www.cad-journal.net
Computer-Aided Design AND APPLICATIONS

Research on Japanese Training Tutoring Method Based on Intelligent CAD Technology and Data Mining

Nian Tong[1*]

[1]College of International Education, Shanghai Jianqiao University, Shanghai, 201306, China

Corresponding author: Nian Tong, svxd12356@163.com

Abstract: From the perspective of positive mental healthcare, the necessity of Japanese skill training is that it can not only improve learners' language ability but also pay more attention to the cultivation of learners' psychological resilience and ability to cope with pressure. In order to improve the effect of Japanese learning, this paper combines artificial intelligence technology and data mining technology to study Japanese training methods and analyzes the shortcomings of traditional acoustic models. This paper proposes to introduce deep learning theory to improve the shortcomings of traditional acoustic models and introduces related technologies. Moreover, this paper builds an intelligent Japanese training model to train the local acoustic model and uses end-to-end technology based on the improved speech feature extraction technology, which greatly reduces the difficulty of building an offline speech recognition system for Japanese training. It can be seen from the simulation teaching that the Japanese training method based on CAD technology and data mining proposed in this paper can effectively improve the effect of Japanese training.

Keywords: Intelligent CAD technology; data mining; Japanese; Tutoring Method;
DOI: https://doi.org/10.14733/cadaps.2025.S8.102-115

1 INTRODUCTION

In the context of globalization, Japanese is an important second language, and its learning and mastering are of great value for international communication, business cooperation, and academic research. However, challenges such as stress, anxiety, and acculturation during language learning can have a negative impact on learners' mental healthcare. Therefore, the integration of positive mental healthcare concepts into Japanese language skills training, by providing psychological support, stress management skills, and positive incentives, can help learners build self-confidence, overcome difficulties, achieve language learning goals, and maintain and promote their psychological

well-being. In addition, this comprehensive training mode can also enhance learners' language learning motivation, improve learning efficiency, cultivate cross-cultural communication ability, and lay a solid foundation for learners' adaptation and development in a multicultural environment. Conditional random field (CRF) is an effective method for sequence labeling and segmentation and has high learning efficiency, which is suitable for complex large-scale computational linguistics. CRF models do not require the strict independence assumptions required by Hidden Markov Models (HMMs) and also overcome the label bias problem of Maximum Entropy Markov Models (MEMM) and other non-generative directed graph models. Moreover, it is able to consider issues from a ten-day perspective. This method has been successfully used in natural language processing tasks, such as Chinese word segmentation, part-of-speech tagging, chunk recognition, named entity recognition, etc., and has achieved good results in these fields. However, there is no relevant report on the analysis of Japanese dependencies [1].

Reference [2] introduced cRF into Japanese dependency analysis, combined it with a cascading blocking algorithm, and took advantage of CRF to consider the problem from a global perspective, transforming the problem of judging whether there is a dependency between adjacent text sections into a sequence annotation problem. This method considers the problem from the point of view of the whole sentence and gives an optimal annotation result to each annotation unit, which overcomes the limitation that the previous classifier does not consider the problem comprehensively. As a useful supplement to statistical methods, rule methods are still widely used in many fields of natural language processing. Traditional rule methods are mostly written by hand based on experience and knowledge. They use linguistic knowledge to summarize rules and manually form a rule base. They can only play a role in limited language phenomena. In the face of rich and colorful natural languages, rules can hardly cover all language phenomena [3]. The traditional rule method has some problems, such as the difficulty of rule acquisition, the tedious manual rulemaking, and the conflict between rules. Aiming at this problem, this paper adopts the error-driven mechanism based on CRF, adds the primary recognition results of CRF as features to the feature template of the secondary recognition of CRF, uses statistical methods to learn the error rules therein automatically, trains the machine recognition model and corrects errors, which further improves the effect of dependency analysis [4].

Compared with the undergraduate courses, the course difficulty of Japanese majors in colleges and universities is slightly lower, which is closer to the actual work and omits more complex academic knowledge points and sentence pattern grammar [5]. Its dynamic nature is the primary feature of Japanese major courses in higher education. The original intention of Japanese courses in higher education is to enable students to use Japanese skillfully and carry out routine homework after entering enterprises. With the progress and development of the times, many levels and processes of Japanese trade are also changing imperceptibly. When teachers teach, they consciously cut and add some of them to meet the needs of the society for Japanese students; in this way, students can be more adaptable and flexible in their work [6]. Not only that, but the dynamic nature of the course is also reflected in the profound changes in the goal of knowledge and ability with the change in employment demand. For example, communication between Japanese enterprises is becoming increasingly complex, and the past mail and text communication can no longer meet the deepening cooperation. Therefore, frequent meetings and negotiations between Japanese enterprises and Japanese enterprises require students to have excellent oral and communication skills. Therefore, The current advanced Japanese course [7] is more about training students to use Japanese conversation in different scenes more smoothly, not limited to the traditional email writing training. Popular speaking, what kind of Japanese students to train, what kind of Japanese course system to use, and what kind of course objectives to establish are changing with the times and market demand, which is a dynamic process; this not only requires students to have flexible adaptability but also requires curriculum builders and developers to have a keen practical vision and set up Japanese

