Towards Personalized Medical Document Classification by Leveraging UMLS Semantic Network

Kleanthi Lakiotaki[1], Angelos Hliaoutakis[1], Serafim Koutsos[1], Euripides G.M. Petrakis[1]

[1]Technical Univ. of Crete (TUC), Dept. of Electronic and Computer Engineering, Chania, Crete, Greece
`klakiotaki@isc.tuc.gr, {angelos, euripides}@intelligence.tuc.gr, skoutsos@gmail.com`

Abstract. The overwhelmed amount of medical information available in the research literature, makes the use of automated information classification methods essential for both medical experts and novice users. This paper presents a method for classifying medical documents into documents for medical professionals (experts) and non-professionals (consumers), by representing them as term vectors and applying Multiple Criteria Decision Analysis (MCDA) tools to leverage this information. The results show that when medical documents are represented by terms extracted from AMTEx, a medical document indexing method, specifically designed for the automatic indexing of documents in large medical collections, such as MEDLINE, better classification performance is achieved, compared to MetaMap Transfer, the automatic mapping of biomedical documents to UMLS term concepts developed by U.S. National Library of Medicine, or the MeSH method, under which documents are indexed by human experts.

Keywords: Health Information System, User profile, MeSH, MEDLINE, AMTEx, MetaMap Transfer, Semantic Network, Recommender Systems, MCDA.

1   Introduction

In the world of medical information literature, two major categories of information seekers are mainly identified. The first, often called "healthcare consumers", or simply consumers, represent those who search in the medical document corpus to find medical information described in simple words, as opposed to "healthcare experts", that often represent medical professionals.

Over the last few decades, consumer involvement in health care has been significantly increased. At the same time, the growing health information available on the Internet offers a valuable tool to healthcare consumers. On February 2011 a research by the Pew Research Center's Internet and American Life Project and the California HealthCare Foundation (CHCF) found that 80% of Internet users look online for

health information, making it the third most popular online search among all those tracked by the Pew Internet Project, following email and search engines [1].

On the other hand, medical information systems such as MEDLINE[1] are designed to serve health care professionals (expert users in general, such as clinical doctors or medical researchers). Typically, expert users are familiar with the type and content of the medical resources (such as the National Library of Medicine - NLM dictionaries and databases) they are using and use medical terminology for their searches. A medical information system must be capable of providing dedicated, domain specific answers to experts or, simple, easy comprehend answers to novice users, respectively. A similar categorization of medical information applies in existing systems such as MedScape[2], MedlinePlus[3], Wrapin[4] and MedHunt[5] (maintained by HON, the Health on the Net Foundation, a non-profit organization aiming at providing authoritative and trustworthy information on the Web) and in other related systems such as Med-Worm[6], a medical RSS feed provider as well as a search engine built on data collected from RSS feeds. PubMed[7] of NLM is of particular interest to us. It provides free access to MEDLINE document abstracts and to articles in selected life sciences journals not included in MEDLINE.

An automatic system able to characterize medical articles as "consumer specific" or "expert specific" and thus appropriately recommend it, is valuable to both cases, by assisting consumers in managing their personal health information and experts in significantly reducing their effort on information seeking task.

In this work, we investigate on potential improvements to the problem of medical document classification by user profile (i.e., consumer users and domain experts). The high classification performance achieved in this work results from the main contributions of this paper, which are: (1) the realization and demonstration that different terms representing a medical document, contribute unequally to its classification; (2) the incorporation of Multiple Criteria Decision Analysis (MCDA), as a method for calculating the significance by which each term participates in the document classification; and (3) the representation of medical documents by term vectors extracted by the AMTEx method [2]. Evaluation results are taken on a subset of MEDLINE documents, the premier bibliographic database of the U.S. National Library of Medicine[8] (NLM). Building upon AMTEx and MetaMap Transfer[9] we show that document representations are semantically compact and more efficient, being reduced to a limited number of meaningful multi-word terms (phrases), rather than by large vectors of single-words (as it is typical in classic information systems work) part of which may be void of distinctive content semantics. Although document contents are summarized

by only a few terms, these terms can be any term in the MeSH[10] with almost 24,000 terms, meaning that MCDA should treat any MeSH term as a separate classification criterion which is prohibitive in practice. In this work, application of MCDA is enabled by mapping MeSH terms to their more abstract category terms in the Semantic Network[11] (SN) of UMLS[12] (Unified Medical Language System).

