



## XGBoost Mini-Project

**Objective:** Use an extreme gradient boosting algorithm to classify human breast tumors as either benign or malignant based on recorded observations and physical measurements. Note: you can read about the data here: <http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/wdbc.names>.

**Architecture:** You are free to define any architecture that you like. The only constraint is that your script must contain a module called **train\_and\_test.py**, which, when run, will do the following automatically:

- 1) train your XGBoost algorithm on the data contained in data/train\_data.txt;
- 2) save the model;
- 3) reload the model and apply it to predict the labels of the data contained in data/test\_data.txt.

Note: because you want your script to run training and testing functions as soon as it's called, you'll need to use an `if __name__ == '__main__':` code block (read about it here if you haven't encountered this yet: <https://stackoverflow.com/questions/419163/what-does-if-name-main-do>).