# Outpatient Department Recommendation Based on Medical Summaries

Hen-Hsen Huang, Chia-Chun Lee, and Hsin-Hsi Chen

Department of Computer Science and Information Engineering,
National Taiwan University
#1, Sec. 4, Roosevelt Road, Taipei, 10617 Taiwan
{hhhuang,cclee}@nlg.csie.ntu.edu.tw, hhchen@ntu.edu.tw

**Abstract.** The family medicine in some regions is not as popular as that in the United States. Most patients choose the outpatient department without professional advice. In this work, we propose a health care aiding system that recommends the outpatient department for a patient according to his/her chief complaint and personal attributes. The recommendation is based on the past medical summaries of a hospital. Three methods including language model, support vector machine, and k-nearest neighbor algorithm along with different features are explored. The experimental results show that the SVM classifier with features selected from chief complaint, as well as personal attributes such as age, gender, and disease information achieves an f-measure of 79.35%.

**Keywords:** Health Care Aiding, IR Application, Medical Informatics, Recommendation.

## 1 Introduction

In some regions such as Taiwan, the family medicine is not popular. Rather than ask the professional advice, most patients have to choose one of outpatient departments from a hospital by themselves when they feel ill. The outpatient departments are classified based on the aspect of medical professionalism, rather than on the individual patient's cognition. For this reason, choosing a proper department is often a challenging task, especially for those patients who lack of medical knowledge or whose disease is imperceptible. Mistaking the proper department not only wastes the patients' time and money and increases their suffering, but also is inherent in the risk of disease progression.

Although the outpatient guide is available in some hospitals, it is still not convenient for many patients because the information on the guide tends to be too brief to help patients rapidly find their most proper departments, especially for the large hospitals including dozens of departments.

Fig. 1 shows a screenshot of the patient guide provided by the website of National Taiwan University Hospital (NTUH), the largest hospital in Taiwan. In this patient guide, the information is organized in the style of frequently asked questions, aka FAQ. All the health problems are classified into nine categories including "Children",

Y. Hou et al. (Eds.): AIRS 2012, LNCS 7675, pp. 518–527, 2012.

"Nerves and Sensory", "Maternity and Urinary", "Skin", "Mental and Psychological", "Health check and Others", "Physique", "Beauty and Body", and "Eye, Ear, Nose, and Throat". The first entry in the guide is a pair of questions, such as "Which department should be taken for the children with development delay?" and the corresponding suggestion, "The pediatric rehabilitation and child psychologist in the department of pediatrics". For the category "Children", only 9 questions are listed in the guide. Even in the largest category "Health check and Others", only 19 questions are listed. It is obvious that this guide is too limited to cover thousands of health issues. In addition, the FAQ-styled data is not ideal for most patients to search the needed information. For instance, the question "Which department should be taken for the patient of stroke?" is aligned in the category "Nerves and Sensory", and the suggestion of this question is "department of neurology". In Chinese, "Nerves" and "Neurology" share the same terminology "神經". If a patient knows the stroke is an issue related to the nerves, and s/he is likely to know the department of neurology is the proper department to take. On the other hand, a patient who does not have any knowledge about the stroke and the nerves will never find this entry in the category "Nerves and Sensory" in this guide. The other instance is observed in the category "Mental and Psychological". The suggested department for all the questions in this category is "department of psychiatry", which is already suggested in the category name. This is a case of non-informative organization because the patient will choose the department of psychiatry without any doubt since s/he knows that her/his health trouble is related to psychology.


臺大醫院 > 就醫指南 > 就醫須知 > 我該看那一科

我該看那一科

| | | |
|---|---|---|
| 兒童方面 | 神經、感覺方面 | 婦產、泌尿方面 |
| 皮膚方面 | 精神、心理方面 | 健檢、其他方面 |
| 筋骨方面 | 美容、美體方面 | 眼、耳 、鼻、喉方面問題 |

