ISSN 2278-3091
Volume 9, No.1.3, 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse6191.32020.pdf
https://doi.org/10.30534/ijatcse/2020/6191.32020

# Intelligent Farming Of The Future: An Integrated Farming Assistance And Crop Prediction Technique Using Isolation Forest Anomaly Detection

**Aromal Benny[1], Aswin A[2], Athena E[3], Basil Mathew[4], Dr. Varghese S Chooralil[5]**
[1]Rajagiri School of Engineering and Technology, India, aromalbenny17@gmail.com
[2]Rajagiri School of Engineering and Technology, India, aswina007007@gmail.com
[3]Rajagiri School of Engineering and Technology, India, athena.ashok@gmail.com
[4]Rajagiri School of Engineering and Technology, India, basil07mathew@gmail.com
[5]Rajagiri School of Engineering and Technology, India, varghesee.kutty@gmail.com

## ABSTRACT

Agriculture is one of the important sectors for every nation and it should be given prior concern, not only because it increases the GDP of a nation but also it feeds all of its citizens. We are now living in a society where people are more attracted to white collar jobs and they hesitate themselves to enter the agriculture sector. Why this happens is because there is a notion about the agriculture sector that it is less safe and is subjected to big losses. Maybe up to an extent this feels reasonable, that is there are a lot of parameters we should be concerned about when entering the agriculture field and some of them are really complex and very difficult to choose. Such instabilities make more people move back from agriculture. But what if there is a system which helps to choose the most suitable crop for someone's land and also guides throughout the agriculture process. In this work we are introducing you to such a system which is implemented as a mobile application. The system considers some parameters and based on these parameters a suitable crop is predicted through isolation forest algorithm as its back-end. Moreover, it also provides certain features required for assisting a farmer in his whole agriculture processes.

**Key words**: Isolation Forest, Anomaly Detection, Crop Prediction, Assistance to Farmers.

## 1. INTRODUCTION

Every sector is being digitized and the whole world encourages it, because it reduces manpower, increases precision, robustness etc. and as a whole the point is that it reduces the complexity and makes everything much more simpler and precise. Based on our study we came across the fact that agriculture has many more areas that can be explored via digitization[9], [11].

The problem started from the base itself, the backbone of agriculture - the farmers. As per the reports and studies the number of farmers and the farmlands are declining at a rapid rate. This happens because of the insecurity in the agriculture sector. The system which being introducing here as a mobile application[13], [15,16] can help the farmers to overcome some of the barriers in farming. For any person who is entering in the agriculture sector, the first question that could be going through his mind is what he should cultivate in his land. He has to consider many parameters before he chooses a crop like what if the crop he chooses don't grow in the current temperature or his land doesn't have the required pH value [12]. Also there are some parameters that we should consider while choosing a crop [10]. Otherwise, it may lead to bad yield, high maintenance cost like the cost for manufacturing and other fertilizers, loss of money and many more.

So, the important thing is that a farmer should choose a crop which is appropriate for his land so that he could pay less for manuring, fertilizers and also he can get good yield out of his land. The system we introduce here will be capable of solving this issue .i.e., the system as a mobile application in the hands of a farmer will be able to predict what crop he should cultivate in his land based on certain parameters which include temperature, altitude, pH value of the soil, soil type, growing season of the crop. The prediction is done on the basis of Isolation forest algorithm[1], [4], which is an anomaly detection technique used in order to classify the anomalies in a dataset provided.

While doing cultivation, we should do proper actions in order to make the land suitable for cultivation, and also there are certain measures to be taken through the crop life cycle. The application guides the farmer to do different measures at different times of the crop life cycle and it also helps new farmers to grow particular crops in his land.

The application is equipped with features as listed below:

- A map showing the location of other farmers who are using the application and information about them like the crop being cultivated in his land and contact information of the person is displayed.

