International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 580-589

https://doi.org/10.70517/ijhsa46452

Research on predictive and preventive mechanisms of athletes' sports injuries and rehabilitation treatment strategies using big data computational analysis methods

Yu Cheng[1,*]
[1] Physical Education School, Hoseo University, Chungcheongnam-do, 31499, South Korea
Corresponding authors: (e-mail: chengyu526113@163.com).

Abstract Sports injuries have become an important factor affecting athletes' competitive performance and career, and traditional prevention methods mainly rely on empirical judgment, which lacks scientificity and precision. Constructing an efficient sports injury prediction model and formulating corresponding rehabilitation strategies are of great significance to improve athletes' health management. In this study, we constructed a sports injury prediction model using Improved Whale Optimization Algorithm Optimized Support Vector Machine (IWOA-SVM), and analyzed it based on 1000 athletes' records in Kaggle dataset. The traditional whale optimization algorithm was improved by Circle chaos strategy initialization, inertia weight adjustment and Cauchy variation strategy, and the prediction model was built by combining with support vector machine. Correlation analysis showed that training intensity was significantly correlated with injury likelihood (p=0.007). The results of model performance evaluation showed that the IWOA-SVM model had an accuracy of 93.92%, a precision of 92.79%, a recall of 93.52%, and an AUC value of 95.45%, which were better than the traditional machine learning methods in all indicators. The feature importance analysis showed that height, weight and training intensity were the key predictors, and the influence weights were more than 0.24. Based on the prediction results, personalized rehabilitation treatment strategies including strains, abrasions, joint sprains and contusions were developed. The prediction model provided a scientific basis for the prevention of sports injuries, and the rehabilitation strategies provided a systematic guide for the athletes' post-injury recovery.

Index Terms Improved whale optimization algorithm, support vector machine, sports injury prediction, training intensity, rehabilitation treatment strategy, machine learning

I. Introduction

Athletes inevitably have sports injury situations in their career. Sports injuries vary by sport type, such as higher rates of arthritic injuries in gymnasts and runners, and higher rates of shoulder injuries in swimmers [1], [2]. In addition to sports injuries caused by physical training, the risk of sports injuries caused by long event cycles and psychological factors is also prevalent, long cycles lead to athlete fatigue and increase the risk of injury, while the risk of sports caused by negative emotions, such as psychological anxiety or tension, is also gradually rising, and psychological factors also affect the effect of post-injury rehabilitation [3]-[5]. In addition, during the training of athletes, long-term and regular participation in sports training activities can effectively improve the physical function of athletes, promote the formation of a good competitive state, and then win the game [6]. However, long-term participation in sports training can easily lead to over-training, and the accumulation of fatigue will bring considerable risk of injury to athletes, resulting in muscle and joint pain, muscle strength and joint instability, which will seriously affect athletes' performance on the field of play, and even the end of their athletic careers [7]-[10]. From the social level, sports injuries cause a huge waste of medical resources. Therefore, scientific and effective prediction, prevention, and rehabilitation of sports injuries are of great significance to prolonging athletes' careers and alleviating the pressure on medical resources.

Due to the high-intensity use of athletes' bodies, the injuries they suffer are often more complex and require longer recovery time than ordinary people, and the science and effectiveness of rehabilitation treatment are directly related to the athletes' comeback effect and future performance [11], [12]. Currently, it is believed that the key to sports injury prediction, prevention and rehabilitation lies in identifying risk factors that are closely related to the formation of sports injuries, and then tracking the changes of risk factors to capture the time window of the occurrence of sports injuries in order to implement injury prevention measures in advance [13]. Therefore, it has been an important part of research in the fields of preventive medicine and biomedical engineering to collect data on indicators closely related to the risk of sports injuries, and to construct an accurate injury prediction and prevention mechanism, and to formulate rehabilitation treatment programs accordingly [14].

Many scholars have used statistical means to analyze the relationship between predictive indicators and the risk of sports injuries and construct statistical models so as to realize the quantitative prediction of athletes' injury risk [15], [16]. However, with the gradual deepening of the research, the statistical model has gradually exposed some problems. Along with the emergence of wearable sports equipment and the construction of digital training venues, a huge amount of data resources on athletes' training and injuries have been produced, which provides data support for the scientific prediction of sports injury risk and the development of preventive mechanisms and rehabilitation programs [17], [18]. However, due to the problems of large volume, low value density and multiple data types of this data resource, the algorithms and arithmetic power of data processing put forward high requirements, while statistical learning models usually use the traditional parameter inference method, which requires the estimation of a fixed number of parameters, and when dealing with large-scale data, the complexity of the model will be significantly increased, and the speed of the feedback will be significantly reduced [19]-[21].

