

Implementasi Algoritma Rough Set untuk Memprediksi Jumlah Pendaftar Siswa Baru pada SMK Swasta Sinar Harapan

Sinta Novianti*¹, Paska Marto Hasugian²

^{1,2}STMIK Pelita Nusantara, Jl. Iskandar Muda No.1 Medan 20154, Indonesia

Email: *¹sintanovianti514@gmail.com, ²paskamarto86@gmail.com

ABSTRAK

Dimana jumlah pendaftar siswa baru di SMK Swasta Sinar Harapan setiap tahunnya tidak stabil dan pihak sekolah mengalami kesulitan untuk memprediksi berapa banyak jumlah pendaftar siswa baru di tahun mendatang serta mempersiapkan kegiatan ekstrakurikuler seperti pelatihan bahasa inggris/komputer ataupun fasilitas terbaik untuk siswa baru seperti kursi, meja, ruang kelas dan sebagainya, serta banyaknya daya saing sekolah swasta di kec. Beringin. Sehingga dibutuhkannya suatu sistem yang bisa menggali data serta membantu dalam memprediksi jumlah pendaftar siswa baru, dengan menggunakan teknik data mining yaitu Algoritma Rough Set yang bertujuan untuk mendapatkan perkiraan rule yang singkat dari sebuah tabel, dalam penelitian ini menggunakan data pendaftaran siswa tahun 2018, 2019, 2020. Dengan menerapkan data mining dengan algoritma rough set dan melakukan tahapan-tahapan raough set dapat menghasilkan keputusan berupa generate rule (knowledge), dan hasil dari analisa perhitungan ialah 152 siswa untuk prediksi jumlah siswa ditahun yang akan datang.

Kata Kunci: Data Mining, Prediksi, Rough Set

ABSTRACT

Where the number of new student registrants at the Sinar Harapan Private Vocational School is unstable every year and the school has difficulty predicting how many new student registrants will be in the coming year and preparing extracurricular activities such as English/computer training or the best facilities for new students such as chairs, tables, classrooms and so on, as well as the competitiveness of private schools in the district. banyan. So we need a system that can dig up data and help predict the number of new student registrants, using data mining techniques, namely the Rough Set Algorithm which aims to get a short rule estimate from a table, in this study using student registration data in 2018, 2019, 2020. By applying data mining with the rough set algorithm and carrying out the rough set stages, it can produce decisions in the form of generating rules (knowledge), and the results of the calculation analysis are 152 students to predict the number of students in the coming year.

Keywords: Data Mining, Prediction, Rough Set

1. PENDAHULUAN

Sekolah Menengah Kejuruan (SMK) Swasta Sinar Harapan yang berlokasi di Jl. Pantai Labu, Gg Bali, Desa Sidodadi R, Kec. Beringin. Dimana jumlah pendaftaran siswa baru di SMK Swasta Sinar Harapan setiap tahunnya tidak stabil. Pada tahun 2020 pendaftaran siswa baru mengalami peningkatan, namun fasilitas dan kegiatan ekstrakurikuler yang disediakan oleh pihak sekolah kurang mencukupi.

Sehingga siswa baru merasa tidak puas dengan sarana maupun kegiatan ekstrakurikuler yang terdapat di sekolah, serta banyaknya daya saing sekolah swasta di

kecamatan beringin, dengan hal tersebut pihak sekolah mengalami kesulitan memprediksi berapa banyak jumlah pendaftar siswa baru di tahun mendatang serta mempersiapkan kegiatan ekstrakurikuler seperti pelatihan bahasa Inggris/komputer ataupun fasilitas terbaik untuk siswa baru seperti kursi, meja, ruang kelas dan sebagainya. Sehingga dibutuhkan suatu sistem yang bisa menggali data serta membantu dalam memprediksi jumlah pendaftar siswa baru, dengan menggunakan teknik data mining.

Data mining merupakan suatu proses untuk menggali nilai tambah berbentuk informasi yang belum diketahui secara manual serta suatu basis data dengan melakukan penggalian pola-pola dari data dengan tujuan untuk memanipulasi data menjadi informasi lebih berarti di peroleh dengan cara mengekstraksi dan mengenali pola yang berarti ataupun menarik dari data yang ada dalam database [1]. Data mining mempunyai sebagian metode yang dapat digunakan yaitu: estimasi, asosiasi, klasifikasi, klastering dan Prediksi. Salah satu metode prediksi dalam data mining adalah Algoritma *Rough Set*.

