

PREDIKSI KELULUSAN JALUR MASUK PERGURUAN TINGGI BANDA ACEH MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE

Marhamah

*Program Studi Pendidikan Teknologi Informasi Fakultas Tarbiyah dan Keguruan,
Universitas Islam Negeri Ar-Raniry Banda Aceh
Email: marhamahm7@gmail.com*

Abstract: Nowadays there are several types of entry points for new student admissions at tertiary institutions. There are many ways that every prospective student can prepare to pass the PTN entrance selection exam. Therefore the research aim to predict the graduation of tertiary entrance by classification method using the SVM algorithm (Support Vector Machine) which is assisted by WEKA machine learning, using data of new students in the 2019 school year. The final results in this study there are two variables that have the most relationship both the PTN entry selection tutoring variable and the interest pathway variable, with the PTN entry selection tutoring variable having a Pearson correlation value of -0.180 ** and a significance value of 0.002, the interest pathway having an accuracy value of 0.311 ** and a significance value of 0,000. Then based on the results of cross-validation testing and percentage split SVM algorithm has very good accuracy with an average accuracy of 99% with an AUC (Area Under Curve) value of 0.9907.

Keyword: PTN Entry Data, Data Mining, Support Vector Machine, Split Percentage

Abstrak : Saat ini ada beberapa jenis jalur masuk penerimaan mahasiswa baru di perguruan tinggi. Banyak cara yang bisa dipersiapkan oleh setiap calon mahasiswa untuk bisa lulus ujian seleksi masuk PTN. Penelitian ini memprediksi kelulusan jalur masuk perguruan tinggi dengan metode klasifikasi menggunakan algoritma SVM yang dibantu dengan *machine learning* WEKA. Hasil akhir penelitian ini terdapat dua variabel yang memiliki hubungan paling baik yaitu variabel bimbel seleksi masuk PTN dan variabel jalur minat, dengan variabel bimbel seleksi masuk PTN memiliki nilai *pearson correlation* sebesar -0,180** dan nilai signifikansi sebesar 0,002, jalur minat memiliki nilai akurasi sebesar 0,311** dan nilai signifikansi sebesar 0,000. Kemudian berdasarkan hasil pengujian *cross-validation* dan *percentage split* algoritma SVM memiliki akurasi yang sangat baik dengan rata-rata akurasi mencapai 99% dengan nilai AUC (*Area Under Curve*) sebesar 0.9907 dan waktu konsumsi yang dibutuhkan untuk pengujian hanya sebesar 0.01-0.02 *second*.

Kata Kunci : Jalur Masuk PTN, Data Mining, Support Vector Machine, Percentage Split

1. Pendahuluan

Banyak cara yang bisa dipersiapkan oleh setiap calon mahasiswa untuk bisa lulus ujian seleksi masuk PTN diantaranya adalah dengan mengikuti bimbel, les privat, mempunyai prestasi akademik, dan nilai UN yang memadai. Akan tetapi ada kasus dimana calon mahasiswa sudah melakukan persiapan tersebut namun hasilnya mereka dinyatakan tidak lulus. Ada juga kasus dimana calon mahasiswa tersebut tidak melakukan persiapan yang matang dinyatakan lulus di jalur yang

diminati. Ini menandakan persiapan-persiapan tersebut belum bisa menjamin calon mahasiswa untuk lulus di jalur yang diminati. Metode data mining adalah salah satu cara untuk menganalisis prediksi pada permasalahan tersebut.

