DOI:10.18686/ahe.v7i34.12137

# Non-linear Mapping Method for Student Internship Evaluation Indicators Based on Deep Learning Algorithms

**Yi Zhan, Cennie Wang**
Chengdu University,Chengdu,610106,China

**Abstract:** This study presents a non-linear mapping method for internship evaluation indicators based on deep learning algorithms. Deploying deep learning algorithms to process student internship evaluation indicators and utilizing a Softmax classifier for feature classification, a non-linear mapping is employed to establish the relationship between evaluation indicator features and evaluation levels.
**Keywords:** Deep learning algorithm; Softmax classifier; Non-linear mapping method; Feature classification

## 1. Introduction

Traditional higher education internship evaluations are qualitative, subjective, and inefficient, with significant manual effort and time costs. A proposed solution is a deep learning-based non-linear mapping method to streamline and objectify student assessment.

## 2. Based on the deep learning algorithm, student internship evaluation indicators selection

The deep learning algorithm (ZHOU Yue et al., 2020) is utilized to obtain the student internship evaluation indicators.

### 2.1 Convolutional Layer

The student internship evaluation indicators are used as input samples in a convolutional neural network, denoted as $X=(x_1,x_2,\ldots,x_n)$, where $x_i\in R^d$ represents the d-dimensional vector of the i-th word in the evaluation indicator $X$. The quantity of text words in the evaluation indicator samples is denoted as $n$, making the indicator sample $x\in R^{n\times d}$. The number of words in the embedded convolution kernel window is denoted as $l$, represented by $m\in R^{l\times d}$. The convolution kernel window needs to slide vertically in the convolution layer. For all positions i in the evaluation indicator sample $X$, there exist $k$ window vectors $w_i$.

$$w_i=\left[x_i,x_{i+1},\cdots x_{i+k-1}\right] \tag{1}$$

Setting up the feature map $s\in R^{n-k+d}$, $s=[s_1,s_2,\ldots,s_{n-k+1}]$, which can be obtained by convolutional vertical sliding of the convolution window, each slide of the window can be described by the following equation:

$$s_i=f\left(w_i*m+b\right) \tag{2}$$

In the equation, $f$ represents the activation function; $b$ represents the bias vector; * denotes the operation of multiplication, and $m$ represents the number of feature maps. Through the embedded operation of the convolution window, $m$ feature maps are obtained in the convolution layer, resulting in the input feature $W=[W_1,W_2,\ldots,W_m]$.

### 2.2 Pooling Layer

The pooling layer can be divided into maximum pooling and average pooling. The advantage of average pooling is to reduce computation, decrease parameters, and data volume.

### 2.3 Gated Recurrent Unit (GRU) Layer

In LSTM, there exists GRU, which introduces connections into simple nodes in the hidden layer and regulates the output of hidden neurons in the convolutional neural network through recurrent units. GRU consists of reset gates and update gates. For time $t$, the update gate $z_t$ is computed using the following equation:

$$z_t=\sigma\left(W^{(z)}x_t+U^{(z)}h_{t-1}\right) \tag{3}$$

The equation contains the following notations: $\sigma$ represents the standardization factor; $U^{(z)}$ denotes the time step; $W^{(z)}$ describes the update feature; $h_{t-1}$ represents the information present at time step $t$-1; and $x_t$ represents the student internship evaluation indicator sample at time t. The amount of information to be forgotten is determined by the reset gate:

$$r_t = \sigma\left(W^{(r)}x_t + U^{(r)}h_{t-1}\right) \tag{4}$$

The equation contains the following notations: $W^{(r)}$ describes the reset feature, and $U^{(r)}$ represents the reset step. Using the reset gate to store new memory content $h_t^{'}$:

$$h_t^{'} = \tanh\left(Wx_t + r_t \odot Uh_{t-1}\right) \tag{5}$$

Describing the process of gathering information from the previous time step $h_{t-1}$ and the current memory content $h_t^{'}$ can be achieved through the following equation:

$$h_t = z_t \odot h_{t-1} + \left(1 - z_t\right) \odot h^{'} \tag{6}$$

### 2.4 Fully Connected Layer

The features of student internship evaluation indicators are obtained through the fully connected layer and then input into the classifier to obtain the student internship evaluation metrics. In the fully connected layer, nodes are interconnected to extract the features of student internship evaluation indicators, and the obtained features are classified using a Softmax classifier.

$$o(i) = \frac{\exp\left(\theta_i^T x\right)}{\sum_{k=1}^{K}\exp\left(\theta_i^T x\right)} \tag{7}$$

The equation contains the following notations: $K$ represents the number of categories; $\theta_i^T x$ denotes the input to the classifier. The student internship evaluation indicators are obtained based on the classification results.

