Evaluation metrics play a crucial role in assessing the performance of regression models, which are used for predicting continuous outcomes. Here are some common regression evaluation metrics and guidelines on when to use them:

1. **Mean Squared Error (MSE):**
   * **Formula:** ���=1�∑�=1�(��−�^�)2*MSE*=*n*1​∑*i*=1*n*​(*yi*​−*y*^​*i*​)2
   * **When to use:** MSE is widely used and provides a good overall measure of model performance. It penalizes larger errors heavily, making it sensitive to outliers.
2. **Root Mean Squared Error (RMSE):**
   * **Formula:** ����=���*RMSE*=*MSE*​
   * **When to use:** Similar to MSE, but the square root is taken to make the metric more interpretable in the same units as the target variable. It is sensitive to outliers.
3. **Mean Absolute Error (MAE):**
   * **Formula:** ���=1�∑�=1�∣��−�^�∣*MAE*=*n*1​∑*i*=1*n*​∣*yi*​−*y*^​*i*​∣
   * **When to use:** MAE is less sensitive to outliers compared to MSE, making it suitable when outliers are present. It provides a more balanced view of the model's performance.
4. **R-squared (R2 or Coefficient of Determination):**
   * **Formula:** �2=1−∑�=1�(��−�^�)2∑�=1�(��−�ˉ)2*R*2=1−∑*i*=1*n*​(*yi*​−*y*ˉ​)2∑*i*=1*n*​(*yi*​−*y*^​*i*​)2​
   * **When to use:** R-squared measures the proportion of the variance in the dependent variable that is predictable from the independent variables. A higher R-squared indicates a better-fitting model. However, it may not be suitable when the goal is prediction rather than explanation.
5. **Mean Squared Logarithmic Error (MSLE):**
   * **Formula:** ����=1�∑�=1�(log⁡(1+��)−log⁡(1+�^�))2*MSLE*=*n*1​∑*i*=1*n*​(log(1+*yi*​)−log(1+*y*^​*i*​))2
   * **When to use:** Useful when the target variable has a wide range and you want to penalize underestimates and overestimates proportionally.
6. **Explained Variance Score:**
   * **Formula:** ���=1−���(�−�^)���(�)*EVS*=1−*Var*(*y*)*Var*(*y*−*y*^​)​
   * **When to use:** Similar to R-squared, it provides a measure of how well the model explains the variance in the target variable. A higher explained variance score is better.
7. **Huber Loss:**
   * **Formula:** Huber Loss is a combination of MSE and MAE, providing a compromise between them.
   * **When to use:** It is less sensitive to outliers than MSE and provides a balance between mean squared and mean absolute errors.

The choice of the metric depends on the specific characteristics of your data and the goals of your analysis. MSE and RMSE are commonly used, but if your data contains outliers or you prefer a more interpretable metric, consider using MAE or other suitable metrics based on the context of your problem. R-squared is useful for understanding the proportion of variability explained, while MSLE and Huber Loss cater to specific scenarios with non-linearities or skewed data. It's essential to consider the application and characteristics of your data when selecting an evaluation metric for regression models.

Top of Form