

Score Standard Deviation:

* **Dataset 1:** 0.029
* **Dataset 2:** 0.013
* **Dataset 3:** 0.187
* **Dataset 4:** 0.011
* **Dataset 5:** 0.004

**HoldOut Vs k-Fold CV:** in holdout approach, the complexity of training and testing Dataset may differ with will lead to bad performance on the dataset. Including that, we Split the dataset which results in less data for training. Whereas In K-Fold CV, Every Data point belongs to training and testing dataset. It increases the training data Volume and It helps in observing over all model performance.

**10x10 Fold CV Entropy without pruning:** For the datasets I chose, this model setting Preforms the best. As it gives approximately high accuracy as compare to the other model settings.

for the HoldOut Approach, error may be caused by the complexity difference between training and testing dataset. And pruning of Decision Tree also effects most of the datasets positively with slight increase in the accuracy. where as the Pruning in Dataset 3, decreases the accuracy more than half. It may be because of detailed features and complex relation ship.