

# Comparative Study for Classification Algorithms Performance in Crop Yields Prediction Systems

1<sup>st</sup> Halbast Rashid Ismael  
Technical College of Informatics-Akre  
Duhok Polytechnic University  
Duhok, Iraq  
[halbast.rashid@dpu.edu.krd](mailto:halbast.rashid@dpu.edu.krd)

2<sup>nd</sup> Adnan Mohsin Abdulazeez  
Research Center  
Duhok Polytechnic University  
Duhok, Iraq  
[adnan.mohsin@dpu.edu.krd](mailto:adnan.mohsin@dpu.edu.krd)

3<sup>rd</sup> Dathar A. Hasan  
Shekhan Technical Institute  
Duhok Polytechnic University  
Duhok, Iraq  
[dathar.hasan@dpu.edu.krd](mailto:dathar.hasan@dpu.edu.krd)

<https://doi.org/10.48161/qaj.v1n2a54>

**Abstract**— The agriculture importance is not restricted to our daily life; it is also an effective field that enhances the economic growth in any country. Therefore, developing the quality of the crop yields using recent technologies is a crucial procedure to obtain competitive crops. Nowadays, data mining is an emerging research field in agriculture especially in the predicting and analysis of crop yield. This paper focuses on utilizing various data mining classification algorithms to predict the impact of various parameters such as area, season and production on the crop yield quality. The performance of the decision tree, naive Bayes, random forest, support vector machine and K-nearest neighbour is measured and compared to each other. The comparison involves measuring the error values and accuracy. The SVM algorithm achieved the highest accuracy value with 76.82% while the lowest is achieved by the KNN algorithm with 35.76%. The highest error value was 111.8855 for KNN. Also, the prediction help farmer to increased and improved the income level.

**Keywords**— Data mining, Classification, Agriculture, Crop Yield.

## I. INTRODUCTION

Data mining is a machine learning technique for analyzing large datasets and identifying useful classifications and patterns. The ultimate aim of the data mining process is to take out information from the dataset and converting it into a usable system that can use in the future [1-4]. Data mining can be used in a variety of fields, from evaluating consumer activity in supermarkets [5][6] to predicting customer phone use. For many years, data mining has been used in agriculture, and this field is important to research [7][8]. The prediction of Crop yield is one of the most precision difficult issues for agriculture and several models were validated and proposed so far [9,10][11].

Agriculture is vital to the economy as well as humanity's future. Job is, first and foremost, essential for survival. It also generates a significant number of jobs. The supply-demand has grown exponentially as time has passed [12,13]. India, in particular, is one of the countries with the greatest interest in agriculture. People are

misusing technology to manufacture in large quantities [14,15].

Every day, new hybrid varieties are created. These hybrids, on the other hand, do not have the same essential nutrients as naturally grown crops. These non-natural methods degrade the soil [16,17]. All of this contributes to moreover environmental damage. The majority of these unnatural methods is used to prevent losses. However, when the farmers of these crops have reliable information on the crop production then the loss of production is reduced [18,19]. The information is based on previous weather, temperature, season, location, and a variety of other factors [20].

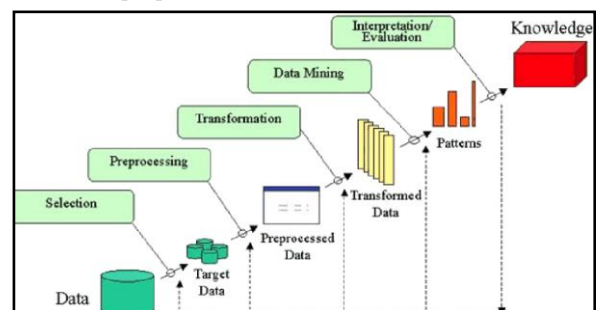


Fig1. Data Mining Process [12]

This paper aims to study the effect of different parameters such as area, season and production of main crops, also to predict crop production to increase yield and benefit for farmers by applying different classification algorithms like Naive Bayes, Decision Tree, SVM, KNN and Random Forest.

