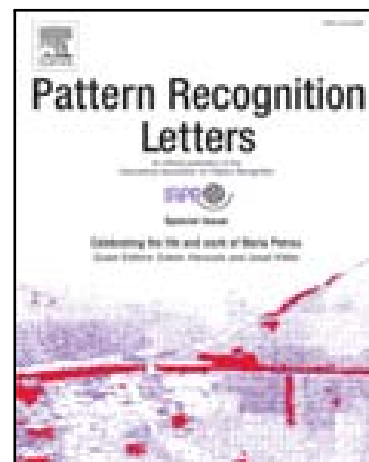


## Accepted Manuscript

A General Extensible Learning Approach for Multi-Disease  
Recommendations in a Telehealth Environment

Raid Lafta, Ji Zhang, Xiaohui Tao, Xiaodong Zhu, Hongzhou Li,  
Liang Chang, Ravinesh Deo

PII: S0167-8655(18)30881-X  
DOI: <https://doi.org/10.1016/j.patrec.2018.11.006>  
Reference: PATREC 7369



To appear in: *Pattern Recognition Letters*

Received date: 19 April 2018  
Revised date: 25 October 2018  
Accepted date: 7 November 2018

Please cite this article as: Raid Lafta, Ji Zhang, Xiaohui Tao, Xiaodong Zhu, Hongzhou Li, Liang Chang, Ravinesh Deo, A General Extensible Learning Approach for Multi-Disease Recommendations in a Telehealth Environment, *Pattern Recognition Letters* (2018), doi: <https://doi.org/10.1016/j.patrec.2018.11.006>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Pattern Recognition Letters  
journal homepage: [www.elsevier.com](http://www.elsevier.com)

## A General Extensible Learning Approach for Multi-Disease Recommendations in a Telehealth Environment

Raid Lafta<sup>a,b</sup>, Ji Zhang<sup>c,a,\*\*</sup>, Xiaohui Tao<sup>a</sup>, Xiaodong Zhu<sup>d</sup>, Hongzhou Li<sup>e</sup>, Liang Chang<sup>e</sup>, Ravinesh Deo<sup>a</sup>

<sup>a</sup>Faculty of Health, Engineering and Sciences, University of Southern Queensland, Australia

<sup>b</sup>Computer Center, University of Thi-Qar, Thi-Qar, Iraq

<sup>c</sup>Nanjing University of Aeronautics and Astronautics, China

<sup>d</sup>Nanjing University of Information Science and Technology, China

<sup>e</sup>Guilin University of Electronic Technology, China

### ABSTRACT

In a telehealth environment, intelligent technologies are rapidly evolving toward improving the quality of patients' lives and providing better clinical decision-making especially those who suffer from chronic diseases and require continuous monitoring and chronic-related medical measurements. A short-term disease risk prediction is a challenging task but is a great importance for telehealth care systems to provide accurate and reliable recommendations to patients. In this work, a general extensible learning approach for multi-disease recommendations is proposed to provide accurate recommendations for patients with chronic diseases in a telehealth environment. This approach generates appropriate recommendations for patients suffering from chronic diseases such as heart failure and diabetes about the need to take a medical test or not on the coming day based on the analysis of their medical data. The statistical features extracted from the sub-bands obtained after a four-level decomposition of the patient's time series data are classified using a machine learning ensemble model. A combination of three classifiers – Least Squares-Support Vector Machine, Artificial Neural Network, and Naive Bayes – are utilized to construct the bagging-based ensemble model used to produce the final recommendations for patients. Two real-life datasets collected from chronic heart and diabetes disease patients are used for experimentations and evaluation. The experimental results show that the proposed approach yields a very good recommendation accuracy and offers an effective way to reduce the risk of incorrect recommendations as well as reduces the workload for chronic diseases patients who undergo body tests most days. Thus, the proposed approach is considered one of a promising tool for analyzing time series medical data of multi diseases and providing accurate and reliable recommendations to patients suffering from different types of chronic diseases.

© 2018 Elsevier Ltd. All rights reserved.

### 1. Introduction

According to the world health organization, the chronic diseases were caused the death for 50% of people worldwide (Kuh and Shlomo, 2004). As patients with chronic diseases require frequent medical attention, the lack of automatically generated medical recommendation systems are currently limiting options for improvements in treatment and care, thus affecting their quality of life. Therefore, an extensive number of attempts have

been made in an attempt to find effective approaches to tackle this issue.

The clinical studies showed that heart disease is one of the major diseases that is considered a life-threatening. It has been ranked as the highest deadly disease among non-infectious diseases which requires a high cost and effort in terms of prevention and treatment (Lavoie et al., 2016). Because of an absence of accurate medical recommendations that can be consequently created for a superior treatment and care, the quality of chronic disease patients' lives has been essentially influenced.

Diabetes is also main public health issue worldwide and cause of mortality and morbidity (Wang et al., 2017b). Based

<sup>\*\*</sup>Corresponding author: Tel.: +61 7 4631 5539  
e-mail: [ji.zhang@usq.edu.au](mailto:ji.zhang@usq.edu.au) (Ji Zhang)

on the obtained data from 2008, the percentage of people with diabetes in China and India is accounted 40% of global population (around 138 million individuals) and accounted 10% in the USA and Russia (approximately 36 million individuals) (Danaei et al., 2011).

In the era of telehealth, many applications have been designed to help patients and medical practitioners deliver their medical information easily (Dewar et al., 2017; Wang et al., 2017a). These digital tools allow patients and practitioners to exchange personalized reminders and medical recommendations, upload information from devices such as a blood pressure cuff or blood glucose monitors, store health records, and share information with health care providers (Van Ma et al., 2016).

Health care providers can interact face-to-face and real-time with their patients using the telehealth services (such as the Internet or video equipment) to provide appropriate recommendations for overall health conditions. Devices such as heart rate monitors and blood pressure monitors can be connected to web-based applications in a telehealth environment to exchange medical information between patients and health care providers (Lafta et al., 2018). Telehealth systems can be particularly useful for people with chronic diseases, as well as those who live in rural or remote areas where fewer visits to health-care organizations such as hospitals and medical centers and are possible access to medical care and medical recommendations can be limited.

Given the importance of disease risk prediction in medical field (Zhang et al., 2017) as well as the urgency of obtaining more effective analytic techniques for disease risk prediction, great efforts are needed to enhance the quality of evidence-based decisions and recommendations in the telehealth environment. Patients with chronic disease need to undertake various daily medical tests to monitor their overall chronic health conditions through the telehealth system. Carrying out various medical tests every day brings lots of inconvenience and even burden to patients and adversely affects their life quality. Producing accurate intelligent recommendations to guide their daily medical tests can effectively reduce the workload associated with taking those tests while keeping the associated health risk in an acceptably low level.

