

Integrating Machine Learning and Color Chemistry: Developing a High-School Curriculum Towards Real-World Problem-Solving

Shiyan Jiang,^{*,1} Jeanne McClure,¹ Hongjing Mao,² Jiahui Chen,² Yunshu Liu,² Yang Zhang,^{*,2}

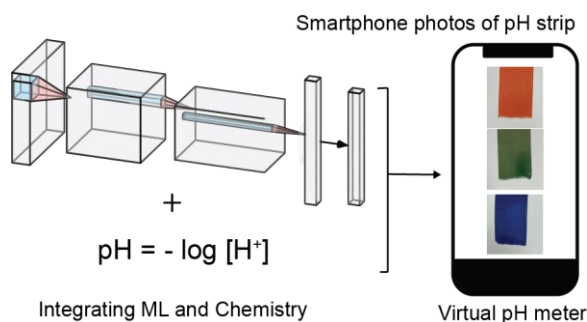
¹Department of Teacher Education and Learning Sciences, North Carolina State University, 2310 Stinson Drive, Raleigh, NC, 27695, USA

²Molecular Analytics and Photonics (MAP) Lab, Program of Polymer and Color Chemistry, Department of Textile Engineering, Chemistry and Science, North Carolina State University, 1020 Main Campus Drive, Raleigh, NC, 27606, USA

ABSTRACT

Artificial intelligence (AI) is rapidly transforming our world, making it imperative to educate the next generation about both the potential benefits and challenges associated with AI. This study presents a cross-disciplinary curriculum that connects AI and chemistry disciplines in the high school classroom. Particularly, we leverage machine learning (ML), an important and simple application of AI to instruct students to build an ML-based virtual pH meter for high-precision pH read-outs. We used a “codeless” and free ML neural network building software – Orange, along with a simple chemical topic of pH to show the connection between AI and chemistry for high-schoolers who might have rudimentary backgrounds in both disciplines. The goal of this curriculum is to promote student interest and drive in the analytical chemistry domain and offer insights into how the interconnection between chemistry and ML can benefit high-school students in science learning. The activity involves students using pH strips to measure the pH of various solutions with local relevancy and then building an ML neural network model to predict the pH value based on color changes of pH strips. The integrated curriculum increased student interest in chemistry and ML and demonstrated the relevance of science to their daily lives and global issues. This approach is transformative in developing a broad spectrum of integration topics between chemistry and ML and understanding their global impacts.

GRAPHICAL ABSTRACT



KEYWORDS

Machine Learning, High School, Analytical Chemistry, pH test, Colorimetric Analysis, Interdisciplinarity

INTRODUCTION

The ways of conducting chemical research and production are reforming with advances in automated instrumentation¹, high-throughput manufacturing², computing power³, and algorithms⁴ which have provided an exponential growth of data and led to the exploration of data-driven approaches for the chemistry field.⁵ A particularly important example is the integration of machine learning (ML) into chemistry which has accelerated the pace of scientific discovery such as developing new drugs and materials⁶, predicting chemical reactions⁷, understanding the mechanisms of chemical reactions⁸, and studying the structure and function of enzymes, proteins, and other biological molecules.⁹ To better integrate ML into the chemistry field, it is critical to expose the next generation to the integration of ML with chemistry to promote their career interests in chemistry and related science, technology, engineering and mathematics (STEM) fields with an AI mindset. To achieve this, students should grasp key knowledge and practices in ML¹⁰ such as data acquisition, preparation, interpretation, and model building with domain knowledge in chemistry. Several publications have reported the implementation of ML and other AI technologies in chemistry classrooms.¹¹⁻¹⁷ The activities include the introductions of (1) basic programming skills in Python¹³ and MATLAB¹⁶ software, (2) building artificial neural networks for a variety of regression-type (chemometric) problems^{13,14,16}, and (3) implementing computer vision for imaging classification^{12,17} and prediction¹¹ of chemical variables. However, most activities are designed for undergraduate and graduate students. To date, there is scarce implementation of integrated ML and chemistry curriculums for lower school grades - from kindergarten (5-6 years old) to the 12th grade at ages of 17-18 (referred to as K-12). To fill this gap, we designed an integrated ML and chemistry curriculum with contents and activities suitable for high-school students who may have limited backgrounds in ML and chemistry.

