

PENERAPAN METODE *ITERATIVE DICHOTOMIZER 3* (ID 3) UNTUK MENENTUKAN BEASISWA BERPRESTASI PADA SMP PGRI CARINGIN SUKABUMI

Rizal Amegia Saputra¹; Lis Saumi Ramdhani²; Supriatman³;

¹Program Studi Sistem Informasi Akuntansi Kampus Kota Sukabumi
Univeritas Bina Sarana Informatika
<http://bsi.ac.id/>
rizal.rga@bsi.ac.id

^{2,3}Program Studi Sistem Informasi Kampus Kota Sukabumi
Universitas Bina Sarana Informatika
<http://bsi.ac.id/>

¹lis.lud@bsi.ac.id, ²supriatmanmamang91@gmail.com



Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi-NonKomersial 4.0 Internasional.

Abstract— Scholarships are assistance from the government to students / students who are less able or have the ability in the academic and non-academic fields that are given individually to reduce the burden in terms of material. Frequently stalling time in selection, the number of students who apply for scholarships, the number of students whose homes are far from school, the number of students who race to come early as one of the criteria eligible to receive scholarships as well as the most scholarship applicants feel disadvantaged by unfavorable decisions. Iterative Dichotomizer 3 (ID3) algorithm is the most basic decision tree learning algorithm (decision tree learning algorithm). This algorithm conducts a thorough search on all possible decisions. In this research, it will be analyzed the application of the iterative dichotomizer 3 method in the case of determining achievement scholarships. In order to make decisions quickly and accurately. From 708 scholarship candidates including 28 eligible and 680 scholarship recipients, 136 scholarship recipients were obtained from ID3 algorithm with 3 eligible and 133 who had not, and obtained an accuracy rate of 97.75% so that it could be concluded that good and can help the school.

Keywords: *Analysis, Algoritma ID3, Achievement Scholarships*

Intisari— Beasiswa merupakan bantuan dari pemerintah kepada siswa/ mahasiswa yang kurang mampu atau memiliki kemampuan dalam bidang akademik dan non akademik yang diberikan secara perorangan agar mengurangi beban dalam segi

materi. Sering mengulurnya waktu dalam penyeleksian, banyaknya siswa yang mendaftar untuk beasiswa, banyaknya siswa yang jarak rumahnya jauh dari sekolah, banyaknya siswa yang berlomba untuk datang pagi sebagai salah satu kriteria yang berhak menerima beasiswa serta yang paling patal pendaftar beasiswa merasa dirugikan oleh keputusan yang kurang baik. Algoritma *Iterative Dichotomizer 3* (ID3) merupakan algoritma *decision tree learning* (algoritma pembelajaran pohon keputusan) yang paling dasar. Algoritma ini melakukan pencarian secara menyeluruh pada semua kemungkinan semua keputusan. Dalam penelitian ini akan dilakukan analisa penerapan metode *iterative dhicotomizer 3* pada kasus penentuan beasiswa prestasi. Agar dapat membuat keputusan dengan cepat dan akurat. Dari 708 calon penerima beasiswa teridiri dari 28 yang layak dan 680 yang belum, maka didapatkan 136 calon penerima beasiswa yang dihasilkan dari pohon keputusan algortima ID3 dengan 3 yang layak dan 133 yang belum, dan diperoleh tingkat akurasi sebesar 97,75% sehingga dapat disimpulkan bahwa baik dan dapat membantu pihak sekolah.

Kata Kunci: *Analisa, Algoritma ID3, Beasiswa Prestasi*

PENDAHULUAN

Pendidikan menjadi kewajiban yang sangat penting bagi setiap warga negara, hal ini tertuang dalam UUD 1945 Pasal 31 ayat 1 “Bahwa setiap warga negara berhak mendapatkan pendidikan”, maka sebab itu pemerintah memberikan solusi

kepada siswa yang kurang mampu atau yang memiliki prestasi dalam bidang akademik maupun non akademik dengan adanya beasiswa Eviciensna, 2016.

Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan kepada perorangan yang memiliki prestasi yang baik tetapi mengalami kesulitan ekonomi bertujuan untuk demi keberlangsungan pendidikan yang ditempuh (Kamal 2017).

Prestasi peserta didik merupakan suatu hal yang penting bagi semua pihak yang terkait secara langsung maupun tidak langsung (Narti 2017).

