

1 Advantages of DPDT over CART with statistical significance.

1.1 Interpretability

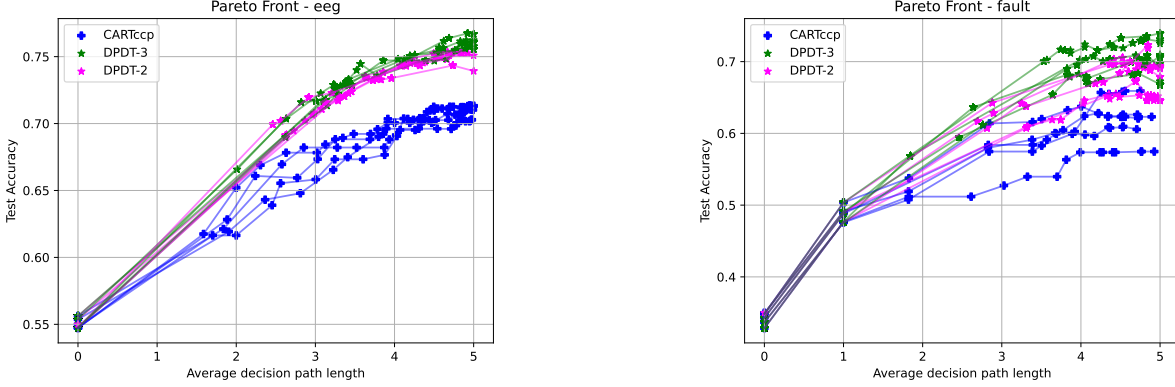


Figure 1: For different regularization values, we plot the decision path length against the test accuracy of DPDT and CART trees. The decision path length of a tree is the number of operation the tree performs to classify a single datum.

1.2 Feature cost

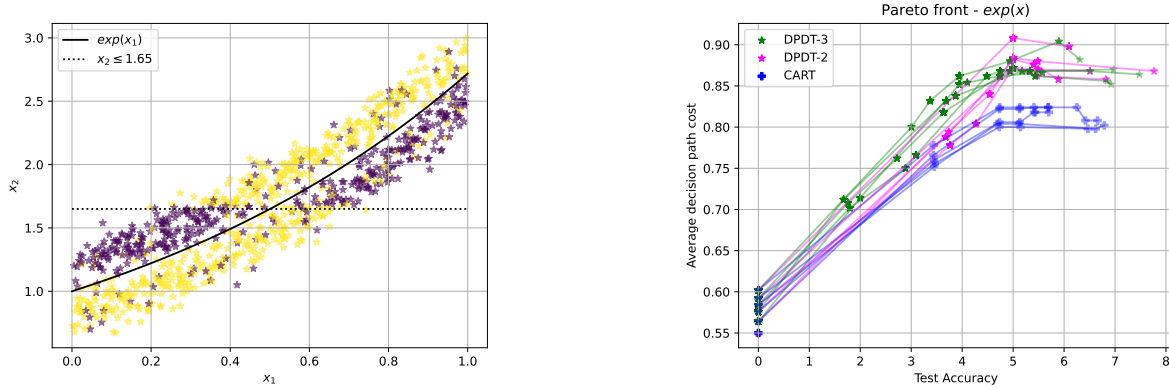


Figure 2: The feature cost problem in supervised classification is a critical issue in many practical applications where the cost of data acquisition is non-negligible. By carefully selecting features, incorporating cost-sensitive methods, and dynamically acquiring features, one can develop models that achieve a balance between performance and cost, making them more viable for real-world deployment. On the left, we plot a toy problem with 2 classes. Each datum has two features x_1, x_2 . We augment those features with costly ones $x_2 - \exp(x_1)$, $x_2 - x_1$. On the right, we plot the the decision path cost against the test accuracy of trees from DPDT and CART.