**Title:** Pavement Performance Prediction using Machine Learning: Supervised Learning with Tree-Based Algorithms

**Manuscript ID:** 5165

**Round:** 1

**RESPONSE TO REVIEWER #1**

We sincerely appreciate your comments and suggestions. Please find below our response.

**Comment 1:**

“It has been seen that almost 1000 rows were deleted and also Boosting algorithm is strong against outliers. In this study, no investigation or cleaning of outliers was performed. So if possible then please check.”

**Reply:**

We value your insightful question regarding the treatment of outliers in our study. Based on your suggestion, we have revised the relevant section of the manuscript to clarify our approach to handling outliers. In the updated text, we now explain that the outliers were not cleaned in this study. Some of the values identified as outliers were actually more degraded pavements with IRI values greater than 3.5. We have also added a note on future work, emphasizing the importance of employing techniques to rebalance the datasets to improve the distribution. This revision highlights the expectation that the models developed on these data will likely yield better results for new or good pavements due to the characteristics of the training data. Finally, we would like to thank you for your evaluation and comment, as it was instrumental in improving the quality of our work.

**RESPONSE TO REVIEWER #2**

We are grateful for your valuable comments and suggestions. Please find below our responses to your questions.

**Comment 1:**

“The introduction should be improved to include the motivation and objectives of the paper more clearly.”

**Reply:**

In response to your recommendation, we have revised the introduction to articulate the motivation and objectives of the paper. The updated introduction now explicitly outlines the aim of understanding the best machine learning techniques for predicting road pavement quality, identifying the most effective algorithms and features, determining the amount of data required to train a good model, and exploring the necessary pre-processing and treatment of the data.

**Comment 2:**

“The manuscript doesn't discuss the results of the study. More discussions regarding the results should be added.”

**Reply:**

We have expanded the discussion of the results in the manuscript, based on your suggestion. The updated text now provides more detailed information on the performance of each model (Decision Tree, Random Forest, and XGBoost), addressing their strengths and limitations, and offers insights into potential improvements for future work.

We believe that these changes have greatly enhanced the quality of our paper and thank you for your constructive feedback.