



Prediction of Hemodialysis Timing Based on LVW Feature Selection and Ensemble Learning

Chang-zhu Xiong¹ · Minglian Su² · Zitao Jiang¹ · Wei Jiang¹

Received: 4 November 2018 / Accepted: 3 December 2018 / Published online: 13 December 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

We propose an improved model based on LVW embedded model feature extractor and ensemble learning for improving prediction accuracy of hemodialysis timing in this paper. Due to this drawback caused by feature extraction models, we adopt an enhanced LVW embedded model to search the feature subset by stochastic strategy, which can find the best feature combination that are most beneficial to learner performance. In the model application, we present an improved integrated learners for model fusion to reduce errors caused by overfitting problem of the single classifier. We run several state-of-the-art Q&A methods as contrastive experiments. The experimental results show that the ensemble learning model based on LVW has better generalization ability (97.04%) and lower standard error (± 0.04). We adopt the model to make high-precision predictions of hemodialysis timing, and the experimental results have shown that our framework significantly outperforms several strong baselines. Our model provides strong clinical decision support for physician diagnosis and has important clinical implications.

Keywords Hemodialysis timing · LVW · Ensemble learning · Prediction · Feature selection · Model fusion

Introduction

Chronic kidney disease is widespread throughout the world. It is a progressive disease with an early incidence of 100 times as late as the late stage. Premature dialysis is not conducive to the preservation and recovery of residual renal function. Therefore, it is necessary to dig out the key feature that guide patients to receive dialysis treatment through data mining techniques, and help doctors to choose the appropriate dialysis timing. In this case, we propose a model fusion scheme based on LVW (Las Vegas Wrapper) embedded model feature extractor to improve hemodialysis prediction accuracy. By selecting key features related to hemodialysis, the physician can decide whether the patient needs hemodialysis or not, to

choose the timing of hemodialysis for the patient and to improve the patient's survival rate. The overall framework of this paper is shown in Fig. 1.

As shown in Fig. 1, the inconsistent and incorrect data is first cleaned in the data preprocessing process; then the key features of the processed data are extracted using the LVW method; the prediction model is created using the ensemble learning method proposed in this paper. Finally, based on the trained model, we can predict whether the patient sample needs dialysis.

During data preprocessing, inconsistent and incorrect data values are cleaned, and data conversion and normalization processes are performed. Data cleaning includes imputation of missing values, reduction of data noise, and elimination of consistency; data conversion includes standardized data, discretized data, and construction of new attributes. The purpose of our dialysis prediction is to assist doctors in diagnosis and treatment and to participate in accurate clinical decision-making. We use data analysis to establish a model that can accurately predict whether a patient needs dialysis treatment. In this way, the burden on the doctor can be reduced and the operating efficiency of the hospital can also be improved.

The current study has confirmed a significant difference between pre-hemodialysis and post-hemodialysis mortality [1], and the timing of dialysis has implications for other

This article is part of the Topical Collection on *Systems-Level Quality Improvement*

