IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

"LULC Classification by Sentinel-2B Data using Machine Learning Techniques in Virajpet Taluk"

Souparnika H Koppal[1], *Student, Dept of Geoinformatics, Karnataka State Rural Development and Panchayat Raj University Gadag.*
Ravi Jadi[1], *Faculty, Dept of Geoinformatics, Karnataka State Rural Development and Panchayat Raj University Gadag.*
Smt. Shivam Trivedi[2], *Scientist/Engineer 'SE', RRSC-South, Dept of Space, ISRO Bengaluru - 560 037.*

ABSTRACT

Vegetation extraction from remote sensing imagery is the process of extracting vegetation information by interpreting satellite images based on the interpretation elements such as the image colour, texture, tone, pattern and association information. The vegetation maps are critical for understanding biodiversity management and planning from local to global scales. There are two essential elements of vegetation maps that are based on classification of vegetation and spatial attribution of that classification. The difference between these vegetation cover can be done with using remote sensing and GIS. Remote sensing classification is an essential technique for mapping forest vegetation attributes like floristic composition, biomass, tree density etc. that addresses the considerable logistical challenge of measuring, mapping and monitoring across very large areas in an accurate, repeatable and cost-effective way. The most basic method is identifying plant communities or species is to create a vegetation map.

Machine learning (ML) is a subdivision of artificial intelligence in which the machine learns from machine-readable data and information. It uses data, learns the pattern and predicts the new outcomes. Its popularity is growing because it helps to understand the trend and provides a solution that can be either a model or a product. This study is carried out to perform various machine learning algorithms for vegetation classification using remote sensing and ancillary spatial data across a heterogenous forest ecosystem in Virajpet Taluk, Kodagu district. The recently developed sentinel-2 satellite imagery holds great potential for improving the vegetation type classification at medium to large scales due to the concurrent availability of multispectral bands with high spatial resolution and quick revisit time. A study is classifying using Supervised classification, a type of machine learning technique. To perform the supervised classification, training samples have been generated using ARCGIS to run the spatial model in ERDAS software. The machine learning algorithm such as support vector machine (SVM), Random Forest (RF), k-Nearest Neighbour (k-NN), Navies Bayes and CART were used to classify. In terms of classification accuracy, a combination of these parameters was tested: The number of trees (NT), the variables per split (VPS) and the Bag Fraction (BF). Thus, this studies key contribution is classification of vegetation types and how the ML techniques are an efficient approach to map different land use and land cover classes, including

different vegetation types and demonstrated that some algorithms perform this task with higher accuracy than the others. In our investigation, the RF algorithm chieved satisfactory results when compared with others and characterized with more accurate rate and less error rate.

keywords: Multispectral bands, Random Forest, Machine Learning, SVM, Forest Classification.

Chapter- 1

INTRODUCTION

This chapter cover features, merits and application of remote sensing and geographical information (GIS). This also covers about the vegetation, remote sensing in vegetation and classification using advanced analytical techniques like machine learning algorithms.

1.1 REMOTE SENSING AND GEOGRAPHICAL INFORMATION SYSTEM(GIS)

Remote sensing provides another tool that can be integrated into GIS. Remote sensing is the art and science of making measurements of the earth using sensors on airplanes or satellites. These sensors collect data in the form of images and provide specialized capabilities for manipulating, analysing and visualizing those images.

Remote sensing is of two types, active remote sensing, and passive remote sensing. The active remote sensing emits their electromagnetic radiation, which then interacts with the object or the area, e.g., RADAR and LIDAR, whereas passive remote sensing using the radiation from the sun. The sensors record the energy which is emitted back from the surface of the object. These sensors are either air-based or space-based.

The radiations recorded by the sensors have to pass through the atmosphere; the quality of the data gets degraded due to the absorption and scattering of radiation. Image processing enhances the image, helps in better visual interpretation, and restores the image if it is subjected to geometric distortion, blurring.

The quality of remote sensing data consists of its spatial, spectral, radiometric and temporal resolution.

Spatial resolution: The pixel size which is recorded in a raster image.

Spectral Resolution: The wavelength of different frequency bands recorded – usually, this is related to the number of frequency bands recorded by the platform.

Radiometric Resolution: The number of different intensities of radiation the sensor is able to distinguish. It depends on the instrument noise.

Temporal Resolution: The amount of time needed to revisit and acquire data for the exact same location.

A Geographic information system (GIS) is a computer-based tool for mapping and analysing feature events on earth. GIS provides tools for display and analysis of various statistics, like vegetation types. The GIS allows capturing, manipulating, calculating, storing, and analysing the data.

GIS gathers concepts from different disciplines like cartography, engineering, surveying, remote sensing, photogrammetry, environmental sciences, geodesy. Data for GIS can be obtained from aerial photographs, satellite imageries, metrological department, field survey. Spatial data and attribute data are the two types of data that can be acquired. Spatial data describes the location and attribute data specify descriptions, measurements, or classifications of the geographical features. The spatial data is either in the form of raster or vector data. Vector data can be made up of points, lines, or polygons, with each object having one or more attribute values. Raster data is in the form of a grid where each cell has an attribute value. GIS show many different kinds of data on one map, such as streets, buildings and vegetation this enables to analyze understand patterns and relationships. GIS works on the location based which is expressed in latitude and longitude, address, or ZIP code. Many different types of information can be compared and contrasted using GIS.

1.2 IMAGE CLASSIFICATION

Image classification is the process of extraction of information from satellite image. It refers to a process in computer vision that can classify an image according to its visual content. Image classification is assigning pixels in the image to categories or classes of interest Examples: built-up areas, water body, green vegetation, bare soil, rocky areas, cloud, shadow etc. in order to classify a set of data into different classes or categories, the relationship between the data and the classes into which they are classified.

Image classification is a procedure to automatically categorize all pixels in an Image of a terrain into land cover classes. Normally, multispectral data are used to Perform the classification of the spectral 2 pattern present within the data for each pixel is used as the numerical basis for categorization. It also refers to the task of extracting information classes from a multiband raster image. The resulting raster from image classification can be used to create thematic maps. Depending on the interaction between the analyst and the computer during classification, there are two types of classification: supervised and unsupervised.

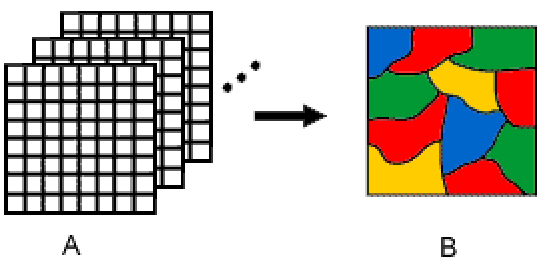


Fig 1.1: Image classicication(source:www.nrcan.ga.ca)

1.3 SUPERVISED CLASSIFICATION

Supervised classification is based on the idea that a user can select sample pixels in an image that are representative of specific classes and then direct the image processing software to use these training sites as references for the classification of all other pixels in the image. Representative sample sites of known cover type, called training sets, are used to compile a numerical "interpretation key," describing the spectral attributes for each feature type. Each pixel in the data set is then compared numerically to each category in the interpretation key, hence classifying the image into different classes of interest.

1.4 UNSUPERVISED CLASSIFICATION

Unsupervised classification is an effective method of partitioning remote sensor image data in multispectral feature space and extracting land-cover information. Unsupervised classifiers involve algorithms that examine the pixels in an image and aggregate them into several classes based on the natural groupings or clusters present in the image values. The classes resulting from the unsupervised classification are spectral classes. The classified data is compared with some form of reference data (such as toposheets/maps) to determine the identity of the spectral classes.

Numerous clustering algorithms can be used to determine the natural spectral groupings called the "K-means" approach. In the K-means approach, each pixel in the image is assigned to the cluster whose arbitrary mean vector is closest. After classifying all the pixels, the revised mean vectors for each of the clusters are computed. The revised mean vectors are then used to reclassify the image data. The procedure continues to take place until there is no substantial change in the location of class mean vectors between any successive iterations of the algorithm.

1.5 HIGH RESOLUTION OPTICAL REMOTE SENSING

Remote sensing systems that acquire images with large spatial extents will generally have a lower resolution, and thereby capture less detail, than images acquired at a higher resolution, which usually depict forest characteristics across smaller spatial extents. Optical remote sensing makes use of visible, near infrared and short-wave infrared sensors to form images of the earth's surface by detecting the solar radiation reflected from targets on the ground. Different materials reflect and absorb differently at different wavelengths. Thus, the targets can be differentiated by their spectral reflectance signatures in the remotely sensed images. Optical remote sensing systems are mainly classified Panchromatic, Multispectral, and Hyper spectral depending on their spectral resolution.

1.6 IMAGE INTERPRETATION

Image interpretation is a powerful technique enable us to identify and distinguish various features in remote sensing images/Aerial photos and allows gaining the knowledge and information about them.

1.6.1 Image Interpretation using visual elements of Interpretation

Tone: Tone is the particular quality of brightness, deepness, or hue of a shade of a colour. Therefore, tone refers to relative brightness or colour of a feature on an image. The tonal variation makes it easier to differentiate between various features on an image. Shapes, patterns and textures on an image are identifiable mainly due tonal variation.

Shape: Shape refers to external form, outline or structure of a particular feature. The man-made features generally have regular, symmetric or sharp in shape while all natural features like forest patches are irregular in shape. Most of the features can solely be identified using the shape element of visual interpretation.

Size: Comparing size of a feature in context with others in an image helps in better understanding and interpretation of an image. A quick approximation of size of a feature makes image interpretation process faster and convenient.

Pattern: Pattern refers to spatial arrangement of features. A repeated sequence of certain form or relationships is characteristic of many natural and constructed features which give an added advantage to the interpreter.

Texture: Texture refers to frequency of tonal changes in a certain area of an image. It is product of shape, size, tone, shadow and pattern of a particular feature.

Association: Association is occurrence of certain feature in relation with other. Certain features are not directly identifiable by its appearance in an image but could be interpreted easily according to its relationship with the surroundings.

1.6.2 Visual Interpretation using colour composite:

Different colour composite images provide an aid for better and much faster interpretation. Colour composite images are displayed using different band combinations on three primary colours (Red, Green, and Blue), associating each spectral band of an image (not necessarily a visible band) to a separate primary colour.

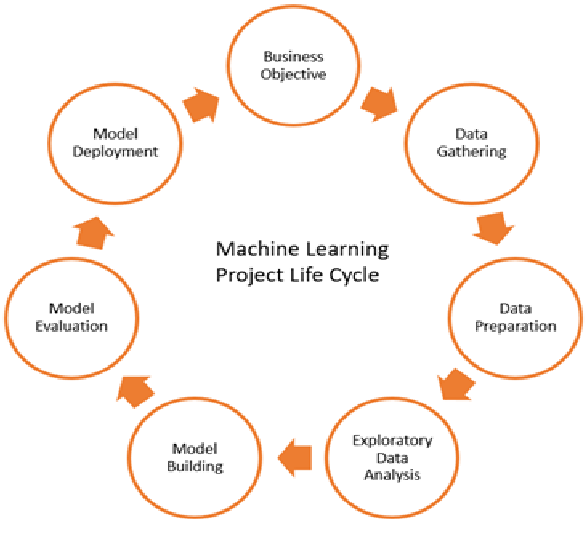
1.7 APPLICATIONS OF REMOTE SENSING AND GIS

- ***Marine and Oceanography-***Coastal zone mapping, Physical Oceanography, Marine resources
- ***Regional/Land development-***Town and country planning, Solid waste disposal
- **Forest, Biodiversity and Environment-**Forest cover, Biodiversity characterization, Environmental Impact Studies, Forest fire mapping
- ***Water-*** Potential drinking water zones, monitoring catchment and reservoir areas, Watershed development.
- ***Land use Monitoring-*** Wasteland mapping, Urban development, Geology, Topography, and Geographic positioning.

