

Case Study Rubric

Assignment Description:

Follow the given Case Study Repository's instructions to create an image classification model that can distinguish between movie posters of different genres. Present your results in an organized GitHub repository.

Purpose:

To synthesize data science techniques and use publicly available resources to build the most efficient convergent neural network image classification model possible.

You will meet the expectations for competing this Case Study when you follow the criteria in the rubric below:

Spec Category	Spec Details
Formatting	<ul style="list-style-type: none">• One GitHub Repository• The top-level page of the repository should contain:<ul style="list-style-type: none">◦ A README.md file (which auto displays)◦ A SCRIPTS folder◦ A DATA folder◦ AN OUTPUT folder
README.md	<ul style="list-style-type: none">• <u>Goal:</u> This file serves as an orientation to everyone who comes to your repository, it should enable them to get their bearings.• Use markdown headers to divide content.• Make an H2 (##) section explaining the contents of the repository• <u>Section 1: Software and platform section</u><ul style="list-style-type: none">◦ The type(s) of software you used for the project.◦ The names of any add-on packages that need to be installed with the software.◦ The platform (e.g., Windows, Mac, or Linux) you used.• <u>Section 2: A Map of your documentation.</u><p>In this section, you should provide an outline or tree illustrating the hierarchy of folders and subfolders contained in your Project Folder, and listing the files stored in each folder or subfolder.</p>• <u>Section 3: Instructions for reproducing your results.</u><p>In this section, you should give explicit step-by-step instructions to reproduce the Results of your study. These instructions should be written in straightforward plain English, but they must be concise,</p>

	but detailed and precise enough, to make it possible for an interested user to reproduce your results without much difficulty.
SCRIPTS folder	<ul style="list-style-type: none"> ● <u>Goal:</u> This folder contains all the source code for your project. ● Include all the scripts you used by building off the given Case Study repository, including any data cleaning and exploratory work. Importantly include all the code needed to create your image classification model. ● All script files should include header comments at the beginning of a script to provide information that anyone working with or executing the script should be aware of. Throughout all your scripts, you should include copious comments explaining what each command or sequence of commands accomplishes and what the purpose is.
DATA folder	<ul style="list-style-type: none"> ● <u>Goal:</u> This folder contains all the data for this project. ● You should AT LEAST the data include the initial data, and the final data analyzed. If needed, the code in the SCRIPTS folder should be able to get you from the initial piece of data to the final one. N.B. If the initial and final data are the same, then just include that dataset. ● If your data fits in GitHub, place all of it here. ● If your data does not fit in GitHub use a single file explaining the process to obtain the dataset. E.g., how to web scrape for movie posters. ● A Data Appendix file as a PDF, which will include text that you type, as well as tables, figures, and other descriptive statistics.
OUTPUT folder	<ul style="list-style-type: none"> ● <u>Goal:</u> This folder contains all the output generated by your project, e.g. figures, tables, etc. Importantly, remember to include a screenshot of the accuracy of your model. ● Use informative names for your files.
References	<ul style="list-style-type: none"> ● All references should be in their own document ● Use IEEE Documentation style

Acknowledgements: Elements taken from DS 4002 MI3 rubric document. This structure is pulled from Streifer & Palmer (2020).