

CROP YIELDING RATE PREDICTION AND ANALYSIS USING DEEP MACHINE LEARNING ALGORITHMS

¹A. GEETHA DEVI, ²T.V.K.P. PRASAD, ³K.VIDYA SAGAR, ⁴M. PREMA KUMAR, ⁵BALA PRASANTHI PAVULURI, ⁶LAKSHMI RAMANI BURRA, ⁷VEERA VASANTH RAO.B.

¹Department of ECE, PVP Siddhartha Institute of Technology, Vijayawada, India

²Department of CSE, SRKR Engineering College, Bhimavaram, India.

³Department of EIE, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad.

⁴Department of ECE, Shree Vishnu Engineering College for Women, Bhimavaram, AP, India.

⁵Department of ECE, RVR & JC College of Engineering, Guntur, A.P., India.

⁶Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India.

⁷Department of ECE, RVR & JC College of Engineering, Guntur, A.P., India.

Email: geetha.agd@gmail.com.

ABSTRACT

The world population rate is increasing. Production of food grains is alarming with rapid change of environmental conditions. Soil parameters are also much influencing on the crop yielding rate. Precision agriculture has much significance in meeting the demands. This paper aimed to predict the yielding rate by considering the crop disease, soil parameters considered are Nitrogen, potassium, phosphorous and environmental conditions considered are ambient temperature, PH and humidity. The disease of the crop is analyzed using machine learning algorithms. ENET regression algorithm, LASSO Regression algorithm, Kernel Ridge Algorithm and stacking algorithm have been considered. The trained data sets are applied to various machine learning algorithms and estimated the infected level of the leaf and eventually the yielding rate of the crop. The results achieved with ENET regression algorithm, LASSO Regression algorithm, Kernel Ridge Algorithm and stacking algorithm are compared to interpret the best fit algorithm for agricultural applications. Mean square error value is considered for comparison. ANET and LASSO algorithms considered maximum pixel values and neglected minimum pixels while predicting the yield rate and disease level over other algorithms. The results obtained with the above approaches are reliable.

Keywords: Crop Yield rate, ENET, LASSO, Kernel Ridge, Stacking, MSE

1. INTRODUCTION

Agriculture is a prominent era for survival of the society. The production of food granules is significant with growing population and to provide humanitarian support. With advancement of technology machine learning and deep machine learning approaches will significantly improve precision agriculture farming. Environment calamities are quite often adding stress on agriculture. Improving the production rate is challenging. Keeping the soil minerals at optimum levels and supervising rapid changes in environment may improve the production rate. Unhealthy state of the crop is seriously diminishing the growth of the plant and production rate. Estimating the healthy state of the crop is more significant to achieve good yield rate. Regression type, Naive bays and random forest type machine

learning algorithms are significant in data processing and prediction applications. This paper considered ENET, LASSO, KERNEL RIDGE and STACKING algorithms to estimate the yield rate with varying soil minerals and environmental parameters.

2. RELEVANT WORK

The writers Niketa Gndhi and Amiya Kumar in [1] Owaiz Petkar, Lisa J. Armstrong, and Tripathi a Support Vector Machine (SVM) is a approach for supervised machine learning. A number of several instances of its application in the agricultural industry. Tripathi described what SVM was utilised to request a decrease in precipitation for scenarios for climate change that reduce the Generalization is achievable with minimal error generalised effectiveness. SVM was utilised to predict to the supply and demand for pulpwood.

Additionally, SVM was used to shed light on crop response. By supplying data on trends connected to climatic conditions the agricultural feature contribution analysis yield forecasting for agricultural categorization datasets employing base support vector discretization Utilizing machinery. Hung informed Hung about the SVM use. Urban land use conversion model. The research revealed a connection between a number of characteristics and rural usage of urban land. SVM also submitted an application for the crop biophysical parameter assessment using the airborne hyper spectral measurements are used.

The writers of [2], Rakesh Kumar, MP Singh, JP Singh and Prabhat Kumar have determined that Indian Summer rains is crucial for agriculture. The relationship between summertime precipitation and production of agricultural products. This essay offers a historical examination of crop-climate relationships information on agriculture. According to correlation analysis, the monsoon precipitation, sea surface temperatures in the Pacific and Indian Oceans, and drawing sea level pressure directly affects India's crop projection. Results indicate that agricultural production at the state level Sub divisional monsoon rainfall data and statistics are congruent with the overall India outcome, with several exceptions. The application of crop sequencing technology enhances net Crop production over a season. It employs a technique known as Method of Crop Selection (CSM).

The authors of [3], Prof. D. S. Zingade, Nilesh Mehra, OmkarBuchade, Chandan Mehta and ShubhamGhodekar have ended. because no mechanism currently in place suggests a variety of variables, including nitrogen, Potassium and phosphorus nutrients in soil elements of the weather, such as temperature and rainfall. The suggested system calls for an android based program that accurately forecasts to the farmer, the most lucrative produce. Users' location is located using GPS assistance. Based on the user site, the appropriate crops that might be grown there is determined using the weather and soil database. These soils were compared to productivity from the previous year. To determine the present crop that will be the most lucrative location. Once this processing has been completed on the server, the output it forwarded to the user's Android app. The sole input for something like the extrapolation is location system. Depending on the various circumstances and more filters based on the user's needs the crop with the highest yield is recommended.

