International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-03, Issue-09, Dec 2017

Fine Grain Knowledge Sharing in Agriculture

[1]Rakhi. R. Bornare, [2]Prof. P. N. Kalavadekar

[1]PG Student, [2]Professor,Dept. of Computer Engineering, SRESCOEK, SPPU, India.

[1]*rakhibornare0l3@gmail.com*,[2]*kprak3004@gmail.com*

Abstract: For most of the people, web interaction is a very common phase to acquire information. It is possible that in a combined environment, more than one person may try to obtain similar information in one domain. One person may like to solve a problem using an unfamiliar Apache Tomcat which he had studied by another person before. Connecting and then sharing with that persons will be more beneficial to get there learned knowledge. Fine grained knowledge sharing is proposed for this combined environment. The system is proposed to classify the surfed data into clusters and summarize the details in fine grained details. For any system the efficiency depends upon the surfing. The framework includes: Data which is surfed, clustered into tasks. Then task is mined in fine grained output. To get proper result, the search method is applied to the output (mined results).The concept of Data Mining in fine grained knowledge is combined with the information gathering and classification to produce efficient data searching technique in agriculture system.

Keywords: Web interaction, Fined-grain, web, mine.

I. INTRODUCTION

Support Vector Machines (SVM) shows excellent accuracy to de-liver data classification, and accuracy in terms of high performance [1] and [2]. Web Interaction and communication with colleagues are very common routine for knowledge acquitting [3]. Data duplication detection is possible when the data are present in the same real world. Finding the same value in more than one file refers to Data Matching. In data integration, the most essential process is duplication detection. The task of finding entries, which is also called record matching [4] refers to similar entity in two or more files. Problems of duplication detection are solved by performing record matching, that is why the requirements of recognizing the appropriate record matching technique follow. Currently used method for duplication detection in known as supervised methods.

For example, several departments in the institution may require the same system software, and the staff of any one department has al-ready have searched about that software. Also in project lab when the new projects come, the developer needs to retrain the previously completed projects background to acquire the background knowledge. When these cases arise, the solution to these can be achieved by restoring the previous data, so that it can save the time and data usages with the benefit of accuracy and minimizing the errors.

In Agriculture system the query matching is proposed to refine the search result as the user trying to retrieve data which are saved in a cluster could be categorize as they could identify them easily for their work .The concept of data mining is used in this proposed system, for efficient searching as well as it provides the user an accurate result within data base. In agriculture system there is a need for some mining tools so that the classification of data based on the trained data could be received in minimum time and with great relatively. In Agriculture system the query matching [5] is proposed to refine the search result as the user trying to retrieve data which are saved in a cluster could be categorise as they could identify them easily for their work [6].

II. RELATED WORK

A]Expert Search*:* Expert search aims at retrieving people who have expertise on the given query topic. Early approaches involve building a knowledge base which contains the descriptions of people's skills within an organization. Expert search became a hot research area since the start of the TREC enterprise track in 2005. Balog*et al.* proposed a language model framework for expert search. Their Model 2 is a document centric approach which first computes the relevance of documents to a query and then accumulates for each candidate the relevance scores of the documents that are associated with the candidate. This process was formulated in a generative probabilistic model. Balog*et al.* showed that Model 2 performed better [1] and it became one of the most prominent methods for expert search. Other methods have been proposed for enterprise expert search, but the nature of these methods is still accumulating relevance scores of associated documents to candidates. Expert retrieval in other scenarios has also been studied, e.g. online question answering communities, academic society. The proposed advisor search problem is different from traditional expert search. (1) Advisor search is dedicated to retrieving people who are most likely

 DOI : 10.18231/2454-9150.2017.0056

possessing the desired piece of fine-grained knowledge, while traditional expert search does not explicitly take this goal. The critical difference lies in the data, i.e. sessions are significantly different from documents in enterprise repositories. A person typically generates multiple sessions for a micro-aspect of a task, e.g. a person could spend many sessions learning about Java multithreading skills. In other words, the uniqueness of sessions is that they contain semantic structures which reflect people's knowledge acquisition process. If we treat sessions as documents in an enterprise repository and apply the traditional expert search methods.

B]Analysis of Search Tasks

Recently, researchers have focused on detecting, modeling and analyzing user search tasks from query logs. Here we name some representative works. Jones and Klinkner found that search tasks are interleaved and used classifiers to segment the sequence of user queries into tasks [15]. Liu and Belkin combined task stage and task type with dwell time to predict the usefulness of a result document, using a 3-stage and 2-type controlled experiment. Ji*et al.* used graph regularization to identify search tasks in query logs. Kotov*et al.* designed classifiers to identify same task queries for a given query and to predict whether a user will resume a task.

