# Chapter 1

# Introduction

In the face of climate change and ecological challenges, conserving and managing forest ecosystems is crucial. Satellite technology plays a key role in understanding and forecasting blossoming patterns in these ecosystems, with a focus on bees' sustainability. Bees are vital pollinators, and the timing of blossoms is critical for their survival. This research leverages high-resolution imagery from Sentinel-2 satellites, offering global coverage and frequent revisits, to study blossoming patterns worldwide.

Historically, universities and government agencies would have to keep track of the blooming season of farmland by manually surveying farmers. This process was not only time-consuming but also resulted in data that was often outdated. Moreover, it didn't provide bee farmers with timely information on when to introduce bees to their farms. The proposed system utilizes Sentinel-2 data, which includes satellite data, to identify farmlands and keep track of the crops growing in different regions. It also monitors the blooming patterns of these regions so that farmers can effectively introduce bees during the blooming season.

This is achieved through the use of the FAIIS algorithm for farmland identification and Convolutional Neural Networks (CNN) for crop type recognition. By combining these advanced technologies with satellite data, the system not only provides accurate and up-to-date information about crop growth and blooming patterns but also assists farmers in making informed decisions about when to introduce bees to their farms, optimizing pollination and crop yields.

The study extends to the biodiverse Indian subcontinent, where satellite data helps predict blossom timing, intensity, and flower characteristics. This knowledge is vital for understanding India's floral diversity, its ecological dynamics, and its impact on local ecosystems and agriculture.

## 1.1 Aims & Objective

The aim of this project is to enhance pollination by identifying and studying the blooming periods of various flowers in different regions across India. This endeavor considers multiple factors, including weather conditions, soil content, and spectral information obtained from satellite data, to determine suitable locations for planting bee hives. This holistic approach seeks to optimize pollination and support the crucial role of bees in agriculture and ecosystems.

**Objective 1: Identifying Farmlands and Obtaining Relevant Features of the Farmland.**

Farmland Detection: Use satellite imagery and image processing techniques to accurately identify and delineate farmlands. Exclude non-agricultural areas, such as human settlements and wastelands, to focus on relevant regions.

Data Integration: Collect and integrate data from various sources, including satellite images, spectral bands, and agricultural institutions. This data should encompass details about soil types, soil contents, and common fertilizers used in the farmlands.

Model Training: Develop machine learning models to learn and classify different features of farmlands. The model should utilize both spectral data and additional information to improve accuracy in identifying farmland types.

**Objective 2: Predicting Blooming Seasons for Different Plants.**

Data Analysis: Utilize the integrated data to understand the relationships between spectral data, soil types, fertilizers, and the blooming seasons of various plants.

Machine Learning Algorithms: Apply machine learning algorithms to predict the blooming seasons of different flowers. The choice of algorithms should depend on the nature of the dataset and the specific features that influence blooming.

Accuracy and Timing: Ensure that the predictions include each stage of the blooming season, from budding to full bloom. Timing is crucial for synchronizing bee installation and pollination efforts with flowering periods.

**Objective 3: Identifying Regions Suitable for Beehives.**

Conditions for Beehives: Define the conditions necessary for the successful placement of beehives, considering factors like altitude, temperature, humidity, and climate conditions.

Machine Learning-Based Suggestions: Use machine learning algorithms to suggest suitable regions for placing beehives. These algorithms should consider all the relevant factors mentioned above.

Time Frames: Provide suggested time frames for installing beehives in these regions to coincide with the blooming seasons.

## 1.2 Scope

The implementation of this project holds significant potential and encompasses a wide scope, with positive implications across various sectors. The predictive system developed has the capacity to efficiently monitor biological changes, which, in turn, can substantially enhance agricultural productivity in the targeted regions. This innovative approach stands to revolutionize traditional farming and beekeeping practices by introducing modern techniques for managing beehives and improving crop cultivation.

In the agricultural sector, the model's adoption promises substantial benefits. Farmers throughout India stand to experience a notable increase in profits and the production of high-quality agricultural products, which ultimately contributes to economic growth. The system's integration with various government apps for farmers creates a convenient one-stop solution, offering farmers access to a comprehensive range of resources and support.

Additionally, the project's implications extend to the monitoring of environmental and biological ecosystems. The system's capabilities enable the efficient tracking of changes in flowering patterns and the influential factors behind them. Bee farmers, in particular, stand to gain valuable insights from this information. Furthermore, the system enhances the monitoring of bee behavior and interactions with the environment, providing a valuable resource for agricultural scientists and government research institutes conducting vital research in these domains.

Furthermore, the project promotes sustainable beekeeping practices. Farmers and beekeepers can leverage the system's data on blooming seasons, enabling them to prepare and plan their activities well in advance. This proactive approach fosters the smooth and efficient management of beehives, further contributing to the sustainability and success of these practices.

# Chapter 2

# Review of Literature

The use of NDVI derived from Sentinel-2 data for farmland and crop identification offers a promising approach in agricultural and land management. NDVI, as a quantitative indicator of photo synthetically active biomass, provides valuable insights into agrotechnical and ameliorative aspects of farmland. Overlaying spectral data on precise farmland maps enhances monitoring accuracy, aiding in the assessment of crop health and conditions. Clustering techniques and classification algorithms such as Random Forest and Support Vector Machine are employed for crop identification using time series data. Additionally, time series optical remote sensing and RADARSAT-2 SAR data are effectively used to classify and monitor crops, offering a dynamic approach to crop monitoring and land cover classification.

## 2.1 Review of Existing Systems

### 2.1.1 Farmland Classification Techniques

The use of NDVI (Normalized Difference Vegetation Index) derived from Sentinel-2 data for farmland and crop type identification is a promising and widely adopted approach in agricultural and land management. This method leverages the capabilities of Sentinel-2 satellites and the unique insights offered by NDVI. The NDVI is a quantitative indicator of photo synthetically active biomass [1] . Interpreting changes in the spectral brightness of vegetation throughout the vegetation period and using the NDVI (Normalized Difference Vegetation Index) offers valuable insights into the agrotechnical and ameliorative status of farmland. The tone or color of the field image can serve as an indicator of these conditions. Additionally, combining this spectral data with current and precise farmland maps enhances the accuracy and objectivity of monitoring results [1]. This synergy streamlines monitoring tasks and significantly reduces associated costs. The information derived from changes in vegetation spectral characteristics and NDVI trends during the growing season can be invaluable for assessing the health and conditions of crops. Different colors or tones in the imagery can reveal variations in crop vigor, stress, or anomalies, providing clues about potential issues like nutrient deficiencies, water stress, or diseases[1]. By monitoring these changes, farmers and land managers can make more informed decisions regarding irrigation, fertilization, and other agrotechnical practices.

There have been numerous clustering techniques used to perform unsupervised classification on satellite image time series (SITS) data for different tasks. Different clustering methods differ mainly in clustering features and similarity measures. For example, the widely used unsupervised clustering algorithm K-means, whose clustering feature may be an original spectral–temporal vector, and whose similarity measured is Euclidean distance; another SITS clustering method, a novel variant of Dynamic Time Warping (DTW) named SC-DTW, introduced by Zhang et al. [2], whose clustering feature is a local shape context vector, but whose similarity measure is dynamic time warping (DTW) distance.

### 2.1.2 Crop Type Classification Techniques

Crop identification is carried out by Google Earth Engine (GEE) cloud platform to

extract time series Sentinel satellite radar and optical remote sensing images combined with simple noniterative clustering (SNIC) multiscale segmentation with random forest (RF) and support vector machine (SVM) classification algorithms to classify and identify major regional crops based on radar and spectral features.

Currently, crop classification and identification methods using time series images

using image-oriented are widely applied [10 - 15]. Time series optical remote sensing images obtained from the HJ-CCD satellite have been employed to develop a decision-tree-based classification model. This model utilizes the temporal variations in spectral vegetation indices to effectively classify and identify multiple crop plantings. The use of time series data and spectral indices offers a dynamic and comprehensive approach to crop monitoring and land cover classification [10]. Multitemporal RADARSAT-2 fully polarized Synthetic Aperture Radar (SAR) time series images have been harnessed to efficiently extract the phenological periods of rice. This is achieved by analyzing the time series curve variations in the polarization feature parameters of rice. This approach leverages SAR technology to monitor rice growth and development stages by capturing the changes in the radar signal's polarization characteristics over time [11]. Phan Thanh Noi et al. [15] conducted a study using Sentinel-2 image data to assess the performance of different classifiers, including Random Forest (RF), k-Nearest Neighbors (kNN), and Support Vector Machine (SVM), for land use cover classification. In this research, various multitemporal and multifeature classification methods were employed, primarily focusing on pixel-based techniques. These methods involved the extraction of temporal optical or microwave features from image elements to classify and recognize different land cover types, particularly for crop classification.

### 2.1.3 BeeKeeping Techniques

The land use and land cover of Sukabumi regency was collected. This land use and land cover was updated by using Landsat 8. The method used was visual image interpretation, based on human vision system to interpret pattern and colors in the image. Spontaneous recognition and logical inference are distinguished [29]. To enhance the image used a pan sharp, where each band was changed from 30m to 15m spatial resolution [26]. The composite band for Landsat 8 was used true color (Red: 4, Green: 3, Blue:2). Validation was done by ground truth and Google earth check. The ground truth was done by checking thirty locations, selected by purposive sampling.

## 2.2 Limitations of Existing Systems

Deep learning is extensively used in a broad range of domains and gains popularity in SITS land cover analysis [16–18]. It produces excellent results but is extremely reliant on the amount of available data, and more precisely labeled data. Remote sensing produces a large volume of data that is typically devoid of annotations. This situation is exacerbated by time series of satellite images. On the one hand, the land cover type of SITS data may change with time, especially for long time series [19], making the class label difficult to determine in some situations. Although there are some early successes of clustering algorithms in SITS data, there have been few attempts to cluster SITS data depending on deep features.

The above mentioned methods for crop classification use multitemporal and multifeature classification methods based on pixels often carry out crop classification and recognition by extracting the temporal optical (or microwave) features of image elements. They used several methods and tested them to determine the best. Although they achieve high classification accuracy, to a certain extent, they usually ignore the spatial correlation between adjacent image elements [20], which is prone to salt-and-pepper noise [21]. Salt-and-pepper noise exists in most pixel based classifications with high-resolution images. The essence of this phenomenon is the misclassification of pixels affected for various reasons.

The land use and land cover mask created for beekeeping was initially designed exclusively for the Sukabumi region, with a notably limited coverage area. To effectively encompass the entire Indian subcontinent, there is a crucial need for more extensive and comprehensive land use and land cover masks.

# Chapter 3

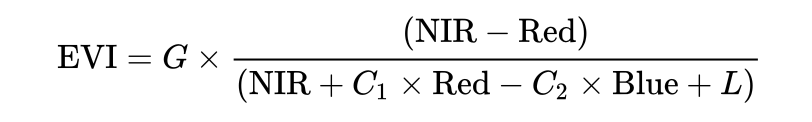
# Proposed System

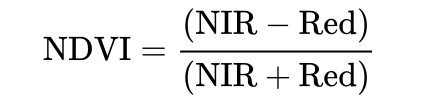
This proposed system offers a comprehensive and data-driven approach to beekeeping in farmlands, helping beekeepers identify the most advantageous locations for beehives using Remote sensing. By leveraging land classification, weather data, and crop characteristics, the system enhances the efficiency and success of beekeeping operations while supporting ecological balance and agricultural productivity.

