World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(REVIEW ARTICLE)

Check for updates

Classification of pH scale based on machine learning approaches

Sasmita Kumari Nayak *

Computer Science and Engineering, Centurion University of Technology and Management, Odisha, India.

World Journal of Advanced Research and Reviews, 2024, 21(01), 235–239

Publication history: Received on 22 November 2023; revised on 01 January 2024; accepted on 03 January 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.21.1.0014

Abstract

This paper provides a demonstration of the idea for a microfluidic, which is based on dry chemicals, stable, and semi-quantitative assay using a larger dataset with a variety of conditions. An optical method is used to measure the color change of pH paper. In order to support the claim and provide evidence, this paper will specifically explore the parameters of an intelligent colorimetric test that satisfies the ASSURED standards. All the time it measures by using traditional manual method. Our goal in this study was to gather objective data by using Red, Green, and Blue values to represent the pH change of the paper. The system under investigation is an intelligent image-based system that performs automatic paper-based colorimetric tests in real-time. This paper classifies the pH scale by using machine learning models.

Keywords: Image processing; Feature Extraction; pH paper image data; Machine Learning; KNN; Decision Tree; Random Forest.

1. Introduction

Machine learning is an artificial intelligence approach and a subfield of computer science. Machine Learning is used in varied fields, which has been discussed in these papers [1-12]. The number of pH changes according to how much of the substance dissolves in water and releases hydrogen ions. In this instance, more hydrogen ions are produced at lower pH values. Consequently, the acid becomes stronger the lower the pH value. A measurement called pH is used to show which solution is more basic or acidic. It is the exponent for hydrogen ion concentration. Definition is the numerical value of the concentration of hydrogen ions (H +) in a given volume of solution. To differentiate between acid and base, pH is measured for the most fundamental purpose. All acidic materials are aqueous solutions that typically dissolve in water to release hydrogen ions (H +) and have a pH of less than 7. It tastes bitter and is slippery as a result. Neutral pH is 7 [1, 2]. This nature gives rise to the sour taste characteristic. Conversely, all bases dissolve in water to produce hydroxide ions and have values greater than pH 7. In paper [13], examined numerical solution of unsteady MHD natural convection flow of a viscous, incompressible, electrically conducting and heat absorbing fluid past an impulsively moving vertical plate with ramped temperature embedded in a porous medium in the presence of thermal diffusion is carried out. In paper [14], considered MHD flow and heat transfer of nanofluid over a stretchable surface with melting, where chemical reaction effects discussed. In paper [15-19], have discussed the effects of hall current and rotation on steady MHD Couette flow of Class-II of a viscous, incompressible and electrically conducting fluid in the presence of a uniform transverse magnetic field. In every field, the machine learning has been used.

In this study, we aimed to obtain objective data by indicating the change of pH paper according to pH as Red, Green, and Blue values. Here, pH scale recognition using machine learning regression model based upon the color model RGB and web application using flask.

* Corresponding author: Sasmita Kumari Nayak

Using a pH paper that changes color based on pH is an easy way to measure pH. These methods are quick and easy to use, but they are not quantitative values; rather, they are subjective results of the measure, which makes it difficult to objectively distinguish minute differences in color and results in poor accuracy.

This article explains the initial attempts to explore the pH scale identification using machine learning algorithms. Section 4 concludes by classifying the pH s of machine learning algorithms in an effort to offer a better solution.

2. Machine Learning Model

Machine learning is an artificial intelligence approach and a subfield of computer science [20-30]. This method has the benefit of allowing a model to address issues that are not amenable to explicit algorithms, and it can be applied in multiple domains. A thorough analysis of various deterministic and machine learning techniques for predicting food, crop yield, weather, and hepatitis is provided in [31-38]. Even in situations where representation is not feasible, machine learning models identify relationships between inputs and outputs; this characteristic allow the use of machine learning models in many cases, for example in data mining and forecasting problems, spam filtering, classification problems, and pattern recognition. Because one must work with large datasets and machine learning models can handle pre-processing and data preparation, the classification and data mining aspects of this field are especially intriguing. Following this stage, forecasting issues can be solved using the machine learning models.

