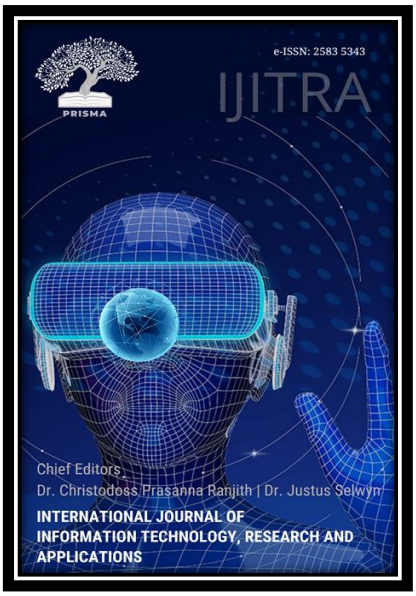


International Journal of Information Technology, Research and Applications (IJITRA)

Mahesh T R, Sindhu Madhuri G.
Prediction of Crop Yield Based-on Soil Moisture using Machine Learning Algorithms. International Journal of Information Technology, Research and Applications, 2(1), 33-41.

ISSN: 2583-5343

DOI: 10.59461/ijitra.v2i1.30

The online version of this article can be found at:
https://www.ijitra.com/index.php/ijitra/issue/archive

Published by:
PRISMA Publications

International Journal of Information Technology, Research and Applications (IJITRA) is a journal that publishes articles which contribute new theoretical results in all the areas of Computer Science, Communication Network and Information Technology. Research paper and articles on Big Data, Machine Learning, IOT, Blockchain, Network Security, Optical Integrated Circuits, and Artificial Intelligence are in prime position.

PRISMA
PUBLICATIONS

https://www.prismapublications.com/

Prediction of Crop Yield Based-on Soil Moisture using Machine Learning Algorithms

Mahesh T R[1], Sindhu Madhuri G[1]
[1] Department of Computer Science and Engineering,
JAIN (Deemed-to-be University), Bangalore, India

Article Info

Article history:

Received October 25, 2022
Revised November 19, 2022
Accepted November 19, 2022

Keywords:

Detection
Machine Learning
Prediction

ABSTRACT

Agriculture planning is playing an important role as the economic growth is very high day by day in our daily life. There is lot of research study going on as there are few important issues like soil nutrients, crop prediction, farming system, crop monitoring in agriculture with modern farming system. Based on various parameters, farming issues and farming system, there is lot of change in production rate and market prices. Crop prediction and crop monitoring is main factor to produce good quality of crops for farmers to predict crop yield based on soil moisture. Prediction of crop yield includes forecasting factors like temperature, humidity, rainfall, etc., and crop yield based on soil moisture includes few measures like pH, NPK (Nitrogen, Phosphorous and potassium) values using various sensors. Farmers can predict or come to a decision the type of soil moisture values, farmers can decide the type of crop to be planted. In this paper, we proposed decision tree supervised machine learning algorithm to improve our results for the prediction of crop yield based on soil moisture parameters to achieve economic growth for achievement of better results.

Corresponding Author:

Mahesh T R
Department of Computer Science and Engineering,
JAIN(Deemed-to-be University),
Bengaluru,
India.
Email: trmahesh.1978@mail.com

I. INTRODUCTION

Agriculture is the primary source of trade which forms the backbone in many different countries. Many challenges currently in agriculture field arises due to traditional farming system necessitate farmers to be equipped with modern farming system. For the improvement of production of crop, soil moisture data are used. Soil moisture nutrients are essential for better enhancement of crop growth and plant tissues [1]. Crop yield prediction which maximizes the production rate of crop is one of the interesting research work carried out by enhancing soil moisture data for producing efficient and accurate results. Crop prediction depends on important parameters of a region like – many harvesting techniques, weather conditions (such as cloud, rainfall, temperature, humidity etc), soil moisture type (such as clay, sandy, saline soil, etc), soil composition (such as pH value, potassium, nitrogen, phosphorous, etc) and many more [2, 3]. The soil moisture data can be enhanced or analyzed by using various machine learning algorithms. The system is created using machine learning algorithm and it predicts appropriate crops to the farmers depending on the quality of the soil in the region. There are many machine learning algorithms for agriculture applications like artificial neural networks to prediction of wheat yield using climate data, linear regression to set the target crop yields, support vector machine for crop yield prediction and so on [4].

