Study on the university students’ satisfaction of the wisdom tree massive open online course platform based on parameter optimization intelligent algorithm

Chou-Yuan Lee¹, Ling-Ming Ruan¹, Zne-Jung Lee², Jian-Qiong Huang¹, Jie Yao¹, Zheng-Yuan Ning¹ and Jih-Fu Tu³
¹School of Big Data, Fuzhou University of International Studies and Trade, China
²School of Intelligent Construction, Fuzhou University of International Studies and Trade, China
³Department of Industrial Engineering and Management, St. John’s University

Abstract

Introduction: Curriculum learning through the wisdom tree massive open online course platform not only gets rid of the limitations of specialty, school and region, eliminates the limitations of time and space in traditional teaching, but also effectively solves the problem of educational equity.

Objectives: This paper proposes an intelligent algorithm combining decision tree, support vector machine, and simulated annealing to obtain the best classification accuracy and decision rules for university students’ satisfaction with the wisdom tree massive open online course platform.

Methods: This study takes the university students in Fuzhou city information management department as the survey object, and adopts the electronic questionnaire survey method. A total of 1136 formal questionnaires were responded, and 1028 valid questionnaires were obtained after data cleaning and deleting invalid questionnaires (the effective rate was 90.49%). In this paper, the reliability and validity of the questionnaire were tested by IBM SPSS-20.0 software, and six explanatory variables including function, achievement, exercise, quality, richness, and interaction were obtained by principal component analysis. Then, the questionnaire data is converted to CSV (comma separated values) format for analysis. This paper proposes an intelligent algorithm combining decision

Corresponding author:
Chou-Yuan Lee, School of Big Data, Fuzhou University of International Studies and Trade, No. 28, Yuhuan Road, Shouzhan New District, Changle, Fuzhou 350202, China.
Email: lqy@fzfu.edu.cn

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (https://us.sagepub.com/en-us/nam/open-access-at-sage).
tree, support vector machine, and simulated annealing to obtain the best classification accuracy and decision rules for university students’ satisfaction with the wisdom tree massive open online course platform. In this paper, the proposed algorithm is compared with decision tree, random forest, k-nearest neighbor, and support vector machine to verify its performance.

**Results:** The experimental results show that training set classification accuracy of decision tree, random forest, k-nearest neighbor, only support vector machine and the proposed algorithm (simulated annealing + support vector machine) are 92.21%, 96.10%, 95.67%, 97.29%, and 99.58%, respectively.

**Conclusion:** The proposed algorithm simulated annealing + support vector machine does increase the classification accuracy. At the same time, the 11 decision rules generated by simulated annealing + decision tree can provide useful information for decision makers.

**Keywords**
Wisdom tree massive open online course platform, simulated annealing, decision tree, support vector machine, k-nearest neighbor

**Introduction**

The massive open online course (MOOC) is a large-scale open online course, and a new learning mode of “Internet plus education.” The outbreak of COVID-19 makes the teaching of universities around the world face a severe challenge. In order to meet the new needs of teaching, the completion of courses is a huge challenge for MOOC.¹ The MOOC has been developing at an alarming speed, which has impacted the global higher education and promoted the construction and sharing of high-quality educational resources in the global scope.² Relying on the well-known EdX, Udacity, and Coursera platform in the United States, countries have successively built MOOC platforms according to their own conditions.³ The MOOC has set off a “digital tsunami” in China’s education field in the way of learning autonomy, universality of knowledge transfer, diversity of courses, and networking of curriculum mode. Platforms such as MOOC of Chinese universities, MOOC of Chinese language, and wisdom tree have emerged as the times require.

Through the wisdom tree MOOC platform to carry out curriculum learning, it is not only free from professional, school, and geographical restrictions, to eliminate the constraints of time and space in traditional teaching, but also to enrich the national education supply, narrow the regional, urban and rural and inter-school differences in education, so that the problem of education equity can be effectively solved.⁴

The wisdom tree teaching platform provides the teacher team function of related courses, abandons the mode of single teacher fighting alone, improves the quality of courses, supports a variety of teaching interaction processes, ensures the high interactivity of online teaching, and provides a full range of operation services for schools or institutions, teachers, students, and teaching alliances, as described below.⁵

1. To provide online university construction scheme for schools or institutions, realize online education operation at the school or institution level, provide corresponding services according to the requirements of curriculum teaching management of schools or institutions, provide course selection and credit certification services, and complete course promotion and enrollment tasks.
2. For teachers, wisdom tree network teaching platform can provide complete online and offline teaching and management of teacher service functions, complete course construction and support course teaching, and provide interactive teaching services based on social network services.

3. For students, online learning service is provided to complete the whole learning process of course selection, class, homework discussion and score credit, and social learning service is provided.

4. Aiming at the teaching alliance, we should provide course exchange services, build a perfect authentication mechanism, and provide public services for alliance members.

Although the wisdom tree teaching platform is popular among learners all over the world, its learning effect is not ideal compared with traditional teaching methods. Due to the large number of learners and the limited interaction time between teachers and students, it is impossible to give specific answers to each person’s situation. In addition, there are still some problems with the teaching effect of the wisdom tree platform. For example: some courses have very low click-through rates, many students give up halfway, and few learners can really gain knowledge.

As the users of wisdom tree platform, learners’ satisfaction with wisdom tree MOOC platform largely determines whether they choose to continue to use this platform. Data classification is an important research topic in machine learning and data mining, because the accuracy of the algorithm depends on the correctness of data classification. With the rapid development of computer, it is particularly important to use related technologies to establish models to analyze data in machine learning and data mining. Some methods for establishing MOOC prediction model are proposed in the literature, including k-nearest neighbor (K-NN), logistic regression, decision tree (DT), random forest (RF), and support vector machine (SVM). This paper proposes an intelligent algorithm combining DT, SVM, and simulated annealing (SA) to obtain the best classification accuracy and decision rules for university students’ satisfaction with the wisdom tree MOOC platform. The SVM has good classification performance and the DT can generate decision rules. The SA has the advantages of jumping off local optimization and reaching global optimization. In this paper, the SA is used to automatically adjust the parameters for SVM and DT to increase the classification accuracy and generate decision rules of university students’ satisfaction with the wisdom tree MOOC platform. The main purpose is to provide an effective analysis method for the satisfaction of the wisdom tree MOOC platform and useful information for decision makers.

