

**Arunella -Data Mining Application for Export Crops**¹M.P.Madushika, ²W.H.I. Wijewantha, ³P.A.D.A.Senarathna, ⁴P.Nawarathnam and ⁵N. Vithana^{1,2,3,4,5} Sri Lanka Institute of Information Technology

Abstract:- *The novel famers faces many difficulties due to the lack of experience. They do not have proper experience in cultivation. Therefore, those different kinds of cultivations face small to large scale of problems because new farmers have insufficient knowledge about agriculture. There are different levels and different types of cultivations. Sri Lanka being an agriculture based country and 39% of the population are farmers. The knowledge about different strategies to be used are gain by the experience. Because of those problems, famers can be in an uncertain state about their cultivation. Insufficient knowledge about lands and crops can be a main reason. Since they have less idea about their lands, soil conditions, non-fertile, crops, natural disease, sales problems, etc. This article explains an android and a web based system the supports decisions making in agriculture field. So, with this system farmers can get a good profit as they expected. The system uses data mining and gives most accurate solution to agricultural problems. This is created for spices cultivation in Sri Lanka.*

Keywords: *Clustering data mining, online chatting, Condiment, Android Application, Web Service, Agriculture, Land Management, Soil condition, Seeds, Crops, Fertilizer, Diseases of plants, Profit from perch, Datamining algorithm, gps*

I. INTRODUCTION

Data mining is an important part of knowledge discovery process that can analyze an enormous set of data and produce useful information. Data mining involves collecting, processing, storing and analyzing data in order to discover (and extract) new information.

At the end of 2011 there were almost 6 billion cellular telephone subscriptions worldwide, and the number is expected to reach 8 million by 2017. Mobile communication technology has become the world's most common way of transmitting data, services, and voice, and no technology has ever spread faster. The dynamic growth of mobile communications technology is creating opportunities for economic growth, social empowerment, and grassroots innovation in developing countries. One of the areas with the greatest potential impact is in the contribution that mobile applications can make to agricultural development, by providing access to information, markets, and services to millions of rural and urban inhabitants.

Insufficient knowledge about lands, crops, soil conditions, non-fertile, crops, natural disease and sales problems can be a main reason for having less harvest. In Sri Lanka, there are different levels and types of cultivations. Therefore, those cultivations face small to large scale of problems. Because of those problems, famers can be in an uncertain state about their cultivation. For both agricultural supply and demand, mobile applications and web services that are related to agriculture can reduce waste, make delivery more efficient, and forge closer links between farmers and consumers.

Since every user doesn't have android systems in their smart mobile phones they can use the web site to have benefits of this system. Users can suggest their problems to the system. System will process given data and display suitable solutions to user through Data Mining technology. "Arunella" is developed to solve above problem. Through this research supposed to give most accurate solution to agricultural problems, using an android mobile application and a web site.

The rest of the paper is organized as follow. Section 2 has provided the background through the Literature Review. Section 3 describes the methodology of this system. Also, the discussion of the developed system has described in Section 4. Section 5 contains results of this research. Finally, the conclusion is discussed in section 6.

II. BACKGROUND

There are lot of researches going on related to agriculture. Agriculture effects on human life cycle directly therefore agriculture is most important for a human. Every farmer supposed to cultivate in a correct manner and earn good harvest from their cultivation. Due to lack of proper knowledge they fail to get profit which they expected. Before cultivating crops, there are few things to consider such as fertilizer, soil condition, diseases and profit changes and identify suitable

crops. And there must be a good knowledge about management, economic rate. If farmers have good knowledge about these things, they can get expected harvest easily.

There are a lot of researches about this concept. But referred researches considered only a few sides of above factors.

“Clustering Algorithms Applied in Education Data Mining” authored by Dutta A et.al [1]. Methodology is Clustering Classification. Advantages in this research they show how data mining techniques apply in education datamining. Helps to know how datamining technology use. Disadvantages in this research consider education fields datamining techniques only. It is hard to get idea how it involves with agriculture side.

