Advanced Science Letters
Vol. 23, 2495–2497, 2017

Decision Support System for Admission Selection and Positioning Human Resources by Using Naive Bayes Method

Dyna Marisa Khairina*, Septya Maharani, Ramadiani, and Heliza Rahmania Hatta

Mulawarman University, Indonesia

Human Resources (HR) are the most important resource for companies that are expected to meet the criteria and become the company's quality standards. HR selection process is important because it is not easy to make a decision effectively. It takes several methods to solve the problem of selection acceptance and positioning of human resources. Naive Bayes is used to classifying methods of probability and statistics. The results of the decision are received or not received with a probability value "yes" is greater than the probability of a "no." Finally, this study resulted in a decision support system for receiving and positioning the selection of human resources, which gives the advice to make the right decision.

Keywords: Criteria, Decision Support System, Human Resources, Naive Bayes, Selection.

1. INTRODUCTION

One of the important things in the management and development of human resources is the process of selection candidates for admission and HR positions in a company. The Company has its own criteria in the standard selection process for HR. The criteria set by the company will be the company's quality standards. Research using the Naive Bayes method in data processing as decision support in the selection process and HR positions. Naive Bayes is a machine learning method that uses probability calculations are expected to help the company to recruit competent human resources objectively in accordance with the standards specified criteria.

2. NAIVE BAYES ALGORITHM

Naïve Bayes is classification with a probability method and statistics expressed by the British scientist Thomas Bayes. Naïve Bayes, for each class of decisions, calculates the probability on condition that the class of decisions is correct, where the vector is object information.[1] A conditional probability is the likelihood of some conclusion, B, given some evidence/observation, A, where a dependence relationship exists between B and A. This probability is denoted as $P(B \mid A)$ where[2]

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}$$

Bayes theorem finds the probability of an event occurring given the probability of another event that has already occurred. B represents the dependent event and A represents the prior event.

To calculate the probability of B given A, the algorithm counts the number of cases where A and B occur together and divides it by a number of cases where A occurs alone. An advantage of Naïve Bayes is that it requires a small amount of training data to estimate the parameters necessary for classification. Since independent variables are assumed, only the variances of the variables for each class need to be determined and not the entire. It can be used for both binary and multiclass classification problems.[3–5]

The Naive Bayes works as follows: Each data sample is represented by an n-dimensional feature vector, $X = (x1, x2, \ldots, Xn)$, depicting n measurements made on the sample from n attributes, respectively $A1, A2, \ldots, An$. Suppose that there are m classes, $C1, C2, \ldots, Cm$. Given an unknown data sample, X (i.e., having no class label), the classifier will predict that X belongs to the class having the highest posterior probability, conditioned on X. That is, the naïve probability assigns an unknown sample X to the class Ci if and only if: $P(Ci \mid X) > P(Cj \mid X)$ for $1 \leq j \leq m$, and $j \neq i$. Thus we maximize $P(Ci \mid X)$. The class Ci for which $P(Ci \mid X)$ is maximized is called the maximum posterior hypothesis. By Bayes' theorem

$$P(Ci \mid X) = \frac{P(X \mid Ci)P(Ci)}{P(X)}$$

As $P(X)$ is constant for all classes, only $P(X \mid Ci)\ P(Ci)$ need be maximized. If the class prior probabilities are not known, then it

*Author to whom correspondence should be addressed.

Adv. Sci. Lett. 23, 2495–2497, 2017

is commonly assumed that the classes are equally likely, that is, $P(C1) = P(C2) = \cdots = P(Cm)$, and we would therefore maximize $P(X \mid Ci)$. Otherwise, we maximize $P(X \mid Ci)P(Ci)$. Note that the class prior probabilities may be estimated by $P(Ci) = |Ci, D|/|D|$, where $|Ci, D|$ is the number of training tuples of class Ci in D. In a simplified way the naïve bayes equation can be written as[3]

$$\text{Posterior} = \frac{\text{prior} * \text{likelihood}}{\text{evidence}}$$

3. RESEARCH METHOD

HR selection process using Naïve Bayes method begins by determining criteria: education, GPA, interview, age, and experience. The rules of the specification table and the resulting were shown in Tables I to V, where the occurrence probability value of each criterion described. The decision results have calculated from the probability of occurrence likelihood value 'yes' and 'no.'

The sample test of two candidates was shown in Table VI and the resulting value of likelihood in Table VII. The probability value obtained must be the same as the first and the greatest probability value was accepted as a new employee.

Table I. Probability of educational value.

	Number of "selected" events		Probability	
Education	Accept	Reject	Accept	Reject
Diploma	8	8	8/16	8/16
Degree	8	8	8/16	8/16
Total	16	16	1	1

Table II. Probability of GPA value.

