UDACITY

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

Machine Learning Capstone Project

REVIEW

CODE REVIEW

HISTORY

Meets Specifications

🏆 Great job completing the capstone project! As you think about applying ML to other use cases, you can also read through this list of ML breakthroughs in 2018 for inspiration.

And if this is your final step to complete the nanodegree program, congratulations and best of luck with whatever projects you decide to work on next! 😎

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.

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.

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

Great work describing the metrics used to evaluate the model's performance.

For more info, you can also see this post on classification metrics.

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.

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.

Excellent job discussing the dataset and including the exploratory visualizations to help readers understand the data. 😎

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

If interested you can also see this Kaggle data science glossary for examples of various ML models and approaches:

- https://www.kaggle.com/shivamb/data-science-glossary-on-kaggle

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

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 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 and documenting your training results. 😅

In general it's a good idea to record the results of all your model training experiments, especially "failures" that might need to be debugged. You can learn more about this and other thoughts from google in their effective ML guidelines.

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.

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.

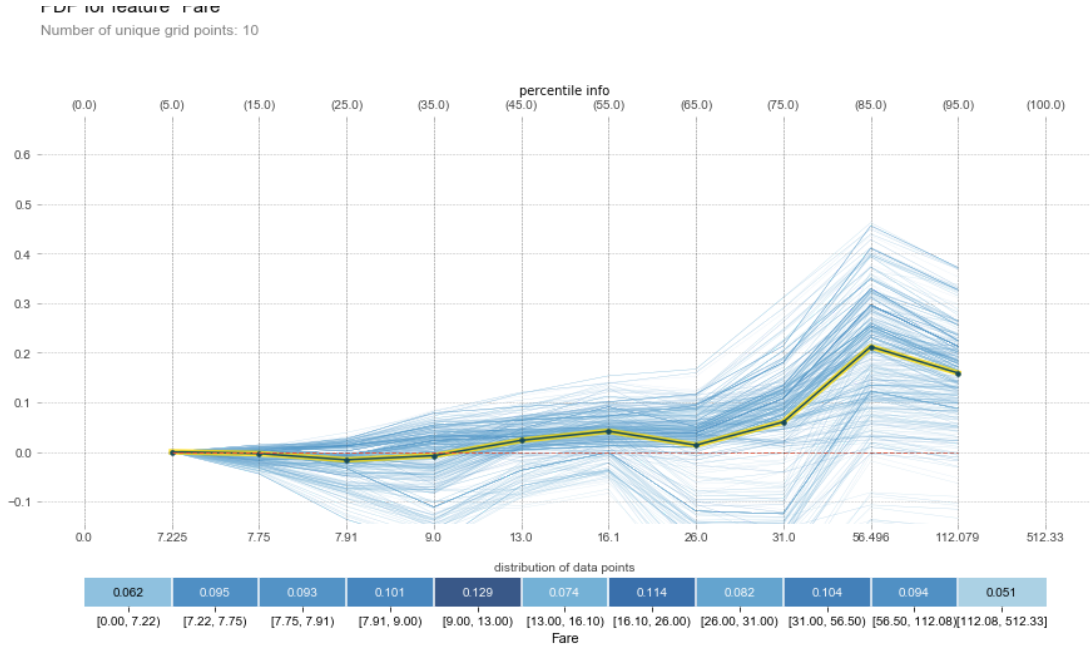
Good job evaluating your final model results, and examining the robustness of the solution with the resampling techniques 😎

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.

Conclusion

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

In order to visualize the effects of individual features on a target variable, you could also try creating Partial Dependence Plots, or Accumulated Local Effects (ALE) Plots.



[PDPBox]

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

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.

Excellent job reflecting back on your experience with the project and discussing improvements that could be made with model tuning and cross validation.

You've clearly expanded your skills beyond the core concepts introduced in the MLND and are ready to tackle new problems like this with machine learning. 😀

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.

DOWNLOAD PROJECT

RETURN TO PATH