UDACITY

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

DISCUSS ON STUDENT HUB ›

Capstone Proposal

REVIEW

HISTORY

Requires Changes

4 SPECIFICATIONS REQUIRE CHANGES

Very strong first submission! Well done! Go through the comments carefully and amend where necessary. You're almost there!

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.

Awesome

- Great job! You did a really good job of introducing the project and getting the reader interested.

Required

- Also required here is some related academic research relevant to this domain. This should be appropriately cited.

Suggestion

- I would suggest adding a bit of a personal touch to this, what's your personal motivation for choosing this problem? is this a domain you're particularly interested in?

✓

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.

Awesome

- The problem is clearly defined

✓

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.

Awesome

- The input data is discussed in detail.

✓

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.

🔄

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.

Required

> Available dataset have 492 frauds out of 284,807 transactions. The dataset is highly unbalanced, where
> the positive class (frauds) account for 0.172% of all transactions. We have to consider this to build our
> model.

- The idea behind a benchmark is a model that we can compare subsequent models to which lets us know if we may consider those subsequent models as "good enough". This could be the

performance of a model worked on by others in research or it could be a simple model we implement e.g A simple Naive Bayes or Logistic Regression model.

https://datascience.stackexchange.com/questions/8785/what-is-a-benchmark-model

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.

Required

- Due to the extreme class imbalance, accuracy is not at all an appropriate metric. It should only be used if, for example, additional steps are being taken to reduce the imbalance. Otherwise, I would strongly suggest excluding it.

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.

Required

- Some more details would need to be provided. What preprocessing steps will be taken? It isn't clear what feature evaluation refers to. Can this be elaborated on? Did you mean feature selection? if so, what techniques will be used? what validation methods will be used? how will the data be split? what models will be considered?

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.

RESUBMIT

DOWNLOAD PROJECT

Learn the best practices for revising and resubmitting your project.

RETURN TO PATH

Rate this review