



COMPSCI 389

Introduction to Machine Learning

Days: Tu/Th. **Time:** 2:30 – 3:45 **Building:** Morrill 2 **Room:** 222

Topic 4.1: Model Evaluation Summary

Prof. Philip S. Thomas (pthomas@cs.umass.edu)

Model Evaluation

- If a model is evaluated using the same data used to train the model, it often results in **over-estimates** of the model's performance.
 - This answers the question “How well does our model predict outcomes for data it has already seen?”
 - We want to answer the question “How well can our model predict outcomes for new, unseen data?”
 - We saw that the *nearest neighbor* (NN) algorithm has zero error on the training data!
- We therefore split the data into two sets, a training set and a testing set.
 - The model is trained (fit) using the training data.
 - We evaluate the model's performance using the testing data.
 - We shuffle the data before splitting it into training and testing sets.

Evaluation Metrics (Regression)

- Mean Error: $\frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i$
 - Rarely what you want.
 - Allows positive and negative errors to cancel each other out.
- Mean Squared Error (MSE): $\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$
 - Very common choice.
 - Gives a higher weight to larger errors, making it sensitive to outliers. It's useful when large errors are particularly undesirable.
- Root Mean Squared Error (RMSE): $\sqrt{\text{MSE}}$
 - Has the same units as the target variable (unlike MSE).

Evaluation Metrics (Regression, cont.)

- Mean Absolute Error (MAE): $\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$
 - Like MSE, but with less emphasis on outliers.
- R-squared (R^2): $1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$, where $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$.
 - Also called the *coefficient of determination*.
 - Indicates the proportion of the variance of the dependent variable (labels) that is predictable from the independent variables (predictions).
 - Larger is better (maximum possible is one).

Intermission

- Class will resume in 5 minutes.
- Feel free to:
 - Stand up and stretch.
 - Leave the room.
 - Talk to those around you.
 - **Write a question on a notecard and add it to the stack at the front of the room.**



