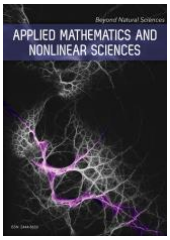

Applied Mathematics and Nonlinear Sciences

https://www.sciendo.com

A study on the influence of the spread of Yangming Studies in Japan on the psychology of the Japanese people based on big data analysis

Hongyan Liu†

Suqian University, School of Foreign Studies, Suqian, Jiangsu, 223800, China

Submission Info

Communicated by Z. Sabir
Received March 19, 2022
Accepted September 7, 2022
Available online April 10, 2023

Abstract

The analysis of the psychological impact of the spread of Yangming studies in Japan on the Japanese people is to enable Yangming studies to be better developed in Japan. Based on big data analysis technology, this paper constructs a hybrid data analysis model using the EM algorithm and proposes performance evaluation indexes for the model. Under the EM data analysis model constructed in this paper, the example indicators of the Japanese people's psychological impact in disseminating Yangming studies by big data analysis are explored, i.e., the psychological acceptability of the dissemination method and the psychological and moral construction impact. Regarding the dissemination method, the Japanese people are more receptive to disseminating Yangming studies in Japan through "learning rules", with an average percentage of 39.37%. Regarding psychological and moral construction, 90.22% of the Japanese people believe that disseminating Yangming studies can promote self-improvement of value standards and correct self-examination. Based on the big data analysis, we can effectively see from the data the impact of Yangming studies on the audience in the process of dissemination, and improve the scope of Yangming studies dissemination according to the data feedback, so that more people can recognize the idea of unity of knowledge and action.

Keywords: Big data analysis; EM algorithm; Yangming studies; Knowledge and action; Japanese people
AMS 2020 codes: 62-07

†Corresponding author.
Email address: yanyanhong20221123@163.com

ISSN 2444-8656
https://doi.org/10.2478/amns.2023.1.00083

1 Introduction

As a member of the "Confucian cultural circle," Japan has been widely and profoundly influenced by Confucianism. In particular, Confucianism, including Yangming, played an important role in the disintegration of feudal thought and the enlightenment of modern thought in Japanese society during the Edo period [1].

Yangming School emerged in China at the end of the 15th century as a philosophical doctrine based on the criticism of Cheng and Zhu, with the tenets of "mind as the reason", "to conscience", and "unity of knowledge and action". It is a philosophy based on the criticism of Cheng and Zhu. Its internal concern for human nature and its pursuit of freedom and equality make Yangming's philosophy more operable and practical []-[3]. At the same time, Yangming's philosophy of mind was spread to Japan at the end of the 16th century and profoundly influenced Japanese thinkers, who adopted Yangming's theories of "mental unity," "self-respect and fearlessness," "conscience," and "the unity of knowledge and action," which emphasize "sincerity" and "practice" []-[5]. In the era of big data, using big data analysis to study the influence of Yangming studies on the psychological aspects of the Japanese people is an alternative approach. By reasonably using the dividends of technological development in the era of big data, we can better create a genuine cultural atmosphere for the development of Yangming studies and let the world feel the profoundness of Chinese Confucianism [6].

The influence of Yangming's school as a classical Confucian doctrine can be considered profound. The literature [7] argues that Wang Yangming's state of mind provides an attempt to explore administrative philosophy from the perspective of the state of mind, combining Confucian patriarchal thought and ethics, and advocates administrative philosophy from Wang Yangming's perspective to study two aspects of administrative philosophy research content of values and mind, administrative philosophy, values and human hypothesis and administrative life, interest relations and philosophy of administrative life. The literature [8] argues that the Taizhou school could make Yangming's theory of mind popular in the world and also make itself an important place in the history of Chinese thought, which benefited not only from its philosophical characteristics of civilian Confucianism but also from the practice of civilian education. The literature [9] compares Yangming's studies with traditional philosophical interpretations, which link Yangming's philosophy with politics, education, law, and psychology as a new and excellent research paradigm with diversity, depth, integration, variety, and fusion.

And the literature [10] argues that Yangming Xinxue coalesces the rich ideological connotation and spiritual value of historical inheritance, which is the power source to enrich and preserve the spirit of Guizhou. In the context of the new era, an innovative interpretation of Yangming's Xinxue and Yangming's spirit, and an openness based on the return to cultivate and promote the spirit of Guizhou in the new era, is the transformation and development of contemporary excellent traditional culture. The literature [11] advocates a new interpretation of the unity of knowledge and action and true knowledge in Yangming Xinxue, according to which the concept argues that true knowledge needs to be free from some form of conflict. The literature [12] argues that from a social-psychological perspective, the Japanese school of Yangming affects the level of social activity in different students and can be effective in improving depression and anxiety, enhancing the positive psychological qualities of students.

This paper examines the impact of the spread of Yangming studies in Japan on the psychology of the Japanese people based on big data analysis in three parts. The first part introduces the concepts related to big data and big data analysis and elaborates on the structure of big data and the algorithms related to big data analysis. Based on the big data algorithm, the EM algorithm, a big data analysis algorithm,

is proposed, the principle of the EM algorithm is explained in detail, and the convergence of the EM algorithm is discussed. And the EM algorithm is iteratively optimized by parallel computing with the Dask parallel computing library, and a hybrid data analysis model is constructed based on the EM algorithm, and the performance evaluation of the model is also carried out using the ten-fold cross-validation method. The second part provides a comprehensive introduction to Yangming Studies, including the three theoretical components of Yangming Studies, i.e., mind as reason, knowledge as action, and to conscience, and illustrates the development process of Yangming Studies in Japan and the specific features of Yangming Studies. The third part discusses a hybrid data analysis model constructed based on the EM algorithm under big data technology and explores two aspects of the psychological impact of Yangming studies on Japanese people in the process of dissemination in Japan, namely, the psychological receptivity of the dissemination method and the psychological, moral construction impact. The data mining of the two aspects is used to explore the psychological impact of Yangming studies on the Japanese people in the dissemination process in Japan by big data analysis.

2 Big data analysis

Big Data, or Big Data, refers to information that is so large that it cannot be captured, managed, processed, and organized in a reasonable amount of time by mainstream software tools to help businesses make more positive business decisions.

2.1 Overview of Big Data

Big data is a data collection that is so large that it greatly exceeds the capability of traditional database software tools in terms of the acquisition, storage, management, and analysis, and has four major characteristics: massive data scale, rapid data flow, diverse data types, and low-value density. The process of data mining analysis carried out by Big Data is shown in Figure 1 [13].

