"Now, I Want to Teach It for Real!": Introducing Machine Learning as a Scientific Discovery Tool for K-12 Teachers

Xiaofei Zhou(✉), Jingwan Tang, Michael Daley, Saad Ahmad, and Zhen Bai

University of Rochester, Rochester, NY 14627, USA
{xzhou50,jtang21,sahmad11,zbai7}@ur.rochester.edu,
mdaley@warner.rochester.edu

Abstract. Machine Learning (ML) is a powerful tool to unveil hidden patterns in data, unearth new insights and promote scientific discovery (SD). However, expertise is usually required to actualize the potential of ML fully. Very little has been done to begin instructing the youth of society in ML, nor utilize ML as an SD tool for the K-12 age range. This research proposes SmileyDiscovery, an ML-empowered learning environment that facilitates SD for K-12 students and teachers. We conducted a 2-session preliminary study with 18 K-12 STEM teachers. Findings confirm the effectiveness of SmileyDiscovery in supporting teachers to (1) carry out ML-empowered SD, (2) design their own curriculum-aligned SD lesson plans, and (3) simultaneously obtain a rapid understanding of k-means clustering. Design implications distilled from our study can be applied to foster more effective learning support in future systems.

Keywords: K-12 Education · Machine learning · Scientific discovery learning · Technology-enhanced learning

1 Introduction

Scientific discovery (SD) learning plays a critical role in K-12 STEM education by mimicking how scientists study the world through data collection, experimental operations, and pattern interpretation [9,10]. SD naturally connects with Machine Learning (ML) which accelerates data analysis by systematically searching hypotheses and revealing complex patterns in big data [16]. With ML becoming increasingly fundamental in generating new findings in astronomy, biology, chemistry [21], and other STEM domains, it is essential to provide opportunities for K-12 students and teachers to apply ML as a new discovery tool.

Imagine a high-school biology teacher encouraging students to discover new knowledge about dynamic ecosystems. The teacher first introduces a dataset containing over 10 ecological attributes collected from hundreds of ecological field sites. By exploring a few field sites, students may raise questions/hypotheses on

I. Roll et al. (Eds.): AIED 2021, LNAI 12748, pp. 486–499, 2021.
https://doi.org/10.1007/978-3-030-78292-4_39

interactions between ecological attributes. Then students can begin probing these initial ideas using pattern recognition with the help of ML. This may, in turn, lead to a cycle of further inquiries with new hypotheses. Such processes would largely promote science practices required by national standards. These practices include asking questions, planning and carrying out investigations, analyzing and interpreting data, engaging in evidence-based argumentation, and so forth [46].

Despite those promising benefits, little effort has been made to understand ML as a data-driven discovery tool for K-12 science learning. One challenge is balancing the support in learning ML and applying ML for novice learners [2,52]. The other is the lack of curriculum-aligned learning activities for K-12 teachers to engage students in ML-empowered SD [29,48].

To address these challenges, we developed a learning environment, SmileyDiscovery, to support low-barrier ML-empowered SD without extra ML training for K-12 teachers and students. SmileyDiscovery integrates three major components aligned with SD learning phases [33]: (1) orientation & initial conceptualization with Smiley-Data mapping, (2) initial investigation with pairwise comparison and manual clustering, (3) further investigation & conceptualization with automatic clustering. Then we evaluated SmileyDiscovery with K-12 teachers due to their essential roles in integrating innovative technology for pedagogy [25]. Our research questions are: **RQ1.** Can SmileyDiscovery support K-12 teachers to carry out ML-empowered SD? **RQ2.** Can SmileyDiscovery support K-12 teachers to design SD learning activities? **RQ3.** Can SmileyDiscovery support learning ML? Our main contributions include:

1. SmileyDiscovery facilitating ML-empowered SD for K-12 STEM learning;
2. A set of ML-SD connections for K-12 teachers to design ML-empowered SD learning activities aligned with STEM curriculum;
3. Design implications for technology designers without SD background.

2 ML-Empowered SD and K-12 STEM Learning

Research shows that ML approaches empower data-driven discovery by enabling hypothesis generation, iterative experimentation with different parameters, and pattern recognition by gradually revealing more refined parameters [30,31]. Various ML techniques have been proposed to automate SD [27]. For example, k-means clustering, an unsupervised ML algorithm, is used to discover laws by grouping similar objects [13,14], identify dependencies of attributes [44], and form taxonomies [51]. Such methods, however, are applied in science at a professional level [16,21] and thus are inappropriate for K-12 teachers and students with limited CS/ML backgrounds. This points out a demand for designing an ML-empowered SD learning environment in K-12 contexts.

