Intelligent Recommendations for Citizen Science

Daniel Ben Zaken, [1] Kobi Gal, [1,3] Guy Shani [1] Avi Segal, [1] Darlene Cavalier. [2]
[1]Ben-Gurion University, [2] Arizona State University,
[3] University of Edinburgh
benzaked@post.bgu.ac.il,
{kobig, shanigu, avise}@bgu.ac.il,
darlene@scistarter.com

Abstract

Citizen science refers to scientKobiific research that is carried out by volunteers, often in collaboration with professional scientists. The spread of the internet has allowed volunteers to contribute to citizen science projects in dramatically new ways while creating scientific value and gaining pedagogical and social benefits. Given the sheer size of available projects, finding the right project, which best suits the user preferences and capabilities, has become a major challenge and is essential for keeping volunteers motivated and active contributors. We address this challenge by developing a system for personalizing project recommendations which was fully deployed in the wild. We adapted several recommendation algorithms to the citizen science domain from the literature based on memory-based and model-based collaborative filtering approaches. The algorithms were trained on historical data of users' interactions in the SciStarter platform - a leading citizen science site - as well as their contributions to different projects. The trained algorithms were evaluated in SciStarter and involved hundreds of users who were provided with personalized recommendations for new projects they had not contributed to before. The results show that using the new recommendation system led people to increased participation in new SciStarter projects when compared to groups that were recommended projects using non-personalized recommendation approaches, and compared to behavior before recommendations. In particular, the group of volunteers receiving recommendations created by an SVD algorithm (matrix factorization) exhibited the highest levels of contributions to new projects, when compared to the other cohorts. A follow-up survey conducted with the SciStarter community confirmed that users felt that the recommendations matched their personal interests and goals. Based on these results, our recommendation system is now fully integrated into the SciStarter portal, positively affecting hundreds of users each week, and leading to social and educational benefits.

1 Introduction

Citizen science engages people in scientific research by collecting, categorizing, transcribing, or analyzing scientific data (Bonney et al. 2009; Funk, Gottfried, and Mitchell 2017; Brossard, Lewenstein, and Bonney 2005). These platforms offer thousands of different projects which rely on the contributions of volunteers to extend scientific knowledge.

Citizen science brings significant scientific, pedagogical and social benefits. It provides new ways for the public to contribute to scientific research: Volunteers can share and contribute to data monitoring and collection programs. Community-based groups can generate ideas and engage with scientists for advice, leadership, and program coordination (Silvertown 2009; Bonney et al. 2014; Ryan et al. 2018). Citizen science can be used as an educational tool, allowing students, educators and scientists to network and promote new ideas to advance our understanding of the world (Vitone et al. 2016). Lastly, citizen science contributes to social inclusion by increasingly engaging volunteers from under-represented and marginalised communities (Haywood and Besley 2014; Sorensen et al. 2019).

SciStarter (scistarter.org) is an online citizen science hub which aggregates over 3,000 projects that are imported through partnerships with federal governments, NGOs, and universities. It is the world's largest catalogue of citizen science projects. Projects span different topics, age groups, location, etc. As of July 2020, SciStarter had 82,014 registered users from diverse age groups, socio-economic and educational backgrounds. Hundreds of the projects use SciStarter supported APIs to collect data from volunteers, who can track their participation across projects in their SciStarter dashboard. SciStarter maintains an active forum where volunteers and researchers can communicate directly. In addition to its scientific contribution, SciStarter also plays an important educational and social role. It collaborates with different institutions (schools, universities, libraries, museums, Girl Scouts, Discover magazine, and more) to customize citizen science pathways, and participates in organized events such as "Citizen Science Month" that promote hundreds of projects all over the world.

Volunteers visit SciStarter in order to discover new projects and keep up to date with community events. Examples of popular projects on the SciStarter platform include iNaturalist [1] in which users map and share observations of biodiversity across the globe; CoCoRaHS [2], where volunteers share daily readings of precipitation; and Stall-Catchers [3], where volunteers identify vessels in the brain as

[1]https://scistarter.org/seek-by-inaturalist
[2]https://scistarter.org/cocorahs-rain-hail-snow-network
[3]https://scistarter.org/stall-catchers-by-eyesonalz

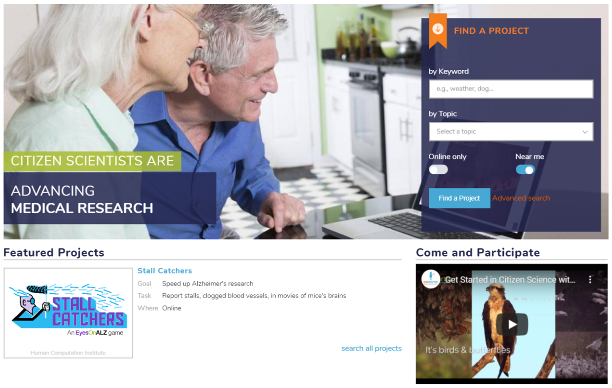


Figure 1: SciStarter User Interface

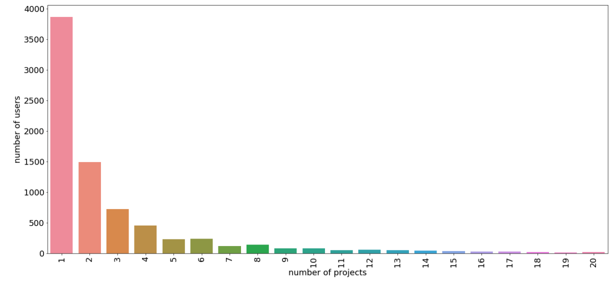


Figure 2: Distribution of user participation in SciStarter projects

flowing or stalled. Figure 1 shows the SciStarter main page window, including a search engine to locate projects, and a featured project that is hand-selected by the SciStarter admins.

