University of Roehampton London

# Machine Learning

Dr Changjiang He, Dr Kuo-Ming Chao
Computer Science| School of Art
University of Roehampton

University of Roehampton London

# Lesson 4.1

# Evaluation of Prediction

# Review of Supervised Learning

University of Roehampton London

- Linear Regression
- Lasso
- Logistic Regression
- Random Forest
- Neural Network (next week)
- …

# Evaluation of Prediction: RMSE

University of Roehampton London

**Root Mean Square Error**

RMSE is a popular formula to measure the error rate of a regression model. However, we can only compare between models whose errors we can measure in the same units

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (p_i - a_i)^2}{n}}$$

# Evaluation of Prediction: RSE

University of Roehampton London

**Relative Square Error**

Unlike RMSE, the relative squared error (RSE) can be compared between models whose errors we can measure in different units

$$RSE = \frac{\sum_{i=1}^{n} (p_i - a_i)^2}{\sum_{i=1}^{n} (\bar{a} - a_i)^2}$$

# Evaluation of Prediction: RAE

University of Roehampton London

**Relative Absolute Error**

Like RSE, the relative absolute error (RAE) can be compared between models whose errors are measured in the different units.

$$RAE = \frac{\sum_{i=1}^{n} |p_i - a_i|}{\sum_{i=1}^{n} |\overline{a} - a_i|}$$

# Evaluation of Prediction: MAE

University of Roehampton London

**Mean Absolute Error**

Mean Absolute Error is the average of the difference between the Original Values and the Predicted Values. It gives us the measure of how far the predictions were from the actual output. However, they don't give us any idea of the direction of the error i.e. whether we are under predicting the data or over predicting the data. Mathematically, it is represented as :

$$MeanAbsoluteError = \frac{1}{N}\sum_{j=1}^{N}|y_j - \hat{y}_j|$$

# Evaluation of Prediction: R2

University of Roehampton London

**Coefficient of Determination**

The coefficient of determination (**R2**) summarizes the explanatory power of the regression model and is computed from the sums-of-squares terms.

R2 describes the proportion of variance of the dependent variable explained by the regression model. If the regression model is "perfect", SSE is zero, and R2 is 1. If the regression model is a total failure, SSE is equal to SST, no variance is explained by regression, and R2 is zero.

# Evaluation of Prediction: R2

University of Roehampton London

$$\textit{Coefficient of Determination} \rightarrow \quad R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

$$\textit{Sum of Squares Total} \rightarrow \quad SST = \sum (y - \bar{y})^2$$

$$\textit{Sum of Squares Regression} \rightarrow \quad SSR = \sum (y' - \bar{y'})^2$$

$$\textit{Sum of Squares Error} \rightarrow \quad SSE = \sum (y - y')^2$$

# Evaluation of Prediction: Plot

University of Roehampton London

- These density plots show the predicted vs. actual values for the regression models to predict the the response variable.
- Note: the yellow/lightest areas represent the highest density of points