Date: 13th March 2025, PR lab (IT517) TU

Assignment 4

Qestion 1: Write a Program

- Calculate Euclidian distance between two 2D points
- Calculate Manhattan (taxicab) distance between two 2D points
- Calculate Chebyshev (chessboard) distance between two 2D points.
- Calculate Minkowski distance with parameter p between two 2D points.
  When p=1: Manhattan distance
  When p=2: Euclidean distance
  When $p \to \infty$: Chebyshev distance
- Calculate Hamming distance between two binary vectors.
- Calculate cosine distance between two vectors
- Calculate Mahalanobis distance between two vectors.

Qestion 2: Data Transformation
(Min max Transformation)

- **Generate sample data**
  np.random.seed(42)
  original_data = np.random.normal(loc=50, scale=15, size=100)

- **Manual min-max scaling (0 to 1)**
  Manually perform min-max scaling on a numpy array or list
  Formula: X_scaled = (X - X_min) / (X_max - X_min)
  Parameters:
  X: numpy array or list to be normalized
  Returns:
  numpy array with normalized values between 0 and 1
- **Custom range min-max scaling (-1 to 1)**
  Scale data to a custom range [new_min, new_max]

Formula: X_scaled = (X - X_min) / (X_max - X_min) * (new_max - new_min) + new_min
Parameters:
X: numpy array or list to be normalized
new_min: minimum value of the new range
new_max: maximum value of the new range
Returns:
numpy array with normalized values between new_min and new_max

- **Using scikit-learn's MinMaxScaler**

- **Display results**
  Min:
  Max:
  Mean:
  Standard Deviation:
- **Create a DataFrame for easy comparison**
  First 5 rows of all datasets
- **Visualize the transformations**

Question 3:
Given the following predictions and actual values:
y_true = [1, 0, 1, 1, 0, 0, 1, 0]
y_pred = [1, 0, 0, 1, 1, 0, 1, 1]
Calculate

- **True Positives (TP):** Correctly predicted positives
- **False Positives (FP):** Incorrectly predicted positives
- **True Negatives (TN):** Correctly predicted negatives
- **False Negatives (FN):** Incorrectly predicted negatives
- **Accuracy**: Overall correctness (TP + TN) / Total
- **Precision**: How many predicted positives were correct TP / (TP + FP)
- **Recall**: How many actual positives were found TP / (TP + FN)
- **F1 Score**: Harmonic mean of precision and recall
- **Mean Squared Error (MSE) Formula:** MSE = (1/n) * Σ(y_true - y_pred)²
- **Root Mean Square Error (RMSE) Formula:** RMSE =sqrt((1/n) * Σ(y_true - y_pred)²)
- **Mean Absolute Error (MAE) Formula:** MAE = (1/n) * Σ|y_true - y_pred|
- **Binary Cross-Entropy Loss :** Formula: -[y_true * log(y_pred) + (1 - y_true) * log(1 - y_pred)]