In many cases, producing accurate recommendations is an essential function in telehealth systems, which is often based on the prediction of patients' disease risk for short-term. Various approaches have been developed as disease risk prediction models in which data mining and statistical analysis tools were used for various healthcare and medical issues (Polat and Güneş, 2007; Yang et al., 2014; Sánchez et al., 2016; Mohktar et al., 2015; Chang et al., 2011; Huang et al., 2012; Lu et al., 2014; Krishnaiah et al., 2013; Yeh et al., 2011; Zhu et al., 2018; Zheng et al., 2017; Zhu et al., 2017b; Zhang et al., 2018; Hu et al., 2017; Zhu et al., 2017a). However, it was found that most of these studies have not particularly dealt with chronic disease issues. In addition, few of the existing studies focus on recommendations based on short-term medical prediction. Short-term prediction, which is studied in our research, is more challenging than long-term prediction as patients' conditions may experience more dramatic and abrupt changes during a short time-

frame. Furthermore, short-term recommendations are equally useful for patients as they provide guidance as to what the patients need to do for the next few days.

From the technological prospective, there has been a growing interest in using transformation techniques such as a Fourier, wavelet and hybrid transformation to analyse time series data such as these generated in the biomedical domain, to generate a good set of features for the subsequent classification/prediction task. One of the promising transformation techniques is called Dual Tree Complex Wavelet Transformation (DTCWT). It is an improved version of wavelet transformation that is designed to tackle some limitations in the wavelet transform. Due to its promising predictive effectiveness, the dual-tree complex wavelet has been successfully utilized in many different research areas such as processing of images and de-noising (Yu et al., 2016; Jung et al., 2017), and detecting epileptic seizures in electroencephalography (EEG) (Swami et al., 2017; Li et al., 2017).

Classification approaches that generate intelligent and accurate medical recommendations are used in our work to produce reliable assessments of the short-term medical risks of patients with chronic diseases. In the literature, the classification problem based on the classification methods in its processes is naturally utilized to predict the actually needing of taking a body test for a given measurement such as heart rate and diabetes. For predication phase, three machine learning techniques – Least Squares-Support Vector Machine, Artificial Neural Network, and Naive Bayes – are used and combined together to design an ensemble classifier in this work. Several important reasons pushed the authors to build the ensemble classifier model. First, the idea of building a single model is an efficient solution for some applications in which the amount of data is huge (Belayneh et al., 2016). Second, due to its ability to improve the overall performance of prediction models, it is an effective tool for many models. Machine learning ensemble approaches have been utilized in several studies in the medical domain with some success (Das et al., 2009; Helmy et al., 2013; Bashir et al., 2015; Verma et al., 2016). However, none of them tackled or solved the problem dealt with in this study to provide accurate recommendations to advise patients with chronic disease whether they should to take or skip the medical tests on the coming day.

We in this paper propose a novel short-term recommendation approach for multi diseases prediction to provide chronic diseases patients in a telehealth environment with accurate recommendations. This method is developed using the dual-tree complex wavelet transformation and a machine learning ensemble model for whether or not a particular patient needs to take the physical body test on the current date using the telehealth facility. Such recommendations are established based on analysing the time series data of chronic disease patients from the past few days. The statistical features of each sliding window of the time series medical data are first extracted using DTCWT which are then input into a bagging-based ensemble model to predict the necessity of taking body test. The bagging-based ensemble model combines the results of three classifiers – Least Squares-Support Vector Machine, Artificial Neural Network and Naive Bayes to produce accurate prediction and rec-

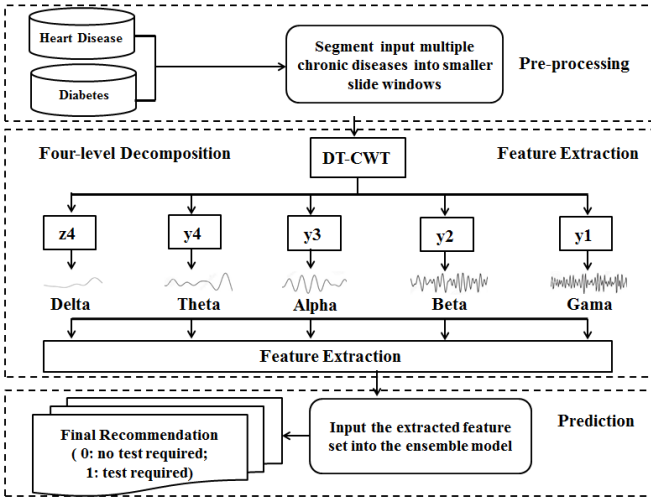


Fig. 1: The architecture of the proposed method

ommendations. We compare our approach with our previous work used to tackle the exactly same problem to prove the superiority of our proposed technique. To this end, we adopted three performance metrics: i.e., *accuracy*, *workload saving* and *risk*, to evaluate the performance of our proposed recommendation method. The experimental evaluations are conducted using two real-life time series datasets collected from a set of heart and diabetes disease patients. The obtained results showed that the proposed approach was effective in providing the accurate recommendations to heart and diabetes disease patients and reducing the workload needed in medical tests for these patients. It also significantly reduced the incorrect recommendations rates for these patients.

The rest of this paper is organized as follows. Section 2 discusses in detail the proposed recommendation approach for chronic disease patients, with a focus on the dual-tree complex wavelet transformation and the ensemble model. Section 3 elaborates on the results of the experiments conducted to evaluate the performance of the proposed method. Finally, Section 5 concludes the paper and highlights our future research directions.

## 2. The Proposed Recommendation Method

This study aims to investigate the effectiveness of proposed method using the dual tree complex wavelet transformation coupled a machine learning-based ensemble model in providing appropriate medical recommendations to patients with chronic diseases as to the necessity of taking a medical test on the coming day. We first present an overview on the architecture of the proposed method. Then the detailed discussion on the dual tree complex wavelet transformation and the machine learning ensemble model, the two major technical components of the proposed method, is presented in this section.

### 2.1. An Overview of Our Methodology

Figure 1 shows the overall architecture of the proposed recommendation method used to provide patients with appropriate

medical recommendations in a telehealth environment. Firstly, the time series data of a given patient is divided into smaller overlapped sliding windows where the size of the slide window is empirically determined. Secondly, the Dual-Tree Complex Wavelet Transformation (DTCWT) is adopted to decompose the segmented time series of two datasets into sub-bands of *delta*, *theta*, *alpha*, *beta* and *gamma*. Employing the DTCWT in our work, which solved the problem of the shift invariance property that found in ordinary Discrete Wavelet Transform (DWT), increases the robustness of the coefficients extracted from the time series data of patients. Finally, the statistical features are extracted from different constituent sub-bands and fed to the ensemble classifier to generate a binary recommendation concerning the patients' condition. The produced medical recommendations are decided whether a given patient needs to take a medical test on the coming day or not. This architecture is general and extensible to deal with many different type of chronic diseases in a multi-task manner as long as compatible time series medical data are generated and fed into the recommendation system.

