Mining Heterogeneous Information Graph for Health Status Classification

Thuan Pham†, Xiaohui Tao†, Ji Zhang†, Jianming Yong†, Wenping Zhang‡, and Yi Cai♯
†*University of Southern Queensland, Australia*
Email: {Thuan.Pham,Xiaohui.Tao,Ji.Zhang,Jianming.Yong}@usq.edu.au
†*School of Information, Renmin University of China, China*
Email: wpzhang@ruc.edu.cn
♯*South China University of Technology, China*
Email: ycai@scut.edu.cn

***Abstract*—In the medical domain, there exists a large volume of data from multiple sources such as electronic health records, general health examination results and surveys. The data contain useful information reflecting people's health and provides great opportunities for studies to improve the quality of healthcare. However, how to mine these data effectively and efficiently still remains a critical challenge. In this paper, we propose an innovative classification model for knowledge discovery from patients' personal health repositories. By based on analytics of massive data in National Health and Nutrition Examination Survey, the study builds a classification model to classify patients' health status and reveal the specific disease potentially suffered by the patient. This paper makes significant contributions to the advancement of knowledge in data mining with an innovative classification model specifically crafted for domain-based data. Moreover, this research contributes to the healthcare community by providing a deep understanding of people's health with accessibility to the patterns in various observations.**

***Index Terms*—Heterogeneous information graph, Classification, Healthcare**

I. Introduction

Improving personalised healthcare has become an important responsibility for many governments. For example, $414.3 billion had been spent on healthcare in 2011 by the United States government [1]. In 2015, the United States government tried to enhance precision medicine by utilising increasingly large amounts of available health data [2]. However, there are many challenges to develop effective models in healthcare applications, which are further complicated by real data.

Data mining has been used widely in many different fields and applications. In healthcare, data mining is used to check treatment effectiveness, health customer relationship management and patient care and to counter fraud and abuse [3]. In the last decades, many research efforts have been invested in the assessment of the risk of diseases in order to support medical practitioners to make more secure and effective decisions. Based on practitioners' experience, clinical decisions are made from medical databases. There is strong evidence that clinical decisions on the basis of risk assessment may improve disease management [4] [5]. To apply data mining to medical diagnosis in healthcare, some researchers have developed predictive models for the classification of clinical risks [6] [7] and for predicting diseases [8] [9]. Recently, Ling (2016) proposed an algorithm called a semi-supervised heterogeneous graph on health(SHG-Health) to predict the risk of patients based on the health examination data. Their work makes a significant result in discovering the different type of the relationship in the heterogeneous graph for health risk prediction by classification.

In this paper, we use a heterogeneous graph to present for the health examination records, which have been given more chance to discover the knowledge in the Medical domain. To build up a heterogeneous graph, we take each of the attributes of the dataset as a node of the graph. Links are set up by using the *Pearson Correlation* coefficient. If the weight of two variables (two nodes) have a coefficient with greater 0.3, A link between two nodes will be set up. And then we build a classification model to predict the high risk of patients. The model has been considered the semantic similar for predicting the health risk. The result has been demonstrated that our model is up to 38,26% more accuracy than the Ling Chen work which is not considered to the semantic similarly for developing the model.

The contributions of our work will be the following:

- A new method is introduced for building a heterogeneous information graph by using the *Pearson Correlation* coefficient to set up the link between two nodes.
- A new binary classification model is proposed, which considers using the semantic similarly for health risk prediction by discovering knowledge from health examination records.
- A solution is also provided as a classification model to assess the personal health status from the heterogeneous information graph.
- The study supplies an evaluation system based on machine learning techniques for healthcare to evaluate how effective the proposed health risk prediction model is.

The remainder of this paper is organised as follows. In Section 2, we review the existing work on measurement for a patient's health status, health risk prediction by classification, and mining a heterogeneous graph by classification. Then, we define the problem of this study, provide some concept about the heterogeneous graph, data correlation, data semantic and propose our model for predicting risk in Section 3. Following by Section 4, we deal with the experiment design, the baseline model and our result. Finally, Section 5 presents our conclusions of this paper.

