DB Learn Space

Section 2022 - 2026

Submitted by:

Ruhab Fatima **2022-CS-69**

Fatima Raees **2022-CS-100**

Supervised by:

Sir Samyan

Department of Computer Science

University of Engineering and Technology Lahore, Pakistan

Abstract

This report presents DB Learn Space, a web-based platform designed to simplify database learning through AI-driven platform. The platform allows users to upload database schemas dynamically and populate these schemas. Key features include schema population, adaptive learning, and performance tracking, providing a comprehensive learning experience for students and educators. By using technologies like Django, MS SQL, and AI models such as LLAMA and Meta, the project ensures accurate question generation and personalized learning paths. The system architecture, implementation strategies, and evaluation results are discussed in detail.

Contents

1 Introduction

1.1 Problem Statement

Learning database concepts can be challenging, especially for beginners, due to the lack of practical platforms that generate real-time queries and provide feedback. Existing platforms often require predefined schemas or lack adaptive learning capabilities. There is a need for a dynamic, user-friendly platform to help students practice and understand database concepts effectively.

1.2 Objectives

The main objectives of the project are:

- Provide an automated platform for database query practice.
- Allow users to dynamically upload schemas for populating schemas.
- Generate database questions and queries automatically using AI models.
- Implement adaptive learning to adjust question difficulty based on user performance.
- Track user progress and provide actionable feedback.

1.3 Scope

The scope of the project includes the development of a web-based platform helping students and teachers. Key functionalities include schema population, query generation, progress tracking, and adaptive learning. The system aims to simplify the learning process for beginners while offering advanced features for experienced users.

2 Product Features

2.1 User Management

Role-based access for students, teachers, and guest users ensures a personalized experience. Registered users can save progress, while guest users can explore the platform without creating an account.

2.2 Schema Population

Users can upload predefined schemas in formats like Json, SQL or MS SQL. The system automatically validates schemas for correctness and then populates it.

2.3 Question Generation

Then there is a module to generate practice questions and queries based on the uploaded schemas. The questions range from basic SELECT statements to complex JOIN operations.

2.4 Adaptive Learning

The platform tracks user performance and adjusts the difficulty of generated questions in real time. This ensures an enhanced learning experience, helping users gradually master database concepts.

2.5 Performance Tracking

Detailed analytics and progress reports allow users to monitor their learning journey. Metrics include accuracy, time taken per query, and overall improvement trends.

2.6 Guest Practice Mode

A practice mode for unregistered users allows them to explore the platform's features and practice database concepts without signing up. These objectives demonstrate how the Smart DB Learn Space simplifies and enhances the learning of database concepts through innovative features and user-friendly tools.

System Architecture

3 System Architecture

3.1 Overview

The system architecture consists of a frontend, backend, and database. A central AI module facilitates dynamic schema validation and question generation. The platform ensures seamless integration between these components to deliver a user-friendly experience.

3.2 Components

Frontend: Built using HTML, CSS and Javascript for an interactive user interface.

Backend: Django serves as the backend framework, handling user requests, schema processing, and AI model integration.

Database: A relational database (MS SQL) stores user data, schemas, generated questions, and performance metrics.

AI Integration: AI models like LLAMA handle query and question generation, ensuring accuracy and adaptability. For population of schema SQL alchemy and faker is used.

3.3 Workflow

Users interact with the platform via the frontend. Input schemas are processed by the backend and validated by the AI module. Questions are generated dynamically and presented to users for practice. Performance data is stored in the database for future analysis.

4 Implementation

4.1 Tools and Technologies

- **Programming Languages:** Python, JavaScript
- **Frameworks:** Django (backend), Html,Css and Javascript (frontend)
- **AI Models:** Meta and LLAMA models for question generation
- **Libraries:** SQLAlchemy for schema validation, Pandas for data analysis

4.2 Features Implementation

Schema Population: Implemented using SQLAlchemy to parse and validate user-provided schemas.
Question Generation: Used LLAMA to create queries and questions based on schema structure.
Adaptive Learning: Integrated an adaptive mechanism to adjust question difficulty dynamically. If a student fails to attempt the question right or it takes too long to solve then the system considers it.

4.3 Key Features

- Populated schema for MySQL and JSON format inputs using SQLAlchemy and Faker.
- Dynamic question generation using the LLAMA model and Meta.
- Query checking by comparing input queries and expected answers.
- Adaptive learning with AI models to personalize user experiences.
- Role-based access for students, teachers, and guest users.
- Real-time performance tracking and practice mode for schema design without login requirements.

5 Evaluation

The correctness of the generated questions and corresponding SQL queries will be evaluated by a database expert (DB Master). This evaluation method was chosen to ensure high-quality validation and address the potential limitations of automated evaluation methods.

The evaluation process involved the following steps:

1. **Question Review**: The expert assessed whether the generated question accurately aligned with the intended database schema and context.
2. **Query Review**: The expert verified that the SQL query generated by the system was correct, executable, and produced the expected results on the provided schema.
3. **Feedback Collection**: The DB Master provided detailed feedback on the issues (if any) in the questions or queries and suggested improvements for incorrect outputs.

5.0.1 Accuracy Based on Expert Evaluation

To quantify the expert evaluations, the feedback was categorized as follows:

- **Correct Questions and Queries:** Questions and their corresponding queries that were deemed accurate and valid by the expert.
- **Partially Correct:** Outputs that required minor modifications to be fully valid.
- **Incorrect:** Questions or queries that were entirely invalid or unrelated to the schema.

The accuracy was then calculated as:

$$\text{Accuracy} = \frac{\text{Number of Correct Outputs}}{\text{Total Number of Outputs Evaluated}} \times 100$$

For example, if 80 out of 100 outputs were correct, the accuracy would be:

$$\text{Accuracy} = \frac{80}{100} \times 100 = 80\%$$

5.0.2 Confusion Matrix from Expert Feedback

Based on the expert's feedback, a confusion matrix was constructed to classify the system's outputs:

- **True Positives (TP):** Correct questions and queries validated as correct by the expert.
- **False Positives (FP):** Incorrect queries or questions falsely marked correct by the system.
- **True Negatives (TN):** Outputs flagged as invalid that were indeed invalid.
- **False Negatives (FN):** Valid outputs flagged as incorrect by the system.

6 Results

6.1 Achievements

The project successfully delivered an AI-driven learning platform with accurate question generation and adaptive learning capabilities. User feedback indicates high satisfaction and improved understanding of database concepts.

6.2 Challenges

- Optimizing AI integration to reduce response time.
- Ensuring compatibility with multiple database formats.

Solutions included using efficient libraries and refining schema validation algorithms.

7 Conclusion and Future Work

7.1 Conclusion

The DB Learn Space project demonstrates the potential of AI in enhancing database learning. The platform provides a comprehensive learning experience, addressing the challenges faced by students and educators.

7.2 Future Work

- Extend support to NoSQL databases like MongoDB.
- Add collaborative features for group learning.
- Enhance the UI for a more intuitive user experience.

References

- *LLaMA: Large Language Model Meta AI Documentation*
- *SQLAlchemy Documentation*
- Pandas Python Library Documentation