Machine Learning Algorithms Implemented From Scratch

1. Introduction
This project focuses on implementing core machine learning algorithms entirely from scratch using NumPy and Pandas. No high-level machine learning libraries were used. The aim was to understand the internal workings of learning algorithms, optimization techniques, and evaluation methods.

2. Libraries Used
NumPy, Pandas, Matplotlib, and Google Colab were used for numerical computation, data handling, visualization, and execution.

3. Datasets
Separate datasets were provided for linear regression, polynomial regression, binary classification, and multiclass classification. Each dataset was divided into training and test sets.

4. Preprocessing
Missing labels were removed. Feature normalization was applied using statistics computed only from training data to prevent data leakage. Multiclass labels were encoded and one-hot encoded manually.

5. Implemented Algorithms
• Linear Regression using Gradient Descent
• Polynomial Regression with feature expansion
• Logistic Regression for binary classification
• KNN and K-Means clustering
• Decision Tree using information gain
• Multi-layer Neural Network with backpropagation, Adam optimizer, and proper weight initialization

6. Training and Evaluation
All models were trained strictly on training data. Test datasets were used only for final evaluation. Appropriate metrics such as MSE and Accuracy were used.

7. Results
The models demonstrated correct learning behavior and reasonable performance on unseen test data, confirming the correctness of the implementations.

8. Challenges
Key challenges included normalization consistency, polynomial feature dimensionality, and debugging shape mismatches. These were resolved through systematic testing and validation.

9. Conclusion
This project strengthened understanding of machine learning fundamentals by building algorithms from the ground up. All implementations complied strictly with the given constraints.

Declaration
This project was implemented entirely from scratch using NumPy and Pandas without using any external ML libraries.