In modern competitive sports, athletes face unprecedented training intensity and competition pressure, and the incidence of sports injuries is on the rise, which not only affects the athletes' competitive status, but also may end their careers. Traditional sports injury prevention is mainly based on coaches' experience and medical staff's clinical observation, which lacks quantitative scientific basis and is difficult to realize accurate personalized prevention. The booming development of artificial intelligence technology has revolutionized the field of sports medicine, especially machine learning algorithms have shown great advantages in handling multi-dimensional and non-linear sports data. By integrating athletes' physiological indicators, training parameters, technical movement characteristics and historical injury records, machine learning models are able to identify potential injury risk factors and predict the probability of injury occurrence, thus realizing the shift from passive treatment to active prevention. However, existing sports injury prediction models generally suffer from improper feature selection, insufficient algorithm optimization, and limited prediction accuracy, making it difficult to meet the demands of practical applications.

In this study, we propose to optimize the sports injury prediction model based on the improved whale optimization algorithm for support vector machines, and improve the traditional whale optimization algorithm by introducing the Circle chaos strategy, inertia weight adjustment and Cauchy's variance strategy to enhance the parameter optimization ability. The feature set was constructed using multi-source athlete data, and correlation analysis was used to screen the key predictors to establish a high-precision injury risk prediction model. On this basis, personalized rehabilitation treatment strategies are formulated by combining the physiological characteristics and rehabilitation needs of different injury types, forming a closed-loop management system of prediction-prevention-rehabilitation.

II. Data sources and analysis

II. A. Data sources

The Kaggle-sourced “injury_data.csv” dataset contains 1,000 records covering 7 key characteristics (age, weight, height, past injuries, training intensity, injury recovery time, and likelihood of injury, etc.), with the goal of predicting the probability of injury in athletes. The data are comprehensive, integrating biological, training and historical health information (Table 1), providing a solid foundation for modeling and facilitating personalized monitoring and strategy adjustment. Preprocessing involved mean-filling missing values and data standardization (mean 0, variance 1) to guarantee accurate and efficient analysis.

II. B. Correlation analysis

Univariate significance analysis is shown in Table 1. The t-test analysis showed that only training intensity was significantly different between the different injury likelihood groups, based on the obtained p-value of 0.007, which is less than 0.05. Partial correlation coefficients between the characteristics and the target variable Likelihood_of_Injury were then measured, and the results of the partial correlation analysis are shown in Table 2. The partial correlation analysis further excludes the interference of other variables and confirms the significant linear association between training intensity and likelihood of injury, which contains the partial correlation coefficients and 95% confidence intervals for each feature, emphasizing the independent importance of training intensity. Based on these correlation analyses, it was tentatively concluded that training intensity was the factor most associated with injury likelihood.

Table 1: Univariate significance analysis

Feature name	T statistic	P value	Pearson
Player_Age	-0.008	0.707	0.0002
Player_Weight	-0.056	0.412	0.0021
Player_Height	-0.919	0.534	0.0297
Previous_Injuries	-1.203	0.528	0.0373
Training_Intensity	-2.813	0.007	0.0852
Recovery_Time	0.496	0.824	-0.0154

Table 2: Partial correlation analysis

Feature name	Pearson	CI95%	P value
Player_Age	-0.005	[-0.073,0.058]	0.384
Player_Weight	-0.006	[-0.075,0.057]	0.513
Player_Height	0.028	[-0.043,0.089]	0.421
Previous_Injuries	0.037	[-0.025,0.098]	0.284
Training_Intensity	0.087	[-0.036,0.151]	0.007
Recovery_Time	-0.017	[-0.072,0.049]	0.395

III. Modeling and evaluation of sports injury prediction

III. A. Sports injury prediction model

In this study, we constructed a multimodal dataset, analyzed the data from multiple dimensions and perspectives, and designed a sports injury prediction model based on IWOA-SVM, which was provided to predict and prevent sports injuries in athletes.

III. A. 1) WOA algorithm

The Whale Optimization Algorithm (WOA) simulates the whale's predatory behavior, which is divided into three phases: encirclement, attack, and search, in order to construct the optimization model. When the whale hunts, it takes the current optimal solution as the goal and guides other individuals to encircle the prey, and this behavior is expressed by Eq:

$$D = \left| CX^*(i) - X(i) \right| \tag{1}$$

$$X(i+1) = X^*(i) - AD \tag{2}$$

where, D is the distance between the whale and the prey, i is the number of iterations, $X^*(i)$ is the optimal position where the whale is located, $X(i)$ is the i th whale position, $X(i+1)$ is the $i+1$ th second whale position, and A and C are the coefficient matrices.

