Education-specific Tag Recommendation in CQA Systems

Peter Babinec
Slovak University of Technology
Bratislava, Slovakia 841 04
babinec.peter@gmail.com

Ivan Srba
Slovak University of Technology
Bratislava, Slovakia 841 04
ivan.srba@stuba.sk

ABSTRACT

Systems for Community Question Answering (CQA) are well-known on the open web (e.g. Stack Overflow or Quora). They have been recently adopted also for use in educational domain (mostly in MOOCs) to mediate communication between students and teachers. As students are only novices in topics they learn about, they may need various scaffolding's to achieve effective question answering. In this work, we focus specifically on automatic recommendation of tags classifying students' questions. We propose a novel method that can automatically analyze a text of a question and suggest appropriate tags to an asker. the method takes specifics of educational domain into consideration by a two-step recommendation process in which tags reflecting course structure are recommended at first and consequently supplemented with additional related tags. Evaluation of the method on data from CS50 MOOC at Stack Exchange platform showed that the proposed method achieved higher performance in comparison with a baseline method (tag recommendation without taking educational specifics into account).

CCS Concepts

•Information systems → Recommended systems; ***Qu estion answering*****; •Human-centered computing →** *Social tagging;* **•Appliedcomputing →** E-learning;

Keywords

tag recommendation, community question answering, MO OC

1. INTRODUCTION

Massive Open Online Courses (MOOCs) brought a significant change to domain of Technologically-Enhanced Learning (TEL). While traditional TEL approaches tackle with relatively small groups of students (usually at a classroom level), typical MOOCs involve large online communities consisting of thousands of students coming from around the whole world. This shit had to be reflected in many concepts and processes including presentation of learning materials, assignments review or communication between students and teachers. Particularly communication became a significant issue since teaching a large number of students may result into many questions and some more text related to various aspects of a course. Without appropriate computer-mediated support, instructors would be overloaded by students' requests as well as students would face great information overload. the most of MOOC platforms (e.g. edX[1], Coursera[2]) use standard discussion forums to facilitate communication. On the one hand, these forums are quite easy to use, on the other hand they provide only limited possibilities how to structure and organize their content. Typically discussion posts are assigned only to categories representing individual parts of a course. Following the specific needs of MOOCs, some instructors started to use alternative external tools, such as social networking sites, real-time challeng rooms or recently Community question Answering (CQA) systems [1]. In this work we focus particularly on CQA systems, e.g. Stack Overflow[3] or quora[4], which are well-known and successful examples of how collective intelligence can be harnessed on the open web. CQA systems provide a possibility to all members of their communities to ask a question as well as provide answers/comments on questions asked by the rest of the community (see our previous survey [16] for an overview of research tasks addressed in CQA systems). their many advantages already motivated researchers to examine their potential in educational domain, what resulted into several kinds of educational CQA systems. At first, there are large-scale open CQA systems (e.g. OpenStudy [13]), that involve students without any particular restriction on MOOC or university. On the opposite side, there are CQA systems for students from a particular course (e.g. Piazza[5] or Green Dolphin [2]). In our previous work, we introduced a novel concept of organizational CQA system Askalot [15]. Askalot involves students from the whole university and thus it allows a better flow of knowledge from students in higher years to their younger peers. In contrast to standard discussion forums, CQA systems provide a more structured way of communication with better content organization - particularly tags are mostly used. Tags, created either individually or collaboratively (so called folksonomy), can serve users for search, navigation or awareness of new content (e.g. users in educational CQA system Askalot get notifications about new questions assigned with tags they watch [15]). In addition, tags may be utilized as a source of meta description, which can be further used for user/domain modeling, personaliza-

[1]https://www.edx.org/
[2]https://www.coursera.org/
[3]http://stackoverflow.com/
[4]https://www.quora.com/
[5]https://piazza.com/

tion or recommendation. In our previous work, we recognized social tagging in the educational system ALEF [5] as a powerful tool for domain model building from the bottom up.

