MADRAS INSTITUTE OF TECHNOLOGY
DEPARTMENT OF COMPUTER TECHNOLOGY
CS6611 CREATIVE AND INNOVATIVE PROJECT

CropForEst - A Machine learning based crop yield estimation and profitability analysis for precision agriculture.

First Review

Team Members:
SUNDARAJAN R B / 2020503549
HAMILTON SAMIC S / 2020503511
SURIYAA V / 2020503550

Project Guide: Dr. C. Valliyammai

PROBLEM STATEMENT:

The agriculture industry faces a significant challenge in predicting crop yield production and crop demand accurately. The traditional methods of predicting crop yield and demand are time-consuming and can be prone to errors, leading to inefficient use of resources and revenue loss. There is a need for predicting the most suitable crop based on various influential factors and forecasting the profitability.

OBJECTIVE:

- To develop an intelligent platform that uses machine learning algorithms to estimate the most suitable crop for cultivation in a given area based on environmental factors.
- To forecast market demand and pricing trends for the identified crop, with the aim of increasing profitability for farmers and stakeholders in the agriculture industry.

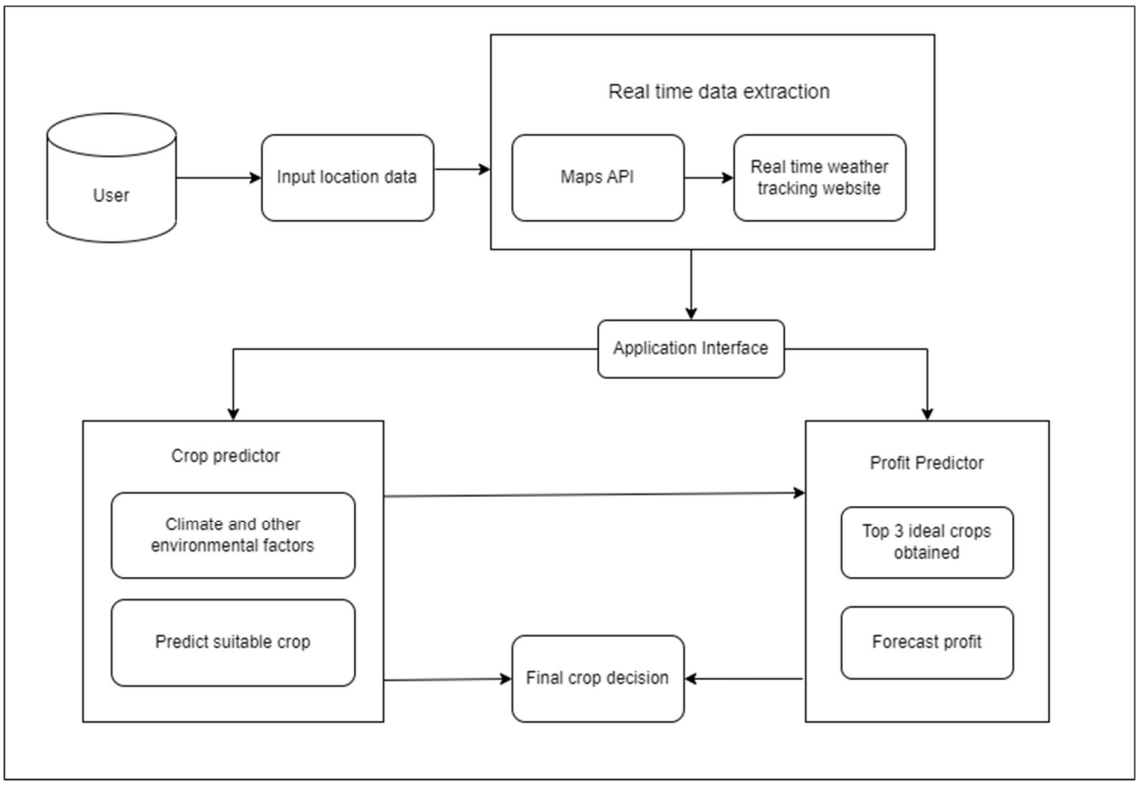
LITERATURE SURVEY:

Publication	Methodology	Outcome
Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications **Authors**: D. Elavarasan and P. M. D. Vincent **Journal**: IEEE Access, vol. 8, pp. 86886-86901, 2020	• An RNN-based feature processing is combined with Deep Recurrent Q-Network model based self-experimental analysis is constructed to forecast the crop yield. • Forecast is done based on major climatic factors, soil parameters dataset extracted from Indian Meteorological department's portal.	• The results of DRQN model were compared with other ANN, BAN models using certain evaluation metrics like error, variance score and it's outperformed all of them with 94% accuracy. • The probability density of actual and predicted yield were also measured.
Estimation of crop yield from combined optical and SAR imagery using gaussian kernel regression **Authors**: Alebele, Yeshanbele and Wang, Wenhui and Yu, Weiguo and Zhang, Xue and Yao, Xia and Tian, Yongchao and Zhu, Yan and Cao, Weixing and Cheng, Tao **Journal**: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing	• Estimation methods selected based on the tradeoff between the performance in terms of given target parameters, interpretability of results, and computational time. • Multiple linear regression, random forest, and neural network are mostly used	• Using Gaussian Linear regression model , a range of crop yield predicted from an unknown distribution. • GPR attempts to approximate the target output f(x) where by interpreting it as a probability distribution function.
Improvised Extreme Learning Machine for Crop prediction. **Authors:** S. Vashisht, P. Kumar and M. C. Trivedi **Journal:** 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2022	• The proposed approach uses a two stage process to predict crop yield. First stage - PSO algorithm is used to optimize the input weights and biases of the EML algorithm, which is a type of feedforward network with a single hidden layer. • Second stage - Optimised EML model is used to predict the crop yield based on the input features.	This technique has been found to outperform traditional methods of crop yield prediction that use only one or two spectral bands resulting in improved accuracy of crop yield

Agricultural Monitoring, an Automatic Procedure for Crop Mapping and Yield Estimation: The Great Rift Valley of Kenya Case **Authors:** R. Luciani, G. Laneve and M. J **Journal:** IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 7, pp. 2196-2208, July 2019,	• Satellite images to extract spectral information and vegetation indices for crop classification. • Random Forest (RF) and Support Vector Machines (SVM) classifiers used to classify the crops and estimate their yield.	Accurate and efficient crop monitoring and yield estimation, which can aid in decision-making for agricultural practices and food security in the Great Rift Valley and other regions with similar characteristics.
Multispectral Crop Yield Prediction Using 3D-Convolutional Neural Networks and Attention Convolutional LSTM Approaches **Authors:** S. M. M. Nejad, D. Abbasi-Moghadam, A. Sharifi, N. Farmonov, K. Amankulova and M. Lászlź **Journal:** IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16	• The method involves the use of 3D-CNNs to extract features. The 3D-CNNs are used to extract features from the multispectral images, which are then fed into the AC-LSTM. • The AC-LSTM captures the temporal dependencies in the data by using attention mechanisms to weight the input at each time step.	This technique has been found to outperform traditional methods of crop yield prediction that use only one or two spectral bands. Potential to improve the accuracy of crop yield prediction.
Exploiting Hierarchical Features for Crop Yield Prediction Based on 3-D Convolutional Neural Networks and Multikernel Gaussian Process **Authors:** M. Qiao et al. **Journal:** IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14	• The proposed approach uses 3D-Convolutional Neural Networks (CNNs) to extract hierarchical features. • Then a Multikernel Gaussian Process (MKGP) which is a non- parametric regression is used to model it. • R-squared value of 0.7 is achieved, indicating a strong correlation between the predicted and actual crop yield.	• The proposed approach can effectively capture the complex relationships between the multispectral data and the crop yield, leading to accurate and efficient crop yield prediction. • The method can aid in decision-making for agricultural practices and food security.