The review of research methods

Decision tree

DT algorithm is a classic data mining algorithm. The structure of DT model is like a tree, including root node, leaf node, and non-leaf node. Each branch represents the direction of prediction, and the leaf node represents the final prediction result. Each node needs to
repeat the above process until it reaches the preset conditions. In this paper, the minimum Gini coefficient is used to select the classification attributes of internal nodes. The Gini index is selected as the splitting attribute, and finally the binary tree is generated. The Gini coefficient is used to represent the impurity of data set. To represent sample set $H$, its Gini coefficient can be expressed as follows:

$$Gini(T) = 1 - \sum_{i=1}^{n} P_i^2$$  \hspace{1cm} (1)

where $P_i$ represents the probability that the data in the sample set $H$ belongs to class $n$. If the sample set $T$ is divided based on the binary of an attribute $H$, it is subdivided into two subsets $T_1$ and $T_2$. Therefore, the Gini index based on the division can be calculated as follows:

$$Gini(T, H) = \frac{|T_1|}{|T|} Gini(T_1) + \frac{|T_2|}{|T|} Gini(T_2)$$ \hspace{1cm} (2)

where the $Gini(T, H)$ represents the uncertainty value of the set $T$ after $H$ partitioning. The larger the Gini index value, the greater the uncertainty result of the sample set. In the DT, the complexity parameter $(CP)$ and the minimum split $(M)$ are two very important parameters, which determine the classification accuracy.

**Random forest**

RF is formed on the basis of DT algorithm, which is composed of many DTs, but there is no correlation between each DT. Every time we encounter samples to be judged, we mainly follow the principle of putting them back, and put the extracted data samples on the root node of the DT to ensure that the relationship between trees is independent. Then, the DT discriminates according to the attribute category, and forms the corresponding result, and obtains the final result by the way of minority subordinate to majority.

**K-nearest neighbor**

The main idea of K-NN algorithm is that when there are $K$ nearest samples in the feature space, most of them are part of a specific category, then this sample is also part of this category. In K-NN algorithm, Euclidean distance is usually chosen as distance measure. The Euclidean distance between two points $A(x_1, y_1)$ and $B(x_2, y_2)$ is calculated as $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.

**Support vector machine**

The SVM finds a hyperplane that classifies a dataset and solves a classification problem well. SVM solves the problem of equation (3) with a given training patterns $(x_i, y_i), \; i = 1, 2, \ldots w, x \in R^z, y_i \in \{-1, +1\}$, the feature input of a multi-dimensional feature vector of $x_i$ in the $i^{th}$ pattern, the number of patterns of $w$, the $z$-dimensional real number space of $R^z$, and the output of $y_i$. The SVM solves the
problem shown in equation (3)

\[
\text{Max } L(\beta) = \sum_{i=1}^{w} \beta_i - \frac{1}{2} \sum_{i,j=1}^{w} \beta_i \beta_j y_i y_j (x_i, x_j)
\]

\[s.t. \ 0 \leq \beta_i \leq C, \text{ and } \sum_{i=1}^{w} \beta_i y_i = 0 \quad (3)\]

where \(\beta_i \geq 0\) denotes the Lagrange multiplier and \(C\) is a parameter of the cost of penalty. The feature space vectors \(x_i, x_j\) are constructed in terms of the kernel \(k\) where \(k(x_i, x_j) = (x_i, x_j)\). Using the feature space \(k(x_i, x_j) = x_i, x_j\), SVM can be expressed in equation (4)

\[
\text{Max } L(\beta) = \sum_{i=1}^{w} \beta_i - \frac{1}{2} \sum_{i,j=1}^{w} \beta_i \beta_j y_i y_j k(x_i, x_j)
\]

(4)

For the radial basis function, it can be expressed as \(k(x_i, x_j) = \exp (-\gamma ||x_i - x_j||^2)\). Two parameters \(C\) and \(\gamma\) must be appropriately set in SVM. It is necessary to set \(C\) and the \(\gamma\) parameters in SVM to achieve a balance between accuracy and brevity.22,23 In this study, the relevant algorithms are modeled and analyzed using R-4.1.0 language.

**Proposed algorithm**

In this paper, the SA is used to achieve the global optimization feature to automatically adjust the parameters of DT and SVM to increase the classification accuracy and generate decision rules. The proposed algorithm includes two stages, as shown in Figure 1.

**Questionnaire design**

In stage 1, because the quality of the course content design is related to whatever it causes students’ learning intent, the performance of the learning effect is also related to the curriculum design, the platform satisfaction is related to students’ willfulness to continue using MOOC.11,24 For the above reasons, the design of the questionnaire assessing the students’ satisfaction of using the wisdom tree MOOC platforms includes the following four aspects: personal basic information, course content design, learning effect, and platform satisfaction. It is described below:

1. Basic personal information includes gender, grade, age, and wisdom tree platform study time per week. (6 questions)
2. The course content design includes curriculum richness, online interaction, course quality, teachers’ teaching attitude, and teachers’ teaching method. (15 questions)
3. The learning effect includes the achievement of learning performance, the design of exercises after class and the learning interest. (8 questions)
4. The platform satisfaction includes the platform function provided by wisdom tree, the clear interface, and the ease of operation. (6 questions)
In addition to the basic personal information, the five point scale was designed for the influencing factors of university students’ satisfaction of the wisdom tree MOOC platform. Through a more intuitive expression, it makes it easier for the respondents to understand and fill in the questionnaire. The scale contains a series of influencing factors of university students’ satisfaction of the wisdom tree MOOC platform. The respondents answered by selecting five items in the five point scale: “very dissatisfied,” “dissatisfied,” “neutral,” “satisfied,” and “very satisfied.” The design of the five point scale has a convenient effect on the collection and processing of the questionnaire, which is convenient for the detailed and accurate analysis of each respondent’s feedback. The options of each questionnaire are 1 to 5. The 1 to 5 represent very dissatisfied, dissatisfied, neutral, satisfied, and very satisfied, respectively.
Questionnaire pre-test and reliability and validity analysis