1.8 MACHINE LEARNING WITH REMOTE SENSING

Machine learning offers the potential for effective and efficient classification of remotely sensed imagery. The strengths of machine learning include the capacity to handle data of high dimensionally and to map classes with complex characteristics. Several machine learning algorithms have been used in the remote sensing community since decades, ranging from basic algorithms such as PCA and K-Means to more sophisticated classification and regression frameworks like support vector machine (SVM), Random Forest (RF), decision tree (DTs) and K-

Fig 1.2: Machine learning life cycle (source: pianetytire.com)



nearest neighbours(K-NN). Machine learning with remote sensing can help to improve prediction about the behaviour of environmental system, improve the automation of data analysis, lead to a better management of resources and the discovery of insights from complex data sets.

1.9 APPLICATION OF MACHINE LEARNING IN REMOTE SENSING

- Machine learning with remote sensing help to improve prediction about the behaviour of environmental systems, improve the automation of data analysis, lead to a better management of resources and discovery of new insights from complex data sets.
- Machine learning algorithm helps to find the patterns in data and uses these patterns to improve prediction.

- It helps to make a decision making more accurately.

1.10 TYPES OF MACHINE LEARNING ALGORITHMS

Supervised Classification*:* Supervised learning provides a clear expectation of outputs after input samples such as classification and regression have been trained through the model.

Type of Machine Learning Algorithm

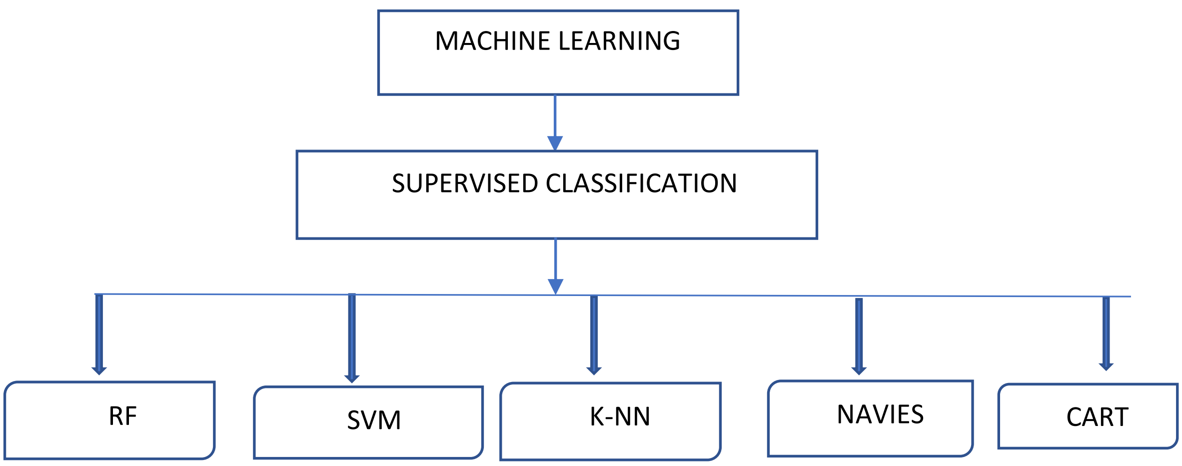


Fig 1.3: Types of machine learning

Unsupervised Classification*:* Unsupervised learning is relatively unpredictable in what type of output is generated after input samples have been trained on the model.

Reinforcement learning*:* Reinforcement learning focuses on regimented learning processes, where a machine learning algorithm is provided with a set of actions, parameters and end values. Reinforcement learning teaches the machine trail and error. It learns from past experiences and begins to adapt its approach in response to the situation to achieve the best possible results.

1.11 Spatial Modeler:

Spatial modeler is a spatial modelling suite that provides robust extensible processing architecture coupled with a responsive and initiative user experience. The system is extensible in both data types and operations on data types, which has enabled it to support a rich set of feature operators and point cloud operators in addition to the traditional raster function. It is the heart of every ERDAS imagine process, making it easy to string together multiple models to increase efficiency and streamline of workflow. Spatial modeler provides with hundreds of functions, algorithms and analytical routines that can easily be chained together into models that solve complex problems.

Spatial model is a collection of operators and is presented as a canvas containing the set of interconnected operators. The main use of the model is to create s self-contained Functional piece that can be re-used in other solutions. Spatial models are by default saved in files using a. gmdx extension model and may also be saved in json format

(. json extension) or in a compiled format (. gmdx extension).

1.12 Accuracy Assessment

Accuracy assessment is an important and essential step in the classification process. Its goal is to quantitatively determine how effectively pixels were grouped into the correct feature classes in the area under investigation. It generally includes three basic components: sampling design, response design, and estimation and analysis procedures.

1.13 Kappa Co- Efficient

It is a discrete multivariate method of use in accuracy assessment. In classification process, where pixels are randomly assigns to classes will produce a percentage correct value. pixels are not assigned randomly during image classification, but there are statistical measures that attempt to account for the contribution of random chance when evaluating the accuracy of a classification. The resulting Kappa measure compensates for chance agreement in the classification and provides a measure of how much better the classification performed in comparison to the probability of random assigning of pixels to their correct categories.

In current work, vegetation mapping is mostly conducted using intelligent classification technologies with multi-source remote sensing images and auxiliary information. Knowledge mining, expert systems, pixel-based classification and targeted classification strategies have been successfully introduced in vegetation classification. Remotely sensed data are extensively and effectively used for vegetation classification. The freely availability of optical satellite data provides more opportunities to use a multitude of images, making it possible to map more thematically detailed classes with high accuracy. This study is classifying the vegetation in an analytical way by using remotely sensing based techniques and analysing its quantitative characters. Sentinel-2 data is used It includes classification of vegetation types to explain the spatial pattern of composition/ Structural/functional biodiversity.

Sentinel-2 data is used to carry out the study with 10m resolution. Recently launched sentinel satellite is receiving much attention due to its fine spatial resolution, fast revisit time, global coverage, last but not least free availability makes it a great choice for various application in the field of remote sensing. Training samples are generated in ARCGIS at a Pixel level. Around thousand samples were taken to run the supervised classification with fifteen classes. Some classes considered as the forest type (evergreen, deciduous and gregarious) and some other to be vegetation types (tea plantation, coffee plantation), waterbodies (rivers, ponds), lands with (crop land, fallow land and open land) and built-up area as settlement.

1.14 VEGETATION

Vegetation is an assemblage of plant species and the ground cover they provide.
The term vegetation is used in ecology to describe overall characteristics of plant cover in referring to dominant plant growth forms or structural characteristics, using either technical description and referring to specific plant communities.

Vegetation integrates the combined influence of variety of environmental factors such as climate, topography, soil parent material and time.

1.15 VEGETATION TYPE DESCRIPTION:

Tropical Evergreen Forest: This forest occupies about 7% of the earth's surface. Tropical evergreen forest is known as rainforest because they receive more than 200cm rainfall per year. An evergreen forest is a forest which remains green throughout the year. There is not a particular season where trees start lose their all leaves.

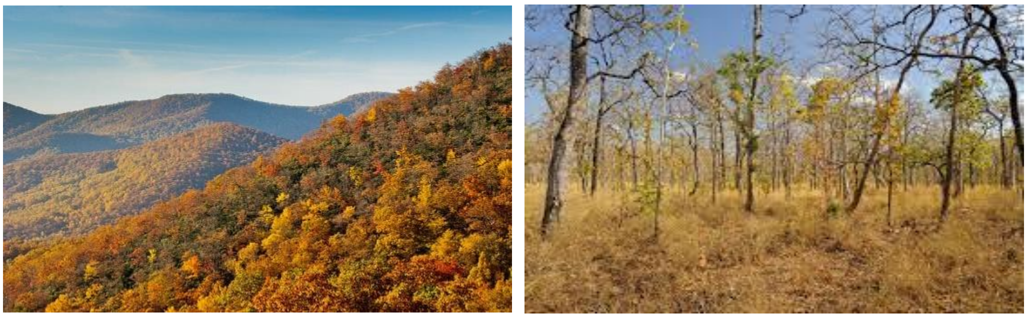
Fig 1.4: Tropical evergreen forest

Tropical Semi-Evergreen Forest: Tropical evergreen forest is the transition type of forest between tropical wet evergreen and tropical deciduous forest. These forests are less dense and characterized by many species. The climbers which are under grown provide an evergreen character to this forest.

Fig 1.5: Tropical semi-evergreen forest

Tropical Decidious Forest: Decidious forest are located in the mid-latitude regions with a temperature climate characterized by the winter season and year-round precipitation. That they lose their leaves every year.

Fig 1.6: Tropical decidioud forest



Tropical dry deciduous forests: These are forests of low or moderate height consisting almost entirely of deciduous species. Their canopy is typically light though it may appear fairly dense and complete during the short rainy season.

Fig 1.7: Tropical dry deciduous forest

Gregarious Forest: Plants growing in an open cluster

Fig 1.8: Gregarious Forest

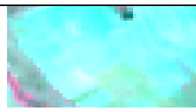
Scrub vegetation : It is a vegetation dominated by shrubs, i.e low ,woody plants which typically forms an intermediate community between grass and health and hifh forest.

Fig 1.9 Scrub Forest

1.16 DEFINITION OF TERMS AND CLASSIFICATION:

The image feature classification based on the tones and texture from the Sentinel 2 image. The parcel of the ground feature as shown below *(table 1.1)*

TABLE 1.1: VEGETATION TYPE WITH LU/LC CLASS

Class code	Symbols	Class
1		Cropland
2		Deciduous Forest-1
3		Deciduous Forest-2
4		Evergreen Forest-1
5		Evergreen Forest-2
6		Fallow-land
7		Gregarious Forest
8		Open-land
9		Others
10		Settlement
11		Tea Plantation
12		Waterbody-1
13		Waterbody-2

14		Coffee-Plantation
15		Shadow

1. ***Plantations:* A** plantation is the large number of the tress, planted for the purpose of afforestation or reforestation or can be for commercial purposes. It includes teak, neem, Sal and eucalyptus.
2. ***Fallow land**:* It is a type of land that has no crop on it, usually for a year. But, it’s a farmland.
3. ***Water-bodies***: This category comprises areas with surface water either in the form of pounds, lakes and reservoirs or flowing as streams, rivers, canals.
4. ***Settlement***: It is the area of human habitation which includes cities, towns, villages, built-up area, transportation facilities and industries.
5. ***Cropland*****:** Area used for the production of adapted crops for harvest.
6. ***Open-land*****:** Non-built up land with no, or with insignificant, vegetation cover.
7. ***Tea plantation*****:** Tea plantation look like huge forest made up of small trees that rarely reach above 1.5m in height.
8. ***Gregarious Forest*****:** Determined as growing in patches among other vegetation, thus contrasting with social species, which is dominating the whole world.
9. ***Evergreen /semi-evergreen Forest:*** These are the areas that comprise of thick and dense canopy of tall trees which remain green throughout the year.
10. ***Deciduous Forest*****:** These are the type of forest, that are mainly Composed of species, which shed their leaves once a year, especially in a summer

CHAPTER-2

OBJECTIVES OF STUDY

- To determine the Major vegetation types using machine learning techniques.
- To evaluate the performance of selected machine learning algorithms.
- To determine the Performance of accuracy assessment to processed algorithms.

CHAPTER-3

REVIEW OF LITERATURE

3.1 CLASSIFICATION OF TREE AND SHRUB SPECIES IN KSU RESEARCH AND APPLICATION FOREST IN KAHRAMANMARAS, TURKEY.

Remote sensing and GIS technology have become increasingly important tools for mapping, inventorying and monitoring forest resource around the world. In remote sensing, the classification is described as the process of separating features into classes (***Raffy, 1994***). Due to advanced capabilities of these technologies to provide, organize and analyze large amount of spatial data, land classification has been used for mapping, inventorying and monitoring purpose in the field of natural resource ***(Carson et.al 2001).***

In the field of forestry, digital interpretation procedures have been widely used for the monitoring forested area based on satellite imagery ***(Gougeoan,1995).*** Large-scale satellite images with high resolution and detailed information have been used for the pixel-based classification ***(Leckie 1990, Beaubien 1994, Meyer et al, 1996).***

Sometimes, large amount of pixel lead to potential classification problems like some pixels in tree crown can be confused with shrub species, this will create variation in background vegetation and soil materials cause high frequency of data variability. To overcome this problem automated image classification method was applied to identify vegetation in KSU research and application in forest in kahramanmaras-Turkey based on satellite imagery.