In contrast to CQA systems on the open web, educational CQA systems are characterized by many specifics. Among them, students may have significant troubles to select appropriate tags during the process of question creation. This problem is even more eminent in educational CQA systems as in non-educational ones, since students are only novices in course topics and thus they may be unfamiliar with domain terms. In addition, students may not be aware of importance of tag selection. This can be illustrated on Askalot CQA system, which was used at Quantum Cryptography course (with about 9 200 enrolled students) in MOOC system edX. In this course, tag selection was only optional during question creation. Only 125 out of 508 questions (24.61) Following this motivation, we can witness in educational CQA systems a compelling need to support students during tag selection. It can be done by tag recommendation. In this paper, we present a novel approach for tag recommendation proposed specifically for questions in educational CQA systems. According to our best knowledge, this is the first approach which tackles with tag recom- mendation specifically designed for educational CQA systems. This paper is organized as follows. Section 2 reviews state-of-the- art approaches to tag recommendation in context of CQA systems and TEL. Section 3 describes a dataset obtained from a selected educational CQA system. Section 4 introduces our method for tag recommendation. Section 5 presents on experimental evaluation.

2. RELATED WORK

Tag recommendation is a specific kind of well-known document classification task which aims to assign a document into one or more categories. Documents may be represented by a text, images, videos or other kinds of multimedia. In our work, we tackle par- ticularly with textual documents - questions (consisting of title and description) and tags at the place of categories that questions should be assigned to. On the basis of state-of-the-art literature review, we recognized that tag recommendation can be approached very miscellaneously.

1. At first, research on tag recommendation can be divided into graph-based and content-based approaches [10]. In the first group, a graph representing users, documents and tags is created (e.g. adaptive probabilistic hypergraph [11]). Afterwards, various graph-based metrics are used to iden- tify similar questions or related tags candidates. The main limitation of graph-based methods is sparsity of the graphs [10]. This problem can be addressed by content-based approaches, which focus primarily on the document content, and thus they can recommend tags also for documents with only limited information in the graph.

2. Secondly, tag recommendation methods differ in information they exploit. The previous works utilize mainly [4]: 1) terms extracted from multiple textual features; 2) term/tag co-occurrence; and 3) tag relevance. Additional data are less commonly utilized, such as user-document-tag rela- tionship or users' tag usage history.

3. Some approaches may consider pre-assigned tags, for ex- ample to infer tag co-occurrence information (e.g. [4]). While other methods work without any information about previously assigned tags (e.g. [10]).

4. Finally, another distinction lies in the scope of tags that methods can suggest to a document. While the majority of approaches work with a set of existing tags, few approaches (e.g. [12]) can also suggest completely new tag candidates by identification of keywords from a document.

2.1 Tag Recommendation in Context of CQA

In the context of CQA systems, two main approaches to question classification exist. At first, questions are organized by a hierarchy of categories (e.g. in Yahoo! Answers). Secondly, questions are organized by a set of tags (e.g. Stack Overflow or Quora). While a notable research effort has been spent on category classification (e.g. [3] or [6]), just a few approaches tackle with tag recommendation. There is only one graph-based approach by [11], who proposed an adaptive probabilistic hypergraph, in which hyperedges can be constructed not only from a question content but also from question answering history of an asker and his/her followees. The remaining approaches can be characterized as content-based. Nishida and Fujimura [12] proposed a hierarchical classification method, in which a hierarchy of tags consists of three abstraction levels: category, theme and keyword. Saha et al. [14] introduced a discriminative model for suggesting tags to Stack Overflow questions. It consists of three main steps: 1) converting questions into vectors (with term frequency weight- ing scheme); 2) training a discriminative model (built with SVM classificators); and finally 3) suggestion of tags with top similarity. Xia et al. [17] proposed a tag recommendation method TagCom- bine which was evaluated on Stack Overflow and Freecode datasets. In contrast to the previous approaches, the proposed method com- bined 3 components: 1) a multi-label ranking component which considers tag recommendation as a multi-label learning problem; 2) a similarity based ranking component which recommends tags from similar objects; 3) a tag-term based ranking component which considers the relationship between different terms and tags.

