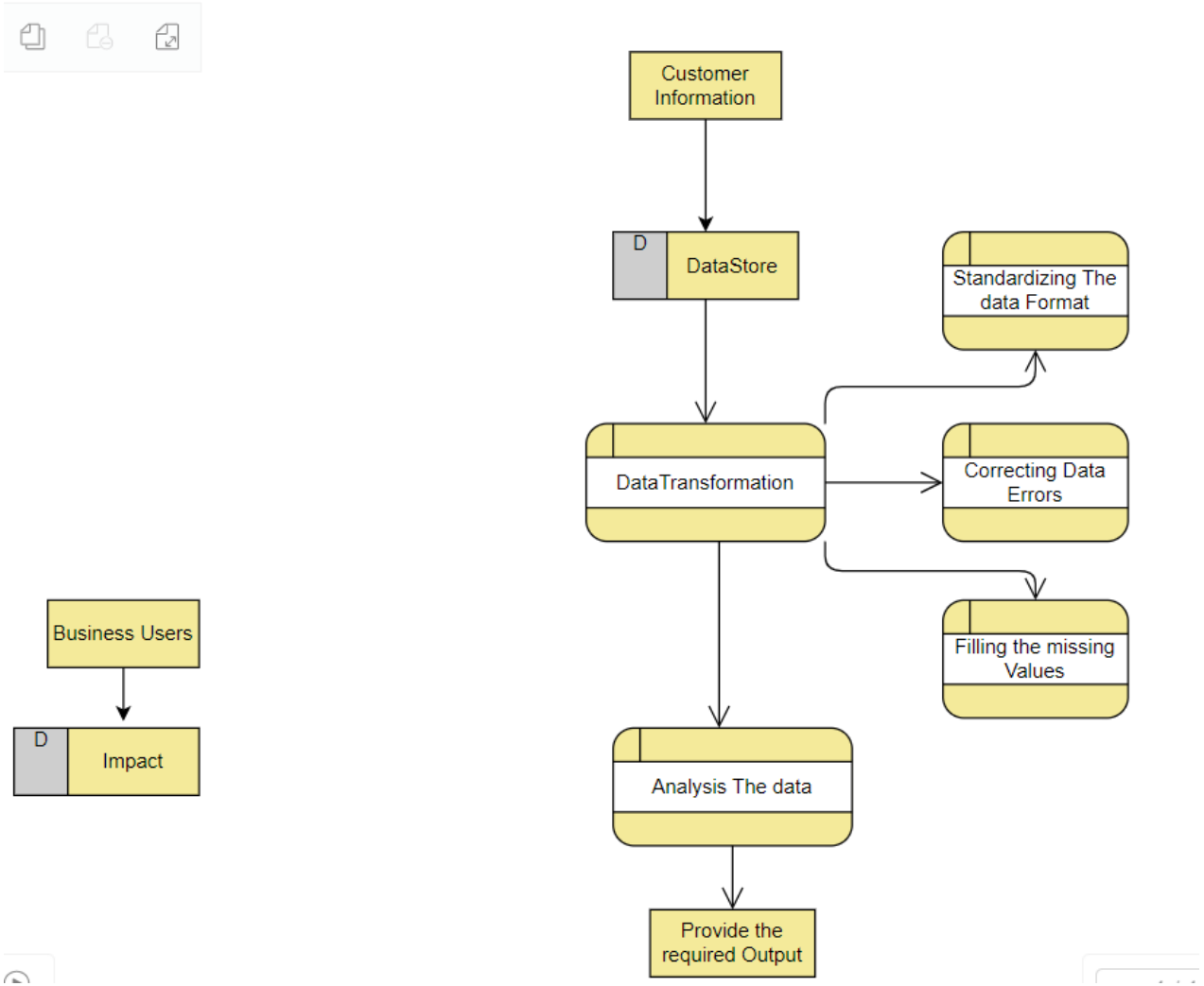
Project Design Phase-II
Data Flow Diagram & User Stories

Date	23 -10-2023
Team ID	PNT2022TMID592946
Project Name	Project – Travel Insurance Predication using Machine Learning
Maximum Marks	4 Marks

Data Flow Diagrams:

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

The data flow diagram (DFD) for the Travel Insurance analysis project illustrates a multi-level structure. At the top level (Level 0), the "Trave Insurance Analysis System" acts as the core, receiving data from "Travel Insurance Data," processing it, training machine learning models, and providing forecasts through a Flask web application. The machine learning models, at Level 1, interact with data stores for configuration and feature importance. The Flask web application, also at Level 1, communicates with users and passes their inputs for sales forecasting. Additionally, the system supports IBM Cloud deployment, allowing users to retrieve predictions. This DFD provides a concise visual representation of how data flows through various processes, entities, and data stores within the project, facilitating the understanding of the system's architecture and data pathways.


Customer Information
D
DataStore
Standardizing The data Format
DataTransformation
Correcting Data Errors
Filling the missing Values
Business Users
D
Impact
Analysis The data
Provide the required Output

User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Data Analyst	Data Preparation	USN-1	As a data Analyst, I will collect the history of travel insurance	The System should be able to store the data in structured format	High	Sprint-1
		USN-2	As a data analyst, I need to preprocess the collected data, including handling missing values and outliers.	The system should successfully clean and preprocess the data, resulting in a high-quality dataset for analysis	High	Sprint-1
	Sales Forecasting	USN-3	As a data analyst, I want to apply machine learning algorithms like Random Forest, Decision Tree, XgBoost, and ARIMA to forecast future sales	The system should train and test these algorithms, providing accurate sales forecasts.	High	Sprint-2
	Deployment and Integration	USN-4	As a data analyst, I want to integrate the analysis and forecasting models into a Flask web application.	The system should create a user-friendly web interface for stakeholders to access the analysis and forecasts.	Medium	Sprint-3
		USN-5	As a data analyst, I need to deploy the Flask application on IBM Cloud for easy access and scalability.	The system should deploy the Flask application on the IBM Cloud platform, ensuring it is accessible to authorized users	High	Sprint-4