There are emerging research efforts to explore the opportunities of making ML concepts and methods accessible for K-12 students [14,28,49,53]. One study shows that data visualization supports students with limited computing knowledge to gain a basic understanding of cluster analysis [49]. Further, it indicates the potentials of applying ML methods for data interpretation by pattern generation. Another study facilitates youth to train and test ML models of their athletic activities [53]. It shows that ML enhances science learning by aligning ML modeling with modeling scientific phenomena, an essential practice of science recommended in curriculum standards [46]. Informed by these, our work aims to design a learning environment connecting ML components with SD practices. We incorporate design guidelines from existing research about introducing ML in K-12 STEM contexts, such as unveiling complex ML concepts step by step [14,28] and visualizing ML models for explainability [13,49,53].

3 The Design of SmileyDiscovery

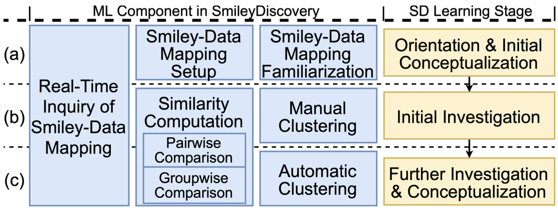

Fig. 1. SmileyDiscovery components: (a) orientation & initial conceptualization by Smiley-Data mapping; (b) initial investigation by pairwise comparison and manual clustering; (c) further investigation & conceptualization by automatic clustering.

We adopted *K-means clustering* to support SD due to its wide application in STEM domains [3,26,35,39]. Compared to supervised learning, unsupervised ML (e.g., clustering) more naturally connects with exploration leading to deeper learning in SD [43], and inductive reasoning through accumulative evidence [15,38], an accessible cognitive skill for young learners [30,40,42]. To make cluster analysis accessible for K-12 students, we used *Smiley* visualization [49], translating each data attribute to a facial feature, to take advantage of people's high processing capacity to human faces [6] and facilitate similarity computation with superposition comparative visualization [17].

SmileyDiscovery enables SD learning stages modified from well-established frameworks [1,5,33], including **orientation & initial conceptualization**, where learners get familiar with the topic and generate hypotheses based on prior knowledge; **initial investigation**, where learners explore dataset for preliminary analysis; **further investigation & conceptualization**, where learners iterate experiments and derive findings. Three components (Fig. 1) support scaffolding for SD [37] to instruct toward higher complexity [24]. This includes introducing from basic (e.g., pairwise comparison) to advanced ML components (e.g., automatic clustering) and from a small subset to the entire dataset. Further, we designed typing boxes to record the generation and refinement of hypotheses.

We collaborated with an experienced science educator and designed three ML-empowered SD learning activities: ecosystems [32], wine chemistry [8], and breast cancer diagnosis [11]. Two ML experts checked the appropriateness of applying k-means clustering in those SD activities. Below we present how SmileyDiscovery supports SD learning across stages with the ecosystem activity.

3.1 Orientation and Initial Conceptualization

First, learners are introduced to multidimensional data about ecological field sites (Fig. 2(a)). Second, they propose initial hypotheses and drag attributes of interest onto facial features (Fig. 2(b)). Such an active construction of Smiley can better engage learners [41]. Third, they manipulate sliders to understand how data attribute values influence corresponding Smiley facial features (e.g., a lower latitude of a field site leads to a smaller mouth). To reduce the cognitive load of memorizing mapping relationships through SD, learners can view Smiley-data mapping in real-time by hovering the cursor over facial features (Fig. 2(d)).

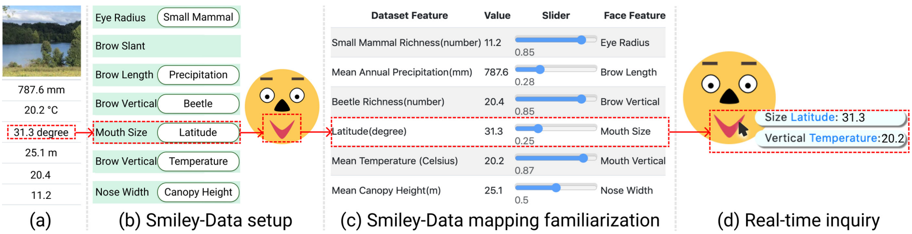

Fig. 2. SmileyDiscovery component supporting orientation & initial conceptualization.