“A Study on Paddy Crops Disease Prediction Using Data Mining Techniques” authored by Ramesh Kumara. S.A. [2]. That advantages are Easy access to specialized Agriculture services by rural, under served, semi urban and in remote areas, early scrutiny and quick medication, reduced visits to fields, reduced travel expenses and reduced trouble of disease. There are two main characteristics of plant disease detection using machine-learning methods that must be fast identify algorithms which will help plant in detecting disease. Work can be done for automatically estimating the severity of the detected disease. Farmers can find diseases and cure using this android application. Two different methods for identify crops diseases. That steps are step by step descriptive method and graphical representation method. Used the association rule mining algorithm for the categorized to data.

“Agriculture Wrapped with Social Networks, Data Mining and Mobile Computing to Boost up Crop Productivity” authored by Haritha Akakineni et.al [3]. This research had proposed a system which help to farmers get information using mobile application. In addition to that they are trying to inform this information over social network. At the end of cropping collect information such as farmer utilized the climatic information, type of soils, type of seeds, pesticides used, yield information using android application. In this paper, they were hope to send notification to farmers to inform information. In this paper, they were proposed a framework for a cost-effective agricultural information propagation system, to propagate expert agricultural knowledge to the farming. Community in order to improve crop productivity.

“Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications” authored by Dutha A.A.G.S et.al [4]. In this paper, they were considered the problem of price prediction of crops. Price Prediction, nowadays, has become very important agricultural problem which is to be solved only based on the available data. Data mining techniques can be used to solve this problem. This work is based on finding suitable data models that helps in achieving high accuracy and generality for price prediction. For solving this problem, different Data Mining techniques were evaluated on different data sets. In this paper, Description and overview of data mining techniques which are applied to agriculture and their applications to agriculture related areas is described. Price prediction is a very important problem for any farmer as he is the one who should know how much cost he would expect for his crops. In past years, price prediction was done by judging farmer’s experience on particular crop and field.

“Agriculture Crop Pattern Using Data Mining Techniques” authored by M. Kaur et.al [5]. This paper discusses how farmers can benefit by using modern data mining methodologies and thereby reduce costs, increase profits. Data mining can help agriculture firms in production practices such as: acquire new farmers, retain current farmers, performing sophisticated classification, correlation between crops scheme. In the agriculture sector, data mining can help government to increase yield advantage mainly to support decision making, reliable and timely information on crop area, crop production and land use is of great importance to planners and policy makers for efficient agricultural development and for taking decisions on procurement, storage, public distribution, export, import and many other related issues to compete in the vend of crop pattern. In this research use following algorithm for data mining. K-means Clustering Algorithm.

III. METHODOLOGY

The “Arunella” project was developed using Iterative Prototype Model which is a System Development Life Cycle (SDLC) concept.

Selecting this model is beneficial because the team hopes to maintain high quality and minimize problems, errors & difficulties. Therefore, many things should be learned to carry out the research and there may be some problems implementing phases. Iterative prototype model will be very helpful because if any error occurs during implementation the team have the opportunity to go back to the previous phases and resolve the errors at the state they occur.

A. Initial Requirement

The first and the most important phase in the SDM is the feasibility study phase. Discussions among group members have been conducted to understand whether the project was financially viable and technically feasible. The team has proved via surveys, interviews, discussions that the results of the research component of this project “Arunella” is applicable for farmers to get information. The research team went to meet farmers to gather requirements by doing a

questionnaire survey. As per the survey the research group mined the data by using data mining methodology and took the best crop details. Through the developed android application and web site and the help of the data mining results the farmers can view the fertilizer according to the crops that they grow, diseases that occurs to the crops while growing, profits that the farmers can take from the lands which the crops are planned and the soil condition of the land they the farmers wish to plan crops. Moreover, the android application applicable of viewing an estimate price for crops that the farmers grow according to the current price in the market.

B. Design

The design phase represents introduction of operation system and diagrammatic representation on hardware, software data stores, files, User Interfaces(UI). In this phase, logical diagrams are converted to physical diagrams and developing and implementing strategies are discussed.