C]Topic Modeling

Topic modeling is a popular tool for analyzing topics in a document collection. topic modeling decomposes a document into topics. After applying topic modeling methods on session data, it is still difficult to find the right advisor by using the mined topics. This is because a person with many sessions containing partially relevant topics would still be ranked unexpectedly high, due to the accumulation of relevance among sessions. Grouping sessions into micro-aspects is important for advisor search.

III. PROPOSED SYSTEM

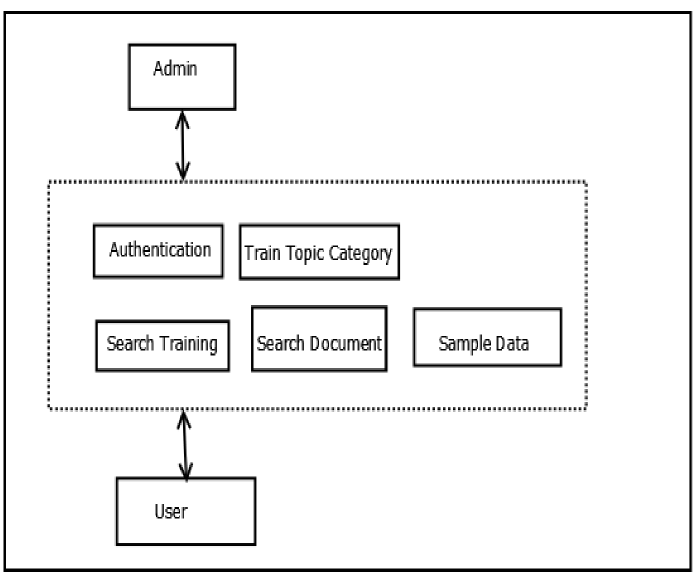

Fig 1.System Diagram

Admin Module : The admin will be authenticated to facilitate access to admin related operations

Train Topic Categories: The selected categories present in the system will trained for associated keywords using sample file uploads specified by the admin

User Module : The user will be authenticated to facilitate access to user operations Search

Training: The training model will be generated here for various search keyword associated to trained topics with respect to various documents in terms of the click ratio generated for each document relative to a particular topic will be mapped

Search Documents: The keywords search will be identified in terms of above mapping to identify the documents more relative in terms of the click ratio identified by the training model. The filtered documents will be then displayed as the final output of the search mechanism.

IV. RESULT AND ANALYSIS

Dataset Used: using standard Agriculture dataset for train nine categories**.**

Soil Testing, Fertilizer, Crops, Weather, Market rates, Equipment's, Seeds, Flowers, Fruits.

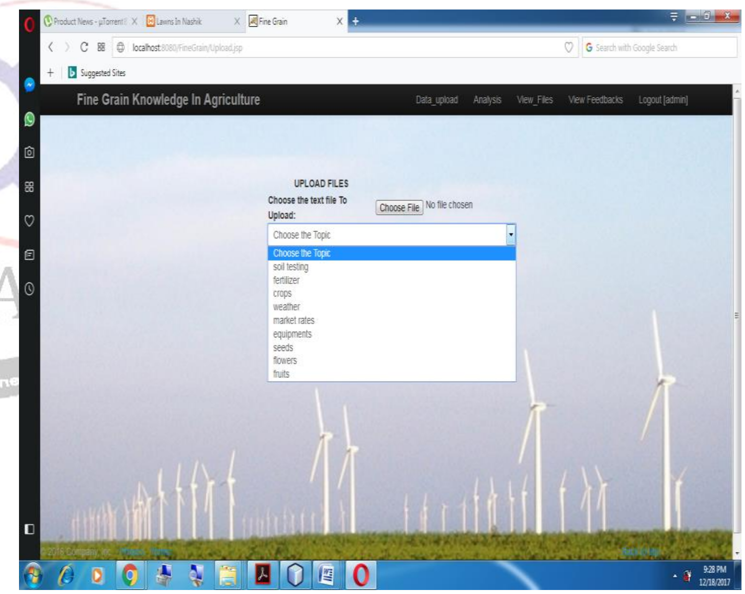

Fig 2.Train Dataset

Train data by our implemented algorithm TF-IDF (Term Frequency –Inverse Document Frequency) admin have authority to train data.

Login admin can train data with selecting any one agriculture category**.**
When Admin train data we send all data to TFIDF algorithm.

We are implemented prediction module on user side. User can input any word relevant to agriculture for find category related to word. We are also getting execution time by our proposed algorithm.

 DOI : 10.18231/2454-9150.2017.0056

1. Give above parameters as input to SVM prediction module of each category
2. Analyze predicted values for each of the category
3. If predicted value == 1 then
4. Add category to category selected list
5. End if
6. Search keyword in category selected list categories

Display **category** associated with each **listed** document.