This paper is organized as follows; The related work is presented in section II, the description of the dataset used to give in section III, all methodology used explained in section IV, the experimental result and discussion are given in section V and the conclusion is presented in section VI.

## II. RELATED WORK

Haque et al. [21] proposed two separate Machine Learning (ML) algorithms to evaluate the yield of crops.

The algorithms Support Vector Regression (SVR) and Linear Regression (LR) have been well suited for validating variable parameters in continuous changeable estimation with 140 data points. The mentioned parameters are important determinants for crop yield. The error rate has been calculated for using Coefficient of Determination (R<sup>2</sup>) and Mean Square Error (MSE) with MSE yielding around 0.005 and R<sup>2</sup> yielding around 0.86. The same dataset has been used to compare the performance of the algorithms quickly.

Nishant et al. [22] proposed web application to forecast the yield of nearly all types of crops grown in India. This study is unique because it uses simple parameters such as state, district, season, and region to predict crop yields in any year the user desires. To predict the yield, advanced regression techniques such as Kernel Ridge, Lasso, and ENet algorithms are used in the paper, as well as the principle of Stacking Regression to improve the algorithms. The root means the square error is the output metric that used in this work. The models when have been implemented individually, ENet had an error of about 4%, Lasso had an error of about 2% and Kernel Ridge had an error of about 1%, and after stacking it was less than 1%.

Kalimuthu et al. [23] suggested a smartphone application for Android, which uses machine learning, which is one of the most advanced crops prediction technologies, to direct beginner farmers in sowing the appropriate crops. The naive Bayes algorithm proposes a method for doing so. The data seed of the crops is collected, along with the suitable parameters such as humidity, moisture, and temperature content, that is aids the crops' development. To begin the prediction process, users are encouraged to input parameters such as their location and temperature this application will take it automatically to start the process of prediction.

Y. J. N. Kumar et al. [24], implemented prediction system on crop production from the collecting of past data. Crop yield is estimated using data mining techniques. They used the Random Forest algorithm to forecast the highest yield crop as a product. Crops yield predictions are often appropriate in the agricultural sector. The higher the accuracy, the higher the benefit on the crop yield. Farmers will use the proposed technique to help them decide which crop to plant in their fields. Under this system would cover the widest range of crops possible. Farmers in India can benefit from accurate forecasting of various crops across various districts.

Gupta et al. [25] used IoT and data mining in monitoring applications to make smart farming possible. Smart farming is a method of providing all of the services needed for a specific time. Soil moisture, light intensity, relative humidity, soil pH reading, and ambient temperature are all resources that are needed. They demonstrated the transformation of a device capable of gathering data from sensors using IoT in the agriculture field. This device successfully senses data and sends it locally to the thing speak cloud, which the user can then access via his or her custom website. The data mining

techniques such as SVM, KNN and Random forest are used to crop-producing with the correct number of resources so that the farmer still has the upper hand, and by comparing the current pattern to the previous one, will be able to determine whether or not a parameter is right. All of these agricultural process values are monitored on a user-defined platform.

Terliksiz & Altlylar [26] used a 3D CNN model that leverages spatiotemporal features, a soybean yield prediction for Lauderdale County, Alabama, USA has been presented. From 2003 to 2016, the yield has been taken from the USDA NASS Fast Stat tool. Google Earth Engine was used to gather satellite data from (NASA's MODIS) land products surface reflectance and land surface temperature. For comparison of the results, the root means squared error (RMSE) has been used as the measurement metric. When 64x64 data frames with more than 20% cropland coverage were used without dropout, the best result for Lauderdale County, Alabama was obtained. The RMSE is 0.81 and the error is 2.70% in this situation.

Khosla et al., [27] The Kharif crops yield were estimated in 2 steps; first, the rainfall has been estimated by using a modular artificial neural network MANN, and then the yield was predicted by using support vector regression SVR. The dataset of experimental included data from the year 2000 to 2016 and generated results for the years 2018 and 2019.

