Exploiting the Knowledge of Domain Experts to Improve e-Learning Recommendation

Blessing Mbipom, Susan Craw, Stewart Massie

School of Computing Science and Digital Media,
Robert Gordon University,
Aberdeen, Scotland,
United Kingdom.

Abstract. Learning materials are increasingly available on the Web making them an excellent source of information for building e-Learning recommendation systems. However, learners often have difficulty finding the right materials to support their learning goals because they lack sufficient domain knowledge to craft effective queries that convey what they wish to learn. The unfamiliar vocabulary often used by domain experts creates a semantic gap between learners and experts, and also makes it difficult to map a learner's query to relevant learning materials. We build an e-Learning recommender system that uses background knowledge extracted from a collection of teaching materials and encyclopedia sources to support the refinement of learners' queries. This allows us to bridge the gap between learners and domain experts. We evaluate our method using a collection of learner queries and a dataset of Machine Learning and Data Mining documents. Evaluation results show our method to outperform benchmark approaches and demonstrates its effectiveness in improving e-Learning recommendation.

1 Introduction

Learners often have difficulty asking an effective query of a search engine for two reasons. First, they lack sufficient knowledge about the domain they are researching, so are unable to assemble effective keywords that identify what they wish to learn [9]. This problem results in an intent gap. Second, the vocabulary used by domain experts is often different from that used by learners, as learners often describe their problems in different terms to how experts present the solutions [14]. This presents a semantic gap.

We use background knowledge extracted automatically from a structured collection of teaching materials as shown in [10]. In our approach, we allow background knowledge which consists of a set of relevant domain concepts, to influence query refinement by providing a vocabulary for refining the queries. WatsonPaths applies a similar approach to reason over domain knowledge sources for answering medical queries [7].

Artificial intelligence (AI) methods have been applied to assist the teaching process in the design of an online course, by using AI agents to provide feedback to learners [5]. AI techniques are also used in our method to assist the learning process by creating an e-Learning recommender system that provides learners with relevant documents. The developed method allows us to bridge both the intent and semantic gap between learners and domain experts. We address the intent gap by placing a learner's query within the space of learning concepts, and identifying the most similar concepts to use for refining the query. The semantic gap is addressed by leveraging the vocabulary associated with the domain concepts to support the refinement of queries. This allows us to refine the learner's query using the vocabulary of the domain. The effect is to focus the search on relevant documents and improve the recommendations made to learners.

An e-Learning recommender system is built to evaluate the effectiveness of our approach using a collection of realistic learner queries and a dataset of Machine Learning and Data Mining resources. Evaluation results show our method to outperform a standard Bag of Words approach. The results demonstrate that using background knowledge to refine learners' queries supports the learning process by helping students to find relevant documents. There are two key contributions from this work. First, an effective method that exploits the knowledge of domain experts for refining learners' queries. Second, an e-Learning recommender system that assists the learning process and helps learners to find relevant learning materials.

2 Related Work

A large amount of e-Learning materials is available to learners on the Web. However, learners are often discouraged by the time spent in finding and assembling relevant resources to support their learning goals [2]. Often, learners are new to the topic they are researching, so they can have difficulty asking effective queries in a search engine. The unfamiliar vocabulary often used in teaching materials poses a challenge to learners trying to find relevant materials. One way of addressing these challenges is by refining queries to improve the recommendation made to learners.

One approach to query refinement is by using internal knowledge from a document collection as a feedback method [16]. This approach is similar to pseudo relevance feedback. In this method, an initial set of documents considered to be relevant are found, then terms from these documents are used to refine the query to improve retrieval performance. A drawback of this approach is that search results may be directed towards a few documents, and this can be harmful if the documents are only about specific topics. Further, the retrieval performance for difficult queries can be affected if the initial retrieval set contains irrelevant documents [8].

Another approach to query refinement involves using external knowledge sources for refining queries [12, 13]. This approach entails using terms from domain sources to refine queries [1]. A source with a good coverage is usually recommended for this task. Domain sources such as Wikipedia [6, 17], and DBpedia [11, 12] have been used to identify potentially relevant terms to use for refining queries. The effectiveness of this approach has been demonstrated in previous work [1, 4]. One potential challenge in adopting this method is the possibility of query drift, where the refined query deviates from the original query [17]. So, one needs to determine how much domain knowledge is sufficient for refining a query.