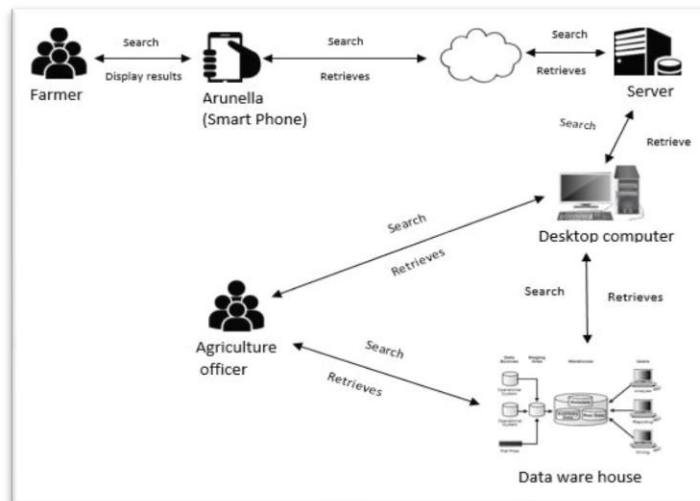


Fig. 1. System architecture diagram

C. Prototyping

Information gathering from quick design is modified to from a prototype. It represents a rough design of the required system.

D. Customer Evaluation

After delivering the system developer needs to provide training about how to use the system. This activity was supported by farmers and agriculture department evaluation, and if any changes are found then updating was done.

E. Review & Updating

There might be also change in input data or system environment. All this require modification of software. After customer is satisfied final design and further process is carried out.

F. Analysis

The group has analyzed the possibility of completing the project successfully considering economic, technological, scheduling and other factors. Rather than looking to the project and wishing for the best, a feasibility study lets the group to get to know the possible negative and positive outcomes of the research project before spending and investing time and money for the project "Arunella".

Requirement analysis phase answers the questions of what are the requirements to each individual component of the system. To implement the system "Arunella", to figure out the basic idea of this research problem, team went through similar systems which have been implemented.

G. Testing

This phase will be used to do the testing part of the application. To identify the bugs in the application it should be tested well. By doing this the team will ensure that the "Arunella" is capable of high performance and reliability. User requirements are mentioned in the document of software requirement specification. So, from the testing it will check whether the application meets that mentioned requirements as well.

After implementing the functions independently unit testing is done for each unit. It will ensure that each function is working correctly in isolation. Integration is done after completing all the functions. Research team used white box testing and black box testing methods. Finally, to ensure the whole application performs as one module, the system testing is carried out again. The team should improvise a method to test if “Arunella” application retrieves predictions as expected.

IV. RESULTS

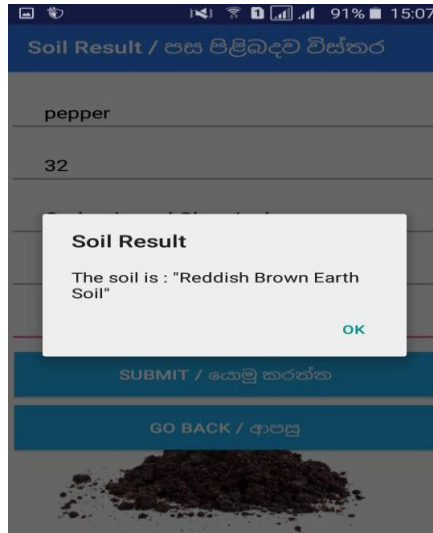


Fig.2. Soil Condition Results (android app)

They have to fill this soil condition form, when farmers want information about crops. Then they can get result according to their information.

