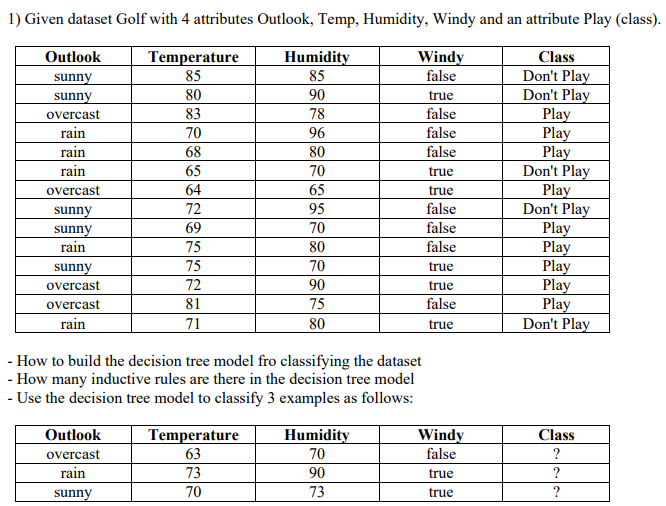
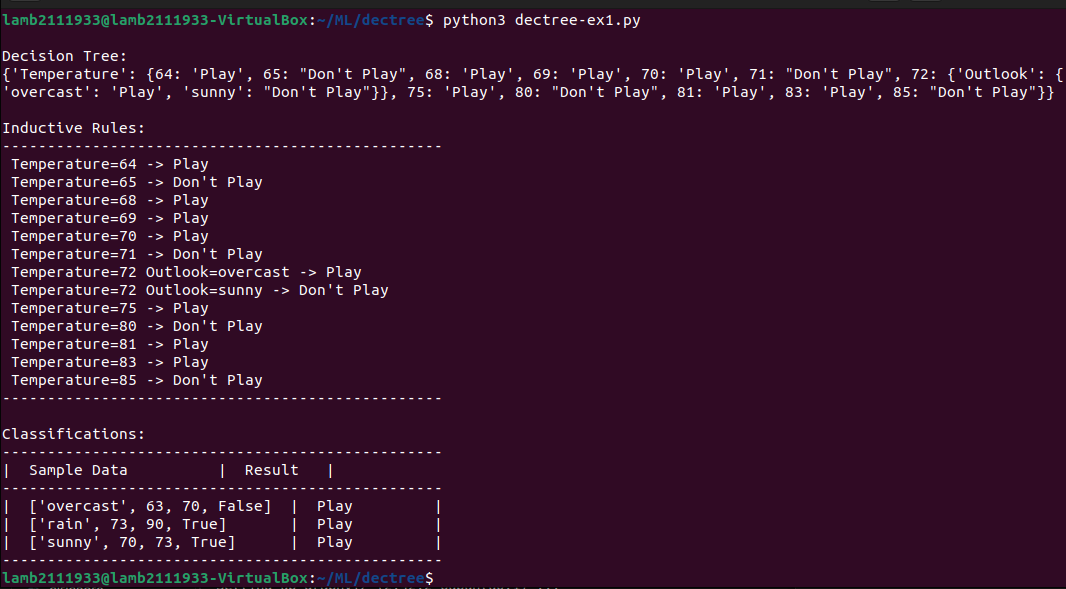
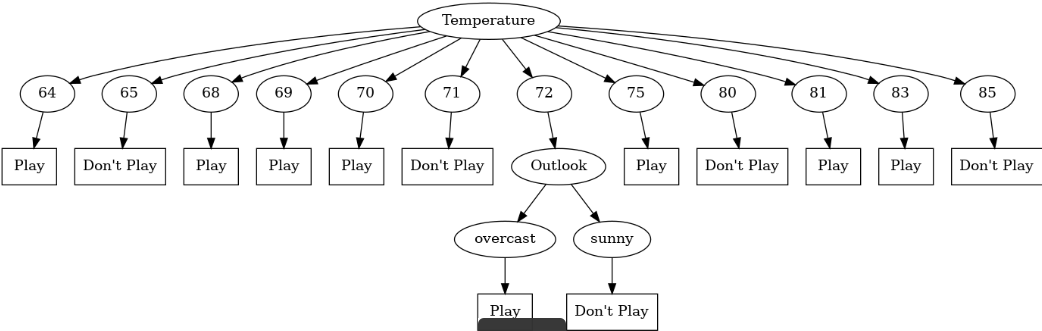
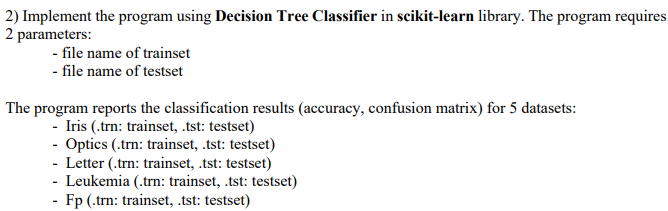
Name: Truong Dang Truc Lam ID: B2111933 Class: CT205H - M04

