Project Hypothesis: Early Detection of Type 2 Diabetes

Simple Hypothesis

Symptoms and demographic features can be used to accurately predict early-stage Type 2 Diabetes using supervised machine learning models.

Expanded Hypothesis

There is a statistically significant relationship between specific clinical symptoms (such as polyuria, polydipsia, sudden weight loss) and the onset of Type 2 Diabetes. By using supervised machine learning algorithms trained on these symptom patterns, it is possible to predict the likelihood of diabetes onset with high accuracy, thereby enabling early intervention and prevention strategies.

Hypothesis Components

Element	Description
Independent variables	Symptoms + demographic features (age, gender, etc.)
Dependent variable	Presence or absence of T2D
Test	Performance of models like Logistic Regression, Random Forest, Decision Tree, etc. on predicting T2D
Measure	Accuracy, Precision, Recall, F1-Score, ROC-AUC