Language Independent Analysis and Classification of Discussion Threads in Coursera MOOC Forums

Lorenzo A. Rossi
Viterbi School of Engineering
University of Southern California
lorenzo.rossi@gmail.com

Omprakash Gnawali
Department of Computer Science
University of Houston
gnawali@cs.uh.edu

Abstract

*In this work, we analyze the discussion threads from the forums of 60 Massive Open Online Courses (MOOCs) offered by Coursera and taught in 4 different languages. The types of interactions in such threads vary: there are discussions on close ended problems (*e.g. *solutions to assignments), open ended topics, course logistics, or just small talk among fellow students. We first study the evolution of the forum activities with respect to the normalized course duration. Then we investigate several language independent features to classify the discussion threads based on the types of the interactions among the users. We use default Coursera subforum categories (Study Groups, Assignments, Lectures, ...) to define the classes of interest and so the labels. We extract features related to structure, popularity, temporal dynamics of threads and diversity of the ids of the users. Text related features, word count aside, are avoided to apply the methods across discussion threads written in different languages and with various technical terminologies. Experiments show a classification performance with ROC AUC between* 0.58 *and* 0.89*, depending on the subforum class considered and with possibly noisy labels.*

1. Introduction

The number of Massive Open Online Courses (MOOCs) has grown rapidly in the past couple of years. For instance, according to the MOOC aggregator class-central.com, during the month of April 2014 there were 287 MOOCs in progress. Coursera is the leading provider, offering almost half of them, followed by Udacity and EdX. There is an increasing number of additional providers also outside the USA. The very large ratio of students to instructors and teaching assistants involved makes the course forums practically the only place for (peer) interaction and question answering. The forum plays the role of the course's crowd-sourced teaching assistant. In large MOOCs, potentially with hundreds of new discussion threads per day, the course staff (instructors and teaching assistants) may be unable to adequately track the forums to find all the issues that need a resolution.

A common aspect of the MOOC forums is the variety of interactions in the threads: they may be dedicated to homework resolution, open ended discussions about topics of the courses, logistics, social/small talk (*e.g.* meetup groups). Hence, only a subset of the discussions may need the participation of the course staff. Another aspect specific to large MOOC providers like Coursera is that the spectrum of the subjects taught is quite broad (Engineering, Humanities, Business, *etc.*) Besides, an increasing number of courses are taught in languages different from English (*e.g.* Chinese, French, Spanish, ...). Thus automatic tools to increase intelligence of the forums need to deal with a broad variety of languages and technical terminologies. For example, the threads to classify could be from a physics course taught in Spanish, while the only data available for training could be from computer science courses taught in English. In this case, typical natural language processing (NLP) tools such as n-grams would not be effective.

We analyze the discussion threads from 60 MOOCs offered by Coursera and taught in 4 different languages. We study the evolution of the discussion threads, looking for example in which categories of subforums the posts are made throughout the courses. Then we define a set of features for the supervised classification of the discussion threads. Such features are related to structure and temporal dynamics of the threads, diversity of user IDs, votes, but not to text/language information except for the word count. Possible applications of this framework are to get insightful analytics on the forums (*e.g.* it could be combined with topic classification find out in which types of threads certain topics are discussed), or develop a component to a real time scalable system to detect discussion threads that need a resolution from the course staff.

The automatic classification of the types of interactions

among forum participants have been studied for instance with purposes such as finding unresolved discussions [12], [2], in online courses. Ours is among the first works, along with [3] and [6], to provide a relatively large scale analysis of MOOCs forums (*i.e.* with datasets ranging from 40 to 80 course forums). This work gives the following insights and contributions:

- we provide the first large public dataset of anonymized Coursera MOOC forums
- we show that the number of active users (and posts) in a Coursera MOOC forum decays exponentially for about the first 6/10 of the course duration (independently from the absolute course duration)
- we define and test (novel) language independent features for the supervised classification of the threads

The rest of this paper is organized as follows. The related literature is reviewed in Sec. 2. Our dataset is described in Sec. 3. Sec. 4 provides insights on some aspects of the Coursera forums. The features used for the supervised classification of the discussion threads and our approach to label the data are presented respectively in Secc. 5 and 6. Experimental results are given in Sec. 7. Conclusions are drawn in Sec. 8.

