Web Intelligence and Agent Systems: An International Journal 0 (2017) 1–12
IOS Press

An Intelligent Recommender System based on Predictive Analysis in Telehealthcare Environment

RaidLafta [a], Ji Zhang [a], Xiaohui Tao [a], Yan Li [a], Vincent S. Tseng [b], Yonglong Luo [c], and Fulong Chen [c]

[a] *Faculty of Health, Engineering and Sciences, University of Southern Queensland , Australia*
E-mail:{RaidLuaibi.Lafta, ji.zhang, xtao, yan.li}@usq.edu.au
[b] *Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan*
E-mail: vtseng@cs.nctu.edu.tw
[c] *School of Mathematics and Computer Science, Anhui Normal University, China*
E-mail: ylluo@ustc.edu.cn, long005@mail.ahnu.edu.cn

Abstract. The use of intelligent technologies for providing useful recommendations to patients suffering chronic diseases may play a positive role in improving the general life quality of patients and help reduce the workload and cost involved in their daily healthcare. The objective of this study is to develop an intelligent recommender system based on predictive analysis for advising patients in the telehealth environment concerning whether they need to take the body test one day in advance by analyzing medical measurements of a patient for the past k days. The proposed algorithms supporting the recommender system have been validated using a time series telehealth data recorded from heart disease patients which were collected from May to January 2012, from our industry collaborator Tunstall. The experimental results show that the proposed system yields satisfactory recommendation accuracy and offer a promising way for saving the workload for patients to conduct body tests every day. This study highlights the possible usefulness of the computerized analysis of time series telehealth data in providing appropriate recommendations to patients suffering chronic diseases such as heart diseases patients.

Keywords: Intelligent system, Recommender system, Heart failure, Time series prediction, Telehealth

1. Introduction

In recent years, the huge volume of data generated by healthcare organizations such as hospitals and medical centres are taken different formats such as texts, numbers, charts, and images [2]. Due to the wealthy information they contain, these data may be used to support evidence-based decisions in healthcare. These days, important clinical decisions are usually made based on specialists experience rather than historical medical data of patients. This practice may lead to adverse biases, human errors and high medical costs and consequently, affects the quality of services provided to patients [9].

The chronical diseases such as heart disease have become the main public health issue worldwide which accounting for 50% of global mortality burden [26]. Excessive efforts have been done to assess and predict chronical diseases using clinical decision support systems, for example, trying to predict heart disease at an early stage [25].

In the last decades, much research efforts have been invested in the assessment of different diseases risk in healthcare domain. Many of these researches have been carried out in data mining and analytic on medical data. The assessment and prediction of various diseases for patients have been widely reviewed by using data mining techniques and statistical tools [4, 5, 6, 7, 8, 25]. Especially, there have been a number of the predictive models which employ artificial neural network and regressions for predicting medical outcomes in various chronical diseases [25, 27, 28, 31]. However, a challenge remains still in securing an effective analytic tool with high accuracy to help support personalized evidence-based decisions.

Telehealth care can be defined as the delivery of health-related services via telecommunication technologies such as telephone or the Internet [29]. Evidence shows that telehealth care could save money and time by reducing emergency department visits, hospital admissions, associated travel and physical limitations [24]. Moreover, telehealth can be used to improve patient health outcomes in health care domain[30]. A great diversity of intelligent techniques including computing techniques have been used to solve complex problems[32][33]. Intelligent technologies can be developed to greatly facilitate the development and deployment of telehealth systems for patients including those who suffer from chronical diseases such as heart disease and require continuous monitoring of their heart-related medical measurements.

When performing remote continuous monitoring of patients' key measurement readings, an abundance of time series data are generated. Most of the existing predictive analytic methods on medical time series data are used to predict the long-term risk (e.g., the chance of survival) or the diagnosis of diseases. Nevertheless, it turns out that short-term prediction is more difficult than long-term projection due to a higher level of short-term uncertainty existing in the readings of various medical measurements. In addition, short-term recommendations will be equally useful for patients as they provide guidance as to what the patients need to do for the next few days.

The main contributions of this paper is offering an intelligent recommender system, supported by several innovative predictive algorithms, for short-term risk assessment on patients in telehealth environment based on analytic of a patient's historical medical data. On the basis of assessment results, the system provides recommendations to patients such as heart diseases patients in relation to the necessity of medical tests taken on a daily basis. The research is conducted with an aim at improving the quality of clinical evidenced-based decisions made by medical practitioners to help reduce the financial and timing cost taken by patients.

To verify the performance of the proposed system, the metrics of accuracy and workload saving are used and experimental evaluations are conducted on a real-life time series data collected from a pilot study on a group of heart failure patients. The experimental results demonstrate that the proposed system yields reasonably good recommendation accuracy and can effectively reduce the workload required in medical tests for the patients. We believe that this system is promising in risk assessment and management associated with heart failure and other similar diseases.

The remainder of this paper is organised as follows. In Section 2, we briefly introduce the related work in predictive data analytic and mining. Section 3 formally formulates the research problem. We present the overall framework of our recommender system In Section 4 and present the details of the system and the predictive analytic algorithms in Section 5. The experiment design, including the experiment environment and the datasets, is discussed in Section 6. Section 7 presents the experimental evaluation results. Finally, Section 8 concludes general conclusions and highlight our future research work.

2. Related Work

Data mining technologies and statistical analysis have been broadly used to different medical and healthcare problems. Several studies have utilized predictive data mining techniques such as classification by decision tree induction, Bayesian classification, Neural Networks, Support Vector Machine (SVM), and classification based on association to predict the various diseases risk in medical domain [23].

One of the prediction methods used in the healthcare domain is prediction the disease risk level in the patients. Genetic algorithm, logistic regression and decision tree have been used to predict the severity level of disease in patients [6, 10, 11, 12]. Prediction models have been designed to assess the risk of different diseases. In these works, laboratory measurements and symptoms were utilized to predict the disease risk.

On other hand, Bayesian classifiers, decision tree, logistic regression and back propagation neural network have been utilized by [4, 5, 7, 13, 14] to predict diseases at an early stage for patients. In these studies, effective predictive models have been built to discover different diseases as early as possible in order to treat them effectively.