Algoritma *Rough Set* merupakan suatu metode matematik yang dikembangkan oleh Pawlack pada tahun (1980). Metode *Rough Set* merupakan metode matematika untuk menangani ketidakjelasan dan ketidakpastian yang di perkenalkan untuk memproses ketidakpastian serta data yang tidak akurat. Teknik ini ialah suatu teknik yang paling efisien untuk proses *knowledge Discovery in Database* (KDD) dan data mining.

Tujuan dari analisis *Rough Set* untuk mendapatkan perkiraan rule yang singkat dari sebuah tabel. Hasil analisis *Rough Set* dapat digunakan dalam proses data mining dan *knowledge discovery* [2].

Berdasarkan penelitian yang dilakukan oleh [1] yang berjudul “Prediksi Jumlah Pendaftar Umroh dan Haji Plus dengan Algoritma *Rough Set* (Studi Kasus PT. Annajwa Islamic Tour & Travel)” dalam penelitian ini sistem prediksi untuk jumlah Jemaah pendaftar umroh dan haji plus di PT. Annajwa Islamic Tour & Travel, ialah sistem informasi yang dirancang untuk prediksi jumlah Jemaah pendaftar umroh dan haji plus menggunakan algoritma *Rough Set* di PT. Annajwa Islamic Tour & Travel untuk menentukan jumlah Jemaah pendaftar umroh serta haji plus ditahun mendatang.

2. TINJAUAN PUSTAKA

2.1 Data Mining

Menurut Arhani dan Nasir (2020:1) “Data mining adalah proses penggalihan informasi yang bermanfaat dari data yang begitu besar. Data mining mencakup pengumpulan data, ekstraksi data, analisis data, dan statistik data.

Menurut Nasution (2014) dalam Indriani (2018) ‘Data mining merupakan proses yang memakai teknik statistik, matematika, kecerdasan buatan untuk mengekstrak serta mengidentifikasi informasi yang bermanfaat serta pengetahuan terkait dari berbagai database besar’ [3].

Menurut Larose (2005) dalam Nofriansyah & Nurcahyo (2019:1) Data Mining adalah analisis dan peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data [4].

Menurut Hermawati (2018:3) Data Mining yaitu beberapa proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstrak pengetahuan (*knowledge*) secara otomatis.

2.2 Algoritma Rough Set

Algoritma *Rough Set* merupakan salah satu teknik pengetahuan data mining untuk memperoleh informasi yang berupa *decision system* (tabel yang dilengkapi dengan atribut kriteria dan keputusan). Keuntungan utama menggunakan *rough set* daripada himpunan fuzzy dalam analisis data tidak membutuhkan data awal atau tambahan tentang data, seperti

probabilitas dalam statistik, tingkatan kenggotaan ataupun nilai probabilitas dalam teori himpunan fuzzy [5].

Berikut ini merupakan skema penyelesaian menggunakan algoritma *Rough Set*, ialah:

1. *Information System*

$$IS = \{U, A\}$$

Information system merupakan Tabel yang terdiri dari baris yang merepresentasikan data serta kolom yang merepresentasikan atribut ataupun variabel dari data. *Information system* dalam data mining diketahui sebagai kumpulan data set. *Information system* dapat dipresentasikan sebagai fungsi:

Keterangan:

$U = \{x_1, x_2, \dots, x_m\}$ merupakan sekumpulan example

$A = \{a_1, a_2, \dots, a_n\}$ merupakan sekumpulan atribut kondisi secara berurutan

2. *Decision System*

$$S = (U, \{A, C\})$$

Decision System merupakan *information system* dengan atribut tambahan yang dinamakan dengan *decision attribute*, dalam data mining dikenal dengan *class* ataupun target. Atribut ini merepresentasikan hasil klasifikasi yang diketahui. *Decision System* merupakan suatu fungsi yang menggambarkan *information system*, sehingga *information system* (IS) menjadi.

Keterangan:

$U = \{x_1, x_2, \dots, x_m\}$ adalah sekumpulan example

$A = \{a_1, a_2, \dots, a_n\}$ adalah sekumpulan atribut kondisi secara berurutan.

$C =$ decision attributes (Keputusan)

3. *Equivalence Class*

Equivalence Class adalah mengelompokkan objek yang sama untuk atribut A (U, A).

