Abstract: First, A review of the literature on academic early warning at home and abroad was summarized. Index system of academic early warning was determined including four factors of daily study time, number of courses selected, number of violation of classroom discipline every semester, academic credit not gained. Two models on academic early warning were constructed based on Fisher Discriminant and Bayes Discriminant. Accuracy reached to 98.1% for the model of Bayes Discriminant. Through research and analysis, the ultimate goal is to enhance the style of study, improve the quality of education and teaching and help college students graduate successfully.

Keywords: Academic Early Warning; Fisher Discriminant; Bayes Discriminant.

INTRODUCTION

When entering university, some students usually behave bad phenomena such as slackness in thought and sloppy behavior due to weak basic knowledge and poor study habits and so on, which will lead to the final exam results not satisfactory. The occurrence of this phenomenon not only causes the students themselves to face the danger of repeating or dropping out of school, but also has a significant impact on the teaching order. This requires us to promptly, scientifically and reasonably warns the students’ academic situation and changes the management methods and means.

From the relevant practical research results, it is found that the students under the attention of early warning mechanism have the characteristics of weak learning motivation and ambiguous goal planning.

The earliest academic warning research started in the 90s of last century, among which, the early warning rate of boy students is much higher than that of girl students, that is, boys are more likely to be in a state of early warning crisis [1]. In August 2007, the Ministry of Education officially promulgated the term “academic early warning”, which indicated that it will play an increasingly important role in the development of higher education in China. Subsequently, some colleges and universities in China began to try to implement the new teaching management mode of academic early warning, to improve the teaching quality and teaching management level [2-7].

Compared with China, the developed countries attach great importance to the impact and role of information technology on education. For the modern information management of teaching system, many foreign universities have been exploring and researching for many years, and are now relatively mature. However, foreign colleges and universities do not have special academic warning for students’ academic situation, but from the school level, for school incidents may occur, for the relevant stakeholders (such as university administrators, parents, legal advisors, etc.) of the school early warning, and emergency management plans for quantitative research, the establishment of emergency handling model and simulation system[8-11].

It is known in a comprehensive way, data mining technology is being widely used in exploring knowledge about student achievement data of educational administration system. It is more mature in foreign countries than in China. Data mining has been used in every link from student enrollment to graduation. However, most of the domestic research is scattered and the effect is limited. Therefore, the data mining research based on the university educational administration management system data can promote the teaching management by revealing the essence of the data.

Model based on Academic Early Warning

Discriminant analysis, also known as "discriminant method", is a multivariate statistical analysis method to determine the classification of a certain object according to its various eigenvalues. The basic principle is to establish one
or more discriminant functions according to certain criteria, and determine the undetermined coefficients of the
discriminant functions with a large amount of data of the object of study, and calculate the discriminant indexes.

The criterion is the theoretical basis and method standard for measuring the degree of similarity between new
samples and known groups. Commonly used distance discriminant criteria, Fisher discriminant criteria, Bayes criterion
and so on. Discriminant criteria can be statistical, such as the use of mathematical statistics when determining the
category of a new sample; can be deterministic, such as determining sample ownership, only considering the size of the
discriminant function value.

**Model based on Fisher Discriminant Analysis**

Discriminatory thought: Projecting the p-dimension data of k groups to a certain direction, the projection groups
and groups are separated as far as possible. Set the k group p-dimension from the k population to observe the values:
G₁, G₂, ..., Gₖ respectively. Let α be any vector in Rᵖ, u(ₓ) = α’ₓ is x, projection to α is normal direction[12].

The sum of squares between groups is:

\[SSG = \sum_{i=1}^{k} n_i (\bar{x}^{(i)} - \bar{x})^2\]

In the form of B = \(\sum_{i=1}^{k} n_i (\bar{x}^{(i)} - \bar{x}) (\bar{x}^{(i)} - \bar{x})'\), \(\bar{x}\) they are group i mean and total mean vector.

