A Neural Network Prediction Method Based on NNSOA

Yong Zhao, Mengjie Xu, Qianhan Wang and Bin Ma*
Research Centre of Urban Management, Anhui Jianzhu University, Hefei 230601 China

*Corresponding author email: mbzj@ahjzu.edu.cn

Abstract. By using a neural network synchronization optimization algorithm (NNSOA) as a search technique, this paper applies a neural network to forecast the promotion of the employees in a home appliances manufacturing enterprise in Hefei, Anhui province. NNSOA not only optimizes the neural network, but also eliminates the unnecessary weight in the neural network structure. The results prove that a neural network trained by ten times of cross-validation experiment design can make a high-accuracy prediction on the promotion of the employees in the enterprise. The prediction efficiency of predictive variables can be identified by the neural network trained with NNSOA. The study provides more targeted suggestions for enterprises to improve the overall employee development environment.

Keywords: Employee promotion; Neural network; NNSOA; Genetic algorithm; Chinese enterprise.

1. Introduction
The internal promotion of an enterprise mainly undertakes two functions in the organization, one is the incentive function, the other is the selection function[Edward, 2000]. It is an effective way to organizational management and motivate the employees, and it is also a good way to employ and retain people [Baker, Gibbs, & Holmstrom, 1994]. Among them, performance-based promotion is a mechanism of selection according to the current performance, which to a large extent reflects the incentive function, but this may destroy the promotion and selection function and can not effectively equip the post with suitable candidates. Because the results of performance evaluation can only reflect the past performance of employees, can not reflect whether employees have the ability to adapt to higher-level positions. Therefore, the results of performance evaluation will be used as the basis for employee promotion, which will fall into the "Peter trap", that is, management posts at all levels will be occupied by incompetent people. How to effectively identify outstanding employees and promote them in time has become the focus of attention of enterprise managers.

In the fields of science and technology, medicine, agriculture, engineering and education, etc, the neural network (NN) has become an effective forecasting tool[Li & Liu, 2015; Cao, Chen, & Li, 2013; Lin & Chen, 2010; Wang, 2004]. Randall et al. adopted neural network synchronization optimization algorithm (NNSOA), an improved genetic algorithm of neural network, as its search technology, and achieved success in accurately predicting employee turnover [Randall, Shannon, Joanna, & Angela, 2005]. This study modifies the genetic algorithm, improves the ability of the algorithm to generalize untrained data, and enables it to identify the relevant and unrelated input variables (NNSOA), can provide researchers or managers with extra information about problems themselves. This method can see which inputs in the dataset are actually helpful to predict employee turnover. In addition, NNSOA automatically determines the optimal number of hidden nodes contained in the NN architecture. This character can save a lot of time and effort without having to experiment repeatedly.
like other neural network programs to find the optimal architecture. This paper uses NNSOA-based neural network technology to predict the possible promotion of employees, for the enterprise talent selection, talent reserve decision-making reference.

2. Neural Network Synchronization Optimization Algorithm

2.1. Overview of NNSOA

NNSOA is only used to search for the weight of input. It is proved that it is more efficient and effective to use ordinary least square (OLS) to determine the output weight. Different from the back propagation (BP) based on gradient information moving from one point to another, NNSOA searches in multiple directions at the same time, which improves the probability of receiving the global optimal. The NNSOA used in this study is summarized as follows.

By extracting random real values from the uniform distribution [-1, 1] of input weights, the population of 12 solutions is created. This happens only once in the training process, and the output weight is determined by OLS.

In order to allot probabilities for each solution to be redrawn in the next generation, based on the objective function (Equation 1) of the square sum error (SSE) setting study, each member of the current population is assessed. To find a concise solution, each non-zero weight (or active connection) adds the SSE with a penalty value.

\[
\text{MinE} = \sum_{i=1}^{N} (O_i - \hat{O}_i)^2 + C \sqrt{\frac{\sum_{i=1}^{N} (O_i - \hat{O}_i)^2}{N}}
\]  

(1)

In the formula, N is the number of data sets to be observed, O is the dependent variable observation value, \( \hat{O} \) is the neural network estimate, and C is the number of non-zero weight in the network. The penalty value added in the search process is variable, which is equal to the current value of the root mean square error (RMSE). The 12 solutions in the population are evaluated based on the objective function. The probability that the next generation will be drawn is equal to the distance between the target value of the current solution and the worst target value of that generation divided by the sum of all the distances of the current generation.

