Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #47571

Experience with an Interdisciplinary Competition-based Cybertraining Workshop

Dr. Bing Wang, University of Connecticut

Dr. Bing Wang is a Professor in the School of Computing at the University of Connecticut (UConn). She earned her Ph.D. in Computer Science from the University of Massachusetts, Amherst in 2005 and has been a faculty member at UConn since. Her primary research areas are computer networking, distributed systems, cyberinfrastructure, and cyber-physical systems.

Prof. Suining He, University of Connecticut

Suining He is an assistant professor in School of Computing at University of Connecticut. Suining He received the NSF CAREER Award in 2023, Google Research Scholar Program Award and NVIDIA Applied Research Accelerator Program Award in 2021, and two UConn Research Excellence Program (REP) Awards in 2022 and 2020, and held the Google PhD Fellowship in Mobile Computing in 2015, HKUST School of Engineering (SENG) PhD Research Fellowship Award in 2015–2016, and Hong Kong Telecom Institute of Information Technology (HKTIIT) Post-Graduate Excellence Scholarship in 2016. His scholarly works appear in WWW, SenSys, UbiComp, INFOCOM, TKDE, and TMC, and received the IEEE MASS Best Paper Runner-up Award in 2020 and IEEE RTSS Outstanding Paper Award in 2021. He was ranked among the Stanford's World's Top 2% Scientists in 2020-2024. He serves as the editor for the IEEE Journal of Indoor and Seamless Positioning and Navigation (J-ISPIN) since 2023.

Prof. Chuanrong Zhang, University of Connecticut

Chuanrong (Cindy) Zhang is a Professor in the Department of Geography and the Center for Environmental Science and Engineering at the University of Connecticut (UConn). With over 20 years of experience, her research interests and expertise lie in GIS, remote sensing, and spatial analysis. She has focused extensively on the spatial analysis of environmental and socioeconomic data, utilizing spatial statistics and GIS methodologies.

Abdul-Wassay Queeshi, University of Connecticut

Abdul-Wassay Qureshi is a graduate from the University of Connecticut with a B.S. from the School of Computing and a B.A. from the Department of Mathematics. His research interests include machine learning optimization, robotics, and the development of explainable machine learning models.

Dr. Weidong Li, University of Connecticut

Weidong Li is an adjunct faculty member and research professor in the Department of Geography at the University of Connecticut (UConn). With over 30 years of research experience, his research interests and expertise focus on spatial statistics, GIS, natural resource and environmental evaluation. He has published extensively in environmental informatics, geostatistics and GIScience.

Dr. Sanguthevar Rajasekaran, University of Connecticut

Sanguthevar Rajasekaran received his M.E. degree in Automation from the Indian Institute of Science (Bangalore) in 1983, and his Ph.D. degree in Computer Science from Harvard University in 1988. Currently he is the Director of the School of Computing, Board of Trustees Distinguished Professor, and Pratt & Whitney Chair Professor of CSE at the University of Connecticut. His research interests include Big Data, AI and Machine Learning, Quantum Computing, Bioinformatics, Algorithms, Data Mining, Randomized Computing, and HPC. He has published over 350 research articles in journals and conferences. He is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE), the American Association for the Advancement of Science (AAAS), the American Institute for Medical and Biological Engineering (AIMBE), and the Asia-Pacific Artificial Intelligence Association (AIAA). He is also an elected member of the Connecticut Academy of Science and Engineering and has been recognized among the world's top 2% of scientists in 2024, according to Stanford/Elsevier's Top 2% Scientist Rankings.

Dr. Wei Wei, University of Connecticut

Dr. Wei Wei is an Associate Professor in Residence in School of Computing at the University of Connecticut. He received a Ph.D. in Computer Science from the University of Massachusetts, Amherst in 2006. His research interests are computer networking, statistical analysis, performance evaluation, and machine learning.

Dr. Elizabeth Howard, University of Connecticut

Liz Howard is a Professor in the Department of Curriculum and Instruction at the University of Connecticut. Her research focuses on dual language education, and she has served as the PI or Co-Investigator of several large-scale projects exploring language and literacy outcomes, the development of sociocultural competence, effective program features, and teacher preparation in this context. She has also served as a senior advisor or external evaluator on a number of federally funded grants.

