

Figure 1: Lasso regularization path for regression coefficients of the input features. Coefficient values deviating from zero suggest that the corresponding feature might be a significant indicator for MI risk. Both plots show the regularization path for the same data set D_1 , but with different estimations for the penalty parameter λ (vertical line). This is caused by the randomness of selecting training and test sets in the cross-validation procedure.

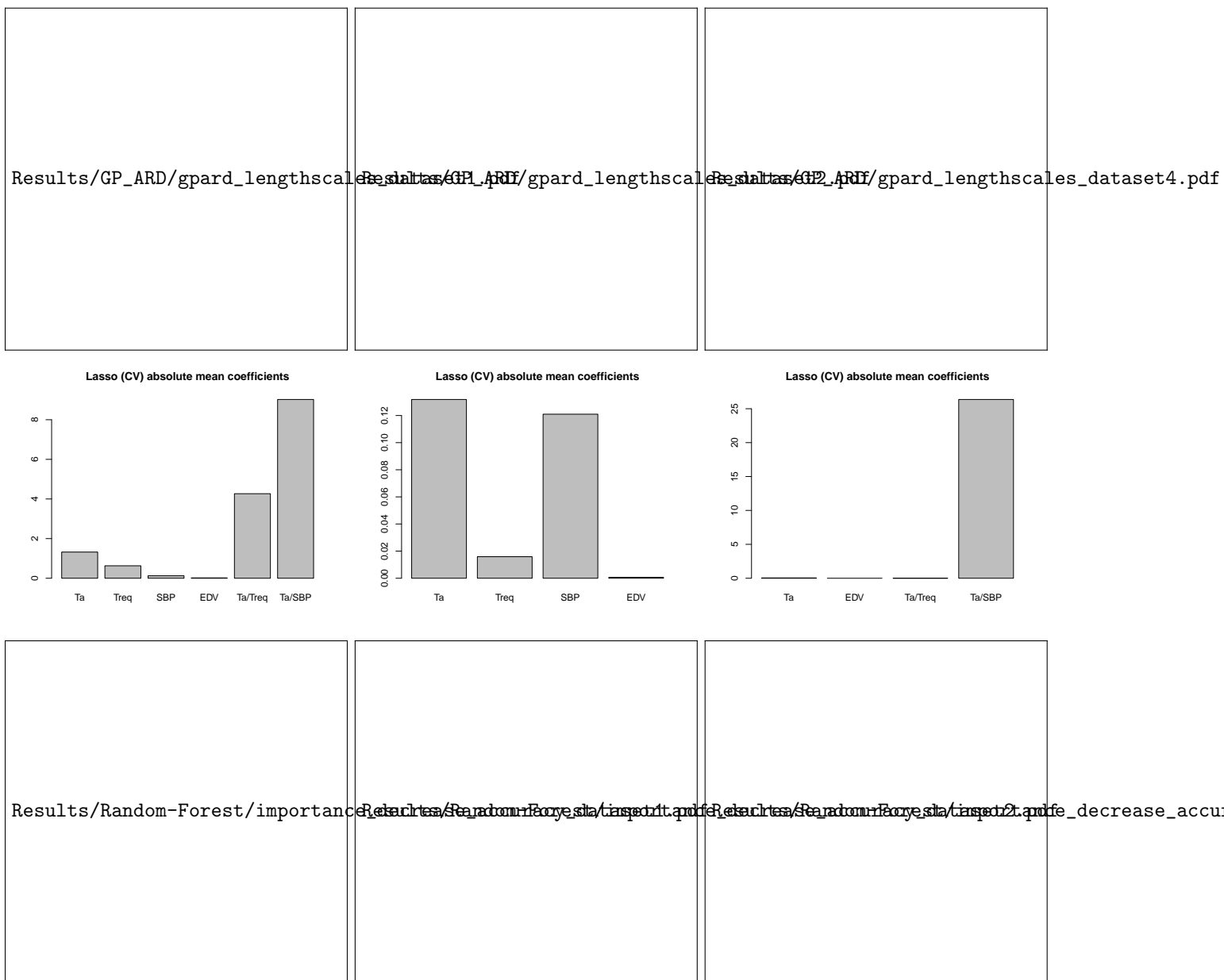


Figure 2: **Importance of features estimated by GP-ARD and Lasso:** The top plots show the inversed relative length scales from GP-ARD and the bottom plots show the optimized regression coefficients from Lasso. Each column corresponds to one data set configuration. The GP-ARD uses a length scale parameter to determine the features with the greatest explanatory power. The length scales were inverted to make a comparison to the Lasso coefficients easier.

