**You will peer evaluate two of your classmates (whose last names directly follow yours).**

Based on the following 7 evaluation metrics, write a reviewer’s report of comments with no more than 200 words in total for each student. Please give a score for each evaluation metric, and add up your total score for each report.

|  |  |  |
| --- | --- | --- |
| Abstract: | 5% | Provide context, motivation, and summary of findings. What questions are being answered? Why are these questions interesting/important? |
| Data: | 5% | Variables descriptions? What cleanups were done to the data? Good Graphics and Visualizations? |
| Models: | 5% | What did you do? What models and techniques did you use? Was any innovation attempted? |
| Results: | 5% | Did you properly evaluate your models performance? What are your conclusions? |
| Code: | 5% | Well documented Python codes with reproducible outputs? Good programming? |
| Quality: | 5% | Clarity of writing/presenting? Good readability of Notebook? |
| Complexity: | 5% | Complexity of your entire data collection, preprocessing, modeling, and analyses process in terms of data size and models sophistications. |

Your name: Chung Yan Wan

The student’s name you are reviewing: Amanda Pavlov Culley

The title of the project you are reviewing:

Your scores and comments for each evaluation metric and the total:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Abstract** | **Data** | **Models** | **Results** | **Code** | **Quality** | **Complexity** | **TOTAL** |
| 5% | 5% | 5% | 5% | 5% | 5% | 5% | 35% |
| 4% | 5% | 5% | 4.5% | 5% | 4.5% | 5% | 33% |
| You didn’t actually have an abstract or intro in your project. It will be nice if you can begin the project with an abstract so that it will look more cohesive. And also, explain why you are interested in this dataset or your motivation of doing this topic. | The variables of the dataset are well explained. I think that the histogram of digit labels was helpful to visualize the proportions of 0’s, 1’s, and so on in the dataset. | Great model analysis! You also evaluated a few models that we were not asked to do on the homework. However, for models that have more than one hyperparameters for tuning, I don’t know if you can tune them one by one, or it would be more appropriate to do a grid search for the best combination of the values of hyperparameters. | For some of the models, like Bagged Decision Tree, Log. Reg., and Rand. Forest, you were not using the optimal values of the hyperparameters for cv model evaluation. Make sure you put in the correct values from the results of hyperparameter tuning. I don’t know if that would lead to different results though. | The code is well written. I tried to reproduce part of your code, and that could be done without a problem. | The layout (flow) of your project can be more organized. For example, the results and conclusion can be placed at the end, in stead of placing at the beginning. Also, your project is missing a title. | Your data is very interesting and I believe it has a good complexity for this project, since it has a multiclass response variable, which is a little different from what we’ve seen in our homework examples. | Very nice job Amanda. I hope that you have time to finish the entire analysis according to you plan and submit it to Kaggle. |