兒童方面

| 問　題 | 回覆與建議 |
|---|---|
| 兒童發展遲緩掛那一科？ | 兒童門診之兒童復健、兒童心理科。 |
| 兒童有皮膚上的問題掛那一科？ | 兒童門診之兒童皮膚科。 |
| 兒童舌繫帶不正常掛那一科？ | 先找兒童門診之兒童復健科醫師檢查，由醫師診斷後處置或建議。 |
| 兒童說話不清楚掛那一科？ | 兒童門診之兒童復健科。 |
| 成人／兒童長不高掛那一科？ | 成人：內科部之內分泌科。<br>兒童：兒童門診→小兒部→內分泌系。 |
| 成人／兒童檢查過敏源掛那一科？ | 成人：家醫科 或 耳鼻喉科。<br>兒童：兒童門診→小兒部→過敏系。 |
| 成人／兒童疝氣掛那一科？ | 成人：外科部之胃腸外科。<br>兒童：兒童門診→兒童外科或 兒童門診→兒童泌尿。 |
| 成人／兒童割包皮掛那一科？ | 成人：泌尿科。<br>兒童：兒童門診→兒童外科或 兒童門診→兒童泌尿。 |
| 成人／兒童食道狹窄掛那一科？ | 成人：外科部之胸腔（食道）外科。<br>兒童：兒童門診→兒童外科。 |

TOP

皮膚方面

| 問　題 | 回　覆 |
|---|---|
| 黑痣掛那一科？ | 皮膚科。 |
| 脂肪瘤掛那一科？ | 皮膚科或外科部→整型外科。 |
| 手汗症掛那一科？ | 皮膚科。 |
| 狐臭掛那一科？ | 皮膚科或外科部→整型外科。 |
| 禿頭掛那一科？ | 皮膚科。 |
| 頭皮上／裏長東西掛那一科？ | 皮膚科或外科部→整型外科。 |
| 掉髮掛那一科？ | 皮膚科。 |
| 肚臍有分泌物掛那一科？ | 皮膚科。 |

TOP

**Fig. 1.** Screenshot of the patient guide on the website of NTUH

From the above observations, the limited coverage and non-informative organization of outpatient guide is useless to many patients, and the patients are still used to ask the staffs in the hospital to get the right outpatient clinic. As a result, how to use the expensive medical resources efficiently becomes an important issue.

Instead of the outpatient guide and the human suggestion, we propose an automatic recommendation system that suggests a patient the most proper outpatient department according to the patient's chief complaint and personal attributes including age, gender, and chronic diseases. The rest of this paper is organized as follows. We introduce the experimental dataset in Section 2. The recommendation methods and the features to be explored are presented in Section 3. Section 4 shows and discusses the experimental results. Finally, we conclude this paper in Section 5.

## 2 Experimental Dataset

In the past, some information retrieval systems use the medical documents to tackle the issues relating to health care [1-2]. Some other work suggests patients useful health information based on learning algorithms [3]. To the best of our knowledge, using the medical documents to suggest patients the proper outpatient department is a new idea.

A set of medical documents from NTUH is employed to develop such a recommendation system. A data entry in the collection is a medical summary of a patient's visit. There are three parts in a medical summary, including a chief complaint, a brief history, and a course and treatment. The chief complaint is mostly a short statement that is said by a patient in the outpatient clinic declaring the purpose of his/her visit, which is usually a description of the patient's physical discomfort. The samples (S1), (S2), (S3), and (S4) are some instances of chief complaints in our dataset.

(S1) Epigastralgia for 10 days
(S2) Tarry stool twice since last night
(S3) Left thyroid goiter for about 1 year
(S4) Cough and dyspnea for 3 days

As shown in the instances, a typical chief complaint consists of two elements, i.e., the symptoms found by the patient and the duration of these symptoms. The brief history is a summary about the physical conditions of the patient. The sample (S5) is the first two paragraphs of a brief history in our dataset.

> (S5) This 61y/o female was a case of (1) schizophrenia, paranoid type, for 1 year and (2) asthma for more than 30 years, and she has taken medicine as instructed at 三總 hospital.
>
> Several months ago, she began to have abdominal fullness especially after drink. Acid regurgitation sensation was found at times. And sleeplessness was caused by abdominal fullness. Abdominal pain, nausea, dysphagia, melena or heart burn sesation were not noted. However, body weight loss 4kg was noticed in recent 3 months (63kg ->59kg). Therefore, she went to 三總 hospital. On

94/12/30, PES revealed gastric mass on body post and her family was told of early gastric cancer. Pathology showed signet-ring adenocarcinoma.

