Factors Influencing Students’ Integration Into English Classrooms in Ecologically Fragile Environments: An Analysis

Yali Zhang, Gansu Agricultural University, China*

ABSTRACT

The internet and informatization have brought a great impact on English classrooms in ecologically fragile areas, bringing convenience to the teaching of some colleges and universities, but also negatively affecting some colleges and universities in ecologically fragile areas that still use traditional lecture-style classroom teaching methods. The entire classroom is an ecological environment, and it is necessary to control the appropriate density (the number of students) and the teaching methods of teachers. How to accurately evaluate whether students in ecologically fragile areas integrate into English classrooms? The authors use the linear regression method in the teaching evaluation model and the density-based outlier detection method to clean the abnormal data. Such independence attributes, a correlation analysis method proposed, which judges dependencies and attributes according to the confidence of the rules, and then combines the correlation coefficients between attributes to determine the feature items with strong correlation.

KEYWORDS

ecological environment, ecology, educational ecology, English classroom, teaching evaluation model

INTRODUCTION

The term “ecology” refers to the dynamic equilibrium relationship established by the interaction between organisms and the living environment in the natural environmental system. The 21st century is the century of ecology. The application fields of ecology have extended from natural sciences to social sciences, and even to the field of education. Lawrence Cremin, dean of the Columbia Normal University Normal School in the United States, proposed the concept of “educational ecology” in his representative work Public Education as early as 1976. He believes that based on ecological principles, various phenomena and causes that occur in education can be studied and explained, in order to grasp the development of laws of education (Joe et al.,2017). Educational ecology uses ecological methods to study the laws of education and human development. Focusing on ecological balance, environment and adaptation, population distribution and composition, interpersonal relations, and other issues,
it attempts to establish a reasonable ecological environment inside and outside the school, improve teaching efficiency, and promote the healthy growth of the young generation. Focusing on education from an ecological perspective is actually a way of thinking that allows us to re-examine education and focus on solving the fundamental problems of education. Undoubtedly, when we use ecological theory to study the fundamental issues of education, the final focus should be on classroom teaching, which is the terminal goal of education. “Ecological classroom” is a classroom teaching model that draws on some basic concepts in ecology, fully utilizes ecological environmental conditions, continuously adjusts the relationships between various factors, and makes them tend to balance, unity, affinity, and harmony (Mliless & Larouz, 2018).

At present, most English classes in ecologically fragile areas are taught using multimedia, which brings certain conveniences to English classroom teaching, such as making the teaching content more intuitive and providing a pleasant classroom atmosphere. Teachers can use multimedia-rich materials to let students learn more knowledge. However, in our actual teaching process, there are often some problems with the use of multimedia equipment, such as when the projector’s display is unclear or not displayed at all, and the computer sound system fails. If the teacher does not carefully check whether the multimedia can work normally before class, or the teacher is not proficient in the use of multimedia equipment, it will inevitably occupy the normal class time and affect the normal teaching. In addition, we also found that multimedia teaching will limit teachers’ on-the-spot performance to a certain extent, because in most cases, teachers just present the courseware content to students step by step (Maxwell et al., 2017).

In addition, the number of English classes has increased sharply, and large class teaching has become the mainstream of the public English teaching model. There are at least 60 or 70 people in a teaching class, and some even reach more than 100 people. The lectures are usually held in large classrooms. Due to the large number of people and the large classrooms, many problems emerge, such as students skipping classes, arriving late, and leaving early. Some even talk, wander, sleep, and play with their phones in the classroom, but teachers find it difficult to constrain the behavior of each student, causing great difficulties in classroom management. Obviously, such a teaching environment will inevitably affect the normal teaching activities (Getie, 2020).

Most of the students in English classes in areas with fragile ecological environment come from rural mountainous areas. Students have differences in intelligence, personality, interests, etc., and their English learning experience and English proficiency are uneven. After entering the classroom, each student’s learning motivation and learning attitude are also different. Individual differences bring many difficulties to teaching, reflected in classroom goal setting, teaching schedule arrangement, teaching activity design, and other aspects. In a relatively unified arrangement it is often difficult to take into account the levels of all students. Students with good language foundation think that the breadth and depth of knowledge they have learned is not enough, while students with weaker language foundation have the problem of not being able to keep up (Mohamadi, 2018).

