UDACITY

‹ Return to "Machine Learning Engineer Nanodegree" in the classroom

DISCUSS ON STUDENT HUB

# Capstone Proposal

REVIEW

CODE REVIEW

HISTORY

## Requires Changes

2 SPECIFICATIONS REQUIRE CHANGES

Dear student

Great start on this proposal! I've noted a couple of areas where you should add a bit more detail, but I think that you've picked a great project and you're definitely on the right track. I think you'll see that most of these things shouldn't take long to update. Almost there...keep going!

Cheers!

### Project Proposal

**Student briefly details background information of the domain from which the project is proposed. Historical information relevant to the project should be included. It should be clear how or why a problem in the domain can or should be solved. Related academic research should be appropriately cited. A discussion of the student's personal motivation for investigating a particular problem in the domain is encouraged but not required.**

Great job giving the reader an introduction to the problem domain!

Suggested:

- Since you've included a link to your data source in this section, you can directly lift it into the `Project Overview` section of your capstone report (to save a bit of time).

**Student clearly describes the problem that is to be solved. The problem is well defined and has at least one relevant potential solution. Additionally, the problem is quantifiable, measurable, and replicable.**

> Here the problem is quantified as classifying the correct stage of Retinopathy and one relevant solution is to use image classification model to predict and assist humans with high accuracy.

Great work! I think that this is a very clear statement of the problem and how it is structured.

**The dataset(s) and/or input(s) to be used in the project are thoroughly described. Information such as how the dataset or input is (was) obtained, and the characteristics of the dataset or input, should be included. It should be clear how the dataset(s) or input(s) will be used in the project and whether their use is appropriate given the context of the problem.**

You've made a really strong start here! The scope and nature of the dataset are pretty clear. Just one small thing:

- Please be sure to also note how the dataset classes are balanced. How many instances are there for each of the 5 classes? Is one class more common than the others or are the classes balanced?

**Student clearly describes a solution to the problem. The solution is applicable to the project domain and appropriate for the dataset(s) or input(s) given. Additionally, the solution is quantifiable, measurable, and replicable.**

I think that this is a pretty solid high-level summary of your proposed solution. Nice job!

**A benchmark model is provided that relates to the domain, problem statement, and intended solution. Ideally, the student's benchmark model provides context for existing methods or known information in the domain and problem given, which can then be objectively compared to the student's solution. The benchmark model is clearly defined and measurable.**

> Doctor/manual intervention for screening can be costly and junior/intern classification results only around 50-60% accuracy

around 50-60% accuracy.

Please be sure to add a citation for this. If you add a citation that indicates that 60% accuracy is a common performance for interns, then this should be appropriate (and you don't need to add or change anything else).

If you can't find a citation for this statistic, please be sure to pick another goal. It's important to choose a specific result that isn't arbitrary. For this reason, it's best to avoid using the Kaggle leaderboard as a benchmark since there are so many different results that could be chosen.

For example, you could choose a relevant result from the academic literature as the benchmark. You could also specify a simple model (that you'll train/test using the same data) to act as the benchmark. If you decided to code a benchmark model yourself, you don't need to have tested it yet...just specify what it will be.

**Student proposes at least one evaluation metric that can be used to quantify the performance of both the benchmark model and the solution model presented. The evaluation metric(s) proposed are appropriate given the context of the data, the problem statement, and the intended solution.**

Great choice! This is a common metric for imbalanced diagnostic classification problems.

**Student summarizes a theoretical workflow for approaching a solution given the problem. Discussion is made as to what strategies may be employed, what analysis of the data might be required, or which**

**algorithms will be considered. The workflow and discussion provided align with the qualities of the project. Small visualizations, pseudocode, or diagrams are encouraged but not required.**

I think that you're definitely meeting the specifications here. Keep in mind that this is a great opportunity to bounce lots of ideas off of the reviewers. You can get feedback without putting in much work. It's also a good idea to build some backup plans into your workflow in case something doesn't work. You don't want to get stuck...

**Proposal follows a well-organized structure and would be readily understood by its intended audience. Each section is written in a clear, concise and specific manner. Few grammatical and spelling mistakes are present. All resources used and referenced are properly cited.**

The template format is followed and the proposal is well written.

RESUBMIT

DOWNLOAD PROJECT


## Best practices for your project resubmission

Ben shares 5 helpful tips to get you through revising and resubmitting your project.

Watch Video (3:01)

RETURN TO PATH