PROJECT DELIVERABLE 05

Group 13

Individual Contribution			
CWID	**Name**	**Contribution (Description)**	**Percent Contribution**
A20593079	Akshada Ranalkar	Worked on the programming part, prepared the video demo.	33.3%
A20563287	Anuja Wavdhane	Worked on the programming part, prepared the video demo.	33.3%
A20560966	Suhasi Gadge	Worked on the programming part, created README file, and documentation.	33.3%

Project Title:

ACADEMIC PUBLICATION DATABASE SYSTEM

(REFERENCE WEBSITE: IEEE Website

https://ieeexplore.ieee.org/abstract/document/344065)

1.1 Objectives of Deliverable # 05

1. Implement advanced SQL queries, including set operations, subqueries, and OLAP features, with meaningful results.
2. Develop a user-friendly interface with interactive elements and an intuitive layout.
3. Record a Loom video showcasing the application’s functionality and explaining complex queries.
4. Add robust error handling with clear messages, handling edge cases, and ensuring data security.
5. Write clean, well-documented code with a detailed README for setup and usage.
6. Organize all files in a zipped folder and include a contribution breakdown for group members.

1.2 Application Overview

1. **Programming Language:** Python

2. **Database:** MySQL

3. **Application Type:** Web-based application using Streamlit for the frontend.

4. **CRUD Operations:**
   - Full implementation of Create, Read, Update, and Delete operations across tables in the IEEE database.
   - Manage records for Publications, Authors, Events, and other entities efficiently through the interface.

5. **Libraries Used:**
   - Streamlit: Provides a simple and interactive web-based frontend.
   - MySQL Connector: Handles seamless interaction with the MySQL database.
   - Pandas: Used for data processing, analysis, and displaying tabular results.
   - Matplotlib/Plotly (if applicable): For creating visual analytics and charts within the dashboard.

6. **Key Features:**
   - Interactive Dashboards: Includes visual analytics to provide insights into IEEE data, such as trends in publications, membership growth, or event participation.
   - Dynamic Search and Filters: Enables users to search and filter records dynamically for ease of use.
   - Error Handling: Provides user-friendly error messages and handles edge cases effectively to maintain a smooth user experience.
   - Responsive Interface: The layout is designed to be responsive and accessible, ensuring compatibility across devices.
   - Data Import and Export: Allows users to upload or download data in standard formats (e.g., CSV).

7. **Purpose and Goals:**
   - Streamline IEEE data management processes.
   - Provide an accessible interface for nontechnical users to interact with the database.
   - Support decision making with analytics and trends derived from stored data.

8. **Future Scalability:**

   The application is built with scalability in mind, allowing for additional tables, relationships, or new features to be added with minimal changes to the existing codebase. Features like authentication and role based access can be implemented for enhanced security.

1.3 CRUD Operations

We have made sure to perform all the CRUD operations for every table present in the database. The video demonstration of these operations has been submitted along with this report.

1.4 Advanced Queries

1. **Set Operations:**
   - Used `UNION`, `INTERSECT`, and `EXCEPT` to combine or differentiate datasets, such as identifying authors who have contributed to multiple publications across different conferences.

2. **Set Membership:**
   - Applied `IN` and `NOT IN` clauses to filter data efficiently, such as listing members who have not participated in any events.

3. **Set Comparison:**
   - Used comparison techniques like `ALL` and `ANY` to find records, such as authors with citation counts exceeding all other authors in their research field.

4. **Subqueries with the `WITH` Clause:**
   - Implemented common table expressions (CTEs) for modular and reusable queries, such as calculating the average citation count for each conference and using it in further queries.

5. **Aggregate Functions:**
   - Utilized advanced aggregate functions like `GROUP BY`, `HAVING`, and window functions (e.g., `ROW_NUMBER`, `RANK`) to analyze data trends, such as ranking members by their number of contributions.

6. **OLAP Queries:**
   - Implemented Online Analytical Processing (OLAP) queries, including:
   - ROLLUP: To calculate subtotal and total publications by topic and year.
   - CUBE: To provide multidimensional analysis of participation across events and locations.

7. **Nested Subqueries:**
   - Designed multilevel subqueries to retrieve complex datasets, such as finding the most cited publication in each topic category.

8. **Join Queries:**
   - Executed advanced join operations (e.g., selfjoins, outer joins) to correlate related data, such as linking authors to their events and publications simultaneously.

9. **Analytics and Insights:**

- Queries to derive insights, such as:
- Top 5 most active members based on event participation.
- Conferences with the highest average number of publications.
- Trends in publication counts over the last five years.

10. **Performance Optimization:**
    - Used indexing and query optimizations to enhance performance for large datasets, ensuring efficient execution of complex queries.

1.5 Features Details

1. **User Interface:**
    - Includes a navigation guide and tablebased layout for seamless transitions between tables.
    - Interactive buttons and dropdowns make operations intuitive and userfriendly.
    - Color-coded notifications for successful operations (e.g., record added, updated, or deleted).

2. **Data Validation:**
    - Implements fieldspecific validation for numeric IDs, dates, and required text fields.
    - Prevents SQL injection and invalid data entry through sanitized inputs and strict type checks.

3. **Error Handling:**
    - Handles errors such as connection issues, duplicate records, or missing fields with detailed and helpful messages.
    - Logs errors for debugging and troubleshooting while maintaining userfriendly responses.

4. **Data Security:**
    - Securely manages database credentials using environment variables or configuration files.
    - Implements restricted access to critical operations by separating admin functionalities (if applicable).

5. **Data Visualization:**
    - Provides charts and graphs (e.g., bar graphs, pie charts) to visualize data trends like publication counts, membership growth, or event participation.
    - Interactive visualizations allow users to filter and analyze data dynamically.

6. **Data Export and Import:**
    - Supports exporting table data to CSV format for offline analysis.

- Allows batch importing of records from CSV files, speeding up data entry processes.

7. **Search and Filters:**
   - Provides advanced search functionality across tables with dynamic filtering based on multiple fields.
   - Users can apply conditions like date ranges, numeric limits, or keyword searches.

8. **Responsiveness and Accessibility:**
   - The interface is designed to work across different screen sizes, ensuring accessibility on mobile, tablet, and desktop devices.
   - Includes tooltips and help sections for easier navigation.

9. **Real-Time Updates:**
   - Displays real time changes in tables without requiring manual refresh, ensuring data consistency and user convenience.

10. **Customizable Views:**
    - Users can customize displayed columns or sorting preferences for a more personalized experience.