



OPEN

Facile and highly precise pH-value estimation using common pH paper based on machine learning techniques and supported mobile devices

Mohamed M. Elsenety[✉], Mahmoud Basseem I. Mohamed, Mohamed E. Sultan & Badr A. Elsayed

Numerous scientific, health care, and industrial applications are showing increasing interest in developing optical pH sensors with low-cost, high precision that cover a wide pH range. Although serious efforts, the development of high accuracy and cost-effectiveness, remains challenging. In this perspective, we present the implementation of the machine learning technique on the common pH paper for precise pH-value estimation. Further, we develop a simple, flexible, and free precise mobile application based on a machine learning algorithm to predict the accurate pH value of a solution using an available commercial pH paper. The common light conditions were studied under different light intensities of 350, 200, and 20 Lux. The models were trained using 2689 experimental values without a special instrument control. The pH range of 1: 14 is covered by an interval of ~ 0.1 pH value. The results show a significant relationship between pH values and both the red color and green color, in contrast to the poor correlation by the blue color. The K Neighbors Regressor model improves linearity and shows a significant coefficient of determination of 0.995 combined with the lowest errors. The free, publicly accessible online and mobile application was developed and enables the highly precise estimation of the pH value as a function of the RGB color code of typical pH paper. Our findings could replace higher expensive pH instruments using handheld pH detection, and an intelligent smartphone system for everyone, even the chef in the kitchen, without the need for additional costly and time-consuming experimental work.

The pH value of different solutions is a particularly important point to determine the optimized conditions and quality control for industrial, biological, chemical, and environmental science either outdoor or indoor applications^{1,2}. Hydrogen ions concentration $[H^+]$ denoted to pH scales from 0 to 14, and the most methods counted for detection are complicated, expensive, and time-consuming as microelectrodes³, acid-based indicator⁴, potentiometric titration¹, colorimetric and fluorescence probes application⁵⁻¹⁰. Currently, potentiometric measurements are the most technique used in pH detection. Where the pH of the solution can be calculated by the measurement of the different voltage between the electrodes of the potentiometric device¹¹. Despite, the high accuracy of conventional potentiometric devices, the operation and calibration process is more complicated and costly, which is not applicable to indoor or outdoor purposes. However, the easy and accessible pH strips are used as an alternative method for visual pH detection, but the strips produce lower precise results.

On the other hand, machine learning (ML) techniques give algorithms the ability to predict novel values from training data derived from experiments using Artificial Intelligence (AI). Thus, there are numerous regression or classification algorithms for ML that depend on hyperparameters and mechanisms to achieve their goals and give high performance for planning¹². ML is being used in chemistry such as chemical discovery¹³, molecular representations¹⁴, synthetic chemistry¹⁵, materials chemistry¹⁶, aquatic chemistry research^{17,18}, and water pollution¹⁹.

Department of Chemistry, Faculty of Science, Al-Azhar University, Nasr City, Cairo 11884, Egypt. ✉email: m.elsenety@azhar.edu.eg

Here, the ML technique was used to improve the precision of common strip pH paper. ML models were trained on the 2689 experimental data which covered the whole pH range. Further, we developed a mobile/web application based on ML algorithms to predict the pH values. Therefore, the developed app could work on mobile which could be used as portable devices for anyone (whether a chemist or not) without additional costs, fast response, and is applicable for different applications.

Materials and apparatus

Acetic acid, phosphoric acid, boric acid, HCl, and NaOH were used without any further purification (Sigma Aldrich). Universal indicator pH paper (1–14) Q/3211821AB001-2002 (China). The pH measurements were carried out on a 3520 pH Meter (JENWAY, England). Mastech MS6612 Digital Luxmeter Illuminometer Light (Range Peak 200,000 Lux) was used for measuring the light intensity in the experimental workplace.

pH buffer solution. Universal Britton–Robinson (B–R) buffer was prepared as reported²⁰. Briefly, the stock aqueous B–R buffer solutions (pH = 2.86) by mixing equal molar ratio (1:1:1) of 0.02 mol/L from acetic acid, phosphoric acid, and boric acid. Dropwise of 0.20 mol/L of NaOH or 0.20 mol/L of HCl was used for adjusting the pH values (interval = 0.10) to cover the whole pH range.

Machine learning algorithms. Regression is a technique used for prediction continues pH values learning and figuring out causal relations between the actual and prediction pH values. Eleven supervised machine learning regression models were applied to the data collected and choose the best model that fits with the selected problem, including Linear Regression (LR), Decision Tree Regressor (DT_R), Random Forest Regressor (RT_R), K Neighbors Regressor (KNN_R), Support Vector Regression (SVR), Lasso regression (L₁), Ridge Regression (L₂), Elastic Net regressor (EN_R), AdaBoost Regression (AB_R), Gradient Boosting Regressor (GB_R), and Artificial Neural Network Regressor (ANN_R). All models can be found in Scikit-learn in the class model²¹. In addition, the data visualization of exploratory data analysis and heatmap figures were created using the seaborn package based on python code²².

