Agripreneurs:An approach to motivate individuals to become entrepreneurs through agriculture

Abstract— **Entrepreneurship is one of the most discussed domains in the 21st century mainly because it not only allows individuals to break boundaries to become exceedingly successful but also because it allows a country's economy to develop with very little effort from the governing side. However, due to many high-risk factors, in developing countries like Srilanka this domain have had reached its peak mainly because of the collapsing economy. On the other hand, agriculture is one of the sectors in Srilanka which provides a high number of employment opportunities for individuals.**

Creating a bridge between the agriculture and the entrepreneurship domain through modern technology is being achieved. This is achieved by the introduction of a motivational gamification model within the system which allows users to not only engaged with the application but also with other users using the application. An IoT module that will analyze soil data which includes humidity, temperature. An integrated auction platform that consists of a machine learning model to analyze user behavior. A robust image processing model to assist users in uncontrollable situations such as plant diseases.

The goal of this research is to allow employed or semi-employed individuals to make a passive source of income, mainly because due to the pandemic situation many individuals are facing low-income issues.

Keywords—IoT (Internet of Things), Prediction Model, Image processing, Gamification Model, Node MCU

I. INTRODUCTION

Entrepreneurship is a high risk, high reward domain. In developing countries like Srilanka, India and other low-income countries, the risk of achieving individual success by becoming an entrepreneur has now become a huge challenge. This is mainly because of the narrow to no opportunities provided by not only the society but also the government as well.

On the other hand, agriculture is a domain which uses many traditional methods to allow unemployed individuals to be engaged in a money-making activity at ease.

Bridging the agriculture and the entrepreneurship domains in developing countries will not only allow the country's economy to boost but also allow unemployed individuals in developing countries like Srilanka to be semi-employed and benefit from this system as well. Due, to the failing economy in developing countries like Srilaka, farmers and other industries are forced to cultivate vegetables, fruits and other crops as it will boost the export output of the country in addition to the boost in economy of the country and feeding the citizens of the country itself. According to other research conducted the GDP for agriculture has be gradually dropping throughout the past decade from 17.2% in 2005 and 11.1% during 2012, and has been dropping even further.[1]

A solution which assists not only medium to small scale farmers but also unemployed or employed individuals to become entrepreneurs through the agriculture domain has not been introduced to Srilanka. Also due to mass globalization and population bursts in Srilanka, the percentage of individuals with low income has gradually increased during the past years.

This research is mainly being focused on both the employed and unemployed individuals, to boost their monthly income by being involved in the agriculture domain through the system introduced. Utilizing spare land of individuals residing in urban areas of the country will benefit as well, since the system will provide the user with the best possible crop for the area they reside at.

The system will be built on a common gamification model which will make users motivated through every aspect of the system they are engaged in. The three main components of the system will be interconnected to each other in a chain format.

The main objective of this system is to motivate individuals to become self-made businesspersons or entrepreneurs by bridging the entrepreneurship and agriculture domains. To motivate users of the system, the three interconnected components will be communicating with each other to provide appropriate results to the end users.

An IoT data analysis model which will take soil data such as humidity, temperature, and other parameters to alert the user with the conditions of the soil in the area they reside at. An integrated auction platform through which the best crop in the market will be analyzed and alerted to the user with. A robust image processing model which will allow users to analyze crop conditions and be alerted with appropriate solutions with, and a success rate prediction figure model according to the location of the user.

II. LITERATURE REVIEW

In the past decade, many research and systems have had been introduced in the agriculture domain. However, most of the systems that are being introduced were primarily targeted to the common farmer. Furthermore, most of the systems that were introduced utilized common deep learning and machine learning algorithms to provide appropriate results to the end users, farmers.

In 2019, a research conducted revealed that if an individual is provided with key opportunities to excel in the entrepreneurship domain, they will be motivated and be involved in the activity in the long run. Furthermore, this research was mainly focused towards the agribusiness sector of the Srilanka[2]. Primarily, 6 major factors affect individuals from entering the entrepreneurship domain through agriculture which includes soil/farm conditions, unavailability of resources, lack of social support, policies, regulations, fluctuating prices in the market and climate changes.