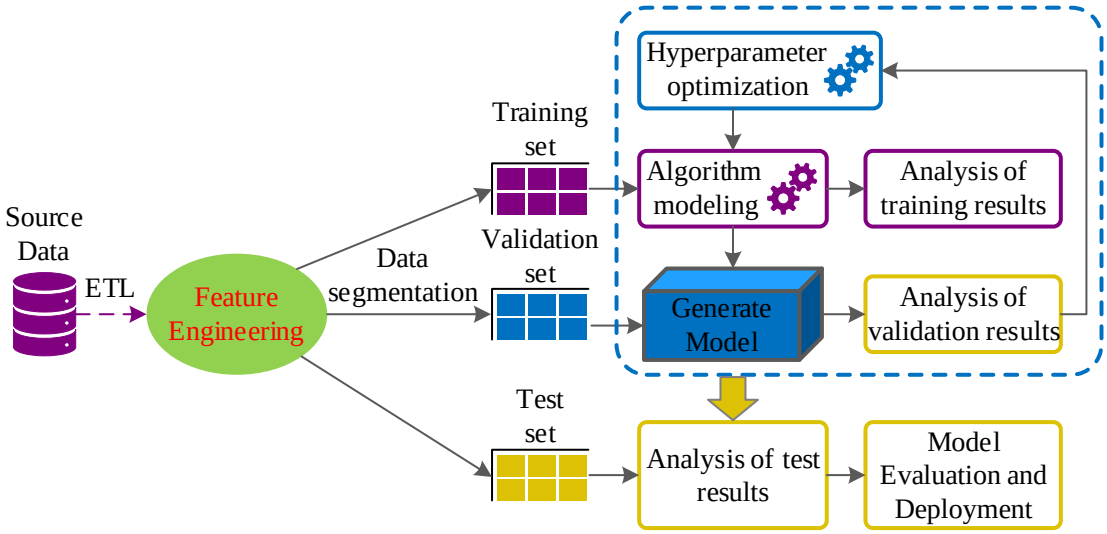

Figure 1. The process of big data mining and analysis

Big data requires special technologies to process large amounts of data within tolerable elapsed time efficiently. Technologies applicable to Big Data include massively parallel processing databases, data mining, distributed file systems, distributed databases, cloud computing platforms, the Internet, and scalable storage systems.

2.1.1 Structure of Big Data

Big Data includes structured, semi-structured and unstructured data, and unstructured data is increasingly becoming a major part of data. To systematically cognize Big Data, it must be comprehensively and meticulously decomposed, starting from three levels:

The first level is the theory, which is the necessary way of cognition and the baseline to be widely recognized and disseminated. Here, we understand the industry's overall depiction and characterization of big data from the definition of its characteristics; we analyze the preciousness of big data from the discussion of its value; we gain insight into the development trend of big data, and we examine the long-lasting game between people and data from the special and important perspective of big data privacy.

The second level is technology, which is the means to embody the value of big data and the cornerstone for moving forward. The process of big data from collection, processing, and storage to result formation is illustrated here from the development of cloud computing, distributed processing technology, storage technology and perception technology, respectively.

The third level is practice, which is the ultimate value of big data. Here are four aspects of big data, namely, big data of the Internet, big data of the government, big data of the enterprises and big data of the individuals, to depict the beautiful scenery that big data has shown and the blueprint that will be realized.

2.1.2 Big Data Algorithms

With a given resource constraint, an algorithm can generate an algorithm that satisfies a given constraint result within a given time constraint, using big data as input.

Big data algorithms cannot be exact, in-memory, serial algorithms that run only on electronic computers, and not only algorithms for cloud computing, parallel programming architecture models, big data analysis and mining [14]. The commonly used big data algorithms are shown in Figure 2.

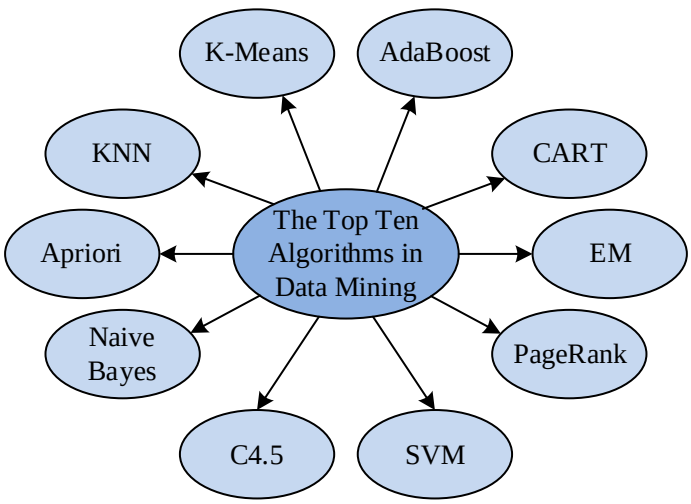

Figure 2. Ten big data analysis algorithms

This paper focuses on the application of EM algorithms in big data analysis in the context of the characteristics of the psychological impact of the spread of Yangming studies in Japan on the Japanese population.

2.2 Big data analysis algorithm - EM algorithm

2.2.1 Principle of EM algorithm

The EM algorithm, the expectation maximization algorithm, was first proposed as an iterative algorithm to solve the problem of estimating probabilistic models that contain observed not only variables but also hidden or latent variables, which cannot be estimated using the maximum likelihood estimation method.

In brief, the iterative step of the EM algorithm consists of an E step, which calculates the expected value of the potential variable from the known parameter values, and an M step, which substitutes the value derived from the E step into the likelihood function and then finds the estimate that enables the likelihood function to reach the maximum value point. Specifically, let $p(\theta|Y)$ denote the density function of the posterior distribution θ given the observed data Y, then $p(\theta|Y,Z)$ denotes the density function of the posterior distribution θ obtained after adding data Z, Y called the incomplete data set, and after adding data Z, (Y,Z) the complete data set is formed. The objective is to calculate the plurality of the posterior distribution $p(\theta|Y)$ and therefore the steps of the EM algorithm are as follows:

Assuming that $\theta^{(t)}$ is the estimate of the posterior plurality at the beginning of the t nd iteration, the step of the $t+1$rd iteration is:

Step E: The conditional distribution of $p(\theta|Y,Z)$ or $\log p(\theta|Y,Z)$ with respect to Z is expected so that Z is producted off, i.e.:

$$\begin{aligned} Q(\theta|\theta^{(t)},Y) \text{ A } & E_Z[\log p(\theta|Y,Z)|\theta^{(t)},Y] \\ & = \int \log[p(\theta|Y,Z)]p(Z|\theta^{(t)},Y)dZ \end{aligned} \tag{1}$$

Step M: Find a point 3 such that $Q(\theta|\theta^{(t)},Y)$ reaches the maximum value of θ, i.e., find a point $\theta^{(t+1)}$ such that:

$$Q(\theta^{(t+1)}|\theta^{(t)},Y) = \max_{\theta} Q(\theta|\theta^{(t)},Y) \tag{2}$$

The above process completes one iteration $\theta^{(t)} \to \theta^{(t+1)}$, and the E and M steps of the above process are repeated until $|\theta^{(t+1)} - \theta^{(t)}\|$ or $\|Q(\theta^{(t+1)}|\theta^{(t)},Y) - Q(\theta|\theta^{(t)},Y)\|$ reach the convergence criteria, then the iteration is stopped [15].

2.2.2 Convergence of EM algorithm

The EM algorithm has the good advantages of simplicity and stability, and using the EM algorithm also requires an understanding of its convergence properties, which facilitates better utilization of the EM algorithm. When the observed data contain latent variables or missing data, the probability density function of the complete data y_c can be expressed as [16]:

$$f(y_c|\theta) = f(y_o, y_m|\theta) = f(y_o|\theta)f(y_m|y_o;\theta) \tag{3}$$

The logarithmic maximum likelihood function for incomplete data y_o is:

$$\log(\theta; y_o) = \log f(y_o \mid \theta) = \log f(y_o, y_m \mid \theta) - \log f(y_m \mid y_o; \theta) \tag{4}$$

Taking the expectation on the conditional probability density function $f(y_m \mid y_o; \theta)$ for both sides of the above equation simultaneously yields:

$$\begin{aligned} \log f(y_o \mid \theta) &= E_\Psi[\log f(y_o, y_m \mid \theta) - \log f(y_m \mid y_o; \theta)] \\ &= Q(\theta \mid \hat{\theta}^{(k)}) - H(\theta \mid \hat{\theta}^{(k)}) \end{aligned} \tag{5}$$

Among them,

$$Q(\theta \mid \hat{\theta}^{(k)}) = \int_\Psi \log f(y_o, y_m \mid \theta) f(y_m \mid y_o; \hat{\theta}^{(k)}) dy_m \tag{6}$$

$$H(\theta \mid \hat{\theta}^{(k)}) = \int_\Psi \log f(y_m \mid y_o; \theta) f(y_m \mid y_o; \hat{\theta}^{(k)}) dy_m \tag{7}$$

Obviously, the Q function in equation (6) exists as follows, namely:

$$Q(\hat{\theta}^{(k+1)} \mid \hat{\theta}^{(k)}) \geq Q(\hat{\theta}^{(k)} \mid \hat{\theta}^{(k)}) \tag{8}$$

Then the right end H function of equation (7) has:

$$\begin{aligned} H(\hat{\theta}^{(k)} \mid \hat{\theta}^{(k)}) - H(\theta \mid \hat{\theta}^{(k)}) &= \int_\Psi \log f(y_m \mid y_o; \hat{\theta}^{(k)}) f(y_m \mid y_o; \hat{\theta}^{(k)}) dy_m \\ &\quad - \int_\Psi \log f(y_m \mid y_o; \theta) f(y_m \mid y_o; \hat{\theta}^{(k)}) dy_m \\ &= \int_\Psi -\log\left(\frac{f(y_m \mid y_o; \theta)}{f(y_m \mid y_o; \hat{\theta}^{(k)})} \right) f(y_m \mid y_o; \hat{\theta}^{(k)}) dy_m \end{aligned} \tag{9}$$

According to the *Jessen* inequality principle, i.e., given a certain convex function $f(x)$, satisfying $E(f(x)) \geq f(E(x))$, this equation (9) can be further expressed as:

$$\begin{aligned} H(\hat{\theta}^{(k)} \mid \hat{\theta}^{(k)}) - H(\theta \mid \hat{\theta}^{(k)}) &= \int_\Psi -\log\left(\frac{f(y_m \mid y_o; \theta)}{f(y_m \mid y_o; \hat{\theta}^{(k)})} \right) f(y_m \mid y_o; \hat{\theta}^{(k)}) dy_m \\ &\geq -\log \int_\Psi \frac{f(y_m \mid y_o; \theta)}{f(y_m \mid y_o; \hat{\theta}^{(k)})} f(y_m \mid y_o; \hat{\theta}^{(k)}) dy_m \\ &= 0 \end{aligned} \tag{10}$$

Obviously,

$$H(\hat{\theta}^{(k)} \mid \hat{\theta}^{(k)}) \geq H(\hat{\theta}^{(k+1)} \mid \hat{\theta}^{(k)}) \tag{11}$$

Therefore, calculating the logarithmic maximum likelihood of the two iterations before and after gives:

$$\log l_M(\hat{\theta}^{(k+1)}; y_o) - \log l_M(\hat{\theta}^{(k)}; y_o) = Q(\hat{\theta}^{(k+1)} \mid \hat{\theta}^{(k)}) - Q(\hat{\theta}^{(k)} \mid \hat{\theta}^{(k)}) - (H(\hat{\theta}^{(k+1)} \mid \hat{\theta}^{(k)}) - H(\hat{\theta}^{(k)} \mid \hat{\theta}^{(k)})) \tag{12}$$

From equations (8), (9), (10), and (11) above, it follows that,

$$\log l_M(\hat{\theta}^{(k+1)}; y_o) \geq \log l_M(\hat{\theta}^{(k)}; y_o) \tag{13}$$

Where the equation holds when and only when $Q(\hat{\theta}^{(k+1)} \mid \hat{\theta}^{(k)}) = Q(\hat{\theta}^{(k)} \mid \hat{\theta}^{(k)})$, that is, when the Q function is maximized. According to equation (12), the maximum likelihood function $l_M(\theta; y_o)$ remains monotonically increasing when updating the parameter values. The EM algorithm ensures that the value of the maximum likelihood function converges to a stable point, which may be the maximum value when converging, but may also be the saddle point value. Therefore, when the likelihood function is multi-peaked, different initial values should be selected, the parameter convergence values under different initial values should be compared, and the parameter value corresponding to the maximum likelihood function value should be selected as the final estimate. The specific flow of the EM algorithm is shown in Figure 3 [17].