2. Related Literature

2.1. Works on online discussions

Joty *et al* study topic classification for online asynchronous discussions in [7], *e.g*, where they propose two unsupervised methods for topic segmentation. Ravi *et al* attempted to detect unresolved discussions in forums for online (non-massive) courses [12]. Lin *et al* in [10] use text mining to classify the genres of asynchronous discussions in online courses. Some of the categories of interest are: announcement, question, clarification, assertion and conflict. Baldwin *et al* in [2] focus on the automatic classification of threads in Linux forums. Some of their goals are assessing whether a thread is focusing on a specific problem, or on a more open ended discussion and then whether such problem has been resolved or not. They mostly use text related features.

The evolution of tree structure and underlying social networks of asynchronous discussion threads across various forums of different online communities (Wikipedia Talk pages, Slashdot forums) have been studied in [5] and [9]. In [5], the authors propose a generative model to analyze the evolution of the structure of discussion threads in different communities.

2.2. Works on MOOCs

MOOCs and in particular the interaction on their discussion forums have been the subject of an increasing number of recent studies. Brinton, Chiang *et al* in [3] study the decline rates of posts and participants in the forums of 73 Coursera MOOCs (171,197 threads, 831,576 posts and comments). The authors also propose a supervised classification method for filtering out small talk type of discussions from the rest of the threads. Huang *et al* in [6] study the behavior of very prolific posters in Coursera MOOC forums and evaluate the quality of their contributions to the discussions. They analyze 44 Coursera MOOCs (70,419 threads, 325,071 posts and comments). Anderson *et al* in [1] study the behavioral patterns of several classes of users for 6 Coursera MOOCs offered by Stanford using data from the forums as well from other sources. They consider 5 categories of users: bystanders, viewers, collectors, all-rounders, solvers. In a work with a similar focus [8], Kizilcec *et al* study the engagement patterns of 4 different subpopulations of users (completing, auditing, sampling and disengaging ones) for three computer science MOOCs offered by Stanford. Wen *et al* perform sentiment analysis on the discussion forums of 3 Coursera courses [14] to infer which students are more likely to drop out.

3. The Dataset

Our dataset consists in the discussion threads from the forums of 60 Coursera MOOCs (99,624 threads, 739,093 posts and comments), downloaded between August 2013 and April 2014. We share an anonymized version of the dataset, with a complete list of the courses, through a GitHub repository (see [13] for link). The discussion forums on Coursera are usually active also after the end of the courses: in our analyses and experiments, we consider threads posted within two weeks from the end of the courses. The 60% of such courses (36) are quantitative, *i.e.* have assignments that require either computer programming and/or the resolution of some numerical problems. Some of the courses (8) are in languages different from English (*i.e.* French, Chinese and Spanish). The number of threads per course goes from 103 to 9,300 (see Fig. 1), with a median of 904.5 and a mean of 1660. The threads are composed of posts and possibly comments to posts. We refer to posts and comments as messages. Figure 2a shows the log-log plot of the number of the threads vs. the thread size (number of messages per thread). Figure 2b shows the log-log plot of the message count vs. the number of users. Both the quantities are characterized by exponential distribution.

The number of unique users per course forum goes from 103 to 11,989, with a median of 1045 and a mean of 2037. Note that students hide their IDs in about the 10% of the posts/comments. Thus the user counts given should be considered underestimates of the actual numbers of active users in the forums. Aside from students and instructors, other categories of Coursera forum users: are Course Staff (teach-