Bubble net predation is another type of predation in the whale population, which narrows the encirclement through a spiral path and releases bubbles to form a bubble net, which is mathematically modeled as:

$$X(i+1) = X^*(i) + D'e^{bl}\cos(2\pi l) \tag{3}$$

$$D' = \left| X^*(i) - X(i) \right| \tag{4}$$

where D' is the distance between the i th whale and its prey, b is the parameter of the helix, and l is a random number between $[-1,1]$.

Whales also perform search behavior during feeding, which is expressed by Eq:

$$D = \left| CX_{rand}(i) - X(i) \right| \tag{5}$$

$$X(i+1) = X_{rand}(i) - AD \tag{6}$$

where, $X_{rand}(i)$ is the i th whale random position vector.

SVM is a linear classification method, and the basic idea is to find an optimal hyperplane for distinguishing the original samples and ensuring that the distance from each type of sample to that hyperplane is as large as possible.

When building a support vector machine model, the penalty factor C and the kernel function parameter g have a significant impact on the accuracy of the model, and this paper applies the improved WOA to optimize the parameters C and g in the SVM to improve the accuracy of sports injury prediction for athletes.

III. A. 2) Support vector machines

The main use of Support Vector Machines (SVMs) is in regression prediction of nonlinear data as well as classification, etc. SVMs build decision surfaces based on classification hyperplanes to maximize the bounding intervals of the positive and negative examples within a certain range, and furthermore to be precise, SVMs reduce structural risk.

With $f(x)$ as the center, a yellow interval region with width ε is constructed on each side, and the prediction is considered to be relatively accurate if the training sample points belong to the above interval region.

(1) Representing the original dataset as equation (7), the nonlinear regression equation of this dataset is $f(x)=\omega\varphi(x)+b$, where ω is the weight vector and b is the bias vector, and:

$$\{(x_1,y_1),(x_2,y_2),\cdots,(x_i,y_i),\cdots,(x_m,y_m)\mid x_i\in R^n, y_i\in R\} \tag{7}$$

(2) The introduction of the penalty constant C, and the relaxation variables ξ_i and ξ_i' for different cases above and below the curve, minimizes the functional value of Eq. (8), which is used to solve for the ω and b values above:

$$\min\left\{\frac{1}{2}\|\omega\|^2+C\sum_{i=1}^{m}(\xi_i+\xi_{i'})\right\} \tag{8}$$

$$s.t.\begin{cases} f(x_i)-y_i\le\varepsilon+\xi_i \\ y_i-f(x_i)\le\varepsilon+\xi_{i'} \\ \xi_i,\xi_{i'}\ge 0,(i=1,2,\cdots,m)\end{cases} \tag{9}$$

(3) According to the Lagrangian dyadic principle, by constructing the Lagrangian function and combining it with the optimization of constraints, we obtain the dyadic problem formula (10) and the equation constraint formula (11):

$$\min\left\{\frac{1}{2}\sum_{i,j}^{m}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)(x_i,x_j)+\sum_{i=1}^{m}\alpha_i(\varepsilon-y_i)+\sum_{i=1}^{m}\alpha_i^*(\varepsilon+y_i)\right\} \tag{10}$$

$$s.t.\begin{cases}\sum_{i=1}^{m}(\alpha_i-\alpha_j)=0 \\ \alpha_i,\alpha_j\in[0,C]\end{cases} \tag{11}$$

(4) Solve to obtain the final nonlinear regression function as in equation (12):

$$f(x_i)=\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)\kappa(x,x_i)+b \tag{12}$$

where, $\kappa(x,x_i)$ is the kernel function.

III. A. 3) Improved WOA

(1) Improvement of population initialization by introducing Circle chaos strategy

In order to make the population in the whale optimization algorithm can appropriately expand the search range of the algorithm when initializing, the introduction of the chaos strategy is used in order to better initialize the population. The structure and chaotic properties of Circle chaotic mapping, which is a typical representative of chaotic mapping, are utilized, and its equations are as follows:

$$x_{i+1}=\mathrm{mod}\left(x_i+a-\left(\frac{b}{2\pi}\right)\sin\left(2\pi x_i\right),1\right) \tag{13}$$

The steps to initialize the population using the inverse learning strategy are as follows: first, multiple initial solutions are generated by Circle chaotic mapping, and then the different initial solutions are used to generate the corresponding inverse solutions as follows:

$$O_i = S\left(X_{\min}^d + X_{\max}^d\right) - X_i \tag{14}$$

where S represents the perturbation factor, randomized between $[0,1]$, O_i represents the inverse solution corresponding to different initial solutions X_i, $X_{\min}^d$ represents the minimum value of the d-dimensional vector in the initial solution, and $X_{\max}^d$ represents the maximum value of the d-dimensional vector.

