Decision Trees

Assignment #4

For this assignment, we will be using the **Tic-Tac-Toe Endgame** dataset located in the Irvine’s Machine Learning Repository located at:

<https://archive.ics.uci.edu/ml/index.php>  
<https://archive.ics.uci.edu/ml/datasets/Tic-Tac-Toe+Endgame>

This **tic-tac-toe data** file under the **Tic-Tac-Toe Endgame** folder encodes the complete set of 958 board configurations at the end of all possible tic-tac-toe games, where **X** is assumed to have played first. The target concept is a win for **X**. Namely, the classification is *positive* if **X** has one of eight possible ways to create a “three-in-a-row”; otherwise, the classification is *negative*.

There are nine attributes, each of which represents one of the nine possible squares of the tic-tac-toe-board. These attributes are (in the order they appear in **tic-tac-toe data**):

1. top-left-square
2. top-middle-square
3. top-right-square
4. middle-left-square
5. middle-middle-square
6. middle-right-square
7. bottom-left-square
8. bottom-middle-square
9. bottom-right-square

Each attribute can take on one of three values:

1. x Player **X** has claimed that square.
2. o Player **O** has claimed that square.
3. b No player has claimed that square. It is blank.

You are tasked with implementing the **ID3** algorithm to create a decision tree to distinguish between the winning and losing Tic-Tac-Toe boards for the player **X**. Your program must additionally output the resulting decision tree in some meaningful and easy to discern way so that it can be seen and evaluated. This does not have to be graphical, but it must be understandable.

You must author your own code using either the C++, Java or Python programming language and without the use of *external* software packages. Your project submission will be run through a source code verifier against previous submissions to check for plagiarism. Any evidence of cheating (CODE NOT WRITTEN SOLEY BY YOU) will result in a failing grade for this course. Treat this seriously. I do!

When submitting your assignment include:

* the output file containing the resulting decision tree
* a copy or your source code
* a README file containing any information required to run your program.

Feel free to consult any relevant literature on decision trees, but **not other ID3 code**!