

ANALISA DAN PERANCANGAN PREDIKSI TINGKAT PRESENTASI MAHASISWA BARU MASUK SEBAGAI MAHASISWA AKTIF DI STIKOM UYELINDO KUPANG MENGGUNAKAN *ROUGH SET*

Erna Rosani Nubatonis

Program Studi Teknik Informatika, STIKOM Uyelindo Kupang
Jln. Perintis Kemerdekaan I - Kayu Putih – Kupang – NTT - Indonesia
Email : ernarosaninubatonis@gmail.com

ABSTRACT

Acceptance of new students is the most important part of STIKOM UYELINDO Kupang as one of the benchmarks for the progress of the campus in the future. In the process of admitting new students (PMB), prospective new students must go through several stages of registration until the stage of filling out the KRS, so that the students concerned are legitimately declared as active students of STIKOM UYELINDO KUPANG. However, many cases occur that not all students arrive at the final stage of filling in the KRS to be declared as active students. Problems that occur result in the division responsible for new students difficult to predict that prospective students concerned in the process of admitting new students, will go through the process until the status of filling KRS or not, and also affect the prediction of the number of new student achievement. This study aims to find out and recognize the pattern of classification of new student registration status so that the level of presentation of new students entering the STIKOM UYELINDO KUPANG can be made by applying the rough set algorithm. In the process of applying Rough Set, it will produce a rule as a rule or pattern for classification of new student registration status data. The data used in this study is the data of new student registration in 2016-2018 with a total record of 579 records. The results of this study are expected to be an important input for the responsibility of new students and high school education institutions, in the strategy of screening new students to achieve the target of better new student admissions.

Keywords: Data Mining, Rough Set, Acceptance of New Students, Prediction, Qualitative Measure

1. PENDAHULUAN

STIKOM UYELINDO KUPANG adalah perguruan tinggi swasta di Nusa Tenggara Timur yang melakukan penerimaan mahasiswa baru setiap tahun. Dalam proses penerimaan mahasiswa baru (PMB), calon mahasiswa baru harus melalui beberapa tahap pendaftaran mulai dari pengisian formulir pendaftaran sampai pada mahasiswa dinyatakan sebagai mahasiswa aktif. Pada kenyataannya setiap kali pendaftaran calon mahasiswa baru, tidak semua

mahasiswa sampai pada tahap akhir pengisian KRS untuk dinyatakan sebagai mahasiswa aktif. Berdasarkan pada kenyataan ini maka penanggung jawab penerimaan mahasiswa baru dalam hal ini Humas & Promosi STIKOM UYELINDO KUPANG, sulit melakukan prediksi bahwa calon mahasiswa yang bersangkutan dalam proses penerimaan mahasiswa baru, akan melalui proses sampai pada pengisian KRS atau tidak. Banyak faktor yang bisa dijadikan acuan untuk memprediksi apakah calon mahasiswa yang bersangkutan akan

melalui semua proses sampai sah menjadi mahasiswa aktif.

Masalah yang dihadapi adalah bagaimana mengetahui dan mengenal pola klasifikasi status pendaftaran mahasiswa baru sehingga dapat dilakukan prediksi tingkat presentasi mahasiswa baru yang masuk di STIKOM UYELINDO KUPANG.

Ada berbagai macam cara yang dapat digunakan untuk menyelesaikan permasalahan diatas, salah satunya adalah dengan teknik penambangan data (data mining). Metode teknik penambangan data yang digunakan dalam penelitian ini adalah Rough Set. Rough Set digunakan untuk melakukan penanganan ketidak jelasan dan ketidak pastian. Rough Set dikembangkan oleh Zdzislaw Pawlak (Akseptor and Vasektomi, 2014). Rough Set merupakan perluasan dari teori set untuk sistem cerdas yang ditandai dengan informasi eksak, pasti, atau samar-samar (Mi, Wu and Zhang, 2004).

