ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

Ranking Adaption Model for Specific Domain Search

A. Sudhakar
PG Student,
Department of CSE,
Narayana Engineering College, Gudur.

V. Ananth Krishna
HOD,
Department of CSE,
Narayana Engineering College, Gudur.

Abstract:

In this paper, we address these difficulties by proposing a regula rization based algorithm called ranking adaptation SVM (RA-SVM), through which we can adapt an existing ranking model to a new domain, so that the amount of labeled data and the training cost is reduced while the performance is still guaranteed. Our algorithm only requires the prediction from the existing ranking models, rather than their internal representations or the data from auxiliary domains. In addition,we assume that documents similar in the domain-specific feature space should have consistent rankings, and add some constraintsto control the margin and slack variables of RA-SVM adaptively. Finally,ranking adaptability measurement is proposed to quantitativelyestimate if an existing ranking model can be adapted to a new domain. Experiments performed over Letor and two large scale datasets crawled from a commercial search engine demonstrate the applicabilities of the proposed ranking adaptation algorithms and theranking adaptability measurement.

Index Terms:

Information Retrieval, Support Vector Machines, Learning to Rank, Domain Adaptation.

I.INTRODUCTION:

Rank is a kind of learning based infor-mation retrieval techniques, specialized in learning a ranking model with some documents labeled with their relevancies to some queries, where the model is hopefully capable of ranking the documents returned to an arbitrary new query automatically. Based on various machine learning methods, e.g., Ranking SVM [12], [14], RankBoost [9], RankNet [4], ListNet [5], LambdaRank [3], etc., the learning to rank algorithms have already shown their promising performances in information re-trieval, especially Web search.

However, as the emergence of domain-specific search engines, more attentions have moved from the broad-based search to specific verticals, for hunting information constraint to a certain domain. Different vertical search engines deal with different topicalities, document types or domain-specific features. For example, a medical search engine should clearly be specialized in terms of its topical focus, whereas a music, image or video search engine would concern only the documents in particular formats.Since currently the broad-based and vertical search engines are mostly based on text search techniques, the ranking model learned for broad-based can be utilized directly to rank the documents for the verticals. For example, most of current image search engines only utilize the text information accompanying images as the ranking features, such as the term frequency (TF) of query word in image title, anchor text, alternative text, surrounding text, URL and so on.

Therefore, Web images are actually treated as text-based documents that share similar ranking features as the document or Web page ranking, and text based ranking model can be applied here directly. However, the broad-based ranking model is built upon the data from multiple domains, and therefore cannot generalize well for a particular domain with special search intentions. In addition, the broad-based ranking model can only utilize the vertical domain's ranking features that are same to the broad- based domain's for ranking, while the domain-specific features, such as the content features of images, videos or music can not be utilized directly. Those features are generally important for the semantic representation of the documents and should be utilized to build a more robust ranking model for the particular vertical. Alternatively, each vertical can learn its own ranking model independently. However, it's laborious to label sufficient training samples and time-consuming to train different models for various verticals, since the number of verticals is large and increasing drastically.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

Ranking adaptation is closely related to classifier adaptation, which has shown its effectiveness for many learning problems]. However, to the best of our knowledge, there are no prior works on the adaptation for the ranking problem. Besides the general difficulties faced by the classifier adaptation, such as covariate shift (or namely sample se- lection bias) [25], [32] and concept drifting [18], ranking adaptation is comparatively more challenging. Unlike classifier adaptation, which mainly deals with binary targets, ranking adaptation desires to adapt the model which is used to predict the rankings for a collection of documents.

Though the documents are normally labeled with several relevance levels, which seems to be able to be handled by multi-class classification or regression, it is still difficult to directly use classifier adaption for ranking. The reason lies in two-fold: (1) in ranking, the mainly concerned is about the preference of two docu- ments or the ranking of a collection of documents, which is difficult to be modeled by classification or regression. (2) the relevance levels between different domains are sometimes different and need to be aligned. The rest of the paper is organized as follows. In Section 2, we formally present and analyze the proposed ranking adaptation algorithm. Section 3 explores the ranking adaptability. We discuss and formulate the ranking adap- tation with the utilization of domain-specific feature in Section 4. The experimental results are shown and discussed in Section 5. Section 6 analyzes the efficiency problem of the proposed method.

