Discovering Elite Users in Question and Answering Communities

Cheng-Yue, Royce
rchengyue@cs.stanford.edu

Hsu, Richard
rhsu@cs.stanford.edu

Stevens, Nicholas
nstevens@cs.stanford.edu

Group 30

Abstract

Question and Answering (Q&A) communities depend on a set of users who have mastery of the topics being discussed and also actively respond to questions; we will refer to these users as "elite users". Identifying these elite users allows general users to identify credible sources. It can also help in allowing community designers to direct unanswered questions to these elite users for a higher reliability of a response and credibility. The goal of our work is to investigate the activity and behavior of users in a particular Q&A community and discover whether or not it is possible to predict early on in a user's career whether or not they will be an elite user in the future.

Keywords.
Question-answering, reputation, value prediction.

General Terms.
Experimentation, human factor, measurement.

1 Introduction

Online education has been growing rapidly over the last few years. Many of them use some aspect of a question and answering forums where anyone can ask a question and anyone in the community can respond. Thus these communities provide a community-driven knowledge portal. However, much of this process relies on the fact that a subset of users in the community have some mastery of the topics; otherwise, the questions would rarely be answered. We refer to these users as "elite users" and are those who contribute actively and are reliable sources. These users form the most important aspect of driving the success of these questions and answering forums. Many websites reward such users with some form of reputation that can be employed through voting or rewards for responding, thereby allowing users to identify those with mastery of the topics. Unfortunately, gaining such status takes time and hides the fact of whether or not the user has mastery and can actively participate.

We believe that identifying elite users will be beneficial by allowing community designers the ability to direct unanswered questions to these elite users for a higher reliability of a response and credibility especially considering the multitude of growth in these communities. It can also help by allowing community designers to grant special privileges to users earlier in his or her career, thus allowing the user to better utilize his or her expertise in building the success of the community. To better assist in the identification of these elite users, we explore attributes in a particular community called Stack Overflow[1] to see if attributes or activities associated with users can help distinguish elite users early on in their careers.

2 Related Works

Question and answering communities have been studied before [1, 5] focusing on the questions and answers and gaining insight about their properties. One particular study looked into predicting long-term value of a question in the Stack Overflow community [1]. We believe that in conjunction with such work being able to identify elite users can provide Q&A communities the ability to bring more attention to these entities sooner. This would provide even more value to both because more exposure to elite users could be given to these questions, thus improving reliability and long-term value to these questions. This study also provides a parallel evaluation for long-term value of an entity in these Q&A communities. A. Anderson, et al. primarily found a high indication of utilizing intrinsic properties of the questions to predict long-term value in questions, which we believe has a similar parallel for users [1]. We believe that elite users have some intrinsic properties in common that help distinguish them from regular users. Although identification of long-term value in users may require longer time frames since creation, there are many properties of users that could act as features to help predict their likelihood to become elite users in a year.

Besides evaluation of long-term value in questions in the communities, the identification of the "expert" set of users in these Q&A communities has been explored in the research community as well [3, 4]. Our work is different in that not only do we want to identify experts, but we also want to identify those expert users who are highly active and can make the biggest impact. Additionally,

[1]http://www.stackoverflow.com

we want to discover value from their behaviors that not only distinguish them in the community but also allow us to predict and identify elite users early on in their careers. Such work is similar to A. Pal, et al. as they argue that evolutionary data of users can be more effective at expert identification than the models that ignore evolution [2]. We want to complement this work with further experimentation on other features expressed by users on Stack Overflow and further discover attributes that separate elite users from regular users.

By continuing the work in the field of gaining value of entities in Q&A communities, we want to show that it is possible to distinguish elite users early on thereby disassociating the temporal aspect of gaining reputation in order to gain credibility in the community. This is important because a past study showed that users who contributed a lot had greater influence than new users [7]. This intuitively makes sense because those already contributing continue to maintain their long-lasting value; however, new users who may be elite users are greatly undermined. We believe that finding a model to accurately predict a user's long-term value will help Q&A communities provide more attention to and encourage a higher quality of responses to all questions.

3 Data Set Description

Stack Overflow employs a targeted model in both domain and question type encouraged on the site. As opposed to the myriad of popular general Q&A sites on the web today (Yahoo! Answers, Quora), Stack Overflow advertises itself as a programming only Q&A site. Furthermore, all questions posed on the site are meant to be looking for a single, 'best' answer. Subjective questions that have no hope for such a definitive response are usually weeded out.