Integrating advanced technology into K-12 classrooms can enhance students' motivation and engagement in science, leading to improved performance and interest in related careers.¹⁸ Particularly, incorporating ML in K-12 classrooms exposes students to AI and ML's power and limitations. However, the way science topics and their connection with ML presented in existing K-12 classrooms is not well-established.¹⁹ Science and ML are sometimes taught as parallel entities rather than integrated subjects, which diminishes students' engagement. To fill this knowledge gap, we propose to develop an integrated ML and chemistry activity for high-school students with implications for real-world challenges to allow the students to better perceive the interplay of ML and chemistry. Consequently, students can learn about AI's role in chemical fields while also developing computational and data literacy skills.

INTEGRATED LEARNING OF ML AND CHEMISTRY

Since chemistry and ML concepts are often abstract, we aim to develop an integrated ML and chemistry curriculum that is (1) within the scope of the high-school chemistry curriculum, (2) engaging by connecting to real-world problems and (3) facilitating the understanding of both ML and

chemistry concepts and their interplays. We propose to leverage the concept of color chemistry²¹ or the observation and understanding of different colors in chemical phenomena as engaging tools for digesting chemistry concepts and linking them to ML. Colored observations as a visual aid²², facilitate teaching and learning chemistry, especially in improving qualitative analysis and critical thinking skills. Further, colored chemical changes can be captured as images, which promotes the understanding of key ML concepts in the visualization and interpretation of imaging data, model comparison, error analysis, and interactive learning. In short, using color changes to represent chemical reactions, students may develop a deeper understanding of the underlying principles of ML and improve their ability to analyze data in the chemistry field.

ML methods are well-known for their unmatched capability in imaging recognition, including the high-accuracy prediction of pH strip readings.²³ Capturing pH strip photos with each color corresponding to a specific pH value not only provides a practical ML dataset but also allows the students to be fully exposed to the entire ML model construction process. Specifically, this practice involves all essential procedures for ML network development and assessment including (1) data acquisition through taking pH strip photos, (2) data preparation by cropping the images, (3) model generation through the construction of an ML network to predict pH values, and (4) model evaluation and optimization by comparing predicted results to pH meter readings. Moreover, this topic has local relevance as students can test samples such as ponds in their community, in-house aquatic tanks, soda drinks, and soils in their backyard. This curriculum also helps students understand the global impact of pH on soil and water quality, climate change, environmental health, plant growth, and human well-being. Thus, students can make a positive impact on building a sustainable future for themselves using the acquired ML and chemistry knowledge.

In terms of educational technology, we chose Orange, an open-source ML and data mining toolkit.²⁴ It provides a visual programming interface and offers a variety of ML algorithms that are suitable for beginners.²⁵ Orange has a large and active community, with many resources and tutorials available, making it an accessible tool for those who may not have extensive programming experience. Additionally, Orange's interactive visualizations can help students understand the relationships between data variables and ML algorithms, promoting teaching data literacy and scientific inquiry and serving as a stepping stone to more advanced ML tools. We have provided a short tutorial (see Step-by-step Orange slides and video instructions in the **Supporting Information**) for creating ML neural networks for the pH prediction tasks.

ACTIVITY

We draw on three theoretical lenses to design the integrated learning activity: systems thinking²⁶, learning by doing²⁷, and engaging by relating.²⁸ By engaging students in a real-world problem-solving activity, they can actively participate in the learning process, fostering a deeper understanding of the concepts and principles introduced in the curriculum. Specifically, the

curriculum was implemented in a high school general chemistry class in the Eastern United States consisting of sixteen students. Students engaged in five lessons in five days (90 minutes per day) (**Figure 1**). The first lesson took place in the traditional Chemistry classroom to cover basic concepts of pH and its measurement using overhead slides and lectures according to Next Generation Science Standards (NGSS) guidelines.²⁹ At the end of the first lesson, the students were asked to collect real-life aqueous samples that were relevant to their daily lives to be measured in lesson two (e.g., drinks, pond/lake water, cosmetics, kitchen ingredients).

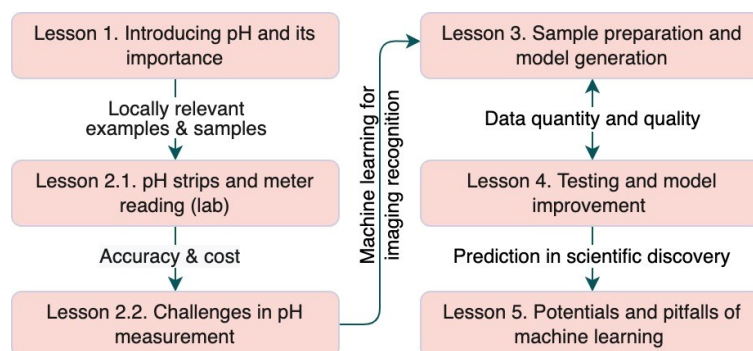


Figure 1. A flowchart highlights the main topics of the 5-lesson curriculum, integrating ML into high-accuracy pH strip prediction.

