1. PROJECT REFERENCE NUMBER: 46S_BE_3526

2. TITLE OF THE PROJECT:

REAL TIME DETECTION AND PREDICTION OF AGRICULTURAL LAND DEGRADATION FROM URBANIZATION INTRUSION USING SATELLITE IMAGERY.

3. COLLEGE NAME: PES INSTITUTE OF TECHNOLOGY AND MANAGEMENT, SHIVAMOGGA

DEPARTMENT: INFORMATION SCIENCE & ENGINEERING.

4. STUDENTS:

ANUSHA A VERNEKAR [4PM19IS053]

SHEETHAL M D [4PM19IS042]

R SRINIDHI [4PM19IS039]

DARSHAN R [4PM19IS011]

GUIDE: MR. AMIT KUMAR K

5. KEYWORDS:

Prediction, Agricultural Land Degradation, Urbanization intrusion, Satellite imagery, Machine Learning, K-Means Clustering Algorithm, Random Forest Regression.

6. INTRODUCTION:

The economic development of any agricultural region directly depends on its land and resource potential. The management of agricultural land use is a complex system that includes theforecasting and planning of land use, land supervision, land monitoring and so on. The forecastingwith agricultural land use is the most important pre-project activity and initial stage of the agricultural land use management process the main goal of which is rational use and protection ofland. To achieve goals regarding the prediction of agricultural land use, it is necessary that the chosen method of forecasting should correspond to the maximum extent. Modern agriculture has to cope with several challenges, including the increasing call for food, as a consequence of the global explosion of earth's population, climate changes, natural resources depletion, alteration of dietary choices, as well as safety and health concerns.

Thousands of hectares of agricultural land are lost by urbanization all over the world every year. This is due to the rapid population growth and the fact that human values change overtime and production demands by humans cannot be achieved without alteration or conversion of land use/cover. But also it is important to save the land for agriculture purpose which is ultimately a sourcefor the human sustention.

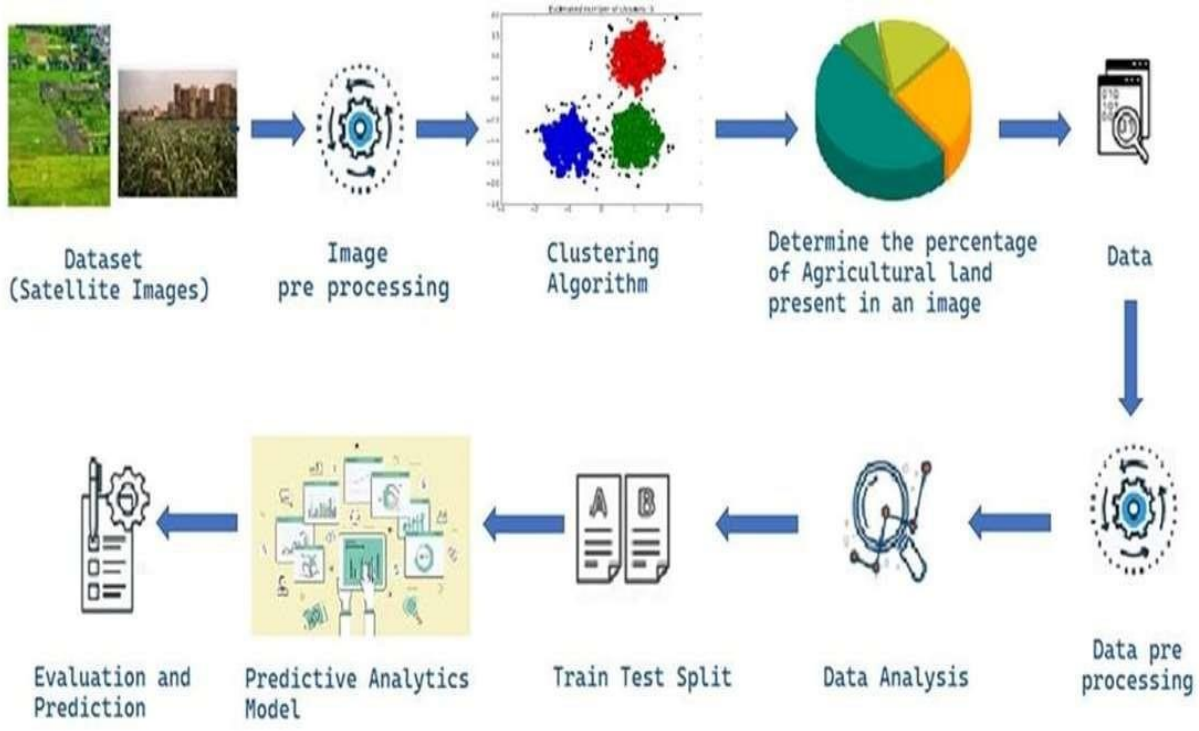
Machine Learning (ML) has emerged, which is a subset of artificial intelligence,by taking advantage of the exponential computational power capacity growth. There is a plethora of applications of ML in agriculture. At present we have lost most of the productive land which was suitable for the agriculture purpose from urbanization intrusion which has resulted in huge loss to the farmers as well the society. It is important to analyze and predict the productivity of the land before the urbanization intrusion so that the loss of the productive land can be avoided thereby helps to choose the unproductive land for urbanization and reduce the loss incurred.

