

# Overfitting

ECE 4200

# Learning

From training set  $S$

Learn a model that **hopefully** models the true process of how the features and labels are associated

This association is **probabilistic**.

# Decision Tree construction

Add nodes to tree

At each node, **hopefully** we learn something about the underlying truth

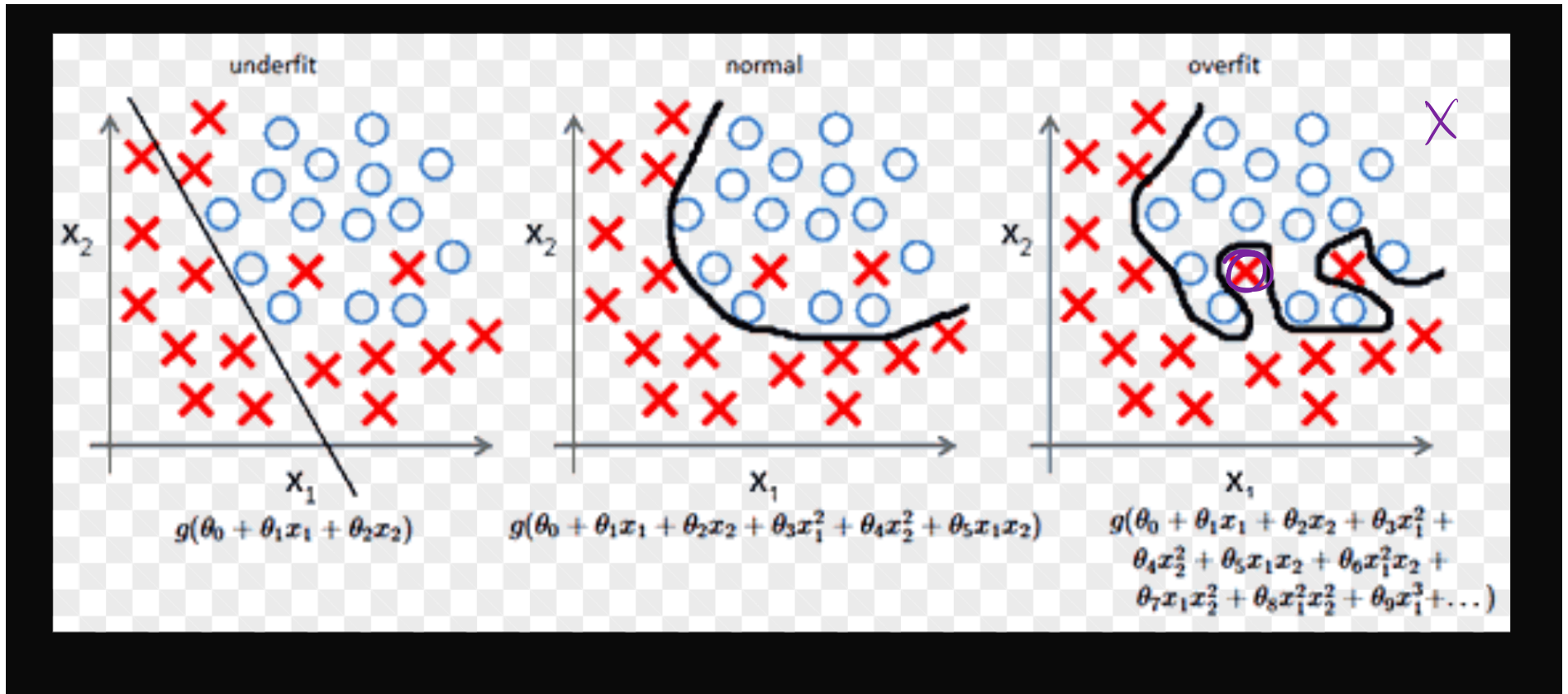
STOP when all examples are correctly classified



Are we really learning at each level??