✉ Chang-zhu Xiong
gongfuxiong93@gmail.com

¹ Department of electronic information, Sichuan University, Chengdu, China

² West China School of clinical medicine, Sichuan University, Chengdu, China

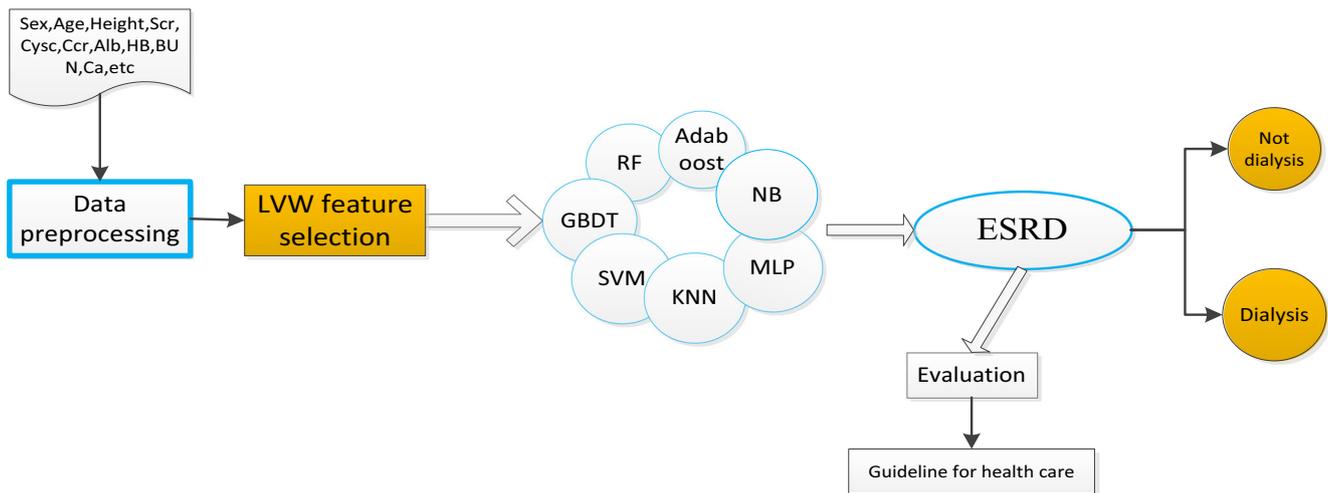


Fig. 1 Model framework of hemodialysis prediction

patient complications [2], so the study of hemodialysis has important practical needs and significance [3]. There have been many studies on the timing of hemodialysis, for example, the diagnosis and treatment of senile renal disease and dialysis; the timing of hemodialysis [4] and so on. Moreover, machine learning techniques have been used extensively in medical assistance, for example, the use of logistic regression to analyze the association between obesity index and chronic kidney disease [5], and the assessment of renal disease [6]; based on support vector machine in the diagnosis of chronic kidney disease [7] and diagnosis of breast cancer; the study of the impact of different feature selection on kidney disease [8]; prediction of breast cancer survival using big data [9]; chronic disease diagnosis based on data mining; predicting end-stage renal disease based on artificial neural networks; identifying diseases using Bayesian network and ROC curve analysis in medical diagnostic evaluation and so on. There are also many researches on feature selection and model application. For example, impulse noise recognition based on artificial neural network classifier, feature

dimensionality reduction [10, 11] based on PCA; combination based on ANFIS and neural network methods to predict electricity demand [12]; using ensemble learning intelligence for accurate data mining [13]; using supervised learning for behavioral prediction [14]. In addition, there are many research work related to feature extraction and dialysis. Feature selection for chest radiographs [15] and hormones in hemodialysis patients [16]; relationship between hemodiafiltration and cardiovascular mortality [17]; machine learning for medical data [18] and mental health [19]; changes in salivary composition after hemodialysis [20] and classification of motor imagery [21]; construction of medical big data warehouse [22] and acquisition system [23]; artificial intelligence systems for interpreting medical datasets [24], application in medical field [25] and predicting customer default in E-commerce [26]. These studies utilize machine learning techniques to make decision analysis on medical data or other data.

Methodology

Feature selection

Features of the medical data set may have redundancy and low correlation features. If the important features can be selected from the data set, the dimension disaster problem will be solved. In addition, removing irrelevant features will often reduce the difficulty of training model tasks, and the interpretability of the models will also be improved.

The feature selection process includes the following four parts: generation process, evaluation function, stop criteria, verification procedure:

- 1) **Generation Procedure:** The generation procedure is a process of searching for a subset of features and is

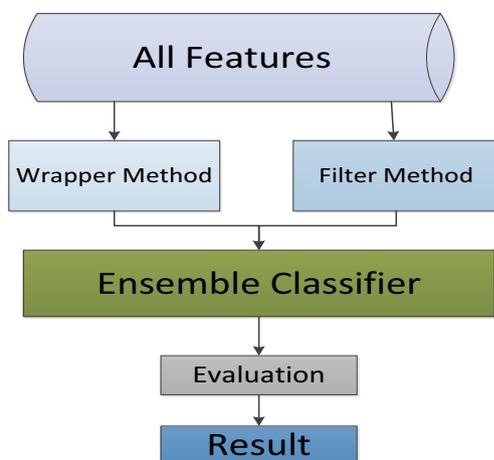


Fig. 2 An overview of the ensemble classifier for Filter and Wrapper

Table 1 The pseudo code of LVW algorithm

Input: Data Set: D ;
 Feature Set A ;
 Learning Algorithm ζ ;
 Stop Condition Control Parameter: T ;

