**AI Based Diabetes Prediction System**

**Problem Definition:** The problem is to build an AI-powered diabetes prediction system that uses machine learning algorithms to analyze medical data and predict the likelihood of an individual developing diabetes. The system aims to provide early risk assessment and personalized preventive measures, allowing individuals to take proactive actions to manage their health.

**Design thinking:**

* **Data collection:**  
  It is crucial for this project as the AI gets trained upon the dataset. More the quality of data, more is the prediction accuracy. Hence dataset is to be collected which should encompass a wide range of information, such as blood glucose levels, blood pressure readings, body mass index (BMI), and details on whether or not each person has been diagnosed with diabetes.
* **Data preprocessing:**

This involves identifying and correcting errors or inconsistencies in the data, such as missing values, outliers, and duplicates. Various techniques can be used for data cleaning, such as imputation, removal, and transformation.

* **Feature selection:**

Not all the data collected would be useful and effective for prediction and not all the data would contribute equally. Hence we need to select features which would affect the prediction the most to improve accuracy.

* **Model Selection:**

In order to perform tasks like classification and regression we would begin with Logistic Regression followed by complex models such as Random Forest, Gradient Boosting or Support Vector Machines. Scikit-learn in python would be used to implement these models.

* **Model Evaluation:**

The performance of the model would be evaluated using approximate metrics like accuracy, precision, recall, F1-score and ROC-AUC. Visualization of the results would be done using ROC curves and confusion matrices.

* **Iterative improvement:**

We would continuously improve the performance and accuracy of our model by refining the data preprocessing steps. We would continuously strive to improve our approach with the suggestions given by mentors and domain experts.

* **Documentation and reporting:**

Documentation of the results play a crucial role in keeping track of all the results obtained. We would prepare a report that holds all the major results obtained. We would deploy our AI model in real-world healthcare setting so that the model would play an important part by contributing to the medical field.