



Predicting Crop Yields: Harnessing IoT Sensor Data for Accurate Forecasting and Sustainable Agricultural Planning

A.Kulesh.B.M.Abhayasundara
(Reg, No.: MS23006584)

A THESIS
SUBMITTED TO
SRI LANKA INSTITUTE OF INFORMATION TECHNOLOGY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

December 2024

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Anuradha Jayakody

Approved for MSc. Research Project:

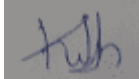
MSc. Programme Co-ordinator, SLIIT

Approved for MSc:

Head of Graduate Studies, FoC, SLIIT

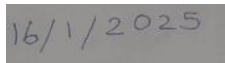
DECLARATION

This is to certify that the work is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources). The work has not previously been submitted in any form to the Sri Lanka Institute of Information Technology or to any other institution for assessment for any other purpose.



Sign:

A.Kulesh.B.M.Abhayasundara



Date:

ABSTRACT

Development of a Framework for Legal Reseach using Text Mining

A.Kulesh.B.M.Abhayasundara

MSc. in Information Technology

Supervisor: Prof. Anuradha Jayakody

December 2024

This research introduces a novel approach to precision agriculture by integrating IoT technology and machine learning algorithms to enhance crop health prediction, suitability forecasting, and crop yield prediction for rice, maize, mango, and watermelon. Utilizing real-time data from in-field sensors monitoring soil moisture, nitrogen, phosphorus, potassium, pH, temperature, humidity, and rainfall, the Agro-Sense system enables continuous and dynamic crop monitoring. This innovative system provides timely, data-driven insights previously unavailable with traditional models. Machine learning algorithms, specifically XGBoost and RandomForestClassifier, predict crop health and forecast suitability with improved accuracy over conventional methods. By incorporating real-time sensor data, the Agro-Sense mobile app optimizes nutrient management, reducing fertilizer waste. This research addresses the limitations of previous models, which often rely on static datasets and fail to account for diverse environmental factors. Unlike conventional methods, Agro-Sense continuously adjusts predictions based on evolving environmental parameters, making it highly adaptable to different crops and regions. The system's specific health assessments and suitability forecasts offer farmers actionable insights for informed, sustainable agricultural decisions. Agro-Sense presents a scalable, IoT-driven framework that bridges gaps in existing crop prediction methods, enhancing both crop health prediction and suitability forecasting for sustainable farming practices. Future advancements should expand predictive capabilities to optimize resource use and enhance crop health across diverse environments.

Keywords - Crop yield prediction, IoT-based agriculture, Machine learning in agriculture, Nutrient management, Precision agriculture, Real-time data, Sustainable farming.

ACKNOWLEDGEMENT

While at the Sri Lanka Institute of Information Technology, I have been fortunate to have exceptional mentors and unwavering support from my family, friends, and colleagues.

I am deeply grateful to Prof. Anuradha Jayakody, who has been an outstanding mentor. His guidance, constructive criticism of my ideas and writing, access to his extensive network of contacts, support, and the freedom to pursue my own projects under his research account have been invaluable throughout this journey.

I would like to extend my heartfelt gratitude to my father, Pulasthi Abhayasundara, and my mother, Disna Weerawardana, for their unconditional love, encouragement, and belief in my potential. Their unwavering support has been the foundation of my success. My uncle, Prof. Praneeth Abhayasundara, and my aunt, Subani Weerawardana, have also been constant sources of inspiration and guidance, and I am incredibly thankful for their contributions to my academic and personal growth.

I am also indebted to my colleague, Mrs. Ramani Jayasooriya, for her collaboration, insights, and encouragement throughout this journey. Her support has been instrumental in helping me navigate challenging times during my research.

To all of you, I extend my deepest thanks for your unwavering support and belief in my work.

TABLE OF CONTENTS

DECLARATION	ii
ABSTRACT	iii
ACKNOWLEDGEMENT	iv
TABLE OF CONTENTS	v
List of Figures	viii
List of Tables.....	x
Chapter 1 Introduction.....	1
1.1 Background.....	1
1.1.1 Importance of Crop Yield Prediction.....	1
1.1.2 IoT usage in Contemporary Agriculture.....	2
1.1.3 AI and ML applied in innovative practices within todays agriculture.....	3
1.1.4 The Influence of Modern Agriculture on Crop Yields	4
1.2 Research Problem.....	7
1.2.1 Why This Research Matters	8
1.3 Objectives And Research Questions	9
1.4 Study Scope.....	10
1.5 Thesis Structure Overview.....	10
Chapter 2 Literature Review	12
2.1 Introduction to IoT in Agriculture	13
2.2 IoT Sensor Technoloies for Agriculture	15
2.3 Data Collection and Integration	18
2.4 Data Analytics for Crop Yield Prediction	20
2.5 Accuracy and Reliability of Predictive Models	22
2.6 Impact of Data on Crop Management Practices.....	25
2.7 Sustainable Agriculture Planning.....	28
2.8 Challenges and Limitations of IoT in Agriculture	30
2.9 Future Trends and Innovations	32
2.10 Addressing Research Novelty: A Comparative Analysis of Existing Case Studies	35
Chapter 3 Methodology.....	38
3.1 Introduction to the Methodology	38
3.1.1 Purpose of the Methodology.....	38
3.1.2 Overview of the Research Design	38
3.1.3 Justification of the Research Design.....	39
3.2 Research Philosophy	41
3.2.1 Epistemology and Ontology	41
3.2.2 Justification of Research Philosophy	41