II. RANKING ADAPTATION:

We define the ranking adaptation problem formally as follows: for the target domain, a query set
Q={q1, q2, . . . , qM}
and a document set
D={d1, d2, . . . , dN}
are given. For each queryqiQ, a list of documents
di={di1, di2, . . . , di,n(qi)}
are returned and labeled withthe relevance degrees
yi={yi1, yi2, . . . , yi,n(qi)}
by hu-man annotators. The relevance degree is usually a real
value, i.e.,
yijR, so that different returned documentscan be compared for sorting an ordered list. For each query document pair < qi, dij>, ans-dimensional querydependent feature vector φ(qi, dij)Rsis extracted, e.g.,the term frequency of the query keywordqiin the title,body, URL of the documentdij . Some other hyperlinkbased static rank information is also considered, such a Pagerank [21], HITS [17] and so on. n(qi)denotes thenumber of returned documents for query qi. The targetof learning to rank is to estimate a ranking functionfRsRso that the documentsdcan be ranked for a given query q according to the value of the prediction f(φ(q, d)).n the setting of the proposed ranking adaptation, boththe number of queries m and the number of the returned documents n (qi)in the training set are assumed to besmall.

They are insufficient to learn an effective rankingmodel for the target domain. However, an auxiliaryranking model fa, which is well trained in anotherdomain over the labeled data QaandDa, is available. Itis assumed that the auxiliary ranking mode lfa containsa lot of prior knowledge to rank documents, so it canbe used to act as the base model to be adapted to thenew domain. Few training samples can be sufficient toadapt the ranking model since the prior knowledge isavailable.Before the introduction of our proposed ranking adap-tation algorithm, it's important to review the formulationof Ranking Support Vector Machines (Ranking SVM),which is one of the most effective learning to rank algorithms, and is here employed as the basis of our proposed algorithm.

Ranking SVM:

Similar to the conventional Support Vector Machines (SVM) for the classification problem [27], the motivation of Ranking SVM is to discover a one dimensional linear subspace, where the points can be ordered into the optimal ranking list under some criteria. Thus, the ranking function takes the form of the linear model f(φ(q, d)) =wTφ(q, d), where the bias parameter isignored, because the final ranking list sorted by the prediction fis invariant to the bias. The optimizationproblem for Ranking SVM is defined as follows:

$$
\begin{aligned}
&\min_{f,\xi_{ijk}} \frac{1}{2}||f||^2 + C\sum\nolimits_{i,j,k}\xi_{ijk} \\
&\text{s.t.} \quad f(\phi(q_i, d_{ij})) - f(\phi(q_i, d_{ik})) \geq 1 - \xi_{ijk} \\
&\qquad \xi_{ijk} \geq 0, \\
&\text{for} \quad \forall i \in \{1, 2, \ldots, M\}, \\
&\qquad \forall j \forall k \in \{1, 2, \ldots, n(q_i)\} \text{ with } y_{ij} > y_{ik},
\end{aligned}
$$

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

Ranking Adaptation SVM:

It can be assumed that, if the auxiliary domain and the target domain are related, their respective ranking func-tions Fa and fshould have similar shapes in the functionspace RsR. Under such an assumption,faactuallyprovides a priorknowledge for the distribution of f in its parameter space. The conventional regularization framework, such aLp -norm regularization, manifold regularization [1] designed for SVM [27], regularized neural network [11] and so on, shows that the solution of an ill-posed problem can be approximated from vari- ational principle, which contains both the data and the prior assumption [11]. Consequently, we can adapt the regularization framework which utilizes the Fa as the prior information, so that the ill-posed problem in the target domain, where only few query document pairs are labeled, can be solved elegantly. By modeling our assumption into the regularization term, the learning problem of Ranking Adaptation SVM (RA-SVM) can be formulated as:

$$\begin{aligned} &\min_{f,\xi_{ijk}} \frac{1-\delta}{2}||f||^2 + \frac{\delta}{2}||f - f^a||^2 + C\sum\nolimits_{i,j,k}\xi_{ijk} \\ &\text{s.t.} \quad f(\phi(q_i, d_{ij})) - f(\phi(q_i, d_{ik})) \geq 1 - \xi_{ijk} \\ &\qquad \xi_{ijk} \geq 0, \\ &\text{for} \quad \forall i \in \{1, 2, \ldots, M\}, \\ &\qquad \forall j \forall k \in \{1, 2, \ldots, n(q_i)\} \text{ with } y_{ij} > y_{ik}. \end{aligned}$$

Optimization Methods:

To optimize Problem (2), we briefly denote xijk=φ(qi, dij)−φ(qi, dik)and introduce the Lagrange multi-pliers to integrate the constraints of (2) into the objective function, which results in the primal problem:

$$\begin{aligned} L_P = &\frac{1-\delta}{2}||f||^2 + \frac{\delta}{2}||f - f^a||^2 + C\sum_{i,j,k}\xi_{ijk} \\ &- \sum_{i,j,k}\mu_{ijk}\xi_{ijk} - \sum_{i,j,k}\alpha_{ijk}(f(\mathbf{x}_{ijk}) - 1 + \xi_{ijk})). \end{aligned}$$

Discussions:

The proposed RA-SVM has several advantages, which makes our algorithm highly applicable and flexible when applied to the practical applications. We'll give more discussions of the characteristics of RA-SVM in the following.

Reducing the computational cost: It has been shown that our ranking adaptation algorithm can be transformed into a Quadratic Programming (QP) prob- lem, with the learning complexity directly related to the number of labeled samples in the target domain. Platt [22] proposed the sequential minimal opti- mization (SMO) algorithm which can decompose a large QP problem into a series of subproblems and optimize them iteratively. The time complexity is around O(n2.3) for general kernels [22]. In [15],cutting-plane method is adopted to solve SVM forthe linear kernel, which further reduces the timecomplexity toO(n).

Adaptation from Multiple Domains:

Our proposed RA-SVM can be extended to a more general setting, where ranking models learned from multiple domains are provided. Denoting the set of auxiliary ranking functions by
F={fa1, fa2, . . . , faR}, the RA-
SVM for the multiple domain adaptation setting can be formulated as:

$$\begin{aligned} &\min_{f,\xi_{ijk}} \frac{1-\delta}{2}||f||^2 + \frac{\delta}{2}\sum\nolimits_{r=1}^{R}\theta_r||f - f_r^a||^2 + C\sum\nolimits_{i,j,k}\xi_{ijk} \\ &\text{s.t.} \quad f(\phi(q_i, d_{ij})) - f(\phi(q_i, d_{ik})) \geq 1 - \xi_{ijk} \\ &\qquad \xi_{ijk} \geq 0, \\ &\text{for} \quad \forall i \in \{1, 2, \ldots, M\}, \\ &\forall j \forall k \in \{1, 2, \ldots, n(q_i)\} \text{ with } y_{ij} > y_{ik}, \end{aligned}$$

III. EXPLORE RANKING ADAPTABILITY:

Though the ranking adaptation can mostly provide ben- efits for learning a new model, it can be argued that when the data from auxiliary and target domains sharem little common knowledge, the auxiliary ranking model can provide little help or even negative influence, to the ranking of the documents in the target domain. Consequently, it is imperative to develop a measure for quantitatively estimating the adaptability of the auxilia ry model to the target domain.

