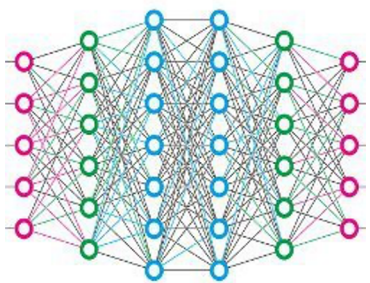
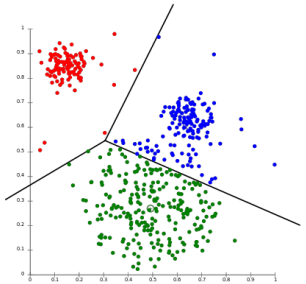
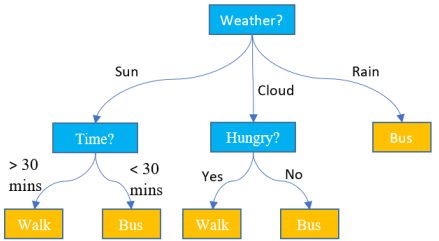
Welcome to CS 445
Introduction to Machine Learning
Model Evaluation, Selection, and Validation

Instructor: Dr. Kevin Molloy







Announcements

- Quiz 2 was due Wednesday by 11:59 pm
- PA 1 is due next Friday (Sept 18th).
- First Homework assignment posted today

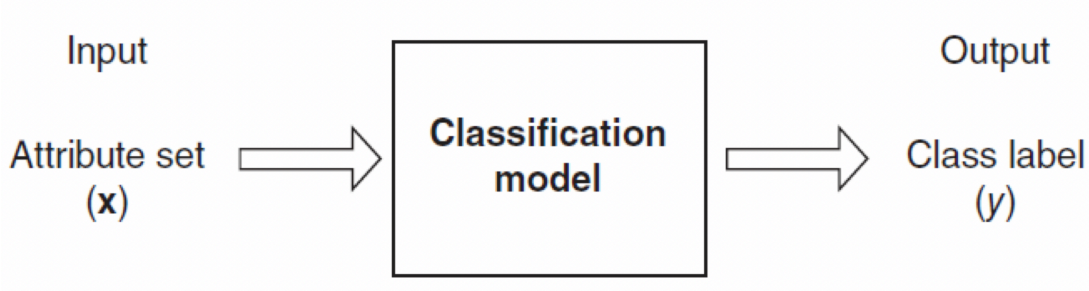
Learning Objectives from Last time

- Define and discuss the differences between model evaluation and model selection
- Define the term **hyperparameter** and make the distinction between it and a **parameter**
- Utilize **training**, **validation**, and **test** sets to design experiments to better characterize your model's performance
- Define and utilize k-fold cross validation

Plan for Today

- Review Using Decision Trees for Regression
- Selecting Hyperparameters
- Model Evaluation Review
- PA1 -- Tree Split warmups

Selecting which Model to Use

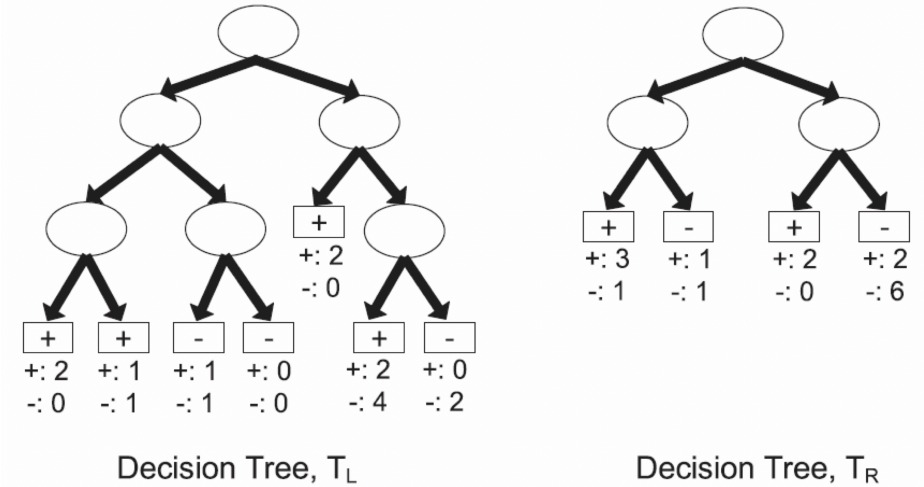


Model Selection

Goal:

Model Evaluation:

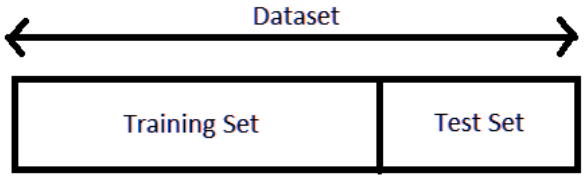
Goal:



Model Selection

The task of finding the model that maximizes the performance of **learning** task is called **model selection**. This involves tuning **hyper-parameters**

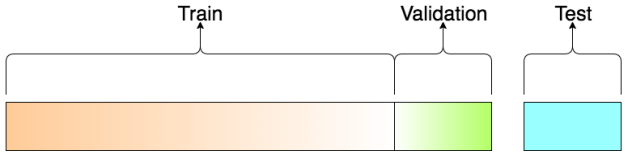
Why can't we just train our model on the training dataset and pick the hyperparameters that do the best on the test dataset?



Max Leaves	Accuracy
6	0.40
7	0.35
8	0.38

1) Best parameter is dataset dependent.
2) Being able to select the hyperparameter that makes the model like the best will **not result** in an unbiased estimate of its performance.

Dataset Purposes



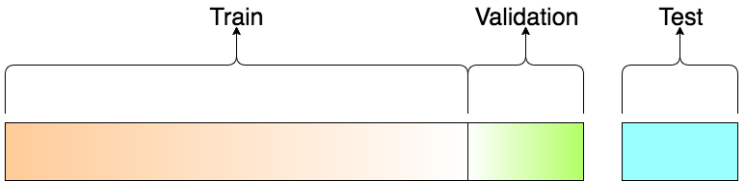
Goal is to minimize **generalization error**.

Model selection

1. For each combination of hyperparameters:
    a) Train the model with this set of hyperparameters on training
    b) Evaluate the model on the validation set
2. Pick the best performing hyperparameters.
3. Retrain model using best performing hyperparameters on the concatenated training and validation datasets (larger == better)
4. Perform model evaluation on the test dataset

Cross Validation

For small datasets, breaking the data up into all these groups is not idea.



Run 1	*D.tr*	*D.tr*	*D.val*
Run 2	*D.tr*	*D.val*	*D.tr*
Run 3	*D.val*	*D.tr*	*D.tr*

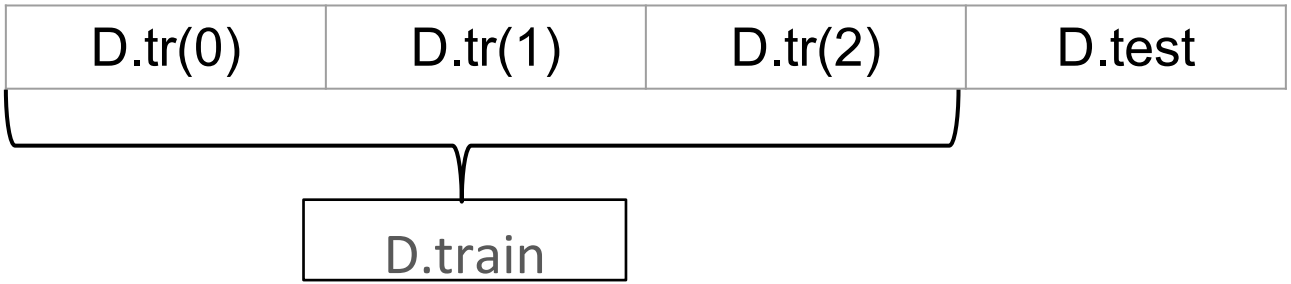
Idea: Divide training set into *k* groups (3 shown here) and perform *k* evaluations (where the validation set changes each time). Take the **average** of the *k* runs as the performance measure for the hyper-parameters being evaluated.

Quiz Question

Assuming 4-fold cross-validation and data D of size 100, what is the size of D.train(0)?