However, statistical analytic tools have been effectively used to estimate lifetime and long-term risk for patients with different diseases [15, 16, 17]. Clinical predictive models have also been developed based on measurements taken on patients in different timely basis, aiming to detect diseases at an early stage and treat them at the right time.

The predictive survival models have been suggested by [18, 19, 20] to analyze patients' survival times and to help develop a superior treatment plan for the diseases. Statistical analytic such as supervised wavelet approximation coefficients and multivariate piecewise Poisson regression method have also been effectively utilized in these works. In all such studies, data in patients' medical profiles such as age, sex, blood pressure and blood sugar, etc. have played an important role and provided informative evidence in decision-making support.

In summary, effective applications of artificial neural network (ANN) and logistic regression to predict medical outcomes have been proved in many works [25, 27, 28, 31]. However, a challenge remains still in securing an effective analytic tool with high accuracy to help make evidence-based decisions with quality efficiently. In this work, we propose an innovative intelligent recommender system to predict and assess the short-term risk of disease and provide appropriate recommendations to the patients for the necessity of taking a medical test in the incoming day.

3. Research Problem Formulation

The ultimate goal of this work is developing an intelligent recommendations system to predict and assess the short-term risk of diseases and provide decision-making support, aiming at reducing the associated cost without sacrificing the quality of healthcare services in the telehealth environment.

The research problem can be formalized as the following. Let D_{med} be the set of observations including measurements and symptoms, and $T = t_1, t_2, t_3, \ldots, t_n$ represents the set of timestamps involved in the dataset D_{med}. D_{med} is thus a $m \times (n+1)$ table, where $D_{med}(i,j)$ means the jth observation at the ith timestamp. Conceptually, the prediction model can also be understood as the following mapping functions:

$$f(D_{med}, k) \longrightarrow \mathbf{Risk} \tag{1}$$

$$f(\mathbf{Risk}) \longrightarrow \mathbf{Recomm} \tag{2}$$

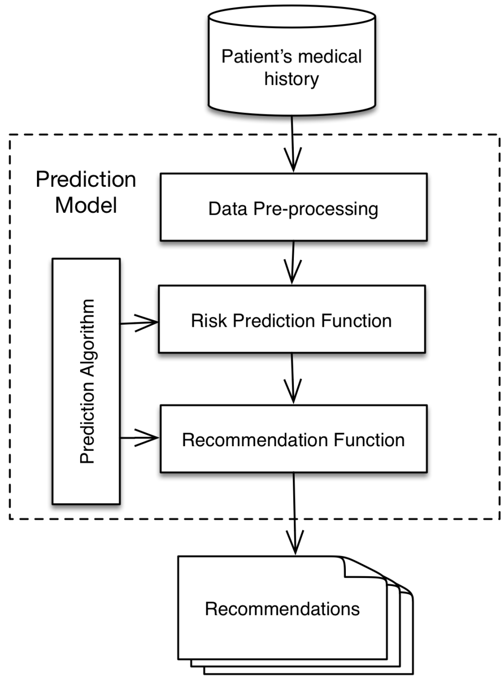

Fig. 1. System Architecture of our recommender system

The first function is the risk prediction function in which observations over the last k timestamps are analyzed to predict the risk of heart disease. $\mathbf{Risk}$ denotes the disease risk level. The second function is the recommendation function that provides recommendations to patients based on the outcome of the risk prediction function, where $\mathbf{Recomm}$ refers to the recommendation(s) produced.

4. Framework

In this paper, we propose an intelligent recommendation system equipped with an innovative predictive algorithm to analyze the medical data of patients, assess their risk and provide them with appropriate recommendations to the patients for the necessity of taking a medical test in the following day based on the outcomes of prediction. The system architecture is illustrated in Figure 1 and the technical details of different major components of the system are discussed in the following subsections.

4.1. Data Preprocessing

Data pre-processing is an important step in the data analysis process that involves transforming raw data into a comprehensible format. Raw data are usually incomplete, duplicated, and/or inconsistent data; therefore we need to prepare them before starting in analysis tasks. The collection of duplicate and missing data would normally result from human or computer errors at the time of entry. Therefore, data pre-processing play a major role in sanitizing the raw data from such inconsistencies and anomalies resulting from data collection.

In this stage, missing data problem, which is caused by the data collection or transmission, is resolved by filling the missing data with a global constant. Also, the noise records with incorrect readings are removed. Another important task of data pre-processing in this work is to extract the information for each individual patient from the original dataset for personalized data analysis and recommendations.

4.2. Time Series Recommendation Algorithm

The key component of the proposed system is the recommendation algorithm based on time series data analysis. The proposed intelligent recommendation system is equipped with a novel prediction algorithm to analyze the medical data of chronic diseases patients, assess the short-term risk of disease for the patients and then provide them with recommendations based on the outcomes of prediction. In this work, we have developed and incorporated several predictive algorithms to provide recommendations in the system. This system will help reduce the workload for patients from the unnecessary medical checkups that they would otherwise have to take every day.

4.3. Human Computer Interaction for the System

Our recommender system involves human-computer interaction to receive input from human users concerning the values of the parameters that are used in the algorithm of our system. The recommendations generated by our system will be returned back to users through different channels and platforms including desktops, laptops and tablets to embrace the latest technological advancements for quick information dissemination. Besides returned back to the patients, the results can also be sent remotely to the practitioners such as doctors and nurses so that they can be informed and keep track of the physical checkups and overall health conditions of the patients. Easy-to-use user-friendly interfaces are developed for the users to supply necessary inputs to the system and receive recommendations from the system.

5. Predictive Algorithms of the System

The predictive algorithm is the key component of the proposed recommender system to analyze the medical data of patients, assess their risk and then provide appropriate recommendations to the patients for the necessity of taking a medical test in the following day based on the outcomes of prediction. Specifically, three different predictive algorithms are developed in this work to provide the recommendation to patients.

The complexity of predictive algorithms are measured using big-O notation. In predictive algorithms, let n be the number of days in our dataset, and k represents the length of the sliding time window, therefore the complexity of each algorithm is $n \times k$. However, we do not focus on the running times of the algorithms in this paper.