In [4], Ramesh Medar and Vijay S. Rajpurohit & Shwetahas came to the conclusion that it explore the numerous machine learning applications gives information on the past industry as well as the issues that farmers confront and the solutions that expanding the countries' farming sectors apply machine learning more broadly. The Artificial Neural Networks are the employed algorithms, Decision tree algorithm, Bayesian belief networks, analysis of clustering regression.

3. METHODOLOGY

This paper considered four different algorithms ENET, LASSO, KERNEL RIDGE, AND STACKING to estimate the crop yielding rate and the precise estimation of the decease rate by considering the soil parameters and environmental parameters. Four different crops are considered apple, grape, maize, and paddy

a) ENET Algorithm

Efficient neural network (ENET) based pixel -wise semantic segmentation is deliberated for Leaf analysis. This algorithm is best fitted to achieve reliable results with minimum FLOPS and minimum parameters than other procedures. To minimize the dimensionality 1X1 Projection is considered

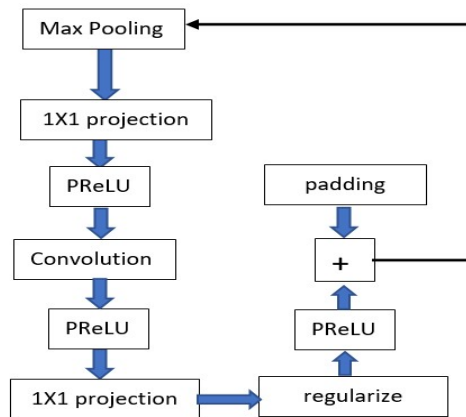


Fig.3.1. ENET Regression algorithm

A 3X3 convolution layer is considered. Parametric rectified linear unit (PReLU) is used between convolution layers for model fitting and for nonlinearities to train extremely deep model networks. Batch normalization is also considered between two convolution layers which impact on unstable regions. Data normalization is imperative before it applies to the next stage of the network.[5][6]. Batch normalization performs operations from the previous layer outputs and

normalizes the values using scaling and shifting operations. Batch normalization is done in three stages. 1. Standardization, 2. Normalization and rescaling of the data, 3. Max pooling layer is considered for down sampling and reconsider 1X1 projection with 2X2 convolution with stride 2. Activation and matching of features of an image is done with Zero pad. The bias terms are not used between the convolution owing to the implication of batch normalization. The network speed is increased with batch normalization.

b) Lasso Regression algorithm

Lasso is least absolute shrinkage and selection operator. This is labelled as L1 norm regularization. The loss function of linear regularization is used. The weights parameter absolute value labelled as penalty is added to this loss function. The result is then multiplied with regularization parameter value. The data points in a data set shrinkage towards the central point [7]. The coefficients with zero value are discarded from the model. Penalties with higher value results the coefficients may be zero.

$$\sum_{i=1}^n (Z_i - \sum_j y_{ij} \beta_j)^2 + \alpha \sum_{j=1}^p |\beta_j|$$

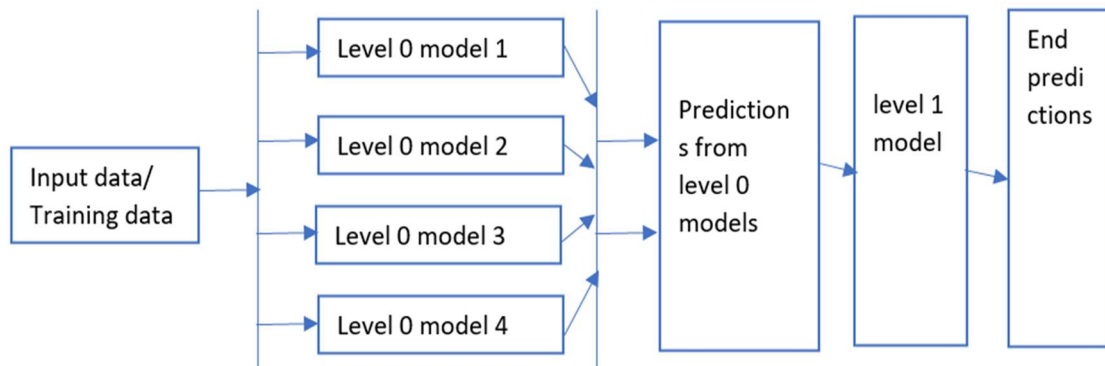


Fig.3.2. Stacking Algorithm

represents

Lasso is adopted for minimization of the squares and feature extraction. The minimization is done by adding the squares (β_j) with some constraint[8]. The data interpretation is simple because few squares are approaching to zero. The penalty strength is controlled with some tuning parameter ‘ α ’ which is the amount of shrinkage. If ‘ α ’ is equal to zero value, then the estimated value and linear aggression value are same then eliminating the parameters can be neglected. With increasing ‘ α ’ value, maximum number of coefficients are approaches to zero value, so the addition of bias value is depending on increasing the ‘ α ’ value, the increase of variance is depending on the decrease of ‘ α ’ value [9],[10],[11]. The decrease prediction is done using

the following equation. The coefficient of j-th term is represented as a_j .