In 2019, motivational models are not being implemented into agriculture focused systems in the past years, a research on mental models affecting individuals psychologically in the farming domain. The research was conducted taking both female and male farmers into account where, female farmers were more motivated on the quality of the products presented and the male farmers were mainly focused on other attributes which included profits, yield, resistance to diseases, long shelf life and tolerance to adverse weather conditions [3]. The gamification model is also a well implemented motivational model which is being applied in various popular software systems. Research has also been conducted on the effect of user experience and design on specific application where the gamification model is being applied [4], this involves in the implementation of a leaderboard system where rewards were given to individuals achieving a certain level score. Furthermore, when the gamification model involves more with interaction between users of the software system, individuals tend to be engaged on the system as well. It has also been revealed, that if a gamification model using agriculture is applied to farmers of rural areas, due to their lack of technological knowledge and education, farmers mostly depend on sharing information amongst them through community interactions [5].

IoT research have been conducted widely on the agriculture domain, most of the modules developed are being widely produced and implemented in many developing countries and is referred to as the smart farming industry. Furthermore, updating data using real time processes and cloud technologies have not yet efficiently advanced in this industry. Wireless based sensors are widely implemented in the agricultural domain to provide farmers easier means to do their activities [11]. Furthermore, data transferring via the modules to cloud or any other data saving platform was done for many years. Throughout the years these technologies were enhanced in the terms of performance by the transfer protocol by decoupling pure data. It was established on technology that termed as "intermediate buffering". This method helps in restoration during the delivery tunnel presenting interruption durations, if these are persistent and are closing in the last moment, the position of the TCP is declined to claim.[12]

Image processing models have being widely used throughout the past for various tasks in the day to day life, similarly in agriculture image processing has being used for analyzing crop conditions. This has mainly being applied to traditional crops with specific models built for each crop. Models developed mainly focus on a specific crop. The main parts of this model consists of a classification, feature extraction and training models where training images are provided at first to a relevant crop and the appropriate results are obtained. Classification accuracy of these models reach up to 94% throughout healthy to infected leaves [6]. It has also been concluded that the morphological shape of plant leaves provides better results when identifying infected leaves of plants when compared to other parameters such as color and texture [7].

Success rate prediction models are widely used in different industries throughout the world, some of the most popular are the sports industry, betting industry and stock market industry. However, models built to predict success rate of business ventures are also developed and implemented. Model's which use machine learning algorithms to predict future trends involve algorithms such as regression mainly which gave results based on unordered or data that isn't in an order. This research had two main areas which were the classification and regression models.[8] The research conducted by [9] analyses a crop prediction methodology where the crop yield is calculated according to the past rainfall of that specific area. ML algorithms including multiple linear regression (MLR) is used to achieve this. Data mining methodology called density-based clustering for the confirmation of predicted results. A system that was proposed, utilized a predicting network which was created using an artificial neural network approach. The most appropriate crop will be given as an output parameter according to the predicted soil type and rainfall datasets. Hence, other effective factors affecting for the yield can be determined, this was only applied for wheat yields.[10]
\

III. OBJECTIVES

The main objective of this research is to provide employed or unemployed individuals make another stream of passive income with their least effort and hence motivate them to become successful entrepreneurs not only supporting themselves but also the economy of the country. Motivating users to be continuously engaged within the system and utilize their spare time efficiently is also an addition to the main objective.

Other objectives include:

1. Utilizing spare land across the country making an income to the landowners.
2. Implementing an integrated bidding platform to allow buyers to identify crops being made in large and lesser amounts hence allowing them to place appropriate bidding values if satisfied.
3. IoT sensor analysis model to detect the humidity, water, temperature, and other external factors to determine the appropriate crop needed to be grown in that environment.
4. Making a game-based/gamification model to motivate the users analyzing psychological criteria with reward schemes.

5. Analyze IoT data and detect that crops that can be grown in specific areas.
6. Prediction model to show users the relative effort needed, and the income/benefit gained for specific crops that is recommended to be grown in that area.
7. Robust image processing model to detect diseases on crops and give the relative approximate solution/fertilizer.

IV. METHODOLOGY

The system proposed will be always focusing on providing a motivational drive to keep individuals engaged in the farming activities through the agriprenuers application.

The system is designed on a main motivational gamification model, where three other main components are connected to each other and to the gamification model.

