English Translation Ability Assessment Based on Neural Network Algorithm of Particle Swarm Optimization

Mu Yuan

Henan Mechanical and Electrical Vocational College, Xinzheng, Henan, China

504911735@qq.com

Keywords: College Students' Employment Problems, Quality Evaluation, Hierarchical Analysis Method, Neural Network, Index Weights, Evaluation Indexes

Abstract: The combination of particle swarm optimization algorithm and neural network, which has emerged in current years, can enhance the capability of global search for optimality, and enhance the speed of convergence. The unity of particle swarm algorithm and neural network is suitable for English teaching. The trained particle swarm optimization neural network pattern is utilized to analyze the correctness of student' English translation ability by learning and preparing the extracted students' translation samples, which helps teachers to estimate the students' translation capability level and offer a reference for the following teaching. The model is based the math pattern of the particle swarm optimization algorithm and the foundational principles of the artificial neural network framework, and the study capability analysis pattern is proposed to identify the topology of the neural network and the number of nodes in the hidden layer of the model. The results of the case application show that the research framework could improve English translation teaching ability and teaching and studying.

1. Introduction

China has the maximal number of English learners all over the world. Whereas, because of the variability of coaches and regional diversity in teaching, Chinese English learners receive a wide range of English education levels. Artificial intelligence-based education and teaching models are an excellent way to solve this problem [1-2]. Artificial intelligence systems can achieve a certain level of intelligence by learning a large amount of collected sample data; for example, in the fields of image recognition, driver lessness, and speech recognition, artificial intelligence has achieved excellent results, and in some respects, especially English translation, it can completely replace human work [3-4]. In recent years, AI-based English translation models have been the industry's focus of the industry. Hameed et al [5] studied an intelligent system that collects and learns a good deal of information about students' studying status, individuality, as well as age to build an analytical model of learning ability, and to develop a model of learning ability.

They utilized the pattern to analyze the features of humans in English translation and to assist English translation algorithms. With the growth of information technique, the kind of English translation method could avoid the difference in education caused by the variability of English learning resources to a certain extent [6]. Hence, it is necessary to study English translation algorithms. In actual English learning environments, Bahamian [7] constructed a database of listening resources based on artificial intelligence, it could mechanically assign listening resources and recognize contextual interaction, thus improving the translation correctness. Moradi et al. [8] used the big data treating and computing power of the cloud platform to introduce artificial intelligence into translation, and the cloud platform approach could trace the translation of human translators and understand the translation features of every linguistic scene in good time and correct way to quantify the translation results [9-10]. Based on this, we explore whether artificial intelligence can be used in English translation activities and accelerate the training of neural networks with particle swarm optimization algorithms to converge faster. The method is tested with actual teaching sample data to check the possibility of the approach.

  DOI: 10.25236/ichess.2023.013

2. Analytical study of particle swarm optimization algorithm

2.1. Math model of actual swarm optimization algorithm

Hypothetical particle coordinates of particle I in the swarm are $X_i = (x_{i1}, x_{i2},..., x_{in})$, the optimal position is recorded as p_1, the greast-recorded position of all particles is p2, and the particle's moving speed is $V_i = (v_{i1}, v_{i2},..., v_{id})$. In the particle study procedure, the trajectories of the particles in space are displayed in Eqs. (1) and (2) for each iteration.

$$v_{id}^{k+1} = v_{id}^{k} + c_1 rand()(p_1^k - x_{id}^k) + c_2 rand()(p_2^k - x_{id}^k) \tag{1}$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \tag{2}$$

In which: c_1 and $c2$ are acceleration constants, separately, there are designed to make particles shift faster to the greatest position and the greatest position of overall particles; and() is a stochastic number [0, 1]; the maximal value of the particle speed is $V_{\max}$. While the speed of the particle studies the maximal value $V_{\max}$, the rate does not enhance, yet remains constant. This purpose of a scene the maximal value of particle speed is to enhance the studied correctness during the whole study procedure. While the particle speed is higher, the particle will loss the best deal of the current solution space; while the particle speed is lower, the particle will drop into the regional best solution. Hence, the scene is essential.