3.2 APPLICATION OF MACHINE LEARNING IN FOREST ECOLOGY: RECENT PROGRESS AND FUTURE CHALLENGES.

About 30% of total earth is covered by forest and they are dominant terrestrial ecosystem on earth ***(Schmitt et al.2009***). The increase in the availability of large amount of data and development of data analysis methods capable of handling large datasets are providing new opportunities to study complex systems ***(Flach 2001: crisci et al.2012.).*** ML is a branch of artificial intelligence, which provides significant advantages over traditional statistical method for analyzing forest ecological data when large number of datasets are available as training model sets. The application of the ML process mainly:(i) selection of relevant data and its processing, (ii) the selection of adequate algorithms, and (iii)its quality assessment solutions ***(Muhamedyev 2015).*** ML approaches powerful and efficient ways to deal with data ***(Bhattacharya 2013; Thessen 2016***). ML models effectively improve the accuracy of models ***(Garzon et al.2008; vaca et al. 2011).*** ML algorithms in forest ecosystem are: Decision Tree (DT) learning, Support Vector Machine (SVM) and Random Forest (RF).

Tree-based learning (DT)

DT learning is a predictive model and a support tool that combines a decision graph with possible outcomes or result. DTs have a simple recursive structure composed of root nodes, internal nodes, leaf nodes and branches that represents the knowledge extracted from data (***Quinlan 1987).*** The root to leaf path represents the value of target variables that is conditional to value of input variables. The logical expression of DTs are as "white box" models (***Breiman et al. 1984***).

Random Forest (RF)

Another, DT approaches is random forest (***Breimen 2001).*** RF method generate and aggregate results of multiple trees using boot strap samples of input data ***(Svetnik et al.2003***). In RF, reliance on multiple trees is used to avoid overfitted trees. It mainly depends on three random processes: (1) The samples that generates DTs are randomly generated (2) the eigenvalues for building a DT are randomly selected, and (3) random direction is chosen for tree fission selection during production process.

Support Vector machine (SVM)

The non-parametric kernel-based techniques that is derived from statistical learning theory used by the SVM, which was primarily invented and developed by vapnik ***(Vapnik 2000; Vapnik and chervenenkis 2015).***

3.3 A SURVEY OF VARIOUS SUPERVISED ALGORITMS

3.3.1 *Narayanan, u., Unnikrishnan, A., paul, V., & Josph, S.,* Study examined five classification methods, navies bayes, neural network, decision tree, support vector machine and K-nearest neighbour and presented a taxonomy of each paper and shortcoming. Additionally, they categorized the paper according to research topic, classification algorithm, survey includes variety of disciplines, including medicine, agriculture, education, business and networking. According to research most frequently used classification algorithms are decision tree and navies bayes.

3.3.2 *Caruana, R., & Niculescu-mizil,* A., Study compares supervised classification method empirically. Use of eight comparison parameter to contrast the fallowing supervised algorithm: SVM, ANN, Logistic regression, Navies Bayes, KNN, Decision tree, Random Forest, Bagged tree, memory-based learning and boosted stumps. Those parameters were: Accuracy, Precision, and recall, F-Score, Cross entropy, Roc Curve, square error, average precision and break-even point. According to their findings, calibrated boosted trees performed all by scoring highly on all comparison parameters. The performance of logistic regression, naive bayes and decision tree was, however, poor. It is worth noting that model's calibration is surprisingly effective at producing an excellent performance.

3.4 TREE SPECIES CLASSIFICATION USING SENTINEL-2 IMAGERY AND BAYESIAN INFERENCE

This paper tells about the freely available optical satellite data. It Provides more opportunities to use a multitude of images, making it possible to map more thematically detailed class with higher accuracy. The use of multi-temporal data for land cover classification has been shown to improve accuracy of tree species classification**.** ***(Amani et al., 2017., Immitzer et al., 2019***). The tree classification is one by using optical data which is using to multi-temporal images where phenological differences used to separate classes ***(Fassnacht et al. 2016; Hagner & Reese , 2007; Hill et al; 2010Reese et al. 2002; wolter et al; 1995***). All the scenes of sentinel data may not be clear so the selection of scenes is done. By this random forest classification came with an 85 percent accuracy. Study tells about the results came across by using random forest classification and with an accuracy of eighty five percent.

3.5 USE OF SENTINEL-2 FOR FOREST CLASSIFICATION IN MEDITERRANEAN ENVIRINMENT.

Forest classification and type of forest are strongly required for addressing a wide range of ecological questions ***(Laurin et al., 2013***), rate of afforestation/deforestation ***(Hirose et al., 2016)*** and global environment changes (***Trumbure et al.,2015).*** All these kinds of application require very fine mapping and monitoring of forest type. Very few studies using multi spectral sensors have evaluated the red edge band of forest classification). As this study tells about the use of sentinel-2 for the differentiation of phenological patterns according to forest categories and types. Accurate temporal resolution data is required for discriminating all the major forest types, sentinel-2 to give all the accurate information as it required. This study proves that sentinel-2 data is suitable for routine, medium to large scale mapping and monitoring of forest changes due to the combination of high spatial resolution and quick revisit time.

CHAPTER 4

4.1 STUDY AREA

The study area comes under a southern dry zone of Agro-Climatic Zone of Karnataka. Virajpet is a town in Virajpet taluk in Kodagu district of karnataka state, India. It belongs to the Mysore division. Which lies between latitude 12.1950N and longitude 75.8040E. It contains area of about 8.26 sq.km. It has a population of about 17,246 as per 2011 census. Virajpet has a total length of forest boundary is 350 kms of which 72 kms constitute inter-state boundary between karnataka and Kerala along the southern and western side of Virajpet division.

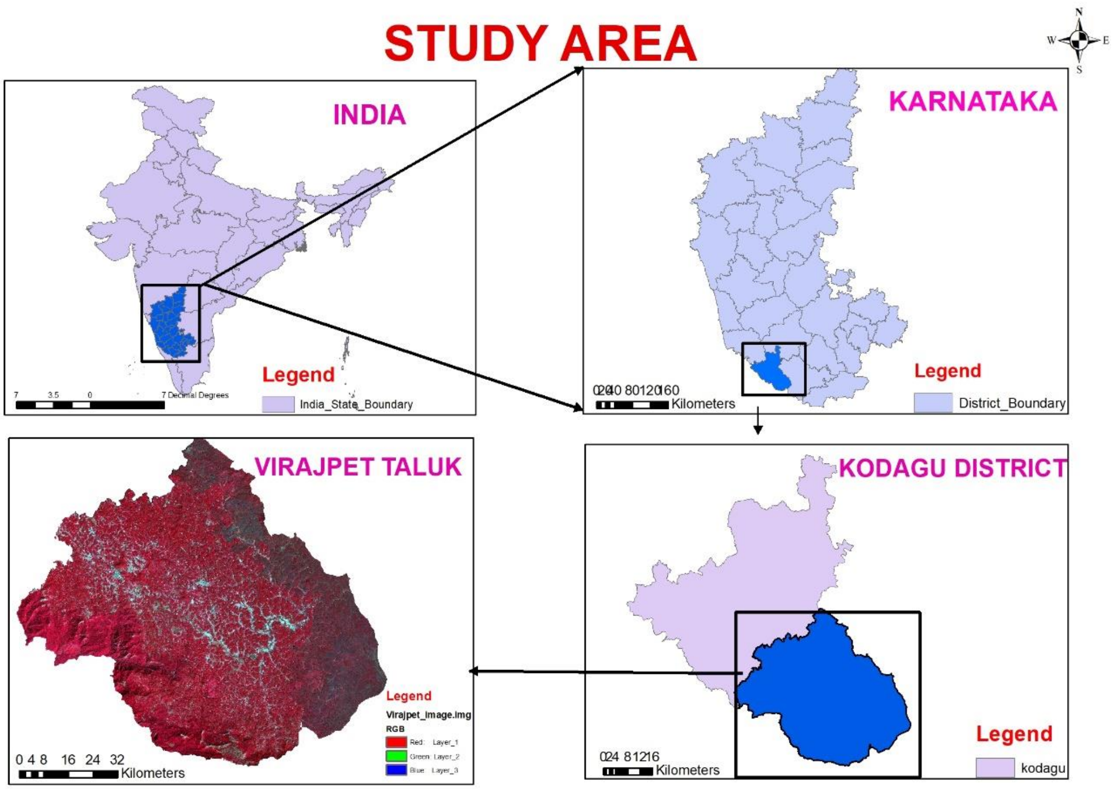


Fig 4.1: Study area

4.1.1 Topography

The tract has mountainous configuration with rivers, valleys, ravines, spurs etc. The tract becomes undulating and plains towards the east, while towards the west, it attains great heights with precipitous slopes. In northwest and southwest directions, it is intersected by a chain of hills, which further project out into innumerable ridges on either side, losing their heights gradually as they recede further but have almost everywhere sharp peaks. The general elevation of the tract varies from 100 m to 1745 m. The highest peak of Kodagu district, Tadiandmol, with an elevation of 1745 m. is situated in this tract. The area is drained primarily towards the east by a number of streams forming tributaries of important rivers like Cauvery and Lakshmanthirtha.

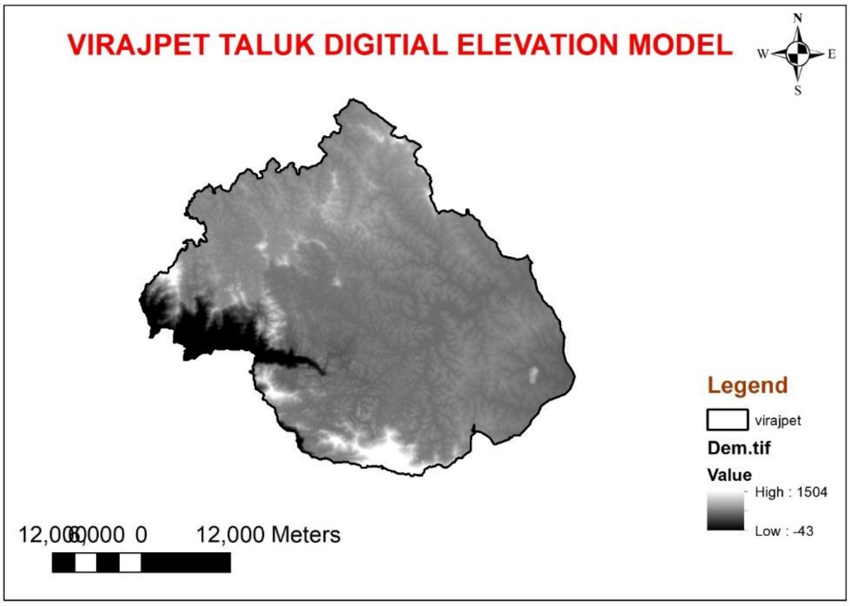


Fig 4.2: Virajpet taluk digital elevation model

4.1.2 Drainage

Major portion of the division drains eastwards into the Cauvery and Laxmanthirtha rivers which ultimately join the Bay of Bengal. As the area is situated in the western Ghats that receive heavy to very heavy rainfall, there are several rivulets and streams of seasonal - 164 - or perennial nature transecting the area and later joining the tributaries of the main rivers Bara pole is the only important river that flows towards the west and joins the Arabian Sea.

The following rivers drain the evergreen tract of the division

1) Matre Hole
2) Kerti Hole
3) Udumbe Hole
4) Pulike Hole
5) Padaimalai Hole and
6) Nadumalai Hole

The following rivers drain the eastern plains and deciduous tracts of this division

1) Cauvery river and its tributaries
2) Laxmanathirtha river and its tributaries.

Most of these rivers originate in the high mountains covered with lush green vegetation and have water form major of the year. But as they flow through highly rugged terrain, they are not suitable for navigation or for assured irrigation round the year. In Virajpet taluk, there are a few natural reservoirs which contain water throughout the year. These are locally called Kolly. Presence of several streams, rivers Kollies and tanks in the region besides the open wells and bore wells mitigates the water problem in the summer season.

