Abstract No: SO-07

A machine learning approach to predict profitable vegetable crop and land requirements in Sri Lankan agriculture

S. A. H. Nethmini[1*] and W. A. C. Weerakoon[1]

[1]Department of Statistics & Computer Science, University of Kelaniya, Sri Lanka
hirushinethmini1999@gmail.com*

More than 80% of the Sri Lankan population depends on agriculture as their livelihood. The main challenges in the agriculture sector are how to select profitable vegetable crops, reduce excess vegetable stocks, and minimise product loss in vegetable varieties. Traditionally, agricultural decisions are made without adequate knowledge of their expected outcomes, causing low productivity. In this context, technology-based decisions in the agriculture sector provide accurate and timely information about weather, market trends, and prices. This research study aims to provide accurate decisions to the challenges by predicting the most profitable vegetable crop and required land extent for cultivation. This analysis considers a range of factors, including weather conditions, cultivation expenses, seasons, location, production, and crop cultivation data. The dataset contains a limited number of records gathered from farmers and agriculture institutes. Multiple machine learning algorithms were employed to develop accurate crop prediction models, namely, random forest (RF), multinomial logistic regression (MLR), and long short-term memory (LSTM). In addition, the extent of land required for cultivation was predicted using linear regression (LR) and random forest regressor (RFR). The multinomial logistic regression archives a higher accuracy of 0.84 for crop prediction, and the random forest regressor algorithm gives an accuracy of 0.83. The LSTM model gives an R^2 score of 0.831. The model provides the profitable crop as an output, and it is used as an input to the land extent prediction system. The application predicts the land extent for selected crops using RFR and LR. RFR gives a higher R^2 score of 0.956 compared to the LR model ($R^2 = 0.45$), demonstrating its superior fit to the data. The. The root mean squared error of the RFR was 23.95, whereas the LR model had 89.29. Using advanced machine-learning techniques, the research provides valuable predictions on profitable crops and land extent to be cultivated in a particular area to increase productivity. This will help the farmers and stakeholders to make accurate decisions to maximise profitability and ensure efficient resource utilisation. Furthermore, the developed model can be expanded with various crop categories to build profitable farming on a broader scale. The impact of this work extends to enhancing food security and reducing poverty in Sri Lanka by allowing farmers with the tools and knowledge to make correct decisions at the correct time.

Keywords: Linear Regression, Long Short-Term Memory, Multinomial Logistic Regression, Random Forest, Random Forest Regressor