UDACITY

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

DISCUSS ON STUDENT HUB

Machine Learning Capstone Project

REVIEW

CODE REVIEW

HISTORY

Requires Changes

1 SPECIFICATION REQUIRES CHANGES

Fantastic project! This was incredibly well-written and well-formatted. There's just one section that needs some more detail, but this is sincerely a fantastic foundation and I have no doubt your next iteration will be sufficient for a pass.
Keep up the great work!

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.

Great start! You've provided a concise explanation of your domain and why it's important.
One thing I would recommend is providing some insight into existing ML research into this domain - what does the current state of the art look like?

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.

Fantastic work in this section. Great problem statement, and nice job with the outline of the proposed solution strategy.

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

Really great work here. You've provided a list of *all* metrics you plan to use to gauge your model performance, and each of them have been explained thoroughly and justified.

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 about the data or input that need to be addressed have been identified.

Excellent job.

- Some discussion of the class imbalance has been provided
- A thorough explanation of how the dataset is represented has been given

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 cool visualization! This is certainly relevant to the project and supplementary comments have been provided to offer additional insights.

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

Solid start here! You've done a good job explaining the architecture and parameters relevant to the algorithm, but you actually need to go a bit broader here as well - what is a CNN and why does it work? You should ensure the reader has some understanding of *why* you want to use a CNN for this problem.
I would also recommend spending some time discussing what transfer learning is and why it's relevant here.

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

This is definitely a solid benchmark. It's always interesting to use human performance as the baseline.

Methodology

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

Excellent work in this section. All preprocessing steps have been thoroughly discussed; a would-be replicator would have no trouble implementing the same preprocessing pipeline based on your description here.

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.

Excellent job. Your attention to detail here is awesome.
Not only have you discussed the actual implementation in detail, but you've made some comments on things you've tried that weren't successful and hypothesized as to why that might be.

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.

Great job. The refinement process is clear and all intermediate scores have been discussed.

Results

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.

Excellent evaluation here. Model parameters have been thoroughly discussed.

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

justification is made as to whether the final model and solution is significant enough to have adequately solved the problem.

Conclusion

A visualization has been provided that emphasizes an important quality about the project with thorough discussion. Visual cues are clearly defined.

I like your choice of visualization here - this is certainly relevant to the final model.

Student adequately summarizes the end-to-end problem solution and discusses one or two particular aspects of the project they found interesting or difficult.

Nice job! It's clear from your reflection that you've learned a lot throughout the project. Kaggle competitions are often very tough, and the winners usually have substantial access to compute power or are willing to spend some money to iterate over and over. Given that, ranking in the 1400s is certainly an accomplishment, and I have no doubt you could've pushed your score even further given more time and resources.

Discussion is made as to how one aspect of the implementation could be improved. Potential solutions resulting from these improvements are considered and compared/contrasted to the current solution.

Image preprocessing is definitely a solid choice. I've found this to be the best bang-for-your-buck avenue for improving the final model performance, and I think you'd find that implementing some of the extra preprocessing steps you mentioned here could really provide a significant boost to your final result.

Quality

Project report 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 to complete the project are cited and referenced.

Code is formatted neatly with comments that effectively explain complex implementations. Output produces similar results and solutions as to those discussed in the project.

RESUBMIT

DOWNLOAD PROJECT

Best practices for your project resubmission

Ben shares 5 helpful tips to get you through revising and resubmitting your project.

Watch Video (3:01)

RETURN TO PATH