4.1.3 Geology

There is wide variability of soil in depth and composition in different parts of the division. In the eastern part, especially in the Devamachi and Mavukal reserved forests, where the rainfall is relatively less and ground is gently undulating, leaching out effect is not there, and soils are clayey loam and deep. But in the western and southwest parts where rainfall is high soil is literate, latrritic bright red or yellow in colour. In some patches. they are loamy but having a sandy top layer. Areas covered with thick vegetation and multistoried forests have deep and fertile soil. In areas receiving heavy rainfall, there exists a delicate balance between the vegetation cover and quality of soil. In the absence of thick vegetal cover, soils are either washed away exposing the underlying rock or have become laterite.

4.1.4 Climate

The climate of the region may be broadly termed as tropical except at the higher reaches of the hills where it is montane sub-tropical. Heavy to very heavy rainfall, high humidity and cold nights characterize it. It may be divided into four seasons The period of June to September is the monsoon period: October and November constitute a post monsoon period; the period between December to February is characterized by clear bright weather with cold night followed by summer season between March and May.

4.1.4.1 Rainfall

There is wide variation of rainfall in the western and eastern parts. It decreases from west to east in the western Part, Makut and Mundrote receive rainfall of about 5290 mm while in the eastern part, Thithimathi receives about 1000 mm rainfall. Eighty percent of the rainfall is received during the period of southwest monsoon. However, in the post monsoon period of October –November, it also gets rainfall by northeast monsoon or due to depression/cyclones formed in the Arabian Sea which moves towards east and crosses over region. Some showers are common during April-May, which are locally called blossom showers.

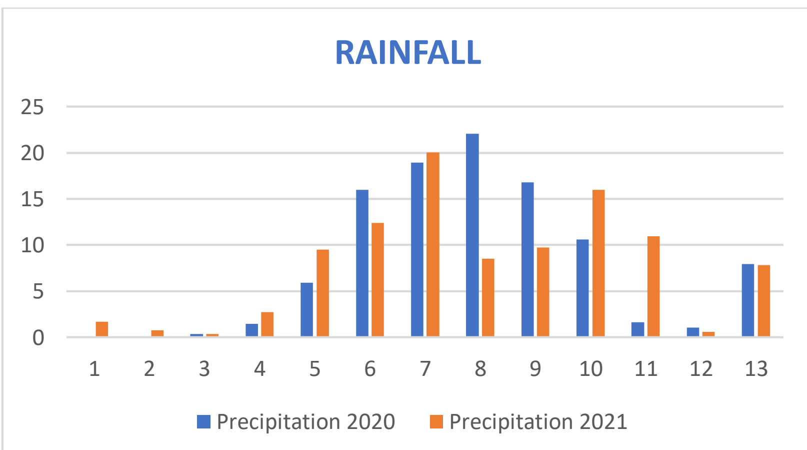


Fig 4.3: Graphical representation of Rainfall (2020-21)

4.1.4.2 Temperature

Temperature also varies considerably from top of the hills to the foothills. During the cold season the temperature drops up to 10.3’C at the upper portion of hill. During the summer season, the temperature at the foothills and especially in the eastern part rises up to 35.0’C. But the Mean maximum and mean minimum temperature is 28.6’C and 12.0’C respectively. In general air remains quite humid all through the year, particularly in rainy season. Fog is also quite common in the region except during the period from March to May.

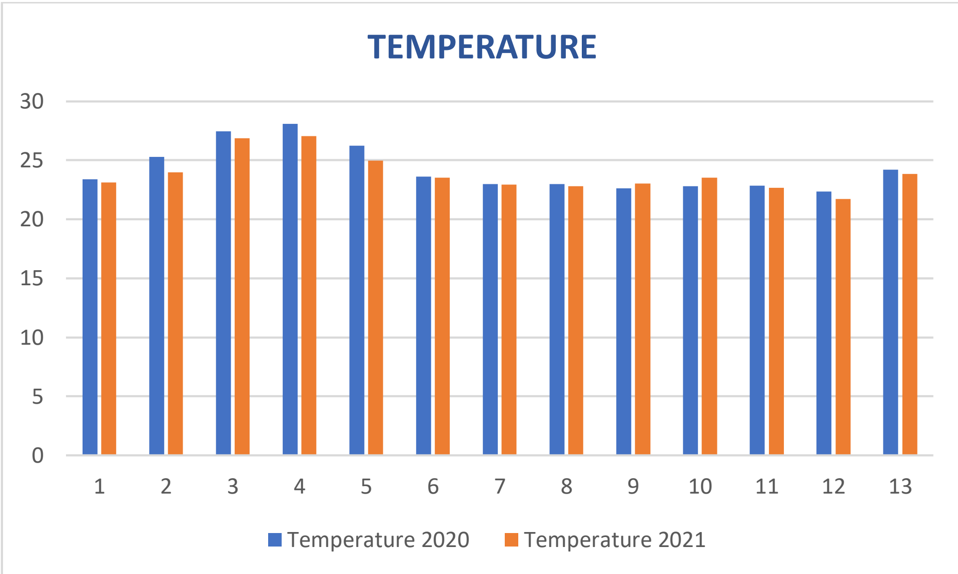


Fig 4.4: Graphical representation of Temperature (2020-21)

4.1.4.3 Winds

The winds are generally light to moderate and sometimes become severe during the southwest monsoon. Winds mainly blow between the southwest and northwest direction during the Southwest monsoon season. In the post monsoon season. winds are mainly north-easterly or easterly. But on some days, north-westerly winds blow in the afternoons. During the rest of the years, winds are from the directions between north and east in the mornings and between southwest and northwest in the afternoons.

4.2 Demography of division

Kodavas are the major population of this area. Gowdas, Muslims, Christians, Jains, Jenukurubas, Kurubas and Yeravas are living together.

Virajpet Division Consists of 120 Villages of Virajpet Taluk and 13 Villages of Madikeri taluk.

4.3 Major Crops

Major Crop is Coffee and Paddy. Inter Crops like Pepper, Areca Nut, Coconut, Palm are also grown

CHAPTER-5

SATELLITE AND SOFTWARE USED

5.1 DATA USED

SENTINEL-2 data were used to forest type classification of southern dry zone. The sentinel-2A satellite was launched on 2015, as part of the European Copernicus program. Sentinel-2 consists of an innovative wide-swath, high-resolution, multispectral imager (MSI). It has a 150mm aperture and a three-mirror anastigmatic design with a focal length of about 600 mm. It consists of 13 bands. Sentinel-2 provides images with spatial resolution varying from 10 m to 60m, and has three bands in the red edge and two bands in the SWIR region. It provides information about a wide range of land and coastal applications. The sentinel-2 data is used in the study area period of February 5, 2022.

The 3,4 and 8 bands of sentinel-2 have the spatial resolution of 10m.The data available for download is in Geo TIFF format. They contain 16-bit data values. The data is layer stack to produce False Colour Composition images. Histogram Equalization of data is done for better visual interpretation.

Sentinel2A-launch	June 2015, by Vega Kourou, Fresh Guiana
Sentinel2B-launch	July 2016, by Rockton from Plesetsk, Russia
Orbit	Sun -Synchronous at altitude 786Km, Mean Local solar time at descending node:10:30
Geometric revisit time	Five days from two-satellite constellation
Design life	Seven years
MSI	MSI covering 13 spectral bands(443-2190nm), with a swath width of 290 km and spatial resolution of 10m ,20m and 60m.
Receiving station	MSI data: Transmitted Via X-band to core sentinel ground stations and Via laser link through EDRS.
Main application	Forest, agriculture, land-use change. Mapping biophysical variables such as leaf chlorophyll, leaf water content, leaf area index.
Mission	Managed, developed, operated and exploited by various ESA establishments.

Table 5.2: Sentinel – 2B satellite

Band1-Coastal aerosol	0.443	60	20
Band2-Blue	0.490	10	65
Band3-Green	0.560	10	35
Band4- Red	0.665	10	30
Band5-Vegetation Red Edge	0.705	20	15
Band6- Vegetation Red Edge	0.740	20	15
Band7- Vegetation Red Edge	0.783	20	20
Band8- NIR	0.842	10	115
Band8A-Narrow NIR	0.865	20	20
Band9- water Vapour	0.945	60	20
Band10- SWIR Cirrus	1.375	60	20
Band11- SWIR	1.610	20	90
Band12- SWIR	2.190	20	180

5.2SOFTWARE USED

The software packages used during the study are listed below:

a) ARCGIS 10.
b) ERDAS 2020

5.2.1 ARC GIS 10.5

ARCGIS is a platform for organization to create, manage share and analyze spatial data. It consists of server components, mobile and desktop application and developer tools. This platform can be deployed on-premises or in the cloud with ARCGIS Enterprise, or used via ARCGIS online which is hosted and managed by Esri.

COMPONENTS OF ARCGIS:

Table 5.3: ARCGIS components

COMPONENTS	USAGE
Arc Map	Main Component-edit, analyse, view geospatial data
Arc Catalog	Data management application
Arc Toolbox	Contains geoprocessing, data conversion and analysis tools
Arc Globe	3D visualization application
Arc Scene	View GIS data in 3D

5.2.2 ERDAS IMAGINE:2020

ERDAS Imagine is an image processing software package that allows users to process both geospatial and other imagery as well as vector data. It use to collect, process, analyze and understand raw geospatial data, and ultimately deliver usable information. ERDAS can also handle hyperspectral imagery and LIDAR from various sensor. ERDAS is integrated within other GIS and remote sensing application and the storage format for the image can be read in many other application (.img files). ERDAS IMAGINE provides true value, consolidating remote sensing, photogrammetry, LiDAR analysis, basic vector analysis, and radar processing into a single product.

- User-friendly ribbon interface
- Spatial modelling with raster, vector and point cloud operators, as well as real-time results preview.
- A variety of change detection tools
- ERDAS ER Mapper algorithm support
- Ability to convert more than 190 image formats into all major file formats, including GeoTIFF, NITF, CADRG, JPEG, JPEG2000, ECW, and MrSID.

