Rice Crop Yield Prediction In India Using Improved Classification Technique

By
SUKHADIA KOLINBEN RAKESHKUMAR
170160702014
Semester IV, ME (Computer Engineering)

Guided By

Prof. M.B. Chaudhary

Department of Computer Engineering

A Thesis Submitted to
Gujarat Technological University in Partial Fulfilment of the Requirement for the Master of Engineering Degree in Computer Engineering

May 2019

Government Engineering College,

Modasa Shamlaji Road, Aravalli District, Modasa, Gujarat 383315

DECLARATION

I hereby declare that this M. E. thesis entitled "Rice Crop Yield Prediction In India Using Improved Classification Technique" was carried out by me for the degree of Master of Engineering in Computer Engineering under the guidance and supervision of Prof. M.B. Chaudhary,Head of Computer Engineering Department, Government Engineering College, Modasa. The interpretations put forth are based on my reading and un- derstanding of the original texts and they are not published anywhere in the form of books, monographs or articles. The other books, articles and websites, which I have made use of are acknowledged at the respective place in the text. For the present thesis, which I am submit- ting to the University, no degree or diploma or distinction has been conferred on me before, either in this or in any other university.

Place: Modasa
Date: 29/4/2019

Signature of Student:

Name of Student: SUKHADIA KOLINBEN RAKESHKUMAR
Enrollment No: 170160702014

i

CERTIFICATE

This is to certify that research work embodied in this report titled **"Rice Crop Yield Prediction In India Using Improved Classification Technique"** was carried out by **SUKHADIA KOLINBEN RAKESHKUMAR (Enrollment No. 170160702014)** at **Government Engineering College, Modasa (Institute Code. 016)** for partial fulfillment of Master of Engineering degree in Computer Engineering to be awarded by Gujarat Technological University. This research work has been carried out under my guidance and supervision.

Place: Modasa

Date: 29/4/2019

Prof. M.B. Chaudhary
(Internal Guide)
Head of Department
Computer Engineering
GEC, Modasa

Prof. M. B. Chaudhari
Head of Department
Computer Engineering
GEC, Modasa

GOVT. ENGG. COLLEGE-MODASA
GUJARAT-INDIA

Dr. A. M. Prabhakar
Principal
GEC, Modasa

PRINCIPAL
Govt. Engg. College
Modasa-383315

Seal of Institute

ii

COMPLIANCE CERTIFICATE

This is to certify that research work embodied in this dissertation titled **"Rice Crop Yield Prediction In India Using Improved Classification Technique"** was carried out by **SUKHADIA KOLINBEN RAKESHKUMAR (Enrollment No. 170160702014)** at **Government Engineering College, Modasa (Institute Code. 016)** for partial fulfillment of Master of Engineering degree to be awarded by Gujarat Technological University. She has complied to the comments given by the Dissertation phase-I as well as Mid Semester Thesis Reviewer.

Place: Modasa

Date: 29/4/2019

Signature of Student:
Name of Student:
SUKHADIA KOLINBEN RAKESHKUMAR
Enrollment No: 170160702014

Signature of Guide:
Name of Guide:
Prof. M.B. Chaudhari
Institute Code: 016

iii

PAPER PUBLICATION CERTIFICATE

This is to certify that research work embodied in this dissertation titled **"Rice Crop Yield Prediction In India Using Improved Classification Technique"** was carried out by **SUKHADIA KOLINBEN RAKESHKUMAR (Enrollment No. 170160702014)** at Government Engineering College, Modasa (Institute Code. 016) for partial fulfillment of Master of Engineering degree to be awarded by Gujarat Technological University has published article "**A Survey on Rice Crop Yield Prediction In India Using Improved Classification Technique**" for publication by the International Journal of Scientific Research in Computer Science ,Engineering and Information Technology (ISSN: 2456-3307), Vol.5, Issue.1, February 2019.

Place: Modasa

Date: 29/4/2019

Signature of Student:
Name of Student:
SUKHADIA KOLINBEN RAKESHKUMAR
Enrollment No: 170160702014

Signature of Guide:
Name of Guide:
Prof. M.B. Chaudhari
Institute Code: 016

GOVT. ENGG. COLLEGE-MODASA GUJARAT-INDIA

Dr. A. M. Prabhakar
Principal
GEC, Modasa

PRINCIPAL
Govt. Engg. College
Modasa-383315

Seal of Institute

iv

THESIS APPROVAL CERTIFICATE

This is to certify that research work embodied in this dissertation titled **"Rice Crop Yield Prediction In India Using Improved Classification Technique"** was carried out by **SUKHADIA KOLINBEN RAKESHKUMAR (Enrollment No. 170160702014)** at Government Engineering College, Modasa (Institute Code. 016) is approved for the degree of Master of Engineering in Computer Engineering by Gujarat Technological University.

Place:

Date:

Examiners Sign and Name:

____________________ ____________________

DECLARATION OF ORIGINALITY

We hereby certify that we are the sole authors of this thesis and that neither any part of this thesis nor the whole of the thesis has been submitted for a degree to any other University or Institution.

We certify that, to the best of our knowledge, the current thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations or any other material from the work of other people included in our thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that we have included copyrighted material that surpasses the boundary of fair dealing within the meaning of the Indian Copyright (Amendment) Act 2012, we certify that we have obtained a written permission from the copyright owner(s) to include such material(s) in the current thesis and have included copies of such copyright clearances to our appendix. We declare that this is a true copy of the thesis, including any final revisions, as approved by thesis review committee.

We have checked write up of the present thesis using the anti-plagiarism database and it is in the allowable limit. Even though later on in case of any complaint pertaining of plagiarism, we are solely responsible for the same and we understand that as per UGC norms, University can even revoke Master of Engineering degree conferred to the student submitting this thesis.

Date: 29/4/2019

Signature of Student:
Name of Student:
SUKHADIA KOLINBEN RAKESHKUMAR
Enrollment No: 170160702014

Signature of Guide:
Name of Guide:
Prof. M.B.Chaudhari
Institute Code: 016

vi

ACKNOWLEDGEMENT

Apart from the efforts of me, the success of any task depends mostly on the encouragement and guidelines of the many others. I take this opportunity to express my sincere gratitude and indebtedness to all the people whose help was very useful to complete my thesis work successfully. This work would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

A sincere thanks to my respected guide and Head of Computer Engineering Department, Prof. M.B Chaudhary for providing me with much needed support and guidance during dissertation phase and pursue it to completion with utmost sincerity. I can't say thank you enough for his tremendous support and help. I feel motivated and encouraged every time I attend his meeting.

And eventually however not least, I might be grateful towards my parents and friends who had supported plenty and provided inspiration and motivation to travel ahead during this project.

-SUKHADIA KOLINBEN RAKESHKUMAR

List of Figures

List Of Tables

TABLE OF CONTENT

Appendix

Appendix A

Appendix B

Appendix C

ABSTRACT

Rice crop production adds to the nourishment security of India, over 40% to all around yield prediction. Its production is dependent on climatic conditions. Changeability from season to season is detrimental to the rancher's salary and livelihoods. Improving the capability of ranchers to predict crop productivity under different climatic situations like vapour pressure rainfall,temperature,cloud cover etc. can assist ranchers and other partners in making important decisions in terms of agronomy and harvest.Agriculture is the single most significant contributor to the Indian economy.The prediction of agricultural yield is a difficult and desirable task for every nation.There are various classification techniques Like Support Vector Machine(SVM),Naïve Bayes,BayesNet,K Nearest Neighbour(KNN),Locally Weighted Learning(LWL),LADTree on rice crop production datasets. This work aimed to increase accuracy of prediction of rice crop yield using improved classification technique.

In our research we have proposed boosting technique(AdaBoost) with BayesNetwork Classification Technique to better prediction of classification accuracy and reduce errors like,RMSE(Root Mean Square Error),RAE(Relative Absolute Error),MAE(Mean Absolute Error).The dataset considered for the rice crop yield prediction is taken from publicly available Indian Government Records.The Parameters considered for the study are Rainfall, reference crop evapotranspiration,minimum temperature,average temperature,maximum temperature,cloud cover,vapour pressure,potential evapotranspiration for the kharif season.

CHAPTER 1

INTRODUCTION

OVERVIEW

Rice crop generation adds to the nourishment security of India, over 40% to by and large yield prediction.Its creation is dependent on great climatic conditions. Improving the ability of farmers to predict crop efficiency in under various climatic situations like rainfall,temperature,cloud cover,vapour pressure and so on can help ranchers and different partners in settling on critical choices regarding agronomy and harvest decision.The prediction of agricultural yield is a difficult and desirable task for every nation.There are many classification techniques like Support Vector Machine(SVM),Naïve Bayes,BayesNet,K Nearest Neighbour(KNN),Locally Weighted Learning(LWL),LADTree on rice crop production datasets. The provision of accurate and timely information can assist farmers to make the best decision for their cropping situations This could benefit them to attain greater crop productivity. This study examines the application of machine learning for the prediction of rice crop yield.The current study use the most rice producing states(Haryana,Punjab,Tamilnadu,Andhra Pradesh,Bihar).The dataset parameters are state name,district name,crop year,Rainfall,maximum temperature,minimum temperature,average temperature,reference crop evapotranspiration,vapour pressure,potproentialevapotranspiration,diurnal temperature range,area,production,yield,class. The dataset was then sorted on the basis of yield to classify the records in categories.

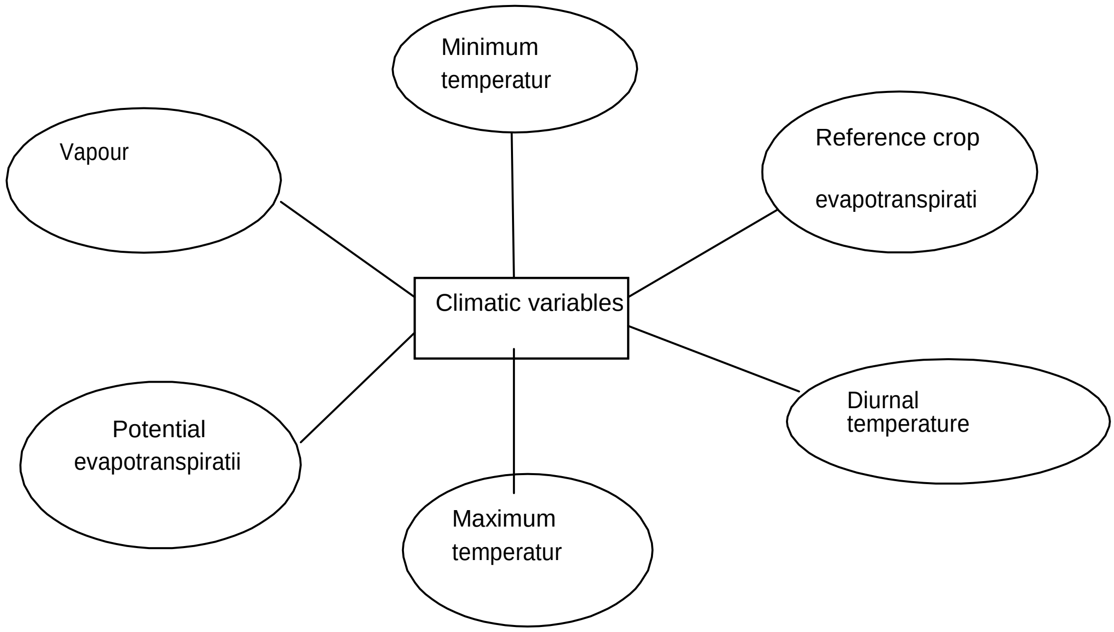


Figure 1.1 : climatic variable of dataset

Applications :

To better predict classification accuracy of rice crop yield using classification technique.

To help farmers to predict rice crop yield in future year by applying machine learning classification technique to historical data[1].

Farmers will be able to knowing how much rice crop yield he is about to expect.

Challenges :

It is challenging that to predict rice crop yield in kharif season in India in earliest time climate was very healthy everything was happened on time. But now most of the things have been changed due to global warming and many other climatic variables. The main problem with rice crop yield in India is lack of rainfall in seasonal time.Cloud cover is also converts as drawback.Many other climatic variables like vapour pressure,ground frost frequency has also affect rice crop yield in India.So this is more difficult to predict better accuracy by data mining classification technique.It is challenging to choose appropriate classification method for prediction of rice crop yield.

