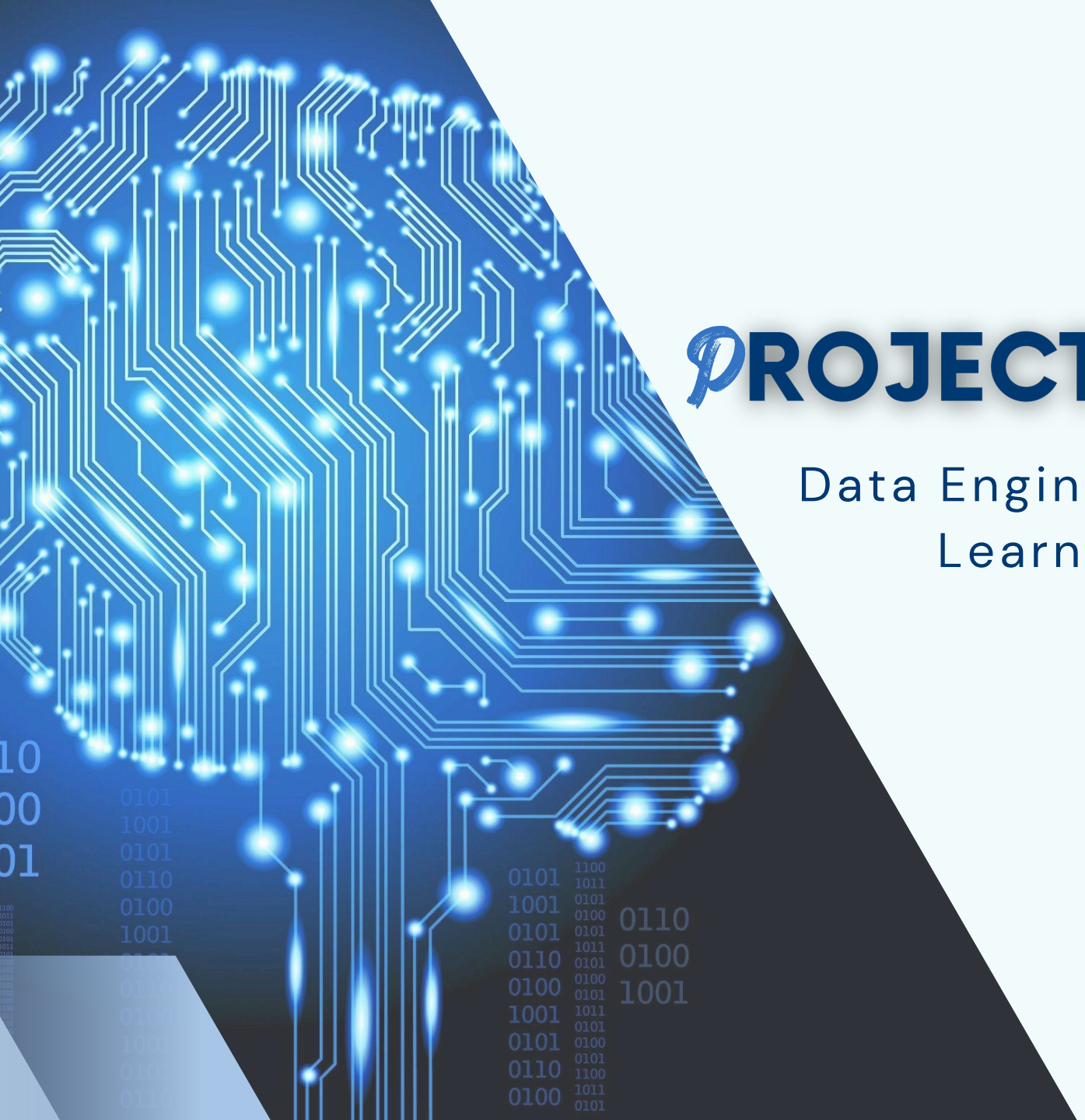
DEPI Technical Project

CUSTOMER DATA MANAGEMENT AND ANALYSIS

Project Proposal



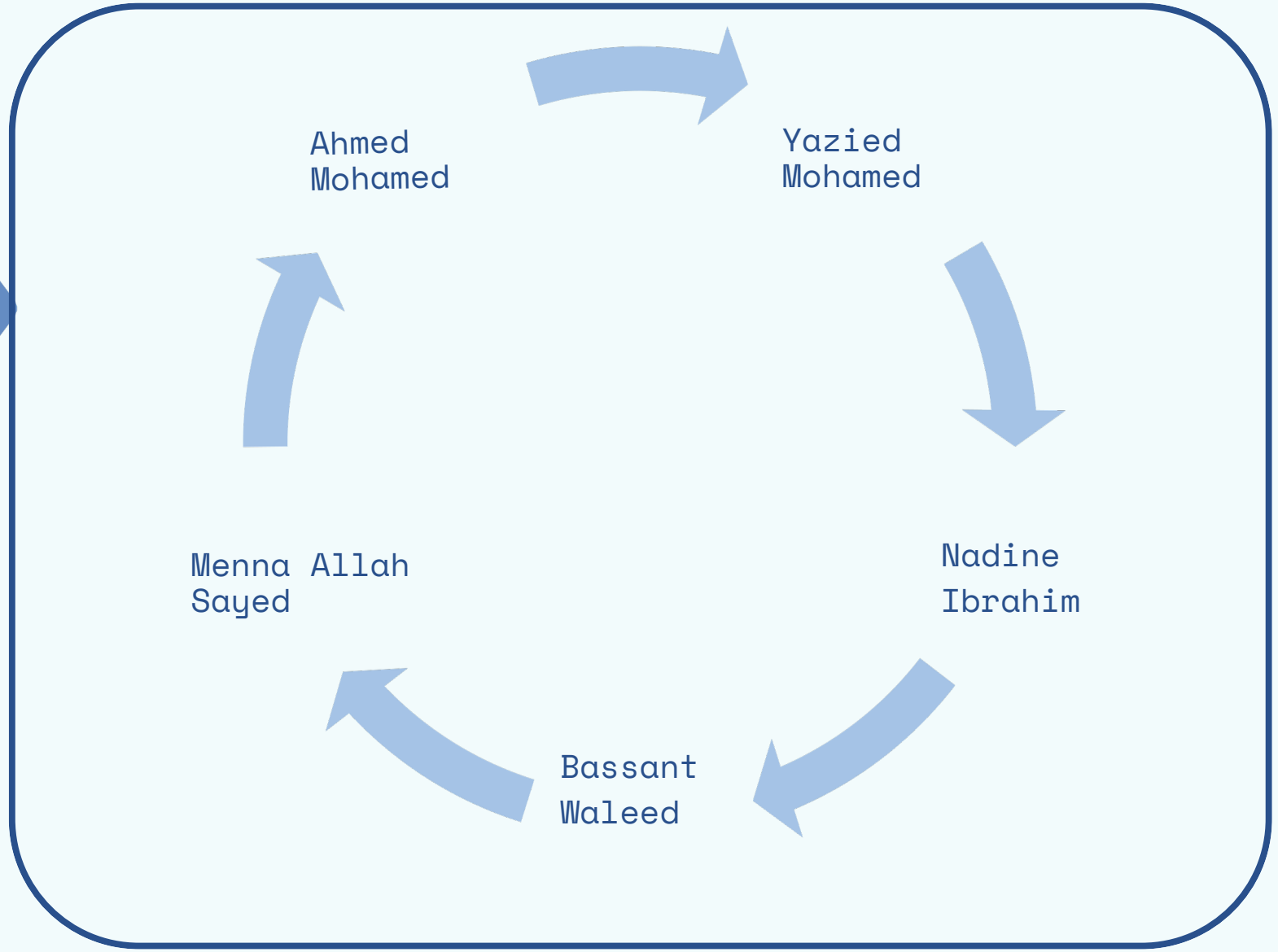
PROJECT TYPE

Data Engineering / Machine Learning Project

Project Purpose and Need

This project aims to design a SQL-based system for managing and analyzing customer data, integrating Python and Azure for data analysis and machine learning deployment. The solution will enable efficient data storage, predictive modeling, and insights generation, benefiting data-driven decision-making and customer engagement strategies.

Team Members



DATA SOURCES:

We generated sample data using the Faker library in Python.