However, given a ranking model and a dataset collected for a particular target domain, it's nontrivial to measure their correlations di-rectly, because neither the distribution of the ranking model nor that of the labeled samples in the target domain is trivial to be estimated. Thus, we present some analysis on the properties of the auxiliary model, based on which the definition of the proposed ranking adaptability is presented.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

Auxiliary Model Analysis:

We analyze the effects of auxiliary models through the loss constraint in the formulation of our RA-SVM. By substituting (4) into (2), we can obtain that:

$$\delta f^a(\mathbf{x}_{ijk}) + \Delta f(\mathbf{x}_{ijk}) \geq 1 - \xi_{ijk}$$
$$\text{with } y_{ij} > y_{ik}, \text{ and } \xi_{ijk} \geq 0.$$

Ranking Adaptability:

Based on the above analysis of f a , we develop theranking adaptability measurement by investigating the correlation between two ranking lists of a labeled query in the target domain, i.e., the one predicted by fa and the ground-truth one labeled by human judges. Intuitively, if the two ranking lists have high positive correlation, the auxiliary ranking model Fa is coincided with the distribution of the corresponding labeled data, therefore we can believe that it possesses high ranking adaptabil-

ity towards the target domain, and vice versa. This is because the labeled queries are actually randomly sampled from the target domain for the model adaptation, and can reflect the distribution of the data in the target domain.

IV RANKING ADAPTATION WITH DOMAIN-SPECIFIC FEATURE:

Conventionally, data from different domains are also characterized by some domain-specific features, e.g., when we adopt the ranking model learned from the Web page search domain to the image search domain, the image content can provide additional information to facilitate the text based ranking model adaptation. In this section, we discuss how to utilize these domain-specific features, which are usually difficult to translate to textual representations directly, to further boost the performance of the proposed RA-SVM.

Margin Rescaling Margin rescaling denotes that we rescale the margin violation adaptively according to their similarities in thedomain-specific feature space. Specifically, the RankingAdaptation SVM with Margin Rescaling (RA-SVM-MR)can be defined as the following optimization problem:

$$\begin{aligned}
\min_{f,\xi_{ijk}} \ & \frac{1-\delta}{2}\|f\|^2 + \frac{\delta}{2}\|f - f^a\|^2 + C\sum\nolimits_{i,j,k}\xi_{ijk} \\
\text{s.t.} \ & f(\phi(q_i, d_{ij})) - f(\phi(q_i, d_{ik})) \geq 1 - \xi_{ijk} - \sigma_{ijk} \\
& \xi_{ijk} \geq 0, \\
\text{for} \ & \forall i \in \{1, 2, \ldots, M\}, \\
& \forall j \forall k \in \{1, 2, \ldots, n(q_i)\} \text{ with } y_{ij} > y_{ik}.
\end{aligned}$$

V. EXPERIMENTS:

In this section, we perform several experiments under two different settings, to demonstrate the effectiveness of the proposed RA-SVM based algorithms and the ranking adaptability measurement.

Datasets and Evaluation Measure:

We firstly conduct the experiments over the Letor bench- mark dataset [20], and adapt the ranking model learned from TD2003 dataset to the ranking of TD2004 dataset. Letor TD2003 and TD2004 datasets are gathered from the topic distillation task of TREC 2003 and TREC 2004, with 50 queries for TD2003 and 75 ones for TD2004. The doc- uments are collected by crawling from the .gov domain. For each query, about 1000 associated documents are re- turned, and labeled with a binary judgment, i.e., relevant or irrelevant.

The features of TD2003 and TD2004 include the low-level features such as term frequency, inverse document frequency, and document length, as well as high-level features such as BM25, LMIR, PageRank, and HITS, for totally 44 dimensional features. However, Letor is a comparatively small dataset, and each document is only labeled with a binary relevance degree, which cannot reflect the practical Web search scenarios with multiple relevance degrees. Also, there are no domain-specific features for the target domain data, where we cannot demonstrate the effectiveness of the proposed ranking adaptation with domain-specific feature algorithms.

TABLE 1
Ranking Adaptation Dataset Information.

Dataset	#Query	#Query-Document	Relevance Degree	Feature Dimension
TD2003	50	49171	2	44
TD2004	75	74170	2	44
Web Page Search	2625	122815	5	354
Image Search	1491	71246	3	354

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

Experiment Settings:

We build the auxiliary ranking model by training Ranking SVM with different parameters over some labeled queries randomly sampled from the auxiliary domain, namely Letor TD2003 dataset and Web page search-dataset, and then select the models that are best performed over the remained data in the auxiliary domain as the auxiliary models for adaptation.