Process:
 1: $E = \infty$;
 2: $d = |A|$;
 3: $A^* = A$;
 4: $t = 0$;
 5: while $t < T$ do
 6: generate feature subsets A' ;
 7: $d' = |A'|$;
 8: $E' = \text{CrossValidation}(\zeta(D^{A'}))$;
 9: if $(E' < E) \vee ((E' = E) \wedge (d' < d))$ then;
 10: $t = 0$;
 11: $E = E'$;
 12: $d = d'$;
 13: $A^* = A'$
 14: else
 15: $t = t + 1$;
 16: end if
 17: end while

Output: Feature Subset A^*

responsible for providing feature subsets for the evaluation function.

- 2) Evaluation Function: The evaluation function is a criterion for evaluating the extent of a feature subset.
- 3) Stopping Criterion: The stopping criterion is related to the evaluation function and is generally a threshold value. When the evaluation function value reaches this threshold, the search can be stopped.
- 4) Validation Procedure: Verify the validity of the selected feature subset on the verification data set.

According to their working principle, common feature selection methods can be roughly divided into two major categories: Filter and Wrapper. Filter feature selection Regardless of the follow-up learner, wrapper feature selection directly uses the performance of the learner as the evaluation criteria of feature subset. It can also be said that the purpose of the wrapper feature selection method is to select the most advantageous feature subset of the learning performance for the learner. Because the wrapper feature selection method directly optimizes the given

Fig. 3 The structure of ensemble learning diagram

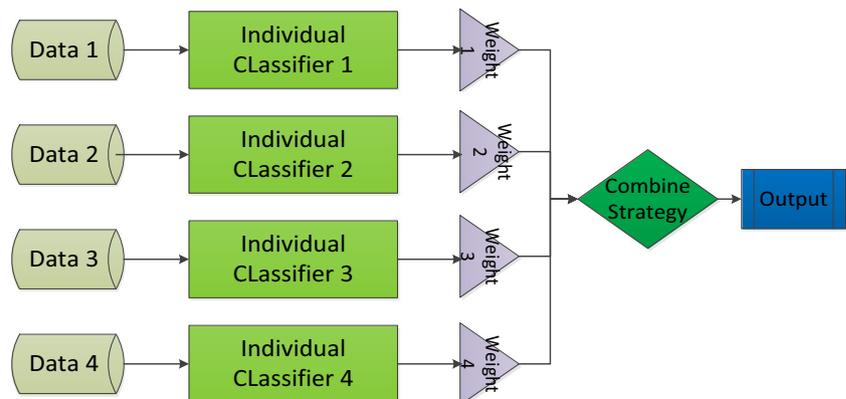


Table 2 The importance coefficient of 25 features based on the PCA

Item	SEX	AGE	HEIGHT	WEIGHT	BPHIGH
Coefficient	-0.052	-0.0	-0.021	0.009	-0.13
Item	BPLOW	SCR	CYSC	CCR	ALB
Coefficient	0.016	-1.126	-1.16	0.001	-0.084
Item	HB	BUN	NS	CA	P
Coefficient	0.181	-0.275	0.156	0.133	0.158
Item	NA	K	CO2	TG	WBC_J
Coefficient	-0.076	0.042	0.086	-0.094	0.279
Item	RBC_J	TP	Q_ALB	GPT	GOT
Coefficient	-0.421	-0.132	0.241	-0.192	0.157

learner, it has better prediction performance than filter feature selection from the perspective of the final learner performance. The sketch that the Filter and the Wrapper for the feature selection embedded in the ensemble classifier was shown in Fig. 2.

In this paper, LVW (Las Vegas Wrapper) method is used for feature selection. LVW uses the random search strategy to search for feature subset under the framework of Las Vegas method. The prediction error of the ensemble classifier which is finally embedded is used as the selected feature subset evaluation standard. In the LVW algorithm, a stochastic strategy is used to select the feature subset. Each evaluation of the feature subset needs to retrain the learner. The pseudo code of LVW algorithm is shown in Table 1. The eighth line of the algorithm utilizes cross-validation on dataset D to estimate the prediction error of learner ζ on feature subset A . If the error of ζ is smaller than the error on the A , the feature subset A is saved.