(2) Improvement of tuning strategy for inertia weight parameters

Humpback whales round up prey through bubble nets in the prey stage, then the mathematical model is established by imitating the behavior of humpback whales in rounding up prey, and in the mathematical model, with continuous iteration, the humpback whale group will constantly shrink the bubble net, which will easily lead to the model falling into the local optimal situation. Therefore, for the above situation is used to add an inertia weight ω, which varies with the optimal fitness of the whales, and the optimal fitness obtained in each round of the iterative process is set as a feedback parameter. In the iterative process of the group of whales, a weight parameter ω is given to each whale, and the whale's behavior is determined by the size of the ω value, when the value of ω is larger, the whale chooses to expand the scope of the bubble net, and vice versa. Then:

$$\omega = \omega_{\max} - \frac{\left(\omega_{\max} - \omega_{\min}\right)\cdot T}{T_{\max}} + \sigma B\left(b_1, b_2\right) \tag{15}$$

$$X_i^{t+1} = \left|X_{best}^t - X_i^t\right| \cdot e^{bl} \cdot \cos(2\pi l) \cdot \omega + X_{best}^t \tag{16}$$

where, ω is the inertia weight, $\omega_{\max}$ is the maximum inertia weight, $\omega_{\min}$ is the minimum inertia weight, T is the number of current iterations, $T_{\max}$ is the maximum number of iterations, which is usually taken as $\omega_{\max} = 0.9$, $\omega_{\min} = 0.4$.

(3) Improvement by Introducing Cauchy Variation Strategy

Aiming at the problem that the original whale optimization algorithm is prone to fall into the local optimal situation during the iteration process, a Cauchy variation strategy that can make the population jump out of the local optimal situation is adopted, which augments the diversity of the population to enhance the global search ability of the whale optimization algorithm in the solution space.

Then the following Cauchy variation formula (17) is used to mutate the global optimal solution at this point:

$$x_{new}^*(t) = x^*(t) + x^*(t) \times cauchy(0,1) \tag{17}$$

where $x_{new}^*(t)$ is the new value produced by the current global optimal solution after Cauchy variation, $cauchy(0,1)$ is the Cauchy operator, and the standard formula for the density function of the Cauchy distribution is given in the following equation:

$$f(x) = \frac{1}{\pi \times (x^2+1)}, x \in (-\infty, +\infty) \tag{18}$$

III. A. 4) IWOA Optimization of SVMs

The following steps are taken in optimizing the parameters of the support vector machine using the improved whale optimization algorithm:

(1) Data preprocessing and segmentation: before starting, the original dataset is first normalized to ensure that the features have the same scale. Next, the dataset is randomly extracted and 80% of the data is assigned to the training set and the remaining 20% is used for the testing set. This ensures that the model is supported by sufficient data in both the training and testing phases.

(2) Initialize the whale population: the whale population is initialized using a combination of chaotic mapping and inverse learning strategy.Circle chaotic mapping and inverse learning strategy can help the whale population to better explore the parameter space. The generated individual whales contain a subset of the SVM's parameters and features. In this stage, the following equation is used:

$$x_{i+1} = \bmod\left(x_i + a - \left(\frac{b}{2\pi}\right)\sin\left(2\pi x_i\right), 1\right) \tag{19}$$

(3) Training and testing the Support Vector Machine model: the generated feature subset and parameters are fed into the SVM classifier for training and testing. In this stage, the following equation is used:

$$O_i = S\left(X_{\min}^d + X_{\max}^d\right) - X_i \tag{20}$$

The fitness value is calculated for each individual whale. The fitness value reflects the performance of the SVM model under the current parameter configuration.

(4) Save the optimal individual whale locations: compare the fitness values of different individuals and save the corresponding whale locations with the smallest fitness values. These positions represent the parameter and feature subset configurations that perform best in the current iteration.

(5) Updating whale positions: use the improved whale optimization algorithm to update whale positions. This involves updating the parameter and feature subset configurations and computing new fitness values.

