

Automatic Discrimination between Biomedical-Engineering and Clinical-Medicine Papers Based on Decision-Tree Algorithms: How Does the Term Usage Differ?

Motoki Sakai

Chiba University of Commerce Faculty of Policy Informatics

Abstract— Biomedical engineering (BM) is a successful example of integrated research. This research area is concerned with solving problems in clinical-medicine (CM) research using techniques such as information engineering. In this research field, novice investigators sometimes have difficulty in searching for and retrieving BM papers, because both BM and CM research papers contain common terms, such as disease names, so a novice researcher cannot retrieve only BM papers from the search results. Thus, this research proposes a decision-tree and random-forest-based method to automatically discriminate between BM and CM papers, and reveals a difference in term usage between BM and CM papers. The discrimination between BM and CM papers was examined by collecting papers containing five common terms: obstructive sleep apnea syndrome (OSAS), T-wave alternans (TWA), late potential (LP), epilepsy (EPY), and event-related potential (ERP). The gathered BM and CM papers were converted into document-term (D-T) matrices, and were discriminated with the decision-tree or random-forest algorithm. Results showed that the decision tree discriminated them with approximately 80% averaged accuracy and sensitivity and approximately 70% specificity, and the random forest discriminated them with approximately 90% averaged accuracy, sensitivity, and specificity. In addition, it was revealed that the terms “signal”, “detection”, “method”, “based”, “patient”, and “with” were effective for discriminating between BM and CM papers.

Keywords— Biomedical engineering, clinical medicine, discrimination, decision tree, random forest

I. INTRODUCTION

Sometimes, researchers in segmented academic fields have difficulty finding new ideas or solutions only from their field. A research field integrating various academic disciplines holds the possibility of discovering new academic insights or offering a new effective solution for a certain problem. Collaborative research among different academic fields has broadened researchers' horizons, connected different logics or techniques, and overcome limitations in a research field.

Cooperation between clinical-medicine (CM) and engineering research areas is a typical successful example of integrated research; this collaborative research field is called biomedical engineering (BM). (To be exact, BM research is also a type of segmented academic field and is an entrenched research area. However, this is actually nothing more than collaborative research between CM and engineering study areas.) BM is a research field that resolves problems in the CM research area using information technology, mechanical engineering, or electrical and electronic engineering techniques. (BM studies relate extensively to several engineering-research areas, but this study focuses only on the information-engineering research field.) Concretely, BM researchers attempt to automate medical diagnoses or extract information about an imperceptible phenomenon using different signal-processing techniques. In general, BM researchers first collect information on the background of a certain CM research area because they have problems in common with CM researchers. Next, they survey and read previous BM research papers related to the CM research to crystallize their study at an early stage.

As it is, novice BM researchers, such as undergraduate students, have trouble surveying related BM research papers, because the BM area is relevant to both information engineering and CM studies, and BM and CM studies have common academic terms. Let us assume that a common academic term is “arrhythmia.” For instance, one CM paper may present a case report on arrhythmia patients. On the other hand, a BM paper may propose an effective signal-processing algorithm to detect an arrhythmia episode from the electrocardiogram (ECG) signal. Thus, a common academic term appears in different contexts in both BM and CM journal papers.

This is problematic for novice BM researchers who must exclusively read BM papers, because the results from a search engine such as Google Scholar may also include many CM papers. Accordingly, novice BM researchers cannot help but spend their time determining whether each collected paper is a BM one, which they need to read. Therefore, it is desirable to automatically determine whether a collected paper is a BM paper or not. Thus, this study applies text-mining techniques to automatically discriminate between BM and CM papers.

Text mining is a general term for quantitatively analyzing documents with a certain purpose. Text-mining techniques are applied for various purposes, such as certifying a hypothesis in literature [1] [2], extracting beneficial information from social network sites [3] [4], or conducting market research in the business science field [5] [6]. A discrimination algorithm is often used to achieve these purposes. Previous studies [7] and [8] also applied discrimination algorithms to automatically retrieve BM papers.