In this research paper, we use decision tree machine learning algorithm which helps to predict crop yield in agriculture that can help framers. It also helps to suggest the crop growth according to soil moisture data. The model predicts and suggests most accurate crop based on the soil attributes. This prediction will help farmer

to predict the yield of the crop before cultivating and thus help them to make appropriate decisions. This machine learning technique provides efficient framework for data driven decision making. This model provides set machine learning techniques to help farmers grow crops depending on their field important attributes [5]. In this research work, few machine learning algorithms are discussed in literature survey chapter, further we highlighted our proposed system in methodology chapter, and compared our analysis in results chapter.

II. LITERATURE SURVEY

As there are many machine learning algorithms based on various applications. In this chapter, we discuss machine learning algorithms used for agriculture applications. Artificial Neural Network (ANN) machine learning algorithm mainly works based on weight processing parameter. It mainly used as topological-based algorithm in a multi-layered and back-propagation weighted technique for the implementation of crop yield prediction based on quality of soil [6, 7]. But as this ANN algorithm is applicable to number of input attributes are very less, so there is no automatic algorithm to predict the suitable crop for sample data space.

Researchers extended working on agriculture application using Support Vector Machine (SVM) algorithm. This algorithm is also known as support vector regression mainly used for crop prediction. The main goal by means of this algorithm is to achieve non-linear function using radial-basis kernel function or polynomial regression function. But it basically tries to avoid complexity of linear function by accepting large input sample space for the optimization of complex problems [8, 9].

K-Nearest Neighbors (K-NN) Machine learning algorithm is also used in agriculture application by predicting and analyzing crop yield sample data. It holds the previous sample space data for prediction of target value. By applying the distance measures between the train sample inputs and corresponding target data, the target results of k-neighbors values are found. But it does not perform training and optimization procedure for prediction of crop yield based on soil moisture values, so time and space complexity are high using this K-NN algorithm [10, 11].

Based on the survey, observations & few limitations, we extended our research work for the prediction of crop yield based on soil moisture using decision tree machine learning algorithm [12] for achieving better accuracy results.

III. METHODOLOGY

In our proposed system, we use efficient machine learning algorithm for crop prediction system based on soil moisture nutrient values.

In our proposed system, to determine the fertility of the given region, we know crops rely on different soil moisture data set. For example: Potatoes are most viable where the soil moisture pH level between 5.5 - 6.5 range, while Maize requires pH level between 6.0 - 7.5 range of optimum growth etc. The quality of the soil moisture can be analyzed using few important nutrient attributes present in the soil such as Nitrogen (N), Potassium (k), Ca (Calcium), Phosphorus (P), pH value. To achieve the nutrient food and different crops to grow on different kind of soil, we consider soil moisture as important parameter to determine what type of crop growth will produce maximum yield in the given region. The block diagram of our proposed system: is shown in figure 1.

Table-1: Soil Moisture Data Set

N	P	K	PH
226.05	18	142	7.86847465
308.25	298	339.35	5.00530698
277.42	48	514.29	6.13613187
236.32	46	496.98	7.52408008
226.05	42	609.52	5.70694325
287.7	31	576.73	5.75850632
226.05	38	273.57	6.34960633
205.5	36	316.11	5.99996903
267.15	64	521.02	5.45359203
308.25	23	369.02	6.37257633
256.77	21	364.64	5.10558836
205.05	108	610.03	5.93534441
267.15	72	578.41	5.17782304
267.15	58	242.89	6.46239161
205.05	28	225.64	7.06917223
308.25	40	365.12	7.47313438
287.7	91	512.4	6.15837697
236.05	53	600.38	7.54987368
256.87	235	469.67	6.02832156
236.32	38	390.47	6.60499348