## 3.1 Design DetailsScreenshot 2023-10-24 at 2.17.46 PM.png

Fig 3.1 Flow diagram for Agricultural Land Identification

In Fig 3.1 we can see the design of how agricultural land is classified based on the input from sentinel data. First the different bands are extracted from sentinel data , Bands useful to us are Blue, Green, Red, Red edge 1, Red edge 2, Red edge 3, NIR, NIR narrow, SWIR, SWIR [31]. Using these bands we calculate NDVI and bm\_diff.

Enhanced Vegetation Index (EVI/bm\_diff): Very similar to NDVI. The only difference is that it corrects atmospheric and canopy background noise, particularly in regions with high biomass[32].

The normalized difference vegetation index (NDVI) is a widely-used metric for quantifying the health and density of vegetation using sensor data.

Now a K-Means clustering integrated with Facebook AI Similarity Search(FAISS) is done to find out the clusters with a different label, representing the land class. Implementation with FAISS helps in efficient similarity search and clustering of dense vectors. Final clustering results are saved as a raster. They represent the pixel clusters of agricultural land.

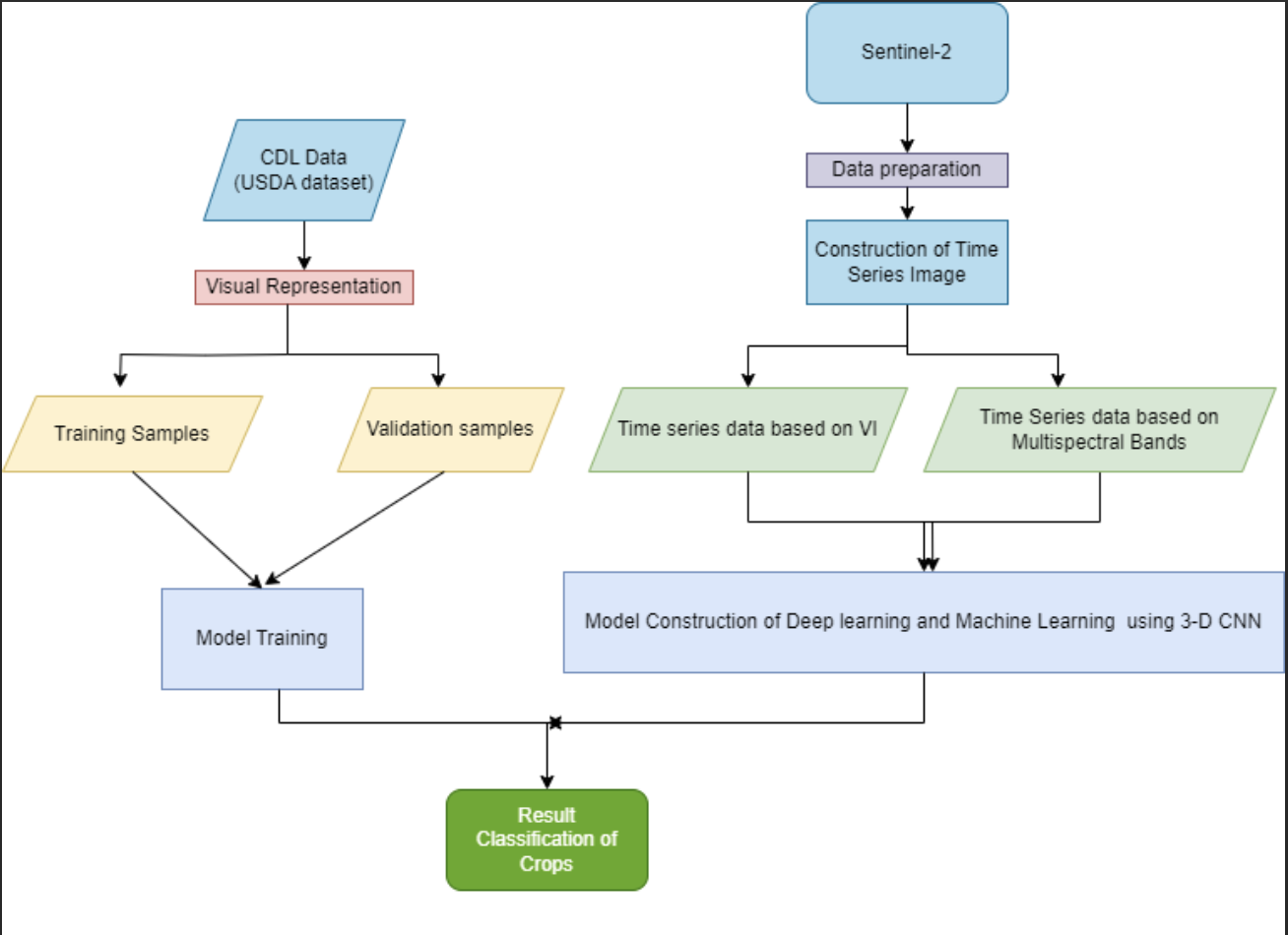


Fig 3.2 Crop Classification using 3D CNN

Fig 3.2 can be broken down into

* + - Data preparation: This step involves preparing the remote sensing data for training and evaluation. This includes preprocessing the data to remove noise and artifacts, and converting it to a format that is compatible with the 3D-CNN model.
    - Construction of time series image: This step involves creating a time series image for each crop type. This is done by stacking multiple images of the same crop from different time periods. This allows the 3D-CNN model to learn the temporal dynamics of crop growth.
    - Model construction: This step involves constructing the 3D-CNN model. This is done by stacking multiple convolutional layers, max-pooling layers, and activation functions. The number and size of the layers will depend on the specific dataset and the desired accuracy.
    - Model training: This step involves training the 3D-CNN model on the prepared data. This is done by feeding the model the time series images and the corresponding crop labels. The model will learn to extract the features from the images that are most discriminative for crop classification.
    - Result classification of crops: Once the model is trained, it can be used to classify new crop images. This is done by feeding the model the image and predicting the crop type.

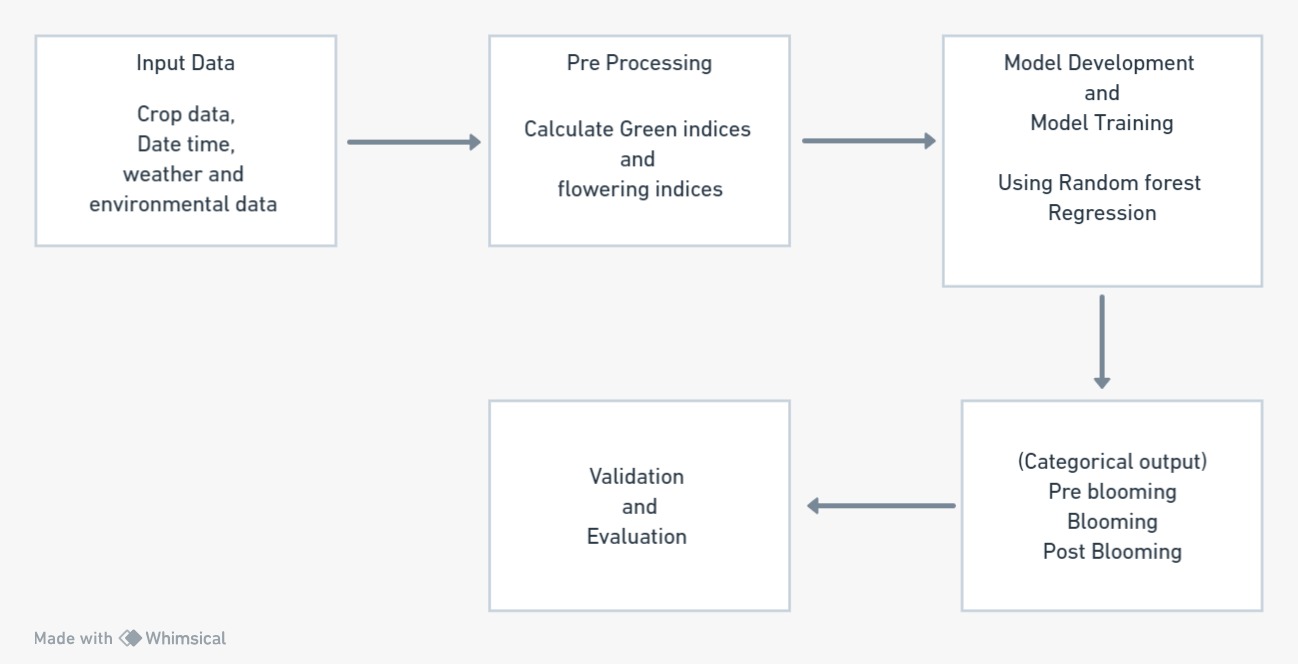


Fig 3.3 Flow diagram of blooming period identification

In Fig 3.3 predicting the blooming status of crops depends significantly on the quality and diversity of the input data collected. Several key pieces of information are essential in this regard, including the specific crop type under consideration, as different crops follow distinct growth patterns. Furthermore, details regarding the soil type play a critical role in understanding how well a particular crop is likely to develop. Weather data, encompassing variables like temperature, precipitation, humidity, and wind, is another crucial aspect, as it directly impacts crop growth and the timing of blooming. Moreover, spectral data derived from remote sensing sources, such as satellite or drone imagery, provides valuable insights into the spectral characteristics of the vegetation, aiding in bloom prediction. Lastly, fertilizer data, which accounts for the type and amount of fertilizers applied to the crops, significantly influences crop development and bloom patterns.

In the preprocessing phase, thorough data cleaning is essential to handle missing values, outliers, and inconsistencies in the input data. This ensures that the subsequent modeling is based on accurate and reliable information. Feature engineering is another critical step, which involves calculating essential vegetation indices like the Normalized Difference Vegetation Index (NDVI) to assess the "greenness" of the crop. Additionally, flowering-related indices are computed based on known bloom characteristics, further refining the dataset. Data integration is necessary to harmoniously combine information from different sources, ensuring that the data is in an appropriate format for modeling.

For the predictive modeling, Random Forest Regression is employed as it is well-suited for both classification and regression tasks. The model is trained using the prepared dataset, with a particular focus on the relationships between input features, such as crop type, soil type, weather conditions, spectral data, and the target variable, which is the bloom status. The tuning of hyperparameters, including the number and depth of trees in the Random Forest model, is carried out to optimize its performance. Additionally, feature importance analysis is conducted to understand which input variables contribute most to the prediction.

The output generated by the Random Forest model is typically in categorical form, assigning each observation to a specific bloom category. Common categories include "pre-blooming," "blooming," and "post-blooming," providing insights into the developmental stage of the crops.