7. OBJECTIVES:

1. To analyze the agricultural land degradation from urbanization.
2. Predict the agricultural land rate over that area in the future.

8. METHODOLOGY:

Workflow:



1. Satellite images of agriculture land, a unlabeled dataset is extracted from google earth pro which is an web application where we can download past satellite images.

2. Images from google earth pro is subjected to some pre-processing in order to make an improvement of the image data that suppresses unwilling distortions or enhances some image features important for further processing, although geometric transformations of images.

3. Once the image pre-processing is completed, the percentage of agriculture land is extracted by using k-means clustering algorithm. Here value of K is 2, i.e., two cluster are considered to group green shades to one cluster and all other colors to another cluster.

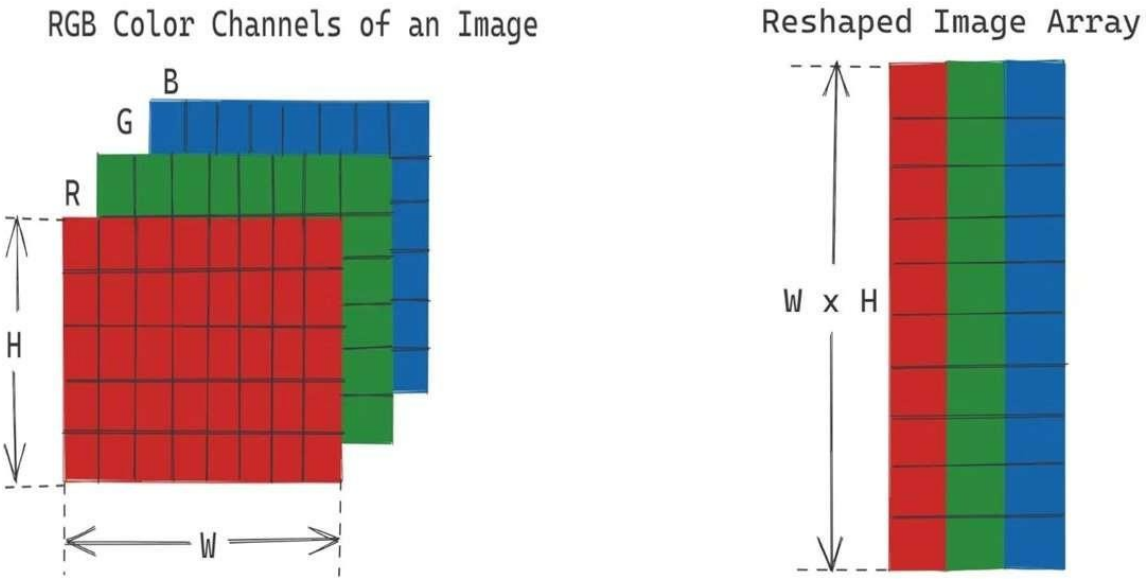
4. Data pre-processing and filtering steps can take considerable amount of processing time. Examples of data preprocessing include cleaning, instance selection, normalization, one hot encoding, transformation, feature extraction and selection.