courses that are more in line with market demand [8]. Accompanying the dynamics are the modular characteristics of the curriculum. The Japanese curriculum is a targeted professional curriculum, which is different from the basic Japanese curriculum. The difficulty of each module is the same; that is, from the beginning of learning to the end of the curriculum, the ability is improving, and the difficulty of the curriculum is relatively stable [9], which results in the learning situation of the curriculum, which is difficult to start and easier to learn, At first, many students felt that this course was too dull, complicated and difficult to start. The modularization feature enables teachers to have a clear concept of modules in teaching and students' learning, and they are able to reflect and summarize in connection with actual life [10].

Japanese courses are all based on real work scenes. Practicality is its distinctive feature. Unlike Japanese listening and Japanese writing, it can improve its level independently. If the teaching and learning of Japanese courses adopt simple mechanical recitation, memorization, and other methods, there will be a big problem in knowledge understanding first; students can't really and deeply understand the way of communication from words to use situations and scenes through books [11] Therefore, in the course of setting up and teaching, on the one hand, we must combine the actual situation, build a simulation scene, and fully mobilize the students' active imagination and thinking so that students can fully understand the meaning of sentences and deepen their memory of knowledge. On the other hand, we must use various auxiliary means to help students participate in various activities that can be simulated to mobilize their enthusiasm because only real participation can help students gain a more profound and vivid experience [12]. The practicality of the advanced Japanese course stems from its course content. At the same time, there is a special form to sublimate and summarize the course content. While students are learning, they are not only limited to the given reference content and knowledge but also enrich it by combining their knowledge in the Japanese learning process. It can be said that in the process of practice, students continue to accumulate rich Japanese knowledge; at the same time, the knowledge learned in various Japanese courses is also used to process and integrate the knowledge of the course. This is a two-way activity construction process, which fully reflects the practicality of the course [13]. It is worth noting that the current Japanese curriculum is not limited to simulation training and practice in the guest room. With the deepening of school-enterprise cooperation, more personnel with rich work experience help full-time teachers in the school to jointly construct the second classroom and teach Japanese-related knowledge in the form of theme activities and lectures. At the same time, students have more opportunities to go to the front-line workplaces, restore the scenes in the textbooks, and deepen understanding and memory. This shows from the side that the practicality of Japanese courses is closely centered around schools [14] and enterprises. It is a long-term process of multilateral exchanges, which requires the active participation of students and full consideration of curriculum implementers and constructors [15].

The Japanese course is an open course. In addition to covering some essential knowledge points of basic business Japanese, most of the content is presented to students in the form of scene modules. Different from the traditional basic Japanese course, its learning and teaching are creative. When students learn, it is essentially a creative process, creating their own Japanese knowledge framework; in the process of learning, knowledge and skills are accumulated continuously without fixed models and requirements [16], which is similar to Dewey's "activity curriculum" to some extent, focusing on the accumulation and development of experience. Learning Japanese should be carried out in a relaxed and active atmosphere. Because in such an atmosphere, students are more likely to spread their thoughts and create more new things. Their creativity also reflects the variability of the learning process. Each student will have a different tendency to understand Japanese knowledge, which does not affect the learning effect but adds to the charm of this course [17]. In the process of teaching, the higher Japanese course usually ignores its creative characteristics and teaches it step by step with a conventional template, which limits the development of the subject and the

potential of students. The Japanese course is not only a course that needs to be carefully studied and understood but also a course that needs to be truly experienced and created with heart. Each student has different communication characteristics and ways of doing things. Long-term instillation of students' stereotyped conversation "routines" not only hinders the improvement of students' Japanese level but also limits their personal development in future posts [18].

This paper combines CAD technology and data mining technology to study Japanese training methods, improve the Japanese learning effect, and promote Sino-Japanese phonetic communication.

2 CONSTRUCTION OF OFFLINE SPEECH RECOGNITION SYSTEM BASED ON DEEP LEARNING

2.1 Disadvantages of Traditional Acoustic Models

This paper assumes that there are n data samples available for observation, as shown in Figure 1, and these points do not conform to the ellipsoid distribution in the d-dimensional space and cannot be described by a single Gaussian density function. Therefore, at this time, we can use the principle of equivalent replacement to treat the points in the figure as approximately obtained from a single Gaussian distribution, which is generated by a total of M models. Because the state of each point is unknown, the state of the model can also be regarded as unknown. Therefore, this model is called a Gaussian mixture model.