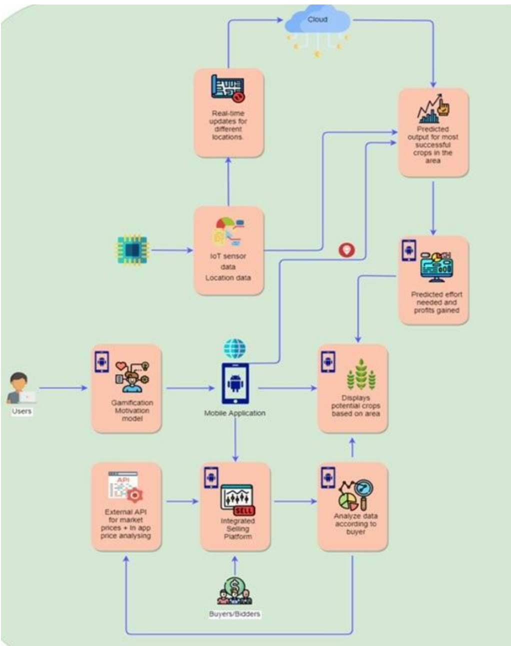


Figure 1: System Overview

As the figure shows, the system is interconnected to each other. In brief there are 3 main models standing on one main model which is the gamification model.

The other there main models include a real-time IoT soil condition analysis data model, an integrated buyer seller platform operating together with a machine learning model, and a robust image processing and success rate predictor model.

A. Gamification Model

This model provides users to be continuously engaged in the system. The gamification model is further divided into 2 main parts, which is the reward scheme implementation with a regional leaderboard and user interfaces designed with a two-toned theme along with psychology principles.

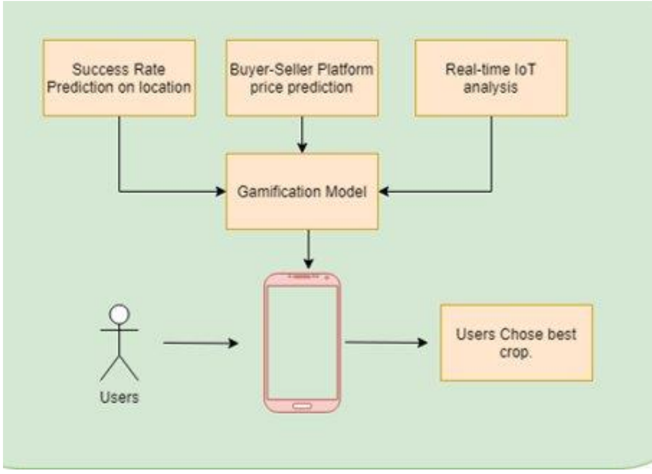


Figure 2: Gamification Model

The model takes input from all the other components and provides the end user with the most appropriate crop result for them to invest their time on.
The gamification model allows users to enter the approximate area of the spare land available to the users, this data is multiplied with market price of the specific crops which is being retrieved by the built-in auction platform.

When moving on to the psychology-based UI component implementation, a theory in the name of Hick's Law is considered. Here, if the application system consists of a lesser number of choices on the user interfaces and crop selections, the time taken for the user to decide decreases. Even if the crop selection is not the most ideal one, they will still be backed up with assistance from the application itself from the rest of the interconnected components.

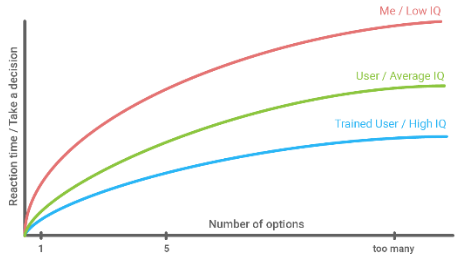


Figure 3: Hick's Law

In addition to this, a regional leaderboard is also being implemented in the system. This utilizes redis caching technology where users are rewarded according to their performance. The reward scheme mainly includes providing discounts to users upon selling crops at the auction platform where a percentage is taken by the developers of the application.

B. Soil Data Analysis IoT model.

Pricing of equipment that will be used will be kept at minimum costs at all times as it will allow more end users to be attracted to the final product. The IoT soil analysis model will be utilizing low cost soil moisture, temperature and humidity sensors. The data is then gathered and sent to a cloud platform as JSON data. Google real-time database is been used for this function. Data is sent to the database via esp8266 Wi-Fi module.