3.2 Initial Investigation

First, learners use pairwise comparison to identify intriguing patterns between two pre-selected field sites representing two distinct ecosystem clusters. This design is informed by contrastive explanation [7] stimulating abductive reasoning [15,40]. E.g., the distinctions between two Smileys (Fig. 3(a)) may trigger learners to wonder if lower latitudes relate to higher temperatures, precipitation, canopy, beetle richness. Second, learners click on Smileys to overlay them on the representative field sites (Fig. 3(b)) to select similar ones. This trial-and-error process supports deeper reasoning [18,20,34] about (dis)similarities unveiled (e.g., some field sites share low latitudes and high canopy & beetle richness, while some share high latitudes and low canopy & beetle richness).

3.3 Further Investigation and Conceptualization

First, learners select a value of k (Fig. 3(c)). Second, learners conduct inductive cluster analysis by investigating (1) shared features within clusters (intra-cluster pattern) via visual inspection of the stack of Smileys belonging to the same group (Fig. 3(d)); (2) differentiating features between clusters (inter-cluster pattern) by overlaying the two average Smileys (centroid) (Fig. 3(e)). To still consider intra-cluster variations while using a centroid representing each cluster, learners can click on each cluster to switch the view between a stack (Fig. 3(d)) and a centroid (Fig. 3(e)). With variations and patterns introduced by the entire dataset, learners are expected to concentrate on fewer ecological attributes than initial investigation. Third, learners synthesize accumulative evidence from intra-&inter-cluster patterns for further conceptualization (Fig. 3(f)). E.g., the first two clusters show that a high canopy may lead to high beetle richness, and field sites with similar latitudes have similar precipitations and temperatures.

The components above naturally open up the black-box of ML by asking students to gradually apply similarity computation, centroid, evaluating values of k with intra-&inter-cluster patterns. The algorithmic process of the k-means clustering is also implicitly embedded in the scaffolding for manual clustering.

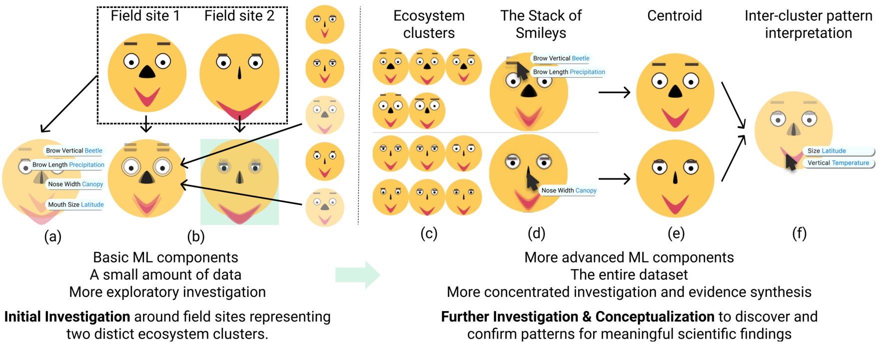

Fig. 3. SmileyDiscovery components supporting the initial investigation and further investigation & conceptualization.

4 Methods

4.1 Study Design

Eighteen in-service K-12 STEM teachers without CS/ML backgrounds were recruited from a teacher education course, *Integrating Technology with STEM*

Table 1. Participant information for each group.

Group	Teaching grades	Subjects
1	Elementary (N = 3), Middle school (N = 1)	Science (N = 2), Math (N = 2)
2	Middle school (N = 4)	Science (N = 3), Math (N = 1)
3	High school (N = 4), Middle school (N = 1)	Science (N = 1), Math (N = 4)
4	High school (N = 5)	Science (N = 5)

Teaching, at a research-based university. They were divided into four groups based on their teaching experience (grades & subjects) (Table 1) and participated in the study via an online meeting platform, Zoom. The study contains two sessions in two consecutive weeks. In the *Teacher-as-Learner (TaL)* session, teachers watched a tutorial video about completing an ML-empowered SD activity in SmileyDiscovery and performed another one with in-time help from researchers. In *Teacher-as-Designer (TaD)*, teachers collaboratively designed ML-empowered SD lesson plans via an online design canvas by specifying each instruction step and selecting SmileyDiscovery components to facilitate corresponding steps.