4. *Discernibility Matrix*

Definisi *Discernibility Matrix* diberikan IS $A = (U, A)$ and B gabung A, *Discernibility Matrix* dari A adalah MB, di mana setiap entri MB(I,j) terdiri dari atribut yang berbeda antara objek X_i dan X_j .

5. *Discernibility Matrix Modulo D*

Discernibility Matrix Modulo D, didefinisikan sebagai berikut dimana Modulo (i,j) adalah sekumpulan atribut yang berbeda antara objek X_i dan X_j dan juga atribut keputusan yang berbeda.

6. *Reduction*

Untuk data dengan jumlah variabel yang sangat banyak, tidak mungkin menemukan semua kombinasi variabel yang ada, karena jumlah indiscernibility yang dicari = $(2^n - 1)$. Oleh karena itu, dibuatlah teknik pencarian kombinasi atribut yang mungkin dikenal dengan istilah *Quick Reduct*, yaitu dengan cara:

1. Nilai *indiscernibility* pertama yang dicari merupakan *indiscernibility* untuk kombinasi atribut terkecil yaitu 1.
2. Setelah lakukan proses pencarian *dependency attributes*. Bila nilai *dependency attributes* yang didapat = 1 maka *indiscernibility* untuk himpunan minimal variable adalah variabel tersebut.
3. Bila pada proses pencarian campuran atribut tidak menemukan *dependency attributes* =1, maka cari kombinasi yang lebih besar, dimana campuran *variable* yang dicari merupakan campuran variabel pada tahap sebelumnya dengan nilai *dependency attributes* terbesar. Lakukan proses (3), sampai didapat nilai *dependency attributes* =1.

7. *Generate Rules*

Generating Rules adalah proses utama menemukan pengetahuan dalam database adalah ekstraksi aturan dari sistem pengambilan keputusan. Metode set kasar dalam

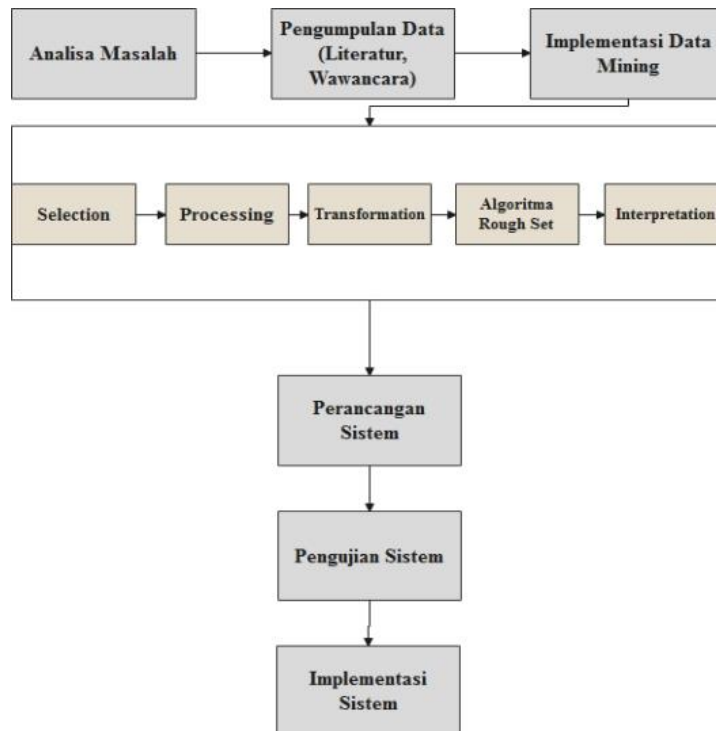
menghasilkan aturan keputusan dari tabel keputusan didasarkan pada perhitungan yang direduksi.

2.3 Prediksi

Memprediksi dapat didasarkan pada metode ilmiah atau subjektif belaka. Prediksi dalam metode adalah proses ilmiah untuk mendapatkan pengetahuan secara sistematis bersumber pada bukti fisik [6].