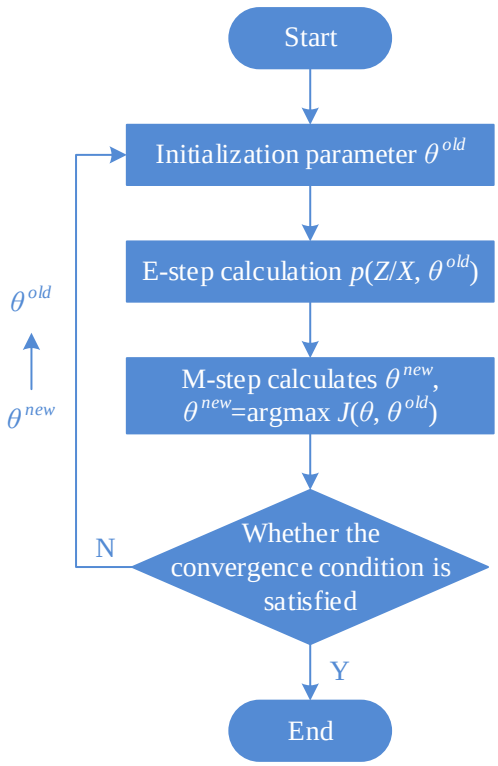

Figure 3. EM algorithm flowchart

2.3 EM algorithm optimization

2.3.1 Parallel computational optimization of EM algorithm

Dask is a parallel computing library in the Python language that efficiently streams data from disk using multi-core CPUs on a single machine, most notably featuring the ability to achieve efficient

parallel computation on a single machine for data sets that far exceed memory. To avoid memory overuse, Dask evaluates low memory usage computations by extracting blocks of data from the disk, performing the necessary processing and discarding intermediate values as soon as possible, when possible. This allows analysts to perform calculations on larger data sets (1000GB+), even on laptops with relatively low power. Since no configuration or setup is required to perform data analysis calculations with Dask, this also means that adding Dask to standalone calculations adds little to no cognitive overhead.

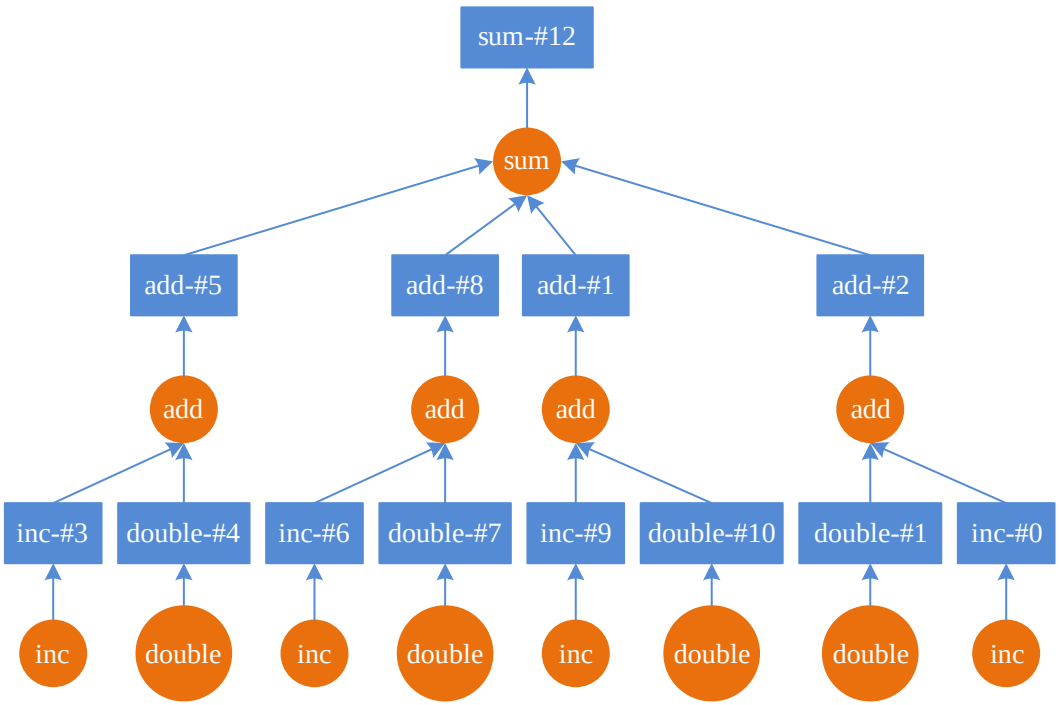

Figure 4. Dask execution task flowchart

As shown in Figure 4, the delayed method in Dask allows us to customize the parallelization algorithm as needed. The parallel EM algorithm implemented in this paper is a custom algorithm based on the Dask library, which provides a parallel dynamic task scheduling and execution task graph. Dask divides a task into multiple small tasks and draws a graph. According to this graph, some tasks can be executed simultaneously to perform dynamic parallel execution. The visualize method allows us to observe the execution flow of the algorithm, the simultaneous execution of the inc function and the double function and then the execution of the add function to obtain the computational results [18].

2.3.2 Hybrid data analysis model based on EM algorithm

A balanced set of samples is more conducive to improving the overall classification performance of the model. Data-level resampling methods can equalize the number of samples in each category through synthetic or extracted samples, but these methods are based on global sampling. Both theoretical and empirical studies have shown that due to the highly unbalanced sample size, characteristics significantly different in most categories may not be manifested in a few categories, also known as intra-category variation. The presence of intra-class variation can also affect the classification effectiveness of the classifier.

Algorithm level by boosting the misclassification cost for a small number of class samples or by selecting significantly different features for each class of samples. Although such an approach improves the model's classification accuracy in a few class samples to a certain extent, it also does not consider intra-class variability. To address the above problems, this paper proposes a hybrid data analysis model based on the EM algorithm by redefining the distribution form of the classification attributes present in the actual problem. The probability distribution of each classification overall is considered from the prior information, and the differences between each class and within the classification overall are described from the perspective of distribution, and the posterior probability under each classification is calculated according to Bayesian theory, and the classification of samples is attributed based on the criterion of maximum posterior probability.