(6) Save the optimal solution: Compare the fitness values of different individuals again to obtain the optimal solution for this iteration and save the position of the previously optimal whale individual. This helps to keep track of the best results throughout the optimization process.

(7) Continue iteration or stop: this step is executed when the algorithm runs to the maximum number of iterations or when the stop iteration condition is met. If iteration is continued, skip to step (2). Otherwise, continue to the next step.

(8) Output the optimal parameter combination: output the minimum fitness value and save the position of the optimal whale individual at this time. The obtained parameter combination (C, g) is the optimal configuration. This is a subset of SVM parameters and features optimized by the IWOA algorithm.

(9) Application to Support Vector Machine Model: finally, the optimal parameter combination (C, g) obtained above is applied to the Support Vector Machine model for training and testing.

III. B. Experimental design and validation methods

All the data is divided into training set and test set in the ratio of 8:2. The training set is used for model training and tuning, and the test set is used for model performance evaluation. In addition, in order to prevent model overfitting, a 5-fold cross-validation method is introduced in this study.

The performance of the model is evaluated by the following metrics. Accuracy: the ratio of the number of correctly predicted samples to the total number of samples. Precision: the proportion of samples predicted to be in the positive category that are actually in the positive category. Recall: the proportion of actual positive samples that are predicted to be positive.F1-score: the reconciled average of precision and recall, used to comprehensively evaluate the model performance. Area Under the Subject Operating Characteristic Curve (AUC): used to assess the classification performance of the model in dealing with different thresholds.

To verify the robustness and stability of the model, cross-validation and leave-one-out validation methods are used. Cross-validation reduces the fluctuation of the model due to the difference in data distribution by dividing the dataset and training the model several times. The leave-one-out method can evaluate the generalization ability of the model with a small number of samples.

III. C. Analysis and Discussion of Results

III. C. 1) Model performance analysis

In order to evaluate the performance of various types of models in sports injury prediction, this study conducted experiments for support vector machine (SVM), random forest RF, decision tree DT, logistic regression LR, convolutional neural network (CNN), long-short-term memory network (LSTM), and its sports injury prediction model (IWOA-SVM) in this paper and analyzed the performance of the models from several perspectives (including accuracy, precision, recall, F1 -Score and AUC value) to comprehensively analyze the performance of the models. The specific performance data of each model is shown in Table 3, and the data visualization results are shown in Figure 1.

(1) Model accuracy analysis

The accuracy rates of CNN and LSTM in multidimensional motion data processing are 91.65% and 90.46% respectively, showing the advantages of deep learning models. Random forest model performs the best among traditional methods with 87.41% accuracy, attributed to its ability to handle nonlinearity and feature interaction. The sports injury prediction model (IWOA-SVM) in this paper has an accuracy of 93.92%, which is better than a single model.

(2) Model Precision and Recall Analysis

The CNN and LSTM model precision rates are 89.15% and 89.73%, respectively, and the IWOA-SVM model improves to 92.79%. In terms of recall rate, the CNN model is the highest at 90.67%, and the IWOA-SVM model reaches 93.52%, showing the advantage of the prediction model in this paper in recognizing athletes' sports injuries.

(3) F1-Score and AUC analysis

F1-Score measures the comprehensive performance of the model, and the F1-Score of the IWOA-SVM model is 92.47%, which is better than the other models. AUC measures the discriminative ability of the model, and the AUC of the IWOA-SVM model is 95.45%, which is significantly higher than that of the other models, showing its high ability to discriminate between positive and negative samples under different thresholds.

The results showed that the IWOA-SVM model performed the best on all indicators, demonstrating its reliability in sports injury prediction, which can be used for the prediction and prevention of sports injuries in athletes.

Table 3: Specific performance data of each model

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
SVM	82.43	80.04	80.09	80.05	85.58
RF	87.41	85.42	84.57	84.87	89.82
DT	80.05	77.88	79.14	78.58	82.75
LR	81.28	80.08	80.29	80.83	83.59
CNN	91.65	89.15	90.51	89.83	94.12
LSTM	90.46	89.73	90.67	90.41	93.62
IWOA-SVM	93.92	92.79	93.52	92.47	95.45

