

Analysis Review

1. Logistic Regression:

Categorical Features: *member_casual* , *season*

Continuous Features: *TMAX*

Target Variable: *rideable_type*

NaN Values: dropped all the rows with null values

Logistic Regression Accuracy: 0.64

Logistic Regression Confusion Matrix: $\begin{bmatrix} 1661458 & 0 \\ 948894 & 0 \end{bmatrix}$

The model correctly predicted the negative class (possibly 'non-bike') 1,661,458 times. The model incorrectly predicted the negative class 948,894 times when it should have predicted the positive class. The model did not correctly predict any true positive cases (possibly 'bike'), indicated by the 0 in this cell.

Categorical Features: *member_casual* , *season*

Continuous Features: *TMAX*

Target Variable: *rideable_type*

NaN Values: Handled NaN Values.

Logistic Regression Accuracy: 0.53

Logistic Regression Confusion Matrix: $\begin{bmatrix} 660353 & 1005389 \\ 615642 & 1134883 \end{bmatrix}$

The model has a fairly balanced distribution of predictions across the positive and negative classes but with substantial error rates. The model struggles to do so accurately, indicating potential issues with feature selection.

Categorical Features: *member_casual* , *day_of_week*

Continuous Features: *TMAX*

Target Variable: *rideable_type*

Logistic Regression Accuracy: 0.64

Logistic Regression Confusion Matrix: $\begin{bmatrix} 1661307 & 151 \\ 948781 & 113 \end{bmatrix}$

This model primarily predicts the negative class, with an extremely low number of true positives, which suggests a significant bias towards the negative class. This could be indicative of class imbalance or an inability of the chosen features (*member_casual*, *day_of_week*) to adequately differentiate between the classes.

Categorical Features: *member_casual* , *day_of_week*, *season*

Continuous Features: *TMAX*, *Elevation_Change*

Target Variable: *rideable_type*

Logistic Regression Accuracy: 0.64

Logistic Regression Confusion Matrix:

1661032	426
948390	504

Still correctly predicting the negative class most of the time. Slightly higher than the previous model, indicating a marginal increase in incorrect positive predictions. An increase in correct positive predictions, but still very low.