Machine Learning in Incident Categorization Automation

Sara Silva, Rubén Pereira
Instituto Universitário de Lisboa (ISCTE-IUL)
Lisboa, Portugal
{satsa, ruben.filipe.pereira}@iscte-iul.pt

Ricardo Ribeiro
Instituto Universitário de Lisboa (ISCTE-IUL), Portugal
L2F, INESC-ID Lisboa, Portugal
ricardo.ribeiro@iscte-iul.pt

Abstract — **IT incident management process requires a correct categorization to attribute incident tickets to the right resolution group and obtain an operational system as quickly as possible, having the lowest possible impact on the business and costumers. In this work, we introduce a module to automatically categorize incident tickets, turning the responsible teams for incident management more productive. This module can be integrated as an extension into an incident ticket system (ITS), which contributes to reduce the time wasted on incident ticket route and reduce the amount of errors on incident categorization. To automate the classification, we use a support vector machine (SVM), obtaining an accuracy of 89%, approximately, on a dataset of real-world incident tickets.**

Keywords - Machine learning, Automated incident categorization, SVM, Incident management, Incident management process, Categorization.

I. INTRODUCTION

Information Technology Service Management (ITSM) is a discipline for managing IT operations and services [1]. Several ITSM frameworks exist to help organizations deal with alignment and management of IT services, in conformity with the business needs. These framework aims at improving business performance through the best IT service delivery. Thus, ITSM focuses on development of methodologies and tools to provide an efficient and high quality service [2], including optimizing IT services and business operations, increasing employees' productivity and costumers' satisfaction.

The Incident Management (IM) process, one of the most important components of ITSM, focuses on tracking and managing all incidents, from opening until closure. Incidents must be resolved as quickly as possible to ensure the minimum business impact for costumers and the correct operation of organizations' IT services [3]. The IM is the process of ITSM that provides most directly visible gains to service quality as well cost reduction [4].

With the exponential usage of IT in companies, it was possible to verify a weakness in support service customers [5]. In companies, a large number of tickets is created every day, and specific IT teams exist to resolve these tickets. However, in many cases this process is not entirely systematic and may be incoherent and inefficient [6]. This process of incidents support and resolution results in a waste of several resources increasing companies' costs [7]. Therefore, to be competitive, companies need an efficient and cost-effective service delivery and support [8]. Consequently, many companies started to adopt tools to help and support teams that are responsible for the IM process [2]. Such tools are software systems used in organizations to record and track all incidents and typically refer to an Incident Ticket System (ITS).

A coordinated ITS provides a positive effect on the efficiency of the IM process, which in turn improves and increases companies' revenue. Most ITS's follow the Information Technology Infrastructure Library (ITIL) [9], the most adopted ITSM framework to facilitate and help the decision-making process [6]. ITIL depicts best practices and standards to IM, helping companies to improve their processes. An ITS represents a significant contribution to an efficient IM process, in order to obtain lower costs and an increased organization growth.

In order to answer companies' needs, this work proposes a module for the automatic classification of incidents on ITS, by developing a method with machine learning (ML) techniques.

The manual classification of incidents implies a wrong delegation of incidents, assigning them to resolution groups that are not capable of solving them, causing delays in the whole process of dispatch [10]. Incidents are forwarded to finally be addressed to the right resolution group, which impacts incident route negatively, which results on the usage of more resources. Consequently, it leads to a waste of time and customer dissatisfaction [11].

In order to attain a right assignment, it is crucial to have an appropriate incident classification, the process that assigns a suitable category to an incident, so they are routed more accurately [9]. Automating incident classification process means avoiding human error, reducing the waste of resources and avoiding incorrect routing due to wrong classification [12].

The remainder of this document consists of six sections that are structured as follows. The second section presents some theoretical background needed to develop the research. The proposed method is presented in section 3. The implementation is described in the fourth section. Section 5 presents the obtained results. Finally, the document closes with the conclusions and some possible future work.

II. THEORETICAL BACKGROUND

A. Incident Management

ITIL defines incident as "an unplanned interruption to an IT service or reduction in the quality of an IT service". These incidents can be related with failures, questions, or queries and should be detected as early as possible [9].