The approach in this paper draws insight from methods that use external knowledge sources for refining queries. However, this challenge is addressed within an e-Learning domain. So, the knowledge used is drawn from learning concepts from the Tables of contents (TOCs) of e-books written by domain experts. The concepts are enriched with descriptive text from DBpedia abstracts. An initial set of similar learning concepts are automatically identified for each query. Highly weighted terms from the identified concepts then form the vocabulary that is used to refine a learner's query.

3 Background Knowledge

Background knowledge refers to specialized information about a domain that can be used for general understanding and problem solving [19]. We attempt to capture background knowledge as a set of domain concepts, each representing an important topic in the domain. For example, in a learning domain, such as Data Mining, you would find topics such as classification, association rules, and regression. Each of these topics is represented by a concept, in the form of a concept label and a pseudo-document which describes the concept. The knowledge extraction process is shown in Figure 1. The input to this process are domain knowledge sources, such as a structured collection of teaching materials and an encyclopedia source. Next, ngrams are automatically extracted from the structured collection to generate a set of potential domain concepts. Then a domain lexicon is used to validate the extracted ngrams to ensure that the ngrams are also being used in another information source. The encyclopedia provides descriptions for the identified ngrams. The output is a set of domain concepts, each having a concept label and a pseudo-document.

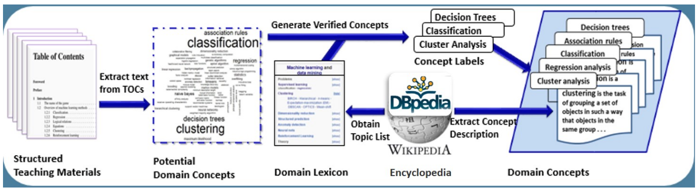

Fig. 1. An overview of the knowledge extraction process

We adopt the approach developed by [10] to capture background knowledge to underpin e-Learning recommendation for a broad learning topic such as Machine Learning and Data Mining. Two knowledge sources are used as initial inputs for discovering domain concepts. First, the TOCs of 20 e-Books are used as a structured collection of teaching materials, which provide a source for extracting important topics identified by teaching experts in the domain. Second, a domain lexicon is used to verify that the concept labels identified from the teaching materials are directly relevant. The lexicon is created from Wikipedia because it contains articles for many learning domains [20], and the contributions of many people [18]. The domain lexicon containing 664 Wiki-phrases provides a broader but more detailed coverage of the relevant topics in the domain. Then, an encyclopedia source, such as DBpedia abstracts is searched and this provides the relevant text to form a pseudo-document for each verified concept label. The final output from this process is the background knowledge containing a set of 150 domain concepts each comprising a concept label and an associated pseudo-document.

The concept vocabulary with terms t_1 to t_c, from concepts, C_1 to C_m is used to create a concept term matrix with TF-IDF weighting [15]. TF-IDF is useful for distinguishing concept terms in the concept space, and for identifying concepts that are relevant to queries hence its use in this method. A set of potentially useful concept terms are selected from the concept vocabulary as a means of scaling up the representation. So, we represent the background knowledge by using the top 10% of concept terms that have the highest TF-IDF values. The selected concept terms are used to create a concept term matrix. The selected terms t_{c1} to t_{cn}, from the concepts, C_1 to C_m are the set of potential terms that would be used for refining a query.

4 Refining Queries using Domain Concepts

The background knowledge representation supports the refinement of queries as a step towards addressing the intent and semantic gap learners face. When a new query is received from a learner, a search is performed on all the domain concepts. A ranked list of domain concepts that are similar to the query is retrieved. The terms from the term-vectors of the most similar concepts are put together to create a potential refined query. Terms with the highest weights are selected from the potential refined query and added to the initial query to create a refined query. The refined query can then be used to search on a document collection, and documents would be retrieved and presented to the learner. We expect the retrieved documents to be relevant because the query used for the search has been generated using domain concepts related to the initial query.

Figure 2 contains an illustration of how a refined query is generated. In this example, Cq_1, Cq_2, and Cq_k are the k most similar concepts to the query, while t_{c1} to t_{cn}, are the selected concept terms. The entries into the matrix are the tf-idf weights of the terms in the respective concepts. While, $SimScore_1$, $SimScore_2$ and $SimScore_k$ are the similarity scores between the query and concepts Cq_1, Cq_2, and Cq_k respectively. The weight of a concept term such as t_{c1} in the potential refined query is generated by computing the weighted sum for that term.