The solution is then selected from the current population according to the specified probability, and a pool of 12 solutions is created. Repeat this process till the total new generation contains 12 solutions drawn. Only solutions from previous solutions are contained in the new generation. Solutions whose probabilities are higher may appear at least once, while solutions whose probabilities are lower may disappear completely.

Once the combination of the previous generation's solutions is generated, the 12 solutions will be randomly paired to construct six sets of parent solutions. Each pair of solutions is randomly selected at a point where the parent solution will swap weights at that point and generate 12 new or next-generation solutions.

For each randomly extracted weight, when it is lower under 0.05, it would be replaced with the one extracted from the whole weight set. To do so, the whole weight set is searched entirely, thus the ability of the algorithm to get the global solution is enhanced.

For the weights of each generation, drawing a random number first, and if it is less than 0.05, using hard zeros instead of their weights. By doing so, you can determine the weights you do not need when you continue to find the best solution. After executing this operator, the next generation of 12 solutions will be evaluated once more, this loop continues until the maximum build set is reached to 70%. Then the best solution till now will replace entire strings in the present build. Every weight in a string population varies with a small random quantity. As the algebra increases to the maximum value it sets, these random numbers are reduced to any size. Finally, the algorithm terminates on the number of builds specified by the user.

2.2. Hidden Node Search
The number of hidden nodes contained in each neural network is automatically determined as follows. For a neural network, generations or MAXHID defined by users are trained begin with hidden node 1. After each MAXHID generation, the best solution will be saved and an additional hidden node will be added to the NN architecture. The neural network is reinitialized by using different random seeds to draw the initial weight, and then trains the MAXHID. The best solution is replacing the weight of the first solution which included in the new generation. The weights found in the best solution are less than that generated by additional hidden nodes, so they will be set to hard zeros. In this way, what has been learned from former generations can be saved. After completing this process, compare the best solution with that of this architecture. If the best solution is worse than this one, this solution now is regarded the best and save it for future assessment. This course will continue until a hidden node is added and there is no better solution than the best solution. Once this happens, the best solution and its corresponding architecture will be trained using an additional number of user-defined generations or MAXGEN to complete the training process. Although the two solutions may get the same objective function value, their architecture may be different.

3. Experimental Process

This study uses the employee data of Hefei Haier Industrial Park (Hefei Haier) to predict employee promotion through NN of NNSOA training, and then verifies the effectiveness of neural network based on NNSOA in employee promotion prediction.

| Input | Input abbreviation | Input description |
|-------|---------------------|-------------------|
| 1     | AGE                 | Age of employee   |
| 2     | SING                | Single employee   |
| 3     | MARD                | Married employee  |
| 4     | DIVORC              | Divorced single employee |
| 5     | MALE                | Male employee     |
| 6     | FEMALE              | Female employees  |
| 7     | HANNAT              | The Han nationality |
| 8     | OTHERNAT            | Other nationalities |
| 9     | COMM part           | Communist         |
| 10    | OTHER PART          | Members of other parties |
| 11    | NON PART            | Non-party         |
| 12    | DOR                 | Doctor            |
| 13    | MAS                 | Master            |
| 14    | BAC                 | Bachelor          |
| 15    | COL                 | Junior college    |
| 16    | UNDER COL           | Under Junior college |
| 17    | GRADE1              | Management Level 1 |
| 18    | GRADE2              | Management Level 2 |
| 19    | GRADE3              | Management Level 3 |
| 20    | GRADE4              | Management Level 4 |
| 21    | GRADE5              | Management Level 5 |
| 22    | GRADE6              | Management Level 6 |
| 23    | GRADE7              | Management Level 7 |
| 24    | TOTTENURE           | Total Tenure      |
| 25    | COMPTENURE          | Tenure in the company |
| 26    | TENURE              | Tenure in present position |
| 27    | TRAITIM             | Annual training time |
| 28    | ANNUAL RRF          | Annual performance |
| 29    | HEAD APPOINT        | Headquarters appointment |
| 30    | LOCAL RECRUIT       | Local recruitment |
Hefei Haier was built and put into production in 2001. It has three product lines, including digital color TV, air conditioner and washing machine, with 5500 employees. The purpose of this study is to verify the effectiveness of the neural network based on NNSOA in predicting employee promotion. In order to reduce interference, we only select the Enterprise Management Sequence position including the staff of the functional management department and the production management position at or above the team leader as the study sample. The data are used for all employees in the above-mentioned management positions who were on duty from 2006 to 2016.