Dalam proses pengolahan data dengan menggunakan data mining, telah banyak dilakukan penelitian sebelumnya, diantaranya penelitian tentang “implementasi data mining untuk memprediksi kelulusan mahasiswa menggunakan metode Naive Bayes, dengan hasil akhir memiliki nilai akurasi sebesar 94% (Syarli & Muin, 2016). Selain itu penelitian tentang “Analisis kinerja Metode Naive Bayes Dan SVM Untuk Penentuan Pola Kelompok Penyakit”, hasil yang didapatkan dengan metode SVM dengan nilai akurasi mencapai 99%, dan metode Naive Bayes dengan nilai akurasi mencapai 93%. Dari nilai akurasi yang didapatkan menunjukkan metode SVM lebih akurat daripada metode naive bayes (SITANGGANG, 2017). Untuk mencapai pengembangan penelitian yang lebih baik, peneliti ingin menggunakan metode Support Vector Machine (SVM) untuk memprediksi kelulusan mahasiswa di jalur masuk PTN di Banda Aceh. Data yang digunakan berkisar antara 500-1000 data dari beberapa perguruan tinggi Banda Aceh.

2. Kajian pustaka

2.1 Jalur Masuk Perguruan Tinggi

Pendidikan tinggi merupakan salah satu lembaga untuk mempersiapkan peserta didik menjadi anggota masyarakat yang bisa memenuhi ke butuhan masyarakat dalam berbagai bidang [1]. Hal tersebut menyangkut dengan perubahan UU pendidikan tinggi tentang penerimaan mahasiswa baru (Abdullah, 2018). Mengenai proses penerimaan mahasiswa di Perguruan Tinggi terdapat beberapa jalur masuk yaitu SNMPTN, SBMPTN, UM-PTKIN, SPAN-PTKIN dan PMB sebagaimana yang telah diterangkan dalam peraturan pemerintah RI Nomor 34 tahun 2010 (Salinggih, 2019).

2.2 Data Mining

Data mining adalah kegiatan mencari dan menggali informasi yang tidak dikenal secara manual dari database. Informasi yang dihasilkan diperoleh dengan mengekstrak dan mengenali pola penting atau menarik dari data yang terkandung dalam database (Teori, n.d.). Data mining merupakan suatu kegiatan dengan menggunakan beberapa teknik yang bertujuan untuk mendapatkan informasi dan pengetahuan yang berhubungan dengan database besar data yang ukurannya lebih besar biasanya di olah menggunakan data mining, kemudian dari data tersebut dilakukan pencarian pola atau trend sesuai dengan tujuan dari penerapan data mining tersebut, selanjutnya hasil dari pengolahan data mining tersebut digunakan untuk pengambilan keputusan maupun hasil prediksi analisis yang dibutuhkan (C et al., 2019).

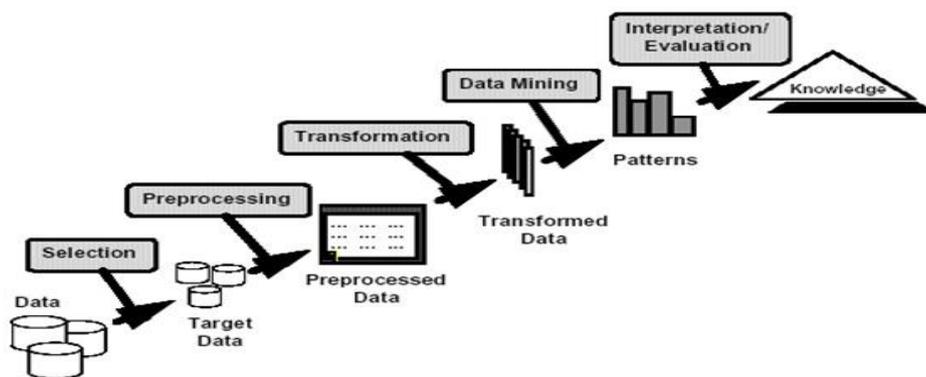
Data mining merupakan kegiatan analisis data dengan memanfaatkan software dan dengan menggunakan tool untuk menemukan pola dengan mengidentifikasi aturan dan karakteristik pada database (Kusumo et al., 2019). Pada

dasarnya, data mining dapat dibagi menjadi dua kategori utama, yaitu (Kunang et al., 2013):

1. Descriptif mining merupakan kegiatan untuk menggali nilai penting dari sebuah database yang tersembunyi dan menemukan pola data tertentu yang belum diketahui sebelumnya.
2. Prediktif, yaitu proses pencarian pola dari data dengan menggunakan beberapa atribut lain di masa akan datang. Klasifikasi termasuk salah satu teknik yang terdapat dalam predictif mining (Defiyanti & Kom, 2013).

