

## Perbandingan Akurasi Algoritma Data Mining dalam Memprediksi Kelulusan Tepat Waktu

Desy Iba Ricoida<sup>1</sup>, Dedy Hermanto<sup>2</sup>, Rusbandi<sup>3</sup>, Desi Pibriana<sup>4</sup>, M. Rizky Pribadi<sup>5</sup>

<sup>1,2,3,4,5</sup>Fakultas Ilmu Komputer dan Rekayasa, Universitas Multi Data Palembang  
email: [desih@mdp.ac.id](mailto:desih@mdp.ac.id)

**Abstract:** On-time graduation is a crucial benchmark for a university to get accreditation scores. To graduate on time, a student must complete their degree within four years, or less than four years if they are at the Strata-1 level. This study utilizes a dataset acquired from a university, containing data from the 2015-2019 academic years. The dataset consists of 1307 rows of data. In this research, 26 attributes were utilized, including entry\_year, college\_time, gender, school\_type, major, Social Studies 1-10, SKS 1-10, and status. The algorithms employed in this study are decision trees, naive Bayes, logistic regression, KNN, and random forest. The findings of this study indicate that the random forest method achieved the most excellent accuracy level of 90.88%. Additionally, the AUC results obtained were 97.2%, and the F1-Score, calculated from the precision and recall values, reached 89.9%, the most among the four other algorithms. The decision tree and logistic regression methods exhibit accuracy rates of 89.12% and 89.47%, respectively. The logistic regression model has a more excellent accuracy value, while the decision tree model outperforms logistic regression in F1-Score, specifically achieving 88.7% compared to 87.6%.

**Keywords:** University, On Time Graduation, Data Mining, Algoritma

**Abstrak:** Lulusan tepat waktu menjadi salah satu poin penilaian sangat penting bagi sebuah perguruan tinggi untuk memperoleh nilai akreditasi. Dikatakan lulusan tepat waktu jika seorang mahasiswa dapat lulus empat tahun atau dibawah empat tahun jika berada pada jenjang Strata-1. Penelitian ini menggunakan dataset yang diperoleh dari universitas dengan data dari angkatan 2015-2019, dimana total data yang digunakan yaitu sebanyak 1307 baris. Sebanyak 26 atribut yang digunakan dalam penelitian ini yaitu tahun\_masuk, waktu\_kuliah, jenis\_kelamin, tipe\_sekolah, jurusan, IPS 1-10, SKS 1-10 dan status. Algoritma yang digunakan dalam penelitian ini yaitu decision tree, naive bayes, logistic regression, KNN dan random forest. Hasil yang diperoleh dalam penelitian ini yaitu algoritma random forest memiliki tingkat akurasi yang paling tinggi sebesar 90.88% dengan hasil dari AUC yang diperoleh yaitu sebesar 97.2% dan perhitungan F1-Score dari hasil nilai precision dan recall diperoleh sebesar 89.9%, tertinggi dari empat algoritma lainnya. Sedangkan untuk algoritma decision tree dan logistic regression memiliki nilai akurasi masing-masing yaitu sebesar 89.12% dan 89.47%. Nilai dari logistic regressing lebih tinggi untuk akurasi, akan tetapi untuk nilai F1-Score decision tree lebih baik dari logistic regression yaitu 88.7% berbanding 87.6%.

**Kata kunci:** Universitas, Lulusan Tepat Waktu, Data Mining, Algoritma

### Pendahuluan

Keberhasilan sebuah perguruan tinggi untuk dapat meluluskan mahasiswanya menjadi sebuah indikator yang disertakan dalam penilaian oleh BAN-PT (Badan Akreditasi Nasional - Perguruan Tinggi). Hal tersebut berkaitan dengan penilaian yang digunakan oleh perguruan tinggi untuk menjadi tolak ukur keberhasilan melaksanakan proses belajar dan

mengajar (BAN-PT, n.d.). Lulusan tepat waktu menjadi satu hal yang menjadi perhatian oleh perguruan tinggi karena menjadi poin terhadap penilaian yang cukup besar (BAN-PT, 2018, 2019).

Akreditasi perguruan tinggi menjadi salah satu penilaian masyarakat untuk melihat sebuah perguruan tinggi lebih baik dari perguruan tinggi lain. Sehingga

berdasarkan hal tersebut peneliti tertarik untuk mengambil tema tentang lulusan tepat waktu. Berdasarkan penelitian terdahulu yang telah dilakukan tentang klasifikasi ketepatan suatu mahasiswa memperoleh kelulusan tepat waktu dimana proses kuliah dikatakan tepat waktu yaitu antara 4 tahun atau kurang dari 4 tahun untuk jenjang S1. Penelitian ini memanfaatkan metode Naive Bayes (Rahayu et al., 2021). Penelitian lain yaitu pemanfaatan data induk mahasiswa untuk memperoleh informasi kelulusan tepat waktu di perguruan tinggi Universitas Stikubank (Jananto et al., 2021).