This focus on both domain and question type is only successful thanks to the users. First, the subset of the population equipped to answer programming questions is small, and the value question askers acquire is largely based on having the right expert answer their question. Second, many of the top users serve as de facto moderators, removing questions that do not fit the mission of Stack Overflow and merging duplicates. To ensure this power is not abused, users are incrementally granted increased abilities based on their own reputation.

The possible and quantifiable actions that occur on Stack Overflow extend beyond simply asking and answering. Users can both comment on questions and answers, all forms of communication can be voted on, and users can favorite questions. Finally, any of the possible answers can be designated as the 'accepted answer' by the original asker. All of these attributes are used to present each question on the site: after displaying the question at the top, the answers and comments are stack ranked based on the number of up minus down votes, with the 'accepted answer' always presented first.

For our project, we are using a complete trace of the

	Total	**Other Statistics**
Users	1,295,620	55.25% asked a question 38.60% answered a question
Questions	3,453,742	62.21% accepted answers
Answers	6,866,609	32.15% accepted
Votes	21,460,580	93.00% positive
Favorites	1,992,831	on 788,991 questions

Table 1: Stack Overflow's Statistics

Action	**Author**	**Action Taker**
Answer is upvoted	`+10`	`+0`
Answer is downvoted	`-2`	`-1`
Answer is accepted	`+15`	`+2`
Question is upvoted	`+5`	`+0`
Question is downvoted	`-1`	`-2`
Answer wins bounty	`+Bounty`	`-Bounty`
Answer marked as spam	`-100`	`+0`
Accepted suggested edit	`+2`	`+0`

Table 2: Stack Overflow's Reputation System

site that extends from the sites inception on July 31, 2008, to August 7, 2012. Some basic statistics of this dataset can be seen in Table 1. We are using MySQL to query the data and the Python library `sklearn` to help with our prediction models. Lastly since our predictions will be relating to evaluation of users being elite or not, we rely on the reputation system that Stack Overflow has developed based on the actions taken by each user (Table 2). We use this evaluation because it is the social evaluation of whether or not a user is of any value to the community and given enough time should be a good indicator of the value of the user to the community.

4 Analysis of Dataset

We began by looking at the dataset to learn more about the community and verify some of our possible intuition about the network. We looked at some various properties but found the dynamic between questions, answers, and reputation to be rathering interesting. We will further develop the relationship between them in this section.

4.1 Reputation Before and After

We began by looking at cohorts of users who joined Stack Overflow in specific time frames. After exploring the data, we found that the correlation coefficient was 0.797 between the reputations of cohort of users who joined 3 months before September 7, 2011 versus their reputations roughly a year later on August 7, 2012. Given the plot of the users' reputations before and after in Figure 1 can look at the top right quadrant and can estimate that the effective loss is smaller than in the bottom right quadrant. This leads us to the intuition that getting higher reputation scores early on is a good predictor of separating out elite users later on.

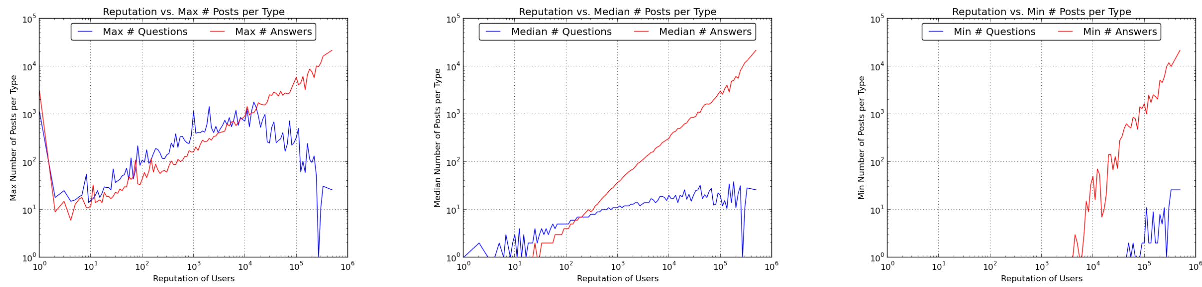

Figure 2: Reputation vs. Number of Posts on Log-Log scale. Number of Answers Given (red) and Number of Questions Asked (blue). Maximum number of posts given a reputation (left). Median number of posts given a reputation (middle). Minimum number of posts given a reputation (right).