About 120 students were randomly selected from the third grade of the information management department of Fuzhou University of International Studies and Trade to conduct a pre-test questionnaire. The main purpose of this pre-test is to test internal consistency when answering questionnaire questions. The IBM SPSS (version 20.0) was used to test the reliability and validity of the questionnaire data. In reliability analysis, Cronbach \( \alpha \) must be greater than 0.7.\(^{25}\) After the reliability analysis is completed, the validity of the questionnaire data needs to be verified. In this study, exploratory factor analysis was used to verify structural validity. In exploratory factor analysis, Kaiser Meyer Olkin (KMO) must be greater than 0.6. The greater the KMO value, the better the validity.\(^{26}\) The criterion for Bartlett’s sphericity test is that the corresponding \( \rho \) value is less than 0.01. The factor load coefficient criterion is higher than 0.4. After passing the verification standards of KMO and Bartlett, it is suitable to use principle component analysis (PCA) to extract factors. The PCA is used to obtain the components with the largest variance of the modified component. The PCA could be described as \( Y_i = \delta_i X \) where \( Y_i \) is the principal component, \( \delta_i \) is the eigenvalue of the sample covariance matrix, \( X = [x_1, x_2, \ldots, x_n]^T \) and \( x_i \) is an observed data vector. PCA could be rewritten as below.\(^{27-29}\)

\[
Y = \delta X
\]

where \( Y = [y_1, y_2, \ldots, y_n]^T \) is the principal component vector and \( \delta = [\delta_1, \delta_2, \ldots, \delta_n]^T \). When PCA extracts factors, the eigenvalues of each factor are set to be greater than 1, and the variance interpretation rate represents the amount of information of a certain factor.

In the pre-test questionnaire data of this study, there are 11 items, and each item is associated with the university students’ satisfaction of the wisdom tree platform. It can

| Aspects               | Items                                      | Cronbach’s \( \alpha \) | Overall reliability |
|-----------------------|--------------------------------------------|--------------------------|---------------------|
| The course content design | Curriculum richness                        | 0.913                    | 0.904               |
|                        | Online interaction                         | 0.909                    |                     |
|                        | Course quality                             | 0.902                    |                     |
|                        | Teachers’ teaching attitude                | 0.739                    |                     |
|                        | Teachers’ teaching method                  | 0.778                    |                     |
| The learning effect    | The achievement of learning performance    | 0.903                    |                     |
|                        | The design of exercises after class        | 0.922                    |                     |
|                        | The learning interest                      | 0.771                    |                     |
| The platform satisfaction | The platform function provided by wisdom tree | 0.921                    |                     |
|                        | The clear interface                        | 0.759                    |                     |
|                        | The ease of operation                      | 0.717                    |                     |
be found from Table 1 that the overall value of Cronbach’s $\alpha$ coefficient of the prediction test data is 0.904, which indicates that the data of the questionnaire pre-test has good reliability.

It can be found in Table 2 that the KMO values of the three aspects are higher than 0.7, and Bartlett’s spherical test results were $\rho < 0.01$, which indicates that the pre-test data of the questionnaire has good construct validity. At the same time, it can be seen in Table 2 that 11 items are extracted through PCA. Among them, six items have relatively high factor load coefficients, which are expressed in bold, including “curriculum richness,” “online interaction,” “course quality,” “the achievement of learning performance,” “the design of exercises after class,” and “the platform function provided by wisdom tree,” ranked in the top 6 and $>0.4$, and there is a strong correlation between the items, so these six items can be regarded as the explanatory variables are convenient for data analysis. The names of these six explanatory variables are richness, interaction, quality, achievement, exercise, and function. Table 3 is the six explanatory variables meaning and options by PCA extraction. The six explanatory variables extracted by PCA can better reflect a large amount of information related to the target variable, retain the internal relationship between the explanatory variables, and avoid analysis difficulties or management problems caused by too many factors.

### Reliability and validity analysis of formal data

This study takes the university students in Fuzhou city information management department as the survey object, and adopts the electronic questionnaire survey method. A total of 1136 formal questionnaires were responded, and 1028 valid questionnaires were obtained after data cleaning and deleting invalid questionnaires (the effective rate was 90.49%). Then, the data containing target variables and explanatory variables are transformed into CSV (comma

---

**Table 2. Validity test of pre-test data.**

| Aspects                  | KMO  | Bartlett’s ($x^2$) ($\rho < 0.01$) | Items                                                                 | Factor load coefficient |
|--------------------------|------|----------------------------------|----------------------------------------------------------------------|-------------------------|
| The course content design| 0.765| 509.682                          | **Curriculum richness**                                              | 0.796                   |
|                          |      |                                  | **Online interaction**                                               | 0.870                   |
|                          |      |                                  | **Course quality**                                                  | 0.811                   |
|                          |      |                                  | Teachers’ teaching attitude                                          | 0.544                   |
|                          |      |                                  | Teachers’ teaching method                                           | 0.589                   |
| The learning effect      | 0.707| 387.501                          | **The achievement of learning performance**                         | 0.848                   |
|                          |      |                                  | **The design of exercises after class**                             | 0.747                   |
|                          |      |                                  | The learning interest                                               | 0.664                   |
| The platform satisfaction| 0.765| 1356.228                         | **The platform function provided by wisdom tree**                   | 0.915                   |
|                          |      |                                  | The clear interface                                                 | 0.602                   |
|                          |      |                                  | The ease of operation                                               | 0.685                   |
separated values) format, which is convenient for modeling and analysis of relevant algorithms using R-4.1.0 language. The 1028 valid specific information are shown in Table 4.