### 2.2. Dual tree complex wavelet transformation

The Discrete Wavelet Transformation (DWT) is a powerful tool for analyzing signals in time-frequency domain. Despite its ability to provide an efficient time-frequency analysis of non-stationary signals, the DWT suffers from several problems such as aliasing, shift variance and lack of directionality (Kingsbury, 1998). The drawbacks of DWT can be mitigated by using the Dual Tree Complex Wavelet Transformation (DTCWT) which offers a better time-frequency representation of signals (Das et al., 2016). Basically, DTCWT employs two real DWT trees (Yu et al., 2016), where the top tree represents the real part of the complex wavelet coefficient whilst the tree on the bottom represents the imaginary part as illustrated in Figure 2.

The DTCWT is adopted to decompose the input time series data of two datasets into sub-bands of *delta*, *theta*, *alpha*, *beta* and *gamma* after four levels of decomposition. These five sub-bands are represented by  $y_1$ ,  $y_2$ ,  $y_3$ ,  $y_4$  and  $z_4$  for each part (tree). Each DTCWT coefficient has two parts: real and imaginary. As a result, ten sub-bands are in total obtained after the four-level decomposition (five sub-bands for each part). In this paper, the DTCWT sub-bands for both real and imaginary parts are presented by  $(y_1,1)$ ,  $(y_1,2)$ ,  $(y_2,1)$ ,  $(y_2,2)$ ,  $(y_3,1)$ ,  $(y_3,2)$ ,  $(y_4,1)$ ,  $(y_4,2)$ ,  $(z_4,1)$  and  $(z_4,2)$ . For example,  $(y_1,1)$ ,  $(y_1,2)$  represent the  $y_1$  sub-band for the real and imaginary parts, respectively. DTCWT decomposes the input data sequence (i.e., the data in the sliding windows of the given time series data of the patients) in order to extract frequency information for the purpose of predicting the patient's condition one day in advance.

In this study, the time series data of patients are represented as  $X = \{y_1, y_2, y_3, \dots, y_n\}$  of  $n$  data. It is divided into a set of overlapped sub-segments where the size of the sliding window is empirically determined based on a predefined value of parameter  $k$  corresponding to each sub-segment.

Then, each sub-segment is subject to a four-level DTCWT decomposition, producing 10 wavelet sub-bands, which repre-

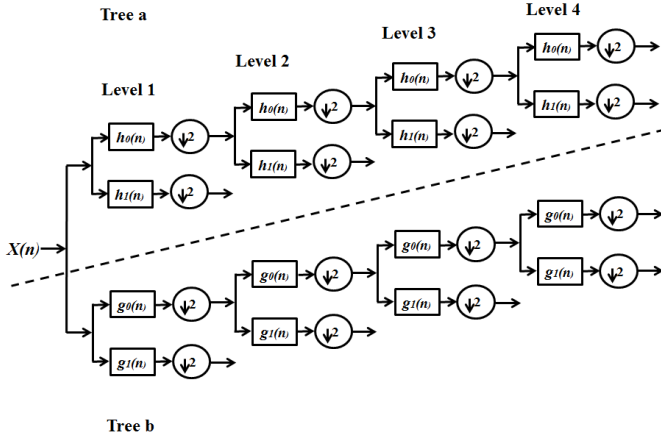


Fig. 2: 1D dual-tree complex wavelet transformation

sent the frequency information to be used for predicting the patient's condition and generating the recommendations.

The conversion of the obtained wavelet coefficients into a reduced set of features is a very important stage in the proposed method. Since these features characterize the behavior of time series data of patients, the process of features selection is crucially important. The following six statistical features are generated and used to predict the patient's condition:

1. Mean of coefficients absolute values in every sub-band,

$$\eta = \frac{1}{N} \sum_{j=1}^N |y_j|$$

2. Average power of the coefficients in every sub-band,  $\mu =$

$$\sqrt{\frac{1}{N} \sum_{j=1}^N y_j^2}$$

3. Standard deviation of the coefficients in every sub-band,

$$\nu = \sqrt{\frac{1}{N} \sum_{j=1}^N (y_j - \eta)^2}$$

4. Ratio of the absolute mean values of coefficients of adjacent sub-bands,  $\xi =$

$$\frac{\sum_{j=1}^N |y_j|}{\sum_{j=1}^N |z_j|}$$

5. Kurtosis of the coefficients in every sub-band,  $\chi =$

$$\sqrt{\frac{1}{N} \sum_{j=1}^N \frac{(y_j - \eta)^4}{\nu^4}}$$

6. Skewness of the coefficients in every sub-band,  $\phi =$

$$\sqrt{\frac{1}{N} \sum_{j=1}^N \frac{(y_j - \eta)^3}{\nu^3}}$$

where  $y$  and  $z$  are two adjacent sub-bands and  $N$  is the length of a sub-band. In literature, the Feature 5 and 6 are considered as a higher order than other statistics features. In this work, three selected sets of features were manually explored as follows:

1. Two-feature set: Features 5-6 (experiment 1)

2. Four-feature set: Features 3-6 (experiment 2)
3. Six-feature set: All features 1-6 (experiment 3)

In the experimental evaluations section, we discuss in more details the impact of different combinations of the extracted features on the final recommendation results generated.

### 2.3. Bootstrap Aggregation (Bagging)

An ensemble approach is an effective method to combine the results of individual base classifiers in order to generate more accurate predictions than individual base classifiers that make them up. One of the popular machine learning ensemble algorithms used to enhance accuracy and stability in the ensemble models is bootstrap aggregation, a.k.a bagging (Gao et al., 2016; Li et al., 2016).

In the last decades, empirical results in the literature showed that machine learning ensembles are often more accurate than single classifiers that make them up (Valentini and Masulli, 2002; Yang et al., 2018). Bagging aggregation, which was proposed by (Breiman, 1996) is a one of machine learning ensemble algorithms. It is designed to improve the performance and stability of machine learning algorithms (Voyant et al., 2017). The ensemble model has been proven its effective and efficient in processing the classification and prediction problems.

In bagging algorithm, the base classifiers are trained independently using a bootstrap method and then the results of the those classifiers are aggregated using an appropriate combination technique.

