Predicting Coronary Atherosclerotic Heart Disease: An Extreme Learning Machine with Improved Salp Swarm Algorithm

Wenming He 1, Yanqing Xie 1, Haoxuan Lu 1, Mingjing Wang 2 and Huiling Chen 1,*

1 Department of Cardiology, the Affiliated Hospital of Medical School; Ningbo University, Ningbo 315020, China; fyhewenming@nbu.edu.cn (W.H.); fxyeyanqing@nbu.edu.cn (Y.X.); fyhuangli@nbu.edu.cn (H.L.)
2 Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam; mingjingwang@duytan.edu.vn
3 College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou 325035, China
* Correspondence: chenhuiling.jlu@gmail.com

Received: 25 August 2020; Accepted: 30 September 2020; Published: 9 October 2020

Abstract: To provide an available diagnostic model for diagnosing coronary atherosclerotic heart disease to provide an auxiliary function for doctors, we proposed a new evolutionary classification model in this paper. The core of the prediction model is a kernel extreme learning machine (KELM) optimized by an improved salp swarm algorithm (SSA). To get a better subset of parameters and features, the space transformation mechanism is introduced in the optimization core to improve SSA for obtaining an optimal KELM model. The KELM model for the diagnosis of coronary atherosclerotic heart disease (STSSA-KELM) is developed based on the optimal parameters and a subset of features. In the experiment, STSSA-KELM is compared with some widely adopted machine learning methods (MLM) in coronary atherosclerotic heart disease prediction. The experimental results show that STSSA-KELM can realize excellent classification performance and more robust stability under four indications. We also compare the convergence of STSSA-KELM with other MLM; the STSSA-KELM model has demonstrated a higher classification performance. Therefore, the STSSA-KELM model can effectively help doctors to diagnose coronary heart disease.

Keywords: salp swarm algorithm; disease diagnosis; swarm intelligence; kernel extreme learning machine; coronary atherosclerotic heart disease

1. Introduction

Coronary atherosclerotic heart disease (CHD), coronary heart disease, is a common cardiovascular disease (CVD) [1]. CHD is caused by abnormal lipid metabolism. The abnormal deposition of lipids and other substances in the blood in the coronary artery can gradually develop into atherosclerotic plaques and cause stenosis and occlusion in the vascular lumen, leading to ischemia, hypoxia, or necrosis of cardiomyocytes in the corresponding blood supply area, and then cause clinical symptoms such as chest tightness and chest pain. In the past decades, the incidence of CVD has been increasing annually and became the biggest threat to human health. Studies have shown that the prevalence of cardiovascular diseases is still on the rise. According to an American Heart Association (AHA) report, ~17.8 million deaths were attributed to CVD globally in 2017, which had an increase of 21.1% compared to the number in 2007. The age-adjusted death rate per 100,000 population was 233.1 in 2017, which decreased by 10.3% from 2007. Overall, the crude prevalence of the cardiovascular disease in 2017 was 485.6 million, an increase of 28.5% over 2007 [2]. It is expected that the mortality rate of CVD will continue to rise in the next 5–10 years. Therefore, the prevention
and control of CVD are in urgent need, and various approaches need to be developed to deal with this public health problem.