There is still little research on factors such as students’ integration into English classrooms in ecologically fragile areas based on classroom evaluation models. This article uses a teaching evaluation model to analyze in depth the factors that influence the integration of students in ecologically fragile areas into English classrooms. It not only analyzes students’ behavior, but also analyzes the importance of teachers in English teaching classrooms. Integrating the theory of educational ecology, it is possible to discover new models of classroom teaching.

**MATERIALS AND METHODS**

The university classroom is a micro ecosystem. It is a basic system composed of the interaction and influence of the classroom ecological environment and the classroom ecological subject (Raygan & Moradkhani, 2022). The environment of the college English classroom ecosystem plays a crucial role in constraining and regulating the emergence, existence, and development of college English classroom
teaching. The ecological environment of college English classrooms includes physical environment, social environment, psychological environment, and physiological environment (Astor et al., 2021).

The physical environment includes classroom layout, seating arrangements, class size, and the use of multimedia equipment (Weatherson et al., 2017). The seating arrangement in classrooms in ecologically fragile areas mostly adopts a fixed determinant. This seating arrangement strengthens the dominant position of teachers and weakens the main role of students. Therefore, it is difficult to carry out two-way communication and is not conducive to mobilizing students’ enthusiasm. Other studies have shown that differences in seating positions can lead to differences in spatial and interpersonal distance, as well as differences in perception and understanding, which to some extent affects students’ learning enthusiasm and attitude (Saltan & Divarci, 2017). For example, students sitting in the front or middle of the classroom can see and hear more clearly during class because they are physically closer to the teacher, making it easier and more active for them to participate in classroom teaching activities. Students who are far away from the teacher are not very engaged in the classroom and are easily distracted.

Social environment refers to the relationship between teachers and students and students’ cooperative learning ability, etc. The social environment of college English classrooms consists of seat arrangement, teacher-student niche, class size and category, the proportion of male and female students, campus language environment, the second classroom, and other elements. A common problem with teaching in multimedia classrooms is that teachers are limited to the podium due to the need to manipulate multimedia at any time. In the eyes of students, teachers rarely step down from the podium to communicate with students, maintaining a sense of distance between them, leading to a lack of teacher-student relationship (Rach & Heinze, 2017). In addition, due to the large number of students in large classes, it is difficult for teachers to take care of each student, resulting in fewer opportunities for students to participate in classroom activities individually (Major et al., 2017). Only a small number of students can exercise skills and practice in the classroom. Over time, other students will feel that they are not gaining the attention of the teacher in the class, and their interest in the course will also decrease. They are also unwilling to cooperate with other students to participate in classroom activities, which seriously affects the teaching effectiveness.

Psychological environment includes students’ learning experience, cognitive style, learning motivation, and learning attitude, and it also refers to the learning atmosphere and classroom culture jointly created by teachers and students. The physiological environment includes physiological characteristics such as students’ age, gender, and physical health (Sriratanaviriyakul & El-Den, 2017). There is little difference in age and physical health among students in English classrooms in areas with fragile ecological environment, but gender is an important factor affecting the physiological environment. Due to the limited teaching resources in schools, the ratio of males to females in some classes is seriously imbalanced. On the other hand, teachers are sometimes not very good at carrying out teaching activities, and it is not easy to mobilize the enthusiasm of students, which affects the teaching to a certain extent.