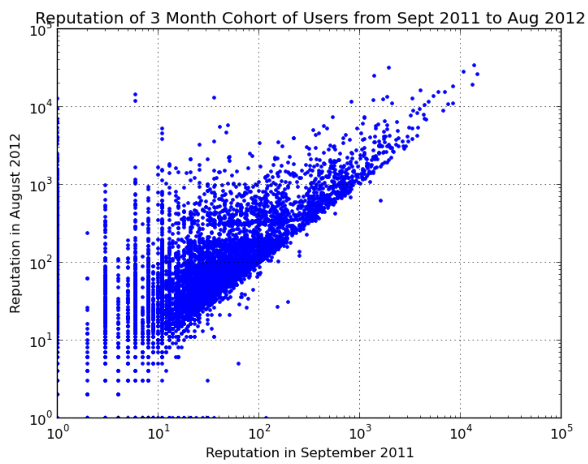

Figure 1: Reputation of a cohort at 3 months compared to a year later.

4.2 Reputation's Social Evaluation

The reputation of Stack Overflow heavily favors those that answer questions. Based on the reputation system, authors of answers stand to gain more reputation than authors of questions (see Table 2). This shows the importance of answers because Question and Answering communities rely heavily on users who have mastered the topics to respond and answer questions. In Figure 2, we have plotted user reputations as of August 7, 2012 and the max, median and minimum number of questions and answers. For the maximum number of post types, the values are very similar for the general population of users, and the trend is that there is a similar number of answers and questions for all reputations below 10,000. After 10,000 reputation, we see a huge separation between the answer and question counts, signifying that top tiered users generally have more answers than questions. This coincides with our intuition of the built in reputation system, which uses reputation to evaluate the "eliteness" of a user over time. Similarly, this holds for the rest of the population as can be seen by the median and minimum graphs in Figure 2. It is interesting to note that after 4,000 reputation all users have at least one answer made, which dramatically increases as the reputation increases. We want to explore the idea of elite users, who are those that can largely contribute to the community through activity and trustworthiness. Since this intuition is consistent with the above reputation analysis, we will utilize reputation to help classify and validate our prediction model. It is also good to note though that, since there is a huge dynamic between answers, questions, and reputation, we will be exploring features related to these entities in order to see if we can discover value in the community.

4.3 Power Law of Posts

The dynamics between questions and answers is rather intruiging. We present a distribution of the frequency of users with a specific number of questions and answers in the community in Figure 3. The calculated power-law exponent α and x_{min} were calculated and shown in Table 3. In general, we see that both distributions follow the power-law extremely well. However, the frequency of users for the number of answers has an exponent of around 3, whereas, for questions, this only lies around 1.8. From the general statistics in Table 1, we see that the number of questions is half that of the number of answers that exist in the community. This shows that in a Q&A community, questions engage the users to participate, and then a higher proportion of users attempt to answer these questions far beyond a 1 to 1 mapping. Given this, we see that the Q&A community has a rather large disproportion to answerers. Naturally, this would suggest that users collaborate to try to give the best responses possible, thereby allowing for a greater chance of answering the question well. However, because each question can only have one accepted answer, we must take into account that a user who answers many questions is not necessarily elite. In our original definition of elite users, we would want to seek users who are not only active, but also reliable. From this stand point, we want users with high acceptance rates for their answers. Because looking at the frequency of a user's answering

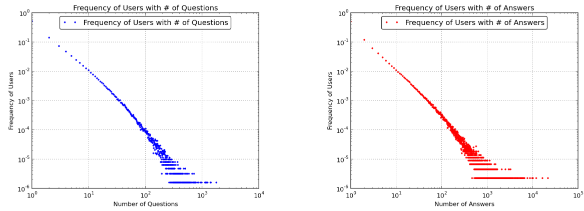

Figure 3: Power Law of the Number of Questions and Number of Posts

	α	x_{min}
Questions	1.80	1
Answers	3.01	1353

Table 3: Power Law α coefficient and x_{min}

activity is not a necessary condition to describe an elite user, we will also look at more inherent properties of answers, including accepted answers that a user may obtain in the beginning of their career.