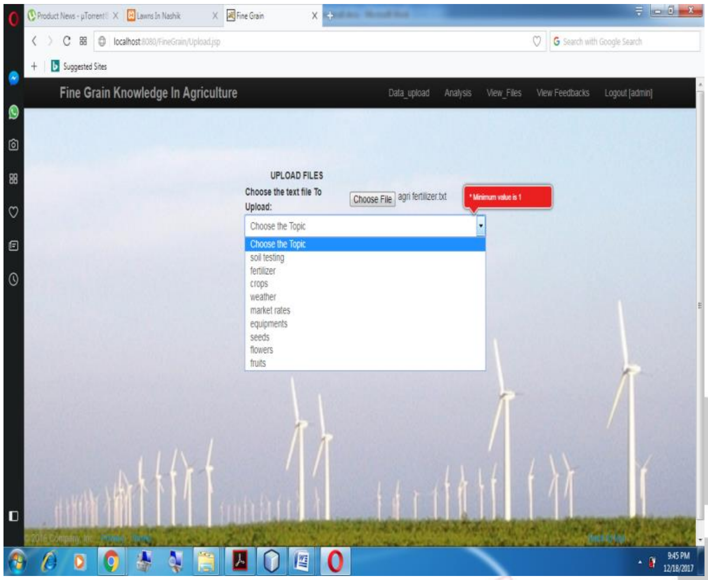

Fig 3.Upload Files

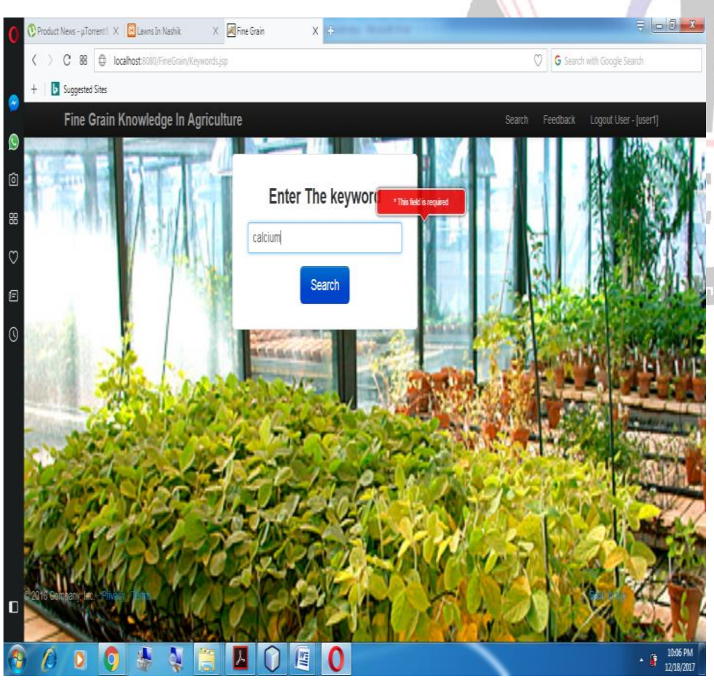

Fig 4.Search Keyword

The keywords search will be identified in terms of above mapping to identify the documents more relative in terms of the click ratio identified by the training model. The filtered documents will be then displayed as the final output of the search mechanism.

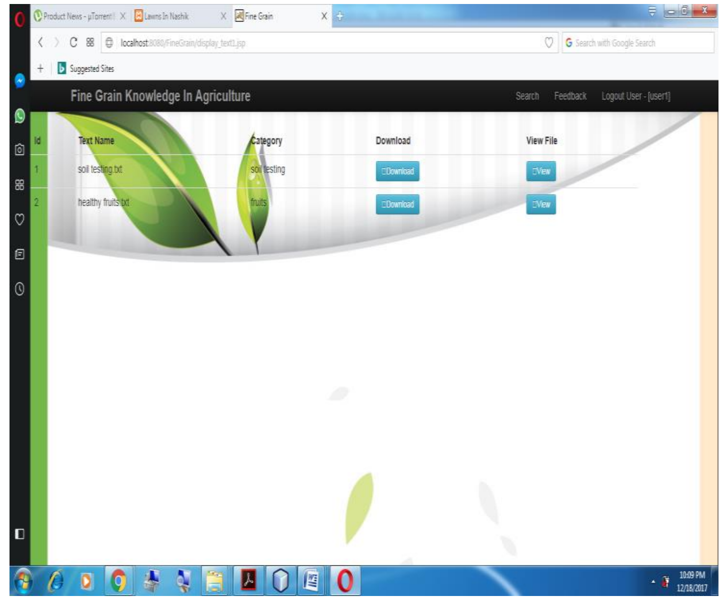

Fig 5.View categories and Files

After input keyword user will get category relevant to that keyword user can also view and download files