To ensure the model's reliability and accuracy, thorough validation and evaluation are crucial. Model validation involves applying the trained model to a separate testing dataset that was not used during training. If the model's performance falls short of expectations, refinements are made, potentially revisiting data preprocessing, hyperparameter tuning, or even exploring different modeling algorithms. This iterative process aims to create a robust and accurate prediction system for crop bloom status.

.

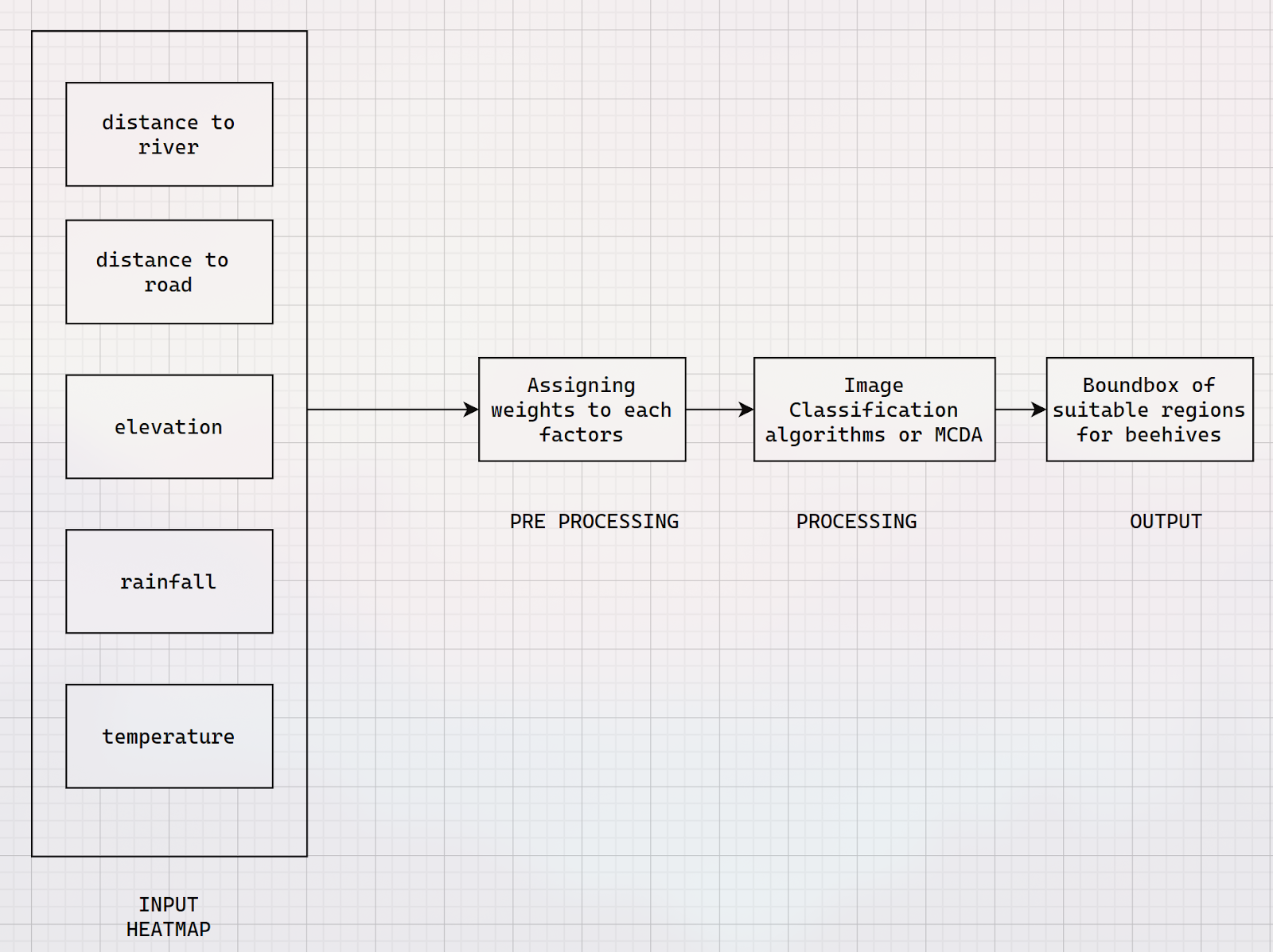


Fig 3.4 Flow diagram of finding suitable regions for beehives

The input data for training the beehive placement model comprises factors delineated in Figure 3.4, encompassing land cover, elevation, temperature, rainfall, distance to the river, and proximity to roadways. The initial phase involves the generation of heatmaps for each of these factors to visualize their spatial distribution and impact. Subsequently, in the preprocessing stage, the relative weights assigned to each factor are determined, a critical step that informs subsequent algorithmic procedures. The model's output is characterized by the identification of regions demonstrating the highest convergence of favorable attributes. This identification is accomplished through the utilization of various image classification algorithms or the application of the Multi-Criteria Decision Analysis (MCDA) algorithm. The outcome is presented in the form of bounding boxes, demarcating regions highly suitable for beehive placement, visually charted on a geographic map.

## 3.2 Methodology

The proposed system is designed to identify ideal locations for placing beehives in farmlands by utilizing a combination of advanced techniques and data sources. At its core, this system aims to distinguish between farmland and forestland in a given region, a crucial step in determining suitable areas for beekeeping.To achieve this, the system employs a fusion of K-Means clustering and Facebook AI Similarity Search (FAISS) to cluster different land classes. These clusters are assigned unique labels, which helps in differentiating farmland from other types of land, such as forests. The utilization of multiple Vegetation Indices (VIs) collected at various times of the year provides a comprehensive view of the land's patterns over time[8].

Beyond land classification, the system integrates weather-related data such as temperature, humidity, sunlight, and wind, as well as crop-specific characteristics, including crop type, age, leaf index, blossoming period, and NDVI flower color characteristics. These parameters collectively paint a detailed picture of the environmental conditions and the specific attributes of the crops in the selected region[5]. By combining all of this valuable dataset, the system is capable of marking hotspot zones where beehives can be optimally placed. These hotspot zones are identified using a specialized algorithm, taking into account a holistic view of factors such as land type, climate, agriculture, and crop characteristics. The goal is to ensure that conditions are suitable for both bee survival and prolific nectar collection. Additionally, this approach contributes to promoting effective pollination, which is essential for crop growth and agricultural sustainability.

### 3.2.1 Identifying Farmland

In our approach to delineating cropland within our dataset, we initiated a series of crucial refinements aimed at improving the precision and efficiency of our classification process. This process commenced with a comprehensive evaluation of the dataset, where we meticulously assessed the indices used in our model, making necessary adjustments to strike a balance between enhancing classification accuracy and optimizing computational resources, a crucial consideration in large-scale remote sensing applications. Once our dataset was reconfigured, we adopted an unsupervised machine learning strategy for the primary phase of land classification. This strategy leveraged K-Means clustering and integrated the Facebook AI Similarity Search (FAISS) framework. It's worth noting that the Sentinel-2 bands in the RE regions, specifically B5, B6 (with a 15 nm spectral width), and B7 (with a 20 nm width), were provided at a 20 m spatial resolution, while the VNIR bands (B2, B3, B4, and B8) exhibited a higher spatial resolution of 10 m alongside lower spectral resolution. This difference in spatial resolution stems from the inherent trade-off in electro-optical system design, considering spatial resolution, spectral resolution, and radiometric sensitivity. (Figure 3.5 depicts these spectral and spatial characteristics.)[27].

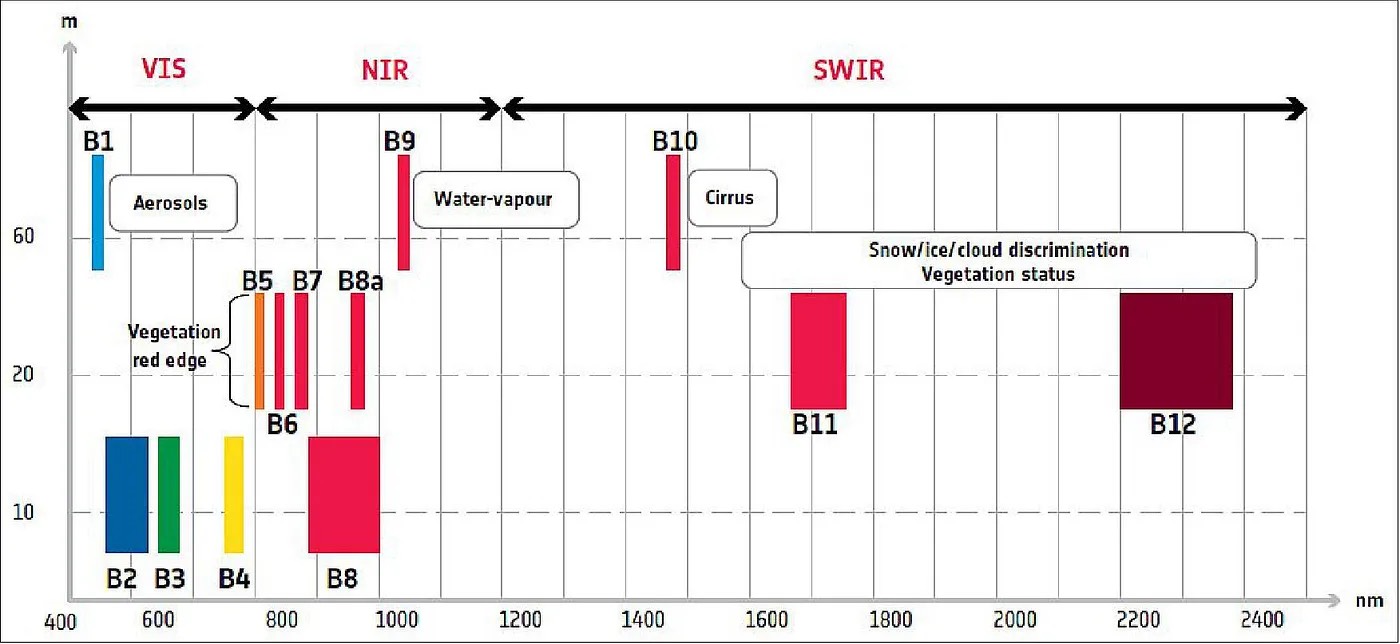
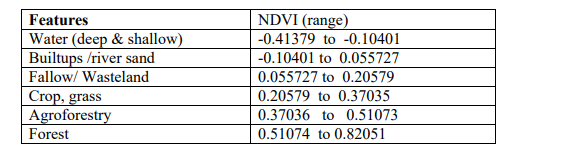


Fig 3.5 Sentinel-2 layout of spectral bands[27].

The NDVI image was examined for distinguishing the NDVI values for different features (Table 3.1). The NDVI values for Agroforestry ranged from 0.37036 to 0.51073, which is lower than natural forest which is almost same as that of previous classification[28].

Table  3.1 - NDVI values for different Features.



  This integration proved to be a game-changer, facilitating efficient similarity search and clustering of dense vectors. K-Means clustering, a data partitioning technique, plays a central role in our methodology. It operates by identifying patterns and similarities among data points based on their intrinsic characteristics, essentially clustering them into distinct groups or clusters. These clusters provide valuable insights into the complex structure of land categories and variations within the dataset.The obtained clusters are then subjected to further analysis, with a specific emphasis on the identification of agricultural cropland. This process involves generating masks that segregate cropland from other land types, creating a clear distinction. By isolating cropland, we can significantly improve the precision and accuracy of land classification, a crucial requirement for applications ranging from precision agriculture and land use planning to environmental monitoring and resource management[8].

