



Research paper

Rekomendasi Strategi Sosialisasi Program Studi Melalui Jalur Undangan Menggunakan Algoritma ID3 dan K-Means

Muh Azhar Hairuddin ¹, Hazriani ², Yuyun Wabula ²

¹ Universitas Islam Negeri (UIN) Alauddin Makassar, Jl. H.M. Yasin Limpo No. 36, Gowa, 92113, Indonesia

² Program Pascasarjana Sistem Komputer STMIK Handayani, Jl. Adhyaksa Baru No. 1, Makassar 90231, Indonesia

ARTICLE INFORMATION

Received: August 13th, 2020
 Revised: January 28th, 2022
 Available online: March 31st, 2022

KEYWORDS

Data Mining, Decision Tree, ID3, K-Means

CORRESPONDENCE

Phone: +6285299158300
 E-mail: muh.azhar.hairuddin@uin-alauddin.ac.id

A B S T R A C T

Based on data obtained from SPAN-PTKIN registrants in 2018 and 2019, the number of interested people through the invitation path who chose the study program at UIN Alauddin as the first choice was 30523 records. Analysis using the ID3 algorithm found that those who interested in the study of religions were more dominant from vocational students. While analysis using the K-Means shows the regions/regencies from which those interested in study programs of religions are spread in 35 regencies/cities. It can be concluded that the socialization of study programs of religions through the invitation path is recommended to be more focused on SMAs that are located in 33 districts/cities as identified in *cluster* 3. The study programs of religions are prioritized, because these study programs experienced the lowest number of registrants. It is expected that by implementing this recommended strategy, the number of interested prospective new students will draw a significant increase in the future.

PENDAHULUAN

Perkembangan teknologi informasi telah memberikan kontribusi pada cepatnya pertumbuhan jumlah data yang dikumpulkan dan disimpan dalam basis data. Data lama yang direkam ini tumbuh dan menjadi data besar universitas. Data besar ini dapat dianalisis dan digali untuk memperoleh informasi penting atau untuk proses lain untuk mendukung manajemen dalam menjaga kualitas universitas [1]. Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu.

UIN Alauddin Makassar sebagai salah satu perguruan tinggi agama islam negeri yang berdiri di Sulawesi Selatan mengelola 61 program studi. Penerimaan mahasiswa baru pada perguruan tinggi adalah sebuah kegiatan yang selalu dilaksanakan setiap tahunnya. Pada tanggal 20 Desember 2018 akreditasi perguruan tinggi telah terakreditasi A [2], namun jumlah pendaftar melalui jalur penerimaan SPAN pada tahun 2018 dan 2019 hanya mengalami sedikit peningkatan dan terjadi perbedaan signifikan antar program studi. Bahkan beberapa program studi jumlah pendaftar lebih rendah dari kuota yang tersedia (mahasiswa yang akan diterima) Hal ini menunjukkan kurang tepatnya sosialisasi yang di lakukan sebelumnya.

<https://doi.org/10.25077/jitce.6.01.14-18.2022>

Penelitian ini menerapkan metode klasifikasi dan klusterisasi digunakan pada data penerimaan mahasiswa baru (PMB) untuk mengetahui strategi penerimaan mahasiswa baru yang tepat. Atribut yang digunakan meliputi asal sekolah, tempat tinggal calon mahasiswa, kemampuan finansial orang tua serta tingkat pendidikan orang tua dengan menggunakan metode klasifikasi ID3. Pada klusterisasi yang digunakan metode K-Means dan atribut kabupaten, program studi, asal sekolah.