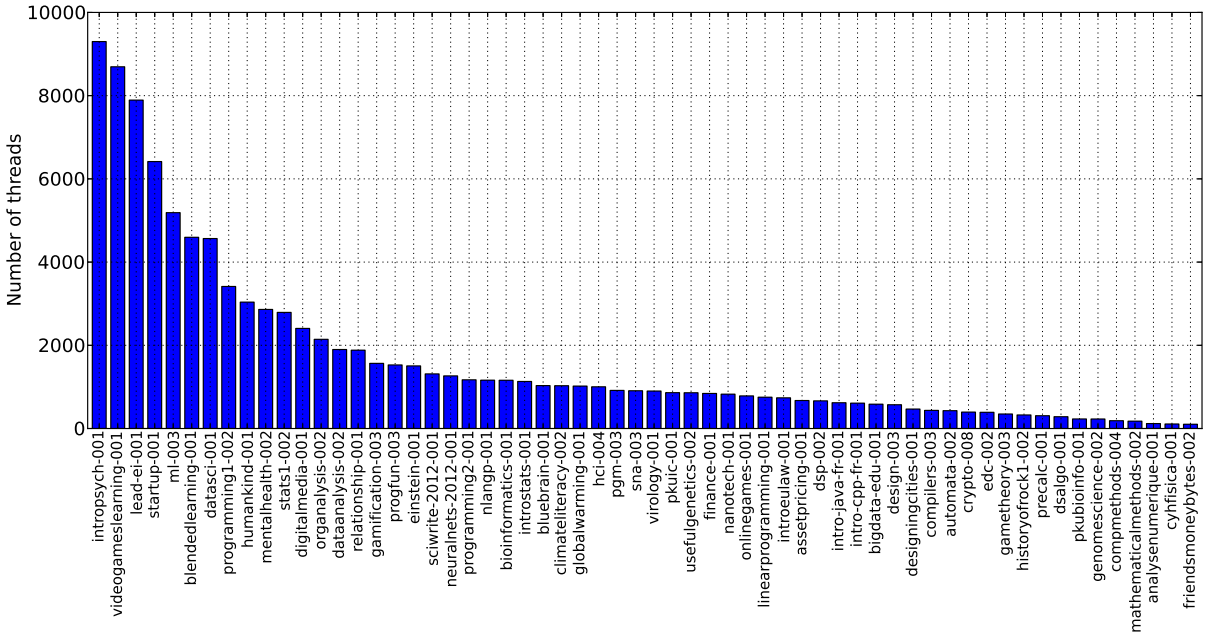
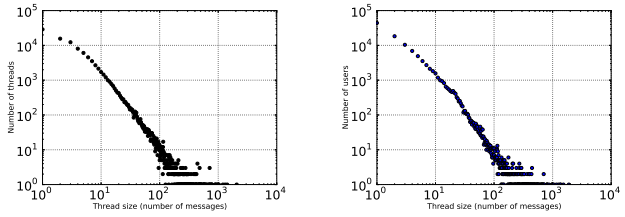


Figure 1. Number of threads vs. course identifiers.



(a) (b)

Figure 2. Log-log plots of: (a) size of the threads v.s. thread count and (b) count of users vs. number of messages.

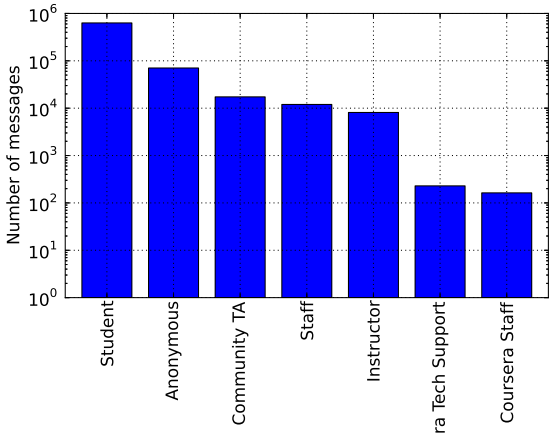


Figure 3. Total number of messages (log scale) by Coursera user type.

ing assistants), Community TAs (who are former students of the courses) and Coursera staff. The total number of messages per user type in our dataset is shown in Fig. 3.

4. Analysis of the Coursera Threads

4.1. Different usages of posts and comments

As we mentioned in Sec. 3, a discussion thread on a Coursera forum is composed of posts and possibly comments. Posts are organized in sequential chronological order order: *i.e.* a new post is placed under the last post added to the thread. Comments can be added to every post in the thread (but not to other comments). This leads to a constrained tree structure (*e.g.* equivalent to the one adopted for question-answers on Stack Overflow) in contrast with discussion threads with completely unconstrained structure (*e.g.* on Slashdot), *i.e.* where a posts can be the child of any other previous post in the thread. Some authors (*e.g.* [3]) consider posts and comments of the Coursera discussions

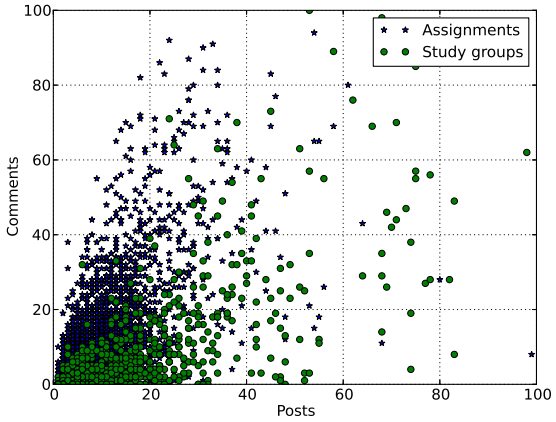


Figure 4. Scatter plot of Assignments vs. Study Groups threads represented only with their number of posts and comments.

