

Individual Peer Evaluation Form

Your name: Jake Meyer

Write the name of your classmate you are preparing this review for in the designated column. Using a scale of 1-4 (1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree) answer each question. If you aren't able to answer the question based on what is posted in the discussion board, reach out to your classmate for more information via the discussion board. Total the numbers in each column. **Make sure to answer the questions on the 2nd page.**

Evaluation Criteria	Peer Name: Joel McMillin
Has plan in place to complete course project.	4
Has found datasets/data sources to support project idea.	4
Has solidified project idea.	4
Has identified resources for project.	4
Topic is related to data science and demonstrates topics learned to date through program.	4
Risks and potential issues have been identified.	4
TOTALS	24

Feedback on Individual's project topic:

1. How clear is the classmate's project topic? What questions does their topic make you consider?

The project is explained very clearly. The initial business problem for the project focused on using accident data in conjunction with insurance rates to classify risk, rate, and potentially predict premiums for policies. The scope of the problem was reconsidered after some difficulty finding relevant data. The revised problem statement focused on predicting accident severity (1- low through 4 -high) using controllable and uncontrollable (environmental) factors. After reading the initial scope of the project, I was curious about the following:

- What type of model or models will be utilized for making the risk prediction?
- What are the most influential factors increasing the risk (and most likely policy premiums) for folks?
- What recommendations will be made based on the results of the model(s) performance?

You answered the questions posed above. You used Decision Tree Classifier and Logistic Regression for your models. Also, you used the Chi-Squared metric to identify the top five features to be severity, starting longitude, zip-code, humidity, and clear condition. The actionable insights from your analysis were intended for insurance company stakeholders. These companies can recommend for drivers to avoid accident prone locations to prevent major accidents. Also, recommendations were made for companies to work with automobile manufacturers or designers to improve safety features for severe weather conditions. Excellent recommendations for action to be taken!

2. What risks or issues should your classmate consider while working on their project?

The risks outlined in the project are all valid concerns. I believe the predictive potential and data integrity concerns are ones that align across a lot (if not all) of our term projects for DSC630. Two potential concerns that come to mind are overfitting (with Decision Trees) and balance for the various categories within the dataset. Otherwise, you outlined a contingency plan and walked through the analysis process very clearly. Great job Joel!

3. Additional suggestions/comments that might be beneficial to your peer?

Joel, overall excellent job on your DSC 630 Term Project up through Milestone 4. It is clear you have thoroughly planned out the steps to approach this project and revisited these prior steps as you went through the analysis process. Two final comments for this peer review. First off, great job handling a dataset with over 2.5 million records. The number of missing values within

your dataset of US Accident Data (~70k) was larger than the entire dataset used for my term project. Second, finding relevant and accurate data towards the business case chosen for these projects tends to be one of the most challenging aspects for these projects. I thought you did a great job explaining this process for evaluating potential data source options as well as what was used for the analysis.

Adapted from a peer evaluation form developed at Johns Hopkins University (October, 2006)