UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

COS 314

Project 2: Decision Trees

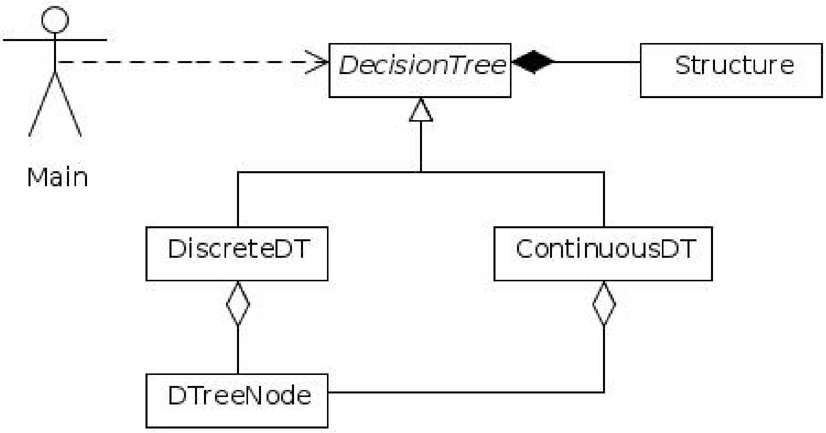
Juan Jaques du Preez
u15189016

Date:
April 29, 2017

Contents

1 Overview

1.1 Overall UML Diagram



1.2 Options Implemented

- -d
- -c
- -md
- not -mc
- -pd
- some of -pc

1.3 Please Note

The code is not as efficient as it could be. For example, I have two functions in DiscreteDT called induceWithMissing() and induceNoMissing(). These two are very much the same, but I wanted to make the different algorithms that were used as clear as possible.

2 Compile and Run Program

For this project I used the language of C++.

2.1 Compile

To compile the program, open a terminal and type in the following commands:

```
cd DecisionTree/
make
```

The makefile compiles the .h and .cpp files into a ”runnable” file in the TestSpace/ directory.

2.2 Run

To run the program, make sure you are in the ”TestSpace/” directory. Copy the two files (data file and spec file) into this directory. Open a terminal and type:

```
./DecisionTree option specfile datafile
```

Where option is one of: -d, -c, -md, -mc, -pd, -pc.

2.3 Contact

If you cannot get the program to compile and run, or have any other issue with regards to anything, please contact me:

Juan du Preez
078 141 0915
u15189016@tuks.co.za
juan.dupreez82@gmail.com

2.4 What does not work

To make your job easier, here is a list of things that do not work:

- There is no implementation for missing values when doing continuous data with missing values, i.e. the -mc option.
- The -pc option, or pruning the tree for continuous data, sometimes goes into an infinite loop.