2. METODE PENELITIAN

Penambangan Data

Istilah penambangan data atau data mining adalah teknik yang sering digunakan dalam melakukan klasifikasi data dari jumlah data yang besar dengan menghasilkan informasi yang dapat digunakan untuk menganalisa dan mengetahui suatu informasi penting. Defenisi data mining dari beberapa penulis sebagai berikut :

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database (Jamaris, 2017).

Data Mining dapat diartikan sebagai pengekstrakan informasi bari yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan (Akseptor and Vasektomi, 2014).

Data Mining atau penambangan data adalah proses atau kegiatan mengekstraksi data untuk mencari pola, keteraturan dan

informasi yang tersembunyi dalam data yang besar atau basis data (Tendy, 2012).

Algoritma Rough Set dan Qualitative Measure

Rough Set dikembangkan oleh Zdzislaw Pawlak (Akseptor and Vasektomi, 2014). Rough Set merupakan perluasan dari teori set untuk studi sistem cerdas yang ditandai dengan informasi eksak, pasti, atau sama-samar (Mi, Wu and Zhang, 2004).

Dalam aplikasi Artificial Intellegenci (AI), Rough Set digunakan untuk menangani masalah Uncertainty Data Imprecision dan Vagueness). Pendekatan roughset menjadi pendekatan yang peting dalam AI dan ilmu kognitif, terutama pada area mechine learning, akuisisi pengetahuan, analisis keputusan, pencarian pengetahuan dari database, sistem pakar, penalaran induktif, dan pengenalan pola (Listiana, Anggraeni and Mukhlason, 2010). Salah satu kelebihan dari roughset adalah proses rough set dapat melakukan penanganan data yang inkonsistensi seperti yang ditampilkan pada tabel 2.1 (Gogoi, Bhattacharyya and Kalita, 2013).

Tabel 2.1 : Inkonsisten Data Set (Gogoi, Bhattacharyya and Kalita, 2013)

Object	Atribut		Atribut Keputusan
	A	C	
P1	Rendah	Tinggi	Yes
P2	Rendah	Low	No
P3	Rendah	Tinggi	No

Dalam roughset, kumpulan objek disebut sebagai information system (IS). Dari IS tersebut data dianalisa dalam area lower approximation, upper approximation, boundary region dan outside region (Listiana, Anggraeni and Mukhlason, 2010).

Information System

Dalam rough set, sebuah set dipresentasikan sebagai sebuah tabel,

dimana baris dalam tabel merepresentasikan objek dan kolom-kolom merepresentasikan atribut dari objek-objek tersebut. Tabel tersebut disebut dengan information system yang dapat digambarkan sebagai berikut(Listiana, Anggraeni and Mukhlason, 2010):

$$S = (U, A) \quad (1)$$

Tabel 2. 2 : Information System (Listiana, Anggraeni and Mukhlason, 2010)

	Headache	Muscle Pain	Temperature
P1	No	Yes	High
P2	Yes	No	High
P3	Yes	Yes	Very high
P4	No	Yes	Normal
P5	Yes	No	High
P6	No	Yes	Very high

Dalam penggunaan Information System, outcome dari klasifikasi adalah Decision System. Decision System dapat digambarkan dengan persamaan (2) (Listiana, Anggraeni and Mukhlason, 2010).

$$S = (U, A \cup \{d\}) \quad (2)$$

Dimana $d \notin A$ adalah atribut keputusan (decision attribute).

Tabel 2.3 Decision System

	Nilai UN	Nilai Tes	Pekerjaan Ortu	Aktif
P1	Cukup	Baik	PNS	Aktif
P2	Cukup	Cukup	Petani	Non Aktif
P3	Baik	Cukup	ABRI	Aktif
P4	Baik	Cukup	PNS	Non Aktif
P5	Baik	Cukup	PNS	Aktif
P6	Baik	Baik	PNS	Aktif

Indiscernibility Relation

Dalam decision system, sebuah objek dapat memiliki nilai yang sama untuk sebuah atribut kondisionalnya. Contohnya, P3, P4, P5, P6 memiliki atribut kondisional Nilai UN yang sama, yaitu

“Cukup”. Hubungan tersebut disebut dengan indiscernible (tidak dapat dipisah)(Listiana, Anggraeni and Mukhlason, 2010).