SMP PGRI Caringin merupakan sekolah swasta yang setiap tahun banyak menerima siswa/siswi baru sehingga dalam menentukan beasiswa banyak sekali menemukan permasalahan yang terjadi diantaranya: Sering mengulur waktu dalam penyeleksian dikarenakan yang menyeleksi terdiri dari guru-guru yang mengajar kelas sehingga tidak memiliki banyak waktu untuk melakukan penyeleksian (Radhitya, Nur Hakim, and Solechan 2016), banyaknya siswa yang mendaftar untuk beasiswa, banyaknya siswa yang jarak rumahnya jauh dari sekolah, banyaknya siswa yang berlomba untuk datang pagi sebagai salah satu kriteria yang berhak menerima beasiswa serta siswa yang mendaftar beasiswa merasa dirugikan oleh keputusan yang kurang baik karena penyaluran beasiswa yang tidak tepat sasaran (Kirom, Bilfaqih, and Effendie 2012).

Algoritma *Iterative Dichotomizer 3* (ID3) merupakan algoritma *decision tree learning* (algoritma pembelajaran pohon keputusan) yang paling dasar (Ispriyanti and Hoyyi 2016).

Dengan *Algoritms* ID3 ini dapat memberikan data yang lengkap sehingga akan lebih dipermudah bagi pihak sekolah untuk menentukan calon penerima beasiswa (Pramadhani and Setiadi 2014).

BAHAN DAN METODE

A. Jenis Penelitian

Didalam penelitian ini digunakan penelitian dengan metode eksperimen dengan menggunakan data siswa calon penerima beasiswa di "SMP PGRI Caringin Sukabumi" sejumlah 708 orang, yang terdiri dari 28 orang siswa yang layak dan 680 orang siswa yang belum layak.

B. Instrument Penelitian

1. Penelitian ini menggunakan data sekunder berupa data siswa yang berhak mencalonkan sebagai penerima beasiswa yang akan digunakan sebagai instrumen guna memperoleh data dalam proses penyeleksian beasiswa.

2. Data disajikan dalam bentuk tabel excel dan variabel masing-masing sebanyak 220 siswa untuk tahun pelajaran 2017/2018, sebanyak 247 siswa untuk tahun pelajaran 2016/2017 dan sebanyak 241 siswa untuk tahun pelajaran 2015/2016, dengan 12 variabel.
3. Perangkat lunak yang digunakan untuk menganalisis adalah *RapidMiner* dan *Graphical User Interface* (GUI) *Visual Basic 6* untuk menguji *rule* algoritma terpilih.

C. Metode Analisa dan Pengujian Data

Teknik analisa data menggunakan data kualitatif berupa matematis terhadap angka atau numerik dan nominal. Pada penelitian ini, analisa data dilakukan melalui data siswa SMP dengan nilai rata-rata siswa yang layak dan belum layak seleksi, data diolah dan di uji dalam pengujian pada algoritma ID3. Kemudian pengujian *Rule* yang diperoleh algoritma ID3 tersebut selanjutnya diuji dengan *Confusion Matrix* dan *Kurva Receiver Operating Characteristic (ROC)* untuk mengukur tingkat akurasi yang akan dihasilkan.

Dengan pengujian diatas dapat diperoleh nilai akurasi dari *rule* algoritma ID3. Sehingga dapat diterapkan pada *Graphical User Interface* (GUI) dengan baik. Selanjutnya aplikasi yang dibuat akan dievaluasi untuk menghasilkan pengetahuan (*knowledge*) baru.

Dalam penelitian ini pengujian data menggunakan metode eksperimen dengan model *Cros Industry Standard Process for Data Mining* (CRISP-DM) yang terdiri dari 6 tahap, yaitu (Pattipeilohy, Wibowo, and Utari 2017):

1. *Business/Research Understanding Phase*
Berdasarkan dari hasil daftar calon penerima beasiswa tiga tahun terakhir sekitar 708 siswa, data dapat dilihat pada lampiran 1 untuk dapat mempecepat penyeleksian harus dilakukan pemeriksaan terhadap data-data yang dikumpulkan, hal ini dapat dibantu dengan memanfaatkan teknik klasifikasi pada data mining model ID3.
2. *Data Understanding Phase* (Fase Pemahaman Data)
Data yang digunakan adalah data sekunder, yang didapat dari wali kelas dan guru piket, didalam data tersebut dapat diketahui siswa layak atau tidak menerima beasiswa yang terdiri dari 12 atribut *predictor* dan 1 atribut hasil. Atribut-atribut yang menjadi parameter terlihat pada tabel 1.