3. DATASET DESCRIPTION

In this paper, we proceed from a dataset from educational CQA system CS50 at Stack Exchange platform[6] . This CQA system was created primarily for students enrolled in CS50 edX course[7] pro- vided by Harvard University. It is an introductory course explaining basics of computer science, such as algorithms, data structures, software engineering and web development. Students solve various tasks assigned to 9 problem sets and one final project. Several languages are used by students including C, Python, SQL, JavaScript or HTML. Our dataset was obtained via Stack Exchange API and covers questions posted between May 2014 and February 2017. It consists of 6365 questions and 796 tags. The more detailed statistics about the dataset are depicted in Table 1. We found out that about 66.4of all questions contain a code snippet. It means that various code constructions present a

[6]https://cs50.stackexchange.com/

[7]https://www.edx.org/course/introduction-computer-science-harvardx-cs50x

Table 1: CS50: Statistical analysis of questions

Name	Tags	Answers	Views	Score
mean	2.40	1.14	206.26	0.26
std	1.06	0.65	605.70	0.85
minimum	1	0	4	-7
25%	2	1	41	0
50%	2	1	80	0
75%	3	1	178	0
maximum	5	8	20268	19

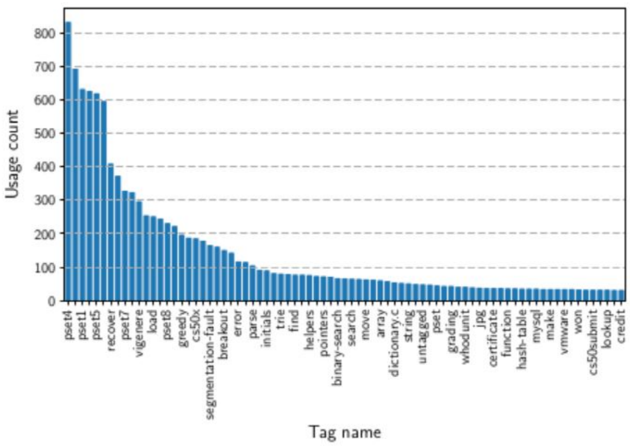

Figure 1: Most frequent tag distribution.

potential to predict tags corresponding to various programming languages. When we analyzed a long-tail distribution of tag usage (see Figure 1), we found out that among the most frequent tags many correspond to problem sets (designated as pset0 to pset8) or final project (final-project). In total, even about 74have a tag referring to a related problem set (almost all of them have exactly one of those tags). The similar categorical tags, which express a relation between questions and parts of the course, are present in the most of educational CQA systems (e.g. also in Piazza or Askalot). We also recognized that despite Stack Exchange mechanism for tags deduplication (manually defined tag synonyms are used to merge redundant tags with the same meaning), there are still some undesirable duplicates. At the same time, we identified incorrectly created tags that mostly join two individual tags, e.g. pset5-load.These inconsistencies had to be addressed in the proposal of our method in order to achieve a truly high-quality tag prediction.

4. TAG RECOMMENDATION METHOD FOR EDUCATIONAL CQA SYSTEMS

In this section, we present an overall framework for tag recommen- dation designed specifically for educational CQ A systems. The proposed method belongs to the group of content-based tag recommendation approaches. It is fundamentally based on terms extracted from a title and a description of a question and it also considers tag co-occurrence. The method works also in situations when no pre-assign tags are available and thus it can be used in three different scenarios:

1. Primarily to recommend tags in real-time during question creation process and thus supporting students with manual tag selection.