Berdasarkan tabel 2.3 maka didapat Indiscernibility Relation sebagai berikut :

$$IND(\text{Nilai UN}) = \{\{P1, P2\}, \{P3, P4, P5, P6\}\}$$

$$IND(\text{Nilai Tes}) = \{\{P1, P6\}, \{P2, P3, P4, P5\}\}$$

$$IND(\text{Pekerjaan ORTU}) = \{\{P1, P4, P5, P6\}, \{P2\}, \{P3\}\}$$

....

$$IND(\text{Nilai UN, Nilai Tes, Pekerjaan ORTU}) = \{\{P1\}, \{P2\}, \{P3\}, \{P4, P5\}, \{P6\}\}$$

Set Approximation

Untuk menentukan approximation (perkiraan) yang ada dalam information system, dimisalkan information system $S = (U, A)$, $B \subseteq A$, dan $X \subseteq U$. Dengan komponen tersebut, approximation dari X dapat dibentuk melalui informasi yang terdapat pada setatribut B dengan mengkonstruksi B -lower dan B -Upper approximation dari X , yang dinotasikan dengan $\underline{B}X$ dan $\overline{B}X$ dimana (Listiana, Anggraeni and Mukhlason, 2010),

$$\underline{B}X = \{x | [x]_B \subseteq X\} \quad (3)$$

$$\overline{B}X = \{x | [x]_B \cap X \neq \emptyset\} \quad (4)$$

Sebagai contoh, sesuai dengan tabel 3.2 maka set approximation-nya adalah sebagai berikut :

$$\underline{B}X = \{\{P1\}, \{P3\}, \{P6\}\}$$

$$\overline{B}X = \{\{P1\}, \{P2, P5\}, \{P3\}, \{P6\}\}$$

$$\overline{B}X - \underline{B}X = \{\{P2, P5\}\}$$

$$U - \overline{B}X = \{\{P4\}\}$$

Dependensi Atribut

Dalam setdata, hal terpenting untuk dicari adalah ketergantungan antar atribut. Secara intuitif, sebuah setatribut D tergantung secara total kepada setatribut C , dinotasikan dengan $C \Rightarrow D$, jika semua nilai dari atribut D secara unik ditentukan oleh nilai dari atribut C (Suraj, 2004). Keterbergantungan D terhadap C

dinotasikan dalam derajat k yang dapat diformulasikan seperti pada persamaan (5)(Listiana, Anggraeni and Mukhlason, 2010).

$$k = \gamma(C, D) = \frac{|POS_C(D)|}{|U|} \quad (5)$$

Jika $k=1$ maka D bergantung secara penuh kepada C , dan jika $k<1$, maka D bergantung sebagian pada C . Misal ingin dicari nilai k dari atribut {Temperature} \Rightarrow {Flu} maka nilai yang didapat adalah $k = 3/6 = 1/3$. Hal tersebut dikarenakan 3 dari 6 pasien yaitu P3, P4, dan P6 dapat diklasifikasikan secara unik sebagai pasien yang terserang flu (Listiana, Anggraeni and Mukhlason, 2010).

Reduksi Atribut

Dalam roughset, sebuah atribut dapat dihilangkan tanpa harus kehilangan nilai yang sebenarnya. Hal tersebut dikarenakan adanya atribut redundant yang tidak akan mempengaruhi hasil klasifikasi jika dihilangkan(Listiana, Anggraeni and Mukhlason, 2010).