- Weather report system which shows the farmer the current and upcoming few week's weather statistics[8], [14,17].
- Market rate window, which enables the farmers to know about the current rate of different crops in different districts.
- Agriculture news window, which helps the farmers to know about different news regarding the agriculture sector.
- A query system where the farmers can ask their questions among an open community and they can reply for those questions hence providing a platform to get proper human guidance and acquire more knowledge.

We believe that this application will help farmers facing many of their problems and it will support them as a helping hand.

### 1.1 Drawbacks of Current System

There are different types of mobile applications available right now to support the farmers regarding different aspects. However, as of now there is no crop prediction system currently in the form of an application that can be accessed by farmers. The absence of such an important system in the current applications is a matter of concern because it is the most relevant problem that many farmer faces when entering the farming sector and it requires a solution [11,18]. And also the applications currently available lack a map system. That is, there is no application that lets farmers know who all the other farmers are near their location.

Implementing such a system will assist the farmers to build a co-operating environment where they would be able to communicate with the others doing the same kind of cultivation in a motive so as to act as a helping hand for each other.

## 2. METHODOLOGY

The project here mainly deals with the providing the users with an updated information as well as the crop prediction mechanism where the users are shown the best possible crop that can be cultivated on the basis of input given by the user. The methodology used here is the Isolation Forest algorithm which comes under the Anomaly detection [6] mechanism.

### 2.1 Anomaly Detection

Anomaly Detection or the unsupervised anomaly detection is a mechanism for identifying the unexpected or rare events in a dataset, more specifically known as the outliers in a specific dataset which can be recognized as an anomaly with respect to other values present in the dataset.

Now as in Figure 1 [5], which shows the points in the feature space. The points represented by blue circles are normal points and the ones classified by red triangle are the anomaly points. Hence the main purpose on the anomaly detection algorithms is to classify the normal points and anomaly points differently in a dataset [5], [6]. One such anomaly detection mechanism is the Isolation Forest algorithm.

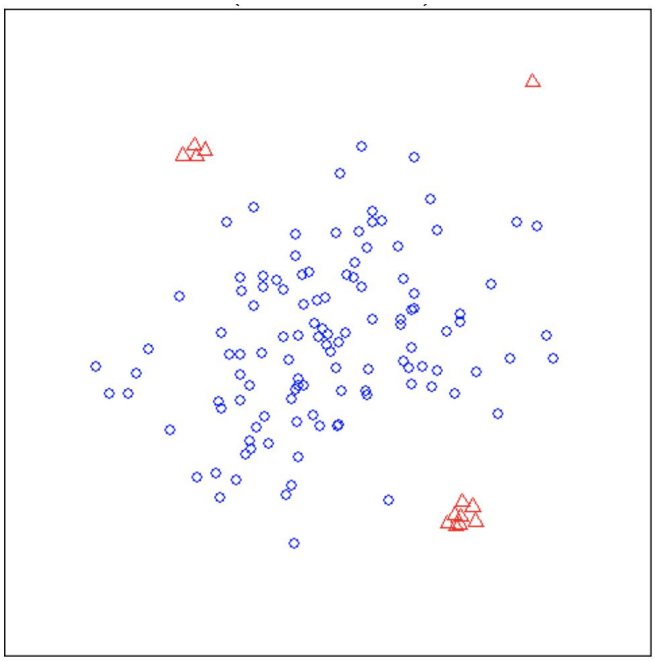

**Figure 1:** Visualizing Outliers

### 2.2 Isolation Forest Algorithm

The isolation forest algorithm [2,19] is an unsupervised machine learning algorithm, which works with anomaly detection as its basic functionality. In a given dataset, its primary aim is to separate the outliers from the inliers. It does such a function by randomly partitioning the domain space, similar to the work done by a decision tree algorithm [2,20].

The randomly partitioning of the domain space on a particular feature from a given sets of features is mainly done in order to obtain a split value between the maximum and minimum value of that feature. This in return will be producing shorter path in the trees in case of anomalous points as compared to that of the normal points.