$$\frac{1}{[2N_{training}]} \sum_{i=1}^{N_{training}} (X_{real}^i - X_{predicted}^i)^2 + \alpha \sum_{j=1}^n |a_j|$$

c) Kernel ridge regression (KRR)algorithm

Linear regression is no best to extract the features for data with unusual features. Kernel trick used to transform the leaf data sets into linear regression in kernel space[12]. Kernel ridge regression is a supervised learning approach. The features of the image is mapped at higher dimensional space. The relationship between the region of interest and the features of the image is linear [13]. The inner product of the features is significant in mapping process.

$$f(n) = \sum_{i=1}^{N_j} c_j q(n, n^{(j)})$$

Where $n(j)$ is the leaf feature descriptor for the jth configuration. Maximum configurations are represented with N_t .

d) Stacking Algorithm

Three level ‘0’ learner models are considered one level ‘1’ meta model merges the base model predictions. The input leaf image data is decomposed into ‘k’ number of folds. These folds represent training data [14][15]. The training data is used by level ‘0’ base models produce an output of compiled different predictions by different level ‘0’ models. The meta model received the data which is not considered by base models [16][17]. The level 1 meta model will train on the data which is not considered by the level 0 base models with various predictions. The meta model (level 1) analyses, learns and merge the best predictions[18][19]. The input data (predicted

outputs) and output data (expected data) formed as a pair to best fit the level 1 meta model.

4. RESULTS AND DISCUSSIONS

The work combines plant leaf diseased image dataset with Crop Yield dataset to predict future crop growth. The soil minerals of Potassium, Phosphorous and Nitrogen parameters are considered and environmental parameters such as Humidity, temperature and PH are considered. Using the input datasets (crop leaf images), Cumulative and average values are estimated from infected i.e., deceased leaves. Apple, grape, maize, and paddy leaf are considered in this paper.

ENET, LASSO, Kernel Ridge and Stacking machine learning algorithms are applied on the trained datasets for comparative study.

iterated and then kept bounding box across infected region of the input data image.

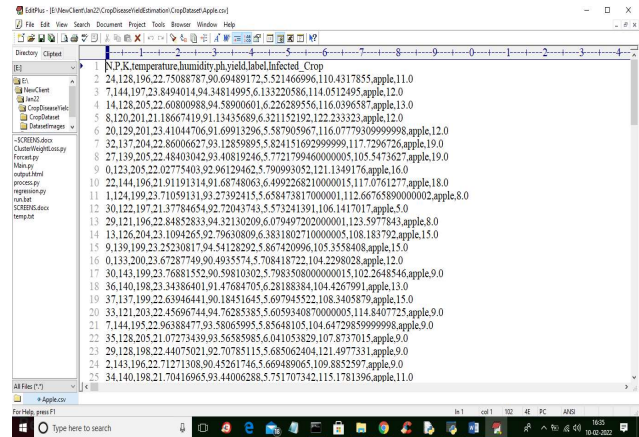


Fig.4.1. Crop Yield Rate Estimation With Soil Minerals

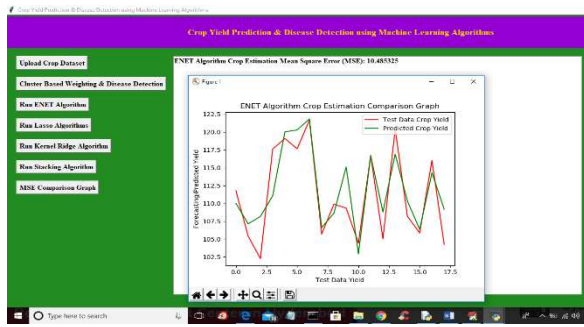
Table 4.1. Crop Yield Rate Estimation With Soil Minerals And Environmental Parameters

Nitrogen	Phosphorous	Potassium	Temperature	Humidity	PH	Yielding rate	Infected region
24	128	196	22.750	90.6948	5.52	110.431	11.0
7	144	197	23.84	94.34	6.133	114.05	12.0
14	128	205	22.60	94.5	6.227	116.03	13.0
8	120	201	21.1866	91.699	5.587	116.077	12.0
20	129	201.	21.18	91.699	5.58	117.72	12.0
32	137	204	22.8600	93.128	5.582	116.07	19.0
27	139	205	22.48	93.408	5.77	105.54	19.0

The mean square error is estimated. i.e., mean squared deviation of estimated value from the actual value