Terms / Concepts	t_{c1}	t_{c2}	...	t_{cn}	SimScore
C_{q1}	tf-idf(t_{c1}, C_{q1})				$SimScore_1$
C_{q2}	tf-idf(t_{c1}, C_{q2})				$SimScore_2$
C_{qk}	tf-idf(t_{c1}, C_{qk})				$SimScore_k$
Potential refined query	t_{c1}	t_{c2}	...	t_{cn}	

Fig. 2. Generating a refined query

An example is shown in Equation 1 of how the weight of term t_{c1} is computed. The weighted sum of t_{c1} is achieved by multiplying the weight $SimScore1$ with the tf-idf scores of terms that appear in concept Cq_1. This is also done for terms that appear in Cq_2 and Cq_k respectively. The column sum for t_{c1} is then computed. Altering the tf-idf weights of concept terms with the respective similarity scores would allow terms from concepts that are more similar to the query to have more influence in the refined query. The output from this process is a potential refined query containing concept terms, t_{c1} to t_{cn} together with their respective weights.

$$WeightedSum(t_{c1}) = \sum_{i=1}^{k} (\text{tf-idf}(t_{c1}, C_{qi}) \times SimScore_i) \tag{1}$$

where t_{c1} is a concept term, and $\text{tf-idf}(t_{c1}, C_{qi})$ is the tf-idf score of term, t_{c1} in the i-th concept, C_{qi}. $SimScore_i$ is the similarity between query, q and the i-th concept C_{qi}.

The weight of a term in the potential refined query gives an indication of the importance of the term within the concept space in relation to the given query. We take advantage of this weight by selecting the highly weighted terms from the potential refined query. These terms are used for generating a refined query. We adopt this approach so that noisy terms would not be included during query refinement [17]. We include the initial query as part of the refined query to maintain the context of the query.

A refined query can be seen as an *initial query* + *generated concept terms*. For example, given an initial query from a learner such as: *"How do you implement gradient descent algorithm?"*. A search is performed on the set of domain concepts and the 3 most similar concepts to this query are: *stochastic gradient descent*, *backpropagation*, and *winnow algorithm*. The terms from these concepts are put together as described above. Ten of the highest weighted terms from an amalgamation of these concepts are: gradient, descent, stochastic, formula, update, momentum, delta, rate, derivative, backpropagation. These terms would then be added to the initial query. So the refined query becomes: *how do you implement gradient descent algorithm gradient descent stochastic formula update momentum delta rate derivative backpropagation*. This refined query is what is used to search on a document collection.

5 User Evaluation

An e-Learning recommender system is developed to demonstrate how refined queries can be used for e-Learning recommendation. Three methods are implemented in the system. First our CONCEPTBASED query refinement method which uses the most similar domain concepts to create a concept based representation of a query. Second, the Bag-Of-Words (BOW) method, which is a standard Information Retrieval method, where a learner's query is represented using the terms in the query only. Finally, a HYBRID method which takes advantage of the features in a query to make a dynamic choice in determining when to apply the BOW or CONCEPTBASED method to refine a query. The evaluation aims to compare the relevance of documents retrieved using the CONCEPTBASED and HYBRID methods against the standard BOW method.

5.1 Data

The data used for evaluation is drawn from 2 collections. First, a document collection for recommendation and second, a collection of learner queries. The document collection contains 504 chapters from 32 Machine Learning and Data Mining (ML/DM) e-books. The collection used is fairly distributed across the domain concepts. The query collection contains learner-focused queries which we use in the evaluation. We used 2 sources to generate our queries. First, postgraduate students in the School of Computing Science and Digital Media took part as learners in generating queries. An e-mail specifying the task was sent to them. In order to allow learners to send anonymous responses, and return their queries without seeing what others had asked, a Google form was created to capture the queries. Second, online sources such as Coursera's Machine Learning MOOC and Quora were used to generate queries. Course specific questions were accessed from Coursera's MOOC, while the open questions in Quora from the Machine Learning and Data Mining topics were chosen.

For the query sources, our aim was to have realistic learner queries, so we used queries where the user wanted to learn about a technique, for example: "what are the various data mining techniques for fraud detection". We did not use generic or career-related queries such as "What is it like to be a data scientist at Amazon?", or queries that were out of scope such as "is there any course on ML?". Overall, 11 queries are from learners and 59 queries from online sources, resulting in 70 queries.