The proposed system steps are as follows:
Step 1: We used two datasets as a sample: the first dataset is a soil moisture dataset (shown in Table -1) and crop dataset (shown in below Table-2):

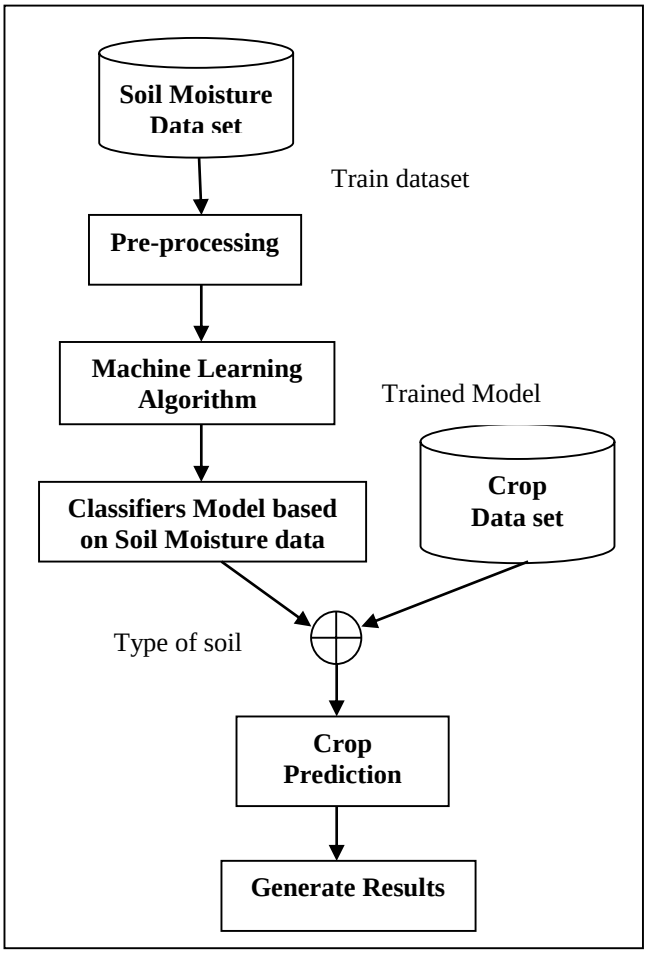


Figure 1: Block Diagram of Proposed System

Table-2: Crop Data Set

Amt of 'N'	Amt of 'P'	Amt of 'K'	Amt of 'pH'	Crop
90	42	43	6.50298529	rice
85	58	41	7.03809636	rice
60	55	44	7.84020714	rice
74	35	40	6.98040091	rice
78	42	42	7.62847289	rice
71	54	16	5.74991442	maize
61	44	17	6.93175656	maize
80	43	16	6.65796475	maize
73	58	21	6.59606065	maize
72	55	17	6.74991442	maize
50	56	76	7.96660503	chickpea
39	71	84	8.14082544	chickpea
25	78	76	7.22896345	chickpea
31	70	77	6.49254605	chickpea
26	80	83	7.52859996	chickpea
18	79	20	5.87734752	kidneybeans
21	63	17	5.97997397	kidneybeans
24	80	22	5.63599397	kidneybeans
34	60	22	5.63523178	kidneybeans
16	75	21	5.67922435	kidneybeans