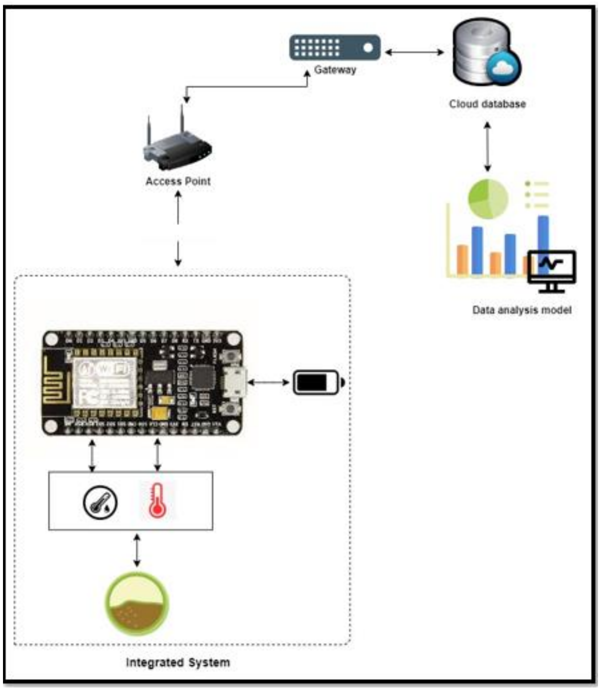


Figure 4: IoT Data Analysis Model

This system is mainly divided into 2 main components,

1. IoT hardware Model
2.Data Analysis Model

The IoT hardware model consists of a NodeMCU development kit (with esp8266) as a microcontroller, a DHT11 moisture sensor and temperature sensor, and a YL-69 soil moisture hydrometer detection sensor. This model will also allow users to use the kit as a portable module.

The sensors are interfaced with the esp8266 microcontroller. The sensor used to get the soil data consists of two probes which are placed in the soil. The probes conduct electricity through the soil. Moisturized soil has less resistance and so the electricity passes easily through it whereas if the soil is relatively dry the electrical resistance of the soil is high hence a lesser number of electrons pass through the soil.

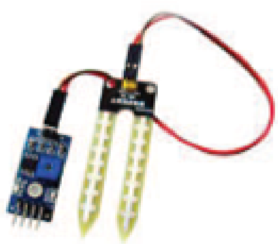


Figure 5: Soil Moisture Sensor

A DHT11 temperature and humidity sensor measures the amount of saturated water vapor in the air by measuring it together with the temperature of the surrounding air. These sensors are small is size as well. The data obtained are easily transferred to the database for further operations.

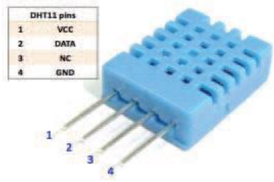


Figure 6: DH11 Temperature and Humidity Sensor

For the transmission of data, an ESP8266 microprocessor module has been used. Here, the data has been updating on a real-time basis on a cloud based database system. The module uses L106 32-bit RISC microprocessor core based on the Tensilica Xtensa Diamond Standard 106Micro running at 80Mhz, 32KiB instruction RAM and 80KiB user-data RAM. It supports upto 16MiB in external QSPI flash and comes up with IEEE 802.11 b/g/n Wi-Fi standards. The data is updated on MongoDB which is a NoSQL database in JSON format.

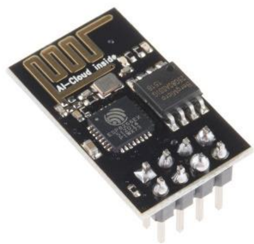


Figure 7:ESP8266 Microprocessor

The data analysis model will receive the data from the database (MongoDB). The data will be preprocessed to fixed threshold values and datasets. The threshold values and datasets will vary according to the crops that are selected and planted from the user. The threshold values and data sets are fixed after considering all the environmental and climatic conditions.