2.3 Tahap-Tahap Data mining

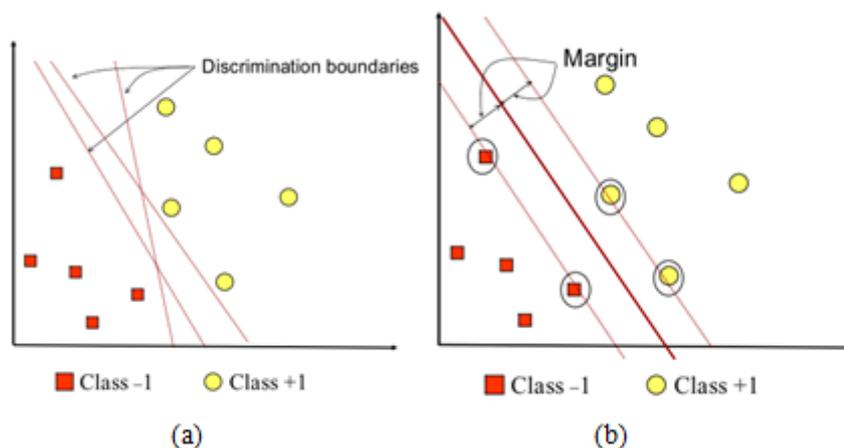
Dalam sejumlah proses, data mining dapat dibagi menjadi beberapa fase (Sitanggang, 2017) :



Gambar 1 Tahap-tahap Data Mining

2.4 Support Vector Machine (SVM)

Model algoritma SVM merupakan salah satu algoritma dari metode klasifikasi, yang bekerja dengan cara mencari suatu garis (hyperlane) untuk memisahkan dua kelompok data. Berikut ini adalah contoh berdasarkan pada Gambar 1 tentang bagaimana SVM mencoba menemukan hyperplane terbaik untuk memisahkan kelas -1 dan +1:



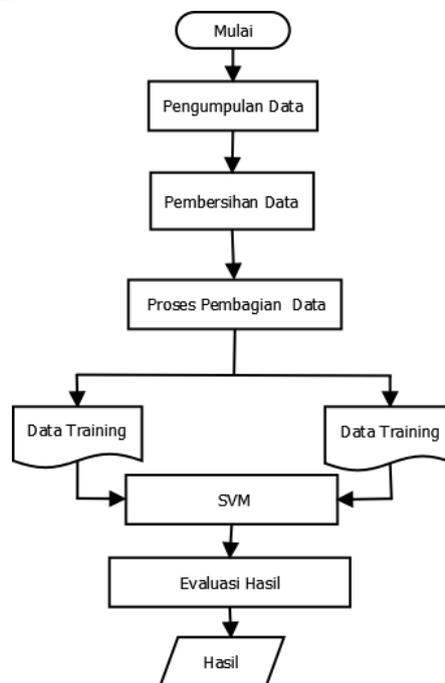
Gambar 2. SVM mencoba menemukan hyperplane terbaik untuk memisahkan kedua class -1 dan +1 (Nugroho et al., 2003).

Hyperplane adalah garis pemisah terbaik antara dua kelas. Untuk mencari hyperplane dapat dilakukan dengan mencari margin hyperplane dan mencari titik maksimum. Margin adalah jarak antara data terdekat di antara dua kelas yang berbeda, yang disebut dengan support vektor(Permana & Sahara, 2019). Garis solid pada gambar 1-b menunjukkan hyperlane yang terbaik, karena terletak tepat diantara kedua class, sedangkan support vector dilambangkan dengan titik merah dan kuning yang berada didalam lingkaran hitam.

