

**PENERAPAN DATA MINING DALAM PENGELOMPOKAN PENYEBARAN
DATA MAHASISWA BARU UNTUK STRATEGI PROMOSI
PROGRAM STUDI TEKNIK INFORMATIKA
(Studi Kasus : IIB Darmajaya)**

Anggie Fiqy Ramadhan , Nisar Zaidal

Program Studi Teknik Informatika Fakultas Ilmu Komputer

Institut Informatika Dan Bisnis Darmajaya Bandar Lampung

E-mail : anggie.fiqy30@gmail.com

Abstract

Accepting new students in a tertiary institution generates a lot of data and happens repeatedly. IIB Darmajaya, in promotional activities, has not fully utilized the data to determine strategies, evaluation of promotions, and promotional strategy. Thus, the promotional activities are running inefficiently and are not on target. This study uses data mining and the K-means Clustering algorithm to group data on new students of the Informatics Engineering Study Program for the last three years based on the similarity of attributes such as gender, city/district of origin, and type of school. Cluster validation was carried out using the Davies-Bouldin Index (DBI). The tools used are Rapidminer as a tool for testing, experimenting, and comparing with the system created. The clusters formed after the K-Means Clustering process are two clusters. The first cluster has 75 student data, and the second has 204. The DBI value obtained is -0.548, indicating that the collection formed has good quality compared to others. The results of this study can utilize cluster data for promotional activities, namely by adjusting appropriate and efficient promotion strategies.

Article History

Submitted 18 januari 2024

Accepted 24 januari 2024

Published 25 januari 2024

Keywords : New Student Admissions, Promotion Strategy, Data Mining, Kmeans Clustering, Davies Bouldien Index, Rapidminer.

Abstrak

Proses penerimaan mahasiswa baru dalam suatu perguruan tinggi menghasilkan banyak data. Hal ini akan terjadi secara berulang di perguruan tinggi. IIB Darmajaya dalam melakukan kegiatan promosi belum sepenuhnya memanfaatkan data-data yang dimiliki untuk kebutuhan penentuan strategi baik evaluasi promosi maupun strategi promosi, sehingga kegiatan promosi yang berjalan tidak efisien dan tidak tepat sasaran. Penelitian ini menggunakan data mining dan algoritma *K-means Clustering* untuk mengelompokkan data mahasiswa baru Program Studi Teknik Informatika 3 tahun terakhir berdasarkan kesamaan atribut seperti jenis kelamin, kota/kabupaten asal, dan jenis sekolah. Validasi cluster dilakukan dengan menggunakan Davies-Bouldin Index (DBI). Tools yang digunakan adalah *Rapidminer* sebagai alat bantu untuk pengujian, percobaan, dan perbandingan dengan sistem yang dibuat. Cluster yang terbentuk setelah proses *K-Means Clustering* adalah dua cluster. Cluster pertama berjumlah 75 data mahasiswa dan cluster kedua berjumlah 204 data mahasiswa. Nilai DBI yang diperoleh -0.548 menunjukkan bahwa cluster yang terbentuk memiliki kualitas yang baik dari cluster yang lainnya. Hasil penelitian ini dapat memanfaatkan data cluster untuk kegiatan promosi, yaitu dengan menyesuaikan strategi promosi yang tepat dan efisien.

Sejarah Artikel

Submitted 18 januari 2024

Accepted 24 januari 2024

Published 25 januari 2024

Kata Kunci : Penerimaan Mahasiswa Baru, Strategi Promosi, Data Mining, *K-means Clustering*, Davies Bouldien Index, *Rapidminer*.

PENDAHULUAN

Mahasiswa merupakan Generasi-generasi penerus bangsa yang diperlengkapi dan mempunyai panggilan untuk menempuh pendidikan tinggi atau universitas sesudah menempuh pendidikan SMA, SMK atau setaranya. Bukan hanya dipersiapkan dalam

hal Intelektual, tetapi pada universitas, mahasiswa pula dibekali dengan pola pikir yang tinggi (Damanik & Sigiro, 2021).

Persaingan penerimaan mahasiswa baru di semua perguruan tinggi negeri dan swasta saat ini semakin meningkat pesat, ada yang menghabiskan banyak uang untuk materi iklan dan ada pula yang menawarkan berbagai penawaran seperti murahnya biaya perkuliahan, setelah lulus cepat mendapatkan pekerjaan dan lain sebagainya (Budiman, 2019).