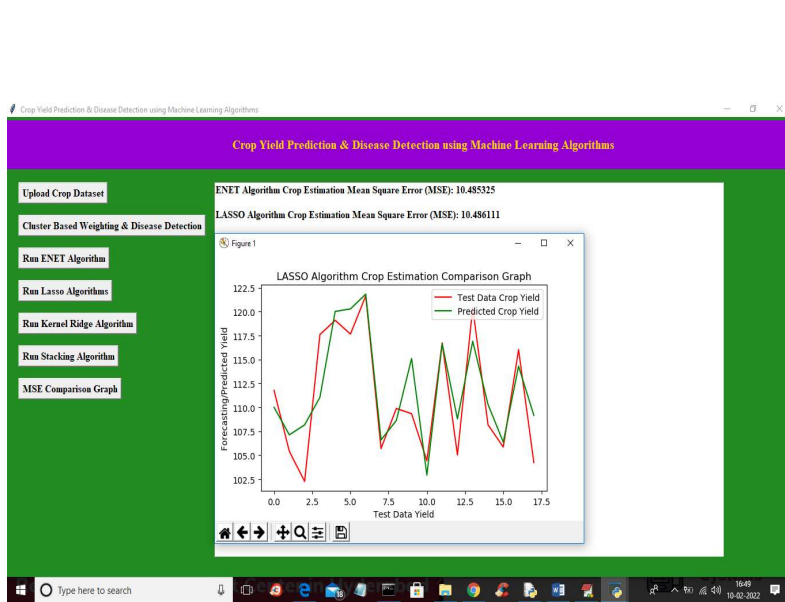
Mean squared deviation is considered for interpreting the unobserved quantity from the input leaves data set. Open CV with inbuilt API's used to cluster the leaf data with different colours. Infected regions of the data are appeared as grey colour and the healthy region of the leaf appeared as green colour. The data values are weighted to discriminate the unhealthy and healthy regions of a leaf dataset. The weight cluster values are estimated as infected black (dark grey) part of the image which considered as a deceased region of the image. Adoptive thresholding is applied to convert the input image into grey scale image. Each pixel value is evaluated to identify decease part and then add selected part to reweight. Each weight value is re-

In the above Fig.4.1. N refers to Nitrogen, K refers to potassium and P refers to phosphorus soil properties and remaining values are the normal environment data with infected and yield crop. The dataset is used to train all machine learning algorithms. Fig.4.2. Represents ENET algorithm achieved MSE value as 10.4853 and in graph red line refers to TEST Crop DATA and green line refers to predicted crop yield.



Test Data Yield	Predicted Yield
111.7763395	110.0079
105.40047509999998	107.13186
102.26485459999998	108.17088
117.61029149999999	111.04815
119.1025189	120.03182
117.66028270000001	120.29226
121.6622761	121.835396
105.6941544	106.599976
109.8852597	108.62111
109.33835520000002	115.10777
104.4267991	102.915794
116.7366261	116.66921
105.02413290000003	108.798256
120.4359949	116.91329
108.183792	110.31094
105.85543510000001	106.3612
116.0396587	114.30057
104.2298028	109.169044

Fig.4.2. Predicted crop yield data Vs Test data using ENET Algorithm



/output.html

Test Data Yield	Predicted Yield
111.7763395	110.00778
105.40047509999998	107.13201
102.26485459999998	108.171135
117.61029149999999	111.048355
119.1025189	120.03269
117.66028270000001	120.29319
121.6622761	121.836975
105.6941544	106.60014
109.8852597	108.62133
109.33835520000002	115.10788
104.4267991	102.91533
116.7366261	116.66918
105.02413290000003	108.797935
120.4359949	116.91325
108.183792	110.31101
105.85543510000001	106.360695
116.0396587	114.300095
104.2298028	109.16947

(a)

(b)

Fig.4.3. Test data and predicted data with LASSO algorithm

The predicted data is approximately following the trained data. So ENET is more reliable in plant leaf disease prediction. Fig.4.3. represents actual test data values and predicted values. From Fig.4.2 (b), first column represents TEST data values and second column represents predicted crop yielding rate in QUINTAL from ENET algorithm.