Metrics for regression. Several metrics were used for evaluating the regression models, coefficient of determination (R²), Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) can be calculated Scikit-learn in class metrics according to Eqs. (1–4).^{23,24}

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (1)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (2)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (4)$$

where N is the number of recorded samples, y_i is the predicted pH value, and \hat{y}_i is the actual pH value.

Automated color-information-extraction from the captured images. To extract the color code (RGB) from images, we used a Python 3.7 code based on the OpenCV package to extract the RGB for each image²⁵. We noted a small deviation of the RGB values at various positions in one image. Thus, the RGB values were estimated at seven distinct (X, Y) positions (10,10;15,15; 20,20; 25,25; 30,30; 35,35; 40,40) to cover the whole image as illustrated in Fig. 1.

pH discrimination with a machine learning model. We exploited a KNN regression model-based machine learning algorithm to study 2689 collected sample data using Python 3.7 and the scikit-learn package^{26,27}. We randomly separated the data into training data (70%, i.e., 1880 samples) and testing data (30%, i.e., 808 samples). In the inference model training phase, the testing data was completely excluded. Furthermore, in machine learning, hyperparameters are those parameters that are explicitly provided by the user to influence the learning process and improve the learning of the model. Thus, we trained our models using a series of integer number [1,2,3,...], and as a result of that the optimal hyperparameters (highest coefficient, and lower errors) was found when we used $K = 5$.

Result and discussion

Figure 2 presents collections of 130 captures of an experimentally colored change of the pH paper (at 350 Lux) in the range of (0–14) by an interval of ~0.1 pH-value. It is worth mentioning that the traditional estimation based on the color change of pH paper is accompanied by a significant variance in pH value (~2). This high variance of pH value led to a noteworthy wrong estimation by eye detection. This finding encourages us to develop a new

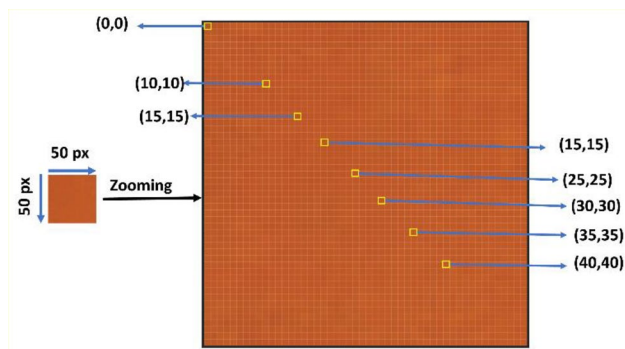


Figure 1. Pixel positions of pH paper image (at pH value = 4, as an example).

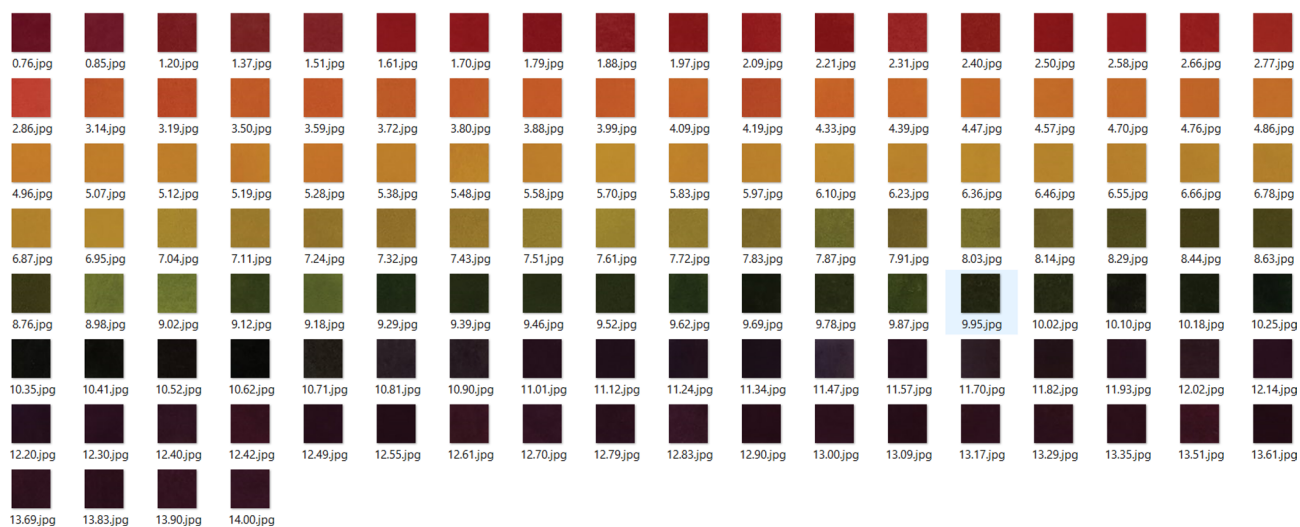


Figure 2. Samples of the captures of an experimentally colored change of the pH-paper at 350 Lux.

simple and more precise method for pH-value detection. Thus, the experiments were extended to cover most of the three different illumination workplaces at 350, 200, and 20 Lux, that the user could work on. Moreover, the homogeneity of the color of the pH paper was emphasized by the collected color RGB code for seven distinct positions per capture. In total, the data set includes 2689 experimental RGB values from different illumination workplaces.