Experience with an Interdisciplinary Competition-based Cybertraining Workshop

Abstract

Cyberinfrastructure (CI) is widely used for data-driven research and scientific discovery. While its importance is well recognized in the scientific community and industry, the education and workforce development in using CI are lagging. In this paper, we report our experience in training students to use CI for conducting research and creating innovative applications through an interdisciplinary competition-based cybertraining workshop. A total of 10 students (5 undergraduate and 5 graduate students) completed the workshop. The participants were divided into four interdisciplinary teams, each with students in two disciplines, computer science and geography. Pre- and post-workshop surveys show that the workshop was well received by the students, who gained experience in CI and developed a greater interest in interdisciplinary collaboration. The experiences we gained from organizing this workshop and the lessons we learned can be helpful for other educators in training students to use CI.

Introduction

Cyberinfrastructure (CI) involves computing systems, data, software, visualization, and people. With the advances of data gathering technologies, large-scale datasets are becoming increasingly common-place. For example, modern satellites equipped with advanced sensors provide extensive imagery data, enabling continuous monitoring capabilities. Drones and Unmanned Aerial Vehicles (UAVs) offer flexible and high-resolution data collection for various applications, such as agriculture, disaster response, and urban planning. Internet of Things (IoT) sensors deployed in various environments collect real-time data on weather, air quality, soil moisture, and more. Smartphones and other portable devices with GPS and various sensors provide vast amounts of location-based data. CI supports the entire cycle of data acquisition, transfer, storage, processing, and visualization of large-scale data[1]. It has become a critical resource for many applications and scientific discovery[2,3]. While the importance of CI is well recognized in the scientific community and industry, the education and workforce development of CI are lagging. In this paper, we report our experience in training students to use CI for conducting research and developing innovative applications through a *interdisciplinary competition-based* cybertraining workshop that lasts for two weeks.

The goal of this workshop is for the participants to use various components in CI to solve an open-ended research problem using *large-scale spatio-temporal data* in an interdisciplinary team. During the workshop, the participants will go through the entire process of finding/downloading relevant data, analyzing data, and visualizing and presenting the results. Through this process, the participants will learn and demonstrate their skills of using CI in solving challenging problems and designing novel applications, as well as interdisciplinary collaboration.

The workshop is *interdisciplinary*, since working on CI often requires interdisciplinary collaboration. Specifically, we target students in Computer Science & Engineering (CSE) and Geography Departments, who have complementary background working with large-scale spatio-temporal data. In the CSE Department, while students have learned different aspects of CI, they have not formed a holistic view of a CI system, and have not fully appreciated the complexity and challenges involved. For instance, although many students have learned big data analytics, they are only familiar with traditional datasets (e.g., images, videos, and text), and have not learned how to analyze spatio-temporal data. In the Geography Department, students similarly do not have a holistic view of a CI system and the complexities involved. While students are trained in spatial thinking and to understand various aspects of spatio-temporal data, they have limited knowledge of computing infrastructure, lack proficiency in high-performance computing, programming, and software development, and are unfamiliar data analytics techniques to process, model, and analyze large-scale spatio-temporal data. One goal of organizing the workshop is to explore the potential of having students with diverse background to work together on a joint project to learn from each other and develop novel ideas.

The workshop is *competition-based*, during which student teams with interdisciplinary backgrounds (CSE and geography) compete with each other to solve an open-ended research problem. We chose a competition-based format for the workshop since it follows a student-centered pedagogy and has been shown to stimulate students interests and enhance learning[4,5].

We organized the workshop in Spring 2024 semester. Four teams completed the workshop. Students' pre- and post-workshop surveys show that they have gained valuable experience in learning and using CI. In addition, they have enjoyed working with students from diverse backgrounds. The results from the teams demonstrate impressive interdisciplinary collaboration and their critical thinking that are far beyond our expectations. Our positive experience indicates that competition-based interdisciplinary workshops are an effective mechanism for students to learn CI and become future workforce who can use CI effectively for their research and discovery. We report the lessons learned from organizing the workshop. Our experience can be beneficial to other educators who plan to organize similar activities.

Related Work

The design of the workshop is inspired by existing studies (see below) that show the benefits of project-based approach and competition, particularly in a team setting, in stimulating students' interest and enhance learning. Specifically, project-based learning is a student-centered education method in which learners acquire knowledge and skills by working on a project over an extended period and the project typically involves exploring and responding to a complex, real-world problem. The learning process in a project-based approach emphasizes active learning, collaboration, critical thinking, and applying theoretical knowledge to practical problems. It has long been recognized as an effective teaching method[6,7,8,9].

Competition is a commonly used mechanism to stimulate students' interest in computer science, with several well-known competitions such as ACM Collegiate Programming Contest, Google Code Jam, Facebook Hacker Cup, Kaggle Competitions. Studies have shown that competitions often bring out the best in individuals and motivate them to achieve more[10,11]. While some studies show negative effects of competition, e.g., causing stress and causing students to focus on winning instead of learning[12,13]. The negative effects are mitigated when competing in team settings, where team members cooperate to leverage their individual strength for the entire team[13,14,15].