The original portion is the basic speed motion of the particle; it demonstrates the motion state of the particle and will stay unaltered without any distractions. The next potion is the cognitive capability of the particle, a procedure that could imitate the mental behaviour of a bird. The third portion is information sharing, in which the particles interact with each other during the optimization process; it may make the overall particle population co-evolve, reflecting the sociological features of particles.

2.2. Process of particle swarm optimization algorithm

Large optimization algorithms are based on gradient knowledge, but particle swarm optimization does not command gradient information. In the actual optimization process, although many estimation functions are required to identify the adaptability of the particles, there are still numerous evident benefits over general evolutionary algorithms.

In particle swarm optimization, the quality of the overall result of the problem is not affected by the individuals, thus providing a high degree of robustness. Additionally, the knowledge interchange between the personnel in the swarm isn't direct; therefore, assuring the scalability of the system is good. In resolving the swarm optimization algorithm, the distributed processing pattern could be used to improve the overall solution efficiency by coordinating multiple processors for parallel computation. Particle swarm optimization algorithms do not command problem-specific continuity and are more scalable compared with standard intelligent algorithms.

Higher scalabilities compared with standard intelligent algorithms. The typical flow of the particle swarm optimization algorithm is displayed in Figure 1.

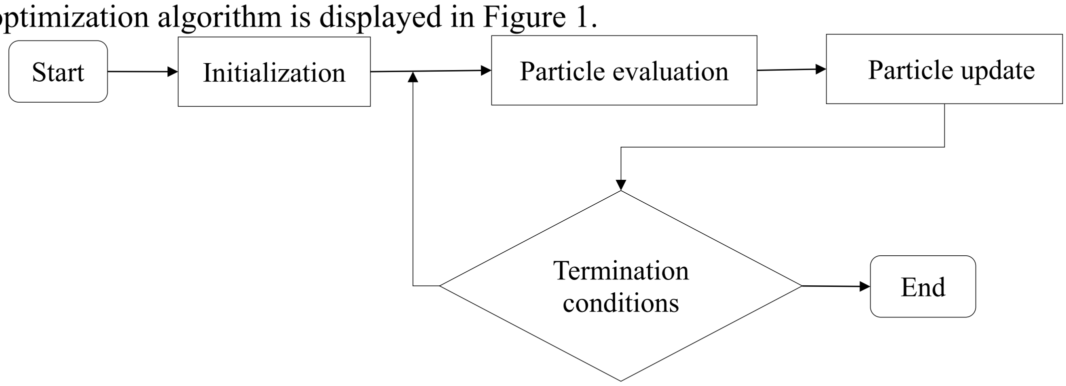

Figure 1 Particle swarm optimization steps

3. Application analysis of English translation based on particle swarm optimization neural network

3.1. Artificial neural network model

A multilayer feedforward neural network pattern is the most generally utilized neural network pattern; it consists of input degree, output degree and hidden degree. The input degree obtains the necessary data from the external, then inputs the acquired data into the neural network for follow-up treatment; The hidden degree stands for the procedure; and the output degree outputs the processing results to the required location.

3.2. English translation application model

In teaching English translation, it is essential to acquire objective information and utilize them for proper analysis. Hence, an application structure is suggested to analyze the studying capability of students in the procedure of teaching English translation, i.e., the studying capability analysis structure, as displayed in Figure 2.

The objective of the studying capability analysis pattern is to analyze many features relative to studying of students in the procedure of studying English translation, and through the analysis, to derive data about student' studying status, and utilize the analysis outcomes to growth-oriented teaching objectives for student, to facilitate the growth of English translation teaching. In the data collection phase, the initial information collection is conducted through questionnaires. In collecting data phase, the raw information is pre-processed to avoid the intervention of useless data in the overall analysis procedure. Since some of the raw data are missing and incomplete, the filling procedure needs to be carried out according to a specific level; then the processed data is entered into the neural network for analysis.