Sum of squares in a group:

\[SSE = \sum_{i=1}^{k} \sum_{j=1}^{k} (\alpha' x_j^{(i)} - \alpha' \bar{x}^{(i)})^2\]

Where, 
\(E = \sum_{i=1}^{k} \sum_{j=1}^{k} (\bar{x}^{(i)} - \bar{x}) (\bar{x}^{(i)} - \bar{x})'\).

According to the variance analysis criterion in Fisher discriminant, the coefficient is c, subject to
\[\lambda = \sum_{i=1}^{k} n_i (\bar{y}^{(i)} - \bar{y})^2 / \sum_{i=1}^{k} (n_i - 1) \delta_i \geq \frac{c^2}{c^2 + E}\]

\(E\) is the sum of squares within the group deviations, and \(B\) is the sum of squared deviations. In order to
maximize the value of \(B = \sum_{i=1}^{k} n_i (\bar{x}^{(i)} - \bar{x}) (\bar{x}^{(i)} - \bar{x})'\), 
\(E = \sum_{i=1}^{k} n_i \eta^{(i)}\).

\(\lambda\) and \(c\) are the generalized eigenvalues of \(B\) and \(E\) matrices and their corresponding eigenvectors. Because \(B\) is
nonnegative definite, the non-zero eigenvalue is the whole root, remember \(\lambda_1 \geq \lambda_2 \geq \ldots \lambda_m > 0\), \(m\) discriminant functions
are constructed by the number of non-zero eigenvalues \(m \leq \min(n - k, p)\), among of \(l = 1, 2, \ldots, m\).

The accuracy of discriminant function is usually expressed by the contribution rate, and the contribution rate is
\(p_l = \lambda_l / \sum_{i=1}^{l} \lambda_i\). In practical applications, when the cumulative contribution rate is \(p > 0.85\).

**Model on Bayes Discriminant Analysis**

Discriminatory thought: assuming that we have a certain understanding of the object of study, we often use a
prior probability distribution to describe this knowledge. Then we take a sample to correct the known knowledge (prior
probability distribution) and obtain a posterior probability distribution. All statistical inferences are carried out by a
posteriori probability distribution. Bias discriminant analysis is used to get Bias discriminant [13].

There are \(k\) total \(G_1, G_2, ..., G_k\). Assuming a prior knowledge of the problem under study, this knowledge is often
described by a priori probability, i.e. the priori distributions of known \(k\) populations are \(q_1, q_2, ..., q_k\), where
\(q_i > 0, q_1 + q_2 + ... + q_k = 1\).

The criteria are: \(D_1, D_2, ..., D_k\) denotes a partition that minimizes the average loss when classified by this partition.

The probability of mis-judgement is:

\[p(j|i) = \int_{D_j} n_i(x)dx\]

The average miscarriage cost is:

\[ECM (D_1, D_2, ..., D_k) = \sum_{i=1}^{k} q_i \sum_{j=1}^{k} c(j|i) p(j|i)\]
Case Analysis

Warning level

Based on the academic early warning work of a college in our school, the early warning mode can be divided into: early-warning of achievement and early-warning of attendance. Academic early warning, as a kind of warning and alarm, sends out different levels of warning information according to the actual situation of students and the degree of influence produced. Academic early warning is divided into five levels: oral early warning, blue early warning, yellow early warning, orange early warning, red early warning, and different levels of early warning information, see Table 1 for details.

<table>
<thead>
<tr>
<th>Grade content</th>
<th>Oral early warning</th>
<th>Blue early warning</th>
<th>Yellow early warning</th>
<th>Orange early warning</th>
<th>Red early warning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early-warning of attendance</td>
<td>Total absences of more than 6 class hours</td>
<td>Total absences of more than 10 class hours</td>
<td>Truancy is more than 20-29 of the class hours and 1/3 is canceled.</td>
<td>Total absences of more than 39 class hours</td>
<td>Continuous total absences of 40 hours or more</td>
</tr>
<tr>
<td>Early-warning of achievement</td>
<td>Failed to pass 1 courses last semester</td>
<td>Failing to pass the course last semester reached 2 or all courses failed, reaching 4-5 courses.</td>
<td>Failing to pass the course last semester reached 3 or all courses failed, reaching 6-7 courses.</td>
<td>Failing to pass the course last semester reached 4-5 or all courses failed, reaching 8 courses.</td>
<td>Failing to pass the course last semester reached 6 or full courses, failing to pass more than 9 courses.</td>
</tr>
</tbody>
</table>

Data Resource

This paper selects 56 students of different grades of Heilongjiang Bayi Agricultural University who received academic early warning to build the model. The collected indicators include daily study time, number of courses selected, number of violation of classroom discipline every semester, academic credit not gained. Among them, four students from $x_1$, $x_2$, $x_3$ and $x_4$ were taken as the samples to be judged.

