

# Training/Testing and Regularization

Il-Chul Moon  
Dept. of Industrial and Systems Engineering  
KAIST

[icmoon@kaist.ac.kr](mailto:icmoon@kaist.ac.kr)

# Weekly Objectives

- Understand the concept of bias and variance
  - Know the concept of over-fitting and under-fitting
  - Able to segment two sources, bias and variance, of error
- Understand the bias and variance trade-off
  - Understand the concept of Occam's razor
  - Able to perform cross-validation
  - Know various performance metrics for supervised machine learning
- Understand the concept of regularization
  - Know how to apply regularization to
    - Linear regression
    - Logistic regression
    - Support vector machine

# Performance Measure of ML

- Is it possible to calculate the bias and the variance?
  - We don't know the target function,  $f(X)$ !
  - We can't compute the average hypothesis,  $\bar{g}(x)$ !
- Therefore, we can't use the bias and the variance as the performance measures.
- Then, what measures to use?
  - Accuracy=  $(TP+FN) / (TP+FP+FN+TN)$
  - Precision and Recall
  - F-Measure
  - ROC curve

|                 |          | Actual Value   |                |
|-----------------|----------|----------------|----------------|
|                 |          | True           | False          |
| Estimated Value | Positive | True Positive  | False Positive |
|                 | Negative | False Negative | True Negative  |

# Precision and Recall

- Consider the two cases
  - Building a classifier
    - Spam filter
    - CRM

|                 |          | Actual Value   |                |
|-----------------|----------|----------------|----------------|
|                 |          | True           | False          |
| Estimated Value | Positive | True Positive  | False Positive |
|                 | Negative | False Negative | True Negative  |

- Goals are slightly different
  - Spam filter: classifying spam
    - Safety is first. You don't want to throw out valid emails estimated as spams
    - Reducing the FP is the priority
  - CRM: classifying VIP customer
    - Reaching out is first. You don't want to miss any VIP customers as ordinary ones
    - Reducing the FN is the priority
- Precision** =  $TP / (TP + FP)$
- Recall** =  $TP / (TP + FN)$
- Then, which metrics to use in each case?

# F-Measure

- Precision and recall are popular metrics, but it has problems in the applications
  - The most safest spam filter == always say 'no spam'
  - The most reaching-out customer filter == always say 'VIP'
- We need a measure that balances the precision and the recall performance
- F-Measure is the derived metric from the precision and the recall
  - $F_b\text{-Measure} = (1+b^2) * (\text{Precision} * \text{Recall}) / (b^2 * \text{Precision} + \text{Recall})$
  - $F_1\text{-Measure} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$
  - $F_{0.5}$  and  $F_2$  are also used.
    - $F_2$  emphasizes recall
    - $F_{0.5}$  emphasizes precision