For each independent observation sample x_i contains a potential variable Z_i (obeying a 0-1 distribution) to label its label, which is called the hidden variable. The mixture model considers that the overall distribution of a k classification problem can be considered as a mixture of k single independent distributions in a certain ratio [19]. It is assumed that a set of observed samples is generated by a mixture distribution p that consists of k independent identical distributions with different parameters, namely:

$$p = \prod_{i=1}^{k} \alpha_k f_k(x;\theta_k) \tag{14}$$

where α_k is the weight of the k nd independent distribution, $\sum_{i=1}^{k} \alpha_k = 1$, $f_k(x;\theta_k)$ is the density function of the k th independent distribution, and θ_k is the distribution parameter.

The mixture model now has two parameter vectors to be estimated, α_k and $\theta_k (k = 1,2,\cdots,K)$. Given a specific distribution, the mixture model is an infinite approximation of the true distribution of the sample data. When the specific form of the distribution is known, the estimate of the unknown parameter is usually given by the method of great likelihood estimation, i.e.:

$$\begin{cases} p(X \mid \theta) = \prod_{i=1}^{N} p(x_i \mid \theta) = L(\theta \mid X) \\ \hat{\theta} = \arg\max L(\theta \mid X) \end{cases} \tag{15}$$

Due to the presence of hidden variables, using the $\hat{\theta}$ equation in Eq. (15) to directly find the partial derivatives of the unknown parameters so that they are zero will result in multiple θ 's of the re-roots. Therefore, in this paper, the EM algorithm is used to solve for the unknown parameter vectors α_k and θ_k.

For the above-mixed data analysis model, the model parameters are randomly initialized, and the learning steps are:

STEP1: Randomly given the initial values of the parameters to be estimated α_0 and θ_0, which are the proportion of each category and the distribution parameters under the given distribution, respectively.

STEP2: Under the condition that the parameters are known, the estimation formula of the hidden variable is obtained, and the expectation of the hidden variable is $\hat{z}_{ik} = E\{z_{ik}|X;\theta\}$. Since Z_i obeys the 0-1 distribution, the expectation of Z_i is equivalent to the probability of $p(Z_i = 1)$, that is:

$$\hat{z}_{ik}^{(m)} = E\{z_{ik} \mid X_i;\hat{\theta}^{(m-1)}\} = p(z_{ik} = 1 \mid X_i;\hat{\theta}^{(m-1)}) = \frac{\hat{\alpha}_k^{(m-1)} f_k(X_i \mid \hat{\theta}^{(m-1)})}{\sum_{k=0}^{K} \hat{\alpha}_k^{(m-1)} f_k(X_i \mid \hat{\theta}^{(m-1)})} \tag{16}$$

STEP3: Under the condition that the estimate of the hidden variable $\hat{z}_{ik}^{(m)}$ is always known, i.e., the classification of the sample is known, the estimate of the parameter vector θ_k is obtained by maximizing the log-likelihood function $l(\theta \mid X_i;\hat{z}_{ik}^{(m)})$, taking the partial derivative of the parameter vector θ_k and making it 0, i.e.:

$$\hat{\theta}_k^{(m)} = \arg\max l(\theta \mid X_i;\hat{z}_{ik}^{(m)}) \tag{17}$$

STEP4: For extreme value problems that need to satisfy the constraint $\sum_{i=1}^{k} \alpha_k = 1$, the Lagrange multiplier method is often used to combine the likelihood function and the constraint to find the partial derivative of α_k and make it 0 to obtain the α_k estimate of the parameter, that is:

$$\hat{\alpha}_k^{(m)} = \sum_{i=1}^{n} \frac{\hat{z}_{ik}}{n} \quad (n \text{ is the number of samples}) \tag{18}$$

STEP5: Given $\varepsilon = 0.00001$, the iteration stops when $|\theta^m - \theta^{m-1}| < \varepsilon$, otherwise go to STEP2.

STEP6: The final obtained parameter estimates are noted as $\hat{\theta}_k$ and $\hat{\alpha}_k$, then the probability estimate of the ird observation sample being classified as class k is given as:

$$p(\hat{z}_{ik} = 1) = \frac{\hat{\alpha}_k f_k(X_i;\hat{\theta})}{\sum_{k=0}^{K} \hat{\alpha}_k f_k(X_i \mid \hat{\theta})} \tag{19}$$

The hybrid model based on the EM algorithm considers the distribution differences between its samples for the majority and minority class samples, respectively. Different weights are assigned to each class of samples, and the intra-class differences of each class are considered to improve the imbalance within each class of data and reduce the impact of intra-class imbalance on the classifier.

2.3.3 Mixed Data Analysis Model Evaluation Metrics

In the imbalance problem, more attention should be paid to the ability of the model to accurately predict the minority class rather than the global model because of the higher cost of misclassification of the minority class. The confusion matrix commonly used in big data analysis is also given, as shown in Table 1.

Table 1. Confusion matrix

	Positive sample	Negative samples
Positive sample	TP	FN
Negative samples	FP	TN

The confusion matrix allows the calculation of each evaluation index with the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{20}$$

$$Precision = \frac{TP}{TP + FN} \tag{21}$$

$$Recall = \frac{FP}{FP + TN} \tag{22}$$

Where, $Accuracy$ represents the accuracy of each class being classified, and the larger the value, the better the classification effect. $Precision$ represents the proportion of true positive samples among the samples predicted as positive classes. $Recall$ represents the proportion of samples that are incorrectly predicted to be correct among all negative samples.

2.3.4 Hybrid data analysis model evaluation analysis

In this paper, the performance of the hybrid data analysis model based on EM arithmetic is validated by the ten-fold cross-validation method with test and validation datasets, and the validation results are shown in Figure 5. The accuracy and precision of the model in this paper are maintained between 85% and 95%, and the recall rate is maintained between 15% and 20%. Among them, the average accuracy of ten data tests is 91.16%, the average precision is 87.73%, and the average recall is 17.66%.