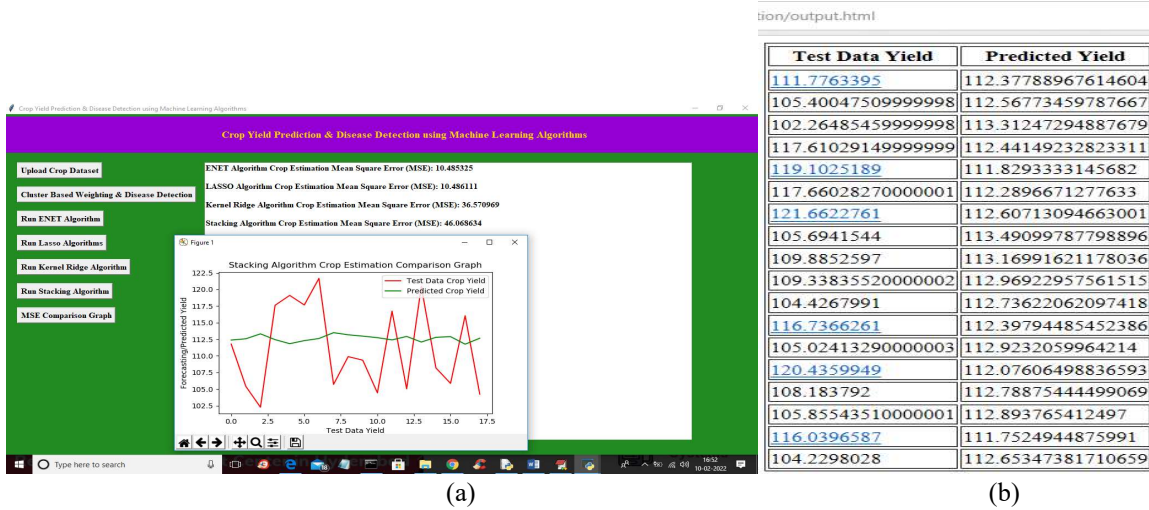


Fig.4.5. Predicted data Vs Test data with Stack Algorithm

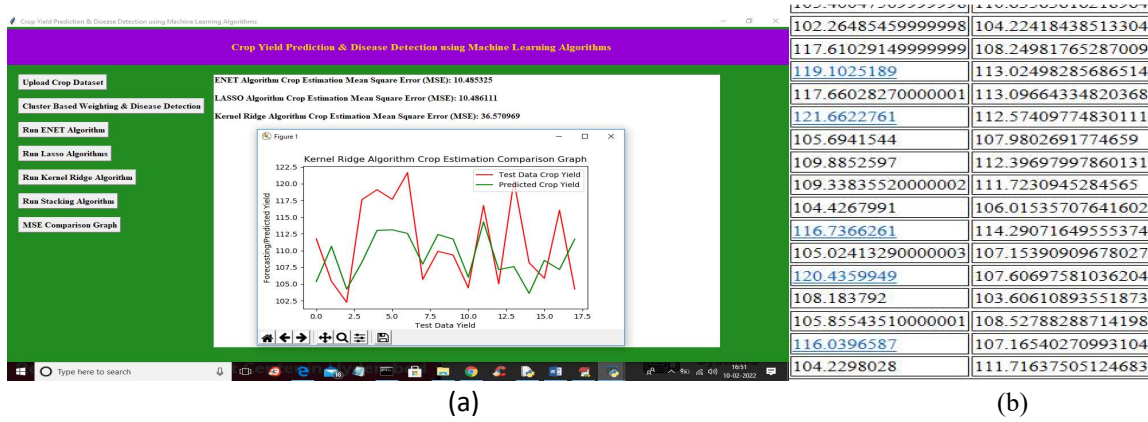


Fig.4.4. Predicted data Vs Test data with Kernel Ridge Algorithm

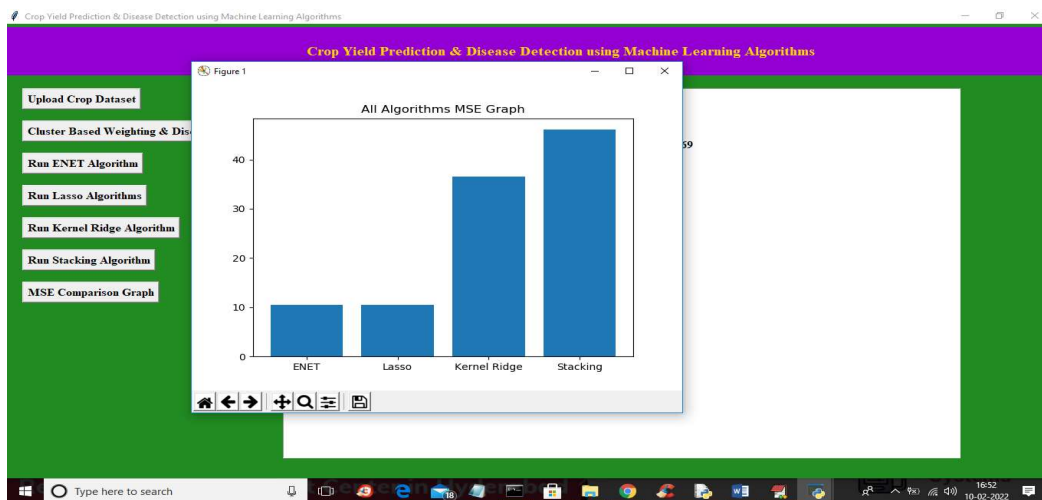


Fig.4.6. Comparison of MSE values with Enet, Lasso, Kernel ridge, And Stack Algorithms.

In Fig.4.10. X-axis represents ENET, LASSO, Kernel ridge, and Stack Algorithms and y-axis represents mean squared error values. ENET, and LASSO produced reliable results over Kernel edge and Stack algorithms.