Consider a feature space as shown in Figure 1 [5]. Here, the normal points will always be clustered together when compared to the anomaly points. Hence when path in decision tree is to be made, outliers will result in giving shorter paths from the root nodes while the inliers will be generating a bigger path as more the number of comparisons required in order to isolate an inliers is large compared to the

number of comparisons required to isolate an outlier. As a result, anomaly points can be detected efficiently using less amount of memory compared to other algorithm performing the same function [2]. This is depicted in Figure 2 [5], where less number of comparisons was done in order to isolate the anomalous points compared to larger number of comparisons done to isolate a normal point as in Figure 3 [5].

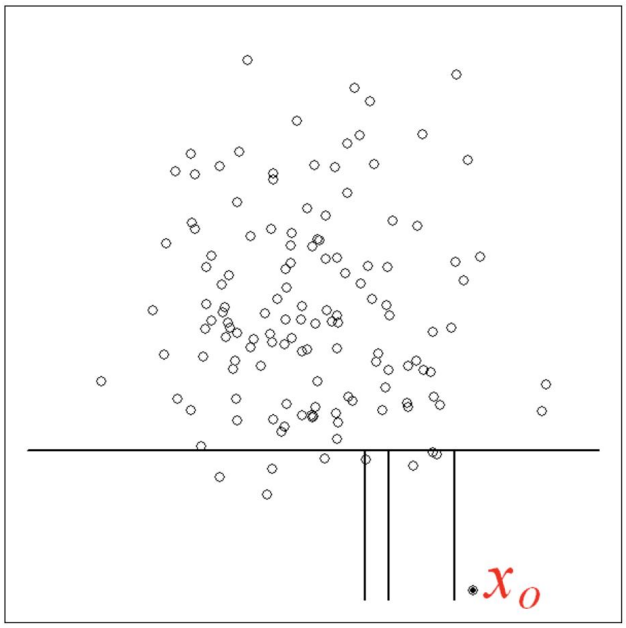


**Figure 2:** Outlier Point Detection

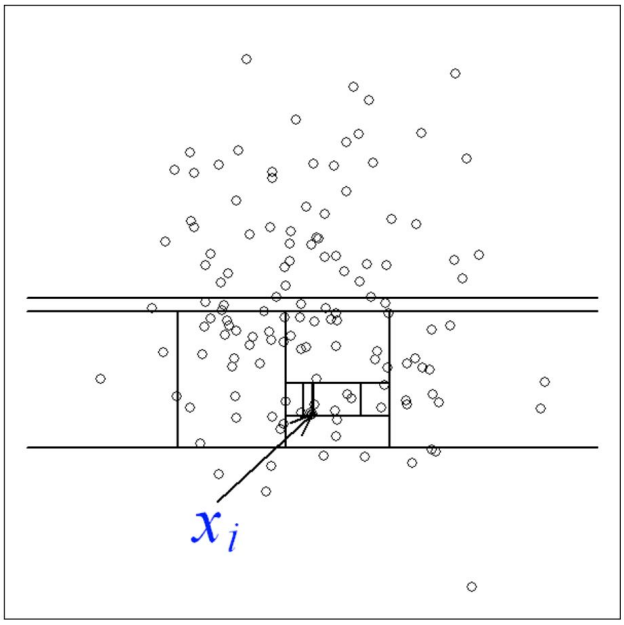


**Figure 3:** Inlier Point Detection

In order to implement it in the project, the isolation forest module [4] is imported from the sklearn package [2] as:
from sklearn.ensemble import IsolationForest

Once the package has been imported, the required dataset is read in the csv format [3] using pandas [7]. Then the isolation function is called with the required parameters and fit() function is called on each dataset one by one in order to generate the model. Once the models are generated and stored, the user's inputs are passed and run on each model using predict() function with the required parameters to generate the output.

