New Generation Computing, 26(2008)49-61
Ohmsha, Ltd. and Springer

SVM-based Interactive Document Retrieval with Active Learning

Takashi ONODA and Hiroshi MURATA
Central Research Institute of Electric Power Industry
2-11-1, Iwado Kita, Komae, Tokyo 201-8511 JAPAN
{onoda, murata}@criepi.denken.or.jp
Seiji YAMADA
National Institute of Informatics, SOKENDAI
2-1-1, Hitotsubashi, Chiyoda, Tokyo 101-8430 JAPAN
seiji@nii.ac.jp

Received 20 March 2007
Revised Manuscript received 22 July 2007

Abstract This paper describes an application of SVM (Support Vector Machines) to interactive document retrieval using active learning. Some works have been done to apply classification learning like SVM to relevance feedback and have obtained successful results. However they did not fully utilize characteristic of example distribution in document retrieval. We propose heuristics to bias document showing for user's judgement according to distribution of examples in document retrieval. This heuristics is executed by selecting examples to show a user in neighbors of positive support vectors, and it improves learning efficiency. We implemented a SVM-based interactive document retrieval system using our proposed heuristics, and compared it with conventional systems like Rocchio-based system and a SVM-based system without the heuristics. We conducted systematic experiments using large data sets including over 500,000 newspaper articles and confirmed our system outperformed other ones.

Keywords: Document Retrieval, Relevance Feedback, Support Vector Machines, Active Learning.

§1 Introduction

As Internet technology progresses, accessible information by end users is explosively increasing. In this situation, we can now easily access a huge doc-

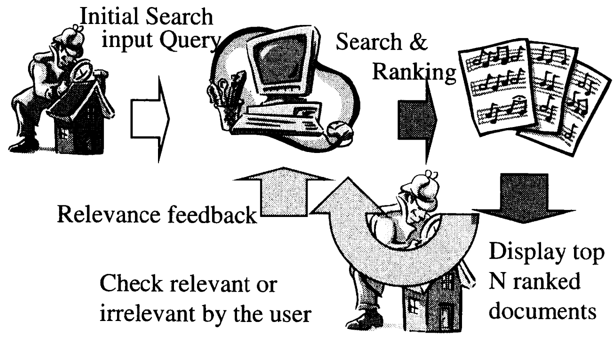


Fig. 1 Overview of Interactive Document Retrieval.

tive IR problem, we deal with not only such learning performance, but also the user's cost to judge displayed documents.

Also some studies applied SVM to the information retrieval problems.[2,19] Tsuge at el.[19] considered the relevance feedback with only one feedback iteration. Because of the single feedback, they did not attack the problem of how to select displayed documents to a user. In contrast, our study will consider the relevance feedback with some feedback iterations and selection of displayed documents. Drucker at el.[2] did not evaluate the useful selection rule for displayed documents at each iteration to a user. In contrast with it, we are interested in what is the most useful selection rule for displayed documents at each iteration to the user.

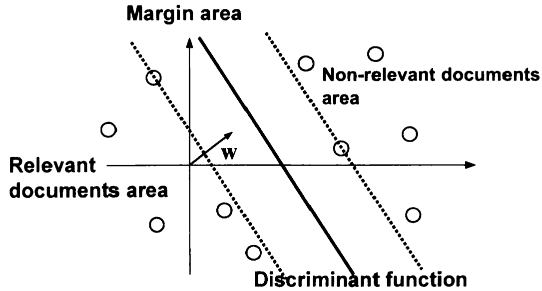
§2 Active Learning with SVM in Interactive Document Retrieval

In this section, we describe the information retrieval system using relevance feedback with SVM from an active learning point of view. Active learning is one of machine learning frameworks in which a leaner actively and selectively obtains useful information like training data from an environment like a teacher. The methodologies of active learning are categorized into two groups: to select the most effective data for pruning of a hypothesis space, and to select data with the most different classes by multiple learners. The former is based on a version space paradigm[17] and expected entropy.[13] The latter is based on the Query By Committee method.[7] We employee the hypothesis space based method in this research, and try to select the most useful data to efficiently obtain the precise discriminant function.