Table 4.2. Predicted crop yield rate Vs Test data using different algorithms

Test data	ENET Predicted data	LASSO Predicted data	Kernel Edge predicted data	Stack Predicted data
111.776	110.007	110.007	105.349	112.377
105.400	107.131	107.132	110.635	112.567
102.264	108.170	108.171	104.224	113.312
117.610	111.048	111.048	108.249	112.441
119.102	120.031	120.032	113.024	111.829
117.660	120.292	120.293	113.096	112.289
121.662	121.835	121.836	112.574	112.607
105.694	106.599	106.600	107.980	113.490
109.885	108.621	108.621	112.396	113.169
109.338	115.107	115.107	111.723	112.969
104.426	102.915	102.915	106.015	112.736
116.736	116.669	116.669	114.290	112.397
105.024	108.798	108.797	107.153	112.923
120.435	116.913	116.913	107.606	112.076
108.183	110.310	110.360	103.606	112.788
105.855	106.361	106.360	108.527	112.893
116.039	114.300	114.300	107.165	111.752
104.229	109.1690	109.169	111.716	112.653

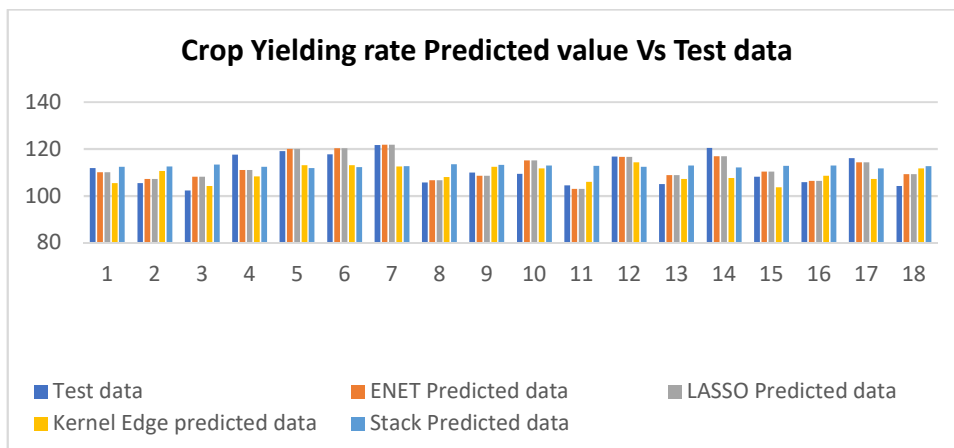


Fig.4.7. Predicted Crop Yield rate Vs test data with various algorithms.

The un-used data set values are minimum with ENET and LASSO algorithms. So that the crop yield rate prediction is reasonably good. Fig.4.7. Shows that Enet and lasso algorithms predicted values are similar to the test and trained data set value. The remaining Kernel edge and stacking algorithms are significantly deviated from the

trained data set values. Increasing the nitrogen, phosphorous, and potassium level of the soil results yielding rate increases. Nitrogen, Phosphorous and Potassium ratio 1:4:7 and humidity levels should be maintained at 90 percent produced maximum yield rate.

