

# Justification for Skipping Cross-Validation in Problem Statement-1

Shubham Kumar Patel

January 27, 2025

## Abstract

Cross-validation is often used to evaluate model performance, but when using XGBoost with default parameters, cross-validation may not always be necessary. This document discusses why, in such cases, additional cross-validation may be redundant, saving both time and computational resources.

## 1 Introduction

XGBoost is a highly effective machine learning algorithm that often performs well with default parameters, such as a learning rate of 0.3, max depth of 6, and 100 estimators. In many scenarios, these default settings lead to satisfactory performance, reducing the need for cross-validation, which is typically used for hyperparameter tuning or model evaluation.

## 2 Default Parameters in XGBoost

XGBoost's default parameters are designed to work well in a wide range of tasks. The default settings are selected based on extensive experimentation across various datasets and are robust enough to give good performance for many problems right out of the box. For instance, a moderate learning rate and a reasonable tree depth generally yield good generalization without overfitting. Therefore, if the model's performance is satisfactory on a hold-out validation set, further tuning through cross-validation may not offer substantial improvements.

## 3 Why Skip Cross-Validation?

### 3.1 Computational Efficiency

Cross-validation involves training the model multiple times on different subsets of the data, which can be computationally expensive, especially with large datasets and models like XGBoost. If the model already performs well with default parameters, repeating this process may not provide new insights. In such cases, skipping cross-validation saves both time and computational resources.

### 3.2 Stable Performance

XGBoost tends to offer stable and consistent performance across a variety of tasks. When the model already demonstrates satisfactory performance on the validation set with default parameters, running cross-validation might only confirm what is already evident: the model is not overfitting and generalizes well. If performance is good on the test set and there is no evidence of overfitting, additional cross-validation is likely unnecessary.

## 4 Conclusion

In conclusion, when using XGBoost with default parameters and achieving satisfactory results, there is often no need for further cross-validation. The default settings are often sufficient for robust model performance,

and skipping cross-validation can save valuable time and computational resources. Thus, in many cases, it is reasonable to rely on XGBoost's default configuration and move directly to deployment.

## References

- [1] ChatGPT, OpenAI. "Justification for Skipping Cross-Validation in XGBoost with Default Parameters."  
*OpenAI, ChatGPT*, 2025.