Research Article

Agricultural Product Sales Prediction of ICM Neural Network Improvement by Sparse Autoencoder

YingHui Li [ID]

Henan Finance University, Zhengzhou 450046, China

Correspondence should be addressed to YingHui Li; liyinghui@hafu.edu.cn

With the rapid development of agricultural product sales data, the traditional prediction model cannot meet the processing needs. Based on deep learning theory, an improved ICM agricultural product sales prediction model using the softmax classifier is proposed. Introducing the sparse autoencoder in ICM can reduce feature loss. The features also can be extracted. In addition, using the pretreatment mode based on fuzzy membership theory, we can obtain the fuzzy correspondence of considerations and grades of agricultural product sales. At the same time, the precision of prediction for the model is further optimized. It can be seen that the agricultural product sales prediction model based on improved ICM can realize the real-time prediction of agricultural product sales. The maximum classification accuracy of the model can reach 80.98%, which means that it has certain practical application value.

1. Related Work

Agricultural products are necessities of daily life. With the increase of sales platform, the agricultural product sales model is more diversified, and the online sales model has gradually become the mainstream sales model of agricultural products. However, because of the particularity of its products, the agricultural products have a strict shelf life. Therefore, to reduce the waste of agricultural product resources during the shelf life of agricultural products, the accurate sales of the online agricultural products sales platform need to be achieved, which is an urgent problem for the online sales of agricultural products. The current online precision sales of agricultural products are mainly through the shallow-layer deep learning network model. Also, it extracts the value information from the massive agricultural product sales data to predict the trend of sales changes, so as to help merchants adjust the sales strategy in real time. For example, Grasman and Kornelis [1] and de Souza et al. [2] used the Bass model and agricultural product sales data to estimate the future agricultural product sales, so as to achieve the inventory management. Meanwhile, the confidence interval is used to process and obtain the required

exact expression. Xiang and Cho et al. analyzed the essential characteristics of the self-service. Also, the importance and significance of building a self-service network marketing system are expounded. Therefore, a new sales model is proposed, which combines independent marketing and online customization. Her-e, merchants, businesses, and consumers collaborate to complete production and sales, so as to realize the precise services for consumers [3, 4]. Tsoumakas and Chen et al. realized accurate short-term sales prediction by reviewing video sales and applying machine learning. It enables the company to minimize in-store inventory and expired products. At the same time, it helps to avoid missing sales [5, 6]. Chen et al. and Eric et al. implemented online sales prediction of agricultural products by using multitask recursive neural network based on trend alignment [7, 8]. However, the prediction performance of the shallow-layer model is limited, so the real-time analysis processing of agricultural product trading data cannot be realized. At the same time, considering that agricultural sales data are growing exponentially every day, the semi-supervised learning model is the main means of predicting such rapid growth data. So, this paper selects the semi-supervised deep learning model as the sales prediction

model of agricultural products. Also, it puts forward a super-imperial crown model based on the ICM to predict the agricultural product sales.

2. Basic Methods

2.1. Introduction to ICM Model. The ICM model is an autoencoder network model, which contains two hidden layers. In addition, its structure of the input node is same with the feature vector structure used for prediction. Because the self-coding network model only learns the predicted feature vectors but does not classify them, this feature enables the classification prediction function of the model by adding a classifier on the top layer. Typically, the ICM classifier can be selected to achieve the classification effect and can get the softmax classifier and the corresponding probability values.

If the input sample set is L, the hypothesis function method is used to analyze each type of j belonging to different classification probability value $p = (y = j \mid x)$, and the hypothesis function can be expressed as [9]

$$h_\theta\left(x^{(i)}\right) = \begin{bmatrix} p\left(y^{(i)} = 1 | x^{(i)}; \theta\right) \\ p\left(y^{(i)} = 2 | x^{(i)}; \theta\right) \\ p\left(y^{(i)} = 3 | x^{(i)}; \theta\right) \\ p\left(y^{(i)} = 4 | x^{(i)}; \theta\right) \end{bmatrix},$$

$$= \frac{1}{\sum_{j=1}^{4} e^{\theta_j^T x(i)}} \begin{bmatrix} e^{\theta_1^T x(i)} \\ e^{\theta_2^T x(i)} \\ \vdots \\ e^{\theta_k^T x(i)} \end{bmatrix}, \quad (1)$$

where $\theta_1, \theta_2, \ldots, \theta_k$ represent the model parameters to be obtained. The cost function is defined as [10]

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^{m} \sum_{j=1}^{k} \left\{ y^1 = j \right\} \log \frac{e^{\theta_j^T x(i)}}{\sum_{l=1}^{k} e^{\theta_l^T x(i)}} \right. \quad (2)$$

Add the self-coding network classified by softmax, and the structure of the ICM model is shown in Figure 1.

The steps for training the ICM model are as follows.

Firstly, the training samples are input into the model. Also, the two hidden layers are adopted to conduct unsupervised training on the training samples. So, the unlabeled output results are obtained. Secondly, the unlabeled output results are input into softmax classifier. At the same time, the labeled samples are used to train the softmax classifier. Thirdly, the values of some functions are solved, such as hidden layer kernel function, reconstruction error function, softmax classifier cost function, and the partial derivatives of all parameters. Finally, the parameters obtained in the above steps are utilized to initialize the model parameters.

Furthermore, the BP algorithm is used to solve the optimum parameters, so as to carry out the supervised training.

2.2. Improved ICM Model. It can be seen from the above analysis that it is easy for the ICM model to ignore the reconstruction of original non-linear data during prediction, which will affect the prediction results. In order to solve this problem, the sparse autoencoder is introduced into ICM and an improved ICM model is proposed.

Sparse autoencoder is an unsupervised learning method. Linear representation of vectors can be achieved by using a set of basis vectors $\varnothing_i$ in the input vector [11]:

$$X = \sum_{i=1}^{k} a_i \varnothing_i \quad (i = 1, 2, \ldots, k). \quad (3)$$

As can be seen from formula (3), k stands for the number of the input nodes and a_i stands for the linear correlation coefficient.

