Using XGBoost to Diagnose Breast Cancer

Run the Model

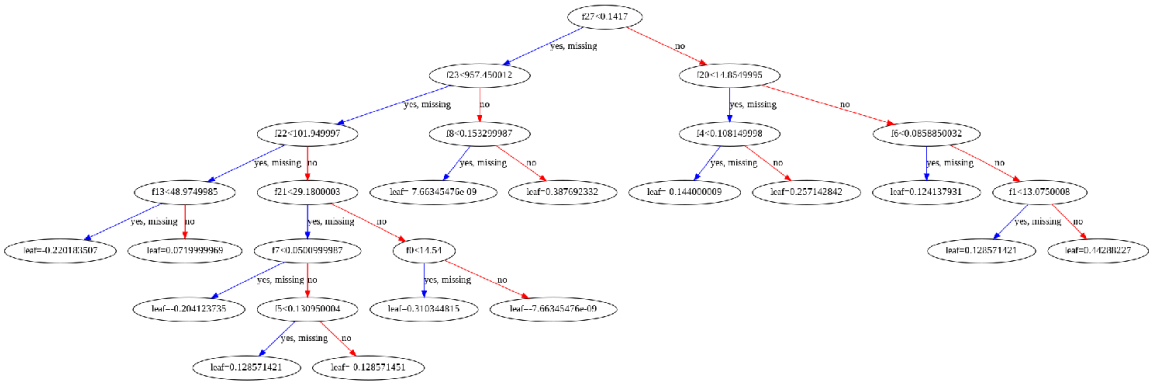
```
[7]  import numpy as np
     from sklearn.metrics import precision_score, recall_score, accuracy_score

     preds = model.predict(D_test)
     best_preds = np.asarray([np.argmax(line) for line in preds])

     print("Accuracy = {}".format(accuracy_score(Y_test, best_preds)))
     print("Precision = {}".format(precision_score(Y_test, best_preds, average='macro')))
     print("Recall = {}".format(recall_score(Y_test, best_preds, average='macro')))
```

```
Accuracy = 0.956140350877193
Precision = 0.9418604651162791
Recall = 0.9671052631578947
```

When a medical diagnosis is important, you mostly want to focus on achieving a high **Recall**, which suggests that the model is producing a small number of false negatives



A graphical depiction of the decision tree provides immediate insights

Attributions:

Original data provided by sklearn is here.

My code is located at my GitHub site.