In this workshop, in addition to leverage the benefits from project-based learning and team-based competition, we further place interdisciplinary cooperation as a critical component in judging rubrics. This is because work on cyberinfrastructure is often interdisciplinary in nature, and one important goal of this workshop is improving students' interests in interdisciplinary work. While there are other workshops on cyberinfrastructure[16,17], the design, goals and scope of this workshop differs significantly from others.

Workshop Outline and Logistics

Workshop Outline

The workshop was organized by six educators, four in the CSE Department and two in the Geography Department at the University of Connecticut (UConn). It was centered on solving an open-ended problem leveraging large-scale spatio-temporal data. We chose a problem on *bike sharing*, which is an interdisciplinary area explored by researchers in both computer science and geography. Specifically, bike-sharing systems generate large volumes of data related to bike usage, docking station status, and user behavior. Computer scientists may develop algorithms and use machine learning techniques to analyze the data for forecasting demand and optimizing bike distribution. Geographers may study the spatial distribution of bike stations and usage patterns. They may analyze how geographical factors such as urban layout, topography, and infrastructure influence bike-sharing usage for optimizing bike routes and station placements to ensure accessibility and efficiency.

The problem studied in the workshop is efficient management of bike sharing systems, which is important since bike sharing has become a major first-mile and last-mile mobility option for urban residents worldwide[18]. This is a challenging problem since bike flow has significant temporal and spatial variations, and are affected by many factors. For better management of bike sharing systems, participants were asked to conduct various geoscience analysis to gain insights of spatial distribution of bike stations and usage patterns (see below), and use the insights to predict bike flow, i.e., the number of pick-ups and returns in each geographic location. Accurate bike traffic prediction can help bike sharing system operators to efficiently rebalance their bike distributions[19]. In addition, the prediction can be further integrated with mobile apps or interactive bike websites to inform the citizens the predicted bike availability in the next few hours/minutes for better planning of their rides.

Participants were asked to use knowledge and approaches in both CSE and Geography, and present results relevant to both disciplines. We provided a dataset, guidelines and suggestions for the students to start the work; the rest of the work is open-ended, allowing the students to decide collaboratively within each team. Students were encouraged to search and download other relevant datasets, and also look into other relevant resources. We briefly describe the dataset and the suggested analysis below.

Dataset. We prepared a dataset based on data downloaded from New York City's open-source Citi Bike Dataset[20]. Specifically, it contains two types of data: (i) raw data (in .csv format), containing the raw data of the bike sharing stations, trips, and rider information for the month of October 2019, and (ii) training data (in .h5 format): containing the processed and formatted spatio-temporal tensor obtained from the raw data. Specifically, it contains a region in Manhattan, divided into a 16×8 grid. The processed data includes the inflow (bike return) and outflow (bike pickup) in each cell of the grid for each 30-minute interval.

Spatio-temporal analysis. We suggested that the participants can geocode the raw data using ArcGIS[21]. Example potential directions that we suggested included: (i) spatial analysis that uses GIS to identify areas with high demands and popular routes, (ii) temporal analysis to understand how bike usage varies throughout the day, week, or year, (iii) network analysis that uses GIS to analyze the connectivity and efficiency of the bike-sharing network to identify gaps in the network, (iv) integration with other transportation modes, i.e., how bike-sharing systems integrate with other modes of transportation, e.g., public transit, (v) accessibility analysis to identify areas with limited access to bike-sharing services, which can inform decisions on where to expand the network, and (vi) demographic and socioeconomic analysis, e.g., by overlaying bike-sharing data with demographic information from the census, to understand the demographic groups (e.g., age groups, income levels, ethnic backgrounds) that use the bike-sharing system.

Machine-learning based prediction. Participants can choose any machine learning models. They were encouraged to develop multiple models and compare their performance. We recommended them to partition the data based on time into two separate periods, one for model training and the other for validation. In addition, we encouraged them to explore other relevant datasets to improve the performance of a basic model that we provided.

The testing data was only released to the participants three days before the deadline. It contained five weeks of data, and the evaluation was based on the prediction accuracy of the last four weeks (in case one planed to use the first week as the input for an end-to-end machine learning model). The performance metric was the mean square error (MSE). While model accuracy was an important judging metric, we emphasized other judging rubrics (e.g., innovation, interdisciplinary collaboration and other technical merits; see judging rubrics later.