In the selected indicators, the number of elective courses and the absence of credits can be inquired through our educational administration system. The number of late, early retirement and truancy can be found in the annual compilation of the college's learning department.

We can download the selected courses and the lost credits to the Excel form to collect the original data.

Construction and resolution of model on academic early warning

Establishment of Fisher discriminant analysis model

(1) Establish discriminant function

The coefficient vectors of unstandardized typical discriminant functions can be obtained from the coefficient tables of typical discriminant functions in SPSS.

<table>
<thead>
<tr>
<th>Function</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>academic credit not gained</td>
<td>0.609</td>
<td>-0.249</td>
<td>-0.133</td>
<td>0.058</td>
</tr>
<tr>
<td>daily study time</td>
<td>0.209</td>
<td>0.049</td>
<td>-0.587</td>
<td>0.317</td>
</tr>
<tr>
<td>number of courses selected</td>
<td>1.557</td>
<td>-0.297</td>
<td>1.638</td>
<td>1.030</td>
</tr>
<tr>
<td>number of violation of classroom discipline</td>
<td>0.452</td>
<td>0.345</td>
<td>0.172</td>
<td>0.310</td>
</tr>
<tr>
<td>(constant)</td>
<td>-16.665</td>
<td>1.558</td>
<td>-0.849</td>
<td>-9.038</td>
</tr>
</tbody>
</table>

According to this coefficient vector, we can write 4 typical discriminant functions which are not standardized.
Wilks' Lambda test for significance of function results

Table 3: Lambda of Wilks

<table>
<thead>
<tr>
<th>Function test</th>
<th>Lambda of Wilks</th>
<th>Chi-square</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 4</td>
<td>0.012</td>
<td>206.029</td>
<td>16</td>
<td>0.000</td>
</tr>
<tr>
<td>2 to 4</td>
<td>0.309</td>
<td>54.579</td>
<td>9</td>
<td>0.000</td>
</tr>
<tr>
<td>3 to 4</td>
<td>0.760</td>
<td>12.782</td>
<td>4</td>
<td>0.012</td>
</tr>
<tr>
<td>4</td>
<td>0.973</td>
<td>1.296</td>
<td>1</td>
<td>0.255</td>
</tr>
</tbody>
</table>

Table 3 shows the validity test of typical discriminant functions. Wilks' Lambda ranges from 0 to 1, with a value close to 0 indicating that the group mean is different, and a value close to 1 indicating that the group mean is not statistically different. Because the distribution table of Wilks' Lambda statistics is not easy to find, it is generally reduced to Chi square statistics. The degree of freedom of df is the degree of freedom used to compute the saliency level. Sig. Significance level is the probability that zero hypothesis is rejected, that is, the probability of making the first kind of error when rejecting zero hypothesis. It can be seen from the table that the sig values of the significance test of discriminant function are 0.000, 0.000, 0.012, 0.255 respectively. Except fourth they are less than 0.05, therefore, the first, second, third discriminant functions are significant and can be used to discriminate the attribution of samples.

Establishment of Bayes discriminant analysis model

(1) The frequencies of students included in oral, blue, yellow, orange and red warnings were used as prior probabilities, i.e.

\[ p_1 = \frac{11}{52} = 0.212; \quad p_2 = \frac{28}{52} = 0.538; \quad p_3 = \frac{4}{52} = 0.077; \quad p_4 = \frac{4}{52} = 0.077; \quad p_5 = \frac{5}{52} = 0.096 \]

Same table 4 Results.