In order to ensure that the formal data collected by the questionnaire is analytically reliable, the reliability and validity of the formal data collected by the questionnaire must be tested again. Table 5 is the reliability test results of the formal data, from which we can find the overall reliability test results of this study is 0.891, and the Cronbach’s $\alpha$ values of all variables are above 0.7, which indicates that the questionnaire item design is better and has good internal consistency. It shows that the formal data of this questionnaire has good reliability.

Table 6 is the reliability test results of the formal data. It can be seen from Table 6 that the validity test results of this study are KMO values $>0.8$, and Bartlett’s test meets the requirements $\rho<0.01$. It shows that the formal data of this questionnaire has good validity.

In stage 2, applying SA provides the best parameter settings for the DT and SVM in the proposed algorithm. The SA algorithm is a heuristic algorithm that simulates the physical process of cooling the classical particle system in thermodynamics. Kirkpatrick et al. first proposed the SA algorithm in 1983 and SA has been widely used in various optimization problems.\textsuperscript{30} Table 7 shows the proposed algorithm pseudo code used for this study. The initial values of parameters are set, and initial solution $N$ is randomly generated. Four parameters, namely $L_{\text{max}}$, $T_0$, $T_{\text{end}}$, and $\lambda$, where $L_{\text{max}}$ denotes the max number of generations, $T_0$ represents the initial temperature, $T_{\text{end}}$ represents the final temperature that stops the proposed algorithm if the current temperature is lower than $T_{\text{end}}$, and $\lambda$ is the coefficient controlling the cooling schedule, respectively. The current temperature $T$ is set the same as $T_0$. The solution represents six features and four variables $C$, $\gamma$, $CP$, and $M$ as shown in Figure 2. The values for each generation, randomly exchange these six features and randomly generate the values of four variables in the current solution $N$ to generate the next solution $N_{\text{new}}$. Run the $L_{\text{max}}$ generation and reduce $T$ according to the formula $T \leftarrow T$, where $0<\lambda<1$. Let $\text{obj}(N)$ denote the testing accuracy of $N$, and $\Delta F$ denote the difference between $\text{obj}(N)$ and $\text{obj}(N_{\text{new}})$; that is $\Delta F = \text{obj}(N) - \text{obj}(N_{\text{new}})$. The probability of replacing $N$ with $N_{\text{new}}$, where $N$ is the current solution and $N_{\text{new}}$ is

| Table 3. The explanatory variables and meaning by principal component analysis (PCA) extraction. |
|---------------------------------------------------------------|
| **Explanatory variables name** | **Meaning (The original question)** |
| Richness | Are you satisfied with the richness of courses offered on the wisdom tree? |
| Interaction | Are you satisfied with the online interaction (mutual evaluation and discussion) of the wisdom tree platform? |
| Quality | Are you satisfied with the course quality of the wisdom tree platform? |
| Achievement | Are you satisfied with the achievement of learning performance on the wisdom tree course? |
| Exercise | Are you satisfied with the design of exercises for after class assessment? |
| Function | Are you satisfied with the functions provided by the current wisdom tree learning platform? |
the next solution. If $\Delta F \leq 0$, the probability of replacing $N$ with $N_{\text{new}}$ is 1. Meanwhile, if $\Delta F > 0$, this is achieved by generating a random number $\theta_{\text{rand}} \in [0, 1]$ and replacing the solution $N$ with $N_{\text{new}}$ if $e^{-\Delta F/T} > \theta_{\text{rand}}$. In the proposed algorithm, SA and SVM are

**Table 4.** The 1028 valid specific information.

| Aspects                          | Variables | The number of samples | Proportion | Average score |
|----------------------------------|-----------|-----------------------|------------|---------------|
| The course content design        | Richness  | 1                     | 81         | 7.88%         | 2.95          |
|                                  |           | 2                     | 146        | 14.20%        |               |
|                                  |           | 3                     | 570        | 55.45%        |               |
|                                  |           | 4                     | 210        | 20.43%        |               |
|                                  |           | 5                     | 21         | 2.04%         |               |
| Interaction                      | 1         | 77                    |            | 7.49%         | 2.99          |
|                                  | 2         | 140                   |            | 13.62%        |               |
|                                  | 3         | 545                   |            | 53.02%        |               |
|                                  | 4         | 245                   |            | 23.83%        |               |
|                                  | 5         | 21                    |            | 2.04%         |               |
| Quality                          | 1         | 85                    |            | 8.27%         | 3.14          |
|                                  | 2         | 127                   |            | 12.35%        |               |
|                                  | 3         | 437                   |            | 42.51%        |               |
|                                  | 4         | 316                   |            | 30.74%        |               |
|                                  | 5         | 63                    |            | 6.13%         |               |
| The learning effect              | Achievement | 1                   | 85         | 8.27%         | 3.14          |
|                                  |           | 2                     | 143        | 13.91%        |               |
|                                  |           | 3                     | 422        | 41.05%        |               |
|                                  |           | 4                     | 300        | 29.18%        |               |
|                                  |           | 5                     | 78         | 7.59%         |               |
| Exercise                         | 1         | 77                    |            | 7.49%         | 2.95          |
|                                  | 2         | 136                   |            | 13.23%        |               |
|                                  | 3         | 574                   |            | 55.84%        |               |
|                                  | 4         | 215                   |            | 20.91%        |               |
|                                  | 5         | 26                    |            | 2.53%         |               |
| The platform satisfaction        | Function  | 1                     | 105        | 10.21%        | 2.93          |
|                                  |           | 2                     | 181        | 17.61%        |               |
|                                  |           | 3                     | 469        | 45.62%        |               |
|                                  |           | 4                     | 208        | 20.23%        |               |
|                                  |           | 5                     | 65         | 6.32%         |               |
| Target(outcome)                  | Satisfaction | 1                   | 78         | 7.59%         | 2.95          |
|                                  |           | 2                     | 137        | 13.33%        |               |
|                                  |           | 3                     | 569        | 55.35%        |               |
|                                  |           | 4                     | 224        | 21.79%        |               |
|                                  |           | 5                     | 20         | 1.95%         |               |

**Figure 2.** The representation of the solution.
used to optimize the parameters ($C$ and $\gamma$) to increase the test accuracy of the selected features, SA and DT are used to optimize the parameters ($CP$ and $M$) to build the decision rules. Repeat the algorithm until $T$ is less than $T_{end}$. Finally, the optimal test accuracy, and decision rules are obtained.