II. Related Work

A. Health Status Measurement

Prognostication has become a very popular task for physicians. Hence, scoring systems have been used widely not only to minimise errors caused from fatigue but also to help physicians make better clinical decisions. For example, for myelodysplastic syndromes, Greenberg et al. [25] introduced a scoring system, called the international prognostic scoring system (IPSS). IPSS has been refined later on in order to get an improvement in analysing the specific impact of marrow blast percentage and depth of cytopenias. as another way of making clinical decisions, Le Gall et al. [26] introduced SAPS II scoring system, which played a role in quantifying the severity of illness in the Intensive care unit area. This scoring system came up with a method of converting the score to the possibility of a hospital's mortality. In accord with SAPS II, the APACHE II scoring system [27] remains the systematic application of clinical judgments about the relative importance of derangement in the physiologic measures. Besides introducing new scoring systems, studies have been conducted to compare the efficiency among these different scoring systems. For example, Keegan et al. [28] have discussed the performance of 4 scoring systems including APACHE III, APACHE IV, SAPS 3 and MPM III. The research showed APACHE III and APACHE IV were similar in discrimination capability and they both performed better compared with SAPS 3 and MPM III. Moreover, the research also showed that the complex models work better than the simple model, and the efficient performance level of these models depends on the number of variables.

B. Health Risk Prediction by Classification

Yeh et al. [10] conducted a study which mainly focused on applying the classification technology to build an optimum cerebrovascular disease predictive model. In this research, 3 attribute input modes, T1, T2, and T3 were built which aimed to build the classification models and compare their efficiencies. Besides that, in order to predict the future health of patients and identify patients at high risk, Neuvirth et al. [11] introduced new methods which apply state-of-the-art methods to explore including emergency care services' needs and the possibility of the treatment producing a sub-optimal result. There were two binary classification algorithms used in this research: logistic regression (LR), and k-nearest neighbour (KNN). Following this, with the essential importance of binary classification, Nguyen et al. [12] introduced a new machine learning approach that used the training phase that accompanies soft labels to refine the binary class information to get a more efficient binary classification model. Label uncertainty is also known as label noise. In order to solve this problem in binary classification, Yang et al. [13] proposed a method which is mainly focused on using the uncertainty information to improve the performance of retraining-based models. The result showed that the new method provided a more efficient demonstration and it was used to reduce human labelling efforts in different applications.

C. Mining Heterogeneous Graphs for Classification

By using the graph-based method, there are more advantaged for discovering the intrinsic characteristics of data, where the vertices and edges of a graph are take-ups as model data points and their relationships, respectively [14]. Particularly, the result of mining these graph will be more improvement if we consider presenting data in a heterogeneous graph. Due to more meaningful results could be generated from the different types among links and objects [15] [16]. In 2010, a classification method for heterogeneous networks is introduced by Ming et.al [17] called GNetMine. Their propose uses only one classification criteria for all of the objects in the network. However, Wang et.al [18] argue that the type differences of objects may have different criteria of classification. they proposed a new method to improve this drawback by providing the concept meta path for mining the heterogeneous graph. In healthcare data, based on the graph-based semi-supervised learning, Hwang et.al [19] introduced a heterogeneous label propagation algorithm to discovery disease gene. they base on homo-subnetwork which links are set up from the same types of objects to build up a heterogeneous disease-gene graph. Recently, Cheng et.al [20] introduced a semi-supervised heterogeneous graph on health algorithm(SHG-Health) to predict for high-risk disease class for the data unlabeled. By mining knowledge from a heterogeneous graph, the model has contributed to a significant improvement for classification problem in healthcare data.