The proposed ensemble model can be divided into two phases: In the first phase, the model is generated a number of training sets using bootstrap sampling. In the second phase, the bootstrap training sets generated through the first phase are used in training the three individual classifiers, i.e., Least Square-Support Vector Machine, Neural Network and Naive Bayes.

Figure 3 illustrates an example of the bagging algorithm for building ensemble model from the three base classifiers.

The training set is divided the training set into multi sub-datasets, and then the three base classifiers are individually applied to these sub-datasets to generate the final prediction. It is noted that different individual classifier in the bagging approach may perform differently. Hence, a specific weight is assigned to each classifier's vote based the performance of the classifier. In this study, we are calculated the classifier's weight based on its error rate. The classifier that has a lower error rate is considered more accurate classifier and is therefore assigned a higher weight.

The weight of classifier  $CS_i$ 's vote is calculated as follows:

$$w(CS_i) = \log \frac{1 - \text{error}(CS_i)}{\text{error}(CS_i)}, 1 \leq CS_i \leq 3 \quad (1)$$

The weighted-bagging ensemble can be easily understood by the following example:

1. Least Square-Support Vector Machine, Neural Network, and Naive Bayes are used as individual base classifier. Suppose that the classifier training is performed on the training data and the error rate is calculated for each base classifier as 0.14 for LS-SVM, 0.25 for NN, and 0.30 for NB;



Table 1: Meta-data attributes of the Dataset

Attribute name	Attribute type
id	Numeric
patient-id	Numeric
hcn	Numeric
visit-id	Numeric
measurement type	Nominal
measurement unit	Nominal
measurement value	Numeric
measurement question	Nominal
date	Numeric
date-received	Numeric

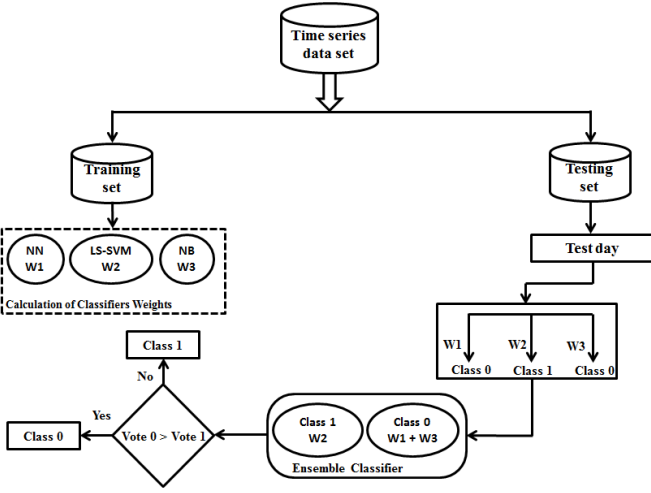


Fig. 3: An example of a bagging algorithm

- Now, according to the formula given in Equation (1), the weight 0.78 is assigned to LS-SVM, 0.47 to NN, and 0.36 to NB;
- Suppose that the three base classifiers generate the following predictions for a coming testing day: LS-SVM predicts 0, NN predicts 1, and NB predicts 1 (0 means no test is required on the testing day for a medical measurement; 1 means a test is required otherwise);
- The ensemble classifier will use the weighted vote to generate the following prediction results:  
Class 1:  $NN + NB \rightarrow 0.47 + 0.36 \rightarrow 0.83$ ,  
Class 0:  $LS-SVM \rightarrow 0.78$ .
- Finally, a Class 1 has a higher value than a Class 0 according to the weighted vote. Thus, the ensemble classifier will classify this testing day as being in Class 1, suggesting that the patient in question needs to take the test for a medical measurement on that day.

### 3. Experimental Results

#### 3.1. The Experimental Setup

To test the practical applicability of the proposed recommendation method for dealing with multiple different types chronic diseases, two real-life time series telehealth datasets, corresponding to two different types of chronic diseases, are used for experimentations and evaluation. The first dataset, called Tunstall dataset, is obtained from a group of chronic heart disease patients while the second dataset is obtained from a set of chronic diabetes disease patients. A real-life dataset obtained from our industry collaborator, Tunstall Healthcare, is used in this work. Ethics clearance was obtained from the Human Research Ethics Committee of the University of Southern Queensland (No: H15REA167) prior to the use of Tunstall dataset for our study.

The Tunstall dataset was collected from a study conducted a set of chronic heart disease patients and the collected data contain the patients' day-to-day medical readings of various medical measurements in a telehealth care environment. The

obtained dataset is a time series and comprises data from six chronic heart disease patients with a total of 7,147 different time series records acquired between May and October 2012. Each record in this data consists of a set different meta-data attributes about the patients, such as patient-id, visit-id, measurement type, measurement unit, measurement question, measurement value, date-received and date. Table 1 shows the characteristics of the meta-data attributes of the dataset. In addition, the dataset contains the numerical readings of several critical medical measurements for each patient on every day during the time period of the study, including Heart Rate, Mean Arterial Pressure (MAP), Diastolic Blood Pressure (DBP), Oxygen Saturation (SO<sub>2</sub>), Blood Glucose and Weight, among which the data of Heart Rate, MAP, DBP and SO<sub>2</sub> are used in the evaluation.

The diabetes dataset used in this study obtained from the Repository of Machine Learning Databases which is available online from the URL link of <https://archive.ics.uci.edu/ml/datasets/diabetes> (Kahn, 1994). The collected data contain measurements taken multiple times per day from 70 patients. Blood glucose measurements, symptoms and insulin treatments were recorded with timestamps for each patient, over the course of several weeks to months. Two different sources: an automatic electronic recording device and paper records have been used in acquiring the records of data. The automatic device had an internal clock to timestamp events, whereas the paper records only provided "logical time" slots (breakfast, lunch, dinner, bedtime) (Moghaddass et al., 2016). The dataset is a time series and comprises data from 70 diabetes patient. Each record in the diabetes dataset consists of four fields about the date of measurement, time of measurement, the code of measurement and the value of measurement. The characteristics of the meta-data attributes of the dataset are shown in Table 2. Each code represents one of a twenty numerical readings about the patients, such as regular insulin dose, NPH insulin dose, ultralente insulin dose, unspecified blood glucose measurement, pre-breakfast blood glucose measurement, post-breakfast blood glucose measurement, pre-lunch blood glucose measurement, post-lunch blood glucose measurement, pre-supper blood glucose measurement, post-supper blood glucose measurement, pre-snack blood glucose measurement, hypoglycemic symptoms, typical meal ingestion, more-than-usual meal ingestion, less-than-usual meal ingestion, typical exercise activity, more-than-usual exercise activity, less-than-usual exer-

Table 2: Meta-data attributes of diabetes dataset

Attribute name	Attribute type
Date	Numeric
Time	Numeric
Code	Numeric
Value	Numeric

cise activity and unspecified special event.