The average activation value of hidden layer unit j is [12]

$$\widehat{\rho}_J = \frac{1}{m} \sum_{i=1}^{m} a \left[\sum_{j}^{(2)} \left(x^i \right) \right]. \quad (4)$$

Here, $i = 1, 2, \ldots, m$, $m = k$; $a_j^{(2)}(x^i)$ represent the hidden layer unit activation amount when the input is X.

Set the sparseness parameter to ρ, make $\widehat{\rho}_J = \rho$, and use KL to optimize the penalty factor, such as formula (5). So, the neuron activity of hidden layer is controlled near ρ.

$$KL(\rho \| \widehat{\rho}) = \rho \log \frac{p}{1 - \widehat{\rho}} + (1 - \rho) \log \frac{1 - \rho}{1 - \widehat{\rho}_J}. \quad (5)$$

Therefore, the global cost function of sparse autocoding networks is [13]

$$J_{\text{spare}}(w, b) = J(w, b) + \mu \sum_{j=1}^{s_2} KL(\rho \| \widehat{\rho}_J), \quad (6)$$

where μ represents the punishment coefficient, and the larger its value is, the greater the punishment intensity is. Here $J(w, b)$ represents the global loss function. It can be calculated by formula (7). λ is responsible for avoiding overfitting [14].

$$J(w, b) = \left[\frac{1}{m} \sum_{i=1}^{m} J\left(W, b; x^i, y^i\right) \right] + \frac{\lambda}{2} \sum_{l=1}^{n_l-1} \sum_{l=1}^{s_l} \sum_{j=1}^{s_l+1} \left(W_{ji}^l\right)^2. \quad (7)$$

$J(W, b; x^i, y^i)$ represents the unit function:

$$J\left(W, b; x^i, y^i\right) = \frac{1}{2} \| h - y \|^2. \quad (8)$$

The output error can be obtained from formula (8).

$$\delta_i^{(3)} = \frac{\partial}{\partial r_i^{(3)}} \frac{1}{2} \| h - y \|^2$$
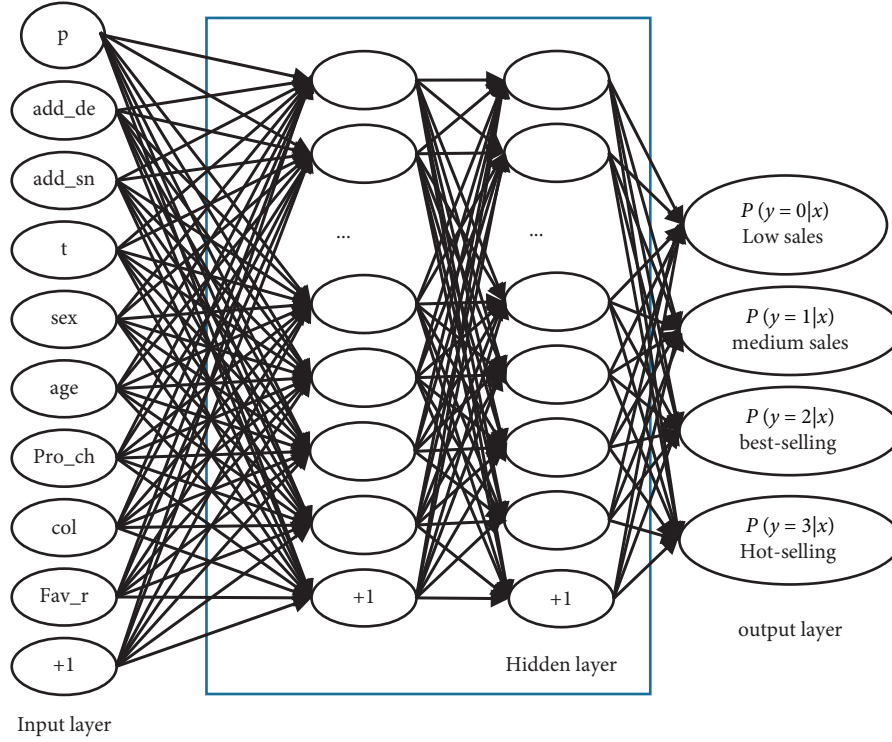$$= -\left(y_i - \alpha_i^{(3)}\right) f'\left(r_i^{(3)}\right). \quad (9)$$

Figure 1: Structure of the ICM model.

It can be seen from formula (8) that $\delta_i^{(3)}$ represents the error of output layer; $a_i^{(3)}$ represents the active function.

$$r_i^{(3)} = W_i^{(2)} a_i^{(3)} + b^{(2)}. \quad (10)$$

Thus, the hidden layer element error can be obtained as

$$\delta_i^{(2)} = \left(\left(\sum_{j=1}^{s_2} W_{ji}^{(2)} \delta_j^{(3)} \right) + \mu \left(\frac{-\rho}{\widehat{\rho}_J} + \frac{1-\rho}{1-\widehat{\rho}_J} \right) \right) f\prime\left(r_i^{(2)}\right). \quad (11)$$

The SICM model introducing sparse autoencoder is implemented as follows:

Step 1. Randomly classify all data into training dataset and testing dataset in a ratio of 4:1

Step 2. Label the testing dataset and randomly select some data to train the SICM model. Furthermore, the remaining data are used to verify the improved ICM model.

Step 3. Use the repeated experiment method to determine the hidden layers and the nodes of each layer in the sparse self-coding network.

Step 4. Select the training dataset as the input vector of sparse autocoding network for unsupervised pretraining and input the labeled and unlabeled data into softmax classifier for training.

Step 5. Adopt BP algorithm to tune network parameters. The global optimal parameters can be obtained. As can be seen from formula (12), when $\lambda \neq 0$, the weight decay function is activated, so as to update the weight [15].

$$\Delta W_{ij} = \frac{\eta \partial J}{\partial \Delta W_{ij}}. \quad (12)$$

The formula shows that $\eta = 0.05$ which represents the learning rate [16].

