Project 2 Questions & Answers

Harlan Wittlieff

Data Science, Bellevue University

DSC 680: Applied Data Science

Dr. Catie Williams

July 24, 2022

Questions and Answers

1. Is the model's accuracy score high enough to be reliable?

- The best model scored in the top 30% so far in the competition on Kaggle. There is some room for improvement, but overall the model is performing well.

2. When will the model be implemented?

- All four of the models have been submitted to the Kaggle competition.

3. What additional steps can be taken to boost the model's accuracy?

- Adding a neural network to the combined model may increase the overall accuracy.
- Additionally, investigating other methodologies for dealing with the nulls in the dataset could provide better results in terms of accuracy.

4. How will we know if the model is working as intended?

- Since a subset of the data was held back for evaluation in the competition, submitting our results to Kaggle will provide immediate feedback on the performance of the model.

5. What additional information may be useful to the model in the future that we should begin tracking now?

- If the project was not fictitious, knowing the height and weight of the passengers may provide some physics insights into the likelihood of transportation.

6. Which feature(s) provided the greatest predictability into the model?

- When reviewing the results, whether or not the passenger was in CrysoSleep was the most predictive feature on its own.

7. Does any bias have the potential to exist in the model?

- There is always potential for bias to exist in any model; however, care was taken to ensure that the risk is minimal.

8. How do these models rank among the other models created for this Kaggle competition?

- The best model scored in the top 30% of the competition.

9. Are there any follow up steps after the competition concludes?

- Reviewing the methodology of the top performing models will provide useful insights and feedback into where the success of this project could have been enhanced.

10. Would investigating any other model types be beneficial?

- Neural networks may be beneficial.