As illustrated in Figure 3.6, various land types exhibit distinct indices. For non-farmland regions, the NDVI indices display diverse colors, reflecting varying values. In contrast, areas designated as farmland or cropland exhibit contrasting indices, often characterized by higher NDVI values. This disparity in index values serves as a reliable basis for distinguishing between non-farmland and farmland areas. Consequently, these distinct indices can be effectively utilized to differentiate and classify farmlands, enhancing the accuracy of land cover classification.

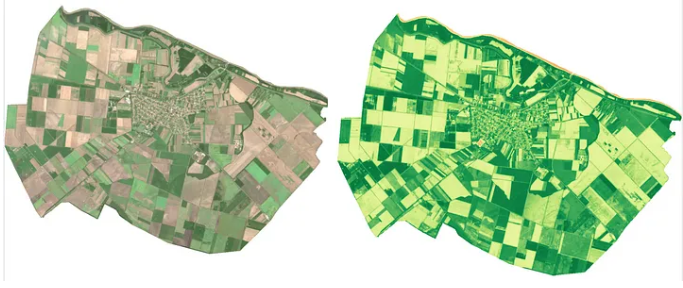


Fig 3.6 - Indices different for different Land[8]

A correlation matrix between all the features and labels is generated to see the trend of relationships between them. Fig 3.7 shows this correlation. From the correlation between all the features and labels in Fig 3.7 ,The analysis reveals that certain features like OSAVI and NDWI exhibit weak correlations with other features and label data. Surprisingly, both NDVI and biomass (bm) not only lack a strong mutual relationship but also display significant correlations with the ground truth data. Subsequently, we conducted an in-depth examination of the relationship between NDVI and biomass to comprehend the trends of these indices in relation to the labels  exploratory data analysis (EDA) suggests that by calculating the differences in NDVI at various time points throughout the year, we can effectively monitor changes in agricultural land density over time and categorize them into distinct classes. To accomplish this, we employ Sentinel-2 satellite imagery to compute NDVI and construct a dataset suitable for clustering. This dataset is created by stacking the time-series NDVI differences, allowing us to identify temporal

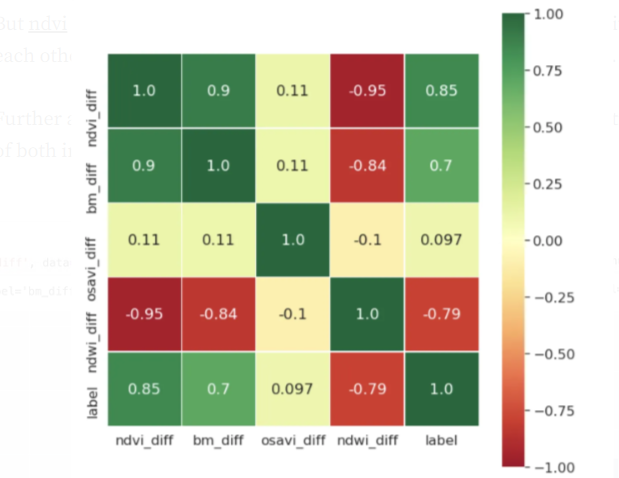


Fig 3.7  Correlation between features and Ground Truth [8]

patterns and changes in agricultural land density[8].Further validation of results is done visually and found that the other classes in lulc are also giving good prediction accuracy using this K-means Clustering Algorithm using FAISS  For eg. For the water bodies identification, an accuracy of more than 78% is achieved[8].

### 3.2.2  Crop Type Identification

Two main strategies are available for multi-temporal crop classification. The first strategy is to stack multi-temporal images by time sequence and classify them with classifiers

such as support vector machine (SVM), random forest and maximum likelihood[5,6]. However, this approach does not model temporal correlations and uses features independently, ignoring possible temporal dependencies[22,23].

Current multi-temporal RS images are multi-spectral, multi-temporal and multi-

spatial. In multi-temporal images, crops are represented via variations in temporal, spectral,

and spatial features. These features can be comprehensively included in four-dimensional

(4D: time, height, width, and band) data that require classification models to learn and

represent temporal, spectral, and spatial features. Multi-temporal images thus pose new

challenges to the models used for data processing, so integrating multi-temporal images

and continuously improving crop classification accuracy requires continued attention[5].

The combination of multi-temporal images and deep learning is an efficient way to obtain

accurate crop distributions and so has drawn increasing attention.

For Training and validation we used The Cropland Data Layer (CDL) is a crop-type distribution product published by the United States Department of Agriculture and the National Agricultural Statistics Service. The 2021 CDL (Figure 1) for Norman County has a spatial resolution of 30 m, and was obtained from the CropScape website portal (https://nassgeodata.gmu.edu/CropScape/(accessed on 20 October 2022)). Although the CDL is not the absolute ground truth, it is the most accurate crop-type product available, especially for corn and soybeans, with over 95% accuracy[5].

This is the overall workflow of Crop Classification using 3D CNN  mentioned in Fig 3.8 Google Earth Engine (GEE) is a versatile tool for remote sensing and offers a comprehensive solution for crop classification. The process begins with the retrieval of multispectral satellite imagery for a specific region of interest. This imagery includes a range of spectral bands like Red, Green, Blue, Near-Infrared (NIR), and Short-Wave Infrared (SWIR). These bands capture various aspects of the Earth's surface, enabling a comprehensive analysis of vegetation and land cover.One crucial step in crop classification is the computation of vegetation indices. GEE provides the tools to calculate these indices, such as the Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), and Enhanced Vegetation Index (EVI). These indices are derived from the spectral bands and provide critical insights into the health and density of vegetation in the target region. NDVI, for instance, is a widely used indicator of vegetation health, with higher values indicating healthier, more abundant vegetation.

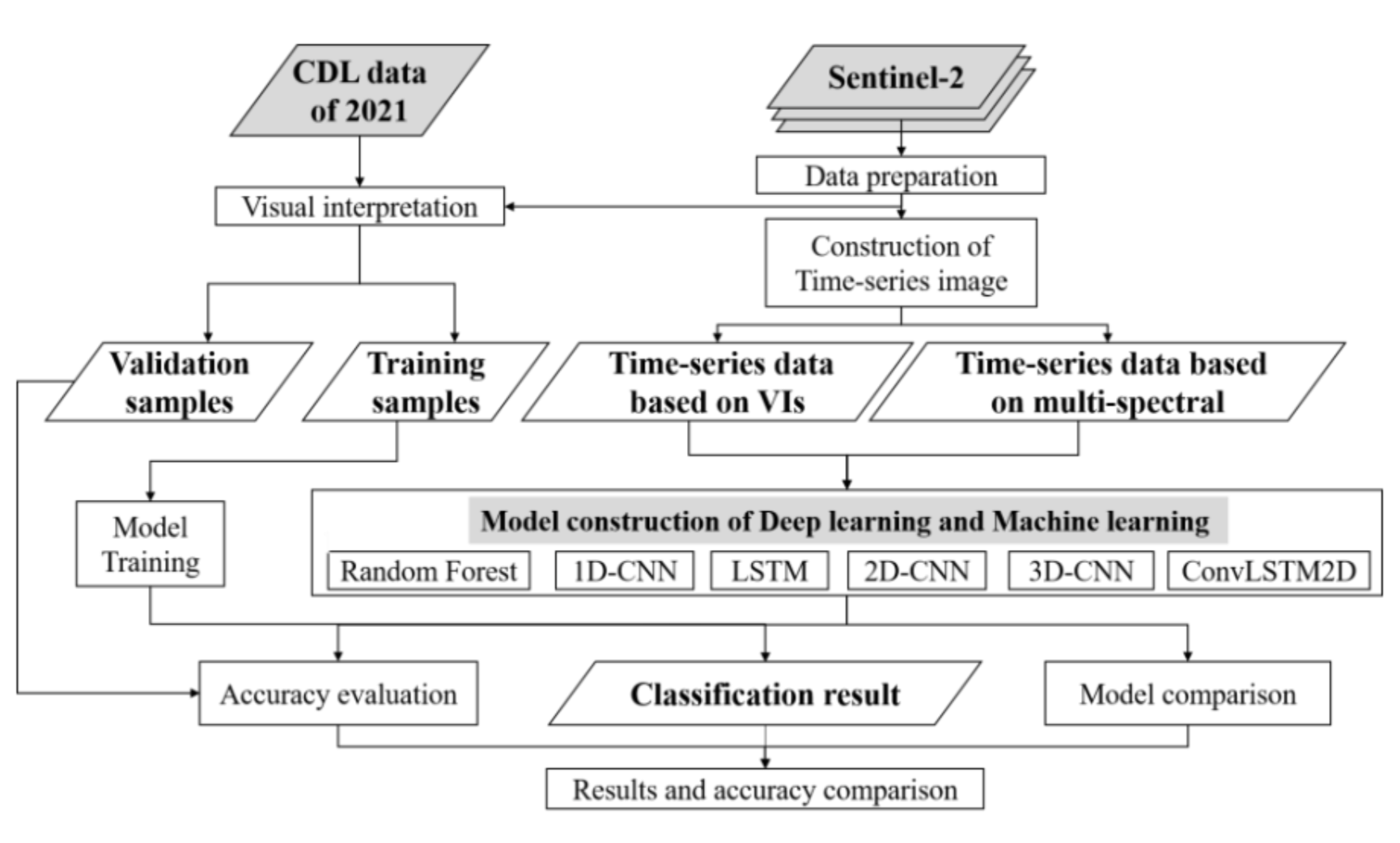


Fig 3.8 -   General workflow for this crop classification[5]

For crop classification process, the calculated vegetation indices are combined to create a 3D dataset. In this dataset, the two spatial dimensions represent the geographical area, while the third dimension represents time. This temporal dimension is crucial for capturing changes in vegetation over time, which is vital for accurate crop classification.After creating the 3D dataset, preprocessing steps are applied, which include data normalization, scaling, and augmentation. These steps are essential to prepare the data for input into a 3D Convolutional Neural Network (3D CNN). The 3D CNN is a deep learning model capable of analyzing spatiotemporal data, making it well-suited for crop classification using multispectral and temporal information.The 3D CNN model is designed to recognize patterns and variations in the 3D dataset and classify them into specific crop types. This model is trained on a labeled dataset, and its parameters are adjusted through iterative optimization. Validation and testing are crucial to ensure that the model's accuracy meets the desired standards.

Table 3.2 The study findings reveal that deep learning outperforms RF-based methods in accuracy, particularly in localized areas, as demonstrated in Figure 6. Overall accuracies (OA) for E3 to E6 spectral bands range from 95.31% to 96.94%. Notably, the inclusion of red-edge (E5) and SWIR (E4) bands significantly enhances crop classification accuracy, with SWIR bands showing a slight advantage. Surprisingly, the LSTM model with E6 proves to be the most accurate configuration, boasting a 1.63% improvement over E3. This underscores the importance of spectral bands in multi-spectral images, which also aids in reducing salt-and-pepper noise and subsequently enhances crop classification precision[5].