Table 4: Prior probability of a group

<table>
<thead>
<tr>
<th>Academic early warning level</th>
<th>Transcendental</th>
<th>Cases for analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unweighted</td>
<td>Weighted</td>
</tr>
<tr>
<td>oral early warning</td>
<td>0.200</td>
<td>11</td>
</tr>
<tr>
<td>blue early warning</td>
<td>0.200</td>
<td>28</td>
</tr>
<tr>
<td>yellow early warning</td>
<td>0.200</td>
<td>4</td>
</tr>
<tr>
<td>orange early warning</td>
<td>0.200</td>
<td>4</td>
</tr>
<tr>
<td>red early warning</td>
<td>0.200</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>1.000</td>
<td>52</td>
</tr>
</tbody>
</table>

(2) Establishing Bayes discriminant function

Table 5: Classification function coefficient

<table>
<thead>
<tr>
<th>Academic early warning level</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>daily study time</td>
<td>4.825</td>
<td>5.451</td>
<td>5.791</td>
<td>8.085</td>
<td>7.871</td>
</tr>
<tr>
<td>number of courses selected</td>
<td>31.249</td>
<td>33.979</td>
<td>42.773</td>
<td>45.717</td>
<td>56.149</td>
</tr>
<tr>
<td>number of violation of classroom discipline</td>
<td>8.321</td>
<td>8.879</td>
<td>12.241</td>
<td>13.806</td>
<td>14.572</td>
</tr>
<tr>
<td>(constant)</td>
<td>-124.055</td>
<td>-153.750</td>
<td>-236.898</td>
<td>-326.298</td>
<td>-443.134</td>
</tr>
</tbody>
</table>

From the above table, 5 discriminant functions in Bayes discrimination are obtained:

\[ y_1 = 8.289x_1 + 4.825x_2 + 31.249x_3 + 8.321x_4 - 124.055 \]
\[ y_2 = 9.978x_1 + 5.451x_2 + 33.979x_3 + 8.879x_4 - 153.750 \]
\[ y_3 = 12.080x_1 + 5.791x_2 + 42.773x_3 + 12.241x_4 - 236.898 \]
\[ y_4 = 14.861x_1 + 8.085x_2 + 45.717x_3 + 13.806x_4 - 326.298 \]
\[ y_5 = 18.295x_1 + 7.817x_2 + 56.149x_3 + 14.572x_4 - 443.134 \]

(3) Verification of accuracy of discriminant function

Available online: [http://saspjournals.com/sjavs](http://saspjournals.com/sjavs)
The accuracy of the discriminant analysis results is verified by the back substitution method and the interactive verification method in SPSS software. The conclusion are: firstly, the accuracy of the discriminant results is 98.1%; second, cross-validate only the cases in the analysis, where each case is classified according to a function derived from all other cases other than that case; thirdly, Third, 96.2% of the cross validation group cases have been correctly classified. Finally, the 56 students were given early warning modeling, and their early warning models were obtained, by discriminant analysis of x1, x2, x3, x4 students, the early warning grades are: red early warning, oral early warning, orange early warning, blue early warning.

CONCLUSIONS
Firstly, this paper analyzes the connotation of academic early warning, and establishes an academic early warning model based on discriminant analysis, taking the students' daily learning time, the number of courses selected, the number of late semester, early quitting, truancy, and not getting credit points as the main indicators. Then, by using the curriculum data of early-warning students in the school educational administration management system as an example, the modeling results based on Fisher discriminant method and Bayesian discriminant method are analyzed. Among them, the accuracy of the Bayes Discriminant model is 98.1%. It has high prediction accuracy and strong reliability. It also provides a reference index for the evaluation of academic early warning grade, thus providing a reliable and credible basis for solving practical problems.

ACKNOWLEDGEMENT
The preparation of this manuscript is supported by the Fund Project:
□ Training Project of Innovation and Entrepreneurship for undergraduates of Heilongjiang Province in 2017 (No.2017LXY04);
□ General Project of higher education reform in Heilongjiang in 2017 (No. SJGY20170441);
□ Key Project of scientific planning for higher education in Heilongjiang in 2018 (No.GBB131807).

REFERENCES