Nigam et al. [28] utilized machine learning techniques such as KNN classifier, Random Forest, Artificial Neural Network, liner Regression and XGboost to predict yield crop. Based on the Mean Absolute Error MAE, the results of these techniques are compared. An algorithm of a machine learning algorithm can be help farmers to determine which crops planting to get the best yield by taking into account factors such as temperature, rainfall, area, ...etc. When all parameters are combined, the results show that the best classifier is Random Forest.

S & R [29] determined most of the important points for accurate Crop's yield Prediction. For improved accuracy, the algorithm of machine learning namely Support Vector Machine, KNN, Regression, Random Forest and Artificial Neural Network have been proposed. The agricultural dataset contains 745 instances, 30% of data is used to test the prediction performance of the models, while 70% of the data is chosen at random and used to train the model. The results show that, by using the same farming training data, the RF algorithm achieves maximum precision employing its error analysis values for all the separated feature sub-sets.

Kumar et al. [30] used the three supervised techniques SVM, KNN and Least Squared Support Vector Machine are. It is a comparative study that shows the training proposed model's accuracy and error rate. They have three different datasets: soil, rainfall, and yield. They then combined the datasets and used the techniques to determine the actual approximate cost as well as the

accuracy of the techniques used. The training model's accuracy will be higher and the error rate will be low. The technique with the highest accuracy rate (LS SVM =90%) and the lowest error rate (LS SVM=0,0362).

### III. METHODOLOGY

#### A. Dataset

The datasets that used in this experiment were collected from Kaggle publicly available datasets, which is a group of data scientists who provide datasets for different analysis or model building [31]. District name, crop, year, location, session, and production. There was a large collection of data for different districts and several years in India, after that the data was pre-processed where the data was used for seven different districts consisting of 152 records. Which was organized in Microsoft Excel. CSV (Comma Separated Values) was the file format.

#### B. WEKA Application

WEKA (Waikato Environment for Knowledge Analysis) [32] has been used to analyze the entire data collection. It is open-source software developed at the Waikato University in New Zealand in the JAVA programming language. It's used to solve problems including machine learning and data mining. It is released beneath the terms of the GNU General Public License. Preprocessing, grouping, regression, clustering, and visualization are some of the tasks it can perform. The data must be entered into the program, as well as selecting the appropriate task. It has a large number of classifiers that can be used to construct models and solve problems analytically it provides an integrated Graphical User Interface GUI, with all of the data analysis options you will need. The dataset for WEKA processing is in Comma Separated Values (CSV) or Attribute Relation File Format (ARFF) format [33].

#### C. Data mining Techniques

Data mining techniques are principally divided into 2 groups; clustering techniques and classification. This study uses Classification algorithms, which use classifiers to classify a group of the same objects into one class, and when a new object is added, predict which class the object may belong to [34-36]. This method aids in the classification of data into various classes. A supervised algorithm is used for classification. K-Nearest Neighbor, Naive Bayes, ZeroR, One-R, Random Forest, Decision Tree, Support Vector Machine and other flavours of classification are available. There are two stages to the classification process. The training set generates the model required to answer our query at first, and test data is then used to assess the model's performance.

In this study, five different classifiers namely (K-Nearest Neighbor-IBK, Decision Tree -J48, Naive Bayes, Random Forest, Support Vector Machine-SGD) were used.

- *Decision Tree*

One of the algorithms used in data mining for classification. The performance of the advanced

Geospatial Decision Support System (GDSS) is demonstrated by the use of mining techniques relevant to risk management [37]. A decision tree is a miniature learning structure that is primarily based on learning a rule that is universal through instances of observations [38,39].

- *Naive Bayes*

Naive Bayes is a probabilistic classification algorithm based on the Bayes theorem and the strong autonomy hypothesis. It is entirely dependent on the precise existence of the probability model. In a supervised learning environment, this classification technique can be trained to a strong level [40,41]. This algorithm has the advantage of only requiring a small amount of training data to evaluate constants such as variances and means of variables that are essential for a classifier [42].

- *Random Forest*

Random Forest is an algorithm used to perfectly systematizing large amounts of data [43]. It's a brand-new approach to regression and categorization learning that increases the number of decision trees running at the same time when executing the category mode of each tree [44,45].