3. Metodologi Penelitian

3.1 Prosedur Penelitian

Metode yang digunakan dalam penelitian ini yaitu model algoritma SVM. Prosedur penelitian yang akan dilakukan pada penelitian ini dapat dilihat pada gambar 3.1 di bawah ini:



Gambar 3 Prosedur penelitian

Keterangan dari prosedur penelitian di atas adalah:

a. Pengumpulan Data

1) Studi pustaka

Peneliti mengumpulkan berbagai referensi dan literatur pendukung penelitian yang berkaitan dengan penggunaan *SVM* dalam memprediksikan kelulusan jalur masuk perguruan tinggi, literatur dapat berupa buku tesk, jurnal dan artikel dan karya ilmiah lainnya.

2) Kuesioner

Kuisisioner adalah teknik pengumpulan data yang menggunakan serangkaian pertanyaan tertulis yang di berikan kepada responden untuk dijawab[5]. Dalam penelitian ini kuisisioner di buat menggunakan *tool google form* yang dibagikan baik secara langsung ataupun melalui sosial media. Kuisisioner terdiri dari 22 variabel serta jawaban yang sudah disediakan oleh peneliti untuk dipilih oleh responden.

b. Proses Pembersihan Data

Data yang diperoleh dari hasil pembagian kuisisioner sebanyak 826 data harus melalui beberapa tahap pemrosesan data awal. Tahap awal dari prepossessing adalah proses pembersihan data. Untuk mendapatkan data berkualitas, beberapa tahapan dapat dilakukan selama pemrosesan data, yaitu (Antar et al., 2017):

1. *Data validation*

Proses *data validation* digunakan untuk mengidentifikasi data membersihkan data, menghapus data ganjil (*outlier/noise*), Normalisasi data mengoreksi data nilai yang hilang atau data yang tidak lengkap informasinya, seperti tidak masuknya asal_sek, nilai UN, prestasi_sek, jalur_lulus.

2. Pemilihan data

Pengambilan dan pemilihan data sesuai dengan kebutuhan penelitian.

3. Tranformasi data

Tranformasi data merupakan proses pengubahan data ke dalam format atau bentuk *Arff* sesuai dengan format perangkat lunak *WEKA*. Data yang di dapatkan melalui microsoft excel dirobah ke dalam format CSV “,” atau data yang menggunakan tanda pemisah dengan koma selanjutnya di robah kedalam format *Arff* melalui *WEKA*.

Setelah melalui proses pembersihan data, data akhir yang didapatkan dalam penenelitian ini berjumlah 809 data, kemudian data tersebut di bagi menjadi dua yaitu data training dan data testing.

c. Proses Pembagian Data

Pada penelitian ini untuk proses pemebagian data menggunakan model pengujian, jenis pengujian yang digunakan adalah *cross validation* dan *percentage split* [11]. Berikut penjelasannya.

1) *Cross-Validation*

Pada *cross-validation*, akan ada inputan user untuk memasukkan nilai fold yang digunakan. Didalam *weka* nilai defaultnya ialah 10. Pada penelitian ini, nilai fold yang digunakan adalah 5 dan 10.

2) *Percentage Split*

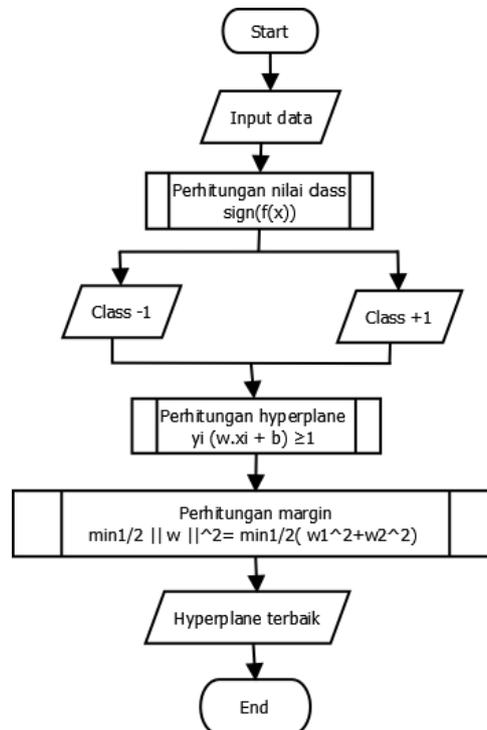
Percentage split merupakan hasil testing dengan menggunakan k% dari data yang dimasukkan. Dimana nilai k adalah input dari user. Pada penelitian ini, percentage