In this work, both the Tunstall and Diabetes dataset are divided into two parts: the training set and the testing set. All models including the base individual classifiers and the ensemble model are trained using the training set and then validated using the testing set as the ground truth result. 75% of the dataset was partitioned as the training data while the remaining 25% was used as testing data in this study.

To evaluate the ability of the proposed method on producing high quality recommendations, the produced recommendations were compared with the obtained actual readings in the testing test. Due to the class-imbalance problem in a patient's historical medical data (i.e., the number of normal data is much larger than that of the abnormal data), we carefully dealt with the class-imbalance problem when training the classifiers. The over-sampling and under-sampling methods have been used as good means to address this problem (Wang et al., 2012).

### 3.2. Performance Metrics

In order to evaluate the performance of the recommendation approach, the following three performance metrics are proposed for this work.

1. Accuracy refers to the percentage of correctly recommended days against the total number of days for which recommendations are provided. It is defined as

$$Accuracy = \frac{N_C}{|D|} \times 100\% \quad (2)$$

where  $N_C$  denotes the number of days with correct recommendations, and  $|D|$  refers to the total number of days in the dataset.

2. Workload saving refers to the percentage of the total number of days when recommendations are provided for skipping the medical test against the total number of days in the training set. It is defined as

$$Saving = \frac{N_{No}}{|D|} \times 100\% \quad (3)$$

where  $N_{No}$  denotes the number of days when the recommendation of skipping the medical test is produced.

3. Risk refers to the percentage of the days with risky recommendation against the total number of days in the training set. It is defined as

$$Risk = \frac{N_R}{|D|} \times 100\% \quad (4)$$

where  $N_R$  denotes the number of days with a risky recommendation that refers to a recommendation that suggests skipping a medical test for a given medical measurement but the actual reading of the measurement in the testing set is abnormal.

Here, a correct recommendation means that the approach produces the recommendation of "test required" for the following day and the actual reading for that day in the dataset is normal. The recommendations other than these two cases are considered incorrect.

All our experiments are conducted using MATLAB on a desktop computer configured with a 3.40 GHz Intel core i7 CPU processor with 8.00 GB RAM.

### 3.3. Recommendation effectiveness under different sub-bands and statistical features

The first experiment was designed to evaluate the recommendation performance of the proposed method under various DTCWT sub-bands. The optimum number of features from those sub-bands were selected and extracted based on the performance of the proposed method. The  $p$ -values were obtained by comparing the features extracted from the sub-bands (Zhu et al., 2017c). The t-test was conducted separately for all sub-bands ( $y_1$ ,  $y_2$ ,  $y_3$ ,  $y_4$  and  $z_4$ ). Table 3 and Table 4 show the corresponding values of accuracy, workload saving, risk and  $p$ -values obtained from each sub-band based on the features. It can be clearly seen from the associated  $p$ -value analysis that the high frequency sub-bands (i.e.,  $y_1$ ,  $y_2$ , and  $y_3$ ) provide a significant difference ( $p < 0.003$ ) towards good recommendation performance, indicating that the high frequency sub-bands (i.e.,  $y_1$ ,  $y_2$ , and  $y_3$ ) are able to produce more accurate recommendations compared to the low-frequency sub-bands (i.e.,  $y_4$  and  $z_4$ ). Hence, features from the high frequency sub-bands (i.e.,  $y_1$ ,  $y_2$ , and  $y_3$ ) were selected to represent the time series data of the patient, while the features from the low-frequency sub-bands (i.e.,  $y_4$  and  $z_4$ ) were ignored.

Different sub-bands sets (combinations) were tested and recorded thoroughly to identify the best combination of sub-bands to design the proposed method. These results are presented in Table 5 and Table 6 that the performance of the proposed method improves significantly when a combination of the three high-frequency sub-bands ( $y_1$ ,  $y_2$ , and  $y_3$ ) was used. In this experiment, the extracted statistical features from the coefficients of  $y_1$ ,  $y_2$ , and  $y_3$  sub-bands were used as the input to the base classifiers as well as to the ensemble model.

However, to study the relationship between the number of the extracted features and the recommendation effectiveness of the proposed method, several experiments were conducted using different sets of extracted features. First, the six statistical features were tested separately using the high frequency sub-bands. Figure 4 shows the order of the statistical features based on their performance. The features were then sorted in ascending order based on their potential patient's condition. In this experiment, the extracted statistical features from the high-frequency sub-bands (i.e.,  $y_1$ ,  $y_2$ , and  $y_3$ ) were used as the input for the ensemble model.

Furthermore, several experiments under different number of features were conducted in this study to produce accurate recommendations. Different statistical feature sets were tested and recorded thoroughly to test the performance of the proposed method. Table 7 shows the performance of recommendation approach under different numbers of features. From the obtained results in Table 7, it is clear that the recommendation

Table 3: Prediction performance using single sub-bands for Tunstall dataset

Names of the used sub-bands	Accuracy (%)	Saving (%)	Risk (%)	<i>p</i> -value
y1	86.18	62.95	05.50	3.3359E-3
y2	89.30	63.85	05.10	2.2058E-5
y3	90.40	64.95	04.95	2.4268E-6
y4	82.45	60.70	06.70	0.01049
z4	81.15	60.80	06.20	0.35026

Table 4: Prediction performance using single sub-bands for Diabetes dataset

Names of the used sub-bands	Accuracy (%)	Saving (%)	Risk (%)	<i>p</i> -value
y1	85.10	61.80	05.80	3.8219E-4
y2	88.40	62.90	05.30	2.3328E-5
y3	91.40	63.95	05.90	2.3251E-5
y4	82.10	60.40	06.90	0.01329
z4	80.10	60.30	06.70	0.25410

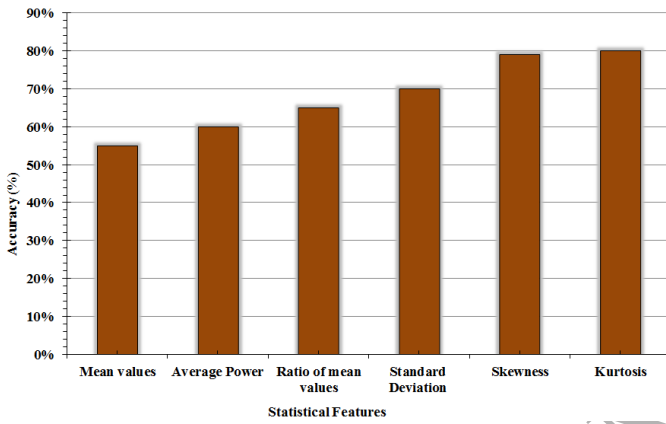


Fig. 4: Ranking of the statistical features based on their accuracy performance

performance of the method in term of the three performance evaluation measurements is clearly improved by increasing the number of features used. For example, using the six statistical features set, the recommendation accuracy, workload saving and risk of the proposed method are approximately 2.7%, 2.4% and 53% better than using four feature set for both datasets. Thus, it is suggested to use all the six statistical features for the three high-frequency sub-bands if necessary for producing accurate recommendations for chronic disease patients. As a result, the recommender method system has achieved a significant accuracy and workload saving with a low risk level. For instance, the recommendation approach can achieve an accuracy over 96%, 65% for a workload saving over while the risk rate is around 1.47%. These results have demonstrated that our recommendation method is highly accurate and is able to significantly reduce the workload for patients to take up their daily medical tests with a very low health risk.