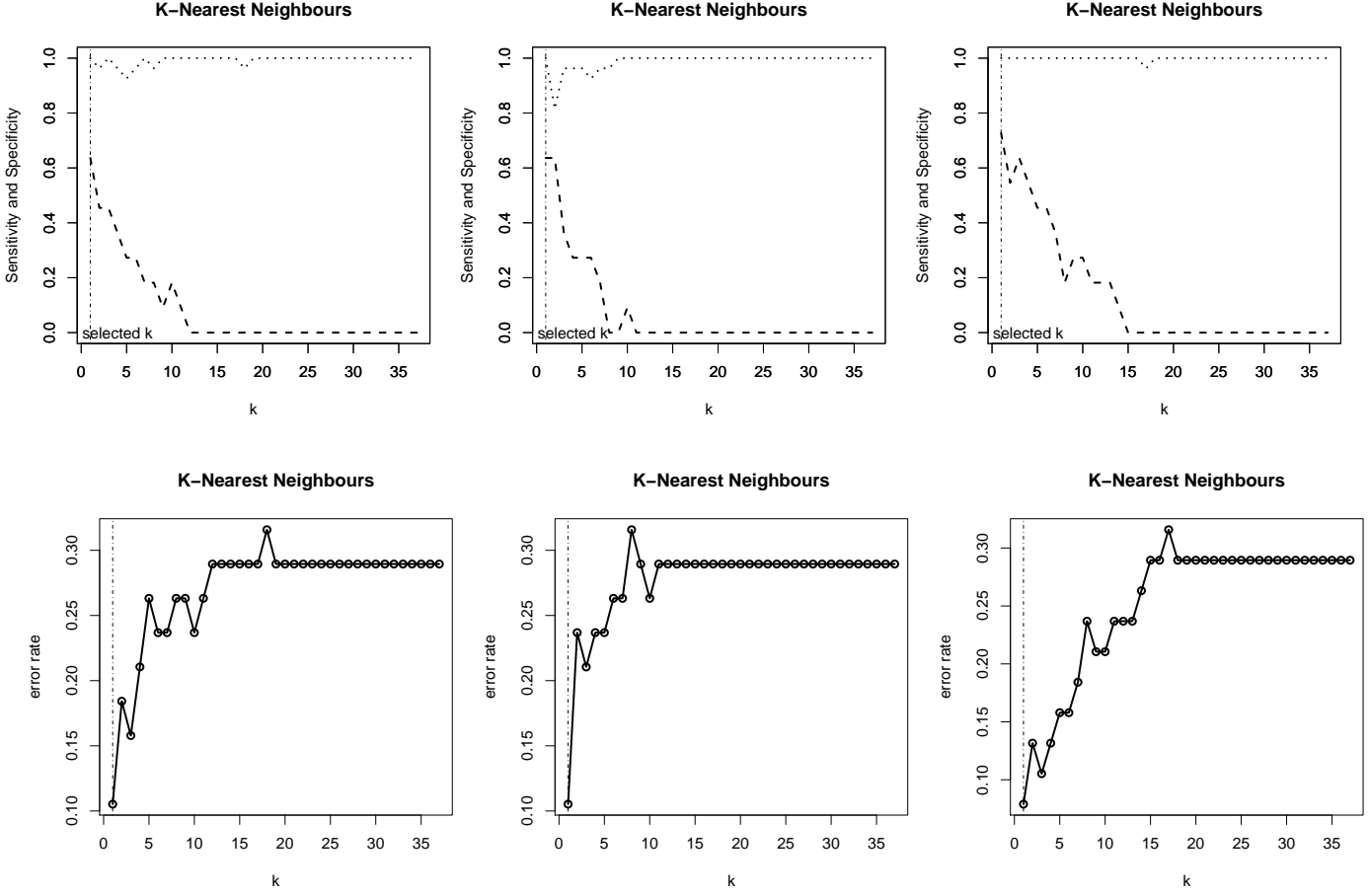


Figure 3: K-nearest neighbour (KNN) with varying parameter $k = (1, \dots, n - 1)$ and corresponding sensitivity (dashed line) and specificity scores (pointed line) in the top panel, and test mean squared errors (MSE) in the lower panel. The scores for each k were determined with out of sample cross-validation using LOOCV (leave one out cross-validation). The left, middle and right columns correspond to the data sets D_1 , D_2 and D_3 , respectively.