Crop yield estimation at field scales by assimilating time series of sentinel-2 data into a modified CASA-WOFOST Coupled model **Authors**: F. Ji, J. Meng, Z. Cheng, H. Fang and Y. Wang **Journal:** IEEE Transactions on Geoscience and Remote Sensing, vol. 60	• A coupled CASA-WOFOST integrated model is proposed to improve the precision, reliability and stability of crop yield estimation. • Ensemble Kalman Filter (EnKF) is used for data assimilation in 2 steps : forecasting and updating. • R2, RMSE, NRMSE, NSE, absolute error and relative error results to evaluate the performance of the models in yield simulation.	The coupled model was better than the individual CASA and WOFOST models based on almost all evaluation metrics.
Prediction of land suitability for crop cultivation based on soil and environmental characteristics using modified recursive feature elimination technique with various classifiers. **Authors:** G. Mariammal, A. Suruliandi, S. P. Raja and E. Poongothai **Journal**: IEEE Transactions on Computational Social Systems, vol. 8	• A new mode of feature selection : Modified recursive feature elimination is proposed which helps to select and rank features while the bagging technique helps accurately predict a suitable crop for the given conditions. • Performance is evaluated using precision metrics like accuracy, F1 score, etc. and the dataset containing the soil and environmental features is preprocessed to remove redundant data.	Training samples are trained with the classifier and unknown samples provided to validate the trained classifier. A major breakthrough in the algorithm is the recursive feature used to eliminate the redundant fields in the dataset.

ARCHITECTURE/TOPOLOGICAL DIAGRAM:



TOOLS:

AccuWeather tracking website:

AccuWeather is a reliable and comprehensive weather forecasting service that provides users with the most up-to-date weather information. With its advanced forecasting techniques, AccuWeather is a valuable resource for anyone looking to stay informed about the weather.

TensorBoard:

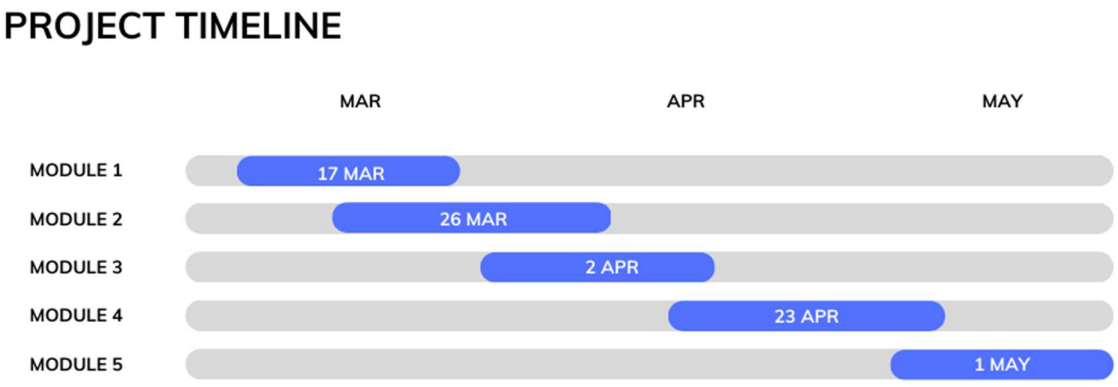
TensorBoard is a visualization tool included in the TensorFlow library. It allows users to visualize the performance of machine learning models during training, including metrics such as loss and accuracy. TensorBoard also provides tools for visualizing the structure of the model itself, such as the graph of operations and the distribution of weights.

OpenStreetMaps API / Google Maps API:

OpenStreetMap (OSM) is a collaborative and free mapping project that provides an open-source and openly licensed database of geographic data. The OpenStreetMap API allows developers to access this data programmatically and use it in their own applications.The OpenStreetMap API provides several different endpoints that allow developers to query the database for different types of geographic data, such as street maps, building footprints, and points of interest.

Visual studio code:

In order to add dynamic and interactive behaviour to the application, scripting in python is required. The scripts are developed in Visual Studio code.