The second lesson was held in a high-school chemistry laboratory for hands-on activities to engage students in pH measurement and to prepare ML datasets. Particularly, the students were grouped in 3-4 and instructed to measure the samples they collected using pH strips with a standard pH-color conversion chart as the reference. Solutions with known pH values prepared by the instructor before class were also distributed to the students. Students were then asked to measure the pH value of samples of their own and took a photo of the wet-colored pH strip immediately for later use as the ML testing dataset. pH strip photos of a series of standard solutions with known pH were captured by the instructors in advance as the training dataset and shown in **Figure 2A** and are available in the **Supporting Information**. Next, the instructor introduced the use of pH meters as high-cost and high-accuracy equipment and instructed the students to repeat the pH measurement of their samples using a pH meter (**Figure 2A** bottom).

Further, the students measured their own samples and compared their readings with their group members to discuss the discrepancies among their visual pH-strip readings in locally relevant samples (**Figure 2B**) and the consistencies of values read from the pH meter. Afterward, the instructor discussed the tradeoff between cost, accuracy, and accessibility, the water quality issue in underdeveloped areas, and the importance of accessible techniques in specific scenarios like preventing food crises and securing human health. Finally, the instructor presented the students with a task to solve this particular problem – developing an ML-based imaging recognition method to create a virtual pH meter for underdeveloped areas. Students were instructed to crop the pH-strip photos while keeping only color-changed areas and a white-paper background (about 20-30% of the sizes of the pH strip) as a simple practice of data cleaning in ML (photos in **Figure 2B**).

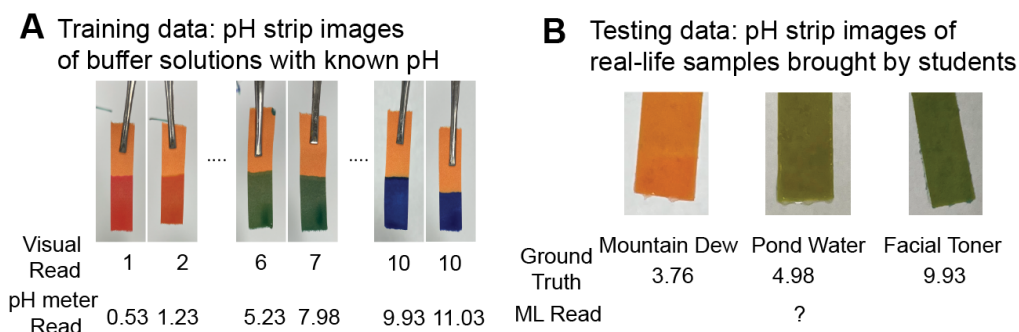


Figure 2. (A) Top: photos of pH strip measured from different buffer solutions as ML training data; bottom: pH strip and meter readings of the buffer solution samples; (B) pH-strip photos and pH meter readings of real-life samples brought by students as the testing dataset for ML.

The following three lessons were conducted in the media lab of the school library to access Orange software for ML practices. Students build upon their prior knowledge using pH training data to develop a model by relating and comparing the accuracy to underscore the importance of data quantity and quality in real-world applications like diabetic testing, lake pollution, and soil nutrient tests, to shape a sustainable future.

The third lesson introduced students to the ML workflow using a short lecture accompanied by a hands-on activity to investigate supervised ML algorithms. Students were regrouped in sizes of 2-3 and worked together to explore the image processing technique of feature extraction. Feature extraction is the conversion of data from the original format (e.g., an image) into a series of quantitative (e.g., numeric) or qualitative (e.g., text description) features that distinguish different objects and are essential for pH identification in a virtual pH meter. To learn about supervised learning and the ML workflow, students used animal sort cards as their data in the unplugged activity. Students picked five animals, used a table template to create labels for the animals as the variable name, and added quantitative and qualitative information. For example, one group may have labeled the animals with ears, tails, legs, and colors. For a black cat, the labels would include four legs, two ears, one tail, and black for the color. Then students would choose one animal and predict which group it would fit into based on features in the table. This unplugged activity was designed to help students understand the concept of feature extraction and how features could be used to classify different objects.