- *K-Nearest Neighbor*

The K-Nearest Neighbor algorithm works based on the Supervised Learning methodology and is one of the most basic Machine Learning algorithms. Assumes that the new case/data and existing cases are identical, and then assigns the new case to the category that is closest to the existing categories. It keeps track of all available data and sorts new data points into categories based on their similarity [46,47]. This means that the K-NN algorithm can easily classify new data into a well-defined group. It can be used for both regression and classification, but classification problems are the most common.

- *Support Vector Machine (SVM)*

Data samples will be divided into two clusters by the Support Vector Machine. SVMs are a form of supervised learning system for classification and regression [48-50]. It can build a model to predict whether or not a new example belongs in the community. It's a concept that refers to a set of combined different learning techniques for displaying rate and patterns data used for classification and regression. A non-probabilistic binary linear classifier [51,52] will predict for each input using a group of inputs and two possible class forms.

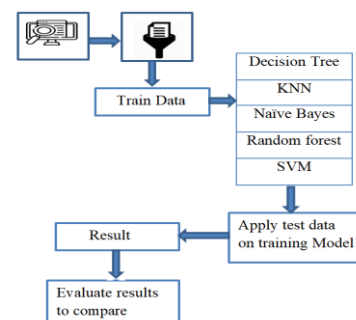


Fig.2: General Architecture of The Prediction Systems

The architecture of the Yield prediction system is depicted in the diagram above. Statistics from raw data used, after that, they are cleaned and sorted. Then the techniques of classification applied to the data that trained. WEKA is used to record the results of each algorithm and compare them.

#### D. Performance Metrics

The various machines learning algorithm is used for prediction of crop yield, as can be seen in the former section. Although there is a variety of machine learning algorithm to choose from the design of the application and the accuracy of the prediction algorithm are the most important factors to consider. Various metrics such (RAE) Relative Absolute Error [8], (RMSE) Root Mean Squared Error [53-55], and (MAE) Mean Absolute Error have been used to validate the prediction accuracy of classifiers.

- *RMSE*

RMSE is a square root of Mean Square Error (MSE). It's the forecast errors' standard deviation. The discrepancy between the model output value and the real target value is determined first [55]. Until computing the root of the mean value, this discrepancy is squared and summed over all data elements. RMSE is expressed as in (1).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\text{predicted} - \text{Actual})^2}{N}} \quad (1)$$

- *RAE*

RAE refers to the overall absolute error between the variables is calculated using this metric [50].

- *MAE*

The average of absolute deviations between the target and predicted values are calculated as the Mean Absolute Error (MAE) [56,57]. The computation of MAE is shown in (2).

$$MAE = \frac{1}{n} \sum_{i=1}^n |\text{predicted value} - \text{actual value}| \quad (2)$$

#### IV. EXPERIMENTAL RESULT AND DISCUSSION

Comparison of experimental results of the classification algorithms that used is shown below in tables and figures. Different classifier gives the different result of the same data. Table 1 shows the accuracy percentage. Table 2 shows the results of the error values for all classifiers, and table 3 presented the comparisons of the time that is taken to build the model.

TABLE 1. Accuracy of The Applied Model

Algorithm	Accuracy (%)
Decision Tree	74.1722
Naïve Bayes	47.0199
Random forest	68.8742
KNN	35.7616
SVM	76.8212

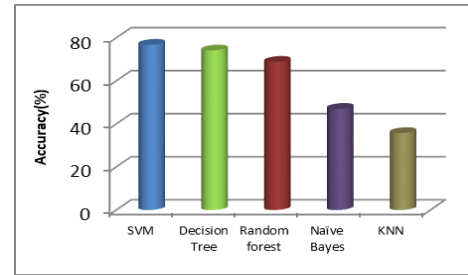


Fig.3: Accuracy of The Applied Models.