Table 3.2  Classification accuracy produced by various models and multi-spectral time-series data.[5]

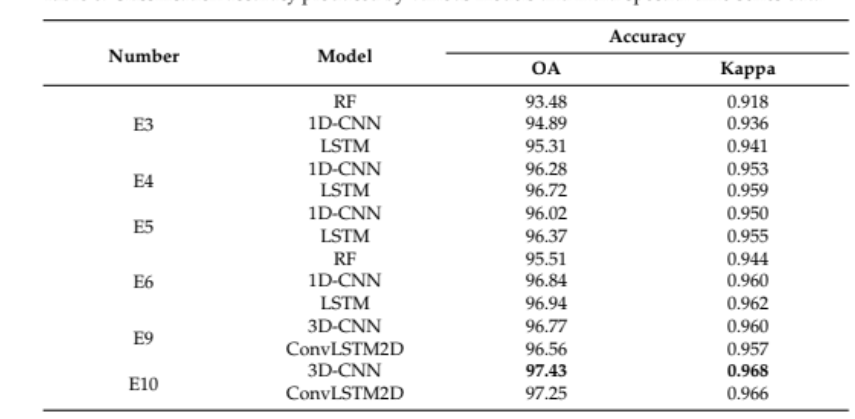


Fig 3.9  represents classification results of E9 and E10 using the 3D-CNN and ConvLSTM2D Models. The OA of 3D-CNN in E9 and E10 was 96.77% and 96.56%, respectively, with kappa coefficients of 0.960 and 0.957. The OA of ConvLSTM2D in E9 and E10 was 97.43% and 97.25%, respectively, with kappa coefficients of 0.968 and 0.966. The accuracy is slightly greater when using the 3D-CNN model than when using the ConvLSTM2D model. The use of the 3D-CNN model on E10 produces the greatest crop classification accuracy of 97.43%[5]. The classification results of the different experiments verify the feasibility of the model constructed herein  for multi-temporal crop classification. The comparison of the results of the different experiments shows that both the construction of the time-series data and that of the classification model influence the crop classification accuracy .

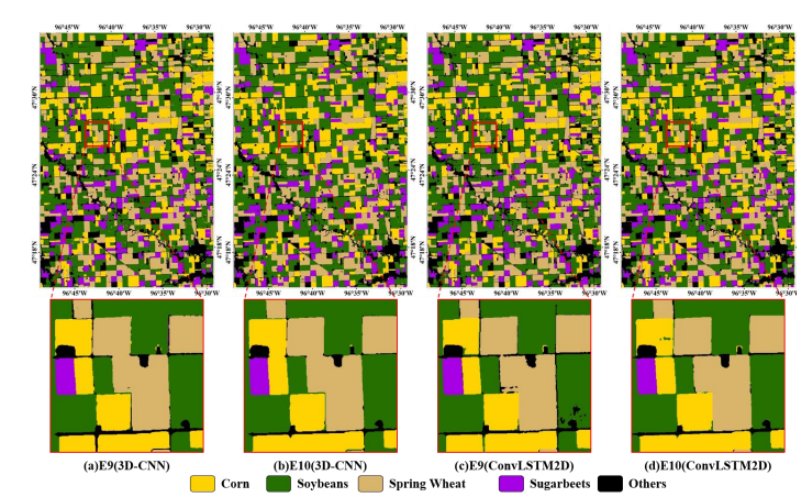


 Fig 3.9 - Crop classification results based on temporal, spectral, and spatial information.[5]

Distinguishing between different crops becomes feasible through the examination of differences in time series curves, as depicted in Figure 3.10 [24]. For instance, the gap observed in the B8 spectral band (as shown in Figure 3.5) during the mid-growing season (approximately DOY 180–220) facilitates the discrimination between spring wheat and sugar beets. Figure 8 illustrates that there is minimal spectral overlap between corn and soybeans in B11 and B12 within a specific timeframe (around DOY 170–200). In the profiles of sugar beets and other crops in bands B6-B8 and B8A, as displayed in Figure 3.8, discernible gaps occur during two distinct periods around DOY 180–220 and 250–270. Spring wheat is distinguishable by examining profiles in B2–B5 around DOY 225, as well as in B11 and B12 from DOY 210–240. Corn and soybeans are more likely to be differentiated during the DOY 170–200 period in B11 and B12. Furthermore, the temporal profile overlap, as observed in the NDVI and other spectra in Figure 3.8, remains similar. The profiles of corn and soybeans show extensive overlap throughout the growing season, explaining the challenge in distinguishing between these two crops. However, the profiles of sugar beets and spring wheat distinctly diverge between DOY 260 and 17[5]. In Conclusion we can use 3D-CNN for crop classification using E9 bands as Multispectral band and Classify the crops with highest Accuracy.

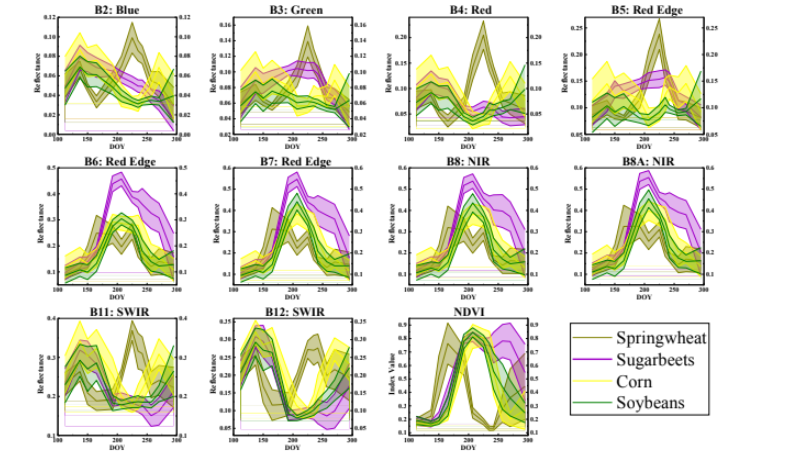
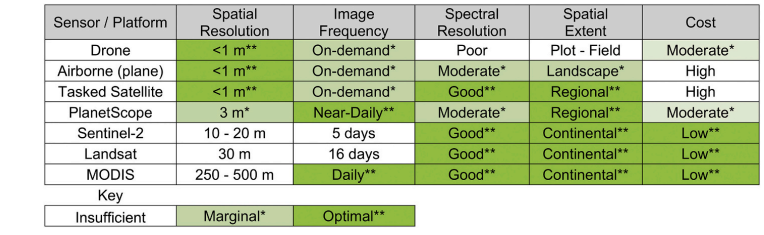


 Fig 3.10 - Time-series spectral band and vegetation indices are aggregated for crop fields. [5]

### 3.2.3 Blooming Period Identification

To effectively determine the optimal timing for placing beehives in a specific region to maximize nectar collection, it is crucial to identify the flowering phases of the crops or plants. These phases typically consist of three key stages: flowering, pre-flowering, and post-flowering. These stages evolve over time, and pinpointing when they occur in a particular region is essential. This knowledge will facilitate informed decision-making for beekeepers and farmers, helping them assess the profitability of hive placement for each crop's specific flowering period.

Table 3.3 Available remote sensing platforms and suggested requirements (\*marginal,    \*\*optimal) to monitor forest flowering phenology[30]



In above Table 3.3 Among the various sensors for remote sensing during the blooming period, Sentinel-2 emerges as a highly attractive choice, primarily due to its combination of remarkable spectral resolution and cost-effectiveness. Sentinel-2's spectral capabilities are noteworthy, featuring a multi-spectral instrument (MSI) that delivers data across 13 spectral bands, encompassing the visible range to the near-infrared and shortwave-infrared regions. This extensive spectral coverage empowers researchers to discern fine details in vegetation and plant health, making it ideal for studying blooming patterns. Moreover, what sets Sentinel-2 apart is its cost efficiency. As part of the European Space Agency's Copernicus program, Sentinel-2 data is made available freely and openly, significantly reducing the financial burden on researchers. The frequent revisit times of Sentinel-2 satellites, approximately every 5 days, are especially beneficial for monitoring dynamic processes like blooming patterns. This enables researchers to capture the various phases of flowering, from pre-flowering to full bloom and post-flowering, with a high degree of temporal granularity. The combination of high spectral and temporal resolution also facilitates multi-temporal analysis, offering insights into how blooming patterns evolve over time, track seasonal changes, and identify trends in plant growth and flowering. Furthermore, Sentinel-2 is well-suited for vegetation monitoring, given its sensitivity to changes in vegetation health and phenology, making it a versatile tool for studying blooming patterns in different types of plants, whether in agricultural, wild, or forested environments. Lastly, its land cover classification capabilities enhance the contextual understanding of flowering events, further contributing to the informed placement of beehives, thus optimizing decision-making for beekeepers, farmers, ecologists, and land managers.

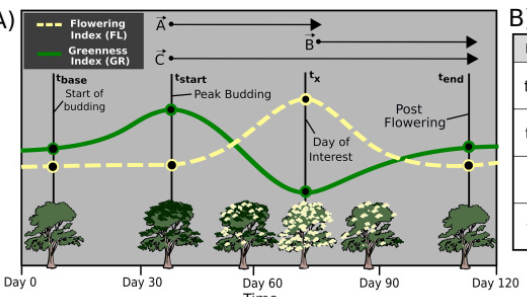


Fig 3.11 Temporal evolution of the GR and FL time series [30]