Motivation :

Rice crop yield prediction is a very issue in agricultural. Any farmer want to know how much crop yield he expect. In the past, yield prediction was performed by considering farmer's experience on particular field. The rice yield prediction is a major issue that remains to be solved based on available rice production data. classifications techniques like naïve bayes,bayes network,support vector machine,locally weighted learning,Ladtree are the better choice for this purpose. Different Data Mining classifications techniques are used and evaluated in agriculture for estimating the future year's rice crop production. This research proposes and implements a system to predict rice crop yield from previous rice crop production data through improved classification technique.

Objective :

The objective of this research is to bettere predicted accuracy of rice crop yield.so this will be helpful to farmers. Any farmer is keen on realizing how much rice yield he is going to anticipate.Each rancher needs to think about the measure of rice crop yield his homestead will deliver to get the assessments about his profit.so for that to estimate better estimation , various classification techniques are applied on rice crop production dataset.This is improved by boosting technique to better predict rice crop yield.

Organization of the Report :

The report is organized as follows: Chapter 2 contains the literature review.Existing methodology are discussed and their respective results as per experiments are shown in chapter

3. Proposed methodology is discussed with block diagram in chapter 4 which is followed by chapter 5 in which experimental work and result analysis is there.conclusion is in chapter 6

The papers are concluded in chapter 7.

CHAPTER 2

LITERATURE REVIEW

Introduction

Horticulture is the most essential area that impacts the economy of India.Rice crop generation adds to the nourishment security of India, over 40% to by and large yield prediction.Its creation is dependent on great climatic conditions. Improving the ability of farmers to predict crop efficiency in under various climatic situations like rainfall,temperature,cloud cover,vapour pressure and so on can help ranchers and different partners in settling on critical choices regarding agronomy and harvest decision.The prediction of agricultural yield is a difficult and desirable task for every nation.There are many classification techniques like Support Vector Machine(SVM),Naïve Bayes,K Nearest Neighbour(KNN),BayesNet,Locally Weighted Learning(LWL),LADTree on rice crop production datasets. The arrangement of exact and convenient data can help farmers to make the best choice for their cropping circumstances This could profit them to accomplish more noteworthy harvest profitability. This examination analyzes the use of data mining for the forecast of rice crop yield.

Overview on Classification Techniques

Bayesnet Classification technique[18]:

graphical model for representational process probabilistic connections among a variables.BN Encodes the conditional independence relationships between the variables in the graph structure. Provides a compact representation of the joint probability distribution over the variables. [18]

Bayesian Network constitutes of

- Directed Acyclic Graph (DAG).
- Set of conditional probability tables for each node in the graph.

A	P(A)
false	0.6
true	0.4

Table 2.1: conditional probability table node A[18]

A	B	P(B|A)
false	false	0.01
false	true	0.99
true	false	0.7
true	true	0.3

Table 2.2: Conditional Probability Table node B[18]

B	D	P(D|B)
false	false	0.02
false	true	0.98
true	false	0.05
true	true	0.95

Table2.3 : Conditional Probability table node D[18]

B	C	P(C|B)
false	false	0.4
false	true	0.6
true	false	0.9
true	true	0.1

Table2.4 : Conditional Probability Table node C[18]

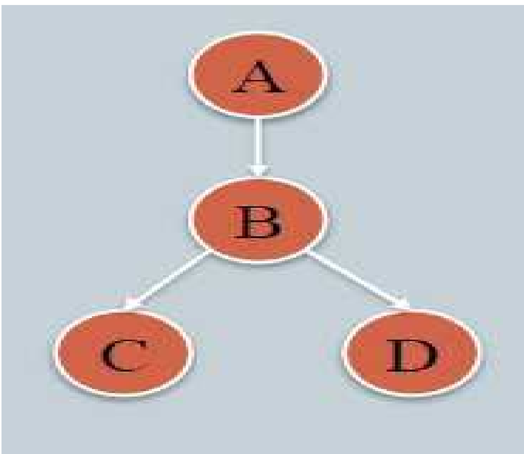


Figure 2.1 : Directed Acyclic Graph [18]

- So BN = (DAG, CPD).
- DAG: directed acyclic graph (BN’s structure)

  Nodes are random variables

  Arcs: indicate probabilistic dependencies between nodes (lack of link signifies conditional independence)
- CPD: conditional probability distribution at each node (BN’s parameters)

Inference in BN

To compute probabilities is called inference using bayesian network. inference involves queries of the form:

$$P(X \mid E)$$

where X is the query variable and E is the evidence variable.

Search procedure for learning structure of bayes network

Hill Climbing Search technique for searching bayesnetwork

- Initialization : some initial structure
- At each search step
- Apply the operator that more increase the score.
- Stop Searching :
- When there is no more improvement of score.
- When it is no possible for add new arc.
- Require : A B-space(S,O) where S is space of possible DAG(Directed Acyclic Graph) and O={addarc,deletearc,reversearc}.
- Initial structure s.
- Dataset D.example of set x={x1,x2,x3,x4} of random variables for domain under study.a scoring function score (s,D).
- Bayesian network structure s with high value of the score.
- While continue do
- compute score (S,D)
- Find best operator op such that op = max score(op(s),D).
- If op exists such that ,score(op(s),D)>score(s,D) then
- s – op(s)……… apply the operator to current structure.
- else
- continue false
- end while
- return s {a structure with high score}

Naïve Bayes Classification technique[19] :

Naive Bayes is a straightforward procedure for developing classifiers: models that relegate class names to problem instances. all Naive Bayes classifiers expect that the estimation of a specific feature is independent of the estimation of some other element, given the class variable.Based on Bayesian theorem the Naive Bayes classification technique is developed. When the value of inputs is very high, this technique is most suitable. Simple Bayes or Idiot Bayes are the other names of Bayes classifiers. The main advantage is that is fast to train and classify. Moreover Naive Bayesian is not sensitive to irrelevant features. Its basis is real and discrete data and also manages streaming data also.[19]

Support Vector Machine classification technique[15]:

- SVM is a learning algorithm to handle prediction and pattern recognition problem as well as for analysis and mapping of both linear and non-linear functions.
- It constructs a hyperplane or set of hyperplanes(classes) in a high dimensional space which can then be used for classification.[15]

It classifies object into categories (above or lower plane) based on the features of the object and using Kernel techniques it can transform nonlinear to linear before partitioning.

- SVM are one of the most popular classification algorithms for handling high dimensional data.This is a supervised learning.

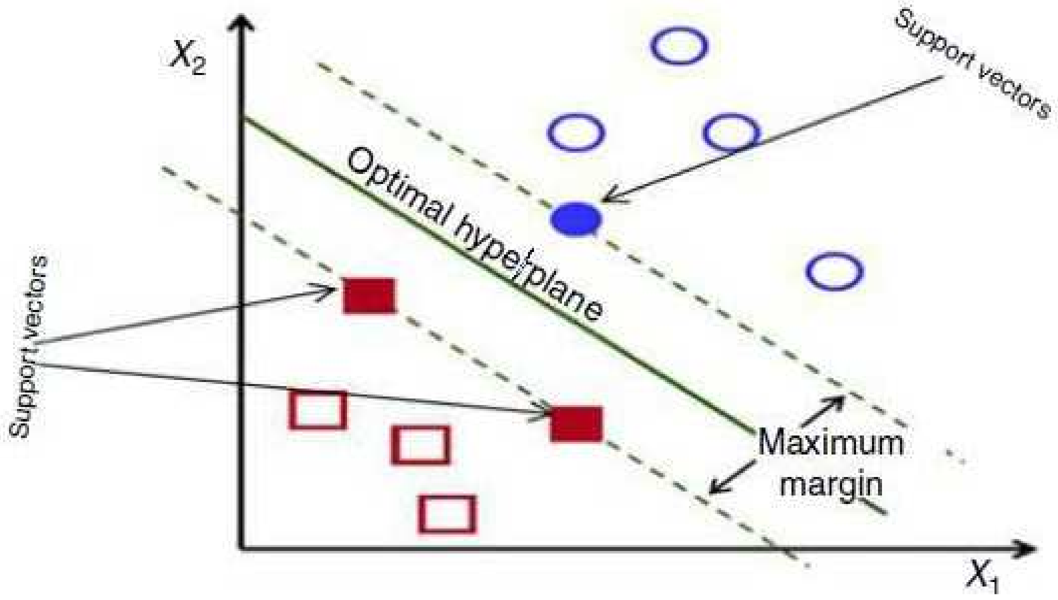


Figure 2.2: Support Vector Machine[15]

K-Nearest Neighbour Classification Technique[16] :

K nearest neighbors is a basic calculation that stores every single accessible case and arranges new cases dependent on a comparability measure (e.g., separate functions).A case is grouped by a dominant part vote of its neighbors, with the case being allocated to the class most regular among its K nearest neighbors estimated by a distance function. Get the test instance whose x value is xt and have to predict yt.Find the training example(x1,y1) whose x value x1 that is closest to xt then predict y1 as output yt.this is singe nearest neighbour. In case of k-nearest neighbour , find k training examples (x1,y1),(x2,y2),(x3,y3) which are closest to xt, then predict the majority class from {y1,y2,y3} as the output yt.[16]

Boosting Technique :

Boosting boosts the performance of the weak classifier to a strong level. It generates sequential learning classifiers using resampling (reweighting) the data instances. Initially equal uniform weights are assigned to all the instances.[17]

During each learning phase a new hypothesis is learned and the instances are reweighted such that correctly classified instance having lower weight and system can concentrates on instances that have not been correctly classified during this phase having higher weights. It selects the wrongly classified instance, so that they can be classified correctly during the next learning step. This process continuous tills the last classifier construction.[17]

Finally the results of all the classifiers are combined using majority voting to find the final prediction. AdaBoost is a more general version of the Boosting algorithm.[17]

AdaBoost :

Consolidating with numerous other learning algorithms, the meta algorithm is called as AdaBoost algorithm. This would improve the execution of classification. AdaBoost utilizes the settled administrator and it has a sub procedure.The better model is generated by sub-processor . More than one classifier is created by model. and generates a better model.By creating more than one classifier,the ac curacy of classification is increased. This model leads to decision making by combining the results of their classification techniques.The accuracy of the given algorithm is improved by boosting method. combining with numerous other algorithms, the meta algorithm is called as AdaBoost algorithm.This would improve the performance of classification. AdaBoost uses sequential process.The better model is generated by sub-processor. By creating more than one classifier, the accuracy of classification is improved by the ensemble model. This model leads to decision making by combining the results of their classification techniques.

adaptably change the distribution of training instances.

Instances of training data (x1,y1),(x2,y2),(x3,y3)….(xn,ym) m training sampels.[18]

Each instance of traning data has equal probability 1/m.

No of learning rounds = T

Initialize D0(i)=1/m   where i=1 to m

Then update D1(i).

increase D1(i) ,For those training examples for which learner classifies them wrongly.

From D1(i) resample the S2 and on S2 apply weak learner.

Now in this way update D2(i),for those training examples for which learner classify them wrongly.[18]

€1 is error over sample1,likewise €2 is error over sample2……

Alpha1,alpha2,alpha3….alpha4 are strengh to each learner.[18]

final classifier is linear combination of these different hypothesis.[18]

Research Paper Literature Review

1) Paper 1: Rice Yield Prediction Model Using Data Mining

Author : Umid Kumar Dey, Abdullah Hasan Masud, Mohammed Nazim Uddin
Publication : International Conference on Electrical, Computer and Communication Engineering(ECCE),IEEE,2017

summary :

This research does to predict the yield of rice in the regions of Bangladesh during the aforementioned seasons using Multiple Linear Regression AdaBoost (Adaptive Boosting), Support Vector Machine Regression and a Modified Nonlinear Regression (MNR) equation and then comparing the modified equation with the other three methods to check its accuracy For Aus and Boro Rice our modified Nonlinear Regression equation not only yields better RMSE, MSE and MAE values, it also has the highest Rsquare value, which proves that the MNR equation is the best fit for this study. this study also proves that weather conditions play very vital role in predicting the yield.

2) Paper 2: Application Of Data amaining Techniques For Predicting Rice Crop Yield Semi-Arid Climate Zone Of India

Author : Niketa Gandhi,Leisa J. Armstrong,Manisha Nandawadekar

Publication : International Conference on Technological Innovations in ICT for agriculture and Rural Development,IEEE,2017

Summary :

The classifiers used for research are J48, LWL, LAD Tree and IBK. The experimental results showed that J48,LADTree achieved the highest accuracy,sensitivity and specificity.semi-arid climate is the climate of a region that receives precipitation below potential evapotranspiration.

