BoolFuncOnDB README

# Introduction

The BoolFuncOnDB program takes in a dataset and a Boolean function to apply that Boolean function on each datapoint of the dataset. If a particular attribute of the dataset is not Boolean, then the program will automatically detect this and ask the user to define a range of two thresholds, or a minimum or maximum threshold, respectively. Values of “yes,” “true,” “t,” “1,” and others are accepted as Boolean. Otherwise, they will need a threshold. In the case of some value of something like “random\_text,” the program will not work.

# Definitions

**Range**: this is defined as an upper or lower bound. Therefore, a threshold of [2, 3] will define any value between 2 and 3 as true (1) and any number outside that threshold will be false.

**Maximum threshold**: for example, a maximum threshold of 2 will define any value less than or equal to 2 as true, and greater than 2 will be false.

**Minimum threshold**: for example, a minimum threshold of 2 will define any value greater than or equal to 2 as true, and lesser than 2 will be false.

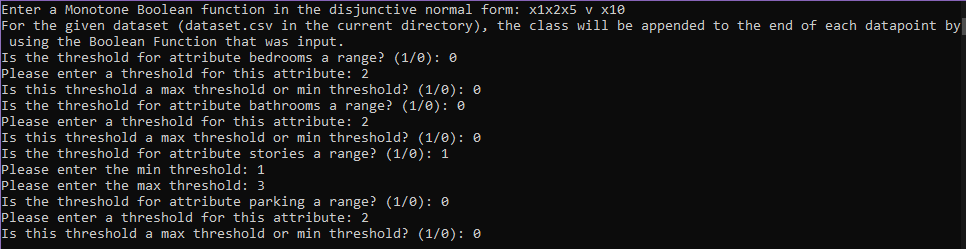
# Input/Output

The program takes in a CSV file as input. It simply has to be labeled “dataset.csv” in the executable’s directory. The first line of the file must be the names of the attributes. If there are more or less attributes than what is in the datapoints, or if the dataset is sparse. The program will likely not work.

The outputted results file will simply be labeled “results.csv.” The results file will have the Boolean function and thresholds (if there are any) located at the top of the file. Then, the datapoints will be listed along with the class. The results file is sorted so that datapoints with a class of 1 are given first. Moreover, the Boolean representation of each datapoint is given and located just to the right of each datapoint.

# Examples

Here is an example of some user input:



Here is an example of some output:

