Journal of Research and Development
Peer Reviewed International, Open Access Journal
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-17, Issue-1 | January - 2025

Original Article

Analysis of Effect of Faculty Satisfaction with Data Extraction Technique: A Perspective on Institutional Performance

Saiyad M. Khaji
Assistant Professor & HOD
Dept. of Computer ApplicationsSri K Puttaswamy First Grade College
VVCE Campus, Gokulam Mysore-02

Manuscript ID:
JRD -2025-170103
ISSN: 2230-9578
Volume 17
Issue 1|
Pp. 12-18
January 2025
Submitted: 02 Dec. 2024
Revised: 27 Dec. 2024
Accepted: 25 Jan. 2025
Published: 31 Jan. 2025

Abstract

This paper tries to analyse the effect on Institutional performance of lecturer satisfaction with help of one of the data extraction techniques. Data extraction is done by Naïve Bayes technique, in this method data is collected through interviews and questionnaires carried out in at least 10 higher education institutions in the Mysore. The qualitative criteria were Promptness, Compassion, reliability and Responsibility. The Survey indicates that the percentage of accuracy with 86.18% and 82.01% Precision, and 93.75% recall value. The data mining technique used is Naïve Bayes which is very much recommended to guess the degree of satisfaction of the faculty with regard to institutional performance based on results of tests using rapid miner application.

Keywords: *data Extraction, faculty satisfaction, prediction, naïve Bayes, Mysore*

Introduction:

One crucial determinant in comparing the service performance received and expected by clients is customer satisfaction. Customer satisfaction is something which can balance both personal and social relations with workers and others. The order of satisfaction perceived to be a requirement of desires and expectations, which in turn derived from perceived satisfaction. It is known that the satisfactory milestones of each available service's academic community can be an interlinked evaluation to find out if services are provided well or not. This observation focuses on assessing the services offered to faculties and their satisfaction which indirectly analyses the performance of an institution (Higher Education).

This can be done by using artificial intelligence and data analysis techniques. There are number of algorithms to analyse the data. The Naïve Bayes algorithm belongs to a group of data extraction Algorithm which turn large volume of unordered data into small set of ordered useful information. The Bayes theorem is known as classification algorithm which works on both text and unrealistic data.

It is highly referred in non-numeric classification. In textual classifier tasks, input has high dimension (as every word refer on indicator data). It is used in spam filtering, sentiment detection, rating classification etc. The advantage of using naïve Bayes is its speed. It is fast and making prediction is easy with high dimension of data. We hope to measure the progress of the institution to provide services for faculty through the naïve bayes algorithm on satisfaction of faculties with institutional performance.

Methodology

The data is collected from observations in the form of questionnaire provided in Google forms for at least 10 Under Graduate level higher education institutes from Mysore and the process follows following stages.

Address for correspondence:
Saiyad M. Khaji, Assistant Professor & HOD Dept. of Computer Applications Sri K Puttaswamy First Grade College VVCE Campus, Gokulam Mysore-02

Website:
https://jrdrvb.org/

DOI:10.5281/zenodo.14949773

Access this article online

How to cite this article:
Khaji, S. M. (2025). Analysis of Effect of Faculty Satisfaction with Data Extraction Technique: A Perspective on Institutional Performance. Journal of Research & Development, 17(1), 12–18. https://doi.org/10.5281/zenodo.14949773

Journal of Research and Development
Peer Reviewed International, Open Access Journal
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-17, Issue-1 | January - 2025

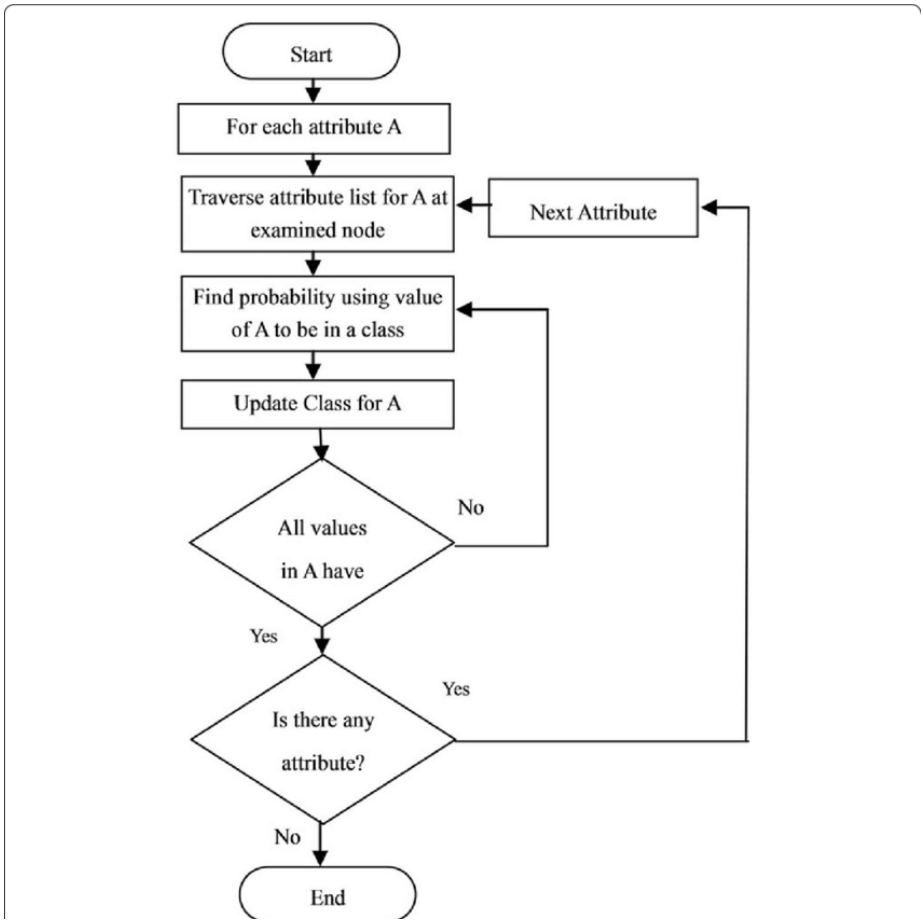


Fig 1 – Naïve Bayes Algorithm Flowchart (Source: Reaserchgate.net)
Based on the flowchart above in fig-1, the following can be illustrated:

1. Collecting and entering questionnaire data from the lecturer concerned at a private institution in the city of Mysore, (UG-Technical courses).
2. Testing of pre-existing training data.
3. Assess the performance of the algorithm. In the testing process, the performance of thealgorithm is tested using a data test set where the data set with the training data set consists of different data.
4. In addition, the implementation of previous data using the naive Bayes algorithm in order to produce accurate accuracy, precision and recall.
5. The system then issues the output in the percentage of satisfaction of the lecturerwith the performance of the institution.

Results and Discussion

In this case, problem-solving is carried out by conducting an assessment of the satisfaction of lecturers on institutional performance by conducting interviews and distributing questionnaires in the form of several questions created in Google forms to lecturers on institutional performance. The solution provided uses the Naïve Bayes method for data mining. The criteria used for the assessment are Promptness, Compassion, reliability and Responsibility The results of the analysis were calculated using software and Microsoft Excel.

Table 1. Average value of each attribute

No	Promptness,	Compassion,		reliability			Responsibility		
1.1	4	5	7	5	5	7	3	4	5
2.1	6	6	4	3	5	5	6	7	5
3.1	6	5	7	5	7	6	5	7	7
4.1	7	6	4	6	6	5	7	6	7
5.1	5	7	5	6	6	7	6	5	6
6.1	6	5	5	6	7	6	5	6	7
7.1	7	6	4	7	5	6	5	7	6

Journal of Research and Development
Peer Reviewed International, Open Access Journal
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-17, Issue-1 | January - 2025

8.1	7	4	5	6	7	7	5	5	7
9.1	5	6	7	7	7	4	7	5	7
10	4	5	4	6	6	5	7	6	5
11	5	7	6	5	7	6	7	7	6
12	6	5	5	5	7	7	6	6	6
13	6	5	7	5	5	5	5	7	7
14	7	6	5	7	5	6	5	5	6
15	6	4	6	5	5	6	6	5	7
16	5	6	5	7	7	7	6	7	6
17	4	7	7	6	6	5	5	6	7
18	5	6	6	5	4	6	6	7	6
19	3	5	5	7	5	7	7	6	5
20	5	7	7	5	7	5	6	7	4
21	6	5	5	7	6	6	7	5	7
22	7	7	6	5	5	7	6	6	4
23	4	5	7	6	5	4	5	7	6
24	5	6	7	4	6	5	7	5	5
25	5	4	5	6	7	6	4	6	5
26	6	6	4	4	7	7	6	5	7
27	7	7	4	6	7	5	6	7	6
28	7	6	5	4	6	5	5	5	6
29	7	7	6	6	4	4	7	7	6
30	6	6	7	4	5	6	6	7	7

The data collected is questionnaire data which is distributed to. The sample data per parameter is grouped into two value intervals as in the following table.

Table 2.

Questionnaire Value Group Per Parameter		
No	**Value Interval**	**Group**
1	3 – 5	Low/little
2	6 -7	High/huge

Classify the results into two value intervals as in the following table:

Table 3. Classification of Questionnaire Value Results

No	Value Interval	Classification
1	3 – 5	Not satisfied/ unhappy
2	6 -7	Satisfied/ happy

Initial data were obtained using Microsoft Excel to produce the data sets in the following table:

Table 4. Training data

No	Promptness	Compassion	Reliability	Responsibility	Result
1	Huge	High	Low	Low	Unhappy
2	Huge	Little	High	Low	Unhappy
3	Huge	Huge	High	High	Happy
4	Huge	Huge	High	High	Happy
5	Huge	Huge	High	Low	Happy
6	Huge	Huge	Low	High	Happy
7	Huge	Huge	Low	Low	Unhappy
8	Huge	Huge	High	High	Happy
9	Huge	Huge	High	High	Happy
10	Low	Low	High	Low	Unhappy
11	High	High	High	High	Happy
12	High	High	Low	High	Happy
13	High	Little	Low	High	Not satisfied
14	High	High	Low	High	Satisfied

Journal of Research and Development
Peer Reviewed International, Open Access Journal
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-17, Issue-1 | January - 2025

15	High	Little	Low	High	Unhappy
16	Low	High	High	High	Satisfied
17	High	High	Low	High	Satisfied
18	Low	Little	High	High	Not satisfied
19	Low	High	High	Low	Not satisfied
20	High	Huge	High	Low	Satisfied
21	High	High	High	High	Satisfied
22	High	Huge	High	Low	Satisfied
23	High	Low	Low	High	Not satisfied
24	High	High	High	Low	Satisfied
25	Low	Huge	Low	Low	Not satisfied
26	High	High	Low	High	Satisfied
27	High	Huge	High	High	Satisfied
28	High	Low	Low	High	Not satisfied
29	High	Little	High	High	Happy
30	High	Low	Low	High	Unhappy

a) **Finding the Mean** ValueSatisfied Class = 24
Class not satisfied = 6
μ(Satisfied)(A1)= (5+6+6+6+4+5+6+5+6+6+6+5+6+6+5+6+6+6+5+7+5+6+7+6)/24
= (136)/24
= 5,6666
μ(Satisfied)(A2)= (4+6+6+6+6+6+7+6+6+6+6+5+6+5+6+6+5+6+7+6+6+6+5+5)/24
= (139)/24
= 5,7916
μ(Satisfied)(A3)= (7+6+7+6+6+7+5+6+7+7+6+6+5+6+7+6+7+7+6+6+6+7+7+7)/24
= (153)/24
= 6,3750
μ(Satisfied)(A4)= (5+7+7+6+7+6+7+7+5+6+6+7+6+7+6+7+6+4+7+4+7+6+6+7)/24
= (149)/24
= 6,2083
μ(Not satisfied)(A1) = (5+4+5+6+5+6)/6
= (31)/6
= 5,1666
μ(Not satisfied)(A2) = (6+6+5+5+7+5)/6
= (34)/6
= 5,6666
μ(Not satisfied)(A3) = (4+7+6+6+5+5)/6
= (33)/6
= 5,500
μ(Not satisfied)(A4) = (5+5+6+5+5+6)/6
= (32)/6
= 5,3333

a) Standard Deviation

σ(Satisfied)(A1) = √(((5-5,6666)^2)+((6-5,6666)^2)+((6-5,6666)^2) + ((6-5,6666)^2)
+((4-5,6666)^2)+((5-5,6666)^2)+((6-5,6666)^2) + ((5-5,6666)^2)
+((6-5,6666)^2)+((6-5,6666)^2)+((6-5,6666)^2) + ((5-5,6666)^2)
+((6-5,6666)^2)+((6-5,6666)^2)+((5-5,6666)^2) + ((6-5,6666)^2)
+((6-5,6666)^2)+((6-5,6666)^2)+((5-5,6666)^2) + ((7-5,6666)^2)
+((5-5,6666)^2)+((6-5,6666)^2)+((7-5,6666)^2) +((6-5,6666)^2))
/(24-1)
= √11,332/23
= √0,4926
= 0,2426
σ(Not satisfied)(A2) =√(((5-5,1666)^2) + ((4-5,1666)^2) + ((5-5,1666)^2)
+((6-5,1666)^2) + ((5-5,1666)^2) + ((6-5,1666)^2))
/(6-1)
= √2,833/5
= √0,5666

Journal of Research and Development
Peer Reviewed International, Open Access Journal
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-17, Issue-1 | January - 2025

= 0,32103556

Table 5. Capacity Probability Value

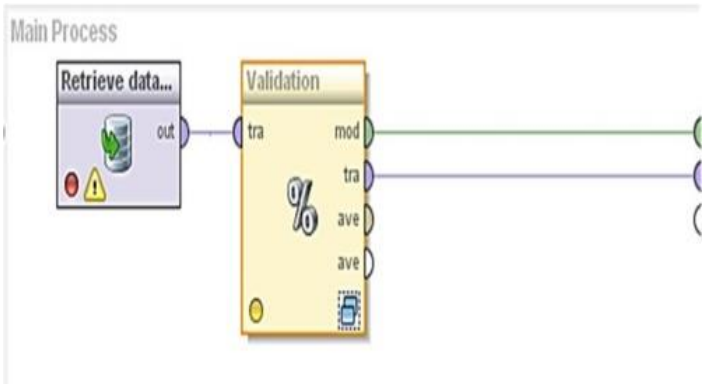
Value	Promptness		Compassion		Reliability		Responsibility	
	Happy	Unhappy	Happy	Unhappy	Happy	Unhappy	Happy	Unhappy
µMean	56,666	51,666	57,916	56,666	6,375	5,5	62,083	53,333
σDeviasi^2	0,4926	0,5666	0,4033	0,6666	0,4184	1,1	0,8676	0,2664
σDeviasi	0,2426	0,321	0,1626	0,4443	0,175	1,21	0,7527	0,0709

The next step is to calculate the probability for the Testing data feature.

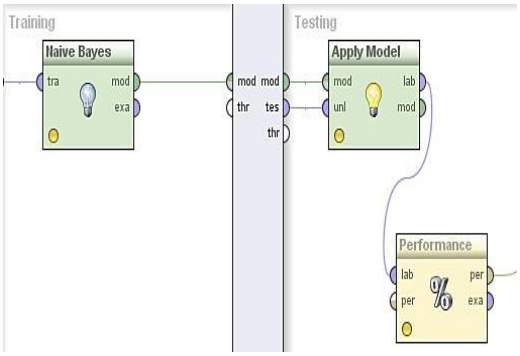
Table 6. Data Testing

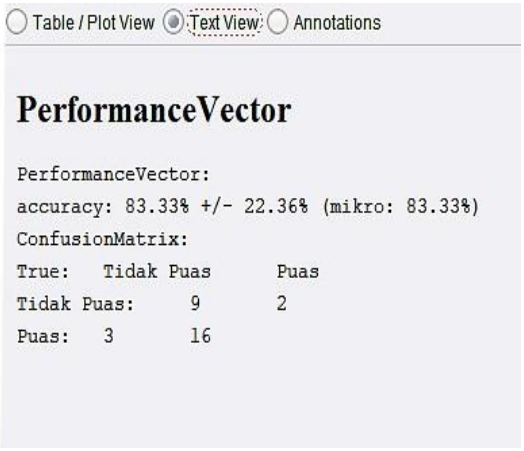
No	Promptnes	Compassion	Reliability	Responsibility	Result
1	Huge	High	Low	Low	?
2	Huge	Low	High	Low	?
3	Huge	High	High	High	?
4	Huge	High	High	High	?
5	Huge	High	High	Low	?
6	Huge	High	Low	High	?
7	High h	High	Low	Low	?
8	Huge	High	High	High	?
9	High	High	High	High	?
10	Little	Low	High	Low	?
11	High	High	High	High	?
12	High	High	Low	High	?
13	Huge	Low	Low	High	?
14	High	High	Low	High	?
15	High	Low	Low	High	?
16	Low	High	High	High	?
17	High	High	Low	High	?
18	Low	Low	High	High	?
19	Little	High	High	Low	?
20	High	High	High	Low	?
21	High	High	High	High	?
22	High	High	High	Low	?
23	High	Low	Low	High	?
24	Huge	High	High	Low	?
25	Low	High	Low	Low	?

In addition, RapidMiner is used to process the data and to test it with Split Validation found in RapidMiner. The results of the data set tests are as follows:

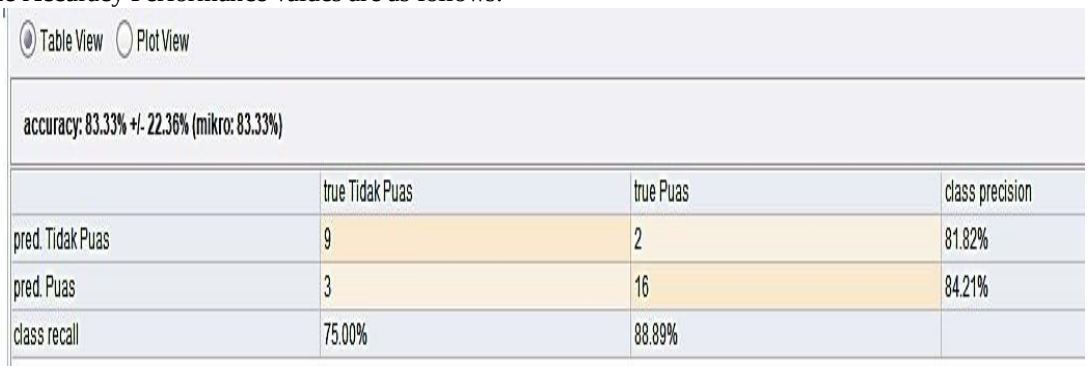


Journal of Research and Development
Peer Reviewed International, Open Access Journal
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-17, Issue-1 | January - 2025





And the Accuracy Performance values are as follows:



	true Tidak Puas	true Puas	class precision
pred. Tidak Puas	9	2	81.82%
pred. Puas	3	16	84.21%
class recall	75.00%	88.89%	

Figure 5. Value of Accuracy Performance

Figure 4. Performance Vector

Information:

a) The number of predictions Eligible and in fact True Eligible is 9 records
b) The number of Predictions of Unfeasible and in fact correct is 3 records.
c) The number of predictions Eligible and in fact true of Unfeasible is 2 records.
d) The number of predictions of Eligible and in fact true of Unreasonable is 16 records

The Accuracy Performance Value shows in the picture that the prediction of the eligible value is 9, with an accuracy rate of 81.82.0%; the incorrect forecast is 10 and its accuracy is 84.25%. The accuracy achieved is therefore 83.33%.

Conclusion

As per the output of tests from the input provided with readiness, empathy, reliability and responsibility attributes, which are based on the Naive Bayes method, an accuracy level of 85.48% with an precision value of 81.08% and a recall value of 93.75% is achieved. Based on the test finding,a lecturer satisfaction with institutional performance can be predicted using the Naïve Bayes Method. Since the accuracy and the recall value of this research are high, it is possible to combine or compare this research with other classification algorithms.

Journal of Research and Development
Peer Reviewed International, Open Access Journal
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-17, Issue-1 | January - 2025

References:

1. Aisyah, S., Hasibuan, H. A., Nadeak, B., & Tohir, A. (2021). Analysis of the effect of the lecturer satisfaction with the naive Bayes data mining technique on institutional performance. *Journal of Physics: Conference Series*, *1933*(1), 012034. https://doi.org/10.1088/1742-6596/1933/1/012034
2. Budiharjo, T. Soemartono, A. P. Windarto, and T. Herawan, "Predicting School Participation in Indonesia using Back-Propagation Algorithm Model," *Int. J. Control Autom.*, vol. 11, no. 11, pp. 57–68, 2018.
3. Budiharjo, T. Soemartono, A. P. Windarto, and T. Herawan, "Predicting tuition fee payment problem using backpropagation neural network model," *Int. J. Adv. Sci. Technol.*, vol. 120, pp.85–96, 2018.
4. B. Supriyadi, A. P. Windarto, T. Soemartono, and Mungad, "Classification of natural disaster prone areas in Indonesia using K-means," *Int. J. Grid Distrib. Comput.*, vol. 11, no. 8, pp. 87–98, 2018.
5. A. Z. Adnan, M. H. Holili, and B. Indrawan, "Analysis of Decision Support System with Analytical Hierarchy Process Method," *J. Phys. Conf. Ser. Pap.*, 2021.
6. M. Darmowiyono, W. Yuliyanto, and K. I. Purnomo, "Application of the Simple Multi Attribute Rating Technique (SMART) Method in the selection of thrush medicine products based on consumers Application of the Simple Multi Attribute Rating Technique (SMART)Method in the selection of thrush medicine pro," *J. Phys. Conf. Ser. Pap.*, 2021.
7. H. Jatnika, M. Huda, R. R. Amelia, and M. A. Manuhutu, "Analysis of Data Mining in the Group of Water Pollution Areas using the K-Means Method in Indonesia Analysis of Data Mining in the Group of Water Pollution Areas using the K-Means Method in Indonesia," *J.Phys. Conf. Ser. Pap.*, 2021.