4.2 Data Collection and Analysis

RQ1. Can SmileyDiscovery Support K-12 Teachers to Carry Out ML-empowered SD? We collected log data of how teachers went through the example activity in TaL, including text input and clicking behaviors (Table 2). We measured successful completions with text input by examining (1) if all questions are answered based on proper ecological attributes, (2) if further conceptualization involves meaningful findings emerging from the data; we then counted clicking behaviors to see if teachers interacted elaborately with ML components. We measured patterns in successful completions by examining hypothesis development and comparing differences in clicking behaviors between successful and unsuccessful completions. Four participants who encountered technical issues were excluded.

Table 2. Log data (text input & clicking behaviors) collected for each SD stage.

Orientation & Initial conceptualization Text input: hypothesis of ecological interactions based on prior knowledge
Initial investigation Clicking behavior: (1) select field sites similar to the two representative field sites for manual clustering; (2) remove less similar field sites from a cluster. Text input: (3) interpretation of shared patterns identified manually; (4) interpretation of differentiating patterns between ecological field site subsets
Further investigation & conceptualization Clicking behavior: (1) conduct automatic clustering with different values of k; (2) switch between Smiley stacks and centroids; (3) compare centroids of ecosystem clusters. Text input: (4) interpretation of shared & differentiating patterns in ecosystem clusters; (5) findings of dynamic interactions between ecological attributes

RQ2. Can SmileyDiscovery Support K-12 Teachers to Design SD Learning Activities? We explored the pedagogical potentials of SmileyDiscovery by asking teachers to (1) post SmileyDiscovery-supported teaching ideas before TaD; (2) collaboratively design ML-empowered lesson plans in TaD; (3) reflect on applying

SmileyDiscovery in teaching in journals after TaD. First, we measured the diversity of teaching ideas. Two researchers independently assessed each teaching idea on whether it includes multidimensional datasets and applies cluster analysis to solve problems, then independently categorized teaching ideas into NGSS disciplinary core ideas [46]. Both achieve near-perfect agreement (Cohen's Kappas: 0.87, 0.91). Second, to measure teachers' fulfillment of ML-SD connections, we identified ML components selected to support each SD phase in teacher-designed lesson plans and counted each connection. Third, we measured teachers' perceptions toward ML-empowered STEM teaching. Two researchers independently coded teachers' reflection journals using thematic analysis, meeting regularly to address disagreements and refine codes.

RQ3. Can SmileyDiscovery Support Learning ML? We administered pre-post tests before and after the TaL to assess teachers' understanding of k-means clustering. Two researchers independently rated the tests, achieving near-perfect agreement with Cohen's Kappas of 0.85 (pre) and 0.83 (post). We measured learning gains by paired t-test as the data satisfies normal distribution. Then we measured the remaining misconceptions by thematic analysis on teachers' answers from post-tests. Two raters coded each incorrect answer independently, reaching near-perfect agreement (Cohen's Kappas above 0.86 for all items).

5 Results

5.1 RQ1. Can SmileyDiscovery Support K-12 Teachers to Carry Out ML-empowered SD?

Completion of ML-empowered SD Learning. 10 out of 14 teachers successfully completed all SD questions and generated meaningful findings of dynamic interactions between ecological attributes through cluster analysis. Two teachers needed to further articulate relationships identified, while the rest two didn't answer the last question for further conceptualization.

The numbers of ecological attributes involved in the investigation show that teachers naturally started with a more exploratory style by looking out attributes as much as possible. Then they reduced the scope as more evidence emerged from the entire dataset. During the initial investigation, 10 out of 14 teachers ended up with clusters sharing high similarity for more than four out of six ecological attributes. After automatic clustering, 10 out of 14 learners narrowed down to fewer attributes most strongly supported by data.

The numbers of different clicking behaviors show that teachers went through all ML components, with more frequent interactions for some of them than others. Specifically, teachers spent much time on manual clustering for initial investigation. On average, they selected 17.43 (SD = 10.21) field sites to compare with two representative field sites, removed 9.07 (SD = 10.54) field sites that are not similar enough, and reserved 8.36 (SD = 2.34) field sites for pattern interpretation. In comparison, teachers roughly played with different values of k for automatic clustering. They tried less than one new k-value ($M = 0.71$, $SD = 0.99$) in addition to the two rounds required by the instruction.

Patterns in Successful Completions. We identified two patterns in hypothesis development. (1) *Iterated initial hypotheses* ($N = 7$): Hypotheses became more specific or more inclusive from initial to further conceptualization. E.g., one teacher initially hypothesized that "latitude and mean temperature are related". In the end, she collected evidence for "different latitudes influence the rest of the ecological attributes a lot". (2) *Generating new findings* ($N = 3$): Original hypotheses were rejected, and new ones were proposed through investigation.