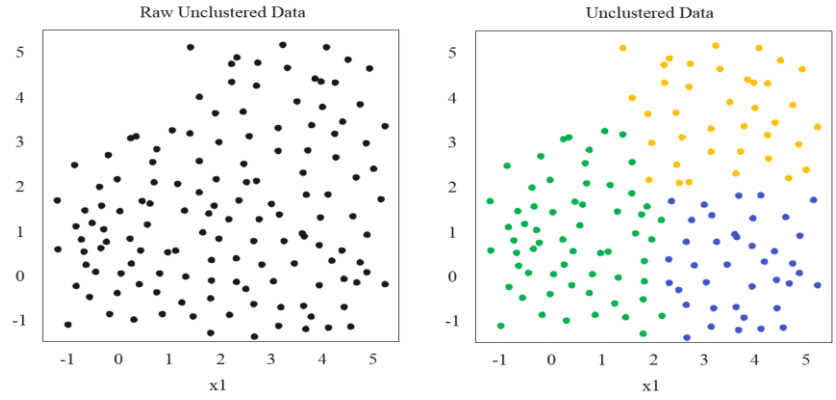

Figure 1: Observation sample

In summary, the probability distribution of the model is:

$$P(\mathrm{x}\left|\theta\right.)=\sum_{k=1}^{K}a_k\varphi(\mathrm{x}\left|\theta_k\right.) \tag{1}$$

In the traditional speech recognition field, GMM-HMM is the most widely used model. The specific flow chart is as follows (Figure 2).

Before people output speech, they need to organize their own language in the brain. This process is unobservable from the outside world, and this process is exactly in line with the Markov chain describing the hidden state in HMM.

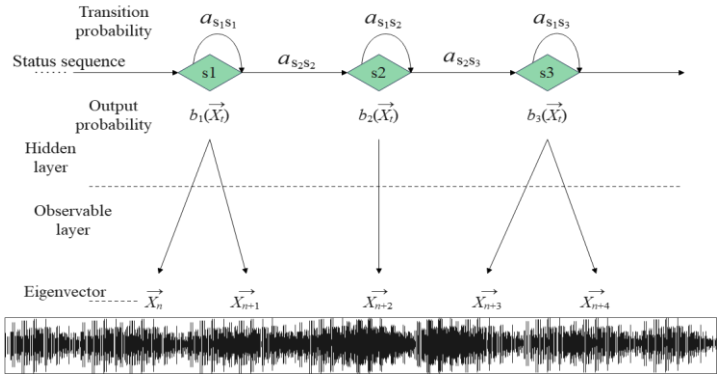

Figure 2: GMM-HMM acoustic structure.

To describe the model λ of the HMM, we assume that the sequence of HMM parameters is as follows:

$$\lambda = \{K, S, \rho, A, B\} \tag{2}$$

After completing the HMM observation sequence, we use the Gaussian mixture model to model the output probability. The principle is as follows:

$$\begin{aligned} b_i\left(\vec{X}_t\right) &= p\left(\vec{X}_t \mid s_t = i\right) = \sum_{j=1}^{M} c_j * N(\vec{X}_t, \mu_{ij}, \sum ij) \\ N(\vec{X}_t, \mu_{ij}, \sum ij) &= \frac{1}{(2\pi)^{p/2}\left|\sum ij\right|^{1/2}} \exp\left\{-\frac{1}{2}\left(\vec{X}_t - \mu_{ij}\right)^{T} \sum_{ij}^{-1} \left(\vec{X}_t - \mu_{ij}\right)\right\} \end{aligned} \tag{3}$$

The emergence of GMM-HMM is a big step in the field of acoustics, and even once dominated the field of acoustic modeling for speech recognition, but any method will have limitations, and GMM-HMM cannot escape this law.

In the current era of CAD, the short speech recognition accuracy and generalization ability of the GMM-HMM model are far from enough, so this topic will introduce a deep neural network to solve the defects of GMM-HMM.

2.2 The Theoretical Background of Deep Learning

Supervised learning is generally used in classification problems and regression problems. It trains the optimal network model by labeling the existing data samples, that is, classifying the test data.

Unsupervised learning is mainly applied to clustering problems.

The gradient descent algorithm is a widely used optimization algorithm in the traditional machine learning field. It is widely used to solve linear or nonlinear optimal solutions. The principle is as follows:

In machine learning, we usually express the objective function in the following form:

$$L(\theta) = \frac{1}{M}\sum_{i=1}^{M} L\left(f\left(x_i, \theta\right), y_i\right) \tag{4}$$

M is the number of training samples. The formula is as follows:

$$\nabla L(\theta) = \frac{1}{M}\sum_{i=1}^{M} \nabla L\ f\ x_i, \theta\ , y_i \tag{5}$$

The updated formula of the model parameters is:

$$\theta_{t+1} = \theta_t - \alpha \nabla L\ \theta_t \tag{6}$$

Therefore, the stochastic gradient descent method is introduced to solve this problem. The algorithm calculates the samples through random probability, which greatly reduces the time to obtain the optimal weights. The objective function and the principle formula of the stochastic gradient descent algorithm are as follows:

$$\begin{aligned} L\ \theta; x_i, y_i\ &= L\ f\ x_i, \theta\ , y_i \\ \nabla L\ \theta; x_i, y_i\ &= \nabla L\ f\ x_i, \theta\ , y_i \end{aligned} \tag{7}$$