As shown in Table 1, the LVW selects features based on the accuracy of subsequent models to select the most appropriate feature subsets that are most beneficial to the performance of a particular learner. LVW firstly randomly selects feature subsets, then uses the selected features to build a prediction model, and feeds back the selected features according to their performance. LVW uses a random search strategy to search for feature subsets. In the stochastic strategy, only several feature subsets are randomly selected first, and the learning model is used for multiple trainings. Then add another feature to train, remove the feature with the smallest weight based on the training result, and then carry out the next training based

on the new feature subset. Thus, by recursively eliminating the feature subset of the minimum weight, the model can be constructed repeatedly to find the best feature subset that conforms to the model.

The LVW feature selection method solves the disadvantages of traditional feature selection methods that need to separate feature selection and model building, and directly combines the subsequent model establishment for its customized feature subset.

Ensemble learning

For a complex task, using multiple learners to predict the proper comprehensive judgment is better than either of them alone. Ensemble learning can accomplish learning tasks by combining and building multiple learner, which can give full play to the advantages of each learner, learn multiple classifiers, and effectively improve classification performance. As shown in Fig. 3, we show the general structure of ensemble learning: first we produce a set of "individual learner", and then combine these learners with some strategies.

In this paper, we proposed ensemble learning such as RF (Random Forest), AdaBoost and GBDT (Gradient Boosting Decision Tree) and seven kinds of classifier such as SVM (Support Vector Machines), KNN (k-NearestNeighbor), NB (Naïve Bayes) MLP (Multi-layer Perceptron), ET (Extremely Randomized Trees), and LR (Logistic Regression) to train and test the ten models with a large number of different data sets. A number of experiments were carried out to determine the optimal weights of the ten models. Each weak learner in the process of establishing a predictive model is a classifier, and ensemble learning is performed by the voting weighting method. According to the different predictive performance of each weak classifier, different voting weights are assigned. And the greater the performance, the bigger the voting values. The final prediction result is the weighted average of the prediction results of the weak classifiers. In this way, the prediction performance of the model can be improved and the overfitting possibility of a single classifier can be reduced.

Among these models, AdaBoost is an algorithm that promotes weak learners to strong learners. The working mechanism of this algorithm is to train the base learner

Table 3 The Ranking of features based on RFE method

(Ranking, Features)

(1.0, 'CCR'), (1.0, 'CYSC'), (1.0, 'SCR'), (2.0, 'BUN'), (3.0, 'HB'), (4.0, 'P'), (5.0, 'CO2')
 (6.0, 'ALB'), (7.0, 'NS'), (8.0, 'CA'), (9.0, 'HEIGHT'), (10.0, 'BPHIGH'), (11.0, 'Q_ALB')
 (12.0, 'WBC_J'), (13.0, 'WEIGHT'), (14.0, 'TP'), (15.0, 'SEX'), (16.0, 'GOT'), (17.0, 'RBC_J')
 (18.0, 'TG'), (19.0, 'AGE'), (20.0, 'K'), (21.0, 'NA'), (22.0, 'BPLOW'), (23.0, 'GPT')

Table 4 The weight table of the feature subset based on LVW

(Weight, Features)
(0.906, 'SCR'), (0.667, 'CCR'), (0.652, 'CYSC'), (0.209, 'BUN'), (0.088, 'P')
(-0.265, 'K'), (-0.335, 'RBC_J'), (-0.413, 'HB'), (-0.417, 'BPHIGH'), (-0.452, 'CO2')
(-0.604, 'SEX'), (-0.706, 'AGE'), (-0.734, 'BPLOW'), (-0.736, 'NA')
(-0.746, 'WEIGHT'), (-0.766, 'HEIGHT'), (-0.77, 'GPT'), (-0.816, 'ALB')
(-0.818, 'CA'), (-0.821, 'GOT'), (-0.869, 'TG'), (-0.879, 'WBC_J')
(-0.891, 'NS'), (-0.904, 'Q_ALB'), (-0.929, 'TP')