The brief history is written by physicians in English with mixing a few Chinese terms such as person names and hospital names. As shown in the sample (S5), the word "三總" is the name of another large hospital in Taiwan. The physicians describe the personal information about the patient and the past medical treatment of the patient in brief history as a reference for advanced treatment. In (S5), the personal attributes including the age and the gender of the patient are mentioned. In addition, the chronic diseases of the patient are also recorded in the brief history. For instance, the first part of the brief history in another medicine summary is provided as the sample (S6).

(S6) The 74 year old woman was a patient with diabetes mellitus, hypertension, chronic renal insufficiency, neurogenic bladder under medical control for six years.

The chronic diseases of the patient such as diabetes mellitus and hypertension are described in the brief history as information for the medical diagnosis and treatment.

The third part of a medical summary is the course and treatment. In this part, the treatment processes such as medication administration, inspection, surgery and the treatment outcomes are described in detail. For instance, the first paragraph of a course and treatment is shown as the sample (S7).

(S7) For her uremic symptoms and hyperkalemia, she started to receive regular hemodialysis three times a week at the day on admission. Peritoneal dialysis was chosen for long term dialysis and Tenckhoff catheter implantation was performed on 95/1/23.

The other instance (S8) shows how the medication administration and the outcomes of inspection described in a course and treatment as follows.

(S8) After admission, he received Gemzar 1920mg (1000mg/m2) and Carboplatin 450mg(AUC=6) on 2006/01/09 smoothly. Besides, chest echo on 1/09 showed minimal left pleural effusion and left supravicular lymph node aspiration was done on the same day. The cytology result was pending. No fever or specific discomforts developed. Under the relative stable condition, he was discharged on 1/10 and followed up at OPD.

All the fields of each medicine summary are listed in Table 1. Besides the three major parts, i.e., chief complaint, brief history, and course and treatment, several metadata are available in each medical summary. The patient's name, the physician's name, and the date of the visit are removed in this dataset because of the privacy. Department is the critical information for our work. After the medical summaries from the smaller departments whose medical summaries are no more than 1,000 are excluded, all the rest medical summaries are classified into 14 classes according to their departments. The statistics of our dataset is shown in Table 2.

From the dataset, we can obtain the department and the chief complaint of each medical summary. In addition, the age, the gender, and the chronic diseases of the

**Table 1.** The fields in a medicine summary in our dataset

| Fields | Explanation | Samples |
|---|---|---|
| Chief Complaint | The statement said by a patient describing the symptoms of the comfort or the illness | See (S1), (S2), (S3), and (S4) |
| Brief History | The personal information and the past medicine events of the patient | See (S5) and (S6) |
| Course and Treatment | The processes and the outcomes of this treatment | See (S7) and (S8) |
| Department | The outpatient department | department of neurology |
| Ward | The patient's ward number | 13D 1201 |
| Year | The year of the patient's visit | 2006 |

**Table 2.** Statistics of the dataset

| Department | Number of entries |
|---|---:|
| Dental | 1,969 |
| Dermatology | 1,144 |
| ENT (Ear, Nose, and Throat) | 6,400 |
| Internal Medicine | 32,160 |
| Neurology | 2,544 |
| Obstetrics and Gynecology | 8,928 |
| Oncology | 5,082 |
| Ophthalmology | 4,332 |
| Orthopedics | 8,759 |
| Pediatrics | 10,904 |
| Rehabilitation | 1,575 |
| Psychiatry | 1,777 |
| Surgery | 27,531 |
| Urology | 7,216 |
| Total | 120,321 |

patient are extracted as features as well. Therefore, we can use this dataset as training data to build a recommendation system that ranks the departments according to a patient's input such as the chief complaint and the personal attributes.

## 3 Methods

We model the recommendation task as a ranking problem, and three ranking approaches and various features are explored in the experiments. The ranking algorithms are introduced as follows.

### 3.1 Language Model

The idea is to find a department such that the chief complaint of a patient is generated most likely by the model trained by the previous patients' complaints of the department. In the training stage, we train each department as a language model based on the patients' medical summaries in the department. In the test stage, the perplexities between the input and all the departments are calculated to rank the departments.

