REVIEW

Meets Specifications

Perfect!
Now is to develop the final project to become a Machine Learning Engineer!
😄 😄 😄

Good luck!

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.

Excellent description of the domain background. You cited references that deal with this kind of problem and also mentioned possible applications and solutions. Awesome!!! 😄

Suggested: Please, in your final report, you might link your intro with references (in the references section, like this ... some reference text [1]... where [1] is the reference number).

✓

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 quantifiable, measurable, and replicable. Great job here! The problem is clearly defined. And you correctly mentioned that it is a regression problem and gave a brief explanation about the inputs and the expected outputs.

✓

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.

Congrats! The dataset to be used in the project are well described here.

Suggested: you could describe in more detail your dataset in the final report. Here is a very complete article on various techniques of the data exploration process. And you could show how many data points are there in the dataset, the values distribution, and so on. This lets your work even more interesting.

✓

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.

You clearly described the solution to the problem, and it 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.

Very cool your choice of benchmark model 😄 This step will be important for you to compare your final model with some of them and see if it got better, same or worse.

✓

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!!! You presented an evaluation metric that can be used to quantify the performance of both the benchmark model and the solution model. And as a future machine learning engineer, it is always important to explain why you chose a specific metric instead other to solve the problem/select the best model 😄

Suggested:

- Here an interesting reference about Choosing the Right Metric for Evaluating Machine Learning Models.
- And here an article that discusses about What metrics should be used for evaluating a model on an imbalanced data set

✓

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.

Excellent strategy to solve the problem. This approach could conduct your work to impressive results 😄

Suggested: You could try the custom models, but you might use then in a cascaded way. That is: you could use a model to predict some some intermediate variable (a temporary target, like deal_demand or deal_quality). Then you could add this intermediate variable as a feature of another prediction model. I used this approach in some problems and helped a lot, but it is an empirical method. Also, in regression problems, it is always important to analyze the **sensitivity** of your final model (the difference in results when the model is tuned with different training sets). It always a good ideia try it.

And remember to use a business language (less technical) to explain the data and the results to your reader.

✓

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 Udacity's template was properly followed, and the proposal is well written. One of the best that I reviewed 😄