The stochastic gradient descent method can only calculate a single sample at a time, but when the sample size M is large, the training speed cannot be guaranteed. Therefore, this topic proposes to use the mini-batch gradient descent algorithm, and its formula is as follows:

$$\begin{aligned} L\ \theta; x_i, y_i\ &= \frac{1}{m}\sum_{j=1}^{m} L\ f\ x_i, \theta\ , y_i \\ \nabla L\ \theta; x_i, y_i\ &= \frac{1}{m}\sum_{j=1}^{m} \nabla L\ f\ x_i, \theta\ , y_i \end{aligned} \tag{8}$$

We use a simple three-layer neural network to understand the principle of the algorithm. First, understand the structure of the neuron. The structure of the neuron is shown in Figure 3.

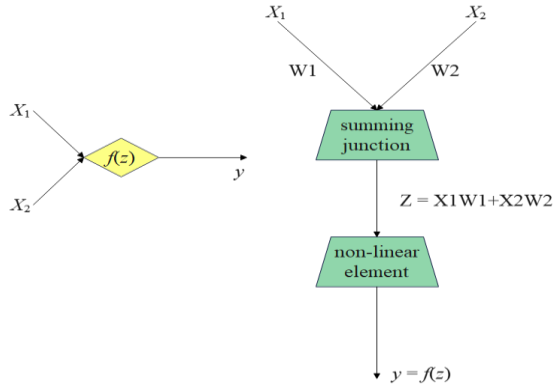

Figure 3: Neuron structure.

A simple three-layer neural network is shown in Figure 4. The forward algorithm from the input layer to hidden layer 1 is calculated as follows:

$$z_j = \sum_{i=1}^{I} w_{ij} x_i \tag{9}$$

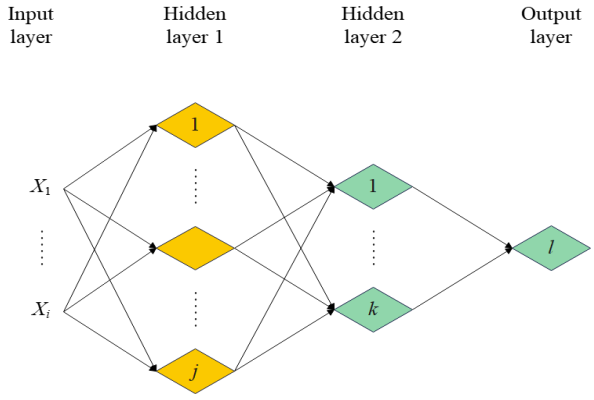

Figure 4: Three-layer neural network.

$$y_j = f\ z_j \tag{10}$$

The forward algorithm from hidden layer 1 to hidden layer 2 is as follows:

$$z_k = \sum_{j=1}^{J} w_{jk} x_j \tag{11}$$

$$y_k = f\ z_k \tag{12}$$

The expression from hidden layer 2 to the output layer is as follows:

$$z_l = \sum_{k=1}^{K} w_{kl} x_k \tag{13}$$

$$y_l = f\ z_l \tag{14}$$

y_l is the final result calculated by the neural network in the forward propagation. This result cannot be guaranteed to be absolutely correct. Therefore, the backpropagation algorithm is used to calculate the loss gradient of each neuron. The calculation flow of the backpropagation algorithm is shown in Figure 5.

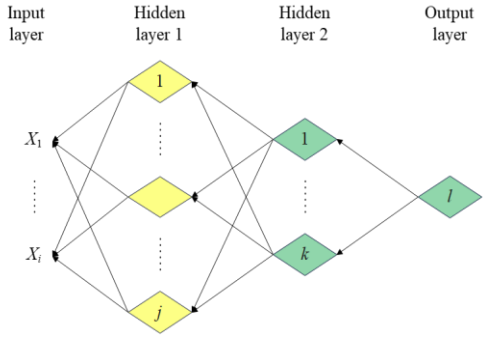

Figure 5: The calculation flow of the back-propagation algorithm.

The calculation method of the backpropagation algorithm is based on the chain derivation rule. Therefore, we assume that the loss function of the output layer l is $E=\frac{1}{2}\ y_l-t_l\ ^2$ and the expected value is t_l, and calculate the partial derivative y_l-t_l with respect to y_l. Because of $y_l=f\ z_l\ $, the partial derivative of the loss function relative to z_l is:

$$\frac{\partial E}{\partial z_l}=\frac{\partial E}{\partial y_l}\bullet\frac{\partial E}{\partial z_l}=\ y_l-t_l\ f'\ z_l \tag{15}$$

The algorithm calculates the error gradient from input layer 1 to hidden layer 2, and the formula is as follows:

$$\frac{\partial E}{\partial y_k}=\frac{\partial E}{\partial z_l}\bullet\frac{\partial E}{\partial y_k}=w_{kl}\frac{\partial E}{\partial z_l} \tag{16}$$