Apparently, many studies have investigated binary classification problems and developed related data mining techniques. Previous researchers like [14] [15] [16] [17] [18] [19] [20] utilised the advantages of the graph-based to reach the point of suggesting the types of disease through traditional diagnosis. However, the advantage of graph-based techniques has not been applied in supporting decision making for health risk predictions by classification [10] [11] [12] [13]. In general, most of these studies used labelled data to develop the prediction algorithm. However, little attention has been paid to use the semantic similarly to develop the classification model in health risk prediction. That motivates us to conduct this study planned as such. Also, *Pearson Correlation* coefficient is related to how strong effect for the relationship between two variables. By combination of the two advantages for mining healthcare data, we aim to improve the understanding of semantics healthcare data and applies such understanding to an innovated classification model for health risk prediction.

III. The Proposed Method

A. Research Problem

The work is focused on resolving the classification problem in healthcare data. The research aims to help assess health risks by using data mining and machine learning techniques, to provide evidence-based decision-making support, instead of experience-based, to doctors, physicians and healthcare practitioners, and to help them reduce human errors. Human brains have limits; medical knowledge updates over time. Doctors and physicians may find it difficult to avoid human errors when they simply rely on their experience. Experience-based decisions may lead to the problem that some critical cases are overlooked. In contrast, data mining in healthcare can help cover the overlooked areas because it does not have the aforementioned limitation. Data mining allows researchers to work with data collected from a huge number of patients, a number that is more than any doctor ever treats. As a result, data mining can provide high-quality evidence covering as many possibilities as possible to support doctors' decision-making by knowledge discovery in healthcare data.

Data mining and machine learning techniques are adopted in this work to help deliver evidence-based decision support. Data analysis will help doctors predict a patient's health status, for example, being healthy or unhealthy, using a binary classification model. Furthermore, classification techniques are not only able tp help doctors categorise patients to the groups of such as "healthy" or "unhealthy" but also provide them suggestions on what kind of diseases they are suffering from. The work presented in this paper is an attempt to the problem utilising data mining and machine learning techniques.

B. Heterogeneous Information Graph

The National Health and Nutrition Examination Survey (NHANES)[1] dataset is used in this study to develop the classification model. NHANES is a program conducted in the United States for studies of the health and nutritional status of adults and children. The survey has thousands of questions to be addressed in interviews and physical examinations. As a result, the NHANES dataset consists of 2585 attributes, covering a wide range of information about an individual, such as personal demographics, observations, laboratory tests, or diagnostic reports. After careful examination, the 2585 attributes were manually categorised into eight classes based on their intuitive semantic relations; *Kidney conditions*, *Hepatitis*, *Diabetes*, *Blood Pressure and Cholesterol*, *Heart disease*, *Respiratory Disease*, *Profile* and *Others*. The eight semantic classes are presented in Table I with brief description.

Consider each attribute in the NHANES dataset a data object, the underlying links connecting objects are discovered adopting the *Pearson Correlation* coefficient, which is a powerful technique to measure the linear correlation between two variables. Denote a *Pearson Correlation* coefficient value by ρ, two objects v_1 and v_2 hold valid connection if $\rho(v_1, v_2) \geq \gamma$, where γ is a threshold defining the validity

[1] https://www.cdc.gov/nchs/nhanes/index.htm

TABLE I
SEMANTIC CATEGORIES

Class	Description
Kidney Conditions	This class will group all of attribute related to kidney disease, especially, most test of urine will be present in this class
Hepatitis	This will include all types of hepatitis such as A,B, ore C. Also some of question related to hepatitis condition. For example, "Have you ever received Hepatitis A vaccine"
Diabetes	Diabetes class will also related some attribute about lab test such as urine or blood.
Blood Pressure and Cholesterol	all of lab tests related to blood will be in this class
Heart disease	heart disease class will contain all of attribute related some of question such as "Has a doctor ever told you that you had a heart attack, coronary heart disease, or congestive heart failure?". Also some kid of symptom as angina, also called angina pectoris
Respiratory Disease	This class will sum of attribute for respiratory disease. For example, "asthma, emphysema, thyroid problem, chronic bronchitis"
Profile	the information such as age, weight, gender will be belong into this class
Others	This class will cover the rest which is unable to identify which class belong to, or some kind of category that we can not obtain the ground truth

of object connection. The range of coefficient value is set between -1 and 1. The relationship between two objects is weakly correlated if the coefficient value is approximately -1. In contrast, the relationship between two nodes is more correlated if the coefficient value is in the possible range nearly 1. Apparently, a larger ρ value reveals stronger connection of the two objects.