Data mining is a deep level method for analyzing data information. By extracting, transforming, analyzing, and modelling data, potential useful information can be extracted from these data (Khlaif, 2018). Data pre-processing is an important link in data mining, including data cleaning, data integration, and data discretization. For a long time, scholars have been continuously optimizing and improving the pre-processing technology. For example, the outlier detection method can be used to judge abnormal data points during data cleaning (Gaias et al., 2019). The outlier detection method adopts the local abnormal factor detection algorithm after K-means cluster pruning, which greatly reduces the complexity of the algorithm. In addition, the classification problem is also a research hotspot in data analysis (Jain et al., 2022). There are many methods for constructing classification models, including statistical methods, machine learning methods, neural network methods, and rough sets (Rana et al., 2020).
Data mining has played a significant role in supporting various domains, particularly in terms of data acquisition. (Larmuseau et al., 2018). The continuous development of the educational field has also promoted the gradual increase of research on the theory. In 2008, the first International Conference on Educational Data Mining was held in Montreal, Canada. Up to now, eight conferences have been successfully held, and the conference has been successfully established (Mei et al., 2018). Educational data mining (EDM) is actually a learning system based on computer systems, interactive learning environments, simulated learning environments, existing school learning systems, etc. to collect user learning behavior data. With the knowledge of psychology and learning science, we can use the knowledge of computer science, data mining, and other fields to find out how students learn.

In foreign countries, the related research of EDM has made some progress. Scholars such as S. Ganesh and A. Binu collect data from social networks and apply Hadoop framework to analyze student behavior (Allen et al., 2018). EDM research results cover all educational environments, student types, and application types, reflecting its research philosophy of “coming from education and returning to education” and its educational philosophy of “student-centered” education (Al Salman et al., 2021). On the other hand, EDM research is still limited by data sources and researchers’ technical capabilities to a large extent.

RESULT ANALYSIS AND DISCUSSION

Data Collection and Feature Analysis

This section mainly elaborates on the data collection, data pre-processing, and correlation analysis among teaching evaluation attributes involved in the one regression analysis and outlier detection algorithms are combined, and correlation analysis, the analysis method of association rules, is proposed, and the method of combining correlation coefficient and association rules is used to judge the correlation of attributes. In data cleaning, regression uses a function to fit the data to smooth the data and identify noise, while the density-based outlier detection method has a better abnormal data detection effect on unevenly distributed data sets. In general, the predicted value obtained by the regression method will also have a certain error, so if all the data deviating from the true evaluation value are treated as abnormal data, the potential value of the data may be lost. The reciprocal of the mean reachability density of the nearest neighbors of MinPts based on $p$ (Zhao & Frank, 2003). The calculation formula is shown in formula (1).

$$lrd_{MinPts}(p) = \frac{N_{MinPts}(p)}{\sum_{O \in N_{reachDis}} \text{reach}_{MinPts}(p,o)}$$

(1)

Formula (2) indicates the degree to which $p$ is an outlier:

$$LOF_{MinPts}(p) = \frac{\sum_{O \in N_{reachDis}} \frac{lrd_{MinPts}(o)}{lrd_{MinPts}(p)}}{\frac{N_{MinPts}(p)}{N_{MinPts}(p)}}$$

(2)

Among them, the one will be in cluster analysis. Divide the dataset into $K$ clusters. In the clustering process, the calculation formula of the class center is as formula (3).
\[ X_0 = \frac{\sum_{k=1}^{ni} X_k}{ni} \quad (3) \]

In the above formula, \( ni \) will be the such formula for calculating the radius of class \( i \) as formula (4).

\[ R_i = \frac{\sum_{k=1}^{ni} |X_k - X_0|}{ni} \quad (4) \]

For each data object \( x_i \), the calculation formula of the distance to the class center \( x_0 \) of its corresponding class is as formula (5).

\[ d_i = \sqrt{\sum_{j=1}^{m} (x_{ij} - x_{0j})^2} \quad (5) \]

\( m \) in formula (5) represents the dimension of each data. The \( K \)-means clustering method can also be used to detect outliers. It mainly regards data that does not belong to any class or data that belongs to a class with few data objects as outliers, but the detection accuracy is highly dependent on the determination of clustering parameters, so the detection effect is often not obvious.