3) paper 3 : PredictingRice Crop Yield Using Bayesian Networks

Author : Niketa Niketa Gandhi,Leisa J. Armstrong,Owaiz Petkar

Publication : 2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 21-24, 2016, Jaipur, India

Summary : The classifiers used in the study were BayesNet and NaiveBayes. The experimental results showed that the performance of BayesNet was much better compared with NaiveBayes for the dataset.

4) paper4 : Predicting Rice Yield For Bangladesh By Exploiting Weather Conditions

Author : Akter Hossain,Mohammed Nazim Uddin,Mohammad Arif Hossain,Yeong Min Jang

Publication : IEEE, 2017

Summary : The proposed approach builds a model to predict the weather parameters applying Neural Networks,then estimates the rice yields applying Support Vector Regression that uses as inputs predicted weather from NN as well as current agricultural data.

5)paper5 : A Model for Prediction of Crop Yield

Author : Manjula, S. Djodiltachoumy

Publication : International Journal of Computational Intelligence and Informatics, Vol. 6: No.4, March 2017

Summary : The pre-processed data was clustered using k-means clustering algorithm. The association rule mining process will apply on clustered data to find the rules. The training phase ends with number of generated rules. In the testing phase, the yield value is predicted based on the generated rules.

6) **paper 6 : Rice Crop Yield Prediction in India using Support Vector Machines**
Author : Niketa Gandhi, Leisa J. Armstrong, Owaiz Petkar, Amiya Kumar Tripathy
Publication : 2016 International Joint Conference on Computer Science and Software Engineering (JCSSE),IEEE,2016
Summary : The parameters like maximum temperature,yield,production,area,average temperature,reference crop evapotranspiration,minimum temperature,precipitation are used in this study for the Kharif season (June- November).experimental results obtained by applying SMO classifier using the WEKA tool on the dataset of 27 districts of Maharashtra state, India in this paper.

7) **paper 7 : Rice Crop Yield Prediction Using Artificial Neural Networks**
Author : Niketa Gandhi, Owaiz Petkar, Leisa J. Armstrong
Publication : 2016 IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development (TIAR 2016)
Summary : The parameters like production,area,precipitation, average temperature, reference crop evapotranspiration minimum temperature, maximum temperature and area, production are use in this study and yield for the Kharif season (June to November) for

the years 1998 to 2002. WEKA tool is used for experiment. A Multilayer Perceptron Neural Network was developed. Validation of data is done using cross validation method.

8) paper 8 : **Smart Farming System: Crop Yield Prediction Using Regression Techniques**

Author : Ayush Shah, Akash Dubey, Vishesh Hemnani, Divye Gala and D. R. Kalbande

Publication : Springer,2018

Summary : The proposed method uses yield and weather data collected from United States Department of Agriculture. The various parameters included in the dataset are humidity,yield temperature and rainfall. regression-based algorithms to predict the crop yield. These were multivariate polynomial regression, support vector machine regression and random forest.support vector machine regression to obtain the best possible results for predicting the crop yield.

9)paper 9 : Crop Production-Ensemble Machine Learning Model for Prediction

Author : Narayanan Balakrishnan, Dr.Govindarajan Muthukumarasamy

Publication : International Journal of Computer Science and Software Engineering (IJCSSE), Volume 5, Issue 7, July 2016

Summary : The classification methods in this paper are the Support VectorMachine (SVM) and Naive Bayes. In this paper, AdaSVM and AdaNaive are the proposed ensemble method. This ensemble model is compared to SVM and Naive Bayes methods. accuracy and the classification error are parameters for prediction of output .The discovering yields that AdaSVM and AdaNaive are pleasant than SVM and Naive Bayes for the dataset analysis..

10) **paper 10 : Adaptive boosting of weak regressors for forecasting of crop production considering climatic variability: An empirical assessment**

Author : Subhadra Mishra, Debahuti Mishra, Gour Hari Santra

Publication : Journal of King Saud University - Computer and Information Sciences

Summary : In this study, regression based adaptive boosting prediction model is presented, using the datasets of Kharif and Rabi seasons along with the climatic features of three districts belonging to Odisha located in India. In this study experiments are performed on various weak regressors, like: SVR regression,linear, ridge,lasso, proposes strong predictors by avoiding the shortcomings of individual weak regressors and propagating the benefits of AdaBoost to improve the predictive accuracy on learning problems.

11) **paper 11: Relationship between rice yield and climate variables in southwest Nigeria using multiple linear regression and support vector machine analysis.**

Author : Philip G. Oguntunde,Gunnar Lischeid,Ottfried Dietrich

Publication : August 2017,springer

Summary : relationships among climate variables and rice yield using various methods like support vector machine,multiple linear regression, principal component analysis in southwest Nigeria. The climate and yield data used was for a period of 36 years between 1980 and 2015

Performance Evaluation Metrics :

Accuracy: It is characterized as the general achievement rate of the classifier[1].

- Accuracy = (TP + TN)/(TP + FP + FN + TN)[1]
- Where , True Positive (TP) portrays the quantity of examples where system recognizes for a condition when it is truly present[1].
- True Negative (TN) delineates the quantity of instances where system does not identify a condition when it is absent[1].
- False Negative (FN) portrays the quantity of instances where system does not recognize a condition when really it is present[1].
- False Positive (FP) portrays the quantity of instances where system distinguishe a condition when it is truly absent[1].
- RMSE(Root Mean Square Error): It is characterized as the contrast between the values predicted by the model and the real values noted[1].
- MAE(Mean Absolute Error): It is another factor in insights which estimates the distinction between two continuous variables[1].
- RAE(Relative Absolute Error): This measure gives the total absolute error between the variables[1].

CHAPTER 3

EXISTING METHOD IMPLEMENTATION

Bayesnet method :

A Bayesian network (BN) is a graphical model for depicting probabilistic relationships among a set of variables. BN Encodes the conditional independence relationships between the variables in the graph structure. Provides a compact representation of the joint probability distribution over the variables A problem domain is modeled by a list of variables X1, …, Xn. Representation of Knowledge about the problem domain is by joint probability P(X1, …, Xn). Directed links represent causal direct influences .Each node has a conditional probability table quantifying the effects from the parents.[18]

Bayesian Network constitutes of

- Directed Acyclic Graph (DAG).
- Set of conditional probability tables for each node in the graph.
- So BN = (DAG, CPD).
- DAG: directed acyclic graph (BN's structure)
- random variables are considered as nodes.
- probabilistic dependencies between nodes is indicated by Arcs.
- CPD: conditional probability distribution (BN's parameters)

  Conditional probabilities at each node, usually stored as a table (conditional probability table, or CPT) [18]

Inference in BN

compute probabilities using a bayesian network is called inference. inference

inference involves queries of the form[18]:

$$P(X \mid E)$$

where X is the query variable and E is the evidence variable.

3.1.1 Existing method Implementation & Result

State_Name	District_Nam	MINIMUM	AVERAGE	maximum	diurnal te	REFERENC	Rainfall	vapour pr	potential	cloud cov	Yield	class
Andhra Pradesh	ANANTAPUR	21.9034	26.2516	30.6348	8.9425	4.604	144.8806	22.11683	5.723333	60.60217	2.527415	high
Andhra Pradesh	ANANTAPUR	21.3504	25.691	30.0762	8.9425	4.536	81.928	21.6475	5.763333	56.00367	2.765971	high
Andhra Pradesh	ANANTAPUR	21.1432	25.4834	29.8688	8.9425	4.512	112.027	21.732	5.693333	59.02767	2.949004	high
Andhra Pradesh	ANANTAPUR	21.481	25.8294	30.2116	8.9425	4.556	86.3842	21.9785	5.666667	62.21083	2.907993	high
Andhra Pradesh	ANANTAPUR	21.8222	26.1618	30.5454	8.9425	4.588	61.2062	21.98833	5.788333	55.69	2.282013	high
Andhra Pradesh	CHITTOOR	22.4848	27.2442	32.0336	9.467167	4.926	149.2236	23.98333	5.85	65.84567	2.509524	high
Andhra Pradesh	CHITTOOR	21.934	26.6734	31.4668	9.467167	4.858	110.2794	23.57233	5.916667	59.27067	2.010468	high
Andhra Pradesh	CHITTOOR	21.7932	26.5494	31.3428	9.467167	4.842	136.8546	23.59217	5.845	63.20617	2.641994	high
Andhra Pradesh	CHITTOOR	22.263	27.0172	31.8114	9.467167	4.9	123.132	23.84383	5.828333	66.093	2.268983	high
Andhra Pradesh	CHITTOOR	22.6244	27.3828	32.173	9.467167	4.944	102.5698	24.12183	5.91	62.18833	2.483985	high
Andhra Pradesh	EAST GODAV	25.7192	29.563	33.4408	7.838333	4.3	199.2762	28.81883	5.57	57.7235	1.508389	low
Andhra Pradesh	EAST GODAV	25.3562	29.2152	33.1088	7.921667	4.646	121.0654	28.06117	5.593333	55.70267	2.569357	high
Andhra Pradesh	EAST GODAV	24.9672	28.833	32.7242	7.9215	4.6	164.339	27.939	5.61	54.08333	2.691002	high
Andhra Pradesh	EAST GODAV	25.5004	29.359	33.2572	7.804	4.664	159.1428	28.85317	5.525	58.38883	2.626001	high
Andhra Pradesh	EAST GODAV	25.5366	29.4252	33.3504	7.975333	4.682	127.9404	28.6655	5.598333	56.80583	2.829002	high
Andhra Pradesh	GUNTUR	25.5474	29.9774	34.4428	8.9775	5.072	148.7	26.72633	5.991667	56.8535	3.087802	high
Andhra Pradesh	GUNTUR	24.806	29.2342	33.699	8.9815	4.978	82.9954	26.00217	6.023333	52.46017	3.20059	high
Andhra Pradesh	GUNTUR	24.5364	28.9636	33.4276	8.975833	4.956	116.7458	26.10217	5.976667	54.151	3.371002	high
Andhra Pradesh	GUNTUR	25.1218	29.5534	34.017	8.937167	5.016	120.0564	26.71017	5.93	57.829	3.431999	high
Andhra Pradesh	GUNTUR	25.2248	29.654	34.1166	8.9815	5.028	105.3838	26.386	6.001667	54.30383	3.186999	high
Andhra Pradesh	KRISHNA	25.5366	29.6964	33.8952	8.4605	4.886	145.9352	27.39533	5.765	56.54417	2.706197	high
Andhra Pradesh	KRISHNA	25.0298	29.1974	33.4032	8.4755	4.834	88.80933	26.69883	5.795	52.9145	2.997288	high
Andhra Pradesh	KRISHNA	24.5936	28.7476	32.9422	8.457667	4.778	162.6517	26.67367	5.768333	53.282	3.087999	high
Andhra Pradesh	KRISHNA	25.2374	29.4054	33.6002	8.404833	4.848	129.8522	27.42833	5.71	57.1235	2.914999	high

Figure 3.1: rice crop dataset

```
=== Summary ===

Correctly Classified Instances         482               94.5098 %
Incorrectly Classified Instances        28                5.4902 %
Kappa statistic                          0.8783
Mean absolute error                      0.0595
Root mean squared error                  0.2252
Relative absolute error                 13.3797 %
Root relative squared error             47.7648 %
Total Number of Instances              510

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.944    0.053    0.973      0.944    0.958      0.879    0.985     0.992     high
                 0.947    0.056    0.894      0.947    0.920      0.879    0.985     0.975     low
Weighted Avg.    0.945    0.054    0.947      0.945    0.945      0.879    0.985     0.986

=== Confusion Matrix ===

   a   b   <-- classified as
 321  19 |   a = high
   9 161 |   b = low
```

Figure 3.2 : bayesnet output

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         449               88.0392 %
Incorrectly Classified Instances        61               11.9608 %
Kappa statistic                          0.7206
Mean absolute error                      0.1196
Root mean squared error                  0.3458
Relative absolute error                 26.8972 %
Root relative squared error             73.3644 %
Total Number of Instances              510

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.947    0.253    0.882      0.947    0.913      0.725    0.847     0.871     high
                 0.747    0.053    0.876      0.747    0.806      0.725    0.847     0.739     low
Weighted Avg.    0.880    0.186    0.880      0.880    0.878      0.725    0.847     0.827

=== Confusion Matrix ===

   a   b   <-- classified as
 322  18 |   a = high
  43 127 |   b = low
```

Figure3.3 : support vector machine output

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         340               66.6667 %
Incorrectly Classified Instances       170               33.3333 %
Kappa statistic                          0
Mean absolute error                      0.4168
Root mean squared error                  0.4417
Relative absolute error                 93.7228 %
Root relative squared error             93.7051 %
Total Number of Instances              510

=== Confusion Matrix ===

   a   b   <-- classified as
 340   0 |   a = high
 170   0 |   b = low
```