The weight updating formula can be expressed as

$$\begin{aligned} \Delta W^{(l)} &= \Delta W^{(l)} + \nabla_{W^{(1)}} J(W, b; x, y), \\ \Delta b^{(l)} &= \Delta b^{(l)} + \nabla_{b^{(1)}} J(W, b; x, y). \end{aligned} \quad (13)$$

Here, $\Delta W^{(l)} = \Delta b^{(l)} = 0$ represents the initial value, and the weight can be obtained according to formulas (14) and (15) [17].

$$W^{(l)} = W^{(l)} - \alpha \left[\left(\frac{1}{m} \Delta W^{(l)} \right) + \lambda W^{(l)} \right], \quad (14)$$

$$b^{(l)} = b^{(l)} - a \left(\frac{1}{m} \Delta b^{(l)} \right). \quad (15)$$

The above process can be represented in Figure 2.

3. Sales Prediction of Agricultural Products Based on Improved ICM

3.1. Building Improved ICM Model

3.1.1. Parameter Selection. The improved ICM model parameter is selected by the characteristic variables of the sales influencing factors of agricultural product. According to literature [18], the characteristic variables affecting the sales of
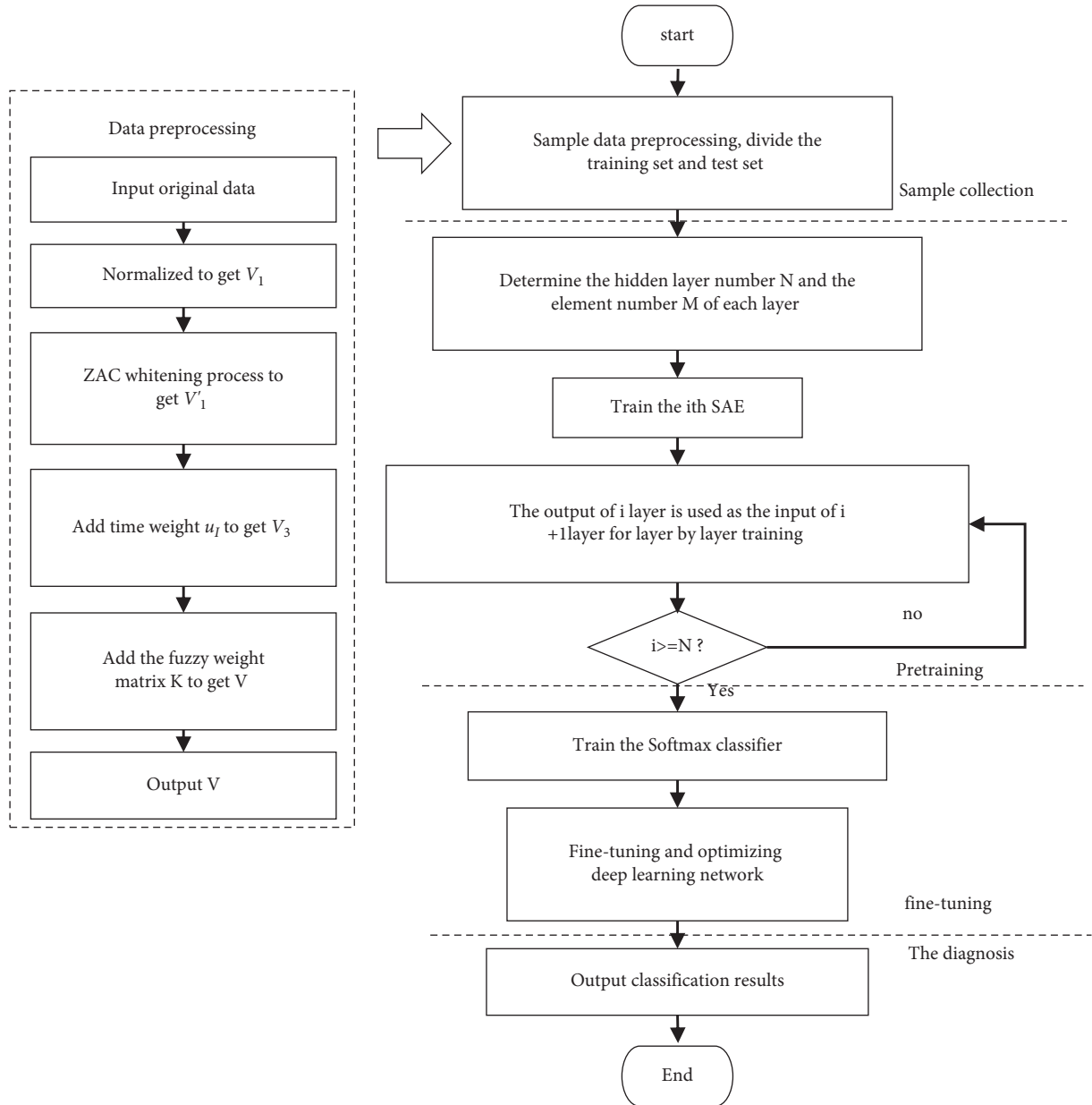
FIGURE 2: Improved ICM model implementation process.

agricultural products mainly include the attribute parameters of agricultural products, the personal factors of buyers, and nine-dimensional characteristic vector formed by promotion channels. Therefore, the parameters of the improved ICM model in this study use $V = [p, \text{add\_}de, \text{add\_}sh, t, \text{sex}, \text{age}, \text{pro\_}ch, \text{col}, \text{fav\_}r]$ to represent the 9-dimensional input vector, which is used for feature learning. Also, the preprocessed vector is expressed as $V = [pI, \text{add\_}de, \text{add\_}sh1, tI, \text{sex}, \text{age}I, \text{pro\_}chI, \text{col}, \text{fav\_}rI]$.

3.1.2. Classifier Design. The improved ICM model classifier uses the softmax classifier of the ICM model. As the cost function of the softmax classifier of ICM is a non-strict convex function, it is easy to fall into local optimum [19]. Therefore, this paper adds a weight attenuation term to the

cost function and penalizes the parameters whose weights are too large, so that the parameters can easily converge to the global optimal. The cost function added weighted attenuation term can be expressed as [20]

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^{m} \sum_{i=1}^{k} \{ y^1 = j \} \log \frac{e^{\theta_j^T x(i)}}{\sum_{l=1}^{k} e^{\theta_l^T x(i)}} \right] + \frac{\lambda}{2} \sum_{i=1}^{k} \sum_{i=1}^{n-1} \theta_{ij}^2. \quad (16)$$

The optimal parameter value can be obtained by minimizing $J(\theta)$, and the classifier can be used for the improved ICM model of agricultural product sales prediction.