Penerimaan Mahasiswa Baru (PMB) merupakan program rutin yang dijalankan oleh setiap lembaga pendidikan seperti universitas untuk menjaring mahasiswa baru setiap tahunnya (Isnanto & Widodo, 2021). Proses penerimaan mahasiswa baru dalam suatu perguruan tinggi menghasilkan banyak data berupa profil mahasiswa baru tersebut. Hal ini akan terjadi secara berulang di perguruan tinggi. Penumpukan data mahasiswa yang konstan akan memperlambat pencarian informasi pada data tersebut. Berdasarkan banyaknya data mahasiswa, informasi yang tersembunyi dapat diketahui dengan cara melakukan pengolahan terhadap data tersebut sehingga bisa berguna bagi universitas (Yunita, 2018).

IIB Darmajaya merupakan salah satu perguruan tinggi swasta yang berada di Kota Bandar Lampung, perguruan tinggi ini memiliki beberapa program studi, salah satunya adalah program studi Teknik Informatika. Dalam melakukan kegiatan promosi penerimaan mahasiswa baru Program Studi Teknik Informatika, PMB IIB Darmajaya belum sepenuhnya memanfaatkan data-data yang dimiliki untuk kebutuhan penentuan strategi baik evaluasi promosi maupun strategi promosi, sehingga kegiatan promosi yang berjalan tidak efisien dan tidak tepat sasaran. Beragamnya asal wilayah mahasiswa baru dari tahun ke tahun pada Program Studi Teknik Informatika ini tentu membuat pihak dari PMB IIB Darmajaya ingin melihat penyebarannya, namun untuk mendapatkan informasi dari data-data yang diperoleh PMB IIB Darmajaya masih menggunakan cara manual sehingga dapat memperlambat pencarian dalam mendapatkan informasi dari data-data tersebut.

Berdasarkan data Program Studi Teknik Informatika yang didapat dari PMB IIB Darmajaya, pada tahun 2020 berjumlah 22 data mahasiswa baru, pada tahun 2021 berjumlah 154 data mahasiswa baru dan pada tahun 2022 berjumlah 180 data mahasiswa baru. Dari data tersebut penulis bermaksud untuk memanfaatkan data historis untuk terus meningkatkan mahasiswa baru dari tahun ke tahun dalam melakukan kegiatan penerimaan mahasiswa baru dengan strategi promosi yang tepat sasaran, data historis sangat berguna dalam menentukan strategi promosi yang tepat apabila data historis tersebut digali pengetahuannya.

Penggunaan data *mining* sangat membantu untuk mendapatkan sebuah informasi berupa pengetahuan baru dengan memanfaatkan data historis yang ada. Secara umum, data *mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstraksi dan mengidentifikasi informasi bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Sutanto et al., 2020). Dalam mendapatkan sebuah informasi berupa pengetahuan baru, perlu dilakukan pengelompokan data mahasiswa baru yang tersebar dari tahun ke tahun, dengan dilakukannya pengelompokan, diharapkan dapat mengetahui kelompok-kelompok wilayah mana saja yang dapat dijadikan prioritas untuk promosi program studi teknik informatika kedepannya.

Melihat dari permasalahan-permasalahan yang sudah dipaparkan diatas, penulis akan melakukan penelitian dengan menerapkan *data mining* dalam pengelompokkan penyebaran data mahasiswa baru untuk strategi promosi Program Studi Teknik Informatika IIB Darmajaya. Data yang diambil dalam penelitian ini adalah data mahasiswa baru Program Studi Teknik Informatika 3 tahun terakhir yaitu, tahun 2020, 2021 dan 2022. Data ini bersumber dari PMB IIB Darmajaya. Dalam permasalahan yang ada penulis bermaksud merancang sebuah sistem yang menerapkan *data mining* untuk pengelompokkan data. *tools* yang digunakan dalam penelitian ini adalah *Rapidminer* untuk membantu pengujian, percobaan dan perbandingan dengan sistem yang dibuat sedangkan algoritma yang digunakan dalam penelitian ini adalah algoritma *K-means Clustering*, dengan algoritma ini, data yang diperoleh dapat dikelompokkan menjadi beberapa kelompok berdasarkan kesamaan datanya, seperti data yang memiliki karakteristik yang sama dikelompokkan dalam satu klaster dan data yang memiliki karakteristik berbeda dikelompokkan ke dalam klaster lain yang memiliki karakteristik berbeda. Dengan adanya pengelompokkan-pengelompokkan data seperti ini, diharapkan bagian PMB IIB Darmajaya dapat melakukan kegiatan promosi dengan strategi yang tepat untuk mendapatkan calon mahasiswa baru.