In the same way, the algorithm calculates the error gradient from hidden layer 2 to hidden layer 2:

$$\frac{\partial E}{\partial y_j}=\sum_k\frac{\partial E}{\partial z_k}\frac{\partial z_k}{\partial y_j}=\sum_k\frac{\partial E}{\partial y_k}\frac{\partial y_k}{\partial z_k}\bullet w_{jk} \tag{17}$$

In the above formula, j represents the neurons of hidden layer 1, and k traverses the neurons of hidden layer 2. The error gradient from hidden layer 1 to input layer i is as follows:

$$\frac{\partial E}{\partial x_i}=\sum_j\frac{\partial E}{\partial y_j}\frac{\partial y_j}{\partial z_j}\frac{\partial z_j}{\partial x_i}=\sum_j\frac{\partial E}{\partial y_j}\frac{\partial y_j}{\partial z_i}\bullet w_{ij} \tag{18}$$

In the above formula, i is the input unit, and j traverses the neurons of hidden layer 1. According to this principle, we can calculate the gradient of each intermediate node, so the i to j error gradient is as follows:

The calculation method of the backpropagation algorithm is based on the chain derivation rule, so we assume that the loss function of the output layer l is $E=\frac{1}{2}\ y_l-t_l\ ^2$, and the expected value is t_l, and calculate the partial derivative y_l-t_l with respect to y_l. Because of $y_l=f\ z_l\ $, the relative partial derivative of the loss function is:

$$\frac{\partial E}{\partial z_l}=\frac{\partial E}{\partial y_l}\bullet\frac{\partial E}{\partial z_l}=\ y_l-t_l\ f'\ z_l \tag{19}$$

2.3 Construction and Implementation of End-to-end Speech Recognition Model Based on Convolutional Neural Network

From the perspective of engineering applications, CNN is more accessible for achieving large-scale operations than RNN/LSTM. Moreover, the optimization and acceleration operation theory of the CNN network is relatively mature. For example, converting small matrices into products of large matrices, reducing the size of convolution kernels, using optimizers, etc., all provide possibilities for CNN in speech recognition applications. This paper collects its own equipment conditions and chooses to learn from the AlexNet network, which has outstanding performance in speech recognition and target

detection, to build the network model required for the project. Compared with traditional convolutional neural networks, AlexNet has the following three advantages:

(1) In this paper, ReLU (Revised Linear Unit) is used as the activation function, which can accelerate the calculation of neural network and greatly reduce the amount of calculation. Formula 20 is the ReLU function formula.

$$\mathrm{Re}\,LU(\mathrm{x})=\begin{cases} x, & \text{if} \quad \mathrm{x}>0; \\ 0, & \text{if} \quad \mathrm{x} \le 0. \end{cases} \tag{20}$$

(2) AlexNet uses Dropout in the fully connected layer for data expansion to prevent overfitting. The principle is shown in Figure 6. The Dropout probability selected by the CNN network built by the system is 0.2.

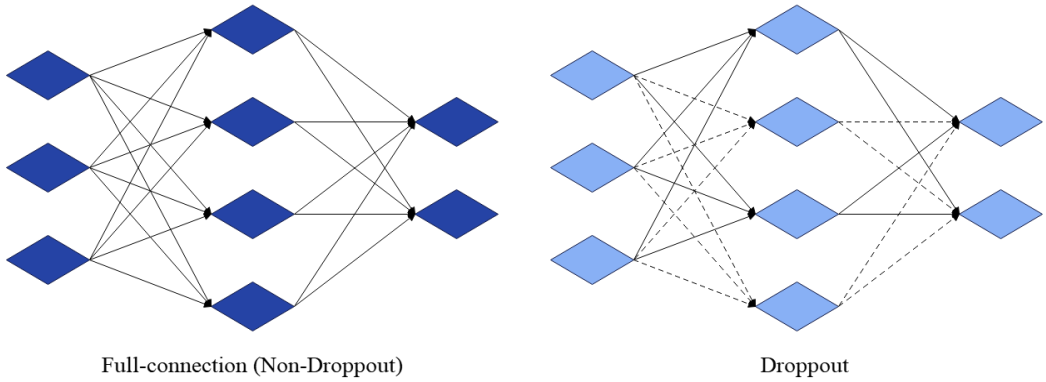

Figure 6: Schematic diagram of dropout.

(3) Local normalization (LRN). The principle is to let the neurons with large responses inhibit the neurons with small responses to form local competition and enhance the generalization ability of the neural network. Its formula is as follows:

$$b_{x,y}^{i}=a_{x,y}^{i}\left(k+\alpha\sum_{j=\max(0,i-n/2)}^{\min(N-1,i+n/2)}{a_{x,y}^{j}}^{2}\right)^{\beta} \tag{21}$$

The CNN network structure constructed by AlexNet in this paper is shown in Figure 7. The input of the CNN network in Figure 7 is a spectrogram with a 200-dimensional eigenvalue sequence, and the simplified structure of the network is represented by a convolutional unit and a fully connected layer (FC). The network includes a total of 12 convolutional layers, four pooling layers, a Reshape layer, and two fully connected layers. Each convolution unit represents two layers of convolution, the size of the convolution kernel is 3 × 3, the activation function is set to ReLU, an LRN is performed after each layer of convolution, and finally, a layer of maximum pooling is performed. The advantage of max pooling is that it reduces the problem of overfitting. If there is a pool: False in the convolution unit, it means that the layer does not perform pooling operations. The dropout operation is used in each fully connected layer to perform data expansion and prevent the network from overfitting. The characters of the output layer are selected as the character length of the data set text, and the softmax classifier is selected to classify the output results.