4.4 Average Views

We want to find the top tier of users who are active and credible. As a result, we explored the average view counts over a user's set of answers and accepted answers. In figure 4, we notice some interesting information. For low reputation users, we see a higher average view count of answers given than that of accepted answers. This occurs for the maximum, median, and minimum; however, as we get higher in reputation at 10,000 reputation, we see that these values begin to converge. This is similar to the threshold where the number of answers given and the number of questions given diverge for maximum number of posts in Figure 2. Consequently, an elite user's accepted answers and given answers are more than likely correlated as compared to a regular user who may have more answers but very few accepted ones. Finally, we look at the max average views for the maximum of these average views (far left graph in Figure 4). This shows that, even with high number of page views, a user is not necessarily going to gain large reputation. Thus, the number of answers and popularity of a post do not coincide with developing intuitions of necessarily having high contributions to the community. Specifically, this does not happen until a user's accepted answers gain more popularity even then we notice that the accepted answer's max average views is staggeringly lower than that of answers given. We also see that the average page views decrease as the reputation of a user increases. As a result, even though a post is popular, it does not necessarily portray the same social measures as being credible or worthy.

4.5 Average Accepted Answers

Our final discussion of our analysis looks at the accepted answer rate of users versus their reputations. Overall, maximum and median do not tell us much except that

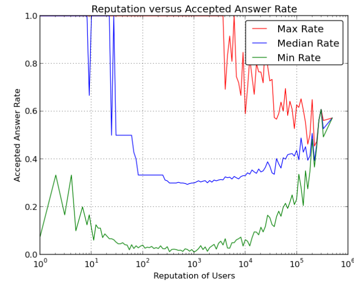

Figure 5: Reputation versus Accepted Answers

there are possibly many users who have a few questions that have been accepted. Reasonably, a new user can have his or her answer accepted, and, as the answer gains reputation, the user gains reputation. The most intruiging part of Figure 5 is the minimum accepted answer rate. Intuitively, a user joining to ask a question or answer probably correlates to him or her answering a question well. In this way, new users can have a decent answer rate if they can achieve few accepted answers. However, we see that, as reputation increases, this minimum average accepted answers decreases to almost 0% and then begins to rise again at around 10,000 reputation (the transition point similar to those seen in previous analysis sections). At this point, we see that the accepted answer rate begins to increase to a point where it converges at almost 60% acceptance rate. The increase is intruiging because it indicates that top tiered users who have lots of reputation tend to have higher average acceptance rates for their answers. There are still many factors for a user to gain reputations such as upvotes and bounty; however, we consistently see from analyzing the answers in relationship to reputation that the acceptance rate of answers being higher leads to a more likely candidate of the user being elite. Finally, to sum up these sections of data analysis, we see that the characteristics of answers help describe whether or not a user will become elite. We aim to not only look at the user's characteristics, but also the answers users give and the types of questions they seem to target. By doing this, we hope to see the significance of these characteristics of a user's authorship to shine light on their value in the community.

5 Methods

To predict whether or not users will be elite the following year, we decided to use two different classification algorithms: logistic regression and SVM with an RBF kernel (RBF-SVM). In general, logistic regression is the standard method for numerous classification tasks and tries to optimally find a linear decision boundary among the data. In the case that our data is not linearly separable, we also decided to use an SVM to generate more complex decision boundaries in the hopes of identifying possible nonlinear structures and achieving a higher performance. A common choice for this purpose is the

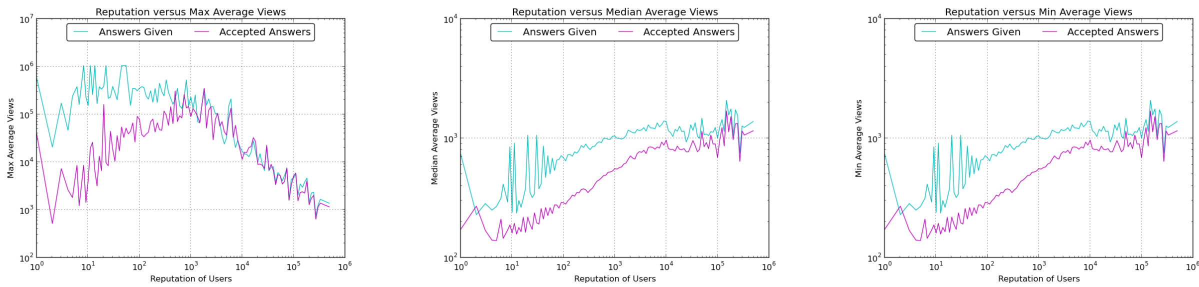

Figure 4: Reputation vs. Average Views on Answers or Accepted Answers. Answers Given(cyan) and Answers Accepted (magenta). Maximum average views (left). Median average views (middle). Minimum average views (right).