from the initial training set first, and then adjust the training sample distribution according to the performance of the base learner. The training samples that were previously misclassified by the base learner are assigned more weights later, and the next base learner is trained based on the adjusted sample distribution. AdaBoost adopts the method of weighted majority voting, which increases the weight of weak classifiers with small classification error rate, making it more effective in voting, and reducing the weak classifier with large classification error rate, making it less effective in voting. When training set D is $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$; weak learner is ζ ; the number of training is T ; we get AdaBoost output $H(x)$:

$$b_1(x) = 1/m \tag{1}$$

$$h_t = \zeta(D, b_t) \tag{2}$$

$$\varepsilon_t = P_{x \sim b_t}(h_t(x) \neq f(x)) \tag{3}$$

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_t}{\varepsilon_t} \right) \tag{4}$$

$$H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right) \tag{5}$$

where the classifier h_t is trained from the data set D based on the distribution b_t , and the error of h_t is estimated. Then the weight of h_t is obtained, and the sample distribution is updated.

The evaluation of the generalization ability of the learners requires not only an effective and feasible experimental evaluation method, but also the evaluation index of the

Table 5 The feature selection method on model enhancement

Feature method	Accuracy	Std
PCA	0.9330	± 0.07
RFE	0.9293	± 0.03
MCMC	0.8957	± 0.05
mRMR	0.9361	± 0.04
LVW	0.9668	± 0.04

generalization ability of the model. In contrast to the performance of different models, the use of different performance evaluation indicators will lead to different performance evaluation results.

Experiments and results

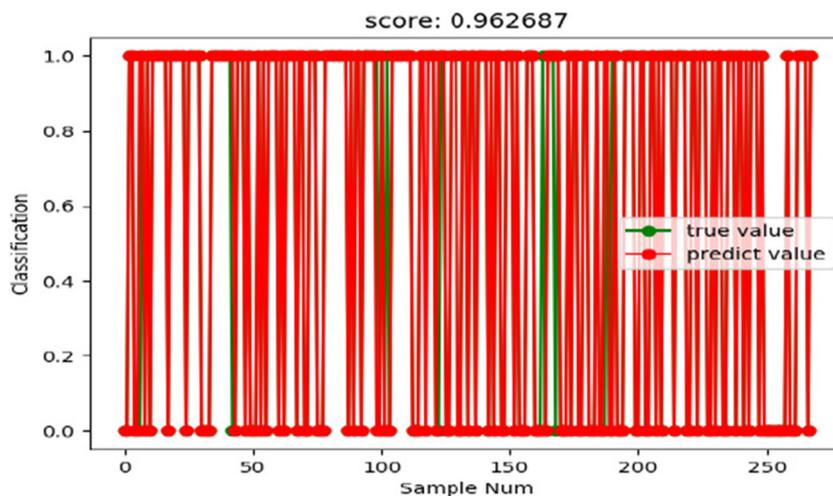
ESRD data preparation and preprocessing

The experimental data that were collected from the HIS system in the Department of Nephrology of Chengdu Military Region General Hospital came from the results of clinical testing for patients with kidney disease from January 2011 to December 2015, including 287,059 incomplete clinical data of all the renal diseases in the hospital at this stage. The prediction model built using this data is suitable for local practice. In the data collection process, the label of each sample of the data is timely and appropriate, that is, the patient's dialysis decision made by the doctor under the current vital signs. So we solved the problem of cognitive bias and expectation influencing the initiation of dialysis.

Data processing consists of the following five steps:

- (1) Unified data storage form: the patient ID as a benchmark, a number of excel file data were fused to form a normative sample of a sample data.
- (2) Remove unrelated attributes: such as patient ID, examination time, etc., and the attributes associated with chronic kidney disease were retained.
- (3) Data cleaning: According to the selected feature attributes, the feature of missing number greater than 15% of the total sample were removed, and samples of Scr&Bun containing missing values were deleted; the missing values of the data were replaced by the corresponding feature mean values.
- (4) Data transformation: In the feature Sex, 0 for men and 1 for women; the category of label attribute Dialysis was replaced by 0 for dialysis, and 1 for non dialysis.
- (5) Data normalization: all features were mapped to the [0,1] according to the maximum and minimum ratio.