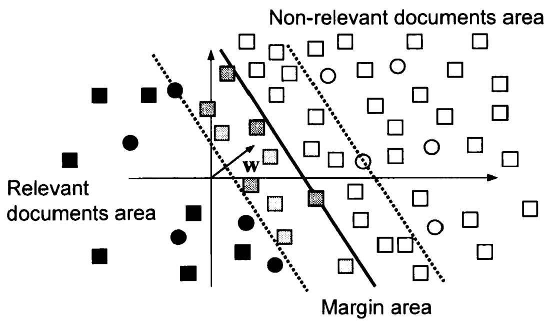
Figure. 1 shows the concept of the interactive document retrieval with relevance feedback. In Fig. 1, the iterative procedure is the gray arrows parts. The SVM have a great ability to discriminate even if the training data is small. Consequently, we have proposed to apply SVM as the classifier in the relevance feedback method.[11] The retrieval steps of the proposed method perform as following procedure.

Step 1: Initial search

The conventional information retrieval system based on vector space



(a)



(b)

Fig. 2 The Discriminant Function and the Displayed Documents.

model displays the top N ranked documents along with a request query to a user. In our method, the top N ranked documents are selected and displayed by using cosine distance between the request query vector and each document vector for the first feedback iteration.

Step 2: Judgment of documents by a user
A user then evaluates these N displayed documents and classifies them into relevant documents or non-relevant documents. After the user's evaluation, the relevant documents have relevance label and the non-relevant documents have non-relevance label. For example, the relevant documents have "+1" label and the non-relevant documents have "-1" label generally, after the user's judgment.

Step 3: Determination of the optimal hyper-plane
The optimal hyper-plane for classifying relevant and non-relevant documents is generated by using a SVM with judged(labeled) documents (see Fig. 2(a)).

Step 4: Selection of displayed documents
The documents, which are retrieved in the Step1, are mapped into the feature space that has the optimal hyper-plane as a discriminant function. The SVM learned by the previous step classifies the non-judged(unlabeled) documents into relevant or non-relevant ones.

ument database through the Web. Although various Web services to retrieve useful information in the Web have been developed and provided to users, the accuracy and efficiency of such systems and services are not sufficient yet. Thus we need to develop a fundamental framework of information retrieval (IR) systems to satisfy such huge needs in the Web. Especially, we consider an interactive IR system is promising because it provides solution to overcome the limitation of a stand-alone and fully automatic IR system.

In general, it is hard for a user to retrieve relevant documents from which he/she can obtain useful information, and a lot of studies have been done in information retrieval, especially document retrieval.[23] Active works for such document retrieval have been reported in TREC (Text Retrieval Conference)[18] for English documents, IREX (Information Retrieval and Extraction Exercise)[5] and NTCIR (NII-NACSIS Test Collection for Information Retrieval System)[8] for Japanese documents.

In most frameworks for information retrieval, a Vector Space Model (called *VSM*) in which a document is described with a high-dimensional vector (called a *document vector*) of term occurrence frequencies is used. [14] An information retrieval system using a vector space model computes the similarity between a query vector and document vectors by cosine of the two document vectors and indicates a user a list of retrieved documents (called a *hit list*).

Since a user hardly describes a precise query in the first trial, interactive approach to modify the query vector by evaluation of the user on documents in a hit list. This method is called *relevance feedback* [12] and used widely in information retrieval systems. In this method, a user directly judges whether a document is relevant or non-relevant in a hit list, and a system modifies the query vector using the user evaluation. A traditional way to modify a query vector is a simple learning rule to reduce the difference between the query vector and document vectors judged as relevant by a user.

In another approach, relevant and non-relevant document vectors are considered as positive and negative examples, and relevance feedback is transposed to a binary classification problem. [9] For the binary classification problem, Support Vector Machines (SVM)[20][21] have shown the excellent ability. Also some studies applied SVM to the text classification problems [17] and the information retrieval problems.[2] Now, we are interested in what is the most useful selection rule for displayed documents in a hit list at each iteration to a user. In this paper, we propose a novel selection rule of displayed documents at each iteration, and then show the comparison results of the effectiveness for the document retrieval.