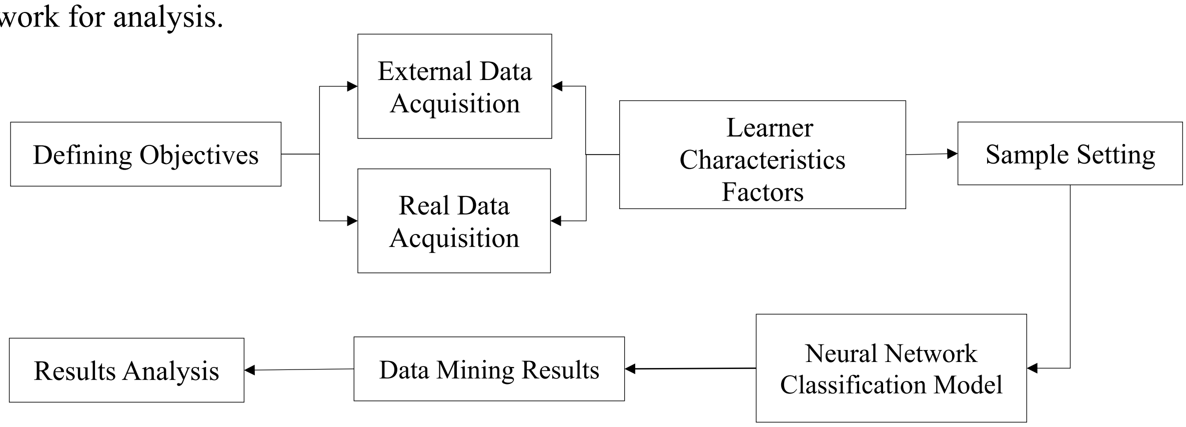

Figure 2 Analysis model of students' learning ability.

In addition, it is essential to identify the topology of the neural network framework, the number of nodes in the input degree, the number of nodes in the output degree, and the number of nodes in the hidden degree. The formulation for calculating the nodes in the hidden degree is displayed in equation (3).

$$J = \sqrt{MN} \tag{3}$$

In equation (3), J is the number of nodes in the hidden degree, M is the number of nodes in the output degree, and N is the number of nodes in the input degree. The relationship between the number of training sessions and the number of nodes in the hidden degree can be received from equation (3). To simplify the connection between the two, the number of training times and the number of nodes in the hidden degree are plotted. As shown in Figure 3, while the number of nodes in the hidden degree of the neural network is 4, the training time of the overall network pattern is the minimum.

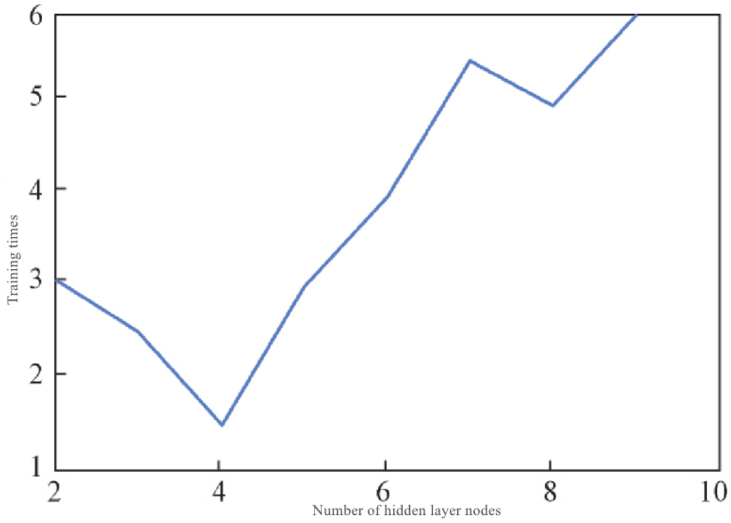

Figure 3 The relationship between the training time and hidden layer node number.