Fig. 4 The prediction point map of ensemble learning



After data preprocessing, we obtained 1338 samples of chronic kidney disease and 25 features, including 507 dialysis samples and 831 non dialysis samples.

Experimental results of feature selection

In this experiment, we used PCA/RFE/MCMC/mRMR and LVW method to get the feature correlation coefficient and weight of each feature from 1338 preprocessing samples. The best feature extraction method is selected by comparing the five feature selection methods to improve the performance of the model.

The importance coefficient of 25 features based on the PCA dimension reduction is shown in Table 2. The larger the absolute value of the coefficient, the greater the significance of the feature. It can be seen that the absolute values of the two feature coefficients of both SCR and CYSC are more than 1 as the dominant feature. The absolute value of the feature coefficients of Q_ALB, BUN and RBC_J is the main feature.

In this experiment, ten fold cross-validation is used in the RFE recursive feature elimination method, and the base learner model is RF. As shown in Table 3, the features are sequentially ordered according to their important priorities, and CCR, CYSC, and SCR are the first 3 key features selected by RFE.

The weight ratio of each feature subset is shown in Table 4. With the reference weight coefficient of SCR as a percentage, we can see that the SCR weight dominated, SYSC, CCR and BUN play an important role. From the largest weight of SCR to the smallest weight of TP, it can be seen that the features of SCR, CYSC, CCR and BUN are the main features according to the LVW method.

In this paper, the optimal feature dimensionality reduction method is selected based on the comprehensive comparison of PCA/RFE/MCMC/mRMR and LVW method. From the above experiment, five dominant features of SCR, CYSC, Q_ALB, BUN and RBC_J can be obtained from PCA dimensionality reduction; three dominant features of CCR, CYSC and SCR can be obtained from RFE recursive feature elimination; three dominant features such as SCR, CYSC and CCR can be obtained from MCMC; four dominant features such as SCR, CYSC, CCR and BUN can be obtained from mRMR; four leading features of SCR, CYSC, CCR and BUN can be obtained from LVW method. Table 5 shows the experimental results of the feature selection method on model enhancement.

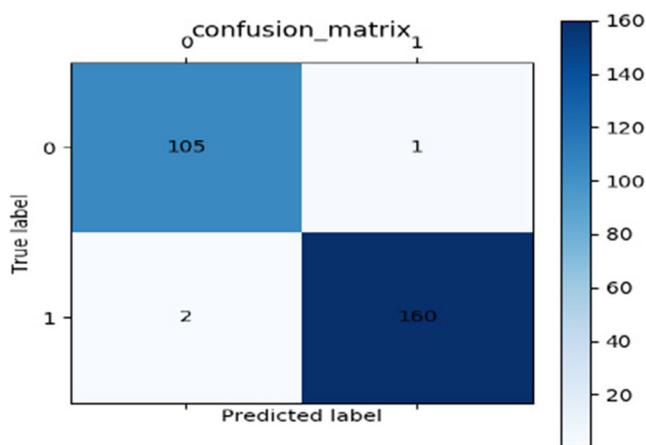


Fig. 5 The confusion matrix of ensemble learning

Table 6 The performance indicator table of ensemble learning

	Precision	Recall	F1-score	Support
Class 0	0.96	0.99	0.98	107
Class 1	0.99	0.98	0.98	161
Avg / Total	0.98	0.98	0.98	268

Table 7 The performance comparison of several state-of-the-art methods

Classifier	Accuracy	Std
Ada	0.9701	± 0.03
SVM	0.9587	± 0.04
KNN	0.9290	± 0.04
RandomForest	0.9479	± 0.04
GBDT	0.6009	± 0.01
NB	0.9473	± 0.04
DT	0.6009	± 0.01
MLP	0.9550	± 0.04
ET	0.9588	± 0.04
LR	0.9550	± 0.04
Ensemble	0.9704	± 0.01

Experimental results of ensemble learning

According to the LVW method, the first five features were selected: SCR, SYSC, CCR, BUN and P, and the experiments used ten classifiers to fuse the model. The experimental data samples were randomly divided into training set 80% (1070 samples), test set 20% (268 samples). After 10 random experiments, the average performance of the prediction model was obtained. In this paper, the model performance obtained by ensemble learning is shown as follows. The integrated model prediction point map is shown in Fig. 4. The confusion matrix of the integration model is shown in Fig. 5, and the performance indicator table of ensemble learning is shown in Table 6.