IM is the process responsible to manage disruptions, a crucial factor to create a high scalable system [13] and is responsible for restoring the normal operation, finding as quickly as possible a resolution for the incident, minimizing business impact [9][13][14]. To obtain the success and efficiency of the process, there are four critical success factors that must be achieved, such as quickly resolving incidents, maintaining IT service quality, improving IT and business productivity, and, finally, maintaining user satisfaction [15]. So, when a disruption on the system is detected, by the system or by users, several activities follow [9]. Table 1 presents the activities that compose IM process.

Table 1: IM Process Activities

Activity	Description
Incident detection and recording	An incident must be recorded as soon as possible after being detected, if possible before of cause impact for users.
Classification and initial support	**Incident categorization:** The incident type should be correctly assigned to the incident
	Incident Prioritization: This process deals with address urgency and an impact on an incident
Investigation and diagnosis	In this step is done the incident escalation, which includes an initial diagnosis to find a resolution. If the resolution is identified, the incident is solved, otherwise the incident is escalated for other support resolution group
Resolution and recovery	In this step, the previously identified resolution must be tested in order to ensure that the system is operational
Incident closure	In the closing of an incident it is necessary to check if the categorization done in step two is correct, if users are satisfied with the respective resolution, and if the documentation related to the incident is correct

Our work focuses on automating the incident categorization process included in the second activity.

During incident recording, there are several fields that are filled and that compose the incident ticket. These fields are the incident attributes and they compose our training data. One of these fields is incident description, which is represented in natural language such as English. To automate incident categorization, it is necessary to process the text of description.

B. Text Categorization

Text categorization (TC), also known as text classification, is one of the applications of text mining [16] and is the process that deals with the assignment of pre-defined categories, topics or labels to natural language texts or documents [17]. Automated TC is a supervised learning task [18] that uses Machine Learning to learn classifiers from examples that perform the categorization automatically [19]. Given a set of documents $D = \{d_1, \ldots, d_n\}$ with assigned categories $C = \{c_1, \ldots, c_n\}$ and a new document d, a learned classifier on these data will predict which category should be assigned to document d.

There are several approaches used in TC, which differ on how they represent documents and decide to assign a category to a document [20]. TC can be divided into two types of classification: binary and multi-class. A binary problem is when a document must be assigned to one of two categories. Multi-class problems can be seen as two types of problem: single-label and multi-label. The first consists on assigning the document to exactly one of a set of pre-defined categories. In multi-label classification, the data are assigned to more than one label at the same time. Our problem is clearly a single-label problem.

The main steps involved in TC are the document pre-processing and feature extraction, model selection, and training and testing the classifier [21]. These steps are presented and described in the next section.

III. RELATED WORK

Over the years, approaches that automate IM process has being studied and developed. One of these approaches is automate the incident classification process that is the purpose of our work. In this section we describe relevant work developed in this area, which algorithms are used and what results were obtained with the respective implementations.

To automate IM, the authors of [4] use ML and information integration techniques to develop an algorithm for correlate incoming incidents. The authors used the incident description to extract keywords and their annotations as features. SVM is the algorithm used to attributes a category to an incoming incident. With this approach was obtained the list of keywords which better identifies each category.

In [22], the authors used for the same problem, incident tickets categorization, SVM, KNN, decision trees and Naïve Bayes. They used 4 different datasets with different categories

to assess which performance is obtained. For this classification, they present three approaches assigned to each algorithm. They present the accuracy results with TF-IDF, with only TF and with Boolean weighting, which attributes the weight of 1 if the term is present in the document, otherwise attributes 0. On average, SVM had accuracy results of 90% approximately, in the three approaches. KNN achieves 75% of accuracy, also with the three approaches. Decision trees have similar results in the three approaches, around of 90%. Finally, Naïve Bayes is the unique which present different results among Boolean weighting with 85% and TF-IDF and only TF, both with 55% of accuracy.

To automate ITS, in [23] the authors resort of two algorithms such as Naïve Bayes (NB) and NN (Neural Network). Related to NB, one of the variants is Multinomial NB, which was used in this research. In NN implementation, they use the Softmax NN. For these two implementations, the authors use for training a dataset composed by 7042 tickets and 717 for test set. The goal is assign a ticket to a tag. The tag indicates which category the ticket should be assigned. Both algorithms use to feature generation, an input word list composed by email subjects and only words present in this list are extracted of the tickets and used as input to Multinomial NB and Softmax NN. Multinomial NB outperformed Softmax NN, achieving 85.8% and 80.7% of accuracy, respectively.