2. To recommend additional tags for questions with incomplete set of existing tags.

3. To automatically identify incorrectly assigned tags.

As we primarily intend to recommend tags for questions imme- diately at their creation time, we cannot rely on additional infor- mation from answers or comments. Finally, we restrict our method on recommendation of existing tags only. Our method is based on machine learning approach and consists of four main phases that correspond to the stages in typical machine learning workflow (see Figure 2).

4.1 Data Preprocessing

As we found out in Section 3, despite manual inspection (by sys- tem moderators) and auto-correct mechanisms (by replacing tags synonyms) the tags in CQA systems may be inconsistent. This may negatively affect tag recommendation. Therefore, at first we attempt to perform a tag correction. It can be done manually or semi-automatically by definition of correction rules, however, in both cases it is highly dependent on particular educational CQA system and MOOC. Besides tag correction, also questions themselves are prepro- cessed. The content of each question (i.e. its title and description) may be enriched with a formatting syntax (markdown). This may negatively affect the following term extraction. So all formatting marks are stripped while the intended code snippets are preserved. The content is afterwards processed by standard NLP methods, namely the text is tokenized and obtained words are lemmatized and transformed to lowercase. Finally stop-words are removed.

4.2 Feature Extraction and Transformation

In the second phase, we convert preprocessed text of questions into high-dimensional vectors, particularly into bag of words (BoW) representation (based on unigrams). As a standard bag of words model weights each term with a simple term frequency weight, we apply a transformation to use a TF-IDF (term frequency - inverse document frequency) weighting scheme instead, which represents the term-document relevance more precisely.

4.3 Model Training

To estimate the appropriateness of a tag candidate to a question, we utilize a supervised multi-label classification. Multi-label classi- fication assigns a document (in our case a question) with a set of labels (in our case a set of tags) which are not mutually exclusive. The problem of multi-label classification is commonly addressed by its transformation into a set of binary classification problems (with one vs. rest strategy), which can be solved by standard binary classification algorithms (e.g. SVM or Random Forest). Following this approach also in our method, we train the model by training a set of binary classifiers for each tag separately.

4.4 Model Application and Evaluation

In the last phase, we apply the trained model (the set of binary classifiers) to predict the most suitable tag candidates. In this step, we take into consideration the educational specifics that distinguish standard CQA system from educational ones. The main concept of our method is that we predict separately: 1) categorical tags; and 2) related tags with taking their co-occurrence with previously predicted categorical tags into account. In the first step, we limit tag recommendation to categorical tags, while the rationale for this decision is twofold. Firstly, presuming that the course categories address different learning concepts, cate- gorical tags are easier to predict, since the corresponding questions are likely to use different terminology. Secondly, as the most of ques- tions should be assigned with at least one categorical tag (except general questions that do not tackle with a particular course part), we explicitly consider their special role and minimize a number of questions when categorical tags are missed.