CHAPTER-6

METHODOLOGY

The Methodology adopted to the vegetation type classification is described below

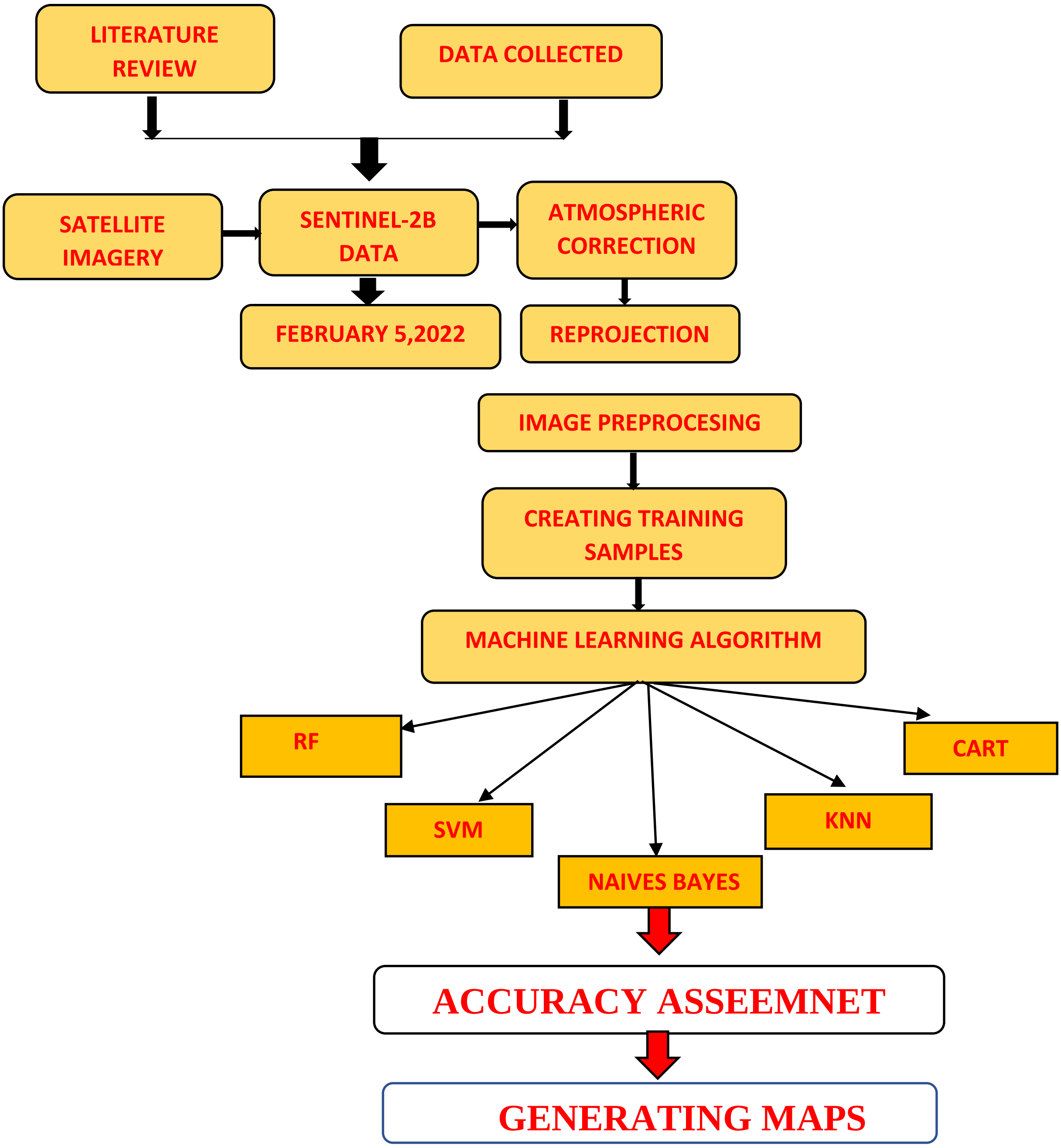


Fig 6.1 Flow chart showing methodology

6.1 Procedures of the methodology

A study is carried by the review of literature through that data is collected. Three satellite images were collected, February 5, 2022 imagery data were used for classification and model running. Other two were used as reference.

6.1.1 Satellite data atmospheric correction

Atmospheric correction is the most important part of the pre-processing of satellite remotely sensed data and any omission produces erroneous results. which is done using snap software.

6.1.2 Reprojection

Orthorectified dataset would be reprojected to the co-ordinate system of the Universe Transverse Mercator (UTM), WGS Datum 1984.

6.1.3 Image pre-processing

The satellite imagery was pre-processed by conducting a mosaic, firstly the area is covered is covered into four tiles, which is made into single imagery by mosaic process. This is done using ERDAS software. Then the study area is turned to false colour composite which is done to easy image interpretation and the image was clipped to the study area.

6.1.4 Creating training samples

A set of polygons was created to collect training sample data by using ARCGIS 10.5 toolbox. The training data were manually sampled in composited RGB imagery of sentinel-2. To enhance the accuracy of training samples, others dated data imagery was used to reference. Different sizes of polygons may differ in the number of pixels per land cover class. A total of 1000 features (polygon) of vegetation type with LULC were annoted on sentinel-2 imagery. Resulting in 1000 polygons with different sizes. Details regarding these samples and their spatial distribution are present in above table.

The vegetation type with LU/LC classes of selected area is divided into 15 classes. Study area Virajpet consists of different ecological diversity which includes evergreen forest, semi-evergreen forest, middle part of the Virajpet is entirely consists of coffee plantation, some part of area is also dominated by the tea plantation and deciduous forest is covering half of the Virajpet. Land cover were classified into five classes and land used is classified into seven class for cropland, fallow land, open-land, tea plantation, coffee plantation, settlements and others. Water area includes rivers, ponds which at image acquired data was partly shallow.

6.2 Machine learning algorithms

Spatial modeller allows to build a flow chart that contain operators strung together logically to produce output results. The collected labelled training samples was used for classification of vegetation type and used to run the spatial modeller in ERDAS software. Spatial modeller is extremely effective way to produce and analyse raster data. To each algorithm, raster input and vector input are given. Feature input are given as the training samples(polygon) and raster input is given as the satellite imagery. To the raster statistics mean is calculated and id value is given for the selected attribute. To each output layer, folder is connected and output are collected.

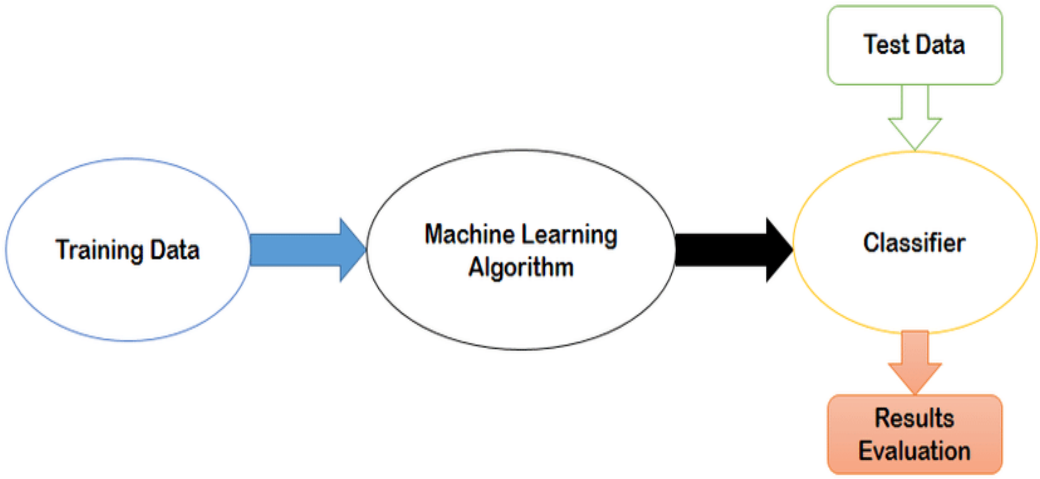


Fig 6.2: Flow chart of machine learning procedure

6.3 RANDOM FOREST (RF)

Random Forest (RF) is a classification and regression tree technique invented by Breiman. A RF randomly and iteratively samples the data and variables to generate a large group, or forest, of classification and regression trees. The classification output from RF represents the statistical mode of many decision trees achieving a more robust model than a single classification tree produced by a single model run. Random forest randomly selects observations, builds a decision tree and the average result is taken. It doesn't use any set of formulas

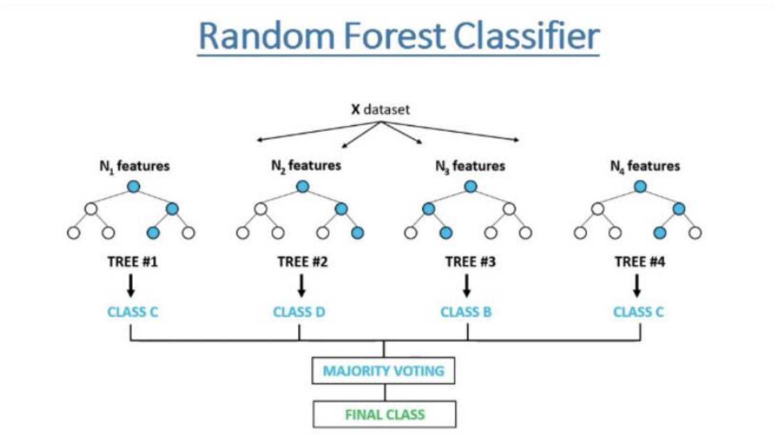


Fig 6.3: Random Forest classification

6.4 SUPPORT VECTOR MACHINE (SVM)

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning. It is a classification and regression technique based on statistical learning theory that has been proved very effective in solving complex classification problems in many different application domains. The success of SVMs is due to the important properties of this approach, which integrated with the effectiveness of the classification procedure and the elegance of the theoretical developments, result in a very solid classification methodology in many different remote sensing data analyses domains.

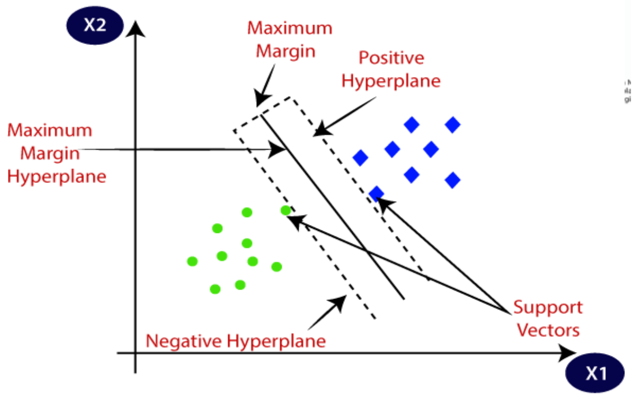


Fig 6.4: Support vector machine classification

6.5 K- NEAREST NEIGHBOUR (KNN)

K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique. K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems. KNN is a lazy learning algorithm because it does not have a specialized training phase and uses all the data for training while classification.

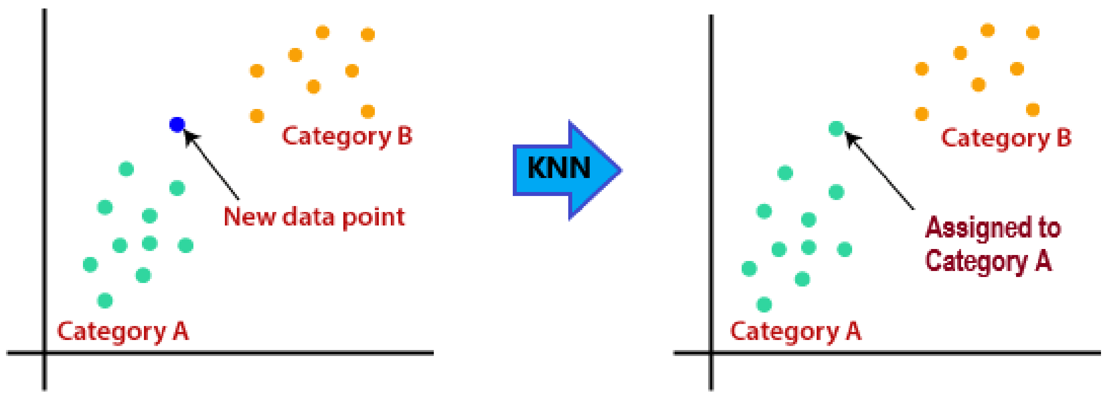


Fig 6.5: K-nearest neighbour Classification

6.6 NAIVES-BAYES

Naïve Bayes algorithm is supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is mainly used in text classification that includes a high-dimensional training dataset. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e., every pair of features being classified is independent of each other.

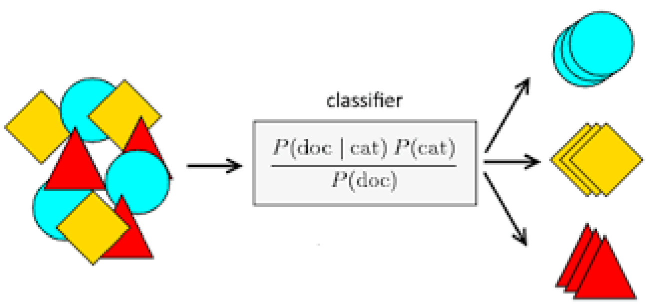


Fig 6.6: Naives Bayes classification

6.7 CART

The CART algorithm is a type of classification algorithm that is required to build a decision tree on the basis of Gini's impurity index. It is a basic machine learning algorithm and provides a wide variety of use cases.