The evaluation system was deployed using Microsoft Azure [3], so the system could be accessible to users online. The evaluation system was available online for 8 weeks. To generate the recommendations for evaluation, the 70 queries were run using the methods, and the top 3 recommendations from each method was stored. A link to the system was shared through mailing lists to users. Each user completed a questionnaire at the start to provide data about their background in the ML/DM domain. The data users provided gave us an idea of the experience and expertise of the users. An analysis of the evaluation results using the questionnaire data allowed us to gain valuable insights into the way different users judged the recommendations made by the system.

There were responses from 22 users. For the user roles, 16 were PhD students, 3 were Researchers, while 3 were Lecturers or Professors. All users had at least an MSc degree or higher. There were 3 users with over 10 years experience in ML/DM, 3 users had over 5 years experience, while 10 users had between 3-5 years experience and 5 users had 1-2 years experience, and only 1 user had less than a year's experience in ML/DM. This level of experience in the ML/DM topic is useful, because the judgements made should be from people who are conversant with the domain. In terms of their expertise in ML/DM, there were 2 experts, 16 competent users and only 4 beginners.

5.2 Design of the User Evaluation

At the start, each user was shown a briefing containing a guide on the evaluation study, the task, a note on confidentiality, and the researcher's contact information. During the study, the user was shown one query each time to evaluate. For each query, the user could choose to skip, if the user had no idea about the query, or proceed to evaluate because the user had some understanding of the query. This allowed each user to evaluate recommendations for queries they were knowledgeable about. When evaluating a query, the user was shown up to 6 retrieved documents in random order. The set of documents were the top 3 documents from the CONCEPTBASED and BOW methods. Since HYBRID applies either CONCEPTBASED or BOW, the documents for HYBRID are already included in the retrieval set shown to users.

It is important that the retrieval set of documents shown to users is presented in a way that avoids any potential bias. Three issues of bias are considered and addressed. First, the users do not know which method produced the recommendations they are evaluating, this prevents a user from favouring one method over the others. Second, the order of documents presented to users is randomized. This rules out the bias of documents shown at the top being considered to be relevant over those lower down the recommendation list. Third, the same user evaluates the recommendations from both BOW and CONCEPTBASED methods for the same query. This ensures that the same user gives an evaluation for both methods at the same time for a given query. This prevents the possibility that we only receive ratings from a positive user for one method, and ratings from a generally negative user for the other method.

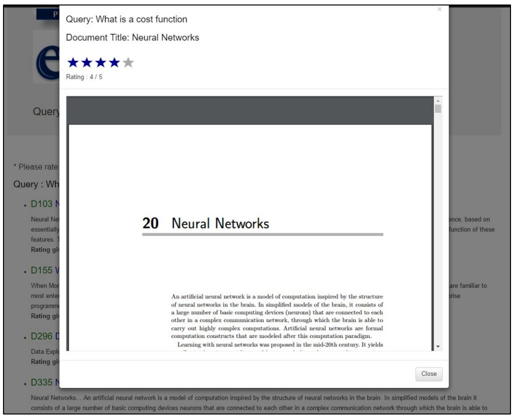

Fig. 3. Selected document and star-rating

Figure 3 shows the page used to display a recommended document that the learner has clicked on. The relevance of each document to the query is captured using a rating scale of 1 to 5 stars where 1 is least likely to be relevant and 5 stars is very relevant. The rating stars as shown in Figure 3 were included in the page that contained the document, so that each user had an opportunity of reading the document before rating it.

5.3 Evaluation Metrics

The evaluation uses the ratings given by users across all query-recommendation pairs to measure the performance of the methods. We compute the $rating$ as the average of ratings from those users who have evaluated the recommendation, r for the query, q.

$$rating\,(q,r) = \sum_{u \in U_q} \frac{R_u(q,r)}{|U_q|} \tag{2}$$

where (q, r) is a query-recommendation pair, U_q are the users that have evaluated a query, q, and R_u is the rating a user, u has given to a (q, r) pair. Performance of a method is computed by taking the average $rating$ across the queries.

5.4 Recommendation Results

Users evaluated 105 queries and provided ratings for 521 query-recommendation (q, r) pairs. There were 6 of the total 70 queries that were not evaluated. Figure 4 shows the spread of ratings for (q, r) pairs that were evaluated as heat maps for the CONCEPTBASED, HYBRID and BOW methods respectively. The ratings range from 5 to 1, and the colours are from green for the highest rating of 5, to red for the lowest score, 1. In plotting the heat map, the average rating values per (q, r) pairs are sorted in descending order. The 3 heat maps are plotted in the same way using the actual average rating value given for each (q, r) pair. Lines are included to show a change in the rating value.