to be completely interchangeable also due to their graphical appearance. However, our experiments show that many users seem to adopt them for distinct purposes. Thus we associate the number of posts in a discussion thread to the maximum depth of the associated tree and the number of comments to aspects of the breadth of the tree. Figure 4 shows a scatter plot of the *Assignments* vs. *Study Groups* threads for the quantitative courses in our dataset, represented only with their respective numbers of posts and comments. Discussions on assignments tend to have more comments than those on *Study Groups*. As a matter of fact, the ROC AUC performance for a linear classifier over the data in Fig. 4 is 0.589, but it is worse (ROC AUC = 0.547) if the number of messages is used as sole feature.

4.2. Number and types of messages over time

In many Coursera MOOCs, the discussion forums are split into the following subforums:

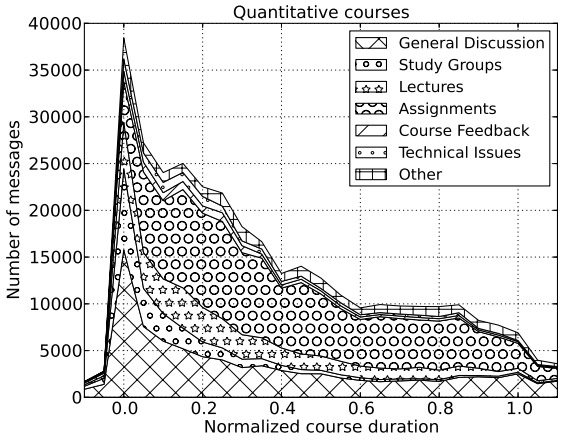
- Study Groups (Meetups)
- General Discussions
- Lectures
- Assignments
- Logistics (Platform Issues)
- (Course Material) Feedback

The instructors can customize the subforum partition in their courses and the one above is the default partition provided by Coursera. Some forum structures differ from the one above just in the names of the subforums. While other courses may adopt partially or completely different forum breakdowns. We manually unified the partition of the forums in our dataset to reflect as much as possible the above categories and to perform comparative analyses. In our dataset, about the 21.9% of the messages (162,070) is in subforums not clearly matching the above categories.

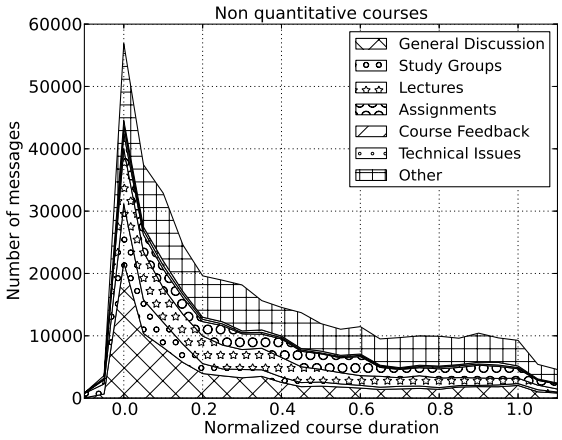
The course duration in our dataset vary from 5 to 17 weeks, with a mean of 8.25 weeks ($sd = 2.55$). In order to aggregate data from courses of different durations, we normalize and quantize the duration of each course to 20 uniform subintervals in $[0, 1]$, where: 0 indicates the beginning of a course and 1 indicates the end of a course. Figures 5a and 5b show the total number of messages per subforums respectively for the quantitative and non quantitative courses in our set. The time axis is the normalized course duration. Note that there is an exponential decay of the number of posts during the first $6/10$ of the course duration and then the total number of posts is more or less steady until the end of the course. The *Assignments* subforums get the relatively largest amount of posts after the beginning of the course and they keep a steadier number of posts throughout the course w.r.t. other subforums. In Fig. 5b, we also notice (as from other analyses) a slight peak of activity around the 85% of the course duration. Usually the activity in the forums continues after the official end of the courses: hence the non zero values for $t > 1$.