2.1. Decision Tree

A decision tree is a type of decision support tool that shows potential decisions, outcomes, or reactions using a model or chart that resembles a tree [39-42].

2.2. Random Forest

The random forest is a model made up of many decision trees. This algorithm combines the output of multiple decision trees to generate the final output [24, 40-42].

2.3. K-Nearest Neighbors (KNN)

One of the most basic machine learning algorithms for regression and classification problems is K-Nearest Neighbors (KNN). Using similarity metrics, KNN algorithms classify new data points using existing data. An efficient machine learning technique for both regression and classification problems is K-nearest neighbors [7]. The idea behind KNN is to use the distance between an unknown sample and the K closest samples in the training set to classify it. The process of classification involves designating the most prevalent class among the K closest neighbors. Because KNN is a lazy learning algorithm, all that needs to be done for training is storing the training data. KNN is quick and memory-efficient because the real classification or regression of fresh samples is done at the prediction stage. KNN can handle both linear and nonlinear data, and it is simple to comprehend and apply. Nevertheless, the selection of K, the size of the features, and features that are not relevant can affect how well it performs. Unlabeled observations are classified using a KNN classifier by putting them in the same class as the labeled examples that are the closest to them. Both the training and test dataset's characteristics are gathered. For instance, the crunchiness and sweetness of fruit, vegetables, and grains can be used to identify them [43-49].

3. Results and Discussion

For this project, gathering datasets is essential. The data has collected from Kaggle, which contains four attributes such as Blue, green, Red and Scale. Our goal in using this dataset is to classify the pH scale, so here tried a number of different algorithms by using python to run the program for experimental purposes. Regression modeling has been used by many to analyze this dataset, but since pH scale images are in categorical format, we considered classification [32]. The accuracy of all machine learning model for classifying the pH scale images is shown in Table 1.

Table 1 Classification Results Using Machine Learning Models

Machine Learning Classification Models	Accuracy of the Model (%)
KNN	77%
Decision Tree	92%
Random Forest	84%

From the above table, we can easily conclude that Random Forest model has the best accuracy in comparison with other implemented machine learning models. The accuracy score of Decision Tree as well as accuracy score also found out to be maximum.

4. Conclusion

The machine learning models were chosen based on how well the model would perform in a binary classification when compared to other experiments under similar conditions. This paper is a companion work to an experiment conducted on a dataset containing pH scale image of interest data. Features seen from the pH image are included in the dataset. Cross-validation was used to run the models flawlessly and determine the ideal meta-parameters for each model. From the result its shows that Decision tree model works well to recognize and classify the pH.

References

[1] Tania, M. H., Lwin, K. T., Shabut, A. M., Najlah, M., Chin, J., & Hossain, M. A. (2020). Intelligent image-based colourimetric tests using machine learning framework for lateral flow assays. Expert Systems with Applications, 139, 112843.

[2] Contreras-naranjo, J. C., Wei, Q., & Ozcan, A. (2016). Mobile phone-based microscopy, sensing, and diagnostics. IEEE Journal of Selected Topics in Quantum Electronics, 22(3), 1–14 [online] Available at: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7265008.

[3] Siddique, M., Panda, D., A hybrid forecasting model for prediction of stock index of tata motors using principal component analysis, support vector regression and particle swarm optimization International Journal of Engineering and Advanced Technology, 2019, 9(1), pp. 3032–3037

[4] Siddique, M; Panda, D., Das, S., Mohapatra, S: Hybrid forecasting model for stock value prediction using soft computing technique, International Journal of Pure and Applied Mathematics, 2019, Vol. 117 (19); 357-363. ISSN:1311-8080

[5] Mishra, S.P., Nayak, S.P., Mishra, S., Siddique, M., Sethi, K.C.: GIS and auto desk modeling for satellite cities around Bhubaneswar, International Journal of Innovative Technology and Exploring Engineering, 2019, 8(11), pp. 297–306.