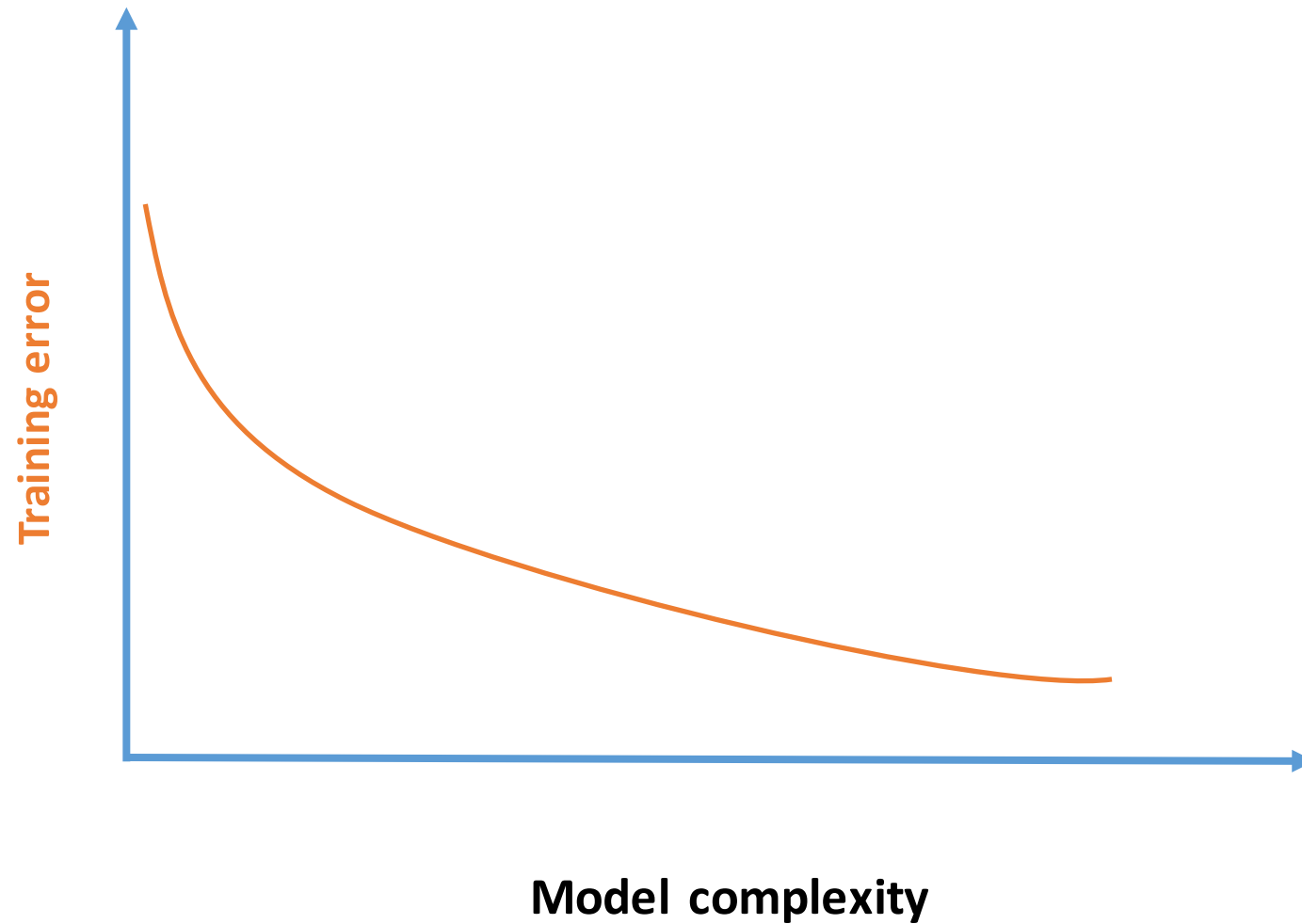
At some point, we stop learning about the underlying process, and try to **learn the noise**