3.1.3. Model Structure. There are two parts for the improved ICM model used for agricultural product sales prediction, namely, data preprocessing and deep learning network. The

specific structure is shown in Figure 3. The input vector of the model refers to the preprocessed data X, and the input cell number of SAE is $N + 1$, where N represents variable and 1 represents offsets. By processing the input vector through the hidden layer, the feature vector for classification can be obtained. Consider that SAE only digs deep features of sample feature vectors to obtain weight parameters of the hidden layer. The softmax classifier is needed to classify the processed data of hidden layers, so as to obtain all deep features [21]. In addition, the softmax classifier is added to the top layer of the model to obtain a complete improved ICM model. The output vector Y is the prediction result of agricultural product sales by the model.

3.2. Agricultural Product Sales Prediction Process Based on Improved ICM.

The improved ICM model is applied to the sales forecasting of agricultural products. First of all, the sales data of agricultural products are preprocessed. Also, the sparse autocoding network is adopted to select representation samples with common features from the sales data of unlabeled agricultural products, so as to realize the adaptive feature extraction. After that use the labeled and unlabeled data training classifier to classify the agricultural product sales level. Furthermore, the parameters of each layer are tuned according to the backpropagation algorithm and strictly convex function to obtain the global optimal parameter values. Thus, we can get an accurate classification prediction improved ICM model which can realize the sales prediction of agricultural products. Finally, the data are input into the prediction model, and the output data are the prediction result. The specific process of agricultural product sales prediction based on Improved ICM is shown in Figure 4.

4. Simulation Experiment

4.1. Simulation Environment and Data Source

4.1.1. Experimental Environment Construction and Data Collection.

This experiment is simulated on 64 bit Windows 7 system and coded on i386 platform. Crawler technology is used to obtain 5000 sets of data from Taobao egg agricultural products trading data; each set of data includes 20 samples.

4.1.2. Data Preprocessing.

Considering that there are noise, incompleteness, and inconsistency in the crawled original data, this paper performs normalization, zero-phase component analysis, and fuzzy preprocessing for the original data before the experiment.

(1) Normalization: carry out normalized preprocessing for experimental data to reduce the influence of data polarization on the final prediction results, which is shown in the following formula [22]:

$$X_n = \frac{Y_n}{z}, \tag{17}$$

where X_n represents the weight of an attribute total parameters in the total samples; Z indicates the total samples; and Y_n indicates sample feature vector with a total number n of an attribute.

(2) Zero-phase component analysis: perform zero-phase component analysis for the normalized data. M sample data are selected to convert the original data into matrix $X_1 = V_1$. Set the single sample dimension to 9 and perform ZCA whitening processing on dataset. The specific process can be shown as follows:

① Set X_1^T as the numerical matrix with original data of $9 * m$. G represents the matrix after normalizing X_1^T, and the average value of each attribute in G matrix is 0.

② Calculate the sample covariance matrix $\sum$ of G to obtain the corresponding eigenvalues and eigenvectors and rank them in order from large to small. The eigenvalues are denoted as $\gamma_1, \gamma_2, \ldots, \gamma_9$, and the corresponding eigenvectors are $U = [u_1, u_2, \ldots, u_9]$.

③ Multiply V_1^T and U^T to obtain the rotation matrix [23]:

$$
\begin{aligned}
V_{1\text{rot}}^{T_r} &= U^T V_1^T \\
&= \begin{bmatrix} u_1^T x_1 & \cdots & u_1^T x_m \\ \vdots & \ddots & \vdots \\ u_9^T x_1 & \cdots & u_9^T x_m \end{bmatrix}.
\end{aligned} \tag{18}
$$

Make the rotation matrix attribute values with unit variance:

$$V_{1\text{rot}}^{T_r} = \frac{1}{\sqrt{\gamma_1 + 0.1}} V_{1\text{rot}}^T \quad (i = 1, 2, \ldots, 9). \tag{19}$$

The unit variance per attribute in the rotation matrix is obtained by multiplying $1/\sqrt{\gamma_i}$ $(i = 1, 2, \ldots, 9)$ with each element of row i $(i = 1, 2, \ldots, 9)$ in $V_{1\text{rot}}'$, and denote the obtained matrix as $V_{1\text{rot}}''$, where $1/\sqrt{\gamma_i + \varepsilon}$ is used to replace $1/\sqrt{\gamma_i}$, which deals with the numerical fluctuation or overflow phenomenon when r_i approaches 0.

④ The results of ZCA whitening processing are as follows:

$$X_1^{T'} = U X_{1rot}^{T'} \tag{20}$$

(3) Fuzzy processing: conduct fuzzy processing for the data by fuzzy membership theory, so the obtained data can be used in the experiment. The method based on fuzzy processing is as follows:

① Let $V_2 = [TS, pI, \text{add\_}$ deI, add_sh1, $t1$, sex1, age1, pro_ch, col, fav_r1]. The time TS in the vector V2 represents the distance between the time of collecting sample data and the time of the system. Data are selected in a random way, and weights are marked for data in different states:
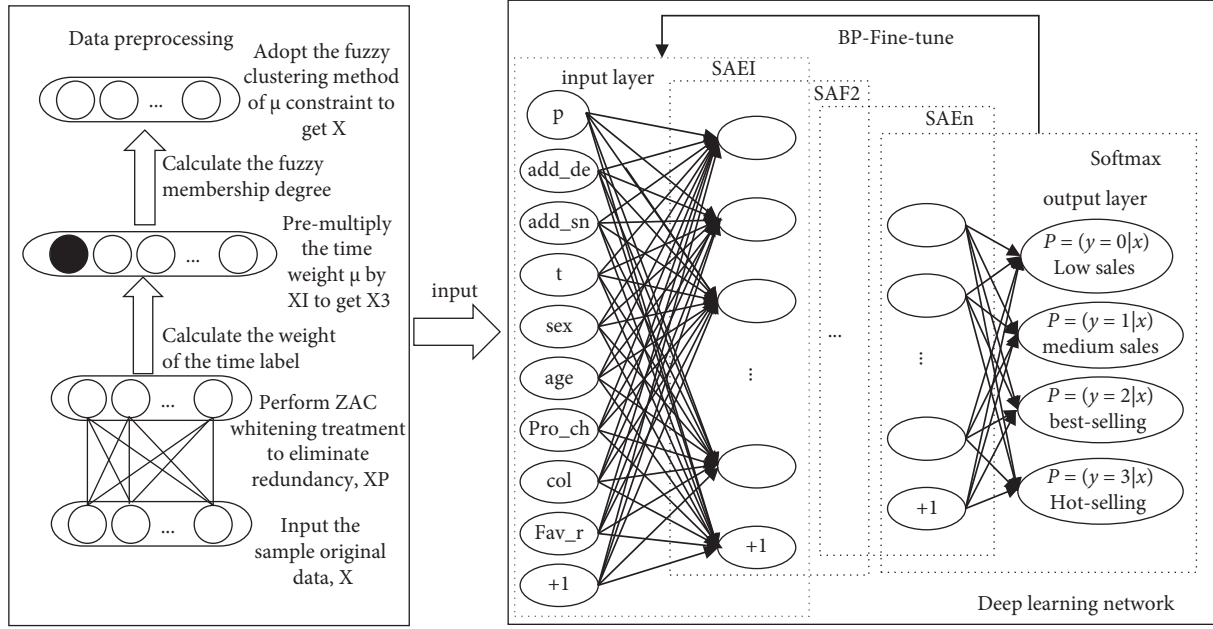
FIGURE 3: The improved ICM model structure.

$$TS = \frac{TS_n - TS_e}{30}, \tag{21}$$

where TS_n represents the current system time when data are processed and TS_c represents the time of data sample collection. The time label format is year-month-day.

② Use the fuzzy subset weight coefficient method to deal with the time weight and use the expert comment set method to set the weight of sample data at different distances:

$$\mu_1 = \begin{cases} 0.4, & TS < 1, \\ 0.31, & 1 \leq TS < 5, \\ 0.20, & 5 \leq TS < 9, \\ 0.08, & 9 \leq TS < 12, \\ 0, & TS \geq 12. \end{cases} \tag{22}$$

③ Determine the membership function. Set V_1' as the weight in the sales prediction of each data obtained according to TS in V2, namely, $V_3 = \mu_1 V_1'$. It can be seen that the factor set V3, the judgment set Y, and the weight matrix K select the membership function of trapezoidal distribution as

$$\mu(x_t) = \begin{cases} 0, & x_t < a, \\ \dfrac{x_i - a}{b - a}, & a \leq x_t \leq b, \\ 1, & x_t > b. \end{cases} \tag{23}$$

Here, x represents the feature vector in V3; t represents the value at t ($t = 1, 2, \ldots, m$) in the eigenvector x; and a and b can be determined by different characteristic attributes.

④ Obtain V3 by applying V_1' in step ①, and the input vector $V = \mu(x_t)_K V_3$ of SAE can be obtained by applyingV3 in step ②, which is as follows:

$$V = \{p, add\_de, add\_sh, t, sex, age, pro\_ch, col, fav\_r\}. \tag{24}$$

The experimental data for this study can be finally obtained by the above preprocessing. Some of the data are shown in Table 1.

As can be seen from the improved ICM model prediction process, part of the sample data is utilized to train the autocoding network, so as to obtain the kernel function, cost function, and derivative function. Also, a part of the sample data is adopted to train the softmax classifier. Therefore, the preprocessed experimental data should be divided into feature learning and classification learning. To ensure general applicability of the improved ICM model, 3000 samples are randomly selected from the experimental dataset to be the feature learning subset, and 1000 samples are selected as classification learning subset, including 600 softmax classifier training sets and 400 test sets.

4.2. Parameter Settings

4.2.1. Determination of Hidden Layers. Set hidden layers to 1~10 and input 500 groups of samples to obtain the classification accuracy of the model [24], as shown in Figure 5. The figure shows that if the number of hidden layers is less than 3, the classification accuracy of the model will be gradually improved with the increase of hidden layers. If the number of hidden layers is greater than 3, the model classification accuracy does not change with the increase of hidden layers. So, it indicates that when there are 3 hidden layers, the model has reached the optimal classification
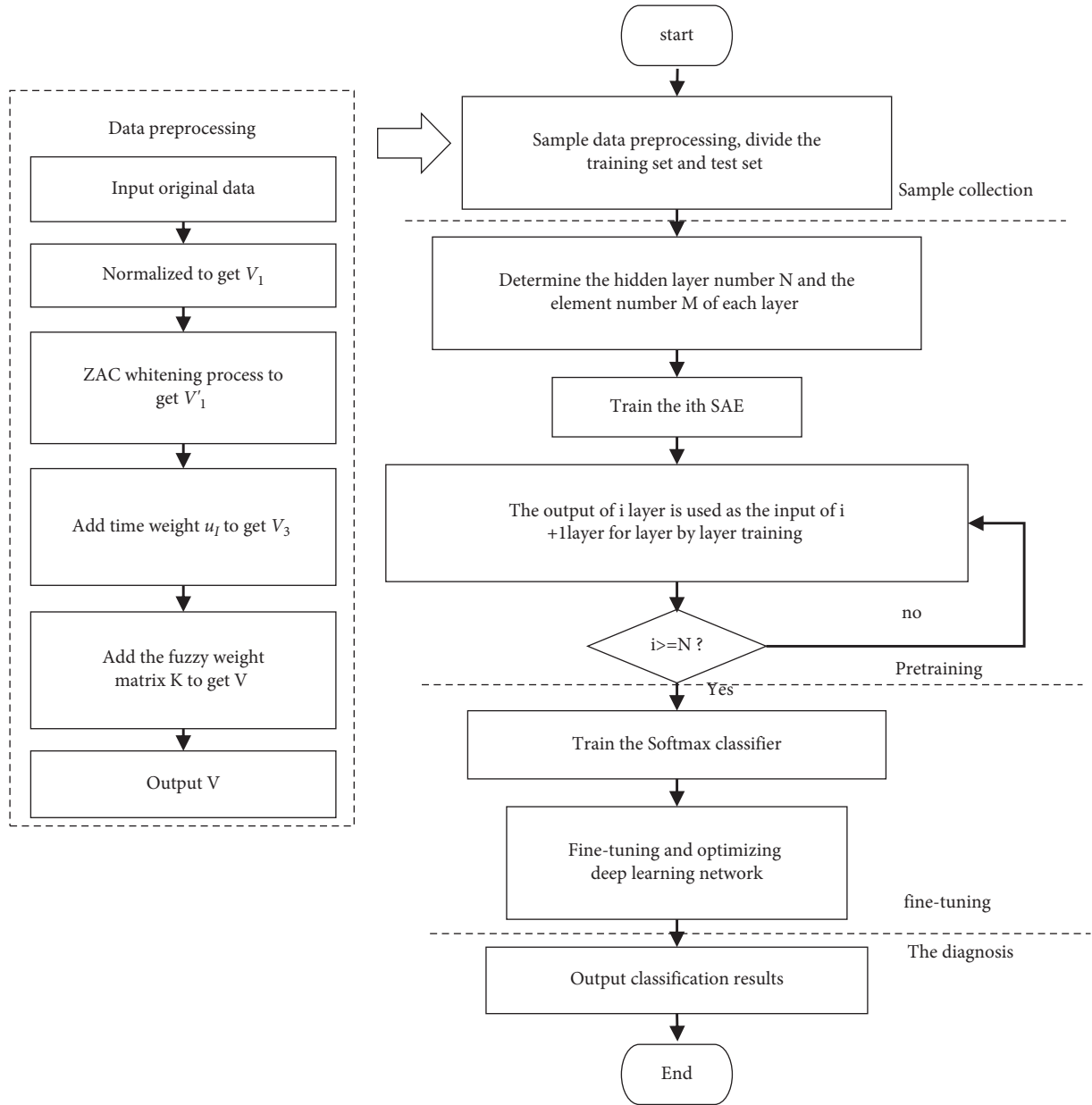
FIGURE 4: The prediction process of agricultural product sales based on improved ICM.

