

Application of Machine Learning with the Concept of Data Mining Rough Set Methods in the Prediction of Lecturer Performance level against Tri Dharma

Wahyu Saptha Negoro^{a,1}, Asbon Hendra Azhar^{b,2}, Ratih Adinda Destari^{c,3}, Fujiati^{d,4},
Ardan Syah^{e,5}

^{abcde} Universitas Potensi Utama, Jl.K.L.Yos Sudarso Km.6,5 No.3 A Medan, Indonesia
wahyusaptha1707@gmail.com, asbon.upu@gmail.com, adindaalkarim0384@gmail.com,
fuji.potensiutama@gmail.com, ardannst26@gmail.com

ABSTRACT

Lecturers are human resources who are very influential in changing thought patterns and in the teaching and learning process in a higher education institution. As a quality lecturer, lecturers must fulfill the tri dharma performance which consists of Education, Research, service and several external duties so as to make lecturers quality in educating and improving knowledge in their field. Data mining is a method in Computer Science that is commonly used in the process of seeking knowledge, some fairly simple data mining algorithms are a rough collection. The rough set method is an efficient technique for. Finding hidden knowledge in large-scale databases, so that what is obtained from lecturer performance can become knowledge as a reference in the future for determining lecturer performance decisions based on past experiences stored in the database. The data consists of 2 attributes where the condition attributes are Education, Research, Service attributes and external task and decision attributes. Research on lecturer performance was carried out at major potential universities, namely college lecturers, thereby producing knowledge about lecturer achievements. The aim of applying the rough set method is to be able to determine the performance achievements of lecturers in implementing the tri dharma each semester based on lecturer data stored for a long period of time, so that you get quality lecturers

Keywords : Lecturer Performance Analysis, Data Mining, Rough Set Method

Article Info :

Submitted:17 Juli

Reviewed:01 Agustus

Accepted :25 Oktober

Copyright © 2022 – IT Journal. All rights reserved.

1. INTRODUCTION

Lecturers are lecturers at a University who must have quality and performance towards Tri Dharma in changing the mindset of students, providing useful knowledge and completing all tasks as lecturers. This can be realized by the efforts that must be fulfilled as a lecturer, namely carrying out tasks in the form of Education, Research, devotion and external tasks.

In the processing of data in the database can use data Mining where data mining is a method in Computer Science commonly used in the process of searching for knowledge [1]. The method used is part of the data mining is Rough Set (RS) method is part of machine learning that analyzes the uncertainty of the dataset used to determine the attributes of important objects (Classification) [2],[3]. The purpose of this study is to extract information from the rough set on the case of the prediction of the level of performance of lecturers against the Tri Dharma.

The advantage of using the Rough Set method is that it does not require initial or additional information from Probability data in statistics, membership levels or probability values in fuzzy set theory [4]. The main purpose of the Rough Set method for extracting rules from the database, reducing the size of data and dealing with inconsistencies in the data [5]. This study tries to explore the knowledge

The steps that must be applied so that research can be carried out in a directed manner and make it easier to analyze existing problems can be seen in the image below.

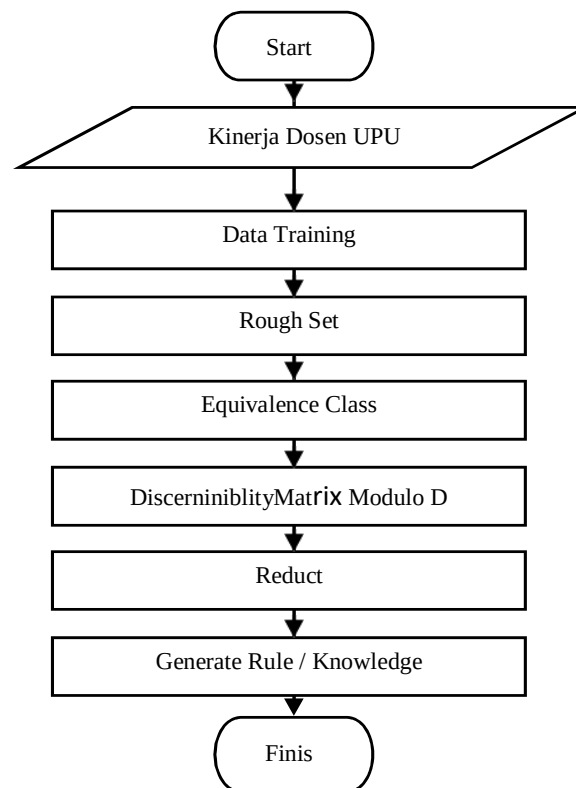


Figure 2. Research Stages

C. Data Mining

Data mining is a process that uses mathematics, statistics, artificial intelligence and machine learning techniques to extract and identify useful information and further get knowledge from the database [9]. Data mining or also known as knowledge discovery in database (KDD), data is analyzed from various different perspectives and then summarizes the data into desired information and can be used both for classification, grouping and for predicting and patting certain patterns or trends as something that new. Knowledge. In the process of searching for information, various techniques and processes of computer study are carried out iteratively and interactively to determine valid, useful, and understandable patterns [10].

Rough Set was built by Zdzislaw Pawlak in the early 1980s. The philosophy of the Rough Set is that information (knowledge, data) can be associated with objects. In a Rough Set, a data set is represented as a table, where the rows in the table represent objects and the columns represent the attributes of the objects [11]. The Rough Set method is one of the data mining techniques that deals with the problem of Uncertainty, Imprecision and Vagueness in Artificial Intelligence (AI) applications. Rough set theory is an efficient technique for Knowledge Discovery in Database (KDD) in the process and Data Mining stages [12].

Stages in using the Rough Set algorithm:

- Data Selection (Selection of data to be used in the database).
- Establishment of a Decision System that contains condition attributes and attributes in decisions.
- Formation of Equivalence Class, namely by eliminating repetitive data.
- Discernibility Matrix Modulo D, is a matrix that contains comparisons between data with different conditions of attributes and attributes of Formation decisions.
- Generate reduct using boolean algebra.

- have a rule (new knowledge stored in the database).

3. RESULT AND DISCUSSION

The following is an explanation of the results of research regarding the prediction of the level of performance of lecturers against the Tri Dharma using the Rough Set method.

A. Data Selection

The data used in this study are lecturers' performance data on the task load of the Tri Dharma, while the lecturers in this study were lecturers at the Main Potential University. Attributes of the conditions used there are 4 attributes including education, research, dedication and external assignments and attribute decisions from the results of lecturer performance.

The components used in this study as the basis for the assessment components are as follows:

- Education (Maximum 20%)
- Research (Maximum 30%)
- Community service (Maximum 30%)
- Outside Task (Maximum 20%)

So that obtained from the assessment component in Education can be seen in Table 1.

Table 1. Criteria For Assessment Of Educational Components

No	Education Criteria	Weight
1.	Number of Face to face	10%
2.	Making UTS and UAS Questions	5%
3.	Evaluation	5%

The assessment component in the study can be seen in Table 2.

Table 2. Education Component Assessment Criteria

No	Research Criteria	Weight
1.	Research Scientific Works, Articles and Books	10%
2.	Research Grant	20%

The components of the assessment on Community Service can be seen in Table 3.

Table 3. Criteria For Assessment Of Community Service Components

No	Community Service Criteria	Weight
1.	PkM Mandiri	10%
2.	Internal Pkm	10%
3.	External PkM	10%

Components of the assessment on External Assignments, can be seen in Table 4.

Table 4. External Assignment Component Assessment Criteria

No	External Task Criteria	Weight
1.	Conference Moderator	10%
2.	Campus Event Organizing Committee	10%

After the weight of the component values of each criterion on the condition attribute is obtained, the next step is Data Transformation.

B. Data Transformation

The results of the assessment in the form of categories with a range can be seen in the following table:

Table 5. Assessment Results With Range

Range	Category	Score
<30	Not enough	1
31≤50	Enough	2
51≤80	Well	3
81≤100	Very good	4

The assessment category in Table 5 is an assessment for the decision systems stage.

C. Formation of Decision Systems

Decision Systems from lecturer performance predictions consist of:

- Condition Attributes: education, research, service and external assignments.
- Decision Attribute: Results from Lecturer.

The following is an example of the data used in the decision system consisting of 8 (eight) data in Table 6:

Table 6. Decision Systems

Lecturer	Value Education	Value Research	Value Devotion	O T	Total Value	Decision
Joni	15	25	21	10	71	Well
Andi	17	26	25	8	76	Well
Roni	18	28	27	9	82	Very good
Ibnu	16	23	29	7	75	Well
Siti	20	22	22	8	72	Well
Ridwan	14	15	10	10	49	Enough
Wahyu	15	17	28	10	70	Well
Ahmad	18	10	15	7	50	Enough

D. Formation of Equality Class

The next stage is the formation of the Equivalence Class because the attributes in the decision system have not been transformed, so the transformation stage must be carried out so that the value of the attribute has a simpler value condition to be carried out in the next step.

At Attribute A (Education), there are 4 groups as follows:

$$0 \leq X \leq 5 = 1$$

$$6 \leq X \leq 10 = 2$$

$$11 \leq X \leq 15 = 3$$

$$16 \leq X \leq 20 = 4$$

In Attribute B (Research), there are 4 groups as follows:

$$0 \leq X \leq 10 = 1$$

$$11 \leq X \leq 15 = 2$$

$$16 \leq X \leq 25 = 3$$

$$26 \leq X \leq 30 = 4$$

In Attribute C (Community service), there are 4 groups as follows:

$$0 \leq X \leq 5 = 1$$

$$6 \leq X \leq 11 = 2$$

$$12 \leq X \leq 20 = 3$$

$$21 \leq X \leq 30 = 4$$

In Attribute D (Outer Tasks), there are 4 groups as follows:

$$0 \leq X \leq 6 = 1$$

$$7 \leq X \leq 10 = 2$$

$$11 \leq X \leq 15 = 3$$

$$16 \leq X \leq 20 = 4$$

After the transformation stage is carried out, the results of the formation of decision systems can be seen in Table 7 below.

Table 7. Value Formation Of Decision Systems After Transformation

Lecturer	Value Education	Value Research	Value Devotion	O T	Total Value	Decision
Joni	3	3	4	2	3	3
Andi	4	4	4	2	3	3
Roni	4	4	4	2	4	4
Ibnu	4	3	4	2	3	3
Siti	4	3	4	2	3	3
Ridwan	3	2	3	2	2	2
Wahyu	3	3	4	2	3	3
Ahmad	4	3	3	2	2	2

The formation of the Decision System value after being transformed in Table 7 above, then the next step is the formation of the Equivalence Class, where the Equivalence Class is grouping objects that have the same condition attribute values based on the decision system table [13].

Tabel 8. Equivalence Class

EC	A	B	C	D	DC
EC1	3	3	4	2	3
EC2. 1	4	4	4	2	3
EC2. 2	4	4	4	2	4
EC3	4	3	4	2	3
EC4	3	2	3	2	2
EC5	4	3	3	2	2

In EC2.1 and EC2.2, they have different decision values while the condition attributes are all the same but in this study it is considered an inconsistent value. So that the results of the Equivalence Class that are formed there are 5 Equivalence Class.

E. Formation of Discernibility Matrix Modulo D

Discernibility matrix is the columns in the matrix filled with a set of condition attributes that have different condition values. The result value of the Discernibility Matrix can be seen in Table 9.

Tabel 9. Discernibility Matrix

EC	EC1	EC2	EC3	EC4	EC5
EC1	X	AB	A	BC	AC
EC2	AB	X	B	ABC	BC
EC3	A	B	X	ABC	C
EC4	BC	ABC	ABC	X	AB
EC5	AC	BC	C	AB	X

While the Discernibility matrix module D is a matrix that contains comparisons between data with different condition attributes and decision attributes. Data with different condition attributes, but the same decision attributes are still considered the same [14][15]. The result value of the Discernibility matrix module D can be seen in Table 10.

Tabel 10. Discernibility Matrix Modul D

EC	EC1	EC2	EC3	EC4	EC5
EC1	X	X	X	BC	AC
EC2	X	X	X	ABC	BC
EC3	X	X	X	ABC	C
EC4	BC	ABC	ABC	X	X
EC5	AC	BC	C	X	X

Based on the results obtained in the Discernibility matrix, module D has 5 classes which will be described in the Reduct formation stage.

F. Formation of Reduct

Reduct is the selection of the minimum attribute (interesting attribute) which from a set of attribute conditions using the Prime Implicant Boolean function. The collection of all Prime Implicants determines the sets of reducts [16].

The Reduction process is used to select the Condition Attributes that will be used to generate Knowledge in a way. Why was it chosen because not all knowledge will be used to generate knowledge, how to increase knowledge by making Boolean Algebra equations based on Discernibility Matrix or Discernibility Matrix Modulo D so the results are used as REDUCT. The resulting reduction can be seen in Table 11.

Table 11. Lecturer Performance Reduct

Class	Boolean Algebra	Results	Reduct
E1	$(B \vee C) \wedge (A \vee C)$	$C \vee (A.B)$	$\{C\}, \{AB\}$
E2	$(A \vee B \vee C) \wedge (B \vee C)$	$B \vee C$	$\{B\}, \{C\}$
E3	$(A \vee B \vee C) \wedge C$	C	C
E4	$(B \vee C) \wedge (A \vee B \vee C) \wedge (A \vee B \vee C)$	$B \vee C$	$\{B\}, \{C\}$
E5	$(A \vee C) \wedge (B \vee C) \wedge C$	C	C

Based on Table 11 above, the Reducts on Lecturer Performance towards Tri Dharma are 3 Reducts, namely:

- $\{C\}$ = {Community service}
- $\{AB\}$ = {Education, Research}
- $\{B\}$ = {Research}

G. Generate Rule / Knowledge

Reduct results obtained are used to generate Knowledge by referring to the Decision System table. So that the resulting Rule by comparing it with a combination of existing attributes, as follows:

- $\{C\} = \{\text{Community service}\}$
If Community service is worth = 4, then the decision is worth = 3 or the decision is worth = 4
If Community service is worth = 3, then the decision is worth = 2
- $\{A,B\} = \{\text{Education, Research}\}$
If Education is worth = 3 and Research is worth = 3, then the decision is worth = 3
If Education is worth = 4 and Research is worth = 4, then the decision is worth = 3 or or the decision is worth = 4
If Education is worth = 4 and Research is worth = 3, then the decision is worth = 3 or or the decision is worth = 2
If Education is worth = 3 and Research is worth = 2, then the decision is worth = 2
- $\{B\} = \{\text{Research}\}$
If Research is worth = 3, then the decision is worth = 3 or the decision is worth = 2
If Research is worth = 4, then the decision is worth = 3 or the decision is worth = 4
If Research is worth = 2, then the decision is worth = 2

4. CONCLUSION

Based on this research, 9 knowledges were produced consisting of 2 knowledge from Reduct Service, 4 knowledge from Reduct Education and research and 3 knowledge from Reduct Research.

The knowledge generated consists of if and then, this knowledge can be used in the future to determine or determine the income of a lecturer based on past experiences stored in the database.

ACKNOWLEDGEMENT

I would like to thank you for your support for this research for completing the research carried out on the performance of lecturers at Potential Main University

REFERENCES

- [1] Enda Esyudha Pratama, Helen Sastypratiwi, Yulianti. (2021). Analisis Kecenderungan Informasi Terkait Covid-10 Berdasarkan Big Data Sosial Media dengan Menggunakan Metode Data Mining. JIP (Jurnal Informatika Polinema). Volume 7, Edisi 2, Februari 2021.
- [2] Mokhamad Ramdhani Raharjo, Agus Perdana Windart. (2021). Application of Machine Learning with the Concept of Data Mining Rough Set (Predicted Level of Student Understanding of the Course). Jurnal Media Informatika Budidarma, Volume 5, Nomor 1, Januari 2021, Page 317-326. DOI 10.30865/mib.v5i1.2745
- [3] Ani Apriani, Iwan Tri Riyadi Yanto, Septiana Fathurrohman, Sri Haryatmi, Danardono. (2018). Variable precision rough set model for attribute selection on environment impact dataset. International Journal of Advances in Intelligent Informatics. Vol. 4, No. 1, March 2018, pp. 70-75. <https://doi.org/10.26555/ijain.v4i1.109>
- [4] Martua Tri Januar Sinaga, Rito Goejantoro, dan Fidia Deny Tisna Amijaya. (2017). The Application of If-Then Method from Rough Set Theory in Handling of Traffic Accidents in Samarinda City 2016. Jurnal EKSPONENSIAL Volume 8, Nomor 2, Nopember 2017.
- [5] Ulfah Indriani. (2018). Penerapan Metode Rough Set Dalam Menentukan pembelian Smartphone Android Oleh Konsumen. urnal Teknik Informatika Kaputama (JTIK) Vol. 2 , No. 1, Jan 2018.
- [6] Jeprianto, RZ Abdul Aziz. (2020). Implementasi Algoritma Rough Set Dan Naive Bayes Untuk Mendapatkan Rule Dalam Menyeleksi Pemohon Bantuan Fasilitas Rumah Ibadah (Studi Kasus : Pemerintah Kabupaten Pringsewu). JTKSI, Vol.03 No.02 Mei 2020. Hal. 74-83.
- [7] Muhammad Ardiansyah Sembiring, Raja Tama Andri Agus, Mustika Fitri Larasati Sibuea. (2021). Analisis Kepuasan Pelanggan Menggunakan Metode Rough Set. Journal of Science and Social Research. June 2021, IV (2): 227 – 230.
- [8] Muheet Ahmed Butt, Sheikh Amir. (2022). Knowledge Discovery in Geographical Sciences—A Systematic Survey of Various Machine Learning Algorithms for Rainfall Prediction. International Conference on Innovative Computing and Communications, Advances in Intelligent Systems and Computing 1388. https://doi.org/10.1007/978-981-16-2597-8_51
- [9] Putu Aryasuta Wicaksana, Ida Bagus Alit Swamardika, Rukmi Sari Hartati. (2022). Literature Review Analysis of Customer Behavior Using RFM Model. Electrical Technology Scientific Magazine, Vol. 21, No.1, Januari-Juni. DOI: <https://doi.org/10.24843/MITE.2022.v21i01.P04>
- [10] Dedy Hartama, Hartono. Analisis Kinerja Dosen Stmik Ibbi Dengan Menggunakan metoderough Set. (2016). Seminar Nasional Teknologi Informasi dan Multimedia 2016
- [11] M. Arif Rahman. Penerapan Metode Rough Set Dalam Memprediksi Penjualan Perumahan (Studi Kasus Di Pt. Anugerah Pasadena Pekanbaru. (2020). Volume 14, Nomor 2: 342-355
- [12] Amat Sofiyan, Ahmedika Azkiya. Penerapan Metode Rough Set Menganalisis Penyakit Yang Sering Dikeluhkan Pasien (Studi Kasus Puskesmas Jaya Mukti Dumai). (2022). Jurnal Informatika, Manajemen dan Komputer, Vol. 14 No. 1 , Mei 2022

- [13] Rahmat Jaya Zega. Penerapan Algoritma Rough Set Untuk Mengidentifikasi Faktor – Faktor Resignnya Karyawan (Study Kasus : PT. Sumber Alfaria Trijaya, Tbk). (2021). Pelita Informatika : Informasi dan Informatika, Volume 9, Nomor 4, April 2021. Hal: 296-302.
- [14] Wita Yulianti, Zikri. Analisis Persepsi Kepuasan Mahasiswa Tentang Kualitas Pelayanan Administrasi Mahasiswa Dengan Metode Rough Set. (2022). Journal of Machine Learning and Data Analytics (MALDA) Volume 01, No. 01, January 2022 Page: 65-74