5.1. Basic Heuristic Algorithm

A basic heuristic algorithm is developed to decide whether a given patient needs to take a medical measurement such as the heart rate test today based on a study of his/her measurement readings for the past k days. If the patient satisfies both of the following two heuristic rules for a measurement, a recommendation of "no test needed" will be generated, and the patient does not need to take the test on the following day for that measurement:

- *She (he) has taken the test for no less than p% of the past k days for this measurement* ($0 <= p <= 100$)*, and*
- *All the readings of this measurement during the past k days are normal.*

The "no test needed" recommendation will be provided to the patient and stored in the backend database as a part of the patient's historical records. If any of the conditions is not satisfied, a recommendation of "test required" will be generated and the patient is suggested to take the medical test on the following day. Again, the recommendation will be stored into the system as a historical record.

Overall, there are four parameters in the recommendation algorithm, i.e., the minimum (min) and maximum (max) of normal values for each measurement, setting up the boundary of healthy range; the length of the sliding time window k, and the minimum percentage (p) of days when medical test is conducted for the measurement in the past k days. The recommendation algorithm is presented in Algorithm 1.

Input : Patient's time series medical testing data (e.g., heart rates).
Output: $\mathbf{Rick} = [0|1]$ (0: low risk; 1: high risk);
$\mathbf{Recomm} = [0|1]$ (0: no test required; 1: test needed.)

```
let k be a limited number from the past days (the length of time window k);
let p% be a value between the range (0 <= p <= 100);
let max and min be the boundary of healthy values in the test;
foreach days for the patient do
    if the patient has taken the test for no less than p% of the past k days (<= k − (p × k)) AND all the measured values during the past k days are normal then
        Risk = 0;
    else
        Risk = 1;
    end
    if Risk == 0 then
        Recomm = 0;
    else
        Recomm = 1;
    end
end
return Recomm.
```

Algorithm 1: Basic Time Series Prediction Algorithm

The algorithm of our recommender system is presented in Algorithm 1. It evaluates the time series data collected for a patient on a continuous basis using a sliding window with a length of k, which is the number of days in the past that the algorithm will look at in support of assessment evaluation for the following day. Two conditions are evaluated in the IF predicate from Line 5 to 14. The first one evaluates the percentage of the actual medical test that has been carried out in the past k days. If a test is skipped for a day, actual reading will be missing and as a result, the certainty and accuracy for risk assessment will drop for future days. Therefore, an upper bound is imposed in this condition on the total number of days when the medical testing is skipped in each sliding window. In addition to this bound, we also require that the readings of all the medical checkups conducted during the past k days are in the normal range for the measurement, as dictated by its corresponding minimum and maximum threshold values. Intuitively speaking, normal readings improve the confidence that the short-term risk is low, whereby a skip of the test on this measurement can be recommended. If both the two conditions are satisfied, then the risk for skipping the physical test for this measurement is deemed low and accordingly a recommendation for skipping a test for the measurement can be made for the following day. Otherwise, the system will provide a recommendation urging patients to take a medical test for the measurement on the following day and the reading of the measurement will be received and stored in the database. It's worthwhile mentioning that the above algorithm will be applied to a single measurement once at the time for risk assessment and recommendation. If the recommendations are to be performed for multiple measurements, then we simply run our system separately to produce recommendations for different measurements.

5.2. Regression-Based Algorithm

The second version of the predictive algorithm is developed based on the linear regression model. This involves modifying the second rule of the basic heuristic algorithm by leveraging the prediction generated by the linear regression model, as presented as follows:

- *She (he) has taken the test for no less than p% of the past k days for this measurement ($0 <= p <= 100$), and*
- *The predicted value for the current day generated by the linear regression based on the past k days is normal.*

Compared with the basic heuristic algorithm, only the second rule has been modified in the regression-based algorithm. Using the data from the past k days a linear regression model is generated for the prediction of the value of the measurement in question for the next day. The predicted measurement value must be normal for producing the "no test needed" recommendation, which will be provided to the patient and stored into the backend database as a part of the patient' historical records. If any of the 2 conditions is not satisfied, a recommendation of "test required" will be generated and the patient is suggested to take the medical test in the following day. Again, the recommendation will be stored into the system as a historical record. The regression-based algorithm is presented in Algorithm 2.

5.3. Hybrid Algorithm

By combining both the basic heuristic and the regression-based algorithms, We can obtain a new hybrid algorithm. The second condition in the basic heuristic and the regression-based algorithms are combined to generate a new condition as follows:

Input : Patient's time series medical testing data (e.g., heart rates).
Output: $\mathbf{Rick} = [0|1]$ (0: low risk; 1: high risk);
$\mathbf{Recomm} = [0|1]$ (0: no test required; 1: test needed.)

let k be a limited number from the past days (the length of time window k);
let $p\%$ be a value between the range $(0 <= p <= 100)$;
let max and min be the boundary of healthy values in the test;
foreach *days for the patient* **do**
  if *the patient has taken the test for no less than p% of the past k days* $(<= k - (p \times k))$ *AND the predicted measurement value for the current day based on the linear regression model is normal* $< max$ **then**
    $\mathbf{Risk} = 0$;
  else
    $\mathbf{Risk} = 1$;
  end
  if $\mathbf{Risk} == 0$ **then**
    $\mathbf{Recomm} = 0$;
  else
    $\mathbf{Recomm} = 1$;
  end
end
return $\mathbf{Recomm}$.

Algorithm 2: Regression-Based Time Series Prediction Algorithm

Input : Patient's time series medical testing data (e.g., heart rates).
Output: $\mathbf{Rick} = [0|1]$ (0: low risk; 1: high risk);
$\mathbf{Recomm} = [0|1]$ (0: no test required; 1: test needed.)

let k be a limited number from the past days (the length of time window k);
let $p\%$ be a value between the range $(0 <= p <= 100)$;
let max and min be the boundary of healthy values in the test;
foreach *days for the patient* **do**
  if *the patient has taken the test for no less than p% of the past k days* $(<= k - (p \times k))$ *AND the predicted measurement value for the current day based on the linear regression model is normal AND all the measured value k days are normal* **then**
    $\mathbf{Risk} = 0$;
  else
    $\mathbf{Risk} = 1$;
  end
  if $\mathbf{Risk} == 0$ **then**
    $\mathbf{Recomm} = 0$;
  else
    $\mathbf{Recomm} = 1$;
  end
end
return $\mathbf{Recomm}$.