In researches [7] and [8], term-term (T-T) matrices created from document-term (D-T) matrices of BM and CM journal papers were used to distinguish between BM and CM papers. Figs. 1 (a) and (b) show the visualized features of T-T matrices of BM and CM papers. The colors shown in Fig. 1 represent the magnitude of the matrix elements, with red and blue indicating larger and smaller magnitudes, respectively. It is clear from the figure that the blue color stands out more in Fig. 1 (a) than (b). This means that the T-T matrix of the BM paper is sparser than that of the CM one. In short, the vocabulary of the CM paper is larger than that of the BM one.

Actually, the BM and CM papers were accurately discriminated using sparsity features to a certain degree. However, previous studies did not reveal the concrete feature differences between BM and CM papers; e.g., a usage difference in the terms used in the papers.

In this paper, the decision-tree and random-forest algorithms are adopted to reveal a difference in term usage. It is helpful for novice BM researchers to know the concrete differences between BM and CM studies because an automatic discrimination may not always be correct. Although this study focuses on BM studies, approaches to determine which papers to read can generally be beneficial for novice researchers in many interdisciplinary or integrated researches.

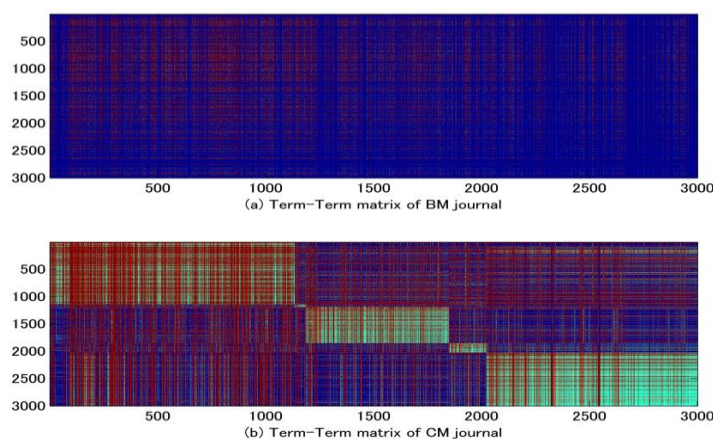


FIGURE 1: VISUALIZED T-T MATRIX OF (A) BM PAPER AND (B) CM PAPER

II. BM AND CM PAPERS USED FOR TEXT MINING

The aim of this research is to identify a concrete term list to discriminate between BM and CM journal papers. In this research, five common topics were selected: obstructive sleep apnea syndrome, T-wave alternans, late potential, epilepsy, and event-related potential. These topics are often found in both BM and CM research. For each of these common topics, 19 BM and 19 CM papers were collected to attempt to discriminate between them. This section explains the backgrounds of BM and CM journal papers for each common topic.

2.1 Obstructive Sleep Apnea Syndrome

Obstructive sleep apnea syndrome (OSAS) is the most general sleep apnea, which is caused by repetitive occlusions of the upper airways. OSAS causes arrested or infrequent respiration, and eventually leads to hypertension, arrhythmia, cardiac arrest, diabetes, or dyslipidemia, which pose a grave life-threatening risk of brain or cardiac infarction. Therefore, many biomedical-engineering and clinical-medicine approaches have been presented for the early detection and treatment of OSAS.

A few of these studies, including [9], [10], and [11], were selected as BM papers, which mainly present signal-processing techniques to automatically detect sleep-apnea episodes from long-duration electrocardiogram (ECG) recordings. On the other hand, papers [12], [13], and [14] were selected as CM papers. These papers present case reports of OSAS patients or the mortality risk due to sleep apnea.

2.2 T-wave alternans

The T-wave is one of the component waves of an ECG, and T-wave alternans (TWA) is defined as a beat-to-beat change in the amplitude of the T-wave. TWA is a promising and important predictor of a sudden cardiac death, but it cannot always be observed with the naked eye, and many BM studies have addressed detecting TWA effectively using signal-processing techniques (e.g., [15], [16]). On the other hand, CM journal papers have reported mortality after myocardial infarction [17] or the importance of a multicenter automatic-defibrillator implantation trial [18].

2.3 Late potential

Like TWA, late potential (LP), a characteristic of ECG waveforms, is also a promising predictor of sudden cardiac death. This is a high-frequency component that occasionally appears in the recorded ECG signal of post-myocardial infarction patients, and has imperceptible amplitude. Thus, LP cannot be detected without signal-processing techniques, and various approaches have been proposed by BM researchers (e.g., [19], [20]). CM papers have presented case reports related to the relationship between LP and ventricular tachycardia, atrial fibrillation, etc. [21] [22].