Workshop Schedule

The workshop was announced to the students in CSE and Geography Departments at the end of the Fall 2023 semester. The competition called for students with CSE or geography background, and the monetary awards range from \$100 (third prize) to \$300 (first prize). The registration was open for a month, until the beginning of the Spring 2024 semester.

After the registration ended, we divided the participants into multiple teams based on their prior experience and academic departments. Each team had 2-5 students covering CSE and geography background. The competition materials and teams were announced on 1/19/2024, Friday, the first Friday of the 2024 Spring semester. Each team was given two weeks, from 1/19/2024, Friday to 2/2/2024, Friday, to work on the project and present their results at the end of the workshop (in the afternoon of 2/2/2024).

Education/Training During Workshop

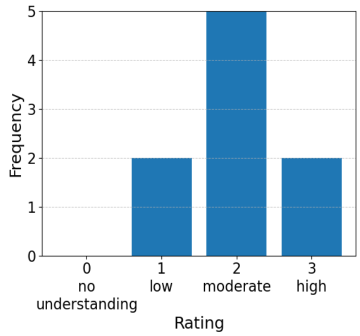
Since this is an education-oriented workshop, we provided the following materials to help the participants.

- Two reference papers on bike sharing systems were released at the beginning of the workshop. One paper[22] was from the discipline of computer science, and the other was from the discipline of geography[23].
- Guidelines on how to work in an interdisciplinary team was released at the beginning of the workshop. The guidelines were based on suggestions from Dr. Elizabeth Howard in the Department of Education at UConn, who had a lot of experience in facilitating interdisciplinary collaboration.
- A basic tutorial on how to read and analyze the data was released on Monday, 1/22/2024, three days after the start of the workshop. This was the time when we anticipated that students in each team had connected with each other and completed preliminary brainstorming. This tutorial was intended as as a reference for the students. In addition, it was meant for the students to focus on research related issues, instead of struggling with understanding the dataset.
- Coaching sessions (each 10-20 minutes) were provided to the students. Specifically, each team had two sessions with the faculty (the organizers): one in the first week to answer questions regarding the problem and provide suggestions on the team's planned approach, and the other in the second week for each team to show their preliminary results and seek further suggestions.

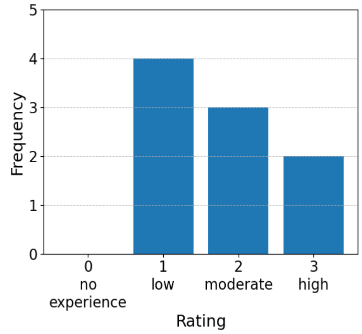
Judging Rubrics

To guide the teams' work, we announced the judging rubrics at the beginning of the workshop. The participants were informed that a committee of four judges (in the areas of CSE and Geography) will rank the teams based on three criteria: technical merit (40%), team collaboration (40%), and presentation quality (20%). We weighted technical merit and team collaboration equally to emphasize the equal importance of technical contributions and interdisciplinary collaboration among the team members with different backgrounds.

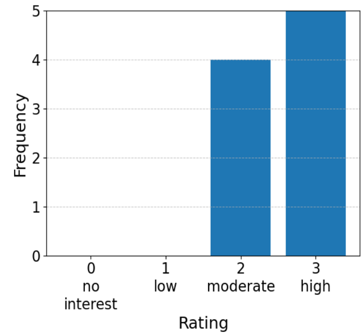
- For technical merit, each team was required to use knowledge and approaches from both CSE and geography, and obtain results relevant to both disciplines. The score was to be based on novelty and rationale of the approaches, as well as the significance and impacts of the results. Bonus points was awarded for finding additional relevant datasets and demonstrating the benefits of these datasets in solving or understanding the research questions.



(a) Understanding of CI.

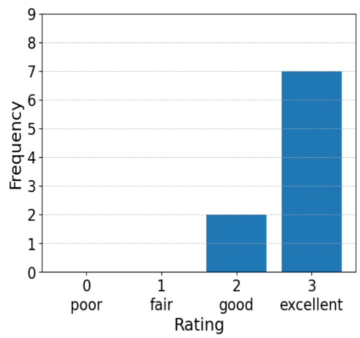


(b) Experience with interdisciplinary work.

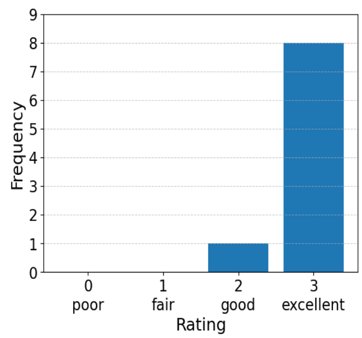


(c) Interests in interdisciplinary work.