SciStarter is one of several large scale citizen science portals, (such as Zooniverse (https://www.zooniverse.org/), CitSci.org, and Anecdata.org) that host between hundreds to thousands of different projects and connect volunteers, researchers and educators. On the one hand, the breadth and size of these portals provide an abundance of opportunities for volunteers to discover and contribute to new projects. There is evidence showing that motivated volunteers can provide quality contributions for several citizen science projects, increasing their social value in addition to their scientific contributions (Larson et al. 2020). On the other hand, the vast majority of citizen science volunteers perform tasks regularly in very few projects (Ponciano and Pereira 2019). To illustrate, Figure 2 shows a histogram of the number of projects that users contributed to on the site between 2017 and 2019. As shown by the figure, the majority of active users in the SciStarter portal do not contribute to more than a single project. A similar pattern of contributions was observed in the Zooniverse citizen science portal (Segal et al. 2018).

To help users discover new projects, SciStarter employs a search engine where users can find projects according to topics (e.g., Archaeology), activities (e.g., can be done online), location (e.g., at a science center or zoo) or age groups. However, recommending projects based on this tool has not been successful. Our analysis shows that about 80% of users do not use the search tool. In addition, data also shows that when users *do* use the search engine, most of them do not visit the projects that are outputted by the tool for their selection. Clearly, a more sophisticated approach is needed in this domain to match people with the right projects.

In this paper, we take an AI approach towards solving the recommendation problem in citizen science: how to match volunteers with new project recommendations in order to increase the number of activities that volunteers contribute to new projects on the SciStarter ecosystem. We match individual volunteers with new projects based on the past history of their interactions on the site (Dwivedi and Roshni 2017; Amatriain 2013).

According to a report from the National Academies of Sciences, Engineering, and Medicine (National Academies of Sciences, Medicine et al. 2018), citizen scientists' motivations are "strongly affected by personal interests," and participants who engage in citizen science over a long period of time "have successive opportunities to broaden and deepen their involvement." Thus our hypothesis was twofold. First, that personalizing recommendations to volunteers will increase their engagement in SciStarter and improve scientific outcomes, as measured by the number of projects that they contribute to, after being presented with recommendations, as well as the extent of their contributions to these projects. Second, that users will be satisfied with the recommendation tool and continue to be motivated contributors to SciStarter.

Recommendation systems have been used in other domains, such as e-commerce, news, and social media (Itmazi and Gea 2006; Fleder and Hosanagar 2007; Kleinerman et al. 2020). However, the nature of interaction in citizen science is fundamentally different than these domains. In citizen science, volunteers are actively encouraged to contribute their time and effort to solve scientific problems (Cohn 2008). Most citizen scientists contribute to few projects, compared to participants in online marketplaces and news sites who consume multiple items (e.g., movie recommendations). Thus solving the recommendation problem for citizen science can be considered a contribution from both scientific and social perspectives.

To address these challenges, we adapted different recommendation algorithms to the citizen science domain. Each algorithm matches a user profile and the user's past history of interactions and outputs a ranking of the most relevant n projects for the user, in decreasing order of relevance.

We conducted a randomized controlled study, in which hundreds of registered SciStarter users were randomly divided into cohorts, and assigned recommendations using different approaches. The first approach personalized projects to participants by using memory-based collaborative filtering algorithms (recommending projects to users based on user or item similarity), and matrix factorization algorithms (predicting the relevance of a new project to a user based on learned latent spaces). These algorithms were compared to two non-personalized algorithms: the first algorithm recommended the most popular projects to users, and the second algorithm recommended promoted projects that were manually determined by SciStarter admins at regular intervals (e.g., promoting a biology based project for Science Day).

The results show that people receiving the personalized recommendations were more likely to contribute to new projects that they had never tried before and participated more often in these projects when compared to participants who received non-personalized recommendations, or compared to behavior before the recommendations. In particular, the cohort of participants receiving recommendations created by the matrix factorization algorithm exhibited the highest levels of contributions to the recommended projects, when compared to the other personalized groups.

In a follow-up survey conducted with the SciStarter community, volunteers expressed high satisfaction with the recommendation tool, providing further support for its positive impact on engagement. Based on the positive results, our recommendation system is now fully integrated with SciStarter, providing recommendations to hundreds of users each day, significantly increasing the number of contributions in new projects by volunteers. This is the first study using AI based recommendation algorithms in a large scale citizen science platform.

2 Related Work

This research relates to past work in using AI to increase participants' motivation in citizen science research as well as work in applying recommendation systems in real world settings. We list relevant work in each of these two areas.