To better understand the observed results in the different workplaces, Exploratory Data Analysis (EDA) of color code RGB against pH values with respect to different light intensities at 20, 200, and 350 Lux, was illustrated in Fig. 3.

The color code points were collected in three parts in a wide pH range. The significant changes in the color code of Red and Green or even Blue were in the range of (2.5: 9) pH values at the three different investigated workplaces of light intensities at (20, 200, and 350 Lux). It is worth mentioning, that the blue color code at low-intensity light of 20 Lux (a little dark workplace) deviates from those obtained in higher or medium light intensity, which suggests avoiding future testing in low light conditions. In contrast, the results revealed no significant difference between the behavior of Red or Green colors at light intensity. The results show the increase in basicity (> 9) or increase in acidity (< 2.5) could interpret the color and may produce less accurate prediction in that part of the pH range. Thus, this finding may encourage the scientific community to prepare higher sensitive material to work in strong acid and/or Strong base medium.

Furthermore, it is critical to recognize and evaluate how dependent each parameter is on the others. This knowledge can aid in the definition of the expectations that these interdependencies provide, leading to the creation of more effective pH devices and color-sensitive materials. Because of this, using a machine learning strategy, the statistical Pearson's correlation coefficients ($r_{x,y}$) between the pH parameters were investigated based on the following Eqs. (5) and (6):

$$\text{cov}_{x,y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N - 1} \quad (5)$$

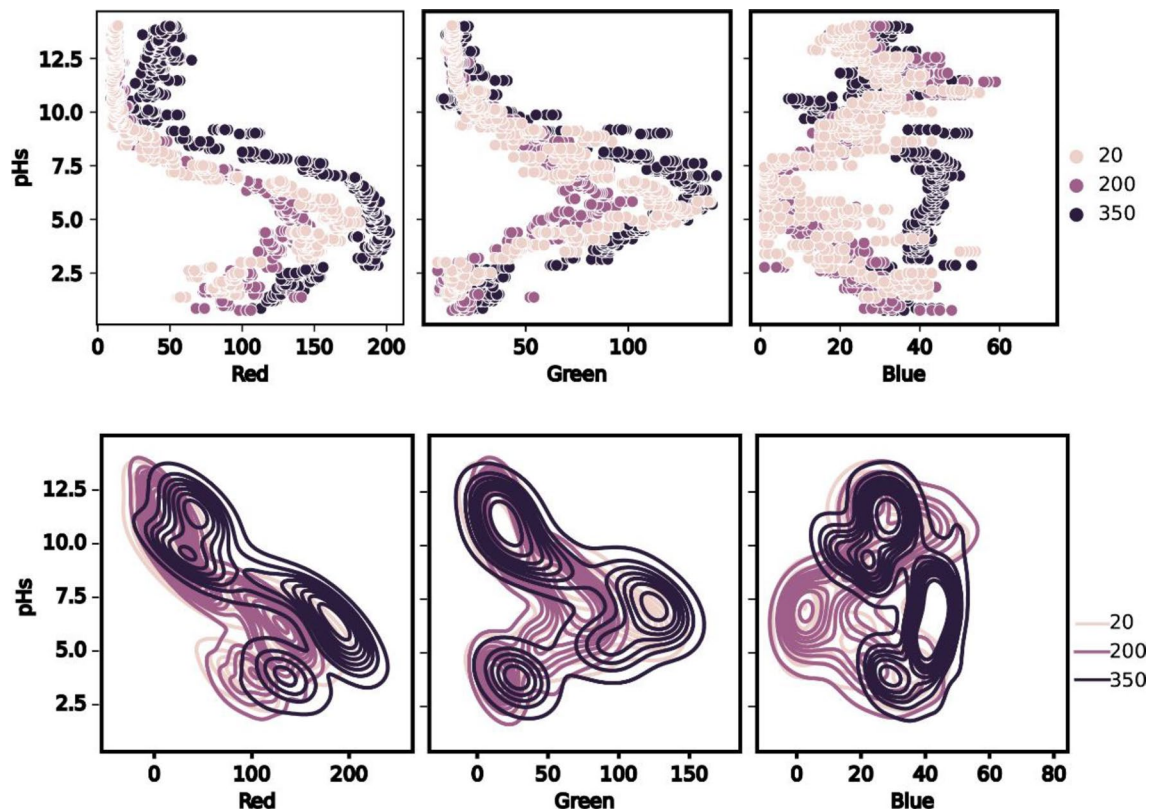


Figure 3. Exploratory Data Analysis of changed RGB code in different illuminated workplaces at (20, 200, and 350 Lux).