In fact, this problem can be effectively addressed if CHD can be diagnosed early and effectively, so as to start a timely intervention. Invasive coronary angiography has long been the gold standard for the diagnosis of coronary heart disease. This operation relies on the support of percutaneous catheterization technology. It is an invasive operation, which has the risk of vascular injury, iatrogenic infection, and other related risks. Therefore, researchers began to develop a non-invasive and efficient diagnostic tool for CHD. Through the health system, a large amount of health data can be collected, including medical history, auxiliary examination, and laboratory examination data [3–5]. This message can be used to establish risk prediction models that can help to predict and prevent future CHD occurrence. Several well-known risk prediction models have been established, including the Framingham risk scores (FRS), the risk of atherosclerotic cardiovascular disease (ASCVD), and diabetes cardiovascular risk assessment in the UK prospective diabetes study (UKPDS). The above model predicts the occurrence of disease through traditional statistical methods, which require that the tested objects must meet the preconditions of the model, such as using the time series model, logistic regression model, and other methods to evaluate the incidence of disease [6,7]. It is generally accepted that current risk prediction models are flawed in assessing CVD risk in specific populations (i.e., young adults and female subjects [8], patients with inflammatory diseases [9], sedentary central obese patients, asymptomatic atherosclerosis patients, chronic kidney disease patients, etc.) and in high-risk populations. Therefore, it is particularly important to develop more effective risk prediction models and to discover more potential risk factors associated with the occurrence of CHD, thereby more effectively predicting the occurrence of CHD.

Diagnosing CHD by machine learning (ML) algorithms are receiving more and more attention from researchers [10]. Artificial neural networks (ANN), decision tree (DT), random forest (RF), and support vector machine (SVM) are the most widely adopted ML methods for CVD research. It applies to almost all data sets reported in the literature with good performance, ease of use, and low computational burden [11]. Dogan et al. developed a CHD risk prediction model by combining single nucleotide polymorphism (SNP) and genome-wide methylation (DNAm) data from 1180 individuals in the training set, using an RF algorithm to predict the risk of symptomatic CHD within five years. The sensitivity and specificity of the test were 0.70 and 0.74, respectively [12]. Donghee Han et al. compared the individual coronary artery calcium score risk prediction model constructed by a boosted ensemble algorithm (LogitBoost) with a traditional risk prediction method, and the result showed that the AUC value in the ML model was obviously more excellent than other models in the test set [13]. The advantage of ML attracts the attention of many researchers in the cardiovascular field. In this study, a new prediction model of CHD is constructed based on the cardiovascular risk factors, and try to incorporate glomerular filtration rate and thyrotropin into the learning model, in order to improve the prediction accuracy further.

In this study, we introduce a space transformation mechanism into salp swarm algorithm (STSSA) [14–17] to optimize the kernel extreme learning machine (KELM) model, and verify that STSSA-KELM can effectively diagnose coronary heart disease. The SSA is one of the metaheuristic methods; hence, it depends on two interactive searching cores [18,19]. In the proposed model (STSSA-KELM), the space transformation mechanism is embedded in SSA to balance exploration and development capabilities further. The improved SSA is used to train an optimal KELM model to predict coronary heart disease. In the experiment, STSSA-KELM was compared with other machine learning algorithms for predicting coronary heart disease. The results show that STSSA-KELM can obtain better sorting results and outstanding stability by observing these four indicators.

The primary contribution of this study can be summarized as follows:

(a) A spatial transformation improved SSA (STSSA) is applied to KELM training.
(b) The established STSSA effectively tackled the parameter turning for KELM in an excellent manner.
(c) The developed STSSA-KELM model is mainly applied to predict coronary heart disease.
The rest of the paper is structured as follows. A brief description of the data is given in Section 2. Section 3 provides a detailed description of the proposed STSSA-KELM. An analysis of the STSSA-KELM on the relevant dataset is presented in Section 4. Section 5 gives some discussions. Conclusions and suggestions for future work are presented in Section 5.

2. Materials and Methods

2.1. DATA Collection

This study collected data from 443 patients who underwent selective coronary angiography during hospitalization, in the Affiliated Hospital of Medical School, Ningbo University, including age, sex, height, weight, history of hypertension, history of hyperlipidemia, history of diabetes, smoking, glomerular filtration rate (eGFR), thyroid-stimulating hormone (TSH). All patients have signed informed consent for coronary angiography. The particulars of the data are shown in Table 1. The data were divided into coronary artery disease group (219 cases) and normal group (224 cases), according to the results of coronary angiography.