Active participation in citizen science projects through the internet is growing fast (Nov, Arazy, and Anderson 2014; Irwin 2018). Yet, most participants participating in citizen science projects perform only a few tasks each before leaving the system (Rotman et al. 2012). For example, less than 10% of all users contribute to more than 10 projects in the SciStarter portal. This reflects a general trend in volunteer-based crowdsourcing, whereby the majority of participants carry out only a few tasks (Segal et al. 2016, 2015).

Ponciano et al. (2019) showed that although volunteers tend to explore multiple projects in citizen science platforms, they perform tasks regularly in just a few of them. They also showed that volunteers recruited from other projects on the platform tend to get more engaged than those recruited outside the platform. This finding motivated our approach to recommend suitable projects for SciStarter .

Several works have studied the motivations of participants in citizen science. Kragh et al. (2016) showed that participants in citizen science projects are motivated by personal interest and a desire to learn something new, as well as their desire to volunteer and contribute to science. Raddic et al. (2009) claimed that citizen scientists mostly exhibit interest in a single topic, such as astronomy and zoology. The user survey we conducted on SciStarter reveals participants' interests to be more diverse and span multiple projects.

Other works have designed interventions for the purpose of increasing participants' engagement in citizen science. Segal et al. (2018; 2016) used AI planning to personalize motivational message policies which significantly increase users' contributions. Laut et al. (2017) showed that participants' contributions can be enhanced through the presence of virtual peers.

Several approaches have used recommendation algorithms to increase participants' engagement in education and social media settings. Labarthe et al. (2016) recommended educational content for students in Massive Open Online Courses (MOOCs) based on student profiles and their online activities. Dwivedi et al. (2017) used collaborative filtering to recommended online courses to students based on their past grades. Freyne et al. (2009) generated recommendations to users during sign-up to social network sites by leveraging aggregated external data from other social media sites. In contrast to these systems, users' profiles in SciStarter do not contain information relating to their project preferences, and we do not have access to their task performance correctness in the projects.

Lastly, we mention works suggesting novel recommendation algorithms to provide recommendations to users that were not evaluated in online settings. Wu et al. (2017) formulated the optimization of long-term user engagement as a sequential decision making problem, where a recommendation is based on both the estimated immediate user click and the expected number of clicks in the future. Lin et al. (2014) developed a recommendation system for crowdsourcing tasks which incorporates negative feedback (tasks that the user chose not to do) into a recommendation system using collaborative filtering. Both of these approaches rely on long term interaction with the user that is absent in our citizen science setting.

3 Methodology

Our approach to solve the recommendation problem in citizen science needed to address the following challenges: Most volunteers contribute to very few projects, and seldom return to the SciStarter portal after having chosen a project. In addition, only 153 projects (out of 3,000) actively report back a clickstream of user behavior to SciStarter. Lastly, many projects do not include content-based information such as topic, location and project description.

To address these challenges we adapted several canonical algorithms from the recommendation systems literature that do not rely on content or project-specific information. We restricted the training of the models to about 6,000 users who contributed to at least 2 projects, and measured the users' project clicks on the SciStarter website in addition to their active contributions to those projects that provide SciStarter with clickstream data. Each algorithm receives as input a target user and the number of recommendations to generate. The algorithm returns a ranking of the recommended projects in decreasing order of relevance for the user.

3.1 User-based KNN Collaborative Filtering

Collaborative filtering assumes that users with a history of contributing to the same projects would prefer to contribute to the same projects in the future. A user-based KNN collaborative filtering algorithm (Ning, Desrosiers, and Karypis

2015), identifies for each user u a set of k "neighbors" (users with similar histories as u, in that they interacted with the same projects). Then, we can recommend to u new projects that other users in her "neighborhood" have interacted with. For example, if both users u_1 and u_2 have contributed in the past to projects CoCoRaHS and Globe at Night, and u_2 has also contributed to project Stall Catchers, then Stall Catchers may be a suitable recommendation for u_1.

To determine whether users belong to the same neighborhood, we need to measure how similar they are in their past interactions with projects. We use the popular cosine similarity (Breese, Heckerman, and Kadie 1998), originating from measuring the angle between vectors. For binary settings, where users either contributed to a project or not, the cosine similarity may be computed using:

$$sim(u_1, u_2) = \frac{|I_{u_1} \cap I_{u_2}|}{|I_{u_1}| \cdot |I_{u_2}|} \quad (1)$$

where I_u is the set of projects that user u has contributed to. Then, we compute a score for each project i that u has not interacted with:

$$\hat{r}_{u,i} = \sum_{u' \in neighbors(u), i \in I_{u'}} sim(u, u') \quad (2)$$

that is, we go over all users in the neighborhood of u who have interacted with i and sum their similarities to u. We order the list of recommendations by decreasing $\hat{r}_{u,i}$.

In our domain, where users interact only with a small number of projects, we optimize the size of the neighborhood differently for each user u to be sufficiently high (threshold set empirically) such that there is a sufficient number of projects to recommend for u.

3.2 Item-based Collaborative Filtering

An orthogonal approach to the user-based KNN approach is an item-based KNN approach, where we compute for each project a neighborhood of other projects that similar users have interacted with (Schafer et al. 2007). For example, our data shows that 83% of users who have contributed to the project Never-Home-Alone (a project surveying wildlife in the home) have also contributed to project iNaturalist. Therefore we may recommend iNaturalist for a user who has already contributed to Never-Home-Alone".