Berdasarkan uraian diatas, maka penulis melakukan penelitian dalam bentuk laporan skripsi ini dengan judul “**PENERAPAN DATA MINING DALAM PENGELOMPOKKAN PENYEBARAN DATA MAHASISWA BARU UNTUK STRATEGI PROMOSI PROGRAM STUDI TEKNIK INFORMATIKA (Studi Kasus : IIB Darmajaya)**” diharapkan hasil penelitian ini dapat membantu kegiatan penerimaan mahasiswa baru dalam mengambil keputusan untuk melakukan strategi promosi yang tepat sasaran.

METODE

PRISMA (Preferred Reporting Items for Systematic Reviews And Meta-Analyses) merupakan alat dan panduan yang digunakan untuk melakukan penelitian terhadap sebuah systematic reviews dan meta analysis (Mohamed et al., 2020). Menurut Mohamed et al., (2020), Prisma diagram adalah Diagram alir menggambarkan aliran informasi melalui fase yang berbeda dari tinjauan sistematis. Ini memetakan jumlah catatan yang diidentifikasi, disertakan dan dikecualikan, dan alasan pengecualian. Templat yang berbeda tersedia tergantung pada jenis ulasan (baru atau diperbarui) dan sumber yang digunakan untuk mengidentifikasi studi . Pembuatan Systematic Literature Review (PRISMA) terdiri dari 4 langkah, yaitu :

1. *Identification* (Identifikasi) journal yang akan disertakan dalam meta-analisis.
2. *Screening* (Penyaringan), penyaringan atau pemilihan data.
3. *Eligibility* (Kelayakan), menentukan artikel yang akan dijadikan bahan literature review
4. *Included* (Inklusi), yakni penggabungan dan pelaporan hasil.

2.1.1 Identification

Identification (Identifikasi) adalah kegiatan yang bertujuan untuk memeriksa dan menganalisa, dalam proses ini penulis mengidentifikasi masalah penelitian

(*research problem*) dikarenakan hal tersebut menentukan kualitas suatu penelitian (Mohamed et al., 2020).

2.1.2 Screening

Screening (Penyaringan) adalah sebuah proses penyaringan untuk memasukkan atau mengecualikan artikel berdasarkan kriteria yang ditentukan oleh penulis (Mohamed et al., 2020).

2.1.3 Eligibility

Eligibility adalah tahap penentuan kelayakan artikel semua artikel diperiksa dengan membaca judul, abstrak, hasil dan pembahasan untuk memastikan memenuhi kriteria inklusi dan sejalan dengan tujuan penelitian saat ini (Mohamed et al., 2020).

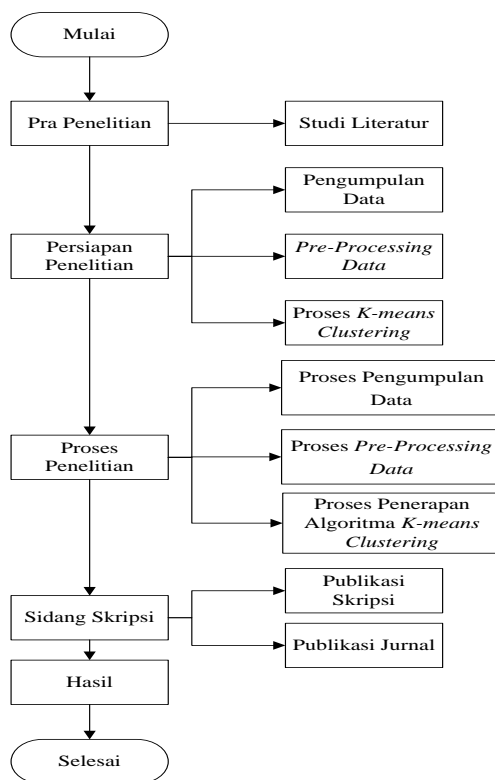
2.1.4 Included

Included merupakan tahap terakhir dari PRISMA (*Preferred Reporting Items for Systematic Reviews And Meta-Analyses*) flow diagram, pada tahap ini kriteria inklusi adalah dimana artikel keluar dan memenuhi persyaratan untuk di analisis (Mohamed et al., 2020).