### 3.4. Effectiveness comparison with previous approaches

In this section, we compare the proposed method with a set our previous work that used the same Tunstall dataset to tackle the exactly same problem as presented in this work. In our previous work, we developed a basic heuristic approach based on some fundamental heuristic rules for producing appropriate recommendations to chronic heart disease patients in a telehealth environment Lafta et al. (2015). Then, we combined the basic heuristic approach with the regression-based prediction algorithm and a hybrid method Lafta et al. (2016) to tackle the same task. We also used a fast Fourier transformation-coupled with a machine learning based ensemble model for providing appropriate recommendations to patients suffering chronic diseases Lafta et al. (2017). Kang et al. (2015) have used SVM as a classification algorithm for the anti-diabetic drug failure prediction problem. The proposed method provides an efficient and effective way to use SVM for large-scale datasets with an accuracy of about 80%.

Table 11 presents the comparisons results. Based on the obtained results, it was noticed that the proposed method achieved the best accuracy among the four approaches. In this method, the accuracy performance was improved from 94% to 96%

performance of the proposed method become more accurate by increasing the number of statistical features used.

Finally, to investigate the recommendation effectiveness of the recommendation approach under different number of statistical features, three different feature sets containing respectively two, four and six statistical features were selected to evaluate the proposed method based on the sorted feature information presented in Figure 4. Only the top two, four and six set features in Figure. 4 were selected into the three sets. The detailed features selected for the three evaluated sets are as follows:

Two-feature set: Kurtosis and Skewness

Four-feature set: Kurtosis, Skewness, Standard Deviation and Ratio of mean values

Six-feature set: Kurtosis, Skewness, Standard Deviation, Ratio of mean values, Average power and Mean values

For each feature set, the performance of the proposed method is evaluate when only the three high-frequency sub-bands are used for each sliding window can indicate the impact of the choice of sub-bands. The three sub-bands used are y1, y2, and y3 respectively.

Table 8, Table 9 and Table 10 show the detailed experiment results for two, four and six-feature sets, respectively. The obtained results from the above tables, it can be seen that the



Table 5: Prediction performance using combinations of sub-bands for Tunstall dataset

Names of the used sub-bands	Accuracy (%)	Saving (%)	Risk (%)	<i>p</i> -value
y1,y2	94.20	64.20	02.50	6.2245E-4
y1,y3	94.90	65.40	02.20	3.4129E-5
y2,y3	95.45	65.55	01.90	4.6328E-6
y1,y2,y3	96.70	66.50	01.25	2.8493E-11

Table 6: Prediction performance using combinations of sub-bands for Diabetes dataset

Names of the used sub-bands	Accuracy (%)	Saving (%)	Risk (%)	<i>p</i> -value
y1,y2	93.80	63.90	03.00	2.3015E-5
y1,y3	94.00	64.80	02.80	3.3254E-4
y2,y3	95.30	65.10	02.95	4.5243E-5
y1,y2,y3	96.00	66.20	01.60	2.6524E-10

Table 7: The average performance of the proposed method under different numbers of features

Feature sets	Accuracy (%)	Saving (%)	Risk (%)
2-feature set	90.10	63.04	04.75
3-feature set	91.00	64.35	04.30
4-feature set	94.05	63.20	02.77
5-feature set	94.70	64.50	02.20
6-feature set	96.82	65.69	01.47

Table 8: The model performance using a two-feature set.

Dataset	Classifiers	Accuracy (%)	Saving (%)	Risk (%)
Tunstall dataset	LS-SVM	87.25	61.35	05.60
	Neural Network	85.80	60.55	06.00
	Naive Bayes	86.82	60.10	05.90
	<b>Ensemble model</b>	<b>90.20</b>	<b>63.50</b>	<b>04.70</b>
Diabetes dataset	LS-SVM	86.80	60.75	06.00
	Neural Network	84.95	60.40	06.10
	Naive Bayes	86.00	60.50	06.05
	<b>Ensemble model</b>	<b>90.00</b>	<b>63.30</b>	<b>04.80</b>

Table 9: The model performance using a four-feature set.

Dataset	Classifiers	Accuracy (%)	Saving (%)	Risk (%)
Tunstall dataset	LS-SVM	91.60	63.55	03.95
	Neural Network	90.30	62.65	04.50
	Naive Bayes	90.75	62.25	04.45
	<b>Ensemble model</b>	<b>94.10</b>	<b>63.40</b>	<b>02.60</b>
Diabetes dataset	LS-SVM	91.00	63.00	04.10
	Neural Network	89.80	62.00	04.75
	Naive Bayes	90.50	62.00	04.60
	<b>Ensemble model</b>	<b>94.00</b>	<b>63.10</b>	<b>02.95</b>

Table 10: The model performance using a six-feature set.

Dataset	Classifiers	Accuracy (%)	Saving (%)	Risk (%)
Tunstall dataset	LS-SVM	94.55	63.95	02.55
	Neural Network	93.15	63.20	02.90
	Naive Bayes	92.30	62.85	03.60
	<b>Ensemble model</b>	<b>96.90</b>	<b>65.88</b>	<b>01.20</b>
Diabetes dataset	LS-SVM	94.00	63.50	02.75
	Neural Network	92.95	63.00	03.10
	Naive Bayes	92.00	62.50	03.90
	<b>Ensemble model</b>	<b>96.75</b>	<b>65.50</b>	<b>01.75</b>

Table 11: Recommendation performance comparison with previous methods

<b>Tunstall dataset</b>				
Method	Techniques used	Accuracy(%)	Saving(%)	Risk(%)
Lafta et al. (2015)	Basic heuristic algorithm	86	10	8
Lafta et al. (2016)	Basic heuristic algorithm, Regression-based algorithm and Hybrid algorithm	91	15	5
Lafta et al. (2017)	Fast Fourier transformation and ensemble model	94	63	3
Kang et al. (2015)	Ensemble model of support vector machines for anti-diabetic drug failure prediction	80	-	-
<b>Proposed method</b>	Dual-tree complex wavelet transformation and ensemble model	<b>96</b>	<b>65</b>	<b>1</b>

while the workload saving has a little improvement from 63% to more than 65%. In addition, the recommendation risk of our method is also lower than the three competitive approaches.