Again, we require a similarity metric between items. The cosine similarity for items can be computed using:

$$sim(i_1, i_2) = \frac{|U_{i_1} \cap U_{i_2}|}{|U_{i_1}| \cdot |U_{i_2}|} \quad (3)$$

where U_i is the set of users who have contributed to project i. We then compute a score for item i that user u has not yet interacted with:

$$\hat{r}_{u,i} = \sum_{i' \in I_u} sim(i, i') \quad (4)$$

and order the recommendation list by decreasing $\hat{r}_{u,i}$.

3.3 Matrix Factorization

A more sophisticated approach attempts to identify latent features that characterize users and items. We compute for each user u and item i a vector of latent features (p_u and q_i, respectively), such that when the inner product between the vectors $p_u \cdot q_i$ is high, then u is likely to prefer i.

A well known approach for computing the latent vectors is the matrix factorization approach (Koren, Bell, and Volinsky 2009; Koren and Bell 2015). We consider the user-item interaction as a matrix $R_{|U| \times |I|}$, where each row represents a user, and each column represent an item, and $r_{u,i}$, the value in a cell, is 1, if the user u has interacted with item i. Then, we compute two matrices $P_{|U| \times k}$ and $Q_{|I| \times k}$, where k is a predefined number of latent features, such that $R \approx PQ^T$.

The matrix factorization approach is very popular in recommendation system research, and there are many methods for computing the matrices P and Q. Here, we chose to use the SVD algorithm (Sarwar et al. 2002) for computing the latent features.

Following the factorization of R into P and Q, which is computed offline, we recommend items for a user u by decreasing $\hat{r}_{u,i}$: $\hat{r}_{u,i} = p_u \cdot q_i$.

4 Results

The first part of the study compares the performance of the different algorithms to predict user behavior on historical SciStarter Data without the recommendation systems. The second part of the study implements a recommendation system in SciStarter, and actively assigns recommendations to users using the different algorithms. IRB approval to run the study was granted by the universities sponsoring the study.

4.1 Offline Study

The training set for all algorithms consisted of historical data collected between January 2012 to September 2019. It included 6353 users who contributed to 153 affiliate projects that use a dedicated API to report back to SciStarter each time a logged in SciStarter user has contributed data or analyzed data on that project's website or app. As data of contributions and participation only existed for the affiliate projects, we only used these projects in the study. We chronologically split the data using cross-set validation into train and test sets such that 10% of the latest interactions from each user are selected for the test set and the remaining 90% of the interactions are used for the train set. We also considered a non-personalized algorithm that recommended projects according to their popularity (the number of users who contribute to projects). Such algorithms have been shown to provide good results in several deployed recommendation systems (Ahn 2006; Jonnalagedda et al. 2016).

We evaluate the prediction accuracy of the top-n recommendation algorithms for $n = 3, 5, 7$ and 10 projects using the precision metric. This range of $3 - 10$ items reflects the range of recommendations presented to the user on the SciStarter site. We also consider the hit rate (percentage of instances when the user visited at least one of the projects

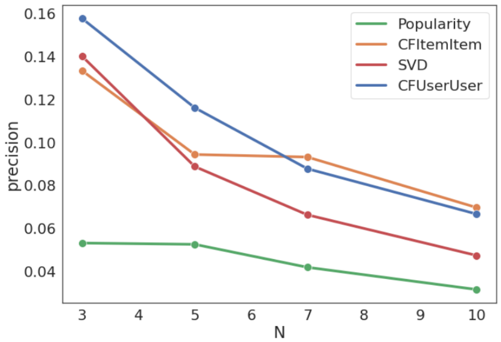


Figure 3: Precision results on offline data

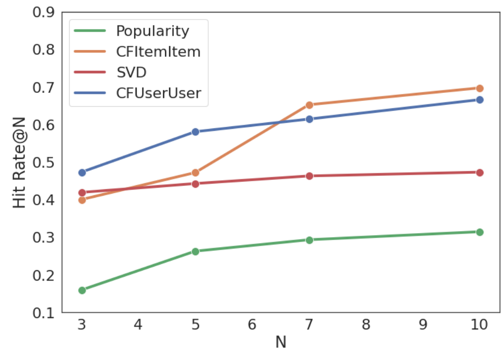


Figure 4: HitRate on offline data

that were predicted), which is commonly used in settings in which users consume few items (Wang, Guo, and Xu 2015)

Fig 3 shows results of the precision metric for the 4 examined algorithms for different numbers of recommended projects. As can be seen from the figure, collaborative filtering Item-based and user-based are the best algorithms and their performance is higher than popularity and SVD for the data points measured. The popularity recommendation algorithm generated the lowest performance.

An interesting result is that for all algorithms, precision drop drastically when recommending $n = 5$ projects and continues to decrease for $n = 7$ etc. The reason for this decline is that in citizen science, volunteers generally contribute to a low number of projects. For example, when $n = 5$, and the volunteer visited one of the new projects recommended by the algorithm, its precision rate will be 1/5, despite the fact that the algorithm was actually successful in getting the user to try a new project.