[12] focus on three big features to automate IM process and one of them is automate ticket classification. In this work, they present an approach based on ML for automate the classification, consisting on analyze incident descriptions writhed in natural language. The approach consists on create automatically a classifier using pre-classified incidents. They denote a sequence of words as incident features and then they use Naïve Bayesian probability to find a feature in an incident that belongs to a specific category. Per this probability, they defined the probability of categorizing incoming incidents. The probability of assign a category to an incident is calculated for all categories. Finally, the incoming incident is assigned to the category that has the maximum value. To ensure that is done a classification so correctly as possible is only assign a category if the maximum probability overcome a defined threshold. The authors have results in 70% accuracy with 1000 features. In agreement with an analysis of IBM internal tools based and built on similar conditions indicates reduction tickets resolution times by over 25%.

IV. Proposed Method

The proposed extension to the IM process consists in assigning the suitable category to an incident, which is the type of a specific incident. In this method it will be necessary to perform text pre-processing, because the incident description is represented in natural language, such as English.

Text pre-processing starts with the text tokenization, which splits the different incidents descriptions in all words that compose them. Then, with the resulting dictionary composed by all different words that are present in descriptions is applied a stop-words list, which is responsible for eliminating the words that are not meaningful for classification. The other applied process is stemming, which reduces the words to their base form, reducing the dictionary.

After the text pre-preprocessing is finished, all descriptions are represented by a feature vector, selecting the keywords that are relevant to identify a document and remove features that are irrelevant for classification. For that we used the term frequency-inverse document frequency (TF-IDF) t2hat consists in assigning to each term a weight based on the frequency of the term in the document. This weight increases with the number of times the term occurs, but is offset by the document frequency of the term in the corpus [22].

This text representation approach is the most used in literature due to the performance achieved with several types of classifier, including SVM [22].

This process leads to a smaller dataset and consequently smaller computational requirements for the TC algorithms, which is crucial to achieve success in the next stage [24].

The learning process is based on Support Vector Machine (SVM) and K-Nearest Neighbor (KNN).

A. SVM Algorithm

An SVM is a method of supervised learning and was introduced in text categorization, between 1998 and 1999 by Joachims [25]. SVM consists in mapping input vectors into a high dimensional space and outputs the creation of a hyperplane [26]. SVM is an appropriate technique for text categorization, proving is more robust than other conventional techniques of text classification [27].

B. KNN Algorithm

The KNN algorithm uses the Euclidean distance as the distance metric to identify which are the K-nearest neighbors of the instances. The class of each instance is determined using a majority vote. This algorithm is considered to be simple, easy to implement [28] and a popular one in text categorization [29]. This algorithm with TF-IDF achieves good performance results in a text categorization case study.

V. Implementation

In this work, we used a dataset provided by a company that due to privacy issues cannot be mentioned. The dataset is composed by 10000 incident tickets.

To train the classifiers, we used a dataset composed by incident descriptions correctly classified with an appropriate category. In this dataset each incident has, in addition to the description, the following information: the caller id, which ordinarily is the person that opens the ticket, the severity and the contact source, which describes the way the incident was reported, by e-mail and telephone.

We explored three approaches of categorization for each algorithm. In the first approach, we use only the attributes previously described: the caller id, severity and the contact source, without description. In the second approach, incidents

are categorized using just the incident description attribute. Then, in the last approach all attributes of each incident ticket are used.

The three approaches make it possible to observe how both algorithms lead with only nominal attributes, with only textual data and finally gathering all attributes. The main goal of select these three approaches is to verify the impact of introducing the description on the incident categorization and consequently verify which are the crucial attributes to obtain a good performance.

Table 2 shows an example of one incident ticket that composes the used dataset.

Table 2: Incident ticket example

Description	Caller id	Severity	Contact Source	Category
"unable to connect on wifi"	client x	3 – "Low"	phone	network

All incidents are assigned to one of the 10 following categories: application, collaboration, enterprise resource planning (ERP), hosting services, network, security and access, output management, software, workplace and support.

The obtained results for the three approaches are presented in the next section and are divided by the algorithm used.

