UDACITY

Back to Machine Learning Engineer Nanodegree

Machine Learning Capstone Project

REVIEW

HISTORY

Requires Changes

5 SPECIFICATIONS REQUIRE CHANGES

This is an interesting choice of project and you have made a good attempt of documenting the steps in the report. However, it needs to be more detailed. Follow the feedback carefully and try to present the implementation in such a way that your results can be reproduced without a prior understanding of these topics. We look forward to your next submission, keep up the hard 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.

Good job giving a compelling introduction to the project and the problem domain, as well as discussing the past work done on the topic.

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 content in the "problem statement" section is better suited for the problem overview. In this section, you need to state the problem in machine learning terms. Eg:

need to state the problem in machine learning terms. Eg:

This is a classification (or any other machine learning task like regression, unsupervised learning, etc.) problem where the model takes X as input and produces the expected output Y. You would then present a brief summary of how the problem will be solved, which would include naming the main algorithms that will be used to arrive at the solution.

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

The evaluation metrics are explained with the required detail, but you also need to justify why these metrics were used for the current problem. Why were other metrics like accuracy score not considered?

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.

Nice work explaining the input data clearly by discussing the important characteristics, abnormalities and the features of the dataset. The origin of the dataset and some descriptive statistics have also been included along with a sample of the training 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.

A few relevant visuals have been used to explore the dataset, and the plots are clear and easy to interpret.

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

RNN and the other algorithm are explained quite clearly and with the required detail.

> In this section, you need to state the problem in machine learning terms. Eg:

This is a classification (or any other machine learning task like regression, unsupervised learning, etc.) problem where the model takes X as input and produces the expected output Y. You would then present a brief summary

where the model takes X as input and produces the expected output Y. You would then present a brief summary of how the problem will be solved, which would include naming the main algorithms that will be used to arrive at the solution.

The use of RNN should be based on the problem and dataset characteristics. Try to explain why this would be a good candidate algorithm specifically for the current case.

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

Using the same model with a smaller set of data doesn't really form a good benchmark. A better alternative would be to use a simpler classifier or use the top 10-15% percentile score from the Kaggle leaderboard.

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.

The pre-processing steps are explained in detail and all abnormalities in the input data have been corrected to prepare it for the main implementation.

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.

Great job describing the implementation with enough detail for a skilled individual to be able to replicate your results. The challenges faced during implementation have also been discussed.

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.

Nice work explaining the refinements that were considered and comparing the results from the initial, intermediate and final implementations.

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.

The final parameters are evaluated and discussed, but you also need to discuss the model's robustness here. For eg: comparing the performance on validation and test data (hold out data that was never used during training or validation). The main idea is that a sensitivity analysis should be conducted which shows that the model is robust to small perturbations in the input data.

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.

The final results are compared to the benchmark to comment on whether the solution is significant enough to solve the problem.

Conclusion

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

This section is missing. You need to include one or more visuals that highlight an important attribute about the final model (like feature importance, learning curve, etc.) or the project. You also need to add insights from these visuals.

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

You've done well in discussing the interesting/challenging aspects of the project. However, you also need to summarize the project end-to-end. This would include stating the problem that you set out to solve, the approach that was taken to solve the problem and the final solution that was obtained. In other words, this would be a brief description of the entire workflow that was followed to implement the project.

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.

These are certainly some good improvements to consider!

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.

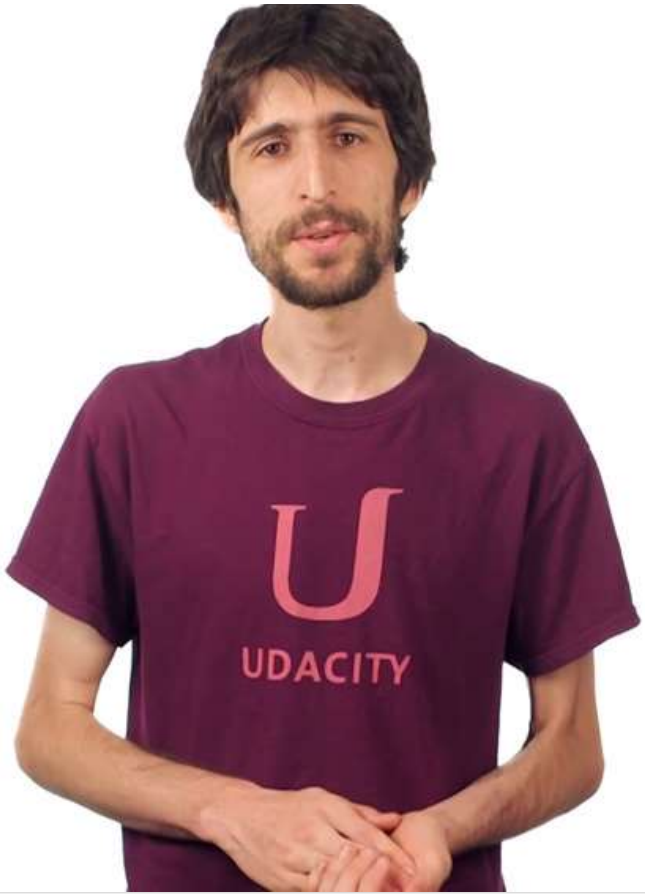
The report is well structured and easy to read. All the sources are cited and referenced.

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

The code produces the described results in the report. It is commented appropriately and formatted clearly to be understood without much difficulty.

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

Rate this review

Student FAQ