Related work is discussed in Section 2. Our method on document categorization by user profile is presented in Section 3. Evaluation results are presented in Section 4 followed by conclusions and issues for further research in Section 5.

2 Background and resources

Recommender Systems (aka Recommenders or Recommendation Systems) have gained increasing popularity on the web, both in research and in industry [3]. Multi Criteria Recommender Systems have been also proved successful in this direction [4]. At the same time, in the healthcare literature, Recommender Systems are still on their infancy. In [5] for example, a health recommendation system architecture is proposed using rough sets, survival analysis approaches and rule-based expert systems. Their main goal was to recommend clinical examinations for patients or physicians from patients' self-reported data. The challenges and opportunities of merging recommender systems with personalized health education can be found in [6].

Medical document repositories, such as MEDLINE and PubMed, contain a huge amount of medical literature and are supported by NLM. Automatic extraction of useful information from these online sources remains a challenge because these documents are unstructured and indexed by human experts by assigning to each one, a number (typically 10 to 12) of terms, based on a controlled list of indexing terms, deriving from a subset of the UMLS Metathesaurus, the MeSH (Medical Subject Headings) thesaurus.

Automatic indexing and categorization of medical documents relies mainly on term extraction for the identification of discrete content indicators, namely index terms. Traditional indexing techniques ignore multi-word and compound terms, which are split into isolated single word index terms. However, compound and multi-word terms are very common in the biomedical domain [7] and are often used in indexing medical documents. Multi-word terms carry important classificatory content information, since they comprise of modifiers denoting a specialization of the more general single-word, head term. For example, the compound term "heart disease" denotes a specific type of disease.

In this work, we focus our attention on multi-word terms and AMTEx for extracting multi-word terms from medical documents. AMTEx and MetaMap Transfer have been shown to be more suitable than single-word term extraction methods not only for document indexing and retrieval, but also, for general concept description and ontology construction tasks [8]. AMTEx, in particular, has been shown to be more selective

---

[10] http://www.ncbi.nlm.nih.gov/mesh/
[11] http://www.nlm.nih.gov/pubs/factsheets/umlssemn.html
[12] http://www.nlm.nih.gov/research/umls/

than the MetaMap Transfer (MMTx) method of NLM which maps arbitrary text to concepts in the UMLS Metathesaurus.

2.1   Unified Medical Language System (UMLS)

The Unified Medical Language System (UMLS) is a source of medical knowledge developed by the U.S. NLM. UMLS consists of the Metathesaurus, the Semantic Network and the SPECIALIST lexicon. Metathesaurus is a large, multi-purpose, and multi-lingual vocabulary database. It integrates about 800,000 concepts from 50 families of vocabularies. The Semantic Network (SN) consists of 134 semantic types categorizing the Metathesaurus concepts. The purpose of the SN is to provide a consistent categorization of all concepts represented in Metathesaurus and a set of useful relationships among these concepts.

2.2   The MeSH Thesaurus

The MeSH Thesaurus (Medical Subject Headings) is a taxonomy of medical and biological terms and concepts suggested by the U.S NLM. The MeSH terms are organized in IS-A hierarchies, where more general terms, such as "chemicals and drugs", appear in higher levels than more specific terms, such as "aspirin".