Many studies have applied SVM to the text classification problems. [3,4,6,17] Although this text classification problem is similar to the interactive IR problem, there is an important difference between them. In the text classification problem, the performance of the text classification like accuracy, precision and recall of learned clusters should be maximized. In these studies,[3,4,6,17] they assume many judged documents are given in a training set, and they are interested in maximizing only the learning performance on the entire test set. In the interac-

Then the system selects the documents by heuristics based on the distance from the optimal hyper-plane and the distribution of the relevant and non-relevant documents. The feature of this distribution can be used as a prior knowledge. The detail of proposed heuristics for the selection rules are described in the next section. The top N ranked documents, which are ranked using the heuristics, are displayed to the user as the information retrieval results of the system.

If the number of feedback iterations is more than m, then go to next step. Otherwise, return to Step 2. The m is a maximal number of feedback iterations and is given by the user.

Step 5: Display of the final retrieved documents

All the documents are ranked by the distance between the documents and the hyper-plane which is the discriminant function determined by SVM. According to this rank, the order of the displayed documents is determined (see Fig. 2(b), in which black, gray squires and circles are considered as relevant documents. Then they are ranked based on the distance from the discriminant hyperplane by SVM.).

§3 Selection Heuristics of Displayed Documents

In general interactive document retrieval, a human user judges the displayed documents and gives feedback to a system repeatedly. The system should obtain relevant documents as much as possible, not only after all feedbacks, but also when the documents are displayed to a user for judge. Thus the following two preconditions should be satisfied in such interactive documents retrieval.

1. *Learning efficiency*: In order to learn classifier efficiently, a system needs to display the near-miss documents for user's judgment.
2. *Retrieval efficiency*: The displayed documents should include relevant documents as much as possible.

With SVM as a classification learner, it satisfies the precondition (1) to select the nearest documents to the learned discriminant function as displayed documents. This has been called a "Simple" method[17)] in active learning. For example, Warmuth et al. reported the successful results to make generation of novel drug more efficient[22)]. This *Simple* method is described graphically like Fig.3, where the black squares are selected as the displayed documents.

In contrast to the simple method, we propose heuristics to *select the displayed documents to a user for judgment.* The heuristics is to select the nearest documents to the region of relevant documents in the margin of SVM. This heuristics is described graphically like Fig.4, where the black squares are selected as the displayed documents.

This *Simple* method assumes the learned discriminant function before convergence of learning is a good approximation to the final optimal one. However, the assumption hardly holds in the problem of interactive document retrieval in which few relevant documents and huge non-relevant documents exist. For example, Fig. 5 shows the distribution of documents vectors for the distance

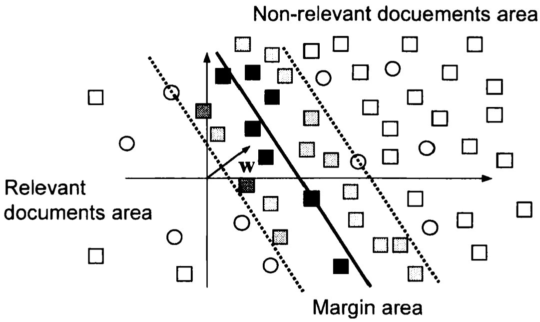


Fig. 3 The Conventional *Simple* Method to Select Displayed Documents.

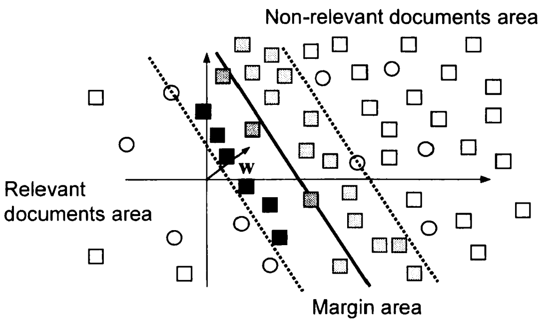


Fig. 4 The Proposed Selection Heuristics for Displayed Documents.

from the learned discriminant function from first user's feedback in a data set of §4. For almost all topics, such distributions were similar to the Fig. 5. In this figure, the x-axis indicates the document vector's distance from the learned discriminant hyperplane, and $x = 0$ corresponds to the discriminant hyperplane itself. Also $x = 1$ and $x = -1$ indicate the boundary hyperplane of relevant documents and that of non-relevant documents, respectively. The y-axis stands for the number of *non-relevant* document vectors in every window with 0.01 width of distance from the discriminant hyperplane. The histogram under the x-axis indicates the number of *relevant* document vectors.