A heterogeneous information graph is then constructed with the semantic classes and the link connecting objects defined by *Pearson Correlation* coefficient.

Definition 1: [Heterogeneous Information Graph] A heterogeneous information graph is a three tuple, $G :=< V, E, M >$ with an object mapping function $\varphi : V \to A$ and a link type mapping function $\psi : E \to R$, where

- $M \times M$ is a matrix built based on all attributes in the NHANES dataset;
- Each object $v \in V$ belongs to one particular class in the semantic class set $A : \varphi(v) \in A$
- Each link $e \in E$ belongs to particular relation type set in the relation type set $R : \phi(e) \in R$, where $|R| = 1$ and e is the link connecting two objects defined by a *Pearson Correlation* coefficient ρ. □

Figure 1 illustrates a subgraph of the heterogeneous information graph constructed using the NHANES dataset. Three types of objects are illustrated, where A is a class for *Kidney condition*, B for *Profile* and C for *Heart disease*.

C. The Classification Model

With the availability of the heterogeneous information graph, a function can be learned from the training data that

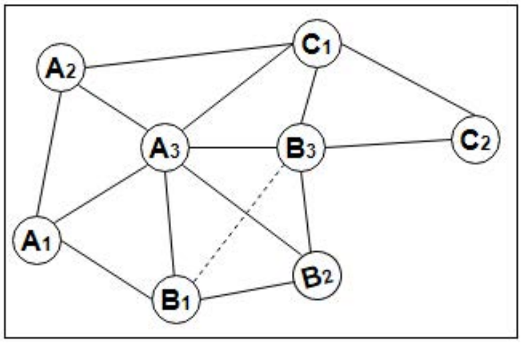

Fig. 1. A sample heterogeneous information graph

formalises the profile of healthy (unhealthy) status of a patient regarding a disease object x:

$$f(x) = \sum_{i=1}^{k} v_i \times \rho(x, v_i) \times \alpha + \sum_{j=1}^{k} v_j \times \rho(x, v_i) \times \beta \quad (1)$$

where $\varphi(x) = \varphi(v_i)$ and $\varphi(x) \neq \varphi(v_j)$. α and β are two different coefficients adopted to balance the impacts of objects that belong to the same and different semantic classes with x.

Based on $f(x)$, a patient's health status can be modelled against a single disease, x:

$$y(x) = \begin{cases} 1, & \text{if } f(x) \geq \theta \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where θ is a threshold to determine the final class, "healthy" or "unhealthy", for the patient. When checking against multiple diseases $x \in \mathcal{X}$, an overall model can then be defined on the basis of Eq. 2:

$$y(\mathcal{X}) = \Pi_{n=1}^{k} y(x_n), \text{where } x \in \mathcal{X}, |\mathcal{X}| = k \quad (3)$$

IV. Evaluation

A. Experimental Design

The survey data of NHANES was used to evaluate the proposed model. The dataset covers a wide range of health assessments, such as lab tests, physical examinations, personal habits, etc. The NHANES dataset contains 9770 participants with more than 2585 attributes. The dataset was manually assessed to generate the ground truth. Five diseases, represented by code DIQ010, KIQ022, KIQ026, HEQ010 and HEQ030, were selected in experiments. Only patients who were clear from all five diseases were considered healthy. As a result, 5144 out of 9770 participants were identified healthy and 4626 participants were unhealthy. Before applying the dataset to experiments, it was pre-processed first as it contained much noisy data. The values of all attributes were normalised into a unified form. Then, the missing data were handled and replaced by the average of all values in the respective attribute. Figure 2 presents the dataflow in experiment design.

