

Variance = (E^2)/n

As number of samples increases variance decreases

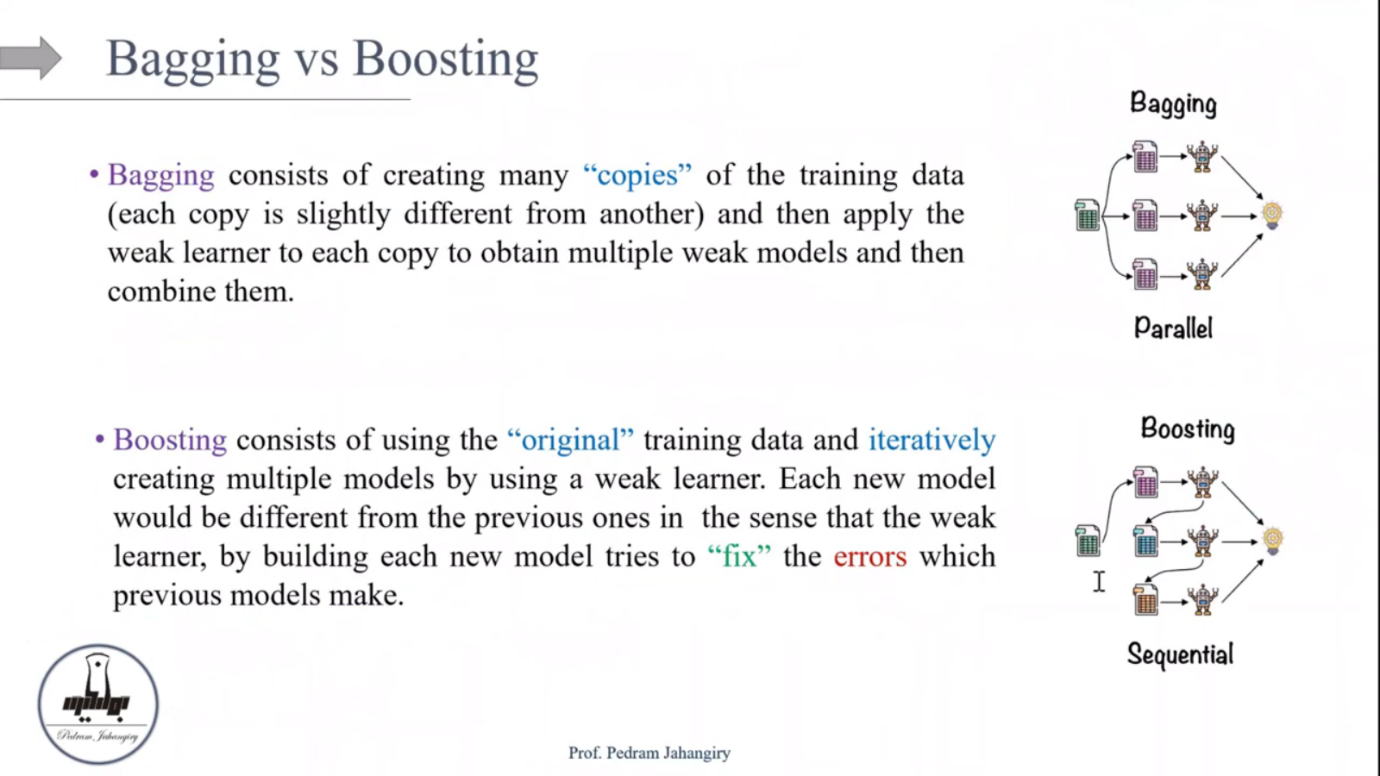
 **Decision Trees and Variance:**

* Individual decision trees are known for their high variance. This means that small changes in the training data can lead to significantly different tree structures and, consequently, different predictions.
* They tend to overfit the training data, capturing noise and specific patterns that might not generalize well to unseen data.