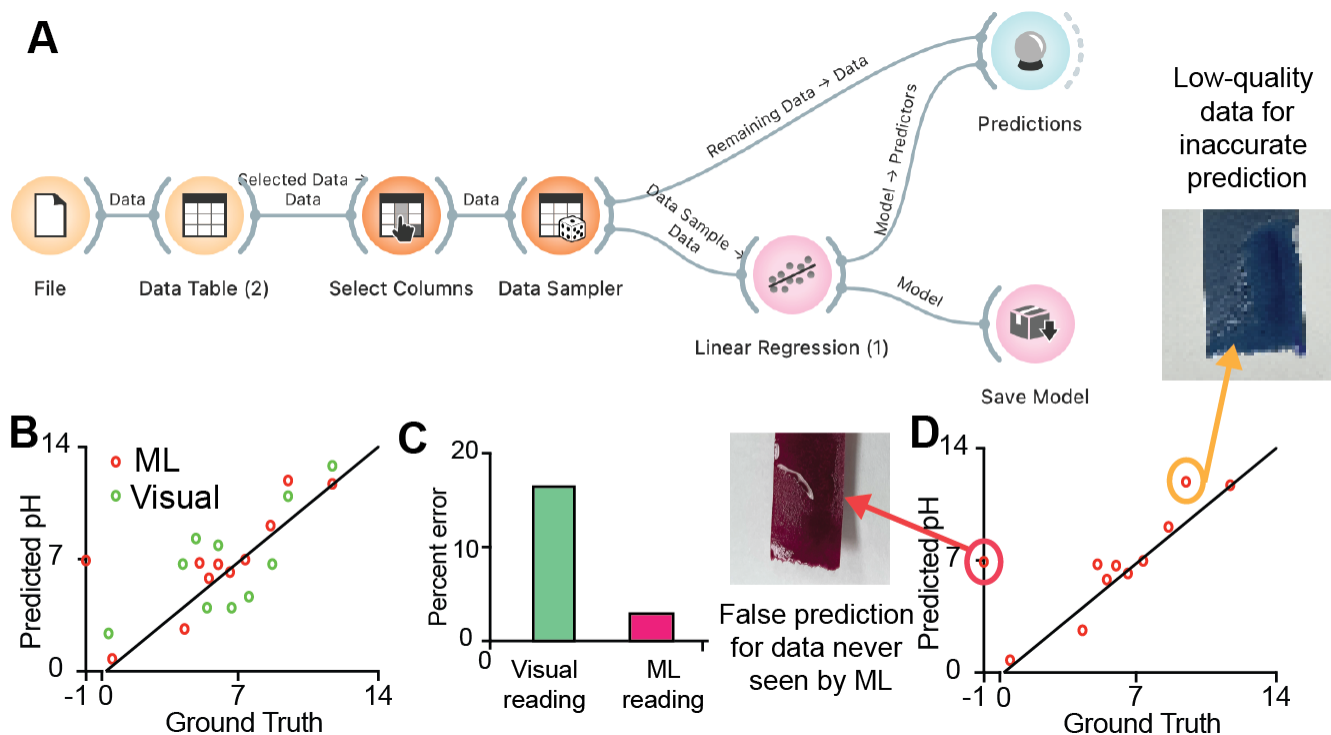


Figure 3. (A) A representative interface of Orange software for generating the ML network for pH strip reading; (B) a scatterplot illustrating pH readout using the student-developed ML network (red dots) or students' visual readings (green dots) versus the corresponding pH meter measurement as the ground truth to demonstrate the ML-enhanced pH strip reading precision as the key component; the dots are closer to the black reference line indicates relatively small errors in pH reading compared with the ground-truth measurement. (C) A comparison of the averaged percent errors in pH strip reading by human visual and by ML. (D) A scatterplot shows ML reading compared with the ground truth pH-meter value of poorly predicted samples; the left image shows the limitation of ML in detecting colors never seen in the training data; the right image shows the influence of poor data quality on ML.

In the fourth lesson, students were introduced to the ML workflow. A representative interface of the ML network generated using Orange is shown in **Figure 3A**. First, the instructor modeled the process of developing clustering models. Then students applied their knowledge by completing a selection of challenges in a scavenger hunt using a traffic light dataset. Challenges included using the tool widgets in Orange to load the dataset, inspecting the dataset to identify the number of instances and attributes, using the image viewer to form a hypothesis, using a regression neural network to create image embedding visualizations (see more details of testdata and modeldata in the **Supporting Information**), and finally communicating results. To stimulate students' critical thinking, the instructor prompted them to engage in discussions regarding how the algorithm clustered the traffic signs and what features were used by the algorithm to make those decisions. This activity was designed for students to get familiar with the Orange platform and its various functionalities for ML workflows. We chose the traffic light dataset to introduce ML workflow as it's an existing dataset in the platform so that students could focus on learning ML workflow instead of hitting the hurdle of creating and loading datasets. Also, high school students in the United States are typically introduced to traffic

rules and regulations as part of their driver's education, making the topic both relevant to their everyday experiences.

