Mean Absolute Error (MAE)

Mean Absolute Error (MAE) is a commonly used regression metric that measures the average magnitude of errors in predictions, ignoring their direction. It provides a straightforward and interpretable way to evaluate the performance of a regression model.

Formula

The formula for MAE is:

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

Where:

- n is the number of data points,
- y_i is the actual value,
- $\hat{y}_i$ is the predicted value,
- $|y_i - \hat{y}_i|$ is the absolute error for each data point.

Characteristics of MAE

1. **Scale-Dependent**
   - MAE is measured in the same unit as the target variable, making it easy to interpret.
2. **Non-Sensitivity to Direction**
   - Since MAE considers the absolute value of errors, it treats over-predictions and under-predictions equally.
3. **Linear Relationship with Errors**
   - MAE increases linearly with the magnitude of errors. Large errors have a proportional impact on MAE, unlike metrics like MSE that penalize larger errors more heavily.

Advantages

1. **Interpretability**:
   - MAE directly reflects the average error in the same units as the dependent variable, making it intuitive for stakeholders.
2. **Robustness to Outliers (Compared to MSE):**
   - MAE is less sensitive to large deviations because it does not square the error terms, unlike Mean Squared Error (MSE).

3. **Applicability Across Domains:**
    - MAE can be applied to various regression tasks due to its simplicity and wide applicability.

Disadvantages

1. **Non-Differentiability:**
    - The absolute value function is not differentiable at zero, which can make optimization using gradient-based methods challenging. However, this is often addressed using subgradient methods or approximations.
2. **Sensitivity to Dataset Scale:**
    - MAE does not account for the scale of the dataset, which can make comparisons across datasets difficult.
3. **Equal Weighting of Errors:**
    - MAE assigns equal weight to all errors, which may not be ideal if larger errors are more critical for the application.

When to Use MAE

1. **Interpretability Is Important:**
    - When you need an easily interpretable metric to communicate the average error to stakeholders.
2. **Outliers Are Present but Not Dominant:**
    - When outliers exist but are not extreme enough to warrant more robust metrics like the Huber Loss.
3. **Balanced Error Impact:**
    - When you want to treat all errors equally without emphasizing larger ones.

Comparison with Other Metrics

1. **MAE vs. MSE:**
    - MAE treats all errors equally, while MSE penalizes larger errors more heavily by squaring them.

    - MSE is more sensitive to outliers, whereas MAE is more robust.
2. **MAE vs. RMSE:**
    - RMSE (Root Mean Squared Error) has the same units as the target variable but emphasizes larger errors more than MAE.
3. **MAE vs. Median Absolute Error:**

- The Median Absolute Error is more robust to outliers than MAE, as it focuses on the median of the absolute errors.

Example Calculation

Suppose we have the following actual (y_i) and predicted ($\hat{y}_i$) values:

- Actual: [3, -0.5, 2, 7]
- Predicted: [2.5, 0.0, 2, 8]

The absolute errors are:

- $|3 - 2.5| = 0.5$,
- $|-0.5 - 0.0| = 0.5$,
- $|2 - 2| = 0$,
- $|7 - 8| = 1.0$.

MAE is calculated as:

$$\text{MAE} = \frac{0.5 + 0.5 + 0 + 1.0}{4} = 0.5$$

Interpretation

An MAE of 0.5 in the above example means that, on average, the predictions deviate from the actual values by 0.5 units.

MAE serves as a practical and interpretable metric for evaluating regression models, especially when equal weighting of errors is appropriate.