split yang digunakan adalah 20%, 30%, 40%, dan 50% dengan hasil uji pengukuran evaluasi yang digunakan ialah precision, recall dan f-measure.

d. SVM

SVM merupakan salah satu algoritma dari metode klasifikasi, yang bekerja dengan cara mencari suatu garis (*hyperplane*) terbaik untuk memisahkan dua kelompok data. Berikut tahapan proses kerja SVM:

Tahapan pertama yaitu menginput data, kemudian data tersebut akan dicari nilai kelasnya yaitu kelas +1 (negatif) dan kelas -1 (positif) dengan menggunakan rumus persamaan (1), setelah diketahui kelasnya kemudian akan dicari dua data dari kelas yang berbeda yang memiliki jarak terdekat (*support vector*) selanjutnya dari dua data tersebut dicari garis pemisah (*hyperplane*) dengan menggunakan rumus persamaan (2), dan terakhir di cari jarak maksimum dari *hyperplane* dengan *support vector* (*margin*) dengan menggunakan persamaan 5 untuk mengetahui *hyperplane* terbaik dari dua data tersebut. Berikut flowchart proses kerja SVM:



Gambar 4. Flowchart SVM

4. Hasil Dan Pembahasan

4.1 Uji Korelasi

Dalam penelitian ini jalur lulus menjadi label akhir untuk menentukan kelulusan mahasiswa di jalur masuk perguruan tinggi. Dalam menetapkan interval kategori kekuatan korelasi penetapannya sebagai berikut :

Tabel 1. Kategori Kekuatan Korelasi

0	Tidak ada korelasi
0,00-0,25	Korelasi sangat lemah
0,25-0,50	Korelasi cukup
0,50-0,75	Korelasi kuat
0,75-0,99	Korelasi sangat kuat
1	Korelasi sempurna

Dari kategori di atas dapat disimpulkan bahwa jika hubungan suatu variabel tidak sama dengan nol (0), maka dapat dikatakan terjadi hubungan pada variabel tersebut. Hasilnya dapat dilihat pada gambar berikut:

Tabel 2. Perbandingan Korelasi Variabel

Variabel	Sig	Pearson Correlation	Keterangan
Jalur minat	0,000	0,311**	Korelasi Cukup
Bimbel seleksi masuk PTN	0,002	-0,180**	Korelasi Sangat lemah
Nilai UN	0,684	-0,014	Korelasi Sangat Lemah
Orgnisasi di SMA	0,304	0,036	Korelasi Sangat Lemah
Les private di SMA	0,607	-0,018	Korelasi Sangat Lemah
Pekerjaan ayah	0,015	0,086*	Korelasi Sangat Lemah
Pekerjaan ibu	0,203	0,045	Korelasi Sangat Lemah
Pendidikan terakhir ayah	0,497	0,024	Korelasi Sangat Lemah
Pendidikan terakhir ibu	0,667	-0,015	Korelasi Sangat Lemah
Orang tua yang dimiliki sekarang	0,165	0,049	Korelasi Sangat Lemah

Setelah melakukan perhitungan di atas dapat diketahui bahwa variabel yang memiliki hubungan positif (artinya variabel X bernilai tinggi dan variabel Y juga bernilai tinggi) dengan variabel jalur lulus yaitu jalur minat, organisasi di SMA, pekerjaan ayah, pekerjaan ibu, pendidikan terakhir ayah dan orang tua yang dimiliki

sekarang, sedangkan yang berhubungan negatif (artinya variabel X bernilai tinggi maka variabel Y bernilai rendah) dengan variabel jalur lulus yaitu bimbel seleksi masuk PTN, les private di SMA dan pendidikan terakhir ibu. Maka dapat disimpulkan bahwa terdapat dua variabel yang memiliki hubungan korelasi yang lebih baik daripada variabel lain yaitu variabel jalur minat dan variabel bimbel seleksi masuk PTN, jalur minat memiliki korelasi sebesar $0,311^{**}$ ($r=0,311$) dan bimbel seleksi masuk memiliki korelasi sebesar $-0,180^{**}$ ($r= 0,180$).