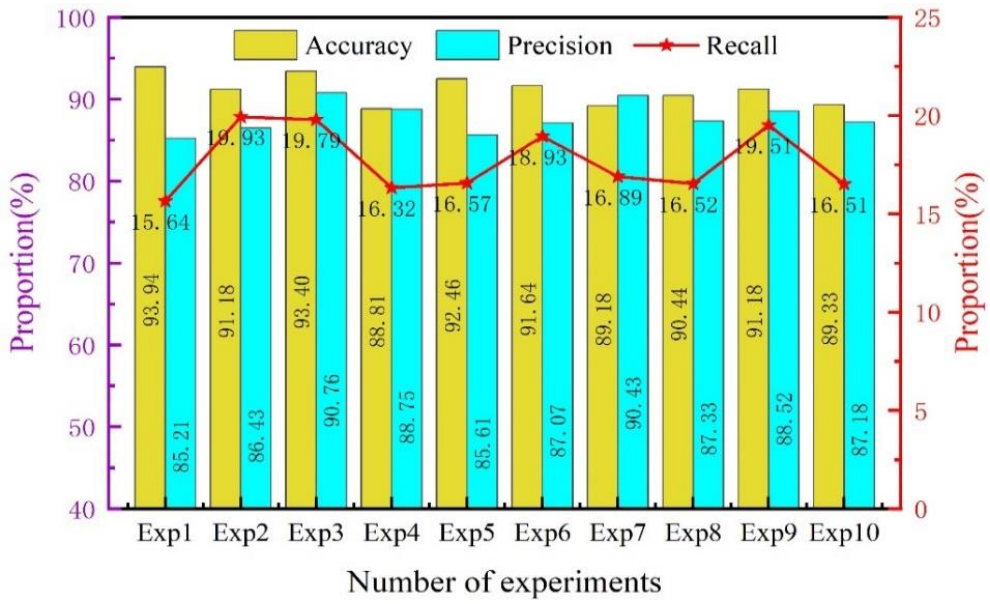

Figure 5. Model performance evaluation

This shows that the mixed data analysis model constructed based on the EM algorithm in this paper has good stability, and all the indexes meet the expected performance requirements of the model and can achieve the purpose of classification and regression for the data.

3 The psychological impact of the spread of Yangming studies in Japan on the Japanese people

3.1 Overview of Yangming Studies

There are three components of Yangming Studies, or Yangming Xinxue, which are mind as reason, knowledge as action, and to conscience. To analyze Yangming studies in depth, it is necessary to understand the components of Yangming studies **Error! Reference source not found.**.

When Wang Yangming was young, he was still in the world of Cheng Zhu's philosophy, and at that time, he advocated the idea of grasping things to get to the truth, believing that the truth comes through grasping things. However, Wang Yangming once spent seven days and seven nights in the bamboo, but he did not get any reasoning and became very sick. Later, when Wang Yangming was in the sarcophagus at Longchang, he realized that all things are in his heart, and all truths do not need to be sought from outside.

Knowing and acting in unity: "The work of knowing and acting cannot be separated from the original". Knowledge and action cannot be separated. If you only know but do not act, you still do not know.

It is easy to be conscientious: it is easy to know oneself. But if you can't get to that conscience, it's 'hard to know, but hard to do'. "It still depends on your sincerity and determination - one can become a saint, depending on whether one wants to become one or not. To be conscientious at all times, in Yangming's words, is called" to follow the matter.

3.2 The Development of Yangming Studies in Japan

It profoundly influenced Japanese thinkers, politicians, and Japanese society as a whole and gave a major impetus to the revolutionary movement and the introduction of Western learning at that time, so much so that it became the most important spiritual weapon of the Meiji Restoration and opened the way for Japan to enter modern times.

It is generally believed that the real pioneer of Yangming studies in Japan should be Nakae Fujiki (1608-1648), who studied Zhuzi studies in his early years, and at the age of 37, he was given the "Wang Longxi Quotations" to read, and later read the "Yangming Quanshu", and made great achievements, so he wrote a poem: "Although the study of knowledge and materialism is new, for eight years it has not yet been true; God bless the Xia Yang to Tai, now the heart seems to return to spring. " So he made his disciples study the Yangming Quanshu. Nakae Toshiki was regarded by later generations as the "Sage of Konoha" and wrote "Weng Q & A". He sought to popularize Yangming's study among the common people and asserted that "the study of the mind is the way from mortals to saints" and that "those who practice Confucianism are the sons of heaven, the vassals, the ministers and the common people. These are the five classes of people who can understand virtue. ..... True Confucianism does not choose between nobility and wealth among the five classes." This shows that Nakae Fujiki is the founding father of the Yangming school in Japan.

After its inception by Fujiki Nakae, the Japanese school of Yangming can be roughly divided into two schools: one school is the Deist school, which focuses on spiritual cultivation and introspective

character, centered on Fuchi Okayama, and faithfully inherits the subjective idealistic tradition of Yangming, without any academic achievements. The other school, centered on Kumazawa Fuyama, focused on the practical and action-oriented character, with the mission of transforming the world as its own. In general, the leftist school is the mainstream of the development of Japanese Yangming studies, and Japanese Yangming scholars do not specialize in theoretical and subtle arguments but have their characteristics in practice, which better represent the characteristics of Japanese Yangming studies. Japanese Yangming scholarship has gradually integrated national characteristics into the development process, as shown in Table 2.

Table 2. Characteristics of Japanese Yangmingology

Characteristics	Explanation
Human subjectivity	Japanese Yangmingology affirmed the role of "people" and encouraged people to eliminate the shackles of feudal ideology dogmatization.
"Time and place" theory	Scholars must open their minds based on their basic judgments of "time and place", be right and wrong, and become self-aware.
"Practice" theory	It is the duty to clarify the truth, practice practical learning, emphasize the unity of consciousness and behavior, and actively fulfill conscience and morality.
The idea of filial piety	All things in heaven and earth are born of filial piety, the ultimate source of the creation and change of all things in the universe.

3.3 Excavation and analysis of the influence of Yangming Studies on the psychology of the Japanese people

Because of the inherent logical requirements in Yangming's thought, it necessarily moved from theory to practice. Especially after its transmission to Japan, it was combined with the spirit of Japanese scholars' rebellion against authority and self-respect and fearlessness to form the unique Japanese Yangmingology. Therefore, the Japanese Yangmingology, which respects humanism, emphasizes practice, and focuses on the implementation of practical work, inspires the courage of its believers to break the shackles of Zhuzi's official thinking, and motivates them to accept and spread "foreign learning" and seek new strategies to save the country and the people. This paper constructs a hybrid data analysis model with the EM algorithm under the big data analysis algorithm and explores the examples of psychological influence on the spread of Yangming studies in Japan.

3.3.1 Analysis of the influence of Yangming Studies on the Japanese people's education and communication methods of excavation

Based on the big data analysis method, this paper collected many data from ten regions in Japan about the educational dissemination methods of Yangming studies in the process of dissemination among Japanese people and analyzed them through a hybrid data analysis model built by the EM algorithm. In disseminating Yangming theory in Japan, the dissemination methods adopted by the ten regions are different because the Japanese people have different psychological choices to accept them. The average percentages of the other two methods of Yangming theory dissemination, "lecture" and "textbook", are 27.82% and 19.33%, respectively. Among the ten places, the average psychological receptiveness of people in Tsuwano, Saga, and Hiroshima is higher, at 31.33%, 30.67%, and 30.94%, respectively, and the overall average psychological receptiveness is between 27% and 32%, indicating that Yangming theory is worth learning and spreading in the minds of Japanese people.