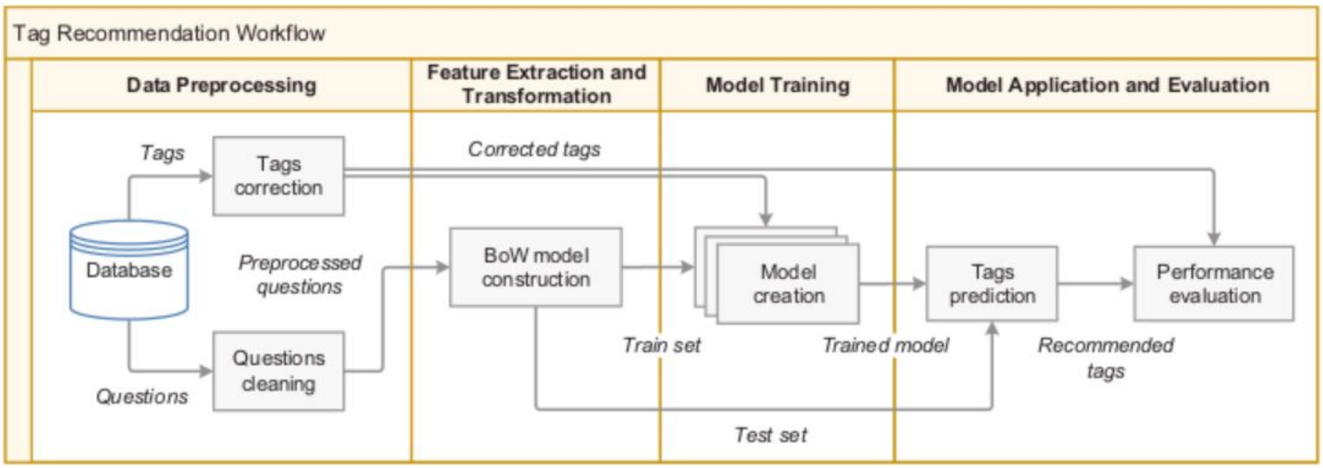

Figure 2: The phases of tag recommendation in the proposed method.

Categorical tags can be easily distinguished from the remaining tags (e.g. in the CS50 dataset, they have a standard naming conven- tion psetn; in Askalot system, categorical tags are predefined by teachers and displayed differently). Nevertheless, there is still an option to identify them automatically by considering their high usage frequency and their close affinity with the learning domain (e.g. number of their occurrence in learning materials). In the second step, we supplement categorical tags with addi- tional question-specific related tags. Similarly as in the previous step, we apply each tag classifier on a question to get tag predictions together with their predict probability (i.e. a classifier certaintythat the prediction is correct). These tags' predict probabilities are boosted by taking co-occurrence with the categorical tags (predicted in the first step) into account (see Algorithm 1). For each categorical tag, the new boosted predict probability Pnew is calculated as:

$$Pnew = Pn + Pn * (1 - Pn) * \frac{count(Tn, CTx)}{max(count(T, CTx))}$$

where Pn is a previous value of predict probability; count (Tn , CTx) is a number of co-occurrences of a related tag Tn and a categorical tag CTx ; max(count(T , CTx)) is a maximal number of co-occurrences between a categorical tag CTx and any tag T (used as a normaliza- tion, which eliminates a popularity bias of the categorical tag). The formula is derived to always keep Pnew within the range $< 0, 1 >$ and works well regardless the number of predicted cate- gorical tags. Also in case of several categorical tags, boost never significantly overcomes the previous probability Pn , since it de- pends on Pn and its strength is decreasing slowly for already strong Pn > 0.5. Maximal potential boost by 0.25 is achieved for the most co-occurring tag and Pn = 0.5. It would be possible to boost predict probability in the same way also by co-occurrence with tags previously used by a student (instead of categorical tags). It would make the tag recommendation more personalized, nevertheless in educational domain, students go through all course topics so their history do not reflect their personal interests as it is in general CQA systems. The final set of predicted tags is created by selecting predicted categorical tags and supplementing them with the most probable related tags (sorted by their boosted predict probabilities) so finally top n tags are returned. In order to prevent recommendation of tags with very low predict probability, the selection of relevant tags should be limited by a threshold for minimal necessary predict probability.

5. EXPERIMENTAL EVALUATION

We evaluated our method in an offline experiment on the CS50 dataset from Stack Exchange platform described in Section 3.

5.1 Experimental Setup

We implemented our proposed method in Python programming lan- guage. For machine learning and data manipulation the following packages were used: scikit-learn, NumPy, SciPy and Pandas.

Algorithm 1 Predicting related tags with boosting of predict probabilities by co-occurrence with categorical tags.