The standard for testing classification methods usually uses classification accuracy. This paper also utilizes precision, recall, F1-score as evaluation indicators.\textsuperscript{31} These performance indicators are calculated based on the confusion matrix. The confusion matrix is shown in Table 8.

The precision rate represents the proportion of positive instances in positive instances determined by the classifier. The precision rate calculation formula is shown in equation (6).

\[
\text{Precision} = \frac{TP}{TP + FP}
\]  

(6)

The recall rate represents the proportion of positive instances predicted to be positive instances. The recall formula is shown in equation (7).

\[
\text{Recall} = \frac{TP}{TP + FN}
\]  

(7)
Table 7. The pseudo code of the proposed algorithm.

Procedure: The pseudo code of the proposed algorithm

Begin
Set the initial temperature $T_0$ and the final temperature $T_{end}$;
Set the maximum number of iteration $L_{max}$ and the temperature cooling rate $\lambda$, $0 < \lambda < 1$;
Randomly generate an initial solution $N$;
While (the number of iteration $L < L_{max}$)
  While (the current temperature $T > T_{end}$)
    For (C loop)
      For (γ loop)
        For (CP loop)
          For (M loop)
            Generate a new solution $N_{new}$ from the current solution $N$;
            $\Delta F = obj(N) - obj(N_{new})$; /* $obj()$ represents classification accuracy */
            If $\Delta F \leq 0$ then
              Accept the new solution $N_{new}$ as the current solution;
            Else
              $P = \exp(-\Delta F/T)$;
              If $P > \theta_{rand}$ then
                Accept the new solution $N_{new}$ as the current solution;
              End if
            End if
            Update the classification accuracy and decision rules;
          End for
        End for
      End for
    End for
  End while
  $L = L + 1$;
End while
Output the best classification accuracy and decision rules
End Begin

Table 8. The confusion matrix.

| Predicted  | Actual positive | Actual negative |
|------------|-----------------|-----------------|
| Predicted positive | TP (true positive) | FP (false positive) |
| Predicted negative   | FN (false negative)  | TN (true negative)  |

where $TP$ is the number of instances that are positive and to be positive, $TN$ is the number of instances that are negative and to be negative, $FP$ is the number of instances that are actually negative but to be positive, and $FN$ is the number of instances that are positive to be negative.
F1-score is a measure of classification problems. It uses the harmonic average method to comprehensively consider the precision rate and the recall rate, the maximum is 1 and the minimum is 0. The F1-score calculation formula is shown in equation (8).

$$F1 \text{ score} = \frac{2 \times \text{Precision } \times \text{Recall}}{\text{Precision } + \text{Recall}} \quad (8)$$

In this paper, classification accuracy reflects the classifier’s ability to judge the entire instance as positive or negative. The classification accuracy is calculated with equation (9).

$$\text{Classification accuracy} = \frac{(TP + TN)}{(TP + FN + FP + TN)} \times 100\% \quad (9)$$

**Experimental results and discussion**

*Comparison of experimental results to other approaches*

In the proposed algorithm, the dataset obtained from university students’ satisfaction of the wisdom tree platform questionnaire was divided into 80% training dataset and 20% testing dataset, and 10 random cross-validation verifications were used to calculate the classification accuracy. In the proposed algorithm, the SA provides the best parameter settings for the DT and SVM. The SA parameters were set to $L_{max} = 5000$, $T_0 = 100$, $T_{end} = 0.1$, and $\lambda = 0.95$. The two parameters of SVM, the search interval of $C$ is between 0.01 and 5000, and that of $\gamma$ is between 0.01 and 5000. The searching range of the parameter $M$ of DT is between 2 and 100, and the searching range of parameter $CP$ of DT is between 0.01 and 0.5. To verify its performance, the proposed algorithm was used with the DT, RF, K-NN, and only SVM approaches, and the simulation results were compared. In this study, it is necessary to set the value of parameters for SVM and DT, and then use all the same values for fair comparison. The only SVM parameter of $C$ is set as 1, and $\gamma$ is set as 0.1 in this study. The only DT parameter of $CP$ is set as 0.1, and $M$ is set as 2 in this study.$^{32}$ The RF is an ensemble learning method for classification that constructs multiple DTs at training time, and outputs the class that depends on the majority of the classes. For RF, the number of the DT classifier is set as 500. The K-NN is a machine learning method to classify according to the distance between different feature values. The K-NN parameters used in this study were $K = 3$, and the Euclidean distance was used.

1. The experimental results in Table 9 include the classification accuracy calculated by only DT, RF, K-NN, only SVM, and the proposed algorithm according to equation (8). It can be seen from Table 9 that the classification accuracy the training set of only DT, RF, K-NN, only SVM, and the proposed algorithm are all greater than the classification accuracy of the testing set and the error is no more than 10%, and there is no over-fitting phenomenon.$^{33}$ From Table 9, it can be found that the classification accuracy of the training set of only SVM is
97.29%, and the classification accuracy of the testing set is 96.57%. In the algorithm proposed in this paper, the classification accuracy of the training set is 99.58%, and the classification accuracy of the testing set is 98.45%, indicating that the SA algorithm has an additive effect on SVM. The classification accuracy can be increased by adjusting the parameters. Because the SA algorithm has the advantage of jumping out of the local optimum according to the probability, it can effectively prevent the search process from falling into the local optimum. It shows that this paper proposed an intelligent algorithm that adds SA to SVM and DT, which can use the advantages of SA to effectively determine $C$ and $\gamma$ for SVM, $CP$ and $M$ for DT. The experimental results show that the classification effect of using SVM only is better than DT alone as SVM is a hyperplane classifier whereas, the DT starts from the root node and classifies from top to bottom, and the RF is the ensemble method of DT. Therefore, it is not surprising that these algorithms outperformed DT in classification predictions. The K-NN was used to calculate Euclidean distance in the two-dimensional plane, and the classification performance was not as good as the SVM classifier.