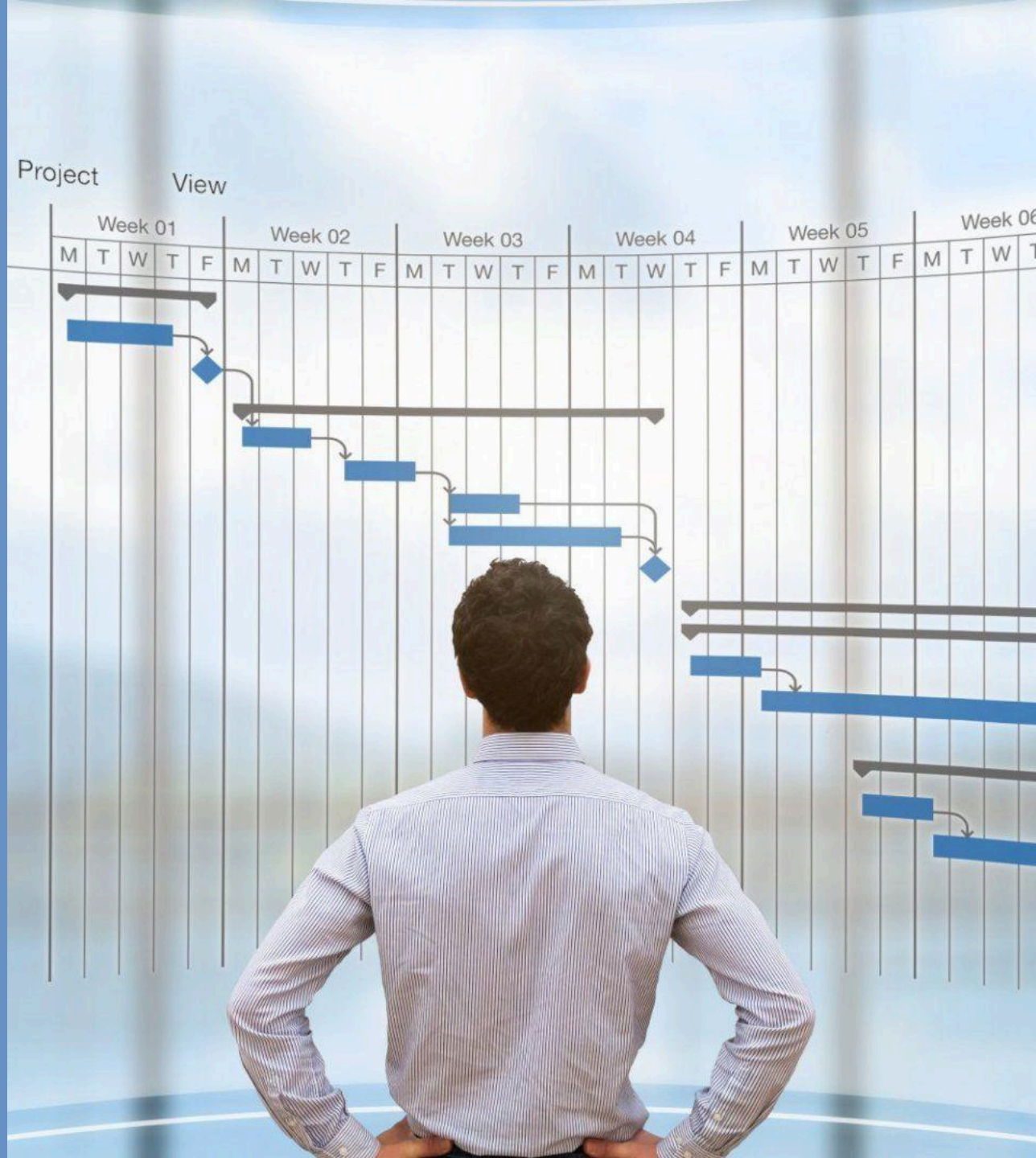
Tools and Technologies:

Cloud: Microsoft Azure (Data Studio, Synapse Analytics, Data Factory, Machine Learning)

- Python (Pandas, NumPy, Scikit-learn, SQLAlchemy)
- SQL (Microsoft SQL Server, SQL Data Warehouse)
- MLflow

- Visualization: Power BI
- Deploymeent: Streamlit(for web deployment)
- Machine Learning Integration Azure Machine Learning, Jupyter Notebooks

Project Plan and Timeline



Week 1: Data Management and SQL Database Setup

Database Design: Design a SQL database schema to manage customer data, including tables for customer information, transactions, and interactions.

Implementation: Create and populate the SQL database using Microsoft SQL Server.

SQL Queries: Write SQL queries to extract, update, and analyze customer data.

Deliverables: A well-designed SQL database schema and populated database. SQL queries for data extraction and basic analysis.

Week 2: Data Warehousing and Python Programming

Data Warehouse Implementation: Implement a SQL Data Warehouse to aggregate and manage large volumes of customer data for analytical purposes.

Data Integration: Load data from various sources into the data warehouse.

Python Programming: Develop Python scripts to interact with the SQL database, perform data extraction, and prepare data for analysis.

Deliverables: A functioning SQL Data Warehouse with integrated data. Python scripts for data extraction and preparation.

WEEK 3: DATA SCIENCE AND AZURE INTEGRATION

Data Science with Python: Perform data analysis and build predictive models (e.g., customer churn prediction) using Python.

Azure Data Fundamentals: Utilize Azure Data services to manage and analyze customer data.

Model Development: Develop and evaluate machine learning models using Azure Machine Learning or similar services.

Deliverables: Analysis report with insights and predictive models. Integrated Azure Data services setup and documentation.

Week 4: MLOps, Deployment, and Final Presentation

MLOps Implementation: Use MLflow to track experiments and manage machine learning models.

Model Deployment: Deploy the machine learning model using Azure services or create a web application for model predictions.

Final Report and Presentation: Prepare a comprehensive report covering data management, analysis, model development, and deployment. Create a presentation to showcase the project results.

Deliverables: Deployed machine learning model or web application. Final report and presentation.

Team Member Responsibilities

Nadine Ibrahim: SQL database design, setup, and writing SQL queries.

Bassant Waleed: Data warehouse implementation and Python scripting.

Menna Allah Sayed: Data analysis, predictive modeling, and Azure integration.

Ahmed Mohamed Yousef: MLOps and machine learning model deployment.

Yazied Mohamed: Data visualization, final report, and presentation.

Instructor Reporting:

Every 5 days, the team leader (....) will provide progress updates to the instructor, ensuring alignment with the project timeline and goals