JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

685.621 Algorithms for Data Science

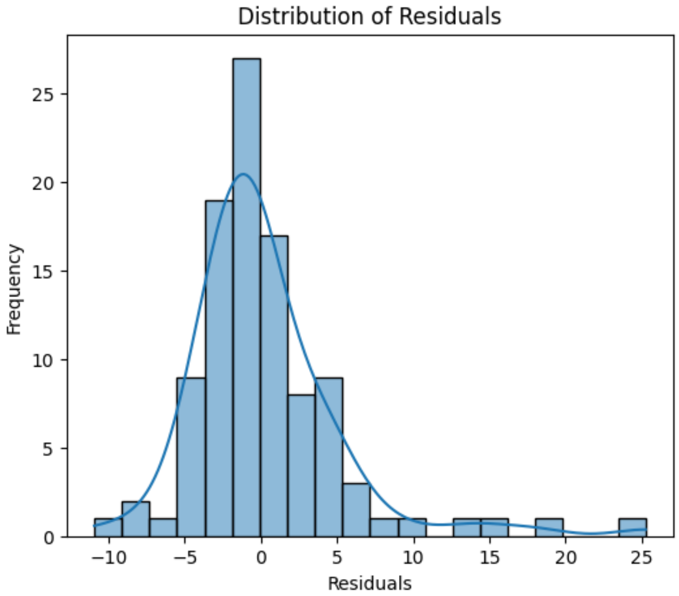
Supervised Learning Regression: Model Evaluation

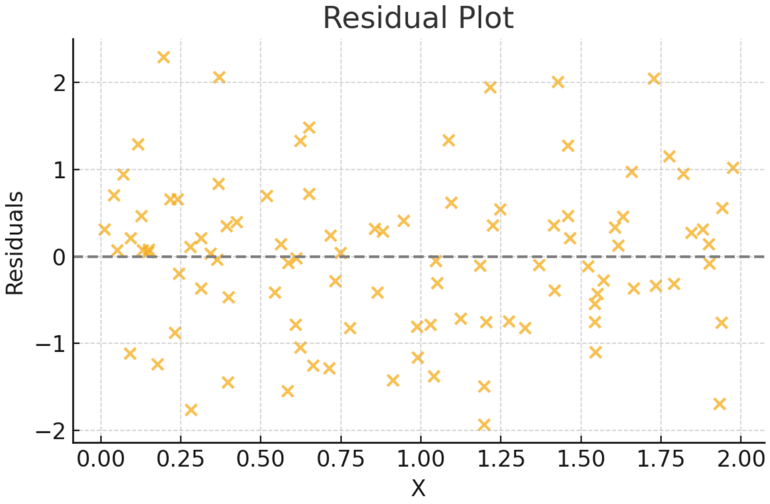
Why Do We Need Model Evaluation

- **Purpose of Evaluation**:
  - Determines how well the model generalizes to unseen data.
  - Helps compare different models objectively.
  - Identifies overfitting and underfitting.
  - Supports model selection and hyperparameter tuning by using consistent metrics.

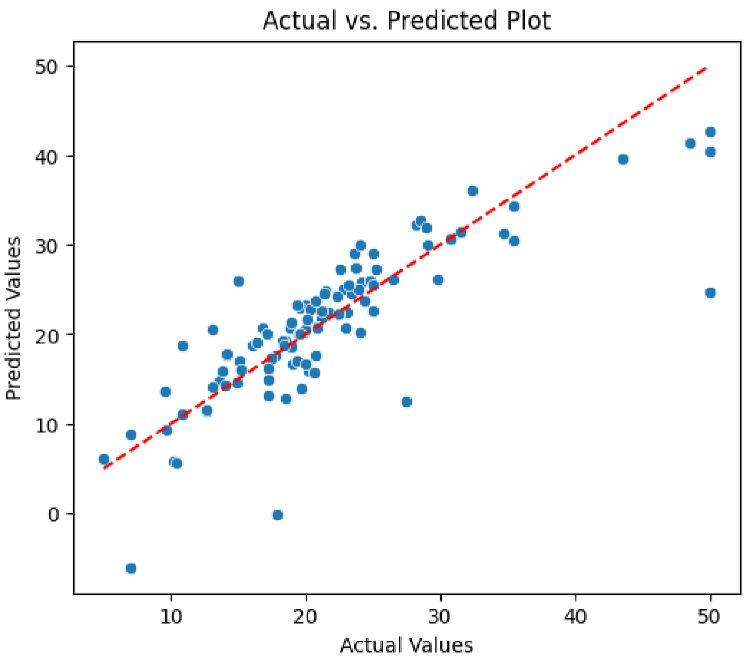
- **Two types of evaluation**:
  - **Training Performance**: How well the model fits the training data.
  - **Generalization Performance**: How well the model performs on unseen data.