C. Robust Image Processing Model.

This model allows users to open the inbuilt camera module or upload individual pictures to the interface.
The image processing model consists of an initial data pre-processing model where the relevant data from the leaves are being obtained. The data preprocessing model consists of 3 main methods which are being used to extract the necessary details and features of the training set of data which are being added to the image processing model.
Hu Moments is used to detect the edges of the leaf image by analyzing the pixels intensities.

$$M = \sum_{x}\sum_{y} I(x, y)$$

Figure 8: Image moments equation

Implementation of Hu Moments to detect leaf edges is first done by converting the image to a gray scale image and analyzing the pixels, this is made easier by the functions given by OpenCV.

```
#feature descriptor1: Hu moments (yellowish/fire)
def fd_hu_moments(image):
    image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
    feature = cv2.HuMoments(cv2.moments(image)).flatten()
    return feature
```

Harlick Textures are used to determine the texture of the images, this is used as the second feature descriptor. Similarly, the image is converted to a grayscale image to deduce the texture of the image which is made easier with OpenCV functions. This is a first order statistic function. The implementation is as follows,

```
#feature descriptor2: Haralick texture(Quantify image according to the tex
def fd_haralick(image):
    gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
    haralick = mahotas.features.haralick(gray).mean(axis = 0)
    return haralick
```

As the third feature descriptor, to detect the yellow to green color variance, the color histogram functionality has been used which is being provided by OpenCV.

```
#featur descriptor3: Color Histogram
def fd_histogram(image, mask = None):
    image = cv2.cvtColor(image,cv2.COLOR_BGR2HSV)
    hist = cv2.calcHist([image],[0,1,2],None,[bins,bins,bins],[0,256,0,256,0,256])
    cv2.normalize(hist,hist)
    return hist.flatten()
```

The above feature descriptors are being used to analyze the images that are fed into the model as training data. The model consists of a train data set to train the model and a test data set to train the model. The total number of images obtained are 1600 which is divided into a ratio of 8:2 which is the training and testing data sets respectively.

OpenCV, Sklearn, joblib and h5py are some of the technologies used in the model to increase the efficiency. The test and training data of the healthy and diseased.

Figure 9: Diseased leaf

Figure 10:Healthy Leaf

Finally the model is being 10 fold validated across different algorithms such as random forest, support vector machine, Naïve Bayes, Decision trees, Logistic Regression and Linear Discriminant. The training model is tested against the algorithms and the algorithm which gives the highest accuracy percentage is been selected and deployed to the application to be used.

D. Auction Platform predicting trending crops.

A buyer seller auction platform will be introduced within the application to make users of the application to seamlessly connect with buyers of the harvest they have made after a few months. The buyer will be able to place a price on crops of their choice, the integrated machine learning model will analyze this data and provide the user with the best trending crops.

The prediction feature of future trending crops could possibly help individuals make better decisions about what vegetables to invest in before growing them initially. This prediction model will use previous data that is collected through the app itself about what buyers mostly seek and then make predictions about the future trending crops that will be most sought after by buyers.
As you can observe from the figure below which shows farm gate and retail prices of vegetable crops throughout the whole island data such as this can be used to train a machine learning model in order to predict the results required.
The machine learning algorithm used for this purpose was the Prophet algorithm (Time Series Model).
Papacharalampous and Tyralis consider the performance of random forests and Facebook's Prophet in forecasting daily streamflow up to seven days ahead in a river in the US. The conclusions that could be drawn from the results were such that random forest algorithm performs better in general terms, while Prophet algorithm outperforms the naïve method for forecast timelines that were longer than three days. The before mentioned facts and many more contributed in making the decision to train the machine learning model using the Prophet algorithm. The dataset considered is displayed below.

Item	Farmgate Price		Retail Price	
	2017	2018	2017	2018
Kurakkan (whole)	134.23	123.40	196.17	191.52
Gingelly (seed)	171.60	204.98	464.11	507.37
Maize (whole - local)	45.74	45.03	146.75	124.01
Cowpea (whole - white)	184.53	174.57	276.46	268.28
Cowpea (whole - red)	194.74	173.11	287.80	276.56
Soya bean (whole)	84.27	104.27	219.99	238.30
Ground nut (with shell)	154.26	166.56	-	-
Ground nut (without shell)	-		431.66	446.74
Black gram	138.98	148.61	348.32	304.13
Green gram (local)	188.80	168.59	225.85	224.50
Potato (local)	96.61	93.80	155.38	151.61
Manioc	49.71	41.95	84.37	69.41
Sweet potato	50.91	52.27	94.42	101.21
Red onion (local)	168.45	116.25	272.08	189.15
Big onion (local)	78.10	73.98	117.89	115.52
Dry chilli	-	209.06	250.10	288.27