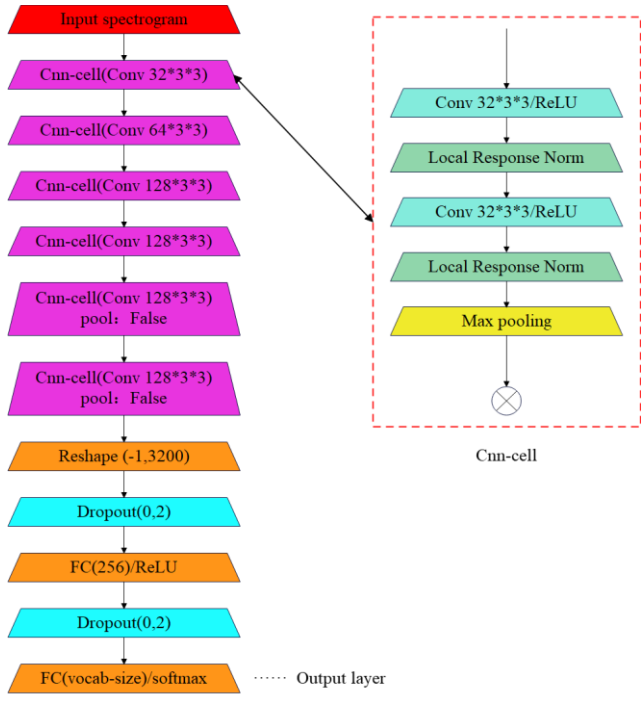

Figure 7: CNN network structure diagram for training the acoustic model.

During traditional acoustic model training, it is necessary to manually mark the labels of each frame of speech data, and iterative alignment of the speech data is required before training, which is very time-consuming. Therefore, this topic uses CTC as the loss function to train the acoustic model.

(1) Forward propagation

We set the forward variable as n(t, q), which represents the forward probability of neuron q at time t, where $q \in [1, 2U+1]$. Variables are initialized:

$$\begin{aligned} n(1,1) &= y_b^1 \\ n(1,1) &= y_1^1 \\ n(1,1) &= y, \forall_u > 2 \end{aligned} \tag{22}$$

From Equation 22, we can deduce:

$$n(t,q) = y_{i_q}^t \sum_{i=f(q)}^{u} n(t-1,i) \tag{23}$$

In formula 23, $f(q) = \begin{cases} q-1, l_q' = blank, or, l_{u-2}' = l_u' \\ q-2, otherwise \end{cases}$.

(2) Backpropagation

Similarly, we initialize the vector:

$$\begin{aligned} & m\ T, Q' = 1 \\ & m\ T, Q' - 1 = 1 \\ & m(T, q) = 0, \forall q < Q' - 2 \\ & n(1, q) = 0, \forall q > 2 \end{aligned} \tag{24}$$

From formula 24, we can deduce:

$$m(t, q) = \sum_{i=q}^{g(q)} m(t+1, i) y_{l_i'}^{t+1} \tag{25}$$

In formula 25, $g(q) = \begin{cases} q+1, l_q' = blank, or, l_{u+2}' = l_u' \\ q+2, otherwise \end{cases}$.

CTCloss uses the maximum likelihood function:

$$\begin{aligned} L(S) &= \sum_{(x,z)=S} L(x, z) \\ L(x, z) &= -\ln p(z \mid x) \end{aligned} \tag{26}$$

According to the above-mentioned forward and backward propagation variables, we can obtain:

$$p(z \mid x) = \sum_{q=1}^{|z|} n(t, q) m(t, q) \tag{27}$$

Thus, we get: $L(x, z) = -\ln \sum_{q=1}^{|z|} n(t, q) m(t, q)$.