4.3. Use of anonymous messages in course forums

Users have the option to anonymize their posts/comments, hiding their IDs. In our data set, almost 10% of the posts (70,531 out 739,093) are anonymized. It seems that no previous works on Coursera MOOCs have studied this aspect. It is reasonable to assume that the authors of anonymous messages are students and that there is some overlap between the identities of the authors of signed and anonymous posts. We study the overall evolution of the rate of anonymous vs. signed student posts, over a normalized segment (where 0 is the beginning of the course and 1 the end). We notice that the number of anonymous posts decreases at a slower rate than the number of signed student posts. In particular the fraction of anonymous messages increases as the course evolves (going from 4% to above 16%). This suggests that a relevant fraction of the students who are completing the course uses the anonymization of the messages as a tool to discuss important matters as assignments. Figure 6 shows the total number of signed student posts vs. the number of anonymized student messages, over a normalized temporal axis. In Fig. 7, we plot the relative fractions of anonymous posts over the normalized course durations for four combinations of cases: quantitative vs non quantitative courses



(a)



(b)

Figure 5. Total number of messages per subforum over normalized course duration, (a) quantitative and (b) non quantitative courses.

and assignment vs other subforums. The *Assignments* subforums in quantitative courses contain the largest fraction of anonymous messages among the four cases, while the other subforums of non quantitative courses have the lowest fraction of anonymous messages. We also notice a peak of the fraction of anonymous messages around 8/10 of the course duration for assignment subforums and quantitative courses. The number of anonymous messages per thread is one of the features used by our classification approach (Sec. 5). One further observation is that the fraction of anonymous posts or comments in our dataset is higher in threads with participation of the course or Coursera staff. In particular, in the *Assignments* subforums, the average fraction of anonymous messages in threads without staff

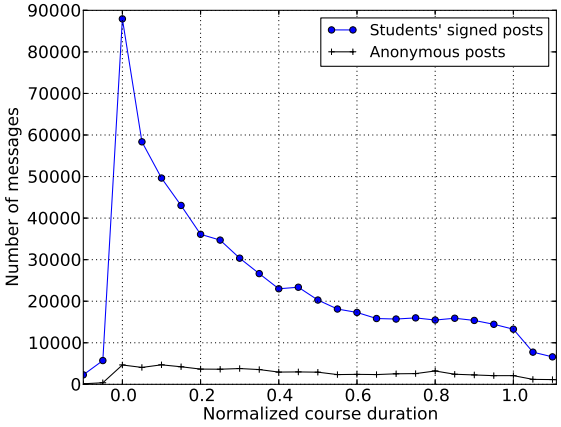


Figure 6. Total Number of signed student messages vs. number of anonymous ones.

participation is 10.33%, but it is 17.68% in threads with posts from staff users.

5. Features

In this section, we present the features for the classification of the discussion threads. A key aspect of our approach is that it is language independent. The motivation is to perform the classification across different subject related vocabularies and languages. For instance, threads from courses in Computer Science could be used to train classifiers then applied on threads from courses in physics. This is a useful characteristic for the rapidly evolving MOOC universe, where courses dedicated to new niches of knowledge and/or taught in an increasing variety of languages keep getting added. Another advantage of this approach is the lower dimensionality of our models (about 20 features), compared the potential much higher dimensionality of approaches based on n-gram features (*e.g.* [12]).

Our basic assumption is that there are some universal aspects of online asynchronous discussions that are independent from the language adopted by the participants, but that depend on the types of interactions associated to the threads. The goal of the feature engineering here is to define those aspects and quantify them for the purpose of training and testing of a classifier.

We consider a set of threads $T = \{X_1, X_2, \ldots\}$, where a thread X_k consists in a set of posts $\{p_1^{(k)}, p_2^{(k)}, \ldots, p_{n_p}^{(k)}\}$ and a possibly non empty set of comments $\{c_1^{(k)}, c_2^{(k)}, \ldots, c_{n_c}^{(k)}\}$. Sometimes we refer to posts and comments simply as messages, $\{m_1^{(k)}, m_2^{(k)}, \ldots, m_{n_m}^{(k)}\}$. For every thread X_j,

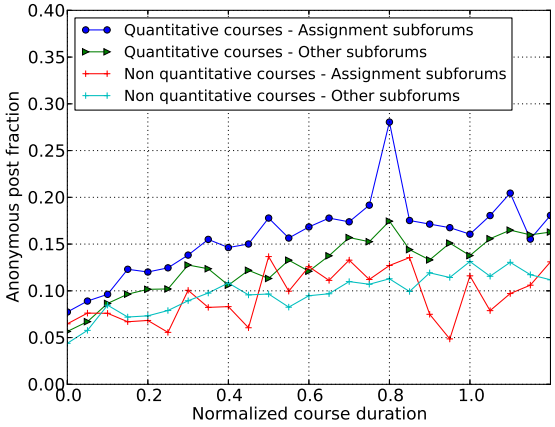


Figure 7. Fraction of students' messages that are anonymous over normalized course duration.