The lower the perplexity is, the more likely the input is related to the department. As a result, the department with the lowest perplexity is the most proper department for the input. The SRI Language Modeling Toolkit (SRILM)[1] is utilized to training and testing the language models. The depth of language models are set as trigram, which is better than unigram and bigram in the experiments.

### 3.2 Support Vector Machine

Department recommendation can be modeled as a text classification task. Support vector machine (SVM), a powerful classification algorithm in many applications, is adopted. In the training stage, we train a multi-class SVM classifier with the training set. In the test stage, the SVM classifier not only predicts the most proper class for the given input, but also reports the confidence measures of all classes for the input. Thus, we can rank the departments by their confidence measures for an input instance. The LIBVSM[2] [4] is used in this work, and the kernel is L2-regularized logistic regression. The parameters are optimized with a grid search algorithm.

### 3.3 k-Nearest Neighbor Algorithm

We can predict the most suitable department based on collective intelligence embedded in medical summaries. The basic idea is that the department selection of patients can be regarded as their votes on the departments. The patients' experiences will give a hint to make decision. The instance-based k-nearest neighbor (kNN) algorithm is explored. We set the k to 5, which is the best value in the preliminary experiments. In contrast to SVM, the kNN algorithm is local sensitive. We will compare the performances of the three different ranking algorithms and show their differences.

### 3.4 Features

For each medical summary in the dataset, the chief complaint and the personal information extracted from the brief history are encoded as features.

***Chief Complaint***: A chief complaint is a sentence or a fragment of a sentence written in English. We first convert all the alphabetic characters in chief complaints to lowercase. All the words are segmented and performed with the Porter2[3] algorithm to obtain their root forms. For the language model, each chief complaint is represented as a sequence of words. For SVM and kNN, the bag-of-words representation is applied.

***Personal Attributes***: The personal attributes of a patient including the age, the gender, and the chronic diseases of the patient are extracted from the brief history. The beginnings of most brief histories are described in a similar pattern like "This X year-old man", "This X years old female", or "This X y/o woman". Therefore, age extraction can be formulated by a pattern matching. Similarly, the gender can also be captured

[1] http://www.speech.sri.com/projects/srilm/
[2] http://www.csie.ntu.edu.tw/~cjlin/libsvm/
[3] http://pypi.python.org/pypi/stemming/1.0

with a dictionary where the gender-related nouns including "female", "man", "girl", etc. are collected. Some instances are shown as follows.

(S9) This 70 year-old gentleman had past history of ....
(S10) This 42 y/o female was healthy before ...
(S11) This 64 y/o lady was well before. No weight loss, general malaise, or other subjective complaints was noted recently...

To extract the chronic diseases, we prepare a dictionary in which 21,030 terms about diseases are collected. The simple pattern matching algorithm does not work well due to the noises from the long patient histories. In other words, many diseases that do not really attack a patient may be also mentioned in the brief history as background information. For instance, the patient's family history is frequently described in the brief history, and the suspected diseases that have been ruled out will be also recorded. In order to avoid introducing these noises, we use a position sensitive algorithm to matching the patterns. The idea is that the real chronic diseases of a patient tend to be described in the first two sentences. For this reason, we only match the diseases that appear in the first two sentences.

# 4 Experiments and Discussion

## 4.1 Experimental Results

The experimental results of using language model, SVM, and kNN with all the features for each department are shown in Table 3. All the models in the experiments are evaluated by 10-fold cross validation. The reported metrics are precision, recall, and f-measure. The t-test is used for significance testing. P (%), R (%), and F (%) in the header denote precision, recall, and f-measure in percentage, respectively. SVM significantly outperforms the other two models at p=0.0001. The departments whose f-measures are more than 90% are dental, obstetrics & gynecology, ophthalmology, and psychiatry. The common property of these four departments is that the health troubles related to them are limited to definite parts of the body or the specific functions of the human being. For instance, the department of dental is related to the troubles of the mouth and teeth, and the department of psychiatry is related to the mental. The department of oncology is the one having the poorest performance. This is understandable because cancer can attack all organs in human body, and various symptoms may point to cancer. Fortunately, the poor performance of the department of oncology is not an issue in practice. In fact, the cases of cancer are usually found in the other specialist outpatients. For this reason, suggesting the patients with cancer potential for a specialist outpatient is not really wrong.