Teachers who successfully completed interacted more with manual clustering (selection: $M = 17.70$, removal: $M = 8.70$) than those who didn't (selection: $M = 4.06$, removal: $M = 3.56$). In further investigation & conceptualization, teachers who successfully completed switched between Smiley stacks and centroids ($M = 9.5$) more than those who didn't ($M = 4$) and compared centroids ($M = 5.2$) more than those who didn't ($M = 1.5$). These indicate the importance of an extensive engagement with similarity computation and sufficient pattern interpretation for generating meaningful findings.

5.2 RQ2. Can SmileyDiscovery Support K-12 Teachers to Design SD Learning Activities?

Diversity of Teaching Ideas. 37 out of 46 teaching ideas were identified as qualified, across science ($N = 31$), mathematics ($N = 4$), and social studies ($N = 2$). For science subjects, we identified 11 out of 13 NGSS [45] core disciplinary ideas, such as biological evolution and engineering design. Three primary learning objectives are identified from the teaching ideas: (1) categorize complex phenomena into groups and describe the patterns (e.g., discover biological patterns in different organisms); (2) understand interactions between different attributes within a system (e.g., investigate relationships between temperatures, humidity, surface types, and bacteria found in different locations); (3) identify the factors most relevant to cause the change/development of a system (e.g., investigate organism traits in different environments and find out which are more critical for survival). These results suggest SmileyDiscovery's pedagogical potential to fulfill a variety of K-12 STEM learning objectives aligned with the curriculum.

Teachers' Fulfillment of ML-SD Connections. Topics of teacher-designed lesson plans are (1) construction materials for flood resistance, (2) biological characteristics & evolution, (3) influential factors to income, and (4) risk factors for heart disease. Two researchers applied the EQuIP rubric [45] and confirmed each lesson plan's alignment with NGSS standards [46]. Patterns in the ML-SD connections applied by teachers are analyzed (Fig. 4). First, similarity computation is used for conceptualization, different from example SD activities. Teachers preferred hypothesis generation through abduction based on a small amount of data rather than prior knowledge. E.g., group 4 asked students to generate initial hypotheses by observing factors' puzzling impacts on heart disease risk. However, accelerating hypothesis generation by ML-revealed patterns is missing from teachers' design. Second, automatic clustering is frequently used for investigation. Two groups designed iterative investigation from small to large datasets. Group 3 proposed to run clustering with different sets of attributes, then compare results from each trial to refine hypotheses of what factors influence a person's income the most [4,51]. Moreover, all groups added a new design for prediction, such as predicting heart disease risk to evaluate the refined hypothesis.

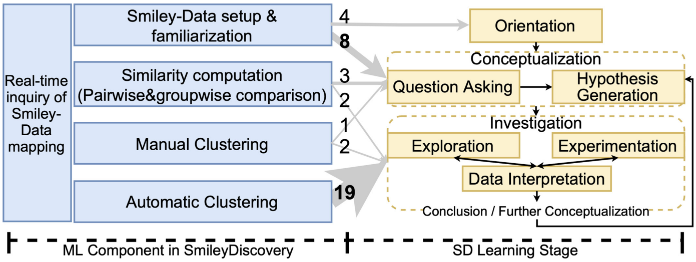

Fig. 4. ML-SD connections identified in four teacher-designed lesson plans.

Teachers' Perceptions Toward ML-empowered STEM Teaching. Teachers appreciated SmileyDiscovery's novelty as a teaching tool as it makes the large data accessible for K-12 students for pattern exploration and interpretation (N = 14), offers a playful learning experience to engage students (N = 10), low barrier to entry (N = 12), and can be applied in various STEM subjects (N = 13). After designing an SD learning activity on what factors influence a person's income, one teacher expressed her wish to conduct the learning activity with the Advancement Via Individual Determination (AVID) program she is teaching: "*If we do create it for real, I can do it with AVID!*" Nevertheless, teachers expected to gain a deeper understanding of ML methods (N = 5) and ML-empowered instruction design (N = 7) before implementing it in actual classrooms.

5.3 RQ3. Can SmileyDiscovery Support Learning ML?

The mean differences of all questions between pre- and post-tests were normally distributed at an alpha level of 0.05. A paired-sample t-test showed significant increases (Table 3) from pre- to post-test for four k-means clustering concepts: similarity computation, centroid, clustering process, evaluating values of k with intra-&inter-cluster pattern interpretation. This suggested that SmileyDiscovery successfully supported teachers to gain a rapid understanding of k-means clustering while applying it for SD. The answers indicate some misconceptions. For *similarity computation*, five teachers only addressed the subjectivity that the decision-making changes based on different criteria. For *evaluating values of k*, nine teachers didn't demonstrate comprehensive procedures, such as using a centroid to represent a cluster without considering intra-cluster variations.