In the adaptation target domain, where the performance of different algorithms are reported, we randomly select several queries as the pool of the labeled data as candidate data for the adaptation, several queries as the validation set todetermine the parameters of different algorithms, and the remaining queries as the test set for the performance evaluation.

We vary the size of adaptation set gradually by selecting different number of queries from the pool of the labeled data, so that we can see the influence of different numbers of labeled samples to the perfor- mance of the adapted ranking model.

For each size of adaptation set, we generate five different adaptation sets by randomly sampling from the labeled adaptation data pool created before. We apply each algorithm over each generated set separately, resulting into five different ranking models. The final performance reported in this paper is the average results of the five ranking models, validated over the identical validation set and evaluated over the identical test set.

Fig. 1. The MAP of TD2003 to TD2004 adaptation results, with 5 and 10 labeled queries in TD2004 respectively.

Adapt from Web Page Search to Image Search To further demonstrate the effectiveness of the proposed RA-SVM algorithm, we perform several experiments by adapting the ranking model trained from Web page search domain to the image search domain. The performances with 5, 10, 20, 30, 40 and 50 labeled queries are shown in respectively. It can be observed that, at each adaptation size, RA-SVM consis- tently outperforms the baseline methods significantly at all truncation levels, while RA-SVM-MR and RA-SVM- SR further improve the performance. In addition, we can derive that for the 5, 10 and 20 queries settings, the performance of Aux-Only model is much better than Tar-only one, because of the insufficient labeled sample problem. On the contrary, for the 40 and 50 queries settings, Tar-only model performs better than Aux-Only one, due to the larger size of training set and the limited performance of the auxiliary model caused by the domain differences.

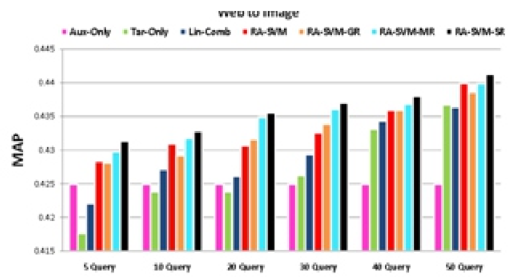

Fig.2.The MAP of Web page search to image search adaptation results, with 5, 10, 20, 30, 40, and 50 labeled queries of image search dataset utilized respectively.

Ranking Adaptability:

In this subsection, we perform several experiments to prove the effectiveness of the proposed ranking adaptability , and the applicability for auxiliary model selection. Firstly, ten ranking models are learned over the train- ing set of the auxiliary domain, i.e., the TD2003 and the Web page search domain respectively, with the same training set used for the experiments and Table 1. We still adopt Ranking SVM to learn the ranking models as the candidate auxiliary models. The ten mod- els are learned by varying the parameter C of Ranking SVM. Then, we apply each model respectively to the target domain for adaptation experiments, using our RA-SVM, RA-SVM-MR and RA-SVM-SR. Finally, according to (12), the ranking adaptabilities of all the

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

models over the adaptation sets from image search domain are cal- culated. The performances and the ranking adaptabilities to be reported are averaged over the five random splits of adaptation sets. To be concise, we only show the results on the adaptation set composed of five labeled queries for TD2004 dataset and twenty labeled queries for image search dataset, while the results of other sizes of adaptation sets are similar.

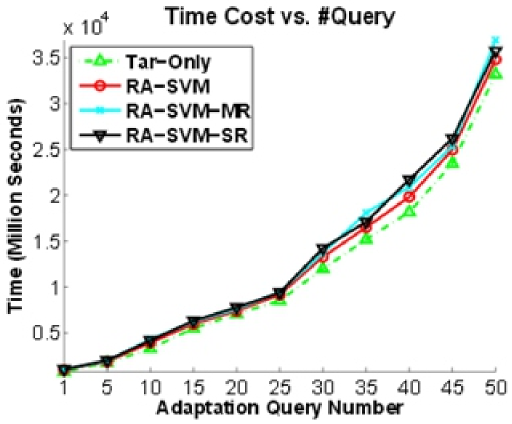

Fig.3.The time cost of each method under different sizes of adaptation sets.