3. METODE PENELITIAN

Untuk membantu dalam penyusunan penelitian ini, maka dibutuhkan kerangka kerja. Kerangka kerja penelitian merupakan tahapan yang dapat diambil dalam proses pemecahan masalah yang hendak dibahas. Berikut ini adalah tahapan kerangka kerja penelitian:



Gambar 1 Kerangka Kerja Penelitian

4. HASIL DAN PEMBAHASAN

4.1 Analisis Algoritma Rough Set

Berikut langkah-langkah dari algoritma rough set:

1. Information System

Information system merupakan tabel yang terdiri dari baris yang merepresentasikan data serta kolom yang mempresentasikan atribut ataupun variable dari data jurusan. *Information system* dapat direpresentasikan sebagai $A = (U, A)$, dimana $U = \{x_1, x_2, \dots, x_m\}$ merupakan sekumpulan atribut *example*, $A = \{a_1, a_2, \dots, a_n\}$ ialah sekumpulan atribut kondisi secara berurutan.

Tabel 1 *Information System*

No	Jurusan	Jumlah 2018	Jumlah 2019	Jumlah 2020
1	BM	33	43	29
2	TR	65	50	85
3	TKJ	35	66	50

Keterangan :

$U = \{e_1, e_2, \dots, e_m\}$: Jurusan BM, TR, TKJ

$A = \{A_1, A_2, \dots, A_N\}$: jumlah 2018, jumlah 2019, jumlah 2020

2. Decision System

Decision system ialah fungsi yang mendeskripsikan *information system*, selanjutnya *information system* (IS) jadi *Decision System* (DS), yang bisa ditafsirkan sebagai $DS = (U, \{A, C\})$ dimana $U = \{x_1, x_2, \dots, x_m\}$ ialah sekumpulan atribut *example*, $A = \{a_1, a_2, \dots, a_n\}$ ialah sekumpulan atribut kondisi secara berurutan serta C merupakan *Decision Attribute* keputusan. Berikut tabel dari *decision attribute* keputusan, dengan data bersumber pada jurusan yang hendak digunakan ialah sebanyak 3 .

Tabel 2 *Decision System*

No	Jurusan	Jumlah 2018	Jumlah 2019	Jumlah 2020	Keputusan
1	BM	33	43	29	Turun
2	TR	65	50	85	Naik
3	TKJ	35	66	50	Stabil

Tabel diatas memperlihatkan suatu *Decision Systems* yang sederhana, terdapat m objek (BM, TR, TKJ) serta n attribute (jumlah 2018, jumlah 2019, jumlah 2020 dan Keputusan). Dalam tabel ini, n-1 jumlah 2018, jumlah 2019, jumlah 2020, merupakan attribute kondisi, sedangkan keputusannya ialah *decision attribute*.

3. Equivalence Class

Equivalence class merupakan penggolompokkan objek yang sama untuk atribut A (U, A). Diberikan *decision system* pada tabel 2 dapat diperoleh dengan *equivalence class* (EC1-EC3).

Tabel 3 *Equivalence Class*

No	Jurusan	Jumlah 2018	Jumlah 2019	Jumlah 2020	Keputusan
1	BM	33	43	29	Turun
2	TR	65	50	85	Naik
3	TKJ	35	66	50	Stabil

Contoh pada tabel 5 dibawah ini dapat disederhanakan kedalam *numerical representasion* untuk memudahkan pengolahan data, berikut ini adalah tranformasi atribut dapat dilihat di tabel dibawah ini:

Tabel 4 Transformasi Atribut

Atribut	Range	Nilai
Jumlah 2018 (A)	$0 \leq X \leq 20$	1
	$21 \leq X \leq 40$	2
	$41 \leq X \leq 60$	3
	$61 \leq X \leq 80$	4
Jumlah 2019 (B)	$0 \leq X \leq 25$	1
	$26 \leq X \leq 45$	2
	$45 \leq X \leq 65$	3
	$66 \leq X \leq 85$	4
Jumlah 2020 (C)	$0 \leq X \leq 30$	1
	$31 \leq X \leq 50$	2
	$51 \leq X \leq 70$	3
	$71 \leq X \leq 90$	4
Keputusan (K)	Turun	1
	Stabil	2

	Naik	3
--	------	---

Pada Tabel 5 merupakan *numerical representation* dari *equivalence class* dari tabel 3.