In the second experiment, we explore the performance of individual features used in the best model, the SVM classifier. The results are shown in Table 4. All the numbers in the table are f-measures in percentage. The CC in the header is the abbreviation of Chief Complaint. The additional gender information does not significantly improve the performance at p=0.05. One reason may be that many diseases are unrelated to a specific gender. On the other hand, the gender information is also useless

**Table 3.** Performance of the three algorithms with all the features

| Departments | Models | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Language Model | | | SVM | | | kNN | | |
| | P(%) | R(%) | F(%) | P(%) | R(%) | F(%) | P(%) | R(%) | F(%) |
| Dental | 80.72 | 90.35 | 85.26 | 91.44 | 85.17 | 88.19 | 82.38 | 79.08 | 80.69 |
| Dermatology | 53.62 | 79.02 | 63.89 | 82.53 | 67.74 | 74.41 | 75.07 | 50.26 | 60.21 |
| ENT | 86.82 | 83.39 | 85.07 | 85.67 | 86.78 | 86.22 | 58.02 | 81.69 | 67.85 |
| Internal Medicine | 73.11 | 75.54 | 74.30 | 73.67 | 82.01 | 77.62 | 69.90 | 73.24 | 71.53 |
| Neurology | 47.90 | 59.28 | 52.99 | 66.10 | 52.12 | 58.29 | 53.87 | 33.14 | 41.03 |
| Obstetrics & Gynecology | 94.34 | 90.40 | 92.33 | 94.84 | 91.83 | 93.31 | 88.10 | 86.42 | 87.26 |
| Oncology | 52.24 | 56.75 | 54.40 | 67.20 | 47.36 | 55.56 | 60.67 | 30.83 | 40.89 |
| Ophthalmology | 95.17 | 95.48 | 95.32 | 95.59 | 94.97 | 95.28 | 89.50 | 88.78 | 89.14 |
| Orthopedics | 84.51 | 85.27 | 84.89 | 85.32 | 87.27 | 86.29 | 68.90 | 84.72 | 76.00 |
| Pediatrics | 87.02 | 85.28 | 86.14 | 87.50 | 88.77 | 88.13 | 87.33 | 84.87 | 86.08 |
| Rehabilitation | 49.39 | 56.57 | 52.74 | 73.09 | 53.46 | 61.75 | 53.81 | 38.98 | 45.21 |
| Psychiatry | 87.45 | 93.36 | 90.31 | 94.38 | 88.86 | 91.54 | 94.77 | 69.27 | 80.04 |
| Surgery | 69.20 | 63.66 | 66.31 | 70.45 | 69.48 | 69.96 | 60.99 | 59.05 | 60.00 |
| Urology | 84.79 | 81.43 | 83.08 | 87.33 | 81.46 | 84.29 | 71.01 | 71.78 | 71.39 |
| Macro-Averaged | 74.73 | 78.27 | 76.22 | 82.51 | 76.95 | 79.35 | 72.45 | 66.58 | 68.38 |

**Table 4.** F-measures of the SVM classifier with different feature combinations

| Departments | Features | | | | |
|---|---|---|---|---|---|
| | CC | CC+Age | CC+Gender | CC+Diseases | All |
| Dental | 84.44% | 82.61% | 84.10% | 89.35% | 88.19% |
| Dermatology | 65.87% | 67.47% | 67.56% | 72.57% | 74.41% |
| ENT | 82.21% | 82.60% | 82.21% | 85.88% | 86.22% |
| Internal Medicine | 70.14% | 73.81% | 70.23% | 76.59% | 77.62% |
| Neurology | 47.92% | 49.18% | 48.45% | 58.06% | 58.29% |
| Obstetrics & Gynecology | 88.00% | 88.72% | 88.50% | 93.39% | 93.31% |
| Oncology | 40.38% | 41.66% | 41.04% | 54.45% | 55.56% |
| Ophthalmology | 93.34% | 93.33% | 93.43% | 95.34% | 95.28% |
| Orthopedics | 84.17% | 84.30% | 84.13% | 86.01% | 86.29% |
| Pediatrics | 66.73% | 85.58% | 67.14% | 84.13% | 88.13% |
| Rehabilitation | 56.40% | 57.34% | 57.08% | 58.55% | 61.75% |
| Psychiatry | 84.54% | 85.75% | 85.20% | 91.01% | 91.54% |
| Surgery | 64.17% | 64.47% | 64.17% | 69.75% | 69.96% |
| Urology | 80.48% | 81.37% | 81.01% | 84.08% | 84.29% |
| Macro-Averaged | 72.06% | 74.16% | 72.45% | 78.51% | 79.35% |

