ALLIANCE UNIVERSITY

Private University established in Karnataka State by Act No.34 of year 2010
Recognized by the University Grants Commission (UGC), New Delhi

Generative AI Project Idea Proposal Documentation

Bachelor of Computer Applications

SEMESTER – IV

Generative AI & Agentic AI Winter Program

Submitted By:

Anuska Ghosh - 2411021240012

Likitha S - 2411021240067

Faculty Signature

HOD Signature

Department of Computer Applications

Alliance University

Chandapura – Anekal Main Road, Anekal

Bengaluru – 562 106

December 2025

Project Idea Proposal Documentation

Project Title

InSighto – An AI-Powered Data Analysis and Reporting System

1. Introduction

In the current data-driven environment, organizations and individuals frequently work with structured datasets but lack the technical expertise required to analyze them effectively. Traditional data analysis tools often require programming knowledge, statistical understanding, or expensive software subscriptions. Additionally, cloud-based AI solutions raise concerns regarding data privacy and cost.

InSighto is a web-based data analysis and reporting system designed to simplify exploratory data analysis by combining automated dataset profiling with locally hosted Generative AI. The system enables users to upload datasets, preview data structure, generate AI-assisted insights, and produce structured reports through an intuitive interface, without relying on external cloud services.

2. Problem Statement

- Data analysis tools often require advanced technical skills
- Manual exploratory analysis is time-consuming and error-prone
- Cloud-based AI solutions pose privacy and cost concerns
- Non-technical users struggle to interpret raw statistical outputs
- Many tools lack simple report-generation capabilities

3. Proposed Solution

The proposed system, InSighto, provides a user-friendly web application that automates key stages of exploratory data analysis. It allows users to upload CSV or Excel datasets, automatically profiles the data, generates AI-assisted insights using a locally hosted language model, and produces a structured analysis report that can be exported as a PDF.

The system emphasizes:

- Simplicity of use
- Data privacy through local processing
- Cost-effectiveness using open-source technologies

4. Objectives

- To design an intuitive web-based data analysis tool
- To automate dataset profiling (rows, columns, missing values)
- To integrate Generative AI for explaining dataset characteristics

- To generate structured analysis reports
- To enable report export in PDF format
- To ensure privacy by performing all processing locally

5. Scope of the Project

Included Scope

- Upload and validation of CSV and Excel datasets
- Dataset preview and metadata display
- Automated dataset profiling
- AI-generated insights and executive summaries
- Structured report generation
- PDF export of reports
- Session-based data handling

Excluded Scope

- Advanced statistical modeling
- Machine learning prediction models
- Real-time collaborative analysis
- User authentication and authorization
- Cloud-based data storage

6. System Architecture Overview

The system follows a modular client–server architecture:

- **Frontend:**
  HTML, CSS, and JavaScript with server-side rendering using Flask templates
- **Backend:**
  Python Flask application managing routing, analysis workflow, and storage
- **Database:**
  SQLite for temporary session-based storage
- **AI Component:**
  Ollama with llama3:8b for local Generative AI inference

7. Working Methodology

1. User uploads a dataset

2. Backend validates and stores the file with a unique session ID
3. Dataset preview is displayed to the user
4. An analysis agent performs dataset profiling
5. AI model generates insights based on profiling results
6. A structured analysis report is generated
7. The report is displayed and can be exported as a PDF

8. Technologies Used

- **Programming Language:** Python
- **Web Framework:** Flask
- **Frontend:** HTML, CSS, JavaScript
- **Database:** SQLite
- **Data Processing:** Pandas
- **Generative AI:** Ollama with llama3:8b
- **Report Export:** PDF (print-based export)

9. Data Privacy and Security Considerations

- All data processing occurs locally
- No dataset is transmitted to external cloud services
- Session-based isolation ensures user data separation
- Temporary storage minimizes long-term data retention

10. Cost Effectiveness

The project is designed to be highly cost-effective:

- Uses open-source tools and frameworks
- No paid API usage
- Local AI inference eliminates recurring costs
- Suitable for students, educators, and small organizations

11. Expected Outcomes

- A fully functional web-based data analysis tool
- Automated dataset profiling and AI-generated insights

- Structured, readable analysis reports
- Ability to export reports as PDF
- Improved accessibility to data analysis for non-technical users

12. Real-World Applications

- Academic learning and teaching of data analysis
- Small business data exploration
- Research dataset exploration
- Non-technical stakeholder reporting
- Privacy-sensitive data analysis environments

13. Conclusion

The updated project proposal for InSighto reflects a practical, privacy-conscious, and cost-effective approach to AI-assisted data analysis. By combining automated dataset profiling with locally hosted Generative AI and structured report generation, the system provides meaningful insights without compromising data privacy or incurring high costs. The project establishes a strong foundation for future enhancements and real-world deployment.

14. Source Code

GitHub Repository Links:

Anuska Ghosh – https://github.com/anuskaGHS/GenAI_DataAnalysis_Assistant

Likitha S - https://github.com/LikithaSrinivas100/GenAI_DataAnalysis_Assistant