#### 4. Conclusions and Future Research Directions

In this work, a novel approach based on the Dual-Tree Complex Wavelet Transformation (DTCWT) with the machine learning ensemble model was developed for short-term disease risk prediction. The proposed method showed that it can be used as an effective medical test recommendation tool in a telehealth environment for patients suffering from chronic disease such as heart and diabetes diseases. Six statistical features were extracted from four-level decomposition of DTCWT, and were used as the key features set. The ensemble model used in the approach involves the use of Least Square-Support Vector Machine, Neural Network, and Naive Bayes to achieve accurate and reliable recommendations for patients with chronic diseases.

The experimental results showed that the proposed method using high-frequency sub-bands,  $y_1$ ,  $y_2$  and  $y_3$  yielded a higher predictive performance compared with the other sub-bands for predicting the patient's condition. Our findings showed that using a combination of six statistical features gives a better performance of the proposed method than the other feature sets. In addition, the proposed method is compared with individual base classifiers and some previous work used to solve the same problem. It was found that the proposed method is more effective than the individual classifiers used in constructing the ensemble model and outperforms the previously proposed approaches. Based on the experimental evaluations, we believe that the proposed method can be used as an effective tool to improve the quality of clinical evidence-based decisions by which the time and cost associated with the daily medical test can be reduced. The experimental evaluation of this approach establishes that the proposed method is an effective in improving the quality of the medical decisions based on clinical evidence and reducing the time costs incurred by the patients suffering from chronic diseases in taking their daily medical tests, whereby improving their generic lives quality.

In future, we are interested in applying other ensemble techniques, such as boosting and Adaboost, to produce accurate recommendations and conducting a comparative study on those different ensemble models. In addition, to make the proposed

method is more generic in dealing with medical time series data, it is possible to apply the proposed method to support telehealth care for patients who suffering from other types of diseases.

#### Acknowledgement

The authors would like to thank the support from Guangxi Key Laboratory of Trusted Software (No. kx201615) and Capacity Building Project for Young University Staff in Guangxi Province, Department of Education, Guangxi Province (No. ky2016YB149).

#### References

- Bashir, S., Qamar, U., Khan, F.H., 2015. Bagmoov: A novel ensemble for heart disease prediction bootstrap aggregation with multi-objective optimized voting. *Australasian physical & engineering sciences in medicine* 38, 305–323.
- Belayneh, A., Adamowski, J., Khalil, B., Quilty, J., 2016. Coupling machine learning methods with wavelet transforms and the bootstrap and boosting ensemble approaches for drought prediction. *Atmospheric Research* 172, 37–47.
- Breiman, L., 1996. Bagging predictors. *Machine learning* 24, 123–140.
- Chang, C.D., Wang, C.C., Jiang, B.C., 2011. Using data mining techniques for multi-diseases prediction modeling of hypertension and hyperlipidemia by common risk factors. *Expert systems with applications* 38, 5507–5513.
- Danaei, G., Finucane, M.M., Lu, Y., Singh, G.M., Cowan, M.J., Paciorek, C.J., Lin, J.K., Farzadfar, F., Khang, Y.H., Stevens, G.A., et al., 2011. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *The Lancet* 378, 31–40.
- Das, A.B., Bhuiyan, M.I.H., Alam, S.S., 2016. Classification of eeg signals using normal inverse gaussian parameters in the dual-tree complex wavelet transform domain for seizure detection. *Signal, Image and Video Processing* 10, 259–266.
- Das, R., Turkoglu, I., Sengur, A., 2009. Effective diagnosis of heart disease through neural networks ensembles. *Expert systems with applications* 36, 7675–7680.
- Dewar, A.R., Bull, T.P., Malvey, D.M., Szalma, J.L., 2017. Developing a measure of engagement with telehealth systems: The mhealth technology engagement index. *Journal of telemedicine and telecare* 23, 248–255.
- Gao, H., Jian, S., Peng, Y., Liu, X., 2016. A subspace ensemble framework for classification with high dimensional missing data. *Multidimensional Systems and Signal Processing*, 1–16.
- Helmy, T., Rahman, S., Hossain, M., Abdelraheem, A., 2013. Non-linear heterogeneous ensemble model for permeability prediction of oil reservoirs. *Arabian Journal for Science & Engineering (Springer Science & Business Media BV)* 38.
- Hu, R., Zhu, X., Cheng, D., He, W., Yan, Y., Song, J., Zhang, S., 2017. Graph self-representation method for unsupervised feature selection. *Neurocomputing* 220, 130–137.

