# Image Recognition Case Study Rubric

**DS 4002 – Fall 2024 - Rithwik Raman**
**Due: TBD**

**Individual Assignment**

**Why am I doing this?** This case study allows you to use your data science knowledge by applying image recognition techniques to a practical and socially relevant problem: detecting face mask usage. As you work through this assignment, you will gain hands-on experience in building and evaluating machine learning models, exposing you to real-world applications of data analysis with implications for public health and safety.

**What am I going to do?** In this case study, you will work with a dataset of 11,800 labeled images sourced from Hugging Face, containing examples of individuals either wearing masks or not wearing masks. Your task is to preprocess the dataset and then develop a convolutional neural network (CNN), such as ResNet or MobileNet, using Python to classify the images accurately. Once the model is built, you will evaluate its performance using key metrics like accuracy, precision, recall, F1 score, and AUC-ROC to assess its ability to classify images correctly and reliably. Finally, you will document your process and reflect on the challenges presented by variations in lighting, orientation, and image quality, while providing insights into potential improvements for real-world applications of this technology.

**Deliverables include:**
- Written description of the case study experience
- Functioning, well-documented source code
- Any plots / tables generated
- Github repository containing all above files

**Link to provided GitHub repository (you will create your own)**:
https://github.com/rithwikraman/DS4002_CS3

**Tips for success:**
- ***Make sure to connect to Rivanna, UVA's high-performance computing system.***
    - Don't fret if your model is taking too long to run - with Rivanna, it will take roughly 1 to 1.5 hours. Without it, however, it will run indefinitely. As a result, **don't wait until the last minute to begin the case study**.
- Use data augmentation techniques to improve model generalization.
- Reflect on how this project could be applied to other image recognition challenges.

**How will I know I have succeeded?** You will meet expectations on this case study when you follow the criteria in the rubric below.

| | |
|---|---|
| Formatting | • Create a new GitHub repository for this assignment<br>    ○ Named as "DS4002_CS2_FaceMaskDetection"<br>    ○ Contains the following files:<br>        ■ README.md<br>        ■ LICENSE.md (use the MIT license)<br>        ■ SCRIPTS directory<br>        ■ OUTPUT directory<br>• Submit the link to this repository on Canvas. |
| README.md | • Goal: A brief, high-level overview of your case study outcomes<br>• It should contain:<br>    ○ 1-2 paragraph summary of the steps you took when reproducing the results, your overall experience doing the case study, and any suggestions for improvement.<br>    ○ Link to the dataset (link to HuggingFace)<br>    ○ List of any external references you used (websites, journal articles, etc.), cited in IEEE format |
| SCRIPTS | • Goal: Contains all source code used in the case study.<br>• Upload the Jupyter notebook used during the case study.<br>• Upload any other scripts you may have used. |
| OUTPUT | • Goal: Contains all plots / tables generated<br>• Upload any EDA (exploratory data analysis) plots that were generated<br>• Upload a table containing model performance metrics.<br>• Use the provided output files as reference. |
| Model Evaluation Metrics | • Goal: Develop a high-performing image recognition model.<br>    ○ Achieve an accuracy of > 90%<br>    ○ Achieve > 85% precision and recall.<br>• Print your results in the Jupyter notebook, and upload a table of results to the OUTPUT directory. |

Acknowledgements: Special thanks to Professor Alonzi for providing the rubric structure.