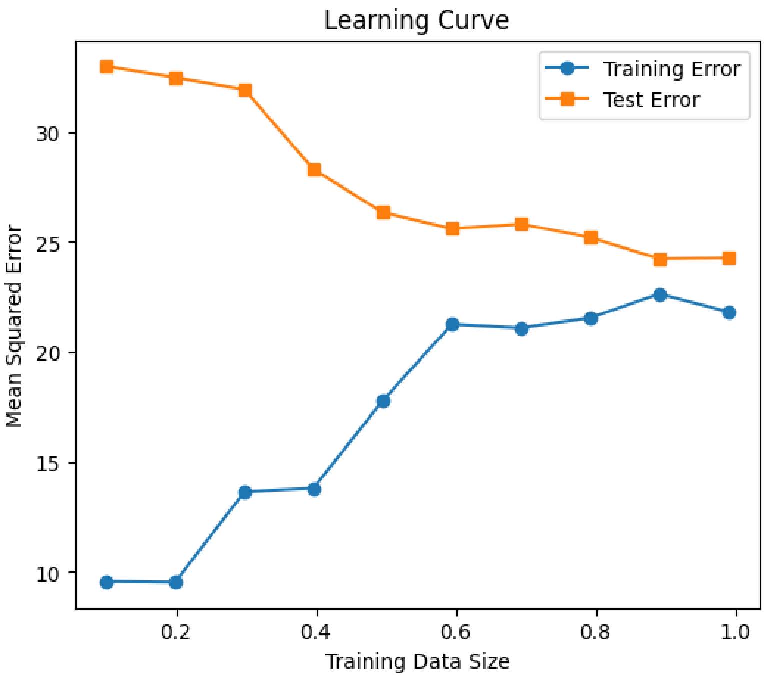
Plots for Evaluating Models





Plots for Evaluating Models





JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Metrics for Evaluating Models

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|$$

$$\text{Adjusted } R^2 = 1 - \frac{\left(1 - R^2\right)(n-1)}{n-p-1}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}$$

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

$$MSLE = \frac{1}{n}\sum_{i=1}^{n}\left(\log\left(1 + y_i\right) - \log\left(1 + \hat{y}_i\right)\right)^2$$