The interpretation of the results was performed jointly by two interventional physicians, and if the results were different, the analysis would be made by the third physician. Through the radial artery selective left and right coronary angiography examination, according to the Judkin’s method of multi-position radiation, through the “diameter method” for coronary stenosis degree interpretation: coronary angiography left the main trunk, left anterior descending branch, left circumflex branch, right coronary artery or its main branch of the stenosis rate ≥50%, defined as CHD.

<table>
<thead>
<tr>
<th>No.</th>
<th>Feature</th>
<th>Detailed Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>Age</td>
<td>Normal group (X, SD) = 60.04 ± 10.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD (Coronary Heart Disease) group (X, SD) = 64.08 ± 9.71</td>
</tr>
<tr>
<td>F2</td>
<td>Marriage</td>
<td>Unmarried = 0; Married = 1; Divorce = 2</td>
</tr>
<tr>
<td>F3</td>
<td>Gender</td>
<td>Female = 0; Male = 1</td>
</tr>
<tr>
<td>F4</td>
<td>Weight (kg)</td>
<td>Normal group (X, SD) = 65.62 ± 10.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD group (X, SD) = 65.89 ± 11.07</td>
</tr>
<tr>
<td>F5</td>
<td>Height (cm)</td>
<td>Normal group (X, SD) = 165.26 ± 7.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD group (X, SD) = 164.15 ± 8.22</td>
</tr>
<tr>
<td>F6</td>
<td>Systolic blood pressure (SBP) (mmHg)</td>
<td>Normal group (X, SD) = 132.93 ± 18.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD group (X, SD) = 134.36 ± 17.39</td>
</tr>
<tr>
<td>F7</td>
<td>Diastolic blood pressure (DBP) (mmHg)</td>
<td>Normal group (X, SD) = 78.69 ± 12.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD group (X, SD) = 77.67 ± 12.14</td>
</tr>
<tr>
<td>F8</td>
<td>Heart rate (HR) (/min)</td>
<td>Normal group (X, SD) = 74.13 ± 12.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD group (X, SD) = 73.78 ± 12.60</td>
</tr>
<tr>
<td>F9</td>
<td>Hypertension (HBP)</td>
<td>No = 0; Yes = 1</td>
</tr>
<tr>
<td>F10</td>
<td>Hyperlipidemia (HPA)</td>
<td>No = 0; Yes = 1</td>
</tr>
<tr>
<td>F11</td>
<td>Diabetes mellitus (DM)</td>
<td>No = 0; Yes = 1</td>
</tr>
<tr>
<td>F12</td>
<td>Renal insufficiency/renal failure (RF)</td>
<td>No = 0; Yes = 1</td>
</tr>
<tr>
<td>F13</td>
<td>History of mental illness (HMI)</td>
<td>No = 0; Yes = 1</td>
</tr>
<tr>
<td>F14</td>
<td>History of vascular diseases (HVD)</td>
<td>No = 0; Yes = 1</td>
</tr>
<tr>
<td>F15</td>
<td>History of pulmonary diseases (HPD)</td>
<td>No = 0; Yes = 1</td>
</tr>
<tr>
<td>F16</td>
<td>Smoking history (SH)</td>
<td>No = 0; Yes = 1</td>
</tr>
<tr>
<td>F17</td>
<td>Drinking history (DH)</td>
<td>No = 0; Yes = 1</td>
</tr>
<tr>
<td>F18</td>
<td>Hemoglobin (HB) (g/L)</td>
<td>Normal group (X, SD) = 138.18 ± 16.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD group (X, SD) = 135.77 ± 16.37</td>
</tr>
<tr>
<td>F19</td>
<td>White blood cell count (WBC) (/L)</td>
<td>Normal group (X, SD) = 6.46 ± 1.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD group (X, SD) = 6.62 ± 1.20</td>
</tr>
<tr>
<td>F20</td>
<td>Platelet count (Plt) (/L)</td>
<td>Normal group (X, SD) = 205.94 ± 58.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD group (X, SD) = 202.75 ± 55.69</td>
</tr>
<tr>
<td>F21</td>
<td>Sodium (Na) (mmol/L)</td>
<td>Normal group (X, SD) = 140.99 ± 2.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD group (X, SD) = 141.12 ± 2.25</td>
</tr>
<tr>
<td>F22</td>
<td>Potassium (K) (mmol/L)</td>
<td>Normal group (X, SD) = 4.53 ± 0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD group (X, SD) = 3.99 ± 0.33</td>
</tr>
<tr>
<td>F23</td>
<td>Calcium (Ca) (mmol/L)</td>
<td>Normal group (X, SD) = 2.81 ± 0.95</td>
</tr>
</tbody>
</table>
2.2. Proposed Stssa-Kelm Method