IDD 3.6.2

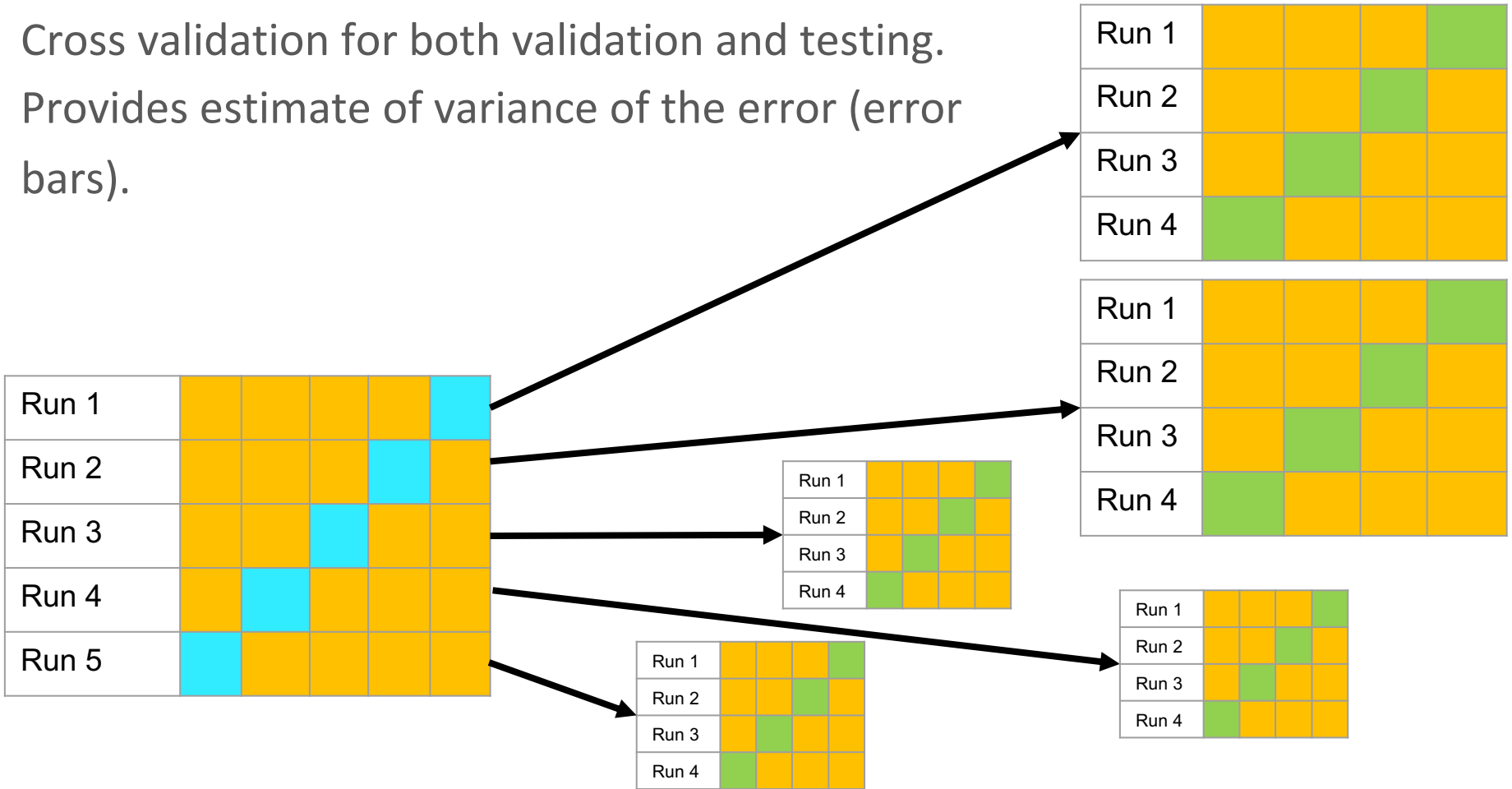
There is only 1 D.train consisting of {D.tr(0), D.tr(1), D.tr(2)}. Thus, the answer was 75, but I went back and also accepted 25.



Nested Cross Validation

Cross validation for both validation and testing.

Provides estimate of variance of the error (error bars).



PA 1 Warm Up

- All features will be numeric. Does this limit what your model building program can do?

- A python *generator* is supplied in the code template that will generate all possible split points for a node.

- Today, we will work on evaluating and selecting the best split point

Split Point Object

Represents one way to split the data in the node.

var name	type	description
dim	int	the dimension/feature used in this split (0..d-1 where d is the number of columns in the feature matrix)
pos	float	the value used to split the data
X_left	ndarray	all X entries that are <= split point
y_left	ndarray	all y labels corresponding to X_left
X_right	ndarray	all X entries that are > split point
y_right	ndarray	all y labels corresponding to X_right

Rate each Split Point

Work in groups of 2 or 3.

1. Download tree_warmup.py from the website.
2. Write and test the impurity function
3. Augment split_demo to:
   a) Prints the weighted impurity
   b) Prints the info gain
   c) At the end of the loop, prints the best split point (identified by dimension and the value (pos)

Submit to Canvas when complete.

Numpy Lab

- Work on the lab in groups of 2 to 3.
- Complete the lab and turn in the ipynb into Canvas.
- Be ready to discuss a few solutions when we return

For Next time

Homework:

- Normalization Homework