Algorithm 3: Hybrid Time Series Prediction Algorithm

- *She (he) has taken the test for no less than p% of the past k days for this measurement* ($0 <= p <= 100$)*, and*
- *all the readings of this measurement during the past k days are normal and the predicted value for the current day generated by the linear regression based on the past k days is normal.*

The hybrid recommendation algorithm is demonstrated in Algorithm 3.

6. Evaluation Design

In this section, we present the details concerning the design of our experimental evaluation including datasets, performance metrics and the experimental platform.

As the predictive performance of the recommendation algorithm is quite important, assessment of potential predictions is critically dependent on the quality of the used dataset. For this reason, telehealth data from Tunstall dataset will be conducted in this work. We use a real-life dataset obtained from our industry collaborator Tunstall to test the practical applicability of the system we propose. A Tunstall dataset obtained from a pilot study has been conducted on a group of heart failure patients and the resulting data were collected for their day-to-day medical readings of different measurements in a telehealth care environment. The Tunstall database employed in the development of the algorithm consists of data from six patients with a total of 7,147 different time series records. Data were acquired between May and January 2012, using a remote telehealth collaborator. The dataset is by nature in a time series and contains a set of measurements taken from the patients on different days. Each record in the dataset consists of a few different meta-data attributes about the patients such as patient-id, visit-id, measurement type, measurement unit, measurement value, measurement question, date and date-received. The characteristics of the features of the dataset are shown in Table 1.

In addition, each record contains a few medical attributes including Ankles, Chest Pain, and Heart Rate, Diastolic Blood Pressure (DBP), Mean Arterial Pressure (MAP), Systolic Blood Pressure (SBP), Oxygen Saturation (SO2), Blood Glucose, and Weight. Ethical clearance was obtained from the University of Southern Queensland(USQ) Human Research Ethics Committee (HREC) prior to the onset of the study. This

Table 1
Characteristic Features of the Dataset

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

dataset is used as the ground truth result to test the performance of our recommendation system. The recommendations produced by our system will be compared with the actual readings of the measurement in question recorded in the dataset to see how accurate our recommendations are.

We devise two performance metrics to evaluate the performance of our proposed system, *accuracy* and *workload saving*. Accuracy refers to the percentage of correctly recommended days against the total number of days that recommendations are provided while workload saving refers to the percentage of the total number of days when recommendations are provided against the total number of days in the dataset. Mathematically, accuracy and workload saving are defined as follows:

$$Accuracy = \frac{NN}{NN + NA} \times 100\% \quad (3)$$

$$Saving = \frac{NN + NA}{|\mathcal{D}|} \times 100\% \quad (4)$$

Where NN denotes the number of days with correct recommendations, NA denotes the number of days with incorrect recommendations and $|\mathcal{D}|$ refers to the total number of days in the dataset. Here, a correct recommendation means that the system produces the recommendation of "no test required" for the following day and the actual reading for that day in the dataset is normal. If this is a case, the recommendation is considered accurate.

We also devise two metrics, namely *Difference* and *Rotating frequency* which will be used to measure data characteristics quantitatively for each medical measurement. We will evaluate their impact on the performance of the recommender system in terms of accuracy and workload saving. Difference refers to the difference in terms of the number of normal and abnormal values for a given medical measurement in the dataset, as dictated by their corresponding minimum and maximum threshold values; whilst rotating frequency refers to the total number of instances where two consecutive values for the medical measurement have different states (one is normal and the other one is abnormal) in the dataset. Both metrics will be normalized against the total number of days in the dataset. Mathematically, difference and rotating frequency for a medical measurement m are defined as follows:

$$Difference(m) = \frac{|NV_m - AV_m|}{|\mathcal{D}|} \times 100\% \quad (5)$$

$$RotatingFrequency(m) = \frac{Rotation}{|\mathcal{D}|} \times 100\% \quad (6)$$

Where NV_m and AV_m denote the number of normal and abnormal values of m in the dataset. Normal and abnormal values are detected based on the corresponding minimum and maximum threshold values of m. $Rotation$ denotes the total number of instances where two consecutive values for the medical measurement have different states. $|\mathcal{D}|$ refers to the total number of days in the dataset. Table 2 demonstrates two examples for difference and rotating frequency.

The prototype of the system has been developed using C/C++ in Microsoft visual studio 2008 on a desktop computer with the configurations of a 3.40 GHz Intel Core i7 CPU processor with 8.00 GB RAM.

7. Result Analysis

In this section, the results of the experimental evaluation conducted on the proposed system are reported. We first evaluate the effect of two key parameters used in our algorithm, k and p. k determines the size of the slide windows (i.e., the number of past days involved in assessing the disease risk of the current day) used for analyzing the time series data, while p is the percentage of the days in the past k days that the patients have

Table 2
Examples of Rotating Frequency

Example1	Exampe2
normal	normal
normal	abnormal
normal	normal
normal	normal
abnormal	abnormal
abnormal	abnormal
abnormal	abnormal
abnormal	abnormal
abnormal	abnormal
abnormal	normal
Difference=2/10=0.2	Difference=2/10=0.2
Rotating Frequency=1/10 =0.1	Rotating Frequency=4/10 =0.4

physically conducted the test for the measurement under study. After evaluating the effectiveness of parameters, investigation of the performance of proposed system from the perspective of different measurements and patients will be conducted. Then, we evaluate the impact of the difference and rotating frequency on different algorithms. After that, a comparative study is undertaken on the three algorithms to evaluate the performance of the proposed system. Finally, correlation coefficient analysis will be quantified to assess the relationships among the performance metrics (*accuracy* and *workload saving*) and dataset characteristics (*differences* and *rotating frequency*). Both accuracy and workload savings are evaluated in all the experiments.