2. The recall reflects the classification model’s ability to recognize positive instances. The higher the recall, the stronger the model’s ability to recognize positive instances. The precision reflects the model’s ability to distinguish negative instances. The higher the precision, the stronger the model’s ability to distinguish negative instances. The F1-score is a combination of the two. The higher the

Table 9. Comparison of classification accuracy using different approaches.

| Machine learning methods | Classification accuracy (%) |
|--------------------------|-----------------------------|
|                          | Training set | Testing set |
| Only DT                  | 92.21        | 91.74       |
| RF                       | 96.10        | 93.22       |
| K-NN                     | 95.67        | 92.41       |
| Only SVM                 | 97.29        | 96.57       |
| The proposed algorithm   | 99.58        | 98.45       |

Table 10. The performance indicators of different approaches.

| Machine learning methods | Training set | Testing set |
|--------------------------|--------------|-------------|
|                          | Precision    | Recall      | F1-score    | Precision    | Recall      | F1-score    |
| Only DT                  | 0.9149       | 0.9226      | 0.9187      | 0.9105       | 0.9133      | 0.9119      |
| RF                       | 0.9337       | 0.9348      | 0.9342      | 0.9355       | 0.9350      | 0.9352      |
| K-NN                     | 0.9509       | 0.9486      | 0.9497      | 0.9262       | 0.9473      | 0.9366      |
| Only SVM                 | 0.9716       | 0.9645      | 0.9680      | 0.9648       | 0.9693      | 0.9670      |
| The proposed algorithm   | 0.9923       | 0.9942      | 0.9932      | 0.9677       | 0.9711      | 0.9694      |
F1-score, the more robust the classification model. It can be seen from Table 10 that the results of recall, precision, and F1-score are higher than of other methods, indicating that the classification effect of the proposed algorithm is generally better, and also shows that the proposed algorithm has good robustness.

3. In 2019, Yu et al. proposed predicting learning outcomes with MOOC Clickstreams while both the K-NN and SVM were used to generate prediction models and their model accuracies were 87.88%, 92.18%. In this paper, for the university students’ Satisfaction of the wisdom tree MOOC platform, the K-NN, SVM training set prediction accuracy rate is 95.67%, 97.29%. Despite using different datasets for comparison, both studies applied the same K-NN and SVM algorithms to build the prediction models. In addition, the proposed algorithm in our study has outperformed all other selected methods with the highest prediction accuracy (99.58%). Apparently, the prior use of SA + SVM provided automatic tuning and optimization of $C$ and $\gamma$ parameters in SVM, the ability to jump out of the local optimal trap, and finally further increase the classification accuracy in our classification prediction task. Given all these additional benefits, the purpose of obtaining a more effective machine learning algorithm for predicting the university students’ satisfaction regarding using the wisdom tree MOOC platform was achieved.

**Analysis of influencing factors**

In the proposed algorithm, the DT generated for the information management department university students in Fuzhou city satisfaction of the wisdom tree MOOC platform is shown in Figure 3. From Figure 3, the DT judges the value of the node according to different attribute values. Start from the root node, encounter branches on the way, until the last leaf node obtains a decision rule. Therefore, there are 11 decision rules for the survey of university students’ satisfaction of the wisdom tree platform. The rules indicate the degree of university students’ satisfaction of the wisdom tree platform and which factor affects satisfaction. The obtained decision rules have a total of 11 DT rules, which are shown in Table 11. The DT is divided with the explanatory variable name is “Function” as the root node, which shows that the platform function provided by the wisdom tree is the most important factor affecting university students’ satisfaction of the wisdom tree MOOC platform. It can be found from Table 11 that according to decision rules 3, 5, 6, 7, and 9, the overall satisfaction of university students’ satisfaction of the wisdom tree platform is four points, which means that they are satisfied with university students’ satisfaction of the wisdom tree MOOC platform. According to decision rules 8, 10, and 11, the overall satisfaction of university students’ satisfaction of the wisdom tree MOOC platform is five points, which means “very satisfied” with university students’ satisfaction of the wisdom tree MOOC platform. In order to illustrate the degree of the influence factors for the information management department university students in Fuzhou city satisfaction of the wisdom tree platform, Table 12 shows the results of using the influence factors value from IncNodePurity (increased node purity) on university students’ satisfaction of the wisdom tree MOOC platform. The IncNodePurity is an...
evaluation method that uses the non-negative sum of squares of the residuals to obtain the value. The size of the value can explain the degree of influence of the explanatory variable on the target variable. As shown in Figure 4, “Function” is the largest among the calculated values of the IncNodePurity. From Table 12 and Figure 4, it is pointed out that the influencing factors are Function>Achievement>Exercise>Quality>Richness>Interaction. It
can also be seen that the value of IncNodePurity of function is the largest, which indicates that the satisfaction of function has the greatest impact on university students' satisfaction of the wisdom tree MOOC platform. This is because for learners, a well-functioning learning platform can greatly reduce cognitive load and improve learning efficiency, so
the satisfaction of function ranks first. Moreover, the satisfaction of achievement ranks second because learners can not only learn knowledge and increase their knowledge level on the wisdom tree MOOC platform, but also hope to have good performance in the exam. In addition, the satisfaction of after-school exercise design ranks third because the after-school exercise can provide learners continuous practices for solving their learning problems and confusions during the class time.

