Asklytics

An LLM for Website Analytics with Sentiment Analysis of Google Analytics 4 (GA-4) & Google Maps Reviews Data

Midterm Video Demonstration:

https://drive.google.com/file/d/1zQoKTNKXytSp2X_AsikXcsxxkwDFTrAR/view

Applied Research Project - COMP 4495

Section: 001

Team Lead

Tamoor Haider Aslam (300367290)

Team Members

Mohamed Nuskhan Mohamed Niyas (300368621)

Date of Submission:

24.02.2025

Table of Contents

INTRODUCTION **1**

Domain *1*
Background & Context *1*
Problems to Solve *1*
Existing Tools *1*
Hypothesis, Assumptions & Benefits *3*

PROPOSED RESEARCH PROJECT **3**

Research Design and Objectives: ..... ***Error! Bookmark not defined.***
Methodology and Justification: *4*
Data Collection: ..... ***Error! Bookmark not defined.***
Technologies: ..... ***Error! Bookmark not defined.***
Expected Results: ..... ***Error! Bookmark not defined.***

RIIPEN EXTERNAL PARTNERS OR AFFILIATES ..... ERROR! BOOKMARK NOT DEFINED.

PROJECT PLANNING AND TIMELINE **4**

PROJECT CONTRACT ..... ERROR! BOOKMARK NOT DEFINED.

WORK DATE/HOURS LOGS **9**

CLOSING AND REFERENCES **9**

Introduction

Domain

In the modern era of data-driven decision-making, organizations increasingly rely on web analytics to understand user behavior, optimize website performance, and enhance customer satisfaction. Google Analytics 4 (GA4) serves as a powerful tool for gathering web traffic data, offering insights into metrics such as bounce rates, user sessions, and engagement patterns. However, despite its utility, interpreting this data can be overwhelming for non-technical users, limiting its accessibility and practical application.

Background & Context

This project addresses the challenge of making website analytics user-friendly by developing an AI-powered chatbot capable of answering natural language queries. Beyond the standard analytical insights, the chatbot will also integrate sentiment analysis capabilities to classify customer feedback into positive, neutral, or negative sentiments. This dual approach not only simplifies data interpretation but also provides actionable insights for businesses to improve their offerings and customer engagement strategies.

Problems to Solve

The key problems this project aims to solve include:

1. The steep learning curve associated with understanding and using GA4 data.
2. The absence of conversational interfaces in current analytics tools.
3. The lack of integrated sentiment analysis in interpreting user feedback along with the real time impact on user interest in business's website.

Existing Tools

Existing research and tools focus heavily on data visualization and automated reporting but rarely incorporate natural language interfaces tailored to specific analytics questions. By leveraging Natural Language Processing (NLP) models, sentiment analysis techniques, and

predictive analytics, this project fills a critical gap in website performance analysis and feedback interpretation.

Literature Review

1. This paper highlights how integrating AI and data analytics is revolutionizing competitive intelligence, enabling businesses to gather, analyze, and act on market insights more effectively. AI-driven technologies like machine learning, natural language processing, and predictive analytics allow organizations to uncover patterns, predict trends, and automate tasks like competitor monitoring and sentiment analysis. Real-time analytics and intuitive visualization tools further enhance decision-making and resource allocation, helping businesses adapt dynamically to changing market conditions. While the adoption of these technologies presents challenges such as data privacy, integration issues, and skill gaps, the paper provides strategic recommendations to address them. Case studies from various industries demonstrate the benefits of AI-powered competitive intelligence, including driving innovation, improving customer engagement, and optimizing strategies

   Jack, Harper. (2024). Leveraging AI and Data Analytics: Revolutionizing Competitive Intelligence for Market Insights. 10.13140/RG.2.2.25148.76166.

2. The rise of big data and advanced analytics has transformed decision-making, enabling organizations to shift from reactive to predictive strategies. Predictive analytics uses machine learning, AI, and statistical models to analyze historical and real-time data, uncover patterns, forecast trends, and optimize decisions. It is widely applied in industries like healthcare, finance, manufacturing, and retail, enhancing outcomes such as improved patient care, fraud prevention, and predictive maintenance. However, challenges such as data quality, integration complexities, and ethical concerns persist. This paper explores the transformative role, applications, challenges, and prospects of predictive analytics, emphasizing its

importance in fostering resilience, sustainability, and competitive advantage in a data-driven world.