According to the relevant research literature\[Zhang, Zhou, & Liang, 2016; Devaro Jed, 2006; Zhang, Liu, & Qi, 2013; Blackwell, Brickley, & Weisback, 1994; Mobbs & Raheja, 2008\], the influencing factors of employee promotion were selected as input variables. Including: age, marital status, gender, nationality, political outlook, the highest academic qualifications, the highest academic qualifications, the total length of service, length of service in the unit, post time, rank, training time in that year, annual assessment results, entry channels and other factors. Because some factors contain 2 or more categories, need to be entered multiple times, for example, the ranks in the organization from team leader (management level 7) to park general manager (management level 1), there are 7 levels, grade this factor needs to be entered seven times. Therefore, there are 30 input variables in this study (Table 1). If the employee is promoted in the current year, the output is set to 1, otherwise it is set to 0.

The annual information needs to be collected, as it is desirable to predict whether an employee will be promoted within the year. So, an employee working for five years in the company, he or she would have five records, each of which represents the year of service. A total of 185 management-class employees in accordance with this study, with a total of 653 observation data collected. Of the 653 observation data, 581 observation data had not been promoted in a particular year, and 72 of them had been promoted. Ten times cross-validation was carried out to increase the study preciseness, and 653 observation data were trained 10 times and tested 10 times. First of all, the order of the observed values was randomized, and then the last 52 observed data were taken out and saved into a file for testing, the rest 601 observed values were saved into another file for training. Then 52 observation data are taken out and set at the top of the former data set. Finally, the latest 52 observation results are extracted and preserved as the second file for testing, and the rest 601 observation data are saved as the second file for training. Do the same for nine data-sets. In the 10th data-set, because the observation number cannot be divisible by ten, in order to ensure each observation appears in a set for testing, the number of sets for training and sets for testing has to be changed. The final set includes 601 training observations and 72 testing observations.

In this paper, the percentage of total classification errors and the percentage of type I and type II classification errors are compared. Using 1.5 GHz computers installed Windows XP operating system, it is written in FORTRAN language with Visual Basic interface. Basically NNSOA has only two parameters defined by users, named MAXHID and MAXGEN, which have been mentioned above. For these 10 training sets, use 100 generations or MAXHID = 100 to perform hidden node searches. Once the appropriate structure is found, 1000 generations are performed, that is, MAXGEN = 1000.

The parameters used by the algorithm are same with the NNSOA mentioned above, but does not use the second mutation operator introduced previously, and eliminates any loss of non-zero weight in the solution.

4. Experimental Result

Once 10 training sets have been trained, the solution is tested in their sets respectively. Because the number of those are promoted in one year is far more than the number those are not promoted, a cut-off point needs to be set to consider the distorted data. In all dataset of 447 observations, 35 observations are about employee promotion (encoded 1). Then divide this number by the total observation value and get 0.0783. So any NN estimate value below 0.0783 will be classified as "not promoted" and any value above 0.0783 will be labeled as "promotion". This cut-off point is intuitive for the NNSOA model because it is very simple to process skewed data.
Table 2. Error Rate for 10 NN Run.

| Run | Overall classification error rate (%) | Class I classification error rate (%) | Class II classification error rate (%) |
|-----|---------------------------------------|---------------------------------------|---------------------------------------|
| 1   | 0.00                                  | 0.00                                  | 0.00                                  |
| 2   | 0.00                                  | 0.00                                  | 0.00                                  |
| 3   | 0.00                                  | 0.00                                  | 0.00                                  |
| 4   | 0.00                                  | 0.00                                  | 0.00                                  |
| 5   | 0.00                                  | 0.00                                  | 0.00                                  |
| 6   | 0.00                                  | 0.00                                  | 0.00                                  |
| 7   | 2.33                                  | 0.00                                  | 32.00                                 |
| 8   | 4.25                                  | 2.80                                  | 21.33                                 |
| 9   | 0.00                                  | 0.00                                  | 0.00                                  |
| 10  | 0.00                                  | 0.00                                  | 0.00                                  |
| Average | 0.58                                  | 0.28                                  | 5.33                                  |

The performance of NNSOA training neural network is evaluated by classification error rate. Since the primary goal of prediction is to generate observation estimates that have never been seen in training, this experiment reports results from only 10 test sets for each classification method. The overall percentage of classification errors of each method for all 10 test sets is showed in table 2.

As Table 2 showing, NNSOA has found a very good solution to predict this problem. In 8 out of 10 runs, NNSOA generated a model that categorized all test observations correctly. For each run, the performance of NNSOA in the best error rate is excellent. For operation 2 and operation 10, the NN only performs three times, the error rate is 0.00%, and the running 8 error rate is 4.25%.