As shown in Fig. 4, the red prediction value basically coincides with the green real value, which shows that the integrated model is better than the single learner.

As can be seen from the confusion matrix in Fig. 5, only 3 out of 268 samples in the test set are misclassified. It can be seen from Table 6 that the prediction accuracy in Class 1 reaches 99% and the overall prediction performance reaches 98%.

In this paper, a large number of different datasets were used to train and test ten models and ten fold cross-validation was used to conduct multiple experiments to determine the best combination strategy of the model. Table 7 shows the performance comparison between the integrated model and the ten other learners.

As shown in Table 7, in comparison with other learner models, the ensemble learning has the best average precision and lower standard deviation, showing that the ensemble learning that we proposed is effective. In general, the test set error is close to the training set, showing that the ensemble learning model has better generalization ability. In addition, we can also see that most integrated learners perform better than single learners, and AdaBoost method performs better than other control methods.

Conclusion

In this paper, we analyzed and predicted the hemodialysis timing of ESRD, and proposed an enhanced prediction model based on LVW and ensemble learning. According to the results of comparison between PCA/RFE/MCMC/mRMR and LVW, this paper discarded the traditional feature selection method based on the relationship between feature subset, and extracted the 4 key features of SCR, CYSC, CCR and BUN using LVW method. In the aspect of model application, we used a variety of learners such as RF, AdaBoost, GBDT, SVM, KNN, NB, DT, MLP, ET, LR and so on to produce multiple classifiers based on data sets. Ten fold cross-validation was utilized to train and test ten models with a large number of different datasets and a number of experiments were carried out to calculate the optimal weight for the ensemble model. Through a large number of comparative experiments of ten other learners, we can see that the model based on LVW and ensemble learning proposed in this paper has the remarkable effect of reducing the hyperthyroidism features and eliminating the noise, and has more excellent generalization ability (97.04%) and AUC value (99%) and lower standard error (+ − 0.04). The study in this paper may have important clinical effects. We apply machine learning techniques to clinical data to assist physicians in making medical decisions that make decisions smarter and more accurate. The results of the study have proved that the proposed model is effective against predicting the timing of hemodialysis.

Funding This study was funded by the Science and Technology Plan Project of Sichuan Province under Grant 2016GZ0092.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

References

1. Kanno, Y., Kanda, E., Comparison of accuracy between pre-hemodialysis and post-hemodialysis levels of nutritional factors for prediction of mortality in hemodialysis patients[J]. *Clinical Nutrition*, 2017. <https://doi.org/10.1016/j.clnu.2017.12.012>.
2. Bakkaloğlu, S.A., Kandur, Y., Serdaroğlu, E. et al., Effect of the timing of dialysis initiation on left ventricular hypertrophy and inflammation in pediatric patients.[J]. *Pediatric Nephrology*, 32(9):1–8, 2017.
3. Daugirdas, J. T., Hemodialysis treatment time: As important as it seems? *Semin. Dial.* 30:93–98, 2017.