Figure3.4 : KStar output

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         449               88.0392 %
Incorrectly Classified Instances        61               11.9608 %
Kappa statistic                          0.7198
Mean absolute error                      0.1018
Root mean squared error                  0.2737
Relative absolute error                 22.9033 %
Root relative squared error             58.0538 %
Total Number of Instances              510

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.950    0.259    0.880      0.950    0.914      0.725    0.940     0.954     high
                 0.741    0.050    0.881      0.741    0.805      0.725    0.940     0.844     low
Weighted Avg.    0.880    0.189    0.880      0.880    0.878      0.725    0.940     0.918

=== Confusion Matrix ===

   a   b   <-- classified as
 323  17 |   a = high
  44 126 |   b = low
```

Figure3.5 :KNN output

No.	Classifier	Classification accuracy
1	Support Vector Machine	88.0392
2	Kstar	66.6667
3	KNN	88.0392
4	Bayesnet	94.5098

```
package weka.api;
import weka.core.Instances;
import java.util.Random;
import weka.classifiers.bayes.net.search.local.HillClimber;
import weka.core.converters.ConverterUtils.DataSource;
import weka.classifiers.bayes.BayesNet;
import weka.core.converters.ArffSaver;
import weka.classifiers.Evaluation;
public class bayesnetimplemented {
    public static void main(String args[]) throws Exception{
        String data = "/C:/Users/LENOVO/Desktop/281118/61218/newdatasetcopy1.arff";
    DataSource source = new DataSource(data);



    Instances trainingData = source.getDataSet();

        ArffSaver arff = new ArffSaver();
        arff.setInstances(trainingData);
        if(trainingData.classIndex() == -1){
            trainingData.setClassIndex(trainingData.numAttributes()-1);
        }
        BayesNet bn = new BayesNet();
        bn.setSearchAlgorithm(new HillClimber());
        bn.buildClassifier(trainingData);
        Evaluation eval = new Evaluation(trainingData);
        Random rand = new Random(1);
        int folds = 10;
        eval.crossValidateModel(bn, trainingData, folds, rand);
```

```
        eval.crossValidateModel(bn, trainingData, folds, rand);
        System.out.println("correct = "+eval.pctCorrect());
        System.out.println("Incorrect = "+eval.pctIncorrect());
        System.out.println("MAE = "+eval.meanAbsoluteError());
        System.out.println("RMSE = "+eval.rootMeanSquaredError());
        System.out.println("RAE = "+eval.relativeAbsoluteError());
        System.out.println("RRSE = "+eval.rootRelativeSquaredError());
        System.out.println("kappa = "+eval.kappa());
        System.out.println("precision = "+eval.precision(1));
        System.out.println("Recall = "+eval.recall(1));
        System.out.println("fMeasure = "+eval.fMeasure(1));
        System.out.println("error rate = "+eval.errorRate());
    }

}
```

Figure3.6 : Bayesnet Implementation

CHAPTER 4

PROPOSED METHODOLOGY

4.1 Problem Statement

The rice crop production in agriculture is affected by several climate factors.Like (rainfall,temperature,reference crop evapotranspiration,potential evapotranspiration,diurnal temperature range,vapour pressure,cloud cover). farmer is interesting to know how much rice crop yield he is about to expect in future year.from that he get estimate about his earning.

The main problem with rice crop production in India is lack of rainfall in seasonal time. To overcome these above issues we need to develop a system which will able to find the hidden facts or results, patterns and insights. The farmer can predict yield so that that he/she can get more benefit.It is challenging to predict better rice crop yied in India.

Research Gap :

It is necessary to use efficient classification method to predict rice crop yield.

It is necessary to reduce classification errors such as Relative absolute error,root mean square error,mean square error.

It is necessary to improve performance by proposing efficient techique.

Proposed Flowchart

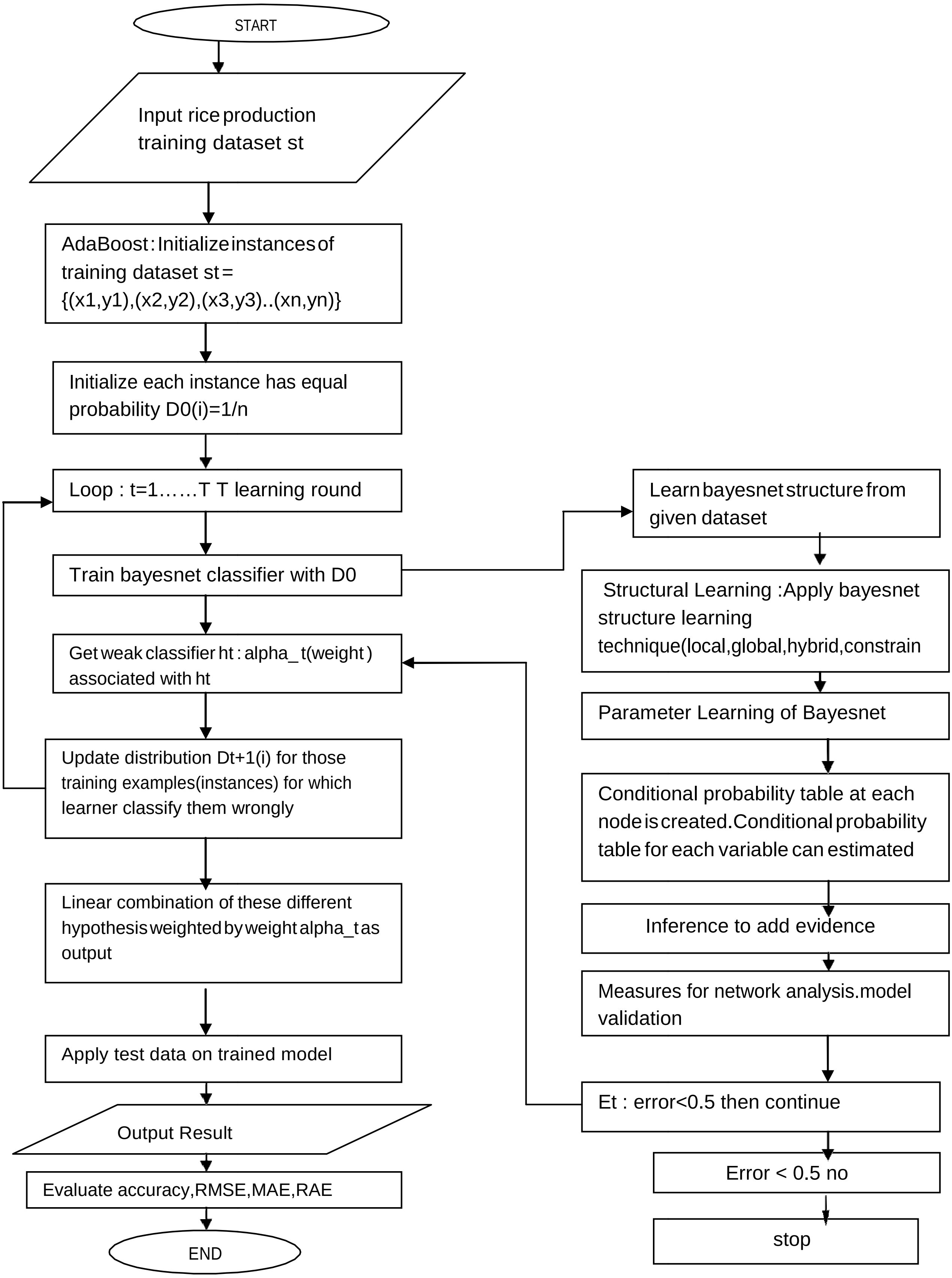

START
Input rice production training dataset st
AdaBoost: Initialize instances of training dataset st = {(x1,y1),(x2,y2),(x3,y3)..(xn,yn)}
Initialize each instance has equal probability D0(i)=1/n
Loop : t=1……T T learning round
Train bayesnet classifier with D0
Get weak classifier ht : alpha_ t(weight) associated with ht
Update distribution Dt+1(i) for those training examples(instances) for which learner classify them wrongly
Linear combination of these different hypothesis weighted by weight alpha_t as output
Apply test data on trained model
Output Result
Evaluate accuracy,RMSE,MAE,RAE
END
Learn bayesnet structure from given dataset
Structural Learning :Apply bayesnet structure learning technique(local,global,hybrid,constrain
Parameter Learning of Bayesnet
Conditional probability table at each node is created.Conditional probability table for each variable can estimated
Inference to add evidence
Measures for network analysis.model validation
Et : error<0.5 then continue
Error < 0.5 no
stop

Figure 4.1 proposed flowchart adaboost with bayesnet

So above is the flowchart of Adaboost with Bayesnet.

4.2 proposed method description

Proposed method Steps :

Step 1 : instances of dataset (x1,y1),(x2,y2)…..(xn,yn).Each instance has equal probability D0(i)=1/n.

Step 2 : T learning rounds

Step 3 : apply classifier bayesnet to rice production dataset.

Step 4 : Learn bayesnet structure from given dataset

Step5:Structural learning: Apply bayesnet structural technique(local,global,hybrid,constrained based)

Step 6 : Parameter Learning of Bayesnet

step 7 : Conditional probability table at each node is created.Conditional probability table for each variable can estimated

Step 8 : Inference to add evidence to network.

Step 9 : Measures for network analysis.model validation. cross validation is a standard way to obtain estimates of models.

Step 10 : Et error < 0.5 if yes then continue else stop

Step 11: get weak classifier ht : alpha_t =(weight associated with ht)

Step 12 : update distribution Dt+1(i) for those examples for which learner classify them wrongly

Step 13 : Linear combination of these different hypothesis weighted by alpha_t

Step 14 : output result of adaboost with bayesnet.

4.3 Proposed method AdaBoost with bayesnet classifier algorithm

Given : (x1,y1),(x2,y2)….(xm,ym)

Initialize weights $D_1(i)=1/m$

For t=1….T

Train weak learner using distribution D_t

Apply Bayesnet base learner,which return base classifier ht with low weighted error

Bayesnet : Learn bayesnet structure from given dataset

Parameter learning of bayesnet by Conditional probability table at each node is created

Inference to add evidence to network

Model validation.cross validation is a standard way to obtain estimates of models.

Calculate $\varepsilon_t = \frac{\sum_{i=1}^{m} D_t(i)\delta(h_t(x_i)\neq y_i)}{D_t(i)}$

Where ε_t is misclassification error for the model.

D is weights of samples weights of samples

y_i is actual output and $h_t(x_i)$ is predicted output. $\delta(h_t(x_i) \neq y_i)$ = 1 if i is misclassified and

$\delta(h_t(x_i) \neq y_i)$ = 0 if i is correctly classified.

Error equals the sum of the misclassification rate, where weight for training sample i and y_i not being equal to our prediction $h_t(x_i)$ (which equals 1 if misclassified and 0 if correctly classified).

Set weight αt based on the error

$$\alpha_t = \frac{1}{2}\ln\left(\frac{1-\varepsilon_t(h_t)}{\varepsilon_t(h_t)}\right)$$

this strength α_t is used to decide weight of voting.

Update the distribution based on the performance

$$D_{t+1}(i) = \frac{D_t(i)\exp\left(-\alpha_t y_i h_t(x_i)\right)}{Z_t}$$

If i is misclassified then $y_i h_t(x_i)$ = -1 else if i is classified correctly then $y_i h_t(x_i)$ = 1

Z_t is normalization factor. Zt is choosen such that sum of probability is equal to 1.

Output the strong classifier :

H(x) = $\text{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$

Final classifier is linear combination of these different learning hypothesis.

Mathematical terms :

In this , there are 10 samples

$D_1(i)=1/m$

So $D_1(i) = 1/10 = 0.1$

$D_1(1)$	$D_1(2)$	$D_1(3)$	$D_1(4)$	$D_1(5)$	$D_1(6)$	$D_1(7)$	$D_1(8)$	$D_1(9)$	$D_1(10)$
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

So each samples have weight equal to 0.1

Now apply classifier bayesnet on weighted data :

Calculate $\varepsilon_t = \frac{\sum_{i=1}^{m} D_t(i)\delta(h_t(x_i)\neq y_i)}{D_t(i)}$

Where ε_t is misclassification error for the model.

D is weights of samples weights of samples

y_i is actual output and $h_t(x_i)$ is predicted output. $\delta(h_t(x_i) \neq y_i)$ = 1 if i is misclassified and

$\delta(h_t(x_i) \neq y_i)$ = 0 if i is correctly classified.

Error equals the sum of the misclassification rate, where weight for training sample i and y_i

not being equal to our prediction $h_t(x_i)$ (which equals 1 if misclassified and 0 if correctly classified).

$$\varepsilon_1 = \frac{0.1*4}{0.1*10} = 0.4$$

1,3,6,9 are misclassified

Set weight αt based on the error

$\alpha_t = \frac{1}{2} \ln \left(\frac{1-\varepsilon_t(h_t)}{\varepsilon_t(h_t)}\right)$ where $\varepsilon_t(h_t)$ = 0.4

$\alpha_1 = = \frac{1}{2} \ln \left(\frac{1-0.4}{0.4}\right) = 0.2027$

Update the distribution based on the performance

$$D_{t+1}(i) = \frac{D_t(i) \exp\left(-\alpha_t y_i h_t(x_i)\right)}{Z_t}$$

Z_t is normalization factor.

$D_2(1) = D_t(i) \exp\left(-\alpha_t y_i h_t(x_i)\right) = \underbrace{D_1(1)}_{0.1} \exp\left(-\underbrace{\alpha_1}_{0.2027} \underbrace{y_1}_{-1} \underbrace{h_1(x_1)}_{1}\right)$

So $D_2(1)=0.1*\exp(-0.2027*-1) = 0.125$

$D_2(2) = 0.1 * \exp(-0.2027 * 1) = 0.083$

$D_2(3) = 0.1 * \exp(-0.2027 * -1) = 0.125$

$D_2(4) = 0.1 * \exp(-0.2027 * 1) = 0.083$

$D_2(5) = 0.1 * \exp(-0.2027 * 1) = 0.083$

$D_2(6) = 0.1 * \exp(-0.2027 * -1) = 0.125$

$D_2(7) = 0.1 * \exp(-0.2027 * 1) = 0.083$

$D_2(8) = 0.1 * \exp(-0.2027 * 1) = 0.083$

$D_2(9) = 0.1 * \exp(-0.2027 * -1) = 0.125$

$D_2(10) = 0.1 * \exp(-0.2027 * 1) = 0.083$

So in this way update distribution :

Now again apply classifier on weighted data :

2,4,5,7 are misclassified

$$\varepsilon_2 = \frac{0.083*4}{0.125+0.083+0.125+0.083+0.083+0.125+0.083+0.083+0.125+0.083} = 0.33$$

Set weight αt based on the error

$$\alpha_2 = \frac{1}{2}\ln\left(\frac{1-0.33}{0.33}\right) = 0.3540$$

$$z_t = \sum_{i=1}^{m} D_t(i)\exp(-\alpha_t y_i h_t(x_i))$$

$$zt = 0.0833*4*\exp(+0.3540) + 0.125*4*\exp(-0.3540) + 0.0833*2*\exp(-0.3540) = 0.942$$

now update distribution

$D_3(1) = D_2(1) * \exp(-0.3540 * 1) = 0.09$

$D_3(2) = D_2(2) * \exp(-0.3540 * -1) = 0.12$

$D_3(3) = D_2(3) * \exp(-0.3540 * 1) = 0.09$

$D_3(4) = D_2(4) * \exp(-0.3540 * -1) = 0.12$

$D_3(5) = D_2(5) * \exp(-0.3540 * -1) = 0.12$

$D_3(6) = D_2(6) * \exp(-0.3540 * 1) = 0.09$

$D_3(7) = D_2(7) * \exp(-0.3540 * -1) = 0.12$

$D_3(8) = D_2(8) * \exp(-0.3540 * 1) = 0.06$

$D_3(9) = D_2(9) * \exp(-0.3540 * 1) = 0.09$

$D_3(10) = D_2(10) * \exp(-0.3540 * 1) = 0.06$

In this way update distribution upto $D_3(10)$.

On this updated distribution apply classifier and calculate error

Instances 2,3,4,5,7 are misclassified,so higher the probability of wrong classified instances and lower the probability of correctly classified instances.

$\varepsilon_3 = 0.5996$

Final classifier is linear combination of these different learning hypothesis.

$H(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$

If the sum is positive, then the first class is predicted, if negative the second class is predicted.

CHAPTER 5

EXPERIMENTAL WORK AND RESULT ANALYSIS

5.1 PROPOSED METHOD IMPLEMENTATION

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         501               98.2353 %
Incorrectly Classified Instances         9                1.7647 %
Kappa statistic                          0.9599
Mean absolute error                      0.0184
Root mean squared error                  0.1149
Relative absolute error                  4.1347 %
Root relative squared error             24.3818 %
Total Number of Instances              510

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.997    0.047    0.977      0.997    0.987      0.960    0.996     0.997     high
                 0.953    0.003    0.994      0.953    0.973      0.960    0.996     0.995     low
Weighted Avg.    0.982    0.032    0.983      0.982    0.982      0.960    0.996     0.996

=== Confusion Matrix ===

   a   b   <-- classified as
 339   1 |   a = high
   8 162 |   b = low
```