2.3   WordNet

WordNet[13] is an on-line lexical reference system developed at Princeton University which attempts to model the lexical knowledge of a native speaker of English. WordNet v.2.0 (2006) contains around 127,361 terms, organized into taxonomic hierarchies. Nouns, verbs, adjectives and adverbs are grouped into synonym sets (synsets). The synsets are also organized into senses (i.e., corresponding to different meanings of the same term or concept).

2.4   MEDLINE and OSHUMED

MEDLINE database is a collection of biomedical articles. It consists of medical publications abstracts together with metadata, which is information on the organization of the data, the various data domains, and the relations between them.
The OHSUMED[14] test collection is a set of 348,566 references from MEDLINE, consisting of titles and/or abstracts from 270 medical journals over a five–year period (1987–1991). OHSUMED is commonly used in benchmark evaluations of IR applications and provides 64 queries and the relevant answer set (documents) for each query.

---

[13] http://wordnet.princeton.edu/
[14] http://ir.ohsu.edu/ohsumed/ohsumed.html

2.5   MetaMap Transfer Technology Transfer and AMTEx

MetaMap Transfer Technology (MMTx) uses the Metathesaurus and SPECIALIST lexicon knowledge resources during the term extraction process. This process maps arbitrary text to Metathesaurus term concepts.

AMTEx [2], implements the C/NC-value [9], a domain-independent method for the extraction of multi-word and nested terms. In this approach, noun phrases are initially selected by linguistic filtering. The subsequent statistical component defines the candidate noun phrase termhood by two measures: C-value and NC-value.

2.6   Multicriteria Decision Analysis

In Decision Sciences, the field of Multiple Criteria Decision Analysis (MCDA) is well established and comes into a large variety of theories, methodologies, and techniques [10]. Here, we adopt the disaggregation-aggregation approach [11], a common approach in the field of MCDA, by exploiting the UTASTAR algorithm, a representative algorithm of this approach. In abstract, the UTASTAR algorithm, considers as input a weak-order preference structure on a set of alternatives, together with the performances of the alternatives (here the medical documents) on all attributes, and returns as output a set of additive value functions based on multiple criteria, in such a way that the resulting structure would be as consistent as possible with the initial structure given by the user. This is accomplished by means of special linear programming techniques. UTASTAR's output involves the value functions associated to each criterion, approximated by linear segments, as well as the criteria significance weights (trade-offs among the criteria values).

3   Medical Document Recommendation by User Profile

We follow a three phase (*data retrieval and term extraction*, *data representation and modeling* and *document classification*) methodology, as described in Figure 1, to prove our assumption that certain Semantic Network sub-category terms are more important that others in medical document classification based on user profile.
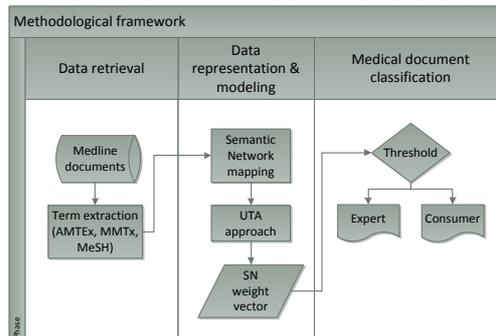

Fig. 1. The proposed methodological framework consisting of three phases (data retrieval, modeling and categorization).

In this work, we advocate that the characterization of a medical document as "for experts" or "for consumers" depends on the expert terms that the document contains, which in turn are mapped to the Semantic Network sub-category terms. The latter, represent the criteria on which the overall percentage of expert terms in a document term vector depends.

To begin, in Sections 3.1, 3.2 and 3.3, we discuss the overall framework of our proposed approach together with all the individual components.

3.1 Data retrieval and expert term extraction

As already stated, our approach relies on popular knowledge and algorithmic resources namely, the UMLS MeSH and Semantic Network for document indexing, MetaMap Transfer and the AMTE$_X$ extension for term extraction and the MEDLINE collection of biomedical articles for testing their performance.