# Example:

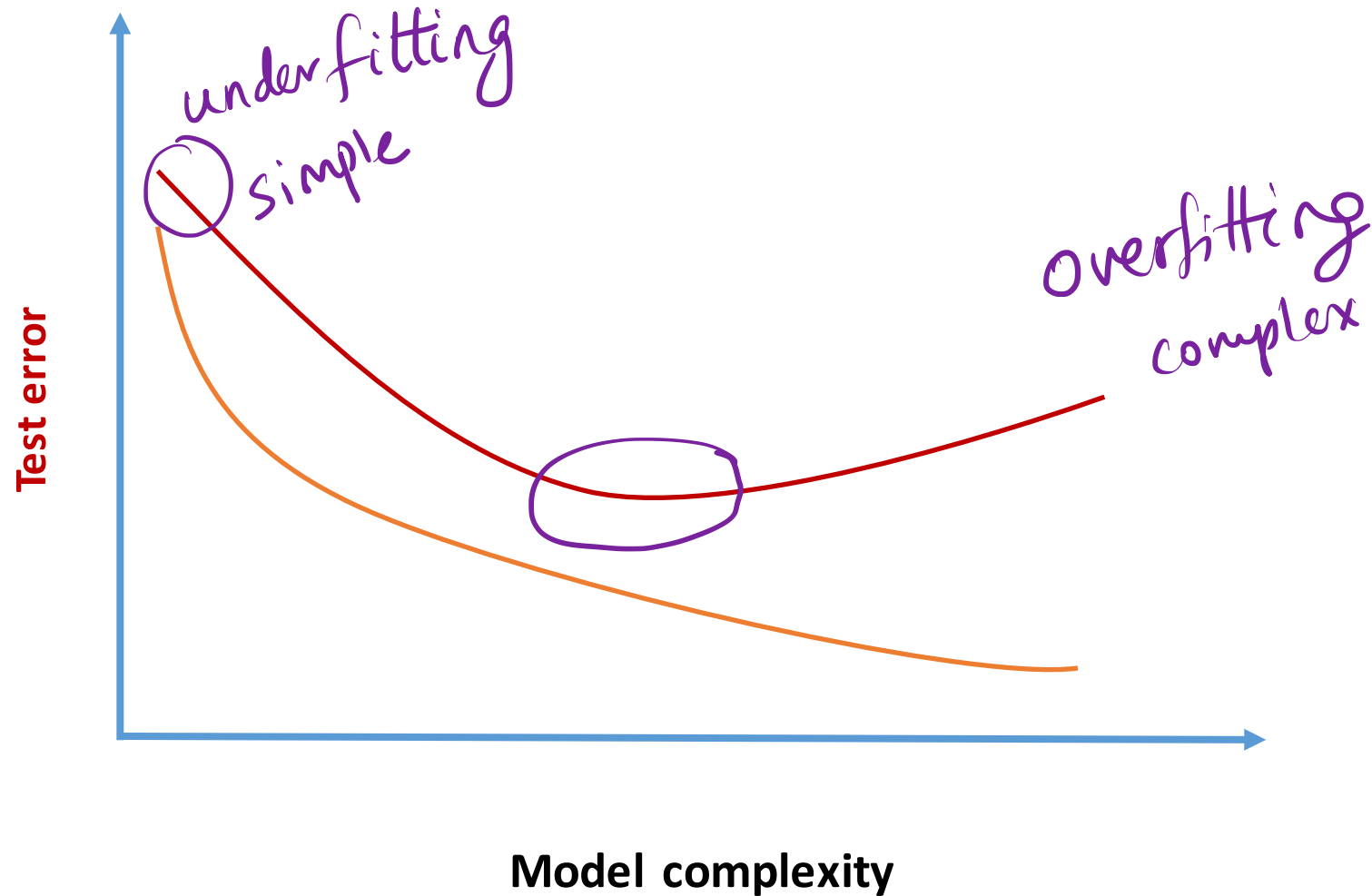


Read WIKI page for overfitting

# What is Overfitting?



# What is Overfitting?



# What is Overfitting?

As you make your model/algorithm more and more complex at some point you **stop** learning about the underlying process, but rather are left trying to learn the **noise**.

More when we do bias variance trade-offs.

# Decision Tree stopping rule

STOP when all examples are correctly classified

Decision trees are very prone to overfitting.

One **noisy** example can make the tree big

((sunny, hot, normal, strong), NO) – how does tree change?



# Overfitting in decision trees?

- STOP growing the tree when IG gets small.
- Grow the tree and post-prune.

# Validation set

Divide the training data into two parts.

Training set and validation set

Make model more complex (train) as long as performance on validation set improves

# Post-pruning

- Obtain the tree, then go and delete nodes that improve the performance on the validation set.
- Rule post pruning.
- Time consuming.

Chapter 3 Mitchell for details