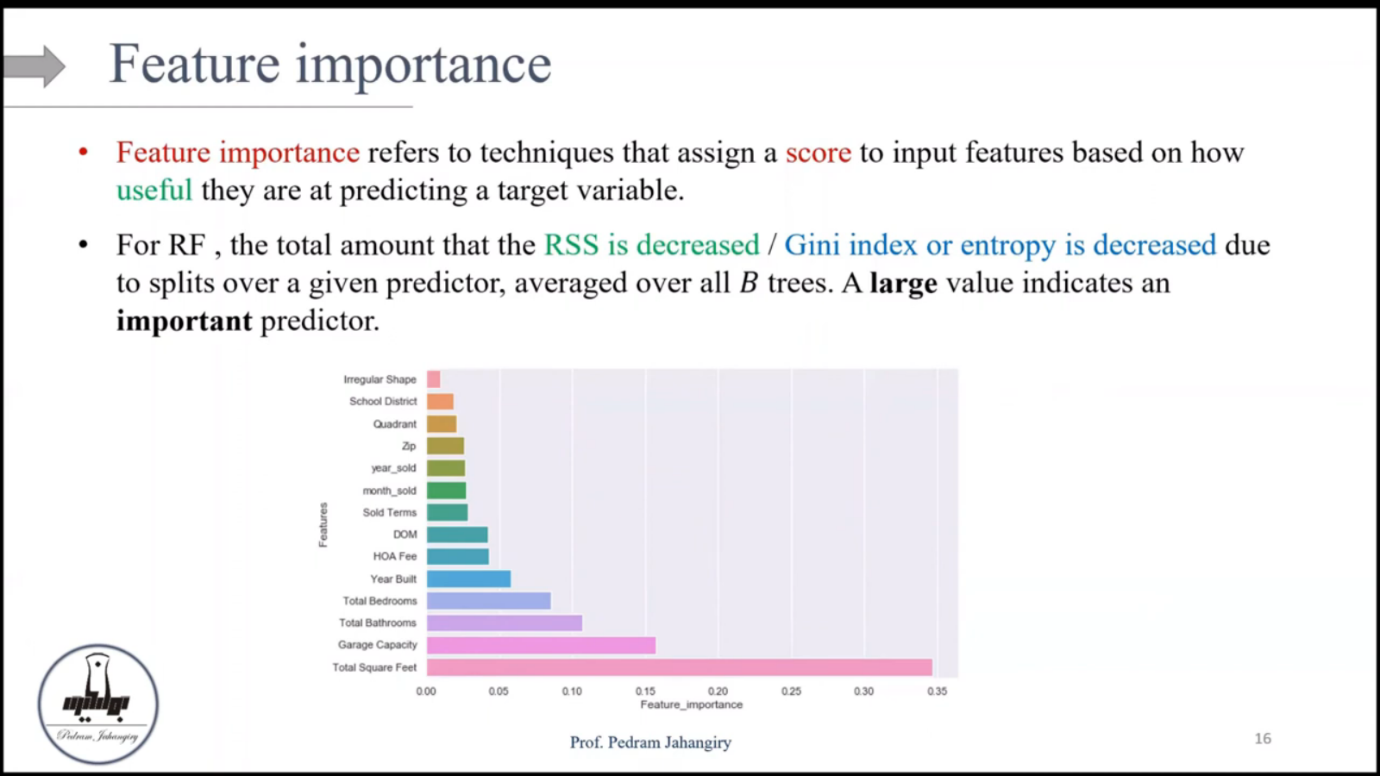
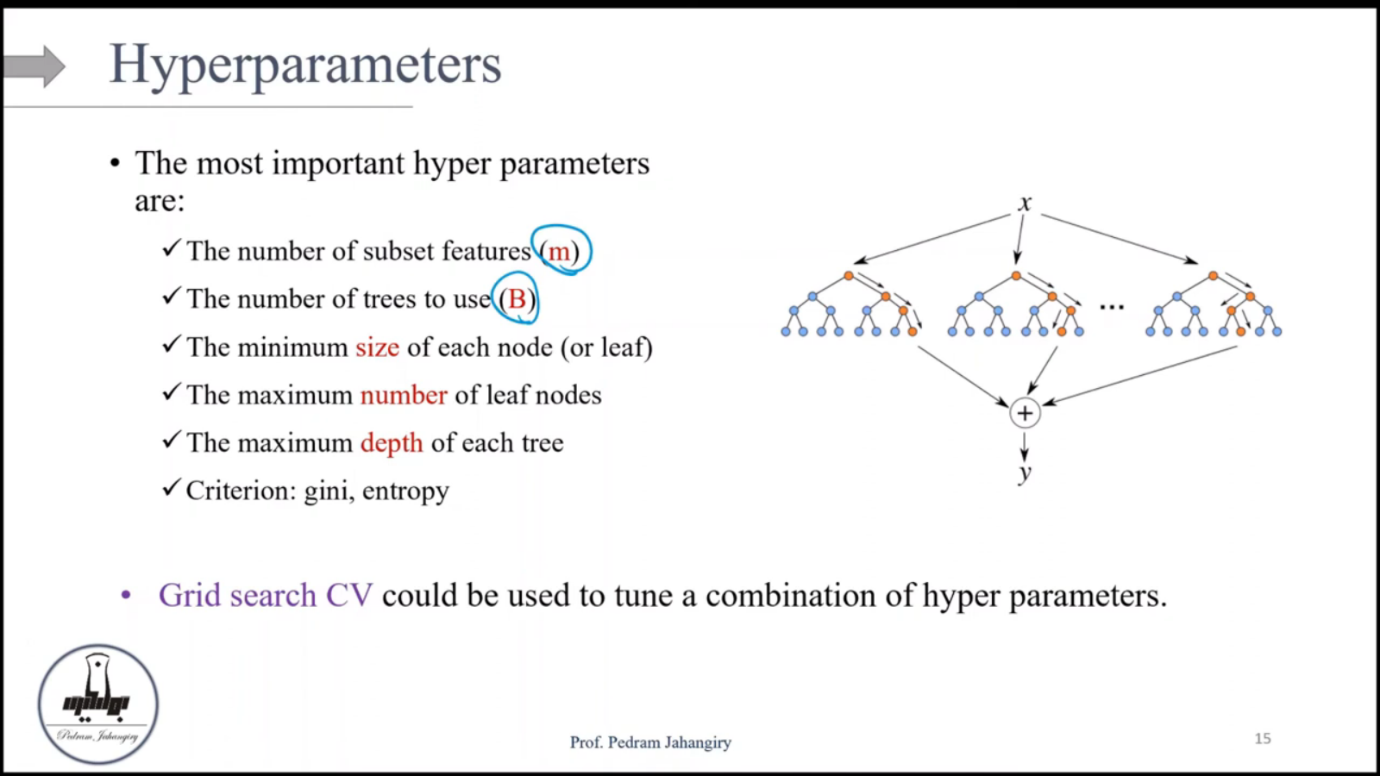
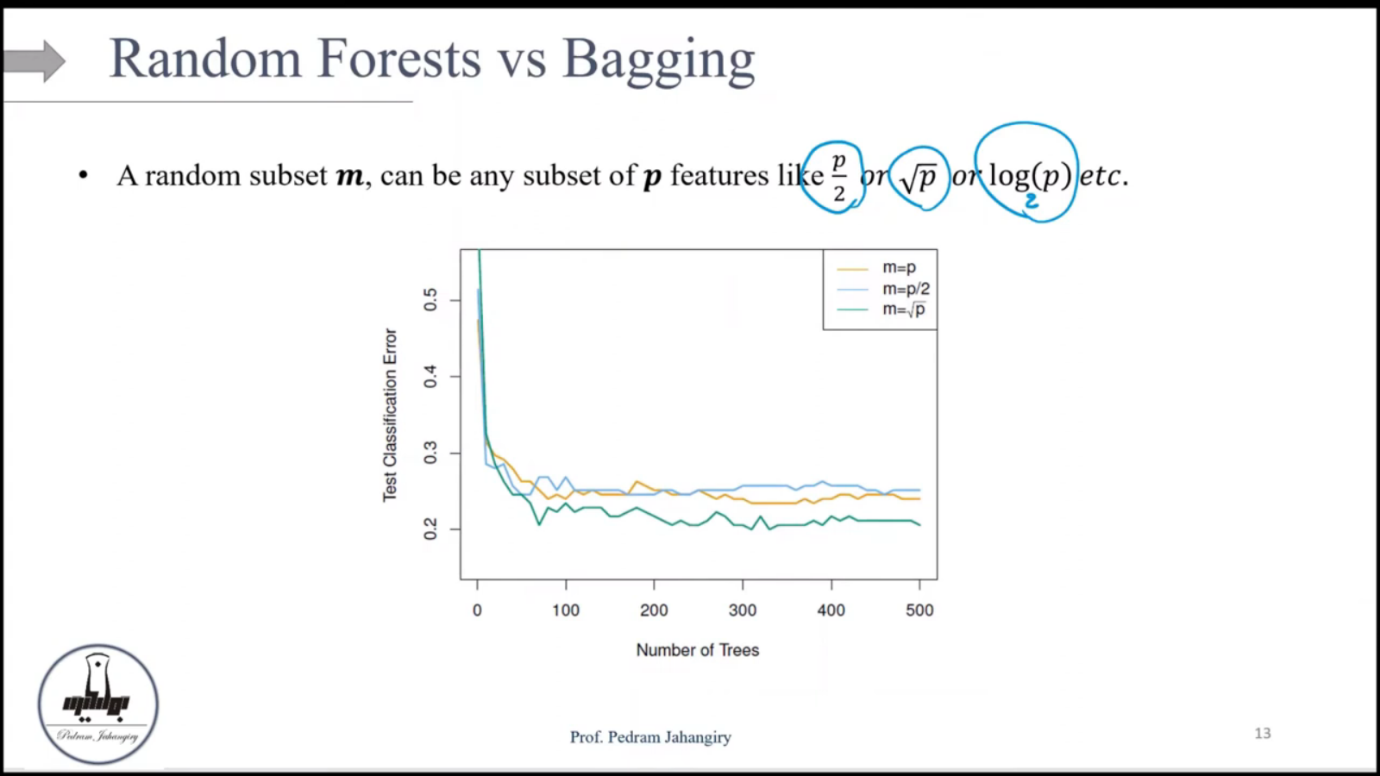
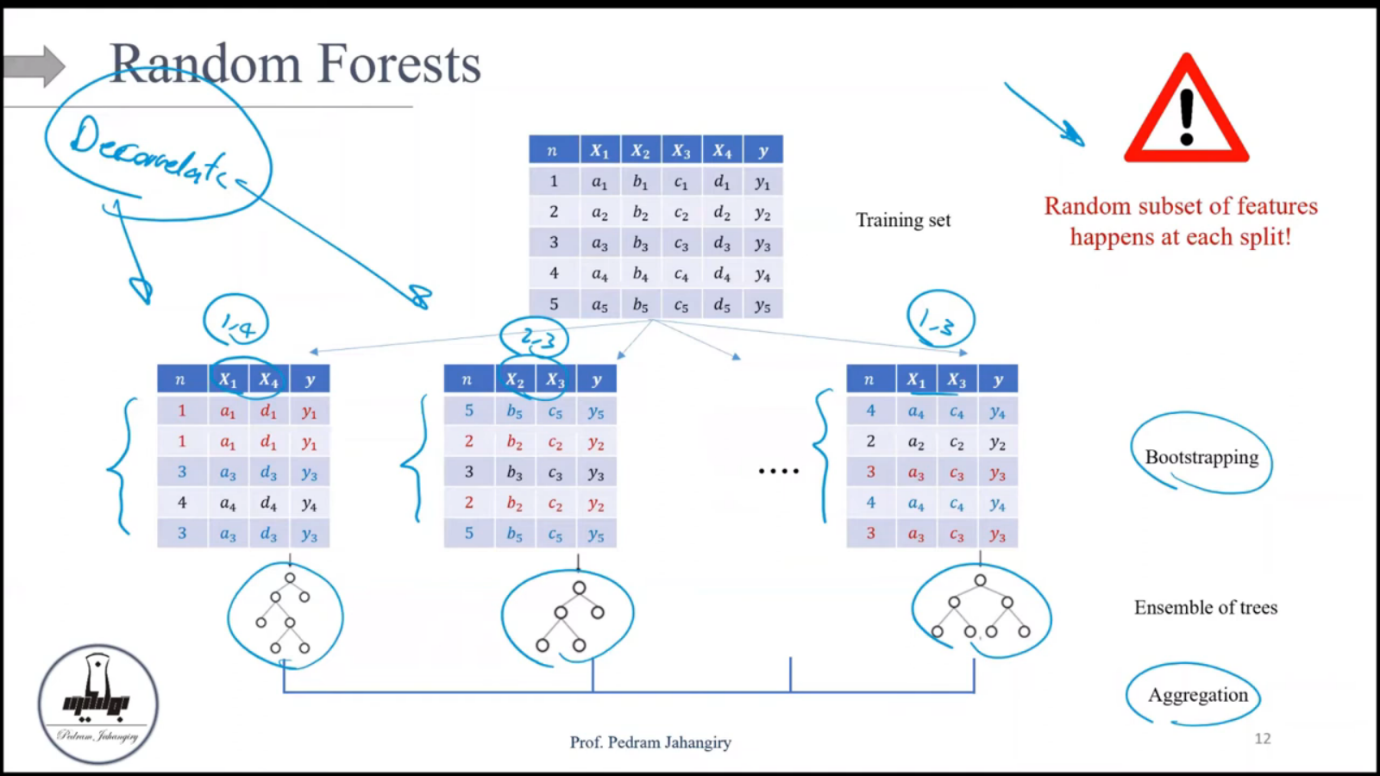
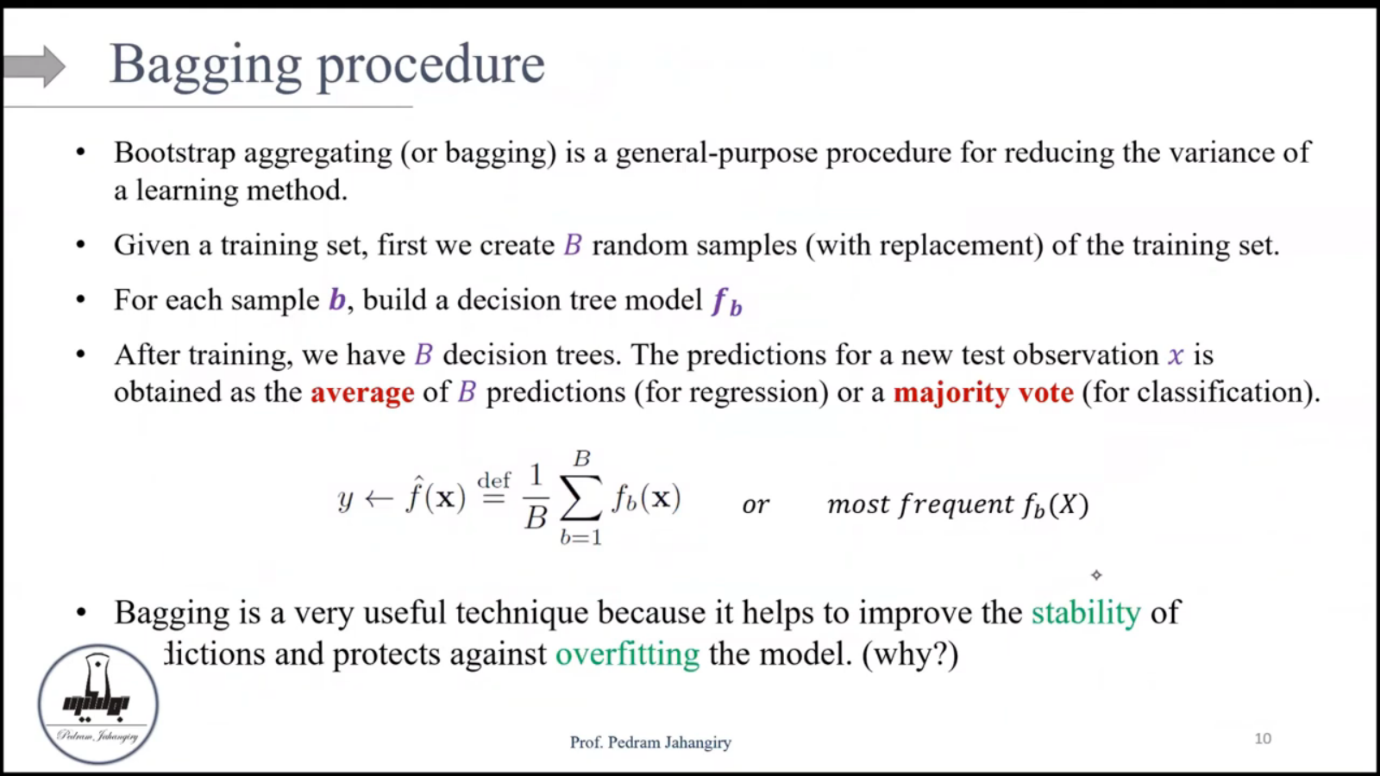