$n_m^{(j)} = n_p^{(j)} + n_c^{(j)}$, $n_p^{(j)} > 0$ and $n_p^{(j)} \geq 0$ (see Ssec. 4.1). Each post p_i (or comment c_j) is associated to a possibly anonymous author, u_i, a vote v_i (a signed integer), a time stamp t_i and textual content. Each thread X_j is associated to a number of views, $\mathcal{V}_j$. We now briefly examine the features we use for the classification of a thread X_j. We consider features related to 5 different aspects of the threads (structure, underlying social network, popularity, time dynamics and content).

Thread structure Features related to structural aspects of the tree associated to a discussion thread.

Number of posts (n_p)**:** number of posts in a thread.

Number of comments (n_c)**:** number of comments in a thread.

Maximum breadth (b_{max})**:** maximum breadth of a thread; computed as the maximum number of comments from unique users associated to the same post.

Index max breadth $(i_{\max b})$: index of the post associated to the maximum breadth of the thread (0, if no comments are made).

Underlying social network Global quantitative parameters of the underlying social network of users involved in the discussion threads.

Number of unique users (n_u)**:** the number of unique user IDs in a discussion thread.

Number of anonymous messages (n_{anon})**:** total number of anonymized posts and comments.

Staff replied (st_{rep})**:** 1, if there is at least 1 post by the staff, 0 otherwise.

User chain (u_{chain})**:** boolean feature to indicate that at least 3 users made more than 1 post each in the thread.

Popularity measures Estimates of the popularity of the thread

Views (vws)**:** the number of views of the thread (at the time the discussion forum was downloaded).

Vote measure $(\|\mathbf{v}\|_2^2)$**:** computed as:

$$\|\mathbf{v}\|_2^2 = \sum_{i=1}^{n_m} |v_i|^2 \tag{1}$$

Index of post with maximum vote (i_{max_v})**:**

$$i_{max\_v} = \operatorname*{argmax}_i v_i \tag{2}$$

Temporal dynamics Aspects related to the temporal dynamics.

Day of the week (week_day): the day of the week for the first post of the thread, integer in $[0, 6]$.

Relative time (t_{rel}):

$$t_{rel} = \frac{t - t_0}{t_{end} - t_0} \tag{3}$$

Average response time (t_{avg})**:**

$$t_{avg} = -n_m^{-1}(t(\text{last post}) - t(\text{first post})) \tag{4}$$

Message rate (msg_{rate})**:** Time for a thread to cumulate at least 60% of its final n_m messages.

Content Simple quantitative aspects of the text of the threads.

Average number of words per thread (w_{avg})**:**

$$w_{avg} = \frac{1}{n_m} \sum_{i=1}^{n_m} w_i. \tag{5}$$

Max words (max_{wrds})**:**

$$max_{wrds} = \max_i w_i,\ i = 1, \ldots, n_m$$

where w_i is the number of words for post i.

Index longest post (i_{max_w})**:**

$$i_{max\_w} = \operatorname*{argmax}_i w_i,$$

Internal links (in_links)**:** number of hyperlinks pointing to other Coursera.

External links (out_links)**:** number of hyperlinks pointing to pages outside Coursera.

6. Supervised Classification of the Threads

Our goal is to classify the discussion threads from the Coursera forums in categories such as social/small talk, open ended topics, (un)resolved close ended problems, course logistics, etc. The assumption is that the staff of a course would be interested in participating to the discussions threads of only a subset of those categories (*e.g.* unresolved close ended problems, but not small talk). However, this requires manual and often expert labeling of (hundreds of) thousands of threads. Therefore in this work, we use the existing Coursera subforums identifiers as class labels (see Ssec. 4.2 for a list of the subforums). We assume that such subforums are good proxies for the aforementioned ideal categories and so that a classification approach able to discriminate among these kinds of classes could also perform well on the ideal categories. Since users can place threads in the wrong subforum (*e.g.* a lecture related thread in the homework subforum), we can assume that the labels are going to be noisy and that such an uncertainty affects training of the classifiers and error computation. However, the disadvantage of having less precise and possibly noisy labels is partly compensated by having a large scale labeled dataset. Other related works relying on *handmade* labeling, *e.g.* [12], [10], used smaller datasets.