The proposed STSSA-KELM method consists of four parts: data normalization, KELM parameter optimization, feature selection, and classification. First, standardize the collected data and scale the data to \([-1, 1]\). Then, the two critical parameters (C and \(\gamma\)) of KELM are optimized by the continuous STSSA, and at the same time, the binary STSSA algorithm is used to search for the best feature subset of the CHD data. At last, the test data set is classified using the tuned KELM and the best feature subset. The use of K-fold cross-validation is a common approach when dividing data into training and test sets. Therefore, this paper uses 10-fold cross-validation for data partitioning. The flowchart of STSSA-KELM is shown in Figure 1.

2.2.1. Parameter Optimization and Feature Selection by Continuous and Binary STSSA

This research developed a hybrid classification model STSSA-KELM based on STSSA and KELM to predict the CHD data. The parametric optimization of KELM uses continuous STSSA, while binary STSSA is used in the feature selection process. From the original thesis, it is known that the original STSSA is all tested on continuous problems and can be used for constrained or non-constrained problems. In order to use continuous STSSA and discrete STSSA to optimize the KELM parameters and data feature subsets at the same time, a search agent, including continuous variables and discrete variables, is adopted.

2.2.2. Classification Based on KELM

The KELM is a machine learning model established on the basis of ELM and inspired by the kernel function in SVM. The stability and generalization of KELM are better than the traditional extreme learning machine [20–28].

The traditional extreme learning machine is similar to feedforward neuron network (FNN). For an ELM with \(L\) hidden layer nodes, its prediction model is as follows:

\[
\min \left( L = \frac{1}{2} \| \beta \|^2 + \frac{1}{2} C \sum_{i=1}^{N} \| e_i \|^2 \right)
\]
s.t. \( f(x_i) = h(x_i)\beta = t_i - \xi_i, i = 1, ..., N \)

where \( \beta \) is the weight matrix between the hidden layer and output layer, \( C \) is the penalty factor, \( \xi_i \) is the training error, \( N \) is the size of training samples \( \{x_i, t_i\}_{i=1}^N \).

Unlike FNN, the weight matrix between the hidden layer and output layer \( \beta \) is obtained by solving pseudo inverse:

\[
\beta = H^T \left( \frac{1}{C} + HH^T \right)^{-1} T
\]

where \( I \) represents the diagonal matrix, \( H \) is the hidden layer output matrix, and \( T \) denotes the target output target matrix.

Based on ELM, KELM improves the stability and generalization of the model, by introducing kernel functions. The random mappings in the ELM are replaced by nuclear mapping. Since the kernel function takes the form of the inner product, it is not essential to either know the feature mapping function \( h(x) \) of the hidden layer node or set the number of hidden layer nodes. When solving the model output, it is only necessary to know the specific form of the kernel function \( K(x, y) \).