7.1. Effectiveness of k

We first explored the effectiveness of k, the size of the sliding window used in time series data analysis. Figure 2 shows the averaged accuracy and saving percentages with time windows in different sizes for all algorithms. Generally speaking, for the datasets collected in the pilot study, the best combination of accuracy and working savings was achieved when k ranged from 3 to 5, the relatively small values, meaning that the risk prediction was more accurate when only the most recent days were considered.

7.2. Effectiveness of p

In this experiment, the proposed system was applied with different values of p on all patients when $k = 5$.

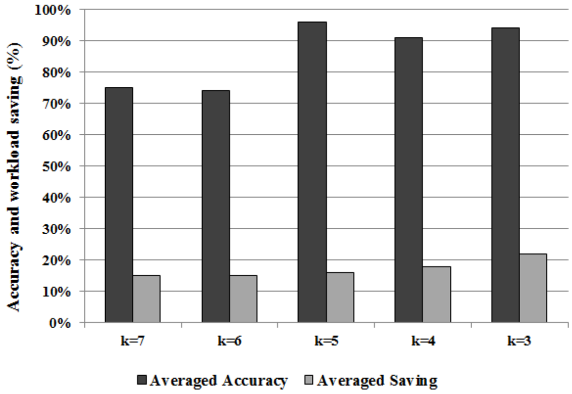

Fig. 2. The Averaged Accuracy and Workload Savings for all algorithms under Varying Values of k

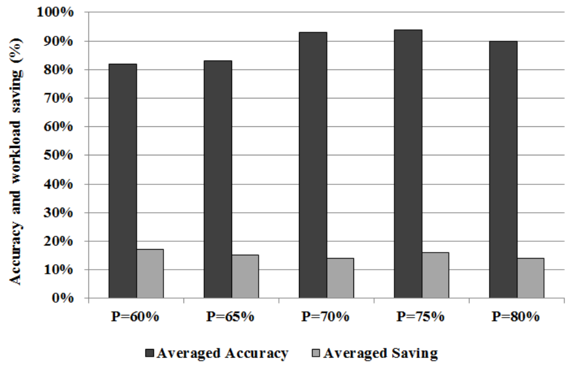

Fig. 3. The Averaged Accuracy and Workload Savings for all algorithms under Varying Values of p

The result is presented in Figure 3. When p takes relatively large values, in other words, when p ranged from 70% to 80%, the system achieves better recommendation accuracy. The reason is that, under a fixed value of k, a larger value of p led to a larger lower-bound for the percentage of the days in the sliding window when a medical test was required. A higher percentage of the days with medical test naturally provided more information to the system, and as a result, made the system capable of making more accurate recommendations. In addition, a higher value of p did not considerably affect the performance of workload saving despite that a better accuracy is achieved, on the basis of observation on the experimental results.

7.3. Performance of Three Algorithms under Different Measurements for all Patients

In this experiment, the performance of the system is evaluated when it is applied to different measure-

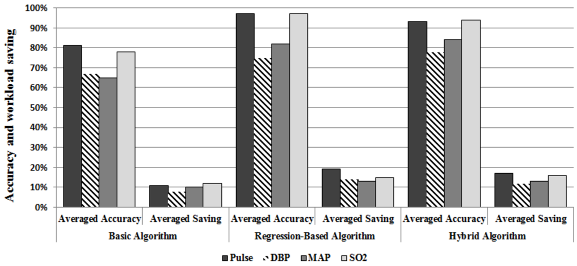

Fig. 4. The Accuracy and Workload Saving for Different Patients

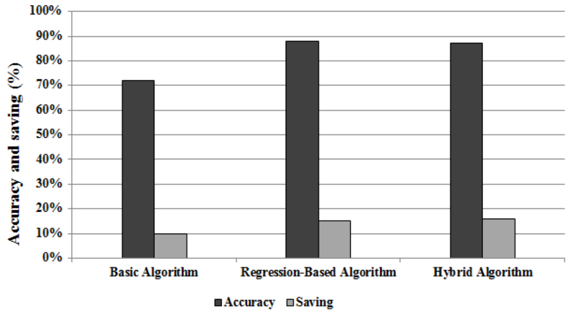

Fig. 5. Average Accuracy and Workload Saving for Each Algorithm

ments for all patients based on $k = 5$. Each algorithm was tested in four rounds for all patients with different medical tests (Heart rate, DBP, MAP and SO2). Figure 4 shows the detailed results we obtained for each algorithm. From the results, we can see that the algorithms yield recommendations with varying degree of accuracy, with the heart rate and SO2 measurements register the highest compared to others. This is because that there are intrinsic stronger correlations for the neighboring readings of heart rate and SO2 measurements so that the short-term prediction of risk becomes more accurate than other measurements.

By further aggregating the results for the measurements in Figure 4, Figure 5 demonstrates the averaged accuracy and workload savings for each algorithm. From the obtained results, it clearly seems that Regression-Based and Hybrid algorithms yield the highest accuracy and saving compared with Basic algorithm. Therefore, the risk prediction of disease provided by the proposed system is more accurate with these two algorithms. The system is capable of helping reduce on average 15% of workload for patients from their daily medical tests.

7.4. Difference and Rotating Frequency Analysis

We also evaluate the relationship between the data characteristics measured by Difference and Rotating Frequency on the performance of the proposed system. Figure 6 shows the performance of our system and data characteristic measurements (Difference and Rotating Frequency) of four measurements. As showed in Figure 6, there is a negative relationship between Accuracy and Rotating Frequency under each measurement for all algorithms. Based on the obtained results, we observe that the recommendations provided by the proposed system are more accurate for a measurement when Rotating Frequency is degraded for this measurement and vice visa. In addition, the value of Difference is positively correlated with that of accuracy, suggesting that the proposed system can produce more accurate recommendations when Difference values are high.

7.5. Algorithms Performance Evaluation

In this experiment, we are tested the three algorithms to evaluate their performance for each patient in our dataset. From Figure 7, we can clearly see that there are relative differences in values of accuracys and savings for each algorithm. General speaking, Regression-Based and Hybrid algorithms yielded a high degree of workload saving compared with the basic algorithm. This means that these two algorithms can help the patients in reducing the workload on more than 15% from their daily medical tests. The recommendations accuracy of proposed system are ranged between 75% and 98% across different patients under the three algorithms.