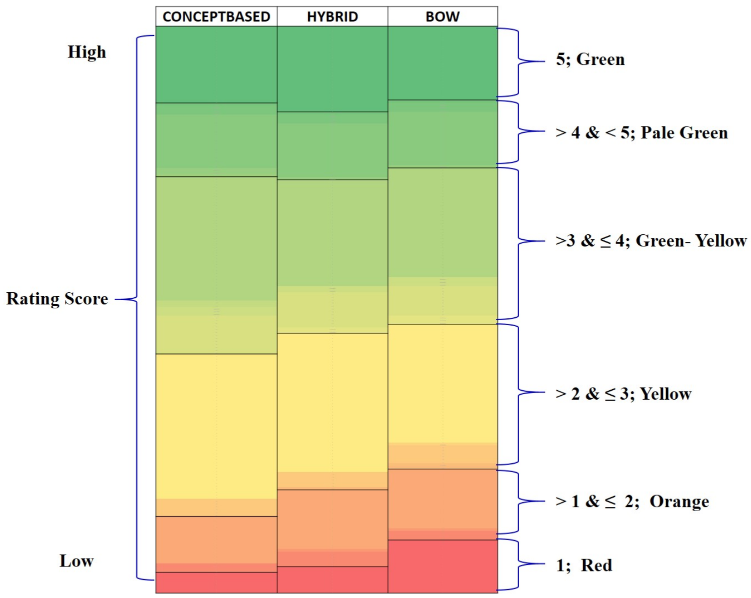

Fig. 4. Spread of ratings for query-recommendation pairs

In considering high ratings, HYBRID does best in producing documents with high quality ratings followed by CONCEPTBASED and then BOW. HYBRID is able to correctly identify when to use either BOW or CONCEPTBASED for refining a query in order to produce such good quality documents. For (q, r) pairs with the lowest ratings, the standard BOW method produces the highest number of documents with very poor ratings, BOW has difficulty preventing poor retrievals from being shown. CONCEPTBASED has the fewest number of (q, r) pairs with poor ratings. In particular CONCEPTBASED is very good at not presenting poor retrievals to users. These results for all the marked areas show that users gave higher ratings to the recommendations made using the HYBRID and CONCEPTBASED methods than those made using the standard BOW method.

The rating for each (q, r) pair is computed using Equation 2. Table 1 contains the rating given by users for each method. We wish to know if a user's expertise affects the ratings they provided. Table 1 captures the average ratings of all users as well as ratings based on expertise of users for each method. This would allow us to confirm if there is some agreement among the users irrespective of their expertise. For all users, we find that CB $>$ HYBRID $>$ BOW. So, using a CONCEPTBASED representation of a query to find learning materials is better than when HYBRID or BOW is used. For experts, the average rating scores for all methods are lower, nonetheless the experts still agree that the best performance is from the CONCEPTBASED method. We are confident in the results received from experts because they know what topics learners should be interested in. The competent users have higher rating scores across all methods, and they also agree that the CONCEPTBASED method performs better. Although the ratings by the beginners for all the methods are very similar, their rating scores also agree with the other users that the CONCEPTBASED method performs best.

Table 1. Average rating

Method	All Users	Experts	Competent	Beginners
CONCEPTBASED	3.54	3.13	3.66	3.29
HYBRID	3.45	2.71	3.58	3.27
BOW	3.33	2.58	3.46	3.25

Relevance judgement is subjective and depends on the users who are providing ratings for the documents seen. We wanted to know how many users preferred the recommendations produced using either our CONCEPTBASED method (CB) or the standard BOW method. We use the rating as given in Equation 2 for BOW and CB and we count how many users rated documents from one method higher than the other. Table 2 contains the results for the preference users had for either CONCEPTBASED or BOW. Half of the experts preferred CONCEPTBASED, while the other half of the experts thought both methods were the same. None of the experts thought that BOW was better than CONCEPTBASED. There were 14 users that preferred the recommendations produced using the CONCEPTBASED method over those of the BOW method.