Misalnya $S = (U, A)$, $B \subseteq A$, dan $\alpha \in B$ maka α adalah dispensable (tidak diperlukan) dalam atribut B jika $IND_S(B) = IND_S(B - \{\alpha\})$; dan sebaliknya jika indispensable maka α sangat diperlukan didalam B . Sebuah set B disebut independent jika semua atributnya sangat diperlukan. Tiap subset B' dari B disebut reduct dari B jika B' independen dan $IND_S(B') = IND_S(B)$. Atribut yang tidak termasuk dalam reduct adalah atribut yang tidak berguna untuk klasifikasi elemen dalam universe(Listiana, Anggraeni and Mukhlason, 2010).

Dari atribut yang ditampilkan pada tabel 3.2 diperoleh kombinasi atribut reduct sebagai berikut :

Muscle Pain & Temperature
Headache & Temperature
Headache, Muscle Pain & Temperature

Decision Rule

Decision rules adalah aturan yang terdiri dari if – then atau if f then g yang dapat dipresentasikan sebagai $f \rightarrow g$.

Bagian f pada rule $f \rightarrow g$ disebut antecedent dan bagian g disebut conclusion. Dalam rough set, decision rule dapat ditarik dari atribut reduct yang telah didapatkan(Listiana, Anggraeni and Mukhlason, 2010).

Sebagai contoh, berdasarkan atribut reduct yang telah didapatkan, maka rule yang dapat ditarik dari kombinasi atribut Muscle Pain & Temperature sebagai berikut :

if(muscle pain = yes) and (temperature = high) then (flu = yes)
if (muscle pain = no) and (temperature = high) then (flu = yes) or (flu = no)
if(muscle pain = yes) and (temperature = very high) then (flu = yes)
if(muscle pain = yes) and (temperature = normal) then (flu = no)

Qualitative Measure

Quantitative measure adalah ukuran yang dapat diekspresikan dalam jumlah atau ukuran yang terbatas. Quantitative measure terdiri dari beberapa pengukuran, yaitu: support, strength, accuracy, dan coverage(Listiana, Anggraeni and Mukhlason, 2010).

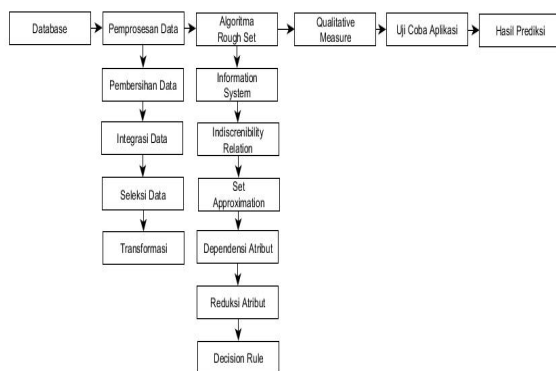
Untuk penentuan bobot (akurasi) pada rule yang dihasilkan, salah satunya dengan menggunakan strength yang dibedakan menjadi left Strength dan right Strength dikarenakan hasil dari rule tidak menutup kemungkinan menghasilkan lebih dari satu nilai. Left Strength dari decision rule menunjukkan frekuensi seringnya sebuah objek memenuhi antecedent rule. Right Strength dari decision rule menunjukkan frekuensi seringnya sebuah objek memenuhi antecedent dan conclusion. Right Strength merupakan implementasi dari strength(Listiana, Anggraeni and Mukhlason, 2010).

Misalkan $I=(U, A, D)$ adalah sistem keputusan, right Strength dalam $f \rightarrow g$ didefinisikan sebagai :

$$\text{RightStrength}(f \rightarrow g) = \frac{\text{RightSupport}(f \rightarrow g)}{\text{card}(U)} \quad (6)$$

3. PROSEDUR PENELITIAN

Studi Kasus penelitian pada penerimaan mahasiswa baru STIKOM UYELINDO KUPANG dari TA. 2016/2017 sampai TA. 2018/2019. Jumlah data yang digunakan adalah 579 record data. Pada metode penelitian untuk Prediksi Tingkat Presentasi Mahasiswa Baru Masuk sebagai Mahasiswa Aktif di STIKOM Uyelindo Kupang menggunakan Rough Set Theory. Skema penelitian digambarkan pada gambar 1 Tahap implementasi dan pengembangan sistem.



Gambar 1. Tahapan implementasi dan pengembangan sistem

Database penerimaan mahasiswa baru yang telah dikumpulkan dari TA. 2016/2017 Semester Ganjil sampai dengan TA.2018/2019 Semester Ganjil, diterapkan pemrosesan data terlebih dahulu untuk menghasilkan database yang siap diolah dan dianalisa menggunakan Algoritma Rough Set untuk selanjutnya dilakukan uji coba pada aplikasi dengan tujuan akhir mendapatkan hasil prediksi tingkat presentasi mahasiswa baru masuk sebagai mahasiswa aktif STIKOM UYELINDO Kupang.

Metodologi penelitian yang digunakan untuk melakukan penelitian ini adalah :

Metode Observasi

Metode observasi dilakukan untuk mengumpulkan data melalui pengamatan proses penerimaan mahasiswa baru dengan tujuan akhir mengumpulkan data penerimaan mahasiswa baru dari tahun

2016 semester ganjil sampai tahun 2018 semester ganjil.

Metode Penelitian Kepustakaan

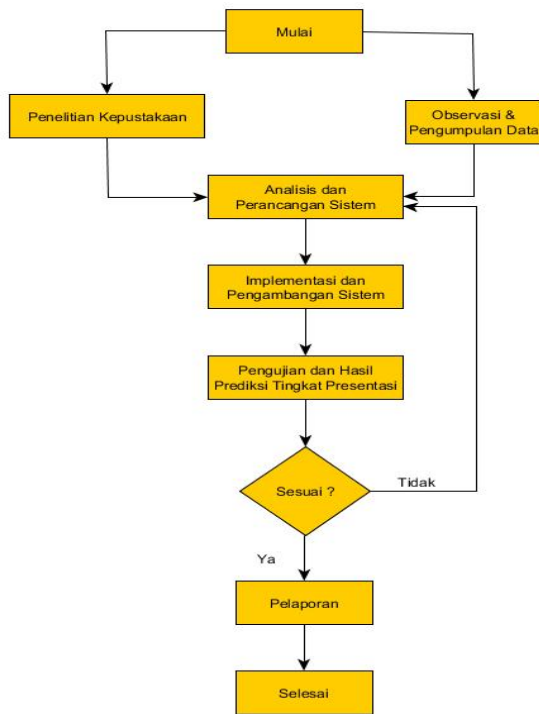
Metode ini dilakukan dengan cara mempelajari literature, buku atau jurnal yang relevan dengan obyek penelitian.

Metode Perancangan Sistem

Perumusan masalah dan penyelesaiannya Masalah yang dirumuskan dalam penelitian ini adalah bagaimana mengetahui dan mengenal pola klasifikasi status pendaftaran mahasiswa baru sehingga dapat dilakukan prediksi tingkat presentasi mahasiswa baru yang masuk di STIKOM UYELINDO KUPANG.

Implementasi dan pengembangan Sistem Tahap implementasi metode rough set dapat dilakukan dengan melakukan pengujian hasil pengolahan data yang telah dikerjakan menggunakan salah satu software aplikasi rough set yaitu Rosetta dan juga untuk aplikasinya dapat dikembangkan dengan pembangunan aplikasi menggunakan tool bahasa pemrograman lainnya. Pengujian dan Hasil Prediksi Tingkat Presentasi

Pada tahap pengujian dan perolehan hasil prediksi tingkat presentasi mahasiswa baru masuk sebagai mahasiswa aktif STIKOM UYELINDO Kupang, dilakukan dengan pengujian data kemudian melakukan perbandingan dari hasil pada data dan hasil pada sistem yang digunakan untuk pengujian atau sistem yang dibangun. Keseluruhan tahapan metodologi penelitian dapat dilihat pada gambar 2. Tahapan implementasi dan pengembangan sistem.