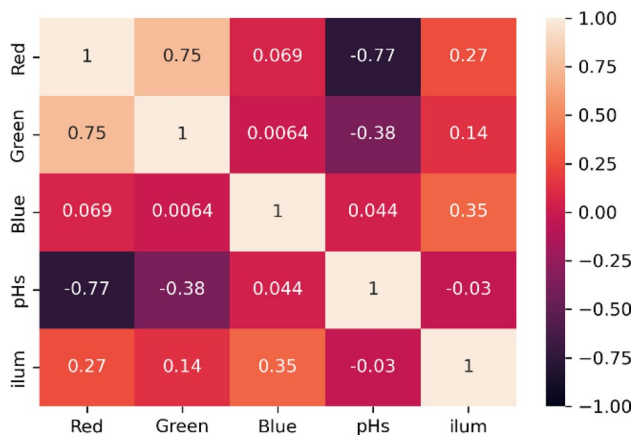


Figure 4. Pearson's correlation coefficients between the pH parameters.

$$r_{xy} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \tag{6}$$

where N number of recorded samples, x_i, y_i are individual elements of RGB and pH predicted values respectively, and \bar{y} the mean value of pH values.

The correlation between the pH parameters was presented with a heatmap in Fig. 4. The obtained results reflect an excellent higher negative correlation between the pH values with Red color (-0.77). In the same way, an acceptable correlation of pH value with the green color by (-0.38). The blue color showed an incredibly low correlation with pH value (0.044) from those observed in the red or green colors. This refers to that the blue color will have a small effect on the machine learning prediction compared to the red and green colors. In the

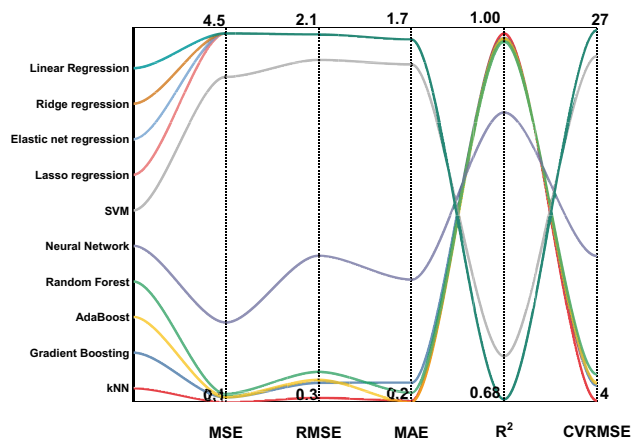


Figure 5. Output results of performed regression algorithms of Linear, Ridge, Lasso, Elastic Net, Polynomial, Support vector machine (SVM) Regressor, Gradient Boosting, AdaBoost, and Random Forest Regressor.

	MSE	RMSE	MAE	R ²	CVRMSE
kNN	0.071	0.266	0.160	0.995	3.385
AdaBoost	0.116	0.341	0.157	0.992	4.332
Gradient Boosting	0.130	0.361	0.242	0.991	4.590
Random Forest	0.164	0.405	0.196	0.988	5.143
Neural Network	0.794	0.891	0.578	0.943	11.321
SVM	3.584	1.893	1.460	0.744	24.058
Ridge regression	4.399	2.097	1.637	0.686	26.651
Linear Regression	4.399	2.097	1.637	0.686	26.651
Elastic net regression	4.399	2.097	1.637	0.686	26.651

Table 1. Regression evaluation metrics of performed algorithms on the experimental data (cross-validation with K-fold = 5).

same way, the illumination of workplaces has no significant effect on the pH value by -0.03 . Thus, the colored pH paper can be safely captured whatever the light intensity.

ML model prediction. Using experimental data, a preliminary analysis of machine learning regression techniques was performed with optimal hyperparameters on K-Nearest Neighbors (KNN), Linear, Lasso, Elastic Net, AdaBoost, Neural Network, Random Forest, and Support vector machine (SVM), and Gradient Boosting Regressor algorithms^{28–30} to estimate coefficients of determination (R^2) and the minimum errors of the corresponding regression evaluation metrics concerning root mean squared error (RMSE), mean absolute error (MAE), and mean squared error (MSE) as shown in Fig. 5 and recorded in Table 1.

It's obvious that the KNN model with optimal hyperparameters of five points performs significant result of R^2 (0.993) combined with the lowest errors of MSE, RMSE, and MAE (0.012, 0.320, and 0.182, respectively) compared to other models. In addition, the coefficient of the variation of the root means square error (CVRMSE) of KNN models shows a higher stability performance of 4.077 compared to other models. Further, the cross-validation with K-fold of (3, 5, 10, and 20) was tested for confirming the stability of the models. However, no significant difference was found between the results, which verified the KNN models.

To deepen understanding, further investigation showed that the results of the model's prediction (based on test data) vs the experimentally obtained pH values are represented in the scatter plot in Fig. 6. The linear regression, elastic net, and Neural network algorithms could not recognize the whole experimental points, especially at the strong acid/base pH range. However, a precise estimate would be placed along a square-diameter line using KNN, Gradient boosting, Random Forest, and AdaBoost algorithms, which could be selected for further steps of deploying the code. Despite the higher performance and exceedingly small deviation of those algorithms, the KNN was chosen for deploying the machine learning mobile application due to having the lowest errors (RMSE; 0.32) and higher stability (CVRMSE; 4.08) as well.