The correlation coefficient reflects the closeness of random variables to each other. According to the correlation coefficient, the degree of mutual connection and mutual influence between attributes can be measured, which is defined as formula (6).

\[ r = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \quad (6) \]

This type of index data has low volatility. For \( n \) different attributes, the correlation coefficient between each pair can be calculated separately, so that all the correlation coefficients form a matrix, as shown in formula (7).

\[ R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & \cdots & \cdots & r_{nn} \end{bmatrix} \quad (7) \]

The magnitude of the correlation coefficient can indicate the degree of mutual dependence between two attributes. If the correlation coefficient between the two attributes is large, the correlation between the two attributes is strong; otherwise, if the correlation coefficient is small, the correlation is weak. Therefore, the independence between attributes can be indirectly analysed through the value range of the correlation coefficient, as shown in formula (8).
\[ s(X \Rightarrow Y) = P(X \cap Y) \] \hspace{1cm} (8)

Confidence degree is as in formula (9).

\[ c(X \Rightarrow Y) = P(Y \mid X) = \frac{P(X \cap Y)}{P(X)} \] \hspace{1cm} (9)

Figure 1 shows the correlation analysis process based on association rules.

**Construction of Classroom Teaching Evaluation Model Based on Weighted Naive Bayes**

Compared with other algorithms, the naive Bayes classification has the smallest error rate, but its precondition is relatively strict. Only when each attribute of the object is independent of each other, the use of a naive Bayes model can get the best classification effect. However, it is difficult to meet the precondition.
this condition among several research attributes of student achievement prediction. In order to make up for this defect, some scholars have proposed a weighted naive Bayes (WNB) classification model, whose main idea is to give different weights to each attribute, so that the naive Bayes method can be extended, reduce the requirements of the algorithm for attribute independence, and improve the efficiency of classification. This section mainly talks about the accuracy as an indicator for evaluating the performance of a classifier, which is shown in formula (10).

\[
\text{pre \_ rate} = \frac{N_c}{N} \tag{10}
\]

Its class label is known, and it can be expressed as \(\{x_1, x_2, \ldots, x_n\}\), to judge the probability that the sample, the calculation formula is shown in formula (11).

\[
p(C_j | X) = \arg \max_{C_j} \frac{p(X | C_j) p(C_j)}{p(X)} \tag{11}
\]

The naive Bayesian classification (NB algorithm) outperforms classifiers such as decision trees and support vector machines. Figure 2 is a naive Bayesian classification model.

The Naive Bayesian classification model can be represented by formula (12) and formula (13).

\[
p(C_j | X) \propto \arg \max_{C_j} p(X | C_j) p(C_j) \tag{12}
\]

\[
p(X | C_j) = \prod_{i=1}^{n} p(x_i | C_j) \tag{13}
\]

The formula for calculating the impact of the assumption of conditional independence of attributes on classifier performance is shown in formula (14) and formula (15).

\[
p(C_j | X) = \arg \max_{C_j} p(C_j) \prod_{i=1}^{n} p(A_i | C_j)^{w_i} \tag{14}
\]

**Figure 2. Naive Bayesian classification model**
\[ p(C_j | X) = \arg \max_{c_j} p(C_j) \prod_{i=1}^{n} p(A_i | C_j) \]

(15)

**Result Analysis of Data Collection and Characteristics**

The abnormal data detection method mentioned above and the data analysis experiments were carried out by regression analysis and local outlier factor (CLOF) algorithm, respectively. The experiment is in the Windows 10 operating environment, using Python as the development language, as shown in Figure 3 (Duursma et al., 2007).

Such real evaluation value is recorded in the data table, and the blue dotted line is the predicted evaluation value given by the model by analysing a series of evaluation attribute values (Hofstede, 1986). Through outlier detection, we found that there is still a certain error in the evaluation of English teaching in areas with fragile ecological environment using the teaching evaluation model, indicating that our model still needs to be improved (Liñán et al., 2011). The improvement and optimization of the teaching evaluation model will help the students better integrate into the English classroom and improve their English learning ability. The experimental results of detecting abnormal data by regression method are shown in Figure 4. The outlier detection results are shown in Figure 5.