The following equation can define the kernel matrix in KELM:

\[
\Omega_{ELM} = HH^T = K(x, y)
\]

where \( K(x, y) \) uses RBF kernel function, its expression with kernel parameter \( \gamma \) is:

\[
K(x, y) = \exp(-\gamma ||x - y||^2), \gamma > 0
\]

Finally, KELM can be expressed by the following formula:

\[
f(x) = \left[ K(x, x_1) \quad \vdots \quad K(x, x_N) \right]^T \left( \Omega_{ELM} + \frac{1}{C} \right)^{-1}
\]

2.2.3. Detailed Procedure of STSSA-KELM

Two parameters in KELM were tuned using STSSA, in order to improve the stability and generalization performance of KELM fully. The selection of CHD features was performed while optimizing these two parameters. The detailed process for parameter optimization and feature selection through STSSA is as follows:

Step 1: Initialize the parameters of STSSA: the maximum number of iterations \( T \), the number of search agents \( N \).

Step 2: Initialize the search agents of STSSA. Use random numbers generated in the solution space to initialize continuous variables in the search agent, and use random 0 or 1 to initialize discrete variables.

Step 3: Calculate the fitness value of each search agent, according to the following formula:

\[
\begin{aligned}
f_1 &= \frac{\sum_{i=1}^{K} acc_i}{K} \\
f_2 &= 1 - \frac{\sum_{i=1}^{n} bin_i}{n} \\
f &= \alpha \times f_1 + \beta \times f_2
\end{aligned}
\]
where \( f_1 \) is the mean accuracy of KELM obtained on K-fold CV, \( K = 5 \). In \( f_2 \), \( bin_j \) denotes the binary value of the \( j \)-th dimension, and \( n \) denotes the number of all features of the CHD data. The \( \alpha \) and \( \beta \) are the weights of \( f_1 \) and \( f_2 \). The values of \( \alpha \) and \( \beta \) are set to be 0.99 and 0.01, respectively.

Step 4: Perform spatial transformation mechanism and select the highest fitness \( N \) search agent updates the current population.

Step 5: Update parameter \( c \), according to Equation (2).

\[
c = 2 \times e^{-\frac{\alpha}{(1)}}^2
\]

Step 6: Update the value of the search agents.

Step 7: If the maximum iterations are satisfied, output the best search agent where the first two dimensions represent \((C, \gamma)\), and the binary values of the other dimensions are used to filter out the selected features. Otherwise, jump to Step 4.

Step 8: Optimize the obtained optimal parameters and optimal feature subsets with the KELM prediction model, and use the optimal model to predict the test set.

Step 9: If the termination condition is met, output the average result. Otherwise, jump to Step 8.

3. Results

The experiment was done with the MATLAB R2018 software. The data are normalized into \([-1,1]\), before processing the classification task. 10-fold cross-validation (CV) was employed to split the date, in which nine parts were used for training data and the last one for the testing phase. STSSA, SSA, PSO, and GWO were implemented from scratch. Regarding the random forest (RF), KELM, and SVM, the related codes shared in public websites (http://www3.ntu.edu.sg/home/egbhuang, https://www.csie.ntu.edu.tw/~cjlin/libsvm/, and https://code.google.com/archive/p/randomforest-matlab) were adopted. Moreover, the trial and error method is used to set parameters. The search
range of the two parameters of SVM and KELM were set the same in the \([2^{-10}, 2^{10}]\). For RF, the default setting was used.

In this experiment, we estimated the effectiveness of the STSSA-KELM model. The detailed results are shown in Table 2. From the table, we can see that the classification accuracy obtained by STSSA-KELM is 84.40%, Matthews correlation coefficient is 69.20%, sensitivity is 87.30%, specificity is 81.70%, and the variance is 0.058, 0.0118, 0.074, and 0.065, respectively. In addition, what we can observe in the experiment is that the STSSA algorithm can automatically obtain the optimal parameters of KELM, which shows that the addition of spatial transformation mechanism makes the new STSSA algorithm have more substantial search capability and accuracy.