## 3. DESIGN

The user interface design of the application [13] is shown in the following subsections.

### 3.1 Home

The home page as shown in Figure 4 provides information on the number of crops cultivated by the user. Initially it is blank, which gets filled up by adding crops using either Crop Prediction Method or simply selecting the desired crop for cultivation. After selection of the crops, the home page will be displaying all the crops being currently cultivated. The user can open any one of them and will be able to see the life cycle of the crop in a weekly manner.

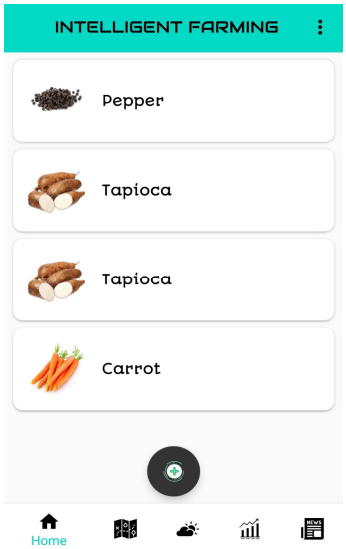


**Figure 4:** Mock up of the Home Tab

### 3.2 Crop Prediction

Initially a data set is made for each of the crop on the basis of different parametric values [10]. This dataset is run through isolation forest algorithm in order to create the model. So when a user provides the relevant data about his/her land, according to the collected data the suitable crop for the user will be predicted. The predicted result may contain more than one result on the basis of land details provided by the user. The user can choose the crop from the result displayed.

After choosing the desired crop, the farmer can navigate through the life cycle of the crop. The life cycle is a week wise representation of procedures which should be carried out by the farmer during the crop cultivation.

As shown in Figure 5, the user is provided with an interface to enter their parameters and finally the predictions are made on the basis of these parameters.

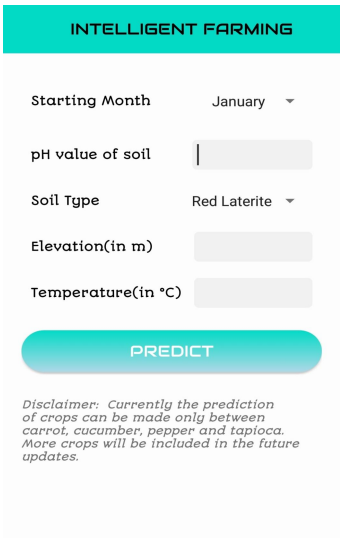


**Figure 5:** Mock up of the Crop Prediction Tab

### 3.3 Map

Map window as shown in Figure 6, is used as a quick access to all the contact details of the nearby farmers and agriculture offices.

It displays the geographic location of the farmer along with different plot marks which indicate either a nearby farmer or a nearby agriculture office. While clicking on the marker, further details about the individual or office will be displayed. If the farmer is selected, his contact info along with the crops he is cultivating will be displayed.

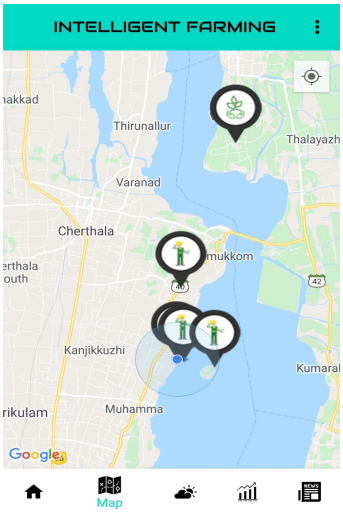


**Figure 6:** Mock up of the Map Tab

### 3.4 Weather

In weather activity as shown in Figure 7, an active weather site was taken where the user can provide the required month and year and the corresponding weather statistics is displayed to the user. The application displays the details about the weather condition of all the days of that particular month. The detailed weather description of each day can also be obtained by clicking on that particular day tag.