2.4 Epilepsy

Epilepsy (EPY) is a chronic brain disorder that causes repetitive seizures, and results from various factors. Generally, EPY is diagnosed by the electroencephalographic (EEG) signal because it can be observed as an abnormal electrical brain activity. In BM papers, various methods, such as a wavelet transformation-based EPY detection method [23] and a bivariate analysis-based EPY prediction method using the EEG signal [24], have been proposed. In contrast, CM papers have reported the prognostic effect of EPY surgery [25] [26], etc.

2.5 Event-related potential

Event-related potential (ERP) is an electrical response in the brain that is evoked by an emotional stimulus such as interest, or an external one such as flashing images. BM and CM researches have mined ERPs for entirely different purposes. In BM research, ERPs are often applied at the brain-computer interface as a technique to control electrical devices, such as interacting with a computer without hands or feet, because ERPs reflect human emotion and decision-making. BM researchers have proposed effective methods to detect ERPs in the EEG signal using methods such as wavelet transformation and support vector machines [27] [28]. Meanwhile, medical professionals have reported differences between the ERPs of healthy people and those of patients with certain disorders, such as Alzheimer's disease [29] [30].

III. METHOD

In general, the random-forest algorithm is more effective than the decision tree for classification. However, this research separately uses a simple decision-tree algorithm to make sure of one example of the generated tree. The processing flow to automatically discriminate BM and CM journals is shown in Fig. 2.

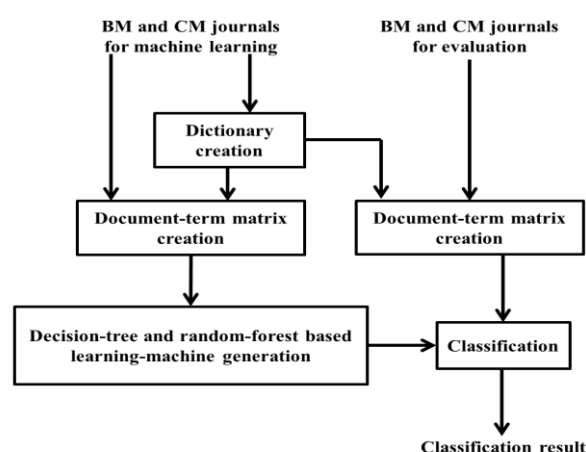


FIGURE 2: PROCESSING FLOW TO DISCRIMINATE BETWEEN BM AND CM PAPERS BASED ON A SIMPLE DECISION-TREE OR RANDOM-FOREST ALGORITHM

3.1 Dictionary creation

As we will discuss later, the dataset of collected BM and CM papers is divided into learning and evaluation data; e.g., some papers related to OSAS are used for evaluation, and others are applied as learning data. First, a dictionary is prepared using terms included only in the learning dataset. In this step, a list of N terms $t = \{t_1, t_2, \dots, t_N\}$ is obtained.

3.2 Document-term matrix creation

A D-T matrix [31] is created with the generated dictionary. Each row of the D-T matrix corresponds to one collected document, and each column corresponds to a term contained in the dictionary. Each matrix-element value indicates the appearance frequency of one term in one document. In general, a D-T matrix is expressed as follows:

$$D = \begin{pmatrix} tf(t_1, d_1) & \cdots & tf(t_N, d_1) \\ \vdots & \ddots & \vdots \\ tf(t_1, d_M) & \cdots & tf(t_N, d_M) \end{pmatrix} \quad (1)$$

where $tf(t_i, d_k)$ is the appearance frequency of term t_i in document d_k . The number of columns N corresponds to the total number of terms contained in a created dictionary. In this study, a $152 \times N$ D-T matrix is created to generate the following learning machine. Additionally, each matrix element is normalized to adjust for the difference in each document's total number of words, using equation (2).

$$tf(t_i, d_k)' = tf(t_i, d_k) / \sum_{i=1}^N tf(t_i, d_k) \quad (2)$$

Likewise, a $38 \times N$ D-T matrix is created from the evaluation dataset.

3.3 Decision-tree and random-forest-based learning-machine generation and classification

The number of terms contained in a generated dictionary is well into the thousands. However, not all terms contained in the dictionary contribute to discriminating between BM and CM papers. Decision tree and random forest are algorithms for selecting a set of terms to classify effectively. In these learning algorithms, the explanatory variables are terms, and the objective variable is the type of paper (BM or CM).