Gambar 2. Tahapan implementasi dan pengembangan sistem

HASIL ANALISA DAN PERANCANGAN

Berdasarkan analisa dan perancangan sistem yang dilakukan mengenai tahapan-tahapan pengolahan data dari tahap pembuatan basis data atau database dalam hal ini pengumpulan data, tahap kedua pemrosesan data yang telah didapat, dilanjutkan dengan penerapan algoritma Rough set untuk menghasilkan rule dan perhitungan akhir dengan Qualitative Measure, tahap selanjutnya adalah Uji Coba Aplikasi yang akan dibangun, dan akan diakhiri dengan mendapatkan hasil prediksi.

Penelitian analisa dan perancangan prediksi tingkat presentasi Mahasiswa Baru Masuk sebagai Mahasiswa Aktif STIKOM UYELINDO KUPANG menggunakan rough set, maka akan dibangun sebuah sistem pengenalan ekspresi wajah dengan menggunakan algoritma dan alur sistem yang telah dihasilkan. Sistem ini diharapkan mampu

4. SIMPULAN DAN SARAN

Simpulan

Berdasarkan hasil analisa dan perancangan sistem prediksi mahasiswa baru masuk sebagai mahasiswa aktif STIKOM UYELINDO KUPANG, dapat

menggunakan metode Rough Set untuk menghasilkan rule yang dapat dijadikan acuan untuk memprediksi mahasiswa aktif pada tahun penerimaan mahasiswa baru berikutnya.

Saran

Diharapkan bahwa sistem ini nantinya akan dikembangkan mejadi aplikasi yang dapat digunakan oleh perguruan tinggi STIKOM UYELINDO KUPANG, dan data rule dapat terus diupdate secara berkala menggunakan aplikasi yang dibangun untuk mendapat rule prediksi mahasiswa mahasiswa aktif secara berkelanjutan.

DAFTAR PUSTAKA

- [1] Laporan Bulanan HUMAS & Promosi Penerimaan Mahasiswa Baru TA. 2016/2017
- [2] Laporan Bulanan HUMAS & Promosi Penerimaan Mahasiswa Baru TA. 2017/2018
- [3] Laporan Bulanan HUMAS & Promosi Penerimaan Mahasiswa Baru TA. 2018/2019
- [4] Tendy, A., 2012. Pengenalan Pola Klasifikasi Stautus Registrasi Calon Mahasiswa Baru Univeristas Sanata Dharma denga Algoritma Reduct Based Decision Tree (RDT). Universitas Sanata Dharma.
- [5] Akseptor, M., and Vasektomi, K.B., 2014. Metode Rough Set Untuk Melihat Perilaku Suami Yang Menjadi Akseptor KB Vasektomi. Informasi dan Teknologi Ilmiah, III, pp.94–99.
- [6] Gogoi, P., Bhattacharyya, D.K., and Kalita, J.K., 2013. A rough set-based effective rule generation method for classification with an application in intrusion detection. International Journal of Security and Networks, 8(2), p.61.
- [7] Jamaris, M., 2017. Implementasi Metode Rough Set Untuk Menentukan Kelayakan Bantuan Dana Hibah Fasilitas Rumah Ibadah. 2(2).
- [8] Listiana, N., Anggraeni, W., and Mukhlason, A., 2010. Implementasi Algoritma Rough Set Untuk Deteksi dan Penangan Dini Penyakit Sapi.
- [9] Mi, J.-S., Wu, W.-Z., and Zhang, W.-X., 2004. Approaches to knowledge reduction based on variable precision rough set model. Information Sciences,

- 159(3/4), p.255.
- [10] Suraj, Z., 2004. An Introduction to Rough Set Theory and Its Applications. ICENCO, Cairo, Egypt.
- [11] Tendy, A., 2012. Pengenalan Pola Klasifikasi Stautus Registrasi Calon Mahasiswa Baru Univeristas Sanata Dharma denga Algortitma Reduct Based Decision Tree (RDT). Universitas Sanata Dharma.