7.6. Correlation Coefficient

7.6.1. The Relationship among Different Measurements

In this experiment, we quantify the values of Pearson's Correlation Coefficient for all the pairs of numerical measurements in the dataset. The purpose of this experiment is to investigate the correlation relationships between different numerical measurements. Since we have five numerical measurements in the datasets, thus we evaluate the correlation coefficient for a total of 10 different pairs. Table 3 presents the values of correlation coefficients of different pairs of measurements for each of the patients and the average value across all the patients. The results show that there are five pairs of measurements (i.e., {HR, SBP}, {HR, DBP}, {DBP, SBP}, {DBP, MAP} and {SBP, MAP}) exhibiting relatively strong positive correlations, ranging from 0.55 to 0.63, where other pairs of

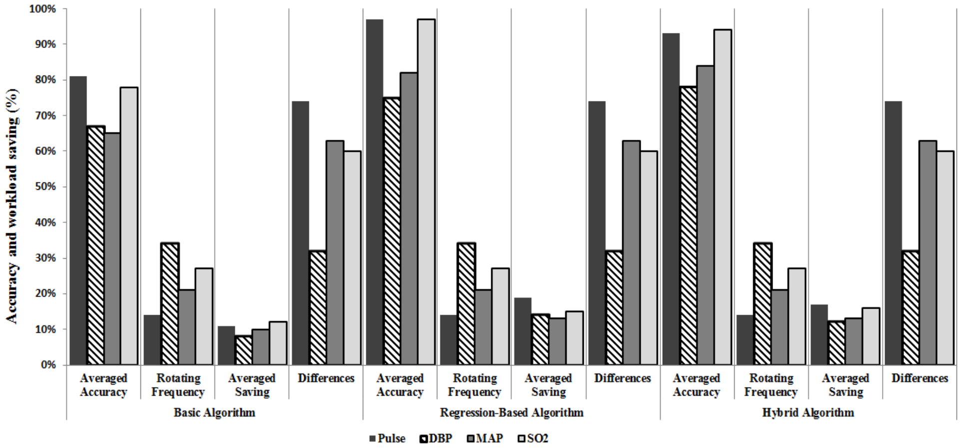

Fig. 6. Rotating Frequency Analysis for each Algorithm under four Measurements

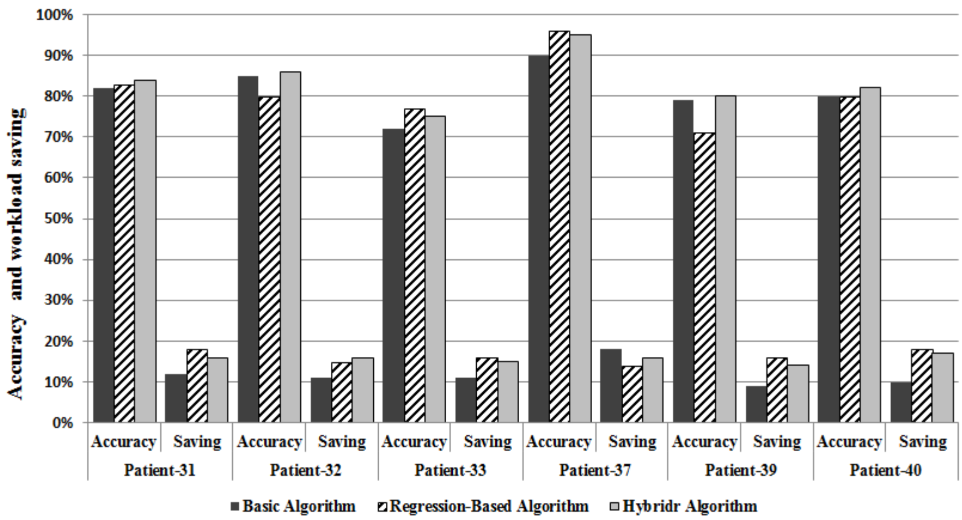

Fig. 7. Comparison the Performance of the Three Algorithms for each Patient

measurements does not show any strong correlations (HR is Heart rate; SBP is Systolic Blood Pressure; DBP is Diastolic Blood Pressure; MAP is Mean Arterial Pressure; SO2 is Oxygen Saturation). This experiment helps unveil the relationship between different measurements and pave the foundation for us to improve our current recommendation system so that risk prediction of a measurement can be performed on not only its own historical data but also those of another measurement that is correlated with it.

7.6.2. The Relationship among Performance Metrics and Dataset Characteristics

In this experiment, we quantify the values of Pearson's Correlation Coefficient for the performance metrics (*Accuracy and Workload Saving*) and dataset characteristics metrics (*Difference and Rotating Frequency*). The purpose of this experiment is to investigate the correlation relationships between different performance metrics and dataset characteristics metrics. Table 4 presents the values of correlation coefficients for different pairs of performance metrics and dataset characteristics metrics for each of the patients and the average value across all the patients. From the obtained results, the interesting observation in Table 4 is a relatively strong positive relationship among the Difference, Accuracy and Workload Saving for all patients, whereas the strong negative correlations clearly appeared with other pairs. From this experiment, we conclude that the system yields a higher Accuracy and Workload Saving for each of the patients for those

Table 3

Correlation Coefficient for 5 Measurements (HR is Heart rate; SBP is Systolic Blood Pressure; DBP is Diastolic Blood Pressure; MAP is Mean Arterial Pressure; SO2 is Oxygen Saturation)

	PID-31	PID-32	PID-33	PID-37	PID-39	PID-40	Average
(HR, SBP)	0.7	0.5	0.4	0.5	0.6	0.8	0.58
(HR, DBP)	0.5	0.4	0.6	0.5	0.7	0.6	0.55
(HR, SO2)	-0.5	-0.6	-0.3	0	-0.4	-0.5	-0.38
(HR, MAP)	0.01	0.02	0.15	0.01	0.07	0.09	0.058
(DBP, SBP)	0.8	0.6	0.5	0.7	0.5	0.7	0.63
(DBP, MAP)	0.7	0.5	0.7	0.4	0.6	0.5	0.57
(DBP, SO2)	0.01	-0.01	-0.01	0.01	0.09	-0.05	0.007
(SBP, MAP)	0.7	0.6	0.5	0.4	0.5	0.7	0.567
(MAP, SO2)	0.03	0.01	0.06	0.03	0	0.04	0.028
(SBP, SO2)	0.02	0.11	-0.1	0.13	-0.4	-0.2	-0.073