Nyoni, Rumbidzai. (2025). Harnessing Data Analytics for Predictive Insights: Advancing Decision- Making with Big Data Innovations. International Journal of Research Publication and Reviews. 6. 2915-2936.

3. Financial institutions are increasingly using data lakes to manage vast and diverse datasets, enabling advanced fraud detection and prevention. These repositories store structured, semi-structured, and unstructured data, facilitating insights into large-scale financial information. By integrating advanced analytics techniques such as machine learning, anomaly detection, predictive modeling, and natural language processing (NLP), data lakes empower institutions to identify patterns, predict fraudulent behavior in real time, and take proactive measures. This paper examines the architecture and technologies behind financial data lakes, their practical applications in fraud prevention, and the challenges of implementation. It also provides a roadmap for successful adoption, emphasizing their role in safeguarding assets and maintaining customer trust.

   Arena, Frank & Paul, Justin. (2024). INNOVATIVE ANALYTICS TECHNIQUES FOR FINANCIAL DATA LAKES.

Hypothesis, Assumptions & Benefits

Our hypothesis is that an intuitive, AI-driven chatbot combined with sentiment analysis will empower users to derive meaningful insights without requiring technical expertise. This project has the potential to enhance decision-making, reduce the time spent on manual analysis, and improve customer-centric strategies for organizations.

Initial Proposed Research Project

The goal of our project is to create an interactive chatbot that simplifies website analytics and enhances customer feedback analysis. The specific objectives include:

1. Implementing a natural language interface to process and respond to queries about website performance metrics.

2. Visualizing Google Analytics 4 (GA4) data in an accessible and interactive format.
3. Classifying user feedback into sentiment categories (positive, neutral, and negative).
4. Providing predictive insights for future website performance and detecting anomalies in traffic patterns.

Changes to Proposal:

Feature: None
Technology Stack: None
Platform: None

Proposed Approach: We were planning to use BERT or R1-DeepSeek Distill Llama 8B but due to GPU constraints we are now using quantized version of Mistral 7B v0.1 from Mistral.AI

Justification: Both R1-DeepSeek & BERT required excessive amount of GPU Compute which made it difficult for local development on our computers at home & even college labs.

Timeline: We are 1 week behind the proposed timelines.

Justification: The model selection and finding ways to fine tune them took more than expected. We are still hopeful to complete the project on time with extra effort.

Project Planning and Timeline

Timeline and Milestones:

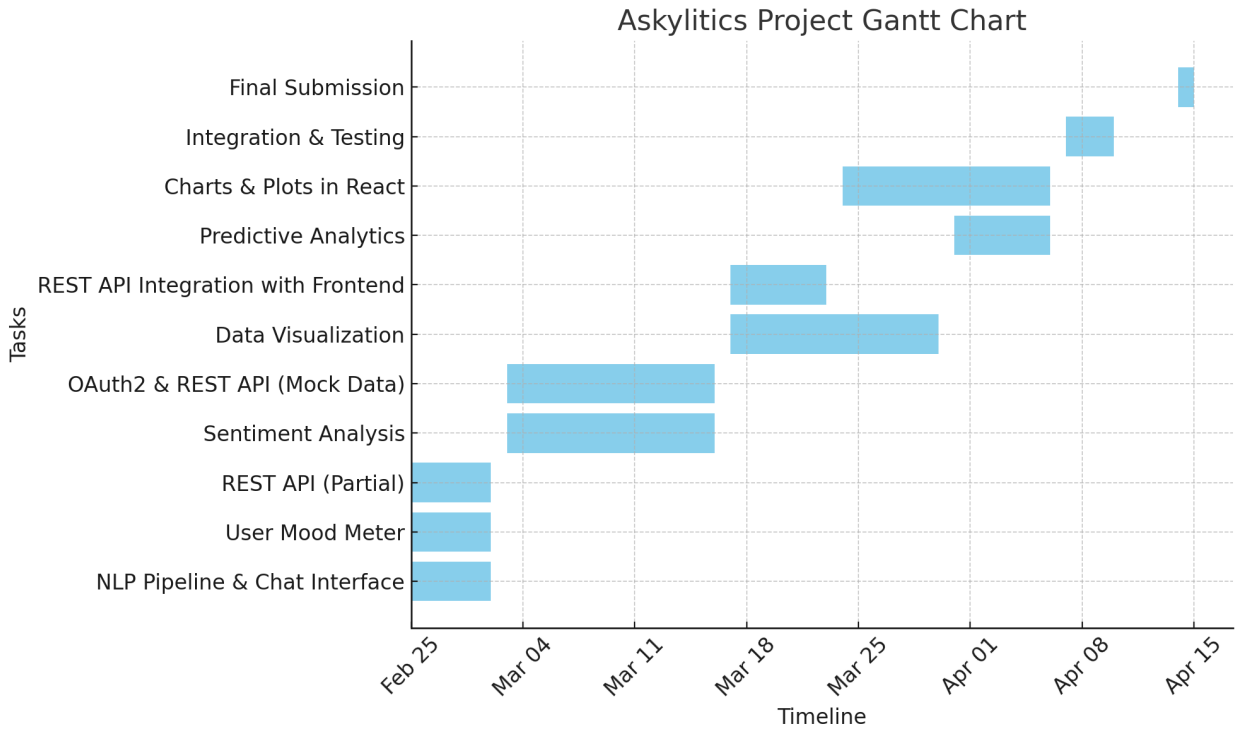
Phase	Milestone	Deliverables	Deadline
2	**Analytics:** • NLP Pipeline Implementation, **Development:** • Chat Interface Complete (React)	Fine-tuned Mistral2 model for natural language queries	Feb 28, 2025