Penelitian ini diharapkan dapat memberikan rekomendasi untuk dapat membantu menemukan strategi penerimaan yang tepat untuk meningkatkan jumlah pendaftar calon mahasiswa baru. Rekomendasi merupakan proses memberitahukan kepada seseorang atau lebih bahwa sesuatu yang dapat dipercaya, dapat juga merekomendasikan diartikan sebagai menyarankan, mengajak untuk bergabung, menganjurkan suatu bentuk perintah disebut sebagai rekomendasi. Dalam Kamus Besar Bahasa Indonesia (KBBI) menyebutkan rekomendasi merupakan saran yang menganjurkan kemudian merekomendasikan memberikan rekomendasi, menasihatkan dan menganjurkan. Dari definisi teori tersebut dapat diambil kesimpulan bahwa rekomendasi adalah suatu proses komunikasi atas suatu produk atau jasa tertentu yang berguna untuk memberikan informasi secara personal [3].

[Attribution-NonCommercial 4.0 International](https://creativecommons.org/licenses/by-nc/4.0/). Some rights reserved

METODE

Penelitian ini menerapkan metode *data mining* untuk menggali informasi yang lebih mendalam (pengetahuan) untuk menghasilkan kebijakan sosialisasi prram studi yang lebih efektif. *Data Mining* (DM) adalah salah satu bidang yang berkembang pesat karena besarnya kebutuhan akan nilai tambah dari database dengan skala besar. Definisi data mining adalah proses ekstraksi suatu data (sebelumnya tidak diketahui, bersifat implisit, dan dianggap tidak berguna) menjadi informasi atau pengetahuan atau pola dari data yang jumlahnya besar [4]. Sesuatu yang awalnya tidak saling berhubungan, tetapi jika dilihat dari sudut yang lebih luas dengan menggunakan kepandaian manusia dan teknologi, akan memiliki informasi yang berguna [5]. Adapun metode yang digunakan adalah algoritma ID3 dan algoritma K-Means.

Algoritma Iterative Dichotomiser 3 (ID3)

Algoritma Iterative Dichotomiser 3 (ID3) merupakan sebuah metode yang digunakan untuk membangkitkan Decision Tree. Algoritma Decision Trees masuk ke dalam kategori data mining klasifikasi. Algoritma Decision Trees mengonstruksi pohon keputusan dari data Algoritma training dari sebuah data training berupa record-record dalam basis data [6]. Metode ID3 mendapatkan informasi berdasarkan entropy yang merupakan sistem pengukuran statistik. mendeskripsikan tiap contoh dan memiliki jumlah nilai yang sudah ditentukan.

Pemilihan atribut yang dijadikan node pada algoritma ID3 berbasis information gain/penambahan informasi [7]. Gain mengukur seberapa baik suatu atribut memisahkan training example ke dalam kelas target. Atribut dengan informasi tertinggi akan dipilih. Dengan tujuan untuk mendefinisikan gain, pertama-tama digunakanlah ide dari teori informasi yang disebut entropi.

$$Entropy(s) = \sum_i^c - p \log_2 p_i \tag{1}$$

Persamaan (1) digunakan untuk mencari entropi yaitu Entropi mengukur jumlah informasi yang ada pada atribut. Dimana c adalah jumlah nilai yang berada pada attribut target (jumlah kelas). Sedangkan pi menyatakan porsi atau rasio antara jumlah sampel di kelas i dengan jumlah semua same pada himpunan data. Setelah mendapatkan informasi dari semua atribut yang dihitung, atribut dengan information gain tertinggi dipilih sebagai atribut node awal (root node) serta cabang-cabangnya di buat sesuai nilai-nilai kemungkinan.

$$Gain(S, A) = Entropy(S) - \sum_{v \in Value(A)} \frac{|S_v|}{|S|} Entropy(S_v) \tag{2}$$

Untuk mendapatkan information gain digunakan persamaan (2). Dimana A adalah attribut, V menyatakan suatu nilai yang mungkin untuk attribut A, Value(A) adalah himpunan nilai-nilai yang mungkin untuk attribut A, |Sv| jumlah sampel untuk nilai v, |S| adalah jumlah seluruh sampel data, Entropy(Sv) entropi untuk sampel-sampel yang memiliki nilai v.

Algoritma K-Means

Algoritma K-Means masuk ke dalam penerapan data mining clustering. Clustering adalah data yang tidak mempunyai label/kelas sehingga sering disebut dengan teknik unsupervised learning [8]. Algoritma K-Means merupakan satu algoritma yang

mudah dan kerap digunakan di dalam teknik pengelompokan karena melibatkan pengiraan yang efisien dan tidak memerlukan banyak parameter. Tujuan dari clustering adalah pengelompokan data ke dalam beberapa *cluster* berdasarkan tingkat kemiripan [9].

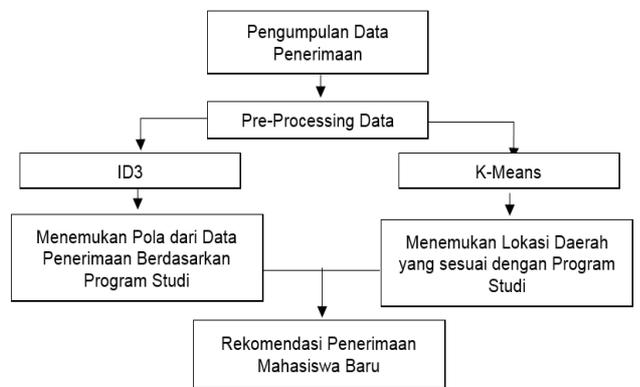
K-Means melakukan pengelompokan data dengan mengambil parameter sejumlah k kluster, dan mempartisi data kedalam kluster tersebut, dengan berpatokan pada kemiripan antar data dalam satu kluster dan ketidak miripan di antar kluster yang berbeda, pusat dari kluster adalah rata-rata dari nilai anggota kluster yang disebut centroid atau center of gravity [10].

$$C_k = \left(\frac{1}{n_k}\right) \sum d_i \tag{3}$$

Persamaan (4) digunakan untuk melakukan pengukuran persamaan atau jarak dalam proses analisa kelompok di mana menggunakan pengiraan matriks jarak (atau perbedaan). Dimana d(x,y) adalah jarak di antara x dan y, y i adalah nilai pemboleh ubah i bagi x dan x i adalah nilai pemboleh ubah i bagi y.

Implementasi Algoritma ID3 dan K-Means

Algoritma akan memeriksa dua hal yang berbeda, pada algoritma ID3 untuk pola dari setiap data penerimaan yang telah di olah sebelumnya dan menghasilkan pola pada setiap program studi. Pola ini akan di gunakan untuk membantu dalam teknik sosialisasi yang akan dilakukan. Selanjutnya dikombinasikan dengan Algoritma K-Means untuk mengetahui program studi dan asal jenjang yang sesuai masing-masing kabupaten. Tahapan implementasi kedua algoritma ini ditunjukkan pada Gambar 1.



Gambar 1. Implementasi Algoritma

HASIL DAN PEMBAHASAN

Sistem yang diusulkan mengikuti tahapan data mining dengan menggunakan algoritma ID3 untuk menentukan karakteristik program studi dan K-Means untuk menentukan daerah yang dapat di jadikan tujuan utama sosialisasi. Sistem ini akan mengolah data peminat yang ada pada UIN Alauddin Makassar selama tahun 2018 sampai tahun 2019.

Implementasi Tahapan Data Mining dengan menerapkan metode-metode tertentu untuk mengungkap pola-pola tersembunyi. Dengan kata lain, proses ini dilakukan untuk penggalian pola-pola dari data. Data yang akan diolah merupakan data penerimaan jalu SPAN yang diperoleh dari UIN Alauddin Makassar. Data

yang tersedia hanya pada tahun 2018 dan 2019. Data tersebut di filter dengan memilih data yang program studi pilihan pertamanya di UIN Alauddin, data ini di pilih agar dapat melihat minat calon mahasiswa ke program studi tersebut.

