**TASK 4:**

**Diabetes Prediction in Healthcare:**

**1. Problem Definition:**

To predict the diabetes for the patient using machine learning

**2. Data Collection:**

Collect the patient data from any resource that include age,Blood pressure,glucose level ,gender,BMI and ect..

**3. Data Cleaning:**

Remove the duplicates and irrelevant coloumns.Handle missing values and replace it bye their mean value.also need to do standardization .by the use of encoding technique converts the categorical values into continuous values.

**4. Exploratory Data Analysis (EDA):**

Visualize data using histograms, scatter plots, and box plots. Identify correlations between features and diabetes occurrences. Identify correlations between features and diabetes occurrences.

**5. Feature Engineering:**

Create new features based on domain knowledge (e.g., age group, BMI category). Encode categorical variables for example one-hot encoding for gender. Normalize numerical features to ensure consistent scaling.

**6. Model Building:**

Select the Algorithm such as logistic regression,random forest and so on.split the data set into training set and testing set .Training data set should be 70% and testing should be 80%.

Next do the Hyperparameter Tuning:Optimize model parameters using techniques like Grid Search or Random Search.

**7. Model Evaluation:**

Evaluate the model with metrices like precision,recall,F1 score and accuracy score. Avoid overfitting in the model and if the accuracy score is less than 70% then neeed to retrain the model.

**8. Deployment:**

Deploy the model in the product environment which is able to use by the user.