MODULES:

Data pre-processing:

This module will involve data extraction, cleaning, and preparation. The main tasks of this module are: data collection from multiple sources, data cleaning to remove duplicates, irrelevant data, and handle missing values, feature engineering to create new features or transform existing ones for better performance, data normalization or scaling to improve the model's convergence and performance, data splitting into training, validation, and testing sets

Model selection:

This module will involve selecting the top 5 ideal ML algorithms based on the problem domain and data characteristics. The main tasks of this module are: researching and selecting the top 5 ML algorithms, implementing and training the selected algorithms on the training dataset, analysing and comparing the performance of the algorithms using appropriate metrics, selecting the best two performing algorithms based on the analysis

Model integration:

This module will involve integrating the best two performing models into a new one. The main tasks of this module are: researching and selecting appropriate techniques to integrate the models, implementing and integrating the models using appropriate techniques, evaluating the performance of the integrated model

Model training:

This module will involve training the newly formed integrated model on the training dataset. The main tasks of this module are: Tuning hyperparameters of the integrated model, training the integrated model on the training dataset, analysing the performance of the integrated model on the validation dataset, repeating the previous two steps until satisfactory performance is achieved.

Application interface development:

This module will involve developing an application interface to provide crop estimation and profitability forecasting functionality. The main tasks of this module are: designing the user interface to provide a seamless experience for the end-users, integrating the trained ML model to the interface to provide crop estimation and and profitability forecasting functionality, testing the interface to ensure usability and functionality deploying the interface to the desired platform (e.g., web, desktop, mobile)

REFERENCES:

1. S. Vashisht, P. Kumar and M. C. Trivedi, "Improvised Extreme Learning Machine for Crop Yield Prediction," 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2022, pp. 754-757, doi: 10.1109/ICIEM54221.2022.9853054.
2. R. Luciani, G. Laneve and M. JahJah, "Agricultural Monitoring, an Automatic Procedure for Crop Mapping and Yield Estimation: The Great Rift Valley of Kenya Case," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 7, pp. 2196-2208, July 2019, doi: 10.1109/JSTARS.2019.2921437.

3. D. Elavarasan and P. M. D. Vincent, "Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications," in IEEE Access, vol. 8, pp. 86886-86901, 2020, doi: 10.1109/ACCESS.2020.2992480.
4. Y. Alebele et al., "Estimation of Crop Yield From Combined Optical and SAR Imagery Using Gaussian Kernel Regression," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 10520-10534, 2021, doi: 10.1109/JSTARS.2021.3118707.
5. S. M. M. Nejad, D. Abbasi-Moghadam, A. Sharifi, N. Farmonov, K. Amankulova and M. Lászlź, "Multispectral Crop Yield Prediction Using 3D-Convolutional Neural Networks and Attention Convolutional LSTM Approaches," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 254-266, 2023, doi: 10.1109/JSTARS.2022.3223423.
6. M. Qiao et al., "Exploiting Hierarchical Features for Crop Yield Prediction Based on 3-D Convolutional Neural Networks and Multikernel Gaussian Process," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 4476-4489, 2021, doi: 10.1109/JSTARS.2021.3073149.
7. F. Ji, J. Meng, Z. Cheng, H. Fang and Y. Wang, "Crop Yield Estimation at Field Scales by Assimilating Time Series of Sentinel-2 Data Into a Modified CASA-WOFOST Coupled Model," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-14, 2022, Art no. 4400914, doi: 10.1109/TGRS.2020.3047102.
8. G. Mariammal, A. Suruliandi, S. P. Raja and E. Poongothai, "Prediction of Land Suitability for Crop Cultivation Based on Soil and Environmental Characteristics Using Modified Recursive Feature Elimination Technique With Various Classifiers," in IEEE Transactions on Computational Social Systems, vol. 8, no. 5, pp. 1132-1142, Oct. 2021, doi: 10.1109/TCSS.2021.3074534.
9. M. Ishak, M. S. Rahaman and T. Mahmud, "FarmEasy: An Intelligent Platform to Empower Crops Prediction and Crops Marketing," 2021 13th International Conference on Information & Communication Technology and System (ICTS), Surabaya, Indonesia, 2021, pp. 224-229, doi: 10.1109/ICTS52701.2021.9608436.
10. B. V. B. Prabhu and M. Dakshayini, "Demand-prediction model for forecasting AGRI-needs of the society," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, India, 2017, pp. 430-435, doi: 10.1109/ICICI.2017.8365388.