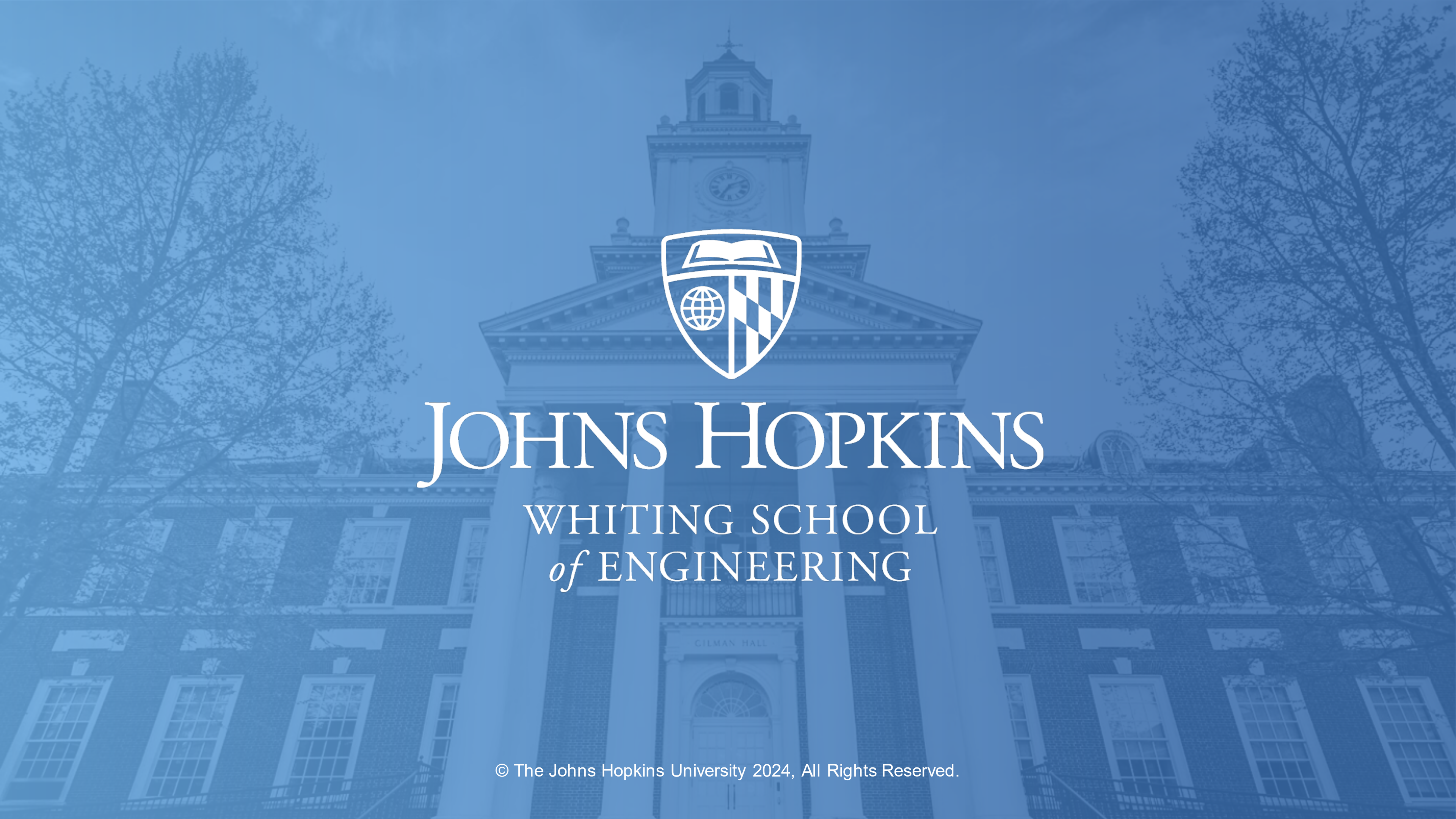
Choosing the Best Models

$$\mathrm{AIC} = 2k - 2\ln(\widehat{L})$$

$$\mathrm{BIC} = k\ln(n) - 2\ln(\widehat{L})$$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:              Head_size   R-squared:                       0.639
Model:                            OLS   Adj. R-squared:                  0.638
Method:                 Least Squares   F-statistic:                     416.5
Date:                Sun, 08 May 2022   Prob (F-statistic):           5.96e-54
Time:                        21:40:40   Log-Likelihood:                -1613.4
No. Observations:                 237   AIC:                             3231.
Df Residuals:                     235   BIC:                             3238.
Df Model:                           1
Covariance Type:            nonrobust
================================================================================
                   coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
Intercept      520.6101    153.215      3.398      0.001     218.759     822.461
Brain_weight     2.4269      0.119     20.409      0.000       2.193       2.661
==============================================================================
Omnibus:                        2.687   Durbin-Watson:                   1.726
Prob(Omnibus):                  0.261   Jarque-Bera (JB):                2.321
Skew:                           0.207   Prob(JB):                        0.313
Kurtosis:                       3.252   Cond. No.                     1.38e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.38e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```



JOHNS HOPKINS
WHITING SCHOOL of ENGINEERING