4.2 Pengujian Klasifikasi

Pada pengujian ini, peneliti menggunakan aplikasi weka untuk menguji keakuratan dari algoritma *support vektor machine*.

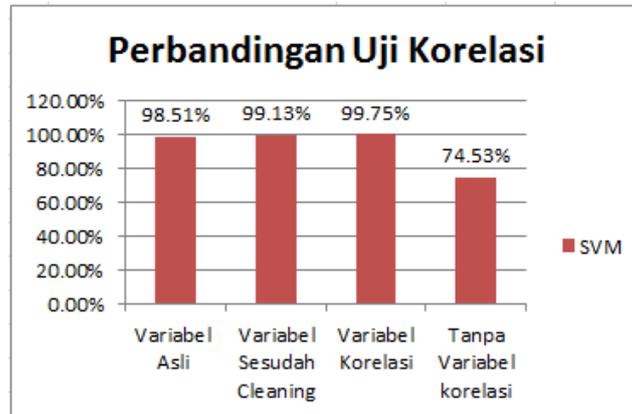
1. Pengujian Korelasi Variabel

Pada penelitian ini peneliti ingin menguji hasil akurasi antara variabel asli dengan variabel yang sudah di *cleaning* dan korelasi.

Tabel 3. Perbandingan Uji Korelasi Variabel

Pengujian	Variabel			
	Variabel Asli	Variabel Sesudah Cleaning	Variabel Korelasi	Tanpa Variabel korelasi
SVM	98.51 %	99.13 %	99.75%	74.53%

Berdasarkan tabel di atas dapat disimpulkan bahwa nilai akurasi variabel asli dengan variabel yang sudah di *cleaning* tidak jauh berbeda karena di kedua variabel tersebut sama-sama terdapat variabel yang sangat berhubungan dengan variabel class, sedangkan antara variabel berkorelasi dengan tanpa variabel korelasi memiliki perbedaan akurasi yang sangat signifikan, dikarenakan pada variabel korelasi tersebut hanya terdapat variabel yang sangat berpengaruh dengan classnya, sedangkan pada tanpa variabel korelasi sama sekali tidak terdapat variabel yang berhubungan dengan variabel class. Berikut adalah grafik dari tabel di atas:



Gambar 5. Grafik Perbandingan Akurasi pada variabel asli, Variabel sesudah *cleaning*, Variabel Korelasi dan tanpa Variabel Korelasi

Dari tabel di atas dapat disimpulkan bahwa variabel korelasi memiliki nilai akurasi yang lebih tinggi dibandingkan dengan variabel lainnya, karena variabel tersebut memiliki hubungan korelasi yang baik dengan variabel class label.

4.3 Waktu Konsumsi

Berikut ini merupakan waktu konsumsi pada SVM berdasarkan pengujian:

Tabel 4. Waktu Konsumsi

Percentage Split	Waktu Konsumsi	Cross Validation	Waktu konsumsi
20%	0.02	5 Fold	0.5
30%	0.02		
40%	0.02	10 Fold	0.5
50%	0.01		

Dari hasil waktu konsumsi antara dua pengujian tersebut dapat disimpulkan pengujian *percentage split* sangat cocok digunakan dalam memprediksi dengan algoritma SVM karena dapat memprediksi dengan waktu yang sangat cepat.