The proposed model was evaluated by comparing with the baseline model, Chen and Ling [20]'s work, representing state-of-the-art related research. The baseline model uses semi-supervised learning algorithm to solve the classification problem with consideration of relationship existing in the data's neighbourhood. Alternatively, our proposed model considers the semantic relations between data attributes and relations underlying from data, which was adopted to constructing the heterogeneous information graph in experiments.

The experimental models' performance would be measured using standard metrics, such as accuracy, recall, precision and F-measure, which are all commonly used by the research community for similar problems [23], [24]. We also adopted k-fold ($k = 5$) validation approach to help assure the reliability of evaluation. The NHANES dataset was randomly separated into five sub-sets. In each experimental run, one of the five sub-sets would be used for training and the other four for testing. The average performance of all five runs over the tests of five diseases then counted as the final performance of the experimental model.

TABLE II
PRECISION RESULT OF FIVE DISEASES

Disease	**Precision**		
	Our model	**BL model**	**Improvement**
DIQ010	**0.903579308**	0.85354816	5.86%
KIQ022	**0.986242126**	0.498513842	97.84%
KIQ026	**1**	0.76075827	31.45%
HEQ010	**0.919895076**	0.231615337	297.17%
HEQ030	**0.808337434**	0.271221332	198.04%
AVG.	**0.923610789**	0.523131388	126.07%

TABLE III
RECALL RESULT OF FIVE DISEASES

Disease	**Recall**		
	Our model	**BL model**	**Improvement**
DIQ010	0.887451802	**1**	-11.25%
KIQ022	0.901529388	**1**	-9.85%
KIQ026	0.90909094	**1**	-9.09%
HEQ010	0.90909094	**1**	-9.09%
HEQ030	0.78088432	**1**	-21.91%
AVG.	0.877609478	**1**	-12.24%

TABLE IV
F-MEASURE RESULT OF FIVE DISEASES

Disease	**F-Measure**		
	Our model	**BL model**	**Improvement**
DIQ010	0.89217736	**0.92092135**	-3.12%
KIQ022	**0.941830372**	0.665199808	41.59%
KIQ026	**0.95238096**	0.86404114	10.22%
HEQ010	**0.91044549**	0.366653288	148.31%
HEQ030	**0.7859063**	0.425336024	84.77%
AVG.	**0.896548096**	0.648430322	56.35%

B. Result and Analysis

We have five diseases being tested in five rounds, as afore-discussed. Tables II to V present the experimental results in precision, recall, and accuracy, respectively, where "BL" refers to the baseline model. Although the percentage of recall and