[6] Siddique, M., Mohanty, S., Panda, D., A hybrid forecasting model for prediction of stock value of Tata steel using support vector regression and particle swarm optimization, International Journal of Pure and Applied Mathematics, 2019, 119 (14): 1719-1727.

[7] Mishra, S.P., Sethi, K.C., Mishra, D.P, Siddique, M., Pre-Monsoon Cyclogenesis over Bay of Bengal, International Journal of Recent Technology and Engineering, 2019, 8 (2):4895-4908.

[8] Pattnaik,R.K, Siddique, M, Mishra, S, Gelmecha, D.J, Singh, R.S, Satapathy,S., Breast cancer detection and classification using metaheuristic optimized ensemble extreme learning machine, International Journal of Information Technology, Sringer Nature Singapore, 2023 15 (8), 4551-4563.

[9] Siddique, M., Panda, D., Prediction of stock index of tata steel using hybrid machine learning based optimization techniques, International Journal of Recent Technology and Engineering, 2019, 8(2), pp. 3186–3193.

[10] Mohapatra, S.K., Prasad, S., Habtemariam, G.M., Siddique, M., Identify determinants of infant and child mortality based using machine learning: Case study on Ethiopia, Big Data Analytics and Machine Intelligence in Biomedical and Health Informatics: Concepts, Methodologies, Tools and Applications, 2022, 21–45

[11] Mohapatra,S., Siddique,M., Paikaray, B.K, Riyazbanu, S., Automated Invasive Cervical Cancer Disease Detection at Early Stage Through Deep Learning, International Journal of Bioinformatics Research and Applications, 2023, 19(4), 306-326.

[12] Pattanaik, R.K., Mishra, S., Siddique, M., Gopikrishna, T., Satapathy, S.Breast Cancer Classification from Mammogram Images Using Extreme Learning Machine-Based DenseNet121 Model, Journal of Sensors, Hindawi Publication, ISSN No: 1687-725X, Volume 2022, Article ID 2731364, page:1-12, 2022.

[13] Seth GS, Sarkar S and Mahato GK. Numerical solution of unsteady hydromagnetic natural convection flow of heat absorbing fluid past an impulsively moving vertical plate with ramped temperature. International Journal of Applied Mathematical Research. 2013; 2 (2): 317-324. DOI: https://doi.org/10.14419/ijamr.v2i2.939

[14] Mahato GK, Mahatha BK, Nandkeolyar R and Patra B. The effects of chemical reaction on magnetohydrodynamic flow and heat transfer of a nanofluid past a stretchable surface with melting. AIP Conference Proceedings. 2020; 2253 (1), 020011-1–020011-13. DOI: https://doi.org/10.1063/5.0019205 , ISBN: 978-0-7354-2012-0

[15] Mahato GK, Mahatha BK and Samal S. Melting Heat Transfer on Magneto- hydrodynamic (MHD) Flow of a Heat Radiating and Chemically Reacting Nano-Fluid past a Stretchable Surface. JP Journal of Heat and Mass Transfer. 2019; 17 (2): 379-398. DOI: http://dx.doi.org/10.17654/HM017020379

[16] Seth GS, Mahato GK and Singh JK. Effects of Hall current and rotation on MHD Couette flow of class-II. Journal of International Academy of Physical Sciences. 2011; 15 (5) Special Issue-1: 213-230.

[17] Seth GS, Singh JK and Mahato GK. Unsteady hydromagnetic Couette flow within a porous channel with Hall effects. International Journal of Engineering. Science and Technology. 2011; 3 (6): 172-183. DOI: 10.4314/ijest.v3i6.14

[18] Seth GS, Singh JK and Mahato GK. Hall effects on unsteady magnetohydrodynamic Couette flow within a porous channel due to accelerated movement of one of its plates. Journal of Nature Science and Sustainable Technology. 2013; 7 (3): 271-290.