As seeing from Fig. 5, most of the not-judged documents are in the margin of the two boundary hyperplanes, and non-relevant documents exist not only in the margin of relevant documents, but also in the region of relevant documents having the x value over 1. We focus that the number of non-relevant documents are sufficiently small in the the neighbor of x =1, and the number is comparative to the relevant documents in that region. Thus, by applying our heuristics: *to select the nearest documents to the region of relevant documents in the margin of SVM*, we can obtain displayed documents that may include near-miss data for both of relevant and non-relevant documents, and the heuristics is considered to

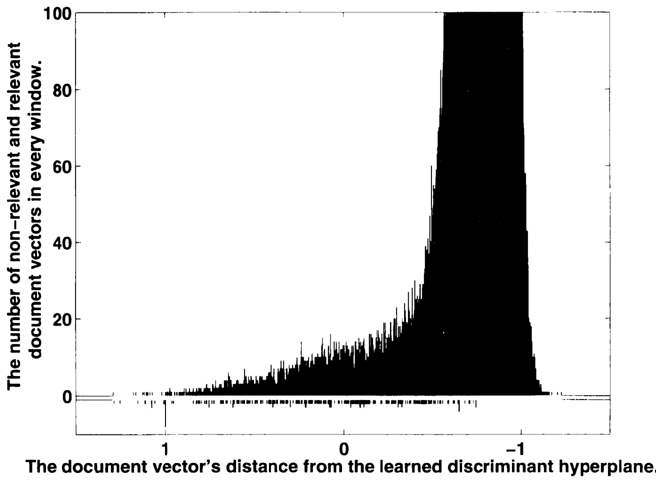


Fig. 5 The Distribution of Document Vectors for a Discriminant Function Learned by One Feedback.

satisfy the precondition (1).

Furthermore, there is the high possibility to display relevant documents by selecting the near documents to the boundary hyperplane of relevant documents in our heuristics. Hence, we can expect the precondition (2) may be satisfied. Thus, we consider that our heuristics satisfy the preconditions (1) and (2). The effectiveness of our heuristics will be experimentally verified in the next section.

§4 Experiments

4.1 Experimental Setting

We conducted experiments to evaluate the effectiveness of our proposed heuristics for active learning of SVM. The document data set we used is a set of newspaper articles in ad hoc task which was widely used in the document retrieval conference 6th, 7th and 8th TREC.[18] The data set has about 530 thousands articles. Each TREC provides 50 retrieval problems (called *topics*) and the information of relevant documents for each retrieval problem.

In our experiments, 150 topics generated by merging each 50 topics of the three TRECs are tested. The topics did not include the same topics. Each topic has three tags, which consist of a title tag, a description tag, and a narrative tag. The title tag has 2 or 3 terms to describe the topic. The description tag introduces the topic, and the narrative tag reports the topic. Our experiments used 2 or 3 terms of the title tag as a query. Also our experiments removed the stopwords and made stemming for documents and queries.

We used TFIDF,[23] which is one of the most popular methods in information retrieval to generate document feature vectors, and the concrete equation[15] of a weight of a term t in a document d w_t^d are in the following.

$$
\begin{aligned}
w_t^d &= L \times t \times u \\
L &= \frac{1 + \log(tf(t,d))}{1 + \log(\text{average of } tf(t,d) \text{ in } d)} \\
t &= \log\left(\frac{n+1}{df(t)}\right) \\
u &= \frac{1}{0.8 + 0.2\frac{uniq(d)}{\text{average of } uniq(d)}}
\end{aligned}
$$

The notations in these equation denote as follows.

1. w_t^d is a weight of a term t in a document d.
2. $tf(t,d)$ is a frequency of a term t in a document d.
3. n is the total number of documents in a document set.
4. $df(t)$ is the number of documents including a term t.
5. $uniq(d)$ is the number of different terms in a document d.

The size N of retrieved and displayed documents in **Step 1** in §2 was set as 10 and 20. We fixed the user's load by limiting the number of judged documents to fifty documents. Thus the feedback iterations m were limited to 5 and 2 so that the total amount of judged documents does not exceed fifty.