VI. EFFICIENCY ANALYSIS:

To analyze the efficiency of the proposed RA-SVM basedmethods, we compare the learning time of different methods by varying the adaptation query number in the Web page search to image search setting. Because the Aux-Only will not spend time learning a new ranking model, and Lin-Comb needs the Tar-Only to be trained beforehand and then linearly combines it with Aux-Only, we only compare the time cost of Tar-Only, RA-SVM, RA-SVM-MR and RA-SVM-SR. The time reported for each method is the summation of the five random splits. All the experiments are done under the same hardware setting, i.e., the Intel Xeon E5440 core with 8GB memory.The results are shown, and we can observe that for small number of adaptation query number, the time costs of different algorithms are very similar. For large adaptation sets, even though Tar-Only is slightly better than RA-SVM based methods, the variance of different methods is not significant. We can conclude that the proposed RA-SVM is quite efficient compared with direct training a model in the target domain.

Also, the results of RA-SVM-MR and RA-SVM-SR show that the incorporation of domain-specific features doesn't brings further learning complexity. These conclusions are consistent with our theoretical analysis mentioned in the previous sections.

CONCLUSION:

As various vertical search engines emerge and the amount of verticals increases dramatically, a global rank- ing model, which is trained over a dataset sourced from multiple domains, cannot give a sound performance for each specific domain with special topicalities, doc- ument formats and domain-specific features. Building one model for each vertical domain is both laborious for labeling the data and time-consuming for learning the model.

In this paper, we propose the ranking model adaptation, to adapt the well learned models from the broad-based search or any other auxiliary domains to a new target domain. By model adaptation, only a small number of samples need to be labeled, and the Based on the regularization framework, the Ranking Adaptation SVM (RA-SVM) algorithm is proposed,which performs adaptation in a black-box way, i.e., only the relevance predication of the auxiliary ranking models is needed for the adaptation.

Based on RA- SVM, two variations called RA-SVM margin rescaling (RA-SVM-MR) and RA-SVM slack rescaling (RA-SVM- SR) are proposed to utilize the domain specific features to further facilitate the adaptation, by assuming that similar documents should have consistent rankings, and constraining the margin and loss of RA-SVM adaptively according to their similarities in the domain-specific feature space.

Furthermore, we propose ranking adaptability, to quantitatively measure whether an auxiliary model can be adapted to a specific target domain and how much assistance it can provide.

REFERENCES:

[1].M. Belkin, P. Niyogi, and V. Sindhwani. Manifold regular ization: A geometric framework for learning from labeled and unlabel ed examples. Journal of Machine Learning Research , 7(Nov):2399–2434, 2006.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

[2] J. Blitzer, R. Mcdonald, and F. Pereira. Domain adaptation with structural correspondence learning. In EMNLP '06: Proceedings of the 2006 Conference on Empirical Methods in Natural Languag e Pro- cessing , pages 120–128. Association for Computational Linguistic s, July 2006.

[3] C. J. C. Burges, R. Ragno, and Q. V. Le. Learning to rank with nonsmooth cost functions. In NIPS '06: Advances in Neural Information Processing Systems , pages 193–200. MIT Press, Cambridge,MA, 2006.

[4] C. J. C. Burges, T. Shaked, E. Renshaw, A. Lazier, M. Deeds, N. Hamilton, and G. Hullender. Learning to rank using gradie nt descent. In ICML '05: Proceedings of the 22th International Conference on Machine Learning , 2005.

[5] Z. Cao and T. yan Liu. Learning to rank: From pairwise appr oach to listwise approach. In ICML '07: Proceedings of the 24th Interna- tional Conference on Machine Learning, pages 129–136, 2007.

