Development of Pavement Performance Prediction Model for Local Roads by Using Statistical Analysis and Machine Learning Techniques

Abdul Nizar Nafrees Ahamed[1], Kelum Sandamal[2]

Abstract

Sri Lankan Road network consists over 117,000 kms of roads which are comprised of national roads and highways (10.4%), provincial roads (16.3%) and the remaining are local roads. In general, local roads are managed by local authorities such as urban or municipal councils. Even though traffic volume is low in these roads, pavement condition is deteriorated with the time hence timely maintenance is a vital factor to consider. There are numerous pavement condition evaluation metrics available to measure the pavement condition such as roughness, distress and skid resistance. Recently, several countries have developed their own pavement performance models to predict roughness progression during lifetime by considering several factors such as distress condition, pavement age, environmental factors etc. Therefore, this research aims to develop an accurate pavement performance prediction model for Sri Lankan local roads.

Moreover, this study aims to close the gap of not having proper pavement performance model for local roads by using statistical analysis and machine learning techniques that consider the distinctive features of Sri Lanka's local roadways. International roughness index (IRI), cracks and raveling density are considered as pavement performance indicators while these indicators compared with pavement age and environmental factors such as temperature and annual precipitation. The research methodology comprised with several steps including data collection, data pre-processing, statistical model development and machine learning model development. The predictor variable is IRI and distress densities (cracking % and raveling %) while using latest rehabilitation year, road inventory data and environmental data as independent variables. IRI data and distress data were collected from local road authorities and using manual data collection methods along with road inventory data. Environmental data such as annual precipitation, average temperature was collected from open-source weather forecasting software and metrological department. Data analysis was carried out using both regression analysis and machine learning techniques. Firstly, a linear regression analysis was conducted between IRI, cracks and raveling with pavement age separately to identify whether there is a significant relationship among those. The results reveal that IRI as the most suitable metric to predict pavement condition against pavement age with a coefficient of determination (R2) value of 0.76. The second-best predictor variable found as raveling with an R2 of 0.65 while cracks shown the least R2 of 0.54. However, all dependent variables show a good relationship with pavement age with R2 more than 0.5. Further, a multiple linear regression analysis has been carried out by using IRI, crack% and ravelling % as the dependent variables while using pavement age, temperature and annual precipitation as independent variables. The multiple regression models shown increment in model predictability compared with individual variables having 0.78, 0.58 and 0.59 of R2 values separately for IRI, crack% and ravel %. Furthermore, a machine learning model was developed using a random forest (RF) regressor as the machine learning algorithm to predict IRI progression against the predictor variables. In the training data, 70% of the data was used while others were used for testing data. The results show that R2 of 0.79 for testing data with a mean absolute error (MAE) of 0.67 for the RF model. Shapley's Additive explanations

(SHAP) analysis showed that pavement age as the most influencing variable with a feature importance of 0.9 while pavement type, temperature and precipitation come afterward with a mean SHAP value of 0.55, 0.35 and 0.27 respectively. Providing precise justifications for the decision-making mechanisms within accurate models through SHAP analysis enhanced the confidence of both road users and domain experts in the machine learning models' predictions. Additionally, this research showcases the superior effectiveness of AI-based techniques over traditional regression analysis in predicting IRI. In summary, by adopting this methodology, road authorities can strategies for prompt maintenance, thus preventing expensive and extensive rehabilitation efforts.

Keywords: *Machine learning, AI based methods, pavement condition evaluation, local roads, international roughness index*

Authors Details;

1. Undergraduate, Sri Lanka Institute of Information Technology. nafreesahamed@gmail.com
2. Lecturer, Sri Lanka Institute of Information Technology. kelumsanadamal@gmail.com