Jumlah data pada tahun record calon mahasiswa pada jalur SPAN tahun 2018 adalah 798.049 dan pada tahun 2019 adalah 867.805. Untuk record calon mahasiswa yang mendaftar di UIN Alauddin pada tahun 2018 berjumlah 12.785 dan tahun 2019 berjumlah 17.738 total semua data 30523. Record tersebut merupakan calon mahasiswa yang memilih UIN Alauddin sebagai pilihan pertama pada jalur SPAN-PTKIN

Implementasi Algoritma ID3

Algoritma ID3 digunakan dalam membangkitkan Decision Tree yang mendapatkan informasi berdasarkan entropy yang merupakan sistem pengukuran statistik. Pada kasus ini data diolah dengan menggunakan algoritma ID3 oleh program berbasis web yang telah dibuat menggunakan bahasa pemrograman PHP. Langkah pertama yang dilakukan sebelum menghitung dengan algoritma ID3 adalah mengelompokkan data penerimaan. Tabel 1 menunjukkan data sampel penerimaan dari 30523 data.

Tabel 1. Data Sampel Penerimaan

Jenis Sekolah Asal	Alamat Tempat Tinggal	Kategori Pendapatan Ortu	Pendidikan Terakhir Orangtua	Pilihan Program Studi
MA	KAB. BULUKUMBA	Rendah	SMP	Perbankan Syariah
SMA	KAB. BONE	Menengah Bawah	SMP	Biologi
MA	KAB. BULUKUMBA	Menengah Atas	Diploma	Manajemen Haji dan Umrah
SMA	KAB. LUWU TIMUR	Rendah	SMA	Biologi
SMA	KAB. SIDENRENG RAPPANG	Menengah Bawah	SMA	Kesehatan Masyarakat
MA	KOTA SORONG	Menengah Atas	SMA	Farmasi
SMA	KAB. WAJO	Menengah Bawah	SD	Komunikasi dan Penyiaran Islam
SMA	KAB. PANGKAJENE KEPULAUAN	Menengah Atas	SMA	Farmasi
SMA	KAB. ENREKANG	Menengah Bawah	SD	Pendidikan Biologi
SMA	KOTA PALOPO	Rendah	SMA	Kesehatan Masyarakat
SMA	KAB. GOWA	Menengah Bawah	SMA	Pendidikan Bahasa Arab
MA	KAB. SOPPENG	Menengah Atas	S-1	Hukum Pidana dan Ketatanegaraan
SMA	KAB. BIMA	Rendah	---	Perbankan Syariah
MA	KAB. WAJO	Rendah	SD	Pendidikan Bahasa Inggris

Pada data transaksi terdapat 30523 total pendaftar dari 37 Program Studi yang tersedia di jalur SPAN-PKIN, ditunjukkan Tabel 2. Pada data yang diperoleh program studi studi agama-agama memiliki jumlah pendaftar terendah yaitu hanya 19

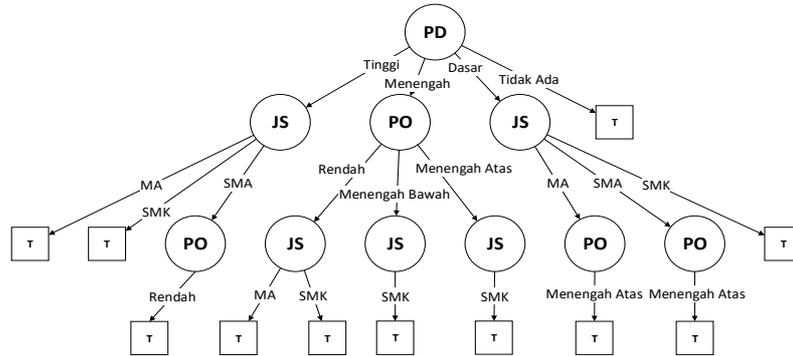
pendaftar, program studi ini yang akan menjadi sampel. Selanjutnya masing-masing data pendaftar tersebut akan dihitung dan akan menghasilkan entropy dan gain sehingga dapat dibuat pohon keputusan.