Conclusions

In this study, IBM SPSS-20.0 software was used to test the reliability and validity of the questionnaire, and six explanatory variables such as function, achievement, exercise, quality, richness, and interaction obtained by PCA. This paper proposes an intelligent algorithm combining DT, SVM, and SA to obtain the best classification accuracy and decision rules for university students’ satisfaction with the wisdom tree MOOC platform. The experimental results show that training set classification accuracy of only DT, RF, K-NN, only SVM and the proposed algorithm are 92.21%, 96.10%, 95.67%, 97.29%, and 99.58%, respectively. At the same time, the 11 decision rules generated by SA + DT can provide useful information for decision makers.

The information management department students of each university in Fuzhou have a different level. Many students refuse to fill out the questionnaire, which limits the number of samples. Respondents fill in the information of satisfaction of the wisdom tree MOOC platform. Only the samples actually received can be analyzed and researched. In order to attract more university’s students to use the wisdom tree MOOC platform and provide decision makers with more useful information, more advanced methods should be used in the future.

This paper proposed methods and experimental results for the information management department university students in Fuzhou city satisfaction of the wisdom tree MOOC platform, hoping to improve university students’ satisfaction of the wisdom tree MOOC platform and promote the construction and development of the wisdom tree MOOC platform, the following suggestions are made to the wisdom tree from the result:

1. With the rapid development of information technology and the comprehensive popularization of Internet technology, university students’ have higher

| Influencing factors | IncNodePurity |
|---------------------|---------------|
| Richness            | 21.10         |
| Interaction         | 22.18         |
| Quality             | 38.41         |
| Exercise            | 58.98         |
| Achievement         | 95.27         |
| Function            | 116.01        |
requirements for the transmission speed of knowledge and information and the diversity of content. For university students’, a faster and more convenient learning platform can greatly reduce cognitive load and improve learning efficiency. The wisdom tree MOOC platform not only needs to provide high-quality learning resources, but also needs to provide better platform functions to increase the interest of university students in learning.

2. At present, to the best of our knowledge, most of the wisdom tree MOOC platforms in China still adopt a lecture-based teaching method. Simply displaying learning materials with text and video may not meet the needs of the majority of university students. In the course development, the wisdom tree MOOC platform should thoroughly explore the needs of the courses and learners; strengthen the content of the courses and improve the quality of the courses.

Acknowledgments

This research was supported in part by Fuzhou University of International Studies and Trade under Grant Nos. FWB19004, 2019-SG-2, 2019020079010, 2018KYTD-05, FWXXS19093, FJTPY-2020009, and BRJF-01.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Chou-Yuan Lee https://orcid.org/0000-0003-2786-2068

References

1. Calvo S, Lyon F, Morales A, et al. Educating at scale for sustainable development and social enterprise growth: the impact of online learning and a massive open online course (MOOC). Sustainability 2020; 12: 3247.
2. Gonda D, Šuriš V, Pavlovičová G, et al. Analysis of factors influencing students’ access to mathematics education in the form of MOOC. Mathematics 2020; 8: 1229.
3. Cirulli F, Elia G, Lorenzo G, et al. The use of MOOCs to support personalized learning, an application in the technology entrepreneurship field. Knowl Manag E Learn 2016; 8: 109–123.
4. Žur A. Two heads are better than one—entrepreneurial continuous learning through massive open online courses. Educ Sci 2020; 10: 62.
5. Sun F, Ji D and Liu X. Research on online and offline mixed teaching model based on “wisdom tree” platform taking course of “installation and debugging of automation production line” as an example. Wirel Internet Technol 2020; 01: 103–104.
6. Li Y. Study on practice of online and offline mixed teaching based on wisdom tree platform: taking course of “automatic production line” as an example. Wirel Internet Technol 2019; 21: 74–75.
7. Li S and Li XF. Application of blended learning based on wisdom tree in design and development of multimedia resources. *Wirel Internet Technol* 2019; 15: 125–127.

8. Shi L. Research on practice of online and offline mixed teaching based on wisdom tree platform — taking course of “principles of electrical appliances” as an example. In: IEEE 2020 International Conference on Artificial Intelligence and Education (ICAIE), 26–28 June 2020. https://doi.org/10.1109/ICAIE50891.2020.00025

9. Wan L, Xie S and Shu A. Toward an understanding of university students’ continued intention to use MOOCs: when UTAUT model meets TTF model. *SAGE Open* July 1, 2020; 10: 1–15.

10. Lu M and Li F. Survey on lie group machine learning. *Big Data Mining Analyt* 2020; 3: 235–258.

11. Kadlecšík M, Munk M and Munková D. The efficacy of MOOC to support students in pedagogical research. *Appl. Sci* 2021; 11: 328.

12. Xiao B, Liang M and Ma J. The Application of CART Algorithm in Analyzing Relationship of MOOC Learning Behavior and Grades. In: IEEE 2018 International Conference on Sensor Networks and Signal Processing (SNSP), 28–31 Oct. 2018. https://doi.org/10.1109/SNSP.2018.00055

13. Sullivan R, Fulcher-Rood K, Kruger J, et al. Emerging technologies for lifelong learning and success: a MOOC for everyone. *J Educ Technol Syst* 2019; 47: 318–336.

14. Liu B, Xing W, Zeng Y, et al. Quantifying the influence of achievement emotions for student learning in MOOCs. *J Educ Comput Res* October 20, 2020; 59: 429–452. https://doi.org/10.1177/0735633120967318

15. Varma AK and Mitra D. Statistical feature-based SVM wideband sensing. *IEEE Commun Lett* 2020; 24: 581–584

16. Liu L, Chen R, Liu X, et al. Towards practical privacy-preserving decision tree training and evaluation in the cloud. *IEEE Trans Inf Forensics Secur* 2020; 15: 2914–2929

17. Davari SA, Nekoukar V, Garcia C, et al. Online weighting factor optimization by simplified simulated annealing for finite set predictive control. *IEEE Trans Ind Inf* 2021; 17: 31–40.