Metode yang digunakan dalam penelitian ini yaitu menerapkan metode *Data Mining* dengan algoritma *K-means Clustering* pada data mahasiswa baru Program Studi Teknik Informatika untuk dianalisa dan dikelompokkan berdasarkan parameter jenis sekolah, asal kota/kabupaten dan jenis kelaminnya.

3.1 Alur Penelitian

Alur penelitian adalah konsep atau gambaran dari penelitian yang akan dilakukan. Penjabaran alur penelitian tersebut dapat dilihat pada gambar 3.1 berikut



Gambar 3.1 Diagram Alur Penelitian

HASIL DAN PEMBAHASAN

Hasil dan Pembahasan pada penelitian ini dijelaskan mengenai hasil dan implementasi program dari berbagai tahapan yang telah dirancang sebelumnya.

4.1 Penentuan Jumlah Cluster Terbaik

Dalam penentuan jumlah cluster terbaik digunakan metode *Davies Bouldin Index*, Nilai *Davies Bouldin Index* untuk setiap ukuran cluster pada k-means dapat dilihat pada tabel 4.1 berikut :

Tabel 4.1 Nilai *Davies Bouldin Index*

Ukuran Cluster (K)	Davies Bouldin Index (DBI)
2	-0.548
3	-0.721
4	-0.972
5	-0.959

Tabel 4.1 diatas menunjukkan hasil dari *Davies Bouldin Index* untuk setiap ukuran cluster. Berdasarkan tabel tersebut, nilai *Davies Bouldin Index* berbanding lurus dengan ukuran cluster, semakin rendah nilai *davies bouldin index* akan memberikan hasil yang baik. *Davies Bouldin index* paling rendah didapatkan pada ukuran cluster 2 sebesar -0.548 dan paling tinggi didapatkan pada ukuran cluster 4 sebesar -0.972. Nilai DBI paling baik yang dihasilkan pada dataset penelitian ini adalah -0.548 pada ukuran cluster 2 sehingga pada penelitian dataset mahasiswa baru ini jumlah cluster terbaik adalah 2 cluster.

4.2 Hasil Penerapan Program

Pada tahapan ini dilakukan penerapan dalam bentuk program. Program dibuat dengan bahasa pemrograman PHP dan menggunakan database MySQL. Berikut adalah tahapan penerapan program yang telah dibuat. 4.2.1 Perbandingan *Rapidminer* dengan Aplikasi Berbasis Website

Dalam penerapan aplikasi yang dibuat, penulis menguji 279 record data yang sesuai dengan penerapan *Rapidminer* sebelumnya. Setelah dianalisa ditemukan perbedaan dalam jumlah proses. Dalam hal ini yang menjadikan perbedaan dalam jumlah proses yaitu nilai awal centroid yang dipilih tidak sama sehingga adanya selisih hasil perhitungan dalam menggunakan *Rapidminer* dan Aplikasi berbasis website. Menentukan centroid awal pada aplikasi berbasis website yang dibuat penulis menentukan secara acak sesuai keinginan. Sedangkan pada aplikasi *Rapidminer* dalam menentukan centroid awal sudah secara otomatis dilakukan oleh aplikasi. Hasil perbandingan perhitungan *Rapidminer* dan Aplikasi berbasis website bisa dilihat pada tabel 4.2 berikut :

Tabel 4.2 Hasil Perbandingan *Rapidminer* Dengan Aplikasi Web

Hasil Perbandingan		
	C1	Total

Rapidminer	64	215	279
Aplikasi Web	75	204	279

4.2.2 Pengelompokan Data Dengan Aplikasi Berbasis Website
 Dalam penerapan aplikasi yang dibuat, penulis menguji 279 record data sehingga hasil tersebut bisa dilihat pada tabel 4.3 berikut :