It is now clear that the KNN model can successfully show the underlying patterns of the color RGB code in the pH value estimation based on experimental data collections. Thus, the machine learning approach based on this model was further expanded and used to develop a versatile platform able to predict the pH value using common pH paper with high accuracy. The online mobile application of the prediction model was developed

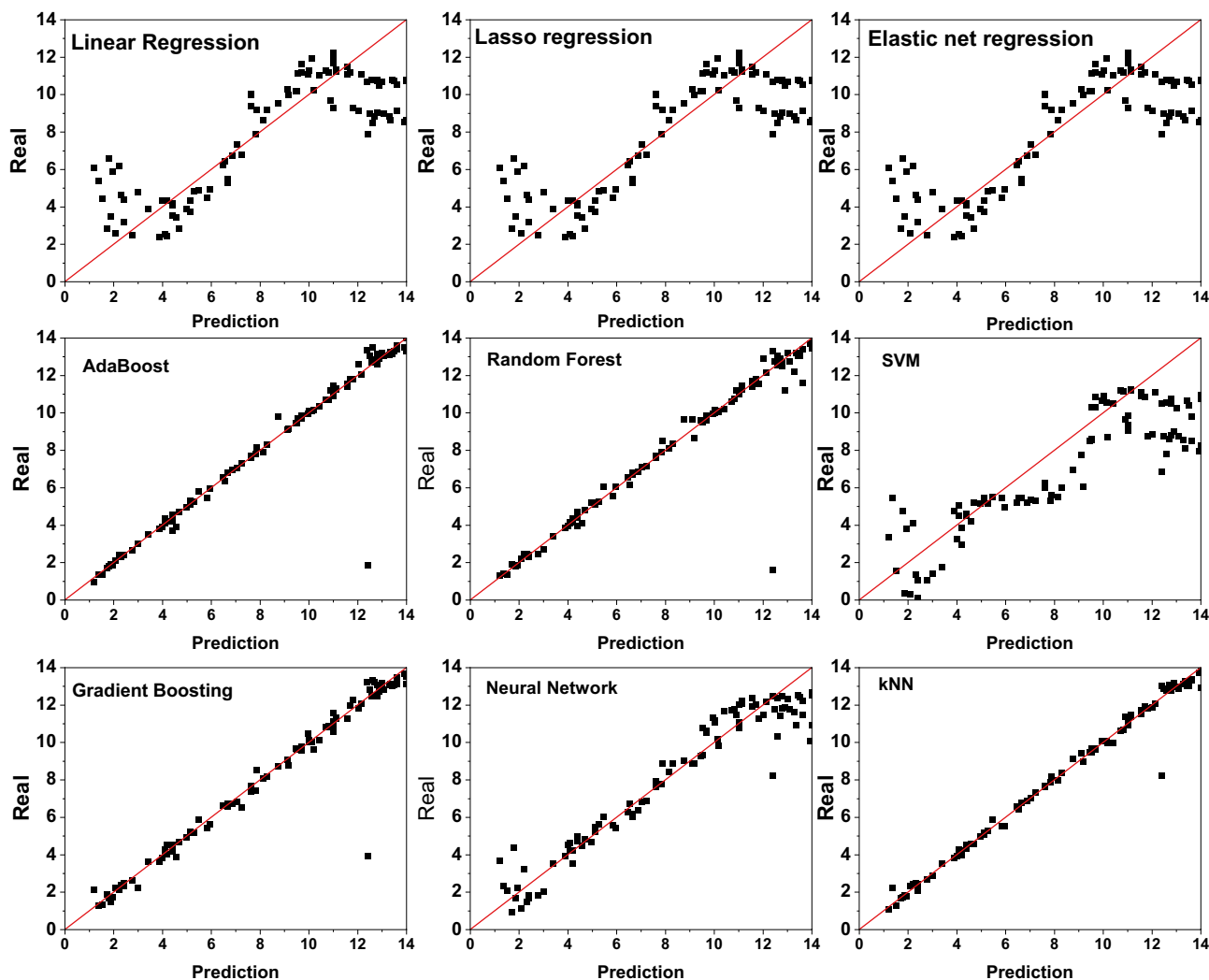


Figure 6. The model's prediction results (based on the test data) vs. experimental results.

using python code and streamlit cloud (freely available) and permits the highly predicted determination of the pH value as a function of the RGB color code of common pH paper.

As illustrated in Fig. 7 the mobile application includes three steps; starting with the input file which could be able to insert the pH paper capture (after being immersed in the target solution immediately). For more facility, we have coded three options (upload a picture, use a mobile camera, or insert a RGB color code). This step is followed by a built-in Machine learning process (without control from the user). Finally, the output of the pH value will appear on the screen.

Our study has a significant advantage over what is already used, Fig. 8 shows the fair comparison of pH instruments, pH paper, and the current study.

Furthermore, Fig. 9 shows the estimated pH value (output results) of the proposed mobile application in comparison with the real pH value. Interestingly, this correlation between real and estimated values in the whole range of pH (acid or base) is related to the higher accuracy of the used ML model.

