UDACITY

PROJECT

Capstone Proposal

A part of the Machine Learning Engineer Nanodegree Program

PROJECT REVIEW

CODE REVIEW

NOTES

SHARE YOUR ACCOMPLISHMENT!

Meets Specifications

Dear student,

Your proposal demonstrates a well-thought out project plan with almost all of the necessary considerations addressed. I hope you enjoy this capstone project and look forward to reading your report.

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 introducing the problem you plan to address with this capstone project. Your discussion includes relevant background concepts to main problem and provides evidence of a literature search prior to embarking on the capstone.

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.

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.

Nice job presenting the dataset and its characteristics. It's important to consider both the size of our data and the distribution. It's clear how you'll use this data in your project.

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.

In general, this is a clear solution statement. However, see my comments for the next section.

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.

You've proposed using derived features from the MRI image data for building classical machine learning models.

- Derived anatomic volumes data include:
  - @Estimated total intracranial volume (eTIV, mm3), Buckner et al., 2004
  - @Atlas scaling factor (ASF), Buckner et al., 2004
  - @Normalized whole brain volume (nWBV, mm3), Fotenos et al., 2004

Moreover, you'd like to benchmark these classical models against a convolutional neural network (I'm assuming you'll use the raw input data here). This seems slightly backwards. For one, I imagine you'll spend just as much time if not more time constructing and training the neural network model. Further, this is a more complex model that removes the need for extracting and engineering features. Thus, it seems like the proposed classical models should serve as a benchmark against using a CNN on the raw image data.

I'm not going to require that you switch your approach (as ultimately if you can build a simpler model that works better, that's ideal), but I would like to leave this as a suggestion for you to consider.

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.

While your classes for the binary classification case are relatively balanced, the classes for the multi-class classification problem will not be balanced. You'll want to use the f1-score for that case.

-- 0=nondemented (341 data points)

-- 0.5 = very mild dementia (193 data points)

-- 1 = mild dementia (69 data points);

-- 2 = moderate dementia (5 data points), from Morris, 1993.

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.

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.

DOWNLOAD PROJECT

RETURN TO PATH

Student FAQ