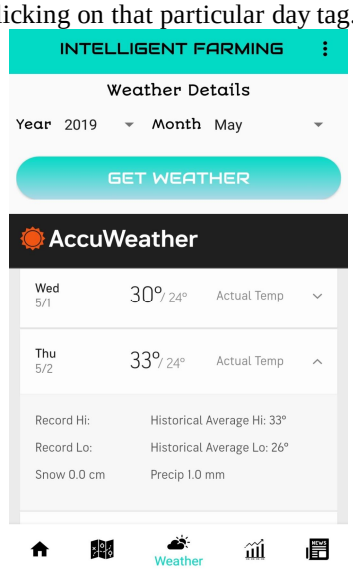


**Figure 7:** Mock up of the Weather Tab

### 3.5 Market

The market rate helps the user to know the price of a particular commodity at a particular period. The user have the options to select the commodity, the year and the month. The application will display the price details of that commodity based on the month and year chosen. According to the month and year the price will be shown as a bar graph, having price on the y-axis and districts on the x-axis. The user can understand the price of a commodity at different districts and different period(i.e. month or year). The main advantage of this market rate is that the user can compare production income with their own previous incomes and also with others which help them to improve.

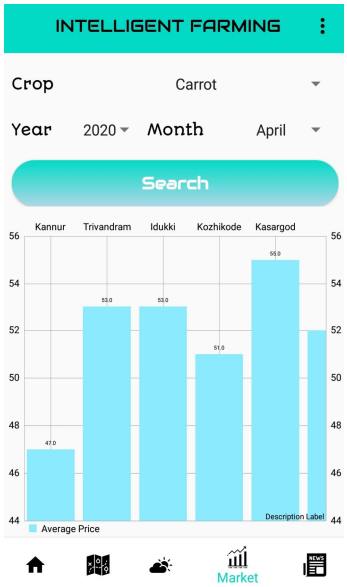


**Figure 8:** Mock up of the Market Tab

### 3.6 News

When we go to the view page of the news we could see the updated news related to agriculture and crop production. The news will be sorted from latest to the oldest and recent news will be displayed on the top of the screen. On the view page the user can see only the headlines and photos of the news. For the detailed news, the user have to click on the news tag and that will direct to the page which have the full details or descriptions regarding that particular news.

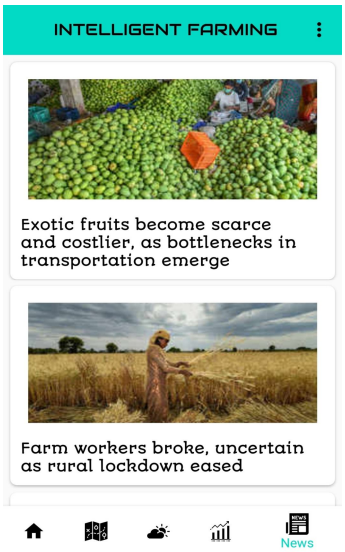


**Figure 9:** Mock up of the News Tab

The updation of the news on the back end is done by the admin. The admin updates news which are taken from various news sites. As an admin, we will just produce a newly released news as subject in the user menu and we will describe about the same including the relevant photos. So when a user clicks on these news feed it will show the already fed data to the users for conveying the information.

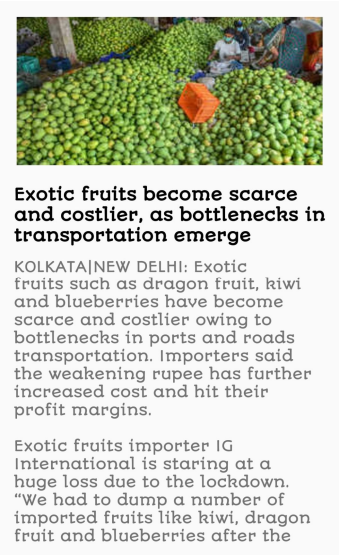


**Figure 10:** Mock up of the News Description Tab

### 3.7 Query Box

The query box system is initially displayed as a set of queries being already asked and posted publicly by other users. They are displayed in the newest to oldest format. The queries are displayed along with specific dates on which they were published. Any user using then application has the privilege to answer the query by clicking on any specific query in the query box. All answers for a particular query will be displayed to the user along with the people who answered it.