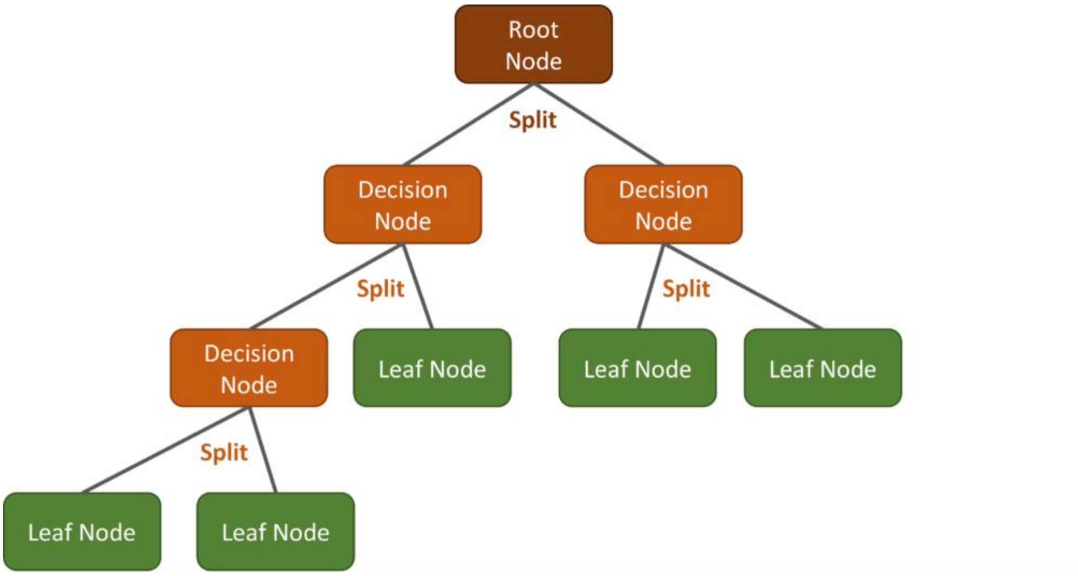


Fig 6.7: CART classification

CHAPTER-7

RESULTS AND DISCUSSION

The present study was envisaged for classification of vegetation types using remote sensing techniques. The classification was carried out using ERDAS spatial modeler feature. In spatial modeler, the sequence of operators that were used to initialize, execute the classification are discussed below:

7.1 RANDOM FOREST:

The main building blocks of spatial modeler is an operator. An operator is a self-describing object that performs a task, usually done on some that. It contains a input or output points that are called ports. These ports allow data to flow through an operator.

Feature Input: One of the most important tasks in building model is to collecting training data. In feature input data, its feature schema contains attribute fields whose name match with non-geometry attribute fields of training data used for training the machine intellect. Manually digitized rectangular boxes (polygon) in the image are taken as feature input which is in .shp format.

Raster input: Raster input is taken as satellite image (FCC) which is in image imagine format.

Raster statistics fer feature: In raster statistics per feature, the operators compute with statistics on pixel of input raster that are inside a feature. The computed statistical value is added as attribute to the feature stream. By default, mean is computed. The constructed attribute name matches with the name of existing field and the existing field is overwritten.

Select Attribute: In the select attribute, the given id, class, or value is used to initialize random forest. The select attribute operator is used to tailor the training data's attribute scheme before the training data is used in the initialize random forest operator.

Initialize random forest: In the initialize random forest it selects attribute operator is used to tailor the training data's attribute scheme. The operator defines and trains a random forest classifier which is used as an input or classifying data using classifying machine learning operator.

Machine intellect output: It saves machine intellect data to file, which is saves in. miz format.

Machine learning input: The saved output file is again taken as machine intellect input that is used in classifying using machine learning operator.

Classifying using machine learning: This operator performs classification on input data using trained classifier specified on the machine intellect port.

Raster Output: Saves raster into a file, by default properties of a output are determined from raster in. It saves the output file with .img format, in a given folder with a specified name.

Random forest is an ensemble of decision tree. In random forest it generates multiple trees with different solution of training samples. Each tree in the forest gives a vote and an object is assigned/ classified to the class that has most votes.

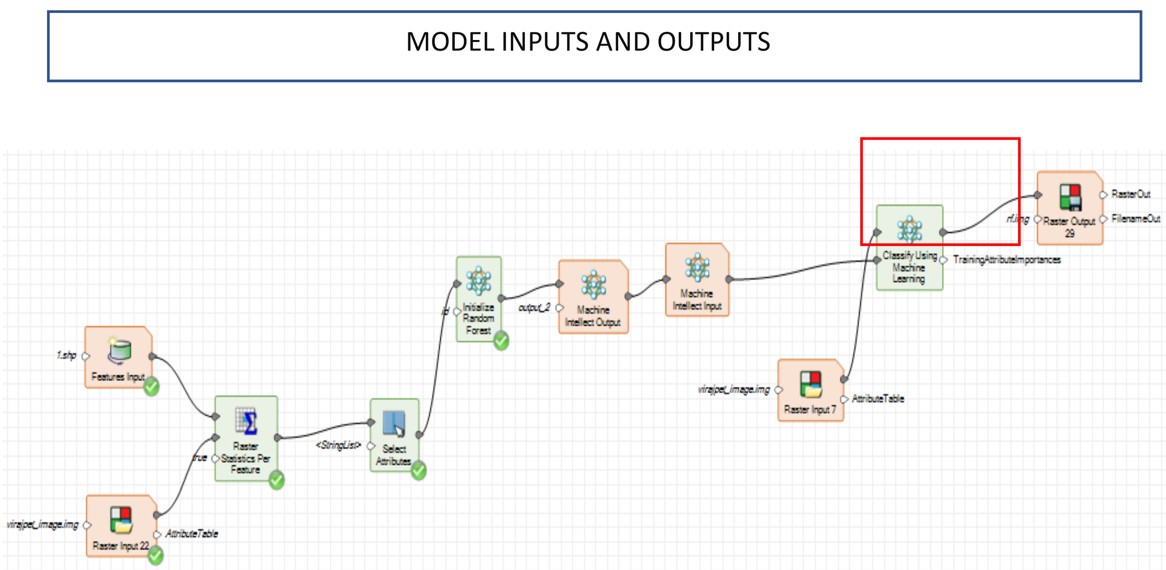


Fig 7.1: Random Forest machine learning inputs and output

7.2 SUPPORT VECTOR MACHINE:

Support Vector machine is a supervised machine learning algorithm that perform classification by finding optimal hyperplanes that separate the classes. The minimum distance between a hyper plane and a class is called margin. The optimal hyper plane is the one which has maximum margin. In feature input training samples are given as input data and raster input is the satellite imagery. In statistics per feature, it computes the mean. In the select attribute id is given for the operator to initialize SVM.

The operator defines and trains a support vector machine classifier which is used as an input or classifying data using classifying machine learning operator. All non-geometry attribute field in training data feature data is used for training. The select attribute operator is used to tailor the training data's attribute schema before the training data is used in the initialize SVM. Machine intellect output it saves machine intellect data to file. In classifying using machine learning operator performs classification on input data using trained classifier specified on the machine intellect port. Saves raster into a file, by default properties of a output are determined from raster in. It saves the output file with .img format, in a given folder with a specified name

MODEL INPUTS AND OUTPUTS

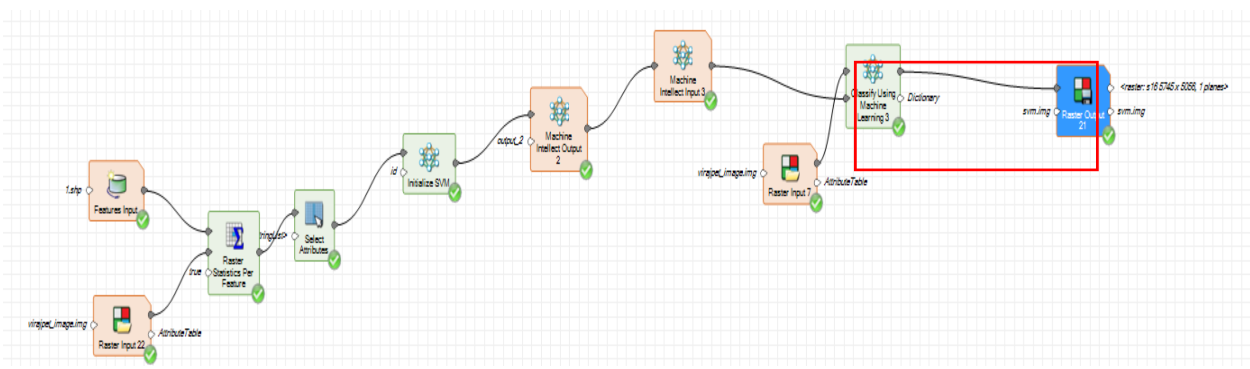


Fig 7.2: SVM machine learning inputs and output

7.3 K- NEAREST NEIGHBOUR

A K-nearest neighbor classifier is a type of non-parametric, instance-based learning classifier. It performs classification by weighted majority votes of its KNN. If all neighbor's have equal weights, the data is assigned to class that is most common among its KNN.

If K=1, the classifier becomes a nearest neighbor classifier, provides basic information about machine intellect.

This operator defines and trains a K-nearest neighbor classifier which is used as an input or classifying data using classifying machine learning operator. All non-geometry attribute field in training data feature data is used for training. The select attribute operator is used to tailor the training data's attribute schema before the training data is used in the initialize KNN.

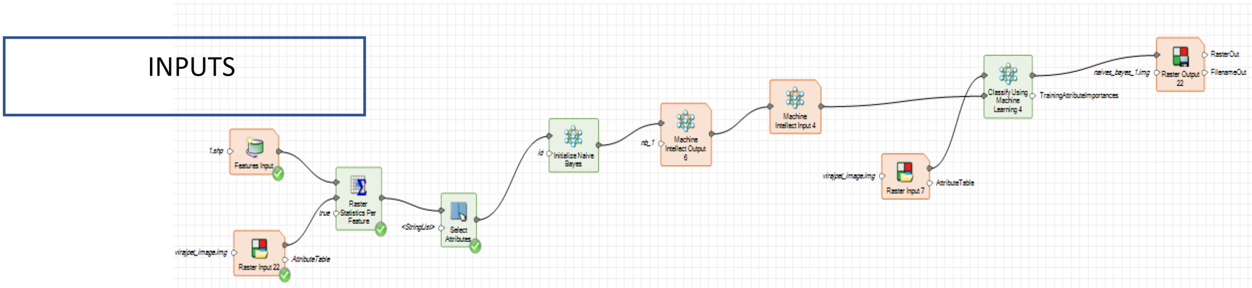


Fig 7.3: K-NN machine learning inputs

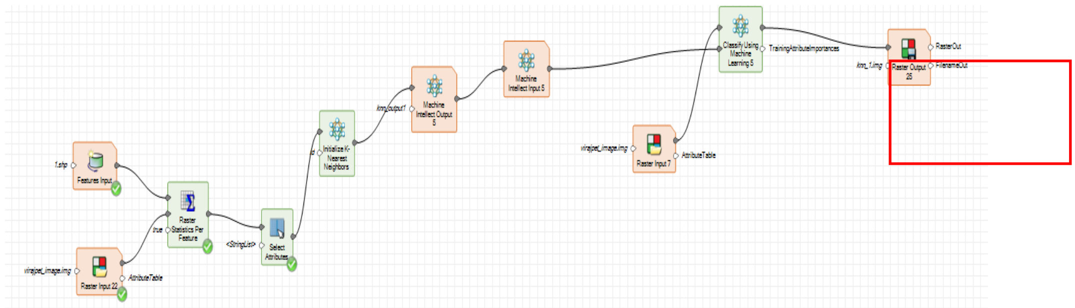
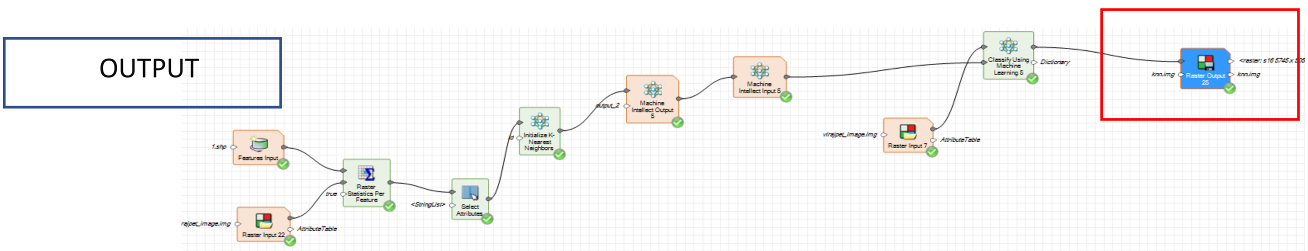