3.3.1 Decision tree:

The decision-tree algorithm uses non-linear discriminant analysis and generates a discriminant tree-structure model by separating the explanatory variables according to certain criteria [32]. This algorithm classifies with a simple IF-THEN rule. In this research, bifurcation is performed with a classification and regression tree (CART) algorithm. The CART algorithm generates an effective decision tree with a Gini coefficient, which is a criterion to measure competitive imbalance. The Gini index ranges from 0 to 1, where 0 represents a perfect equality. Simply stated, the CART algorithm makes a decision tree to attempt to reduce the Gini index to 0.

3.3.2 Random forest:

Random forest is a bootstrap learning algorithm based on multiple decision trees [32]. First, a number of subsets are randomly selected from learning datasets; then, decision trees are generated for the selected sub-data. Finally, the final discrimination is performed by a majority vote of the decision trees' leaves (= selected explanatory variables). The contribution rate of the selected explanatory variables is defined as the Gini decrease rate averaged on all decision trees (mean Gini decrease).

IV. EVALUATIONS AND RESULTS

As described above, one of the five types of papers was selected as evaluation data, and the remaining four types were used to create a learning machine; e.g., evaluation data = {OSAS}, learning data = {TWA, LP, EPL, EDR}. Similarly, five types of journal papers were evaluated by replacing the learning data.

The precisions of the decision-tree and random-forest-based discrimination were examined in terms of accuracy, sensitivity, and specificity. These precisions are defined as follows:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (3)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (4)$$

$$Specificity = \frac{TN}{TN + FP} \quad (5)$$

In Eqs. (3) through (5), TP , TN , FP and FN represent the quantities of true positives, true negatives, false positives, and false negatives, respectively. In this research, TP means the number of correct classifications of BM papers and TN means that of the CM ones.

Tables 1 and 2 show the discrimination results obtained by the simple decision tree and random forest, respectively. Figs. 3 through 7 show decision trees actually created by the CART algorithm. They indicate effective terms to discriminate between BM and CM papers related to OSAS, TWA, LP, EPL, and ERD, respectively.

TABLE 1
EVALUATION RESULTS FOR DECISION-TREE ALGORITHM

Papers	Accuracy [%]	Sensitivity [%]	Specificity [%]
OSAS	86.8	78.9	94.7
TWA	84.2	73.7	94.7
LP	71.1	100	42.1
EPL	81.6	78.9	84.2
ERP	76.3	94.7	57.9
Average	80.0	85.2	74.7

TABLE 2
EVALUATION RESULTS FOR RANDOM-FOREST ALGORITHM

Papers	Accuracy [%]	Sensitivity [%]	Specificity [%]
OSAS	92.1	84.2	100
TWA	89.5	89.5	89.5
LP	94.7	100	89.5
EPL	94.7	94.7	94.7
ERP	86.8	100	73.7
Average	91.6	93.7	89.5

