ST JOSEPH COLLEGE OF ENGINEERING

CSE-DEPARTMENT

TOPIC

AI-BASED DIABETIS PEDICTION MODEL

PROJ_227128_TEAM_1

PHASE-1

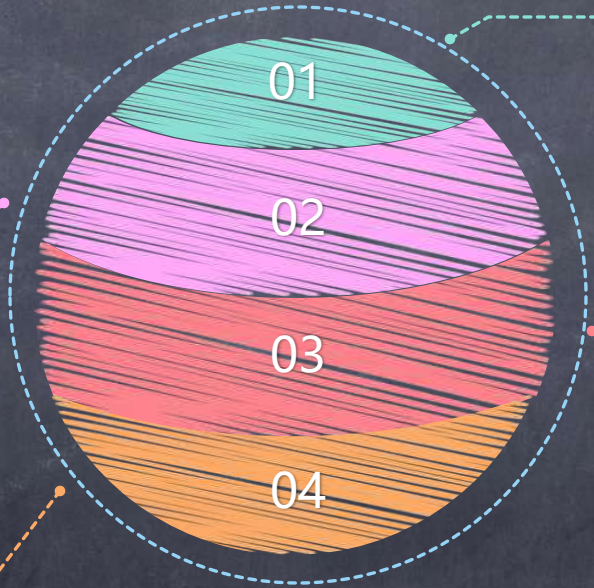
01 | PROBLEM DEFNITION

02 | DESIGN THINKING

01

PROBLEM DEFINITION

01 PROBLEM DEFINITION



DATA QUALITY

Performance

Model Selection

TESTING

01

DATA QUALITY

The performance of AI models heavily depends on the quality of the data used for training. If the data is incomplete, inconsistent, or imbalanced, it can lead to inaccurate predictions. For example, if a dataset has more instances of non-diabetic patients than diabetic ones, the model might be biased towards predicting a non-diabetic outcome.

MODEL SELECTION

Choosing the right machine learning model for prediction can be challenging due to the variety of models available. Each model has its strengths and weaknesses, and the choice of model can significantly impact the prediction accuracy.
these are also used for analysis in predictive and aknowledgeble way.

PERFORMANCE

Despite the advancements in machine learning techniques, their performance in predicting new-onset diabetes has not been superior to conventional risk stratification models that use statistical approaches.

01 TESTING

Splitting the Data: The dataset is typically split into a training set and a test set. The training set is used to train the model, while the test set is used to evaluate its performance.

Cross-Validation: Cross-validation is often used to assess how the results of the model will generalize to an independent dataset. It involves partitioning the data into subsets, training the data on a subset, and then validating the analysis on the other subset.
Performance Metrics: Several metrics can be used to evaluate the performance of the model, including:
Accuracy: This measures the proportion of true results (both true positives and true negatives) in the data.
Precision: This measures the proportion of true positive results in the data.
Recall (Sensitivity): This measures the ability of the classifier to find all positive instances.
F1 Score: This is a weighted average of precision and recall.
ROC Curve: This is a plot that illustrates the diagnostic ability of a binary classifier.
AUC Score: This tells how much the model is capable of distinguishing between classes.
Confusion Matrix: This is a table that is often used to describe the performance of a classification model on a set of data for which the true values are known.
Overfitting & Underfitting Check: Overfitting occurs when your model learns too much from training data and performs poorly on unseen data. Underfitting occurs when your model cannot capture underlying trends of the data.

02

DESIGN THINKING

02 EMPATHIZE

The first step in design thinking is to understand the needs and challenges of the end-users, which in this case are patients and healthcare providers. This could involve conducting interviews or surveys to gather insights about their needs and expectations from an AI-based diabetes prediction system.

02 DEFINE

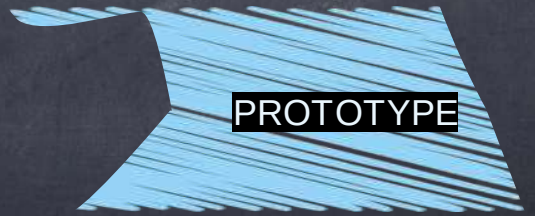
The next step is to clearly define the problem that needs to be solved. This could be predicting the likelihood of a patient developing diabetes based on their health data, or it could be identifying early warning signs of diabetes.

02 Ideate

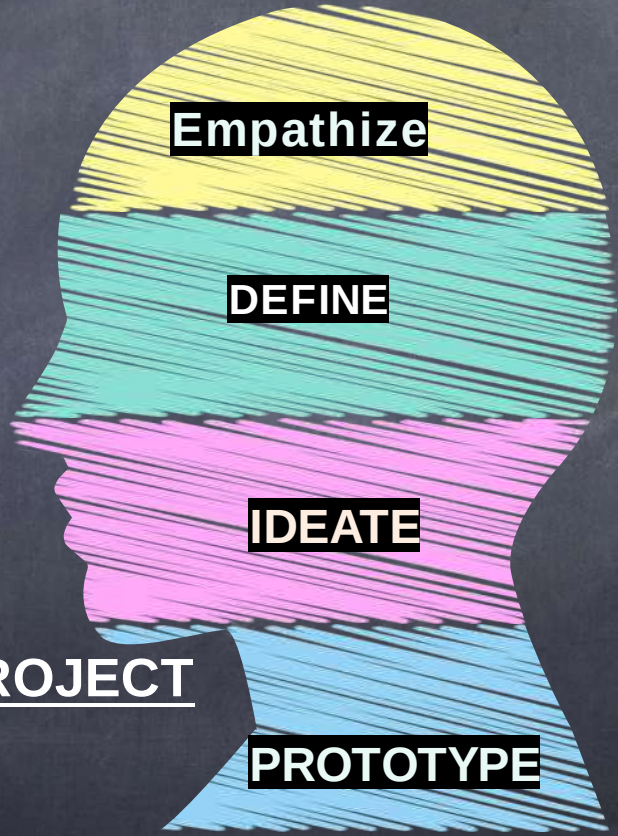
Once the problem is defined, the next step is to generate a wide range of ideas for potential solutions. This could involve exploring different machine learning algorithms (like decision trees, neural networks, or support vector machines), feature selection methods (like mutual information, chi-square test, or recursive feature elimination), and data preprocessing techniques (like handling missing values, outlier detection, or data normalization).

02 PROTOTYPE

After settling on a promising solution, a prototype of the AI system is developed. This prototype can then be tested using real-world data to see how well it performs.this will act as a source of project and deleberately the object will outsource the define result



02 OVERVIEW



THIS OUTSOURCES THE OVERALL PROJECT

THANK YOU !

Reported by : SYED UZAIR AHMED.Q
212921104052

PROJECT 227128 TEAM 1