# Decision Tree

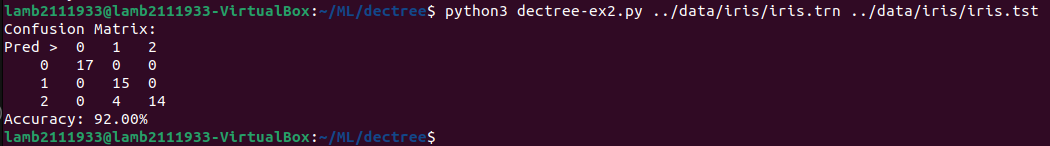




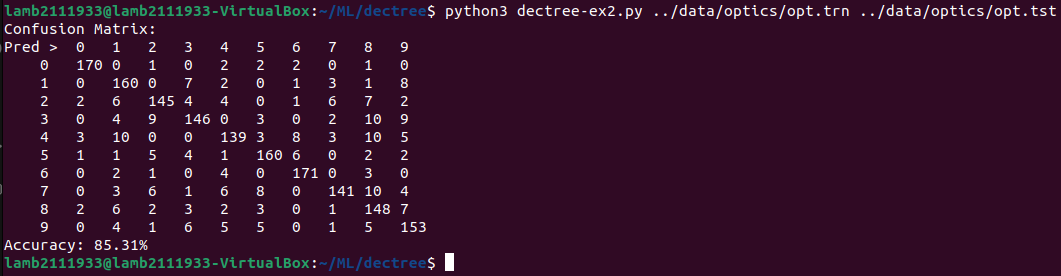




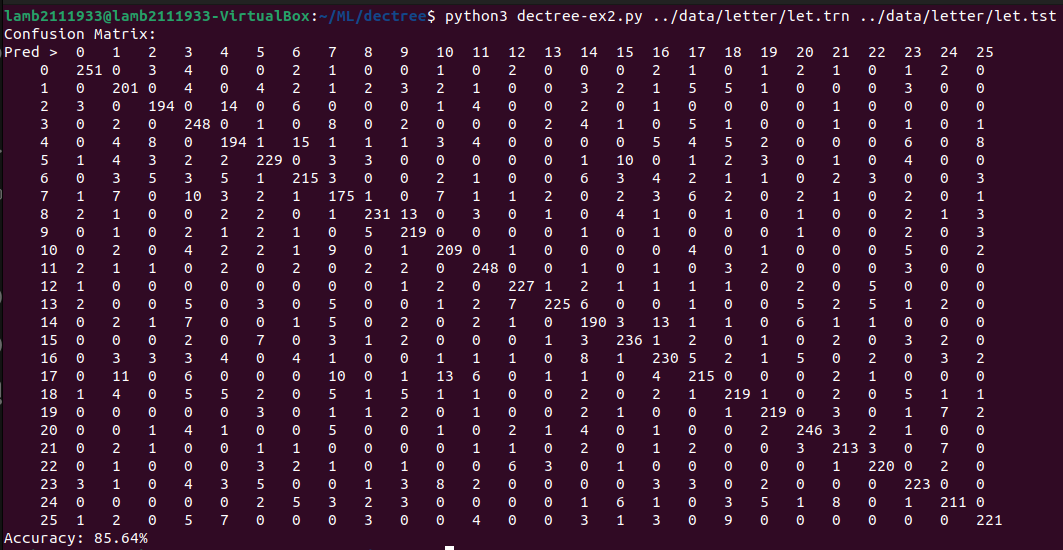
Iris dataset:



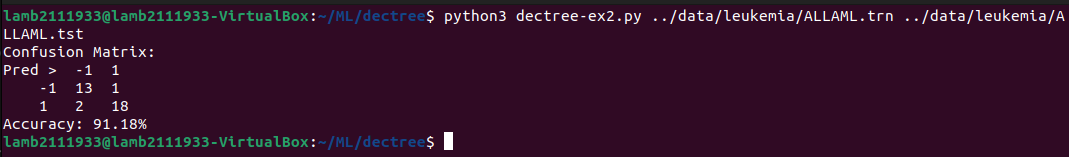
Optics dataset:



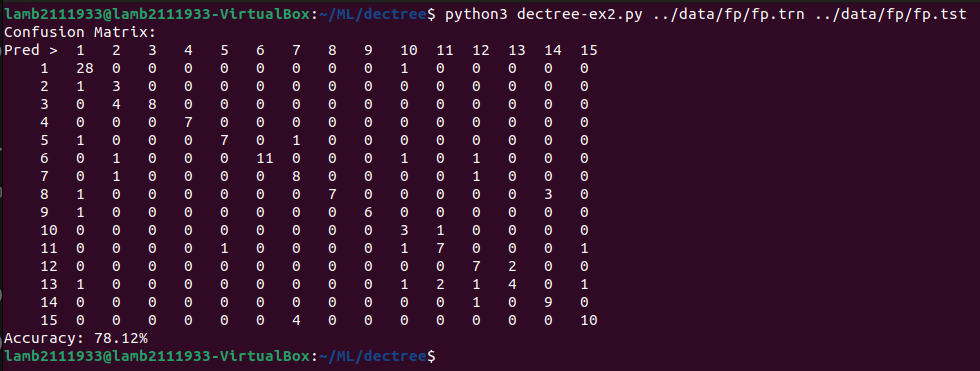
Letter dataset:



Leukemia dataset:



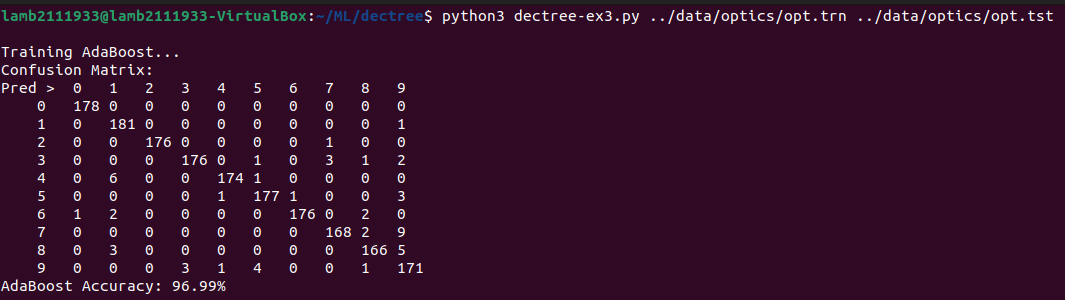
Fp dataset:

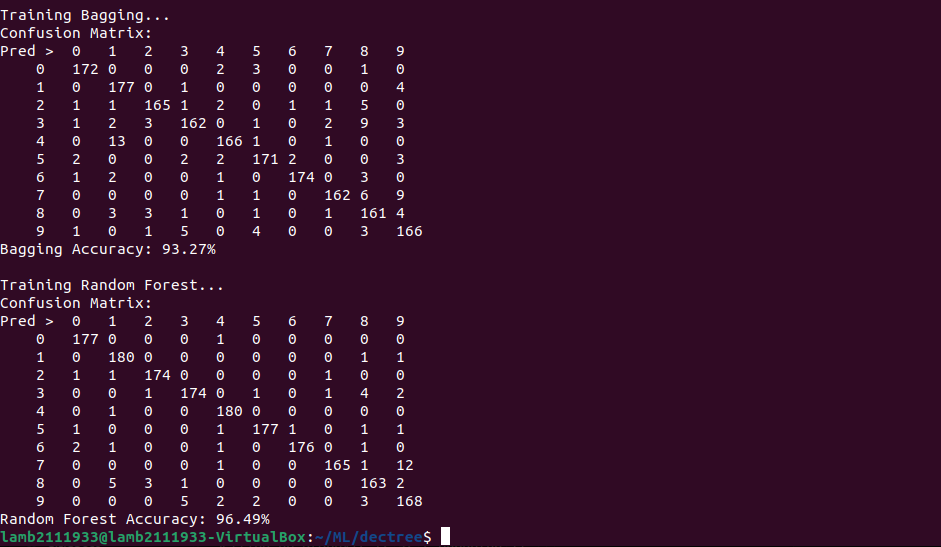




Iris dataset:

Optics dataset:





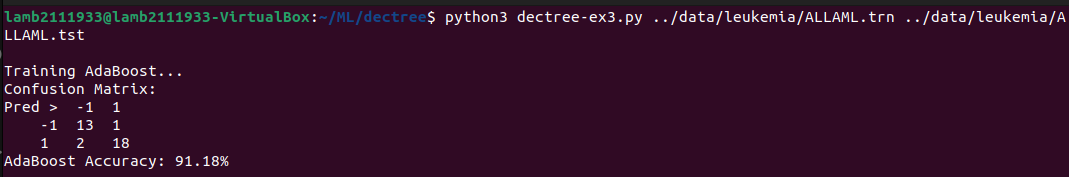
Letter dataset:





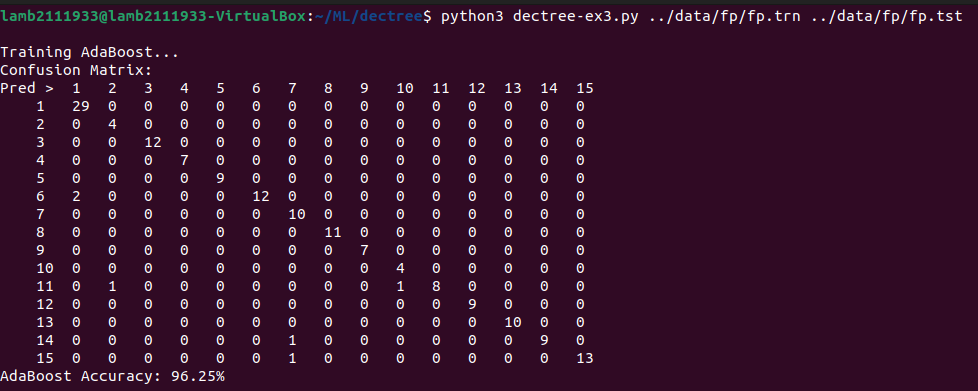


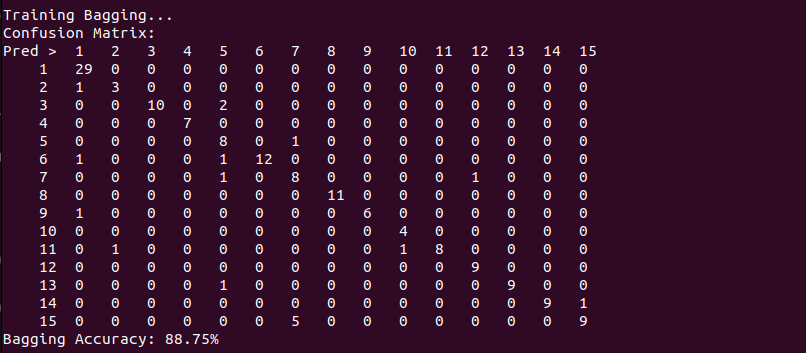
Leukemia dataset:

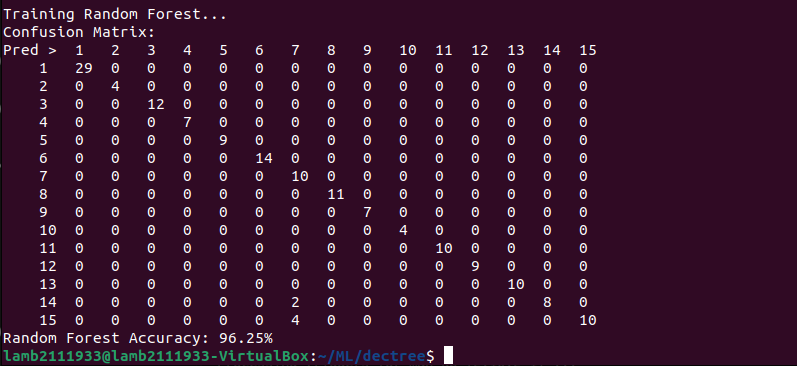




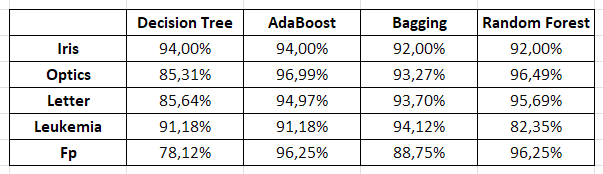
Fp dataset:











Ensemble methods like AdaBoost, Bagging, and Random Forest are usually more accurate than single decision trees. Because they are combined of multiple trees, leading to better predictive performance overall.



Decision Trees will offer a simple approach for classification but they often get overfitted, which can cause it exhibit high variance on new data.

Ensemble-based model (such as AdaBoost, Bagging, and Random Forest) will combine multiple decision trees to solve these problems and improve overall performance:

* AdaBoost will build a sequence of weak learners by emphasizing misclassified instances, which allows it to progressively focus on challenging cases.
* Bagging can reduce variance by training several trees on different subsets of the data and calculating the average of their predictions to create a more stable model.
* Random Forest can further refine Decision Trees approach by adding random feature selection at each split, which can help to reduce the similarity between trees and make the model more stable.