RBF-SVM.

In addition to applying these algorithms with a full set of features, we also decided to have baselines using logistic regression and RBF-SVM on the following three features: reputation, upvotes, and downvotes. From our data anaylsis on early reputation values (Figure 1), we found that there is a high correlation between the early reputation of a user and the reputation of a user the following year. Because Stack Overflow uses upvotes and downvotes as main sources of reputation and direct evaluations of a user's questions and answers, we decided to add the number of upvotes and the number of downvotes to the baseline features. We used these baselines to determine whether or not our chosen models and predictions provide meaningful insights into the development of elite users.

To evaluate our performance, we used area under the ROC curve (AUC) and F_1-scores. Although classification accuracy is a standard metric in determining the success of the classification algorithms, this measurement usually provides misleading results for highly skewed data. A common metric that addresses this issue is AUC, which measures the relationship between true positive rates and false positive rates. F_1-scores can also be used for this purpose:

$$2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

Because the number of elite users is considerably less than the size of the general population, AUC and F_1-scores will provide better metrics in determining the quality of our classification algorithms.

In order to train our models, we calculated and used the following normalized features, which we've chosen from what we thought best to help characterize a user's behavior and activity:

- **Characteristic of the User:** Number of questions asked, number of answers given, number of accepted answers, number of comments, number of personal favorites, amount of bounty, number of badges, number of user profile views, number of upvotes, number of downvotes, reputation.
- **Characteristics of All of a User's Questions:** Accumulated score, accumulated views, accumulated favorites, accumulated number of answers, accumulated length of questions, accumulated length of title of questions.
- **Characteristics of All of a User's Answers:** Accumulated score, accumulated number of comments, accumulated length of answers, accumulated score of the corresponding questions, accumulated views of the corresponding questions, accumulated favorites of the corresponding questions, accumulated number of other answers to corresponding questions, accumulated length of the corresponding questions, accumulated length of the titles of the corresponding questions.

Our classification task is to predict whether or not users will become elite users based on what they have accomplished in a time frame since they joined Stack Overflow. To simplify our model, we also defined an elite user to be in the top 10% percentile of a cohort based on reputation. While this is an arbitrary choice, we viewed any larger percentiles (top 25%, etc) would lower the difficulty in achieving high evaluation scores in this prediction task and would make it too easy for a user to be classified as elite. Also we believe only the top performing users should really be characterized as those who contribute enough that attention should be given to them.

For our predictions, we decided to sample users at time frames of one week, one month, and three months since their inceptions and look at their features after this time frame. We use a user's features as examples and whether or not their reputation is in the top 10% of their cohorts the following year as labels. Specifically, we trained our classification algorithms on September 2010 and September 2011 data, and tested our classification algorithms on September 2011 and August 2012 data.

After running our classification algorithms on our baseline and full feature sets, we decided to use feature selection to determine a set of essential features and focus our analysis on these imperative features.

	Baseline		Full Features	
	Logistic	**SVM**	**Logistic**	**SVM**
1 Week	0.70	0.56	0.68	0.63
1 Month	0.80	0.68	0.81	0.80
3 Months	0.88	0.84	0.89	0.89

Table 4: Area under the ROC curve for logistic regression and SVM for both baseline and full feature set.

	Baseline		Full Features	
	Logistic	**SVM**	**Logistic**	**SVM**
1 Week	0.16	0.17	0.28	0.30
1 Month	0.42	0.39	0.47	0.41
3 Months	0.62	0.57	0.63	0.59

Table 5: F-1 scores for logistic regression and SVM for both baseline and full feature set.

6 Results

Table 4 and 5 provide the results found through logistic regression and RBF-SVM on our baseline and full feature sets. From these results, we find that our prediction model for the one week cohort is less accurate than that for the one month and three month cohorts. Intuitively, this result makes sense as it shows that the longer a user spends in the community, the better prediction we can make in their future community contribution level. In terms of the specific algorithm performances, we find that logistic regression provided similar or better results than RBF-SVM across all time frames. However, after users have three months in the community, RBF-SVM performs almost as well as logistic regression. We also notice that, although logistic regression performed similarly for the full feature set and the baseline feature set in terms of AUC scores, logistic regression performed much better with the full feature set than with the baseline logistic regression in terms of F_1 scores. On the other hand, RBF-SVM performed much better with the full feature set than the baseline feature set on all time frames and evaluation metrics.