Tabel 4.3 Hasil Cluster 1 Penerapan Algoritma *K-Means* (Berbasis Web)

Jenis Kelamin		Kota atau Kabupaten		Jenis Sekolah	
Laki-Laki	58	Pringsewu	11	SMA Negeri	35
Perempuan	17	Metro	9	SMK Negeri	17
		Lampung Timur	9	SMA Swasta	7
		Tulang Bawang	9	LAINNYA	10
		Tanggamus	8	SMK Swasta	6

Kesimpulan

Kesimpulan dalam penelitian ini adalah sebagai berikut:

Jumlah *cluster* yang didapatkan berdasarkan hasil validasi DBI (*Davies Bouldien Index*) yaitu terbentuk *cluster* $k=2$ dengan nilai -0.548 .

Strategi promosi bagi calon mahasiswa baru IIB Darmajaya Program Studi Teknik Informatika yang tepat sasaran adalah dengan mempertimbangkan kota/kabupaten asal mahasiswa dan jenis sekolah mahasiswa berdasarkan hasil cluster yang sudah terbentuk sesuai jenis kelamin mahasiswa. Jadi untuk sasaran promosi Program Studi Teknik Informatika IIB Darmajaya ke depannya bisa dilakukan berdasarkan Cluster 1 pada SMA Negeri yang berada di wilayah Kabupaten Pringsewu, Metro, Lampung Timur, Tulang Bawang. Cluster 2 diperoleh sasaran untuk promosi dengan jenis sekolah sekolah SMA Negeri, SMK Negeri yang berada di kota/kabupaten asal terbanyak yaitu Kota Bandar Lampung dan Kabupaten Lampung Tengah. Dari penelitian yang telah diterapkan menggunakan metode *K-means Clustering* ini didapatkan hasil bahwa mahasiswa angkatan 2020 sampai 2022 yang dibagi menjadi 2 cluster dimana pada cluster 1 terdapat mahasiswa dengan asal kota/kabupaten dan jenis sekolah terendah. Jumlah mahasiswa terendah di Program Studi Teknik Informatika ini berasal dari Kabupaten Rejang Lebong, Kabupaten Sarolangun dan Kabupaten Bogor dengan asal sekolah paling sedikit dari SMA Swasta dan SMK Swasta. Kemudian pada cluster 2 berisikan mahasiswa dengan jumlah frekuensi kota/kabupaten asal lebih banyak dibandingkan dengan cluster 1.

Saran

Saran yang dapat penulis berikan untuk penelitian lebih lanjut yaitu :

1. Agar menyempurnakan kekurangan dalam aplikasi tersebut untuk proses *K-means Clustering* agar lebih sempurna kedepannya.
2. Pembuatan website ini masih sangat sederhana, terutama dari segi penampilan dan segi keamanan, sehingga ada baiknya di pengembangan sistem

selanjutnya dapat membuat website ini menjadi lebih menarik dan sistem keamanannya lebih terjamin.

3. Menambahkan sumber data dalam penelitian ini tidak hanya mencangkup data pada Program Studi Teknik Informatika tetapi mencangkup pada data Fakultas ataupun mencangkup semua data di IIB Darmajaya itu sendiri.

