

Decision Tree

How to run the project

This is a simple python project including 3 parts of machine learning algorithm –tree, parser and main. In the zip file, there will be an extra data directory containing the dataset for learning and testing. In the main python file, it has some validation functions and it is also the entrance for the whole project. The tree file provides two basic classes of the tree and node. The Parser file is designed to parser the data file into other form.

How to meet requirements

1. As requirement, I use iris-discrete dataset and another one from UCI web for testing. And it adapts the cross-validation as method. Then the algorithm will automatically print the correct rate.
2. If you want to use the program to test other dataset files, you can move the file into data directory and correctly input document name in main file. But you need to make sure all the attributes in dataset are discrete, otherwise the program is meaningless.

Notes

If you want to know how to design the algorithm and the idea of function, you can check these in comments

I hope this doc will help you use the algorithm.