Phase	Milestone	Deliverables	Deadline
	• User Mood Meter Complete (React) • REST API design (Partial)		
3	**Analytics:** • Sentiment Analysis **Development:** • OAuth2 flow (Complete) • Implement RESTAPI endpoints with mock data except OAuth2 (complete)	Sentiment classification module for user feedback	Mar 14, 2025
4	**Analytics:** • Data Visualization **Development:** • Integrating RESTAPI with Frontend (Complete). • Adding Plots/Charts generation in React (Partial)	Integrated Chart.js and Plotly components	Mar 28, 2025
5	**Analytics:** • Predictive Analytics **Development:** • Adding Plots/Charts generation in React (Complete)	Traffic forecasting and anomaly detection models	Apr 05, 2025
6	Integration & Testing	Fully functional system with end-to-end testing	Apr 10, 2025
7	Final Submission	Complete documentation and video demonstration	Apr 15, 2025

Team Responsibilities:

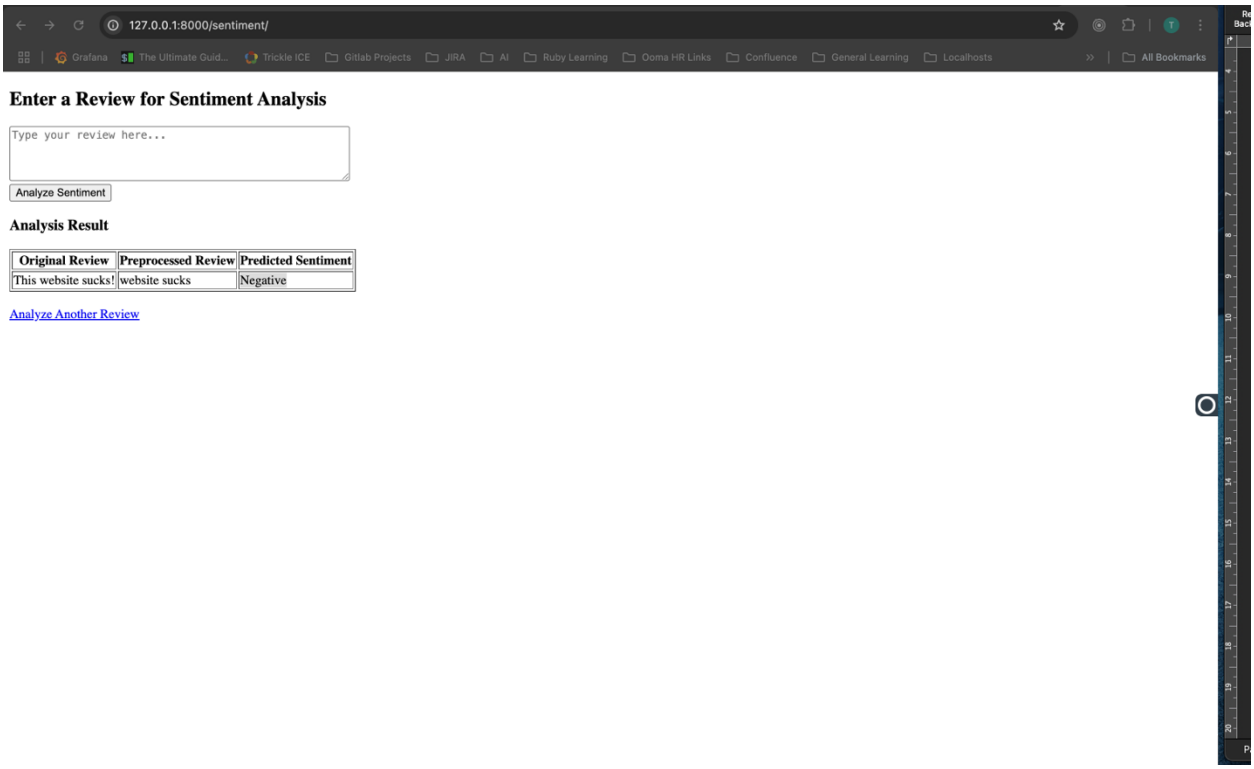
- **Mohamed Nuskhan Niyas:** Responsible for leading the LLM training, testing, improving & creating NLP pipeline implementation.
- **Tamoor Haider Aslam:** As a Teamlead, Tamoor will be responsible for project management, communication, giving direction, removing blockers etc. Also he will take part in LLM testing, frontend and backend architecture design and development.

