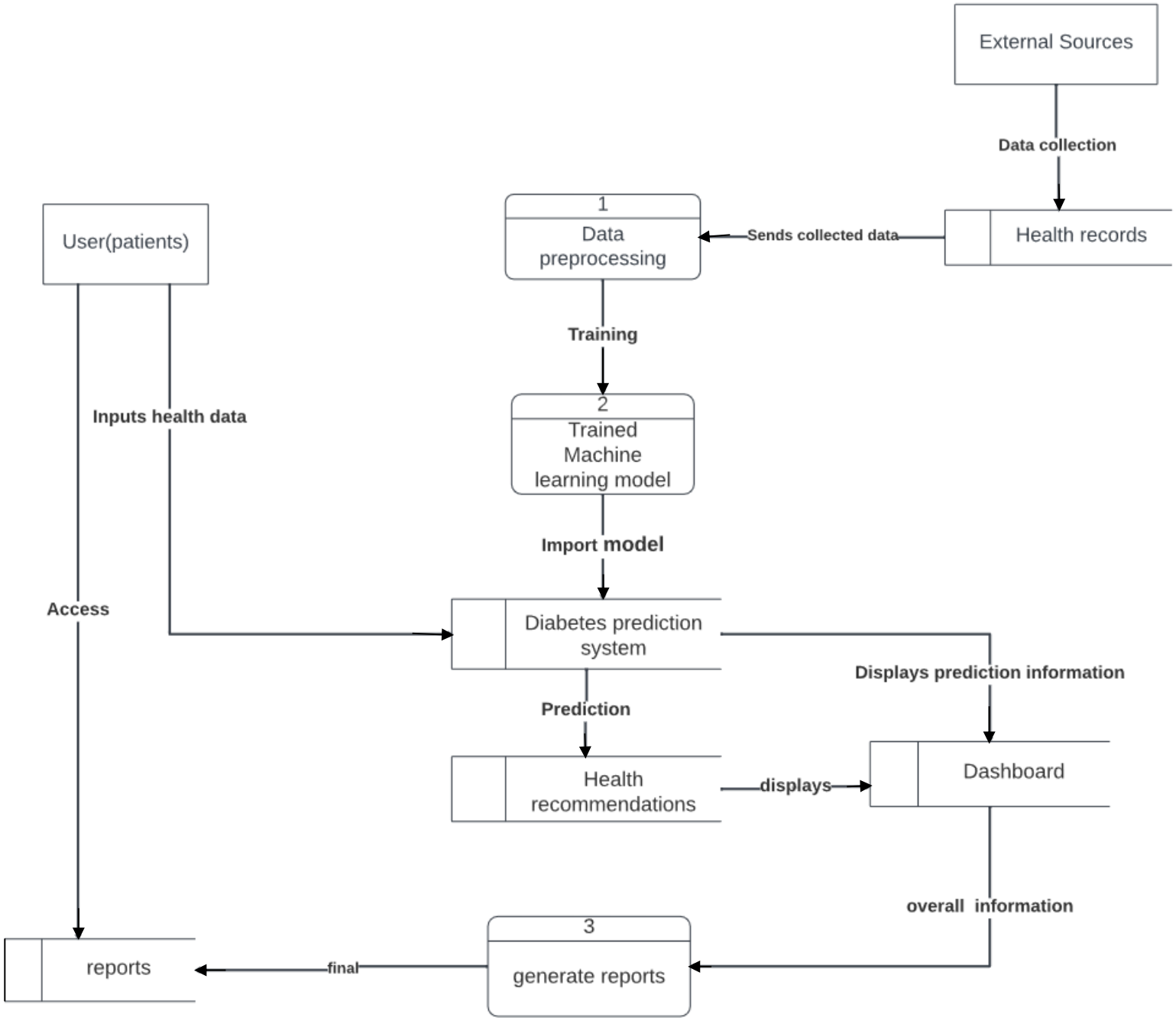
Project Design Phase-II
Data Flow Diagram &
User Stories

Date	**23 October 2023**
Team ID	**PNT2022TMID593014**
Project Name	**Diabetes Prediction Using Machine Learning**
Maximum Marks	**4 Marks**

Data Flow Diagram for Diabetes Prediction using ML:

Level 0 (Industry Standard)


External Sources
Data collection
Health records
Sends collected data
1
Data preprocessing
User(patients)
Training
2
Trained Machine learning model
Inputs health data
Import model
Access
Diabetes prediction system
Displays prediction information
Prediction
Health recommendations
displays
Dashboard
overall information
3
generate reports
final
reports

User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Healthcare Professionals	Predict Diabetes Onset	USN-1	As a healthcare professional, I want to input patient health records and relevant parameters to predict the onset of diabetes.	The system should accept input data including blood pressure, BMI, heart diseases, cholesterol levels, age, family history, and lifestyle habits.	High	Sprint 1
Healthcare Institutions and local government	Data collection	USN-2	As a health care institutions, I want to collect and Gather a comprehensive dataset of health records and relevant parameters for training the diabetes prediction model.	Collect a diverse and representative dataset containing information such as blood pressure, BMI, heart diseases, cholesterol levels, age, family history, and lifestyle habits	High	Sprint 1
Researchers and Academics	data preprocessing	USN-3	Preprocess the collected dataset by cleaning, normalizing, and splitting it into training and validation sets.	Successfully clean and preprocess the dataset, handling missing values, outliers, and data inconsistencies.	High	Sprint 2
Healthcare Professionals	Model Development & Training	USN-4	select the most suitable model for predicting diabetes onset and Train the selected machine learning model using the preprocessed dataset.	Train the model using the preprocessed dataset. Monitor and optimize the model's performance on the validation set	High	Sprint 3
System Administrators	Model Deployment & Integration	USN-5	As a system Administrator, I want to Deploy the trained machine learning model as a service or API and integrate it into a user-friendly interface.	Develop a user interface for individuals to input their health records and receive diabetes prediction results.	medium	Sprint 4
Individuals/Patients	Personalized Risk Assessment	USN-6	As an individual, I want to input my health data into the system to receive a personalized risk assessment for diabetes onset.	The report should explain the factors that contribute to their risk of diabetes, and provide recommendations for reducing their risk	medium	Sprint 5
Researchers And Academics	Model Evaluation and Enhancement	USN-7	As a researcher, I want tools to evaluate the effectiveness of the diabetes prediction model and continuously enhance its performance.	Implement model evaluation metrics (e.g., accuracy, precision, recall).	medium	Sprint 5