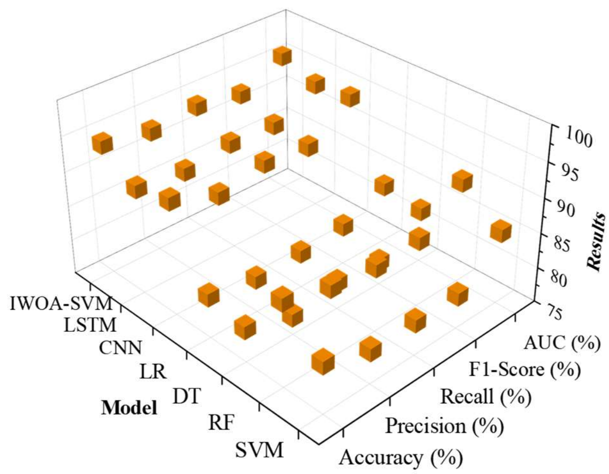

Figure 1: Data visualization results

III. C. 2) Characteristic importance analysis

The results of the importance of each feature are shown in Figure 2. The feature importance analysis highlights athlete height, weight and training intensity as the most critical factors for sports injury prediction, with influence weights all exceeding 0.24. This result emphasizes the importance of training intensity and reveals the significant effect of height and weight and their combination with training intensity to form a high-risk scenario.

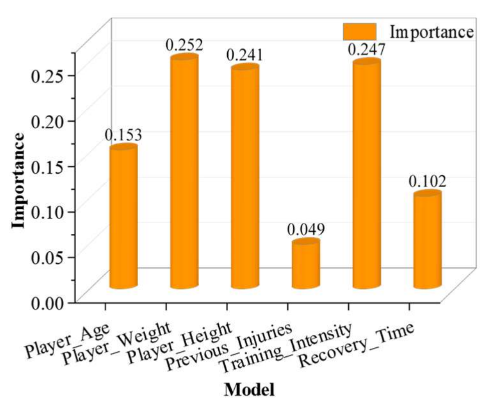

Figure 2: The importance of each feature

The interaction effect of raw characteristics with the inclusion of height and weight is shown in Figure 3. The raw analysis indicated that height and weight had a 50.4%h 49.6% equal effect on training intensity. However, with the addition of the height-weight cross term, the importance of the two was adjusted to 39.5% and 31.3%, respectively, with a cross term importance of 29.2%, indicating that the interaction effect of the two had a significant impact on training intensity and contributed to a more balanced distribution of feature importance.

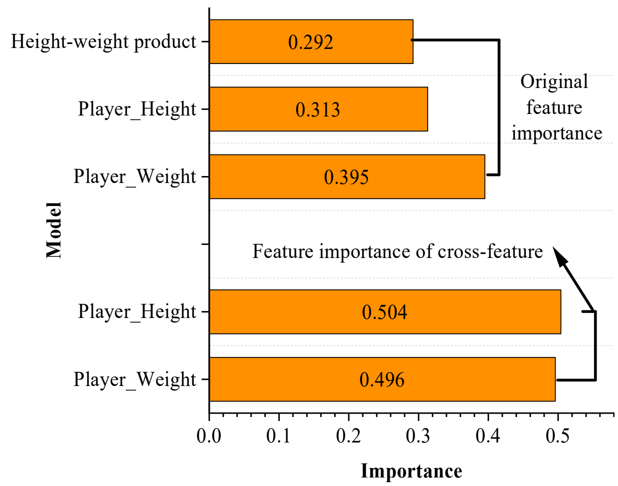

Figure 3: The interaction effect of original feature and adding height weight

IV. Rehabilitation strategies for sports injuries in athletes

The occurrence of sports injuries is closely related to sports programs, training program arrangements, sports environment, the sportsman's own condition and technical skill level. This paper uses big data computational analysis methods to construct a sports injury prediction model for the prevention of sports injuries in athletes, based on which the rehabilitation treatment strategies for sports injuries are discussed.

IV. A. Strain

Flexibility should be practiced within the limits of strength, and if an injury has occurred, it should be treated according to the specific circumstances of the injury. For a small part of the muscle fiber tear students, immediately stop training to give cold compresses, local pressure bandage and elevate the injured part, for most of the muscle fiber tear students, after pressure bandage should be immediately sent to the hospital for further detailed treatment. When there is a mild ligament injury, the treatment mainly focuses on pain relief and swelling reduction. The key to the treatment of moderate ligament injuries is to make the injured leg braked, avoiding the position of pulling at the ligament, in order to accelerate the healing, also can be used to fix the injury with elastic bandage. For severe injuries, the broken ends of the ligaments should be well aligned in the early stages of the injury. When ligament injuries have already occurred, secondary injuries should be avoided as much as possible.

