Regression Analysis Results Discussion

I employed Poisson regression to predict the expected count of ride-sharing trips for both bikes and e-scooters.

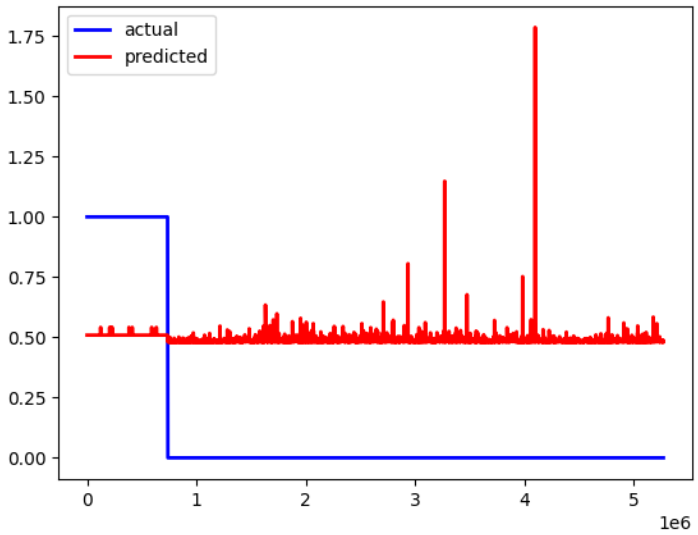
Since I'm ultimately aiming for binary classification (0 or 1 - ride or no ride), the predicted values represent probabilities between 0 and 1.

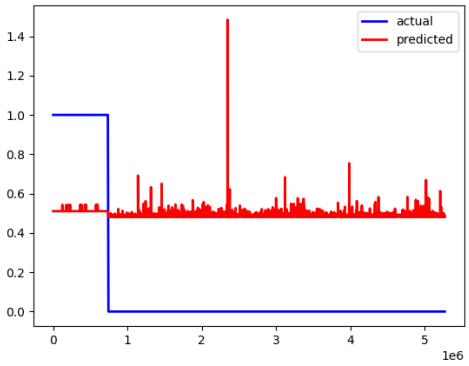
However, it's important to understand how to interpret these probabilities to determine when a ride is more likely to occur.

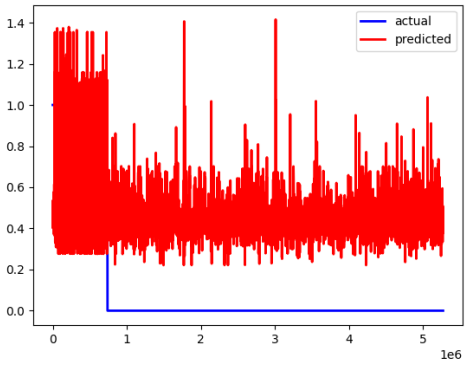
Feature Analysis: I analyzed the impact of various features on ride-sharing modality (bike vs. e-scooter):

- **Season:** (Fall, Spring, Summer, Winter):
- RMSE value: 0.62 (indicating low accuracy)
- Season appears to have a weak correlation with transportation preference.
- **Distance:** RMSE value: 0.47 (indicating better accuracy)
- A clear correlation exists between distance and chosen mode of transport.
- **Duration (Minutes):** RMSE value: 0.6498 (like Season)
- Duration exhibits a weaker correlation compared to distance.

Note: Distance is a stronger predictor of chosen transportation mode (bike vs. e-scooter) compared to Season and Duration.


actual
predicted
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
1
2
3
4
5
1e6


actual
predicted
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
5
1e6


actual
predicted
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
5
1e6