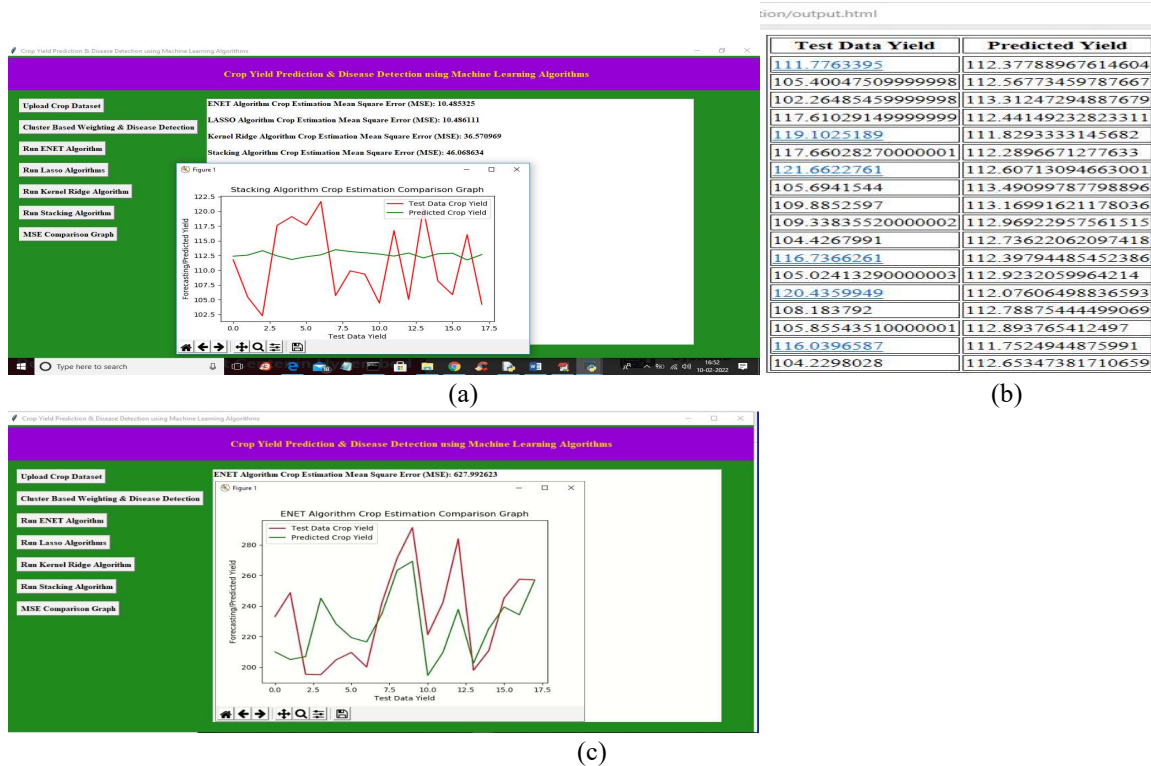


Fig.4.8. predicted Yield rate of paddy leaf Vs Test data using ENET Algorithm



Fig.4.9. Predicted Yield rate of paddy leaf Vs Test data Using Lasso Algorithm

With Nitrogen, Phosphorous and Potassium ratio 1:5:8 achieved very low yielding rate. With increasing potassium and phosphorous minerals of the soil and keeping the humidity levels at 90

percentage much influence on the yield rate of the crop than nitrogen mineral levels. The achieved results show that, this paper will add significant

support to the agriculture era to meet the constant demands from society.

Data set 2

The test image is extracted from paddy crop



Fig.4.10. Predicted Yield rate of paddy leaf Vs Test data using Kernel Ridge Algorithm

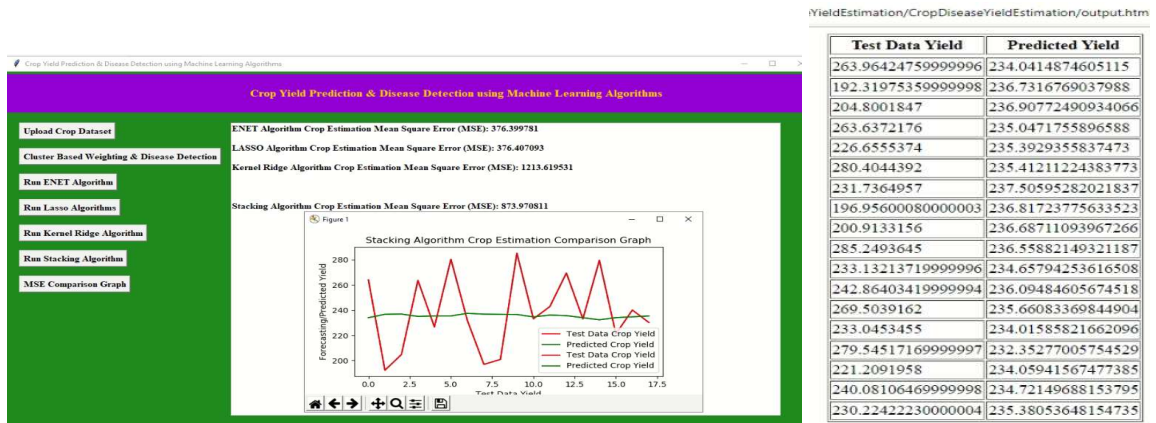


Fig.4.11 Predicted Yield rate of paddy leaf Vs Test data Using Stacking Algorithm

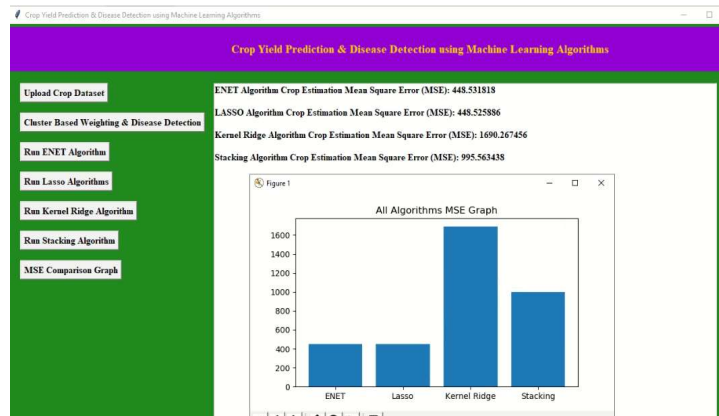


Fig.4.12. MSE Comparison with different algorithms.

Based on data set 1 and data set 2, Both ENET and LASSO produced significantly reliable results and similar mean squared error value is produced over other Kernel ridge and stacking algorithm. The stacking algorithm is not supporting to predict the yield rate of the crop.