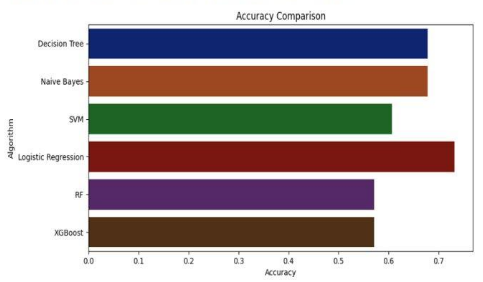
Source: Department of Census and Statistics

Figure 11: FARMGATE AND RETAIL PRICES OF OTHER FIELD CROPS - All Island (Rs/kg)

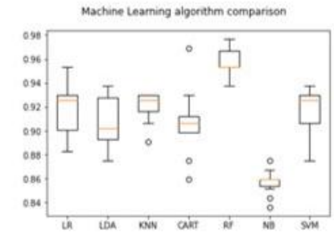
V. RESULTS

The gamification model was cross tested across various ages groups and genders to check their feedback, for this we used a google form to obtain the results as acceptance testing was hindered external causes. The survey allowed interfaces color schemes to be altered according to the end user's preference. Furthermore, the IoT module was successfully deployed and appropriate data was obtained and processed by the data analysis model. The Naïve Bayes algorithm provided the highest accuracy percentage in the logistic regression algorithm.

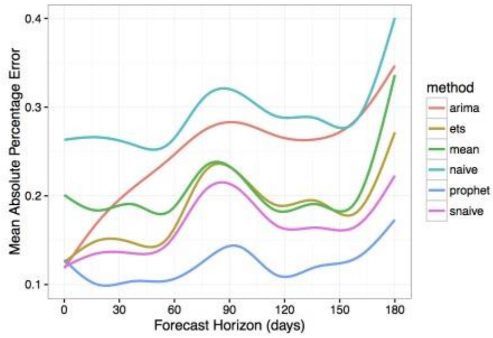
```
In [24]: from sklearn.naive_bayes import GaussianNB

         NaiveBayes = GaussianNB()

         NaiveBayes.fit(Xtrain,Ytrain)

         predicted_values = NaiveBayes.predict(Xtest)
         x = metrics.accuracy_score(Ytest, predicted_values)
         acc.append(x)
         model.append('Naive Bayes')
         print("Naive Bayes's Accuracy is: ", x)

         print(classification_report(Ytest,predicted_values))

         Naive Bayes's Accuracy is:  0.6785714285714286
                       precision    recall  f1-score   support

         Bitter Gourd       0.50      0.71      0.59         7
              Brinjal       1.00      0.75      0.86         8
             Capsicum       0.94      0.85      0.89        20
               Chilli       0.53      0.64      0.58        14
               Tomato       0.20      0.14      0.17         7

             accuracy                           0.68        56
            macro avg       0.63      0.62      0.62        56
         weighted avg       0.70      0.68      0.68        56

In [25]: # Cross validation score (NaiveBayes)
         score = cross_val_score(NaiveBayes,features,target,cv=5)
         score

Out[25]: array([0.60714286, 0.69642857, 0.89090909, 0.76363636, 0.61818182])

<AxesSubplot:title={'center':'Accuracy Comparison'}, xlabel='Accuracy', ylabel='Algorithm'>
```



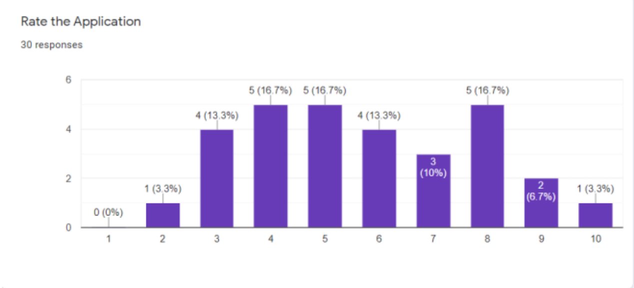
The robust image processing model was successfully trained and obtained to have an accuracy score of 98% when using the random forest algorithm.