After running feature selection, we find that there are nine essential features for each time frame. Performing logistic regression and RBF-SVM on these nine essential features shows about a 0.01 decrease on AUC scores and 0.03 decrease on F_1 scores from using the full feature set on all time frames. However, if we remove a feature from this set of essential features, our evaluation metrics take a considerable performance hit of around 0.05 decrease on AUC scores and 0.1 decrease on F_1 scores.

The relative importance of the essential features with respect to each time frame is shown in Table 6, you can also see the top 9 chosen essential features for each time frame as well. By analyzing the weights generated from our nine essential features, we find numerous intriguing points. First, we notice that there are four consistent features across all the time frames: accumulated number of comments for user's answers, accumulated question favorite count, number of accepted answers, and reputation.

In the early stages of a user, reputation is not a strong predictor of whether or not the user will eventually become elite. This was what we believed from our previous analysis of the data and it still holds true. We know intuitively that reputation requires time and therefore, over time, we expected reputation to become a stronger feature for helping predict elite users. Thus, once we allow users a month to contribute to the Stack Overflow community, we began to see that reputation became a stronger predictor of whether or not they will be an elite user the following year. This analysis follows closely to Figure 2, in which we found that we begin to see divergences once a user has begun accumulating more reputation. Another similar feature is that of number of accepted answers. Elite users are very important to a community, as they are the ones that actively contribute as well as help provide support for being credible resources. Thus with our results we see that the number of accepted answers a user has as a feature begins to become a stronger predictor as we increase the time frame. Both reputation and number of accepted answers are both intuitive features that we hope would shine in the prediction model.

The accumulated number of comments for a specific user's answers and the accumulated question favorite count are two other features that contribute to the prediction of an elite user. We notice that these two features are related in the sense that they involve community engagement. Specifically, if a user answers an important or thought provoking question in the community, it is likely that the user's answers will generate discussion among other users, causing a large number of comments for the user's answers. Similarly, we find that another common feature of elite users is the importance of their questions, which can be estimated by the number of favorites on their questions. By asking interesting questions, elite users welcome other users to exchange ideas and participate in these discussions. As a result of this analysis, we find that elite users not only contribute meaningful answers, but they also involve other members of the community in their questions and answers. This can be further seen in Figure 2 where, for the highest of reputation users, the minimum number of questions elite users ask is nonzero. This signifies that even the top tiered users are not solely authors of answers, but they also ask questions.

Apart from these four common features in all the time frames, there are four other features that play an important part in determining whether or not a user will eventually become elite: accumulated score of the corresponding questions, number of comments, accumulated length of answers, and the number of downvotes. Interestingly, we find that the accumulated scores of questions a user answers has a negative weight. This result means

Feature	Coefficients 1 Week	1 Month	3 Month
Accumulated number of comments for user's answers	-0.551	+0.857	+0.572
Accumulated question favorite count	+0.579	+0.046	+0.052
Number of accepted answers	+0.129	+0.313	+1.186
Reputation	+0.665	+4.412	+8.845
Accumulated score of the corresponding questions	-0.261	–	-0.035
Number of comments	+0.053	–	+0.392
Accumulated length of answers	–	+0.252	+0.156
Number of downvotes	–	-0.239	-0.977
Accumulated number of answers for a user's questions	+0.230	–	–
Accumulated score of the answers	+4.450	–	–
Number of answers given	+0.268	–	–
Accumulated favorites of the corresponding questions	–	-0.263	–
Accumulated number of other answers to corresponding questions	–	+0.343	–
Accumulated question views	–	+0.020	–
Number of personal favorites	–	–	+0.138

Table 6: Top coefficients using logistic regression with nine essential features on the 1 week, 1 month, and 3 month cohorts.

that elite users do not answer only questions with large number of upvotes, but they answer a variety of questions, including questions that may not be as popular or heavily upvoted. The number of comments and the accumulated length of the user's answers are two features that relate back to the idea that elite users not only provide thoughtful insights to questions, but they also generate community engagement among their peers. The number of downvotes is also expected to have a strong negative weight; however, this is intuitively a high-precision, low-recall feature. Specifically, if a user is downvoted many times, the user is most likely not an elite user; however, having low number of downvotes does not necessarily indicate that the user will eventually become an elite user.