18. Yuan Y, Wu L and Zhang X. Gini-Impurity index analysis. *IEEE Trans Inf Forensics Secur* 2021; 16: 3154–3169.

19. Wang H and Wang G. Improving random forest algorithm by lasso method. *J Stat Comput Simul* 2021; 91: 353–367.

20. Onyezewe A, Kana AF, Abdullahi FB, et al. An enhanced adaptive K-nearest neighbor classifier using simulated annealing. *Int J Intell Syst Appl (IJISA)* 2021; 13: 34–44.

21. Liu W, Shen J and Yang X. Rolling bearing fault detection approach based on improved dispersion entropy and AFSA optimized SVM. *IJEEE* 17 Jul 2020; 8: 1–11. https://doi.org/10.1177/0020729092040584.

22. Zan T, Liu Z, Wang H, et al. Prediction of performance deterioration of rolling bearing based on JADE and PSO-SVM. *Proc IMechE, Part C: J Mechanical Engineering Science* 2020; 235: 1684–1697. https://doi.org/10.1177/0954406220951209.

23. Bansal N, Defo M and Lacasse MA. Application of support vector regression to the prediction of the long-term impacts of climate change on the moisture performance of wood frame and massive timber walls. *Buildings* 2021; 11: 188.

24. Yang Q and Lee YC. The critical factors of student performance in MOOCs for sustainable education: a case of Chinese universities. *Sustainability* 2021; 13: 8089.

25. Iloh C. Not non-traditional, the new normal: adult learners and the role of student affairs in supporting older college students. *J Stud Aff* 2018; 27: 25–30.

26. Huang H, Lee ZJ and Su WS. A Study on Social Support, Participation Motivation and Learning Satisfaction of Senior Learners. In: International Conference on Frontier Computing, 2018, pp.1979–1984. Singapore: Springer.
27. Lee ZJ, Huang H, Lee CY, et al. A study on the sustainable development for senior learners. *Sustainability* 2019; 11: 6338.

28. Muhammed HH, Ammenberg P and Bengtsson E. Using feature-vector based analysis, based on principal component analysis and independent component analysis, for analysing hyperspectral images. In: Proceedings of the IEEE 11th International Conference on Image Analysis and Processing, Bydgoszcz, Poland, 11–13 September 2001, pp.309–315.

29. Daher D, Deracinois B, Daniel A, et al. Principal component analysis from mass spectrometry data combined to a sensory evaluation as a suitable method for assessing bitterness of enzymatic hydrolysates produced from micellar casein proteins. *Foods* 2020; 9: 1354

30. Lin DY and Huang TY. A hybrid metaheuristic for the unrelated parallel machine scheduling problem. *Mathematics* 2021; 9: 768.

31. Jiang JR, Kao JB and Li YL. Semi-supervised time series anomaly detection based on statistics and deep learning. *Appl Sci* 2021; 11: 6698

32. Lee CY, Lee ZJ, Huang JQ, et al. Urban air quality analysis and forecast based on intelligent algorithm with parameter optimization and decision rules. *Appl Sci* 2019; 9: 5445.

33. Ghasemian A, Hosseinmardi H and Clauset A. Evaluating overfit and underfit in models of network community structure. *IEEE Trans Knowl Data Eng* 2020; 32: 1722–1735.

34. Yu CH, Wu J and Liu AC. Predicting learning outcomes with MOOC clickstreams. *Educ Sci* 2019; 9: 104.

35. Liu S, Liang T, Shao S, et al. Evaluating localized MOOCs: the role of culture on interface design and user experience. *IEEE Access* 2020; 8: 107927–107940.

36. Qu S, Li K, Wu B, et al. Predicting student performance and deficiency in mastering knowledge points in MOOCs using multi-task learning. *Entropy* 2019; 21: 1216.

37. Stracke CM and Trisolini G. A systematic literature review on the quality of MOOCs. *Sustainability* 2021; 13: 5817.

**Author biographies**

**Chou-Yuan Lee** received a BS degree in electronic engineering, an MS degree in automatic control engineering from Feng-Chia University (FCU), Taichung, Taiwan, in 1989 and 1991, respectively, and a PhD degree in electrical engineering from National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan, in 2008. He is a full-time professor with the Department of Big Data Management and Application, School of Big Data, Fuzhou University of International Studies and Trade, China. His research interests include big data, data mining and computational intelligence.

**Ling-Ming Ruan** graduated from the Department of Information Management and Information, School of Big Data, Fuzhou University of International Studies and Trade in 2020 with a BS degree. He is mainly engaged in data mining and computational intelligence.

**Zne-Jung Lee** received his BS degree in electronic engineering, his MS degree in automatic control engineering from Feng-Chia University (FCU), Taichung, Taiwan, in 1986 and 1988, respectively, and his PhD degree in electrical engineering from National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan, in 2002. He is a full-time professor with the Department of Intelligent Construction, School of Intelligent Construction, Fuzhou University of International Studies and Trade, China. His research interests include data mining and computational intelligence.
Jian-Qiong Huang graduated from Fuzhou University in 2012 with an MS degree in computer technology. Now she is a full-time professor with the Department of Big Data Management and Application, School of Big Data, Fuzhou University of International Studies and Trade. Her current research interests include data mining and database technology.

Jie Yao graduated from Beijing University of Posts and Telecommunications in 2011 with an MS degree in electronic and communication engineering. Now she is an associate professor with the Department of Big Data Management and Application, School of Big Data, Fuzhou University of International Studies and Trade. Her main research direction is artificial intelligence.

Zheng-Yuan Ning is a professor with the Department of Information Management and Information, School of Big Data, Fuzhou University of International Studies and Trade, Fujian, China. He is mainly engaged in intelligent computing, algorithmic analysis, big data, cloud computing and so on.

Jih-Fu Tu is a full-time professor with Department of Industrial Engineering and Management, St. John’s University, New Taipei City, Taiwan. He is mainly engaged in intelligent computing, big data, and data analysis.