Milestone 2 | Project Proposal and Plan

Please reference the rubric for grading purposes. This is located at the bottom of the page.

The goal of the milestone is to draft a project proposal and plan that you will complete during the course. The proposal and plan you develop are flexible and should be thought of as a guide, not a strict set of actions you have to follow. If your plan changes as you create your project, it is encouraged to go back to your proposal and plan document and update it.

If possible I encourage you to use existing/previous projects you have worked on or are actively working on in a different class. I believe it is more valuable to dig deeper with one project than have 2 separate projects. If you plan to combine this project with another from a different class or research please state that in your proposal, including the other classes/research project requirements.

Please note these proposals and plans are expected to be made public - to the DSA, your classmates, and future employers/research institutions.

The Proposal

Your proposal should contain the following sections:

- A problem (this can be related to your interests you completed in milestone 1): **Creating an algorithm that improves sensory vision for color vision deficiency.**
- A proposed method of solution involving a machine learning algorithm**: Creating an ML algorithm that will detect whether the person is looking at a red, blue, etc image.**
- **I was aso thinking of using https://onlinepngtools.com/convert-png-to-rgb-values to turn pictures into 2d vector of RGB values. Eventually learn how to use that data to create an image color classification.**
- **Would use small sized images, simple images with mostly one color, but eventually use some that use an array of colors.**
- Potential Stakeholder - who could benefit from using your model? Is it more than just yourself?**: People with impaired vision, biotechnology companies.**
- Potential Obstacles - what are some obstacles you think you might encounter, these can be technical problems (ei I may not find the data set that describes my problems statement), or general problems.**: How to clean the data, not enough data that supports my problem statement.**

- What is novel about your problem and your potential approach?**: Can help optometrists and engineers create a neural device that helps color blind people potentially see the right colors.**

Please see the rubric for length and quality metrics for your proposal grading.

The Plan

Your plan should be an outline of the following bullet points:

- Where do you plan to get your data sets?**: Will create my own data by collecting photos of images that range from blue, green, red, etc.**
- If you plan to create your own dataset, please include a detailed plan of how you are going to collect, where you are going to collect it and where you plan to store the data. - This data will be accessible publicly. It is necessary that you have the correct permissions/rights to use the data you collect.**: I will take images from the internet that range in colors, I plan on storing this data in excel files including citations where everything is pulled from.**
- Take some time to sift through your datasets, how do you plan to organize your data?**: I will classify the images by color.**
- How do you plan to analyze your data? What are you looking for? What might you find? **I am going to look where the image lies on the RGB spectrum, see how accurate it is in classifying the images/color. It is possible I find shading, color, shape may affect how humans perceive the color.**
- Based on your problem statement, what type of model do you plan to use? Is it one of the ones we have discussed/will discuss in class?
    - If we have not done a model overview in class yet it may be necessary to do a quick google of what models we will talk about in class and how they may be used to solve your problem statement. - You may also use models not discussed in class.
    - **CNNS: a type of neural network specifically designed to process grid-like data, such as images. They are mainly used for prediction or classification of objects. I believe that while the focus of my project isn't with identifying the object, being able to comprehend the object with its color provides the beneficial data.**
- How will you determine if your model is accurate? **All the correct colors should be classified together, could look at the R^2 the coefficient of determination and it determines whether this regression model is working or not. Evaluating metrics such as precision, recall, and F1 score.**
- Are there similar models/paper/code online that are trying to do what you are doing? If there are, use them as a baseline for improvement

- **Yes!**

Project Expectations

- All data sources should be cited in the final report and in code.
- All reference material should be cited in the final report
- Try your best! A “failed” model can still earn you a passing grade. In your final report there will be a section on future improvements.

Submission of this milestone:

Create a template and clone the following repo to your local machine:
https://github.com/NCState-Exploring-Machine-Learning/DSC412-project.git

This repo should be public, have the organization owner be **2024-FALL-DSC-412-001**

Once cloned add the pdf of your proposal and plan to the repo and push it up to github. This will be the repo you will use for your project for the semester.

The file should have the format: DSC412_001_FA24_PR_unityID.pdf

For example mine would be:
DSC412_001_FA24_PR_tjembry.pdf

Feedback:

Once your proposal and plan is submitted, I will read and re-scope (if necessary) your project, so it is obtainable within this one credit hour course. Re-scopes that suggest less work is optional, re-scopes that suggest more work are mandatory. I highly encourage you to use existing material or combine this project with other class projects, or existing research. In my opinion, it is more valuable to have one project where you go deeper, than two projects that are surface level.

It is not recommended to start your project until you have received feedback from your instructor. The goal is to have feedback 1 week after the due date.

Rubric:

Criteria	Failing	Passing	Reaching Full Credit
Submitted as PDF and through github repo	Not submitted as PDF or github repo	Submitted as PDF in github repo with link as submission	
Submitted with correct file name format DSC412_001_FA24_PR_unityID.pdf	Submitted with incorrect file name	Submitted with correct filename	
Proposal Quality	Not all bullets addressed in proposal section	All bullets addressed in the proposal section with rational responses.	All bullets addressed thoroughly and thoughtfully in the proposal section.
Plan Quality	Not all bullets addressed in plan section	All bullets addressed in plan section with rational responses	All bullets addressed thoroughly and thoughtfully in the plan section.
Late Submission*	Submitted greater than 3 days late (flexible with email to instructor)	Submitted less than 3 days of being late	Submitted before due date

* Due to this assignment having feedback it is very important this is submitted in a timely manner.

Grading:

If any criteria are “Failing” | 50% Maximum
All criteria “Passing” | 70%
For each “Reaching Full Credit” criteria met | + 10%
Total Possible Percent - 100%

Addendum:

- 09/19 9:10pm - Updated criteria names on rubric to be more reflective of end score. - Thank you for the feedback!