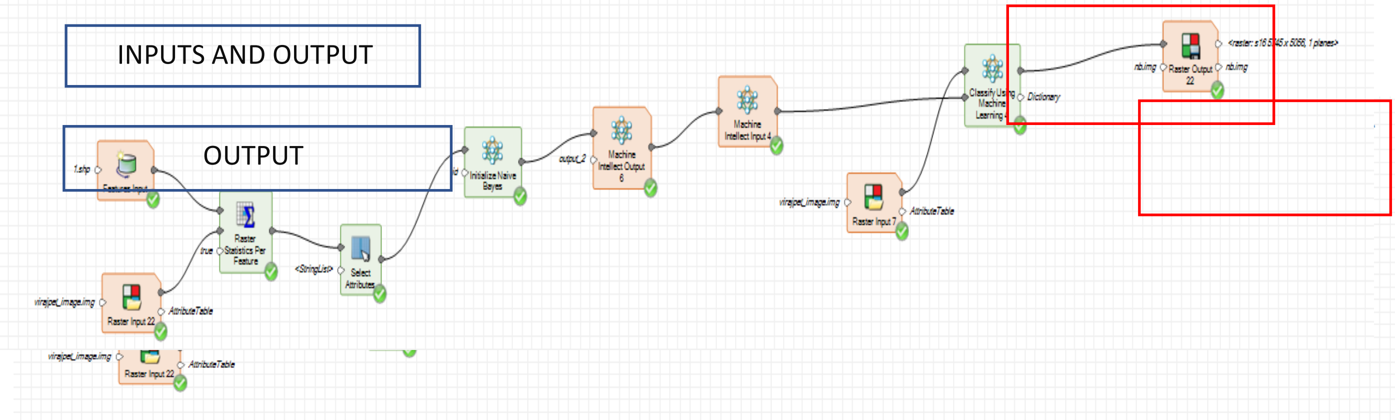
Fig 7.4: K-NN machine learning output



7.4 NAIVES BAYES

Naive bayes classifiers are part of family of simple probabilistic classifiers based on bayes theorem. The naive part if naive bayes comes from strong but somewhat naives assumption of independence between different attributes or parameters used for classification.

Naive bayes classification are based on supervised machine learning algorithm thtat needs to be trained to perform classification. This operator defines and trains a naïve bayes classifier which is used as an input or classifying data using classifying machine learning operator. All non-geometry attribute field in training data feature data is used for training. The select attribute operator is used to tailor the training data's attribute schema before the training data is used in the initialize naives bayes.



Fig

7.5: Naives Bayes machine learning inputs and outputs

7.5 CORRELATION AND REGRESSION

CART operator defines and trains a cart classifier which is used as an input or classifying data using classifying machine learning operator. All non-geometry attribute field in training data feature data is used for training. The select attribute operator is used to tailor the training data's attribute schema before the training data is used in the initialize cart.

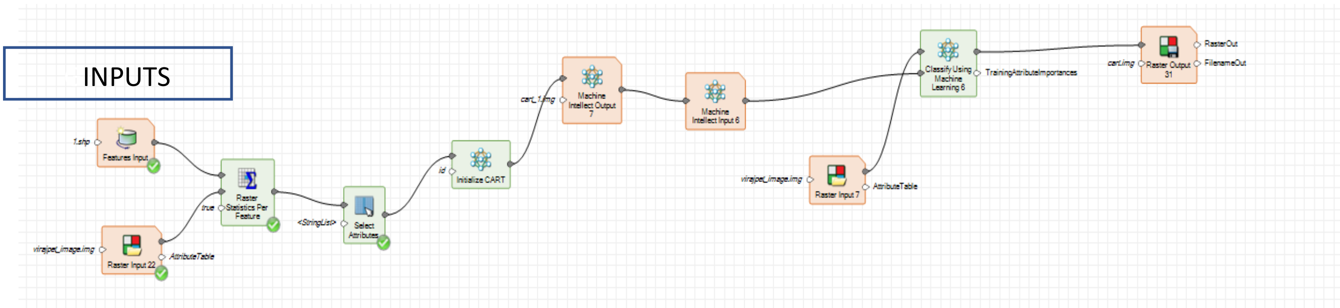


Fig 7.6: CART machine learning input

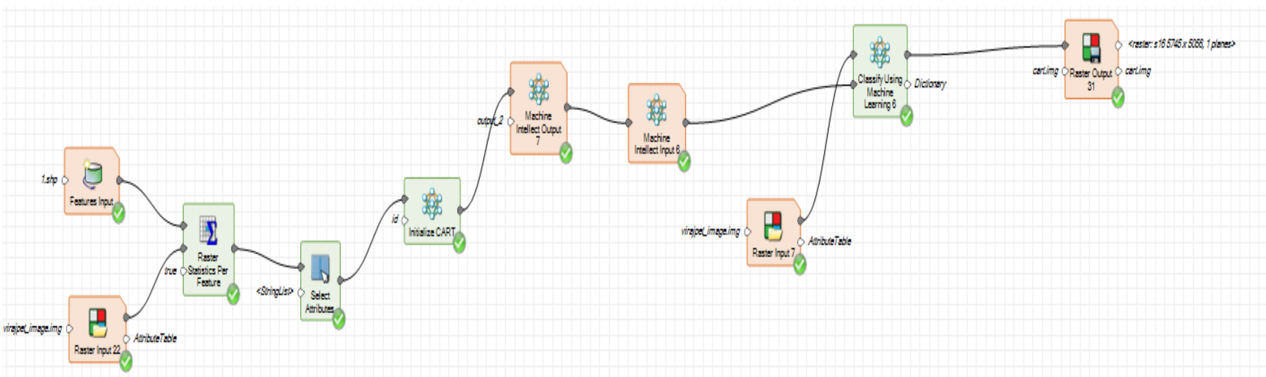


Fig 7.7: CART Machine learning output

7.6 ACCURACY ASSESMENT

Accuracy assessment is an important and essential step in the classification process. It compares the classified image to ground truth data. The basic principle for accuracy assessment is to compare estimate with reality, and to quantify the difference between the two. In the context of remote sensing-based classification, the classification, the 'estimate', are the classes mapped for each pixel and 'reality' is the actual area corresponding to each pixel. In our study, accuracy assessment is carried output in ERDAS software. Accuracy assessment is done to each machine learning algorithm classified image outputs. A major difference can be seen in each algorithm due to its sample size and quality of training data had a large impact on classification accuracy.

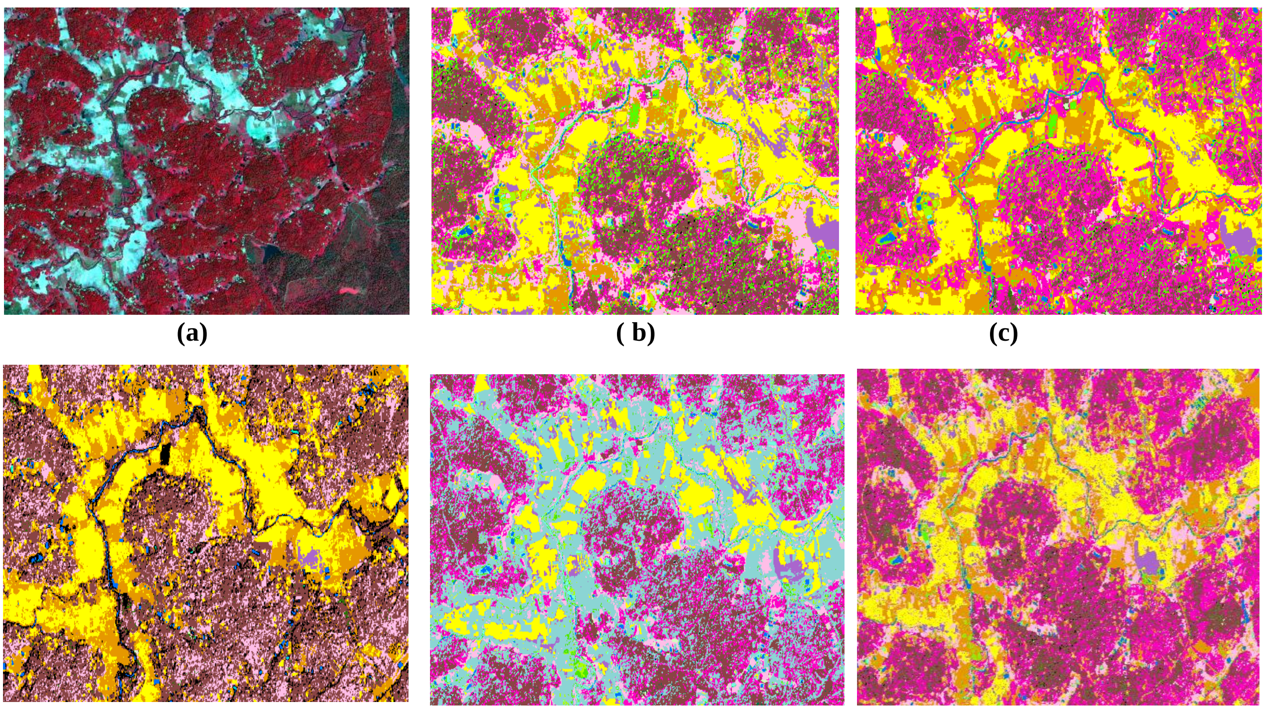
Accuracy assessment is done by using on screen digitized features. The created feature area called training samples. The selection of the training sites was based on those are clearly identified in all sources of images. In the study, thousands labelled training samples was used. The Accuracy Assessment Cell Array Reference column was filled according to the best guess of each reference point.

TABLE 7.1: CLASSIFICATION ACCURACY ASSESSMENT

ALGORITHM	OVERALL ACCURACY	AVERAGE ACCURACY	KAPPA CO-EFFICIENT
RANDOM FOREST	85.5%	81.60%	0.81
SVM	45.16%	40.26%	0.51
K-NN	70.11%	70.1%	0.70
NAÏVES BAYES	43.12%	41.11%	0.41
CART	65.26%	62.12%	0.62

Although the five evaluated machine learning algorithm are assumed with different performance visually, the quantitative analysis has demonstrated that best learners in terms of kappa value is random forest. Random forest comes with a overall accuracy of 85.5% and with a average accuracy of 81.16 with a kappa value of 0.8%.

RESULTS OF IMAGE CLASSIFICATION OF EACH ALGORITHM



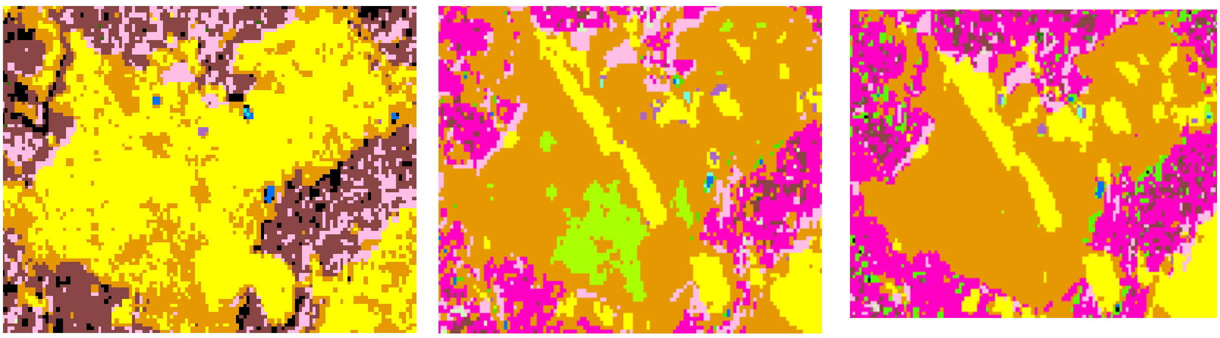
(a) (b) (c) (d) (e) (f)

Fig 7.8: Results of the image classifications of Fallow land with each algorithm

(a) FCC image (b) RF (c) KNN (d) SVM (e) NB (f) CART

Image (a) is a false colour composite with a band combination of 11,8,4 this data is compared with all other classified images, (b) shows the random forest classifier with a good accuracy classified image , fallow land and open land are correctly classified , water bodies are exactly classified as same compare with ground data , (c)In K-nearest classified image some fallow lands are merged with open land and coffee plantation is merged with gregarious forest , cropland is merged with coffee plantation. (d) in support vector machine some of the fallow-land are considered as the shadow, coffee planted area is correctly classified (e) In naives bayes large part of the fallow-land is merged with another feature , fallow-land are considered as the settlement, (f) CART and RF results came with a same accuracy level, fallow-land is correctly classified, and some of the crop lands and coffee plantation are also correctly classification.