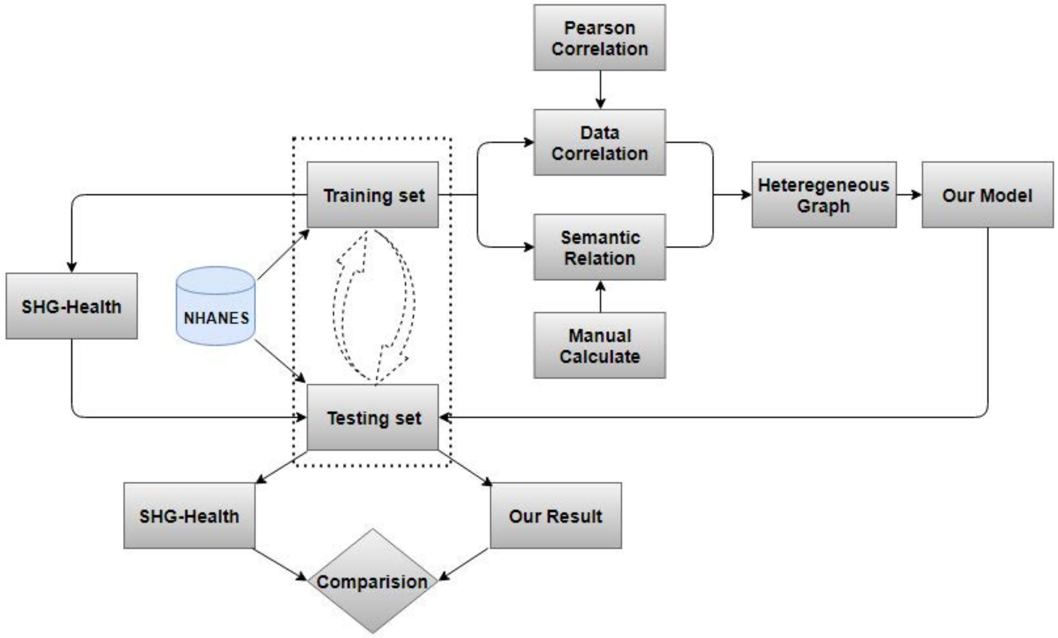

Fig. 2. Experiment Data-flow

accuracy did not bring in more improvement for our model, the percentage of precision which shows the accuracy of classification for the possible class, has a significant improvement, as shown in Table II. To evaluate the model, F-measure is adopted and the related results are presented in Table IV. In overall, our model has been improved 38.26% compared to baseline model, as shown in Table VI.

In Table IV, we can see in the first disease, the F-measure performance of our model is lower than the baseline model. However, for the rest of all diseases our model is better than the baseline model. Especially, in one of the diseases our model has a strong improvement compared to the baseline model with 148.31% of improvement. This is because the training samples are different and data are sparse. Non-balanced data may be the main reason for this. we suppose that all models work better if the data is clear. However, when data is sparse with a lot of noise, the performance drops down quickly. Also, in the baseline model, links between nodes are set up based on the sequence of time. In this experiment, we only use data for one year of the participant in training the model. The results of our experiment show that if we only use the records of the patient in one year, the baseline model will have limited capacity.

TABLE V
ACCURACY RESULT OF FIVE DISEASES

Disease	Accuracy		
	Our model	**BL model**	**Improvement**
DIQ010	0.983077512	**0.986845888**	-0.38%
KIQ022	**0.996416836**	0.967378568	3.00%
KIQ026	**0.991553382**	0.970836194	2.13%
HEQ010	**0.99856984**	0.97543062	2.37%
HEQ030	**0.996178244**	0.97630493	2.04%
AVG.	**0.993159163**	0.97535924	1.83%

TABLE VI
OVERALL COMPARISON OF OUR MODEL TO THE BASELINE MODEL

	Precision	**Recall**	**Accuracy**	**F-Measure**
Our model	**0.9236107**	0.8776094	**0.9931591**	**0.8965480**
Baseline model	0.5231313	**1**	0.9753592	0.6484303
Improvement	126.07%	-12.24%	1.83%	56.35%

V. CONCLUSIONS

By mining a heterogeneous information graph constructed using health examination data, we developed a novel classification model. The proposed model introduces a new method for building the heterogeneous information graph by using the *Pearson Correlation* coefficient to set up the link between a pair of nodes. Also, the model adopted the semantic similarly to discover the relations among categories. As a result, our model has been improved the precision of information retrieval by more than 38% compared to the base line model. The research delivers significant contributions to knowledge advancement in data mining and helps improve the design of healthcare systems.

REFERENCES

[1] L. B. Mirel and K. Carper, Trends in Health Care Expenditures for the Elderly, Age 65 and Older: 2001, 2006, and 2011, 2014.
[2] F. S. Collins and H. Varmus, A new initiative on precision medicine, New England Journal of Medicine, vol. 372, no. 9, pp. 793−795, 2015.

[3] H. C. Koh, G. Tan, and others, Data mining applications in healthcare, Journal of healthcare information management, vol. 19, no. 2, p. 65, 2011.

[4] P. A. James et al., 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), Jama, vol. 311, no. 5, pp. 507−520, 2014.

[5] M. M. Hunink et al., Decision making in health and medicine: integrating evidence and values. Cambridge University Press, 2014.

[6] 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, Cluster computing, vol. 17, no. 3, pp. 881−891, 2014.

