Programming Assignment 2 Part 1 - Decision Tree

Results (Accuracy Difference):

- **Without Pruning**:
    - Noiseless Data: **71.4%**
    - Noisy Data: **48.1%**
    - **Accuracy Difference**: $71.4\% - 48.1\% = 23.3\%$
- **With Reduced Error Pruning**:
    - Noiseless Data: **76.6%**
    - Noisy Data: **57.83%**
    - **Accuracy Difference**: $76.6\% - 57.83\% = 18.77\%$
- **With Pre-Pruning (Model 3)**:
    - Noiseless Data: **73.37%**
    - Noisy Data: **62.7%**
    - **Accuracy Difference**: $73.37\% - 62.7\% = 10.67\%$

Conclusion:

By modifying the decision tree through **pre-pruning** (max depth, min samples split, min samples leaf), the **accuracy difference between noiseless and noisy data decreased from 23.3% to 10.67%**, a significant reduction.