Figure5.1 : proposed method output

```java
package weka.api;
import weka.core.Instances;
import java.util.Random;
import weka.classifiers.bayes.net.search.local.HillClimber;
import weka.core.converters.ConverterUtils.DataSource;
import weka.classifiers.bayes.BayesNet;
import weka.core.converters.ArffSaver;
import weka.classifiers.meta.AdaBoostM1;
import weka.classifiers.Evaluation;


public class hi {
    public static void main(String args[]) throws Exception{
        //String data = "/C:/Users/LENOVO/Desktop/281118/61218/newdatasetcopy1.arff";
        //String data = "D:/eclipsericedatabasic.arff";
        String data = "C:/Users/LENOVO/Desktop/finaldataset.arff";

    DataSource source = new DataSource(data);



    Instances trainingData = source.getDataSet();

        ArffSaver arff = new ArffSaver();
        arff.setInstances(trainingData);



        if(trainingData.classIndex() == -1){
            trainingData.setClassIndex(trainingData.numAttributes()-1);
```

```
        }
        AdaBoostM1 m1 = new AdaBoostM1();
        BayesNet bn = new BayesNet();
        bn.setSearchAlgorithm(new HillClimber());
        m1.setClassifier(bn);
        m1.setNumIterations(10);
        m1.setSeed(1);
        m1.setUseResampling(false);
       //m1.setBatchSize(100);
        m1.setDebug(true);
        m1.setDoNotCheckCapabilities(false);
        m1.setNumDecimalPlaces(2);
        m1.setWeightThreshold(100);

        m1.buildClassifier(trainingData);
        Evaluation eval = new Evaluation(trainingData);
        Random rand = new Random(1);
        int folds = 10;
        eval.crossValidateModel(m1, trainingData, folds, rand);

       System.out.println("correct = "+eval.pctCorrect());
       System.out.println("Incorrect = "+eval.pctIncorrect());
       System.out.println("MAE = "+eval.meanAbsoluteError());
       System.out.println("RMSE = "+eval.rootMeanSquaredError());
       System.out.println("RAE = "+eval.relativeAbsoluteError());
       System.out.println("RRSE = "+eval.rootRelativeSquaredError());
       System.out.println("kappa = "+eval.kappa());
       System.out.println("precision = "+eval.precision(1));
       System.out.println("Recall = "+eval.recall(1));
       System.out.println("fMeasure = "+eval.fMeasure(1));
       System.out.println("error rate = "+eval.errorRate());
    }
}
```

Figure5.2 : proposed method implementation

Accuracy Comparision graph :

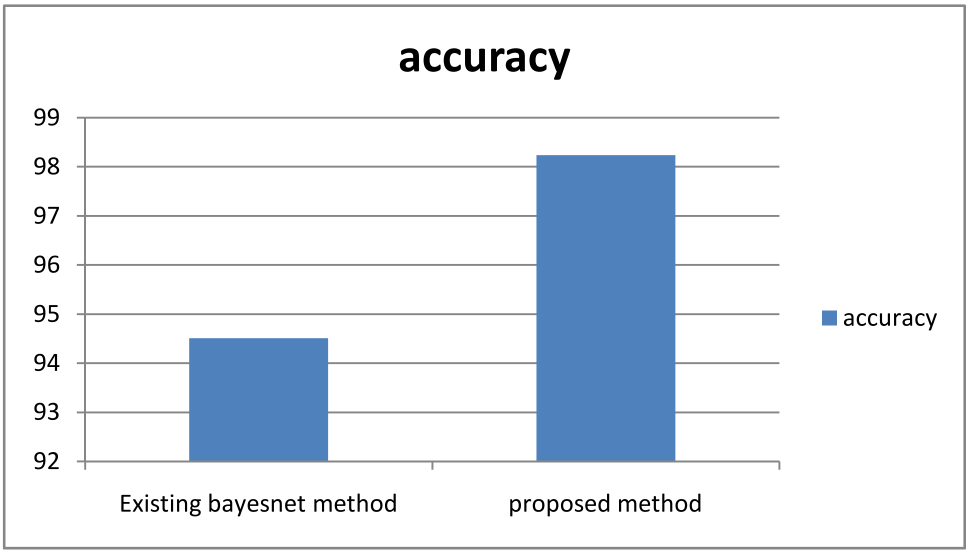


Figure5.3 : Accuracy Comparision graph

Root Mean Square Error comparision graph :

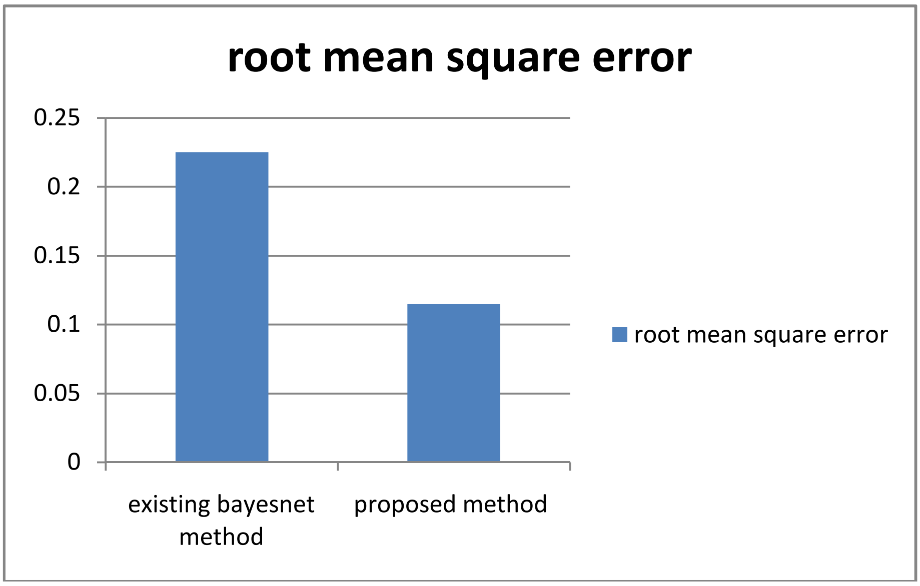


Figure5.4 : Root Mean Square Error Comparision graph

Relative Absolute Error comparision graph :

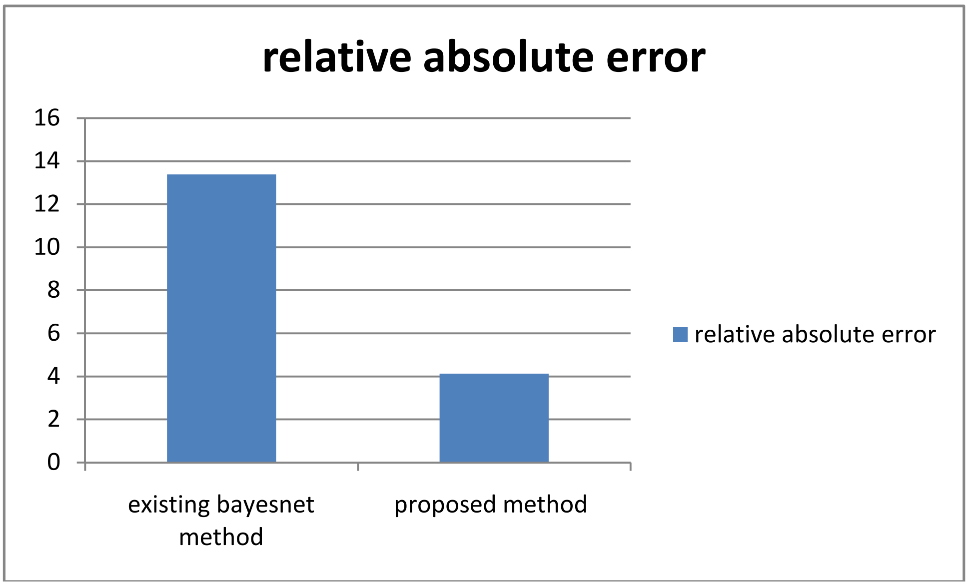


Figure5.5: Relative Absolute Error Comparision graph

Mean Absolute Error comparision graph :