In our experiments, we used the linear kernel for SVM, and found a discriminant function for the SVM classifier in this feature space. Since the VSM of documents is high dimensional space, we do not need to use the kernel trick and the regularization parameter.[1]

For comparison with our approach, two information retrieval methods were adopted. The first is an information retrieval method that uses the selection rule *Simple*, which is described in §3. The second is an information retrieval method using conventional Rocchio-based relevance feedback[12] which is widely used in information retrieval research.

The Rocchio-based relevance feedback modifies a query vector Q_i by evaluation of a user using the following equation.

$$Q_{i+1} = Q_i + \alpha \sum_{x \in R_r} x - \beta \sum_{x \in R_n} x \qquad (1)$$

where R_r is a set of documents which were judged as relevant documents by a user at the ith feedback, and R_n is a set of documents which were judged as non-relevant documents at the ith feedback. α and β are weights for relevant and non-relevant documents respectively. In this experiment, we set $\alpha = 1.0$, $\beta = 0.5$ which are known adequate experimentally.

In order to compare the usefulness of our proposed method with the other methods, we evaluated the following criteria which are widely used in information retrieval.[23]

P10: Precision within the top 10 documents, which is a ratio of relevant documents in the top 10 documents.

P30: Precision within the top 30 documents, which is a ratio of relevant documents in the top 30 documents.

MAP: The average value of all precisions at every relevant document for a topic. When a relevant document is not retrieved, this value is 0.

R05P: Recall when precision first becomes less than 0.5 from the top document. 10 documents need to be checked at least.

P10 and P30 evaluate the precision of the document retrieval method and are the user oriented measures for the retrieval efficiency in §3. MAP and R05P evaluate precision and recall of the document retrieval method, and are the system oriented measures for the learning efficiency in §3.

4.2 Experimental Results

Table 1 shows the experimental results when the number of displayed documents at each iteration is 10. Table 2 shows the experimental results when the number of displayed documents at each iteration is 20. In these tables, F# denotes the number of feedback iterations. SVM-A, SVM-S, and Ro are the compared interactive document retrieval methods. SVM-A denotes an interactive document retrieval method based on our proposed heuristics. SVM-S denotes an interactive document retrieval method based on a *Simple* method. Ro denotes Rocchio-based interactive document retrieval method. Each value is an average value for topics in Table 1 and Table 2. The value with underline denotes the best performance.

From these tables, we can understand that our proposed interactive document retrieval method shows better performance than the other methods for all criteria. Therefore, our proposed interactive document retrieval method achieves more effective retrieval performance and better learning performance than the others.

Table 1 Experimental Results (A)
The number of displayed documents is 10.

F#	P10			P30		
	SVM-A	SVM-S	Ro	SVM-A	SVM-S	Ro
1	<u>0.368</u>	0.169	0.301	<u>0.247</u>	0.140	0.214
2	<u>0.436</u>	0.291	0.282	<u>0.317</u>	0.231	0.218
3	<u>0.407</u>	0.304	0.255	<u>0.301</u>	0.261	0.200
4	<u>0.350</u>	0.319	0.223	<u>0.268</u>	0.261	0.181
5	<u>0.357</u>	0.303	0.225	<u>0.275</u>	0.233	0.170
F#	**MAP**			**R05P**		
	SVM-A	SVM-S	Ro	SVM-A	SVM-S	Ro
1	<u>0.151</u>	0.080	0.146	<u>0.139</u>	0.079	0.129
2	<u>0.182</u>	0.137	0.144	<u>0.153</u>	0.097	0.113
3	<u>0.172</u>	0.142	0.131	<u>0.126</u>	0.092	0.102
4	<u>0.154</u>	0.146	0.121	<u>0.112</u>	0.095	0.085
5	<u>0.150</u>	0.135	0.112	<u>0.103</u>	0.078	0.079

Table 2 Experimental Results (B)
The number of displayed documents is 20.

F#	P10			P30		
	SVM-A	SVM-S	Ro	SVM-A	SVM-S	Ro
1	0.404	0.189	0.266	0.285	0.141	0.201
2	0.427	0.314	0.271	0.310	0.253	0.206
F#	MAP			R05P		
	SVM-A	SVM-S	Ro	SVM-A	SVM-S	Ro
1	0.177	0.087	0.130	0.158	0.075	0.119
2	0.182	0.150	0.125	0.138	0.113	0.096

Note that when the number of feedback iterations is small in Table 1 and Table 2, we can understand that the Rocchio-based method outperformed the *Simple* method. Especially, the tendency is very clear when the number of iterations is one. This implies that straightforward application of SVM can not improve the performance, and our proposed heuristics played effective role to improve the performance of SVM-based interactive document retrieval.

