Jorge Argueta

Southern New Hampshire University

CS-499 Computer Science Capstone 24EW4

Professor Brooke Goggin

06 April 2024

5-2 Enhancement Three: Databases

This project, developed as part of the course DAD 220: Introduction to Structured Database Environments, involved integrating advanced MySQL features and data mining techniques, specifically K-means clustering. The primary components included setting up and connecting to a MySQL database, implementing complex database functionalities such as stored procedures and triggers, applying the K-means algorithm for data analysis, and visualizing the results.

**Justification for Inclusion in the ePortfolio**

This artifact is included in my ePortfolio to showcase my proficiency in several key areas:

* *Database Management:* Demonstrated by the setup and optimization of a MySQL database, including advanced features like stored procedures and triggers.
* *Data Analysis and Mining:* Illustrated through the application of K-means clustering to analyze employee salary data to uncover patterns.
* *Software Engineering Principles:* Highlighted by my ability to design, implement, and test software components rigorously.

Enhancements to the project included refining the database schema, optimizing query performance with indexes, and extending data analysis techniques, which improved both efficiency and analytical depth.

**Reflection on Course Objectives and Outcome-Coverage Plans**

The primary course objective met through this enhancement was the application of advanced database management and data mining techniques to address real-world problems, demonstrating both theoretical and practical understanding.

**Course Outcomes and Skills Demonstrated**

Employ strategies for building collaborative environments: I facilitated collaboration by using version control systems and regular team meetings, which helped align our project with organizational decision-making needs.

* *Design, develop, and deliver professional-quality communications*: I ensured that all project documentation was clear, technically sound, and tailored to our stakeholders, which included both technical and non-technical audiences.
* *Design and evaluate computing solutions using algorithmic principles:* The use of stored procedures and triggers, along with the K-means clustering algorithm, exemplifies designing solutions that adhere to computer science practices. I also managed trade-offs such as computation time versus accuracy.
* *Demonstrate use of innovative techniques in computing practices*: By optimizing the database schema and applying indexing strategies, I employed innovative and effective techniques that improved performance and scalability.
* *Develop a security mindset*: I anticipated potential vulnerabilities by implementing robust access controls and data validation within the database to enhance security and protect privacy.

**Learning and Challenges**

* *Advanced MySQL Features:* I gained a deeper understanding of encapsulating business logic within the database, improving performance, and reducing application-side load.
* *Data Mining Techniques:* K-means clustering offered practical insights into segmentation and pattern recognition, crucial for data-driven decision-making.

**Challenges:**

* *Database Optimization:* Designing an efficient schema that handled large data volumes while maintaining quick response times was challenging. Indexing strategies were crucial here.
* *Data Mining Application:* Initially simplistic, applying K-means directly to salary data required refinement for meaningful cluster analysis.

**Conclusion**

This enhancement plan significantly bolstered my technical skills, problem-solving abilities, and understanding of database and data mining applications in business contexts. The educational process of enhancing this project demonstrated my competencies in a real-world scenario, making it a valuable addition to my ePortfolio.