Tabel 1. Atribut dan kategori layak dan belum layak menerima beasiswa

Atribut	Kategori
Nilai Semeseter 1 (S1)	>=80 <=80
Nilai Semester 2 (S2)	>=80

Atribut	Kategori
	<=80
Sakit	<=3 >=4
Izin	<=2 >=3
Alpa	<=2 >=3
Jumlah piagam	>=2 <=1
Jumlah do'a	>=90 <=89
Jumlah Surat	>=90 <=89
Kesiangan	<=2 >=3
Kabur	<=2 >=3
Merokok	=0 >=1
Jahil	=0 >=1

Sumber: (Saputra, Ramdhani, and Supriatman 2018)

3. Tahapan data persiapan (*Data Preparation*)

Berdasarkan data latih yang diperoleh, maka sebelum pemodelan dilakukan, diperlukan beberapa teknik *pre-processing* yang digunakan, yaitu:

a. Data Validation

Untuk mengidentifikasi dan menghapus data yang ganjil (*outlier/noise*), data yang tidak konsistensi dan data yang tidak lengkap (*missing value*).

b. Data integration and transformation

Untuk meningkatkan akurasi dan efisiensi algoritma. Data yang digunakan dalam penulisan ini bernilai kategorikal.

c. Data size reduction and discretization

Untuk memperoleh data set dengan jumlah atribut dan *record* yang lebih sedikit tetapi bersifat informatif. Proses data *training* yang digunakan dalam penelitian ini, dilakukan seleksi atribut dan penghapusan data duplikasi menggunakan software *RapidMiner*. Setelah dilakukan pengidentifikasian serta menghapus

dan menggabungkan beberapa data yang lengkap maupun tidak lengkap juga dilakukan *reduction and discretization* untuk data *training*.

4. Tahapan pemodelan (*Modelling*)

Tahap ini juga disebut tahap *learning* karena pada tahap ini data *training* diklasifikasikan oleh model dan kemudian menghasilkan sejumlah aturan. Pada penelitian ini, pembuatan model menggunakan algoritma ID3.

5. Tahap evaluasi (*Evaluation*)

Pada tahap ini dilakukan pengujian terhadap model-model untuk mendapatkan informasi model yang akurat. Evaluasi dan validasi menggunakan metode *Confision Matrix* dan Kurva ROC.

6. Tahap Deployment

Pada tahap ini diterapkan model yang memiliki akurasi tinggi atau yang paling baik pada SMP PGRI Caringin yang relevan untuk penyeleksian data calon penerima beasiswa.

HASIL DAN PEMBAHASAN

Dalam membuat pohon keputusan terlebih dahulu kita hitung jumlah *class* yang layak dan tidak belum layak dan *entropy* dari masing-masing *class* berdasarkan atribut yang telah ditentukan dengan menggunakan data *training*

Dari data *training* diketahui jumlah kasusnya ada 220 *record*, siswa yang layak berjumlah 1 *record* dan yang belum layak sebanyak 219 *record* untuk tahun pelajaran 2017/2018, jumlah *record*nya adalah 708 sehingga nilai *entropy*nya yaitu:

$$\begin{aligned} \text{Entropy (S)} &= -(28/708) \times \log_2 (28/708) + \\ &\quad -(680/708) \times \log_2 (680/708) \\ &= 0,240216038 \end{aligned}$$

Untuk mendapatkan *gain* tiap atribut, maka harus dihitung *entropy* atribut berdasarkan tiap-tiap kasus. Hasil dari perhitungan dapat dilihat pada tabel 2 dibawah ini:

Tabel 2. Perhitungan *Gain* Node 1

atribut	nilai	total	layak	belum	entropy	gain
S1	>=80	558	28	530	0,287157747	0,013896797
	<=79	150	0	150	0	
S2	>=80	500	28	472	0,31135737	0,02033089
	<=79	208	0	208	0	
SAKIT	<=3	541	28	513	0,293810032	0,015708655
	>=4	167	0	167	0	
IJIN	<=2	662	28	634	0,252722327	0,003913524
	>=3	46	0	46	0	
ALPA	<=2	393	28	365	0,370559404	
	>=3	315	0	315	0	

atribut	nilai	total	layak	belum	entropy	gain
PIAGAM	>=2	94	28	66	0,878674493	0,034524165
	<=1	614	0	614	0	
						0,123555865
DO'A	>=80	99	28	71	0,859275981	0,120063041
	<=79	609	0	609	0	
						0,114273009
SURAT	>=80	108	28	80	0,825626526	0,006434155
	<=79	600	0	600	0	
						0,004001654
KESIANGAN	<=2	634	28	606	0,261068728	0,004622411
	>=3	74	0	74	0	
						0,003039625
KABUR	<=2	661	28	633	0,253010263	0,004622411
	>=3	47	0	47	0	
						0,004001654
MEROKOK	=0	654	28	626	0,255046312	0,004622411
	>=1	54	0	54	0	
						0,004622411
JAHIL	=0	672	28	644	0,249882293	0,004622411
	>=1	36	0	36	0	
						0,004622411

Sumber: (Saputra, Ramdhani, and Supriatman 2018)

Dari hasil perhitungan *entropy* dan *gain* yang terdapat pada tabel 2, terlihat bahwa atribut Piagam yang diberi warna kuning mempunyai nilai tertinggi **0,123555865**. Oleh karena itu atribut piagam menjadi akar atau node pertama dari pohon keputusan yang terbentuk. Pada atribut piagam terdapat 2 cabang, yaitu = <=1 dan = >=2

sudah terklarifikasi yaitu dengan = <=1 bernilai **"Belum"** dan = >=2 bernilai atribut **Surat** sehingga atribut Surat belum terklarifikasi sehingga perlu dilakukan perhitungan entropi dan gain untuk masing-masing cabang tersebut. Hasil perhitungan untuk node 1.1 (= >=2) dapat dilihat pada tabel 3 berikut:

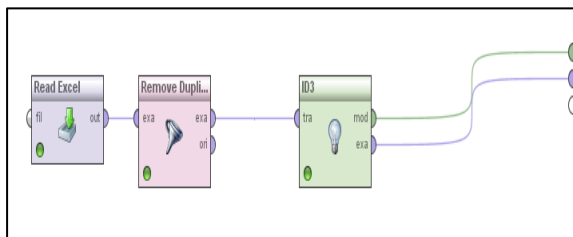
Tabel 3. Node 1.1 Hasil Konversi

atribut	nilai	total	layak	belum	entropy	gain
total kasus		136	3	133	0,152849	
S1	>=80	88	3	85	0,21451	0,014049
	<=79	48	0	48	0	
						0,016735
S2	>=80	81	3	78	0,228538	0,015944
	<=79	55	0	55	0	
						0,007429
SAKIT	<=3	83	3	80	0,224327	0,022418
	>=4	53	0	53	0	
						0,054212
IJIN	<=2	108	3	105	0,183122	0,073225
	>=3	28	0	28	0	
						0,075685
ALPA	<=2	68	3	65	0,260863	0,01226
	>=3	68	0	68	0	
						0,009274
PIAGAM	>=2	26	3	23	0,515947	0,075685
	<=1	110	0	110	0	
						0,01226
DO'A	>=90	15	3	12	0,721928	0,073225
	<=89	121	0	121	0	
						0,073225
SURAT	>=90	14	3	11	0,749595	0,075685
	<=89	122	0	122	0	
						0,075685
KESIANGAN	<=2	93	3	90	0,205593	0,01226
	>=3	43	0	43	0	
						0,01226
KABUR	<=2	102	3	99	0,191433	0,01226
	>=3	34	0	34	0	
						0,01226

atribut	nilai	total	layak	belum	entropy	gain
MEROKOK	=0	103	3	100	0,189991	0,008959
	>=1	33	0	33	0	
JAHIL	=0	116	3	113	0,173196	0,005124
	>=1	20	0	20	0	

Sumber: (Saputra, Ramdhani, and Supriatman 2018)

Hasil dari Node 1.1 diketahui nilai yang paling tinggi yaitu atribut Surat dengan nilai **0,075685**. Setelah dilakukan hasil perhitungan *entropy* dan *gain*, maka pohon keputusan akan terbentuk, dimana pohon keputusan tersebut terbentuk dengan memanfaatkan *Framework RapidMiner versi 3*, dengan desain model sebagai berikut:

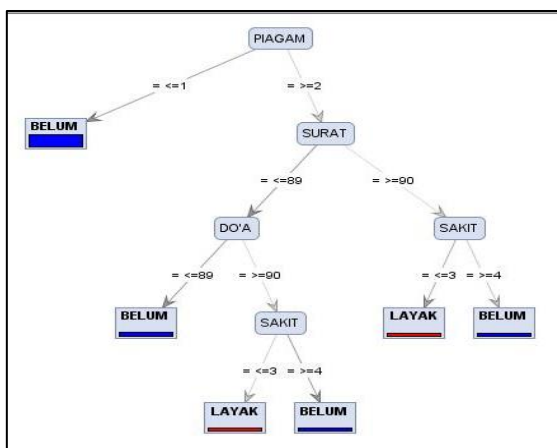


Sumber: (Saputra, Ramdhani, and Supriatman 2018)

Gambar 1. Desain Model ID3

Tahapan awal dalam *Ekperimen* adalah memasukkan dataset yang akan digunakan kedalam modul *Read Excell*, kemudian dihubungkan dengan modul *Remove Duplicate* untuk menghilangkan data yang sama, kemudian dihubungkan dengan modul algoritma ID3.

Sehingga jika model tersebut dijalankan maka akan didapatkan hasil pohon keputusan



Sumber: (Saputra, Ramdhani, and Supriatman 2018)

Gambar 2. Pohon Keputusan Penentuan Beasiswa

Pada gambar 2 diatas maka diperoleh aturan-aturan sebagai berikut, yaitu:

```

PIAGAM = <=1: BELUM {BELUM=110, LAYAK=0}
PIAGAM = >=2
|   SURAT = <=89
|   |   DO'A = <=89: BELUM {BELUM=19, LAYAK=0}
|   |   DO'A = >=90
|   |   |   SAKIT = <=3: LAYAK {BELUM=0, LAYAK=1}
|   |   |   SAKIT = >=4: BELUM {BELUM=3, LAYAK=0}
|   |   SURAT = >=90
|   |   |   SAKIT = <=3: LAYAK {BELUM=0, LAYAK=2}
|   |   |   SAKIT = >=4: BELUM {BELUM=1, LAYAK=0}

```

Evaluasi Model *Confusion Matrix*

Pada table 4 adalah perhitungan berdasarkan data *training* pada table 4, diketahui dari 136 data, 3 diklasifikasikan layak mendapatkan beasiswa sesuai dengan prediksi yang dilakukan dengan metode ID3, lalu 133 data diprediksi belum layak mendapatkan beasiswa.

Tabel 4. Model *Confusion Matrix* untuk Metode ID 3
Accuracy 97.75% +/-3.44% (mikro:97.79%)

	True BELUM	True LAYAK	Class precision
Pred BELUM	132	2	98.51%
Pred LAYAK	1	1	50.00%
Class recall	99.25%	33.33%	

Sumber: (Saputra, Ramdhani, and Supriatman 2018)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{132+1}{132+1+1+2} = 0,977941$$

$$Sensitivity = \frac{TP}{TP+FN} = \frac{132}{132+2} = 0,985075$$

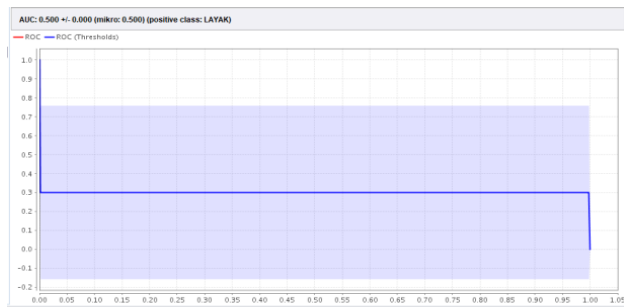
$$Specificity = \frac{TN}{TN+FP} = \frac{1}{1+1} = 0,5$$

$$PPV = \frac{TP}{TP+FP} = \frac{132}{132+1} = 0,333333$$

$$NPV = \frac{TN}{TN+FN} = \frac{1}{1+2} = 0,992481$$

Model Kurva ROC

Hasil perhitungan kurva ROC untuk algoritma ID 3 dengan menggunakan data *training* hasilnya sebesar 0,500



Sumber: (Saputra, Ramdhani, and Supriatman 2018)

Gambar 3. Kurva ROC ID3

Dilihat bentuk dari kurvanya seperti gambar 3 Hasil perhitungan kurva ROC untuk algoritma ID3 dengan menggunakan data *training* hasilnya sebesar 0,500.