VI. Performance Results

To train and test the method we use cross-validation, which consists on dividing the whole training set into *n* subsets of equal size. One subset is used to test the classifier, which is trained with the remaining *n* – 1 subsets of the complete dataset. This process prevents overfitting, due to the fact that training sets are independent of the test set [30] [31].

We used a dataset composed by 10000 instances, where each 1000 are assigned to a specific category.

Figure 1 presents the accuracy after applying the SVM algorithm.

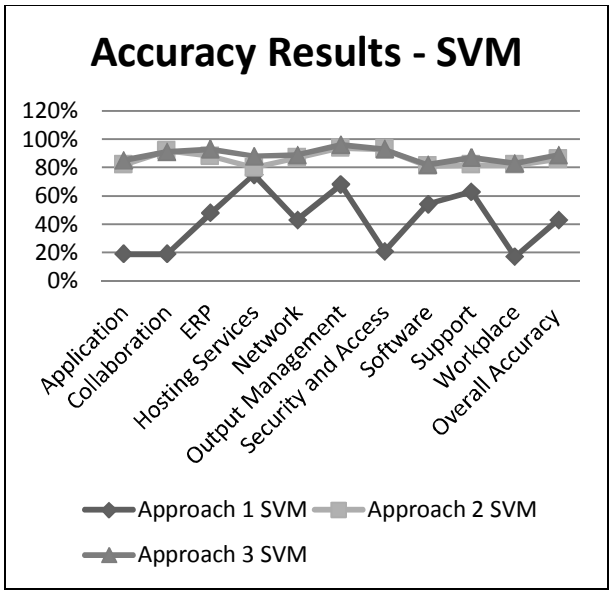

Figure 1 : Accuracy results (SVM)

Considering the results presented in Figure 1, it is possible to conclude that the incident description has an important role in the categorization of incidents. Training the data with only the incident description has an accuracy result of 86%. Using the other attributes, caller id, severity and contact source, leads to worse results, with the accuracy dropping to 43%. However, there are categories for which were obtained high accuracy values like output management and hosting services. The model trained with all attributes achieved an accuracy of 89%, a 3% p.p. increase comparatively with the model trained only with incident description.

Table 2 : Overall Accuracy - SVM

Overall Accuracy		
Approach 1	**Approach 2**	**Approach 3**
43%	86%	89%

Figure 2 shows the results for the KNN algorithm.

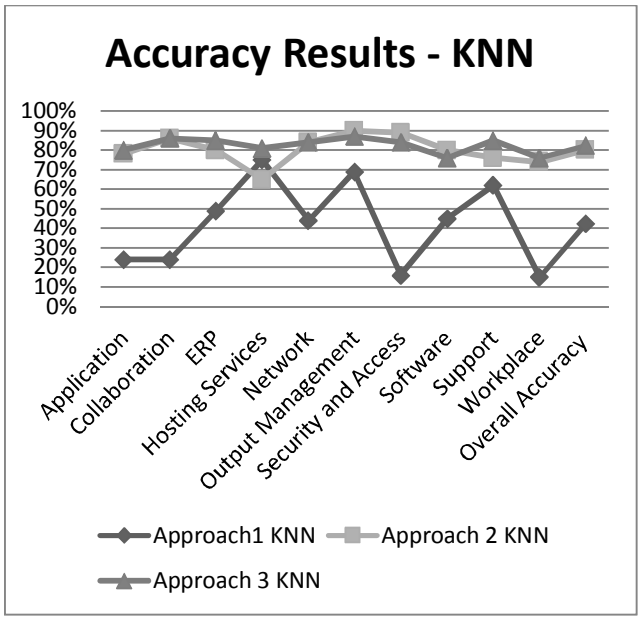

Figure 2: Accuracy Results (KNN)

With KNN algorithm, the results regarding the first two approaches (other attributes only; description only) are very similar to the results achieved with the SVM, especially for approach 1. When using only the incident description, the results increase to 80%, however they are not as high as with the SVM algorithm.

Table 3: Overall Accuracy k-NN

Overall Accuracy		
Approach 1	**Approach 2**	**Approach 3**
42%	80%	82%

Using all attributes in the KNN categorization lead to an accuracy of 82 %.