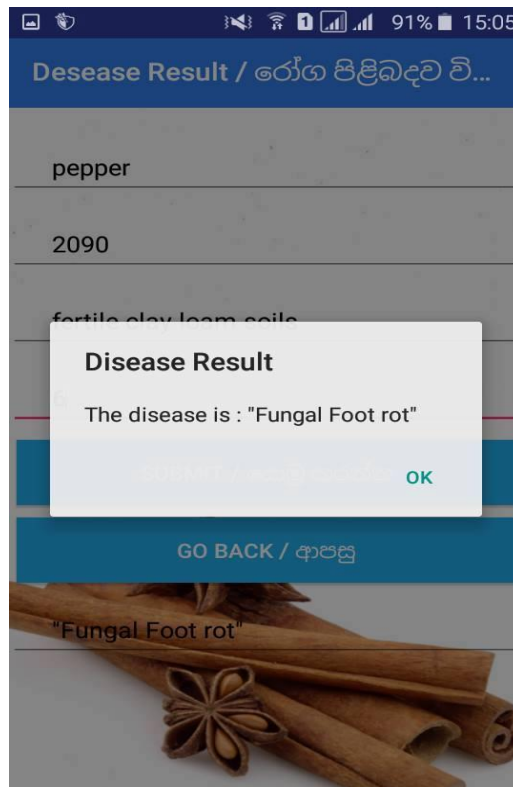


Fig. 3. Diseases Results (android app)

They have to fill this diseases form, when farmers want information about crops. Then they can get result according to their information.

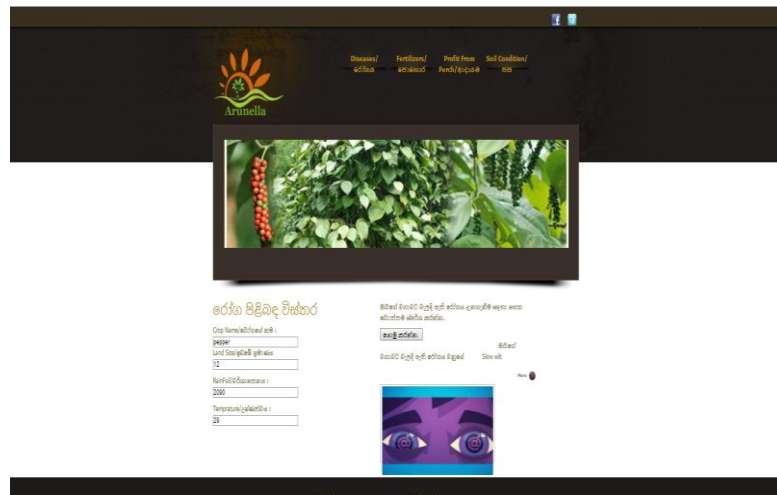


Fig. 4. Diseases Results (web)

They have to fill this diseases form, when farmers want information about crops. Then they can get result according to their information.

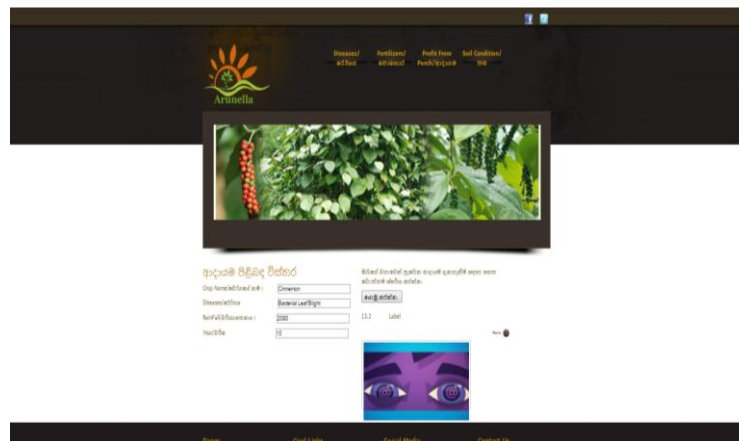


Fig. 5. Profit Results (web)

They have to fill this profit form, when farmers want information about crops. Then they can get result according to their information.



Fig. 6. Fertilizer Results (web)

They have to fill this fertilizer form, when farmers want information about crops. Then they can get result according to their information.