Table 3. Paired t-test results for pre- and post-tests (N = 18)

Questions (Scores range 0–3)	Pre-test		Post-test		*t-test*	*p*
	M	SD	M	SD		
What makes two multidimensional datapoints similar or dissimilar?	0.31	0.49	1.50	1.23	−4.26	0.001
What is the centroid of a cluster of data points?	0.89	1.08	1.75	1.19	−2.67	0.031
Order the major steps for the K-means clustering algorithm	0.97	0.60	1.56	0.78	−3.58	0.002
How to decide which value of k gives better clustering results?	0.25	0.49	1.14	1.04	−4.05	0.001

6 Discussion and Future Work

SmileyDiscovery aims to bridge the gaps in ML-enhanced & curriculum-aligned STEM learning [29,48] for K-12 students and teachers with limited computing backgrounds [2,52]. Results show that K-12 teachers applied ML to discover meaningful scientific findings and simultaneously understood related ML concepts and methods. Teaching ideas and lesson plans show SmileyDiscovery's pedagogical potential in diverse K-12 STEM subjects. Teachers also reported that SmileyDiscovery is an innovative and playful way with a low entry barrier to "*explore data and draw connections with visualization*".

Informed by the study findings, we identified three key design implications for more effective ML-empowered SD. First, it's critical to design efficient scaffolding for ML visual analytics [12,36], as teachers novice to ML tended to carry out less sufficient investigation and synthesis of ML-generated patterns. For example, immediate feedback [23] can be designed to address common challenges in analyzing ML-generated results, such as outlier interpretation and considering intra-cluster variations while interpreting inter-cluster patterns. Second, advanced design to support converting visual representation (e.g., Smiley) to data is needed to support efficient sense-making in the context of subject matter, as teachers reported that the frequent manual Smiley-data translation was overwhelming when interpreting the ML-generated patterns. The advanced design may involve automating such non-salient & routine tasks [37] to reduce cognitive load for SD, which already requires high working memory [22]. Third, trial-and-error should be encouraged by a more inviting design for exploratory ML-enhanced investigation [20], as teachers with better SD performance experimented with more Smileys for similarity computation during manual clustering.

In the teacher-designed learning activities, no teacher applied automatic clustering for conceptualization, indicating certain biases introduced by example SD activities. A customizable authoring system can be designed to provide personalized recommendations of a list of potential ML-SD connections for teachers to select from based on their teaching objectives. Besides, teachers' after-study reflection shows a need to reveal more advanced mathematical knowledge about ML methods: "*While I can conceptualize the process, the mathematical computations in the analytic is a bit abstract to me.*" Technical tutorials, such as an interactive workbook [47], can be embedded as supplementary supports.

Limitations and Future Work. As a preliminary study to explore an innovative system [19], our work has several limitations. First, COVID-19 interruption and remote participation constrained data collection and undermined teachers' engagement. Second, the study didn't include a control condition. Thus, our next step is to evaluate the educational effectiveness of an improved SmileyDiscovery on students' learning of scientific knowledge and skills, compared to traditional computer-supported SD learning environments. For more effective and accurate science learning, a component to review the main takeaways can be added at the end of an ML-empowered SD learning activity. Besides, we plan to extend SmileyDiscovery with other similarity-based and supervised ML algorithms [50], engaging learners to derive evaluable scientific laws through SD.

