in performance and hardware utilization. **Mistral** is the fastest, completing tasks in an average of 10 seconds, with moderate memory and CPU usage, but relies heavily on the GPU. **Llama 3** takes 27 seconds, using more memory and CPU resources while moderately relying on the GPU. **Solar**, the slowest model, takes 1 minute 15 seconds, consuming the most memory and CPU resources but the most minor GPU usage. Overall, Mistral offers the fastest performance due to its heavy GPU reliance, while Solar is more CPU-bound and slower, making it less efficient.

The results indicate that integrating generative AI into academic record management can dramatically streamline the evaluation process, reducing faculty workload and improving operational efficiency. This can lead to quicker decision-making and more timely academic interventions for students. While the AI system showed promising results, the study was limited to a specific dataset of 100 student records from the 2020 curriculum of third-year IT students only. The system's performance might differ with a more extensive and diverse dataset from different academic programs. Also, it may vary depending on the hardware used, particularly in environments with higher processing power or other configurations.

Future research should explore the system's scalability by including a broader range of student records from different programs and academic years. Investigating the performance of generative AI on varied hardware configurations and exploring the potential of integrating more advanced models or hybrid AI models could further enhance the efficiency and adaptability of the system across diverse educational settings. Further study should delve into the nuances of prompt construction and examine how variations in prompt design influence the performance of AI models.

References

[1] A. Falolo, T. Kim, and J. Lee, "Enhancing Decision-Making in University Administration Using AI-Powered Data Retrieval," Journal of Artificial Intelligence in Education, vol. 29, no. 3, pp. 259–276, 2022.

[2] N. Mahajan, "Challenges in Higher Education Record Management: A Case Study on AI Implementation," in Proc. 2023 IEEE Symposium on Educational Technology (SET), 2023, pp. 88–94.

[3] R. Duruin and G. Siddayao, "Development of Student Records Management System of Magalalag National High School," AIDE Interdisciplinary Research Journal, vol. 8, pp. 84–94, 2024.

[4] S. Siino and I. Tinnirello, "Exploring Natural Language Processing for Academic Research: A Review of SemEval 2024 Contributions," in Proc. 2024 International Conference on Computational Linguistics, 2024, pp. 302–310.

[5] Jing, A. Sablayrolles, et al., "Mistral 7B: A High-Performance Large Language Model," arXiv preprint, 2023. doi: 10.48550/arXiv.2310.06825.-D. Darwish, , "Advancements in Machine Learning and Deep Learning for Education Systems," DOI: 10.4018/979-8-3693-2913-9.ch006

[6] S. Feuerriegel, J. Hartmann, C. Janiesch, et al., "Generative AI," Business & Information Systems Engineering, vol. 66, pp. 111–126, 2024. doi: 10.1007/s12599-023-00834-7.
[7] Y. Jiang, B. Liu, and H. Zhao, "Performance Analysis of Lightweight AI Models: Mistral 7B vs. Llama 3," IEEE Access, vol. 12, pp. 10245–10258, 2023