Of course, these results significantly depend on the topic based on user's search intention. This is one of the most important problems in IR, and it is hard to determine how to measure the influence of the user's search intention to retrieved results. Thus we could not deal with this problem here. We consider this user's intention problem as an open problem.

The CPU time for learning was 0.18∼4.05(sec) (CPU: Xeon 3.6GHz, RAM: 2GB, OS: Linux 2.6.8, Library: libsvm-2.71). Since the computational cost depends on the number of data, not the dimension of data, learning of our proposed system with a small number of data is quick.

§5 Discussion

5.1 Coverage of Proposed Heuristics

We verified the effectiveness of our proposed heuristics: *to select the nearest documents to the region of relevant documents in the margin of SVM* by systematic experiments in §4. We consider that the experimental results supported our heuristics is effective over interactive document retrieval for general and huge text data because the newspaper articles has sufficient generality as a text data set. Furthermore our heuristics depends only on the distribution of relevant/non-relevant document vectors: huge non-relevant documents and a small number of relevant documents, not on any structure, term occurrence frequencies of text and representation of text like TFIDF, Latent Semantic Indexing and a probabilistic model. Thus we consider our heuristics has sufficient coverage to document retrieval.

In order to confirm the coverage experimentally, we are preparing to conduct experiments for document search in the Web. The initial search is done by using Google search engine API and a large hit list is obtained in general. Then, with the hit list as a document set, the same procedures in §1 are applied.

We need to improve the bag-of-words representation because the Web pages are semi-structured text with HTML-tags, and this is the issue we are currently attacking.

5.2 Dealing with Non-relevant Documents

Seeing from the equation in §4.1, a query vector can be updated only with non-relevant documents in the conventional Rocchio-based system. However, the SVM-based interactive document retrieval system can not learn a discriminant hyperplane only with such negative examples(or examples of one class), and this is the limitation of general classification learning. Unfortunately we observed that such situations often occurred in interactive document retrieval.

In order to overcome this drawback of SVM-based interactive document retrieval, we are currently developing a non-relevance feedback method[10] which can learn a discriminant hyperplane only with non-relevant documents by employing one-class SVM.[16] It is also a heuristics-based approach as well as the system proposed in this paper, and we are improving the heuristics and try to give intuitive understanding to it.

§6 Conclusion

In this paper, we proposed SVM-based interactive document retrieval using the special selection rule, which can select the effective displayed documents both for a IR system and a user. In our experiments, our proposed interactive document retrieval method showed more effective document retrieval and better learning performance than the two conventional methods using some kind of criteria.

The proposed heuristics, where the documents that are near the bound of the relevant documents area and in the margin area of SVM, are displayed to a user, show better performance of document retrieval. Generally, data mining applications based SVM to drug discovery,[22] bioinformatics and so on have been discussed from active learning point of view. In the discussion, the learning efficiency of computers is very important. In the interactive documents retrieval between human and computers, however, the retrieval efficiency is also important as well as the learning efficiency of computers. Therefore, interactive documents retrieval systems should display the documents which are interesting for users at each iteration and keep the learning performance of the systems. In this paper, we proposed the heuristics to select displayed documents in order to satisfy both of learning and retrieval efficiency. Eventually, we confirmed the effectiveness of our proposed heuristics through systematic experiments with a large document data set.