We perform the classification via a support vector machine (SVM) with linear kernels, [4], with the implementation provided by the *scikit-learn* Python module, [11]. SVM, with linear kernels has the advantage of being scalable and to provide the weights of the features.

7. Classification Results

For our experiments we consider discussion forums from 25 different Coursera quantitative courses, a training set with 12 quantitative courses and test set with 13 quantitative courses. We placed 3 courses taught in French (2) and Spanish (2) in the test set. We semi-randomly separated the rest of the courses between training and test set to keep the total number of threads in the training set 60% larger than the number of threads in the test set. We share an anonymized version of the dataset (without the text of the messages) and the Python code to produce the results given in this paper via *GitHub*: github.com/elleros/courseraforums

We first study the performance of the framework with two classes where the positive class is a specifc subforum, while the negative class is the rest of the subforums. We consider one class of threaded discussion at a time and evaluate the classification performance vs. the rest of the classes of threads lumped into one single negative class.

Table 1 shows the classification performance for the threads of one class at a time versus the rest. We compute the ROC AUC for threads with at least 3 and 5 posts. The

Table 1. Classifier Performance (ROC AUC)

	$n_p \geq 3$	$n_p \geq 5$	Top features
Gen. Discc.	0.582	0.600	$\|\mathbf{v}\|_2^2$, vws, outlinks
Assignments	0.664	0.679	vws, n_u, $\|\mathbf{v}\|_2^2$
Meetups	**0.890**	**0.917**	w_{avg}, $\|\mathbf{v}\|_2^2$, t_{rel}
Lectures	0.624	0.635	n_p, vws, n_u
Logistics	0.608	0.611	vws, $\|\mathbf{v}\|_2^2$, n_u
Feedback	0.630	0.660	n_p, $\|\mathbf{v}\|_2^2$, vws

Table 2. Confusion Matrix $(n_p \geq 3)$

	Assignments	Meetups	Lectures
Assignments	2120	134	253
Meetups	23	176	10
Lectures	616	74	292

threads of the class *Meetups* (*Study Groups*) are the easiest to classify (.89 ROC AUC), while those of the class *General Discussions* are the hardest. Clearly *General Discussions* are a mix of many types of different threads therefore lead to noisy labels and poor classification performance. The classification performance improves as the size of the thread increases. We also display the top 3 features for each class.

For the multiple class problem we consider only the classes of *Assignments*, *Lectures* and *Study Groups*. An SVM classifier with linear kernels gave the values of 0.686 and 0.700 respectively for the average precision and recall. The confusion matrix is show in Tab. 2, while the related feature weights are in Fig. 8.

7.1. Discussion

Overall, our experiments show the classification performance for *Study Groups* type of discussions is quite good (almost 0.9 ROC AUC) and that for most of the classes the ROC AUC is greater or equal than .6 without language related features. If we consider *Study Groups* to be a proxy for small talk type of threads, our approach can be used to remove small talk threads, independently from their language, with a probability of success of about 0.9. The classification performance improves as the number of posts of the threads increases. In general, features related to popularity measures (views and votes) seem to be the most effective, followed by features related to structure (number of posts, comments, maximum breadth of the threads), underlying social network (number of users) and time. We believe that there is still potential to improve the language free classification performance by experimenting with more features related to the underlying social network and temporal dynamics of the discussion threads.

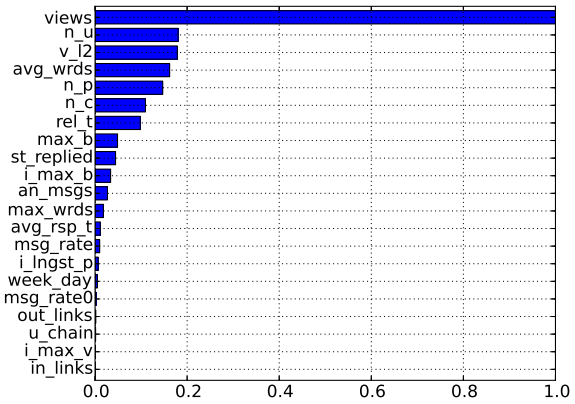


Figure 8. Feature weights for the 3 class classification (Tab. 2).