The fifth lesson focused on supervised learning with vectorized image embedding (see Student Handout of Lesson 5 in the **Supporting Information**). Students participated in a code-along building of the training model. The code-along guided students through the process of uploading pH training data, creating image embeddings, and observing the data dimensions and features. Students also followed the workflow identifying the association of the image and pH values as inputs (image descriptions) and outputs or targets (pH value) to create a linear regression for prediction. After developing the training model, students validated their model performance by testing the original pH strip pictures they took before by comparing the ground truth (pH meter) vs. predicted value (virtual pH). In this activity, we also emphasized the important concept of ground truth, which refers to the accurate and objective data that serves as a reference for training and evaluating ML models³⁰, provided by direct measurement. We showed the pH prediction results of the ML network the students developed compared with their visual readings (**Figure 3B**). Furthermore, we guided the students to compare both ML and visual readings against the ground truth measurements to validate that the ML methods would have a more accurate prediction than the visual readout. We further instructed the students to measure the percent errors of ML and visual reading using the formula $[\text{ML or visual value} - \text{pH meter value}] / \text{pH meter value}$. The averaged percent errors (**Figure 3C**) of visual readings (18%) are 5.5 times higher than that of the ML reading (3.3%), which further demonstrates the power of ML in these prediction tasks.

To underscore the importance of data quantity and quality in model prediction performance, students identified ML's limitations due to outliers (left photo in **Figure 3D**) and misclassification of data resulting from noise and human error (right photo in **Figure 3D**). To expand the discussions, students were instructed to think about the tradeoff between accuracy and cost when detecting important chemical substances/pollutants, such as heavy metals. The students discussed the potential of colorimetric analyses in chemistry to address real-world problems, such as COVID-19 detection which requires quick, economical, and accurate testing. These discussions were designed to help the students understand how the concepts they were learning could be applied to a sustainable future.

RESULTS AND DISCUSSION

The student-developed ML methods provide an average percent error of 3.3% in reading the pH strip values compared with the pH meter value, which is six times lower than those measured by the visual readings (18%) (**Figure 3**). The expenses of generating this virtual pH meter are only about the use of a smartphone camera to take photos while more data can be collected to further improve the ML reading accuracy. We used the pH strips from the same vendor throughout the study to avoid different pH-color responses. We intentionally left a white background using a clean white letter paper underneath the wetted pH strip as a normalization factor for ML to compensate for minute differences in white balance, camera mode, and lighting conditions in the imaging acquisition process. Overall, the high precision of an ML-based pH meter indicates its potential for practical uses in the school

environment with reduced cost, in field studies where sophisticated pH meters are more prone to be broken. The commercial standards for calibrating the pH meters have limited shelf life and are relatively expensive if intensive use is required.

220 We collected pre-and post-survey responses, classroom video recordings, activity-specific assignments, and screen recordings, and conducted student interviews (see Student Learning Outcome Examples and Student Interview Protocol in the **Supporting Information**) after the project implementation. We conducted three evaluation constructs (i.e., career motivation, intrinsic motivation, and self-determination) from the Science Motivation Questionnaire II.³¹ Nine out of the
225 sixteen students completed the pre- and post-survey. All three constructs are statistically different in pre-versus post-intervention showing that the intervention fosters students' science motivation. **Figure 4A-C** respectively showed changes in students' science motivation from three representative items from each construct: "I am curious about discoveries in science" (intrinsic motivation), "Knowing science will give me a career advantage in the future" (career motivation), and "I spent a lot of time
230 learning science" (self-determination). These results further highlight that the integrated ML and chemistry curriculum motivates the students in relevant careers.

