5 specifications require changes

Dear student,

This was a solid submission; you've done a great job articulating your thought processes and implementation details. You clearly have a good grasp of the foundations of machine learning. I hope you continue to explore new machine learning projects - it's a huge field, and there's a ton more to learn.

Please correct the fields where the requirements aren't met and resubmit. Thanks.

Definition

Student provides a high-level overview of the project in layman's terms. Background information such as the problem domain, the project origin, and related data sets or input data is given.

Nice work here with your opening section, as you have given good starting paragraphs to outline the project and have provided background information on the problem domain. A solid introduction to the problem you're solving, with a good focus on its origin.

Suggestions and comments

An important part of preparing for a project is conducting some form of literature review. This gives you a sense of what work has been done in the field, what techniques were used, and what kind of performance you might expect to see. Is there any interesting academic research that has been published on this topic? If so, it should be cited here. If not, make a note of that.

Metrics used to measure the performance of a model or result are clearly defined. Metrics are justified based on the characteristics of the problem.

Excellent choice of metrics. Congrats!

Suggestions and comments

I think these links might be useful:

- https://towardsdatascience.com/20-popular-machine-learning-metrics-part-1-classification-regressionevaluation-metrics-1ca3e282a2ce
- https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234

The problem which needs to be solved is clearly defined. A strategy for solving the problem, including discussion of the expected solution, has been made.

The input and output are well defined, which makes for a solid problem statement. Your approach to the problem is clear and certainly suitable.

Analysis

If a dataset is present, features and calculated statistics relevant to the problem have been reported and discussed, along with a sampling of the data. In lieu of a dataset, a thorough description of the input space or input data has been made. Abnormalities or characteristics of the data or input that need to be addressed have been identified.

Nice work here describing the features in the dataset.

Algorithms and techniques used in the project are thoroughly discussed and properly justified based on the characteristics of the problem.

This is a bit confusing. If you did not use DeepAR, why is it so extensively mentioned in the introduction? Which algorithms did you use? Only curve fitting from scipy? You need to provide more information on the algorithms.

Also, if you only have 151 points, you could augment data in between and add a bit of noise. Go by half days or third days and add noise to the augmented data.

A visualization has been provided that summarizes or extracts a relevant characteristic or feature about the dataset or input data with thorough discussion. Visual cues are clearly defined.

Very good visualizations. It captures your data well.

Student clearly defines a benchmark result or threshold for comparing performances of solutions obtained.

I do not believe you understand the point of the benchmark. You can use a model like ARIMA if you are not using it in your analysis.

As a suggestion, I leave you with some links:

https://medium.com/levvel-consulting/define-benchmark-deploy-6a8d0fb0decd

https://towardsdatascience.com/benchmarking-simple-machine-learning-models-with-featureextraction-against-modern-black-box-80af734b31cc

Methodology

The process for which metrics, algorithms, and techniques were implemented with the given datasets or input data has been thoroughly documented. Complications that occurred during the coding process are discussed.

You only implemented a curve fit? You will also have to implement a benchmark to compare.

The process of improving upon the algorithms and techniques used is clearly documented. Both the initial and final solutions are reported, along with intermediate solutions, if necessary.

All preprocessing steps have been clearly documented. Abnormalities or characteristics of the data or input that needed to be addressed have been corrected. If no data preprocessing is necessary, it has been clearly justified.

Nice work documenting the preprocessing steps. I would remove the code from the documentation; I feel it looks tacky.

Results

The final results are compared to the benchmark result or threshold with some type of statistical analysis. Justification is made as to whether the final model and solution is significant enough to have adequately solved the problem.

You need to resubmit with the proposal review link in the submission notes, as the directions state.

Learning Capstone project rubric. Be sure to review this rubric thoroughly and self-evaluate your project before submission. All criteria found in the rubric must be *meeting specifications* for you to pass.

Submission Files

At a minimum, your submission will be required to have the following files listed below. If your submission method of choice is uploading an archive (`*.zip`), please take into consideration the total file size. You will need to include

- Your **capstone proposal** document as **proposal.pdf** if you have completed the pre-requisite *Capstone Proposal* project. Please also include your review link in the *student submission* notes.
- A project report (**in PDF format only**) addressing the five major project development stages. The recommended page length for a project report is approximately *nine to fifteen pages*. **Please do not export an iPython Notebook as PDF for your project report.**
- All development Python code used for your project that is required to reproduce your implemented solution and result. Your code should be in a *neat and well-documented format*. Using iPython Notebooks is strongly encouraged for development.
- A `README` documentation file which briefly describes the software and libraries used in your project, including any necessary references to supporting material. If your project requires setup/startup, ensure that your `README` includes the necessary instructions.
- Any additional supporting material such as datasets, images, or input files that are necessary for your project's development and implementation. If these files are too large and you are uploading your submission, **instead provide appropriate means of acquiring the necessary files in your included** `README` .

I will wait to grade this section after a proper benchmark.

The final model's qualities—such as parameters—are evaluated in detail. Some type of analysis is used to validate the robustness of the model's solution.

I will wait to grade this section after a proper benchmark.