IV. B. Abrasions and strains

When there are abrasions, you should immediately stop training, check the injured part, if the trauma is not big, you can first use alcohol or iodine to disinfect the wound, put a band-aid on it, and carry out slow-motion activities on the injured part, if there is no big problem, just a slight abrasion, then continue to carry out training. If the wound is large, use alcohol or iodine to disinfect the wound first, and then use clean gauze to bandage up the wound. If the injured area is red, swollen and painful, use ice or cold towels to apply cold compresses to prevent further bleeding. Strain injuries occur mainly as a result of prolonged stress on the area. When pain occurs in the area where the strain injury occurs, the amount of activity in the area should be reduced, appropriate local massage can also be carried out, local physical therapy, acupuncture, massage, and topical application of traditional Chinese medicine.

IV. C. Joint sprains

When there is swelling in the wrist joint but the pressure and pain are not obvious, it can be treated by pressing, moistening, kneading and pinching, and then flexing and stretching the wrist to unblock the meridians. When the injury is more serious, seek medical attention immediately. When mild pain occurs in the ankle joint, exercise should be stopped immediately to minimize secondary injuries. Minor sprains will improve in a few days and usually heal on their own. When the pain is significant and the ankle joint cannot be stressed, immediate medical attention should be sought. Ultrasound, massage, acupuncture, and anti-inflammatory medication should be used at that stage. When pain occurs in the lower back, exercise should be stopped immediately and can be treated externally, such as topical rubbing of oil of all flowers or topical application of pain relieving ointment. When the injury is more severe, other treatments such as prescription anti-inflammatory medications, muscle relaxants, physical therapy, osteopathic massage, and acupuncture can also be incorporated for recovery. Neck injuries are mainly caused by the back and forth rolls in skill maneuvers. Therefore, physical therapy, such as massage, heat, electrical stimulation and ultrasound, can be used to reduce muscle spasm at the site of injury and help to regain a pain-free range of motion in the neck and strengthen the neck muscles.

IV. D. Contusions

When a contusion occurs, the athlete should stop training immediately. If the skin is found to be bleeding, immediately disinfect the wound with alcohol or iodine, followed by sprinkling some topical anti-inflammatory powder and bandaging it up with clean gauze. If you find that the injured part is red, swollen and painful, you can first use ice or cold towels to apply a local cold compress to prevent further bleeding. In 24 ~ 48 hours after the switch to hot towels on the injured area, for blood, pain, swelling, but also on the injured part of the rubbing, conditions can also be coated with alcohol or turpentine in the injured area. After treatment to reduce the injury, we should move the injured joints or muscles in time, so that it restores the function, to avoid the injured parts of the joints after the good activity is not flexible, or even muscle atrophy phenomenon. It is worth noting that contusion generally occurs in the double bar, single bar action, so more attention should be paid when practicing the action.

V. Conclusion

The support vector machine model based on the improved whale optimization algorithm shows excellent performance in sports injury prediction. The experimental results show that the IWOA-SVM model significantly outperforms traditional machine learning methods in key indicators such as accuracy, precision, recall and AUC value, in which the AUC value reaches 95.45%, which reflects the model's strong ability to distinguish between positive and negative samples under different thresholds. The feature importance analysis reveals the status of athletes' height, weight and training intensity as the most critical factors for sports injury prediction, with the influence weights of the three exceeding 0.24, which provides a scientific basis for the development of personalized training programs. By introducing the height-weight interaction effect, the model further optimized the distribution of feature importance, in which the importance of the cross term reached 29.2%, confirming the important influence of multifactor synergy on injury risk. Correlation analysis showed that training intensity was the only single factor with significant linear association with injury likelihood, providing an important reference for training program development. The constructed rehabilitation treatment strategy covers four major injury categories: strains, abrasions, joint sprains and contusions, which provides athletes with systematic post-injury recovery guidance and realizes the whole process management from prediction and prevention to rehabilitation treatment.

References

[1] Zhang, X. (2023). Control of Joint Injuries in Aerobic Gymnasts. Revista Brasileira de Medicina do Esporte, 29, e2022_0522.

[2] McKenzie, A., Larequi, S. A., Hams, A., Headrick, J., Whiteley, R., & Duhig, S. (2023). Shoulder pain and injury risk factors in competitive swimmers: a systematic review. Scandinavian journal of medicine & science in sports, 33(12), 2396-2412.

[3] Soligard, T., Steffen, K., Palmer, D., Alonso, J. M., Bahr, R., Lopes, A. D., ... & Engebretsen, L. (2017). Sports injury and illness incidence in the Rio de Janeiro 2016 Olympic summer games: a prospective study of 11274 athletes from 207 countries. British journal of sports medicine, 51(17), 1265-1271.