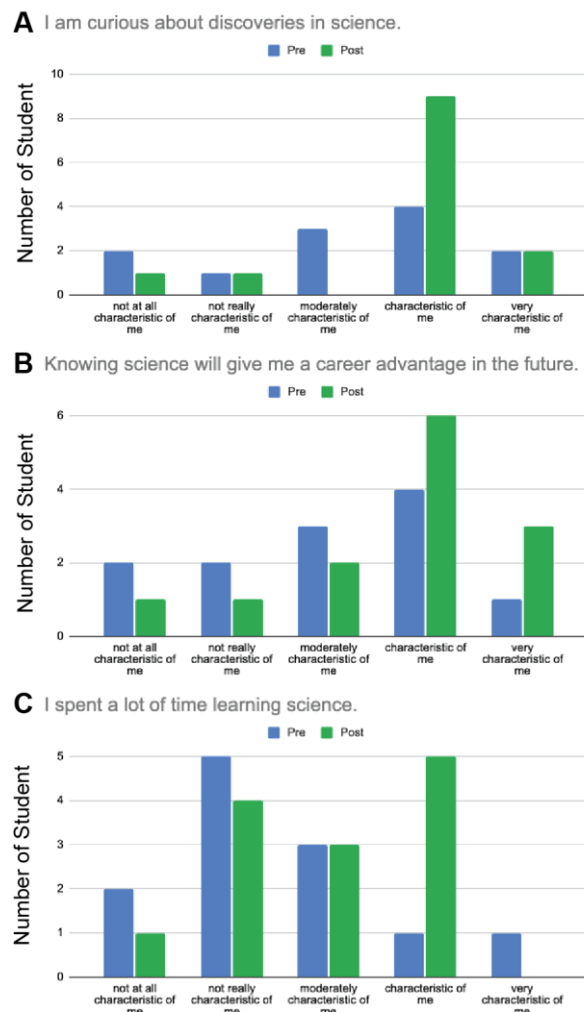


Figure 4. Bar charts evaluating students' science motivation from three representative items of (A) intrinsic motivation, (B) career motivation and (C) self-determination (blue bars) and after (green bars) the implemented integrated ML-Chemistry curriculum.

EXPERIMENTAL SETUP

The buffer solutions with known pH were prepared in a phosphate buffered saline (PBS) solution, pH 7.4 with additions of different amounts of 5 M HCl and 5 M KOH diluted in PBS. The ground truth pH values of the solutions were measured using a pH meter (Thermo Scientific Orion Star A120) one day before the lecture. The pH meter was calibrated using Thermo Scientific Orion All-in-One pH Buffer kits containing standard buffer solutions of pH 10.01, 7.00, and 4.01, respectively. The instructors took about 100 pH strip photos of solutions with pH values ranging 1-13. A pH strip photo of 10 N sulfuric acid was also taken to demonstrate the outlier in ML. All the remaining pH strip measurements of buffer solutions and real-life samples were done in the lab class.

HAZARDS

Hazards in this activity arise from the use of hydrochloric acid (5 M), sulfuric acid (10 N), and potassium hydroxide (5 M) in PBS during the standard solution preparation process by the instructor wearing appropriate laboratory protective equipment.

CONCLUSION

We developed an integrated ML and chemistry curriculum for high-school students by introducing colorimetric pH strip measurement to build an ML neural network to predict pH value. The student was instructed on the basic pH concept, the connection of pH to real-world examples and issues, and the implementation of an ML model to predict pH strip values. The student-developed ML-network predicts pH 5.5 times more precisely than human visual reading. Our integrated ML and chemistry curriculum makes chemistry learning more accessible, engaging, and interconnected with real-world applications and also fosters science motivation, encourages critical thinking about ML applications, and expands students' knowledge and understanding of AI and ML concepts. We envision that this integrated approach is transformative to developing a broad spectrum of integration topics between color chemistry in food, textile and apparel, medical, energy, and cosmetic fields and how ML can enhance signal readout precision in these fields along with understanding their global impacts.

ASSOCIATED CONTENTS

Supporting Information

The Supporting Information is available at <https://pubs.acs.org/doi/10.1021/acs.jchemed.xxx>. Detailed student handouts, additional student learning outcomes, student interview protocols (DOCX), lecture notes (PDF), tutorials for the use of Orange Software (PDF, MP4) and photos of pH strips (ZIP).

AUTHOR INFORMATION

Corresponding Author

*E-mail: sjiang24@ncsu.edu (Shiyan Jiang); yang.zhangfl@gmail.com (Yang Zhang)

ACKNOWLEDGMENTS

We acknowledge the generous support from the National Science Foundation (CHE-2246548, DRL-1949110, and DRL-2025090) and the National Institutes of Health (R21GM141675 and R01GM143397)

REFERENCES

- (1) Shi, Y.; Prieto, P. L.; Zepel, T.; Grunert, S.; Hein, J. E. Automated Experimentation Powers Data Science in Chemistry. *Acc. Chem. Res.* **2021**, 54 (3), 546–555. <https://doi.org/10.1021/acs.accounts.0c00736>.
- (2) Selekman, J. A.; Qiu, J.; Tran, K.; Stevens, J.; Rosso, V.; Simmons, E.; Xiao, Y.; Janey, J. High-Throughput Automation in Chemical Process Development. *Annu. Rev. Chem. Biomol. Eng.* **2017**, 8 (1), 525–547. <https://doi.org/10.1146/annurev-chembioeng-060816-101411>.
- (3) *Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems | Chemical Reviews*. <https://pubs.acs.org/doi/10.1021/acs.chemrev.1c00107> (accessed 2023-04-10).