This experiment has shown that NNSOA has a good performance in the overall classification, but it has not proven that its model is correct in terms of the error percentage of type I and type II. The first type of error is that the employee is misclassified as a promotion, and in fact, they are not promoted. The second type of error is that the employee is misclassified as unpromoted, and in fact they have been promoted. Table 2 also shows the percentage of errors for type I for the 10 test runs.

The experimental results show that the prediction effectiveness of the neural model is very high, the average error is less than 1%. 92% of the employees have not been promoted within 10 years of the collected data, which can not only predict that their positions have not been promoted, but also explain the good performance of type I errors. From Table 2, you can view the error percentage of type II for mutual verification.

Table 2 also showing, the accuracy of NNSOA prediction is 94.6% (5.33%) for the employees who have been promoted in the current year, and the accuracy of NNSOA prediction is as high as 99% (0.28%) for employees who have not been promoted in the current year, and the effectiveness of the neural network prediction model based on NNSOA is proved. In addition, NNSOA is able to compress the network into a thin architecture while maintaining the high validity of the forecast (Table 3).

Table 3. Optimal number of hidden nodes and non-zero weights.

| Run | Optimal hidden nodes | Optimal non-zero weights |
|-----|----------------------|--------------------------|
| 1   | 2                    | 3                        |
| 2   | 2                    | 5                        |
| 3   | 2                    | 3                        |
| 4   | 2                    | 3                        |
| 5   | 3                    | 5                        |
| 6   | 3                    | 6                        |
| 7   | 3                    | 4                        |
| 8   | 3                    | 6                        |
| 9   | 3                    | 4                        |
| 10  | 2                    | 3                        |
| Average | 2.5                  | 4.2                      |
Intuitively, the opportunity to introduce additional errors in the estimation is also reduced because all NNSOA networks have significantly reduced unwanted weights. This may be why the NNs trained by NNSOA are doing so well in the test observations. The input weights were removed 94% in 10 runs averagely. The percentage decrease per network indicates that there may be several unrelated variables in our data. We need to accurately determine which of the 30 inputs have been deleted from our model later. Furthermore, input variables remain in the model so that we have a higher degree of predictability in the promotion of personnel.

Table 4. Frequency of the relevant variables found in 10 runs.

| Frequency | Relevant variables |
|-----------|-------------------|
| 10        | ANNUPORF, TRAITIM |
| 5         | DOR, MAS, BAC, COL, UNDERCOL, |
| 4         | COMMPART, OTHERPART, NONPART, |
| 3         | TENURE, COMPTENURE |
| 0         | AGE, SING, MARD, DIVORC, HANNAT, OTHERNAT, GRADE1, GRADE2, GRADE3, GRADE4, GRADE5, GRADE6, GRADE7, TOTTENURE, HEADAPPOINT, LOCALRECRUIT, MALE, FEMALE |

Table 4 lists at least one input variable with non-zero weights that are used to connect to hidden nodes. The list is ordered according to the frequency of the number of times that a particular input variable has at least one non-zero weights in the final solution. If the weights of a particular network and input are all set to zero, we can ascertain that the input or variable is unnecessary for the network to predict. It can be seen from the table that the annual assessment results (ANNUPORF) and the annual training time (TRAITIM) of the variable category were relevant in all 10 runs, and the level of education (DOR, MAS, BAC, COL, UNDERCOL,) was related 5 out of 10 operations. These predictors are consistent with the conclusions of the relevant literature [Liao, Jiang, Wen, & Wang, 2015; Zhang, Zhou, & Liang, 2017; Alberto & Pedro, 2006; Matthew, 2011], and are also consistent with the intuitive cognition of most people. The political landscape (COMMOART, OTHERPART, NONPART) was found to be related 4 out of 10 times, and this is a very Chinese discovery. The post-holding time of the post (TEAVE) and the length of service of the unit (COMPETTE) was related 3 out of 10 times, which is consistent with the working experience in the career theory on the promotion of career development [Gibbons & Murphy, 1992; Tremblay & Toulouse, 1995].

![Figure 1. Sensitivity analysis of variables.](image-url)
Sensitivity analysis provides a deeper understanding of the extent to which the actual estimate can be changed by a particular input. In order to achieve this, a set for testing is constructed in this experiment, which will give an indication of the degree and direction of variable change. To know a particular input, when its value changes from a minimum to a maximum, how to change the estimate, a test set is built in the way below. For each of the 30 inputs, two manual observations were carried out. To do this so, the minimum value in the whole dataset is put in one observation and the maximum one is put in another observation. The other 29 input values for these two observations are set to the mean of the whole data-set. Make two observations with the help of the network, we can see the two estimates generated by NN, see the extent it has changed, and in which direction the change is. Figure 1 shows that as the annual performance, annual training time, length of service in the unit, length of service in the post, education level and political outlook increase, the estimated value will increase. This is meaningful, but it should be pointed out that interaction between variables in this analysis is not be taken into account. That means the analysis is only a guide to determine trends.