accuracy. Therefore, the hidden layer number of the model is set as 3 in this paper.

4.2.2. Determination of the Unit Number of Hidden Layers.

Set iterations to 100, hidden layers to 3, and the sparsity parameter to 0.05. Meanwhile, input 1200 sets of data and increase units of the three hidden layers in a ratio of 3 : 2 : 2, and the results are shown in Figure 6. Here, when the unit of hidden layer in the second and third layers changes, the classification accuracy changes accordingly. When the units of each hidden layer are 6, 4, and 4, the classification accuracy of the model is the highest. Therefore, the units of hidden layer is set to 6, 4, and 4, respectively.

4.3. Evaluation Index.

Precision and recall are used to evaluate model classification performance in this paper, and false positive rate (FPR) and false negative rate (FNR) are used to evaluate risk. The specific calculation methods are shown in formulas (25)–(28).

$$P = \frac{\text{the number correctly divided into this class}}{\text{the number actually divided into this class}}, \quad (25)$$

$$R = \frac{\text{the number correctly divided into this class}}{\text{the number actually divided into this class}}, \quad (26)$$

$$\text{FPR} = \frac{\text{the number incorrectly divided into this class}}{\text{the number actually divided into this class}}, \quad (27)$$

TABLE 1: Examples of preprocessed experimental specification data.

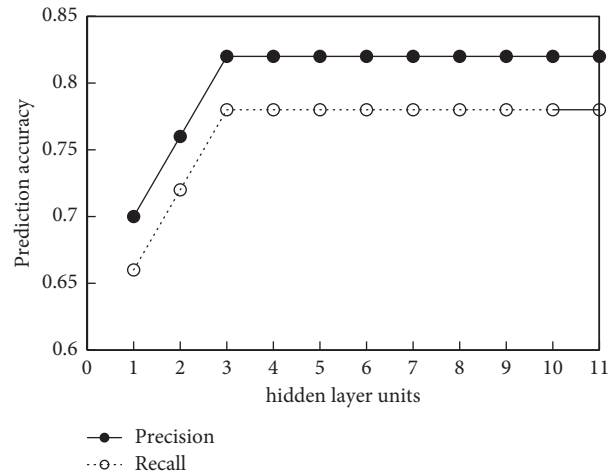| Id	P	Add_de	add_sh	Sex	Age	pro_ch	Col	fav_r
1	0.17	0.61	0.25	0.1	0.21	0.25	0.12	0.09
2	0.45	0.61	0.25	0.9	0.21	0.09	0.12	0.09
...	...	...	...	...	...	...	...	...
30	0.29	0.61	0.04	0.9	0.3	0.47	0.24	0.11

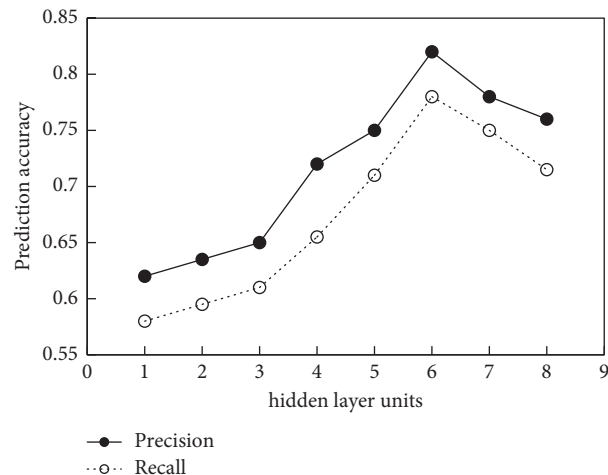FIGURE 5: Influence of hidden layers on model classification accuracy.