Tabel 5 *Equivalence Class* (Transformasi)

Class	Jumlah 2018 (A)	Jumlah 2019 (B)	Jumlah 2020 (C)	Keputusan (K)
EC1	2	2	1	1
EC2	4	3	4	3
EC3	2	4	2	2

4. *Discernibility Matrix*

Tabel *Equivalence Class* merupakan tabel rujukan untuk menciptakan *Discernibility Matrix*. Dalam mencari *Discernibility Matrix*, diberikan sebuah IS $A=(U,A)$ serta B gabung A, *discernibility matrix* dari A merupakan MB, dimana tiap-tiap entry MB (I,j) terdiri dari sekumpulan atribut yang berbeda antara xi dan xj. Setelah itu lakukan perbandingan antar objek ataupun baris berdasarkan atribut kondisi. Bila terdapat kesamaan antara atribut kondisi yang dimiliki objek, maka atribut kondisi ditulis dengan tanda kali (X), sebaliknya bila tidak terdapat kesamaan antara atribut kondisi yang dimiliki objek, maka atribut kondisi tersebut diambil/ditulis.

Atribut dimodelkan dengan:

1. Jumlah 2018 : A
2. Jumlah 2019 : B
3. Jumlah 2020 : C

Contoh pada Tabel 6 EC1 dengan EC1, seluruh atribut sama sehingga hasilnya X (baris 2 kolom 2), kemudian EC1 dengan EC2, ada perbedaan ialah atribut jumlah 2018, jumlah 2019, jumlah 2020, sehingga pada tabel 6 (baris 2 kolom 3) hasilnya adalah ABC, jadi selanjutnya. Tabel 6 menampilkan *discernibility matrix* dari tabel 5.

Tabel 6 Hasil *Discernibility Matrix*

Class	EC1	EC2	EC3
EC1	X	ABC	BC
EC2	ABC	X	ABC
EC3	BC	ABC	X

5. *Discernibility Matrix Module D*

Dapat didefinisikan di mana MB (I,j) merupakan sekumpulan atribut yang berbeda antara objek Xi dan Xj, juga keputusan atribut yang berbeda. Berdasarkan tabel 4.13, selanjutnya bandingkan setiap *class* berdasarkan keputusan, Bila atribut keputusan (K) mempunyai kesamaan, hingga atribut kondisi yang ditulis dengan tanda kali (X), sebaliknya bila atribut keputusan (K) berbeda, tuliskan perbedaan atributnya berdasarkan tabel 4.13.

Contoh : EC2 dengan EC1 Keputusan (K) berbeda, sehingga hasilnya adalah : ABC (baris 3 kolom 2).

Tabel 7 Hasil *Discernibility Matrix Module D*

Class	EC1	EC2	EC3
EC1	X	ABC	BC
EC2	ABC	X	ABC
EC3	BC	ABC	X

6. Proses *Reduction*

Proses *reduct* merupakan proses pemilihan atribut minimal (*Interesting attribut*) dari sekumpulan kondisi dengan menggunakan *prime implicant* fungsi *Boolean*. Pengumpulan

semua *prime Implicant* mendeterminasikan *sets of reduct*. *Discernibility matrix modulo D*. Langkah-langkah dalam pencarian *reduct* ialah:

1. Membuat persamaan Aljabar *Boolean* berdasarkan pada Tabel 4.14 hasil *Discernibility Matrix Module D*, misalnya $EC1 = CD.C$ persamaan Aljabar *Boolean* ialah $(C \vee D) \wedge (C)$.
2. Setelah itu hasil penyederhanaan dengan prinsip ataupun konsep persamaan Aljabar *Boolean*, dengan menggubah tanda \wedge dengan tanda kali (*) dan tanda \vee dengan menggubah tanda tambah (+). Jadi hasilnya yaitu $(C + D) * (C)$. setelah itu lakukan perkalian distributif, hingga $CC + CD = C + CD = C(1+D)=C$
3. Setelah itu hasil penyederhanaan tersebut digunakan sebagai *reduct*. Bila hasilnya memakai tanda \wedge , maka *reduct* nya satu kurung kurawal, misal $(A \wedge C)$ *reduct* nya $\{A, D\}$, bila tanda \vee , maka *reduct* nya terpisah. Misal $A \vee C$, maka *reduct* nya adalah $\{A\}, \{C\}$. Tetapi, sebab jawaban permasalahan di atas ialah C, maka *reduct* nya adalah $\{C\}$.
4. Jika nilai 1 + atribut itu sama dengan 0, maka dapat dihapus.