Fig 4 shows the hit rate of the different algorithms, defined as the percentage of instances in which users accessed at least one project that was recommended to them (Wang, Guo, and Xu 2015). The hit rate for all algorithms rises consistently as the number of recommended projects increase. The difference between user-based collaborative filtering and SVD was statistically significant for each n using Mann-Whitney tests. (for $n = 3$, the default number of recommendations provided by the recommendation system in the online study, Mann-Whitney parameters were $U = 376424.0, p < 0.05$)

4.2 Online Study

The second part of the study was an online experiment. Users who logged on to SciStarter starting on December 2nd, 2019 were randomly assigned to one of 5 cohorts, each providing recommendations based on a different approach: (1) Item-based collaborative filtering, (2) User-based collaborative filtering, (3) Matrix factorization, (4) Most popular projects, (5) Promoted projects. Projects in this category were manually determined by SciStarter and often aligned with social initiatives and current events. Examples of such projects included FluNearYou (flunearyou.org), in which individuals report flu symptoms online, and was one of the promoted projects during the initial COVID-19 outbreak.

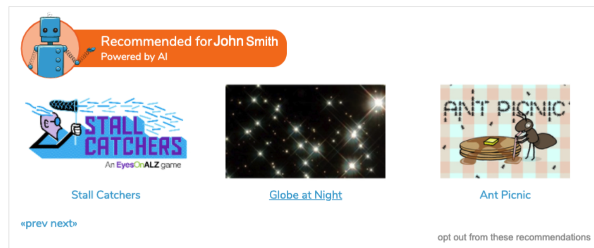


Figure 5: Screenshot of recommendation tool for user “John Smith”.

These projects are changed periodically by the SciStarter administrators.

The recommendation tool was active on SciStarter for 3 months. Users who logged on during that time were randomly divided into cohorts, each receiving a recommendation from a different algorithm. Each cohort had 42 or 43 users. The recommendations were embedded in the user's dashboard in decreasing order of relevance, in sets of three, from left to right. Users could scroll to reveal more projects in decreasing or increasing order of relevance. Figure 5 shows the top three recommended projects for a target user.

All registered users in SciStarter received notification via email about the study, stating that the “new SciStarter AI feature provides personalized recommended projects based on your activity and interests.” A link to a blog post containing more detailed explanations of recommendation algorithms and their role in the study was supplied.[4] Additionally, the study privacy policy was explained and users were given the option to opt out of receiving recommendations at any point in the experiment. In practice, none of the participants selected the opt out option at any point in time.

Figure 6 (top) shows the average hit rate (defined as the percentage of instances in which users accessed at least one project that was recommended to them) and Figure 6 (bottom) shows the average click trough rate (defined as the ratio of recommended projects that the users accessed). As shown by the Figure, both measures show a consistent trend,

[4] https://blog.scistarter.org/2019/09/smart-project-recommendations-on-scistarter/

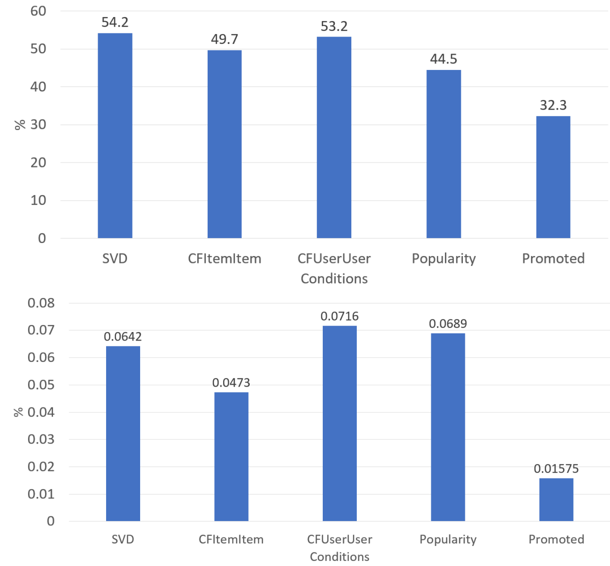


Figure 6: Hit rate (top) and Click through rate (bottom) and measures for online study

in which the user-based collaborative algorithms achieved the best performance, while the promoted projects method achieving worse performance. Despite the trend, the differences between conditions were not statistically significant in the $p < 0.05$ range. We attribute this to measuring clicks on recommended projects rather than actual contributions which is the most important aspect for citizen science.

To address this gap we defined two new measures that consider the contributions made by participants to projects, which constitutes the system utility as identified by Gunawardana and Shani (Gunawardana and Shani 2009). The measures include the average number of activities that users carried out in recommended projects (RecE), and the average number of activities that users carried out in non-recommended projects (NoRecE). Figure 7 compares the different algorithms according to these two measures. The results show that users assigned to the intelligent recommendation conditions performed significantly more activities in recommended projects than those assigned to the popularity and promoted projects conditions. Also, users in the SVD algorithm performed significantly less activities in non-recommended projects than the popularity and promoted projects conditions. These results were statistically significant according to Mann-Whitney tests ($p < 0.05$).