[4] Dacus, L., Castagno, C., Castagno, C., Gontre, G., & Weiss, W. M. (2023). Impact of traumatic sports injury on an athlete's psychological wellbeing, adherence to sport and athletic identity. Journal of Sports Medicine and Therapy, 8(3), 036-046.

[5] Goddard, K., Roberts, C. M., Byron-Daniel, J., & Woodford, L. (2021). Psychological factors involved in adherence to sport injury rehabilitation: a systematic review. International Review of Sport and Exercise Psychology, 14(1), 51-73.

[6] Talibov, A. K., Gushchina, N. V., Lyashenko, O. V., & Dmitrieva, E. S. (2019). Dynamics of athletes' health and physical development in long-term training process. Theory and Practice of Physical Culture, (4), 35-35.

[7] Walters, B. K., Read, C. R., & Estes, A. R. (2017). The effects of resistance training, overtraining, and early specialization on youth athlete injury and development. The Journal of sports medicine and physical fitness, 58(9), 1339-1348.

[8] Liu, G., Sun, H., Yang, T., & Sun, M. (2023). Rehabilitation of Chronic Exercise Injuries: Overtraining, Muscle Imbalances, and Prevention. Studies in Sports Science and Physical Education, 1(2), 54-59.

[9] Armstrong, L. E., Bergeron, M. F., Lee, E. C., Mershon, J. E., & Armstrong, E. M. (2022). Overtraining syndrome as a complex systems phenomenon. Frontiers in network physiology, 1, 794392.

[10] Eklund, R. C., & DeFreese, J. D. (2021). Fatigue, overtraining, and burnout. Sport, Exercise and Performance Psychology: Research Directions To Advance the Field, 278.

[11] Karpov, V. Y., Zavalishina, S. Y., Dorontsev, A. V., Skorosov, K. K., & Ivanov, D. A. (2019). Physiological Basis of Physical Rehabilitation of Athletes after Ankle Injuries. Indian Journal of Public Health Research & Development, 10(10).

[12] Buckthorpe, M. (2019). Optimising the late-stage rehabilitation and return-to-sport training and testing process after ACL reconstruction. Sports Medicine, 49(7), 1043-1058.

[13] Fonseca, S. T., Souza, T. R., Verhagen, E., Van Emmerik, R., Bittencourt, N. F., Mendonça, L. D., ... & Ocarino, J. M. (2020). Sports injury forecasting and complexity: a synergetic approach. Sports medicine, 50, 1757-1770.

[14] Guelmami, N., Fekih-Romdhane, F., Mechraoui, O., & Bragazzi, N. L. (2023). Injury prevention, optimized training and rehabilitation: how is AI reshaping the field of sports medicine. New Asian Journal of Medicine, 1(1), 30-34.

[15] Whittaker, J. L., Booysen, N., De La Motte, S., Dennett, L., Lewis, C. L., Wilson, D., ... & Stokes, M. (2017). Predicting sport and occupational lower extremity injury risk through movement quality screening: a systematic review. British journal of sports medicine, 51(7), 580-585.

[16] Rhon, D. I., Teyhen, D. S., Collins, G. S., & Bullock, G. S. (2022). Predictive models for musculoskeletal injury risk: why statistical approach makes all the difference. BMJ Open Sport & Exercise Medicine, 8(4), e001388.

[17] Zadeh, A., Taylor, D., Bertsos, M., Tillman, T., Nosoudi, N., & Bruce, S. (2021). Predicting sports injuries with wearable technology and data analysis. Information Systems Frontiers, 23, 1023-1037.

[18] Ota, S., & Kimura, M. (2023). Statistical injury prediction for professional sumo wrestlers: Modeling and perspectives. PLoS one, 18(3), e0283242.

[19] Luu, B. C., Wright, A. L., Haeberle, H. S., Karnuta, J. M., Schickendantz, M. S., Makhni, E. C., ... & Ramkumar, P. N. (2020). Machine learning outperforms logistic regression analysis to predict next-season NHL player injury: an analysis of 2322 players from 2007 to 2017. Orthopaedic journal of sports medicine, 8(9), 2325967120953404.

[20] Zhu, D., Zhang, H., Sun, Y., & Qi, H. (2021). Injury risk prediction of aerobics athletes based on big data and computer vision. Scientific Programming, 2021(1), 5526971.

[21] Dong, X., & Sun, Y. (2024). Injury Risk Prediction and Prevention Algorithm for Athletes Based on Data Mining. Journal of Electrical Systems, 20, 1717-1728.