



Predicting associated statutes for legal problems



Yi-Hung Liu^{a,*}, Yen-Liang Chen^a, Wu-Liang Ho^b

^a Department of Information Management, National Central University, Chung-Li 320, Taiwan, ROC

^b Department of Legal Service, Straits Exchange Foundation, Taipei 105, Taiwan, ROC

ARTICLE INFO

Article history:

Received 16 October 2013

Received in revised form 29 May 2014

Accepted 9 July 2014

Available online 10 August 2014

Keywords:

Text mining

Statute

Criminal judgment

Normalized Google Distance (NGD)

Support vector machines (SVM)

Apriori algorithm

ABSTRACT

Applying text mining techniques to legal issues has been an emerging research topic in recent years. Although a few previous studies focused on assisting professionals in the retrieval of related legal documents, to our knowledge, no previous studies could provide relevant statutes to the general public using problem statements. In this work, we design a text mining based method, the three-phase prediction (TPP) algorithm, which allows the general public to use everyday vocabulary to describe their problems and find pertinent statutes for their cases. The experimental results indicate that our approach can help the general public, who are not familiar with professional legal terms, to acquire relevant statutes more accurately and effectively.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The law represents social norms, which protect civilian rights and maintain social order. To achieve these objectives, the legislature enacts legal provisions preserving people's rights to life, property, and so on. As an increasing number of individuals' rights have been violated, more and more litigation has occurred.

When people have legal issues, it is critical to know which statutes are involved. Usually, because of a lack of legal knowledge, they may seek help from legal experts, such as attorneys, or from automatic systems. Since legal consultation is very costly, automatic systems are a much more affordable form of legal support. They involve utilizing search engines on the Internet or searching legal databases, such as [Westlaw International \(2013\)](#) and [LexisNexis \(2013\)](#). Although these automatic systems provide query methods, users cannot obtain pertinent statutes through simple case statements. This motivated us to propose a new approach that will help laypeople obtain relevant statutes by simply stating their problem or case using daily vocabulary, and without the help of legal experts. In [Fig. 1](#), we show the proposed statute retrieval approach.

The purpose of this research is to provide a statute retrieval method that will help people deal with their legal problems more effectively. It can be helpful to at least two types of people. First, for specialists, this method can reduce workloads and be used as a reference when dealing with legal cases. As the sheer number of legal cases increase, legal experts need to expend more time and energy in their work. This research provides an aid for efficiently processing cases. Second, for laypeople, this approach can reduce searching and consultation needs. When laypeople have legal issues, it is difficult for them to acquire related statutes through existing automatic systems because of their insufficient knowledge of professional legal terms. This results in incorrect search results and lengthens the searching process. With the help of our approach, this insufficient knowledge problem can be alleviated.

* Corresponding author.

E-mail address: lyh0315@gmail.com (Y.-H. Liu).

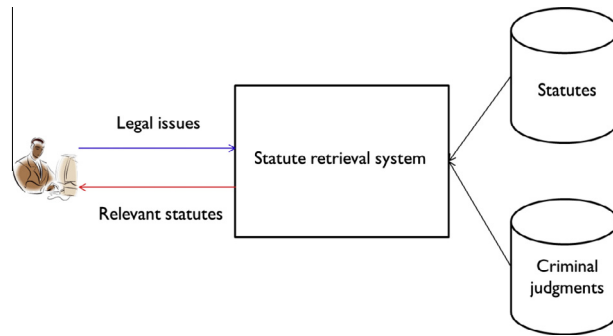


Fig. 1. Framework of relevant statute retrieval.

In recent years, text mining research has gotten more and more attention. Basically, text mining is the procedure of uncovering salient features and information from textual data. Since most human knowledge is stored in text, abundant text mining applications and methods have been developed. Examples of these text mining applications include patent retrieval (Chen & Chiu, 2011; Tikk, Biró, & Töröcsvári, 2007), e-mail security (Bergholz et al., 2010; Wei, Chen, & Cheng, 2008), news categorization (Calvo, 2001; Zheng, Miliotis, & Watters, 2002), authorship identification (Stamatatos, 2009; Zheng, Li, Chen, & Huang, 2006), scientific document retrieval (Kaur, Yusof, Boursier, & Ogier, 2010), document sentiment analysis (Li & Wu, 2010; Schumaker, Zhang, Huang, & Chen, 2012), document summarization (Goldstein, Mittal, Carbonell, & Kantrowitz, 2000; Li, Du, & Shen, 2013; Wang, Zhu, Li, & Gong, 2009), online advertisement recommendations (Thomaidou & Vazirgiannis, 2011; Wang, Wang, Duan, Tian, & Lu, 2011), search engines (Kawai, Jatowt, Tanaka, Kunieda, & Yamada, 2011; Yin, 2007), etc.

Text mining has been applied in various areas. Although a few past studies applied text mining techniques to the legal domain (Chen & Chi, 2010; Chou & Hsing, 2010; Conrad & Schilder, 2007; Moens, 2001), all of them focused only on helping professional users retrieve or classify legal documents. None of them considered how to help laypeople retrieve relevant statutes from a case statement using daily customary terms. Therefore, this research aims to develop a framework for statute prediction that will remedy this problem. This framework is built with judgments and statutes. Judgments are included in the framework because they contain the facts of the crime and the cited statutes from the judge’s adjudication. From them, we can find the connections between the problem and the cited statutes. In turn, these connections help us to determine the most relevant statutes with respect to the user’s problem.

The prediction method process is shown in Figs. 2 and 3, Batch and Online, respectively. In the Batch process, three outputs are generated to be used in the Online process. The first output is a classification model that classifies cases to statutes, and is produced by adopting a SVM (support vector machine) classifier. In the second output, all statutes are represented as statute vectors. The last output is a set of association rules, which show what statutes frequently occur together, and is generated from the training collection of judgments. In the Online process, the classification model is adopted to acquire the prediction of top k_1 statutes for the user query. Then, the NGD (Normalized Google Distance) method (Cilibrasi & Vitanyi, 2007; Evangelista & Kjos-Hanssen, 2006) is used to perform terms transformation between the statutes and user query, and the top k_2 most similar statutes are selected. Finally, by applying associative statute rules to the top k_2 statutes, the statute weight computation metric is defined, so as to obtain the most relevant statutes for the user query.

The advantages of our approach are that (1) ordinary users can express their cases using daily vocabulary, (2) a bridge is created between laypeople and legal statutes, and (3) the most pertinent statutes are recommended to users. This work acquires relevant statutes by developing a three-staged algorithm. In the first stage, we utilize the multi-label SVM text

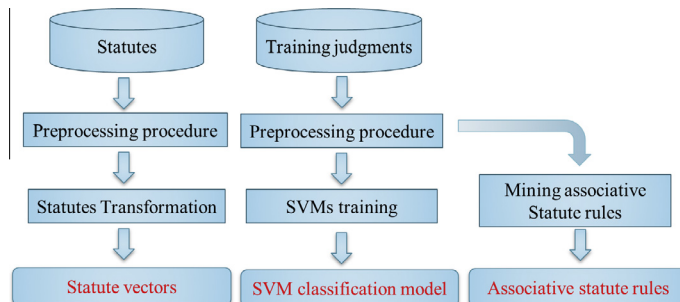


Fig. 2. Batch process of the prediction approach.

Download English Version:

<https://daneshyari.com/en/article/515385>

Download Persian Version:

<https://daneshyari.com/article/515385>

[Daneshyari.com](https://daneshyari.com)