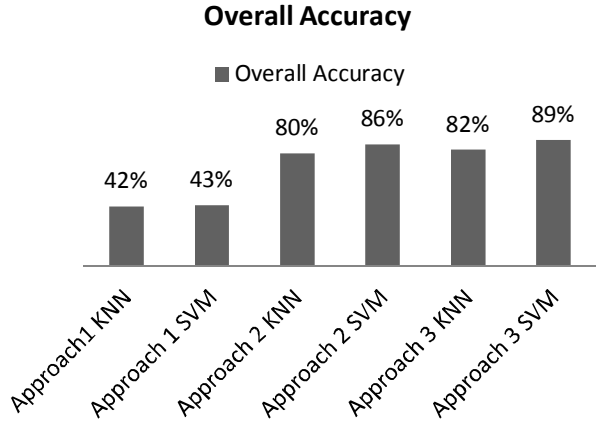

Figure 3: Overall accuracy with SVM & KNN

In approach 1, the obtained results are very similar between the two algorithms, demonstrating almost the same accuracy. Comparing the results obtained with SVM and KNN in approach 2, is possible to verify that with only the incident description SVM it overtakes KNN, with 6% difference. Finally, when the textual data is introduced in approach 1 (approach 3), SVM has one more time better performance than KNN.

The results obtained in approach 2 note that SVM leads better with textual data, therefore it is possible to understand why SVM represents a better performance on the categorization.

VII. Conclusions and Future Work

The categorization of incidents plays an important role in the IM process and to all teams, which are responsible to manage the process. The manual categorization is a process that takes time, which can be reduced with automated categorization. Besides of taking time, most of the times incidents are assigned to an incorrect category. With this work, it is possible to conclude which are the most critical attributes of incidents that determine how to obtain a good performance on categorization. The results obtained in this research are the base to progress with future work.

We believe that the proposed module will have a positive impact in the categorization process, reducing the errors of incident categorization. This is determinant to obtain a correct assignment and consequently reduce the time wasted and improve the whole incident route.

As future work, we will integrate the module into an ITS in a specific company. With the integration, we hope to verify by interviews to the IT teams responsible for the IM process, the impact of the proposed extension. It is also intended to extend the categorization to the whole activity of classification and initial support in IM process, which includes automating the assignment of a priority and urgency to incidents. Moreover, we pretend to automate the resolution and recovery activities, finding and suggesting automatically a possible resolution to an incoming incident.

References

[1] S. D. Galup, R. Dattero, J. J. Quan, and S. Conger, 'An overview of IT service management', *Commun. ACM*, vol. 52, no. 5, p. 124, 2009.

[2] P. Marcu, G. Grabarnik, L. Luan, D. Rosu, L. Shwartz, and C. Ward, 'Towards an optimized model of incident ticket correlation', *2009 IFIP/IEEE Int. Symp. Integr. Netw. Manag. IM 2009*, pp. 569–576, 2009.

[3] R. Gupta, K. H. Prasad, and M. Mohania, 'Information integration techniques to automate incident

management', *NOMS 2008 - 2008 IEEE Netw. Oper. Manag. Symp.*, pp. 979–982, 2008.

[4] R. Gupta, K. H. Prasad, and M. Mohania, 'Automating ITSM incident management process', *5th Int. Conf. Auton. Comput. ICAC 2008*, vol. 1, pp. 141–150, 2008.

[5] T. Dias Freire de Mello and E. C. Lopes, 'Using case-based reasoning into a decision support methodology for the incident resolution control in IT', in *2015 10th Iberian Conference on Information Systems and Technologies (CISTI)*, 2015, pp. 1–6.

[6] S. Salah, G. Maciá-Fernández, J. E. Díaz-Verdejo, and L. Sánchez-Casado, 'A Model for Incident Tickets Correlation in Network Management', *J. Netw. Syst. Manag.*, vol. 24, no. 1, pp. 57–91, 2016.

[7] Y. Song, A. Sailer, and H. Shaikh, 'Problem classification method to enhance the ITIL incident and problem', *2009 IFIP/IEEE Int. Symp. Integr. Netw. Manag. IM 2009*, pp. 295–298, 2009.

[8] W. Zhou, W. Xue, Q. Wang, and L. Shwartz, 'STAR : A System for Ticket Analysis and Resolution', *KDD 2017 Appl. Data Sci. Pap.*, pp. 2181–2190, 2017.