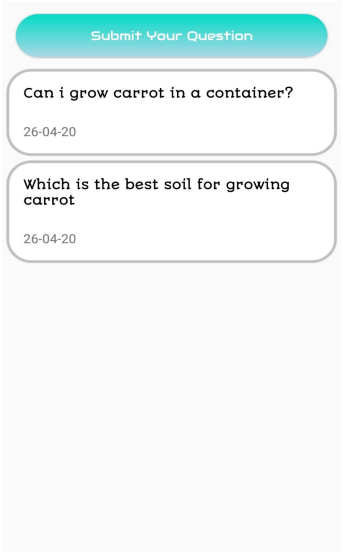


**Figure 11:** Mock up of the All Queries Tab

Now in the main query page itself, an option is provided to the user to submit a question. Once the user clicks on that button, an alert box will be displayed where our question can be submitted as shown below.

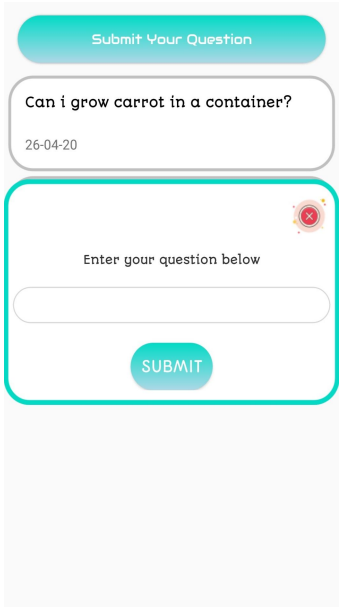


**Figure 12:** Mock up of the Ask Query Tab

## 4. CONCLUSION AND FUTURE SCOPE

Currently we have limitations to access all the data required for our project in a much more precise manner. Most of the data we acquired was in the form of hard copies. For example the soil, climate, weather, market data are not so precise and some relevant data needed are currently not available. By the time when there is a precise mechanism to identify these data, more precise prediction of crops and better features can be implemented in this application. Currently this project is constrained to Kerala and all its districts only. As part of a future plan we would like to implement this application all over India and include more native languages in order to make it more user friendly.

By creating an algorithm to predict weather data precisely the application could be helpful to the farmers in taking appropriate decisions. As of now we have considered four crops only. So as an extension to our work, more number of crops can be added to it. Most of the data regarding crops are in hard copy format and the dataset we created has its own time constrained limitations. By the availability of more datas, our dataset can be much more elaborated and with the extension in the parameters, the prediction can be made much more precise.

The map system which we have generated containing different information about other users or farmers have great scope .i.e., the map can be further modified with implementing additional features where each user can create room for chatting, posting videos and images.

Moreover, it is possible to implement a payment system where the farmers can directly deal with the collectors than with intermediates to sell their crops. This will definitely increase the commerciality of the application.

In addition to all the features as listed above, if we are able to implement a crop disease identification system on the basis of crop diseases dataset and image processing, it will make up for much a greater functionality of the current system.