for the gender-specific departments such as obstetrics & gynecology. As expected, the feature of chronic diseases is very useful, especially for certain departments such as oncology and psychiatry. The overall performance of using the features Chief Complaint and chronic diseases (denoted as CC+Diseases in Table 4) is close to using all the features (denoted as All). However, using all the features still yields a significant improvement at p=0.0001.

The confusion matrix of SVM classification is shown in Table 5. Each row of Table 5 shows the percentages of instances of a department that are classified into the 14 departments. The two largest departments, Internal Medicine and Surgery, are the two departments that most instances are misclassified into. The reason is that these two departments are more general than the specific organ specialist departments. It is a common situation that a patient would be referred to the department of surgery from the department of internal medicine for surgery, and then would be referred to the internal medicine back for postoperative follow. Another common situation is that a patient of a specialist department would be referred to the department of surgery for surgery, and then would be referred to the department of oncology for chemotherapy after the surgery. For this reason, the specialist departments, the department of internal medicine, the department of surgery, and the department of oncology share a number of common patients.

A special case is that 10.22% of Rehabilitation instances are misclassified into Neurology. This is the most frequent misclassification besides the cases that Internal Medicine and Surgery are involved. A major disease of the department of neurology is the stroke (cerebrovascular accident, CVA), which might cause the inability to move limbs. To recover the loss body function of stroke patients, rehabilitation treatment is usually used on those patients. Thus, the patients of Neurology and Rehabilitation are higher overlapped. (S12) is an instance of a stroke patient who accepted the treatment from the department of rehabilitation.

(S12) Sudden onset right limb weakness and speech disturbance on 98/3/6.

**Table 5.** Confusion matrix of SVM classification

| Actual Class | Predicted Class | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Dent | Derm. | ENT | Med | Neur. | O&G | Onc | Ophth | Ortho | Pedi | Reha | Psyc | Surg | Urol |
| Dent | 85.17 | 0.10 | 2.84 | 1.83 | 0.05 | 0.10 | 0.36 | 0.05 | 0.46 | 0.76 | 0.00 | 0.00 | 7.87 | 0.41 |
| Derm | 0.52 | 67.74 | 0.87 | 16.78 | 0.17 | 0.26 | 0.44 | 0.26 | 1.66 | 1.05 | 0.00 | 0.00 | 9.88 | 0.35 |
| ENT | 0.23 | 0.05 | 86.78 | 3.73 | 0.06 | 0.05 | 0.97 | 0.19 | 0.13 | 0.61 | 0.00 | 0.00 | 7.03 | 0.17 |
| Med | 0.05 | 0.28 | 0.44 | 82.01 | 0.82 | 0.39 | 2.02 | 0.19 | 0.71 | 1.31 | 0.27 | 0.13 | 10.48 | 0.90 |
| Neur | 0.00 | 0.08 | 0.94 | 23.70 | 52.12 | 0.16 | 0.67 | 0.79 | 1.02 | 0.63 | 3.69 | 1.26 | 14.31 | 0.63 |
| O&G | 0.04 | 0.02 | 0.08 | 2.26 | 0.00 | 91.83 | 0.17 | 0.00 | 0.21 | 0.48 | 0.00 | 0.01 | 4.05 | 0.83 |
| Onc | 0.10 | 0.08 | 3.92 | 28.87 | 0.35 | 0.61 | 47.36 | 0.08 | 1.38 | 0.26 | 0.14 | 0.02 | 15.72 | 1.12 |
| Ophth | 0.07 | 0.02 | 0.23 | 1.04 | 0.07 | 0.12 | 0.07 | 94.97 | 0.39 | 0.35 | 0.00 | 0.00 | 2.45 | 0.23 |
| Ortho | 0.17 | 0.06 | 0.13 | 1.78 | 0.06 | 0.05 | 0.25 | 0.07 | 87.27 | 0.53 | 0.13 | 0.00 | 9.37 | 0.15 |
| Pedi | 0.13 | 0.07 | 0.29 | 4.23 | 0.14 | 0.22 | 0.23 | 0.08 | 0.34 | 88.77 | 0.05 | 0.06 | 5.25 | 0.15 |
| Reha | 0.00 | 0.00 | 0.32 | 16.70 | 10.22 | 0.19 | 0.32 | 0.00 | 3.62 | 0.76 | 53.46 | 0.06 | 12.70 | 1.65 |
| Psyc | 0.06 | 0.11 | 0.28 | 5.46 | 1.07 | 0.17 | 0.11 | 0.00 | 0.06 | 0.51 | 0.11 | 88.86 | 2.98 | 0.23 |
| Surg | 0.27 | 0.15 | 1.50 | 18.80 | 0.68 | 0.58 | 1.20 | 0.24 | 2.92 | 2.61 | 0.33 | 0.04 | 69.48 | 1.17 |
| Urol | 0.06 | 0.03 | 0.21 | 6.75 | 0.03 | 1.05 | 0.44 | 0.10 | 0.28 | 0.33 | 0.15 | 0.00 | 9.12 | 81.46 |