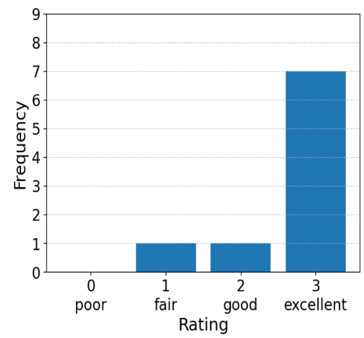
Figure 1: Pre-workshop survey: ratings on understanding of CI, experience with interdisciplinary work, and interests in interdisciplinary work before the workshop.



(a) Overall experience.



(b) Content.



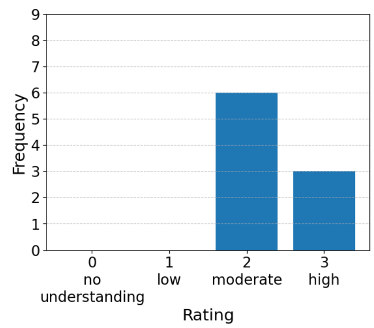
(c) Teamwork experience.

Figure 2: Post-workshop survey: rating on overall experience from the workshop, content of the workshop, and their experience within the team.

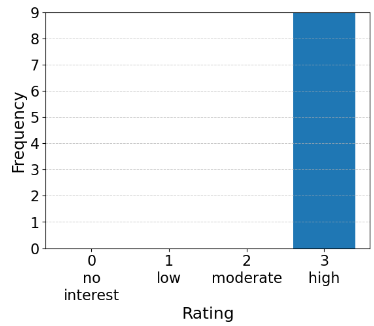
- For team collaboration, the score was based on the synergy of the work and the team members. Each team was required to reflect on the collaboration among the team members in terms of what worked and what did not work for their interdisciplinary collaboration.
- For the final presentation, we suggested a presentation structure and outlined the main components in the presentation. Each team member was required to participate in the presentation. In addition, each member was asked to present part of the materials that was outside their background (i.e., the work of other team members). As a result, participants in each team needed to communicate with each other and understand each other's work.

Participants and Workshop Outcome

A total of 12 students registered in the workshop. We assigned them into five interdisciplinary teams, each with 2-3 students. Later on, two students dropped out of the workshop, causing one team to have no students with CSE background. As a result, we had to merge two teams into a single team so that the combined team had both CSE and geography backgrounds. In summary, a total of four teams, comprising 10 participants, completed the workshop. Three teams consisted



(a) Understanding of CI.



(b) Interests in interdisciplinary work.

Figure 3: Post-workshop survey: ratings on understanding of CI and interests in interdisciplinary work after the workshop.

of two students each, while one team consisted of four students. The unbalanced team sizes were not intentional; rather were caused by the dropout of the two participants. Among these 10 participants, five were undergraduate students (all senior students), one was a Master's degree student, and four were doctoral students (ranging from their 2nd to 4th's year of PhD study). One participant was female and the rest were male. Among the four teams, two teams had a mixture of undergraduate and graduate students, one only had undergraduate students (due to the preference of these two students), and one only had graduate students (one undergraduate student who was originally assigned to the team dropped out).

Overall, each team did an excellent job, achieving impressive results in a short two-week period, far exceeding our expectations. The content and format of the workshop were also well received by the students. In the following, we briefly present the responses from the participants and results from their team project.

Survey Results

We asked participants to complete pre- and post-workshop surveys. For both surveys, 9 out of the 10 participants responded to the survey. The surveys were anonymous. So we do not know whether the same participants completed both sets of surveys. This is intended—for privacy reasons, we did not intend to link the responses of the pre- and post-workshop surveys. Rather, our goal was to obtain a general understanding on the effectiveness of the workshop and perception of the participants. Although we do not know for sure whether the two students who dropped out the competition participated in pre-workshop survey or not, we believe it is unlikely that they participated since they dropped out very early in the process.

Pre-workshop survey. Fig. 1a-c plot the responses to the three questions in pre-workshop: 'How would you rate your current understanding of cyberinfrastructure?', 'How would you rate your current level of experience with interdisciplinary work?', and 'How would you rate your level of interest in interdisciplinary work?'. We see that most students had low to moderate understanding of CI, and low to moderate experience with interdisciplinary work. In terms of their interests in interdisciplinary work, approximately half of the students showed moderate interests, while the other half showed high interests.