Table 2. Classification performance of STSSA-KELM, in terms of ACC (Accuracy), MCC(Matthews correlation coefficient), sensitivity, and specificity.

<table>
<thead>
<tr>
<th>Fold</th>
<th>ACC</th>
<th>MCC</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>0.9550</td>
<td>0.9130</td>
<td>1.0000</td>
<td>0.9130</td>
</tr>
<tr>
<td>#2</td>
<td>0.8000</td>
<td>0.6010</td>
<td>0.8180</td>
<td>0.7830</td>
</tr>
<tr>
<td>#3</td>
<td>0.8000</td>
<td>0.6010</td>
<td>0.8180</td>
<td>0.7830</td>
</tr>
<tr>
<td>#4</td>
<td>0.8670</td>
<td>0.7370</td>
<td>0.9090</td>
<td>0.8260</td>
</tr>
<tr>
<td>#5</td>
<td>0.8410</td>
<td>0.6830</td>
<td>0.8180</td>
<td>0.8640</td>
</tr>
<tr>
<td>#6</td>
<td>0.9090</td>
<td>0.8220</td>
<td>0.9550</td>
<td>0.8640</td>
</tr>
<tr>
<td>#7</td>
<td>0.8180</td>
<td>0.6360</td>
<td>0.8180</td>
<td>0.8180</td>
</tr>
<tr>
<td>#8</td>
<td>0.7500</td>
<td>0.5050</td>
<td>0.8180</td>
<td>0.6820</td>
</tr>
<tr>
<td>#9</td>
<td>0.8410</td>
<td>0.6830</td>
<td>0.8180</td>
<td>0.8640</td>
</tr>
<tr>
<td>#10</td>
<td>0.8640</td>
<td>0.7400</td>
<td>0.9550</td>
<td>0.7730</td>
</tr>
<tr>
<td>Mean</td>
<td>0.8440</td>
<td>0.6920</td>
<td>0.8730</td>
<td>0.8170</td>
</tr>
<tr>
<td>STD.</td>
<td>0.0580</td>
<td>0.1180</td>
<td>0.0740</td>
<td>0.0650</td>
</tr>
</tbody>
</table>

To validate the usability of this approach, we present a comparative study with five other effectiveness machine learning models, including SSA-KELM, PSO-KELM, GWO-KELM, SVM, and RF. The comparison results of the six methods are shown in Figure 2. The results show that the STSSA-KELM model is superior to the other five models in four evaluation indexes. This signifies that the STSSA-KELM model obtains better performance and stability by adding a spatial conversion mechanism. In the ACC evaluation index, the STSSA-KELM model has the best evaluation effect, which is 1.80 percentage points higher than GWO-KELM. SSA-KELM, RF, and PSO-KELM are in the second place, SSA-KELM is 2.00 percentage points lower than STSSA-KELM, the effect of SVM is the worst, GWO-KELM variance is the largest, 0.079. In the MCC evaluation index, the STSSA-KELM model still acquired the best results, followed by GWO-KELM. GWO-KELM was 3.80% lower than STSSA-KELM, SSA-KELM, RF, and PSO-KELM were behind, SSA-KELM was 4.00% lower than STSSA-KELM, SVM effect was the worst; GWO-KELM variance was the largest, 0.158. From the perspective of sensitivity evaluation indexes, the STSSA-KELM model gets the best evaluation effect, followed by the SSA-KELM model, with a difference of only 2.00 percentage points, followed by RF, GWO-KELM, and SVM. PSO-KELM model has the worst effect and the largest variance, reaching 0.089. The results of STSSA-KELM and GWO-KELM were the best. PSO-KELM, RF, and SSA-KELM are the second, in which the difference between SSA-KELM and STSSA-KELM is only 2.30 percentage points; the worst is SVM, and the variance of PSO-KELM is the largest, reaching 0.109.
We have conducted the Friedman test to validate the significance of the proposed STSSA-KELM as shown in Table 3. As shown in Figure 2, the results show that STSSA-KELM has the best performance, and SSA-KELM is the second. From the table, we can see that the mean level of STSSA-KELM is 1 higher than that of the second SSA-KELM, which further proves the effectiveness of STSSA-KELM.