We can trust the judgement of experts and competent users as they are more knowledgeable of the domain, and they know what documents should be relevant to learners. We note that no expert preferred the standard BOW method. Recall from Table 1, that the scores provided by beginners for all methods were very similar. This is because beginners are not so reliable when deciding if materials are relevant or not. So we cannot rely totally on the judgements provided by beginners. There were 4 users who rated documents seen from both methods equally. This could come from users that gave equal scores to all documents seen for a query. Overall, majority of competent users preferred the recommendations produced by the CONCEPTBASED method over BOW.

Table 2. Preference of methods

Preference	Users	Demographics
CB > BOW	14	1 expert, 12 competent, 1 beginner
BOW > CB	4	1 competent, 3 beginners
CB = BOW	4	1 expert, 3 competent

5.5 Coverage of Relevant Topics

Having recommendations with high ratings is good, but recommending documents that cover topics relevant to the query is important. We capture the coverage by asking users to provide feedback after they evaluate each query. This feedback is optional for users. Users were asked what extent they thought the documents they were shown covered the topics relevant to the query. The user could make a selection from 4 options: Complete, Good, Partial and Limited coverage. There were 50% of entries from users which stated that the documents had good coverage. An additional 19% of entries said the documents had complete coverage, while 21% of entries said the documents had partial coverage. Only 10% of entries said the documents had limited coverage. So, most recommendations covered topics that were relevant to the queries.

Figure 5 contains a heat map showing a broader view of the spread of coverage scores for the queries evaluated. The heat map is plotted by converting the coverage options to numeric values where Complete is 4, Good is 3, Partial is 2, and Limited is 1. The colour ranges from green for complete coverage to red for limited coverage. The heat map is sorted twice. First based on the average coverage scores per query and second based on the average coverage scores per user. Figure 5 captures the queries with the best coverage at the top, as seen by the dark green slots nearer the top left, and those with least coverage as seen by the red slots nearer the bottom left of the diagram.

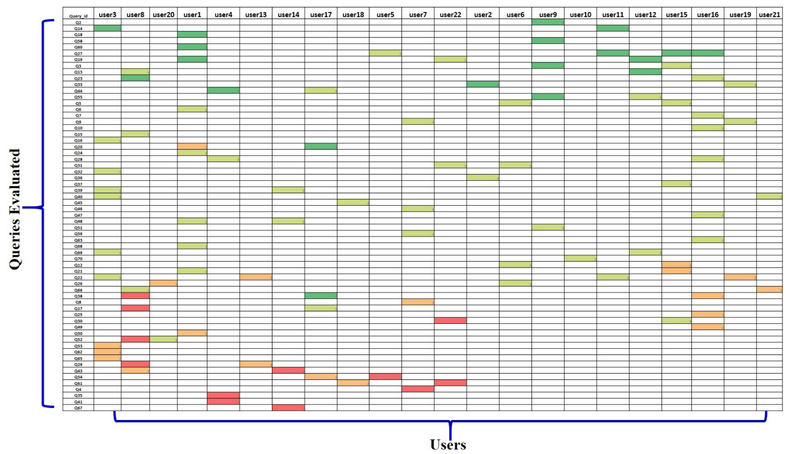

Fig. 5. Spread of scores for the coverage

The queries rated had consistent coverage scores from more than one user, so this allowed us to gain more insight to the queries. For example, a query such as: *"How does cluster analysis work?"* had "complete coverage" from 3 users and "good coverage" from 1 user. While the query: *"Is it possible to use reinforcement learning to solve any supervised or unsupervised problem?"* had partial coverage from 1 user and limited coverage from 1 user. We draw the following insights from the coverage. First, some queries may not be well written and so can be hard to understand. Second, the topic contained in some queries may be very specialized. These results show potential in exploring the features of queries when refining them for e-Learning recommendation.

6 Conclusions

There are large amounts of e-Learning materials available to learners on the Web. However, learners often have difficulty finding relevant materials. Learners are often new to the topic they are researching, and so are unable to create effective queries for a search engine. We have created an e-Learning recommender system that uses background knowledge extracted from teaching materials from domain experts, to support the refinement of learners' queries. The rich vocabulary from the background knowledge allows us to focus the search on documents that are relevant to learners.

We use a collection of realistic queries and a dataset of Machine Learning and Data Mining documents for evaluation. Relevance judgement is subjective and depends on the opinions of individual users taking part in the evaluation. However, we had a good level of consensus on the relevance judgements provided for each method by users with different levels of expertise. Results from experts, competent users and beginners all showed that using a CONCEPTBASED representation of a query to search produced documents that were consistently more relevant to learners than when the standard method was used. User evaluation results demonstrate the effectiveness of our approach in exploiting the knowledge of domain experts to improve e-Learning recommendation. In future, the background knowledge can be developed to provide a guided view of a learning domain and support intelligent browsing of documents.