Input : categorical Predictions map of key-value pairs, keys are questions and values are predicted categorical tags; train Samples set for training containing questions associated with all tags but categorical; test Samples set for testing.
Output: related Predictions map of key-value pairs, keys are questions and values are predicted related tags.

```
Predict Related Tag
 (categorical Predictions, train Samples, test Samples)
model M ⟵ fit(trainSamples)
relatedPredictions ⟵ newMap(key, value)
foreach question Q  testSamples do
    probabilities ⟵ M.predictProbabilities(Q)categ
     oricalTags ⟵ categoricalPredictions.get(Q)
    foreach tag CT  categorical Tags do
        probabilities ⟵ boost(probabilities, CT)
    end
    related Tags
     decide(probabilities, count, threshold)
     relatedPredictions.add(Q, relatedT as)
end
return relatedPredictions
```

Tag correction was performed semi-automatically by definition of correction rules describing which tags should be merged together or on the contrary split into separate tags. Markdown from ques- tions was removed by BeautifulSoup library. For performing NLP tasks, NLTK library was used. For experimental evaluation we selected only those tags that have been assigned to at least 20 questions (what corresponds to an average number of questions per tag). There are two motives for this selection. At first, we prevent a possible cold start problem that can cause an over-fitting for rare tags. Moreover, we want to rec- ommend only those tags that are meaningful for search, navigation or obtaining awareness of other classmates. After tag correction and selection, we got 118 unique tags in total. We divided the dataset into a train and test set chronologically so the train set covers first two iterations of the course from 2014 to 2015 (4841 questions) and the test set covers the last iteration of the course starting in 2016 (1302 questions). By this division, we achieved an ecologically valid experiment as we simulated the situation when the method is used across all problem sets. Similarly as in the previous studies (e.g. [14], [17]), we consider tags assigned by users as predicted labels. For performance evalua- tion, we utilized standard metrics commonly used in information retrieval and document classification, namely precision, recall and F1 score. In addition, we evaluated a number of questions without any tag recommendations (i.e. a number of cases when a method cannot support an asker). As the baseline, we used a tag recommendation method without educational specifics (i.e. without two-step tag recommendation and co-occurrence boosting of predict probability), which corre- sponds to the approach proposed in [14]. This baseline recommends all tags which are labelled with a positive class by their correspond- ing binary classifiers and sorted by their predict probability. In our method, we recommended categorical tags supplemented with maximum of 3 related tags with the threshold for minimal predict probability set at 0.5 (the threshold controls a trade-off between precision and recall, therefore it is possible to set a higher value to improve precision or a lower value to improve recall).

Table 2: Comparison of the proposed method and baseline

Method used	Precision	Recall	F1 score
Our method (SVM)	0.7038	0.5979	0.6465
Baseline (SVM)	0.7003	0.5559	0.6198
Our method (RF)	0.4708	0.6403	0.5426
Baseline (RF)	0.6038	0.5717	0.5873

5.2 Results Description

The design of our method allows us to use any binary classifier. We selected particularly Support Vector Machines (SVM) and Ran- dom Forest (RF) that achieved in the existing approaches the most promising results. For both classifiers, we employed for a hyper- parameter tuning a gridsearch to systematically explore the param- eter space. Consequently, we applied a random search around the best combination of parameters found by the gridsearch, what even improved the performance of our as well as the baseline method. In case of our method, the hyper-parameter tuning was done sep- arately for prediction of categorical and related tags (the selected parameters were different because in the case of related tags, the classifiers need to tackle with a larger number of classes and a bigger similarity between BoW models). At first, we examined a performance of categorical tag recom- mendation since its high precision is a prerequisite to successful recommendation of related tags (in the case of incorrect categorical tag assignment, incorrect tags would be boosted). SVM classifier achieved precision at 92.87 and recall at 90.71, what sufficiently satisfies our requirements. Table 2 presents the comparison of overall performance (for cat- egorical and related tags) achieved by the baseline method and our proposed method. We can see a significantly better results for SVM (in case of both methods), what confirms the conclusion from the previous works that SVM is generally more successful for document classification and high-dimensional vectors. While the precision is approximately the same, we can see an improvement in recall and F1 score of our education-specific method over the baseline (our further analyses confirmed that the proposed co-occurrence boosting leads to this improvement). In addition, our method also significantly outperforms the base- line in a number of questions with no predicted tags. Our further analyses showed that the baseline failed to predict categorical tags for 164 questions (14.7(4.9one categorical tag. It means that two-step recommendation, which addresses categorical tags separately, actually helped to reduce a number of cases, when important categorical tags are missed.