Table 4

Correlation Coefficient for Performance Metrics and Dataset Characteristics

	PID-31	PID-32	PID-33	PID-37	PID-39	PID-40	Average
(Differences, Accuracy)	0.8	0.6	0.9	0.7	0.4	0.9	**0.7**
(Differences, Saving)	0.7	0.4	0.6	0.08	0.9	0.9	**0.6**
(Accuracy, Saving)	0.7	0.8	0.6	0.4	0.4	0.9	**0.6**
(Differences,Rotating)	-0.8	-0.9	-0.9	-0.9	-0.8	-0.6	**-0.8**
(Rotating, Accuracy)	-0.7	-0.8	-0.7	-0.9	-0.8	-0.4	**-0.7**
(Rotating, Saving)	-0.8	-0.7	-0.6	-0.17	-0.8	-0.7	**-0.6**

medical measurements with a higher value of Difference and a lower value of Rotating Frequency and vice visa.

8. Conclusions and Future Work

This work has illustrated that it is possible to develop an intelligent recommender system that predicts and assesses the short-term disease risk for patients suffering from chronical diseases such as heart disease. This work is considered one of the first studies to use medical measurements of patients in the prediction of the short-term disease risk in the telehealth environment. The system is developed aiming at improving the quality of clinical evidence-based decisions and helping reduce financial and timing cost taken by patients. The system is based on a time series prediction algorithm which predicts short-term disease risk for patients and facilitates the generation of recommendations whether a particular patient needs to take the physical body test on the current date or not using the telehealth facility. In addition, this system is believed to contribute to a better monitoring, assessment and management the disease risk of for patients suffering chronical diseases such as heart disease. The work makes theoretical contribution by the time series prediction algorithm and applicable contribution by an intelligent system to improve the quality of health care services.

In the future, we are interested in further improving the predictive efficacy of our system. We will develop and evaluate more predictive models such as SVM, KNN classifier, decision tree, etc in this recommender system. Cross-measurement prediction, which is the prediction of the health risk of one medical measurement by making use of one or more other measurements, will be also studied based on the relationships that have been unveiled among different medical measurements in this work. Finally, further experimental evaluations will be also conducted by using additional large data sets for experiment of evaluation and validation.

Acknowledgement

The authors would like to thank the support from National Science Foundation of China through the research projects (No. 61370050, No. 61572036 and No. 61672039).

References

[1] K. Ahmed, T. Jesmin, and M. Z. Rahman. Early Prevention and Detection of Skin Cancer Risk using Data Mining. In *International Journal of Computer Applications*, (0975ÂąÂğC8887) Volume, 2013.

[2] A. Barsheshet, A. J. Moss, D. T. Huang, S. McNitt, W. Zareba, and I. Goldenberg Applicability of a risk score for prediction of the long-term (8-year) benefit of the implantable cardioverter-defibrillator. In *Journal of the American College of Cardiology*, vol. 59, no. 23, pp. 2075-2079, 2012.

[3] C.-D. Chang, C.-C. Wang, and B. C. Jiang. Using data mining techniques for multi-diseases prediction modeling of hypertension and hyperlipidemia by common risk factors. In *Expert systems with applications*, vol. 38, no. 5, pp. 5507-5513, 2011.

[4] M. Farhadian, P. J. Lisboa, A. Moghimbeigi, J. Poorolajal, and H. Mahjub. Supervised Wavelet Method to Predict Patient Survival from Gene Expression Data. In *The Scientific World Journal*, vol. 2014, 2014.

[5] N.-C. Hsieh, L.-P. Hung, C.-C. Shih, H.-C. Keh, and C.-H. Chan. Intelligent postoperative morbidity prediction of heart disease using artificial intelligence techniques. In *Journal of medical systems*, vol. 36, no. 3, pp. 1809-1820, 2012.

[6] F. Huang, S. Wang, and C.-C. Chan. Predicting disease by using data mining based on healthcare information system. pp. 191-194, 2012.

[7] P. A. James, S. Oparil, B. L. Carter, W. C. Cushman, C. Dennison-Himmelfarb, J. Handler, D. T. Lackland, M. L. LeFevre, T. D. MacKenzie, and O. Ogedegbe. 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the Eighth Joint National Committee (JNC 8) In *Jama*, vol. 311, no. 5, pp. 507-520, 2014.

[8] T. W. Joo, and S. B. Kim. Time series forecasting based on wavelet filtering. In *Expert Systems with Applications*, 2015.

[9] M. Kantardzic. In *Data mining: concepts, models, methods, and algorithms*, John Wiley & Sons, 2011.

[10] J.-K. Kim, J.-S. Lee, D.-K. Park, Y.-S. Lim, Y.-H. Lee, and E.-Y. Jung. Adaptive mining prediction model for content recommendation to coronary heart disease patients. In *Cluster Computing*, vol. 17, no. 3, pp. 881-891, 2014.

[11] H. C. Koh, and G. Tan. Data mining applications in healthcare In *Journal of healthcare information management*, vol. 19, no. 2, pp. 65, 2011.

[12] V. Krishnaiah, D. G. Narsimha, and D. N. S. Chandra Diagnosis of lung cancer prediction system using data mining classification techniques. In *International Journal of Computer Science and Information Technologies*, vol. 4, no. 1, pp. 39-45, 2013.

[13] M. Kurosaki, N. Hiramatsu, M. Sakamoto, Y. Suzuki, M. Iwasaki, A. Tamori, K. Matsuura, S. Kakinuma, F. Sugauchi, and N. Sakamoto. Data mining model using simple and readily available factors could identify patients at high risk for hepatocellular carcinoma in chronic hepatitis C In *Journal of hepatology*, vol. 56, no. 3, pp. 602-608, 2012.