However, Solmaz et al.³¹ studied pH strips colorimetric detection using ML, as presented in Table 2.

However, four different types of smartphones were used to check the accuracy of pH value predictions for three buffer solutions (pH = 3, 7, and 10). The default setting was used to avoid any smartphone effects. As shown in Fig. 10 and Table 3, the various smartphones do have no significantly different pH value estimations with an accuracy of more than 90% for each type.

Furthermore, Table 4 shows recommended conditions and limitations for using the application to achieve more accurate predictions.

Overall, the present findings solve the problem of pH accuracy using common pH paper without the need for additional costly and time-consuming experimental work. However, our approach solves the problems of excessive cost and maintenance required for traditional pH meters.

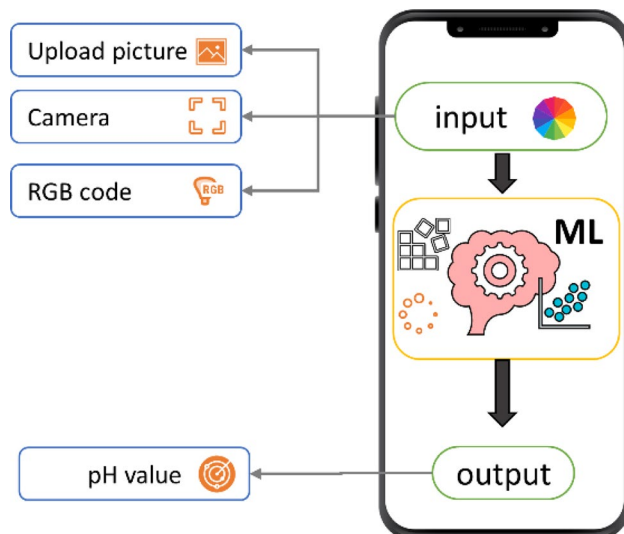


Figure 7. Schematic process of pH detection using ML.

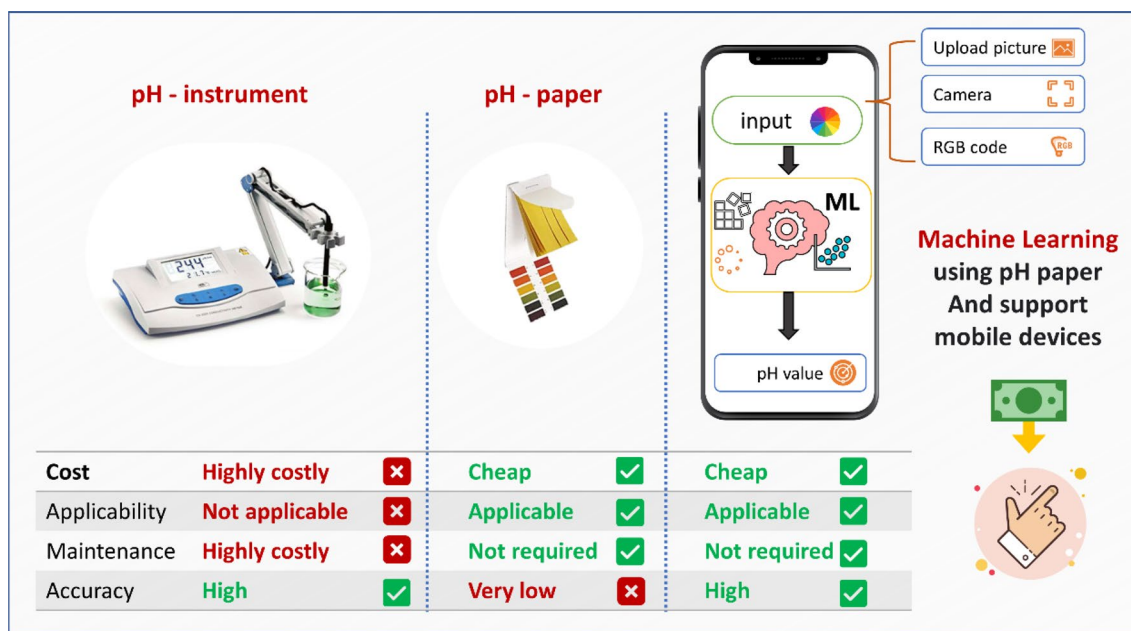


Figure 8. Comparison of pH instrument, pH paper, and the current study.

Conclusion

The findings demonstrate a strong negative association between pH values and both the red color (−0.77) and the green color (−0.38). The blue color will have an insignificant impact on machine learning prediction which revealed a low correlation (0.044). The KNN model exhibits significant R^2 (0.993) results along with the lowest MSE, RMSE, and MAE errors (0.012, 0.320, and 0.182, respectively). This paper also demonstrated the potential of the ML approach to estimate the pH value of solutions using common pH paper. We developed a freely available application that supported mobile devices to predict the pH value based on ML and using common pH paper with precise results. Future research should consider the preparation of new optical material with extremely sensitive color changes in a strong acid/base medium.





