FIGURE 6: Influence of hidden layer units on model classification accuracy.

$$\text{FNR} = \frac{\text{the number not divided into this class}}{\text{the number actually divided into this class}}. \quad (28)$$

4.4. Experimental Result

4.4.1. Model Performance Verification. To verify the performance of the proposed model, the improved ICM is used to classify experimental datasets of different sizes, and its classification accuracy is compared with that of SVM and softmax classifier. The results are shown in Table 2. The table shows that with the increase of data scale, the classification precision and recall of each model are improved, and the

classification time increases accordingly. Compared with SVM and softmax classifier, the improved ICM has higher accuracy and shorter classification time under the same data scale. Furthermore, the advantage becomes more obvious with the increase of data scale. Therefore, there is strong self-learning ability for the proposed model to the massive and variable agricultural product sales data, where the high classification accuracy can be achieved, and the classification prediction advantage is obvious.

To verify the classification precision of the proposed model before and after optimization, in this paper, the classification accuracy of the model before and after optimization is determined by the number of iterations, which is shown in Table 3. The table shows that the classification

TABLE 2: Comparison of classification accuracy of different models.

Number of samples	Algorithm	Precision	Recall	Time
	Improved ICM	0.721	0.695	7.5
1000	SVM	0.646	0.628	8.6
	Softmax	0.425	0.406	8.8
	Improved ICM	0.763	0.743	12.1
2000	SVM	0.691	0.667	15.5
	Softmax	0.546	0.528	14.8
	Improved ICM	0.813	0.789	20.5
3000	SVM	0.715	0.709	25.2
	Softmax	0.551	0.538	25.9

TABLE 3: Precision changes of model classification.

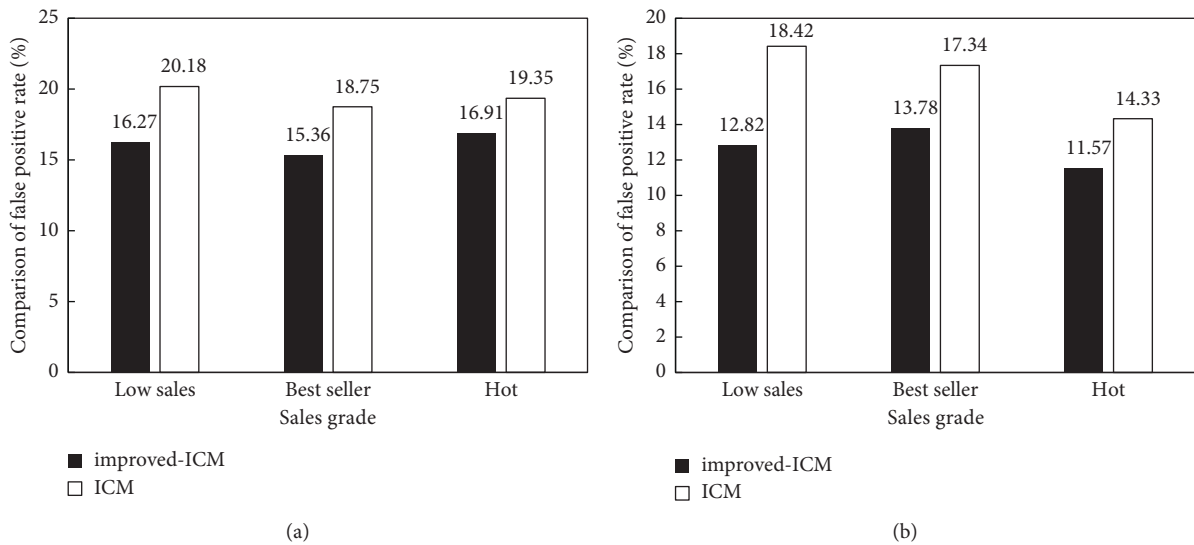| Softmax iteration (times)	SAE1 iteration (times)	SAE2 iteration (times)	SAE3 iteration (times)	Fine-tuned iteration (times)	Precision (%)
10	—	—	—	—	62.9
100	—	—	—	—	64.1
100	100	—	—	—	68.8
100	100	100	—	—	73.4
100	100	100	100	—	79.3
100	100	100	100	100	81.3

FIGURE 7: Comparison of model classification results. (a) Comparison of false positive rate. (b) Comparison of false negative rate.

accuracy before the proposed model optimization is 62.98%, and the classification accuracy after optimization is 80.98%, which increases by 18%. Therefore, the classification accuracy can be improved be fine-tuning the model.

4.4.2. Model Comparison. To verify the processing ability of the proposed model for fuzzy information, the improved ICM model is used to classify and predict 4000 sets of data, which are randomly selected from experimental samples. The false positives and false negatives of the two models are obtained by comparing the classification results with the ICM model, which are shown in Figure 7. The figure shows

that the false positive rate and false negative rate of proposed model are obviously lower than those of the ICM model, which indicates that the proposed model can improve the classification and recognition ability of the model through fuzzy data preprocessing.

In order to further verify the classification performance of the improved ICM model, the improved ICM and ICM models are used for classification under different data scales, which are shown in Table 4. The table shows that the improved ICM model has better classification accuracy, recall, and classification time indexes than ICM model under different data scales. Also, this advantage becomes more obvious with the increase of data scale. The reason is that the

TABLE 4: Comparison of model classification accuracy under different data sizes.

Number of samples	Algorithm	Precision	Recall	Time
1000	Improved ICM	0.73	0.71	7.11
	ICM	0.65	0.62	7.93
2000	Improved ICM	0.75	0.73	12.3
	ICM	0.72	0.69	13.8
3000	Improved ICM	0.82	0.79	22.9
	ICM	0.72	0.71	23.7

improved ICM model adopts sparse autocoding on the basis of the ICM model to reduce the loss of data features as much as possible. Compared with ICM autoencoder, the improved ICM model has stronger feature learning ability. In addition, the theory of fuzzy membership degree is added to the data preprocessing of the improved ICM model. It solves the problem that the ICM model ignores the fuzzy correspondence between the influencing factors and the grades of agricultural product sales. Therefore, the improved ICM model is superior to the ICM model in classification prediction of agricultural product sales.