	Number of "selected" events		Probability	
GPA	Accept	Reject	Accept	Reject
< 2.75	0	8	0/16	8/16
≥ 2 75 < 3 00	8	0	8/16	0/16
≥3.00	8	8	8/16	8/16
Total	16	16	1	1

Table III. Probability of interview value.

	Number of "selected" events		Probability	
Interview	Accept	Reject	Accept	Reject
Good	6	4	6/16	4/16
Enough	6	4	6/16	4/16
Less	4	8	4/16	8/16
Total	16	16	1	1

Table IV. Probability of age value.

	Number of "selected" events		Probability	
Age	Accept	Reject	Accept	Reject
Productive	10	8	10/16	8/16
Un productive	6	8	6/16	8/16
Total	16	16	1	1

Table V. Probability of experience value.

	Number of "selected" events		Probability	
Experience	Accept	Reject	Accept	Reject
Yes	12	4	12/16	4/16
No	4	12	4/16	12/16
Total	16	16	1	1

Table VI. Data sample test.

Criteria	HR candidate A	HR candidate B
Name	Thomas	Mike
Education	Degree	Degree
GPA	≥ 2 75 < 3 00	< 2.75
Interview	Enough	Good
Age	Unproductive	Productive
Experience	Yes	Yes

Table VII. Value likelihood of each candidate.

Likelihood	Likelihood Yes	Likelihood No
HR candidate A	0.0132	0
HR candidate B	0	0.0039

Probability value calculation HR candidate A

$$\text{Probability Yes} = \frac{\text{Likelihood value Yes}}{\text{Likelihood value Yes} + \text{Likelihood value No}} = \frac{0{,}0132}{0{,}0132+0} = 1$$

$$\text{Probability No} = \frac{\text{Likelihood value No}}{\text{Likelihood value Yes} + \text{Likelihood value No}} = \frac{0}{0{,}0132+0} = 0$$

Based on the probability value HR candidate A, probability value 'Yes' is greater than the probability value 'No,' so the result of HR candidates with these criteria is accepted.

Probability value calculation HR candidate B

$$\text{Probability Yes} = \frac{\text{Likelihood value Yes}}{\text{Likelihood value Yes} + \text{Likelihood value No}} = \frac{0}{0+0{,}0039} = 0$$

$$\text{Probability No} = \frac{\text{Likelihood value No}}{\text{Likelihood value Yes} + \text{Likelihood value No}} = \frac{0{,}0039}{0+0{,}0039} = 1$$

Based on the probability value HR candidate B, probability values 'No' is greater than the probability value 'Yes,' so the result of HR candidates with these criteria is rejected.

4. RESULTS AND DISCUSSION

The system aims to make effective decisions. Measurement of effectiveness depending on the time indicator, accuracy in data analysis and output, as well as the relevance of the benefit. The indicators have been developed into a number of statements using

Table VIII. Indicators value and effectiveness.

No.	Indicators	Mean indicator
1.	Time	3,72
2.	Output	3,83
3.	Relevance	3,80
Indicator value		11,35
Mean effectiveness		3,78

a Likert scale. There are 12 statements for the three indicators set, then spread to 20 HR in each company as a respondent. The test results obtained from the mean value of each indicator, see Table VIII, where the mean value obtained 3,78 at intervals of 3,41 to 4,2 is the criteria of effective (E).

5. CONCLUSION

Based on the results of research and discussion can be concluded that the selection process and the position of HR took several criteria into consideration. A simulation test of the sample showed that the probability of "Yes" should be greater than the probability of "No" to produce an acceptable decision. Results of testing the effectiveness of the system are obtained by using a Likert scale that produces a value of 3.78 so that the system can be declared effective.

References and Notes

1. D. L. Olson and D. Dursun, Advanced Data Mining Techniques, Springer-Verlag, Berlin Heidelberg, USA **(2008)**.
2. G. Subbalakshmi, et al., *Indian Journal of Computer Science and Engineering IJCSE* 2 **(2011)**, ISSN: 0976–5166.
3. D. Ratnam, et al., *International Journal of Computer Science and Information Technologies IJCSIT* 5, 2384 **(2014)**, ISSN: 0975–9646.
4. D. M. Khairina, F. Ramadhani, S. Maharani, and H. R. Hatta, Department recommendations for prospective student vocational high school of information technology with naïve bayes method, *2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering ICITACEE* , IEEE **(2015)**, pp. 92–96.
5. S. Maharani, N. Dengen, G. Y. Saputra, D. M. Khairina, and H. R. Hatta, Expert system applications for early diagnosis teeth and oral disease in children, *2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering ICITACEE* , IEEE **(2015)**, pp. 87–91.

Received: 12 October 2016. Accepted: 2 November 2016.