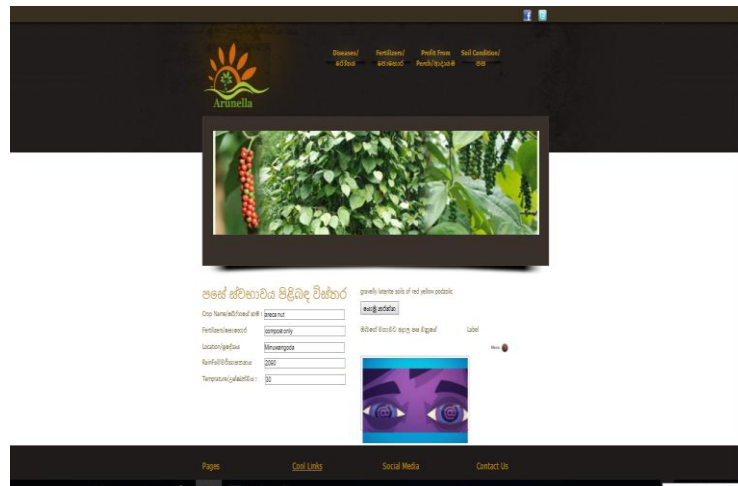


Fig. 7. Soil Condition Results (web)

They have to fill this soil condition form, when farmers want information about crops. Then they can get result according to their information.

V. DISCUSSION

The system we started to implement was planned to retrieve agriculture information from database and to give prediction results. Each component has been tested in isolation. After stabilizing all the modules each of them were integrated together and tested again until the complete system is error free.

Arunella contains functioning components such as Select diseases, fertilizer, soil condition, profit through android mobile application, online chatting, notification and the website. Already there are many data mining research, but very less of it in agriculture field. Because of this reason the team decided to complete our research in data mining for agriculture. After having meetings with lecturers who have done higher studies on data mining, the idea of creating both website and android application which goes along with our data mining technology was proposed. The task of creating four different types of websites was not easy as it was since the team could not find an information source regarding how to create websites and android application which will go along with data mining technology. Finally after all the discussions we decided to create a Clustering algorithm as the data mining model, for the application.