The proposed approach for categorizing MEDLINE documents by user profile relies on the observation that MeSH terms are distinguished into i) general medical terms expressing known concepts (e.g., "pain","headache") which are easily conceived by all users, ii) domain specific terms which are used mainly by experts, iii) general - non medical terms. The more expert terms a document contains, the higher its probability to be a document for experts [12]. Moreover, since the amount of expert terms in a document is low, even for expert documents, we ignore consumer terms during the modeling process in our experiments and we represent all documents based on expert terms. Consequently, we assume mutual exclusion, meaning that any medical document that is not expert is presumably a consumer document.

We combine information from WordNet and MeSH to construct the following three term vocabularies and exploit them to characterize terms as expert or consumer:

Vocabulary of General Terms (VGT): these are terms that belong to WordNet but not to MeSH:

$$VGT = (WordNet) - (MeSH)$$

It follows that VGT contains 105,675 general (WordNet) terms.

Vocabulary of Consumer Terms (VCT): these are terms that belong to both, WordNet and MeSH:

$$VCT = (WordNet) \cap (MeSH)$$

It follows that VCT contains 7,165 consumer (MeSH) terms.

Vocabulary of Expert Terms (VET): these are MeSH terms that do not belong to WordNet:

$$VET = (MeSH) - (Wordnet)$$

It follows that VET contains 16,719 expert (MeSH) terms. Next, documents are represented by terms, extracted by applying three different approaches: AMTEx, MetaMap Transfer and the original MeSH terms.

3.2 Data representation and modeling

Since the Disaggregation-Aggregation approach of MCDA has mainly focused on the development of comprehensible decision models from small data sets, whose

main objective is to support decision aiding through an interactive model calibration process, the total number of expert terms found in the document is too large in order to consider all the initial MeSH terms extracted from this document as criteria in the MCDA process. Furthermore, the entire set of alternatives (here medical documents) should be evaluated on the same number of criteria. For these reasons, after the term extraction process, every term originating from either AMTEx or MetaMap Transfer is mapped by the two-layered indexing structure to the UMLS Semantic Network category terms and these sub-categories are considered as criteria. MeSH terms as given by Medical Subject Headings Section staff for every document are also similarly mapped.

Only expert terms, as described in Section 3.1, count in this process and the simple term frequency measure is applied. Hence, a document in the dataset is represented by a 130-dimensional vector of expert term frequency as:

$$d_i=\{tf_1,tf_2,...tf_n\}, \qquad where\ n=1,2,...130$$

and

$$tfi = \frac{\sum_{j=1}^{k} t_j \rightarrow t_j \in VET}{N} \tag{1}$$

where k is the number of expert terms that belong to the ith SN category and N is the total number of expert terms in d_i. For example, consider that for a document d_i five different expert MeSH terms are extracted by AMTEx, two of which belong to the sub-category "Molecular Function", one to the sub-category "Cell" and the remaining two to "Disease or Syndrome". Then, the value of sub-categories "Molecular Function" and "Disease or Syndrome" will be 2/5, while of "Cell", 1/5. Therefore, the smaller the number of a sub-category, the less this sub-category contributes to the classification of d_i as expert document. A zero value here means that no expert term from this SN sub-category was extracted. The question that arises at this point and we also try to address in this work is: *Do all sub-category of the Semantic Network contribute identically to the characterization of a document as expert? And if not, what is the significance of each sub-category for medical document classification as expert or consumer?*

MCDA methods usually assume that only a small reference set is available, since it is difficult for decision makers to express their global preferences on too many alternatives. Therefore, during the 10-fold stratified cross validation that was performed to estimate the performance of our predictive model, which was built based on the aforementioned Multicriteria Decision Analysis approach, the training set was every time randomly split into 10 segments of size n with replacement and the average values of the estimated UTASTAR parameters were calculated. By applying the so called UTASTAR algorithm in the training set, a vector of significance weights for the UMLS Semantic Network category terms is calculated, indicating the different role of each category in the characterization of a MEDLINE document as "for consumers" or "for experts".