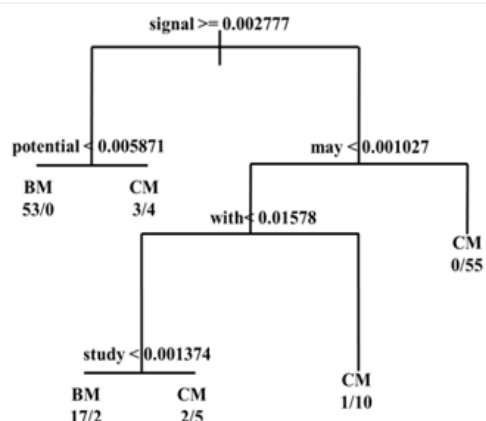


FIGURE 3: DECISION TREE TO DISCRIMINATE BETWEEN BM AND CM PAPERS THAT INCLUDE “OSAS” AS A COMMON TERM

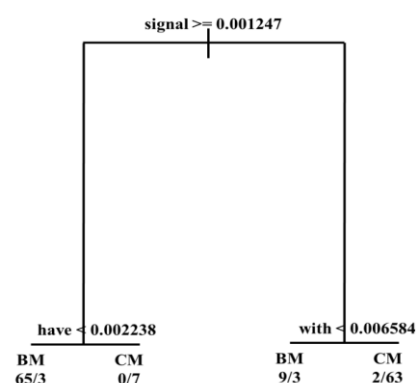


FIGURE 4: DECISION TREE TO DISCRIMINATE BETWEEN BM AND CM PAPERS THAT INCLUDE “TWA” AS A COMMON TERM

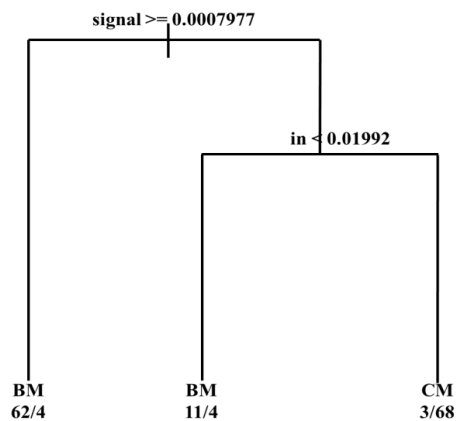


FIGURE 5: DECISION TREE TO DISCRIMINATE BETWEEN BM AND CM PAPERS THAT INCLUDE “LP” AS A COMMON TERM

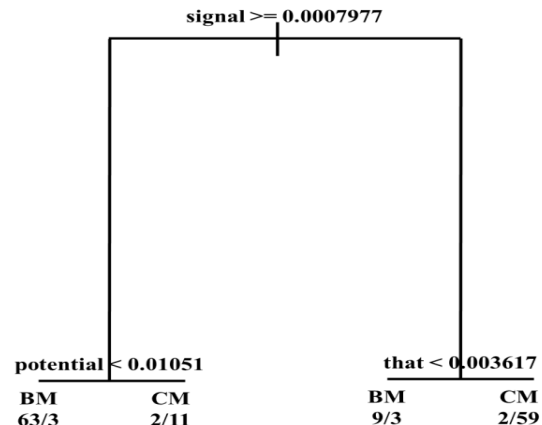


FIGURE 6: DECISION TREE TO DISCRIMINATE BETWEEN BM AND CM PAPERS THAT INCLUDE “EPL” AS A COMMON TERM

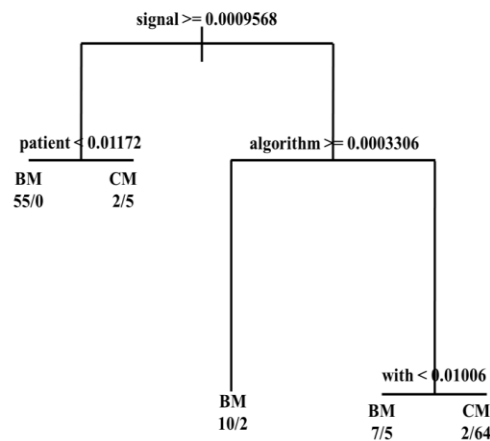


FIGURE 7: DECISION TREE TO DISCRIMINATE BETWEEN BM AND CM PAPERS THAT INCLUDE “ERP” AS A COMMON TERM

Table 3 shows the list of terms that the random-forest algorithm selected to discriminate BM and CM papers. These terms were contributable for all five academic topics: OSAS, TWA, LP, EPL, and ERP. In Table 3, *RAF* (Ratio of Appearance Frequency) means the ratio between the appearance frequency of certain terms in BM papers and that in CM ones. *RAF* is computed as follows:

$$RAF = \frac{\left(\sum_{d_{BM}=1}^{M'} AF_{t,d_{BM}} \right) / M'}{\left(\sum_{d_{CM}=1}^{M'} AF_{t,d_{CM}} \right) / M'} \quad (6)$$

where $AF_{t,d_{BM(CM)}}$ means the appearance frequency of term t in BM (or CM) paper d , and M' means the number of BM and CM papers used as learning data; in this study, $M' = 76$. When $RAF > 0$, it means that one term is used more frequently in BM papers.

V. DISCUSSION

In this research, BM and CM journal papers were discriminated using the decision-tree and random-forest algorithms. The evaluation results showed that the random forest classified more accurately than the decision tree. Tables 1 and 2 show that the precisions obtained by the decision tree reached approximately 80%, and the ones obtained by the random forest were approximately 92%. These results show that BM and CM papers can be discriminated with a high degree of accuracy.