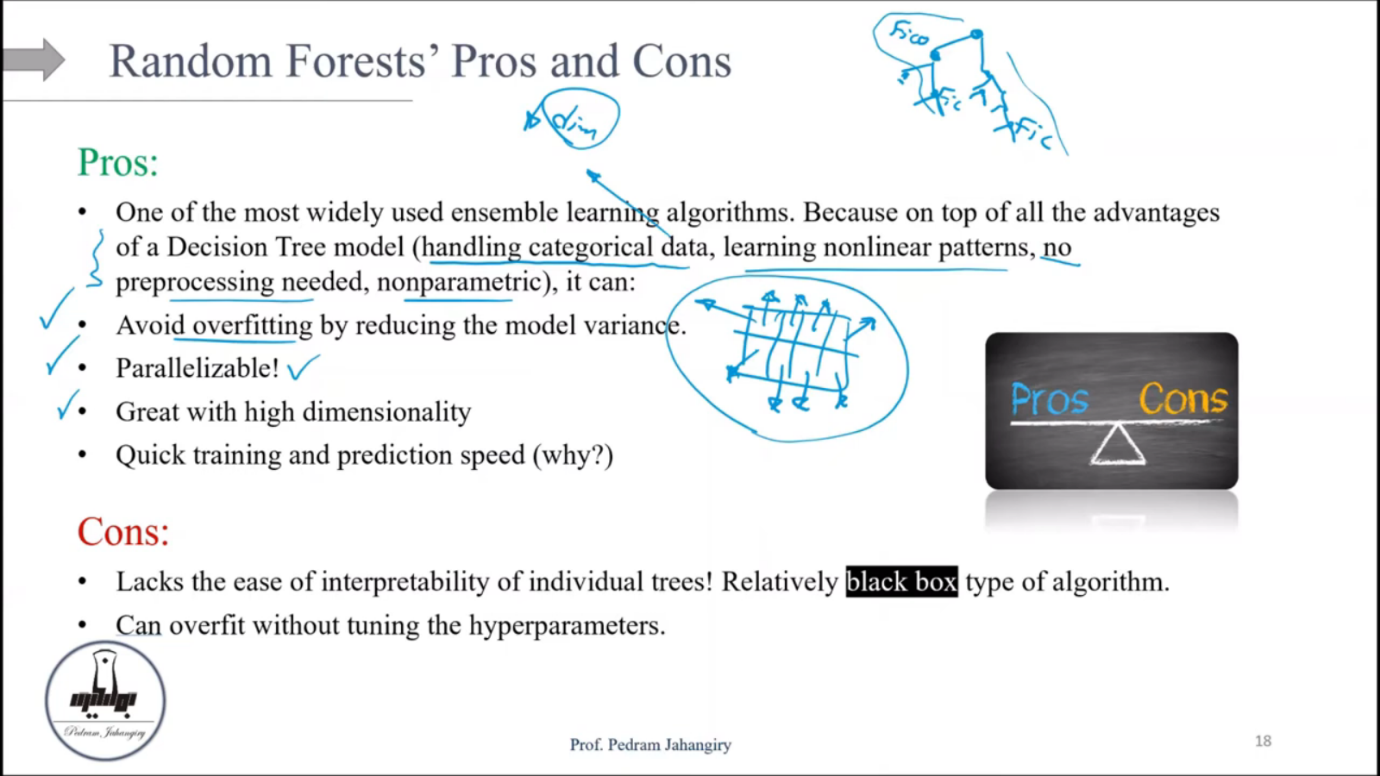
 **Random Forests and Variance Reduction:**

* A random forest creates an ensemble of many decision trees, each trained on a different subset of the training data (bootstrapping) and using a random subset of features at each split.
* By averaging (for regression) or voting (for classification) the predictions of these individual trees, the random forest smooths out the individual trees' errors and reduces the overall variance.
* Essentially, the random forest leverages the "wisdom of the crowd" principle.

The individual trees may be noisy and have high variance, but their combined prediction is more stable and less prone to overfitting.



Bootstrap: sampling from a dataset with replacement

Fast training is because they train on subset of features and not on full features  