3.3. Application of particle swarm optimization neural network-based solution

The best theory of particle swarm algorithm is referred to the neural network to promote the worldwide optimization capability of the algorithm; the hybrid algorithm utilizes the motion and updating of particles to discover the best result of the neural network in the original phase phase.

Step 1 Pre-processing of feature data.

In the study procedure of artificial neural networks, the data are normalized to avoid the effect of large values on model-based prediction or diagnosis.

Normalization procedure. The data change is restricted to a specific scope, always in (0, 1). Due to the Sigmoid role being utilized as a transformation role for the output degree of the neural network, the Sigmoid transfer role has the characteristic that as x approaches active values or negative infinity, the output value will come 0 or 1. Hence the output unstable ranges from (0, 1). No normalized information, the effect of small-valued neurons on the network could be more minor compared with that of large-valued neurons, hence influencing the training outcomes.

$$x_i = \lambda_1 + (\lambda_2 - \lambda_1)\frac{z_i - z_i^{\min}}{z_i^{\max} - z_i^{\min}} \tag{4}$$

The output value is denormalized, and the denormalization formula is shown in equation (5).

$$z_i = (\frac{x_i - \lambda_1}{\lambda_2 - \lambda_1})(z_i^{\max} - z_i^{\min}) + z_i^{\min} \tag{5}$$

In Eq. (5): x_i is the normalized value; z_i is the denormalized value; λ_1 is the inferior restrict; λ_2 is the higher restrict; $z_i^{\max}$ is the maximal value in the source information; $z_i^{\min}$ is the smallest value in the source information.

Step 2 Set the parameters of the neural network. The BP (backpropagation) neural network is used to estimate the cost module, which has several vital parameters, including the number of hidden nodes, hidden degrees, activation role, studying rate, and momentum coefficient.

(1) Number of hidden nodes: the more hidden nodes, the slower the convergence speed, but the smaller the mistake could be, particularly the "training sample" mistake. However, while the number increases to a certain level, increasing the number will not reduce the error, and the execution time may suddenly increase. Equation (3) calculates the number of nodes in the hidden degree.

(2) Number of hidden degrees: The number of hidden degrees has a major impact on the convergence velocity of the network. Usually, 1 to 2 layers are ideal and can solve most application problems. One layer is used for training.

(3) Activation function: Sigmoid role is utilized as the activation role, and its formulation is

displayed in equation (6).

$$\delta(t) = \frac{1}{1+e^{-t}} \tag{6}$$

In Eq. (6): t is the independent variable of the Sigmoid function.

(4) Learning rate and momentum coefficient: The learning rate has a major influence on the convergence velocity of the network, usually η=0.1~1.0. In general, the learning algorithm will add a momentum coefficient, i.e., adding a certain proportion of the past weight changes to weaken the oscillations in convergence and accelerate the convergence. Set $\eta+\mu$=1.0.

Step 3 Casually cause the actual speed and position of particle swarm Set $X_i = (x_{i1}, x_{i2}, ..., x_{in})$ as the initial position of particle II, set $V_i = (v_{i1}, v_{i2}, ..., v_{id})$ as the initial speed of particle I, and $P_g = \min\{P_0, P_1, ..., P_s\}$ is the optimal position record of all particles, i.e., the optimal local function, and the speed and position vectors of the particle population in the n-dimensional space are randomly generated in (0, 1).

Step 4 Computes the output vector of the forward of the neural network.