TABLE 2. Comparison of Error Values for The Applied Models

Algorithm	RMSE	MAE	RAE(%)
Decision Tree	0.2608	0.1184	40.8276
Naïve Bayes	0.4538	0.2671	92.1478
Random forest	0.3025	0.1662	57.3174
KNN	0.5595	0.3243	111.8855
SVM	0.3525	0.277	95.57

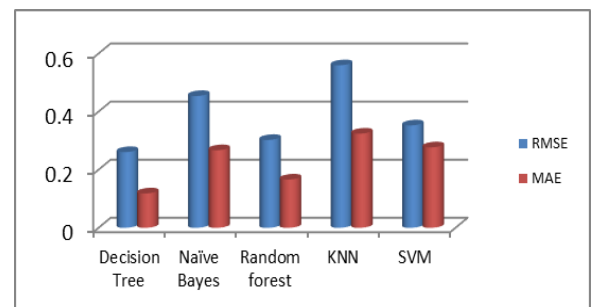


Fig.4: RMSE & MAE Values of The Applied Algorithms.

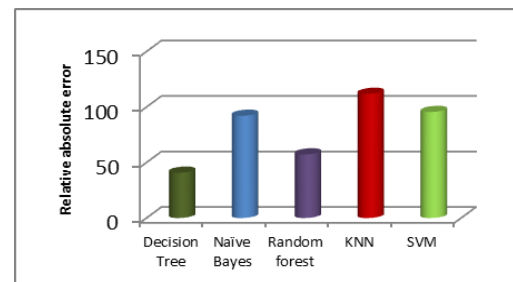


Fig.5: The RAE of The Applied Models

According to the use of the WEKA tool, the results shown in the tables were obtained. It is shown in the first table that SVM has 76% accuracy, Decision Tree has 74.17%, Random forest has 68.87 accuracies, and the lowest accuracy value was for KNN and Naïve Bayes algorithm which is 35.76%, 47.01%. Based on these results of accuracy, we note that SVM has the highest accuracy value compared to other techniques; therefore, we can say that the SVM classification is the best according to accuracy and can depend on it, while KNN has the lowest accuracy as shown in Figure 3 as well. In contrast, if we notice in the second table, which shows error values, that KNN has a higher value of errors 111.88, and Decision Tree has the lowest error value 40.82 compared to other algorithms used, as the difference in errors for all algorithms is clearly shown in Figures 4, 5. If we compare the time that taken to build each model that is shown in table 3. The time for each algorithm is converging, except Random forest that is taken longer than all models, which is 0.25 seconds out of all models

used. Finally, we can say that the decision tree is the best in terms of time and errors because it contains the lowest error value and the time it takes is 0. These values obtained are not fixed and can be changed. The result depends on the type and nature of datasets.

TABLE 3. Comparisons of Time Taken to Build Model

Algorithm	Time per second
Decision Tree	0
Naïve Bayes	0.01
Random forest	0.25
KNN	0
SVM	0.19

## V. CONCLUSION

In this paper, the data mining classification techniques have been used as a predictive modelling technique for the prediction of crop yield. Five classifier algorithms namely Decision Tree, Naive Bayes, SVM, Random Forest and KNN were used. The outcome of these techniques is compared based on error values and accuracy. The result of the experiment showed that the SVM algorithm gets the highest accuracy value 76.8212, the lowest accuracy was for KNN 35.7616. And the highest error value 111.8855 was for KNN. However, using these techniques, improved the farmer income level will be increased and the crop yield can be increased. Hence data mining techniques could be used in the agriculture sector for tacking better decisions.

## REFERENCES

[1] Sulaiman, M. A. (2020). Evaluating Data Mining Classification Methods Performance in the Internet of Things Applications. *Journal of Soft Computing and Data Mining*, 1(2), 11-25.

[2] Medar, R. A., & Rajpurohit, V. S. (2014). A survey on data mining techniques for crop yield prediction. *International Journal of Advanced Research in Computer Science and Management Studies*, 2(9), 59-64.

[3] Nathgosavi, V., & Patil, S. (2021). A Survey on Crop Yield Prediction using Machine Learning (No. 5238). EasyChair.

[4] Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., & Zeebaree, S. R. (2017). Combination of K-means clustering with Genetic Algorithm: A review. *International Journal of Applied Engineering Research*, 12(24), 14238-14245.