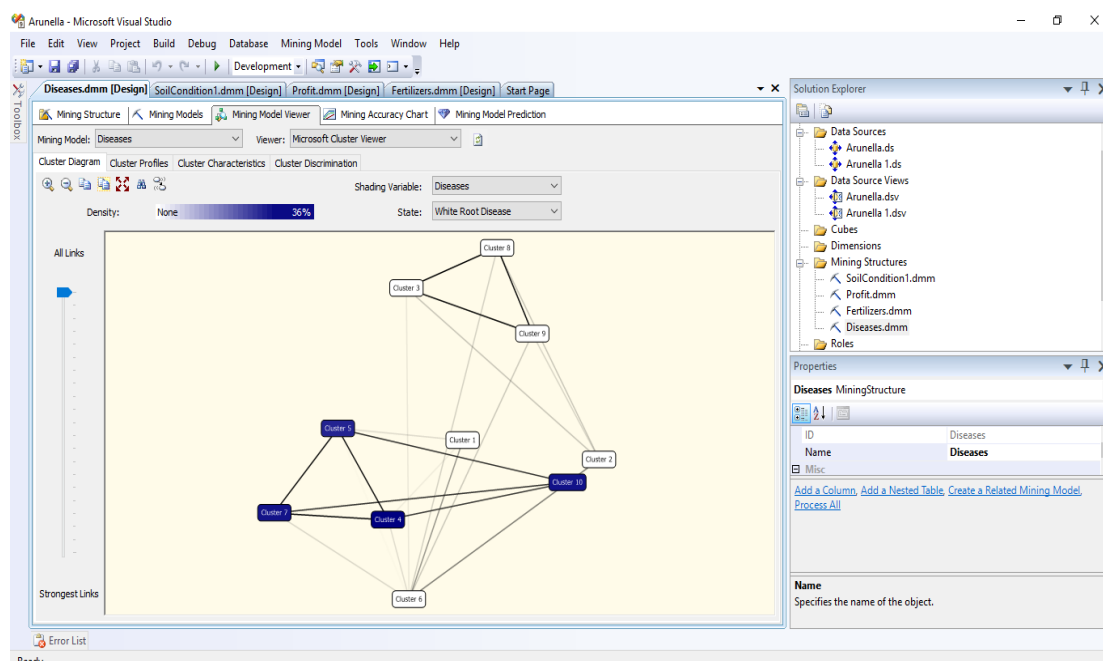


Fig. 8. Data mining model

After doing further research on this proposed idea the team managed ourselves to create the schema and retrieve the information from the Agriculture Department. After discussing with Agriculture Department we found hard copies of data sets with relevant to our data mining. Then our team entered that data into MS SQL databases manually. Fortunately, our team found least 5000 of records of dataset. Then we mine that data set with SQL Server Business Intelligence tools. After that found suitable model to do data mining our dataset. Clustering tree is the most suitable method to mine this data set.

The team was to develop a semantic web service to run the internal application. Though the team did many researches on how to implement a web service for this kind of situation, team was unable to find. The internal application was not supposed to run in the mobile device because of memory requirement is not enough for that kind of operation to take place. As for this reason, a normal web service was implemented. The mining part runs in this web service. In this web service SQL queries are implemented to retrieve information from the mining which is embedded to the web service. Then those information is passed as string values to the mobile application to display it to the user. Jason is the intermediate language to pass the data. Web site is created by asp.net framework. It also use same method

In order to have full functionality of the system the user must use the application in an android mobile device or device to browse in web (PC, Laptop, etc.) and having internet connection.

VI. CONCLUSIONS

Arunella tries to get an understanding on the nature of export crops of Gampaha District, SriLanka based on the predictions over them through datamining. Our study discusses how a crop's geographical location, fertilizers, weather conditions, diseases, price etc. can affect the production and economy of the crop. The paper states how a website is built for export department and a software application is built for farmers to use that in their mobile phones. So farmers can communicate with export department and among them whenever they have queries.

Our research suggests an enormous amount of time is taken to produce datasets on crops for datamining. Crops' information is collected and entered in EXCEL sheets by filling the required data. We include translation in English from Sinhala in order to fill data to specific columns of our tables.

This paper discusses on clustering data mining algorithms used in datamining models in predicting results using Microsoft SQL Server R2. We could able to predict results, utilized the processed results and use them in website. So that information we have in website can be provided to farmers' mobile phones through notifications or while they are querying.

In future work, people can give a different kind of treatments to their collected datasets with their designed or any advanced data mining algorithms using various tools. So that they can enhanced their percentage of reliability from prediction results.

Our research helps farmers to communicate their needs to their agricultural officials in terms of mobile devices with android OS assuming that they can afford to have one mobile phone and they are skilled to operate. Even though we have come up with solutions to deliver our application interfaces user friendly with their native language, additional work is to be done for farmers to communicate their queries in native language to officials.

Here our challenge is to integrate our functionalities we wrote in Android, PHP, ASP.NET and our databases MySQL and MS SQL Server R2 to be packed to find a path to deliver as one mobile application and a website application with its flexibility to communicate each other.

Since our research is limited to Gampaha in SriLanka, we emphasize that whole export crops of the country to be studied to predict their nature, to improve their production and economy, by giving chances to our farmers to address their problems with high authorities and to get solutions over them.

- *Future Research*

Arunella system can be developed in several ways since the concept is applicable to this kind of situation. Adding more functionality to the data mining models would be a better idea so that the information prediction can be done with any kind of website. As data mining is still in developing level and when the technology of data mining gets evolved this idea could be resourceful in the future ahead. The same technology and the concept can be adapted to a different type of a system in the future.