References

1. Understanding science. http://www.understandingscience.org (2021). Accessed 15 Jan 2021
2. Agassi, A., Erel, H., Wald, I.Y., Zuckerman, O.: Scratch nodes ML: a playful system for children to create gesture recognition classifiers. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–6 (2019)
3. Ay, M., Kisi, O.: Modelling of chemical oxygen demand by using ANNs, ANFIs and k-means clustering techniques. J. Hydrol. **511**, 279–289 (2014)
4. Boutsidis, C., Drineas, P., Mahoney, M.W.: Unsupervised feature selection for the k-means clustering problem. In: Advances in Neural Information Processing Systems, pp. 153–161 (2009)
5. Bybee, R.W., et al.: The BSCS 5E instructional model: origins and effectiveness. In: BSCS, vol. 5, pp. 88–98. Colorado Springs, Co. (2006)
6. Chernoff, H.: The use of faces to represent points in k-dimensional space graphically. J. Am. Statist. Assoc. **68**(342), 361–368 (1973)
7. Chin-Parker, S., Bradner, A.: A contrastive account of explanation generation. Psychon. Bull. Rev. **24**(5), 1387–1397 (2017)
8. Cortez, P., Cerdeira, A., Almeida, F., Matos, T., Reis, J.: Modeling wine preferences by data mining from physicochemical properties. Decis. Support Syst. **47**(4), 547–553 (2009)
9. National Research Council: National Science Education Standards. National Academies Press (1996)
10. De Jong, T., Sotiriou, S., Gillet, D.: Innovations in STEM education: the Go-Lab federation of online labs. Smart Learn. Environ. **1**(1), 1–16 (2014)
11. Dua, D., Graff, C.: UCI Machine Learning Repository. School of Information and Computer Science, University of California, Irvine, CA (2019)
12. Endert, A., et al.: The state of the art in integrating machine learning into visual analytics. In: Computer Graphics Forum. vol. 36, pp. 458–486. Wiley Online Library (2017)

13. Essinger, S.D., Rosen, G.L.: An introduction to machine learning for students in secondary education. In: 2011 Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE), pp. 243–248. IEEE (2011)
14. Evangelista, I., Blesio, G., Benatti, E.: Why are we not teaching machine learning at high school? a proposal. In: 2018 World Engineering Education Forum-Global Engineering Deans Council (WEEF-GEDC), pp. 1–6. IEEE (2018)
15. Folger, R., Stein, C.: Abduction 101: reasoning processes to aid discovery. Hum. Resource Manage. Rev. **27**(2), 306–315 (2017)
16. Gil, Y., Greaves, M., Hendler, J., Hirsh, H.: Amplify scientific discovery with artificial intelligence. Science **346**(6206), 171–172 (2014)
17. Gleicher, M., Albers, D., Walker, R., Jusufi, I., Hansen, C.D., Roberts, J.C.: Visual comparison for information visualization. Inf. Visual. **10**(4), 289–309 (2011)
18. Hitron, T., Orlev, Y., Wald, I., Shamir, A., Erel, H., Zuckerman, O.: Can children understand machine learning concepts? The effect of uncovering black boxes. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–11 (2019)
19. IES: Common guidelines for education research and development. https://www.nsf.gov/pubs/2013/nsf13126/nsf13126.pdf (2013)
20. Käser, T., Schwartz, D.L.: Modeling and analyzing inquiry strategies in open-ended learning environments. Int. J. Artif. Intell. Educ. **30**(3), 504–535 (2020)
21. Kitano, H.: Artificial intelligence to win the nobel prize and beyond: creating the engine for scientific discovery. AI Mag. **37**(1), 39–49 (2016)
22. Klahr, D., Dunbar, K.: Dual space search during scientific reasoning. Cognit. Sci. **12**(1), 1–48 (1988)
23. Van der Kleij, F.M., Feskens, R.C., Eggen, T.J.: Effects of feedback in a computer-based learning environment on students' learning outcomes: a meta-analysis. Rev. Educ. Res. **85**(4), 475–511 (2015)
24. Koedinger, K.R., Booth, J.L., Klahr, D.: Instructional complexity and the science to constrain it. Science **342**(6161), 935–937 (2013)
25. Koehler, M., Mishra, P.: What is technological pedagogical content knowledge (TPACK)? Contemp. Issues Technol. Teach. Educ. **9**(1), 60–70 (2009)
26. Kupfer, J.A., Gao, P., Guo, D.: Regionalization of forest pattern metrics for the continental united states using contiguity constrained clustering and partitioning. Ecol. Inform. **9**, 11–18 (2012)
27. Langley, P.: The computational support of scientific discovery. Int. J. Hum.-Comput. Stud. **53**(3), 393–410 (2000)
28. Lin, P., Van Brummelen, J., Lukin, G., Williams, R., Breazeal, C.: Zhorai: designing a conversational agent for children to explore machine learning concepts. In: AAAI, pp. 13381–13388 (2020)
29. Marques, L.S., Gresse von Wangenheim, C., Hauck, J.C.: Teaching machine learning in school: a systematic mapping of the state of the art. Inform. Educ. **19**(2), 283–321 (2020)
30. McAbee, S.T., Landis, R.S., Burke, M.I.: Inductive reasoning: the promise of big data. Hum. Resource Manage. Rev. **27**(2), 277–290 (2017)
31. Muller, M., Guha, S., Baumer, E.P., Mimno, D., Shami, N.S.: Machine learning and grounded theory method: Convergence, divergence, and combination. In: Proceedings of the 19th International Conference on Supporting Group Work, pp. 3–8 (2016)
32. NEON: National ecological observatory network data products: [neondp1.100220.001, dp1.10072.001]. http://data.neonscience.org (2020). Accessed 15 Jan 2021

33. Pedaste, M., et al.: Phases of inquiry-based learning: definitions and the inquiry cycle. Educ. Res. Rev. **14**, 47–61 (2015)
34. Penner, D.E., Giles, N.D., Lehrer, R., Schauble, L.: Building functional models: designing an elbow. J. Res. Sci. Teach.: Official J. Natl. Assoc. Res. Sci. Teach. **34**(2), 125–143 (1997)
35. Perini, F.: High-dimensional, unsupervised cell clustering for computationally efficient engine simulations with detailed combustion chemistry. Fuel **106**, 344–356 (2013)
36. Pike, W., et al.: The scalable reasoning system: lightweight visualization for distributed analytics. Inf. Visual. **8**(1), 71–84 (2009)
37. Quintana, C., et al.: A scaffolding design framework for software to support science inquiry. J. Learn. Sci. **13**(3), 337–386 (2004)
38. Rodrigues, C.T.: The method of scientific discovery in peirce's philosophy: deduction, induction, and abduction. Logica Universalis **5**(1), 127–164 (2011)
39. Romesburg, C.: Cluster Analysis for Researchers. Lulu (2004)
40. Ross, J.M.: Informatics creativity: a role for abductive reasoning? Commun. ACM **53**(2), 144–148 (2010)
41. Sawyer, R.K.: The Cambridge Handbook of the Learning Sciences. Cambridge University Press (2005)
42. Schulz, L.: The origins of inquiry: inductive inference and exploration in early childhood. Trends Cognit. Sci. **16**(7), 382–389 (2012)
43. Schwartz, D.L., Chase, C.C., Oppezzo, M.A., Chin, D.B.: Practicing versus inventing with contrasting cases: the effects of telling first on learning and transfer. J. Educ. Psychol. **103**(4), 759 (2011)
44. Skapa, J., Dvorsky, M., Michalek, L., Sebesta, R., Blaha, P.: K-mean clustering and correlation analysis in recognition of weather impact on radio signal. In: 2012 35th International Conference on Telecommunications and Signal Processing (TSP), pp. 316–319. IEEE (2012)
45. Achieve and National Science Teachers Association: EQuIP rubric for lessons and units: Science (2014)
46. NGSS Lead States: Next Generation Science Standards: For states, by states. National Academies Press, Washington, DC (2013)
47. Tang, D., et al.: Empowering novices to understand and use machine learning with personalized image classification models, intuitive analysis tools, and MIT App Inventor. Ph.D. thesis, Massachusetts Institute of Technology (2019)
48. Vazhayil, A., Shetty, R., Bhavani, R.R., Akshay, N.: Focusing on teacher education to introduce AI in schools: Perspectives and illustrative findings. In: 2019 IEEE Tenth International Conference on Technology for Education (T4E), pp. 71–77. IEEE (2019)
49. Wan, X., Zhou, X., Ye, Z., Mortensen, C.K., Bai, Z.: Smileycluster: supporting accessible machine learning in k-12 scientific discovery. In: Proceedings of the Interaction Design and Children Conference, pp. 23–35 (2020)
50. Wang, D., Yang, Q., Abdul, A., Lim, B.Y.: Designing theory-driven user-centric explainable AI. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–15 (2019)
51. Wang, D., Nie, F., Huang, H.: Unsupervised feature selection via unified trace ratio formulation and K-means clustering (TRACK). In: Calders, T., Esposito, F., Hüllermeier, E., Meo, R. (eds.) ECML PKDD 2014. LNCS (LNAI), vol. 8726, pp. 306–321. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44845-8_20

52. Zhang, Y., Wang, J., Bolduc, F., Murray, W.G., Staffen, W.: A preliminary report of integrating science and computing teaching using logic programming. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 9737–9744 (2019)
53. Zimmermann-Niefield, A., Turner, M., Murphy, B., Kane, S.K., Shapiro, R.B.: Youth learning machine learning through building models of athletic moves. In: Proceedings of the 18th ACM International Conference on Interaction Design and Children, pp. 121–132 (2019)