

# Assignment 6

---

0.5 points - What is the 90% confidence interval for validation set accuracy for this model.

lower bound: 0.8486  
upper bound: 0.8955

0.5 points - At the 80% confidence level (using a two sided bound) is the logistic regression model better than the most common class model? In 2-3 sentences explain how you came to the answer.

We cannot conclude that the logistic regression model is better at the 2-sided 80% confidence interval because the upper bound of the common model (0.861) is greater than the lower bound of the logistic model (0.854) at this confidence. We thus cannot reject the null hypothesis that these models are of equivalent accuracy at this confidence level.

0.5 points - At the 75% confidence level (using a one sided bound) is the logistic regression model better than the most common class model? In 2-3 sentences explain how you came to the answer.

Using a one-sided bound at 75% confidence, we can conclude that the logistic model is more accurate than the common model because the lower bound of the logistic model (0.863) is greater than the upper bound of the common model (0.851). We can thus reject the null hypothesis and conclude that the logistic model is more accurate at this confidence level.

1.0 points - Among the following possibilities for one-sided bounds: 75%, 90%, 95%, 97.5%, 99.5% Which is the highest level of confidence where we can say the logistic regression model is better than simply predicting the most common class? In 3-5 sentences explain how you came to that conclusion.

Two-sided accuracy bounds:

Validation set accuracy Logistic Regression: 0.9003.  
0.50% accuracy bound: 0.8975 – 0.9032  
0.80% accuracy bound: 0.8949 – 0.9058  
0.90% accuracy bound: 0.8933 – 0.9073  
0.95% accuracy bound: 0.8920 – 0.9087  
0.99% accuracy bound: 0.8893 – 0.9114

Validation set accuracy Common Class: 0.8660.  
0.50% accuracy bound: 0.8628 – 0.8693  
0.80% accuracy bound: 0.8598 – 0.8723  
0.90% accuracy bound: 0.8581 – 0.8740  
0.95% accuracy bound: 0.8565 – 0.8756  
0.99% accuracy bound: 0.8535 – 0.8786

Based on the data above, we can say the logistic regression model is better than the common class model at 99.5% accuracy with one-sided bounds. Looking at the 99% two-sided bounds above for both the Logistic and

Common model (which translates to the 99.5% one-sided bounds), we can see that the lower bound of the logistic regression model accuracy at 99% is higher than the upper bound of the common class model at 99%. Thus, we can reject the null hypothesis and conclude that the logistic regression model has higher accuracy than the common class model at 99.5% accuracy.