According to the methodological requirements of the Disaggregation-Aggregation approach, a weak preference order of the alternatives (here, the medical documents),

is required to apply ordinal regression. In our experiments, the probability of a document to be considered as "expert" or "consumer" is calculated as the number of expert terms that the specific document contains divided by the total number of terms extracted from the specific document. More specifically, this probability is calculated as the percentage of terms that belong to the VET. For example, a document with VET% = 0.62 has 62% probability of being a document suitable for experts. Therefore, we assume that this probability represents the global estimation of a document as expert and based on that, we transform the initial global ratings into a ranking order for the training set.

3.3 Medical document classification

During the last phase, the classification phase, a medical document is labeled as either expert or consumer. This choice is based upon the utility score calculated based on the final solution that corresponds to the marginal value functions (criteria weights). A linear transformation of the form $\sum_i w_i \cdot b_i$ where i denotes the SN sub-categories and b_i the value of the document under consideration for the specific SN sub-category provides the utility score for every document that belongs to the test set. The main concern that arises at this point is threshold selection. Statistical classifiers, such as a Naive Bayes classifier or Neural Network classifiers, usually calculate a score, representing the degree to which an instance is a member of a class. These scores represent probabilities, in which case standard theorems of probability can be applied. In our case however, utility scores, are not strict probabilities, in which case the only property that holds is that a higher score indicates a higher probability. Thus, in such "scoring classifiers" a threshold is necessary to produce a discrete (binary) classifier.

In diagnostic studies, the ROC curve, a plot of a test's sensitivity versus (1-specificity) for every possible threshold value, and the area under the ROC curve (AUC) are important tools in assessing the diagnostic utility of classifiers [13]. Still, finding an optimal cut-point for discriminating between binary classes is also of paramount importance.

Although the area under the ROC curve (AUC) is the most commonly used global index of diagnostic accuracy the Youden Index [14] has also frequently used by researchers. This index can be defined as $J=max_i\{Sensitivity(i) + Specificity(i) - 1\}$ and ranges between 0 and 1.

4 Evaluation

As a proof of concept we designed a series of experiments whose purpose is twofold: First, to study and compare the effectiveness of AMTEx, MetaMap Transfer and MeSH, in classifying medical documents and second, to prove that Semantic Network categories contribute differently in classifying medical documents.

Initially, documents are retrieved from a subset of the OHSUMED TREC collection consisting of 10% of OHSUMED (i.e., 34,000 documents). Both, data store and

access mechanisms are implemented using Lucene. The retrieved documents were evaluated manually by users, as consumer or expert documents and this categorization is considered as the ground truth in our experiment. Subsequently, we extracted only expert documents that contain at least 2 Semantic Network category terms resulting in a subset of 237 different expert documents. To avoid any inconsistencies originating from our dataset, the same number of consumer documents is selected. Therefore, our experimental data set consists of the above 237 expert documents and 237 consumer documents and is used in all our experiments below.

As already stated, we advocate that not all the Semantic Network sub-categories contribute identically to the classification of medical documents as expert, or consumer. On the contrary, we identify different significance weights for those subcategories. It is crucial to mention at this point that the significance weights do not follow the frequency appearance of the sub-categories in the document corpus, meaning that the most frequent sub-category is not necessarily the most important.

To evaluate classification performance we calculated several classification accuracy measures (precision, recall, F-measure, classification accuracy), in 10-fold crossvalidation, as well as ROC Analysis, which altogether help us to identify the best classification performance.

Table 1 summarizes the results of best-case, worst-case and average case values of all the evaluation measures calculated on AMTEx, MetaMap Transfer and MeSH term vectors from the 10 folds. The best and worst values are assigned in bold and italics, respectively.

It is obvious from Table 1 that AMTEx outperforms all the other methods in the majority of performance measures, whilst MeSH cannot adequately separate expert and consumer documents. Moreover, it is crucial to mention here, that UTASTAR weights achieve much better evaluation scores compared to a relative equal weight scenario, by simply adding the SN category values into a utility score and apply exactly the same evaluation measures.