The auction platform will also be gathering data to train the machine learning model in order to allow the users to select the most appropriate crop to be grown for the maximum profit. The results are as follows,



As the application was provided to random users in a specific area, their responses were as follows:



Many of the users were satisfied with the overall experience through the user interfaces and user experience of the applicatiom. Issues faced by other users who responded with

VI. CONCLUSION & FUTURE WORK

This application is ideal to be used by individuals residing in the urban areas as the relative effort they have to provide is lesser than that of a regular full-time farmer. However, if a disease or sudden climate changes occur most of the results might turn out unfavorable. This can be avoided with the help of the meteorology department of the country to obtain more reliable data on which crop would be appropriate to be grown in which region. Furthermore, the auction platform can be further extended by providing transportation needs for the crops bought from the buyers. The seeds and other fertilizers utilized by the end users can also be provided by the application itself in the future. Machine learning models can also be developed according to the success rate in each of the processes relative to the effort needed. This research in building up an application to motivate individuals to make a passive income can also be utilized by farmers as a reliable utility to understand climate changes and other alterations in the market and to adapt as needed.

VII. REFERENCES

[1] Ft.lk. (2019). Sri Lanka heading for an agricultural issue in 2017? [Online] http://www.ft.lk/columns/sri-lanka-heading-for-anagriculturalissue-in-2017, 2017.

[2] K. M. V. Sachitra1, "Entrepreneurial Opportunities and Role of Capability Approach in Agribusiness: Evidence from Sri Lanka", 1Department of Commerce, University of Sri Jayewardenepura Sri Lanka, 2019..

[3] J. Okello, Y. Zhou, I. Barker, and E. Schulte-Geldermann, "Motivations and Mental Models Associated with Smallholder Farmers' Adoption of Improved Agricultural Technology: Evidence from Use of Quality Seed Potato in Kenya," *Eur. J. Dev. Res.*, vol. 31, no. 2, pp. 271–292, 2019.

[4] I. Nuritha, V. P. Widartha, and S. Bukhori, "Designing gamification on Social Agriculture (SociAg) application to increase end-user engagement," *Proc. 2017 4th Int. Conf. Comput. Appl. Inf. Process. Technol. CAIPT 2017*.

[5] E. Beza *et al.*, "What are the prospects for citizen science in agriculture? Evidence from three continents on motivation and mobile telephone use of resource-poor farmers," *PLoS One*, vol. 12, no. 5, pp. 1–26, 2017.

[6] Wenjiang Huang, Qingsong Guan, Juhua Luo, Jingcheng Zhang, Jinling Zhao, Dong Liang, et al., "New Optimized Spectral Indices for Identifying and Monitoring Winter Wheat Diseases", *IEEE journal of selected topics in applied earth observation and remote sensing*, vol. 7, no. 6, June 2014.

[7] Monica Jhuria, Ashwani Kumar and Rushikesh Borse, "Image Processing For Smart Farming: Detection Of Disease And Fruit Grading", *Proceedings of the 2013 IEEE Second International Conference on Image Information Processing, 2013.*

[8] Bukun Afolabi, T. Cordelia Ifunaya, Funmilayo G. Ojo, Chinonye Moses, "A model for Business success prediction using Machine learning," Department of Computer and Information sciences, Covenant University, Ota, Nigeria, 2019.

[9] D Ramesh and B Vishnu Vardhan, "ANALYSIS OF CROP YIELD PREDICTION USING DATA MINING TECHNIQUES", International Journal of Research in Engineering and Technology eISSN: 2319-1163 | pISSN: 2321- 7308, 2015

[10] Miss.Snehal S.Dahikar and Dr.Sandeep V.Rode, "Agricultural Crop Yield Prediction Using Artificial Neural Network Approach", International journal of innovative research in electrical, electronics, instrumentation and control engineering, Vol. 2, Issue 1, January 2014, ISSN (Online) 2321 ,2004.

[11] N. Kaewmard and S. Saiyod, "Sensor data collection and irrigation control on vegetable crop using smart phone and wireless sensor networks for smart farm," in 2014 IEEE Conference on Wireless Sensors (ICWiSE), 2014.

[12] Luzuriaga, Cano, Calafate, Manzoni, Perez and Boronat, "Handling mobility in IoT applications using the MQTT protocol," 2015 Internet Technologies and Applications (ITA), Wrexham, UK, 2015.