

Decision Supporting System for Enhancing Upcountry Vegetable Cultivation in Sri Lanka

U. R. Samaratunge

Index Number : 129168 K

Dissertation submitted to the Faculty of Information Technology,
University of Moratuwa, Sri Lanka for the partial fulfillment of the requirements of
the M. Sc. in Information Technology.

March 2015

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

U.R. Samaratunge
Name of Student

Signature of Student

Date:

Supervised by:

Mr. S.C. Premaratne
Name of Supervisor

Signature of Supervisor

Date:

Acknowledgements

From the initiation there were many people behind me to support and encourage the research, development and the evaluation of this project. This research on agro based data mining and content search and the implemented tool for upcountry vegetables was not an isolated achievement of just me but an outcome of many people who helped me in many other ways and means.

First and foremost the Dean of the faculty of Information Technology, Department Heads, the kind and friendly Lecturers and Instructors were a big help for me to successfully complete this project. I must also thank my Supervisor for being very helpful and inspective throughout the project and helping me every time I was in trouble. Also there were many friends and batch mates and friends who helped me to complete this project both from their knowledge and encouragement.

Last but not least I also thank the editors of the various contents I used to gain knowledge to complete this project. Without their knowledge and generous thoughts I would have found it much difficult to find knowledge this easy and complete this task this much comprehensively. So, all in all I would like to thank all the people who helped this attempt become fruitful in many ways and means.

Abstract

Agriculture is considered to be the most widely spread and most commonly practiced form of employment in Sri Lanka. The Sri Lankan society has been dependent on agriculture since the beginning of its early civilization. Agriculture is the most important sector of the economy which fulfills the basic food needs of the citizens. This project considers upcountry vegetables; particularly selected four crops of which the information related to these are so important for development of this sector. The areas where upcountry vegetables have been cultivated fall into different agro-ecological zones varying from rain fall, elevation, land type, height of the top soil, slope of the land and soil erosion etc. The potential of information technology (IT) can be assessed as a tool for direct contribution to agricultural productivity and as an indirect tool for empowering farmers to take informed and quality decisions which will have positive impact on the agricultural activities. In this project, it is attempted to build an indirect tool for all interested parties including farmers aiming to develop an information system to that facilitates accurate and reliable forecasting through data mining to enhance upcountry vegetable cultivation in Sri Lanka. This information system going to use two sets of data; particularly one from Data Bank of Hector Kobbekaduwa Agrarian Research and Training Institute and the other from a sample survey conducted by this tools web interface developed using a content management system and using the sample data that recorded harvest will be forecast through data mining and creating appropriate modules and finally these modules will be incorporated in to the information system. The final phase of the project is the evaluation of the web-based information system with several different data mining algorithms. The Project has successfully implemented dynamic filtering, classification and prediction of data using the latest technologies and tools available in the field of IT to uplift the agricultural field in Sri Lanka and the evaluation of its performance looks promising for practical implementation.

Contents

	Page
Chapter 1 - Introduction	
1.1 Introduction	01
1.2 Background and Motivation	02
1.3 Aim and objectives	03
1.4 Purpose of the document	04
1.5 Structure of the thesis	04
Chapter 2 - Review of Others Work	
2.1 Introduction	05
2.2 A Prediction Model Framework for Crop Yield Prediction	05
2.3 Early Prediction of Crop Yield	06
2.4 Operational Prediction of crop yields using MODIS	08
2.5 Summary	09
Chapter 3 - Methodology	
3.1 Introduction	10
3.2 Outline of the Methodology	10
3.3 Methodology for Implementing the System	11
3.4 Summary	12
Chapter 4 – Data Analysis and System Design	
4.1 Introduction	13
4.2 Analysis of the Optimal Technologies	13
4.3 Architecture of the Prediction Tool	14
4.3.1 Training Module	14
4.3.2 Prediction Module	16
4.3.3 User Interface through Content Management System	17
4.5 Summary	18

Chapter 5 – Requirement Analysis

5.1 Introduction	19
5.2 User Requirements	19
5.3 Functional Requirements	21
5.4 Non-functional Requirements	22
5.5 Summary	24

Chapter 6 – Technology Adopted for the System

6.1 Introduction	25
6.2 Reasons for selecting Proposed Technology	25
6.3 Background of the Technology Used	26
6.4 Data Mining Technologies	28
6.5 Data Representation Technologies	32
6.6 Summary	34

Chapter 7 - Implementation

7.1 Introduction	35
7.2 Data Preprocessing Implementation	35
7.3 Classification Implementation using Weka and Java	37
7.4 Joomla Based Web Interface	39
7.5 Design Challenges and Resolutions	40
7.6 Summary	40

Chapter 8 - Evaluation

8.1 Introduction	41
8.2 Achievements of the objectives	41
8.3 Classifier Algorithm Evaluations	42
8.4 Web-based Interface for the System	45
8.5 Assessment on Results	49
8.6 Summary	49

Chapter 9 – Conclusion & Further Work	
9.1 Introduction	50
9.2 Assessment of the system	50
9.3 Problems and Limitations in development of project	51
9.4 Future Developments	52
9.5 Summary	53
References	28

List of Figures

	Page
Figure 2.1 - Actual vs. Predicted Data	
(Current week data- next week yield prediction)	07
Figure 2.2 – The 2005 classification for Iowa using the 8-day MODIS	
250 m composite imagery	08
Figure 3.1 – User Interface Snapshot of Weka	11
Figure 4.1 – Component diagram of the System	14
Figure 4.2 – Layered Architecture of the Training Module	15
Figure 4.3 – Layered Architecture of the Prediction Module	16
Figure 4.4 – Layered Architecture of the Prediction Module	17
Figure 5.1 – Business Activity Diagram of the Home page	23
Figure 6.1 – User Interface Snapshot of Weka	27
Figure 6.2 – Admin Interface Snapshot of Joomla	33
Figure 7.1 – Discretization Process in Data Preprocessing	36
Figure 7.2 – Flowchart of Naïve Bayes decision tree algorithm	37
Figure 8.1 – Classifier Accuracy Levels	44
Figure 8.2 – Classifier Error Levels	45

List of Tables

	Page
Table 2.1 - Performance Evaluation of the Models	06
Table 5.1 – User requirements	21