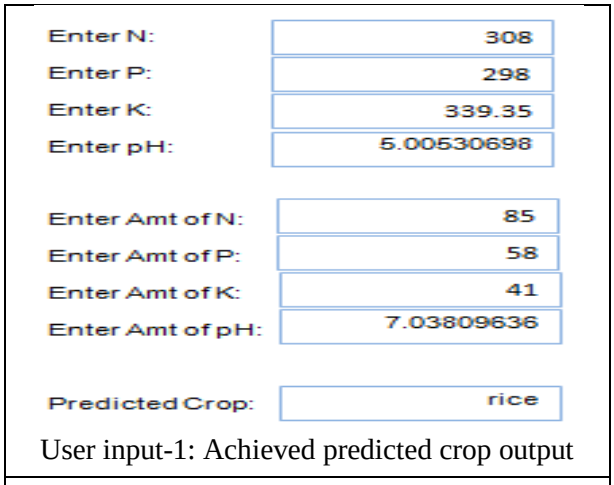
The soil moisture dataset shown in Table-1 consists of input nutrient attribute parameters like: nitrogen (N), Phosphorus (P), potassium (K), pH values dataset. And the crop dataset shown in Table-2 consists of important factors like: type of crop yield prediction like rice, maize, etc based on amount of N, P, K, pH values. The main aim of our proposed system is to build a user-friendly farming system based on soil parameters soil and crop sample dataset are taken from online sources:

Step 2: The sample soil moisture and crop dataset is cleaned and noise data or missing data is been removed by converting into integer. In pre-processing step, the dataset is loaded to extract the noisy data, duplicate and replaces with the converted data into integers. This pre-processed data will be further stimulated to remove duplicate data from the dataset and will split the dataset into training and testing dataset. When splitting the dataset into test data and train pre-processed data, we consider test dataset as 0.2 (i.e., 20%) and train dataset as 0.9 (i.e., 90%).

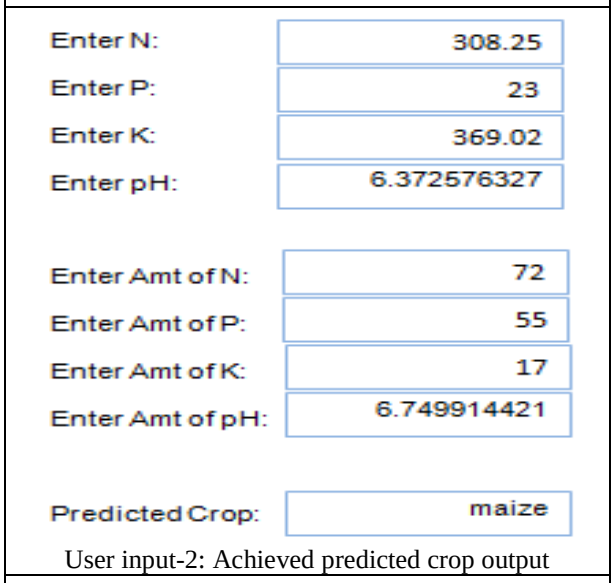
- Step 3: Our datasets are in .CSV format and once we fetch or import the soil moisture dataset and crop dataset, further preprocessed, then our dataset is ready to train and test the model. We perform decision tree machine learning algorithm for our sample datasets to achieve effective prediction of crop yield. This decision tree algorithm forms many decision trees based on count value and maximum number of trees our algorithm measures accurately to provide same output results. During this testing and training process, we divide the dataset to 90% for training data and 20% for test the dataset. Decision Tree algorithm belongs to the family of supervised learning algorithms. The goal of using a decision tree machine learning algorithm is to create a training model that can use to predict the class or value of the target variable by following simple decision learning rules taken from previous dataset as training set. In this algorithm, for predicting a class label for a record, we start from the root of the tree. We compare the values of the root attribute with the record's attribute. On the basis of comparison, we follow the branch corresponding to that value and jump to the next node. This algorithm observes the difference between expected outputs with the sample dataset and updates the weights between the nodes. This decision tree algorithm model is trained with much many iterations until we achieve desired output. After training the dataset, the generalized model based on the test data, are compared to check the prediction of crop yield based on soil moisture data is accurate or not and the results are discussed in the next section..