[6] J. Cui, F. Wen, and X. Tang. Real time google and live image search re-ranking. In ACM Multimedia , pages 729–732, 2008.

[7] W. Dai, Q. Yang, G.-R. Xue, and Y. Yu. Boosting for transfer learning. In ICML '07: Proceedings of the 24th international conference on Machine learning , pages 193–200, 2007.

[8] H. Daume, III and D. Marcu. Domain adaptation for statistical classifiers. Journal of Artficial Intelligence Research , 26:101–126, 2006.

[9] Y. Freund, R. Iyer, R. E. Schapire, Y. Singer, and G. Diett erich. An efficient boosting algorithm for combining preferences. Journal of Machine Learning Research , 4:933–969, 2003.

[10] B. Geng, L. Yang, C. Xu, and X.-S. Hua. Ranking model adaptation for domain-specific search. In CIKM '09: Proceeding of the 18th ACM conference on Information and knowledge management , pages 197–206, 2009.

[11] F. Girosi, M. Jones, and T. Poggio. Regularization theo ry and neural networks architectures. Neural Computation, 7:219–269,1995.

[12] R. Herbrich, T. Graepel, and K. Obermayer. Large margin rank boundaries for ordinal regression. In Advances in Large Margin Classifiers , 2000.

[13] K. J¨arvelin and J. Kek¨al¨ainen. Ir evaluation methods for retrieving highly relevant documents. In SIGIR '00: Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval , pages 41–48, 2000.

[14] T. Joachims. Optimizing search engines using clickthr ough data. In KDD '02: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining , pages 133–142, 2002.

[15] T. Joachims. Training linear svms in linear time. In KDD '06: Proceedings of the 12th ACM SIGKDD international conference onKnowledge discovery and data mining, pages 217–226, 2006.

[16] M. G. Kendall. A new measure of rank correlation. Biometrika,30(1/2):81–93, June 1938.

[17] J. M. Kleinberg, S. R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. The web as a graph: Measurements, models and methods. In Proceedings of the International Conference on Combina- torics and Computing, no. 1627 in Lecture Notes in Computer SCience,pages 1–18, 1999.

[18] R. Klinkenberg and T. Joachims. Detecting concept drift with support vector machines. In ICML '00: In Proceedings of the Seventeenth International Conference on Machine Learning , pages 487–494. Morgan Kaufmann, 2000.

[19] J. Lafferty and C. Zhai. Document language models, query models, and risk minimization for information retrieval. In SIGIR '01: Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval, pages 111–119, 2001.

[20] T.-Y. Liu, T. Qin, J. Xu, W. Xiong, and H. Li. Benchmark dataset for research on learning to rank for information retrieval. InSIGIR'07: Proceedings of SIGIR 2007 Workshop on Learning to Rank for Information Retrieval (LR4IR 2007) , 2007.

[21] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. In Technical report, Stanford University , 1998.

[22] J. C. Platt. Fast training of support vector machines usIng sequential minimal optimization. Advances in kernel methods: support vector learning , pages 185–208, 1999.

[23] J. M. Ponte and W. B. Croft. A language modeling approach to information retrieval. InSIGIR '98: Proceedings of the 21st annual international ACM SIGIR conference on Research and develop ment in information retrieval, pages 275–281, 1998.

[24] S. Robertson and D. A. Hull. The trec-9 filtering track final report.In Proceedings of the 9th Text Retrieval Conference , pages 25–40, 2000.

[25] H. Shimodaira. Improving predictive inference under cOvariate shift by weighting the log-likelihood function. Journal of Statistical Planning and Inference, 90:227–244(18), 2000.

Author Details:

Mr. A. Sudhakar is a student of Narayana Engineering College, Gudur, presently he is pursuing M.Tech (C.S) and his project on Ranking Adaption Model for Specific Domain Search.

Mr. V. Ananth Krishna, excellent teacher. He received M.Tech, Ph.D. and working as H.O.D for the department of Computer Science & Engineering in Narayana Engineering College Gudur.