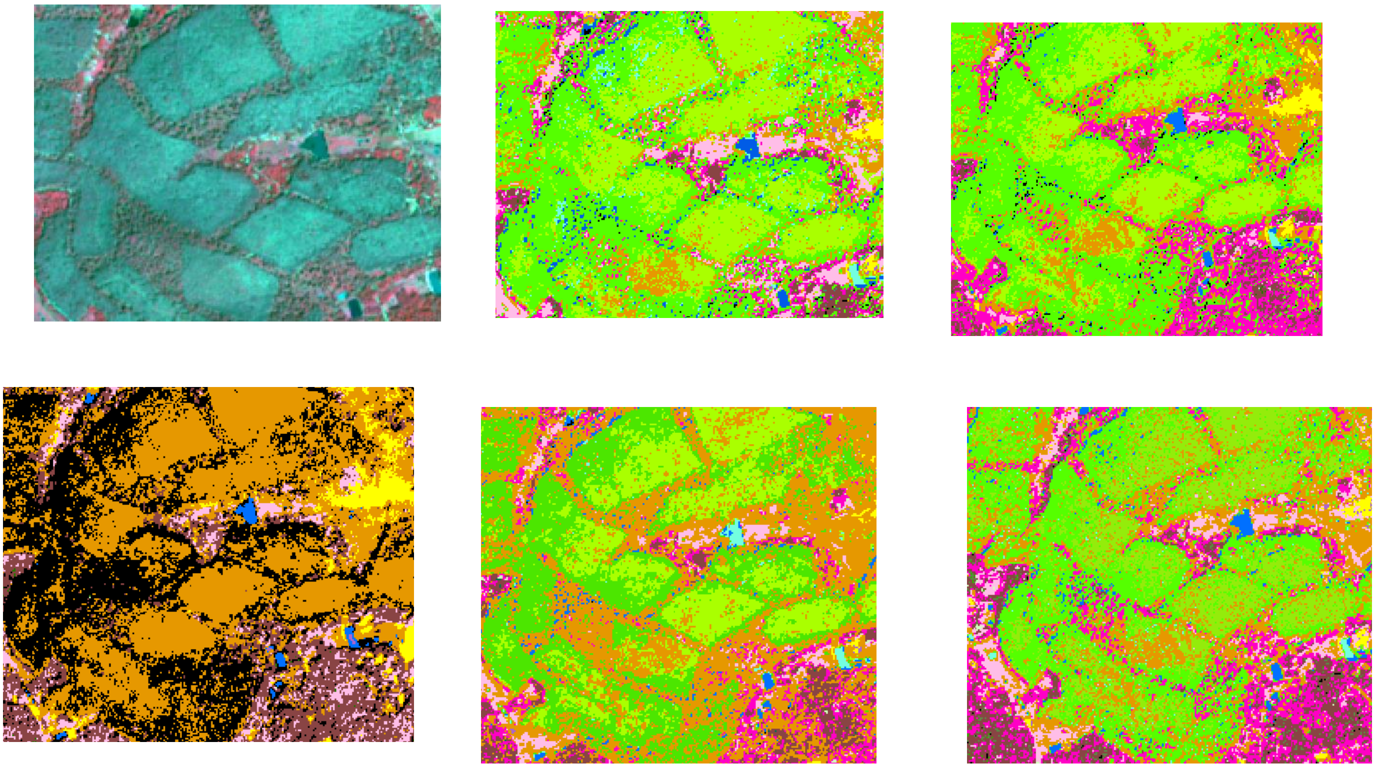
(a) **(b)** **(c)**



(d) **(e)** **(f)**

Fig 7.9: Results of the image classification of open land with each algorithm

Image (a) is showing the original image with a features containing large portion of open land and some part with fallow land and covering some part of coffee plantation (b) random forest classified with a good results open land and fallow land are correctly classified when compared with ground data, some open area are classified as cropland , waterbody are classified correctly, (c) in k- nearest neighbor some of the cropland is merged with gregarious forest , fallow land and open land are correctly classified, (d) in support vector machine all the open land is considered as the fallow land and all the fallow land is considered as the open land, large part of the area is wrongly classified (e) in naives bayes open land is merged with deciduous forest and some of the fallow land is merged with open land, (f) in CART classification some part of the fallow land are considered as the settlement because of the pixel colour and coffee planted area is considered as the cropland due to which accuracy level decreases.



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Fig 7.10 Results of image classification of forest area with each algorithm

The image showing the deciduous forest, the fcc of this image is shown in (a) an also it is covered with cropland an d waterbody. (b) in random forest, deciduous forest area is classified with accuracy when compared with ground data and water body are correctly classified , crop land perfectly classified when compared with ground data, (c) in KNN some part of cropland is classified as gregarious forest, open land correctly classified with much accuracy, (d) in support vector machine entire deciduous forest is considered as the open land and the edge of the forest area is considered as the shadow , water bodies are classified correctly and in (e) naives bayes classification, some part of the forest area is merged as the open land and cropland is taken as the open land, (f) it is same as the naives bayes classified image , deciduous forest is merged with open land and fallow land is merged with cropland.

(a) **(b)** **(c)**

(d) **(e)** **(f)**

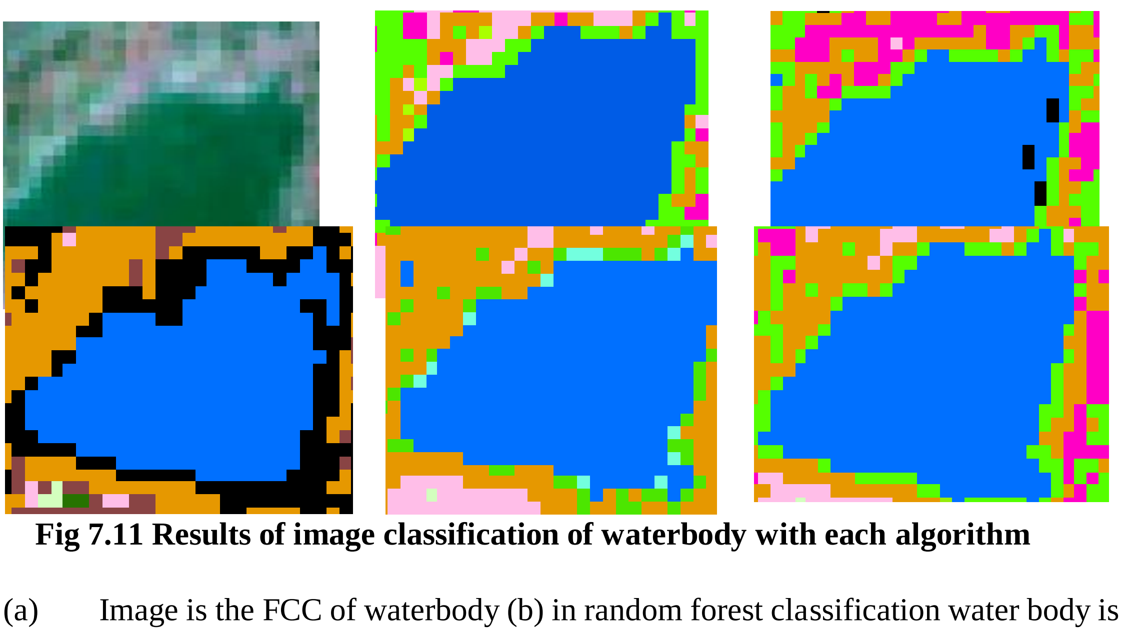


Fig 7.11 Results of image classification of waterbody with each algorithm

(a) Image is the FCC of waterbody (b) in random forest classification water body is correctly classified and in (c) K- nearest neighbor some of the shadows are merged with waterbody due to which accuracy level decreases,

(d) In support vector machine boundary of the waterbody is taken as the shadow, (e) in naives bayes and (f) CART classified image with good accuracy.

CONCLUSION 8

The study concludes that ML techniques are an efficient approach to map different land use and land cover classes, including different vegetation types and demonstrated that some algorithms perform this task with higher accuracy than the others. Here, evaluated that the performance of multiple machine learning models for mapping vegetation type with LULC classification in Virajpet taluk using multispectral images collected by an orbital sensor, embedded in the Sentinel-2 platform. In our investigation, the RF algorithm achieved satisfactory results when compared with others and characterized with more accurate rate and less error rate.

The results came with a overall accuracy rate is 85.5%, and average accuracy of 81.60% with a kappa co-efficient 0.81. Studies framework is appropriate to accurately map vegetation-types with land use and land cover class in Virajpet taluk and future research may benefit from the information presented here.

Acknowledgment

In this research, the data and software are used in this project was provided by *RRSC-South, Dept of Space, ISRO Bengaluru - 560 037*

REFERENCE

1. Narayanan, U., Unnikrishnan, A., Paul, V., & Joseph, S., "A survey on various supervised classification algorithms,"2017 International Conference on Energy Communication, Data Analytics and Soft Computing (ICECDS), IEEE, pp.2118-2124, August 2017.
2. Caruana, R., & Niculescu-Mizil, A., "An empirical comparison of supervised learning algorithms," Proceedings of the 23rd international conference on Machine Learning, pp.161-168, June 2006.
3. Sébastien Rapinel and Laurence Hubert-Moy**, "**One-Class Classification of Natural Vegetation Using Remote Sensing".
4. Zelin Liu, Changhui Peng, Timothy Work, Jean-Noel Candau, Annie DesRochers and Daniel Kneeshaw Application of machine-learning methods in forest ecology: recent progress and future challenges.
5. Fabian Ewald Fassnacht, Hooman Latifi, Krzysztof Stereńczak, Aneta Modzelewska, Michael Lefsky, Lars T. Waser, Christoph Straub, Aniruddha Ghosh, "Review of studies on tree species classification from remotely sensed data".
6. Use of Sentinel-2 for forest classification in Mediterranean environments.
   (https://www.researchgate.net/publication/321213981)

7. Arvid Axelsson, Eva Lindberg a, Heather Reese b, Håkan Olsson, "Tree species classification using Sentinel-2 imagery and Bayesian inference".
8. Biodiversity characterisation at community level in India using Earth Observation data. (Department of space and Department of Biotechnology).
9. Dino Dobrini´, Mateo Gašparovi´c, Damir Medak, ´ "Sentinel-1 and 2 Time-Series for Vegetation Mapping Using Random Forest Classification: A Case Study of Northern Croatia".
10. Heikki Astola, Tuomas Hame, Laura Sirro, Matthieu Molinier, Jorma Kilpi,"Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region".
11. Bogdan Zagajewski , Marcin Kluczek, "Comparison of Random Forest, Support Vector Machines, and Neural Networks for Post-Disaster Forest Species Mapping of the Krkonoše/Karkonosze Transboundary Biosphere Reserve."
12. Jan Svoboda**,** Natalia Kobliuk, "Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data—A Case Study of Czechia."
13. "Land Cover and Forest Type Classification by Values of Vegetation Indices and Forest Structure of Tropical Lowland Forests in Central Vietnam".
14. Gang Jiang, Quanshun Zheng, "Remote Sensing Recognition and Classification of Forest Vegetation Based on Image Feature Depth Learning."
15 Sophia S. Rwanga, "Accuracy Assessment of Land Use/Land Cover Classification Using Remote Sensing and GIS."