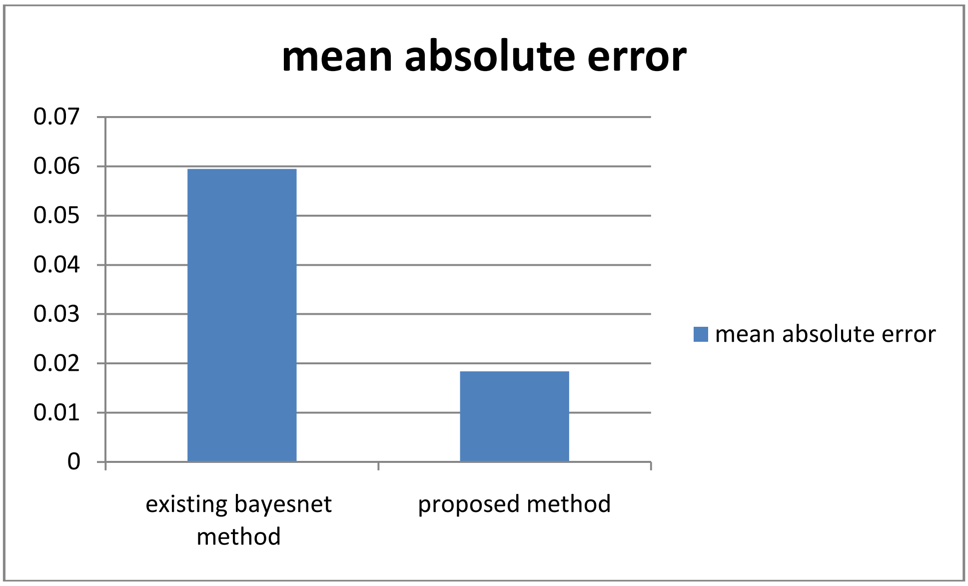


Figure5.6: Mean Absolute Error Comparision graph

I. add more parameters in rice crop dataset…………diurnal temperature range

Figure5.7 : bayesnet classifier output after adding diurnal temperature range parameter

```
    ---------- ----- ----------
=== Summary ===

Correctly Classified Instances         484               94.902  %
Incorrectly Classified Instances        26                5.098  %
Kappa statistic                          0.887
Mean absolute error                      0.055
Root mean squared error                  0.2165
Relative absolute error                 12.3788 %
Root relative squared error             45.9284 %
Total Number of Instances              510

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.947    0.047    0.976      0.947    0.961      0.888    0.988     0.993     high
                 0.953    0.053    0.900      0.953    0.926      0.888    0.988     0.980     low
Weighted Avg.    0.949    0.049    0.951      0.949    0.949      0.888    0.988     0.989

=== Confusion Matrix ===

   a   b   <-- classified as
 322  18 |   a = high
   8 162 |   b = low
```

Performance measures :

Accuracy = 94.902

RMSE = 0.2165

RAE = 12.3788

MAE = 0.055

Figure5.8 : proposed method output after adding diurnal temperature range parameter

```
=== Summary ===

Correctly Classified Instances         500               98.0392 %
Incorrectly Classified Instances        10                1.9608 %
Kappa statistic                          0.9556
Mean absolute error                      0.0202
Root mean squared error                  0.1358
Relative absolute error                  4.5476 %
Root relative squared error             28.8043 %
Total Number of Instances              510

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.991    0.041    0.980      0.991    0.985      0.956    0.997     0.999     high
                 0.959    0.009    0.982      0.959    0.970      0.956    0.997     0.994     low
Weighted Avg.    0.980    0.030    0.980      0.980    0.980      0.956    0.997     0.997

=== Confusion Matrix ===

   a   b   <-- classified as
 337   3 |   a = high
   7 163 |   b = low
```

Performance measure

Accuracy = 98.0392

RMSE = 0.1358

RAE = 4.5476

MAE = 0.0202

Add one more parameter in Rice Crop Dataset

Reference Crop

Evapotranspiration

Figure5.9 : proposed method Output after adding one more parameter

```
=== Summary ===

Correctly Classified Instances         501               98.2353 %
Incorrectly Classified Instances         9                1.7647 %
Kappa statistic                          0.9599
Mean absolute error                      0.0184
Root mean squared error                  0.1149
Relative absolute error                  4.1347 %
Root relative squared error             24.3818 %
Total Number of Instances              510

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.997    0.047    0.977      0.997    0.987      0.960    0.996     0.997     high
                 0.953    0.003    0.994      0.953    0.973      0.960    0.996     0.995     low
Weighted Avg.    0.982    0.032    0.983      0.982    0.982      0.960    0.996     0.996

=== Confusion Matrix ===

   a   b   <-- classified as
 339   1 |   a = high
   8 162 |   b = low
```

Accuracy : 98.2353

RMSE : 0.1149

RAE : 4.1347

MAE : 0.0184

II. State wise Rice crop yield prediction

State : Maharastra

District : Ahmednagar ,Amravati ,Aurangabad Beed/Bid Bhandara Bhandara ,Buldhana ,Chandrapur, Dhule ,Gadchiroli, Gondia ,Hingoli ,Jalana ,Jalgaon ,Kolhapur,Latur ,Nagpur ,Nanded ,Nasik ,Osmanabad ,Parbhani ,Pune ,Sangli ,Satara Solapur,Wardha.

Figure5.10 : Maharastra State Rice Crop Yield Prediction using Existing method

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         119               95.2    %
Incorrectly Classified Instances         6                4.8    %
Kappa statistic                          0.9002
Mean absolute error                      0.0475
Root mean squared error                  0.179
Relative absolute error                  9.7081 %
Root relative squared error             36.2028 %
Total Number of Instances              125

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.887    0.000    1.000      0.887    0.940      0.905    0.998     0.998     high
                 1.000    0.113    0.923      1.000    0.960      0.905    0.998     0.999     low
Weighted Avg.    0.952    0.065    0.956      0.952    0.952      0.905    0.998     0.998

=== Confusion Matrix ===

  a  b   <-- classified as
 47  6 |  a = high
  0 72 |  b = low
```

Figure5.11 :Maharastra state rice crop yield prediction using proposed method

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         120               96      %
Incorrectly Classified Instances         5                4      %
Kappa statistic                          0.9171
Mean absolute error                      0.0381
Root mean squared error                  0.1782
Relative absolute error                  7.787  %
Root relative squared error             36.0386 %
Total Number of Instances              125

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.906    0.000    1.000      0.906    0.950      0.920    0.990     0.991     high
                 1.000    0.094    0.935      1.000    0.966      0.920    0.990     0.989     low
Weighted Avg.    0.960    0.054    0.963      0.960    0.960      0.920    0.990     0.990

=== Confusion Matrix ===

  a  b   <-- classified as
 48  5 |  a = high
  0 72 |  b = low
```

Accuracy comparision graph for State wise Rice Crop Yield Prediction

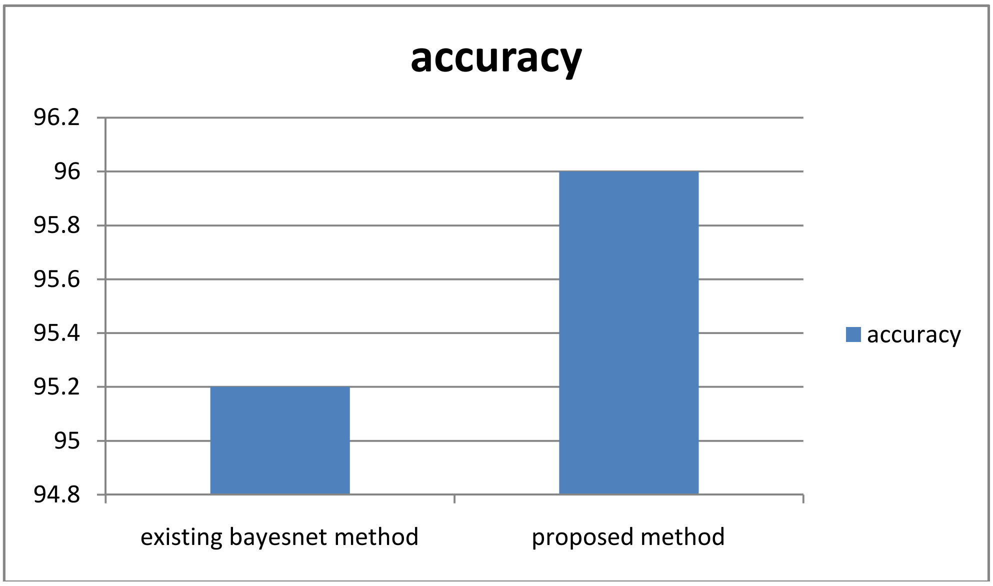


Figure5.12 :Maharastra state Rice crop yield prediction accuracy comparision graph using proposed method

Root Mean Square Error comparision graph for State wise Rice Crop Yield Prediction:

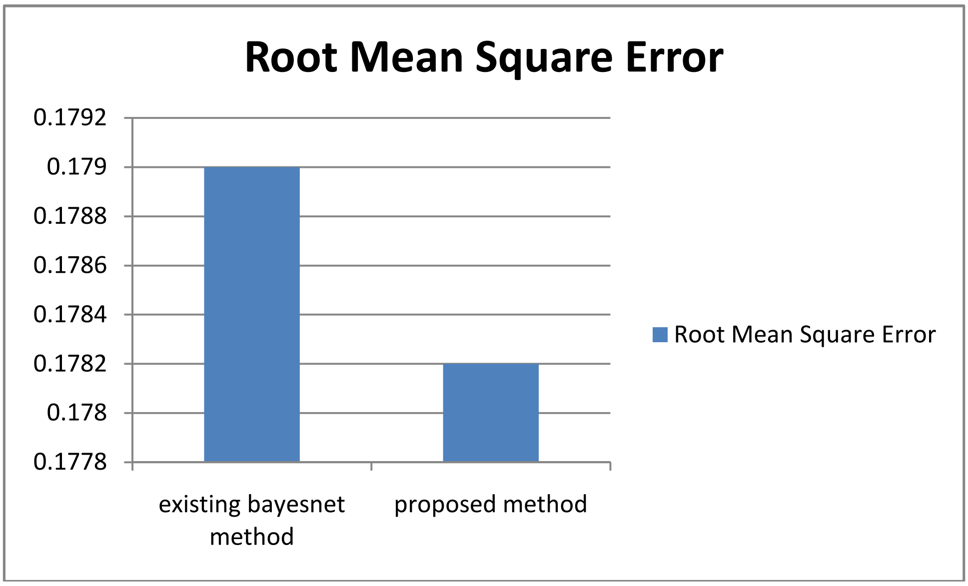


Figure5.13:Maharastra state Rice crop yield prediction RMSE comparision graph using proposed method

Relative Absolute Error comparision graph for State wise Rice Crop Yield Prediction:

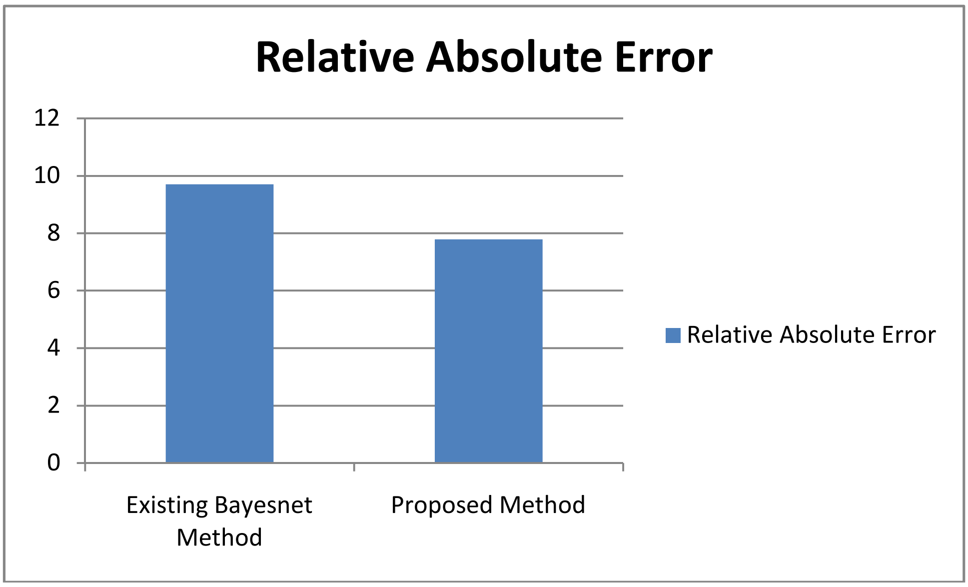


Figure5.14:Maharastra state Rice crop yield prediction RAE comparision graph using proposed method