The abnormal points detected by the two methods are not exactly the same, indicating that although some data points detected by regression analysis deviate from the overall data distribution, they still have potential in local research. The value of information cannot be directly ignored, and the data detected at the same time have a very high possibility of abnormality, so these data points need to be properly processed. The presence of outliers in areas with fragile ecological environments can impact the analysis of students’ integration into English classrooms, leading to potential errors.
Figure 4. Outlier detection results based on regression analysis

Figure 5. Outlier detection results based on CLOF algorithm
in the findings. Therefore, we need to strengthen the detection of outliers in order to better analyze the factors that help students integrate into the English classroom. The reason for the error may be that the students in the fragile ecological environment come from backward mountainous areas, and their learning abilities are also different, so there will be a certain measurement error in the results. Through the detection of outliers, we can make suitable learning plans for students with different abilities and effectively improve their ability to learn English.

According to the experimental data, the correlation coefficient between the evaluation attributes is calculated, and the maximum correlation coefficient between the evaluation attributes.

Analysis of Table 1 shows that the smaller the correlation coefficient value is, the stronger the independence of the attribute is, and there is a strong correlation between attributes with a larger correlation coefficient.

For this data set, from the analysis of the correlation coefficient value, the correlation coefficient between attributes D and E, H and I, M and N is relatively high, and there is a strong mutual influence relationship between them. Combining the strong rules, it can be seen that the attributes such as D, H, K, M and other attributes have a high degree of confidence, which verifies that there is indeed a strong correlation between these attributes and other attributes. If other properties exist, they are likely to exist, and these properties do not need to be reconsidered. Therefore, when designing the evaluation index system, consider removing the attributes D, H, K and M, and only keep the attributes A, B, C, E, F, G, I, J, L, N, and the original 14 evaluation attributes. The teaching evaluation index system is optimized to a system with only 10 independent attributes. Effective analysis of the English classroom can optimize the teaching level in areas with fragile ecological environment, so that students can experience a better English teaching experience, and proper research can also improve the teaching level of teachers according to the needs of students, achieving a win-win situation. It is a very important foundation for students in areas with fragile ecological environments to better integrate into English classrooms.

This section mainly studies and analyzes the data collection, data pre-processing, and feature analysis involved in the construction of the teaching evaluation model. Pre-processing is the process of transforming the original data into a data format that meets the requirements of data mining. Because

<table>
<thead>
<tr>
<th>Attribute</th>
<th>The most relevant property</th>
<th>Phase relation value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>H</td>
<td>0.336</td>
</tr>
<tr>
<td>B</td>
<td>J</td>
<td>0.348</td>
</tr>
<tr>
<td>C</td>
<td>D</td>
<td>0.581</td>
</tr>
<tr>
<td>D</td>
<td>E</td>
<td>0.676</td>
</tr>
<tr>
<td>E</td>
<td>D</td>
<td>0.676</td>
</tr>
<tr>
<td>F</td>
<td>D</td>
<td>0.636</td>
</tr>
<tr>
<td>G</td>
<td>E</td>
<td>0.580</td>
</tr>
<tr>
<td>H</td>
<td>I</td>
<td>0.656</td>
</tr>
<tr>
<td>I</td>
<td>H</td>
<td>0.656</td>
</tr>
<tr>
<td>J</td>
<td>K</td>
<td>0.483</td>
</tr>
<tr>
<td>K</td>
<td>N</td>
<td>0.603</td>
</tr>
<tr>
<td>L</td>
<td>K</td>
<td>0.555</td>
</tr>
<tr>
<td>M</td>
<td>N</td>
<td>0.639</td>
</tr>
<tr>
<td>N</td>
<td>M</td>
<td>0.639</td>
</tr>
</tbody>
</table>
the quality of data directly affects the results of data mining, pre-processing plays an important role in data mining. Pre-processing mainly includes data cleaning, whose goal is to eliminate the noise in the data, and data integration, whose goal is to merge data from multiple data sources. The goal of data reduction is to reduce the size of the original data, so as to improve the efficiency of data mining. The goal of data transformation is to transform numerical values into the form required by data mining algorithms. Through the research on the teaching evaluation data, redundant teaching evaluation indicators are removed, and the evaluation index system is optimized, in which outlier detection and feature correlation analysis of data sets are described in detail. By constructing a suitable teaching evaluation model, the singular value of the factors of students’ integration into the English classroom can be detected, and the factors of students’ integration into the classroom can be accurately analyzed. By optimizing the teaching evaluation model, the teaching quality of teachers in ecologically fragile areas can be improved. It can also enhance students’ interest in learning English and be good for a country’s delivery of high-quality English talents.