Project Management Chart:

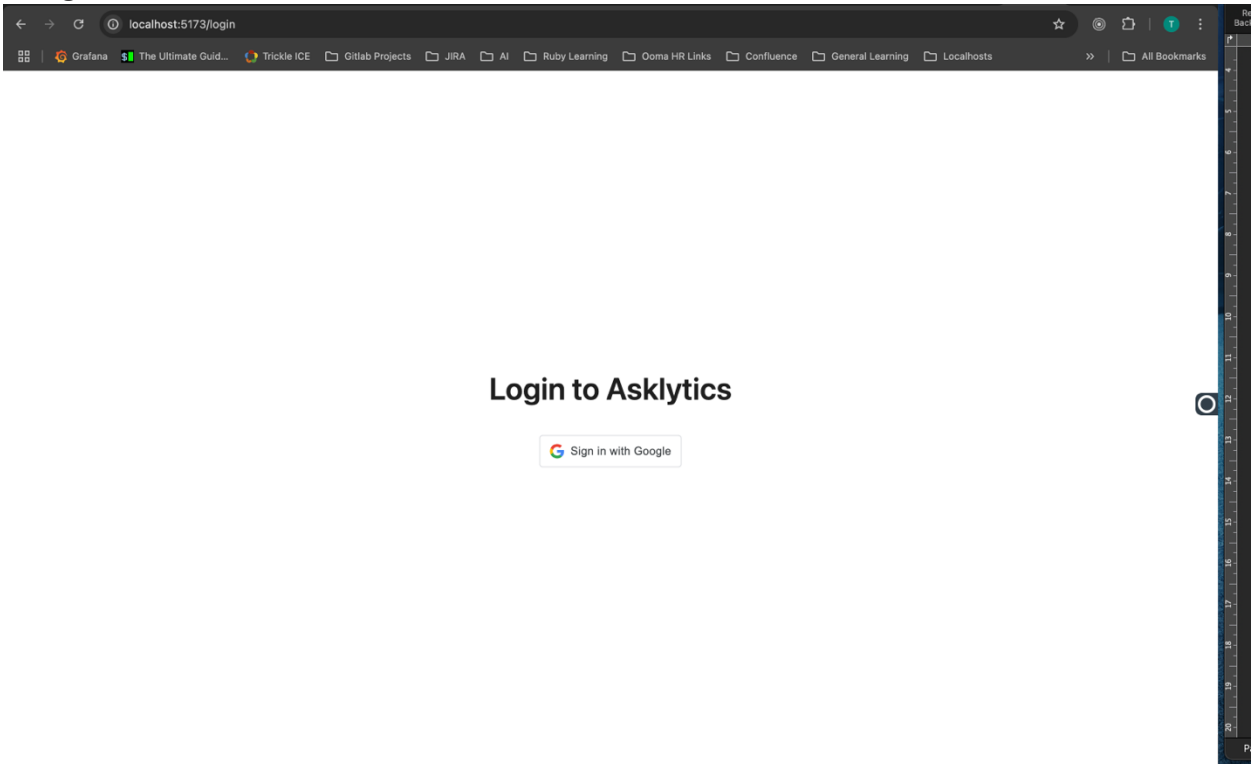