Mean Absolute Error comparision graph for State wise Rice Crop Yield Prediction:

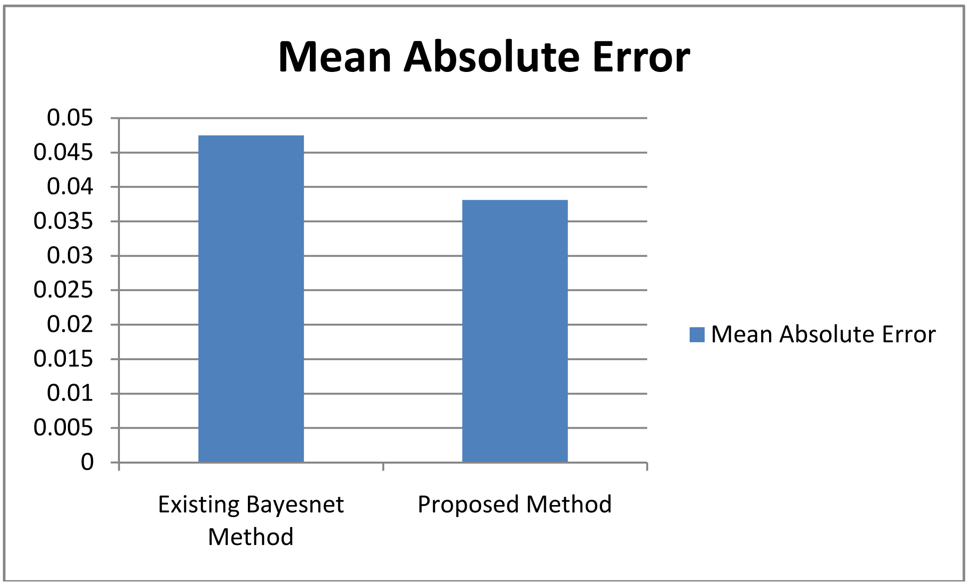


Figure5.15:Maharastra state Rice crop yield prediction MAE comparision graph using proposed method

CHAPTER 6

CONCLUSION

In rice crop yield prediction,various classification techniques are applied like bayesnet,support vector machine,naïve bayes,bayesnet,locally weighted learning,ladtree etc.they all have some drawbacks like low accuracy and high classification errors like RMSE(Root Mean Square Error),MAE(Mean Absolute Error),RAE(Relative Absolute Error) on rice crop production datasets.To get more significant result and to get better classification accuracy,boosting technique Adaptive boosting technique is used with classification method bayesnet and that helps us to determine better accuracy of prediction of rice yield**.**

CHAPTER 7

BIBLIOGRAPHY

1. Shruti Mishra,Priyanka Paygude,Snehal Chaudhary,Sonali Idate"Use of Data Mining in Crop Yield Prediction", Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018),pp. 796-802,2018

2. Umid Kumar Dey,Abdullah Hasan Masud,Mohammed Nazim Uddin,"Rice Yield Prediction Model Using Data Mining",International Conference on Electrical, Computer and Communication Engineering (ECCE),pp. 321-326,2017

3. Niketa Gandhi,Leisa J. Armstrong,Manisha Nandawadekar,"Application Of Data amaining Techniques For Predicting Rice Crop Yield Semi-Arid Climate Zone Of India",2017 IEEE International Conference on Technological Innovations in ICT for agriculture and Rural Development,pp. 116-120,2017

4. Ayush Shah, Akash Dubey, Vishesh Hemnani, Divye Gala,D. R. Kalbande,"Smart Farming System: Crop Yield Prediction Using Regression Techniques",Proceedings of International Conference on Wireless Communication,pp. 49-56,2018

5. Niketa Gandhi, Leisa J. Armstrong, Owaiz Petkar, Amiya Kumar Tripathy,"Rice Crop Yield Prediction in India using Support Vector Machines",2016 International Joint Conference on Computer Science and Software Engineering (JCSSE)

6. Niketa Gandhi, Leisa J. Armstrong, Owaiz Petkar,"Predicting Rice Crop Yield Using Bayesian Networks",2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI),pp. 795-799,2016

7. Akter Hossain,Mohammed Nazim Uddin,Mohammad Arif Hossain,Yeong Min Jang,"Predicting Rice Yield For Bangladesh By Exploiting Weather Condition",2017 International Conference on Information and Communication Technology Convergence (ICTC),pp.589-594,2017

8. Niketa Gandhi, Owaiz Petkar, Leisa J. Armstrong,"Rice Crop Yield Prediction Using Artificial Neural Networks",2016 IEEE International Conference on Technological

Innovations in ICT For Agriculture and Rural Development (TIAR 2016),pp.105-110,2016

9. Niketa Gandhi, Leisa J. Armstrong,"Applying Data Mining Techniques to Predict Yield Of Rice in Humid Subtropical Climatic Zone Of India",2016 International Conference on Computing for Sustainable Global Development,pp.1901-1906,2016

10. Sagar S. Nikam ,”A Comparative Study of Classification Techniques in Data Mining Algorithms”,ISSN April 2015.

11. Narayanan Balakrishnan, Dr.Govindarajan Muthukumarasamy,”crop Production-Ensemble Machine Learning Model for Prediction”,International Journal of Computer Science and Software Engineering (IJCSSE), Volume 5, Issue 7, July 2016,pp. 148-153.

12. Subhadra Mishra, Debahuti Mishra, Gour Hari Santra,"Adaptive boosting of weak regressors for forecasting of crop production considering climatic variability: An empirical assessment",Journal of King Saud University - Computer and Information Sciences,pp. 1-23.

13. Philip G. Oguntunde,Gunnar Lischeid,Ottfried Dietrich,"Relationship between rice yield and climate variables in southwest Nigeria using multiple linear regression and support vector machine analysis",August 2017,springer.

14. Haresh L. Siju,Pinal J. Patel,"Review on Crop Yield Prediction using Data Mining Focusing on Groundnut Crop and Naive Bayes Technique",International Journal of Scientific Research in Computer Science, Engineering and Information Technology,Volume 3 | Issue 1,ISSN : 2456-3307,pp.1947-1954

15. Olugbenga Wilson Adejo,Thomas Connolly,"Predicting student academic performance using multi-model heterogeneous ensemble approach",Journal of Applied Research in Higher Education,Vol. 10 Issue: 1, pp.61-75

16. ADNAN MASOOD,"Bayesian Networks",A BRIEF INTRODUCTION

17. Mrinal Pandey, S. Taruna,” A Comparative Study of Ensemble Methods for Students’ Performance Modeling”, International Journal of Computer Applications, Volume 103 – No.8, October 2014,pp.26-32.

18. https://www.youtube.com/watch?v=MRD67WgWonA

19. Jiawei Han,Micheline Kamber,"Data Mining:Concepts and Techniques"

20. https://www.youtube.com/watch?v=PNglugooJUQ

21. ADNAN MASOOD,"Bayesian Networks",A BRIEF INTRODUCTION

22. https://en.wikipedia.org/wiki/Naive_Bayes_classifier

23. Friedman, N., Geiger, D. and Goldszmidt M. (1997),” Bayesian network classifiers”, Machine Learning 29, 139–164.

24. Castillo, G. (2006) ,”Adaptive Learning Algorithms for Bayesian Network Classifiers”, PhD Thesis, Chapter 3.

25. Bouckaert, R. (2007),” Bayesian Network Classifiers in Weka”,Technical Report

26. Heckerman, D. (1996),” A Tutorial on Learning with Bayesian Networks”, Microsoft Technical Report MSR-TR-95-06.

27. G. Cooper, E. Herskovits (1992),”A Bayesian method for the induction of probabilistic networks from data,”Machine Learning. 9(4):309-347.

28. V. B. Nikam and B. B. Meshram, "Modeling Rainfall Prediction Using Data Mining Method: A Bayesian Approach", Fifth International Conference on Computational Intelligence, Modelling and Simulation, Seoul, 2013, pp. 132-136.

29. Niketa Gandhi, Leisa J. Armstrong, “Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India using data mining Techniques”, 2016 IEEE International Conference on Advances in Computer Applications (ICACA)

30. Rupanjali D. Baruah, R.M. Bhagat, Sudipta Roy, L.N. Sethi, ”Use of Data Mining Technique for Prediction of Tea Yield in the Face of Climate Change of Assam, India”, 2016 International Conference on Information Technology, pp. 265-269.

31. Akanksha Verma, Aman Jatain, Shalini Bajaj, ” Crop Yield Prediction of Wheat Using Fuzzy C Means Clustering and Neural Network”, International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 11 (2018) pp. 9816-9821.

32. Tirtha Ranjeet, Leisa Armstrong, ” An Artificial Neural Network for Predicting Crops Yield in Nepal”, Proceedings of AFITA 2014 - Copyright ASICTA Inc, pp. 376-386.

[33]. B. J I , Y. SUN, S.YANG, J. WAN1, " Artificial neural networks for rice yield prediction in mountainous regions", Journal of Agricultural Science (2007), pp. 249–261.

33. Ankita Bissa, Meena Kushwaha , Mayank Patel, "A Survey of Machine Learning Appliacations In Decision Making To Improve Farming", International Journal of Computer Sciences and Engineering, Vol.-6, Issue-9, Sept. 2018, pp. 669-683.

34. Saisunee Jabjone, Sura Wannasang, " Decision Support System Using Artificial Neural Network to Predict Rice Production in Phimai District, Thailand", International Journal of Computer and Electrical Engineering, Vol. 6, No. 2, April 2014, pp. 162-166.

35. Robert E. Schapire," Explaining AdaBoost",pp.1-16.

36. Saisunee Jabjone, Sura Wannasang, " Decision Support System Using Artificial Neural Network to Predict Rice Production in Phimai District, Thailand", International Journal of Computer and Electrical Engineering, Vol. 6, No. 2, April 2014, pp. 162-166.

37. Veenadhari, B. Misra, and C. D. Singh, "Machine learning approach for forecasting crop yield based on climatic parameters, " 2014 Int. Conf. Comput. Commun. Informatics Ushering Technol. Tomorrow, Today, ICCCI 2014, pp. 1-5, 2014.

38. E. Manjula, S. Djodiltachoumy, " A Model for Prediction of Crop Yield", International Journal of Computational Intelligence and Informatics, Vol. 6: No. 4, March 2017, pp.298-305.

39. Sellam and E. Poovammal, "Prediction of Crop Yield using Regression Analysis, " vol. 9, no. October, 2016.

40. B. H. Dhivya, R. Manjula, S. B. S, and R. Madhumathi, "A Survey on Crop Yield Prediction based on Agricultural Data, " pp. 4177-4183, 2017.

41. Savla and A. Mandholia, "Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture, " 2015.

42. Yildirak, Z. Kalaylioglu, and A. Mermer, "Bayesian estimation of crop yield function : drought based wheat prediction model for tigem farms, " Environ. Ecol. Stat., vol. 22, no. 4, pp. 693-704, 2015.

43. Mrs. K. E. Eswari, Ms. L.Vinitha, " Crop Yield Prediction in Tamil Nadu using Baysian Network", International Journal of Intellectual Advancements and Research in Engineering Computations, Volume-6, Issue-2, pp.1572-1576.

44. Niketa Gandhi, Owaiz Petkar, Leisa J. Armstrong, "Rice Crop Yield Prediction Using Artificial Neural Networks", 2016 IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development (TIAR 2016), pp.105-109.

Appendix A

Plagiarism Report

Plagiarism Checker X Originality Report

Similarity Found: 24%

Date: Monday, May 06, 2019

Statistics: 1460 words Plagiarized / 6126 Total words

Remarks: Medium Plagiarism Detected - Your Document needs Selective Improvement.

