### CPSC 330 Lecture 10: Regression Metrics

Firas Moosvi (Slides adapted from Varada Kolhatkar)

#### **Recap: Ridge and RidgeCV**

- **Ridge Regression**: alpha hyperparameter controls model complexity.
- **RidgeCV**: Ridge regression with built-in cross-validation to find the optimal alpha.

### Recap: alpha hyperparameter

#### • Role of alpha:

- Controls model complexity
- Higher alpha: Simpler model, smaller coefficients.
- Lower alpha: Complex model, larger coefficients.

## Regression metrics: MSE, RMSE, MAPE

- Mean Squared Error (MSE): Average of the squares of the errors.
- Root Mean Squared Error (RMSE): Square root of MSE, same units as the target variable.
- Mean Absolute Percentage Error (MAPE): Average of the absolute percentage errors.

# Applying log transformation to the targets

- Suitable when the target has a wide range and spans several orders of magnitude
  - Example: counts data such as social media likes or price data
- Helps manage skewed data, making patterns more apparent and regression models more effective.
- TransformedTargetRegressor
  - Wraps a regression model and applies a transformation to the target values.

### iClicker Exercise 10.1

iClicker cloud join link: https://join.iclicker.com/YJHS

Select all of the following statements which are TRUE.

- a. Price per square foot would be a good feature to add in our X.
- b. The alpha hyperparameter of Ridge has similar interpretation of C hyperparameter of LogisticRegression; higher alpha means more complex model.
- c. In Ridge, smaller alpha means bigger coefficients whereas bigger alpha means smaller coefficients.

### iClicker Exercise 10.2

iClicker cloud join link: https://join.iclicker.com/YJHS

Select all of the following statements which are TRUE.

- a. We can use still use precision and recall for regression problems but now we have other metrics we can use as well.
- b. In sklearn for regression problems, using r2\_score() and .score() (with default values) will produce the same results.
- c. RMSE is always going to be non-negative.
- d. MSE does not directly provide the information about whether the model is underpredicting or overpredicting.
- e. We can pass multiple scoring metrics to GridSearchCV or RandomizedSearchCV for regression as well as classification problems.

8

### Group Work: Class Demo & Live Coding

For this demo, each student should click this link to create a new repo in their accounts, then clone that repo locally to follow along with the demo from today.

9