Figs. 3 through 7 show the decision trees generated to discriminate between BM and CM papers. These figures reveal that the term “signal” was the most effective for discriminating. This result indicates that biological information, such as an EEG or ECG, loses its medical feature and can be regarded as a signal defined by mathematical or morphological features in the text of a BM paper. Moreover, two of the five decision trees showed that the term “potential” was more frequently used in CM papers.

Usage examples include “evoked potential”, “potential population”, “potential risk”, and “potential morbidity and mortality”. The usage example “evoked potential” clearly indicates electronic phenomena. However, many of them were used in a context where implicit risks for patients were discussed, based on statistical facts. The fact that “potential” was selected as an effective discrimination term is reasonable, because the statistical perspective is important for investigating every complaint possibility in a medical procedure.

Table 3 shows the effective terms selected by random forest. It also indicates the effectiveness of the term “signal”, but the random-forest algorithm selected “detection” as the most effective term. Specifically, the term “detection” appeared 13 times more frequently in BM papers than in CM ones. This result is also reasonable because BM studies use biological signals to “detect” specific conditions, such as a disease. In short, “detection” is a goal for BM research. (In a sense, a CM study begins after “detection” occurs.) In addition, Table 3 shows that the terms “method”, “base”, and “use” appeared more frequently in BM papers than in CM ones. In BM papers, these terms were used to explain techniques for the detection of vital phenomena, such as TWA; e.g., “This method is based on XXXX”, “using the XXXX method”, or “based on features extracted from”. This result is also rational considering that BM research must solve problems on the medical front based on engineering methods.

The remaining terms in Table 3 were used more frequently in CM papers than in BM ones ($RAF < 1$). Especially, the terms “age”, “who”, “patient”, and “with” relate to one another. The core word in this term group is “patient”. The others are often used to append the patient’s condition or add information to their complaint; e.g., “patient with epilepsy”, “patient who undergoes revascularization procedures”, or “patients aged XXXX or older”. In addition, the term “with” is often used to present an instrument or avenue; e.g., “the electrode was cleaned with alcohol”, “a catheter with a XXXX-mm tip”, or “with lamotrigine”. The term “increase” is often used to compare a patient’s group with a control group; e.g., “The mortality rate of patients with XXXX disease increased over that of the control group.” Comparing with a control group is a standard tactic in medical research, so this result also makes sense.

The usage of the other terms shown in Figs. 3 through 7, e.g., “algorithm”, “have”, “in”, “may”, “study”, and “that”, were difficult to generalize because they were selected by only one decision tree. However, their usage may become clearer if more papers are analyzed.

As stated in Sections 1 and 2, a BM researcher focuses on engineering methods to detect biological phenomena, while a CM researcher reports clinical cases about patients. These differences in contexts between BM and CM papers could be confirmed in this research, and it was revealed that their differences were shown in their use of standard terms, such as “base”, “use”, “with”, and “in”. Simultaneously considering precisions, it can be concluded that the automatic discrimination between BM and CM papers is feasible, and the basis for discrimination is clear enough to understand intuitively.

TABLE 3
CONTRIBUTIVE TERMS TO DISCRIMINATE BETWEEN BM AND CM PAPERS

Term	RAF
detection	13.177
signal	10.276
method	6.631
base	4.386
use	1.086
increase	0.605
age	0.183
who	0.169
patient	0.149
with	0.055

VI. CONCLUSION

The goal of this study was to discriminate between BM and CM journal papers. The decision-tree and random-forest algorithms were adopted for this purpose. The results showed that BM and CM papers were discriminated with approximately 80% through 90% precision. In addition, the proposed methods revealed effective terms to discriminate between BM and CM papers, e.g., “signal”, “detection”, “patient”, “method”, etc.