5. Data Analysis is the process of systematically applying statistical and/or logical techniques to describe and illustrate, condense and recap, and evaluate data.

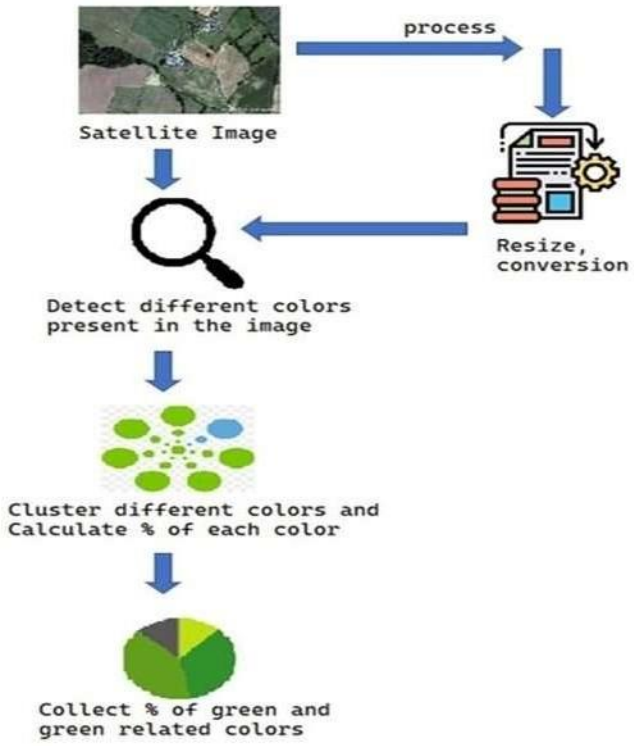
6. The simplest way to split the modelling dataset into training and testing sets is to assign 2/3 data points to the former and the remaining one-third to the latter. Therefore, we train the model using the training set and then apply the model to the test set. In this way, we can evaluate the performance of our model.

7. After splitting the dataset, we train the machine using training dataset. Random Forest Regression is used because of its high accuracy to do the predictive analysis.

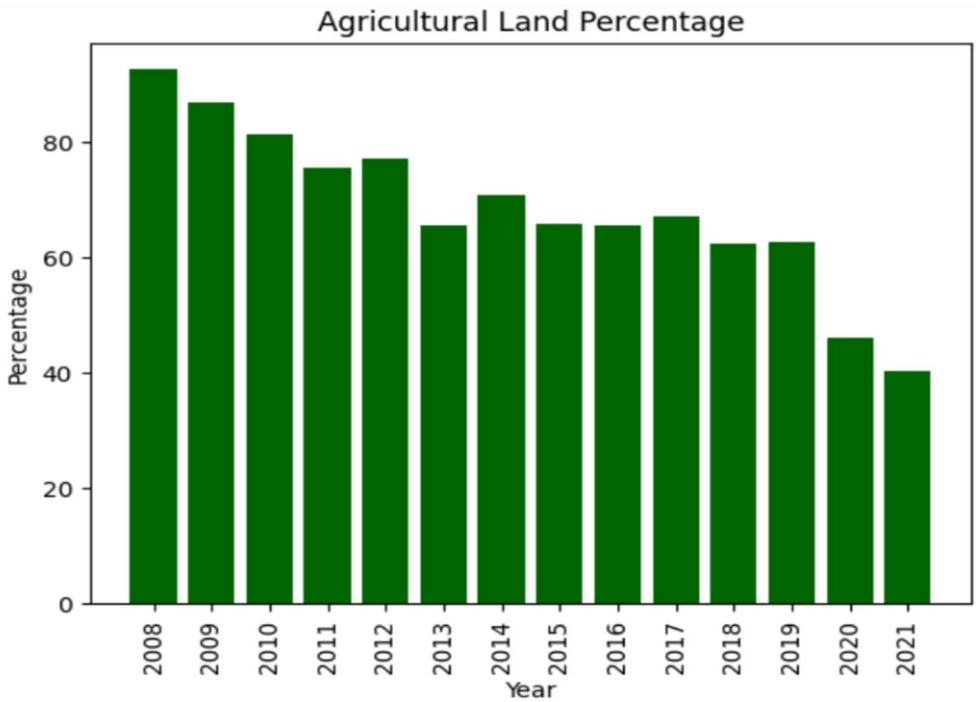
Image pre-processing:



Steps to collect the percentage of agricultural land:



9. RESULTS:



Satellite images of a particular land are processed with respect to year. Image is converted to numpy array and resized to 2-D array. These represents image pixels. These are subjected to K-Means clustering where we have 2 clusters one cluster represents green and green related colors and another cluster for non-green. According to pixel intensity and RGB values the pixels(data

points) are clustered. The value of number of agricultural land area pixels present in an image is derived by K-Means clustering algorithm, also total pixels are already calculated. Hence, percentage of agricultural land present in that image is calculated.

Similarly, the agricultural land percentage with respect to year for a particular land is calculated and a graph is plotted to analyze the percentage deviation.

Year	Agricultural Land Percentage
2008	43.23581
2009	41.49853
2010	48.45729
2011	31.90777
2012	64.0136
2013	36.60494
2014	41.64959
2015	39.87559
2016	40.06275
2017	50.36363
2018	41.00115
2019	28.28214
2020	36.77822
2021	45.00411

The agricultural land percentages derived by using K-Means clustering is stored in a excelfile automatically. It is stored as percentage with respect to the year. This is used as a dataset to pass through the prediction model.

```
Predicted percentage of agricultural land in future: 66.93071075779034
Predicted percentage of agricultural land degradation in future: 33.06928924220966
Accuracy of the model: 97.57981806840216
```

Prediction model is build using Random Forest Regression algorithm to study the past and present percentage of agricultural land of a particular area and predict the percentage of agricultural land in that area in future.

CONCLUSION:

Agriculture is crucial in terms of employment, food security and budgetary allocation,

eventhough the share of agriculture income has decreased in the national market. In the growing population it's important to conserve and protect potential farmlands. Vertical construction shouldbe promoted. The government should ensure the developmental projects in certain areas other thanproductive agriculture land, strengthen the zoning regulation and enact policies to reduce agriculture land losses.

10. SCOPE FOR FUTURE WORK:

Increase investment in agricultural capital and technology, and accurately implement agricultural subsidies. Considering the significant impact of terrain conditions on nature-induced marginalization, increase investment in agricultural capital and technology, improve the quality of cultivated land through the implementation of cultivated land consolidation, carry out terrace construction in mountainous and hilly areas according to local conditions, improve agricultural production conditions, and build high-standard farmland, so as to increase the productivity of arable land. Speed up the progress of land consolidation in an orderly manner and promote the intensive utilization of cultivated land. Unfavourable farming conditions are the main reason for the transformation of cultivated land to other types of land, and the fragmentation of cultivated land limits the large-scale operation of cultivated land. Considering the significant impact of per capita arable land, total power of agricultural machinery, and average land plot area on the farmland marginalization, it is necessary to speed up land consolidation, integrate scattered arable land in suburban areas, and adjust arable land structure, so as to promote large-scale arable land management.