CONCLUSIONS

This paper focused to estimate the crop yielding rate for apple, paddy leaves. The crop yielding rate is influenced by the soil parameters and environmental parameters and specifically these crops are very sensitive to harmful diseases impacted by insufficient soil minerals and abnormal environmental conditions. Predicting the yield rate and quantifying the amount of infection will add supportive mechanism to the agriculture era. Four machine learning algorithms are considered for meticulous estimation of the yielding rate and quantity of crop infection. ENET and LASSO algorithms produced favourably precise results over Kernel Ridge and stacking algorithms. Mean square error value is considered to estimate the performance of each algorithm. Mean square error will determine the unused pixel values of an image to determine the yielding rate and infected region of the crop image. Increasing the nitrogen, phosphorous, and potassium level of the soil results yielding rate increases. Nitrogen, Phosphorous and Potassium ratio 1:4:7 and humidity levels should be maintained at 90 percent produced maximum yield rate. With Nitrogen, Phosphorous and Potassium ratio 1:5:8 achieved very low yielding rate. With increasing potassium and phosphorous minerals of the soil and keeping the humidity levels at 90 percentage much influence on the yield rate of the crop than nitrogen mineral levels. The achieved results show that, this paper will add significant support to the agriculture era to meet the constant demands from society.

References

- [1]. Niketa Gandhi, OwaizPetkar, Leisa J Armstrong "Rice crop yield prediction using Support Vector Machines" 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development.
- [2]. J.P. Singh, M.P. Singh, Rakesh Kumar, and Prabhat Kumar "Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique", International Journal on Engineering Technology, May 2015.
- [3]. Prof. D. S. Zingade, OmkarBuchade, NileshMehra, ShubhamGhodekar, ChandanMehta "Crop Prediction system using machine Learning". International Research Journal of Engineering and Technology (IRJET), Vol8, Issue 7 July 2021.
- [4]. Ramesh Medar, Vijay S.Rajpurohit, Shweta "Crop yield Prediction using Machine Learning Techniques" 2019 IEEE 5th International Conference for Convergence in Technology (I2CT)
- [5]. S. S. Kale and P. S. Patil, "A Machine Learning Approach to Predict Crop Yield and Success Rate," 2019 IEEE Pune Section International Conference (PuneCon), Pune, India, 2019.
- [6]. Gour Hari Santra, Debahuti Mishra and Subhadra Mishra, Applications of Machine Learning Techniques in Agricultural Crop Production, Indian Journal of Science and Technology, October 2016
- [7]. Karan deep Kauri, Machine Learning: Applications in Indian Agriculture, International Journal of Advanced Research in Computer and Communication Engineering, April 2016.
- [8]. S. Djodiltachoumy, A Model for Prediction of Crop Yield, International Journal of Computational Intelligence and Informatics, March 2017..
- [9]. Nishit Jain, Amit Kumar, SahilGarud, Vishal Pradhan, Prajakta Kulkarni, Crop Selection Method Based on Various Environmental Factors Using Machine Learning, Feb -2017..
- [10]. D.Sindhura, B.Navaya Krishna, K.SaiPrasanna Lakshmi, B.Mallikarjun Rao, Dr. J Rajendra Prasad, Effects of Climate Changes on Agriculture International Journal of Advanced Research in Computer Science and Software Engineering, March 2016.
- [11]. Lohit V K, L. Vijayalakshmi, Brunda. G, Sanjay M D, Rashmi K T," Crop Yield Prediction using Machine Learning", International Journal of Engineering Research & Technology, Vol.10(12): pp 59-62,2022
- [12]. Madhuri Shripathi Rao, Arushi Singh, N.V. Subba Reddy, and Dinesh U Acharya," Crop prediction using machine learning", Journal of Physics: Conference Series: Conference series,2022
- [13]. Potnuru Sai Nishant, Pinapa Sai Venkat, Bollu Lakshmi Avinash, B. Jabber," Crop Yield Prediction based on Indian Agriculture using Machine Learning", International Conference for Emerging Technology, June,2020
- [14]. Surya Prasada Rao Borra et al.,"Pragmatic Investigations To Smart Dusts Location Appraisal Precisely Using Machine Learning", Journal of Theoretical and Applied

- Information Technology, 2023, 101(16), pp. 6301–6309.
- [15] Surya Prasada Rao et. al.,“Unhealthy Plant Region Detection in Plant Leaves Using Adaptive Genetic Algorithms”, Journal of Theoretical and Applied Information Technology, 2023, 101(8), pp. 3133–3137.
- [16] K. V. Sagar et. Al.,“Saline fluid flow Supervision in Intensive Care Unit using PrecisionAlgorithm”, Journal of Theoretical and Applied Information Technology, 2023, 101(23), pp.7769–7775.
- [17] A.Geetha devi, et.al “An Improved CHI 2 Feature Selection based a Two-Stage Prediction of Comorbid Cancer Patient Survivability”, Revue d’Intelligence Artificielle, 2023, 37(1), pp. 83–92.
- [18] V. S. M. , C. R. G. , R. V. D. , G. K. D. and P. K.M., "Automatic Fault Detection in Industrial Smart Grids Using KNN and Ensemble Classifiers", El-Cezeri, vol. 10, no. 2, pp. 240-252, May. 2023,
- [19] K. Phani Rama Krishna, Duvvuri Eswara Chaitanya, M. Prema Kumar, Venkata Bhujanga Rao Madamanchi, Balaji. Tata, Kurra Upendra Chowdary “Detection of Casing Laceration Using Dense CNN”, Journal of Theoretical and Applied Information Technology (JATIT) 31st December, 2023. Vol.101. No 24 pp 8277-8282.