REFERENCES

- [1] P. Sallis and S. Shanmuganathan, “A Blended Text Mining Method for Authorship Authentication Analysis,” In *Proceedings of Second Asia International Conference on Modelling & Simulation 2008*, pp. 451–456, 2008.
- [2] H. Hosoi, T. Yamagata, Y. Ikarashi and N. Fujisawa, “Visualization of Special Features in ‘The Tale of Genji’ by Text Mining and Correspondence Analysis with Clustering,” *Journal of Flow Control, Measurement & Visualization*, vol. 2, no. 2, pp. 1–6, 2014.
- [3] W. Black, R. Procter, S. Gray and S. Ananiadou, “A data and analysis resource for an experiment in text mining a collection of micro-blogs on a political topic,” In *Proceedings of the International Conference on Language Resources and Evaluation*, pp. 2083–2088, 2012.
- [4] S. Stieglitz and L. Dang-Xuan, “Social media and political communication: a social media analytics framework,” *Social Network Analysis and Mining*, vol. 3, no. 4, pp. 1277–1291, 2013.
- [5] P. Gupta and B. Narang, “Role of Text Mining in Business Intelligence,” *Gyan Jyoti E-Journal*, vol. 1, no. 2, 2012.
- [6] O. Netzer, R. Feldman, J. Goldenberg and M. Fresko, “Mine Your Own Business: Market-Structure Surveillance Through Text Mining,” *Marketing Science*, vol. 31, no. 3, pp. 521–543, 2012.
- [7] M. Sakai, “Text Mining Technique to Distinguish between Clinical Medicine and Biomedical Engineering Research Articles,” *International Journal of Engineering Science and Innovative Technology*, vol. 4, issue 2, pp. 290–296, 2015.
- [8] M. Sakai, “Differences between the Quantitative Features of Articles from Biomedical Engineering and the Medical Sciences Aimed at Effective Retrieval of Literature,” *International Journal of Electronics Communication and Computer Engineering*, vol. 6, issue 6, pp. 664–670, November 2015.
- [9] J. Corthout, S. Van Huffel, M.O. Mendez, A.M. Bianchi, T. Penzel and S. Cerutti, “Automatic Screening of Obstructive Sleep Apnea from the ECG based on Empirical Mode Decomposition and Wavelet Analysis,” *Conf Proc IEEE Eng Med Biol Soc 2008*, pp. 3608–3611, 2008.
- [10] T-P. Kao, J-S. Wang, C-W. Lin, Y-T. Yang and F-C. Juang, “Using Bootstrap AdaBoost with KNN for ECG-based automated obstructive sleep apnea detection,” *Neural Networks (IJCNN), 2012 International Joint Conference on*, pp. 1–5, 2012.
- [11] M. Sakai, X. Zhu, Y. Yoshida and D. Wei, “Development of Lead System for ECG-Derived Respiration Aimed at Detection of Obstructive Sleep Apnea Syndrome,” In *Proceedings of the 9th International Conference on Signal Image Technology & Internet Based Systems*, pp. 971–975, 2013.
- [12] J. He, M.H. Kryger, F.J. Zorick, W. Conway and T. Roth, “Mortality and apnea index in obstructive sleep apnea. Experience in 385 male patients,” *Chest*, vol. 94, no. 1, pp. 9–14, 1988.
- [13] H.K. Yaggi, J. Concato, W.N. Kernan, J.H. Lichtman, L.M. Brass and V. Mohsenin, “Obstructive Sleep Apnea as a Risk Factor for Stroke and Death,” *The New England Journal of Medicine*, vol. 353, pp. 2034–2041, 2005.
- [14] G. Jean-Louis, F. Zizi, L.T. Clark, C.D. Brown and S.I. McFarlane, “Obstructive Sleep Apnea and Cardiovascular Disease: Role of the Metabolic Syndrome and Its Components,” *Journal of Clinical Sleep Medicine*, vol. 4, no. 3, pp. 261–272, 2008.
- [15] A.D. Bakhshi, A. Ahmed, S. Gulfam, A. Khaqan, A. Yasin, R. Riaz, K. Alimgeer, S. Malik, S. Khan and A.H. Dar, “Detection of ECG T-wave Alternans Using Maxima of Continuous-Time Wavelet Transform Ridges,” *Przegląd Elektrotechniczny*, vol. 88, no. 12b, pp. 35–38, 2012.
- [16] V. Monasterio, P. Laguna, I. Cygankiewicz, R. Vázquez, A. Bayés-Genís, A.B. de Luna and J.P. Martínez, “Average T-wave alternans activity in ambulatory ECG records predicts sudden cardiac death in patients with chronic heart failure,” *Computing in Cardiology*, vol. 38, pp. 621–624, 2012.
- [17] L. Burattini, S. Man, G. Ottaviano, S. Fioretti, F.D. Nardo and C.A. Swenne, “T-wave alternans rate of change with exercise for cardiac risk assessment,” *Computing in Cardiology*, vol. 41, pp. 177–180, 2014.
- [18] V. Monasterio, J.P. Martínez, P. Laguna, S. McNitt, S. Polonsky, A.J. Moss, M. Haigney, W. Zareba and J.P. Couderc, “Prognostic value of average T-wave alternans and QT variability for cardiac events in MADIT-II patients,” *Journal of Electrocardiology*, vol. 46, no. 6, pp. 480–486, 2013.
- [19] M. Kalra, M. Kumar and R. Mittal, “Continuous Wavelet Transform for the Analysis of Arrhythmia,” *International Journal of Latest Research in Science and Technology*, vol. 4, issue 3, pp. 80–84, 2015.
- [20] S. Subramanian, G. Gurusamy and G. Selvakumar, “A New Ventricular Late Potential Classification System Using Ant Colony Optimization,” *Journal of Computer Science*, vol. 8, no. 2, pp. 259–264, 2012.

