Testing Sensitivity of Results to Hyperparameters in Deep Learning

Sajad Sabzi
MohammadReza AhmadiTeshnizi

December 18, 2023

Introduction

Deep learning models are inherently **sensitive to hyperparameters**, and understanding this sensitivity is crucial for model optimization and robustness. This exploration addresses the importance of testing the sensitivity of results to hyperparameters in deep learning and the challenges associated with it.

Testing Sensitivity of Results to Hyperparameters in Deep Learning

Testing the sensitivity of deep learning models to hyperparameters is an essential aspect of model development. Classical hyperparameter optimization methods focus on finding the best hyperparameters for a given model but often do so without providing insight into the relative importance of these parameters. This lack of interpretability is a significant drawback as understanding the impact of each hyperparameter can be crucial, especially when dealing with multiple objectives such as accuracy, execution speed, and memory consumption.

A promising approach to address these concerns is the integration of hyperparameter optimization with **hyperparameter analysis**. Sensitivity analysis, a method of assessing the effect of input variables on a function's output, can be utilized to understand the impact of hyperparameters on a neural network's performance. This approach can offer insights into the relative importance of hyperparameters, help in selecting values that balance multiple objectives, and identify key areas for optimization.

Challenges and Limitations

One of the primary challenges in testing the sensitivity of hyperparameters in deep learning is the complexity of the models and the hyperparameter space. Hyperparameters can be of different natures (categorical, discrete, boolean, continuous), interact with each other, and some may not be directly involved in every configuration. This complexity can lead to a 'curse of dimensionality,' making the optimization process more challenging.

Furthermore, certain sensitivity analysis methods, like those based on variance analysis, can be computationally expensive and may not always provide the required detailed information. Hyperparameter spaces often require normalization due to the diverse nature of hyperparameters, adding another layer of complexity to the analysis. Additionally, the interactions between hyperparameters can significantly affect the outcome, necessitating a more nuanced approach to analysis.

Conclusion

Testing the sensitivity of results to hyperparameters in deep learning is not only beneficial but necessary for developing robust and efficient models. However, the complexity of deep learning models and their hyperparameter spaces presents significant challenges. Advanced techniques like sensitivity analysis, particularly methods like the **Hilbert-Schmidt Independence Criterion (HSIC)**, provide a promising pathway forward. These methods can offer insights into the relative importance of hyperparameters, balance multiple objectives, and reduce unnecessary complexity in the hyperparameter space.

References

1. arXiv - Goal-Oriented Sensitivity Analysis of Hyperparameters in Deep Learning. Link

2. IEEE Xplore - Sensitivity Analysis for Deep Learning. Link

3. Springer Link - Hyperparameter Optimization for Deep Neural Network Models. Link