Proses pencarian *Prime Implicant* dengan fungsi *Boolean* terhadap hasil penyeleksi sebelumnya, yaitu:

1. $A+1 = 1+1=1$
2. $AA=A$

Class EC1 terdiri dari X,ABC,BC menjadi $(A \vee B \vee C) \wedge (B \vee C)$

$$\begin{aligned}
 &= (A+B+C)*(B+C) \\
 &= (AB+AC+BB+BC+CB+CC) \\
 &= (AB+AC+B(1+C)+C(1+B)) \\
 &= (AB+AC+B+C) \\
 &= (AB+B+AC+C) \\
 &= (B(1+A)+C(1+C)) \\
 &= (B+C) = \{B\}, \{C\}
 \end{aligned}$$

Tabel 8 Hasil *Reduction*

Class	CNF of <i>Bolean Function</i>	<i>Prime Implicant</i>	<i>Reduct</i>
EC1	$(A \vee B \vee C) \wedge (B \vee C)$	$(B+C)$	$\{B\}, \{C\}$
EC2	$(A \vee B \vee C) \wedge (A \vee B \vee C)$	$(A+B+C)$	$\{A\}, \{B\}, \{C\}$
EC3	$(B \vee C) \wedge (A \vee B \vee C)$	$(B+C)$	$\{B\}, \{C\}$

7. Generate Rule (*Knowledge*)

Ada pula *rule* yang dihasilkan berdasarkan *Reduct* terdiri dari kombinasi atribut yaitu:

1. $\{B\} = \text{Jumlah 2019}, \{C\} = \text{jumlah 2020}$
2. $\{A\} = \text{Jumlah 2018}, \{B\} = \text{Jumlah 2019}, \{C\} = \text{jumlah 2020}$
3. $\{B\} = \text{Jumlah 2019}, \{C\} = \text{Jumlah 2020}$

Tahap selanjutnya adalah mengetahui cara membuat *generate rules* berdasarkan tabel *reduct* yang dihasilkan. Membuat *Generate rules* hendak memakai kata *And* bila memiliki lebih dari satu atribut kondisi, namun bila atribut kondisi sama tetapi atribut keputusan lebih dari satu maka hendak digunakan kata *or*. Sehingga *rule* yang dihasilkan ialah berdasarkan pada *equivalence class* dengan membandingkannya dengan kombinasi atribut yang ada, sehingga didapatkan hasil yaitu:

1. $EC1 \text{ Reduct} = \text{If } B=2 \text{ Then } K=1, \text{ If } C=1 \text{ Then } K=1$
If Tahun 2019 = 43 Then Keputusan Turun
If Tahun 2020 = 29 Then Keputusan Turun
2. $EC2 \text{ Reduct} = \text{If } A=4 \text{ Then } K=3, \text{ If } B=3 \text{ Then } K=3, \text{ If } C=4 \text{ Then } K=3$
If Tahun 2018 = 65 Then Keputusan Naik
If Tahun 2019 = 50 Then Keputusan Naik
If Tahun 2020 = 85 Then Keputusan Naik
3. $EC3 \text{ Reduct} = \text{If } B=4 \text{ Then } K=2, \text{ If } C=2 \text{ Then } K=2$

If Tahun 2019 = 66 Then Keputusan Stabil

If Tahun 2020 = 50 Then Keputusan Stabil

Sehingga berdasarkan hasil *rules* diatas dapat menghasilkan informasi berupa *knowledge*, dapat dijadikan acuan dalam menentukan keputusan, dan dapat mempermudah kepala sekolah untuk memprediksi jumlah pendaftar siswa tahun yang berikutnya.

4.2 Analisa Perhitungan Prediksi Jumlah Pendaftar Siswa Baru

Pada tahap analisis ini, untuk menghitung jumlah pendaftar siswa baru pada tahun selanjutnya hendak menggunakan data 3 tahun terakhir dengan tujuan untuk mendapatkan hasil perhitungan yang akurat. Dengan menggunakan rumus rata-rata: $Y1 + Y2 + Y3 \dots Yn = i/n$

Dimana:

Y adalah tahun, n adalah jumlah total tahun yang digunakan

Contoh perhitungan rata-rata:

Tabel 9 Jumlah Data Siswa

2018	2019	2020
133	159	164

Dari data sebelumnya dilakukan dengan perhitungan *rough set* yang akan menghasilkan prediksi dalam bentuk *knowledge*. Karena hasil akhir berupa *knowledge*, sehingga user dapat memahami hasil prediksi, sistem ditransfer dalam bentuk angka dengan perhitungan nilai rata-rata 3 tahun yang diambil dari jumlah pendaftaran setiap tahunnya. Sehingga output yang dihasilkan ialah prediksi untuk tahun yang akan datang. Dengan hasil: Rata-rata jumlah pendaftar = $(133+159+164) / 3 = 152$

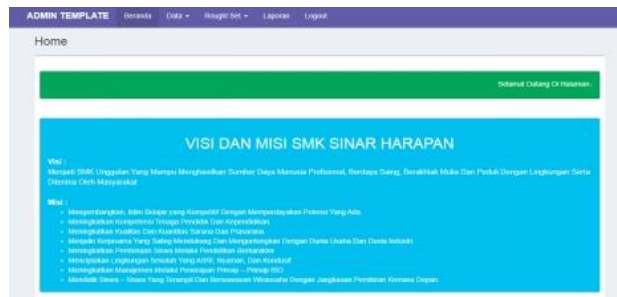
4.3. Implementasi Sistem

Menu *Login* merupakan tampilan menu utama dari program dimana *user* akan mengetik *username* dan *password* untuk *login* ke menu berikutnya.



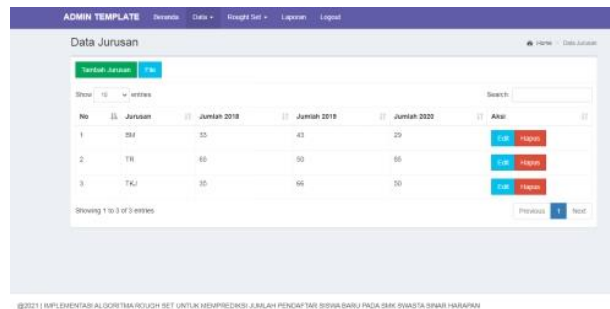
Gambar 2 Tampilan *Login*

Menu utama merupakan tampilan berikutnya setelah *user* melakukan *login*, pada menu utama terdapat beberapa sub menu, berikut ini adalah tampilan dari menu utama:



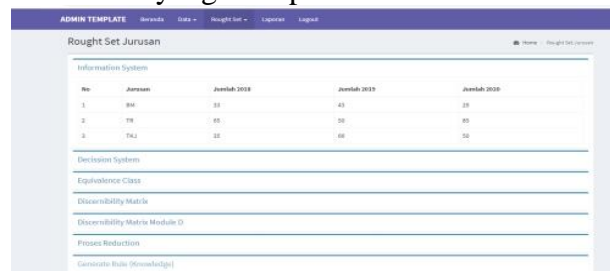
Gambar 3 Tampilan Menu Utama

Dimana pada menu ini terdapat data siswa dari masing-masing jurusan yang akan di proses *rough set*. Pada bagian menu data jurusan, *user* bisa menambahkan data, mengubah dan menghapus.



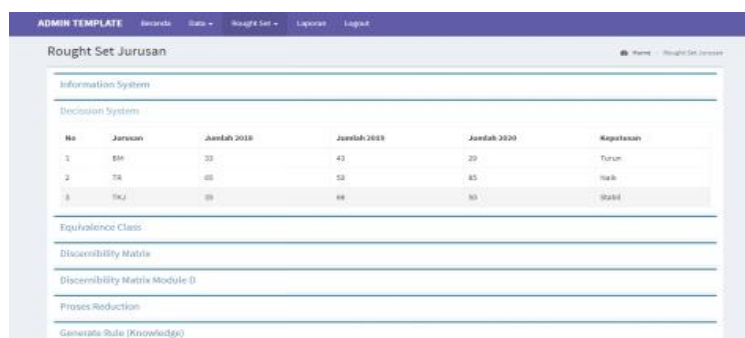
Gambar 4 Tampilan Data Jurusan

Pada menu ini terdapat proses *rough set* untuk jurusan, berikut adalah penjelasan dari menu *rough set* untuk jurusan:
 Pada gambar di bawah ini terdapat *form information system* jurusan terdiri dari baris dan kolom dapat memperlihatkan data yang merepresentasikan atribut atau variabel dari data.