Acidic medium										
Real pH value	1.61	2.21	2.77	3.59	4.19	4.86	5.19	5.97	6.46	6.95
pH paper color										
Estimated pH value	1.72	2.23	2.72	3.54	4.19	4.83	5.21	5.95	6.50	6.95
Basic medium										
Real pH value	7.11	7.83	8.03	8.63	9.29	9.78	10.62	11.34	12.14	13.17
pH paper										
Estimated pH value	7.08	7.89	7.99	8.59	9.27	9.71	10.84	11.36	12.41	13.26

Figure 9. Estimated pH value from mobile application in comparison with real one.

	Solmaz et al. work ³¹	Current work
Type of pH	pH strips	Universal pH paper
Cost	A little higher	Very cheap
Type of ML techniques	Classification	Regression
Input data	450	2688
Illumination source	Special tools required	In ambient condition
Illumination effect	Not studied	Studied at 350, 200, 20 Lux of light intensities
pH range studied	4: 9	Whole range 0: 14
Accuracy	Low (estimated pH value ± 1)	Very high (estimated pH value ± 0.1)

Table 2. The advantages of the current work that related studies.

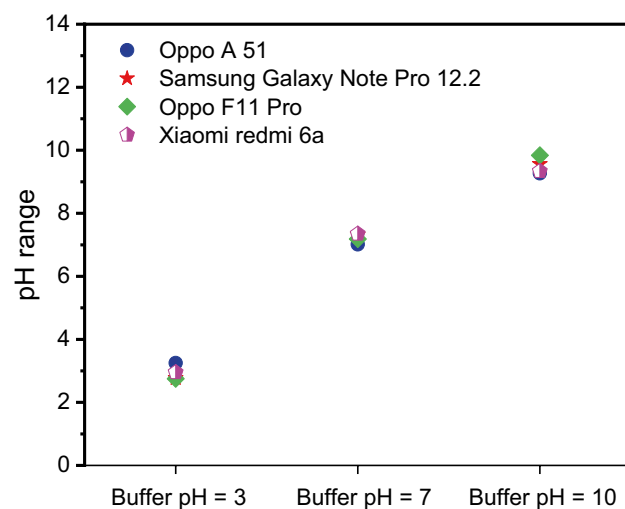


Figure 10. Estimated pH value from different smartphones.

Xiaomi redmi 6a	Oppo F11 Pro	Samsung Galaxy Note Pro 12.2	Oppo A 51	Expected value	
2.951	2.75	2.77	3.25	3	
98%	91%	92%	92%		Accuracy
7.3461	7.1845	7.32	7.013	7	
95%	97%	96%	100%		Accuracy
9.3348	9.84	9.56	9.27	10	
93%	98%	95%	92%		Accuracy

Table 3. Accuracy and estimated pH value from different smartphones. Significant values are in bold.

Light intensity workplace	Normal brightness (200:400 Lux)
pH -paper	Universal indicator pH paper (1–14) Q/3211821AB001-2002
Distance between smartphone and pH paper	15:20 cm
Smartphone camera resolution	More than 5 MP
Smartphone camera operation	Default capture without effects or using different software
Smartphone capture centering	The pH paper (under test) should be at the center position of the smartphone camera

Table 4. Recommended mesurment conditions for application users.

Data availability

The web application and mobile application are freely available “<https://elsenety-ph4-ph-app4-pazg10.strea.mlittapp.com>” or “<https://soft.yallascience.com/2018/06/researcher-tools-software.html>”. The ML code and RGB data are available on request due to privacy/ethical restrictions from the corresponding author [Mohamed M. Elsenety; m.elsenety@azhar.edu.eg].