[7] M. Sabibullah, V. Shanmugasundaram, and R. Priya, Diabetes patients risk through soft computing model, International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), vol. 2, no. 6, pp. 60−65, 2013.

[8] 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, Expert systems with applications, vol. 38, no. 5, pp. 5507−5513, 2011.

[9] F. Huang, S. Wang, and C.-C. Chan, Predicting disease by using data mining based on healthcare information system, in 2012 IEEE International Conference on Granular Computing, 2012, pp. 191−194.

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

[11] H. Neuvirth et al., Toward personalized care management of patients at risk: the diabetes case study, in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, 2011, pp. 395−403.

[12] Q. Nguyen, H. Valizadegan, and M. Hauskrecht, Learning classification models with soft-label information, Journal of the American Medical Informatics Association, vol. 21, no. 3, pp. 501−508, 2014.

[13] Y. Yang and M. Loog, Active learning using uncertainty information, in Pattern Recognition (ICPR), 2016 23rd International Conference on, 2016, pp. 2646−2651.

[14] A. Guillory and J. A. Bilmes, Label selection on graphs, in Advances in Neural Information Processing Systems, 2009, pp. 691−699.

[15] B. Long, Z. M. Zhang, X. Wu, and P. S. Yu, Spectral clustering for multi-type relational data, in Proceedings of the 23rd international conference on Machine learning, 2006, pp. 585−592.

[16] Y. Sun, Y. Yu, and J. Han, Ranking-based clustering of heterogeneous information networks with star network schema, in Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, 2009, pp. 797−806.

[17] M. Ji, Y. Sun, M. Danilevsky, J. Han, and J. Gao, Graph regularized transductive classification on heterogeneous information networks, in Joint European Conference on Machine Learning and Knowledge Discovery in Databases, 2010, pp. 570−586.

[18] C. Luo, R. Guan, Z. Wang, and C. Lin, Hetpathmine: A novel transductive classification algorithm on heterogeneous information networks, in European Conference on Information Retrieval, 2014, pp. 210−221.

[19] T. Hwang and R. Kuang, A heterogeneous label propagation algorithm for disease gene discovery, in Proceedings of the 2010 SIAM International Conference on Data Mining, 2010, pp. 583−594.

[20] L. Chen et al., Mining Health Examination RecordsA Graph-Based Approach, IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 9, pp. 2423−2437, 2016.

[21] M. Ji, J. Han, and M. Danilevsky, Ranking-based classification of heterogeneous information networks, in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, 2011, pp. 1298−1306.

[22] Y. Sun, Y. Yu, and J. Han, Ranking-based clustering of heterogeneous information networks with star network schema, in Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, 2009, pp. 797−806.

[23] D. Bowes, T. Hall, and D. Gray, Comparing the performance of fault prediction models which report multiple performance measures: recomputing the confusion matrix, in Proceedings of the 8th International Conference on Predictive Models in Software Engineering, 2012, pp. 109−118.

[24] G. P. Visa and P. Salembier, Precision-recall-classification evaluation framework: Application to depth estimation on single images, in European Conference on Computer Vision, 2014, pp. 648−662.

[25] P. L. Greenberg et al., Revised international prognostic scoring system (IPSS-R) for myelodysplastic syndromes, Blood, p. blood2012, 2012.

[26] P. Prakash, K. Krishna, and D. Bhatia, Usefulness of SAPS II scoring system as an early predictor of outcome in ICU patients, J Indian Acad Clin Med, vol. 7, no. 3, pp. 202−5, 2006.

[27] D. P. Wagner and E. A. Draper, Acute physiology and chronic health evaluation (APACHE II) and Medicare reimbursement, Health care financing review, vol. 1984, no. Suppl, p. 91, 1984.

[28] M. T. Keegan, O. Gajic, and B. Afessa, Comparison of APACHE III, APACHE IV, SAPS 3, and MPM0III and influence of resuscitation status on model performance, Chest, vol. 142, no. 4, pp. 851−858, 2012.