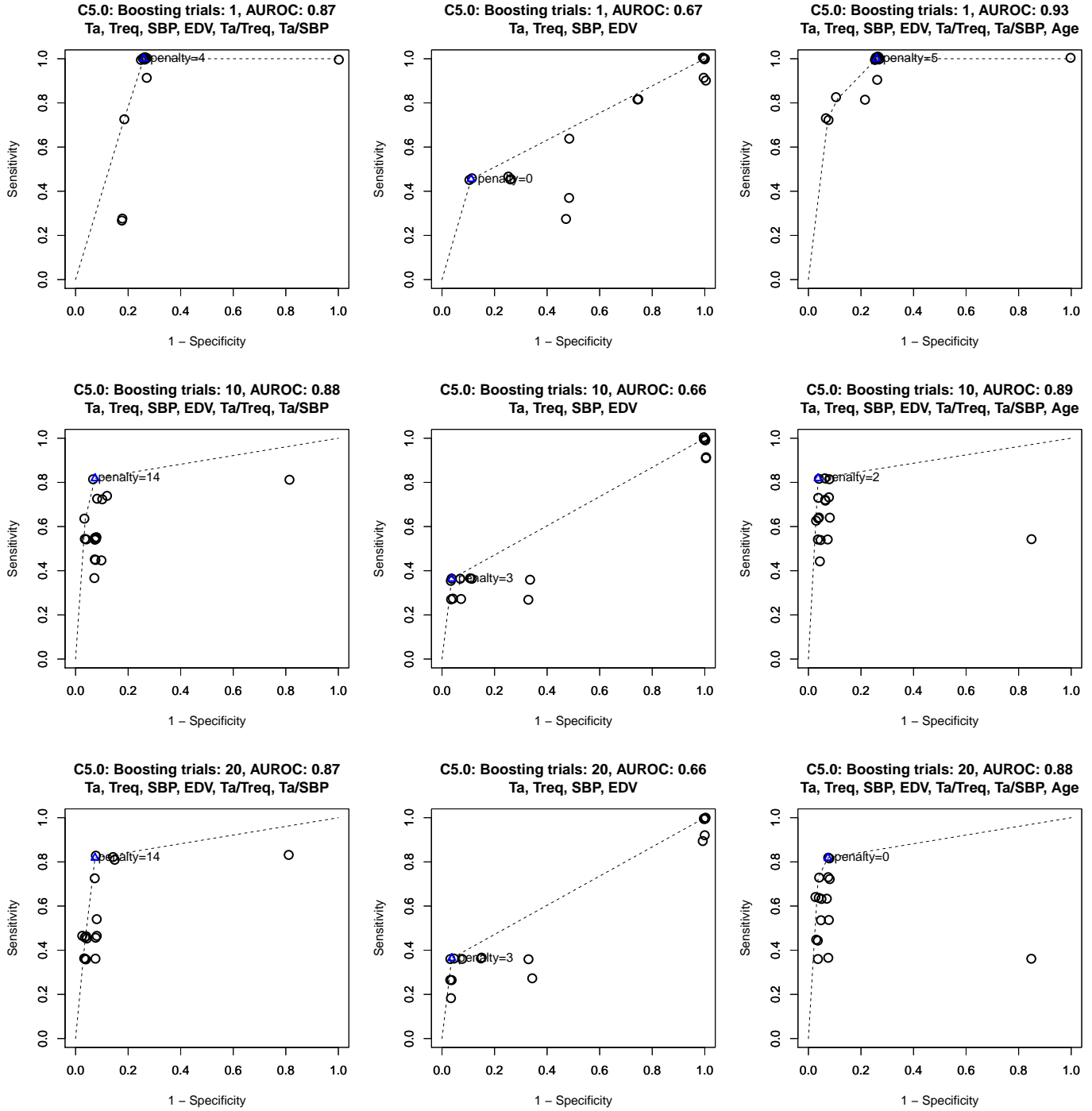


Figure 4: **C5.0 Decision Tree configurations:** Different settings for boosting trials and the cost penalty for three different data setups displayed in each column. The upper row shows ROC plots that lack boosting, and the lower row has a boosting of 10 trials. Each symbol in the ROC plots corresponds to cost penalty setting. The left top most point is marked with a blue triangle. LOOCV was used for each setting of boosting and error to estimate sensitivity and specificity.