Table 1. Classification evaluation measures for AMTEx, MetaMap Transfer and MeSH

		AMTEx	**MetaMap Transfer**	**MeSH**
F-measure	Best -case	**0.8000**	0.3175	*0.2615*
	Worst -case	**0.5333**	0.2581	*0.1977*
	Average-case	**0.6527**	0.2814	*0.2261*
Precision	Best -case	**0.7097**	0.1942	*0.1604*
	Worst -case	**0.3636**	0.1481	*0.1149*
	Average-case	**0.5085**	0.1660	*0.1313*
Recall	Best -case	**1.0000**	**1.0000**	**1.0000**
	Worst -case	**0.9130**	0.8696	*0.6957*
	Average-case	**0.9449**	0.9322	*0.8391*
AUC	Best -case	**0.9732**	0.7293	*0.6578*
	Worst -case	**0.9528**	0.6407	*0.5359*
	Average-case	**0.9629**	0.6882	*0.6047*

Accuracy	Best -case	**0.9580**	0.6705	*0.6336*
	Worst -case	**0.8397**	0.4733	*0.3359*
	Average-case	**0.9037**	0.5656	*0.4743*

Additionally, Figure 2 illustrates the ROC curves for the best precision values in all cases allowing us to compare the classification performance of our model built based on AMTEx terms, MetaMap Transfer terms and MeSH terms.
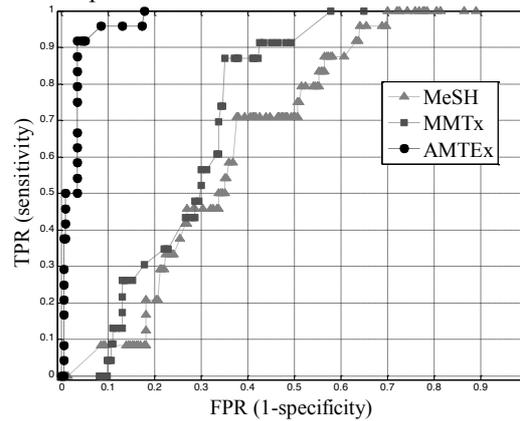

Fig. 2. ROC curves for AMTEx, MetaMap Transfer and MeSH approaches.

Clearly, AMTEx shows the best performance than MetaMap Transfer or MeSH.

We also studied the classification performance of the three medical document representation methods by applying Decision Tree analysis (DTA), as an alternative of Multiple Criteria Decision Analysis. Decision trees are a simple, but powerful form of multiple attribute analysis. Decision trees are produced by algorithms that identify various ways of splitting a data set into branch-like segments. Here, we consider two segments, experts and consumers, as the dependent class variables and the Semantic Network sub-categories as the independent variables. We performed 10 fold cross-validation on the J48 classifier with default parameter values (confidence threshold for pruning at 0.25) for inducing classification trees. The J48-algorithm is a Java re-implementation of the C4.5-algorithm and is a part of the machine learning package WEKA[15]. Table 2 shows the results of Decision Tree Analysis on the same document collection.

Table 2. Decision Tree Analysis results

	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)	AUC
AMTEx	83.75	87.3	83.8	80.8	0.812
MetaMap Transfer	78.9	79.5	78.9	78.8	0.807
MeSH	65.82	66.0	65.8	65.7	0.674

---

[15] http://www.cs.waikato.ac.nz/ml/weka/

Clearly, the Decision Tree analysis results, also confirm that AMTEx is the most appropriate method for term extraction and representation for the MEDLINE documents that we studied. This result becomes more apparent in the MCDA experiments, probably due to the fact that AMTEx terms seem to be more representative for the classification of medical documents, than MetaMap Transfer or MeSH terms assigned by humans, leading thus to a better identification of the significance of SN sub-categories in this classification. Moreover, AMTEx achieves better results in most evaluation measures when MCDA is applied compared to DTA.