[5] Alpaydin, E., 2010. Introduction to Machine Learning, 2nd ed. Retrieved from [https://books.google.nl/books?hl=en&lr=&id=TtrxCwAAQBAJ&oi=fnd&pg=PR7&dq=introduction+to+machine+learning&ots=T5ejQG\\_7pZ&sig=0xC\\_](https://books.google.nl/books?hl=en&lr=&id=TtrxCwAAQBAJ&oi=fnd&pg=PR7&dq=introduction+to+machine+learning&ots=T5ejQG_7pZ&sig=0xC_)

[6] Abdulkareem, N. M., Abdulazeez, A. M., Zeebaree, D. Q., & Hasan, D. A. (2021). COVID-19 World Vaccination Progress Using Machine Learning Classification Algorithms. *Qubahan Academic Journal*, 1(2), 100-105.

[7] Witten, I.H., Frank, E., Hall, M.A., Pal, C.J., 2016. Data Mining: Practical Machine Learning Tools and Techniques. *Data Mining: Practical Machine Learning Tools and Techniques*. <https://doi.org/10.1016/c2009-0-19715-5>.

[8] Chicho, B. T., Abdulazeez, A. M., Zeebaree, D. Q., & Zebari, D. A. (2021). Machine Learning Classifiers Based Classification For IRIS Recognition. *Qubahan Academic Journal*, 1(2), 106-118.

[9] Mahmood, M. R., & Abdulazeez, A. M. (2019, April). A different model for hand gesture recognition with a novel line feature extraction. In 2019 International Conference on Advanced Science and Engineering (ICOASE) (pp. 52-57). IEEE.

[10] Haque, F. F., Abdelgawad, A., Yanambaka, V. P., & Yelamarthi, K. (2020, June). Crop yield analysis using

machine learning algorithms. In 2020 IEEE 6th World Forum on Internet of Things (WF-IoT) (pp. 1-2). IEEE.

[11] Taher, K. I., Abdulazeez, A. M., & Zebari, D. A. (2021). Data Mining Classification Algorithms for Analyzing Soil Data. *Asian Journal of Research in Computer Science*, 17-28.

[12] Paul, M., Vishwakarma, S. K., & Verma, A. (2015, December). Analysis of soil behaviour and prediction of crop yield using data mining approach. In 2015 International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 766-771). IEEE.

[13] Patel, H., & Patel, D. (2014). A brief survey of data mining techniques applied to agricultural data. *International Journal of Computer Applications*, 95(9).

[14] D Ramesh, B Vishnu Vardhan, "Data Mining Techniques and Applications to Agricultural Yield Data", *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 2, Issue 9, September 2013.

[15] Zeebaree, D. Q., Haron, H., & Abdulazeez, A. M. (2018, October). Gene selection and classification of microarray data using convolutional neural network. In 2018 International Conference on Advanced Science and Engineering (ICOASE) (pp. 145-150). IEEE.

[16] Ramesh, D., & Vardhan, B. V. (2015). Analysis of crop yield prediction using data mining techniques. *International Journal of research in engineering and technology*, 4(1), 47-473.

[17] Jambekar, S., Nema, S., & Saquib, Z. (2018, August). Prediction of Crop Production in India Using Data Mining Techniques. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (pp. 1-5). IEEE.

[18] Kumar, Y. J. N., Spandana, V., Vaishnavi, V. S., Neha, K., & Devi, V. G. R. R. (2020, June). Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector. In 2020 5th International Conference on Communication and Electronics Systems (ICCES) (pp. 736-741). IEEE.

[19] Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., & Zebari, D. A. (2019, April). Machine learning and region growing for breast cancer segmentation. In 2019 International Conference on Advanced Science and Engineering (ICOASE) (pp. 88-93). IEEE.

[20] Van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture*, 177, 105709.

[21] Haque, F. F., Abdelgawad, A., Yanambaka, V. P., & Yelamarthi, K. (2020, June). Crop yield analysis using machine learning algorithms. In 2020 IEEE 6th World Forum on Internet of Things (WF-IoT) (pp. 1-2). IEEE.