## 3. Nonlinear Mapping Method for Evaluating Student Internship Evaluation Indicators

The non-linear mapping of student internship evaluation indicators (XIONG Ning et al., 2020) is performed to establish the relationship between student internship evaluation indicators and evaluation levels. Using $D$ to represent the dataset, its expression is as follows:

$$D = \left\{\left(x_1, y_1\right), \left(x_2, y_2\right), \ldots, \left(x_i, y_i\right)\right\} \tag{8}$$

The equation contains the following notations: $x_i$ represents the training data, and $y_i$ is the class label corresponding to $x_i$. Through the non-linear function $\varphi(x){:}R^n \rightarrow RN$, the input vector is mapped in an $N$-dimensional space, and the expression of the non-linear function is as follows:

$$\varphi(x) = \left[\varphi_1(x), \varphi_2(x), \ldots, \varphi_N(x)\right]^T \tag{9}$$

Based on the mapping result, the network output $y(x)$ is obtained:

$$y(x) = sign\left[\varphi^T(x)\omega + b\right] \tag{10}$$

In the equation, $b$ represents a scalar constant, and $\omega$ describes the weight output vector. Using the input evaluation information sample in the above equation, the following is obtained:

$$y\left(x_i\right) = \varphi^T(x)\omega + b, i \le l \tag{11}$$

The results of the above equation are described in matrix form:

$$\Phi(X)\omega = Y - B \tag{12}$$

Where the parameters $\Phi(X)$, $Y$, and $B$ can be calculated using the following formulas:

$$\begin{cases} \Phi(X) = \left[\varphi^T\left(x_1\right), \varphi^T\left(x_2\right), \ldots, \varphi^T\left(x_l\right)\right]^T \\ Y(x) = \left[y\left(x_1\right), y\left(x_2\right), \ldots, y\left(x_l\right)\right]^T \\ B = bE \in R^l \end{cases} \tag{13}$$

In the equation, $E$ represents the training error.It is necessary to satisfy $y(x)(\varphi T(x)\omega+b)\geq 1-\xi$, where $\xi$ is a non-negative constant. During the non-linear mapping process of student internship evaluation indicators, it is necessary to solve the following problem:

$$\min J = \frac{1}{2}\left(\left\|\Phi(X)\omega - Y\right\|^2 + \left\|\omega\right\|^2\right) + CE_1^T\xi \tag{14}$$

In the equation, $C$ represents the penalty coefficient. For the const raint in equation (14), it is expressed as $D_y\Phi(X)\omega \geq E_l - \xi$, where $\xi_i \geq 0$, and $i \in l$, with the parameter $D_y = diag(y_1, \ldots, y_l)$.

Based on the principle of Lagrange multipliers, the following equation is constructed (YAN Zizong et al., 2023).

$$L = J - a^T\left(D_y\Phi\omega - E_1 + \xi\right) - \gamma^T\xi \tag{15}$$

In the equation, the parameter $\gamma = [\gamma^1, \ldots, \gamma^l]^T$, and $a$ is the Lagrange multiplier, given by $a = [a^1, \ldots, a^l]^T$. At this point, the following equation exists:

$$\omega = \left(\Phi^T\Phi\right)^{-1}\Phi^T\left[D_y a + Y\right] \tag{16}$$

Combining the above equations, the Lagrange equation can be simplified to the following:

$$L = -a^T Ha + a^T E_1 + Y^T Y \tag{17}$$

In the equation, the parameter $H = \frac{1}{2}D_y\Phi(\Phi^T\Phi)^{-1}\Phi^T D_y$ is positive semi-definite, leading to the construction of the dual equation:

$$\max W\left(a\right) = \max_a\left(L\right) = -a^T Ha + a^T E_1 \tag{18}$$

By solving the aforementioned dual equation, the non-linear mapping of student internship evaluation indicators is completed, establishing the relationship between the performance of student internship evaluation indicators and the level of student internship evaluation, thus concluding the student internship evaluation.

# 4. Experiments and Results

When using this method to compare the obtained evaluation results with the actual results, an assessment was conducted on 30 students who completed their internships.The actual mean score was 83.1 with a standard deviation of 5.3, while the mean score of the new method was 83.7 with a standard deviation of 4.2. The result of the paired t-test was $t(29) = -1.37$, $p = 0.18$. At the significance level of $\alpha = 0.05$ the p-value was more than 0.05, which indicates that the difference is not significant, proving the validity of the proposed method of evaluating student internships.

Before applying non-linear mapping to student internship evaluation indicators, the proposed method uses a deep learning algorithm to process the student internship evaluation indicators and utilizes a Softmax classifier for feature classification, thereby reducing the evaluation time and improving efficiency.

The evaluation times of the proposed method, method B (ZHOU Yue et al., 2020), and method C (XIONG Ning et al., 2020) for student internship evaluation are compared. With the same number of students (30 students), the proposed method demonstrates lower evaluation time.

ANOVA revealed the proposed evaluation method had a 95% precision, outperforming method B (80%) and method C (78%). Significant differences in precision were confirmed ($p < 0.001$). Post hoc tests showed the proposed method was significantly more precise than both methods B and C, but no notable difference was found between methods B and C. Moreover, the proposed method consistently surpassed 90% user satisfaction, exceeding that of method B (58% in second iteration) and method C (63% in third iteration).

# 5. Conclusion

This method has the advantages of high satisfaction, short processing time and strong nonlinear mapping ability. This approach is closer to the student's actual internship performance and provides more valuable feedback to the student.

# References:

[1]XIONG Ning,ZHU Wenguang,ZHONG Shiyuan,SHU Jiao,LI Weiwei,WANG Guang. (2020). Evaluation Method of Area Distribution Networks Based on Nonlinear Mapping and Kernel Principal Component Analysis.Modern Electric Power(05), 463-469(in Chinese).

[2]YAN Zizong and SHU Zhijia. (2023). A New Mathematical Explanation of Lagrange Multiplier.Studies in College Mathematics,26 (02): 52-54+84(in Chinese).

[3]ZHOU Yue,TIAN Wei,LIAO Wenhe,ZHANG Lin,LI Bo. (2020). Method for Predicting Burr of Hole Made Based on Convolutional Neural Network.Machine Building & Automation(02), 64-68(in Chinese).