[14] M. H. Lee, H. I. Yang, J. Liu, R. Batrla-Utermann, C. L. Jen, U. H. Iloeje, S. N. Lu, S. L. You, L. Y. Wang, and C. J. Chen. Prediction models of long-term Cirrhosis and hepatocellular carcinoma risk in chronic hepatitis B patients: Risk scores integrating host and virus profiles. In *Hepatology*, vol. 58, no. 2, pp. 546-554, 2013.

[15] D. M. Lloyd-Jones, E. P. Leip, M. G. Larson, R. B. d'Agostino, A. Beiser, P. W. Wilson, P. A. Wolf, and D. Levy Prediction of lifetime risk for cardiovascular disease by risk factor burden at 50 years of age. IN *Circulation*, vol. 113, no. 6, pp. 791-798, 2006.

[16] H.-Y. Lu, C.-Y. Huang, C.-T. Su, and C.-C. Lin. Predicting Rotator Cuff Tears Using Data Mining and Bayesian Likelihood Ratios. In *PloS one*, vol. 9, no. 4, pp. e94917, 2014.

[17] J. F. Lu, M. J. Chi, and C. M. Chen. Advocacy of home telehealth care among consumers with chronic conditions. In *Journal of clinical nursing*, vol. 23, no. 5-6, pp. 811-819, 2014.

[18] S. I. Nihtyanova, B. E. Schreiber, V. H. Ong, D. Rosenberg, P. Moinzadeh, J. G. Coghlan, A. U. Wells, and C. P. Denton. Prediction of Pulmonary Complications and Long-Term Survival in Systemic Sclerosis. In *Arthritis & rheumatology*, vol. 66, no. 6, pp. 1625-1635, 2014.

[19] L. Nobel, N. E. Mayo, J. Hanley, L. Nadeau, and S. S. Daskalopoulou. MyRisk_Stroke Calculator: A Personalized Stroke Risk Assessment Tool for the General Population. In *Journal of Clinical Neurology*, vol. 10, no. 1, pp. 1-9, 2014.

[20] B. A. Payne, J. A. Hutcheon, J. M. Ansermino, D. R. Hall, Z. A. Bhutta, S. Z. Bhutta, C. Biryabarema, W. A. Grobman, H. Groen, and F. Haniff. A risk prediction model for the assessment and triage of women with hypertensive disorders of pregnancy in low-resourced settings: the miniPIERS (Pre-eclampsia Integrated Estimate of RiSk) multi-country prospective cohort study. In *PLoS medicine*, vol. 11, no. 1, pp. e1001589, 2014.

[21] S. Roessler, E. L. Long, A. Budhu, Y. Chen, X. Zhao, J. Ji, R. Walker, H. L. Jia, Q. H. Ye, and L. X. Qin. Integrative genomic identification of genes on 8p associated with hepatocellular carcinoma progression and patient survival. In *Gastroenterology*, vol. 142, no. 4, pp. 957-966. e12, 2012.

[22] M. Sabibullah, V. Shanmugasundaram and R. Priya. Diabetes patient's risk through soft computing model. In *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, vol. 2, no. 6, pp. 60-65, 2013.

[23] F. Siraj, and M. A. Abdoulha. Mining enrolment data using predictive and descriptive approaches. In *Knowledge-Oriented Applications in Data Mining*, pp. 53-72, 2007.

[24] S. Tuffery. *Data mining and statistics for decision making*. John Wiley & Sons, 2011.

[25] D.-Y. Yeh, C.-H. Cheng, and Y.-W. Chen. A predictive model for cerebrovascular disease using data mining. In *Expert Systems with Applications*, vol. 38, no. 7, pp. 8970-8977, 2011.

[26] D. Kuh, and Y. B. Shlomo A life course approach to chronic disease epidemiology. In *Oxford University Press*, 2004.

[27] J. Myers, C. R. de Souza, A. Borghi-Silva, M. Guazzi, P. Chase, D. Bensimhon, M. A. Peberdy, E. Ashley, E. West, and L. P. Cahalin. A neural network approach to predicting outcomes in heart failure using cardiopulmonary exercise testing. In *International journal of cardiology*, vol. 171, no. 2, pp. 265-269, 2014.

[28] M. S. Mohktar, S. J. Redmond, N. C. Antoniades, P. D. Rochford, J. J. Pretto, J. Basilakis, N. H. Lovell, and C. F. McDonald. Predicting the risk of exacerbation in patients with chronic obstructive pulmonary disease using home telehealth measurement data. In *Artificial intelligence in medicine*, vol. 63,

no. 1, pp. 51-59, 2015.

[29] A. L. Hawkes, T. A. Patrao, J. Atherton, R. S. Ware, C. B. Taylor, A. OÂąÂŕNeil, R. Foreman, and B. F. Oldenburg. Effect of a Telephone-Delivered Coronary Heart Disease Secondary Prevention Program (ProActive Heart) on quality of life and health behaviours: primary outcomes of a randomised controlled trial. In *International journal of behavioral medicine*, vol. 20, no. 3, pp. 413-424, 2013.

[30] H. C. Noel, D. C. Vogel, J. J. Erdos, D. Cornwall, and F. Levin. Home telehealth reduces healthcare costs. In *Telemedicine Journal and E-Health*, vol. 10, no. 2, pp. 170-183, 2004.

[31] M. Parsaeian, K. Mohammad, M. Mahmoudi, and H. Zeraati. Comparison of logistic regression and artificial neural network in low back pain prediction: second national health survey. In *Iranian journal of public health*, vol. 41, no. 6, pp. 86, 2012.

[32] Z. Zhang, and C. Zhang. Building agent-based hybrid intelligent systems: A case study. In *Web Intelligence and Agent Systems: An international journal*, vol. 5, no. 3, pp. 255-271, 2007.

[33] J. Ji, C. Liu, J. Yan, and N. Zhong. An improved Bayesian network structure learning algorithm and its application in an intelligent B2C portal. In *Web Intelligence and Agent Systems: An International Journal*, vol. 5, no. 2, pp. 127-138, 2007.