5. Kesimpulan Dan Saran

Adapun kesimpulan dari penelitian ini adalah:

1. Dari berbagai cara yang dipersiapkan oleh calon mahasiswa hanya variabel bimbel yang bisa menjamin calon mahasiswa untuk lulus di jalur yang diminati dengan nilai korelasi sebesar $-0,180^{**}$ dan nilai signifikansi sebesar 0,002.
2. Dari hasil implementasi algoritma SVM dalam memprediksi kelulusan calon mahasiswa di setiap seleksi jalur masuk perguruan tinggi berdasarkan hasil evaluasi dan validasi dapat disimpulkan bahwa algoritma SVM memiliki akurasi dan kinerja sangat baik yaitu rata-rata sebesar 99% dengan nilai AUC (*Area Under Curve*) sebesar 0.9907 dan waktu konsumsi yang dibutuhkan hanya sebesar 0.01-0.02 *second*.

Saran yang dapat diberikan untuk pengembangan penelitian selanjutnya, sebagai berikut:

1. Dalam penelitian ini hanya menggunakan aplikasi machine learning weka sebagai alat bantu, diharapkan pada penelitian selanjutnya, dapat melakukan penelitian menggunakan aplikasi tambahan untuk mendukung dan memperkuat hasil akurasi dari klasifikasi dokumen.
2. Untuk pengembangan selanjutnya bisa melakukan perbandingan dengan berbagai algoritma, dan menggunakan alat pengujian dengan nilai yang lebih tinggi.

Daftar Pustaka

- Abbas, P. D. S. 2008. (n.d.). *MANAGEMEN PERGURUAN TINGGI 2008*.
- Abdullah, F. (2018). *TINGGI AGAMA ISLAM NEGERI DI SULAWESI SELATAN Fahri Abdullah Institut Agama Islam Negeri Bone , Indonesia. 17, 683–693*.
- Antar, S., Vol, B., No, V. I., & Supriyadi, E. (2017). *Metode SVM Berbasis PSO untuk Meningkatkan Prediksi Ketepatan Waktu Kelulusan Mahasiswa. 2, 113–120*.
- C, D. A., Bayes, N., & Svm, D. A. N. (2019). *PERBANDINGAN METODE DATA MINING UNTUK PREDIKSI NILAI DAN WAKTU KELULUSAN MAHASISWA PRODI TEKNIK INFORMATIKA. 13(1), 16–25*.
- Defiyanti, S., & Kom, M. (2013). *Analisis dan Prediksi Kinerja Mahasiswa Menggunakan Teknik Data Mining. 2, 1–8*.
- HARDIYANTI, W. F. (2019). *KLASIFIKASI DOKUMEN SITUS WEB BERITA ONLINE TENTANG KECELAKAAN LALU LINTAS MENGGUNAKAN NAÏVE BAYES CLASSIFIER (NBC) DAN SUPPORT VECTOR MACHINE (SVM)*.
- Kunang, Y. N., Andri, & Murniati, S. (2013). *Implementasi Teknik Data Mining Untuk Memprediksi Tingkat. 2013(A-56-A-63), 1–8*.
- Kusumo, D. S., Bijaksana, M. A., & Darmantoro, D. (2019). *Data Mining Dengan Algoritma Apriori Pada Rdbms Oracle. TEKTRIKA - Jurnal Penelitian Dan Pengembangan Telekomunikasi, Kendali, Komputer, Elektrik, Dan Elektronika, 8(1), 1–5. https://doi.org/10.25124/tektrika.v8i1.215*
- Nugroho, A. S., Witarto, A. B., & Handoko, D. (2003). *Support Vector Machine*.
- Permana, R. A., & Sahara, S. (2019). *Metode Support Vector Machine Sebagai Penentu Kelulusan Mahasiswa pada Pembelajaran Elektronik. VII(1), 50–58*.
- Salinggih, R. fajar. (2019). *IMPLEMENTASI NAIVE BAYES UNTUK MEMPREDIKSI CALON MAHASISWA BARU YANG AKAN LANJUT MELAKUKAN PEMBAYARAN ATAU TIDAK (Studi Kasus : PMB UIN Bandung). 2019*.
- SITANGGANG, R. (2017). *Analisis Kinerja Metode Naïve Bayes dan SSVM untuk Menentukan Pola Kelompok Penyakit*.
- Sitanggang, R. U. S. U. (2017). *Analisis Kinerja Metode Naïve Bayes dan SSVM untuk Menentukan Pola Kelompok Penyakit*.
- Syarli, S., & Muin, A. A. (2016). *Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi). Jurnal Ilmu Komputer, 2(1), 22–26*.