DAFTAR PUSTAKA

- Adhitama, R., Burhanuddin, A., & Ananda, R. (2020). Penentuan Jumlah Cluster Ideal Smk Di Jawa Tengah Dengan Metode X-Means Clustering Dan K-Means Clusterin. *JIKO (Jurnal Informatika Dan Komputer)*, 3(1), 1–5. <https://doi.org/10.33387/jiko.v3i1.1635>
- Adisasmitho, W. (2007). Faktor Risiko Diare pada Bayi dan Balita di Indonesia. *Jurnal Makara Kesehatan*, 11(1), 1–10.
- Budiman, R. (2019). Penerapan Data Mining Untuk Menentukan Lokasi Promosi Penerimaan Mahasiswa Baru Pada Universitas Banten Jaya (Metode K-Means Clustering). *ProTekInfo (Pengembangan Riset Dan* <http://ejurnal.lppmunsera.org/index.php/ProTekInfo/article/view/1691>
- Cahyono, E. A., Sutomo, & Harsono, A. (2019). Literatur Review: Panduan Penulisan dan Penyusunan. *Jurnal Keperawatan*, 12.
- Chasanah, T. T., & Widiyono, W. (2017). Penentuan Strategi Promosi Penerimaan Mahasiswa Baru dengan Algoritma Clustering K-Means. *IC-Tech*. <https://ejournal.stmik-wp.ac.id/index.php/ictech/article/view/17>
- Damanik, N., & Sigiro, M. (2021). PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING PADA PENERIMAAN MAHASISWA BARU SEBAGAI METODE PROMOSI. *JUTISAL Jurnal Teknik Informatika Universal*. <https://universal.ac.id/jurnal/index.php/jutisal/article/view/4>
- Hairani, H., Susilowati, D., Puji Lestari, I., Marzuki, K., & Mardedi, L. Z. A. (2022). Segmentasi Lokasi Promosi Penerimaan Mahasiswa Baru Menggunakan Metode RFM dan K-Means Clustering. *MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 21(2), 275–282. <https://doi.org/10.30812/matrik.v21i2.1542>
- Isnanto, S., & Widodo, S. (2021). PENERAPAN DATA MINING PADA PENERIMAAN MAHASISWA BARU DENGAN ALGORITMA K-MEANS CLUSTERING. *Jurnal Tekinkom (Teknik Informasi Dan* <http://jurnal.murnisadar.ac.id/index.php/Tekinkom/article/view/367>
- Mahmudi, I., Indriyanti, A. D., & ... (2020). Penerapan Algoritma K-Means Clustering Sebagai Strategi Promosi Penerimaan Mahasiswa Baru pada Universitas Hasyim Asy'ari Jombang. *Inovate: Jurnal Ilmiah* ...

<http://ejournal.unhasy.ac.id/index.php/inovate/article/download/692/555>

Mohamed, R., Ghazali, M., & Samsudin, M. A. (2020). A Systematic Review on Mathematical Language Learning Using PRISMA in Scopus Database. *Eurasia Journal of Mathematics, Science and Technology Education*, 16(8), 1–12. <https://doi.org/10.29333/ejmste/8300>

Sarman, S. N., & Soebagyo, J. (2022). Analisis Bibliometrik terhadap Kemampuan Berpikir Kritis Matematika Berdasarkan Pemecahan Masalah Berbasis VOS Viewer. *Vygotsky: Jurnal ...*, 117–128. <http://jurnalpendidikan.unisla.ac.id/index.php/VoJ/article/view/590%0Ahttp://jurnalpendidikan.unisla.ac.id/index.php/VoJ/article/download/590/407>

Setiawan, R. (2017). PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI PROMOSI MAHASISWA BARU (Studi Kasus : Politeknik LP3I Jakarta). *Jurnal Lentera Ict*. <http://journal-computing.org/index.php/journal-sea/article/view/10>

Sutanto, F. A., Yulianton, H., & Razaq, J. A. (2020). ANALISA SEBARAN MAHASISWA UNIVERSITAS STIKUBANK MENGGUNAKAN METODE K-MEANS. [www.unisbank.ac.id. https://www.unisbank.ac.id/ojs/index.php/sendu/article/view/8018/2988](http://www.unisbank.ac.id/ojs/index.php/sendu/article/view/8018/2988)

Wulandari, S., Iqbal, M., & Gunadarma, U. (2021). Analisis Kelayakan Lokasi Promosi Dalam Penerimaan Mahasiswa Baru (PMB) Dengan Algoritma Naïve Bayes & Decision Tree C4. 5. In vol. [pdfs.semanticscholar.org. https://pdfs.semanticscholar.org/9506/99af997ec6b9775915c5f2f5e533de0630bb.pdf](https://pdfs.semanticscholar.org/9506/99af997ec6b9775915c5f2f5e533de0630bb.pdf)

Yulmaini, Y., & Kom, S. (2007). Desain Sistem Jaringan Pengontrol Aktivitas Praktikum Mahasiswa Di Laboratorium Komputer STMIK Darmajaya. *Jurnal Informatika*, 7(1).

Yunita, F. (2018). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru. *Sistemasi: Jurnal Sistem Informasi*. <http://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/view/388>

