

Assignment: Algorithm Comparison

Objective:

The objective of this assignment is to help students understand the situations where specific machine learning algorithms—Logistic Regression, KNN, Decision Tree, and SVM—are most suitable. Students will explore the strengths, weaknesses, and suitability of each algorithm for different datasets.

Instructions:

Complete the following tasks by comparing the algorithms based on their characteristics, performance, and application scenarios. The assignment consists of 2 parts.

Part 1: Algorithm Overview

For each of the algorithms below, write a brief overview that includes:

- How the algorithm works.
- Two key strengths and two limitations.

Algorithms:

1. Logistic Regression
2. K-Nearest Neighbors (KNN)
3. Decision Tree
4. Support Vector Machine (SVM)

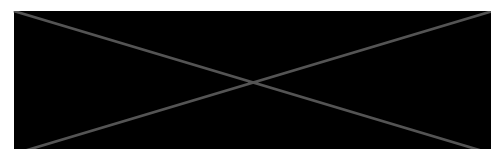
Expected Output:

- 50-75 words per algorithm.

Part 2: Application Scenarios

For each of the following dataset scenarios, recommend the most suitable algorithm (Logistic Regression, KNN, Decision Tree, or SVM). Provide a brief explanation for your choice.

1. **High-Dimensional Data** (e.g., text or gene expression data)
2. **Imbalanced Dataset** (e.g., fraud detection, rare disease prediction)
3. **Small Dataset with Many Features** (e.g., medical or genetic data)
4. **Non-linear Data Separation** (e.g., complex shapes like spirals or circles)
5. **Dataset with Noise** (e.g., data with many irrelevant or misleading features)



Expected Output:

- 50-100 words for each scenario explaining your choice and why the selected algorithm is suitable.

Deliverables:

1. **Submission Format:** PDF report covering both parts.