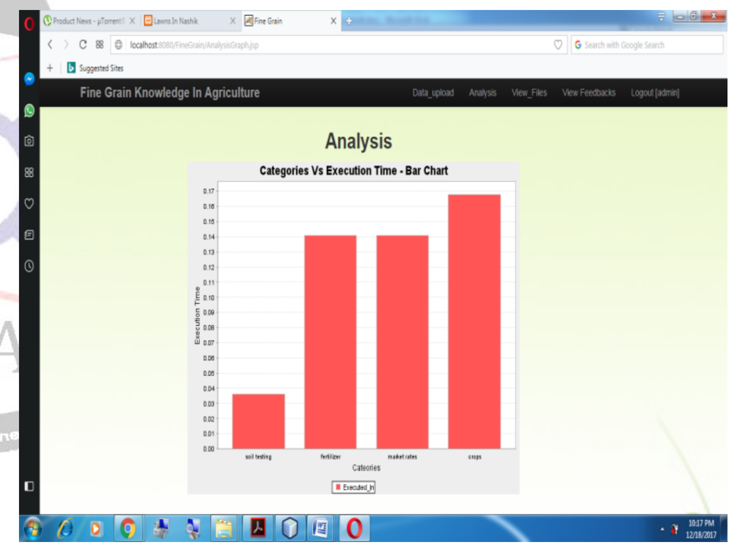

Fig 6.Analysis

Analysis module implemented on admin side for getting execution time for getting word.

V. CONCLUSION

We introduced a novel problem, fine-grained knowledge in Agriculture system, which is desirable in practice. We identified digging out fine grained knowledge reflected by peoples interactions with the outside world as the key to solving this problem. We proposed a two-step framework to mine fine-grained knowledge and integrated it with the classic expert search method for finding right ad-visors. Experiments on real Web surfing data showed encouraging results. There are open issues for this problem. (1) The fine grained knowledge could have a hierarchical structure. For example, Java IO can contain File IO and Network IO as sub-

 DOI : 10.18231/2454-9150.2017.0056

knowledge. We could iteratively apply diHMM on the learned micro-aspects to derive a hierarchy, but how to search over this hierarchy is not a trivial problem. (2) The basic search model can be refined, e.g. incorporating the time factor since people gradually forget as time flows. (3) Privacy is also an issue. In this work, we demonstrate the feasibility of mining task micro-aspects for solving this knowledge sharing problem. We leave these possible improvements to future work.

REFERENCES

[1] K. Balog, L. Azzopardi, and M. de Rijke.*Formal models for expert finding in enterprise corpora*. In SIGIR 2006, pages 43–50.

[2] M. J. Beal, Z. Ghahramani, and C. E. Rasmussen. *The infinite hidden markov model*. In Advances in neural information processing systems 2001, pages 577–584.

[3]D. Blei and M. Jordan. *Variational inference for dirichlet process mixtures*. Bayesian Analysis 2006, 1(1):121–143.

[4]D. M. Blei and J. D. Lafferty.*Dynamic topic models*. In ICML 2006, pages 113–120.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan. *Latent dirichlet allocation*. Journal of machine Learning research 2003.

[6] P. R. Carlile.,*Working knowledge: how organizations manage what they know*. Human Resource Planning 1998, 21(4):58–60.

[7] N. Craswell, A. P. de Vries, and I. Soboroff.*Overview of the trec*2005 enterprise track.In TREC, 2005.

[8] H. Deng, I. King, and M. R. Lyu.*Formal models for expert finding on dblp bibliography data*. In ICDM 2009, pages 163–172.

[9]Y. Fang, L. Si, and A. P. Mathur. *Discriminative models of integrating document evidence and document-candidate associations for expert search*. In SIGIR 2010, pages 683–690.

[10]H. Wang, Y. Song, M.-W. Chang, X. He, R. White, and W. Chu.*Learning to extract cross-session search tasks*. In WWW 2013, pages 1353–1364.

[11] R. White, P. Bailey, and L. Chen.*Predicting user interests from contextual information*. In SIGIR 2009, pages 363–370.

[12] Prof. Dalvir Singh, "*Economic Perspectives of Organic Agriculture: A Review of Literature*", International Journal for Research in Engineering Application & Management (IJREAM), Vol-03, Issue-08, Nov 2017, pp 95-98.

[13] Y. Zhao, G. Karypis, and U. Fayyad.*Hierarchical clustering algorithms for document datasets*. Data Mining and Knowledge Discovery 2005, 10(2):141–168.

 DOI : 10.18231/2454-9150.2017.0056