As user-assigned tags may not be complete for some questions, we conducted an additional posterior evaluation of tags recom- mended by our method (with SVM predictors). We randomly se- lected 100 questions (with 211 user-assigned tags in total) and for each predicted tag, an expert annotator evaluated whether it is correct for the corresponding question, incorrect or cannot decide. Out of 173 predicted tags,

Table 3: Table 3: Tag recommendations by the proposed method

User asigned		Additionally predicted	
Matched	**Missed**	**Correct**	**Incorrect**
Baseline (SVM)			
Our method (RF)	0.4708	load	0.5426
pset2, vigenere	0.6038	0.5717	0.5873

121 (70%) matched user-assigned tags. Out of 52 remaining predicted tags, 27 tags (52%) were evaluated by an annotator as correct. It confirms that our method can be used also in the scenario when the recommended tags are used as suggestions to complement the existing user-assigned tags. To illustrate tags recommended by our method, we randomly selected few examples that represent various cases of correct as well as incorrect tag predictions (see Table 3).

6. DISCUSSION AND CONCLUSIONS

Tagging was previously recognized as essential and beneficial for efficient learning, nevertheless there is a significant lack of research addressing tag recommendation in educational domain. In this paper, we proposed a novel method for tag recommendation specifi- cally designed for educational CQA systems, that recently started to appear as an alternative to standard discussion forums in MOOCs. The novelty of our method lies in education-specific two-step tag recommendation approach, in which we predict separately important categorical tags (i.e. tags that are used to interconnect the questions with particular parts of the MOOC). At the same time, we are aware of some limitations of the proposed method in terms of cold-start problem and robustness. In order to train the model for existing tags, a sufficient number of previously correctly tagged questions must exist. It means that the method is not applicable to recommend very rare tags or at the beginning of the first iteration of the course (when there are no previous questions). Secondly, the course structure and learning materials can change in time (e.g. the order of problem sets can change). These changes will naturally influence the correctness of predicted tags, however, they will also invalidate tags previously assigned to questions. So if these tags will be readjusted to a new course structure, also all tag classifiers can be retrained afterwards. To evaluate the performance of our proposed method, we con- ducted the offline experimental evaluation on the dataset from CQA system CS50 at Stack Exchange platform. The results showed that the proposed method achieved significantly better results in comparison with the baseline. The precision of categorical tag rec- ommendation was even about 93of all tags was about 70plies that we can provide students with reliable recommendation of categorical tags and suggestions of top n additional related tags. In addition, recommended tags can serve also as a source of metadata for user modelling, content recommendation or personalization.

ACKNOWLEDGMENTS

This work was partially supported by grants No. APVV-15-0508, VG 1/0646/15, KEGA 028STU-4/2017 and it is the partial result of collaboration within the SCOPES JRP/IP, No. 160480/2015.

REFERENCES

[1] Carlos Alario-Hoyos, Mar Perez-Sanagustin, Carlos Delgado-Kloos, Hugo A. Parada G., and Mario Munoz-Organero. 2014. Delving into Participants' Profiles and Use of Social Tools in MOOCs. IEEE Transactions on Learning Technologies 7, 3 (jul 2014), 260–266.

[2] Chulakorn Aritajati and N. Hari Narayanan. 2013. Facilitating Students' Collab- oration and Learning in a Question and Answer System. In Proceedings of the 2013 Conference on Computer Supported Cooperative Work Companion - CSCW '13. ACM Press, New York, New York, USA, 101–106.