[22] Nishant, P. S., Venkat, P. S., Avinash, B. L., & Jabber, B. (2020, June). Crop Yield Prediction based on Indian Agriculture using Machine Learning. In 2020 International Conference for Emerging Technology (INCET) (pp. 1-4). IEEE.

[23] Kalimuthu, M., Vaishnavi, P., & Kishore, M. (2020). Crop Prediction using Machine Learning. 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 926-932. <https://doi.org/10.1109/ICSSIT48917.2020.9214190>

[24] Kumar, Y. J. N., Spandana, V., Vaishnavi, V. S., Neha, K., & Devi, V. G. R. R. (2020, June). Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector. In 2020 5th International Conference on Communication and Electronics Systems (ICCES) (pp. 736-741). IEEE.

[25] Gupta et al., (2020) Gupta, G., Setia, R., Meena, A., & Jaint, B. (2020, June). Environment Monitoring System for Agricultural Application using IoT and Predicting Crop Yield using Various Data Mining Techniques. In 2020 5th International Conference on Communication and Electronics Systems (ICCES) (pp. 1019-1025). IEEE.

[26] Terliksiz, A. S., & Altıylar, D. T. (2019, July). Use of deep neural networks for crop yield prediction: A case study of soybean yield in Lauderdale county, Alabama, USA. In 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics) (pp. 1-4). IEEE.