- (4) Wang, J.; Wang, X.; Sun, H.; Wang, M.; Zeng, Y.; Jiang, D.; Wu, Z.; Liu, Z.; Liao, B.; Yao, X.; Hsieh, C.-Y.; Cao, D.; Chen, X.; Hou, T. ChemistGA: A Chemical Synthesizable Accessible Molecular Generation Algorithm for Real-World Drug Discovery. *J. Med. Chem.* **2022**, 65 (18), 12482–12496. <https://doi.org/10.1021/acs.jmedchem.2c01179>.
- (5) Stringer, B. *Big data and the chemical industry*. ICIS Explore. <https://www.icis.com/explore/resources/news/2013/12/13/9735874/big-data-and-the-chemical-industry> (accessed 2023-04-10).
- (6) Ashdown, G. W.; Dimon, M.; Fan, M.; Sánchez-Román Terán, F.; Witmer, K.; Gaboriau, D. C. A.; Armstrong, Z.; Ando, D. M.; Baum, J. A Machine Learning Approach to Define Antimalarial Drug Action from Heterogeneous Cell-Based Screens. *Sci. Adv.* **2020**, 6 (39), eaba9338. <https://doi.org/10.1126/sciadv.aba9338>.
- (7) Kayala, M. A.; Azencott, C.-A.; Chen, J. H.; Baldi, P. Learning to Predict Chemical Reactions. *J. Chem. Inf. Model.* **2011**, 51 (9), 2209–2222. <https://doi.org/10.1021/ci200207y>.
- (8) Peiretti, F.; Brunel, J. M. Artificial Intelligence: The Future for Organic Chemistry? *ACS Omega* **2018**, 3 (10), 13263–13266. <https://doi.org/10.1021/acsomega.8b01773>.
- (9) Townshend, R. J. L.; Eismann, S.; Watkins, A. M.; Rangan, R.; Karelina, M.; Das, R.; Dror, R. O. Geometric Deep Learning of RNA Structure. *Science* **2021**, 373 (6558), 1047–1051. <https://doi.org/10.1126/science.abe5650>.
- (10) *Rules of Machine Learning*. Google Developers. <https://developers.google.com/machine-learning/guides/rules-of-ml> (accessed 2023-04-10).
- (11) Kosenkov, Y.; Kosenkov, D. Computer Vision in Chemistry: Automatic Titration. *J. Chem. Educ.* **2021**, 98 (12), 4067–4073. <https://doi.org/10.1021/acs.jchemed.1c00810>.
- (12) Zhang, B.; Frkonja-Kuczin, A.; Duan, Z.-H.; Boika, A. Workshop on Computer Vision for Bioanalytical Chemists: Classification and Detection of Amoebae Using Optical Microscopy Image Analysis with Machine Learning. *J. Chem. Educ.* **2023**, 100 (2), 539–545. <https://doi.org/10.1021/acs.jchemed.2c00631>.
- (13) Lafuente, D.; Cohen, B.; Fiorini, G.; García, A. A.; Bringas, M.; Morzan, E.; Onna, D. A Gentle Introduction to Machine Learning for Chemists: An Undergraduate Workshop Using Python Notebooks for Visualization, Data Processing, Analysis, and Modeling. *J. Chem. Educ.* **2021**, 98 (9), 2892–2898. <https://doi.org/10.1021/acs.jchemed.1c00142>.
- (14) Revignas, D.; Amendola, V. Artificial Neural Networks Applied to Colorimetric Nanosensors: An Undergraduate Experience Tailorable from Gold Nanoparticles Synthesis to Optical Spectroscopy and Machine Learning. *J. Chem. Educ.* **2022**, 99 (5), 2112–2120. <https://doi.org/10.1021/acs.jchemed.1c01288>.
- (15) St James, A. G.; Hand, L.; Mills, T.; Song, L.; Brunt, A. S. J.; Bergstrom Mann, P. E.; Worrall, A. F.; Stewart, M. I.; Vallance, C. Exploring Machine Learning in Chemistry through the Classification of Spectra: An Undergraduate Project. *J. Chem. Educ.* **2023**, 100 (3), 1343–1350. <https://doi.org/10.1021/acs.jchemed.2c00682>.
- (16) Joss, L.; Müller, E. A. Machine Learning for Fluid Property Correlations: Classroom Examples with MATLAB. *J. Chem. Educ.* **2019**, 96 (4), 697–703. <https://doi.org/10.1021/acs.jchemed.8b00692>.
- (17) Sharma, A. K. Laboratory Glassware Identification: Supervised Machine Learning Example for Science Students. *J. Comput. Sci. Educ.* **2021**, 12 (1), 8–15. <https://doi.org/10.22369/issn.2153-4136/12/1/2>.
- (18) Vaino, T.; Vaino, K.; Rannikmäe, M. Enhancing Students' Interests in Science and Technology Related Careers Through a Specially Designed Optional Course. *Procedia - Soc. Behav. Sci.* **2015**, 177, 331–335. <https://doi.org/10.1016/j.sbspro.2015.02.352>.
- (19) Shamir, G.; Levin, I. Teaching Machine Learning in Elementary School. *Int. J. Child-Comput. Interact.* **2022**, 31, 100415. <https://doi.org/10.1016/j.ijcci.2021.100415>.
- (20) Sanfeliz, M.; Stalzer, M. Science Motivation in the Multicultural Classroom. *Sci. Teach.* **2003**, 70 (3), 64–66.
- (21) Christie, Robert. *Colour Chemistry*; Royal Society of Chemistry, 2001.

