Comparison between Glmertree (Gtree), ctree (conditional inference tree) and CART. The following uses sim\_3. I use a training data set with n=100 and a test data set with n=30. The error is mean square error. They all use default parameters

The following uses the code “Glmertree\_comparison”

1. I use all the features for ctree, CART, and in Gtree I use all the features as splitting variable and use t as regressor
   1. Time: Gtree(6s)>ctree(2s)>CART(0.4s)
   2. Variable selection: ctree (5)>Gtree(3)=CART(3)
   3. Error: ctree(87)>Gtree(79)>CART(73)
2. When I use true important features (1,2,3,301,302,303) for ctree, CART and in Gtree I use true important features for regressor and all the features for splitting:
   1. Time: Gtree(3s)>ctree(0.3)>CART(0)
   2. Error: CART=ctree(53)>>Gtree(18)
3. When I use part of true important features (2,302,303), Gtree still does much better than others in terms of error. Even when I only keep part of true important features and use some other irrelevant ones, Gtree is still better (or comparable) than other methods using the exactly correct features.

Also, in terms of prediction, Gtree is also better or comparable then RF and FF.

**Therefore, if Gtree happens to use all/some of important features, the prediction will improve a lot!**