Deployment

Dalam penerapan pada GUI yang dibuat, data yang digunakan sebanyak 708 *record* yang diambil secara random, hasil dari perhitungan tersebut didapat 28 siswa yang layak dan 680 siswa yang belum layak, Berikut Gambar 4 tampilan GUI yang dibuat:

Sumber: (Saputra, Ramdhani, and Supriatman 2018)

Gambar 4. Tes Uji yang Layak

Gambar 4 menggambarkan GUI untuk memprediksi penentuan beasiswa, dengan hasil layak mendapatkan beasiswa.

KESIMPULAN

Kesimpulan dari penerapan metode *iterative dichotomizer 3* (ID3) untuk penentuan beasiswa berprestasi. Dalam upaya membantu pihak sekolah khususnya tata usaha dalam menangani seleksi beasiswa, metode ini dapat menjadi alternatif pemecahan masalah diantaranya: aplikasi ini dapat membantu tata usaha sebagai pengolah dalam penentuan beasiswa

untuk lebih cepat lagi dalam menentukan yang layak beasiswa dan penentuan beasiswa dengan menggunakan metode ID3 ini, memiliki tingkat akurasi 97,75% sehingga akurat dalam analisa penyeleksian calon penerima beasiswa.

REFERENSI

- Ispriyanti, Dwi, and Abdul Hoyyi. 2016. "Analisis Klasifikasi Masa Studi Mahasiswa Prodi Statistika Undip Dengan Metode Support Vector Machine (SVM) Dan ID3 (Iterative Dichotomiser 3)." *Media Statistika* 9(1): 15–29.
- Kamal, Ahmad. 2017. "Sistem Penunjang Keputusan Penentuan Penerima Beasiswa Dengan Metode Multiple AHP." *Jurnal Sains dan Informatika* 3(2): 97. <http://ejournal.kopertis10.or.id/index.php/sains/article/view/2787>.
- Kirom, Dalu Nuzlul, Yusuf Bilfaqih, and Rusdhianto Effendie. 2012. "Sistem Informasi Manajemen Beasiswa LTS Berbasis Sistem Pendukung Keputusan Menggunakan Analytical Hierarchy Process." *Jurnal Teknik Pomits* 1(1): 1–6. <http://digilib.its.ac.id/public/ITS-paper-24313-2207100183-Paper.pdf>.
- Narti. 2017. "Sistem Pendukung Keputusan Pemilihan Siswa Berprestasi Menggunakan Metode AHP Dan Topsis." *Jurnal Informatika* 4(2): 196–205.
- Pattipeilohy, William Frado, Arief Wibowo, and Dyah Retno Utari. 2017. "Pemodelan Dan Prototipe Sistem Informasi Untuk Prediksi Pembaharuan Polis Asuransi Mobil Menggunakan Algoritma C.45." In *Seminar Nasional Teknologi Dan Informatika 2017 (SNATIF)*, Kudus, 791–99. <https://www.neliti.com/id/publications/173500/pemodelan-dan-prototipe-sistem-informasi-untuk-prediksi-pembaharuan-polis-asuran>.
- Pramadhani, Aline Embun, and Tedy Setiadi. 2014. "Penerapan Data Mining Untuk Klasifikasi Penyakit ISPA Dengan Algoritma Decision Tree." *Jurnal Sarjana Teknik Informatika e-ISSN: 2338-5197* 2(1): 831–39.
- Radhitya, Yogha, Fitro Nur Hakim, and Achmad Solechan. 2016. "Rancang Bangun Sistem Pendukung Keputusan Penentuan Penerima Beasiswa Dengan Metode SAW." *Jurnal Speed - Sentra Penelitian Engineering dan Edukasi* 8(2): 23–32.
- Saputra, Rizal Amegia, Lis Saumi Ramdhani, and Supri Supriatman. 2018. "Laporan Penelitian-Penerapan Metode Iterative Dichotomizer 3 (ID 3) Untuk Menentukan Beasiswa Berprestasi Pada SMP PGRI."