Table 3. Comparison results of STSSA and other five models, based on the Friedman test.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>STSSA-KELM</th>
<th>SSA-KELM</th>
<th>PSO-KELM</th>
<th>GWO-KELM</th>
<th>SVM</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-level</td>
<td>1.5</td>
<td>3</td>
<td>4.75</td>
<td>5</td>
<td>2.5</td>
<td>4.25</td>
</tr>
<tr>
<td>Rank</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

To characterize the convergence of the proposed STSSA-KELM model, we also record the tendency that the accuracy of three kinds of KELM models, which are based on swarm intelligence algorithm, changes with population iteration. It is revealed in Figure 3 that after many iterations, the STSSA-KELM model can quickly and continuously jump out of the local optimum to reach the optimal accuracy, indicating that the STSSA-KELM method has strong local search ability and global search ability. The main reason is that the spatial transformation mechanism enhances the local search ability and global search ability of the SSA, and makes the hybrid algorithm better balanced between the exploration and the exploitation. By observing the curve in Figure 3, it is found that the SSA-KELM model needs more iterations to converge, and the accuracy is not as high as STSSA-KELM. GWO-KELM has the second-highest accuracy, which can continuously jump out of the local optimum, and the improvement with iteration accuracy is not apparent. PSO-KELM has the lowest accuracy among all algorithms, far less than the STSSA-KELM model, and the improvement with iteration accuracy is not apparent, so it is easy to sink into the local optimum.
In this procedure, STSSA not only achieves the best setting of KELM, but also chooses the best feature set. We used ten times of CV technology. Figure 4 illustrates the frequency of the main characteristics identified by the STSSA-KELM during the 10-fold CV procedure. As shown in the figure, F33 (thyrotropin) and F26 (estimated glomerular filtration rate) are the two characteristics with the highest frequency, 10 and 9 times, respectively. Other clinically widely validated indicators, such as F1 (age), F3 (gender), F4 (weight), F5 (height), F9 (hypertension), F10 (hyperlipidemia), F11 (diabetes mellitus), F16 (smoking history), also showed a higher frequency.

4. Discussion

According to the 2019 Guidelines for Primary Prevention of Cardiovascular Diseases issued by ACC/AHA, hypertension, smoking, dyslipidemia, diabetes mellitus, overweight and obesity, inadequate physical activity, unreasonable diet, metabolic syndrome and air pollution are nine
crucial risk factors for CVD [29]. Since the 1970s, many countries and regions have carried out large-scale studies on cardiovascular risk factors, and have successively introduced a variety of cardiovascular risk assessment methods, such as the FRS [30], ASCVD [31], UKPDS [32], and so on. Many clinical researchers have recognized the established traditional risk factors of CVD and its risk prediction model. However, it cannot wholly and accurately predict CHD. Therefore, researchers are also committed to exploring new predictors and biomarkers to help better predict the risk of CHD.

Studies have found that patients with chronic kidney disease have a high incidence of cardiovascular events [33]. Potential mechanisms include the accumulation of harmful biological factors due to endothelial function, decreased renal clearance, activation of the renin-angiotensin system, oxidative stress, disorders of lipid metabolism, chronic inflammation caused by the accumulation of inflammatory factors, etc. [34]. Tonelli et al. compared the incidence of cardiovascular events between patients with chronic kidney disease and those without chronic kidney disease through a population cohort study. They found that the incidence of cardiovascular events in patients with chronic kidney disease was significantly higher than that in the latter, revealing that chronic kidney disease may be a potential cardiovascular risk factor [35]. Matsushita et al. explored the independent association of changes in eGFR with CHD and all-cause death, and the results showed that changes in eGFR were related to a higher risk of CHD and all-cause death [36].