Tabel 2. Hasil Iterasi 0 – Prodi Studi Agama – Agama

0	Variabel	Pendaftaran	Studi Agama-Agama	Prodi Lain	Entropy
	Pendaftar	30523	19	30504	0.00753
	Jenis Sekolah				
	SMA	21911	15	21896	0.00819
	MA	7062	4	7058	0.00693
	SMK	1550	0	1550	0.00000
					GAIN = 0.00005
	Pendapatan Orangtua				
	Rendah	9727	8	9719	0.00962
	Menengah Bawah	13030	7	13023	0.00661
	Menengah Atas	7766	4	7762	0.00637
					GAIN = 0.00002
	Pendidikan Orangtua				
	Pendidikan Tinggi	9310	3	9307	0.00420
	Pendidikan Menengah	11449	8	11441	0.00833
	Pendidikan Dasar	9400	8	9392	0.00991
	Tidak Ada	364	0	364	0.00000
					GAIN = 0.00007

Program Studi Studi Agama-Agama adalah program studi dengan pendaftar terendah. Hasil analisa dengan algoritma ID3 nilai yang menjadi root adalah yang memiliki gain tertinggi, dari data Jenis sekolah, pendapatan orangtua dan pendidikan orangtua, yang

memiliki gain tertinggi. Proses selanjut yang memiliki gain tertinggi tersebut yang akan menjadi root, proses ini akan di ulang sampai semua node terpartisi.



Gambar 2. Hasil Pohon Keputusan Program Studi Studi Agama-Agama

Kode PD pada pohon keputusan Gambar 2 adalah Pendidikan Orangtua, selanjutnya JS adalah Jenis Sekolah, PO adalah Pendapatan Orangtua, dan T adalah tidak mendaftar. Dari pohon keputusan dapat disimpulkan bahwa jenis sekolah SMK dan Pendidikan Orangtua pada tingkat tidak pernah mengambil pendidikan formal, tidak tertarik mendaftar pada prodi Studi Agama-Agama. Pendaftar terbanyak berasal dari jejang SMA hal ini cukup menarik, karena program studi keagamaan, pendaftar terbanyak dari SMA.

Implementasi Algoritma K-Means

Algoritma K-Means merupakan satu algoritma yang melakukan pengelompokan data dengan mengambil parameter sejumlah kemudian mempartisi data kedalam kluster tersebut, dengan berpatokan pada kemiripan antar data dalam satu kluster dengan kluster lainnya. Nilai yang digunakan dalam perhitungan kemiripan dari seluruh data yang ada dengan menggunakan nilai entropy yang di peroleh menggunakan persamaan (1). Data tersebut dikelompokkan menjadi 3 cluster, dan kemudian dipilih cluster yang paling berpengaruh.

Tabel 3. Data Sampel Penerimaan

Data Ke	X	Y	Z
1	4.36953	0	4.48176
2	4.36953	0.9183	4.48176
3	4.36953	2.17347	4.13386
...
30523	4.36953	4.39525	4.48176

Pada Tabel 3 terdapat variabel X, Y dan Z, dimana X adalah entropy jenis sekolah, Y adalah entropy jenis sekolah, dan Z adalah adalah entropoy pendapatan orangtua. Data pada tabel tersebut, akan di hitung jarak pada setiap cluster. Proses selanjutnya akan dicari jarak antara centroid dengan data yang ada. Hasil dari Algoritma dapat dilihat pada Tabel 4 berikut.

