Regression vs Classification

1. Basic Definitions

1.1 Regression

- Predicts continuous numerical values
- Output is a quantity or magnitude
- Example outputs: price, temperature, height, weight
- Measures accuracy using error metrics (MSE, RMSE, MAE)

1.2 Classification

- Predicts discrete categories or classes
- Output is a label or category
- Example outputs: yes/no, spam/not spam, type of flower
- Measures accuracy using classification metrics (accuracy, precision, recall)

2. Key Differences

2.1 Output Type

Regression:

- Continuous numbers
- Infinite possible values
- Represents quantities
- Example: House price ($250,756)

Classification:

- Discrete categories
- Finite number of possibilities
- Represents classes/labels
- Example: Email type (spam/not spam)

2.2 Evaluation Metrics

Regression Metrics:

- Mean Squared Error (MSE)
- Root Mean Squared Error (RMSE)
- Mean Absolute Error (MAE)
- R-squared (R^2)
- Adjusted R-squared

Classification Metrics:

- Accuracy
- Precision
- Recall
- F1-Score
- ROC Curve
- AUC Score
- Confusion Matrix

3. Common Algorithms

3.1 Regression Algorithms

1. Linear Regression
2. Polynomial Regression
3. Ridge Regression
4. Lasso Regression
5. Support Vector Regression (SVR)
6. Decision Tree Regression
7. Random Forest Regression

3.2 Classification Algorithms

1. Logistic Regression
2. K-Nearest Neighbors (KNN)
3. Support Vector Machines (SVM)
4. Decision Trees
5. Random Forests
6. Naive Bayes
7. Neural Networks

4. Applications

4.1 Regression Applications

- Stock price prediction
- Sales forecasting
- Temperature forecasting
- Age prediction
- Salary estimation
- Population growth prediction
- Resource consumption prediction

4.2 Classification Applications

- Email spam detection
- Credit card fraud detection
- Disease diagnosis
- Image recognition
- Customer churn prediction
- Sentiment analysis
- Document categorization