3.3 Research Approach.....	43
3.3.1 Inductive Vs. Deductive Approach.....	43
3.3.2 Justification of the Approach.....	44
3.4 Research Strategy.....	46
3.4.1 Experimental Design.....	46
3.4.2 Case Study.....	46
3.4.3 Longitudinal Vs. Cross-sectional.....	47
3.5 Data Collection Methods.....	49
3.5.1 Primary Data Collection.....	49
3.5.2 Secondary Data Collection.....	50
3.5.3 Instrument Design.....	51
3.5.4 Sampling Methods.....	51
3.5.5 Pilot Study.....	52
3.6 Data Analysis Techniques.....	53
3.6.1 Quantitative Data Analysis.....	53
3.6.2 Qualitative Data Analysis.....	76
3.7 Research Tools and Techniques.....	77
3.7.1 IoT Devices.....	77
3.7.2 Mobile Application connected to Firebase.....	80
3.7.3 Machine Learning Algorithms.....	80
3.7.4 Traing and Validation.....	81
3.8 Reliability and Validity.....	82
3.8.1 Reliability.....	82
3.8.2 Validity.....	82
3.8.3 Strategies for Addressing Data Transmission and Environmental Challenges.....	83
3.9 Limitations of the Methodology.....	84
3.9.1 Challenges Encountered.....	84
3.9.2 Limitations of the Study Design.....	84
3.9.3 Steps to Mitigate Limitations.....	84
Chapter 4 Results.....	86
4.1 Results of Primary and Secondary Data Collections for Agro-Sense.....	86
4.1.1 Iot Sensors used for Data Collection in Agro-Sense Native Android APP.....	86
4.1.2 Primary Data Collected from IoT Sensors was fruitful.....	87
4.1.3 Use of Secondary Data from Kaggle online dataset store for Enhanced Analysis.....	88
4.1.4 Comparison of Datasets (Primary and Secondary).....	89
4.2 Agro-Sense Equipment and its profound infrastructure.....	95
4.3 Linking Firebase Real-Time Databaseto Agro-SenseAndroid Native APP.....	101

4.3.1 Linking Firebase Real-Time Database to Agro-Sense Android Native APP.....	101
4.3.2 Real-time Updates advantage	105
4.3.3 Android APP Integration (Agro-Sense).....	106
4.4 Crop Health Prediction and Forecast Model Outcomes	114
4.5 Challenges and Limitations in Real-Time Crop Yield Prediction faced with Agro-Sense APP	117
4.5.1 Sensor failures and Data inconsistencies that happend.....	117
4.5.2 Latency and Network Issues in the system aroused.....	118
4.5.3 Integration Challenges.....	118
4.5.4 Challenges in Creating and Configuring Arduino Equipment for Crop Yield Prediction	119
Chapter 5 Discussions	121
5.1 Interpretation of Results.....	121
5.1.1 Aligning with Literature on IoT and Machine Learning for Agriculture.....	121
5.1.2 Advancing the Use of Real-Time Data.....	122
5.1.3 Addressing Gaps in Model Accuracy and Scalability	123
5.1.4 Filling the Gap in Sustainable Agricultural Planning	124
5.1.5 Enhancing Machine Learning Techniques for Agricultural Prediction	125
5.2 Connecting Research Questions with Outcomes.....	126
5.2.1 Research Question 1.....	126
5.2.2 Research Question 2.....	128
5.3 Technical and Model Restrictions.....	132
5.4 Developing Future Directions	134
5.4.1 Extension to Different Areas and Regions	134
5.4.2 Advances required in the IoT turf and models shaped with ML	135
Chapter 6 Conclusion	137
References	140
Appendix	145
Appendix A: Detailed Sensor Specifications.....	145
Appendix B: Qualitative Analysis Proof	147