Lastly, we measure the average number of sessions for users in the different conditions, where sessions are defined as a continuous length of time in which the user is active in a project. Figure 8 shows the average number of sessions for users in the different cohorts, including the number of sessions for the historical data used to train the algorithms, in which no recommendations were provided. The results show that users receiving recommendations from the personalized algorithms performed more sessions than the number of sessions in historical data. These results are statistically significant in the $p < 0.05$ range using Mann-Whitney tests. Although there is a clear trend that users in the SVD condition achieved the highest number of sessions, these results were not significant in the $p < 0.05$ range.

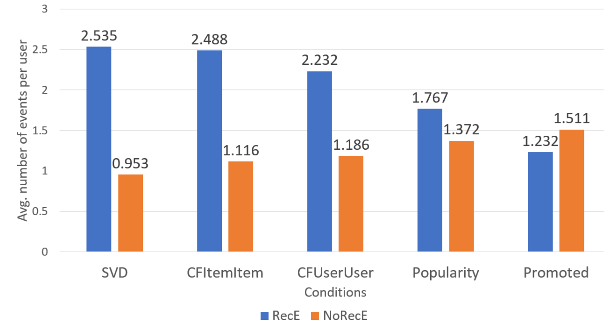


Figure 7: Average activities on recommended projects (RecE), and on non-recommended projects (NoRecE) for each condition

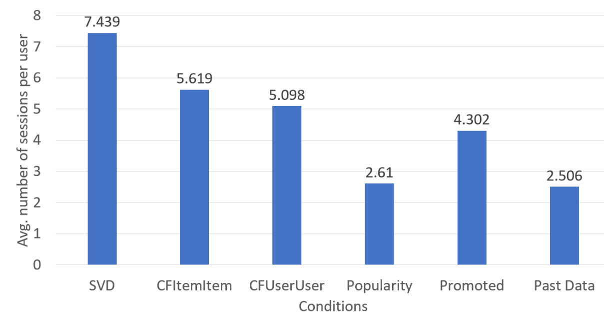
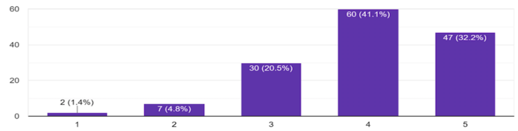


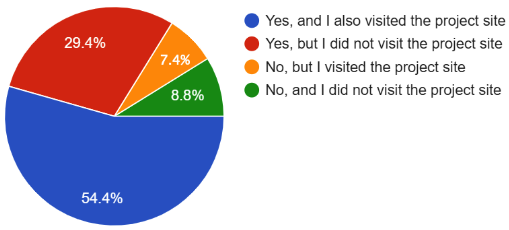
Figure 8: Average number of sessions for each condition

To explain SVD's good performance in the online study, we first note that SVD is considered as a leading algorithm in the domain of recommendation systems (Sadek 2012). Second, in our setting, SVD tended to generate recommendations that participants had not heard about before which seemed to resonate with many participants. As one participant remarked: "I am more interested in projects I didn't know existed before".

Lastly, we note the obstacles we encountered when carrying out the study. The first obstacle we encountered was the small number of relevant projects that could be recommended. Out of 3000 projects that SciStarter offers, we restricted ourselves to 153 affiliate projects which actively provide data of users' interactions. Another obstacle was that we were constrained to a subset of users who log on to the SciStarter platform and use it as a portal for contributing to the project, rather than accessing the project directly. Out of the 65,000 registered users of SciStarter, only a small percentage are logged in to both SciStarter and an affiliate project. As a result, we have relatively few users getting recommendations. In addition, some of SciStarter's projects are location-specific and can only be performed by users in the same physical location. (e.g collecting a water sample from a particular lake located in a particular city). Therefore, we kept track of users' location and restricted our recommen-



How satisfied are you with the recommendation tool?



Did you click on one (or more) of the recommended projects?

Figure 9: User satisfaction with recommendation tool (top) and User self report on clicking on recommended project (bottom)

dation system to be a location-based system, which recommends users with projects they are able to participate in.

4.3 User Study

In order to learn the users' opinion on the recommendations and their level of satisfaction, we conducted a survey with SciStarter's users. Our survey invitation was sent to all SciStarter community users. One hundred and thirty eight users have filled the survey, where each user was asked about the recommendations presented to them by the algorithm they were assigned to. The survey included questions about users' overall satisfaction with the recommendation tool as well as questions about their pattern of behavior before and after the recommendations.

The majority of users (73.3%; Responses 4 and 5 on a five-level Likert scale) were satisfied with the recommendation tool (Figure 9 top) and claimed that the recommendations matched their personal interests and goals. The majority of users (54%) reported they have clicked on the recommendations and visited the project's site, while only 8.8% of users did not click the recommendation nor visited the project site (Figure 9 bottom).

Interestingly, users who were not familiar with the recommended projects before, clicked more on the recommendations, as well as users who previously performed a contribution to a project. Users who did not click on the recommendations can be divided into 3 main themes: (1) Users who didn't have the time "right now" but planned to click the project in the future. (2) Users who felt that the recommendations were not suitable for their skills and materials: "Seemed out of my league", "I didn't have the materials to participate". This behaviour was also discussed in (Segal et al. 2015), and was named "classification anxiety". (3) Users who felt that the recommendations were not suitable for their interests: "No interest in stall catchers", "The photos and title didn't perfectly match what I am looking for".