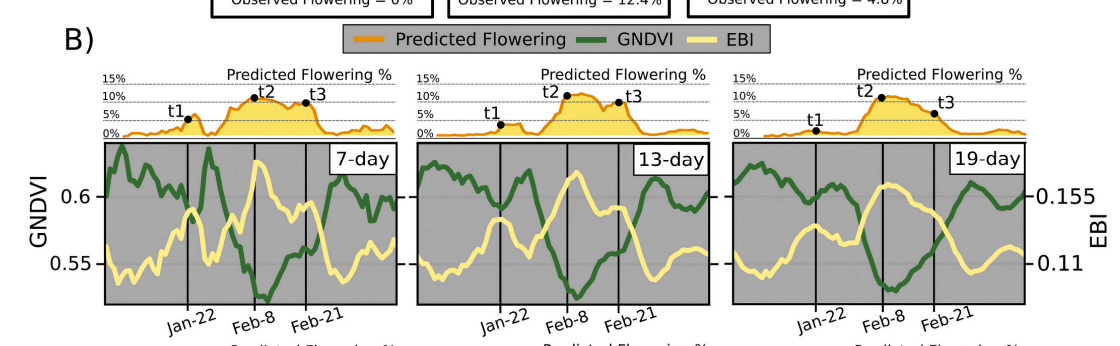


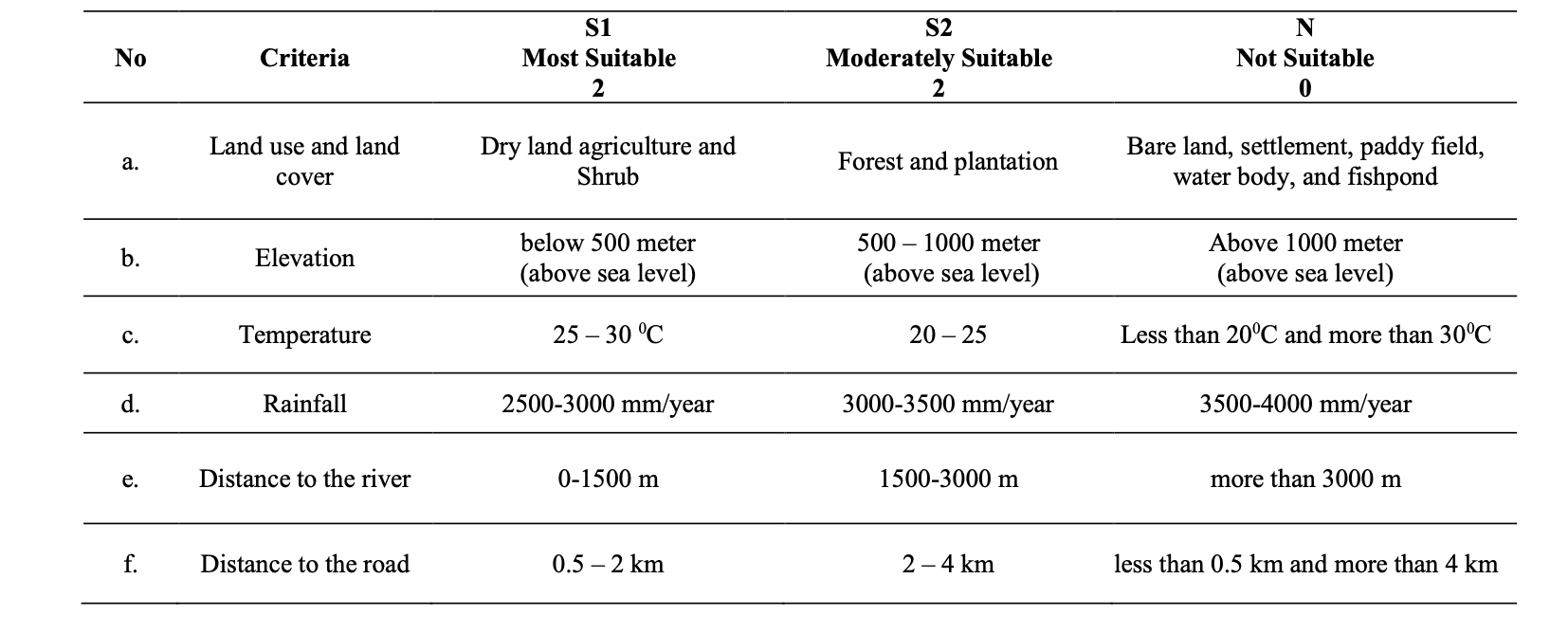
 Fig 3.12  Relation between GNDVI,EBI and predicted Flowering[30]

In fig 3.12 we can depict that this model can be used to generate pixel-level time series of flowering to analyze highly local and regional-level phenology environment relationships.

This model's potential applications extend to the generation of pixel-level time series data representing flowering patterns, allowing for the analysis of phenology-environment relationships at both localized and regional scales. Flowering is a fleeting and highly diverse ecological event that mirrors shifts in environmental conditions. this model can be deployed to create regional maps illustrating the dynamics of flowering in forest ecosystems, even in areas lacking ground truth data. The model relies on predictor variables derived from satellite images' time series of vegetation indices, which are sensitive to the presence of green vegetation and cream-colored flowers.the results demonstrated promising performance on specific dates, such as during the pre-flowering, flowering, and late-flowering periods[30]. Importantly, the model exhibited resilience in avoiding false predictions of flowering outside the actual flowering season and aptly captured the temporal progression of flowering events throughout the growing season. The predictions generated by this model serve as valuable tools for monitoring the condition of forest ecosystems and facilitating further research into the environmental factors influencing the phenology of eucalypt forests.Thus,we can use GNDVI and EBI for identifying blooming period across different regions and different crops.

### 3.2.3 Finding Regions To Place Bee Hives

The biophysically suitable area in this analysis consist of six criterias, namely land use, distance to the river, distance to the road, temperature, elevation and rainfall [29]. Elevation has strong relation with temperature and bee activity. When elevation is too high, temperature will decrease that would Affect The Activity of bee. This means that the productivity of honey will decrease. The elevation that is too high will create difficulty for farmers in maintaining, overseeing and moving the hive to other place [29].

Table 3.4 -  Classification of the biophysically suitable criteria for beehives [29].

Hence, the input for the model training is the factors mentioned in the Table 3.4 above such as land cover, elevation, temperature, rainfall, distance to the river and distance to the road. Heatmap with respect to each factors would be plotted and observed and regions with maximal overlap of suitable characteristics would be identified and marked as a suitable region using classification algorithms. The output would be  boundbox of regions which are more suitable for beehives plotted on the map.

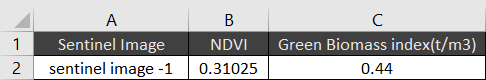
# Chapter 4

# Implementation Details

## 4.1 Experimental Setup

### 4.1.1 Farmland identification dataset

Table 4.1 Farmland identification dataset header



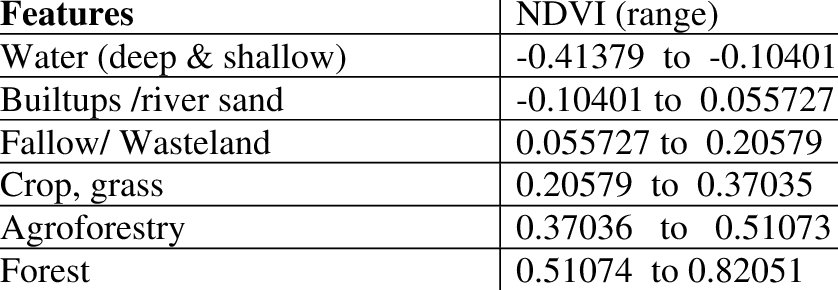
The dataset consists of relevant input features and gives corresponding results of clustering, which are used for farmland identification. The three columns include:

* The "Sentinel Image" is a critical source of high-resolution, multispectral satellite imagery captured by the Sentinel-2A satellite. These images contain valuable information about the land cover and vegetation in the study area. Sentinel images provide a rich source of data, including visible, near-infrared, and shortwave-infrared spectral bands that can reveal variations in land cover and vegetation health.
* The "NDVI" values are derived from the Sentinel images. NDVI is a widely used vegetation index calculated from the near-infrared and red spectral bands. It provides a quantifiable measure of vegetation density and health. NDVI values range from -1 to 1, where higher values indicate healthy vegetation, and lower values correspond to non-vegetated or stressed areas.
* The "Green Biomass Index" is another vegetation-related metric calculated from the Sentinel images. It measures the amount of green plant material, which is a key indicator of crop density and health. Higher values of the Green Biomass Index typically signify more abundant and healthy vegetation.

Unsupervised clustering is a data analysis technique that groups data points into clusters based on their similarity. In the context of farmland identification, the Sentinel Image, NDVI, and Green Biomass Index data are used collaboratively within an unsupervised clustering algorithm [8]. These features, particularly NDVI and Green Biomass Index, provide critical information about the land's vegetation. Unsupervised clustering aims to group areas with similar vegetation characteristics, effectively segregating farmland from other land cover types. By analyzing the combined information from NDVI and Green Biomass Index in the clustering process, the algorithm can automatically group regions exhibiting similar vegetation patterns. These clusters are significant in identifying areas with potential agricultural activity [9].

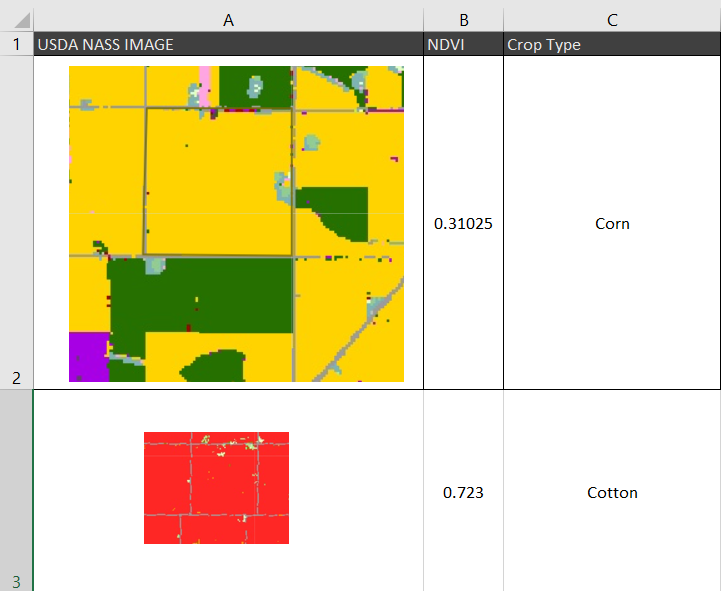
Table 4.2 presents the NDVI values, which provide valuable insights into the clustering of farmland [8].

Table 4.2 NDVI range for different features [8]



### 4.1.2 Crop type identification dataset

Table 4.3 Crop identification dataset header [25]



The dataset is built on the integration of NDVI values derived from satellite imagery and USDA NASS Cropland Data Layers. Table 4.3 show the columns included in the dataset [25].

* NDVI (Normalized Difference Vegetation Index): The "NDVI" column contains NDVI values extracted from satellite imagery. NDVI serves as a vital indicator of vegetation health and density in the target agricultural areas.
* USDA NASS Cropland Data Layers: These layers are obtained from the United States Department of Agriculture's (USDA) National Agricultural Statistics Service (NASS) and provide valuable information about cropland distribution and types.
* Crop Type (Output label) : The "Crop Type" column represents the categorical output label, specifying the type of crop grown in the corresponding agricultural areas. Each sample is assigned a specific crop type label.

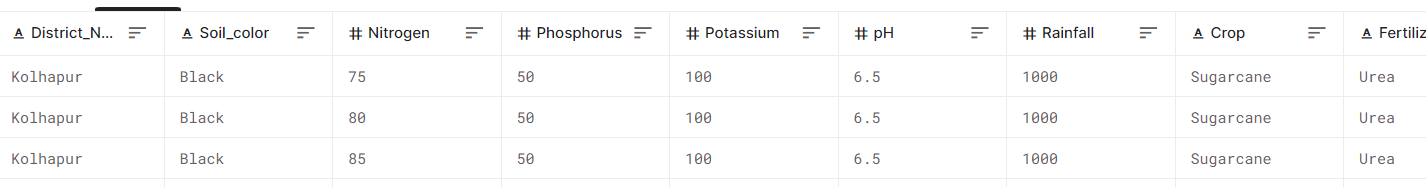
### 4.1.3 Crop and Fertiliser dataset

This section provides a detailed description of the dataset used for analyzing the impact of fertilizer and soil properties on plant growth. The dataset encompasses various features that influence plant health and crop yield, including "District Name," "Soil Color," and various soil and environmental parameters. The columns shown in Table 4.4 are as follows:

* District Name: The "District Name" column contains the geographic location of the data points, which may represent different agricultural regions or districts.
* Soil Color: The "Soil Color" column categorizes the color of the soil, which can provide insights into the soil type and composition.
* Soil Nutrients: The "Nitrogen", "Phosphorus", "Potassium” column contains information about the soil nutrient content in the soil, a key nutrient for plant growth.
* pH: The "pH" column records the soil pH level, which is crucial for nutrient availability and plant growth.
* Rainfall: The "Rainfall" column records the amount of precipitation or rainfall in the region. Rainfall significantly influences plant water supply and crop growth.
* Fertilizer: The "Fertilizer" column includes information about the type and quantity of fertilizer applied to the soil. Fertilizer is an important factor that influences plant growth and nutrient availability.