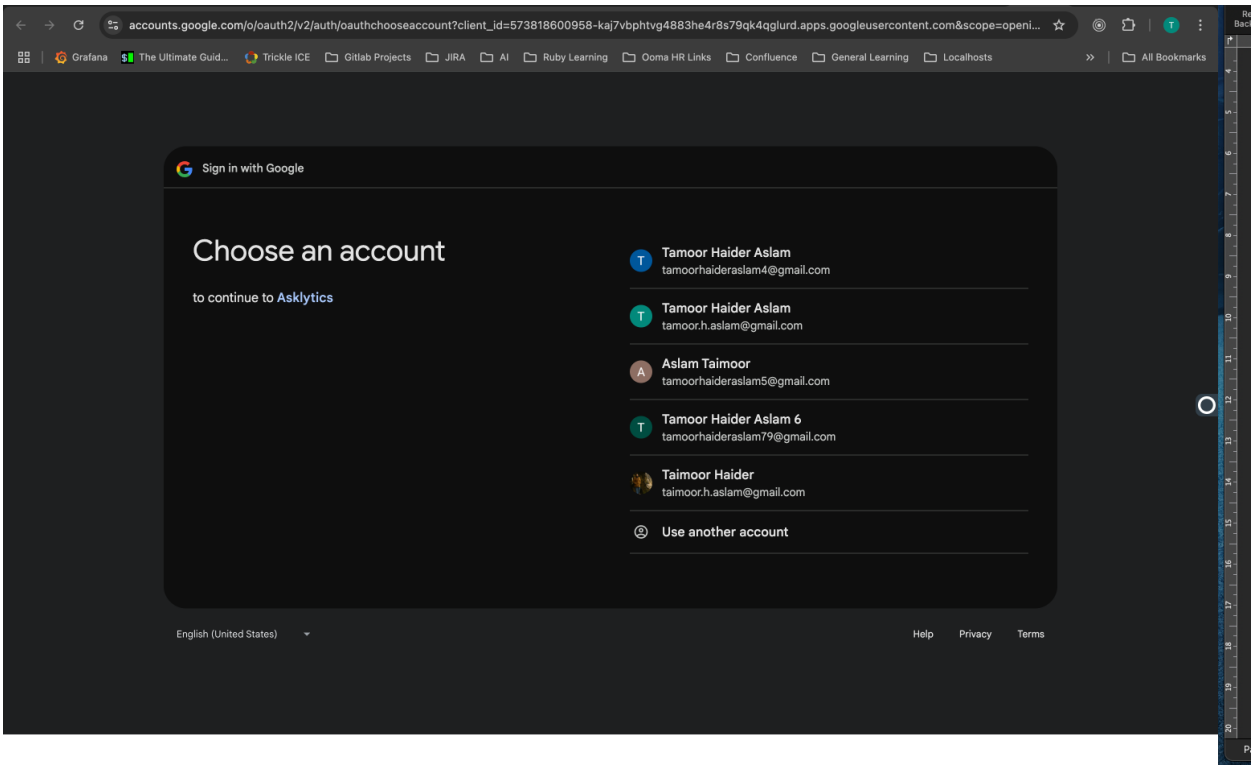
Features

Sentiment Analysis:

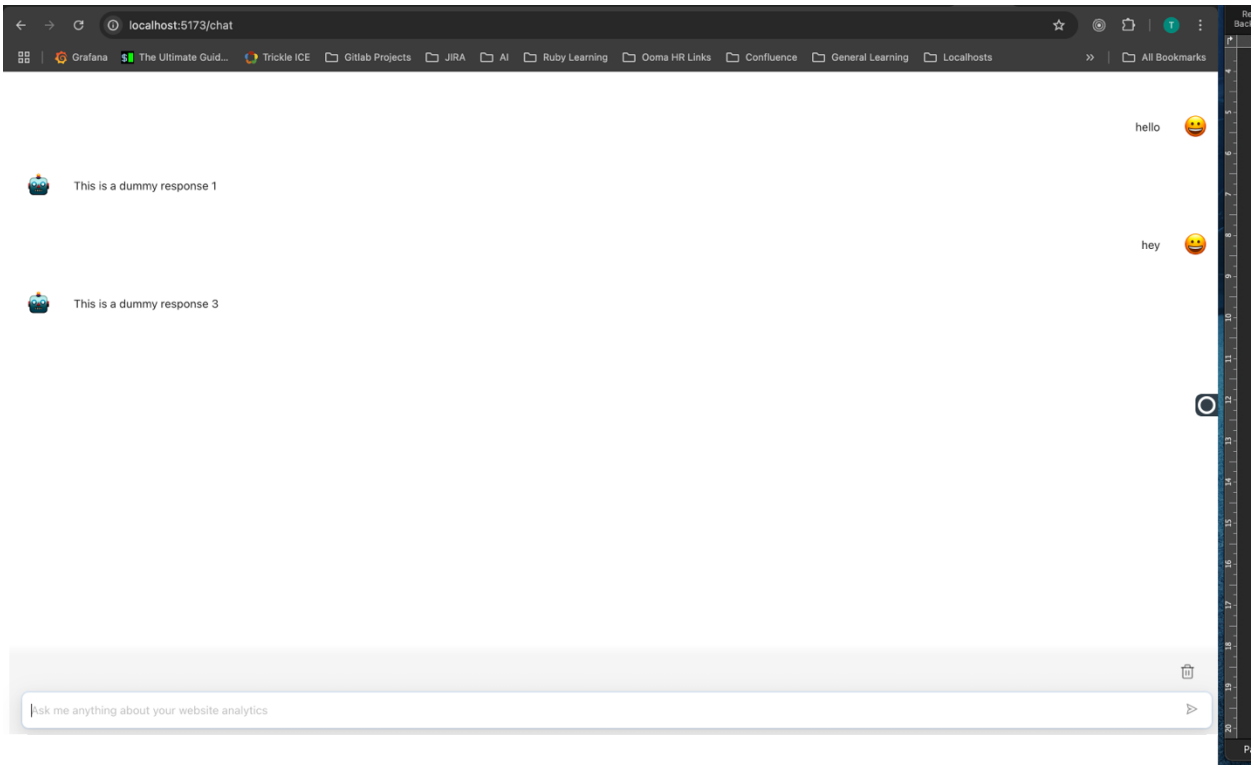


Google OAuth2

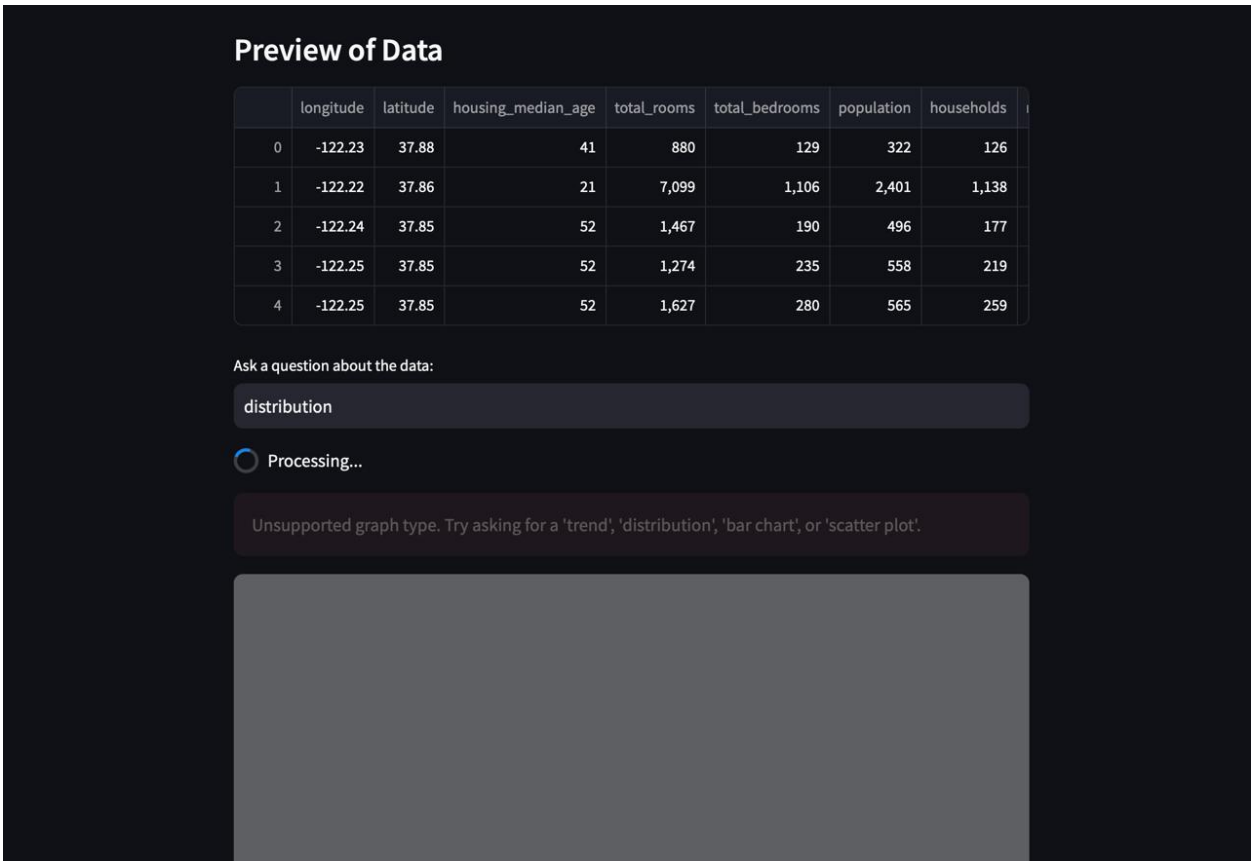




Chat UI



Mistral2 Fine Tuned:



Work Date/Hours Logs

Student Name	Date	Hours	Description of Work Done
Mohamed Nuskhan	Jan 23, 2025	2	Continued refining NLP pipeline for natural language queries.
Tamoor Aslam	Jan 26, 2025	2	Set up React & Django projects, worked on architecture.
Mohamed Nuskhan	Jan 28, 2025	3	Testing and improving LLM fine-tuning methods.
Tamoor Aslam	Jan 30, 2025	2	Designed chat UI wireframes, API interaction draft.

Student Name	Date	Hours	Description of Work Done
Mohamed Nuskhan	Feb 2, 2025	2.5	Implemented initial sentiment classification logic.
Tamoor Aslam	Feb 5, 2025	2.5	Integrated OAuth2 flow for authentication.
Mohamed Nuskhan	Feb 8, 2025	3	Analyzed GA4 data integration challenges, proposed solutions.
Tamoor Aslam	Feb 12, 2025	3	Developed REST API endpoints (Partial).
Mohamed Nuskhan	Feb 15, 2025	2	Fine-tuned Mistral2 model for analytics queries.
Tamoor Aslam	Feb 18, 2025	2.5	Connected frontend with REST API for chat interactions.
Mohamed Nuskhan	Feb 21, 2025	3	Worked on testing sentiment analysis module with sample data.
Tamoor Aslam	Feb 24, 2025	2	Reviewed and documented API workflow, prepared midterm report.

Closing and References

Acknowledgments:

Special thanks to Douglas College for providing access to resources, tools, and guidance essential for this project.

References:

1. Devlin, J., et al. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*.

2. Google Developers. (2025). *Google Analytics Data API Documentation*. Retrieved from https://developers.google.com/analytics/devguides/reporting/data/v1.
3. Hutto, C., & Gilbert, E. (2014). *VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text*.
4. Arena, F., & Paul, J. (2024). Innovative analytics techniques for financial data lakes.
5. Nyoni, R. (2025). Harnessing data analytics for predictive insights: Advancing decision-making with big data innovations. International Journal of Research Publication and Reviews, 6, 2915–2936.
6. Jack, H. (2024). Leveraging AI and data analytics: Revolutionizing competitive intelligence for market insights. https://doi.org/10.13140/RG.2.2.25148.76166