4. Rivara, M. B., and Mehrotra, R., Timing of dialysis initiation: What has changed since IDEAL? *Semin. Nephrol.* 37:181–193, 2017.
5. Low, S., Lim, S. C., Zhang, X., Zhou, S., Yeoh, L. Y., Liu, Y. L., Tavintharan, S., and Sum, C. F., Development and validation of a predictive model for chronic kidney disease progression in type 2 diabetes mellitus based on a 13-year study in Singapore. *Diabetes Res. Clin. Pract.* 123:49–54, 2017.
6. Neves, J., Martins, M. R., Vilhena, J., Neves, J., Gomes, S., Abelha, A., Machado, J., and Vicente, H. A., Soft computing approach to kidney diseases evaluation. *J. Med. Syst.* 39:131, 2015.
7. Polat, H., Danaei Mehr, H., and Cetin, A., Diagnosis of chronic kidney disease based on support vector machine by feature selection methods. *J. Med. Syst.* 41:55, 2017.
8. Sedighi, Z., Ebrahimpour-Komleh, H., Mousavirad, S. J., Feature selection effects on kidney disease analysis[C]//Technology, Communication and Knowledge (ICTCK), 2015 International Congress on. IEEE, 2015:455–459, 2015.
9. Lotfnezhad Afshar, H., Ahmadi, M., Roudbari, M., and Sadoughi, F., Prediction of breast cancer survival through knowledge discovery in databases. *Global J. Health Sci.* 7:392–398, 2015.
10. Caceres, C. A., Roos, M. J., Rupp, K. M. et al., Feature Selection Methods for Zero-Shot Learning of Neural Activity[J]. *Frontiers in Neuroinformatics*, 11:41, 2017.
11. Wang, L., Wang, Y., and Chang, Q., Feature selection methods for big data bioinformatics: A survey from the search perspective. *Methods (San Diego, Calif)* 111:21–31, 2016.
12. Yang, Y., Chen, Y., Wang, Y. et al., Modelling a combined method based on ANFIS and neural network improved by DE algorithm: A case study for short-term electricity demand forecasting[J]. *Applied Soft Computing*, 49:663–675, 2016.
13. Chowdhury, A. K., Tjondronegoro, D., Chandran, V., and Trost, S. G., Ensemble methods for classification of physical activities from wrist accelerometry. *Med. Sci. Sports Exerc.* 49:1965–1973, 2017.
14. Ladds, M. A., Thompson, A. P., Slip, D. J., Hocking, D. P., and Harcourt, R. G., Seeing it all: Evaluating supervised machine learning methods for the classification of diverse otariid behaviours. *PLoS One* 11:e0166898, 2016.
15. Vajda, S., Karargyris, A., Jaeger, S. et al., Feature selection for automatic tuberculosis screening in frontal chest radiographs. *J. Med. Syst.* 42(8):146, 2018.
16. Ulusoy, S., Ozkan, G., Guvercin, B., and Yavuz, A., The relation between variability of intact parathyroid hormone, calcium, and cardiac mortality in hemodialysis patients. *Artif. Organs* 40:1078–1085, 2016.
17. Maduell, F., Varas, J., Ramos, R., Martin-Malo, A., Perez-Garcia, R., Berdud, I., Moreso, F., Canaud, B., Stuard, S., Gauly, A., Aljama, P., and Merello, J. I., Hemodiafiltration reduces all-cause and cardiovascular mortality in incident hemodialysis patients: A propensity-matched cohort study. *Am. J. Nephrol.* 46: 288–297, 2017.
18. Li, H., Zhu, L., Meng, S. et al., Blockchain-based data preservation system for medical data. *J. Med. Syst.* 42(8):141, 2018a.
19. Srividya, M., Mohanavalli, S., and Bhalaji, N., Behavioral modeling for mental health using machine learning algorithms. *J. Med. Syst.* 42(5):88, 2018.
20. Khanum, N., Mysore-Shivalingu, M., Basappa, S., Patil, A., and Kanwar, S., Evaluation of changes in salivary composition in renal failure patients before and after hemodialysis. *J Clin Exp Dent* 9: e1340–e1345, 2017.
21. Kirar, J. S., and Agrawal, R. K., Relevant feature selection from a combination of spectral-temporal and spatial features for classification of motor imagery EEG. *J. Med. Syst.* 42(5):78, 2018.
22. Sebaa, A., Chikh, F., Nouicer, A. et al., Medical big data warehouse: Architecture and system design, a case study: Improving healthcare resources distribution. *J. Med. Syst.* 42(4):59, 2018.
23. Li, B., Li, J., Lan, X. et al., Experiences of building a medical data acquisition system based on two-level modeling. *Int. J. Med. Inform.* 112:114–122, 2018b.
24. Altman, R., Artificial intelligence (AI) systems for interpreting complex medical datasets. *Clinical. Pharmacol. Ther.* 101(5): 585, 2017.
25. Xiang-Yi K , Ren-Zhi W , Neurosurgery D O . Artificial Intelligence and Its Application in Medical Field[J]. *Journal of Medical Informatics*, 2016.
26. Vanneschi, L., Hom, D. M., Castelli, M. et al., An artificial intelligence system for predicting customer default in E-commerce. *Expert Syst. Appl.* 104:1–21, 2018.