Tabel 4. Hasil Klasterisasi K-Means

Variabel	Value	Jumlah Data		
		Cluster 1	Cluster 2	Cluster 3
Jenis Sekolah	SMA	483	656	20772
	MA	314	1023	5725
	SMK	131	55	1364
Penghasilan Orangtua	Menengah Bawah	419	708	11903
	Menengah Atas	374	583	6809
	Rendah	135	443	9149
Jumlah Kabupaten		370	60	35
Jumlah Data		928	1734	27861

Berdasarkan perhitungan titik centroid, pada iterasi ke 15, di dapatkan hasil centroid yang sama dengan nilai titik centroid baru. Proses ini membutuhkan waktu yang cukup lama yaitu 39 jam. Pada data pada cluster 3 adalah 27861 tetapi memiliki jumlah kabupaten berbeda yaitu 35 berbanding terbalik dengan cluster 1 jumlah data hanya 928 tapi jumlah kabupaten 370, sehingga rekomendasi sosialisasi dapat di fokuskan pada cluster ke 3 yang memiliki jumlah pendaftar terbanyak dan jumlah kabupaten terkecil.

Tabel 5 menunjukkan daftar kabupaten yang termasuk pada cluster 3 untuk semua program studi yang ada, kabupaten di urutkan berdasarkan total pendaftar tertinggi ke terendah.

Tabel 5. Hasil Klasterisasi K-Means pada Cluster 3 seluruh program studi

No.	Nama Kabupaten	Total Pendaftar
1	KOTA MAKASSAR	2967
2	KAB. GOWA	2487
3	KAB. BULUKUMBA	2442
4	KAB. BONE	2059
5	KAB. SINJAI	1646
6	KAB. ENREKANG	1491
7	KAB. TAKALAR	1395
8	KAB. WAJO	1330
9	KAB. PINRANG	1175
10	KAB. SOPPENG	1120
11	KAB. JENEPONTO	948
12	KAB. PANGKAJENE DAN KEPULAUAN	910
13	KAB. LUWU	881
14	KAB. POLEWALI MANDAR	864

Tabel 5. Hasil Klasterisasi K-Means pada Cluster 3 seluruh program studi (*lanjutan*)

No.	Nama Kabupaten	Total Pendaftar
15	KAB. BARRU	742
16	KAB. SIDENRENG RAPPANG	696
17	KAB. LUWU TIMUR	680
18	KOTA PALOPO	563
19	KAB. BANTAENG	494
20	KAB. KEPULAUAN SELAYAR	423
21	KAB. MAROS	399
22	KAB. MAJENE	360
23	KAB. LUWU UTARA	343
24	KAB. MAMUJU	244
25	KOTA PAREPARE	216
26	KAB. BIMA	191
27	KAB. TANA TORAJA	189
28	KAB. MAMUJU TENGAH	116
29	KAB. KOLAKA	116
30	KOTA BAUBAU	114
31	KOTA PALU	79
32	KAB. SUMBAWA	60
33	KOTA JAKARTA SELATAN	52
34	KOTA BONTANG	48
35	KAB. PANGANDARAN	21

Untuk daftar kabupaten khusus program studi keagamaan, memiliki hasil yang sama dengan semua prodi, yaitu *cluster* 3 memiliki jumlah pendaftar terbanyak dan jumlah kabupaten terendah, jumlah kabupaten pada *cluster* 1 adalah 87 kabupaten dengan jumlah pendaftar 128, untuk *cluster* 2 ada 52 kabupaten dengan total pendaftar 514 dan *cluster* 3 ada 33 kabupaten dengan jumlah pendaftar 9605. Adapun total seluruh pendaftar khusus keagamaan ada 10247. Tabel 6 menunjukkan kabupaten khusus program studi keagamaan yang di urutkan berdasarkan total pendaftar tertinggi ke terendah.