It is important to note that the impact of fertilizer data on blooming patterns can vary based on the specific crop, soil conditions, and the region's climate. Therefore, it is crucial to analyze the data in the context of the specific agricultural setting and crop types under consideration. By combining fertilizer data with other relevant factors, such as weather conditions and soil properties, we can build more accurate models for estimating blooming pattern timelines, which is valuable for agricultural planning and management.

Table 4.4 Crop and Fertilizer dataset header [33]



### 4.1.4 Weather Dataset

The combination of these weather-related parameters allows for a more comprehensive analysis of environmental conditions that influence plant growth and blooming patterns. Incorporating this data into the predictive model can enhance its accuracy and effectiveness in estimating blooming timelines for different crops and agricultural regions.

Table 4.5 Temperature dataset header [34]

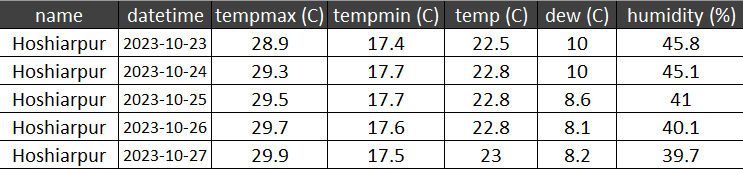


Table 4.5 is described as follows:

* Place Name: The "Place Name" column specifies the geographic location for which temperature data is recorded.
* Date: The "Date" column represents the timestamp of the recorded weather data.
* Temp Max: "Temp Max" denotes the maximum temperature observed on the given date.
* Temp Min: "Temp Min" represents the minimum temperature recorded on the given date.
* Temp: The "Temp" column provides the average temperature for the specified location and date.
* Dew: "Dew" signifies the dew point temperature, an important metric for understanding humidity and potential condensation.
* Humidity: "Humidity" indicates the relative humidity, which influences plant transpiration and growth.

Table 4.6 Precipitation and cloud dataset header [34]

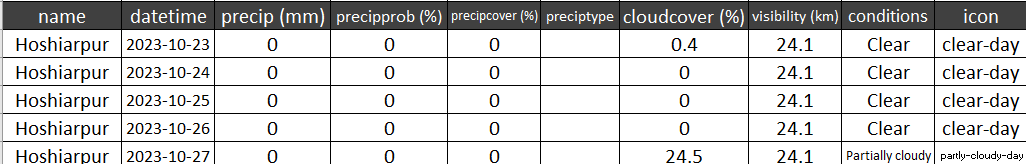


Table 4.6 is described as follows:

* Precipitation: "Precipitation" denotes the amount of rainfall or other forms of precipitation on the given date.
* Precipitation Probability: "Precipitation Probability" provides the likelihood of precipitation occurring.
* Precipitation Type: "Precipitation Type" specifies the form of precipitation (e.g., rain, snow).
* Precipitation Cover: "Precipitation Cover" describes the extent of the area affected by precipitation.
* Cloud Cover: "Cloud Cover" quantifies the portion of the sky obscured by clouds.
* Visibility: "Visibility" indicates the distance at which objects can be clearly seen, which can affect plant growth.
* Conditions: "Conditions" describe the general weather conditions, providing additional context.

Table 4.7 Wind dataset header [34]

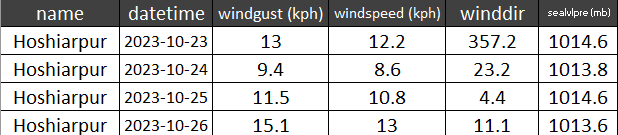
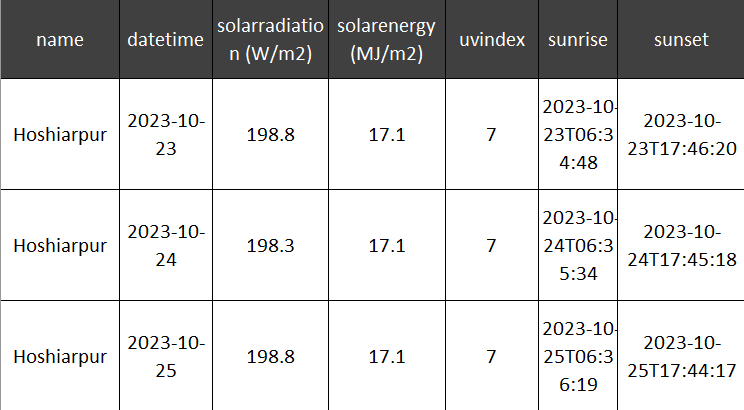


Table 4.7 is described as follows:

* Wind Gust: "Wind Gust" represents the highest wind speed observed during the day.
* Wind Speed: "Wind Speed" records the average wind speed.
* Wind Direction: "Wind Direction" indicates the prevailing wind direction.
* Sea Level Pressure: "Sea Level Pressure" provides information on atmospheric pressure at sea level, influencing weather patterns.

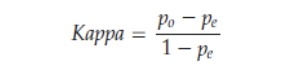
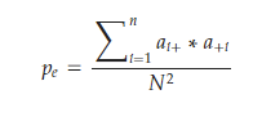
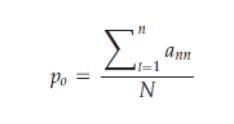
Table 4.8 is described as follows:

* Solar Radiation: "Solar Radiation" measures the incoming solar energy, which is vital for photosynthesis and plant growth.
* Solar Energy: "Solar Energy" represents the solar energy available for plant utilization.
* UV Index: "UV Index" quantifies the level of ultraviolet radiation, which can affect plant health.
* Sunrise and Sunset: "Sunrise" and "Sunset" times provide information about daylight duration.

Table 4.8 Solar dataset header [34]

## 4.2 Performance Evaluation

### 4.2.1 Performance Evaluation for crop identification

The cross-validation was used in the study. Some samples were selected in the training set to test the model. A part of the training set data was retained as the test set, and the parameters generated by the training set were tested to determine the degree of conformity of these parameters to the data outside the training set relatively objectively. The whole dataset was fixed into a training set and a test set. The samples were randomly selected in the proportion of 7:3 in this study. That is, 70% of all samples were used for the training set and 30% for the test set. After all classifications were completed, a confusion matrix was used to analyze the accuracy, and the kappa coefficient was used to test the consistency. Finally, producer accuracy, user accuracy, and overall accuracy were calculated in this study [6].

### 4.2.2 Performance Evaluation using ground truth

Ground truth evaluation is essential for assessing the accuracy and reliability of predictive models, especially in projects involving remote sensing and data analysis for agriculture, such as predicting blooming patterns.

**MAE (Mean Absolute Error)**

Calculate Absolute Errors: For each data point in ground truth dataset, absolute error is calculated by taking the absolute difference between the predicted date and the actual date of blooming. The formula for calculating the absolute error (AE) for a single data point is:

{ AE = |Actual Date - Predicted Date| }

Calculate MAE: Once calculated the absolute errors for all data points, calculate the Mean Absolute Error (MAE) by taking the average of these absolute errors. The formula for calculating MAE is: { MAE = (Σ AE) / n }. Σ AE represents the sum of all absolute errors. n is the total number of data points in ground truth dataset.

Interpretation: A lower MAE indicates that model's predictions are, on average, closer to the actual blooming dates, suggesting better accuracy. A higher MAE indicates that model's predictions are, on average, farther from the actual blooming dates, suggesting lower accuracy.

**RSME (Root Mean Squared Error):**

Calculate Squared Errors: For each data point in ground truth dataset, squared error is found by the square of the difference between the predicted date and the actual date of blooming. The formula for calculating the squared error (SE) for a single data point is:

{ SE = (Actual Date - Predicted Date)^2 }

Calculate Mean Squared Error (MSE): After calculating the squared errors for all data points, the Mean Squared Error (MSE) is computed by taking the average of these squared errors. The formula for calculating MSE is: MSE = { (Σ SE) / n }. Σ SE represents the sum of all squared errors. n is the total number of data points in ground truth dataset. Calculate RMSE: Finally, the Root Mean Squared Error (RMSE) is calculated by taking the square root of the MSE. The formula for calculating RMSE is: { RMSE = √(MSE) }.

Interpretation: A lower RMSE indicates that model's predictions are, on average, closer to the actual blooming dates, suggesting better accuracy, and smaller errors. A higher RMSE suggests that model's predictions are, on average, farther from the actual blooming dates, indicating lower accuracy and larger errors.

## 4.3 Software and Hardware Setup

FAISS library: FAISS (Facebook AI Similarity Search) is a library developed by Facebook for efficient similarity search and clustering of high-dimensional vectors. It is widely used in the field of machine learning and allows you to perform tasks such as nearest neighbor search on large datasets, making it particularly useful for applications like recommendation systems and image retrieval.

scikit-learn: Scikit-learn is an open-source machine learning library for Python. It provides a wide range of tools for tasks related to classification, regression, clustering, dimensionality reduction, and more. Scikit-learn is known for its ease of use and extensive documentation, making it a popular choice for machine learning practitioners.

TensorFlow: TensorFlow is an open-source deep learning framework developed by Google. It is widely used for building and training deep neural networks. TensorFlow offers both high-level and low-level APIs, making it suitable for a range of machine learning tasks, from simple models to complex deep learning architectures.

PyTorch: PyTorch is another open-source deep learning framework, developed by Facebook's AI Research lab (FAIR). It is known for its dynamic computation graph and flexibility, making it a popular choice for researchers and developers in the field of deep learning. PyTorch is often used for tasks like natural language processing and computer vision.

Google Earth Engine: Google Earth Engine is a cloud-based geospatial analysis platform developed by Google. It provides a vast amount of Earth observation data and tools for analyzing and visualizing geospatial data. Researchers and scientists use it for a wide range of applications, including environmental monitoring, deforestation tracking, and land use analysis.

Copernicus: Copernicus is the European Union's Earth Observation program. It consists of a series of Earth-observing satellites, known as the Copernicus Sentinels, and a range of services for environmental and security monitoring. Copernicus data and services are invaluable for tasks such as climate monitoring, disaster management, and urban planning.

NumPy: NumPy, short for "Numerical Python," is a fundamental library for scientific computing in Python. It provides support for large, multi-dimensional arrays and matrices, along with a wide variety of mathematical functions to operate on these arrays. NumPy is the foundation for many other scientific and data analysis libraries in Python.