IV. RESULTS

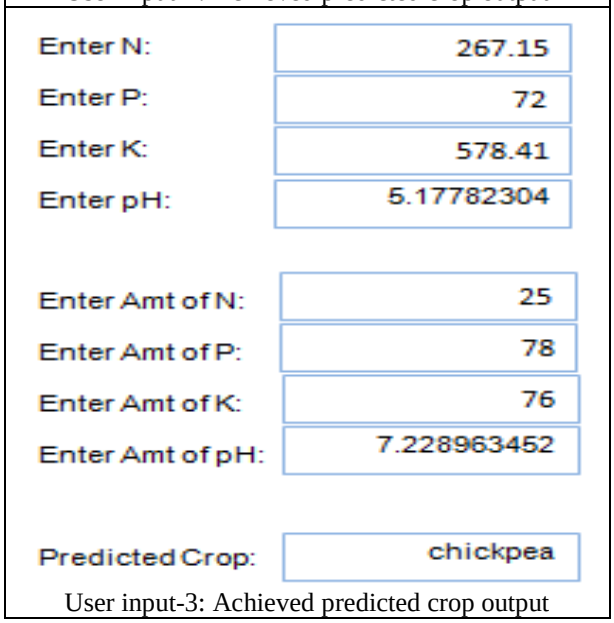
Our research effort is carried out to predict the crop yield production by analyzing and processing the datasets by implementing in Python programming using Anaconda software. The results are shown below Figure 2:



User input-1: Achieved predicted crop output



User input-2: Achieved predicted crop output



User input-3: Achieved predicted crop output

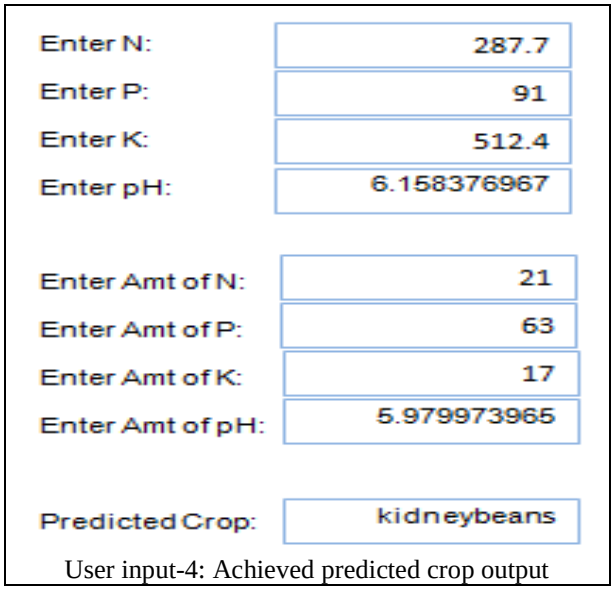


Figure 2: Achieved results with different user inputs

Using machine learning decision tree algorithm for prediction of crop yield based on soil moisture data gives better results. We achieved performance measures error rate as 0.5 and accuracy as 95% in our proposed system is shown in below figure 3:

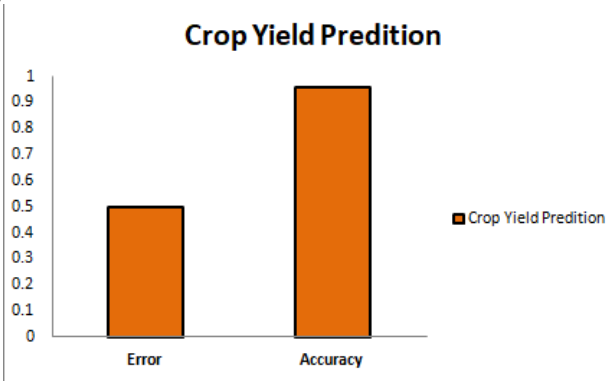


Figure 3: Performance measures: error and accuracy

V. CONCLUSION

This proposed model helps farmers prioritize a crop yield that is most viable in the land they own, so they get maximum yield of the crops. India is a country where 50% of the workforce is directly working in the field of agriculture but its contribution to the GDP is only about 17%. So, using this system, we will be directly helping the economic level of farming in the country. We extend our work for automatic detection of crop yield based on soil moisture data by considering more parameters like weather forecast, soil testing and so on to help farmers for better growth of crop yield.

