Machine Learning Characterization of a Novel Panel for Metastatic Prediction in Breast Cancer

(Supplementary)

1. Then the precision, recall, F1-score and accuracy are calculated as shown below.

In Equations 1-4, TP (true positive) denotes the correctly classified metastasis classes that have assigned as 1 and TN (the true negative) implies the correctly classifies non-metastasis classes assigned as 0. On the other hand, FP (false pos- itive) indicates the number of false classifications of metastasis that are actually non-metastasis. FN (false negative) indicates the number of false classifications of non-metastasis that are actually metastasis.

1. According to Figure S1, a difference between genes with and without metastasis can be clearly seen in the heatmap.

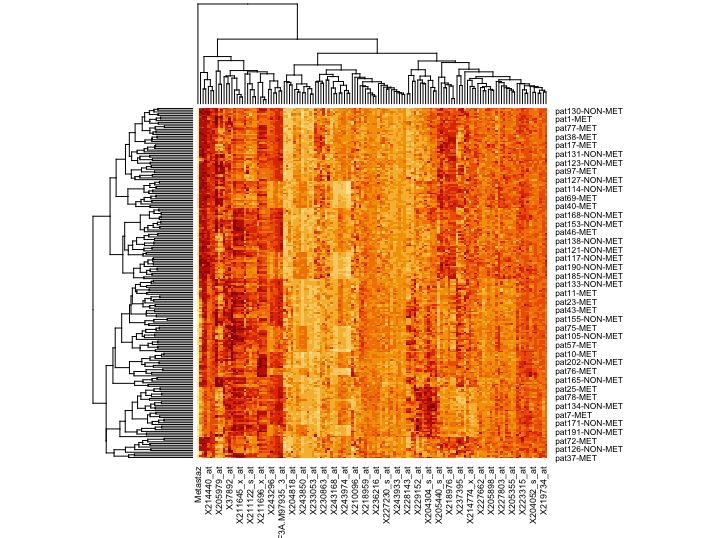


Figure S1: Heatmap of the dataset

1. According to Figure S2, network analysis of the dataset can be seen.

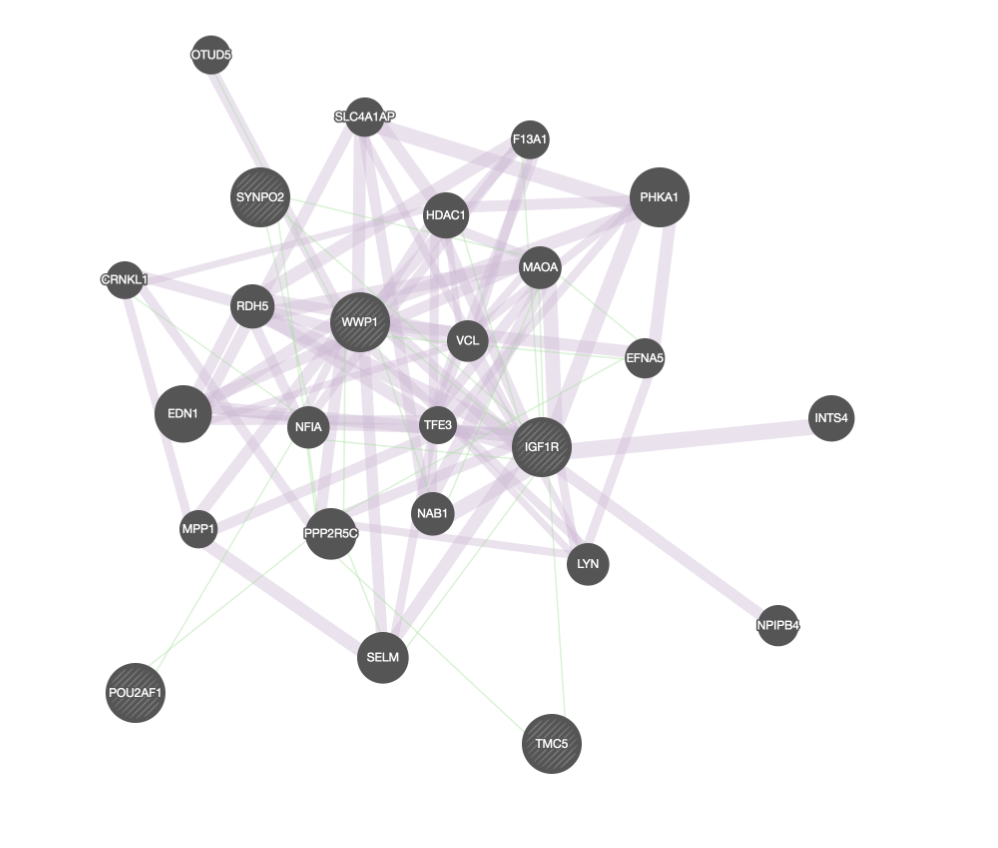


Figure S2: Network analysis of the output of the Gradient Boosting results.