

Acquisition and Determination of Correctness of Answers in Educational System Using Crowdsourcing

Marek LÁNI*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
mareklani@gmail.com*

In past years, the Web began to be used largely for education purposes. There are many technology enhanced learning (TEL) or community question answering (CQA) portals and web sites, which are being used to gain knowledge and information. Therefore, the systems are beneficial to the users, however the users can be beneficial to the systems too. We can say, that it is a win-win relationship. The benefit for the systems comes from the content, which is often crowdsourced, i.e. generated by users themselves. Representatives of such systems are for instance portal stackoverflow.com¹ or answers.yahoo.com². Thanks to these systems, it is possible to get an answer to almost any question. Nevertheless, it is important to filter and rate the answers, because the correctness of the user generated content cannot be guaranteed. Filtering and rating are activities, which users can perform in the majority of CQA systems. These activities consist, e.g. from assigning thumbs up/down or marking the answer as correct.

A basis of this CQA principle can also be used as an exercise in TEL systems. Teaching processes many times comprise scenarios, in which teachers ask questions and students (crowd) try to provide the correct answer. By this activity the students can clarify their understanding or gain new knowledge and the teacher can get an overview of the knowledge of the students. The whole process can be either part of a lessons or part of a TEL system as an exercise. In the exercise, it is possible to add additional features to this process and enhance it. Such a feature can be the evaluation of answer correctness. In case the students are involved in process of evaluation of correctness of answers, i.e. peer-grading principle is used, students can learn not only by answering, but also by evaluating and according to Sadler and Good [2] it brings

* Supervisor: Jakub Šimko, Institute of Informatics and Software Engineering

¹ <http://stackoverflow.com/>

² <http://answers.yahoo.com/>

different pedagogical advantages. This principle of interactive exercise based on answer correctness evaluation was used in the related work [3] on which our work is built. In our work, we took the exercise and enhanced it with a new features as question answering, discussion etc., while we tried to provide the best possible contributions for improvement of the learning experience. We have also qualitatively verified this exercise with several users and they agreed, that the exercise is helpful in learning process.

Nevertheless aim and research contribution of our work is focused not only on providing the new interactive learning exercise, but also on analysis of collected data. We use different methods to interpret crowd evaluations of the answers correctness and to determine, if the crowd is capable to evaluate the answers similarly to expert. There are many works in context of CQA, which tries to determine answer correctness, but only few of them use crowd evaluations of answer correctness, however they are not giving much importance to these evaluations and they use them just as one factor in determination of answerer expertise level [1][2]. As we think, that the crowd has big potential in answer correctness determination, we have created methods of interpretation of the crowd evaluations to determine answer correctness. These methods are:

Filtering of outliers – in clear cases (cases with the small distribution of evaluations) we determine outlying evaluations resp. outliers and filter their evaluations out in unclear cases (cases with the big distribution of evaluations)

Evaluations weighting according to their distribution – we have created more methods, which focuses on examination of distribution of evaluations and which try to prove different assumptions resulting from evaluations distribution.

Machine learning – this method uses neural network, where the input is vector of evaluations and output is estimation of expert evaluation.

We plan to evaluate these interpretation methods by comparison with expert evaluation assigned to every answer and as a reference values, we plan to use the simple average of the evaluations of the answers.

Acknowledgement. This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. 009STU-4/2014.

References

- [1] Kao, W. Ch., Liu, D. R., Wang S. W.: Expert finding in question-answering websites: a novel hybrid approach. In: SAC '10: Proceedings of the 2010 ACM Symposium on Applied Computing, ACM, 2010, s. 867-871.
- [2] Sadler, P., Good, E.: The Impact of Self- and Peer-Grading on Student Learning. In: Educational Assessment, Volume 11, Routledge, 2006, s. 1–31.
- [3] Šimko, J., Šimko, M., Labaj, M., Bieliková, M.: Vzdelávacie objekty typu otázka-odpoveď : kolaboratívna validácia pomocou davu študentov. In: Znalosti 2012 : Sborník príspevků 11 ročník konference, 2012, s. 11-20.