The survey result provide evidence for the positive impact of using the recommendation systems in SciStarter. We note some additional comments by users: "I am very impressed by the new Artificial Intelligence feature from SciStarter! Your AI feature shows me example projects that I didn't know before exist", and "I like how personalized recommendations are made for citizen science users".

5 Conclusion and Future Work

This work reports on the use of recommendation algorithms to increase engagement of volunteers in citizen science, in which volunteers collaborate with researchers to perform scientific tasks. These recommendation algorithms were deployed in SciStarter, a portal with thousands of projects, and were evaluated in an online study involving hundreds of users who were informed about participating in a study involving AI based recommendation of new projects. We trained different recommendation algorithms using a combination of data including users' behavior in SciStarter as well as their contributions to specific projects. Our results show that using the new recommendation system led people to contribute to new projects that they had never tried before and led to increased participation in SciStarter projects when compared to groups that were recommended projects using non-personalized recommendation approaches, and compared to behavior before recommendations.

This project has transformed how SciStarter helps projects recruit and support participants and better respond to their needs. It was so successful in increasing engagement, that SciStarter has made the recommendation system a permanent feature of their site. This will help support deeper, sustained engagement to increase the collective intelligence capacity of projects and generate improved scientific, learning, and societal benefits.

An important avenue for future work is to provide users with explanations to the recommendations in order to increase the system's reliability and user's satisfaction with it. We also plan to extend the recommendation system to include content based algorithms, and test its performance as compared to the existing algorithms. We believe that integrating content in citizen science domain (such as the project description, its location, age group etc.) will enable us to capture more intrinsic characteristic of the projects, such as required effort or type of task.

Acknowledgments

Na'ama Dayan's work on this paper is equivalent to that of a co-author and is not recognized due to an oversight. This work has received funding from the NESTA Collective Intelligence Grants 1.0 and from the European Union's Horizon 2020 WeNet research and innovation program under grant agreement No 823783. Thanks to Daniel Arbuckle and Caroline Nickerson for their programming assistance and support.

References

Ahn, H. J. 2006. Utilizing popularity characteristics for product recommendation. *International Journal of Electronic Commerce* 11(2): 59–80.

Amatriain, X. 2013. Big & personal: data and models behind netflix recommendations. In *Proceedings of the 2nd international workshop on big data, streams and heterogeneous source Mining: Algorithms, systems, programming models and applications*, 1–6. ACM.

Bonney, R.; Cooper, C. B.; Dickinson, J.; Kelling, S.; Phillips, T.; Rosenberg, K. V.; and Shirk, J. 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience* 59(11): 977–984.

Bonney, R.; Shirk, J. L.; Phillips, T. B.; Wiggins, A.; Ballard, H. L.; Miller-Rushing, A. J.; and Parrish, J. K. 2014. Next steps for citizen science. *Science* 343(6178): 1436–1437.

Breese, J. S.; Heckerman, D.; and Kadie, C. 1998. Empirical Analysis of Predictive Algorithms for Collaborative Filtering. In *UAI*, 43–52. Morgan Kaufmann.

Brossard, D.; Lewenstein, B.; and Bonney, R. 2005. Scientific knowledge and attitude change: The impact of a citizen science project. *International Journal of Science Education* 27(9): 1099–1121.

Cohn, J. P. 2008. Citizen science: Can volunteers do real research? *BioScience* 58(3): 192–197.

Dwivedi, S.; and Roshni, V. K. 2017. Recommender system for big data in education. In *2017 5th National Conference on E-Learning & E-Learning Technologies (ELELTECH)*, 1–4. IEEE.

Fleder, D. M.; and Hosanagar, K. 2007. Recommender systems and their impact on sales diversity. In *Proceedings of the 8th ACM conference on Electronic commerce*, 192–199. ACM.

Freyne, J.; Jacovi, M.; Guy, I.; and Geyer, W. 2009. Increasing engagement through early recommender intervention. In *Proceedings of the third ACM conference on Recommender systems*, 85–92. ACM.

Funk, C.; Gottfried, J.; and Mitchell, A. 2017. Science news and information today. *Pew Research Center* .

Gunawardana, A.; and Shani, G. 2009. A survey of accuracy evaluation metrics of recommendation tasks. *Journal of Machine Learning Research* 10(12).

Haywood, B. K.; and Besley, J. C. 2014. Education, outreach, and inclusive engagement: Towards integrated indicators of successful program outcomes in participatory science. *Public understanding of science* 23(1): 92–106.

Irwin, A. 2018. No PhDs needed: how citizen science is transforming research. *Nature* 562: 480–482.

Itmazi, J.; and Gea, M. 2006. The recommendation systems: Types, domains and the ability usage in learning management system. In *Proceedings of the International Arab Conference on Information Technology (ACIT'2006), Yarmouk University, Jordan*.

Jonnalagedda, N.; Gauch, S.; Labille, K.; and Alfarhood, S. 2016. Incorporating popularity in a personalized news recommender system. *PeerJ Computer Science* 2: e63.

Kleinerman, A.; Rosenfeld, A.; Ricci, F.; and Kraus, S. 2020. Supporting users in finding successful matches in reciprocal recommender systems. *User Modeling and User-Adapted Interaction* 1–49.