(3) BP training of CTC

y_k^t represents the probability of outputting k at time t, and n_k^t represents the value of the corresponding output neuron k at time t before softmax transformation:

$$\frac{\partial L(x, z)}{\partial y_k^t} = -\frac{1}{p(z \mid x)} \frac{\partial p(z \mid x)}{\partial y_k^t} \tag{28}$$

At this time, we only need to calculate the propagation path through neuron k at time t to obtain:

$$\frac{\partial P(x, z)}{\partial y_k^t} = \sum_{q \in B(z,k)} \frac{\partial n(t, q) m(t, q)}{\partial y_k^t} \tag{29}$$

Among them, B(z, k) represents the set of neurons with k, and we can get:

$$n(t, q) m(t, q) = \sum_{\pi \in X(t,q)} \prod_{t=1}^{T} y_{\pi_t}^t \tag{30}$$

From this, we can get:

$$\frac{\partial P(x,z)}{\partial y_k^t} = \sum_{q \in B(z,k)} \frac{n(t,q)m(t,q)}{y_k^t} \tag{31}$$

Therefore, we can get the partial derivative of the loss function with respect to y_k^t :

$$\frac{\partial L(x,z)}{\partial y_k^t} = -\frac{1}{p(z \mid x) y_k^t} \sum_{u \in B(z,k)} n(t,q)m(t,q) \tag{32}$$

Similarly, we can find the partial derivative of n_k^t according to the loss function:

$$\frac{\partial L(x,z)}{\partial n_k^t} = y_k^t - \frac{1}{p(z \mid x)} \sum_{u \in B(z,k)} n(t,q)m(t,q) \tag{33}$$

3 RESEARCH ON JAPANESE TRAINING METHODS BASED ON CAD TECHNOLOGY AND DATA MINING

Japanese mobile learning system users have three roles: students, teachers, and administrators. This paper integrates the functions of the three roles to design four major modules: resource management, learning space, public modules, and system management, as shown in Figure 8.

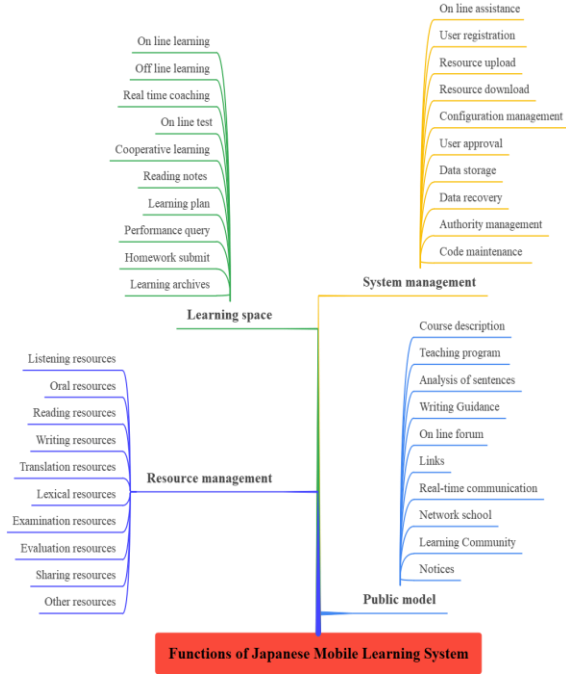

Figure 8: Japanese mobile learning system function.

The Japanese training method based on CAD technology and data mining proposed in this paper is combined with simulation teaching to verify the system performance, and the results shown in Figure 9 are obtained.

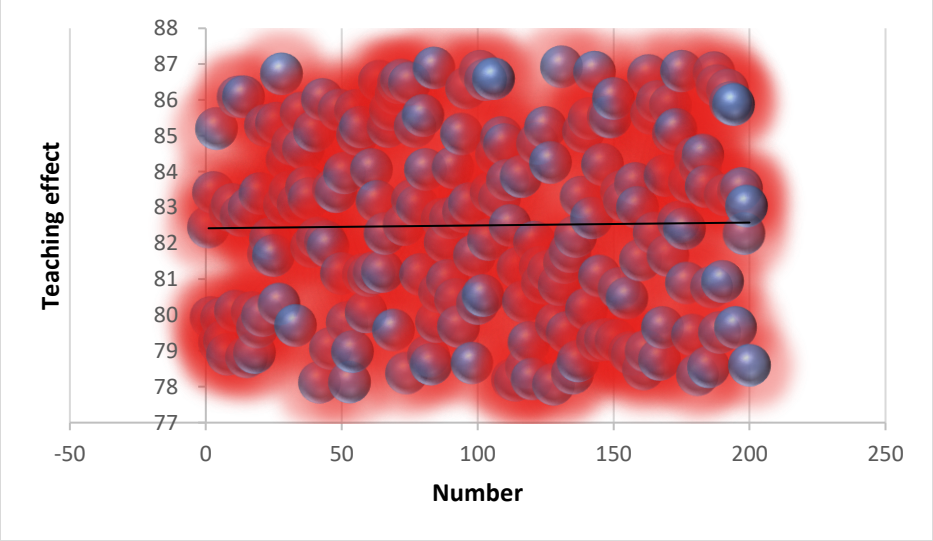

Figure 9: Verification of the effect of the Japanese training method based on CAD technology and data mining.

It can be seen from the above simulation teaching that the Japanese training method based on CAD technology and data mining proposed in this paper can effectively improve the effect of Japanese training.