- 335 (22) Kaspar, F.; Crameri, F. Coloring Chemistry—How Mindful Color Choices Improve Chemical
Communication. *Angew. Chem. Int. Ed.* **2022**, *61* (16), e202114910.
<https://doi.org/10.1002/anie.202114910>.
- (23) Elsenety, M. M.; Mohamed, M. B. I.; Sultan, M. E.; Elsayed, B. A. Facile and Highly Precise pH-
340 Value Estimation Using Common pH Paper Based on Machine Learning Techniques and
Supported Mobile Devices. *Sci. Rep.* **2022**, *12* (1), 22584. <https://doi.org/10.1038/s41598-022-27054-5>.
- (24) Demšar, J.; Curk, T.; Erjavec, A.; Gorup, Č.; Hočevár, T.; Milutinovič, M.; Možina, M.; Polajnar,
M.; Toplak, M.; Starič, A.; Štajdohar, M.; Umek, L.; Žagar, L.; Žbontar, J.; Žitnik, M.; Zupan, B.
Orange: Data Mining Toolbox in Python. *J. Mach. Learn. Res.* **2013**, *14* (1), 2349–2353.
- 345 (25) Godec, P.; Pančur, M.; Ilenič, N.; Čopar, A.; Stražar, M.; Erjavec, A.; Pretnar, A.; Demšar, J.;
Starič, A.; Toplak, M.; Žagar, L.; Hartman, J.; Wang, H.; Bellazzi, R.; Petrovič, U.; Garagna, S.;
Zuccotti, M.; Park, D.; Shaulsky, G.; Zupan, B. Democratized Image Analytics by Visual
Programming through Integration of Deep Models and Small-Scale Machine Learning. *Nat.*
Commun. **2019**, *10* (1), 4551. <https://doi.org/10.1038/s41467-019-12397-x>.
- 350 (26) Arnold, R. D.; Wade, J. P. A Definition of Systems Thinking: A Systems Approach. *Procedia*
Comput. Sci. **2015**, *44*, 669–678. <https://doi.org/10.1016/j.procs.2015.03.050>.
- (27) *The learning-by-doing principle*. <https://psycnet.apa.org/fulltext/2014-55719-001.html> (accessed
2023-04-25).
- (28) Kearsley, G.; Shneiderman, B. Engagement Theory: A Framework for Technology-Based Teaching
355 and Learning. *Educ. Technol.* **1998**, *38* (5), 20–23.
- (29) *Next Generation Science Standards: For States, By States*; National Academies Press: Washington,
D.C., 2013. <https://doi.org/10.17226/18290>.
- (30) Kotsiantis, S. B. Supervised Machine Learning: A Review of Classification Techniques. In
Proceedings of the 2007 conference on Emerging Artificial Intelligence Applications in Computer
360 *Engineering: Real World AI Systems with Applications in eHealth, HCI, Information Retrieval and*
Pervasive Technologies; IOS Press: NLD, 2007; pp 3–24.
- (31) Glynn, S. M.; Brickman, P.; Armstrong, N.; Taasoobshirazi, G. Science Motivation Questionnaire
II: Validation with Science Majors and Nonscience Majors. *J. Res. Sci. Teach.* **2011**, *48* (10),
1159–1176. <https://doi.org/10.1002/tea.20442>.
- 365