Regression-based Modeling of the Relationship between Weather and Tea Production in Sri Lanka

M.D.C.De.S.Jayatilake
(Reg. No.: MS20902216)

M.Sc. in IT

Supervisor: Dr.Windhya Rankothge
Assistant Professor
Computer Systems Engineering Faculty of Computing

December, 2021

Faculty of Graduate Studies
Sri Lanka Institute of Information Technology

Acknowledgements

Leading, I would like to convey my sincere gratitude to my supervisors Dr.Windhya Rankothge, Assistant Professor, Computer Systems Engineering Faculty of Computing, Sri Lanka Institute of Information Technology and Dr. Dharshana Kasthurirathna, Assistant Professor, Computer Science & Software Engineering Faculty of Computing, Sri Lanka Institute of Information Technology (SLIIT) for the continuous help of my MSc research and study. Supervisors endurance, motivation, keenness, and enormous of knowledge guidance helped me a lot in all the time of research.

My sincere thanks also go to the Mr. Samantha Rajapaksha, MSc Coordinator (IT), Senior Lecturer, Information Technology Faculty of Computing, Sri Lanka Institute of Information Technology. I would express my special great to Prof. Rahula Attalage , Dean - Faculty of Graduate Studies & Research.

I would like to thank to my loving parents and family members for giving birth to me at the first place and behind me spiritually throughout the life. Last but not the least; unfortunately, my Father was passed away 08th April 2021 during my research.

Further , I would like to greatly thankful to my loving wife, Thilina Pinidiya who had introduce me the first glance of SLIIT - MSc degree and my daughter, baby Mithuki Methara for encouraging me to achieve this journey success-fully.

Table of Contents

List of Figures

List of Tables

Abstract

The research study shows the implementation of regression-based tea–weather prediction models for the tea production in Sri Lanka based on six weather parameters such as rainfall, , temperature minimum, temperature maximum, relative humidity, average wind speed, and sunshine hours. The statistics of the eight agro climate geographical regions of the UVA province, which contribute to the more productive contribution tea production of the country. They were used for this research and the significance of weather parameters on the tea production was explored by employing Random Forest algorithm and the variable importance of each of them was determined and further correlation in the middle of the climate parameters was examined. The results indicate that the Minimum Relative Humidity, Rainfall and the maximum Temperature throughout the tea plantation period are the most influential weather parameters. Furthermore, regression analysis implementations were applied for the Random Forest (RF), Linear Regression (LR), Multiple Linear Regression (MLR), and Support Vector Machine (SVM). According to the results, RF is the most reliable and accurate model for the prediction of tea production in Sri Lanka. UVA province prediction model accuracy (R square score) is 88.58% of the eight agro climate districts of the UVA province and region-wise prediction tea-production model accuracy is low parentage according to the results. Further RF, MLR and SVM, Machine Learning implementation already trained and validated for the same dataset. Although the results indicated that other models low percentage in comparison to the RF implantation model. The research regression analysis already applied for RF and LR for the region-wise of the UVA province Sri Lanka. Concluded that the final output of the research same process can be applied to the tea- weather prediction model of all the tea growing areas of the country.

Key words: Agro climate, Regression model, Random Forest, Tea production, Tea-weather