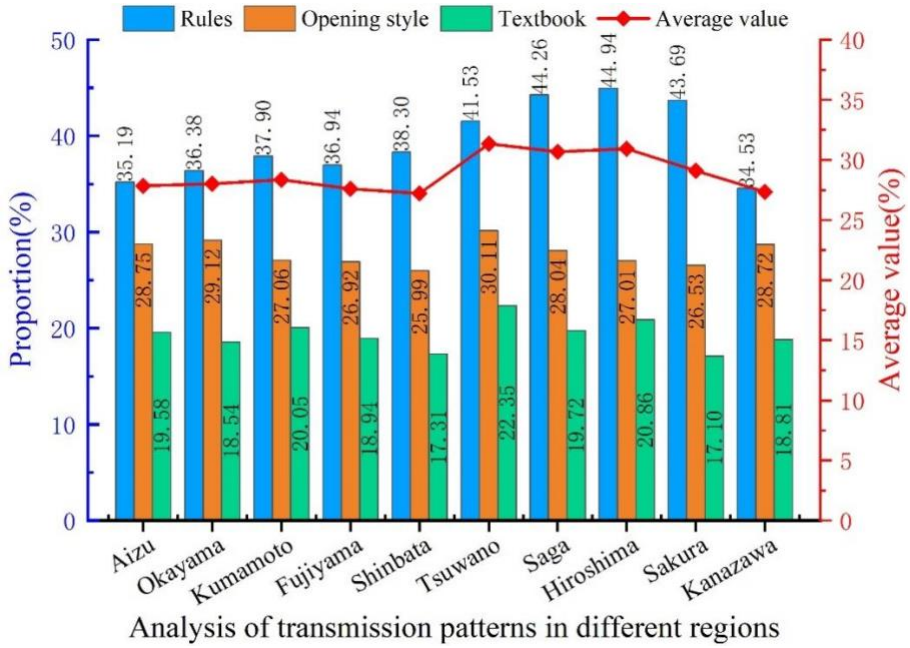

Figure 6. Excavation and analysis of the influence of Yangmingology on the psychology of Japanese people's education and study

The hybrid data analysis model based on the EM algorithm for Yangming theory under big data analysis can more clearly see the communication method with strong psychological acceptance of the people and choose the correct communication method more conducive to the dissemination of Yangming theory and more suitable for the psychological effect of the Japanese people.

3.3.2 Analysis of the influence of Yangming Studies on the construction of psychological and moral aspects of the Japanese people by excavation

Yangming studies helped Japan seek a balance between traditional cultural thought and the Western spirit of freedom and independence in the process of modernizing and reforming capitalism, and began to apply relevant ideas of Yangming studies in national moral education and patriotic education to achieve the purpose of regulating national thought and guiding the trend of national social values and moral development. In this paper, the influence of Yangming Studies on the psychological and moral construction of the Japanese people is analyzed by nine excavation items, as shown in Table 3.

Table 3. Indicators for the configuration of the structure of ideological and political courses

First-level indicators	Secondary indicators	Three-level indicator	Evaluation ratings				The metric number
			A	B	C	D	
Psychomoral construction	Discover conscience	This prompted the Japanese to begin to examine society and themselves					PC1
		Prompting Japanese people to discover their inner "heavenly reason."					PC2
		Prompting the Japanese to begin to return to their hearts					PC3
	Know the conscience	Improve the moral system of the individual's ontology					PC4
		Form the right criteria for value judgment					PC5
		Promote the sublimation of personal views of good and evil and values					PC6
	Practice conscience	Get rid of decadent negative thoughts					PC7
		Promote the harmonious development of people and conform to the principles of heaven					PC8
		The ideological trend of "harmony and foreign compromise" has been formed					PC9

Based on the examples of indicators provided in Table 3, a hybrid data analysis model based on the EM algorithm under big data analysis was applied to evaluate the impact of the development and dissemination of Yangming theory in Japan on the psychological and moral construction of the Japanese people, and the results of the psychological and moral construction analysis are shown in Figure 7.

Among the indicators of the psychological impact of the dissemination of Yangming theory on the Japanese people, 90.22% of the overall ratings are at grade C and above, with 36.87% of the A-grade ratings, indicating that the Japanese people are very satisfied with the ideological and psychological changes brought about by the theory of Yangming in the process of its dissemination. Among the A-level indicators, 46.92% believe that Yangming's theory has a significant role in forming correct value judgment standards among the Japanese and can help them establish correct mental and moral standards. Among the B-level indicators, 37.65% believe that the dissemination of Yangming theory has begun to psychologically motivate the Japanese to examine society and themselves, enabling them to define their values more clearly. Among the C-level indicators, 41.18% believe that Yangming's theory has not been effective enough in helping Japanese people to develop the trend of "harmony and compromise" and that the dissemination of this theory should be strengthened in the future. Among the D-level indicators, 17.49% and 14.60% believe that the dissemination of Yangming theory has not been effective in helping the Japanese people develop psychological harmony, nor has it helped them to return to their original minds.

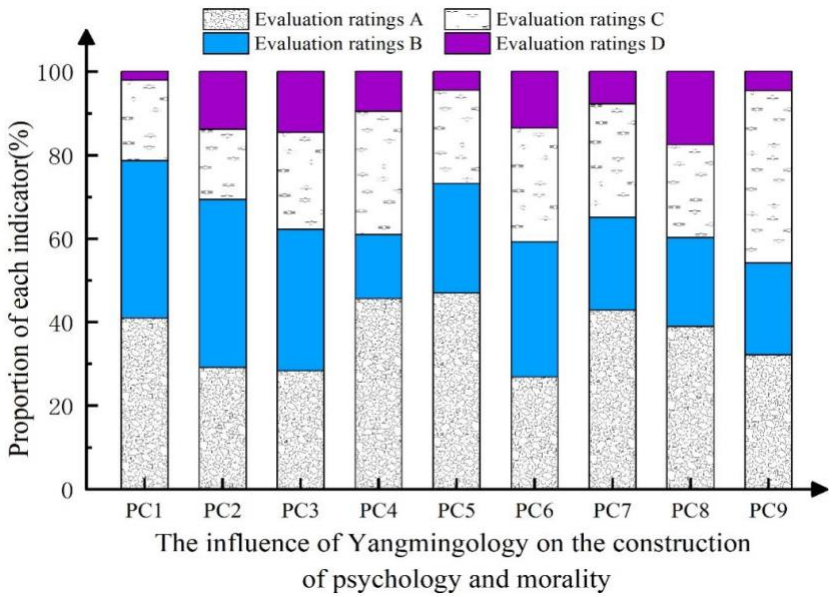

Figure 7. Analysis of the influence of Yangmingology on the construction of psychology and morality

The dissemination of Yangming studies in Japan based on big data analysis has had a greater psychological impact on the Japanese people as a whole, and some of the problems shown by the data need to be reformed in the subsequent dissemination process to promote the better dissemination of Yangming studies among the Japanese people.