[19] Seth GS, Mahato GK and Singh JK. Combined Free and Forced Convection Couette-Hartmann Flow in a Rotating System with Hall Effects. Journal of Nature Science and Sustainable Technology. 2012; 6 (3): 125-150.

[20] Nayak, S. K., Padhy, S. K., & Panigrahi, S. P. (2012). A novel algorithm for dynamic task scheduling. Future Generation Computer Systems, 28(5), 709-717.

[21] Nayak, S. K., Panda, C. S., & Padhy, S. K. (2018). Efficient multiprocessor scheduling using water cycle algorithm. Soft Computing Applications, 131-147.

[22] Nayak, S. K., Panda, C. S., & Padhy, S. K. (2018). Efficient multiprocessor scheduling using water cycle algorithm. In Soft Computing Applications (pp. 131-147). Springer, Singapore.

[23] Nayak S.K., Padhy S.K., Panda C.S. (2018) Efficient Multiprocessor Scheduling Using Water Cycle Algorithm. In: Pant M., Ray K., Sharma T., Rawat S., Bandyopadhyay A. (eds) Soft Computing: Theories and Applications. Advances in Intelligent Systems and Computing, vol 583. Springer, Singapore.

[24] Nayak, S. K., Panda, C. S., & Padhy, S. K. (2019, February). Dynamic Task Scheduling Problem Based on Grey Wolf Optimization Algorithm. In 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP) (pp. 1-5). IEEE.

[25] S. K. Nayak, C. S. Panda and S. K. Padhy, "Dynamic Task Scheduling Problem Based on Grey Wolf Optimization Algorithm," 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), Gangtok, India, 2019, pp. 1-5, doi: 10.1109/ICACCP.2019.8882992.

[26] Nayak, S. K., & Panda, C. S. (2021). Dynamic task scheduling using nature inspired algorithms. J. Math. Comput. Sci., 11(1), 893-913.

[27] Nayak, S. K. Multiprocessor Scheduling Using Nature Inspired Optimization. Thesis. Sambalpur University, Odisha, 2021. Web. http://hdl.handle.net/10603/464925

[28] Nayak, S. K., Panda, C. S., & Padhy, S. K. (2018). Multiprocessor Scheduling using Krill Herd Algorithm (KHA). International Journal of Computer Sciences and Engineering, 6(6), 7-17.

[29] S.K. Nayak, C.S. Panda, S.K. Padhy, "Multiprocessor Scheduling using Krill Herd Algorithm (KHA)," International Journal of Computer Sciences and Engineering, Vol.6, Issue.6, pp.7-17, 2018. CrossRef-DOI: https://doi.org/10.26438/ijcse/v6i6.717

[30] Nayak, S. K., & Panda, C. S. (2019). Multiple Processor Scheduling with Optimum Execution Time and Processor Utilization Based on the SOSA. International Journal of Recent Technology and Engineering, ISSN: 2277-3878. Volume-8, Issue-2, July, 5463-5471.

[31] Swain, S., Nayak, S. K., & Barik, S. S. (2020). A review on plant leaf diseases detection and classification based on machine learning models. Mukt shabd, 9(6), 5195-5205.

[32] S Sucharita, S Nayak, S Panda et al., "Human Face Recognition using LBPH", International Journal of Recent Technology and Engineering, vol. 8, no. 6, pp. 3208-3212, 2020.

[33] Stitiprajna Panda*, Swati Sucharita Barik, Sasmita Kumari Nayak, Aeisuriya Tripathy, and Gourav Mohapatra. 2020. Human Face Recognition using LBPH. International Journal of Recent Technology and Engineering (IJRTE) 8, 6 (2020), 3208-3212. DOI:https://doi.org/10.35940/ijrte.f8117.038620

[34] Jena, T. R., Barik, S. S., & Nayak, S. K. (2020). Electricity consumption & prediction using machine learning models. Acta Tech. Corviniensis-Bull. Eng, 9, 2804-2818.