VI. REFERENCES

[1] H. K. Shashikala, T. R. Mahesh, V. Vivek, M. G. Sindhu, C. Saravanan and T. Z. Baig, "Early Detection of Spondylosis using Point-Based Image Processing Techniques," 2021 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), 2021, pp. 655-659, doi: 10.1109/RTEICT52294.2021.9573604.

[2] Gowramma, G. S., Mahesh, T. R., & Gowda, G. (2017). An automatic system for IVF data classification by utilizing multilayer perceptron algorithm. *ICCTEST-2017*, *2*, 667-672.

[3] M. R. Sarveshvar, A. Gogoi, A. K. Chaubey, S. Rohit and T. R. Mahesh, "Performance of different Machine Learning Techniques for the Prediction of Heart Diseases," 2021 International Conference on

Forensics, Analytics, Big Data, Security (FABS), 2021, pp. 1-4, doi: 10.1109/FABS52071.2021.9702566.

[4] Dharahas, R. T., & Mahesh, T. R. (2021). A Pragmatic Approach for Detecting Brain Tumors Using Machine Learning Algorithms. *BIOSCIENCE BIOTECHNOLOGY RESEARCH COMMUNICATIONS Special Issue*, *14*(11).

[5] Pinaki, G., & Mahesh, T. R. (2015). Smart city: Concept and challenges. *International Journal on Advances in Engineering Technology and Science*, *1*(1).

[6] P. Chaitanya Reddy, R. M. S. Chandra, P. Vadiraj, M. Ayyappa Reddy, T. R. Mahesh and G. Sindhu Madhuri, "Detection of Plant Leaf-based Diseases Using Machine Learning Approach," 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), 2021, pp. 1-4, doi: 10.1109/CSITSS54238.2021.9683020.

[7] R. Pasumarty, R. Praveen and M. T. R, "The Future of AI-enabled servers in the cloud- A Survey," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2021, pp. 578-583, doi: 10.1109/I-SMAC52330.2021.9640925.

[8] Mahesh, T. R., & Naik, D. M. K. (2020). Analysis of Academic Performance in massive Open Online Courses (Moocs) Using Process mining. *International Journal of Computer Trends and Technology*, *68*(12), 21-25.

[9] Mahesh, T.R., Vinoth Kumar, V., Vivek, V. *et al.* Early predictive model for breast cancer classification using blended ensemble learning. *Int J Syst Assur Eng Manag* (2022). https://doi.org/10.1007/s13198-022-01696-0

[10] T. R. Mahesh, V. Vivek, V. V. Kumar, R. Natarajan, S. Sathya and S. Kanimozhi, "A Comparative Performance Analysis of Machine Learning Approaches for the Early Prediction of Diabetes Disease," 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), 2022, pp. 1-6, doi: 10.1109/ACCAI53970.2022.9752543.

[11] P. Shrestha, A. Singh, R. Garg, I. Sarraf, T. R. Mahesh and G. Sindhu Madhuri, "Early Stage Detection of Scoliosis Using Machine Learning Algorithms," 2021 International Conference on Forensics, Analytics, Big Data, Security (FABS), 2021, pp. 1-4, doi: 10.1109/FABS52071.2021.9702699.

[12] K. K. Jha, R. Jha, A. K. Jha, M. A. M. Hassan, S. K. Yadav and T. Mahesh, "A Brief Comparison On Machine Learning Algorithms Based On Various Applications: A Comprehensive Survey," 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), 2021, pp. 1-5, doi: 10.1109/CSITSS54238.2021.9683524.