5. Conclusion

To sum up, the proposed prediction method of agricultural product sales based on deep learning uses sparse autocoding to reduce the loss of data feature as much as possible. Meanwhile, we use the preprocessing agricultural product sales data based on fuzzy membership theory to build the improved ICM classification prediction model, which can effectively achieve the sales prediction of agricultural products. Compared with the ICM model, SVM, and softmax classifier, the model has higher sales prediction accuracy. When the number of hidden layers of the model is 3 and the units of hidden layers are 6, 4, and 4, the maximum classification accuracy of the model can reach 80.98%, which can be used to predict the actual sales of agricultural products. Although some achievements have been made, there are still some shortcomings, such as limitations in data sample acquisition, small size of data, and no verification of data processing capability with a lot of noise. So, the next step is to further discuss the classification and prediction performance of the improved ICM for massive data, which makes the model have stronger universality.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

References

[1] J. Grasman and M. Kornelis, "Forecasting product sales with a stochastic Bass model," *Journal of Mathematics in Industry*, vol. 9, no. 1, pp. 1–10, 2019.

[2] M. P. de Souza, M. Hoeltz, P. D. Gressler, L. B. Benitez, and R. C. S. Schneider, "Potential of Microalgal Bioproducts: general Perspectives and main Challenges," *Waste and Biomass Valorization*, vol. 10, no. 8, pp. 2139–2156, 2019.

[3] Y. Xiang, "Set self-service sales and online customization in one of the product network marketing system construction and management research," *Cluster Computing: the Journal of Networks, Software Tools and Applications*, vol. 22, no. 12, pp. 8803–8809, 2019.

[4] Y. S. Cho, W. S. Na, and S. C. Moon, "Periodicity analysis using weighted sequential pattern in recommending service," *Cluster Computing*, vol. 22, no. 4, pp. 1049–1056, 2019.

[5] G. Tsoumakas, "A survey of machine learning techniques for food sales prediction," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 441–447, 2019.

[6] K. Chen, P. Luo, and H. Wang, "Investigating transitive influences on WOM: from the product network perspective," *Electronic Commerce Research*, vol. 17, no. 1, pp. 149–167, 2017.

[7] T. Chen, H. Yin, H. Chen, and H. Wang, "Online sales prediction via trend alignment-based multitask recurrent neural networks," *Knowledge and Information Systems*, vol. 62, no. 7, pp. 1–29, 2019.

[8] W. K. Eric, See-To, and E. W. T. Ngai, "Customer reviews for demand distribution and sales nowcasting: a big data approach," *Annals of Operations Research*, vol. 270, no. 1-2, pp. 415–431, 2018.

[9] S. K. Sharma, S. Chakraborti, and T. Jha, "Analysis of book sales prediction at Amazon marketplace in India: a machine learning approach," *Information Systems and E-Business Management*, vol. 17, no. 2-4, pp. 261–284, 2019.

[10] H. Yuan, W. Xu, Li Qian, and R. Lau, "Topic sentiment mining for sales performance prediction in e-commerce," *Annals of Operations Research*, vol. 270, no. 1-2, pp. 553–576, 2018.

[11] L. Wen and X. Yuan, "Forecasting the annual household electricity consumption of Chinese residents using the DPSO-BP prediction model," *Environmental Science and Pollution Research*, vol. 27, no. 1, pp. 22014–22032, 2020.

[12] Y. Zhang and G. Pan, "A hybrid prediction model for forecasting wind energy resources," *Environmental Science and Pollution Research*, vol. 27, no. 2, pp. 19428–19446, 2020.

[13] Yu-J. Choe and J.-H. Yom, "Improving accuracy of land surface temperature prediction model based on deep-

learning," *Spatial Information Research*, vol. 28, no. 8, pp. 377–382, 2020.

[14] V. Koolwal and K. K. Mohbey, "A comprehensive survey on trajectory-based location prediction," *Iran Journal of Computer Science: International Journal*, vol. 3, no. 1, pp. 65–91, 2020.

[15] Z. Lei, "Research and analysis of deep learning algorithms for investment decision support model in electronic commerce," *Electronic Commerce Research*, vol. 20, no. 1–2, pp. 275–295, 2020.

[16] H. H. Elmousalami and M. Elaskary, "Drilling stuck pipe classification and mitigation in the Gulf of Suez oil fields using artificial intelligence," *Journal of Petroleum Exploration and Production Technology*, vol. 10, no. 10, pp. 2055–2068, 2020.

[17] K. Priya, N. Goel, and V. K. Garg, "Human age prediction using DNA methylation and regression methods," *International Journal of Information Technology: An Official Journal of Bharati Vidyapeeth's Institute of Computer Applications and Management*, vol. 12, no. 2, pp. 373–381, 2020.

[18] T. Izui and G. Venture, "Correlation analysis for predictive models of Robot User's Impression: a study on Visual medium and Mechanical noise," *International Journal of Social Robotics*, vol. 12, no. 5, pp. 425–439, 2020.

[19] Y. Gao, "Forecast model of perceived demand of museum tourists based on neural network integration," *Neural Computing & Applications*, vol. 33, no. 4, pp. 1–11, 2020.

[20] D. Li, T. Nian, H. Wu, F. Wang, and L. Zheng, "A predictive model for the geometry of landslide dams in V-shaped valleys," *Bulletin of Engineering Geology and the Environment*, vol. 79, no. 2, pp. 1–14, 2020.

[21] J. N. Swerdel, J. M. Reps, D. Fife, and P. B. Ryan, "Developing predictive models to determine Patients in End-of-life Care in Administrative datasets," *Drug Safety: The Official Journal of the International Society of Pharmacovigilance [ISoP]*, vol. 43, pp. 447–455, 2020.

[22] L.-Y. Xue, Z.-Y. Jiang, T.-T. Fu et al., "Transfer learning radiomics based on multimodal ultrasound imaging for staging liver fibrosis," *European Radiology*, vol. 30, no. 5, pp. 2973–2983, 2020.

[23] W. Zhao, R. Zhang, Y. Lv, and J. Liu, "Quantile regression and variable selection of single-index coefficient model," *Annals of the Institute of Statistical Mathematics*, vol. 69, no. 4, pp. 761–789, 2017.

[24] M. Kim, K. M. Jang, J.-H. Kim et al., "Differentiation of mass-forming focal pancreatitis from pancreatic ductal adenocarcinoma: value of characterizing dynamic enhancement patterns on contrast-enhanced MR images by adding signal intensity color mapping," *European Radiology*, vol. 27, no. 4, pp. 1722–1732, 2017.