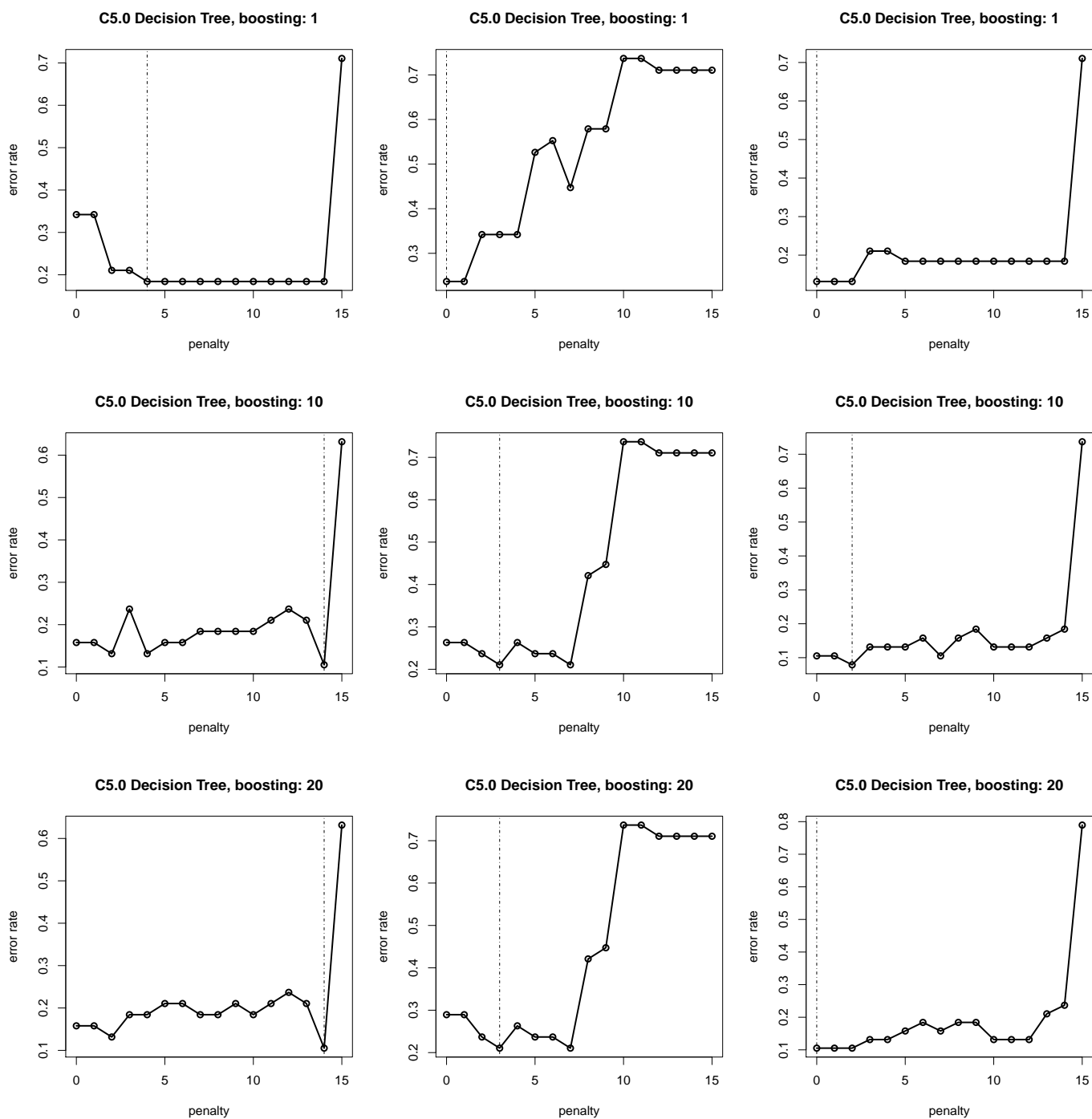


Figure 5: Mean squared error for C5.0 Decision Trees with varying error penalty.