Post-workshop survey. The post-workshop survey asked the participants: ‘How would you rate your overall experience with the workshop?’, ‘How would you rate the content of the workshop?’. In addition, it inquired: ‘How would you rate your group process during the workshop?’ and ‘How would you rate your understanding of cyberinfrastructure after participating in the workshop?’. Fig. 2a-c show that the participants enjoyed the workshop: 7 participants rated the workshop as 3 (excellent), and 2 participants rated the workshop as 2 (good). They also rated the the content of the workshop positively, with 8 responses as 3 (excellent) and 1 response as 2 (good). Their ratings of their experience within the team were mixed: two ratings were fair and good, while 7 ratings were excellent; we will return to this point later in Discussion section.

Fig. 3 shows the responses to the two questions, ‘How would you rate your current understanding of cyberinfrastructure?’ and ‘How would you rate your level of interest in interdisciplinary work?’, that were asked in the pre-workshop. The goal was to evaluate whether students in general gained better understanding of CI after the workshop, and whether the workshop helped improving their interests in interdisciplinary work. Comparing Fig. 3a and Fig. 1a, while most students’ reported levels of prior understanding already in the ‘moderate’ or ‘high’ ranges, there was a positive shift towards the higher end of the scale, with all students reporting ‘moderate’ or ‘high’ levels of understanding by the end of the workshop, and higher percentages in each of those categories. This positive shift was particularly noteworthy in light of the very short 2-week timeframe of the workshop. Comparing Fig. 3b and Fig. 1c, we see post-workshop responses were unanimous in reporting ‘high’ levels of interest, whereas pre-workshop responses were divided across the ‘moderate’ and ‘high’ levels. As was the case with the increased reporting of cyberinfrastructure understanding, this positive shift was noteworthy in light of the very short 2-week timeframe of the workshop.

Results of Team Projects

Each team submitted their code and developed models for evaluation. In addition, each team presented their results at the end of the workshop. All the teams performed spatio-temporal analysis and machine learning based prediction of bike flows. We next briefly describe the highlights from their projects.

In terms of spatio-temporal analysis, students used ArcGIS tools to visualize the data in 2D maps (e.g., using heat maps and flow maps) and 3D space (e.g., using space-time cubes). For example, one team used space-time cubes to visualize both the spatial and temporal dimensions of bike-borrowing and bike-returning events simultaneously. This provided a comprehensive view of how these events evolved over time across different locations. By visualizing the data in a three-dimensional format, it became easier to identify trends, patterns, and anomalies in bike-borrowing and bike-returning that might not have been apparent in traditional 2D maps or time series plots. One team further developed a web-based dashboard to display various statistics and made several findings, including the age range of the primary user group, and characteristics of short and long trips.

In addition to the datasets that were provided, the teams explored other datasets. One team incorporated crime rate data, using clustering to identify unsafe bike station locations, particularly those with high frequency of bike usage and high crime rates. Another team examined bike lane data, road network, and census data, and used their analysis results to identify several features (e.g., population density, weather conditions, average speed, speed limit, bike lane length) for developing bike flow prediction models using machine learning.

In terms of machine learning based prediction models, all the teams first experimented with the basic model that we provided in the basic tutorial. After that, they all developed and evaluated various deep learning models using LSTM, GRU, or dense layers, and explored different structures of the models[24]. Several teams incorporated features obtained from the spatio-temporal analysis and conducted ablation study to demonstrate the benefits of using these additional features.

In addition to their analysis and prediction models, some teams gained a higher level understanding of interdisciplinary work and data analysis. One team summarized that they gained a better understanding of GeoAI[25], and found that spatio-temporal analysis and AI provided bidirectional benefits to each other: AI excels at recognizing complex bike-sharing trip patterns within large datasets, while spatio-temporal analysis provides AI models with critical geographical context that enhances the accuracy and relevance of predictions. Another team explored the benefits of analyzing data at multiple spatio-temporal scales. For example, they showed understanding peak usage times during the day and how bike usage changes over different seasons might help urban planners make informed decisions. Additionally, their multi-scale analysis helped identify anomalies or outliers that might have been missed when looking at a single scale.

All the teams reflected on their collaboration. They felt that they learned a lot from their team members with different backgrounds, which was consistent with the post-workshop survey showing increased interests in interdisciplinary work. Teams also reflected on what worked and what did not work in team collaboration. A common reflection was that that they would like more time for team work, a point we will return to later in Discussion section.

Discussion and Lessons Learned

Workshop outcome. Students learned different aspects of CI and data-driven research using CI. They learned about the importance of finding and incorporating relevant data sources for their analysis. They also found that visualization is a powerful tool not only for presenting the final results, but also in framing their approaches for analyzing the problem. For example, students found that spatio-temporal analysis using ArcGIS and the resultant visualization were very informative, helping them quickly grasp the main characteristics of the dataset. All the teams incorporated results from spatio-temporal analysis with writing Python code to develop software systems for bike flow prediction. All the students benefited from the process. The competition format engaged all team members to contribute to the outcome based on their individual background and skill set.