- [21] M. Masuda, K. Inoue, K. Iwakura, A. Okamura, Y. Toyoshima, A. Doi, Y. Sotomi, I. Komuro and K. Fujii, "Impact of pulmonary vein isolation on atrial late potentials: Association with the recurrence of atrial fibrillation," *Europace*, vol. 15, pp. 501–507, 2013.
- [22] M. Efremidis, E. Prappa, C. Charalampous, N. Protonotarios, K. P. Letsas and A. Sideris, "Endocardial Late Potentials During Sinus Rhythm Define the Re-Entry Circuit of Ventricular Tachycardia in Arrhythmogenic Right Ventricular Cardiomyopathy," *Hellenic Journal of Cardiology*, vol. 54, pp. 394–396, 2013.
- [23] D. Wang, D. Miao and C. Xie, "Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection," *Expert Systems with Applications*, vol. 38, pp. 14314–14320, 2011.
- [24] P. Mirowski, D. Madhavan, Y. LeCun and R. Kuzniecky, "Classification of patterns of EEG synchronization for seizure prediction," *Clinical Neurophysiology*, vol. 120, pp. 1927–1940, 2009.
- [25] D.J. Costello, D.C. Shields, S.S. Casha, E.N. Eskandar, G.R. Cosgrove and A.J. Colea, "Consideration of epilepsy surgery in adults should be independent of age," *Clinical Neurology and Neurosurgery*, vol. 111, pp. 240–245, 2009.
- [26] I. Najm, L. Jehi, A. Palmini, J. Gonzalez-Martinez, E. Paglioli and W. Bingaman, "Temporal patterns and mechanisms of epilepsy surgery failure," *Epilepsia*, vol. 54, no. 5, pp. 772–782, 2013.
- [27] V.N. Raju, I. Ra and R. Sankar, "A P300-Based BCI Classification Algorithm Using Least Square Support Vector Machine," *International Journal of Software Engineering and Its Applications*, vol. 9, no. 5, pp. 247–254, 2015.
- [28] V. Bostanov, "BCI Competition 2003—Data Sets Ib and IIb: Feature Extraction From Event-Related Brain Potentials With the Continuous Wavelet Transform and the t-Value Scalogram," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1057–1061, 2004.
- [29] V. Taler and N.A. Phillips, "Event-related potential measures of lexical activation in Alzheimer's disease," *Brain and Language*, vol. 103, pp. 177–178, 2007.
- [30] G.D. Tillman, C.S. Calley, T.A. Green, V.I. Buhl, M.M. Biggs, J.S. Spence, R.W. Briggs, R.W. Haley, J. Hart Jr. and M.A. Kraut, "Event-related potential patterns associated with hyperarousal in Gulf War illness syndrome groups," *NeuroToxicology*, vol. 33, pp. 1096–1105, 2012.
- [31] J. Shawe-Taylor and N. Cristianini, "Kernel Methods for Pattern Analysis," *Cambridge University Press*, 2004.
- [32] J. Ali, R. Khan, N. Ahmad and I. Maqsood, "Random Forests and Decision Trees," *IJCSI International Journal of Computer Science Issues*, vol. 9, issue 5, no. 3, pp. 272–278, 2012.