## REFERENCES


1. Benefits of Anomaly Detection Using Isolation Forests, https://blog.easysol.net/using-isolation-forests-anamoly-detection/
2. Anomaly Detection Using Isolation Forest in Python, https://blog.paperspace.com/anomaly-detection-isolation-forest/
3. Reading and Writing CSV Files in Python, https://realpython.com/python-csv/
4. Isolation Forest Algorithm, https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.IsolationForest.html
5. Anomaly Detection Using Isolation Forest, https://lambda.grofers.com/anomaly-detection-using-isolation-forest-80b3a3d1a9d8
6. Machine learning for anomaly detection and condition monitoring, https://towardsdatascience.com/how-to-use-machine-learning-for-anomaly-detection-and-condition-monitoring-6742f82900d7
7. Pandas Data Frame, https://www.w3resource.com/pandas/dataframe/dataframe-dropna.php
8. M. N. Wibisono, A. S. Ahmad. **Weather forecasting using Knowledge Growing System (KGS),**2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, pp. 35-38, 2017 doi: 10.1109/ICITISEE.2017.8285526.
9. Y. Gandge, Sandhya. **A study on various data mining techniques for crop yield prediction,** International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), Mysuru,pp.420-423, 2017 doi: 10.1109/ICEECCOT.2017.8284541.
10. S. Veenadhari, B. Misra and C. Singh, **Machine learning approach for forecasting crop yield based on climatic parameters,**International Conference on Computer Communication and Informatics, Coimbatore, pp. 1-5,2014 doi: 10.1109/ICCCI.2014.6921718.
11. M. Paul, S. K. Vishwakarma and A. Verma, **Analysis of Soil Behaviour and Prediction of Crop Yield Using Data Mining Approach,** International Conference on Computational Intelligence and Communication Networks (CICN), Jabalpur, pp. 766-771, 2015 doi: 10.1109/CICN.2015.156.
12. S. Ramesh et al., **Plant Disease Detection Using Machine Learning**, International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), Bangalore, pp. 41-45, 2018 doi: 10.1109/ICDI3C.2018.00017.
13. J. Liu and J. Yu, **Research on Development of Android Applications,** 4th International Conference on Intelligent Networks and Intelligent Systems, Kunming, 2011, pp. 69-72, 2011 doi: 10.1109/ICINIS.2011.40.
14. Keerthana R, Dr. Varghese S Chooralil. **Forecasting of the Air Pollution Based on Meteorological Data and Air Pollutants using Deep Learning: A Novel Review**, Vol. 9 No.1, pp. 801-807, January – February 2020
15. Thomas, Monica, and V. S. Chooralil. **Security and Privacy via Optimised Blockchain.** International Journal of Advanced Trends in Computer Science and Engineering, Vol. 8 No. 3, pp. 415-418, 2019
16. Vinodh P Vijayan, Deepti John, Merina Thomas, Neetha V Maliackal, Sara Sangeetha Varghese **"Multi Agent**

**Path Planning Approach to Dynamic Free Flight Environment"**, International Journal of Recent Trends in Engineering (IJRTE), ISSN 1797-9617 Volume 1, Number 1, May 2009, Page(s): 41-46.

17. Juby Joseph, Vinodh P Vijayan**" Misdirection Attack in WSN Due to Selfish Nodes; Detection and Suppression using Longer Path Protocol"** International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 7, July 2014, pp. 825-829, ISSN: 2277 128X
18. V P Vijayan, Biju Paul **"Multi Objective Traffic Prediction Using Type-2 Fuzzy Logic and Ambient Intelligence"** International Conference on Advances in Computer Engineering 2010, Published in IEEE Computer Society Proceedings, ISBN: 978-0-7695-4058-0, Print ISBN: 978-1-4244-7154-6
19. Vijayan V P, Gopinathan E **"Improving Network Coverage and Life-Time in a Cooperative Wireless mobile Sensor Network "** Fourth International Conference on Advances in computing and communications (ICACC) Aug, 2014. Published in IEEE Computer Society Proceedings. Print ISBN: 978-1-4799-4364-7, INSPEC Accession Number:14630874,DOI:10.1109/ICACC.2014.1 6 PP 42-45.
20. Vinodh P Vijayan, Biju Paul **" Traffic scheduling for Green city through energy efficient Wireless sensor Networks "** International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, No.4, July – August 2019, ISSN 2278-3091, https://doi.org/10.30534/ijatcse/2019/81842019.