

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 still ran the data through former models, but removed any output values that could change the results of the machine learning models.

I looked at the different machine learning types of tasks within this model. I ultimately chose the results after boosting it. I used the target amount to help determine the binary factor of whether they got the claim or not.

I think for this dataset, it is better to use the classification models that we used in a previous module. Although the insurance company cares most about whether their drivers get into an accident or not (a binary value) the actual insurance rates should be built on a linear scale.