UDACITY

< **Return to "Machine Learning Engineer Nanodegree"**   DISCUSS ON STUDENT HUB
   in the classroom

Capstone Proposal

REVIEW
HISTORY

Meets Specifications

This is a well written and interesting project proposal of a real world image classification problem.

I'm looking forward to see what you will do in the second stage of the project.

Once more excellent job :) Keep it up :)

Project Proposal

> **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 description of the problem that is to be solved is simple and easy to understand as well as quantifiable, measurable, and replicable. Good job :)

> **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.**
>
> Excellent job explaining the Domain Background. :)

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.

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.

You are correct accuracy should not be used for imbalanced datasets. The defined metric on Kaggle is appropriate for this problem.

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.

As the rubric states 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.

You have written: For the benchmark model, it will be used the CNN model with an accuracy of more than 10%.

Which is similar but not exactly the the same thing. Perhaps it can be rephrased as:

The benchmark model will be a custom CNN trained from scratch.

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.

Detailed description of the dataset has been provided.

You could have provided more information on the fact that the data is imbalanced. How many examples are there for each class. This can be nice graph summarizing this information.

Student summarizes a theoretical workflow for approaching a solution given the problem. A 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.

Well written project design stages. :)

Tips:
One can use cross validation to check the robustness of the solution. Take a look here:
https://machinelearningmastery.com/k-fold-cross-validation/

Visualizing and interpreting model results:
SHAP (SHapley Additive exPlanations) is a game theoretic approach to explain the output of any machine learning model. It connects optimal credit allocation with local explanations using the classic Shapley values from game theory and their related extensions see the following link:
https://github.com/slundberg/shap

The 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.

Good structure of the whole report. Great job :)

⤓ DOWNLOAD PROJECT

RETURN TO PATH

Rate this review