# Assignment-5

**Name: Shukla Kushang Akshay**

**Email ID: kushangashukla1@gmail.com**

**Mobile No.: +91 8511350065**
**Enrollment No.: 221130107024**
**College Name: 113-Sal College of Engineering**
**Branch: Computer Engineering**
**Semester- 6$^{th}$**
**Faculty Code: T00111- Mikin Dagli Sir**

- **Project Report:**
  - Regression Evaluation Metrics & Confusion Matrix Analysis.

- **Evaluation Metrics for Regression Algorithms:**

  - **Mean Absolute Error (MAE):**
    - **Description:** MAE measures the average absolute difference between actual and predicted values. It gives a direct interpretation of how much error exists in the predictions.
    - **Example Use Case:** Suitable for predicting house prices where equal importance is given to all errors, whether small or large.

  - **Mean Squared Error (MSE):**
    - **Description:** MSE calculates the average of the squared differences between actual and predicted values, penalizing more significant errors than smaller ones.
    - **Example Use Case:** Used in weather forecasting models where large deviations should be penalized heavily to improve accuracy.

  - **Root Mean Squared Error (RMSE):**
    - **Description:** RMSE is the square root of MSE, making it more interpretable in terms of the original data units.
    - **Example Use Case:** Ideal for energy consumption prediction models where a lower RMSE means better forecasting accuracy.

- **R-squared ($R^2$) Score:**
  - **Description:** $R^2$ represents the proportion of variance explained by the model. It ranges from 0 to 1, with higher values indicating better fit.
  - **Example Use Case:** Used in stock market prediction to measure how well historical data explains future trends.

- **Mean Percentage Error (MPE) & Mean Absolute Percentage Error (MAPE):**
  - **Description:** MPE calculates the percentage difference between actual and predicted values, while MAPE takes absolute values to avoid cancellations.
  - **Example Use Case:** Used in sales forecasting to measure percentage-based errors in revenue predictions.

## Understanding the Confusion Matrix:

- **What is a Confusion Matrix?**
  - A confusion matrix is a table used to evaluate classification models by displaying actual vs. predicted classifications. It helps assess model performance using four key values:
    - **True Positives (TP):** Correctly predicted positive instances.
    - **True Negatives (TN):** Correctly predicted negative instances.
    - **False Positives (FP):** Incorrectly predicted positive instances (Type I Error).
    - **False Negatives (FN):** Incorrectly predicted negative instances (Type II Error).

- **Example AI Model & Predictions:**
  - Consider an AI system predicting whether a machine part is defective (1) or not defective (0). Below are the model's predictions on a fictitious dataset:

| Actual Value | Predicted Value |
|---|---|
| 1 | 1 |
| 0 | 0 |
| 1 | 0 |
| 0 | 1 |
| 1 | 1 |
| 1 | 0 |

- **Calculating Confusion Matrix Metrics:**
  - **Precision = TP / (TP + FP):**
    - Precision measures how many of the predicted positives were correct.
  - **Recall = TP / (TP + FN):**
    - Recall determines how well the model identified actual positives.
  - **F1 Score = 2 × (Precision × Recall) / (Precision + Recall):**
    - F1 Score balances precision and recall, especially useful when dealing with imbalanced datasets.

- **Interpretation of Results:**
  - If precision is high but recall is low, the model is good at avoiding false positives but may miss actual defective parts. If recall is high but precision is low, the model catches most faulty parts but also predicts too many non-defective ones as defective. A good model balances both.