This system we implemented to Gampaha district only. By adding more other districts in next versions and develop value of this system and helpful for farmers to know how to cultivate and take care of plants properly. As well as for people who would like to have a good crop. As a research mainly focus in four crops (cinnamon, turmeric, paper and areca nut). As a future work insert more crops and develop the range of the system by adding more features such as vegetables and fruits.

REFERENCE

- [1] Dutt A., AG S., IS M., MA H., "Clustering Algorithms Applied in Educational Data Mining", International Journal of Information and Electronics Engineering, Vol. 5, No. 2,(2015):n. pag. Web 2 March 2015.<Available URL: <http://repository.um.edu.my/99092/1/Clustering%20Algorithms%20Applied%20in%20Educational%20Data%20Mining.pdf>>
- [2]Lausch A., SC A., TH L., "Data Mining and linked opened data- New perspectives for data analysis in environmental research", Article Ecological Modelling·(2014):n pag. Web 30 Sep 2015.<Available URL: <http://www.sciencedirect.com/science/article/pii/S0304380014004335>>
- [3]Sadilek A., KA H., DI L., LA B., PO E., TE J.,SI V., "Deploying nEmesis: Preventing Foodborne Illness by Data Mining Social Media", International Journal of Information and Electronics Engineering (2015) :n pag. Web 3 Dec 2015.<Available URL: <http://192.53.208/u/kautz/papers/sadilek-kautz.pdf>>
- [4]Guijalba F., MA T., SA M., "Dynamical Approach for Real-Time Monitoring of Agricultural Crops", IEEE Transaction of Geoscience Remote Sensing.(2015) :n pag. Web 05 Feb 2015.<Available URL: http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6987323&url=http%3A%2F%2Fieeexplore.ieee.org%2Fexpl%2Fabs_all.jsp%3Farnumber%3D6987323>
- [5]Mahato,A.,"Climate Change and Climate Change and Climate Change and Climate Change and its Impact on Agriculture n Agriculture n Agriculturen", International Journal of Scientific and Research Publications, Volume 4, Issue 4, April 2014[online] [Access Date 15-feb-2016] <Available : <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.639.7437>>
- [7]Bhargavi,P. and Jyothi,S.,"Applying Naive Bayes Data Mining Technique for Classification of Agricultural Land Soils", IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.8, August 2009[Online] [Access Date : 15-jan-2016] <Available : http://paper.ijcsns.org/07_book/200908/20090817.pdf>
- [8]Moore1,P. et.al, "Drugs Highly Associated with Infusion Reactions Reported using TwoDifferent Data-mining Methodologies" , J Blood Disorders Transf 5: 195, jan 2014[OnlineAccess data : 15-feb-2016]. <Available : <http://dx.doi.org/10.4172/2155-9864.1000195a> >
- [9]Bhargava, R, R Pramoda, and PremjyotiPatil. "AGRORATES: Location Based Agricultural Yields Rate Prediction By Using Android". International Journal of Recent Advances in Science Engineering 1.8 (2015): n. pag. Web. 15 Feb. 2016.<<http://israse.com/ijrase/assets/papers/V1-I8-IJRASE-CS-01.pdf>>

Author Biography

M.P.Madushika, W.H.I. Wijewantha, P.A.D.A.Senarathna, P.Nawarathnamare final year students at Sri Lanka Institute of Information Technology. They will be graduating with a BSc in Information Technology by the end of this year. This research paper is based on the year long, final year project they conducted as requisite for the degree course. M.P. Madushika email: pavanimadarasinghe@gmail.com

Nipunika Vithana is a senior lecturer at SLIIT Computing Pvt Ltd, a subsidiary of Sri Lanka Institute of Information Technology (SLIIT). She has been an academia in the field of Information Technology from 2004. Currently she is reading for her PhD from Management and Science University (MSU), Malaysia. She obtained her BSc in IT from Curtin University, Australia in 2004 and her Masters in Information Management from SLIIT in 2007. Her main research interests are Database technologies and empirical research in Software Engineering. She has presented papers in conferences and published articles in various journals.

Email: nipunika.v@sliit.lk