**Decision Trees**

The first feature returns the player whose piece is in the bottom left corner of the board. This feature is not very useful because having a piece in the bottom left corner of the board is not something that should determine if a player wins. This feature was implemented because it was required by the project guidelines.

The second feature returns which player has more pieces that aren’t at the edge of the board. This feature should be useful because when a piece is at the edge of the board it has a lot less potential to be connected to. In general a piece toward the center would have a higher chance of being used to win than a piece on the side. This feature predicts that the player with the least pieces at the edge will win.

The third feature calculates a heuristic value based on how close all of a players pieces are to the center. This feature is similar to feature two because it also values pieces that are away from the edge of the board. This feature goes beyond feature two because each piece closer to the center has more weight in the heuristic calculation. The heuristic is calculated by adding one for every piece in an edge column. For every column closer to the center, a piece in that column is worth twice as much as a piece in the adjacent column closer to the edge of the board. Therefore, the number of pieces in a column is multiplied by 1, 2, 4, 8, 4, 2, or 1 respectively. The total value of all of the columns is evaluated and the player with the highest total is returned by the feature.

The fourth feature uses the same heuristic as feature three, but instead of returning whichever player has a higher heuristic, the feature returns the value of player one’s heuristic minus player two’s heuristic. This feature therefore returns a continuous function that represents how much of an advantage one player has over another because of a more centered positioning. This feature’s usefulness will be compared to the previous feature’s usefulness to determine whether it is better for this data set to use a continuous value to represent positioning advantage, or a categorized set of values.

The fifth feature tests who has “top control” of the columns of the board. “Top control” is defined as whoever has the highest piece on any given column. For every column where player 1 has top control, 1 is added to the top control variable and for every column that player 2 has top control, 1 is subtracted from this variable. Columns with no pieces in them do not count for top control at all and do not effect the top control score. When doing the first project, we found that often that your ability to make advantageous moves was dependent on how much control your had over the top edge of the board, because you’re only able to make plays that will become new edges on the top, interacting with the old top of the board.