Interdisciplinary teams. During the final presentation, students reflected on their experience working with peers from different backgrounds. All the students found that exchanging knowledge and ideas with their teammates was beneficial to them. They felt that they had learned a lot from their teammates. Some students particularly commented on the importance of *in-person discussion*. They found that in-person meetings were particularly helpful for better exchanging ideas and fostering a deeper understanding among team members, ultimately leading to improved outcomes. The students' positive experience with interdisciplinary work was reflected in their stronger interests in interdisciplinary work in the post-workshop survey compared to the pre-workshop survey, which was very encouraging, particularly considering the two-week short duration of the workshop. Our design of enforcing interdisciplinary teams and facilitating collaboration among team members led to positive results.

Timing of setting up teams. Our post-workshop survey showed that some students felt that their teamwork experience could be improved. Our understanding is that this is mainly due to the insufficient amount of time that students had to work together with their teammates. During our coaching sessions with students, some teams mentioned the difficulty of finding common meeting times due to different schedules. One feedback we got was setting up the team earlier (instead of when releasing the problem), so that team members could have more time to know each other and brainstorm together. This is an excellent suggestion that can be helpful for future offerings of such workshops.

Duration of workshop. Another timing related issue is the duration of the workshop. We set it to be two weeks, considering that the workshop was at the beginning of the semester, and students might get busy soon. Indeed, for some courses, the first midterm exam might be in the fourth week of the semester. Feedback from the students indicated that they would like the workshop to be longer. In retrospect, we agree that it would indeed be better to set the workshop to be of a longer duration (e.g., three to four weeks). Having a longer duration would allow the students to spend more time doing background research and brainstorm the directions for solving the problems. It would also reduce the stress when certain analyses and models did not work as expected and allow them to further fine-tune their results.

Research-themed workshop for undergraduate students. While open-ended research is an important component of graduate education, it is typically not included in the undergraduate curriculum. Hence many undergraduate students do not have prior experience in research. Before hosting the workshop, we were uncertain how this research-themed workshop would be received by undergraduate students. It turned out that, with the various training materials we provided, the undergraduate participants were able to quickly learn the basics of doing research (in terms of reading and understanding the reference papers, framing the questions and steps in answering the questions) and had thoroughly enjoyed the workshop. This positive experience indicates that undergraduate students can quickly ramp up in learning challenging materials, and are not intimidated by the open-ended nature of research problems. In fact, the challenges can be stimulating to them, making them excited about the problem and motivating them to learn the necessary background to work on it.

Long-term and continuous post-workshop collaboration. Ideally, the enthusiasm that the students had shown toward the topic (CI, spatio-temporal analysis, machine learning, and other problems using CI) can continue after the workshop. One team expressed their interests in continuing their project and would like to extend their competition results for future publications. They mentioned that it would be achieved through an independent study that one student (from

Geography Department) would conduct during the semester, with the other student (in CSE Department) helping out as needed to continue the project. Finding effective approaches to extending the benefits of the workshop to achieve longer-term impact is important. We think one mechanism is to couple it with other course work or research activities. This can be arranged for the graduate students through their research activities. For undergraduate students, it might be helpful to arrange other curriculum activities (e.g., independent studies or design labs) or research programs (e.g., Research Experience for Undergraduates[26]) for the students.

Larger-scale offering of cybertraining workshops. Our results are based on one offering of the cybertraining workshop, from a relatively small number of 10 participants. More offerings of the workshop is needed to further validate our findings. In addition, further exploration is needed to expand the scope of the offering to a larger number of students, which may involve interesting challenges, e.g., how to team the students, how to coach the students and provide helpful guidance to maximize their experience from the workshop. We hope our experience will be helpful to educators in organizing such workshops or events in the future.

Conclusion

In this paper, we have reported our experience in organizing an interdisciplinary competition-based workshop to train students in understanding CI and using CI for research and developing novel applications. The workshop led to outcomes that far exceeded our expectations. We believe the competition-based format of the workshop is an effective mechanism for stimulating interests in learning CI and tackling open-ended research problems. Our efforts in enforcing and facilitating interdisciplinary collaboration led to positive experience for the participants, and improved their interests in interdisciplinary collaboration for their future work. We further presented lessons that we learned in the process, which can be helpful to other educators.

References

[1] National Science Foundation, Cyberinfrastructure Council, Cyberinfrastructure Vision for 21st Century Discovery (Washington, D.C.: National Science Foundation, March 2007). http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf.