The output vector H of the hidden layer is shown in Equation (7) and Equation (8).

$$t_h = \sum_i (a_{ih}) x_i - \theta_h \tag{7}$$

$$H_h = f(t_h) = \frac{1}{1+e^{-t_h}} \tag{8}$$

In which: t_h denotes the weighted input sum of layer h; a_{ih} is the linkage parameter between layer i and degree h; x_i is the activation value of layer i; and θ_h represents the bias of layer h.

The vector Y of the output layer is calculated as shown in equations (9) and (10).

$$t_j = \sum_h (a_{hj}) H_h - \theta_j \tag{9}$$

$$Y_j = f(t_j) = \frac{1}{1+e^{-t_j}} \tag{10}$$

In which: t_j is the input weighted sum of the jth layer; Y_j is the output of the jth layer.

Step 5 Calculates the inverse difference δ.

$$\delta_i = Y_j(1-Y_j)(T_j - Y_j) \tag{11}$$

$$\delta_h = H_h(1-H_h)\sum_j w_{hj}\delta_j \tag{12}$$

Step 6 Calculates the weight matrix Δw and the bias vector $\Delta\theta$ variation.

$$\Delta w_{hj} = -\eta\delta_j H_h \tag{13}$$

$$\Delta\theta_j = -\eta\delta_j \tag{14}$$

The weight matrix and bias vector are updated.

$$w_{hj} = w'_{hj} + \Delta w_{hj} \tag{15}$$

$$\theta_j = \theta'_j + \Delta\theta_j \tag{16}$$

In which: w_{hj}, θ_j denote the new weight matrix and bias vector; w'_{hj} , θ'_j are the original weight matrix and bias vector, separately.

Step 7 Calculates the fitness of every particle. The soundness of every particle is estimated based on objective function for a particular problem, and the fitness role value is compared with the best role value in the memory, and the particle modifies the search speed in the following stage according to the best value in the memory. The sum of squared error (SSE) is shown in equation (17).

$$f(i) = SSE = \sum_{i=1}^{n}(T_i - E_i)^2 \tag{17}$$

In Eq. (17), T_i is the data fitted by the algorithm, and E_i is the original data. Step 8 records the particle position and updates its position velocity vector.

If the particle optimal value is greater over the worldwide optimal value, the worldwide optimal value in the memory is modified, and the position velocity is corrected for each particle to prepare for the next global search.

$$V_i^{k+i} = WV_i^k + c_1 rand()(s_i^{k^*} - s_i^k) + c_2 rand()(s_i^{\#} - s_i^k) \tag{18}$$

$$s_i^{k+1} = s_i^k + V_i^{k+1} \tag{19}$$

In which: c_1, c_2 are learning factors, also known as acceleration constants; W is the inertia factor; rand( ) is a uniform, casual number of [0, 1]; $s_i^{k^*}$ denotes the k*th dimension of the extreme personal value of the i-th variable; $s_i^{\#}$ denotes the i-th size of the worldwide best result; and s_i^k denotes the k-th size of the unique extreme importance of the i-th variable.

Step 9: Apply the neural network algorithm model of particle swarm optimization to verify the influence of English translation teaching.

While the particle swarm attains the worldwide best state, the network training is completed, the average mistake information of the training set is received, and the mistake data of the sample set is tested.

4. Case Application and Result Analysis

A neural network algorithm pattern with particle swarm optimization was applied to identify the effectiveness of teaching English translation. The first step is to collect samples of students' English translation studying features, and then to train the neural network with the particle swarm optimization algorithm in two stages, and the second phase is to assess and check the effectiveness of the pattern.

As a significant part of implementing of the neural network algorithm, the information set plays a crucial role in the dependability of the pattern output. To enhance the quantity of data and improve the generalization capability of the pattern without affecting the validity of the design, data augmentation operations, such as adding noise, were performed to make the data set contain the most critical data. To enhance the quantity of information without influencing the validness of the pattern and to enhance the generalization capability of the framework, information augmentation operations such as adding noise were performed to make the data set contain 4,500 testable data. The data set received by the above procedure may ensure the correctness and difference of the English translation feature samples to the maximum extent.