## 4.2 Discussion

Our department recommendation system is an aiding tool for the patients seeking the proper outpatient department. Even if the accuracy of the recommendation is not perfect, some incorrect suggestions will not cause serious risk in practice. If a patient takes a wrong department, the physician will find this error and give an appropriate treatment. Such a change can also be a feedback to refine the recommendation system. In general, the proposed system provides decent and efficient suggestions to most patients.

Although the chief complaints are said by patients, all the chief complaints in our dataset are recorded by physicians. The physicians try to record the patients' words as realistic as possible, but the gap between experts and the public cannot be ignored. Different choices of words by the general public may decrease the performance of this system in real world if the inputs are done by patients themselves. How to narrow down the gap is a future work.

Furthermore, all the chief complaints in our datasets are written in English. In other words, the knowledge extracted from the medical summaries is in terms of English. For serving the patients speaking in other languages, e.g., Chinese, a medical summary translation system [5] should be integrated into the recommendation system. We will develop a multilingual outpatient department recommendation system in the future. The chief complaints in Chinese will be translated into English, and submitted to our outpatient department recommendation system. Vocabulary gap and translation accuracy should be tackled at the same time.

## 5 Conclusion

In this paper, we introduce the collective intelligence embedded in medical summaries to design an outpatient department recommendation system. This system suggests a patient an outpatient department according to the patient's chief complaint and personal attributes. Three different learning models and various feature combinations are proposed and evaluated. In the experiments, the SVM classifier with all features achieves an f-measure of 79.70%. This system aids the patients without medical knowledge to seek the proper outpatient departments. That reduces the risk of misdiagnosis and helps save the expensive medical resources. In the future work, we will recommend outpatient departments of much finer grain, in particular, for the two largest departments, i.e., Internal Medicine and Surgery. Besides, resolving the information gap in practical uses and introducing the multilingual capabilities into the proposed system will be investigated.

**Acknowledgment.** Research of this paper was partially supported by National Science Council (Taiwan) under the contract NSC 101-2221-E-002-195-MY3.

## References

1. Mondal, D., Gangopadhyay, A., Russell, W.: Medical decision making using vector space model. In: ACM SIGHIT 2010, pp. 386–390 (2010)
2. Zhang, Y., Wang, P., Heaton, A., Winkler, H.: Health information searching behavior in MedlinePlus and the impact of tasks. In: ACM SIGHIT 2012, pp. 641–650 (2012)
3. Sondhi, P., Vydiswaran, V.G.V., Zhai, C.: Reliability Prediction of Webpages in the Medical Domain. In: Baeza-Yates, R., de Vries, A.P., Zaragoza, H., Cambazoglu, B.B., Murdock, V., Lempel, R., Silvestri, F. (eds.) ECIR 2012. LNCS, vol. 7224, pp. 219–231. Springer, Heidelberg (2012)
4. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2, 27:1–27:27 (2011)
5. Chen, H.-B., Huang, H.-H., Tan, C.-T., Tjiu, J., Chen, H.-H.: A statistical medical summary translation system. In: ACM SIGHIT 2012, pp. 101–110 (2012)