References

1. Bendersky, M., Metzler, D., Croft, W.B.: Effective query formulation with multiple information sources. In: Proceedings of the 5th ACM International Conference on Web Search and Data Mining, pp. 443–452 (2012)
2. Chen, W., Niu, Z., Zhao, X., Li, Y.: A hybrid recommendation algorithm adapted in e-learning environments. World Wide Web **17**(2), 271–284 (2014)
3. Copeland, M., Soh, J., Puca, A., Manning, M., Gollob, D.: Microsoft Azure: Planning, Deploying, and Managing Your Data Center in the Cloud (2015)
4. Diaz, F., Metzler, D.: Improving the estimation of relevance models using large external corpora. In: Proceedings of the 29th ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 154–161 (2006)
5. Goel, A.K., Joyner, D.A.: Design of an online course on Knowledge-Based AI. In: AAAI, pp. 4089–4094 (2016)

6. He, B., Ounis, I.: Combining fields for query expansion and adaptive query expansion. Information Processing & Management **43**(5), 1294–1307 (2007)
7. Lally, A., Bagchi, S., Barborak, M.A., Buchanan, D.W., Chu-Carroll, J., Ferrucci, D.A., Glass, M.R., Kalyanpur, A., Mueller, E.T., Murdock, J.W., et al.: WatsonPaths: Scenario-based question answering and inference over unstructured information. AI Magazine **38**(2) (2017)
8. Li, Y., Luk, W.P.R., Ho, K.S.E., Chung, F.L.K.: Improving weak ad-hoc queries using Wikipedia as external corpus. In: Proceedings of the 30th ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 797–798 (2007)
9. Liu, J., Kim, C.S., Creel, C.: Why do users feel search task difficult? Proceedings of the Association for Information Science and Technology **50**(1), 1–4 (2013)
10. Mbipom, B., Craw, S., Massie, S.: Harnessing background knowledge for e-learning recommendation. In: M. Bramer, M. Petridis (eds.) Research and Development in Intelligent Systems XXXIII: Incorporating Applications and Innovations in Intelligent Systems XXIV, pp. 3–17. Springer (2016)
11. Meij, E., Bron, M., Hollink, L., Huurnink, B., De Rijke, M.: Learning semantic query suggestions. The Semantic Web-ISWC pp. 424–440 (2009)
12. Meij, E., Bron, M., Hollink, L., Huurnink, B., de Rijke, M.: Mapping queries to the linking open data cloud: A case study using DBpedia. Web Semantics: Science, Services and Agents on the World Wide Web **9**(4), 418–433 (2011)
13. Meij, E., de Rijke, M.: Supervised query modeling using Wikipedia. In: Proceedings of the 33rd ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 875–876 (2010)
14. Millard, D.E., Gibbins, N.M., Michaelides, D.T., Weal, M.J.: Mind the semantic gap. In: Proceedings of the Sixteenth ACM Conference on Hypertext and Hypermedia, HYPERTEXT, pp. 54–62. ACM (2005)
15. Salton, G., Buckley, C.: Term-weighting approaches in automatic text retrieval. Information Processing & Management **24**(5), 513 – 523 (1988)
16. Wu, H., Fang, H.: An incremental approach to efficient pseudo-relevance feedback. In: Proceedings of the 36th ACM SIGIR Conference on Research and development in Information Retrieval, pp. 553–562 (2013)
17. Xu, Y., Jones, G.J., Wang, B.: Query dependent pseudo-relevance feedback based on wikipedia. In: Proceedings of the 32nd ACM SIGIR conference on Research and development in Information Retrieval, pp. 59–66 (2009)
18. Yang, H.L., Lai, C.Y.: Motivations of Wikipedia content contributors. Computers in Human Behavior **26**(6), 1377 – 1383 (2010)
19. Zhang, X., Liu, J., Cole, M.: Task topic knowledge vs. background domain knowledge: Impact of two types of knowledge on user search performance. In: Advances in Information Systems and Technologies, pp. 179–191. Springer (2013)
20. Zheng, Z., Li, F., Huang, M., Zhu, X.: Learning to link entities with knowledge base. In: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp. 483–491 (2010)