5. Conclusions

In the training process of employee promotion problem, NNSOA not only optimizes the neural network, but also eliminates the unnecessary weight in the neural network structure. This results in a compact NN architecture that performs well when testing data sets. By reducing the structure of neural networks, the generalization ability may be improved, because the lack of weights forces the algorithm to develop general rules to distinguish input patterns, rather than memory special cases. Another advantage of identifying weights that are not needed in the solution is that you can identify unrelated variables in the neural network model. This recognition provides researchers with more information about problem behavior. Because unnecessary weights are eliminated from the model, the neural network runs efficiently. In addition, the NNSOA for implicit node search has been modified in an additional way, allowing the researcher to essentially insert the data set and make it converge to the final solution without the tedious task of determining the implied node's trial-and-error method. The research shows that the neural network based on NNSOA performs well in the prediction of personnel promotion, and the research results can provide effective help for enterprises to establish reserve talent pool, identify outstanding employees and make employee promotion decisions.

Acknowledgements

This work was supported by Humanities and Social Sciences Project of Anhui Education Department (SK2019A0632; SK2019JD05).

References

[1] Lazear Edward P 2000 Performance pay and productivity The American Economic Review vol 90 pp 1346-1361
[2] Baker George Michael Gibbs and Bengt Holmstrom 1994 The Internal economics of the firm: Evidence from personnel data The Quarterly Journal of Economics vol 109 pp 881-919
[3] Li Jingwen and Liu Gang 2015 The prediction model of the boom based on the BP neural network Statistics & Decision vol 43 pp 70-74
[4] Cao Baoyu Chen Guoan and Li Wei 2013 Multi-rotor fault type diagnosis based on neural network Modern Manufacturing Engineering vol 5 pp 126-130
[5] Lin Meiqin and Chen Yanfeng 2010 Fault diagnosis of analog circuits based on wavelet packet analysis and neural network Science, Technology and Engineering vol 10 pp 1996-1998
[6] Wang Zongjun 2004 The application of neural network in urban evaluation Exploration of Nature vol 23 pp 48-56
[7] Randall S Sextona Shannon McMurtreyb Joanna O Michalopoulasc and Angela M Smith 2005 Employee turnover: A neural network solution Computers & Operations Research vol 32 pp 2635-2651
[8] Zhang Hong Zhou Li-an and Liang Jianzhang 2016 Promotions within the Firm: Empirical evidence from a personnel data Management World vol 4 pp 127-188
[9] DeVaro Jed 2006 Internal promotion competitions in firms The RAND Journal of Economics vol 37 pp 521-542
[10] Zhang Zhaoguo Liu Yawei and Qi Xiaolin 2013 Research on management background characteristics, promotion incentive and over-investment Nankai Business Review vol 16 pp 32-42
[11] Blackwell D W Brickley J A Weisback M S 1994 Accounting information and internal performance evaluation: Evidence from Texas banks Journal of Accounting and Economics vol 17 pp 331-358
[12] Mobbs S Raheja C G 2008 Executive promotions: Compensation,CEO influence and firm valuation Working Paper
[13] Liao Jianqiao Jiang Zhensi Wen Xing and Wang Xingdong 2015 Promotion opportunity stagnation and job performance: The mediating-role of intrinsic motivation and perceived insider status Industrial Engineering and Management vol 20 pp 15-21
[14] Zhang Hong Zhou Li-an and Liang Jianzhang 2017 Insider advantage: Empirical evidence from firm’s personnel data Management World vol 12 pp 141-188
[15] Bayo-Moriones Alberto and Pedro Ortín-Ángel 2006 Internal promotion versus external recruitment in industrial plants in Spain Industrial and Labor Relations Review vol 59 pp 451-470
[16] Bidwell Matthew 2011 Paying more to get less: The effects of external hiring versus internal mobility Administrative Science Quarterly vol 56 pp 369-407
[17] Gibbons R and Murphy K J 1992 Optimal incentive contracts in the presence of career concerns: Theory and evidence Journal of Political Economy vol 100 pp 468-505
[18] Tremblay M Roger and Toulouse J M 1995 Career plateau and work attitudes: An empirical study of managers Human Relations vol 48 pp 221-237