List of Figures

Figure 2.1 IoT Sensor Technologies for Agriculture.....	15
Figure 2.2 Impact of IoT Data on Crop Management Practices	26
Figure 2.3 Sustainable Agricultural Planning.....	28
Figure 2.9 AI powered Precision Agriculture.....	34
Figure 2.10 Conceptual Diagram of the whole process.....	37
Figure 3.4 N, P, K values compared with the crops	47
Figure 3.5 Chosen field for testing.....	52
Figure 3.6.1 Time frames of Collected data	53
Figure 3.6.2 Phosphorus (P) Requirements for Different Crops	55
Figure 3.6.3 Scatterplot for the Crops	55
Figure 3.6.4 Variance of the four Crops.....	56
Figure 3.6.5 The workflow for Health Prediction.....	56
Figure 3.6.6 The workflow for Forecasting suitable Crop	57
Figure 3.6.7 Outlier code extract	58
Figure 3.6.8 Normalization code extract	59
Figure 3.6.9 Missing Value adjustment code	59
Figure 3.6.10 Heatmap of factor relationships	60
Figure 3.6.11 Cross-Validation	61
Figure 3.6.12 Performance Metric.....	61
Figure 3.6.13 Hyperparameter Tuning.....	62
Figure 3.6.14 Nitrogen factor bar chart for crops	63
Figure 3.6.15 Phosphorus factor bar chart for crops	63
Figure 3.6.16 Potassium factor bar chart for crops.....	64
Figure 3.6.17 Mango Sinhala extract with English translation.....	65
Figure 3.6.18 Mango temperature, rainfall and pH	65
Figure 3.6.19 Mango NPK values	66
Figure 3.6.20 Maize temperature and pH.....	67
Figure 3.6.21 Maize NPK values	67
Figure 3.6.22 Watermelon Sinhala extract with English translation.....	68
Figure 3.6.23 Watermelon temperature, rainfall and pH.....	69
Figure 3.6.24 Watermelon NPK values.....	69
Figure 3.6.25 Rice soil temperature, rainfall and pH.....	70
Figure 3.6.26 Former Model Performance Comparison.....	71
Figure 3.6.27 Latest Model Performance Comparison.....	72
Figure 3.6.28 Predicted Vs. Actual Health Status	72
Figure 3.6.29 Time Series chart 1	73
Figure 3.6.30 Time Series chart 2	73
Figure 3.6.31 Crop Confusion matrices.....	74
Figure 3.6.32 Crop Suitability Forecasting.....	75
Figure 3.7.1 NPK sensor got from internet store	77
Figure 3.7.2 pH sensor got from internet store	78
Figure 3.7.3 Humidity sensor got from market	78
Figure 3.7.4 Soil Moisture sensor bought from store	79
Figure 4.1.1 Primary Data from the scientific equipment	88
Figure 4.1.2 Secondary Dataset from Kaggle.....	89

Figure 4.1.3 Kaggles Dataset with 21 crop data	91
Figure 4.1.4 Modified Primary Dataset	92
Figure 4.1.5 Summary statistics for the distinct elements	93
Figure 4.1.6 Key Correlations Affecting Crop Yield Predictions.....	93
Figure 4.1.7 The figure shows the 4 crops nutrients.....	94
Figure 4.1.8 Ratio of nutrients valid for the fruits	94
Figure 4.2.1 The IoT Equipment prototype no.1	95
Figure 4.2.2 OLED screen.....	98
Figure 4.2.3 Agro-Sense Block Diagram	99
Figure 4.3.1 How the database looks.....	101
Figure 4.3.2 QR code checker	103
Figure 4.3.3 History Filter in Agro-Sense	103
Figure 4.3.4 Dataflow from sensors to APP	104
Figure 4.3.5 Flash Screen and Login Screen.....	106
Figure 4.3.6 SignUp and Reset.....	106
Figure 4.3.7 The live veiw of the crop stats.....	107
Figure 4.3.8 QR code register and Device GUI.....	108
Figure 4.3.9 Unhealthy Crops Look and Statistics	109
Figure 4.3.10 Healthy Crops Look and Statistics	109
Figure 4.3.11 History statistics visualization.....	110
Figure 4.3.12 Mango suitability forecast.....	111
Figure 4.3.13 Rice suitability forecast.....	112
Figure 4.3.14 Watermelon suitability forecast.....	112
Figure 4.3.15 Maize suitability forecast	113
Figure 4.4.1 Error Distribution in Predictions	114
Figure 4.4.2 Model Comparison Classification Heatmaps	114
Figure 4.4.3 Model Performance Metrics Bar Charts.....	115
Figure 4.4.4 Performance Matrix for crops	115
Figure 4.4.5 Older Time series comparison charts	115
Figure 4.4.6 New Time series comparison charts.....	116
Figure B1 Survey conducted for validation.....	147
Figure B2 Participants Chart	147
Figure B3 Sensor feedback.....	147
Figure B4 App suitability for farmers	148
Figure B5 App preference for farmers	148
Figure B6 ML model accuracy.....	148
Figure B7 Factor sufficiency	148
Figure B8 App for end users	149
Figure B9 App for making decisions.....	149
Figure B10 Feedback for approach	149
Figure B11 Improvements needed.....	149
Figure B12 Suggestions.....	150
Figure B13 Final response.....	150

List of Tables

Table 4.1 Datasets' Comparison.....	90
Table 5.2 Research Questions and Key Outcomes.....	130
Table 5.3 Model Restrictions	132
Table A.1 Distinct sensors used to get primary data records.....	145