[3] Kyoungman Bae and Youngjoong Ko. 2012. An effective category classification method based on a language model for question category recommendation on a cQA service. In Proceedings of the 21st ACM International Conference on Information and Knowle

[4] Fabiano Bel m, Eder Martins, Tatiana Pontes, Jussara Almeida, and Marcos e Goncalves. 2011. Associative tag recommendation exploiting multiple textual features. In Proceedings of the 34th Int. ACM Conf. on Research and Development in Information Retrieval - SIGIR '11. ACM Press, New York, New York, USA, 1033.

[5] M ria Bielikov , Mari n Simko, Michal Barla, Jozef Tvaro ek, Martin Labaj, a a a z R bert M ro, Ivan Srba, and Jakub Sevcech. 2014. ALEF: from Application to o o Platform for Adaptive Collaborative Learning. In Recommender Systems for Technology Enhanced Learning: Research Trends Applications. Springer, Berlin, Heidelberg, 195–225.

[6] Wen Chan, Weidong Yang, Jinhui Tang, Jintao Du, Xiangdong Zhou, and Wei Wang. 2013. Community Question Topic Categorization via Hierarchical Kernelized Classification. In Proceedings of the 22nd ACM International Conference on Information and Knowledge management - CIKM '13. ACM Press, New York, New York, USA, 959–968.

[7] Ernesto Diaz-Aviles, Marco Fisichella, Ricardo Kawase, Wolfgang Nejdl, and Avar Stewart. 2011. Unsupervised Auto-tagging for Learning Object Enrichment. e Springer, Berlin, Heidelberg, 83–96.

[8] Simone Kopeinik, Elisabeth Lex, Paul Seitlinger, Dietrich Albert, and Tobias Ley. 2017. Supporting collaborative learning with tag recommendations. In Proceedings of the 7th International Learning Analytics Knowledge Conference - LAK '17. ACM Press, New York, New York, USA, 409–418.

[9] Emanuel Lacic, Dominik Kowald, Paul Seitlinger, Christoph Trattner, and Denis Parra. 2014. Recommending Items in Social Tagging Systems Using Tag and Time Information. (jun 2014). arXiv:1406.7727

[10] Marek Lipczak and Evangelos Milios. 2010. Learning in efficient tag recom- mendation. In Proceedings of the 4th ACM Conference on Recommender Systems - RecSys '10. ACM Press, New York, New York, USA, 167.

[11] Liqiang Nie, Yi-Liang Zhao, Xiangyu Wang, Jialie Shen, and Tat-seng Chua. 2014. Learning to Recommend Descriptive Tags for Questions in Social Forums. ACM Transactions on Information Systems 32, 1 (jan 2014), 1–23.

[12] Kyosuke Nishida and Ko Fujimura. 2010. Hierarchical Auto-Tagging: Organizing QA Knowledge for Everyone. In Proceedings of the 19th ACM International Conference on Information and Knowledge Management - CIKM '10. ACM Press, New York, New York, USA, 1657–1660.

[13] Ashwin Ram, Hua Ai, Preetha Ram, and Saurav Sahay. 2011. Open Social Learning Communities. In Proceedings of the International Conference on Web Intelligence, Mining and Semantics - WIMS'11. ACM Press, New York, New York, USA.

[14] Avigit K. Saha, Ripon K. Saha, and Kevin a. Schneider. 2013. A Discriminative Model Approach for Suggesting Tags Automatically for Stack Overflow Ques- tions. In Proceedings of 10th Working Conference on Mining Software Repositories - MSR '13. IEEE, 73–76.

[15] Ivan Srba and Maria Bielikova. 2015. Askalot: Community Question Answering as a Means for Knowledge Sharing in an Educational Organization. In Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work - CSCW '15 Companion. ACM Press, New York, New York, USA, 179–182.

[16] Ivan Srba and Maria Bielikova. 2016. A Comprehensive Survey and Classification of Approaches for Community Question Answering. ACM Transactions on the Web 10, 3 (aug 2016), 1–63.

[17] X Xia, D Lo, X Wang, and B Zhou. 2013. TagCombine: Tag Recommendation in Software Information Sites.(2013). 10th Working Conference on Mining Software (2013), 287–296.