8. Conclusion

In this work, we analyze the threads from the forums of 60 Coursera MOOCs (taught in 4 languages), for a total of almost 100,000 discussions. We give insights on different usages of post, comments and anonymization. We show how the subforum partition evolves over the normalized course duration. Our preliminary analysis suggests that the number of active users (and so the number of messages) decays exponentially for the first 6/10 of a course (independently from the absolute value of its duration), but then it is more or less steady until the end of the course. We then study the supervised classification of the discussion threads. We propose several (novel) non-textual features of the threads, except for the word count. Our approach is independent from languages and specif terminologies used in the discussions. Besides, it relies only on about 20 features and therefore is easily scalable. The ROC AUC measure to classify the main standard categories of Coursera threads (*Lectures*, *Assignments*, ...) is between 0.59 and 0.89, with the best performances for *Study Groups* and *Assignments*.

Future work includes the estimation of the errors in the labels and the investigation on more complex models for the underlying social networks of users and the temporal dynamics of the discussion threads. We are also looking into labeling the threads in a more meaningful way (*e.g.* with unresolved vs. unresolved types of categories) by means of crowdsourcing.

9. Acknowledgments

We want to thank Ricky Sethi and Yolanda Gil for the essential logistic support given to us via the MadSci Network and the USC Information Sciences Institute and for their inputs and encouragements at the beginning of our research on discussion forums for online education. We also thank Vicenç Gómez, Andreas Kaltenbrunner, David Laniado and Riccardo Tasso for having shared with us their Slashdot and Wikipedia datasets which, although not used in this work, were very helpful to our research.

References

[1] A. Anderson, D. Huttenlocher, J. Kleinberg, and J. Leskovec. Engaging with massive online courses. In *Proceedings of the 23rd international conference on World Wide Web*, pages 687–698. WWW, 2014.

[2] T. Baldwin, D. Martinez, and R. B. Penman. Automatic thread classification for linux user forum information access. In *Proceedings of the Twelfth Australasian Document Computing Symposium (ADCS 2007)*, pages 72–79, 2007.

[3] C. G. Brinton, M. Chiang, S. Jain, H. Lam, Z. Liu, and F. M. F. Wong. Learning about social learning in MOOCs: From statistical analysis to generative model. *arXiv preprint arXiv:1312.2159*, 2013.

[4] C. J. Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167, 1998.

[5] V. Gómez, H. J. Kappen, N. Litvak, and A. Kaltenbrunner. A likelihood-based framework for the analysis of discussion threads. *World Wide Web*, 16(5-6):645–675, 2013.

[6] J. Huang, A. Dasgupta, A. Ghosh, J. Manning, and M. Sanders. Superposter behavior in mooc forums. In *Proceedings of the First ACM Conference on Learning @ Scale Conference*, L@S '14, pages 117–126. ACM, 2014.

[7] S. Joty, G. Carenini, and R. T. Ng. Topic segmentation and labeling in asynchronous conversations. *Journal of Artificial Intelligence Research*, 47(1):521–573, 2013.

[8] R. F. Kizilcec, C. Piech, and E. Schneider. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the III LAK*, pages 170–179. ACM, 2013.

[9] D. Laniado, R. Tasso, Y. Volkovich, and A. Kaltenbrunner. When the wikipedians talk: Network and tree structure of wikipedia discussion pages. In *ICWSM*, 2011.

[10] F.-R. Lin, L.-S. Hsieh, and F.-T. Chuang. Discovering genres of online discussion threads via text mining. *Computers & Education*, 52(2):481–495, 2009.

[11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[12] S. Ravi and J. Kim. Profiling student interactions in threaded discussions with speech act classifiers. *Frontiers in Artificial Intelligence and Applications*, 158:357, 2007.

[13] L. A. Rossi. Anonymized Coursera Discussion Threads Dataset, *GitHub*. `http://github.com/elleros/courseraforums`, July 2014.

[14] M. Wen, D. Yang, and C. P. Rosé. Sentiment analysis in MOOC discussion forums: What does it tell us? *Proceedings of Educational Data Mining*, 2014.