- Teori, D. (n.d.). *PENGGUNAAN METODE NAÏVE BAYES CLASSIFIER PADA APLIKASI PERPUSTAKAAN*.
- Abbas, P. D. S. 2008. (n.d.). *MANAGEMEN PERGURUAN TINGGI 2008*.
- Abdullah, F. (2018). *TINGGI AGAMA ISLAM NEGERI DI SULAWESI SELATAN Fahri Abdullah Institut Agama Islam Negeri Bone , Indonesia. 17, 683–693*.
- Antar, S., Vol, B., No, V. I., & Supriyadi, E. (2017). *Metode SVM Berbasis PSO untuk Meningkatkan Prediksi Ketepatan Waktu Kelulusan Mahasiswa. 2, 113–120*.
- C, D. A., Bayes, N., & Svm, D. A. N. (2019). *PERBANDINGAN METODE DATA MINING UNTUK PREDIKSI NILAI DAN WAKTU KELULUSAN MAHASISWA PRODI TEKNIK INFORMATIKA. 13(1), 16–25*.
- Defiyanti, S., & Kom, M. (2013). *Analisis dan Prediksi Kinerja Mahasiswa Menggunakan Teknik Data Mining. 2, 1–8*.
- HARDIYANTI, W. F. (2019). *KLASIFIKASI DOKUMEN SITUS WEB BERITA ONLINE TENTANG KECELAKAAN LALU LINTAS MENGGUNAKAN NAÏVE BAYES CLASSIFIER (NBC) DAN SUPPORT VECTOR MACHINE (SVM)*.
- Kunang, Y. N., Andri, & Murniati, S. (2013). *Implementasi Teknik Data Mining Untuk Memprediksi Tingkat. 2013(A-56-A-63), 1–8*.
- Kusumo, D. S., Bijaksana, M. A., & Darmantoro, D. (2019). *Data Mining Dengan Algoritma Apriori Pada Rdbms Oracle. TEKTRIKA - Jurnal Penelitian Dan Pengembangan Telekomunikasi, Kendali, Komputer, Elektrik, Dan Elektronika, 8(1), 1–5. <https://doi.org/10.25124/tektrika.v8i1.215>*
- Nugroho, A. S., Witarto, A. B., & Handoko, D. (2003). *Support Vector Machine*.
- Permana, R. A., & Sahara, S. (2019). *Metode Support Vector Machine Sebagai Penentu Kelulusan Mahasiswa pada Pembelajaran Elektronik. VII(1), 50–58*.
- Salinggih, R. fajar. (2019). *IMPLEMENTASI NAIVE BAYES UNTUK MEMPREDIKSI CALON MAHASISWA BARU YANG AKAN LANJUT MELAKUKAN PEMBAYARAN ATAU TIDAK (Studi Kasus : PMB UIN Bandung). 2019*.
- SITANGGANG, R. (2017). *Analisis Kinerja Metode Naïve Bayes dan SSVM untuk Menentukan Pola Kelompok Penyakit*.
- Sitanggang, R. U. S. U. (2017). *Analisis Kinerja Metode Naïve Bayes dan SSVM untuk Menentukan Pola Kelompok Penyakit*.
- Syarli, S., & Muin, A. A. (2016). *Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi). Jurnal Ilmu Komputer, 2(1), 22–26*.
- Teori, D. (n.d.). *PENGGUNAAN METODE NAÏVE BAYES CLASSIFIER PADA APLIKASI PERPUSTAKAAN*.