**Individual Peer Evaluation Form**

Your name: Akila Selvaraj

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:  Brett Keller |
| 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. | 3 |
| TOTALS | 23 |

Feedback on Individual’s project topic:

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

Brett Keller has explained the project, he is going to work on very well. The dataset he has taken has considerable number of attributes needed for his analysis. Will these attributes be good enough to predict the survival of the patients? Does any other health related data sample need to be included as well for better prediction? Some of the things that I feel need to be considered are, if the dataset has enough volume of data to determine the risks. Does the dataset have details of both patients who were affected by heart failure as well as the information on healthy patients?

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

There could be lot of variables which may affect patient’s possibility to get heart problem. Out of all the variables, he needs to identify which ones have the greatest impact and which ones are controlling the dependent variable. At the same time, model may not work well if large number of features are considered. So, while working on this project, he needs to consider right number of features to predict the target variable.

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

It is good to analyze the correlation of each independent variable to the dependent variable before building the model and select only the important features and drop unwanted features to improve the accuracy of the classification model. Missing data and outliers should be treated appropriately using proper method.

Feature engineering is one of the best ways to increase the accuracy of the model. Scaling of data before applying the model could yield better results. Also, remove skewness of the data as much as possible before fitting the model. If creation of new variables using the exiting variables would generate new insights and if it helps in understanding the data well, that can also be considered.

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