[35] Nayak, S. K., Barik, S. S., & Beura, M. (2020). Analysis of Infectious Hepatitis Disease with High Accuracy Using Machine Learning Techniques. TEST Engineering & Management, 83, 83.

[36] Nayak, S. K. (2020). Analysis and high accuracy prediction of coconut crop yield production based on principle component analysis with machine learning models. International Journal of Modern Agriculture, 9(4), 359-369.

[37] Sasmita Kumari Nayak, Swati Sucharita Barik, Mamata Beura," Weather Forecasts Based on Rainfall Prediction Using Machine Learning Methodologies," Adalya Journal 9 (6), Page No: 72 – 80, ISSN NO: 1301-2746.

[38] Mishra, S. P., Siddique, M., Beura, M., & Nayak, S. K. (2021). Analysis of Indian Food Based on Machine learning Classification Models. Journal of Scientific Research and Reports, 27(7), 1-7.

[39] Nayak, S. K. (2023). Classification of cyclones using machine learning techniques. World Journal of Advanced Research and Reviews, 20(2), 433-440.

[40] Sasmita Kumari Nayak. (2020). CONSTRICTING THE BIVARIATE ANALYSISOFHIGHEST CROP YIELD PRODUCTION BASED ON DIFFERENT ZONES OF INDIA. International Journal of Modern Agriculture, 9(4), 216 - 226. Retrieved from https://modern-journals.com/index.php/ijma/article/view/205

[41] Sasmita Kumari Nayak, Mohammed Siddique. (2020). EFFECT OF STOCK INDEX PARAMETERS ON FORECASTING THE HIGH STOCK VALUE OF VISA STEEL USING DEEP LEARNING NEURAL NETWORK MODEL. International Journal of Modern Agriculture, 9(4), 227 - 236. Retrieved from https://modern-journals.com/index.php/ijma/article/view/207

[42] Sowmya Jagadeesan, B. B. (2022). A Perishable Food Monitoring Model Based on IoT and Deep Learning to Improve Food Hygiene and Safety Management. IJFANS International Journal of Food and Nutritional Sciences, 11(8), 1164-1178.

[43] Jagadeesan, S., Barman, B., Agarwal, R. K., Srivastava, Y., Singh, B., Nayak, S. K., & Venu, N. A Perishable Food Monitoring Model Based On Iot And Deep Learning To Improve Food Hygiene And Safety Management. interventions, 8, 9.

[44] Rajesh, J., Ashraf, M. S., Kaur, L., Rout, S., Nayak, S. K., Kaur, G., & Saikanth, D. R. K. APPLICATION OF FUZZY LOGIC IN SMART AGRICULTURE TO RECOGNISE TOMATO FRUIT RIPENESS. IJFANS International Journal of Food and Nutritional Sciences, 11(1), 2360-2367.

[45] N. Dash, S. K. Nayak and J. Majumdhar, "Detection of Cut Transition in Videos Using Optical Flow and Clustering," 2021 Asian Conference on Innovation in Technology (ASIANCON), PUNE, India, 2021, pp. 1-7, doi: 10.1109/ASIANCON51346.2021.9544553.

[46] Majumdhar, J., & Nayak, S. K. (2021, August). A Novel Method on Summarization of Video Using Local Ternary Pattern and Local Phase Quantization. In 2021 2nd International Conference on Range Technology (ICORT) (pp. 1-6). IEEE.

[47] Chakravarthy, A., Panda, B. S., & Nayak, S. K. (2023). Review and Comparison for Alzheimer's Disease Detection with Machine Learning Techniques. International Neurourology Journal, 27(4), 403-409.

[48] Nayak, S. K. (2023). Nature inspired algorithms in dynamic task scheduling: A review. World Journal of Advanced Research and Reviews, 20(2), 829–833.

[49] Nayak, S. K. (2023). Exploring and forecasting of solar radiation with machine learning methods. World Journal of Advanced Research and Reviews, 20(2), 824–828.