Received: 10 September 2022; Accepted: 23 December 2022

Published online: 30 December 2022

References

- Wilson, G. S. *et al.* Measurement of pH. Definition, standards, and procedures. *Pure Appl. Chem* **74**, 2169–2200 (2002).
- Jensen, W. B. & Ault, B. The symbol for pH. *J. Chem. Educ.* **81**, 21 (2004).
- Cha, C. S., Li, C. M., Yang, H. X. & Liu, P. F. Powder microelectrodes. *J. Electroanal. Chem.* **368**, 47–54 (1994).
- Zoromba, M. S. Novel and economic acid-base indicator based on (p-toluidine) oligomer: Synthesis; characterization and solvatochromism applications. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **187**, 61–67 (2017).
- Elmorsi, T. M., Aysha, T. S., Machalický, O., Mohamed, M. B. I. & Bedair, A. H. A dual functional colorimetric and fluorescence chemosensor based on benzo[f]fluorescein dye derivatives for copper ions and pH; kinetics and thermodynamic study. *Sens. Actuators B Chem.* **253**, 437–450 (2017).
- Mohamed, M. B. I. *et al.* Colorimetric chemosensor and turn on fluorescence probe for pH monitoring based on xanthene dye derivatives and its bioimaging of living *Escherichia coli* bacteria. *J. Fluoresc.* **30**, 601–612 (2020).
- Aysha, T. S., El-Sedik, M. S., Mohamed, M. B. I., Gaballah, S. T. & Kamel, M. M. Dual functional colorimetric and turn-off fluorescence probe based on pyrrolinone ester hydrazone dye derivative for Cu²⁺ monitoring and pH change. *Dye. Pigment.* **170**, 107549 (2019).
- Aysha, T. S., Mohamed, M. B. I., El-Sedik, M. S. & Youssef, Y. A. Multi-functional colorimetric chemosensor for naked eye recognition of Cu²⁺, Zn²⁺ and Co²⁺ using new hybrid azo-pyrazole/pyrrolinone ester hydrazone dye. *Dye. Pigment.* **196**, 109795 (2021).
- Elsayed, B. A., Ibrahim, I. A., Attia, M. S., Shaaban, S. M. & Elsenety, M. M. Highly sensitive spectrofluorimetric analysis and Molecular Docking using benzocoumarin hydrazone derivative doped in sol-gel matrix as optical sensor. *Sens. Actuators B Chem.* **232**, 642–652 (2016).
- Elsenety, M. M., Elsayed, B. A., Ibrahim, I. A. & Bedair, M. A. Photophysical, DFT and molecular docking studies of Sm(III) and Eu(III) complexes of newly synthesized coumarin ligand. *Inorg. Chem. Commun.* **121**, 108213 (2020).
- Kim, H. *et al.* Fluorescent sensor array for high-precision pH classification with machine learning-supported mobile devices. *Dye. Pigment.* **193**, 109492 (2021).
- Mekonnen, Y., Namuduri, S., Burton, L., Sarwat, A. & Bhansali, S. Review—machine learning techniques in wireless sensor network based precision agriculture. *J. Electrochem. Soc.* **167**, 037522 (2020).
- Qu, X., Latino, D. A. R. S. & Aires-De-sousa, J. A big data approach to the ultra-fast prediction of DFT-calculated bond energies. *J. Cheminform.* **5**, 34 (2013).
- Raghunathan, S. & Priyakumar, U. D. Molecular representations for machine learning applications in chemistry. *Int. J. Quantum Chem.* **122**, e26870 (2022).
- Pflüger, P. M. & Glorius, F. Molecular machine learning: The future of synthetic chemistry?. *Angew. Chemie Int. Ed.* **59**, 18860–18865 (2020).
- Raccuglia, P. *et al.* Machine-learning-assisted materials discovery using failed experiments. *Nature* **533**, 73–76 (2016).
- He, L. *et al.* Applications of computational chemistry, artificial intelligence, and machine learning in aquatic chemistry research. *Chem. Eng. J.* **426**, 131810 (2021).
- Li, L., Rong, S., Wang, R. & Yu, S. Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: A review. *Chem. Eng. J.* **405**, 126673 (2021).
- Chen, H. *et al.* Kernel functions embedded in support vector machine learning models for rapid water pollution assessment via near-infrared spectroscopy. *Sci. Total Environ.* **714**, 136765 (2020).

- 20 Britton, H. T. S. & Robinson, R. A. Universal buffer solutions and the dissociation constant of veronal. *J. Chem. Soc.* <https://doi.org/10.1039/jr9310001456> (1931).
21. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
22. seaborn. Available at: <https://seaborn.pydata.org/generated/seaborn.heatmap.html>.
- 23 Awan, M. J. *et al.* Cricket match analytics using the big data approach. *Electronics* **10**, 2350 (2021).
24. Zakeri-Nasrabadi, M. & Parsa, S. Learning to predict test effectiveness. *Int. J. Intell. Syst.* **37**, 4363–4392 (2022).
25. Yu, Q., Cheng, H. H., Cheng, W. W. & Zhou, X. Ch OpenCV for interactive open architecture computer vision. *Adv. Eng. Softw.* **35**, 527–536 (2004).
26. Guo, G., Wang, H., Bell, D., Bi, Y. & Greer, K. Using kNN model for automatic text categorization. *Soft Comput.* **10**, 423–430 (2005).
27. Pedregosa FABIANPEDREGOSA, F *et al.* Scikit-learn Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
28. Tao, Q., Xu, P., Li, M. & Lu, W. Machine learning for perovskite materials design and discovery. *NPJ Comput. Mater.* **7**, 1–18 (2021).
29. Yilmaz, B. & Yildirim, R. Critical review of machine learning applications in perovskite solar research. *Nano Energy* **80**, 105546 (2021).
30. She, C. *et al.* Machine learning-guided search for high-efficiency perovskite solar cells with doped electron transport layers. *J. Mater. Chem. A* **9**, 25168–25177 (2021).
31. Mutlu, A. Y. *et al.* Smartphone-based colorimetric detection: Via machine learning. *Analyst* **142**, 2434–2441 (2017).

Author contributions

M.M.E.: Conceptualization, Data curation, Investigation, coding, Formal analysis, Methodology, Writing—original draft, Formal analysis, Supervision. MB. M.: Data curation, Formal analysis, Methodology, Investigation, Writing—original draft, coding. M.E.S.: Resources, Investigation, Methodology. B.A.E.: Writing—review & editing, Supervision.

Funding

Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB). Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.M.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022