Project Design Phase-2
Technology Stack

Date	27 October 2023
Team ID	PNT2023TMID592731
Project Name	Walmart Sales Analysis For Retail Industry With Machine Learning
Maximum Marks	4 Marks
Team Members	Akshara Daram Dereddy Venkata Yogitha S. Chethan Manish

Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2

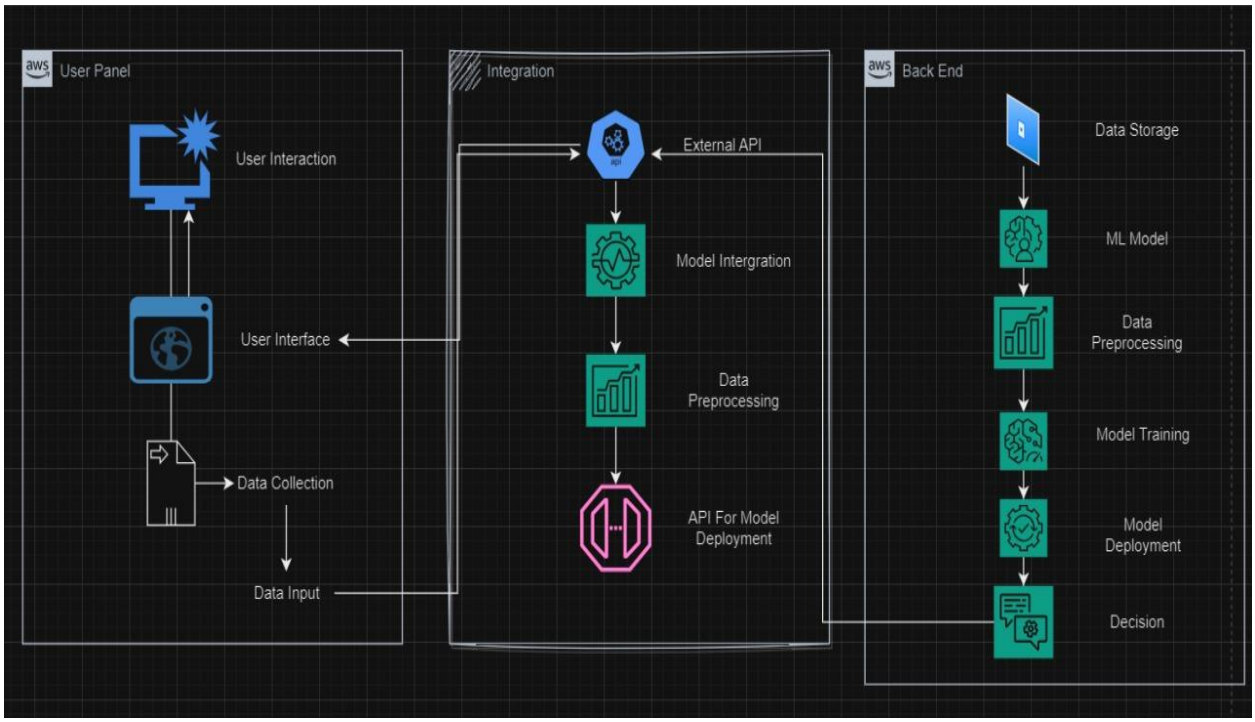

Table-1 : Components & Technologies:

Component	Description	Technologies
User Interaction	Interface for user interaction with the application along with creating an user friendly interface.	Web-based UI (HTML, CSS, JavaScript) or mobile app (React Native, Swift/Android).
Application Logic-1	Core logic responsible for handling user requests.	Python, Flask, FastAPI, or Node.js (Express).
Data Collection	Gathering raw sales data from various sources or mainly from the user about the sales of the Walmart Store.	Web scraping, ETL tools, data warehouses, and data pipelines (e.g., Apache Nifi).
Data Input	Handling and processing user-provided input data.	Forms, file uploads, APIs, or command-line input.
External API	Integration with external data sources or services.	RESTful APIs (e.g., Walmart's API, weather API).
Cloud Database	Storage and management of structured data.	Amazon RDS, Google Cloud SQL, or Azure SQL.
Model Integration	Interface for integrating machine learning models.	RESTful API endpoints (Flask, FastAPI,JSON,XML).
API Model Deployment	Responsible for deploying machine learning models as APIs, enabling real-time predictions and external interaction. It ensures model accessibility and scalability.	Docker, Kubernetes, or serverless (AWS Lambda).
Machine Learning Model	The predictive model for sales analysis.	Scikit-Learn, TensorFlow, PyTorch, XGBoost, RandomForest, Adaboost etc.
Data Preprocessing	Data preparation and feature engineering.	Pandas, NumPy, scikit-learn, or custom scripts.
Model Deployment	Hosting and serving the machine learning model.	Flask, FastAPI, TensorFlow Serving, or AWS Sagemaker.
Infrastructure (Server/Cloud)	Underlying cloud infrastructure and resources.	AWS, Google Cloud, Azure, or on-premises servers like Local, Cloud Foundry, Kubernetes etc.

Table-2: Application Characteristics:

Component	Description	Technologies
Open-Source Frameworks	Utilizing open-source frameworks for model development and deployment, ensuring cost-efficiency and flexibility which makes easy for the user to get the sales impact on holidays.	- Scikit-Learn, TensorFlow, PyTorch for model development. - Flask or FastAPI for API deployment. - Kubernetes for container orchestration. - Jupyter Notebook for model prototyping and development.
Security Implementations	Implementing robust security measures to protect sensitive data, model APIs, and ensure user data privacy which helps the walmart stores to protect its professional data from the hackers.	- OAuth 2.0 or JWT for user authentication. - Encryption (HTTPS/SSL) for data in transit. - Role-based access control. - Regular security audits and updates. - Compliance with industry standards (e.g., GDPR).
Scalable Architecture	Designing a scalable architecture that can handle growing data volumes and user demands which can manage the huge inflow of user demands assuming as a big data.	- Microservices architecture for modularity and scalability. - Containerization with Docker and orchestration with Kubernetes. - Load balancers for distributing traffic. - Auto-scaling based on resource usage.
Availability	Ensuring high availability and minimal downtime for the application to support continuous data analysis which helps in the growth of sales for walmart.	- Redundancy in database and API deployment. - Geographically distributed data centers or cloud regions. - Monitoring and alerting systems (e.g., Prometheus, Grafana). - Failover mechanisms for fault tolerance.
Performance	Optimizing application performance to provide quick insights and predictions.	- Caching mechanisms for frequently accessed data. - Model optimization (e.g., quantization) for faster inference. - Load testing and performance tuning.
User-Friendly Interface	Creating an intuitive and user-friendly interface for data input, visualization, and interaction.	- HTML, CSS, JavaScript for web-based UI. - React or similar frameworks for responsive design. - Data visualization libraries (e.g., D3.js). - User experience (UX) testing and design principles.
Interoperability and Accuracy	Ensuring seamless integration with external systems and maintaining high prediction accuracy.	- RESTful API design for interoperability. - Integration with external data sources (e.g., weather data). - Continuous model monitoring and retraining for accuracy improvement. - Data preprocessing techniques to enhance model accuracy.
Data Transparency	Enhancing transparency in Walmart sales analysis by making data sources, processing, and insights accessible and accountable. This promotes informed decision-making and trust.	Data documentation tools. - Metadata management systems. - Data catalog solutions. - Access control mechanisms. - Data lineage and provenance tracking. - Data visualization tools. - Data governance frameworks. - Compliance and auditing tools.