- Huang, F., Wang, S., Chan, C.C., 2012. Predicting disease by using data mining based on healthcare information system, in: *Granular Computing (GrC)*, 2012 IEEE International Conference on, IEEE. pp. 191–194.
- Jung, C., Yang, Q., Sun, T., Fu, Q., Song, H., 2017. Low light image enhancement with dual-tree complex wavelet transform. *Journal of Visual Communication and Image Representation* 42, 28–36.
- Kahn, M., 1994. AIM-94 data set provided by Michael Kahn, MD, PhD, Washington University, St. Louis, MO, US.
- Kang, S., Kang, P., Ko, T., Cho, S., Rhee, S.j., Yu, K.S., 2015. An efficient and effective ensemble of support vector machines for anti-diabetic drug failure prediction. *Expert Systems with Applications* 42, 4265–4273.
- Kingsbury, N.G., 1998. The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters, in: *Proc. 8th IEEE DSP workshop*, Utah. p. 86.
- Krishnaiah, V., Narsimha, D.G., Chandra, D.N.S., 2013. Diagnosis of lung cancer prediction system using data mining classification techniques. *International Journal of Computer Science and Information Technologies* 4, 39–45.
- Kuh, D., Shlomo, Y.B., 2004. A life course approach to chronic disease epidemiology. 2, Oxford University Press.
- Lafta, R., Zhang, J., Tao, X., Li, Y., Abbas, W., Luo, Y., Chen, F., Tseng, V.S., 2017. A fast fourier transform-coupled machine learning-based ensemble model for disease risk prediction using a real-life dataset, in: *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Springer. pp. 654–670.
- Lafta, R., Zhang, J., Tao, X., Li, Y., Diikh, M., Lin, J.C.W., 2018. A structural graph-coupled advanced machine learning ensemble model for disease risk prediction in a telehealthcare environment, in: *Big Data in Engineering Applications*. Springer. pp. 363–384.
- Lafta, R., Zhang, J., Tao, X., Li, Y., Tseng, V.S., 2015. An intelligent recommender system based on short-term risk prediction for heart disease patients, in: *Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 2015 IEEE/WIC/ACM International Conference on, IEEE. pp. 102–105.
- Lafta, R., Zhang, J., Tao, X., Li, Y., Tseng, V.S., Luo, Y., Chen, F., 2016. An intelligent recommender system based on predictive analysis in telehealthcare environment, in: *Web Intelligence*, IOS Press. pp. 325–336.
- Lavoie, L., Khoury, H., Welner, S., Briere, J.B., 2016. Burden and prevention of adverse cardiac events in patients with concomitant chronic heart failure and coronary artery disease: a literature review. *Cardiovascular therapeutics* 34, 152–160.
- Li, M., Chen, W., Zhang, T., 2017. Automatic epileptic eeg detection using dt-cwt-based non-linear features. *Biomedical Signal Processing and Control* 34, 114–125.
- Li, S., Tang, B., He, H., 2016. An imbalanced learning based mdr-tb early warning system. *Journal of medical systems* 40, 164.
- Lu, H.Y., Huang, C.Y., Su, C.T., Lin, C.C., 2014. Predicting rotator cuff tears using data mining and bayesian likelihood ratios. *PloS one* 9, e94917.
- Moghaddass, R., Rudin, C., Madigan, D., 2016. The factorized self-controlled case series method: An approach for estimating the effects of many drugs on many outcomes. *Journal of Machine Learning Research* 17, 1–24.
- Mohktar, M.S., Redmond, S.J., Antoniadis, N.C., Rochford, P.D., Pretto, J.J., Basilakis, J., Lovell, N.H., McDonald, C.F., 2015. Predicting the risk of exacerbation in patients with chronic obstructive pulmonary disease using home telehealth measurement data. *Artificial intelligence in medicine* 63, 51–59.
- Polat, K., Güneş, S., 2007. Breast cancer diagnosis using least square support vector machine. *Digital Signal Processing* 17, 694–701.
- Sánchez, A.S., Iglesias-Rodríguez, F.J., Fernández, P.R., de Cos Juez, F., 2016. Applying the k-nearest neighbor technique to the classification of workers according to their risk of suffering musculoskeletal disorders. *International Journal of Industrial Ergonomics* 52, 92–99.
- Swami, P., Gandhi, T.K., Panigrahi, B.K., Bhatia, M., Santhosh, J., Anand, S., 2017. A comparative account of modelling seizure detection system using wavelet techniques. *International Journal of Systems Science: Operations & Logistics* 4, 41–52.
- Valentini, G., Masulli, F., 2002. Ensembles of learning machines, in: *Italian Workshop on Neural Nets*, Springer. pp. 3–20.
- Van Ma, L., Kim, J., Park, S., Kim, J., Jang, J., 2016. An efficient session-weight load balancing and scheduling methodology for high-quality telehealth care service based on webrtc. *The Journal of Supercomputing* 72, 3909–3926.
- Verma, L., Srivastava, S., Negi, P., 2016. A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data. *Journal of medical systems* 40, 1–7.
- Voyant, C., Notton, G., Kalogirou, S., Nivet, M.L., Paoli, C., Motte, F., Fouilloy, A., 2017. Machine learning methods for solar radiation forecasting: A review. *Renewable Energy* 105, 569–582.
- Wang, J., Qiu, M., Guo, B., 2017a. Enabling real-time information service on telehealth system over cloud-based big data platform. *Journal of Systems Architecture* 72, 69–79.
- Wang, N., Cheng, J., Han, B., Li, Q., Chen, Y., Xia, F., Jiang, B., Jensen, M.D., Lu, Y., 2017b. Exposure to severe famine in the prenatal or postnatal period and the development of diabetes in adulthood: an observational study. *Diabetologia* 60, 262–269.
- Wang, S., Li, Z., Chao, W., Cao, Q., 2012. Applying adaptive over-sampling technique based on data density and cost-sensitive svm to imbalanced learning, in: *Neural Networks (IJCNN)*, The 2012 International Joint Conference on, IEEE. pp. 1–8.
- Yang, D., Wang, S., Li, Z., 2018. Ensemble neural relation extraction with adaptive boosting. *arXiv preprint arXiv:1801.09334*.
- Yang, J.G., Kim, J.K., Kang, U.G., Lee, Y.H., 2014. Coronary heart disease optimization system on adaptive-network-based fuzzy inference system and linear discriminant analysis (anfis-lda). *Personal and ubiquitous computing* 18, 1351–1362.
- Yeh, D.Y., Cheng, C.H., Chen, Y.W., 2011. A predictive model for cerebrovascular disease using data mining. *Expert Systems with Applications* 38, 8970–8977.
- Yu, B., Jia, B., Ding, L., Cai, Z., Wu, Q., Law, R., Huang, J., Song, L., Fu, S., 2016. Hybrid dual-tree complex wavelet transform and support vector machine for digital multi-focus image fusion. *Neurocomputing* 182, 1–9.
- Zhang, J., Lafta, R.L., Tao, X., Li, Y., Chen, F., Luo, Y., Zhu, X., 2017. Coupling a fast fourier transformation with a machine learning ensemble model to support recommendations for heart disease patients in a telehealth environment. *IEEE Access* 5, 10674–10685.
- Zhang, S., Li, X., Zong, M., Zhu, X., Wang, R., 2018. Efficient knn classification with different numbers of nearest neighbors. *IEEE transactions on neural networks and learning systems* 29, 1774–1785.
- Zheng, W., Zhu, X., Zhu, Y., Hu, R., Lei, C., 2017. Dynamic graph learning for spectral feature selection. *Multimedia Tools and Applications*, 1–17.
- Zhu, X., Li, X., Zhang, S., Ju, C., Wu, X., 2017a. Robust joint graph sparse coding for unsupervised spectral feature selection. *IEEE transactions on neural networks and learning systems* 28, 1263–1275.
- Zhu, X., Li, X., Zhang, S., Xu, Z., Yu, L., Wang, C., 2017b. Graph pca hashing for similarity search. *IEEE Transactions on Multimedia* 19, 2033–2044.
- Zhu, X., Zhang, J., Li, H., Fournier-Viger, P., Lin, J.C.W., Chang, L., 2017c. Froid: A deeply integrated feature-rich interactive system for effective and efficient outlier detection. *IEEE Access* 5, 25682–25695.
- Zhu, X., Zhang, S., Hu, R., Zhu, Y., et al., 2018. Local and global structure preservation for robust unsupervised spectral feature selection. *IEEE Transactions on Knowledge and Data Engineering* 30, 517–529.