TSH is a hormone secreted by the adenohypophysis, which promotes the proliferation of thyroid follicular epithelial cells and the synthesis and release of thyroid hormones. TSH responds significantly and accurately to small changes in thyroid hormone concentrations in circulating blood. Besides, there is increasing evidence that thyrotropin has extrathyroid effects [37]. Studies have shown that hypothyroidism is associated with an increased incidence of CHD, with potential mechanisms including changes in lipid profiles caused by thyroid hormones, and abnormal lipid metabolism (increased levels of total cholesterol and low-density lipoprotein cholesterol), leading to atherosclerosis [37]. Onat et al. studied the correlation between serum TSH level and the risk of CHD in the normal range, and the results showed that the TSH level might be associated with the incidence of CHD [38]. A prospective study conducted by Li et al. explored the effect of TSH on the risk of CHD in patients with normal thyroid function, and analyzed the correlation and potential value of TSH in predicting CHD. The results showed that TSH could be used as a biomarker for predicting CHD [39].

Based on the above analysis results, in addition to the traditional risk factors of CHD, such as age, sex, height, weight, history of hypertension, history of hyperlipidemia, history of diabetes mellitus, and smoking history, we also incorporated TSH and eGFR into the proposed ML model. According to the experimental results, we can find that the accuracy can be improved significantly after incorporating the factors of eGFR and TSH, compared with the model constructed by traditional risk factors of CHD.

It should be noted that there are also some deficiencies in this study. Firstly, in the original data, we tried to count the symptoms of patients with CHD. However, in the end, because the clinical manifestations of patients with CHD are too complex to make statistics more complicated, we removed these data; there are also different degrees of missing sample values in the data set, among which the missing laboratory test indicators are more serious. More missing data are BNP, which may lead to a decline in the actual diagnostic performance of the model because the BNP value filled with the mean is lower than the BNP level in patients with heart failure, which is considered as a possible complication of CHD. In addition, traditional risk scores such as FRS and ASCVD are largely derived from data from non-Spanish and American African populations, while our data are derived from Asian populations, and there is some racial interference in the comparison of the results.

In conclusion, in this research, we put forward a new diagnostic model of CHD based on electronic health data using a machine learning algorithm. The model has certain risk assessment ability, which can provide a personalized reference for risk assessment of CHD in the clinic, and help to reduce the workload of clinicians and provide a new method for the clinical diagnosis of CHD.
5. Conclusions and Future Perspectives

In this study, an efficacious STSSA-KELM model was developed to predict coronary heart disease. The main originality of this method is to introduce the space transformation mechanism into SSA, which is used to better balance the global search capability and the local search capability. STSSA-KELM has superior prediction accuracy and a more consistent performance than other machine learning algorithms in the diagnosis of coronary heart disease.

In the future, it is our goal to utilize the STSSA-KELM model to diagnose other diseases, such as pleural effusion, diabetes, appendicitis, and so on. Additionally, the core method STSSA can also be used for parameter tuning for the artificial intelligence models, such as support vector machines [40–44], and convolutional neural networks [45–49]. Moreover, it can also be applied to tackle the feature selection problems [28,41,50–51]. Other problems including social evolution modelling [52], video coding optimization [53], parameter estimation of solar cell [54–57], and others [58–62] are also interesting topics that are worthy of investigation in the near future.

Author Contributions: W.H. and Y.X. conceived and designed the experiments; M.W. performed the experiments; H.L. and H.C. analyzed the data; H.C. contributed reagents/materials/analysis tools; W.H. wrote the paper.

Funding: This research received no external funding.

Acknowledgments: We acknowledge the efforts of Ali Asghar Heidari (http://aliasgharheidari.com) during this research.

Conflicts of Interest: The authors declare no conflict of interest.

References:


© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).