- [27] Khosla, E., Dharavath, R., & Priya, R. (2019). Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression. *Environment, Development and Sustainability*, 1-22.
- [28] Nigam, A., Garg, S., Agrawal, A., & Agrawal, P. (2019, November). Crop yield prediction using machine learning algorithms. In 2019 Fifth International Conference on Image Information Processing (ICIIP) (pp. 125-130). IEEE.
- [29] Maya Gopal P. S. & Bhargavi R. (2019) Performance Evaluation of Best Feature Subsets for Crop Yield Prediction Using Machine Learning Algorithms, *Applied Artificial Intelligence*, 33:7, 621-642, DOI: [10.1080/08839514.2019.1592343](https://doi.org/10.1080/08839514.2019.1592343)
- [30] Kumar, A., Kumar, N., & Vats, V. (n.d.). 2018 EFFICIENT CROP YIELD PREDICTION USING MACHINE LEARNING ALGORITHMS. *International Research Journal of Engineering and Technology (IRJET)* 05(06), 9.
- [31] S. I. Saleem, A. M. Abdulazeez and Z. Orman, "A new segmentation framework for Arabic handwritten text using machine learning techniques," *Computers, Materials & Continua*, vol. 68, no.2, pp. 2727–2754, 2021.
- [32] WEKA 3: Data Mining Software in Java, Machine Learning Group at the University of Waikato, Official Website: <http://www.cs.waikato.ac.nz/ml/weka/index.html>, accessed on 12th October 2017.
- [33] Mishra, S., Paygude, P., Chaudhary, S., & Idate, S. (2018, January). Use of data mining in crop yield prediction. In 2018 2nd International Conference on Inventive Systems and Control (ICISC) (pp. 796-802). IEEE.
- [34] Charbuty, B., & Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, 2(01), 20-28.
- [35] Saeed, J. N., & Abdulazeez, A. M. (2021). Facial Beauty Prediction and Analysis Based on Deep Convolutional Neural Network: A Review. *Journal of Soft Computing and Data Mining*, 2(1), 1-12.
- [36] Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., & Zebari, D. A. (2019, April). Trainable model based on new uniform LBP features to identify the risk of breast cancer. In 2019 International Conference on Advanced Science and Engineering (ICOASE) (pp. 106-111). IEEE.
- [37] Ibrahim, I., & Abdulazeez, A. (2021). The Role of Machine Learning Algorithms for Diagnosing Diseases. *Journal of Applied Science and Technology Trends*, 2(01), 10-19.
- [38] Georg Ruß, Rudolf Kruse, Martin Schneider, and Peter Wagner. "Estimation of neural network parameters for wheat yield prediction" In Max Bramer, editor, *Artificial Intelligence in Theory and Practice II*, volume 276 of IFIP International Federation for Information Processing, pages 109–118. Springer, July 2008.
- [39] Suganya, M. (2020). Crop Yield Prediction Using Supervised Learning Techniques. *International Journal of Computer Engineering and Technology*, 11(2).
- [40] Sujatha, R., & Isakki, P. (2016, January). A study on crop yield forecasting using classification techniques. In 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16) (pp. 1-4). IEEE.
- [41] Gadge, Y. (2017, December). A study on various data mining techniques for crop yield prediction. In 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICECCOT) (pp. 420-423). IEEE.
- [42] Bhosale, S. V., Thombare, R. A., Dhemey, P. G., & Chaudhari, A. N. (2018, August). Crop yield prediction using data analytics and hybrid approach. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (pp. 1-5). IEEE.
- [43] Bhanumathi, S., Vineeth, M., & Rohit, N. (2019, April). Crop yield prediction and efficient use of fertilizers. In 2019 International Conference on Communication and Signal Processing (ICCSP) (pp. 0769-0773). IEEE.
- [44] Sahu, S., Chawla, M., & Khare, N. (2017, May). An efficient analysis of crop yield prediction using the Hadoop framework based on random forest approach. In 2017 International Conference on Computing, Communication and Automation (ICCCA) (pp. 53-57). IEEE.
- [45] Abdulkareem, N. M., & Abdulazeez, A. M. (2021). Machine Learning Classification Based on Radom Forest Algorithm: A Review. *International Journal of Science and Business*, 5(2), 128-142.
- [46] Abbas, F., Afzaal, H., Farooque, A. A., & Tang, S. (2020). Crop yield prediction through proximal sensing and machine learning algorithms. *Agronomy*, 10(7), 1046.
- [47] D. Q. Zeebaree, A. M. Abdulazeez, D. A. Zebari, H. Haron and H. Nuzly, "Multi-level fusion in ultrasound for cancer detection based on uniform LBP features," *Computers, Materials & Continua*, vol. 66, no.3, pp. 3363–3382, 2021.
- [48] Abdulqader, D. M., Abdulazeez, A. M., & Zeebaree, D. Q. (2020). Machine Learning Supervised Algorithms of Gene Selection: A Review. *Machine Learning*, 62(03).
- [49] Nathgosavi, V., & Patil, S. (2021). A Survey on Crop Yield Prediction using Machine Learning (No. 5238). *EasyChair*.
- [50] Jahwar, A. F., & Abdulazeez, A. M. (2020). Meta-Heuristic Algorithms For K-Means Clustering: A Review. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 17(7), 12002-12020.
- [51] Ronan Collobert, Samy Bengio, and C. Williamson. "Svm torch: Support vector machines for large-scale regression problems". *JMLR, Journal of Machine Learning Research*, 1:143–160, 2001.
- [52] Yahia, H. S., & Abdulazeez, A. M. (2021). Medical Text Classification Based on Convolutional Neural Network: A Review. *International Journal of Science and Business*, 5(3), 27-41.
- [53] Veenadhari, S., Misra, B., & Singh, C. D. (2014, January). Machine learning approach for forecasting crop yield based on climatic parameters. In 2014 International Conference on Computer Communication and Informatics (pp. 1-5). IEEE.
- [54] Dahikar, S. S., & Rode, S. V. (2014). Agricultural crop yield prediction using an artificial neural network approach. *International journal of innovative research in electrical, electronics, instrumentation and control engineering*, 2(1), 683-686.
- [55] Safa, M., Samarasinghe, S., & Nejat, M. (2015). Prediction of wheat production using artificial neural networks and investigating indirect factors affecting it: a case study in Canterbury province, New Zealand. *J. Agr. Sci. Technology*, 2015, Vol. 17, 791-803.
- [56] Palanivel, K., & Surianarayanan, C. (2019). An approach for prediction of crop yield using machine learning and big data techniques. *International Journal of Computer Engineering and Technology*, 10(3), 110-118.
- [57] Zeebaree, D. Q., Abdulazeez, A. M., Hassan, O. M. S., Zebari, D. A., & Saeed, J. N. (2020). Hiding Image by Using Contourlet Transform.