7 Conclusion

Our results show that we can predict whether or not users will become elite users in the future from their initial behaviors. Although our prediction model does not improve a huge amount over our baseline analysis, we believe we have gained a great amount of insight on the interaction between the current evaluation system of reputation and what we would want to see reputation indicate. We can see that, over time, reputation is a great indication of whether or not a user is or will be elite; however, we also know that other interactions and behaviors provide clues on whether or not the user will eventually contribute and provide value to the community.

Based on our AUC and F-1 scores, we see that the one week time frame is not enough to accurately predict whether or not users will eventually be elite. This is due to the fact that user activity in the beginning is most likely the same. A user is more likely to join the community to do one of two things: ask a question, or answer a question. Due to this dynamic and our previous analysis, we see that, in the beginning, question and answer counts do not have a large divergence from each other. It is likely that users are generally active in their first week, and, over time, either become inactive or maintain their activity to eventually become elite users.

We initially approached this as an implementation exercise, but ended up learning far more. First, examining the data and deciding on an approach proved to be very important. We were lucky to have a large dataset to work with, but its size made initial data analysis before exploring possible features paramount. Next, we learned a lot about the varying methods of both producing the necessary features and running our algorithms efficiently. When dealing with such large cohorts of users who each take thousands of actions, the efficiency of our queries was very important in enabling us to examine the large cohorts and our large list of features. Finally, we also learned about the importance of a good evaluation metric. We initially thought that classification accuracy would be sufficient, but we soon learned that the nature of our prediction task led to high classification accuracy regardless of our features. In exploring AUC and F_1 scores and their advantages, we ended up being exposed to a much more telling indicator of accuracy. Through our analysis of the data set and user features, we also gained deep insights into early indicators of whether a user will eventually become an elite user.

7.1 Possible Future Work

Next steps for this task involve possible new features and new algorithms. In particular, we could look into temporal information, such as trying to recreate the user's history. For instance, the time to the first answer of a new question or the time it takes for the highest scoring answer to arrive could be important for a user's behavior. To better understand the meaning of an elite user, we could look at other metrics or methods of determining whether or not a user will become elite. Also, we could explore other algorithms or models. For instance, we thought about looking at this data with a modification of the PageRank algorithm, possibly defining the nodes as users and the edges as high quality interactions between them, and compute a PageRank score for each user. We could also look at betweenness centrality to find the most important and connected users in the question and answer network. To generate discussion on the overall structures of the network, we can form clusters and examine the elite user communities. In particular, these methods can either be represented as new models to discover elite users or be used as features in similar prediction tasks to the task in this paper.

Overall, nurturing the most productive users is paramount to a site like Stack Overflow, where users run, moderate, and produce most of the content. With the large amount of data Stack Overflow collects, we found that simply using some select features could help the site accurately identify the best users within a month of joining the site. In consequence, beyond the normal social evaluators such as reputation, upvotes, and downvotes, we see that there is definitely predicative power in the features of a user.

8 References

1. A. Anderson, D. Huttenlocher, J. Kleinberg, J. Leskovec. Discovering Value from Community Activity on Focused Question Answering Sites: A Case Study of Stack Overflow . *In Proc. KDD*, 2012.
2. A. Pal, S. Change, J. A. Konstan. Evolution of Experts in Question Answering Communities. *AAAI*, 2012.
3. X. Liu, W. B. Croft, M. Koll. Finding Experts in Community-Based Question-Answering Services. *CIKM*, 2005.
4. P. Jurczyk, E. Agichtein. Discovering Authorities in Question Answer Communities by Using Link Analysis. *CIKM*, 2007.
5. C. Shah, J. Pomerantz. Evaluating and predicting answer quality in community QA. *SIGIR*, 2010.
6. L. A. Adamic, J. Zhang, E. Bakshy, and M. S. Ackerman. Knowledge sharing and Yahoo Answers: everyone knows something. *WWW*, 2008.
7. B. Keegan, and D. Gergle. Egalitarians at the gate: One-sided gatekeeping practices in social media. *In Proc. CSCW*, 2010, ACM Press, 131-134.