4 Conclusion

This paper is based on a big data analysis algorithm and constructs a hybrid data analysis model based on the EM algorithm. The evaluation indexes of accuracy, precision and recall are proposed for the model, and the model's performance is verified by the ten-fold cross-validation method. Two indicator example mining analyses are also proposed for the psychological impact of the spread of Yang Ming theory in Japan on the Japanese people.

1) From the evaluation data, the accuracy and precision of the model in this paper are 91.16% and 87.73%, respectively, and the average recall rate is 17.66%, and all the indexes meet the expected performance requirements of the model. This shows that the mixed data analysis model constructed based on the EM algorithm in this paper has good stability and can achieve the purpose of classification and regression for the data.

2) As an example of the dissemination method, the psychological acceptance of Yangming theory in Japan is 39.37%, with the average percentage of psychological acceptance being 39.37%. This indicates that the development of Yangming theory in Japan can be better facilitated through the designation of rules.

3) In terms of the psychological, moral construction index, the overall percentage of those who rated C and above was 90.22%, among which the average percentages of A, B, C, and D ratings were 36.87%, 27.93%, 25.42%, and 9.78%, respectively. This indicates that the Japanese people psychologically recognize Yang Ming theory more and believe it can better help them form correct standards of value judgment and eliminate negative and decadent thoughts, enabling them to recognize better and discover themselves.

In summary, the process of Yangming Studies dissemination in Japan based on big data analysis can have a great impact on the psychology of the Japanese people, both in terms of the acceptance of the dissemination method and the promotion of the psychological and moral construction of the Japanese people, and Yangming Studies has indispensable benefits. Therefore, in the era of big data we should know how to use technical means to analyze the various influences of Yangming studies in the process of dissemination, which in turn can make the dissemination of Yangming studies wider and form a real atmosphere for the development of Yangming studies culture.

References

[1] Wang, Z. (2022). East Meets West: The New Gnoseology in Giordano Bruno and Wang Yangming. Religions, 13(9), 854.
[2] Ding, Q. (2022). Ideological characteristics of yangming studies in japan and its application in mental health from the perspective of social psychology. Psychiatria Danubina, 34(suppl 2), 148-148.
[3] Li, Y,. Zhang, H., Zhang, H. (2021). Analysis on The Combination of Yangming Psychology and Guizhou Primary School Moral Education Teaching. International Journal of Education and Economics, 13.
[4] Jia-Hui, X.U., Ding-Qian, L. I. (2014). Wang Yangming's Developing Track of Psychology in his Poems. Journal of Shanxi Datong University(Social Science Edition).
[5] Zhang, W. L., Yi-Duo, Y. E. (2017). A New Approach to the Development of Higher Vocational College Students' Psychological Capital-On Application of "Yangming" in Students' Cultural Education in Higher Vocational Colleges. Journal of Yulin Normal University.
[6] Yi, L. I., Wang, J. T., University T. (2019). Wang Yangming Xinxue and the Study of "Spirituality" in the Ming Dynasty. Journal of Mudanjiang University.
[7] Tang, C., University, G. N. (2018). Discussion on Administrative Philosophy in thePerspective of Psychology. Journal of Liuzhou Vocational & Technical College.
[8] Lai, L. (2020). On the Characteristics of Taizhou School's Educational Thought[J]. International Journal of Social Science and Education Research, 23(5), 163-167.
[9] Wei, C. (2017). The Study on Yangming Philosophy in the Republic of China. Journal of Guizhou Normal University(Social Sciences).
[10] Qinqin, L. I., Mengyi, H. E., Marxism, S. O., et al. (2019). Analysis on the Fit Degree and Value of Yangming Culture in the Spirit of Guizhou in the New Era. Journal of Xingyi Normal University for Nationalities.
[11] Lederman, H. (2022). The Introspective Model of Genuine Knowledge in Wang Yangming. Philosophical Review, 131(2), 169-213.
[12] Dong, J., Li K. (2022). The influence of the application of architectural environmental psychology in college architectural design on college students'mental health. Psychiatria Danubina, 34(suppl 2), 150-150.
[13] Song, C. W., Jung, H., Chung, K. (2019). Development of a medical big-data mining process using topic modeling. Cluster Computing, 22(1), 1949-1958.
[14] Davidson, R. (2020). Cyber-physical production networks, artificial intelligence-based decision-making algorithms, and big data-driven innovation in Industry 4.0-based manufacturing systems. Economics, Management, and Financial Markets, 15(3), 16-22.
[15] Beesley, L. J., Taylor, J. M. G. (2019). EM algorithms for fitting multistate cure models. Biostatistics, 20(3), 416-432.
[16] Guirao, J. L. G., Sabir, Z., Saeed T. (2020). Design and numerical solutions of a novel third-order nonlinear Emden–Fowler delay differential model. Mathematical Problems in Engineering.
[17] Ossevorth, F., Schegner, P. (2021). Approximating stochastic loads using the em-algorithm. IFAC Journal of Systems and Control, 18, 100175.
[18] Sabir, Z., Amin, F., Pohl, D., et al. (2020). Intelligence computing approach for solving second order system of Emden–Fowler model. Journal of Intelligent & Fuzzy Systems, 38(6), 7391-7406.
[19] Sabir, Z., Umar, M., Guirao, J. L. G., et al. (2021). Integrated intelligent computing paradigm for nonlinear multi-singular third-order Emden–Fowler equation. Neural Computing and Applications, 33(8), 3417-3436.

[20] Lederman, H. (2020). Conceptions of genuine knowledge in Wang Yangming. Oxford Studies in Epistemology.