Koren, Y.; and Bell, R. 2015. Advances in collaborative filtering. In *Recommender systems handbook*, 77–118. Springer.

Koren, Y.; Bell, R.; and Volinsky, C. 2009. Matrix factorization techniques for recommender systems. *Computer* 42(8): 30–37.

Kragh, G. 2016. The motivations of volunteers in citizen science. *environmental SCIENTIST* 25(2): 32–35.

Labarthe, H.; Bouchet, F.; Bachelet, R.; and Yacef, K. 2016. Does a Peer Recommender Foster Students' Engagement in MOOCs?. *International Educational Data Mining Society* .

Larson, L. R.; Cooper, C. B.; Futch, S.; Singh, D.; Shipley, N. J.; Dale, K.; LeBaron, G. S.; and Takekawa, J. Y. 2020. The diverse motivations of citizen scientists: Does conservation emphasis grow as volunteer participation progresses? *Biological Conservation* 242: 108428.

Laut, J.; Cappa, F.; Nov, O.; and Porfiri, M. 2017. Increasing citizen science contribution using a virtual peer. *Journal of the Association for Information Science and Technology* 68(3): 583–593.

Lin, C. H.; Kamar, E.; and Horvitz, E. 2014. Signals in the silence: Models of implicit feedback in a recommendation system for crowdsourcing. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*.

National Academies of Sciences, E.; Medicine; et al. 2018. *Learning through citizen science: enhancing opportunities by design*. National Academies Press.

Ning, X.; Desrosiers, C.; and Karypis, G. 2015. A comprehensive survey of neighborhood-based recommendation methods. In *Recommender systems handbook*, 37–76. Springer.

Nov, O.; Arazy, O.; and Anderson, D. 2014. Scientists@ Home: what drives the quantity and quality of online citizen science participation? *PloS one* 9(4).

Ponciano, L.; and Pereira, T. E. 2019. Characterising volunteers' task execution patterns across projects on multi-project citizen science platforms. In *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems*, 16. ACM.

Raddick, M. J.; Bracey, G.; Gay, P. L.; Lintott, C. J.; Murray, P.; Schawinski, K.; Szalay, A. S.; and Vandenberg, J. 2009. Galaxy zoo: Exploring the motivations of citizen science volunteers. *arXiv preprint arXiv:0909.2925* .

Rotman, D.; Preece, J.; Hammock, J.; Procita, K.; Hansen, D.; Parr, C.; Lewis, D.; and Jacobs, D. 2012. Dynamic changes in motivation in collaborative citizen-science projects. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*, 217–226.

Ryan, S.; Adamson, N.; Aktipis, A.; Andersen, L.; Austin, R.; Barnes, L.; Beasley, M.; Bedell, K.; Briggs, S.; Chapman, B.; et al. 2018. The role of citizen science in addressing grand challenges in food and agriculture research. *Proceedings of the Royal Society B* 285(1891): 20181977.

Sadek, R. A. 2012. SVD based image processing applications: state of the art, contributions and research challenges. *arXiv preprint arXiv:1211.7102* .

Sarwar, B.; Karypis, G.; Konstan, J.; and Riedl, J. 2002. Incremental singular value decomposition algorithms for highly scalable recommender systems. In *Fifth international conference on computer and information science*, volume 1. Citeseer.

Schafer, J. B.; Frankowski, D.; Herlocker, J.; and Sen, S. 2007. Collaborative filtering recommender systems. In *The adaptive web*, 291–324. Springer.

Segal, A.; Gal, K.; Kamar, E.; Horvitz, E.; Bowyer, A.; and Miller, G. 2016. Intervention strategies for increasing engagement in crowdsourcing: platform, predictions, and experiments. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, 3861–3867.

Segal, A.; Gal, K.; Kamar, E.; Horvitz, E.; and Miller, G. 2018. Optimizing Interventions via Offline Policy Evaluation: Studies in Citizen Science. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Segal, A.; Gal, Y.; Simpson, R. J.; Victoria Homsy, V.; Hartswood, M.; Page, K. R.; and Jirotka, M. 2015. Improving productivity in citizen science through controlled intervention. In *Proceedings of the 24th International Conference on World Wide Web*, 331–337.

Silvertown, J. 2009. A new dawn for citizen science. *Trends in ecology & evolution* 24(9): 467–471.

Sorensen, A. E.; Jordan, R. C.; LaDeau, S. L.; Biehler, D.; Wilson, S.; Pitas, J.-H.; and Leisnham, P. T. 2019. Reflecting on efforts to design an inclusive citizen science project in West Baltimore .

Vitone, T.; Stofer, K.; Steininger, M. S.; Hulcr, J.; Dunn, R.; and Lucky, A. 2016. School of ants goes to college: integrating citizen science into the general education classroom increases engagement with science. *Journal of Science Communication* 15(1): A03.

Wang, X.; Guo, Y.; and Xu, C. 2015. Recommendation algorithms for optimizing hit rate, user satisfaction and website revenue. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.

Wu, Q.; Wang, H.; Hong, L.; and Shi, Y. 2017. Returning is believing: Optimizing long-term user engagement in recommender systems. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 1927–1936. ACM.