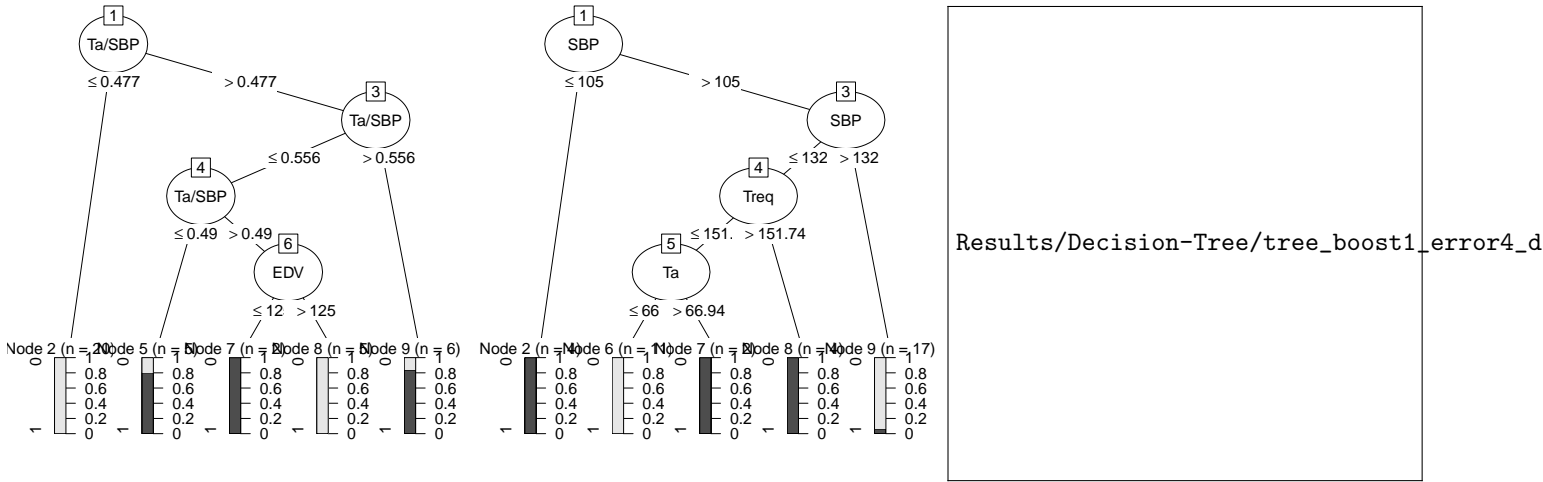


Figure 6: **C5.0 Decision Trees with lack of boosting:** Trees plotted for the three different data sets corresponding to the one shown in Fig. 8.



Figure 7: **Posterior probabilities plots of Gaussian Process - Automatic Relevance Determination (ARD):** The two plots show the posterior probabilities mesh given the most important features estimated by GP-ARD from Fig. 2, left column (data set 1): $Ta/Tref$, $Tref$, and SBP .

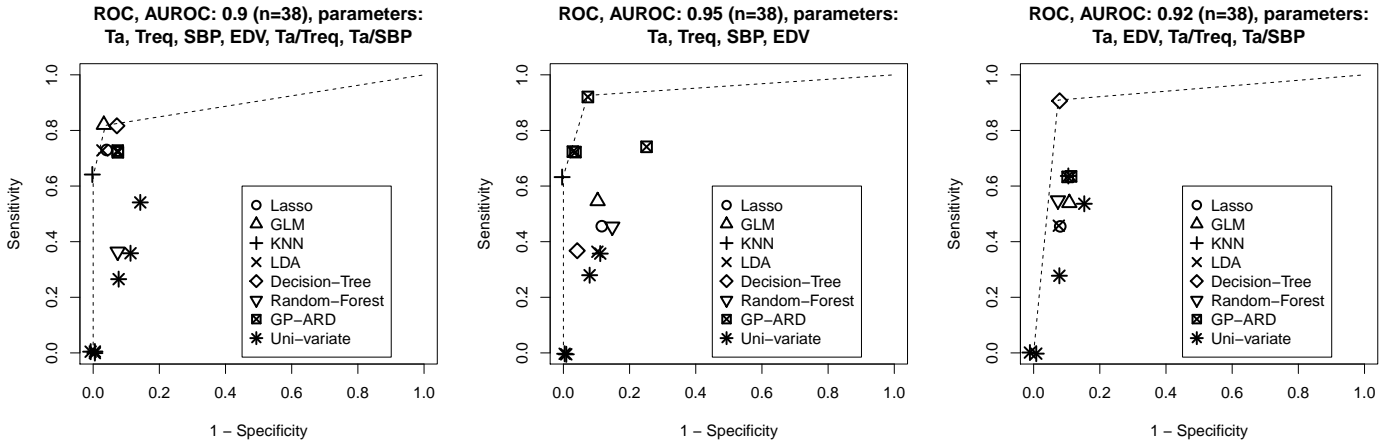


Figure 8: **Overall classification performance given three different data setups:** The area under the ROC curve (AUROC) is calculated from the area under the convex hull that connects the best performing methods. The specificity and sensitivity score for each method was estimated with LOOCV. The left plot uses the features from data set D_1 , the middle plot shows the performance for data set D_2 , and the right plot for data set D_3 .