4 CONCLUSIONS

The core problem of Japanese dependency analysis is to analyze whether there is a dependency between two adjacent texts. In the previous method, whether there is a dependency relationship is regarded as a binary classification problem, and the binary classifier of Support Vector Machines (SVM) is used for classification. Support vector machines have high accuracy for binary classification, nephew. There are some limitations to the problem of dependency parsing, such as the low classification accuracy of vectors near the classification hyperplane and the difficulty of considering the entire sentence information when extracting features. In recent years, conditional random field methods have begun to be applied in the field of natural language processing. This paper combines CAD technology and data mining technology to study Japanese training methods to improve the effect of Japanese learning. It can be seen from the simulation teaching that the Japanese training method based on CAD technology and data mining proposed in this paper can effectively improve the effect of Japanese training.

Nian Tong, https://orcid.org/0009-0006-4901-9986

REFERENCES

[1] Põldvere, N.; Johansson, V.; Paradis, C.: On the London–Lund Corpus 2: design, challenges and innovations, English Language & Linguistics, 25(3), 2021, 459-483. https://doi.org/10.1017/S1360674321000186

[2] Suárez, Z. B.; Gallardo-del-Puerto, F.; Gandón-Chapela, E.: The Primary Education Learners' English Corpus (PELEC): Design and compilation, Research in Corpus Linguistics, 8(1), 147-163. https://doi.org/10.32714/ricl.08.01.09
[3] Wang, C.; Zhao, Y.; Sun, D.: Research on Design and Sharing of Yi Language Corpus Resources Database Based on Syntactic Rules, Solid State Technology, 63(5), 2020, 10563-10574. https://doi.org/10.23977/jsoce.2021.030404
[4] Snaith, M.; Conway, N.; Beinema, T.; De Franco, D.; Pease, A.; Kantharaju, R.; Pelachaud, C.: A multimodal corpus of simulated consultations between a patient and multiple healthcare professionals, Language resources and evaluation, 55(4), 2021, 1077-1092. https://doi.org/10.1007/s10579-020-09526-0
[5] Esplà-Gomis, M.; Sentí, A.: Presentació del monogràfic «Spoken Corpus Linguistics in Romance: thoughts, design and results.» Caplletra, Revista Internacional de Filologia, (69), 2020, 117-123. https://doi.org/10.7203/caplletra.69.17266
[6] Alkhalifa, A. A. A.; Elhassan, I. B. M.: Corpus-Based, Genre-Analytic Approach to Discipline-Specific Materials Design and Development, Bulletin of Advanced English Studies–Vol, 3(1), 2019, 34-43. https://doi.org/10.31559/baes2019.3.1.4
[7] Etter, M.; Colleoni, E.; Illia, L.; Meggiorin, K.; D'Eugenio, A.: Measuring organizational legitimacy in social media: Assessing citizens' judgments with sentiment analysis, Business & Society, 57(1), 2018, 60-97. https://doi.org/10.1177/0007650316683926
[8] Trottier, D.: Scandal mining: political nobodies and remediated visibility, Media, Culture & Society, 40(6), 2018, 893-908. https://doi.org/10.1177/0163443717734408
[9] Assumpção, T. H.; Popescu, I.; Jonoski, A.; Solomatine, D. P.: Citizen observations contributing to flood modelling: Opportunities and challenges, Hydrology and Earth System Sciences, 22(2), 2018, 1473-1489. https://doi.org/10.5194/hess-22-1473-2018
[10] Chatterjee, S.; Kar, A. K.; Mustafa, S. Z.: Securing IoT devices in smart cities of India: from ethical and enterprise information system management perspective, Enterprise Information Systems, 15(4), 2021, 585-615. https://doi.org/10.1080/17517575.2019.1654617
[11] Grossi, V.; Rapisarda, B.; Giannotti, F.; Pedreschi, D.: Data science at SoBigData: the European research infrastructure for social mining and big data analytics, International Journal of Data Science and Analytics, 6(3), 2018, 205-216. https://doi.org/10.1007/s41060-018-0126-x
[12] Pangrazio, L.; Sefton-Green, J.: The social utility of 'data literacy'. Learning, Media and Technology, 45(2), 2020, 208-220. https://doi.org/10.1080/17439884.2020.1707223
[13] Walsh, J. P. (2020). Social media and moral panics: Assessing the effects of technological change on societal reaction, International Journal of Cultural Studies, 23(6), 840-859.
[14] Wiggins, A.; Wilbanks, J.: The rise of citizen science in healthcare and biomedical research, The American Journal of Bioethics, 19(8), 2019, 3-14. https://doi.org/10.1080/15265161.2019.1619859
[15] Budiharto, W.; Meiliana, M.: Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis, Journal of Big Data, 5(1), 2018, 1-10. https://doi.org/10.1186/s40537-018-0164-1
[16] Berendt, B.; Littlejohn, A.; Blakemore, M.: AI in education: learner choice and fundamental rights. Learning, Media and Technology, 45(3), 2020, 312-324. https://doi.org/10.1080/17439884.2020.1786399
[17] Ferguson, K.; Caplan, A.: Love thy neighbor: allocating vaccines in a world of competing obligations, Journal of Medical Ethics, 47(12), 2021, e20-e20. https://doi.org/10.1136/medethics-2020-106887