While training the neural network utilizing the particle swarm optimization algorithm, not only do we need to process the existing data, but we also need to establish many parameters. The prematureness element is set to 0.01, and other parameters are generated randomly. According to the convention of particle swarm algorithm, three different particle numbers, i.e., 6, 12 and 24, are set, and the operational efficiency of the three diverse particle numbers is analyzed. While the number of iterations attains 254, 456 and 283, the mistake between the algorithm output and the training and trial samples are small that no better result has emerged for some time. Hence, the consequence is the best solution, and then the iteration is stopped, and the related infomration are recorded, as displayed in Table 1.

Table 1 Case consequences of neural network models for particle swarm optimization.

Particles number	Iterations number	Optimal particles number	The average error of training samples	Evaluation error of test samples
6	254	4	0.03	0.22
12	456	8	0.27	0.30
24	283	5	0.25	0.04

According to the practice of English translation, it could be seen through Table 1 the neural network trained via the particle swarm optimization algorithm can obtain the best solutions for diverse particle populations with small error values. According to the mentioned studying analysis framework, the trained particle swarm optimization neural network framework could be utilized to analyze the correct level of student' English translation capability and assist teachers in evaluating the extent of students' translation capability for further teaching reference.

5. Conclusion

Based on the neural network pattern of particle swarm optimization, we proposed a learning ability analysis model. We determined the topology of the neural network and the number of nodes in the hidden degree of the model. Through the analysis of the experimental outcomes, it is discovered the approach could assist teachers evaluate students' translation capability and offer references for student to enhance English translation skills. The proposed method of applying artificial intelligence algorithms to help English teaching could be commonly used to many scopes of English teaching with the further growth of computer technique.

References

[1] Rao P, Jana P K, Banka H. A particle swarm optimization-based energy efficient cluster head selection algorithm for wireless sensor networks[J]. Wireless Networks, 2017, 23(7):2005-2020.

[2] Am bursa F U, Latip R, Abdullah A, et al. A particle swarm optimization and min–max-based workflow scheduling algorithm with QoS satisfaction for service-oriented grids[J]. Journal of Supercomputing, 2017, 73(5):2018-2051.

[3] Guilherme A . AI and education: the importance of teacher and student relations[J]. AI & society, 2019, 34(1):47-54.

[4] Zhang X F, Sui G F . Quantum-Behaved Particle Swarm Optimization Algorithm for Solving Nonlinear Equations[J]. Advanced Materials Research, 2013, 756-759:2926-2931.

[5] Hameed M,Sharqi S S,Yaseen Z M,et al.Application of artificial intelligence techniques in water quality index prediction: a case study in a tropical region, Malaysia[J]. Neural Computing and Applications,2017,28(1):893-905.

[6] Lee J H,Kim J W,Song J Y,et al.A novel memetic algorithm using modified particle swarm optimization and mesh adaptive direct search for PMSM design[J].IEEE Transactions on Magnetics, 2017, 52(3):1-4.

[7] Javad B.Heterogeneous networked cooperative scheduling with anar chic particle swarm optimization[J].IEEE Transactions on Engineering Management,2017,64(2):166-178.

[8] M H,Abedini M.A combination of genetic algorithm and particle swarm optimization for optimal distributed generation location and sizing in distribution systems with fuzzy optimal theory[J]. International Journal of Green Energy,2012,9(7):641-660.

[9] Xiang H,Xiao H,Yuan W.Research optimization on logistics distribution center location based on adaptive particle swarm algorithm[J]. Optik-International Journal for Light and Electron Optics, 2016,127(20):8443-8450.

[10] Xu J,Yao L,Zhao X.A multi-objective chance-constrained network optimal model with random fuzzy coefficients and its application to logistics distribution center location problem[J]. Fuzzy Optimiza tion&Decision Making, 2011, 10(3):255-285.