Tabel 6. Hasil Klasterisasi K-Means pada Cluster 3 Program Studi Keagamaan

No.	Nama Kabupaten	Total Pendaftar
1	KOTA MAKASSAR	974
2	KAB. BULUKUMBA	944
3	KAB. GOWA	887
4	KAB. BONE	665
5	KAB. TAKALAR	523
6	KAB. ENREKANG	517
7	KAB. SINJAI	511
8	KAB. PINRANG	452
9	KAB. WAJO	442
10	KAB. POLEWALI MANDAR	410
11	KAB. SOPPENG	372
12	KAB. JENEPONTO	328
13	KAB. PANGKAJENE DAN KEPULAUAN	289
14	KAB. BARRU	253
15	KAB. SIDENRENG RAPPANG	245
16	KAB. LUWU	237
17	KAB. BANTAENG	202
18	KAB. LUWU TIMUR	194
19	KOTA PALOPO	178
20	KAB. MAROS	165
21	KAB. MAJENE	132
22	KAB. KEPULAUAN SELAYAR	121
23	KAB. LUWU UTARA	120
24	KAB. MAMUJU	83
25	KOTA PAREPARE	78
26	KAB. TANA TORAJA	77
27	KOTA BAUBAU	39
28	KAB. MAMUJU TENGAH	38
29	KOTA PALU	38
30	KAB. KOLAKA	35
31	KOTA JAKARTA SELATAN	27
32	KAB. SUMBAWA	15
33	KOTA BONTANG	14

KESIMPULAN

Penerapan metode klasifikasi ID3 dan klasterisasi K-Means dapat memberikan rekomendasi untuk meningkatkan pendaftaran pada prodi studi agama-agama. Analisis menggunakan algoritma ID3 diperoleh bahwa peminat program studi agama-agama lebih dominan adalah siswa SMA. Sedangkan analisis menggunakan K-Means menunjukkan daerah/kabupaten asal peminat prodi agama-agama yang tersebar pada 33 kabupaten/ kota. Sehingga dapat disimpulkan bahwa sosialisasi prodi agama-agama melalui jalur undangan direkomendasikan agar lebih difokuskan pada SMA yang berada pada 33 kabupaten/ kota sebagaimana yang teridentifikasi pada *cluster* 3. Prodi agama-agama menjadi prioritas, dikarenakan prodi ini merupakan prodi yang memiliki jumlah pendaftar terendah. Sehingga diharapkan dengan menerapkan strategi ini kedepannya jumlah peminat akan mengalami peningkatan yang signifikan.

REFERENSI

- [1] H. Muttaqien, M. Lutfi, M. KH, A. Muis, and H. Zainuddin, "Recommendation of Student Admission Priorities Using K-Means Clustering," in *Proceedings of the 1st International Conference on Science and Technology, ICOST 2019, 2-3 May, Makassar, Indonesia*, 2019.
- [2] [2] N. Bahar, "Tutup Tahun 2018, UIN Alauddin Makassar Raih Akreditasi A - UIN Alauddin Makassar," *UIN Alauddin*, 21-Dec-2018. [Online]. Available: <https://uin-alauddin.ac.id/berita/detail/tutup-tahun-2018-uin-alauddin-makassar-raih-akreditasi-a>. [Accessed: 29-May-2020].
- [3] [3] "KBBI Daring." [Online]. Available: <https://kbbi.kemdikbud.go.id/entri/rekomendasi>. [Accessed: 30-May-2020].
- [4] [4] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining (Fourth Edition)*. 2017.
- [5] [5] A. Berson and S. J. Smith, *Data warehousing, data mining, and OLAP*. McGraw-Hill, Inc., 1997.
- [6] [6] D. T. Larose and C. D. Larose, *Discovering knowledge in data: an introduction to data mining*, vol. 4. John Wiley & Sons, 2014.
- [7] [7] S. Yang, J.-Z. Guo, and J.-W. Jin, "An improved Id3 algorithm for medical data classification," *Comput. Electr. Eng.*, vol. 65, Aug. 2017.
- [8] [8] H. He and Y. Tan, "Neurocomputing A two-stage genetic algorithm for automatic clustering," *Neurocomputing*, vol. 81, pp. 49–59, 2012.
- [9] [9] W. Xiang, N. Zhu, S. Ma, X. Meng, and M. An, "A dynamic shuffled differential evolution algorithm for data clustering," *Neurocomputing*, vol. 158, pp. 144–154, 2015.
- [10] [10] J. Han and M. Kamber, "Data Mining: Concepts and Techniques 2nd Edition, Second Edi." San Francisco: Elsevier, 2006.