[13] S C, Asha and S C, Asha and TR, Mahesh and V, Vivek and M B, Suresh, The Importance of Teacher's Mental Health and Wellness for Quality Learning in Classrooms during COVID-19 Pandemic (April 30, 2021). Available at SSRN: https://ssrn.com/abstract=3837304 or http://dx.doi.org/10.2139/ssrn.3837304

[14] Mahesh, T. R., & Naik, D. M. K. (2021). A Comprehensive Review of Behavioral Customer Segmentation For A Better Understanding. *International Journal of Computer Science and Engineering*, *8*(1), 1-4.

[15] Tarun, R. R., Sahana, J. S., Sadvik, B. S., Shashank, S., & Mahesh, T. R. (2019). Context based Sentiment Analysis of Twitter using Hadoop Framework. *International Journal of Computer Science and Mobile Computing*, *8*(5), 193-202.

[16] Subashini, S., & Mahesh, T. R. Web Mining: Prominent Applications and Future Directions. *International Journal of Computer Science and Information Technology & Security*, 825-830.

[17] S. Roopashree, J. Anitha, T.R. Mahesh, V. Vinoth Kumar, Wattana Viriyasitavat, Amandeep Kaur, An IoT based authentication system for therapeutic herbs measured by local descriptors using machine learning approach, Measurement, Volume 200, 2022, 111484, ISSN 0263-2241, https://doi.org/10.1016/j.measurement.2022.111484.

[18] Mahesh, T. R., Kumar, D., Vinoth Kumar, V., Asghar, J., Mekcha Bazezew, B., Natarajan, R., & Vivek, V. (2022). Blended Ensemble Learning Prediction Model for Strengthening Diagnosis and Treatment of Chronic Diabetes Disease. *Computational Intelligence and Neuroscience*, *2022*.

[19] A. Srivastava, V. V. Kumar, M. T. R and V. Vivek, "Automated Prediction of Liver Disease using Machine Learning (ML) Algorithms," 2022 Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), 2022, pp. 1-4, doi: 10.1109/ICAECT54875.2022.9808059.

[20] S. Surana, K. Pathak, M. Gagnani, V. Shrivastava, M. T. R and S. Madhuri G, "Text Extraction and Detection from Images using Machine Learning Techniques: A Research Review," 2022 International Conference on Electronics and Renewable Systems (ICEARS), 2022, pp. 1201-1207, doi: 10.1109/ICEARS53579.2022.9752274.

[21] Mahesh, T. R., Krishna, G. V., Sathwik, P., Chowdary, V. A., & Hemchand, G. (2022). Providing Voice to Susceptible Children: Depression and Anxiety Detected with the Help of Machine Learning. In *Integrated Emerging Methods of Artificial Intelligence & Cloud Computing* (pp. 444-450). Springer, Cham.

[22] Mahesh, T. R., Ram, M. S., Ram, N., Gowtham, A., & Swamy, T. V. (2022). Real-Time Eye Blinking for Password Authentication. In *Integrated Emerging Methods of Artificial Intelligence & Cloud Computing* (pp. 428-434). Springer, Cham.

[23] K. K. Jha, A. K. Jha, K. Rathore and T. R. Mahesh, "Forecasting of Heart Diseases in Early Stages Using Machine Learning Approaches," 2021 International Conference on Forensics, Analytics, Big Data, Security (FABS), 2021, pp. 1-5, doi: 10.1109/FABS52071.2021.9702665.

[24] TR, Mahesh and V, Vivek, Recommendation Systems: The Different Filtering Techniques, Challenges and Review Ways to Measure the Recommender System (April 14, 2021). Available at SSRN: https://ssrn.com/abstract=3826124 or http://dx.doi.org/10.2139/ssrn.3826124

[25] V. Vivek; T. R. Mahesh; C. Saravanan; K. Vinay Kumar, "5 A Novel Technique for User Decision Prediction and Assistance Using Machine Learning and NLP: A Model to Transform the E-commerce System," in Big Data Management in Sensing: Applications in AI and IoT , River Publishers, 2021, pp.61-76.

