CS - Licence Plate Recognition Project Case Study

DS 4002 Spring 2024 - Rishi Raghavan
Due: May 2025
Submission Format: GitHub repository and Powerpoint Presentation submitted to Canvas

Individual Assignment

General Description: Submit to Canvas a link to your Github Repository and a PPT of your presentation.

Why am I doing this? Case studies are an excellent way to apply technical and theoretical knowledge in real world settings. They provide students opportunities to demonstrate problem-solving skills alongside critical-analysis. You will practice developing a methodology to a problem and utilizing relevant data-science techniques to reach conclusions. Working with real data sets will challenge you to perform good data cleaning and preliminary exploratory data analysis before actually using your data.

- Course Learning Objective: Apply theoretical data science knowledge to create functional results
- Course Learning Objective: Apply critical thinking to adapt models to a real world situation
- Course Learning Objective: Analyze findings to find conclusions
- Course Learning Objective: Create a presentation to frame your work in an easy-to-digest manner for general audiences to consume.

What am I going to do? You will utilize image-recognition techniques to determine from an image of a license plate which state the plate is from. This will involve creating code that processes an image so it is usable for image-to-text recognition, and then fitting this code within a framework that allows batches of images to be processed. You will then run several tests to assess the accuracy of your model and determine accuracy relative to an 80% threshold. The method of assessing accuracy is up to your choice, and should be statistically justifiable. You will then combine code and results of this in a GitHub repository with a high level of organization that makes it easy for anyone to follow your work. Finally, your findings will be compiled into a presentation that you can give to your peers.

Tips for success:

- Be bold. This is your chance to try different techniques and expand your capabilities. There are many successful ways to complete this project!
- Don’t overthink it. A clear presentation of fundamentals is more valuable than an unclear presentation of cutting edge techniques.
- Talk to the instructors. This is a creative assignment, and you are allowed to show ideas to people for comment.
- Talk to your fellow students. This is a creative assignment and you are allowed to show ideas to people for comment.

How will I know I have Succeeded? You will meet expectations on Case Study when you follow the criteria in the rubric below.

Spec Category	Spec Details
GitHub Repository	● One Github Repository (submitted via link on canvas) ● **A README.md File** o Goal: This file serves as an orientation to everyone who comes to your repository, it should enable them to get their bearings. ● Section 1: Software and platform section ○ 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. o Section 2: A Map of your documentation. ▪ 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. o Section 3: Instructions for reproducing your results. ▪ 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, but detailed and precise enough, to make it possible for an interested user to reproduce your results without much difficulty. ● **LICENSE.md File:** ○ Goal: This file explains to a visitor the terms under which they may use and cite your repository. ● Select an appropriate license from the GitHub options list on repository creation. ● Usually, the MIT license is appropriate. ● **SCRIPTS Folder:** o Goal: This folder contains all the source code for your project. o Include all the scripts you used. Try to name each script according to the order it needs to be executed to reproduce the results. o 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:** ○ Goal: This folder contains all of 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. ○ 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:** ○ Goal: This folder contains all of the output generated by your project, e.g. figures, tables, etc. ○ Importantly, any information like tables, figures shown in your presentation should be here. ○ Use informative names for your files. ● **References:** ○ All references should be listed at the end of the document ○ Use IEEE Documentation style (link)
Presentation	● About 7 slides ● PDF format for submission to collab ● Generate the slides through the program of your choice ● Slide numbers (except for title slide) ● Order ○ Title & Outline ○ Motivation/Context/Hypothesis/Research Question/Modeling Approach/Goal/Etc. ○ Data Explanation/Acquisition ○ Analysis Plan and Justification ○ Tricky Analysis Decision ○ Bias and Uncertainty Validation ○ Results/Conclusions ○ Next Steps ○ References/Resources/Acknowledgements ○ Closing Slide ● General note: ○ Each section can take as many slides as it needs to unless specifically indicated as 1 slide. ○ Do not take the exact slide count as critical. When you practice you can time out the talk to hit the 7 minute mark. Some slides need more, and some slides need less time.