References

1) Cortes, C. and Vapnik, V., "Support vector networks," *Machine Learning, 20*, pp. 273–297, 1995.
2) Drucker, H., Shahrary, B. and Gibbon, D. C. "Relevance feedback using support vector machines," in *Proc. of the 18th Int'l Conf. on Machine Learning*, pp. 122–129, 2001.
3) Drucker, H., Wu, D. and Vapnik, V.N., "Support vector machines for spam categorization," *IEEE Transaction on Neural Networks, 10*, pp. 1048–1054, 1999.
4) Dumais, S.T., Platt, J., Heckerman, D. and Sahami, M., "Inductive learning algorithms and representations for text categorization," in *Proc. of the 17th Int'l Conf. on Information and Knowledge Management*, pp. 148–155, 1998.
5) IREX. http://cs.nyu.edu/cs/projects/proteus/irex/.
6) Joachim, T., "Text categorization with support vector machines: Learning with many relevant features," in *Proc. of the 10th European Conf. on Machine Learning*, pp. 137–142, 1998.
7) Melville, P. and Mooney, R.J., "Diverse ensembles for active learning," in *Proc. of the 21st Int'l Conf. on Machine Learning*, pp. 584–591, 2004.
8) NTCIR. http://www.rd.nacsis.ac.jp/˜ntcadm/.
9) Okabe, M. and Yamada, S., "Learning filtering rulesets for ranking refinement in relevance feedback," *Knowledge-Based Systems, 18(2-3)*, pp. 117–124, 2005.
10) Onoda, T., Murata, H. and Yamada, S., "Non-relevance feedback document retrieval based on one class svm and svdd," in *Proc. of 2006 IEEE World Cong. on Computational Intelligence*, pp. 2191–2198, 2006.
11) Onoda, T., Murata, H. and Yamada, S., "Support vector machines based active learning for the relevance feedback document retrieval," in *Proc. of the Int'l Workshop on Intelligent Web Interaction*, pp. 393–396, 2006.
12) Rocchio, J., "Relevance feedback in information retrieval," in *The smart system-experiments in automatic document processing*, pp. 313–323, Prentice Hall, Englewood Cliffs, N.J., 1971.
13) Roy, N. and McCallum, A., "Toward optimal active learning through sampling estimation of error reduction," in *Proc. of the 18th Int'l Conf. on Machine Learning*, pp. 441–448, 2001.
14) Salton, G. and McGill, J., *Introduction to modern information retrieval*, McGraw-Hill, 1983.
15) Schapire, R.E., Singer, Y. and Singhal, A., "Boosting and rocchio applied to text filtering," in *Proc. of the 21st Annual Int'l ACM SIGIR*, pp. 215–223, 1998.
16) Schölkopf, B., Platt, J., Shawe-Taylor, J., Smola, A. and Williamson, R., "Estimating the support for a high-dimensional distribution," *Technical Report MSRTR-99-87, Microsoft Research*, 1999.
17) Tong, S. and Koller, D., "Support vector machine active learning with applications to text classification," *Journal of Machine Learning Research, 2*, pp. 45–66, 2001.
18) TREC Web page. http://trec.nist.gov/.
19) Tsuge, S., Shishibori, M., Kuroiwa, S., Tanaka, Y., Hirai, T., Okamoto, R. and Kita, K., "Relevance feedback with support vector machine for information retrieval," in *Proc. of Int'l Conf. on Computer Processing of Oriental Languages*, pp. 35–40, 2001.

20) Vapnik, V.N., *The Nature of Statistical Learning Theory,* Springer, 1995.

21) Vapnik, V.N., *Statistical Learning Theory,* Wiley, New York, 1998.

22) Warmuth, M.K., Rätsch, G., Mathieson, M., Liao, J. and Lemmen, C., "Active learning in the drug discovery process," in *Advances in Neural Information Processing Systems, Vol. 14,* 2002.

23) Yates, R.B. and Neto, B.R., *Modern Information Retrieval,* Addison Wesley, 1999.

Takashi Onoda, Ph.D: He received the M.S. degree in nuclear engineering from Tokyo Institute of Technology, Tokyo, Japan in 1988. He works at Central Research Institute of Electric Power Industry from 1988. He received the Dr. Eng. degree in mathematical engineering from University of Tokyo, Tokyo, Japan, in 2000.

Hiroshi Murata: He received M.S. degree in electronics from Tokyo University of Science in 1993. Since then, he has been with Central Research Institute of Electric Power Industry (CRIEPI), Japan. He is currently pursuing the Ph.D. degree in the Graduate University for Advanced Studies (SOKENDAI), Japan. His current research interests include image processing and machine learning.

Seiji Yamada, Ph.D: He is a professor at the National Institute of Informatics. Previously he worked at Tokyo Institute of Technology. He received B.S., M.S. and the Ph.D. degrees in artificial intelligence from Osaka University. His research interests are in the design of intelligent interaction including Human-Agent Interaction, intelligent Web interaction and interactive machine learning.