[2] Daniel E. Atkins. Revolutionizing science and engineering through cyberinfrastructure: Report of the national science foundation blue-ribbon advisory panel on cyberinfrastructure. 2003. URL `https://api.semanticscholar.org/CorpusID:59671021`.

[3] James R. Bottum, James F. Davis, Peter M. Siegel, Brad Wheeler, and Diana G. Oblinger. Cyberinfrastructure: In tune for the future. *EDUCAUSE Review*, 43(4), July/August 2008.

[4] Juan C. Burguillo. Using game theory and competition-based learning to stimulate student

motivation and performance. *Computers Education*, 55(2):566–575, 2010. ISSN 0360-1315.

[5] S.A. Sukiman, H. Yusop, R. Mokhtar, and N. H. Jaafar. Competition-based learning: Determining the strongest skill that can be achieved among higher education learners. *Regional Conference on Science, Technology and Social Sciences*, 2016.

[6] Sylvia Chard. *The Project Approach: Developing the Basic Framework (Practical Guide 1)*. Scholastic Inc., USA, 1998.

[7] John W. Thomas. A review of research on project-based learning, 2000. `http://www.bobpearlman.org/BestPractices/PBL`$_Research.pdf$.

[8] John Howard. Technology-enhanced project-based learning in teacher education: Addressing the goals of transfer. *Journal of Technology and Teacher Education*, 10(3): 343–364, 2002.

[9] Houghton Mifflin. Project-based learning: Background knowledge & theory, 2003. `http://www.college.hmco.com/education/pbl/background.html`.

[10] M. Fasli and M. Michalakopoulos. Supporting active learning through game-like exercises. In *Proceedings of the 5th IEEE International Conference on Advanced Learning Technologies*, pages 730–734. IEEE, 2005. doi: 10.1109/ICALT.2005.159.

[11] R. Lawrence. Teaching data structures using competitive games. *IEEE Transactions on Education*, 47(4):459–466, 2004. doi: 10.1109/TE.2004.825053.

[12] S. Lam, P. Yim, J. Law, and R. Cheung. The effects of classroom competition on achievement motivation. In *Proc. of Annual Conference of the American Psychological Association*, 2001.

[13] Ghassan Issa, Shakir M. Hussain, and Hussein Al-Bahadili. Competition-based learning: A model for the integration of competitions with project-based learning using open source lms. *International Journal of Information and Communication Technology Education*, 10(1): 1–11, 2014. ISSN 1550-1876. doi: 10.4018/ijicte.2014010101.

[14] Ignacio Cantador and José M. Conde. A simple e-learning system based on classroom competition. In *Lecture Notes in Computer Science*, volume 6383, pages 488–493. Springer, 2010.

[15] Fu-Yun Yu, Liang-Jen Chang, Yi-Hsuan Liu, and Tak-Wai Chan. Learning preferences towards computerized competitive modes. *Journal of Computer-Assisted Learning*, 18(3): 341–350, 2002. doi: 10.1046/j.0266-4909.2002.00238.x.

[16] Cyberaccelerate workshop, October 2024. Available at: `https://tinyurl.com/9rnwmepe`.

[17] Cyberinfrastructure for education and learning for the future: A vision and research agenda. Technical report, Computing Research Association, 2005. Available at: `https://tinyurl.com/5c226se8`.

[18] Paul DeMaio. Bike-sharing: History, impacts, models of provision, and future. *Journal of public transportation*, 12(4):3, 2009.

[19] Mauro Dell'Amico, Eleni Hadjicostantinou, Manuel Iori, and Stefano Novellani. The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega*, 45:7–19, 2014.

[20] Citi Bike System Data. https://citibikenyc.com/system-data.

[21] ArcGIS Online. https://www.arcgis.com.

[22] Junbo Zhang, Yu Zheng, and Dekang Qi. Deep spatio-temporal residual networks for citywide crowd flows prediction. In *Proc. of AAAI*, 2017.

[23] Yunhe Cui, Xiang Chen, Xurui Chen, and Chuanrong Zhang. Competition, Integration, or Complementation? Exploring Dock-Based Bike-Sharing in New York City. *The Professional Geographer*, 75(1):65–75, 2023.

[24] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[25] Krzysztof Janowicz, Song Gao, Grant McKenzie, Yingjie Hu, and Budhendra Bhaduri. GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *International Journal of Geographical Information Science*, 34(4): 625–636, 2020.

[26] Research Experiences for Undergraduates (REU). https://www.nsf.gov/crssprgm/reu/.