[9] D. Cannon and D. Wheeldon, *ITIL Service Operation*. 2007.

[10] S. Agarwal, R. Sindhgatta, and B. Sengupta, 'SmartDispatch: enabling efficient ticket dispatch in an IT service environment', *Proc. 18th ACM ...*, pp. 1393–1401, 2012.

[11] Q. Shao, Y. Chen, S. Tao, X. Yan, and N. Anerousis, 'Efficient Ticket Routing by Resolution Sequence Mining', pp. 605–613, 2008.

[12] R. Gupta, K. H. Prasad, L. Luan, D. Rosu, and C. Ward, 'Multi-dimensional knowledge integration for efficient incident management in a services cloud', *SCC 2009 - 2009 IEEE Int. Conf. Serv. Comput.*, pp. 57–64, 2009.

[13] M. L. Abbott and M. T. Fisher, *The Art of Scalability: Scalable Web Architecture, Processes, and Organizations for the Modern Enterprise*. 2009.

[14] S. Operation, *ITIL Version 3 Service Operation*. .

[15] R. A. Steinberg, *Measuring Itsm: Measuring, Reporting, and Modeling the It Service Management Metrics That Matter Most to It Senior Executives*. Trafford Publishing, 2013.

[16] S. Vijayarani, J. Ilamathi, and M. Nithya, 'Preprocessing Techniques for Text Mining - An Overview', *Int. J. Comput. Sci. Commun. Networks*, vol. 5, no. 1, pp. 7–16, 2015.

[17] F. Sebastiani, 'Machine learning in automated text categorization', *ACM Comput. Surv.*, vol. 34, no. 1, pp. 1–47, 2002.

[18] Y. Yang and X. Liu, 'A re-examination of text categorization methods', in *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '99*, 1999, pp. 42–49.

[19] T. Joachims, 'Text categorization with Support Vector Machines: Learning with many relevant features BT - Machine Learning: ECML-98', 1998, pp. 137–142.

[20] A. Cardoso-Cachopo and A. L. Oliveira, 'An Empirical Comparison of Text Categorization Methods', in *Proceedings of SPIRE-03, 10th International Symposium on String Processing and Information Retrieval*, 2003, pp. 183–196.

[21] M. K. Dalal and M. A. Zaveri, 'Automatic Text Classification: A Technical Review', *Int. J. Comput. Appl.*, vol. 28, no. 2, pp. 37–40, 2011.

[22] M. Altintas and A. C. Tantug, 'Machine learning based volume diagnosis', *Proc. Int. Conf. Artif. Intell. Comput. Sci. (AICS 2014)*, no. September, pp. 195–207, 2014.

[23] G. Son, V. Hazlewood, and G. D. Peterson, 'On Automating XSEDE User Ticket Classification', 2014.

[24] M. Ikonomakis, S. Kotsiantis, and V. Tampakas, 'Text classification using machine learning techniques', *WSEAS Trans. Comput.*, vol. 4, no. 8, pp. 966–974, 2005.

[25] T. Joachims, 'Text categorization with support vector machines: Learning with many relevant features', *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1398, pp. 137–142, 1998.

[26] V. N. Vapnik, *The Nature of Statistical Learning Theory*, vol. 8, no. 6. 2000.

[27] T. Joachims, '1 Introduction 2 Text Categorization 3 Support Vector Machines', *Mach. Learn.*, vol. 1398, no. LS-8 Report 23, pp. 137–142, 1998.

[28] Y. Song, J. Huang, D. Zhou, H. Zha, and C. L. Giles, 'IKNN: Informative K-Nearest Neighbor Pattern Classification', *Proc. Eur. Conf. Princ. Pract. Knowl. Discov. Databases*, pp. 248–264, 2007.

[29] B. Trstenjak, S. Mikac, and D. Donko, 'KNN with TF-IDF based framework for text categorization', *Procedia Eng.*, vol. 69, pp. 1356–1364, 2014.

[30] S. Arlot and A. Celisse, 'A survey of cross-validation procedures for model selection', vol. 4, pp. 40–79, 2009.

[31] and C.-J. L. Chih-Wei Hsu, Chih-Chung Chang, 'A Practical Guide to Support Vector Classification', *BJU Int.*, vol. 101, no. 1, pp. 1396–400, 2008.