Alzheimer's Case Study Rubric

DS 4002 – Spring 2023 – Ashley Kim
Due: May 10th, Noon
Submission format: Upload link to GitHub repository on UVA Canvas

Individual Assignment

Why am I doing this?

This case study allows you to leverage your data science knowledge by producing an image classification model that helps detect Alzheimer stages using MRI scans of the brain. It is the hope that by working through this assignment, it can encourage you to see the impactful ways data can be utilized in a real-world context – especially in the field of medicine and early detection of fatal neurogenerative conditions, data science could very much save a person's life.

What am I going to do?

The link to the GitHub repository for this case study is https://github.com/ayk2ea/DS4002-CS2_Alzheimers. Use the provided dataset on Kaggle that contains four datasets of MRI images separated by Alzheimer stages to create an image classification model in either R or Python utilizing a convolutional neural network (CNN). In addition, after you build your model, you will need to pass through several images (that are provided on the GitHub repository) through your model to see how well it can predict with new data. A standard workflow would be to first familiarize yourself with the images you're analyzing within the dataset through some exploratory data analysis, then conduct preprocessing/standardizing with the images, and finally building out and evaluating your machine-learning based classification system.

Your final deliverables should include:

- Well documented source code for your image classification model
  - Well commented, well documented and <u>clean</u>!
- A markdown file citing any resources (journal articles, websites, etc.) referenced in helping you create your model
- All these above should be included in a GitHub repository that is submitted to the class Canvas site

Tips for success:

- There are a lot of different ways to go about producing the model. I recommend looking at the initial references provided before doing your own research. These references have been selected to start you off and build your knowledge about the convolutional neural network algorithm that can hopefully provoke you into making astute decisions in how to go about building the classification system
- Be sure to split your data into training and testing sets, as the dataset given is not split. A general best practice is to do an 80/20 training testing split.

- Epochs are the number of cycles/iterations the training data is put through in a neural network, and more epochs means that the model has better generalizations given new data. Be sure to set the number of epochs to what's appropriate for your model as overfitting can be a potential problem you run into while building it.

How will I know I have succeeded? You will meet expectations on this case study when you successfully follow and complete the criteria in the rubric below:

Spec Category	Spec Details
Formatting	- One GitHub repository (submitted via link on Canvas) - Create a new GitHub repository for this assignment titled 'CS2_LastName-FirstName' that contains: - README.md - LICENSE.md - Source Code File - REFERENCES.md
README.md	Brief summary of what you've produced for the case study, this does not have to be super detailed but provide enough information to orient people to your repository
Source Code File	Well documented R Markdown or Jupyter Notebook file that contains the code used to execute your song recommendation system. What you must include in the source code are as follows: - Training/testing separation - Exploratory data analysis visualizations/plots - Preprocessing/standardization of images - CNN Model - Model accuracy plots - Predictions with new data
REFERENCES.md	Markdown File titled "REFERENCES.md" with citing any resources (journal articles, websites, etc.) referenced in helping you create your model in IEEE Documentation style. Also include brief annotations under

	each citation on how each reference informed/helped you for this case study.

Acknowledgements: Thank you to Professor Alonzi for providing the rubric structure!