

Course Name: Programming and Problem Solving for Data Analytics

Course Code: 04802-E

Instructor: George Okeyo.

TAs: [Faustin, Janvier, Pacifique]

Assignment Title: Assignment IV- Build ML Models.

Report Title: Building a Credit Approval Model using Neural Networks

Nmaes: Yves Himbaza Mugiraneza

Andrew: ymugiran

Submission Date: 09/12/2021

Hyperparameter tuning model vs random model in neural network

I. Background and Problem Description

Many models incorporate parameters that, while crucial, cannot be calculated directly from data. The tuning parameters (also known as hyperparameters²⁷) are significant because they frequently determine the model's complexity and hence impact any variance-based trade-off that may be made^[1]. This report discusses the comparison between two neural network models, one being the model without hyperparameter tuning and other one with hyperparameter tuning.

In machine learning , hyperparameter optimization or tuning is the problem of choosing a set of optimal hyperparameters for a learning algorithm. A hyperparameter is a parameter whose value is used to control the learning process^[2]. Neural networks on the other hand are Artificial neural networks (ANNs), usually simply called neural networks (NNs), are computing systems inspired by the biological neural networks that constitute animal brains .

II. Methods

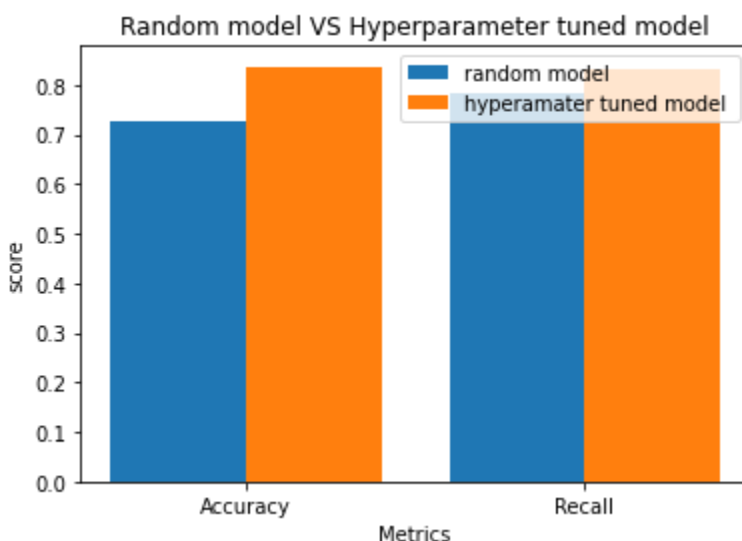
Several methodologies have been employed in this model comparison. Python programming language and its well-known machine learning packages such as pandas, numpy, scikit learn, and others have been used for data preparation, feature engineering, hyperparameter tuning , and building and evaluating ML models. A credit approval dataset was also utilized. Finally, the jupyter notebook was used as well because of its ease of use and extensive feature set.

The procedure began with data preparation, in which data was read and loaded. Second, exploratory data analysis (EDA) is used to comprehend the data, identify missing data, and detect outliers. Preprocessing, Feature Selection, and Engineering

were all carried out. At this point mainly , missing data is handled by removing the rows with missing values, dealing with outliers with z-score, and categorical variables encoding. Finally, model building and assessment were carried out. At this point, the Python package scikit-learn was used to build a multi-layer neural network model using random parameters (parameters picked at random by the author). This model was then tested using 10-fold cross validation, and the performance recall and accuracy metrics were calculated. A second model was also generated with scikit-learn, but with hyperparameter tuning. This second model was likewise tested using 10-fold cross validation and its performance was determined using recall and accuracy measures. Recall measures were employed because they avoid false negatives, which are subsequently used to evaluate the model's capacity to grant credit cards solely to deserving individuals. Accuracy, on the other hand, was chosen as the optimum metric to utilize when the target class is balanced.

III. Results and Discussion

After building both a random model (the model whose parameters were chosen at random) and a model that was subjected to hyperparameter tuning, the performance results were gathered and reported in graph1.



graph 1

The graph:1 clearly illustrates that the performance of the model that has been hyperparameter tuned is slightly better than the other models due to improved recall and accuracy.

IV. Conclusion

In conclusion, this paper compared the hyperparameter tuned neural network model to the random neural network model. As predicted, the model using hyperparameter tuning won the race. This is because hyperparameter tuning using a greedy search cycled through multiple parameter combinations and chose the optimal one.

V. Reference

- [1]M. Johnson, "3.5 Tuning Parameters and Overfitting | Feature Engineering and Selection: A Practical Approach for Predictive Models", *Feat.engineering*, 2021. [Online]. Available: <http://www.feat.engineering/tuning.html>. [Accessed: 09- Dec- 2021].
- [2]"Hyperparameter optimization - Wikipedia", *En.wikipedia.org*, 2021. [Online]. Available: https://en.wikipedia.org/wiki/Hyperparameter_optimization. [Accessed: 09- Dec- 2021].
- [3]"Artificial neural network - Wikipedia", *En.wikipedia.org*, 2021. [Online]. Available: https://en.wikipedia.org/wiki/Artificial_neural_network. [Accessed: 09- Dec- 2021].