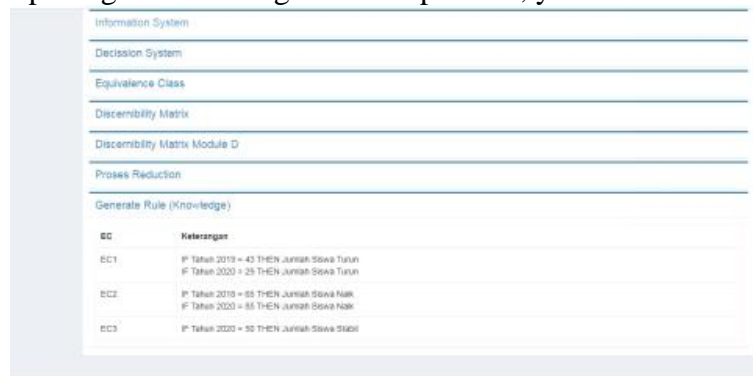
Gambar 5 Tampilan *Information System*

Selanjutnya terdapat *form decision system* pada *rough set* jurusan dengan atribut tambahan yaitu keputusan, berikut tampilan pada *decision system*.



Gambar 6 Tampilan *Decision System*

Selanjutnya proses yang terakhir adalah *generate rule* yaitu proses yang menghasilkan *knowledge* yang dapat digunakan sebagai hasil keputusan, yaitu:



Gambar 11 Tampilan *Generate Rule*

Pada menu laporan terdapat hasil cetak laporan prediksi jumlah siswa, berikut tampilan *form laporan*:

Laporan Prediksi Jumlah Pendaftar Siswa SMK Swasta Harapan Bangsa			
2018	2019	2020	Prediksi
133	159	164	152

Diketahui, 05-06-2021
(Admin)

Gambar 12 Tampilan Cetak Laporan

5. KESIMPULAN

1. Dalam memprediksi jumlah pendaftar siswa baru dengan menggunakan data – data pada masa lampau, ialah data siswa tahun 2018, tahun 2019 dan tahun 2020.
2. Dengan menerapkan data mining dengan algoritma rough set dan melakukan tahapan-tahapan raough set dapat menghasilkan keputusan berupa *generate rule (knowledge)*, dan hasil dari analisa perhitungan ialah 152 siswa untuk prediksi jumlah siswa ditahun yang akan sistem.
3. Dalam membangun sistem algoritma rough set untuk memprediksi jumlah pendaftar siswa baru dilakukan dengan menggunakan aplikasi pemrograman berbasis web dan menggunakan *database Mysql*, maka aplikasi dapat di implementasikan sehingga memudahkan pihak sekolah dalam memprediksi jumlah siswa baru.

DAFTAR PUSTAKA

- [1] A. Z. Hasibuan, G. Ginting, and K. Tampubolon, “Prediksi Jumlah Jamaah Pendaftar Umroh dan Haji Plus dengan Algoritma Rough Set (Studi Kasus: PT Annajwa Islamic Tour & Travel),” *Inf. dan Teknol. Ilm.*, vol. 13, pp. 187–191, 2018.
- [2] A. Putra, Z. A. Matondang, N. Sitompul, I. Pendahuluan, and A. Prediksi, “Implementasi Algoritma Rough Set Dalam Memprediksi Kecerdasan Anak,” *J. Pelita Inform.*, vol. 7, no. 2, pp. 149–156, 2018.
- [3] U. Indriani, “Penerapan Metode Rough Set Dalam Menentukan,” vol. 2, no. 1, pp. 85–

- 92, 2018.
- [4] D. Nofriansyah and G. W. Nurcahyo, “Algoritma Data Mining Dan Pengujian,” *Algoritma Data Mining dan Pengujian*. 2019.
 - [5] S. Sulaiman, N. A. A. Rahim, and A. Pranolo, “Generated rules for AIDS and e-learning classifier using rough set approach,” *Int. J. Adv. Intell. Informatics*, vol. 2, no. 2, pp. 103–122, 2016, doi: 10.26555/ijain.v2i2.74.
 - [6] S. Kasus, P. Sumut, R. Tasya, E. Buulolo, and P. G. M, “Menggunakan Metode Rough Set,” vol. 13, pp. 157–161, 2018.
 - [7] Arhami, M., & Nasir, M. (2020). *Data Mining Algoritma dan Implementasi*. Yogyakarta: Andi
 - [8] Hermawati, F. J. (2018). *Data Mining*. Yogyakarta: C.V Andi Offset