Government Engineering College,Modasa Page 1 CHAPTER 1 INTRODUCTION OVERVIEW Rice crop generation adds to the nourishment security of India, over 40% to by and large yield prediction.Its creation is dependent on great climatic conditions. Improving the ability of farmers to predict crop efficiency in under various climatic situations like rainfall,temperature,cloud cover,vapour pressure and so on can help ranchers and different partners in settling on critical choices regarding agronomy and harvest decision.The prediction of agricultural yield is a difficult and desirable task for every nation.There are many classification techniques like Support Vector Machine(SVM),Naïve Bayes,BayesNet,K Nearest Neighbour(KNN),Locally Weighted Learning(LWL),LADTree on rice crop production datasets.

The provision of accurate and timely information can assist farmers to make the best decision for their cropping situations This could benefit them to attain greater crop productivity. This study examines the application of machine learning for the prediction of rice crop yield.The current study use the most rice producing states(Haryana,Punjab,Tamilnadu,Andhra Pradesh,Bihar).The dataset parameters are state name,district name,crop year,Rainfall,maximum temperature,minimum temperature,average temperature,reference crop evapotranspiration,vapour pressure,potproentialevapotranspiration,diurnal temperature range,area,production,yield,class.

The dataset was then sorted on the basis of yield to classify the records in to low, moderate and high. The low yield was from 0.1 to 2 tonnes/hectare,and high from 2 to 5 tonnes/hectare. Government Engineering College,Modasa Page 2 Figure 1.1 : climatic variable of dataset Applications : To better predict classification accuracy of rice crop yield using classification technique.

To help farmers to predict rice crop yield in future year by applying machine learning classification technique to historical data[1]. Farmers will be able to knowing how much rice crop yield he is about to expect. Challenges : It is challenging that to predict rice crop yield in kharif season in India in earliest time climate was very healthy everything was happened on time.

But now most of the things have been changed due to global warming and many other climatic variables. The main problem with rice crop yield in India is lack of rainfall in seasonal time.Cloud cover is also converts as drawback.Many other climatic variables like vapour pressure,ground frost frequency has also affect rice crop yield in India.So this is more difficult to predict better accuracy by data Minimum temperatur Vapour Reference crop evapotranspirati Potential evapotranspiratii Diurnal temperature Maximum temperatur Climatic variables Government Engineering College,Modasa Page 3 mining classification technique.It is challenging to choose appropriate classification method for prediction of rice crop yield. Motivation : Rice crop yield prediction is a very issue in agricultural.

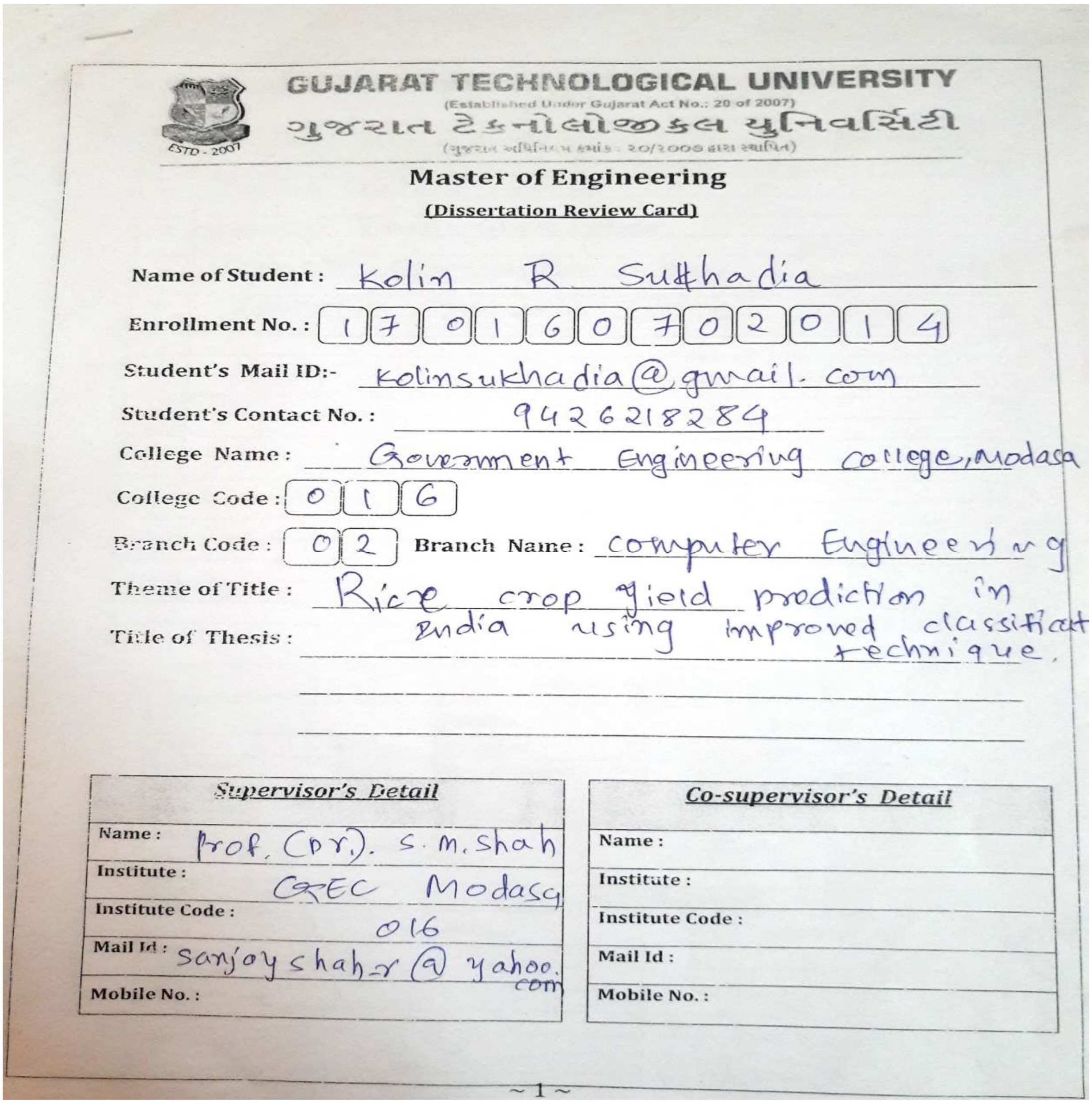
Any farmer want to know how much crop yield he expect. In the past, yield prediction was performed by considering farmer's experience on particular field. The rice yield prediction is a major issue that remains to be solved based on available rice production data. classifications techniques like naïve bayes,bayes network,support vector machine,locally weighted learning,Ladtree are the better choice for this purpose.

Different Data Mining classifications techniques are used and evaluated in agriculture for estimating the future year's rice crop production. This research proposes and implements a system to predict rice crop yield from previous rice crop production data through improved classification technique. Objective : The objective of this research is to bettere predicted accuracy of rice crop yield.so this will be helpful to farmers.

Any farmer is keen on realizing how much rice yield he is going to anticipate.Each rancher needs to think about the measure of rice crop yield his homestead will deliver to get the assessments about his profit.so for that to estimate better estimation , various classification techniques are applied on rice crop production dataset.This is improved by boosting technique to better predict rice crop yield.

Appendix B

Review Card



GUJARAT TECHNOLOGICAL UNIVERSITY

(Established Under Gujarat Act No.: 20 of 2007)

ગુજરાત ટેક્નોલોજીકલ યુનિવર્સિટી

ESTD - 2007

Master of Engineering

(Dissertation Review Card)

Name of Student : Kolin R Sukhadia

Enrollment No. : 1 7 0 1 6 0 7 0 2 0 1 4

Student's Mail ID:- kolinsukhadia@gmail.com

Student's Contact No. : 9426218284

College Name : Government Engineering college, Modasa

College Code : 0 1 6

Branch Code : 0 2 Branch Name : computer Engineering

Theme of Title : Rice crop yield prediction in

Title of Thesis : India using improved classification technique.

Supervisor's Detail	
Name :	Prof. (Dr.). S. M. Shah
Institute :	GEC Modasa
Institute Code :	016
Mail Id :	sanjayshahr@yahoo.com
Mobile No. :	

Co-supervisor's Detail	
Name :	
Institute :	
Institute Code :	
Mail Id :	
Mobile No. :	

~ 1 ~

❖ **Comments For Internal Review (2730002)** (Semester 3)

Exam Date : ___/___/_____

Sr. No.	Comments given by Internal review panel (Please write specific comments)	Modification done based on Comments
*	Search latest papers.	Done.
•	Specify base classifier.	Proposed approach
✓	Overall presentation is good.	is ready
*	Improve reference writing.	

(Guide Sign.)

Particulars	Internal Review Panel	
	Expert 1	Expert 2
Name :	M. C. Zala	M.M. Goyani
Institute :	Gec, Modasa	GEC, Modasa
Institute Code :	016	016
Mobile No. :	8460274518	
Sign :		

Particulars	Internal Guide Details	
	Expert 1	Expert 2
Name :		
Institute :		
Institute Code :		
Mobile No. :		
Sign :		

~ 2 ~

Enrollment No. of Student : 170160702014

❖ **Comments of Dissertation Phase-1 (2730003)** **(Semester 3)**

Exam Date : 13 / 12 / 18

Hall No : 09

Title : Rice Crop Yield Prediction in India Using Improved classification Technique

1. Appropriateness of title with proposal. (Yes/ No) Yes
2. Whether the selected theme is appropriate according to the title ? (Yes / No) Yes
3. Justify rational of proposed research. (Yes/ No) Yes
4. Clarity of objectives. (Yes/ No) Yes

~ 3 ~

Enrollment No. of Student : 1 7 0 1 6 0 7 0 2 0 1 4

Hall No. : 09

Exam Date : 13 / 12 / 18

Sr. No.	Comments given by External Examiners (DP-I) : (Please write specific comments)	Modification done based on Comments
→	Literature Review: Satisfactory	—
→	Background knowledge: Good	—
→	Implement your Proposed work and compare the results with related work.	Satisfactory completed
→	Do more experementation	partially completed in progress

(Internal Guide Sign.)

- Approved ✓
- Approved with suggested recommended changes
- Not Approved

Please tick on any on. If approved/approved with suggession then put marks ≥ 50 %.

➢ ***Details of External Examiners :***

Particulars	Full Name	University / College Name & Code	Mobile No.	Sign.
Expert 1	Prof. S. J. Malik	VGEC, Chandkheda (017)	9824343342	
Expert 2	Prof. J. J. Dhobi	GEC Gandhinagar (013)		

~ 4 ~

Enrollment No. of Student : 1 7 0 1 6 0 7 0 2 0 1 4

❖ **Comments of Mid Sem Review (2740001)** (Semester 4)

Hall No : 03

Exam Date : 8/3/19

Sr. No.	**Comments given by External Examiners :** i) The appropriateness of the major highlights of work done; State here itseif if work can be approved with some additional changes. ii) Main reasons for approving the work. iii) Main reasons if work is not approved.	**Modification done based on Comments**
		Done
	Good work done. yet Explanatory of Mathematical Terms & formula should be clear.	
		Internal Guide Sign.

- *Approved* ✓
- *Approved with suggested recommended changes*
- *Not Approved*

Please tick on any on. If approved/approved with suggession then put marks ≥ 50 %.

➢ ***Details of External Examiners :***

Particulars	*Full Name*	*University / College Name & Code*	*Mobile No.*	*Sign.*
Expert 1	S. J. Molia	VGEC (017)	9824347742	
Expert 2	Tejal Upadhyay	Nirma Uni.	9825507689	

~ 5 ~

Appendix C

Paper publication certificate

Scientific Journal Impact Factor = 4.032
Online ISSN : 2456-3307
UGC Journal No : 64718

International Journal of Scientific Research in Computer Science, Engineering and Information Technology

CERTIFICATE OF PUBLICATION

Ref : IJSRCSEIT/Certificate/Volume 5/Issue 1/4254 26-Feb-2019

This is to certify that **Kolin Sukhadia** has published a research paper entitled 'A Survey on Rice Crop Yield Prediction in India Using Improved Classification Technique ' in the International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), Volume 5, Issue 1, January-February 2019.

This Paper can be downloaded from the following IJSRCSEIT website link
http://ijsrcseit.com/CSEIT1951122
DOI : https://doi.org/10.32628/CSEIT1951122
IJSRCSEIT Team wishes all the best for bright future

Editor in Chief
IJSRCSEIT
website : http://ijsrcseit.com

Peer Reviewed and Refereed International Journal