Analysis of the Results of English Teaching Evaluation Based on Weighted Naive Bayes

The unified into one in the [0,1] interval through data normalization processing, and a model is formed by setting an error threshold, which is used for new evaluations (Haug et al., 2008). The sample data are used to predict the evaluation level. In the back propagation (BP) algorithm, ones are randomly selected as the training set, and 70 data records are used as the test set. In the BP algorithm, randomly select one as the training set and 70 data records as the test set. After experimental debugging, the optimal experimental parameter is set to 10000.

The test results are obtained by training the neural network. For the data it is easy to overfit, causing a generally high prediction level. Therefore, it is considered to discretization the percentile scores into five level evaluation values after pretreatment, and to randomly select some levels to mix into the training data set. Figure 6 shows the comparison of classification accuracy between the BP algorithm and WNB algorithm.

Compared to the neural network model, the WNB English teaching evaluation model has better accuracy, and the evaluation time is shorter, which can reduce the time occupied by teachers in evaluations and let them devote more time to facilitate learning. In the process of English teaching, it helps to improve the quality of English classrooms, and also to provide more rest time for teachers, so that teachers can be more creative to innovate in English classrooms. In ecologically fragile areas, students’ integration into English classrooms requires the input of teachers. Teachers’ teaching methods, textbook selection, evaluation methods, and many other aspects in the classroom can have an impact on students’ integration into the English classroom (Sidiropoulos, 2018). This running time comparison chart will be Figure 7.

In ecologically fragile areas, due to limitations in transportation and resources, students may not be able to enjoy high-quality educational resources. Therefore, teachers can consider using digital tools such as the internet and online educational platforms to push educational resources to a wider range of student groups, promoting fair sharing of global educational resources. In English classrooms, teachers can increase students’ interest in the subject by introducing multicultural elements, creating an environment that respects and tolerates cultural differences, thus promoting students’ integration into the classroom. For example, design teaching content and activities related to students’ daily lives.

CONCLUSION

This article uses a teaching evaluation model to analyze in depth the factors that influence the integration of students in ecologically fragile areas into English classrooms. It not only analyzes students’ behavior, but also analyzes the importance of teachers in English teaching classrooms. Introducing the theory of educational ecology, it is possible to discover the effectiveness of classroom
Figure 6. Comparison of classification accuracy between BP algorithm and WNB algorithm

Figure 7. Time comparison chart
teaching. This article uses data mining technology to preprocess the obtained data and mainly uses outlier detection algorithms to clean the data. Then the classification algorithm in machine learning is introduced to construct a teaching evaluation model, and a classifier based on WNB algorithm is proposed. The WNB algorithm has been verified to have better evaluation accuracy and shorter evaluation time for English teaching models compared to traditional BP neural network algorithms through runtime and classification accuracy, which can reduce the time occupied by teachers in evaluation and allow them to invest more time in encouraging student learning.

This article mainly focuses on the factors that affect the integration of students in ecologically fragile areas into English classrooms, but the environmental and cultural backgrounds of different regions may also affect the needs and differences in English education. In the future, we can gain a deeper understanding of the needs and differences of English education in different regions through cross regional comparisons and onsite research, providing a basis for formulating targeted English educational policies.

**AUTHOR NOTE**

Data availability: The figures and tables used to support the findings of this study are included in the article.

Conflicts of interest: The author declares that there are no conflicts of interest.

Funding statement: This paper was not supported by any funds.

Acknowledgements: The author would like to express sincere thanks to those whose techniques have contributed to this research.

Yali Zhang (1985-), female, Han nationality, born in Baiyin, Gansu province, master lecturer, major in translation theory and practice, English education, etc.
REFERENCES


