

Thomas Courtney  
Machine Learning, I  
Professor Wilck  
2/14/2021

Heading back to the insurance data using a classification model excited me. This is a very realistic example of how machine learning models are used in typical data science jobs. I applied to a job at Geico where I would be running similar models.

First, I loaded in the data. I then made sure the data was understood correctly. Next, I gained greater understanding of the data through various charts. I then cleared out NA's and binned different groups of data. Then I developed my model. I used a similar model to models 1-3, multiplying, adding and subtracted by certain values.

I ultimately chose model four because it had the best confusion matrix results for the predicted results of the dataset  $(1726/(1726+4282))$ . I then did this on the test data and it classified individuals that are flagged with a "1" 28% of the time, which is on average for the logistic regression problem in earlier modules.