Case Study Rubric – Caption Generation

DS 4002 - Fall 2023
Due: See Canvas
Submission format: GitHub repository (submitted by link to canvas)

Individual Assignment

General Description: Submit to Canvas a link to your GitHub repository for this project.

Preparatory Assignments

- In class sessions on case studies

Why am I doing this? The goal of this assignment is to apply your knowledge of data science in a real world scenario. You are doing this to practice thinking like a data scientist and to use the skills and knowledge you already have to take on a challenge. This is a great opportunity for you to use your problem solving skills to explore different techniques and try to find a solution to a unique challenge.

- Course Learning Objective: Problem Solving & Applied Thinking
- Course Learning Objective: Analysis
- Course Learning Objective: Documentation

What am I going to do? You will start off by reading the prompt that you've been assigned. Take a moment to understand the context behind the prompt, now would also be a good time to explore the provided articles giving more background information. Afterwards, you should take a good look at the data that you will be using and try to brainstorm next steps, read over any technical articles that were provided and in addition do your own research to determine what direction you want to take to produce your deliverable. Have a goal in mind and lay out steps that you should follow to reach your goal. Make sure to thoroughly look over all the resources that have been provided to you as well as researching on your own. All of your source code and anything related to your analysis process should be in your github repository that you will turn in as your final deliverable. Your presentation with your conclusions should be completed as part of your deliverable and should be linked in the readme of your github repository.

Tips for success:

- Stay focused and don't procrastinate!
- Think of a strategy and take time to brainstorm before jumping into the code.
- Don't be afraid to try new ideas, explore as much as you can.

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

Spec Category	Spec Details
Formatting	● Github Repository ● Create a github repo for this assignment consisting of: ○ README.md

	○ LICENSE ○ SRC Folder ○ DATA Folder ○ FIGURES Folder
README.md	● Goal: This file should give an overview of every folder in your repo. ● SRC section ○ Describe how to access and install your code ● DATA section ○ Summarize the data you used to train and test your model with ● FIGURES section ○ Describe any figures you created or used ● PRESENTATION section ○ Provide the link to your slideshow ● REFERENCES section ○ All references should be listed in IEEE Documentation Style
LICENSE	● Goal: This file explains to a visitor the terms under which they may use and cite your repository. ● Use the MIT license
SRC Folder	● Goal: This folder contains all the code for your project. ● Include any code you used to train and test the model or clean data
DATA Folder	● Goal: This folder contains all of the data you used to train and test the model, if it can't fit then include a link to a google drive folder ● Include csv files of generated captions for the test image sets
Figures Folder	● Goal: This folder contains all of the figures you made or used ● Include figures that you plan on using in your presentation
Presentation	● Goal: Demonstrate your analysis process and conclusions ● The research question and any relevant findings must be discussed ● Address which test set had more accurate captions generated and what this could mean in terms of machine learning and accessibility ● Provide an example of a caption generated for each test set by your model ● Describe the criteria you used to judge the accuracy of your model ● Your presentation should be 5-6 minutes and can follow this general order: Title, Motivation, Research Question/Hypothesis/Model Approach, Data Acquisition, Analysis Plan, Results/Conclusions, Next Steps

Acknowledgements: Special thanks to Jess Taggart from UVA CTE for coaching on making this rubric. This structure is pulled directly from Streifer & Palmer (2020). Special thanks to Professor Alonzi for his guidance on this rubric.