[26] G. Sindhu Madhuri; T. R. Mahesh; V. Vivek, "7 A Novel Approach for Automatic Brain Tumor Detection Using Machine Learning Algorithms," in Big Data Management in Sensing: Applications in AI and IoT , River Publishers, 2021, pp.87-102.

[27] G. Sindhu Madhuri; T. R. Mahesh; V. Vivek, "7 A Novel Approach for Automatic Brain Tumor Detection Using Machine Learning Algorithms," in Big Data Management in Sensing: Applications in AI and IoT , River Publishers, 2021, pp.87-102.

[28] V, Vivek and TR, Mahesh and Das, Shilpa, A Security Framework for Application Instance Data to Enhance the Privacy in Decentralized Cloud Environment (October 21, 2020). Available at SSRN: https://ssrn.com/abstract=3715953

[29] V, Vivek and TR, Mahesh and Das, Shilpa, A Security Framework for Application Instance Data to Enhance the Privacy in Decentralized Cloud Environment (October 21, 2020). Available at SSRN: https://ssrn.com/abstract=3715953

[30] Ghosh, P., & Mahesh, T. R. (2019). Untraceable privacy-preserving authentication protocol for RFID tag using salted hash algorithm. International Journal of Advanced Intelligence Paradigms, 13(1-2), 193-209.

[31] Aja-Fernández S., Curiale A. H., Vegas-Sánchez-Ferrero G., A local fuzzy thresholding methodology for multiregion image segmentation. Knowledge-Based Systems, 83:1–2. 2015.

[32] Patel K., Jha J., Brain tumor image segmentation using adaptive clustering and level set method, image, 9, 2014.

[33] Jaganathan P., Kuppuchamy R. A threshold fuzzy entropy based feature selection for medical database classification. Computers in Biology and Medicine, 43, 2222–2229, 2013.

[34] Sujan M., Alam N., Noman S. A., Islam M.J.. A segmentation based Automated System for Brain Tumor Detection. International Journal of Computer Applications, 153:41–9, 2016.

[35] Ilhan U., Ilhan A.. Brain tumor segmentation based on a new threshold approach, Procedia Computer Science, 120:580–587, 2017.

[36] DongjuLiu, JianYu, “Otsu method and K-means”, 978-0-7695-3745-0/09, IEEE. DOI 10.1109/HIS.2009.74, 2009.

[37] Aimi Salihah Abdul-Nasir1 , Mohd Yusoff Mashor2 , Zeehaida Mohamed3 , " Colour Image Segmentation Approach for Detection of Malaria Parasites Using Various Colour Models and k-Means Clustering", Issue 1, Volume 10, January 2013.
[38] Juanying Xie, Shuai Jiang, "A simple and fast algorithm for global K-means clustering", 978-0-7695-3987-4/10, IEEE, 2010.
[39] Rafael C. Gonzalez & Richard E. Woods, "Digital Image Processing using Matlab", Third edition: Pearson education, 2005.
[40] Du, Cheng-Jin, and Da-Wen Sun. 2004. "Recent developments in the applications of image processing techniques for food quality evaluation." Review of. Trends in food science & technology, 15, (5):230-49.
[41] Du, Cheng-Jin, and Da-Wen Sun. 2004. "Recent developments in the applications of image processing techniques for food quality evaluation." Review of. Trends in food science & technology 15 (5):230-49.
[42] Shashikala H K, Sindhu Madhuri G, "Image pre-processing techniques for X-ray medical images: A Survey", International Journal of Creative Research Thoughts (IJCRT), ISSN: 2320-2882, Vol 9, Issue 1, January 2021.
[43] Sindhu Madhuri G, Shashikala H K, "Image Processing techniques for Detecting Extra Growth of Teeth in Medical Images", Solid State Technology, Vol 64, Issue 2, Jan 2021.