## **Understanding the Problem**

### ****Regression****

* Predicts a **continuous numerical value.**

### ****Classification****

* Predicts a **discrete class label.**
* Example: Predicting if a person has diabetes (Yes = 1, No = 0).

### ****Regression as Classification****

We can use a **linear regression model** to output a continuous score, and then **convert** that score into a class label using decision boundaries.

For example:

* y<0.5y ​<0.5 → Class 0 (No diabetes)
* y≥0.5y ≥0.5 → Class 1 (Diabetes present)

## **Dataset Description:-**

|  |  |  |
| --- | --- | --- |
| **Feature** | **Type** | **Description** |
| gender | Integer | 0 = Female, 1 = Male |
| age | Float | Age in years |
| bmi | Float | Body Mass Index |
| diabetes | Integer | 0 = No diabetes, 1 = Diabetes present |

## **Process to Implement:-**

**Step 1 – Train a Multi-Variable Linear Regression Model**

* Features: gender, age, bmi
* Target: diabetes (binary, 0/1)

**Step 2 – Convert Regression Output to Class**

* Use threshold = 0.5:
  + If prediction < 0.5 → Class 0
  + Else → Class 1

## **Changes Needed in Multi-Variable Linear Regression Code for Classification:-**

1. **Target Variable Encoding**
   * Regression: y is continuous.
   * Classification: Encode categories (0, 1, 2, …).
2. **Post-processing Predictions**
   * Regression: Use the continuous output.
   * Classification: Apply threshold or mapping function to assign class labels.
3. **Evaluation Metrics**
   * Regression: MSE, RMSE, R2.
   * Classification: Accuracy, Precision, Recall, Confusion Matrix.
4. **Optional Upgrade**
   * Instead of LinearRegression + threshold, use Logistic Regression for better performance in binary classification.

## **Conclusion:-**

**In this assignment, we learn** that a regression model can be adapted to perform classification by converting continuous predictions into discrete categories using thresholds. While this method works for basic binary classification tasks, logistic regression or other dedicated classifiers are more suited for classification problems.