Pandas: Pandas is an open-source data manipulation and analysis library for Python. It provides data structures, such as DataFrames, for handling structured data and tools for data cleaning, transformation, and analysis. Pandas is commonly used in data science and data analysis projects.

Jupyter Notebook: Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text. It's widely used in data science and scientific research for interactive and reproducible data analysis. Jupyter supports various programming languages, including Python and R.

# Chapter 5

# Implementation

## 5.1 Timeline Chart for Term 1 and Term 2

### 5.1.1 Term 1

**Semester 7 (August 2023 - December 2023)**

**Week 1 - 3 : Project Research**

* + - Conduct an extensive literature review to identify existing methodologies and technologies for farmland and crop identification using satellite imagery.
    - Explore different remote sensing datasets, including Sentinel 1 , Sentinel 2 and Landsat, and assess their suitability for the target region.
    - Identify potential algorithms, such as FAISS for farmland identification and CNN for crop type identification, based on the literature review.

**Week 4 - 5 : Data Collection and Preparation**

* + - Acquire high-resolution satellite imagery for the target region, ensuring sufficient coverage over multiple seasons.
    - Collect ground truth data, either through field surveys or by collaborating with local agricultural agencies and farmers, to create a labeled dataset for training and validation.

**Week 6 - 8 : Implement Farmland detection Algorithm**

* + - Implement the FAISS clustering algorithm for farmland identification, using features extracted from the satellite imagery.
    - Train the algorithms using the labeled dataset, fine-tuning parameters, and optimizing model performance.

**Week 9 - 11 : Implement Crop Type detection Algorithm**

* + - Develop a Convolutional Neural Networks (CNN) model for crop type identification, incorporating relevant spectral and textural features.
    - Train the algorithms using the labeled dataset, fine-tuning parameters, and optimizing model performance.

**Week 12 : Documentation**

* + - Prepare comprehensive documentation of the methodology, model performance, and any refinements made during Term 1.
    - Present the findings and progress to professor incharge for feedback and future planning.

### 5.1.2 Term 2

**Semester 8 (January 2024 - May 2024)**

**Week 1 - 2 : Integration**

* + - Integrate the farmland and crop identification models into a cohesive system that can process and analyze satellite imagery for the target region.
    - Test the cohesive system and ensure its proper working.

**Week 3 - 4 : Research**

* + - Conduct extensive literature review to find out different methodologies to implement Blooming period prediction.
    - Conduct extensive literature review to find out different methodologies to implement Bee Hive placement prediction.

**Week 4 - 5 : Data Collection**

* + - Acquire dataset about blooming period of different crop types.
    - Acquire dataset of different factors affecting bee hive placement and their particular effects on the surrounding crops.

**Week 6 - 8 : Implement Blooming pattern prediction Algorithm**

* + - Implement GNDVI and EBI for identifying blooming period across different regions and different crops.

**Week 9 - 11 : Implement Bee Hive placement Algorithm**

* + - Plot the heat map based on multiple factors mentioned above and predict the bound box for bee hive placement.

**Week 12 : Documentation**

* + - Prepare comprehensive documentation of the methodology, model performance, and any refinements made during Term 2.
    - Present the findings and progress to professor incharge for feedback and future planning.

**Week 13 - 16 : Presentation**

* + - Finalize the project, addressing any remaining issues or bugs.
    - Create a professional presentation and rehearse for the final presentation.

# Chapter 6

# Conclusion

In conclusion, this project serves a vital purpose by addressing the critical issue of pollination enhancement through a multifaceted approach. By meticulously identifying and studying the blooming periods of diverse plant species across India, considering a myriad of environmental factors, and leveraging satellite data for farmland analysis, it strives to optimize pollination. The project's three core objectives, encompassing farmland identification, blooming season prediction, and suitable beehive placement, synergize to promote sustainable agriculture and support the pivotal role of bees in ecosystems. Through the integration of machine learning and remote sensing, this holistic initiative offers a promising path toward the betterment of both agricultural practices and ecological well-being on a national scale.

## 

# References

[1] R Hamdamov and H Rakhmanov 2019 J. Phys.: Conf. Ser. 1260 102005. Remote monitoring of agricultural land using multispectral satellite imagery Sentinel 2 by

contour analysis.

[2] Zhang, Z.; Tang, P.; Duan, R. Dynamic time warping under pointwise shape context. Inf. Sci. 2015, 315, 88–101. [CrossRef].

[3] Guo, W.; Zhang, W.; Zhang, Z.; Tang, P.; Gao, S. Deep Temporal Iterative Clustering for Satellite Image Time Series Land Cover.

[4] Felegari, S.; Sharifi, A.; Moravej, K.; Amin, M.; Golchin, A.;Muzirafuti, A.; Tariq, A.; Zhao, N. Integration of Sentinel 1 and Sentinel 2 Satellite Images for Crop Mapping. Appl. Sci. 2021, 11,10104.

[5] Li, Q.; Tian, J.; Tian, Q.Deep Learning Application for Crop Classification via Multi-Temporal Remote Sensing Images. Agriculture 2023, 13, 906.

[6] Xue, H.; Xu, X.; Zhu, Q.;Yang, G.; Long, H.; Li, H.; Yang, X.;Zhang, J.; Yang, Y.; Xu, S.; et al.Object-Oriented Crop Classification Using Time Series Sentinel Images from Google Earth Engine. RemoteSens. 2023, 15, 1353.

[7] https://write.agrevolution.in/satellite-imagery-downloading-and-processing-dfc6035e18de

[8] https://write.agrevolution.in/agriculture-area-identification-using-sentinel-2-time-series-part-1-b04ce0959c2a

[9] https://write.agrevolution.in/identifying-agricultural-land-using-satellite-imagery-and-unsupervised-ml-511507bb2c5c

[10] Li, X.; Xu, X.; Wang, J.; Wu, H.; Jin, X.; Li, C.; Bao, Y. Crop classification recognition based on time-series images from HJ satellite.Trans. Chin. Soc. Agric. Eng. 2013, 29, 9, (In Chinese with English abstract).

[11] Li, H.; Li, K.; Shao, Y.; Zhou, P.; Guo, X.; Liu, C.; Liu, L. Retrieval of Rice Phenology Based on Time-Series Polarimetric SARData. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia,Spain, 22–27 July 2018; pp. 4463–4466. [CrossRef].

[12] Abdikan, S.; Sanli, F.; Ustuner, M.; Caló, F. Land cover mapping using sentinel-1 SAR data. ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2016, XLI-B7, 757–761.

[13] Hao, P.; Zhan, Y.; Wang, L.; Niu, Z.; Shakir, M. Feature Selection of Time Series MODIS Data for Early Crop Classification Using Random Forest: A Case Study in Kansas, USA. Remote Sens. 2015, 7, 5347–5369. [CrossRef].

[14] Liu, Z.; Liu, D.; Zhu, D.; Zhang, L.; Zan, X.; Tong, L. Research progress and prospect of fine recognition and automatic mapping of crops by remote sensing. Trans. Chin. Soc. Agric. Mach. 2018, 49, 1–12. (In Chinese) [CrossRef].

[15] Thanh Noi, P.; Kappas, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. Sensors 2018, 18, 18. [CrossRef] [PubMed].

[16] Ndikumana, E.; Ho Tong Minh, D.; Baghdadi, N.; Courault, D.; Hossard, L. Deep recurrent neural network for agricultural classification using multitemporal SAR Sentinel-1 for Camargue, France. Remote Sens. 2018, 10, 1217. [CrossRef].

[17] Cai, Y.; Guan, K.; Peng, J.; Wang, S.; Seifert, C.; Wardlow, B.; Li, Z. A high performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. Remote Sens. Environ. 2018, 210, 35–47.[CrossRef].

[18] Moskolaï, W.R.; Abdou, W.; Dipanda, A. Application of deep learning architectures for satellite image time series prediction: A review. Remote Sens. 2021, 13, 4822. [CrossRef].

[19] Pekel, J.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. Nature 2016, 540, 418–422. [CrossRef].

[20] Fauvel, M.; Tarabalka, Y.; Benediktsson, J.A.; Chanussot, J.; Tilton, J.C. Advances in spectral-spatial classification of hyperspectral images. Proc. IEEE 2013, 101, 652–675. [CrossRef].

[21] Yu, Q.; Gong, P.; Clinton, N.; Biging, G.; Kelly, M.; Schirokauer, D. Object-based Detailed Vegetation Classification with Airborne High Spatial Resolution Remote Sensing Imagery. Photogramm. Eng. Remote Sens. 2006, 72, 799–811. [CrossRef].

[22] Cai, Y.; Guan, K.; Peng, J.; Wang, S.; Seifert, C.; Wardlow, B.; Li, Z. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. Remote Sens. Environ. 2018, 210, 35–47.[CrossRef].

[23] Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object- based time-weighted dynamic time warping analysis. Remote Sens. Environ. 2018, 204, 509–523. [CrossRef]

[24] Wang, H.; Zhao, X.; Zhang, X.; Wu, D.; Du, X. Long time series land cover classification in China from 1982 to 2015 based on Bi-LSTM deep learning. Remote Sens. 2019, 11, 1639. [CrossRef]

[25] USDA National Agricultural Statistics Service Cropland Data Layer. 2023. Published crop-specific data layer [Online].

[26] Triantomo, Varian, Widiatmaka Widiatmaka and Asnath Maria Fuah. “LAND USE PLANNING FOR BEEKEEPING USING GEOGRAPHIC INFORMATION SYSTEM IN SUKABUMI REGENCY, WEST JAVA.” Journal of Natural Resources 6 (2016): 168-168.

[27] Alparone, L.; Garzelli, A.;Zoppetti, C. Fusion of VNIR Optical and C-Band Polarimetric SAR Satellite Data for Accurate Detection of Temporal Changes in Vegetated Areas. Remote Sens. 2023, 15, 638.

[28] TY - BOOK AU - Rizvi, Raza AU - Yadav, Ram AU - Singh, Ramesh AU - Datt, Keshav AU - Khan, I AU - Dhyani, S.PY - 2009/09/17 T1 - Spectral Analysis of Remote Sensing Image for Assessment of Agroforestry Areas in Yamunanagar district of Haryana.

[29] Janssen, L. L. F, 2000. Visual Image Interpertation. Principles of Remote Sensing (ITC Educational Textbook Series 2) pp.125-139.

[30] Dan J. Dixon, J. Nikolaus Callow, John M.A. Duncan, Samantha A. Setterfield, Natasha Pauli,Satellite prediction of forest flowering phenology,Remote Sensing of Environment,Volume 255,2021,112197,ISSN 0034-4257.

[31] Li,C.;Zhou,L.;Xu,W. Estimating Aboveground Biomass Using Sentinel-2 MSI Data and Ensemble Algorithms for Grassland in the Shengjin Lake Wetland, China.

[32] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, L. G. Ferreira. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment 83(2002) 195-213.

[33] Kaggle : https://www.kaggle.com/datasets/sanchitagholap/crop-and-fertilizer-dataset-for-westernmaharashtra

[34] https://www.visualcrossing.com/weather/weather-data-services