5   Conclusions

This work brings together ideas from document information management and Recommender Systems and shows how these ideas can applied for effectively classifying medical documents by user profile. We investigated the problem of automatic categorization of medical information on two common types of users (consumers and experts). Medical documents were represented by term vectors extracted from three different approaches (AMTEx, MetaMap Transfer and MeSH). Based on our experiments, we conclude that when documents are represented by AMTEx, a medical document indexing method, specifically designed for the automatic indexing of documents in large medical collections, such as MEDLINE, the categorization performance is significantly increased compared to when the same documents are represented by MetaMap Transfer, the automatic mapping of biomedical documents to UMLS term concepts developed by U.S. National Library of Medicine, or the MeSH method, under which documents are indexed by human experts, based on a controlled list of indexing terms, deriving from a subset of the UMLS Metathesaurus. We also proved that the UMLS Semantic Network sub-category terms can act as criteria for the categorization of a medical documents, however their performance play an important role in their classification ability. Moreover, our experiments show that Multiple Criteria Decision Analysis, as a method for identifying the significance of Semantic Network sub-category terms in their classification ability, achieves better results compared Decision Tree Analysis, when documents are represented by AMTEx terms.

Future work involves extending our studies in discovering mechanisms for classifying medical documents into several thematic categories. For example, our intention is to be able to recommend medical documents to expert users for "breast cancer".

References

1. Fox, S.: 80% of internet users look for health information online. (2011).
2. Hliaoutakis, A., Zervanou, K., Petrakis, E.G.M.: The AMTEx approach in the medical document indexing and retrieval application. Data & Knowledge Engineering. 68, 380-392 (2009).
3. Ricci, F., Rokach, L., Shapira, B., Kantor, P.B.: Recommender Systems Handbook. Springer US, Boston, MA (2011).
4. Lakiotaki, K., Matsatsinis, N., Tsoukiàs, A.: Multicriteria User Modeling in Recommender Systems. IEEE Intelligent Systems. 26, 64-76 (2011).
5. Pattaraintakorn, P., Zaverucha, G.M., Cercone, N.: Web Based Health Recommender System Using Rough Sets, Survival Analysis and Rule-Based Expert Systems. Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (LNCS). pp. 491-499 (2007).
6. Luis, F.-L., Karlsen Randi, Vognild Lars: Challenges and opportunities of using recommender systems for personalized health education. Medical Informatics in a United and Healthy Europe. pp. 903-907. European Federation for Medical Informatics (2009).
7. Maynard, D., Ananiadou, S.: TRUCKS: A model for automatic multi-word term recognition. Journal of Natural Language Processing. 8, 101–125 (2000).
8. Divita, G., Tse, T., Roth, L.: Failure analysis of MetaMap Transfer (MMTx). Studies in health technology and informatics. 107, 763-7 (2004).
9. Frantzi, K., Ananiadou, S., Mima, H.: Automatic recognition of multi-word terms : the C-value / NC-value method. International Journal on Digital Libraries. 3, 115-130 (2000).
10. Figueira, J., Greco, S., Ehrgott, M.: Multiple criteria decision analysis: state of the art surveys. Springer Verlag, London (2005).
11. Siskos, Y., Grigoroudis, E.: UTA methods. In: J. Figueira, S.G. and M.E. (ed.) Multiple criteria decision analysis: State of the art surveys. pp. 297-344 (2005).
12. Petrakis, E.G.M., Hliaoutakis, A.: Automatic Document Categorisation by User Profile in Medline. 1-10 (2011).
13. Fawcett, T.: ROC Graphs : Notes and Practical Considerations for Data Mining Researchers, HP Laboratories Palo Alto. (2003).
14. Fluss, R., Faraggi, D., Reiser, B.: Estimation of the Youden Index and its associated cutoff point. Biometrical Journal. 47, 458-72 (2005).