Berdasarkan hal tersebut peneliti akan melakukan penelitian tentang kelulusan tepat waktu dengan menggunakan algoritma yang terdapat dalam data mining yaitu Decision Tree, Naive Bayes, Logistic Regression, KNN dan Random Forest. Hal ini dilakukan agar membedakan penelitian yang dilakukan dengan penelitian sebelumnya. Setelah hasil diperoleh, maka akan dilakukan perbandingan hasil yang diperoleh dari setiap nilai yang diperoleh dari setiap algoritma yang digunakan.

## Landasan Teori

### Decision Tree (DT)

Algoritma klasifikasi yang memiliki 3 buah node yaitu root node, internal node dan leaf node, dimana untuk memperoleh root node diperoleh dari nilai gain tertinggi dari atribut yang digunakan (Batubara et al., 2022). Proses perhitungan menggunakan Decision Tree menggunakan Entropy (S) dan Gain (S,A).

$$Entropy(S) = \sum_{i=1}^n -p_i \cdot \log_2 p_i$$

Keterangan:

S = himpunan kasus

n = jumlah partisi S

$p_i$  = proporsi  $S_i$  terhadap S

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Keterangan:

S = Himpunan kasus

A = Fitur

n = jumlah partisi atribut A

$|S_i|$  = proporsi  $S_i$  terhadap S

$|S|$  = jumlah kasus dalam S

### Naive Bayes (NB)

Sebuah metode yang digunakan untuk menentukan suatu keputusan menggunakan pohon keputusan dimana proses yang dilakukan yaitu melakukan klasifikasi dengan mencari probabilitas dan nilai yang dihasilkan pada saat perhitungan (Situmorang & Sirait, 2020). Rumus umum yang digunakan dalam Naive Bayes yaitu:

$$P(A | B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Keterangan:

$P(A|B)$  = Probabilitas posteriori dari B

$P(A)$  = Apriori dari A

$P(B)$  = Apriori dari B

### Logistic Regression (LR)

Algoritma klasifikasi yang memiliki fungsi logistik yang digunakan untuk mengambil nilai diskrit dari sebuah kelas yang akan digunakan dalam sebuah observasi (Veluchamy et al., 2018).

### K-Nearest Neighbor (KNN)

Sebuah metode klasifikasi yang menggunakan data latih melalui jarak paling dekat terhadap sebuah objek atau biasa disebut dengan tetangga (neighbor) (Manullang et al., 2021).

### Random Forest (RF)

Metode ini merupakan bagian metode klasifikasi dengan teknik bootstrap aggregating (bagging) dimana proses dalam menghasilkan pohon melalui pengambilan bagian data dari dataset (Nurhikam et al., 2023).

**Penelitian Terdahulu**

Penelitian yang dilakukan, peneliti menggunakan tahapan yang dimulai dengan inisiasi dan pengumpulan masalah melalui sebuah studi literatur yaitu dilakukan dengan proses mengumpulkan penelitian sebelumnya yang berhubungan dengan tema yang akan diambil dalam penelitian. Proses ini tersaji pada Tabel 1.

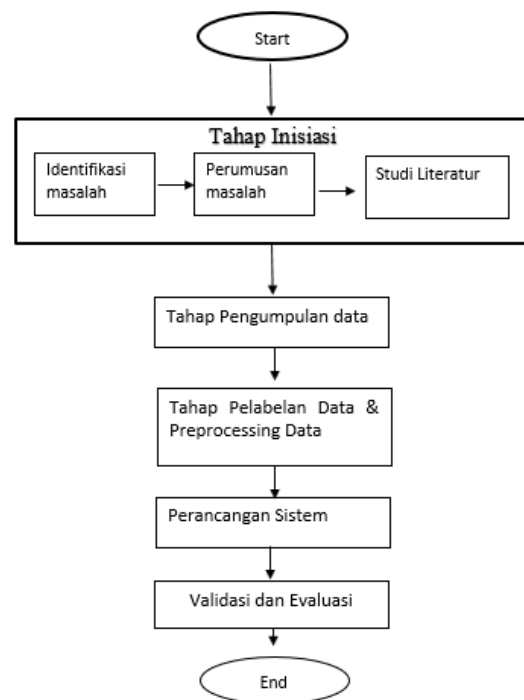
Tabel 1. Algoritma dalam Penelitian

Ref	DT	NB	LR	KNN	RF
(Rahayu et al., 2021)		✓			
(Jananto et al., 2021)	✓				
(Rahmayanti et al., 2022)	✓	✓			
(Yatimah, 2021)	✓	✓			
(Endang Etriyanti, 2021)	✓			✓	
(Nurhikam et al., 2023)			✓		
(Puspa, 2021)			✓		
(Hasibuan & Mahdiana, 2023)	✓				
(Qisthiano et al., 2023)	✓				
(Gunawan et al., 2021)		✓		✓	
Ours	✓	✓	✓	✓	✓

Dari penelitian terdahulu yang dilakukan bahwa dapat disimpulkan dalam penelitian ini, peneliti melakukan penelitian tentang lulusan tepat waktu dengan memanfaatkan algoritma klasifikasi yaitu Decision Tree, Naive Bayes, Logistic Regression, KNN dan Random Forest. Hasil selanjutnya akan membandingkan algoritma yang terbaik antara 5 algoritma klasifikasi yang digunakan.

**Metodologi**

Tahapan awal yang dilakukan peneliti dalam penelitian ini untuk memperoleh permasalahan yang diangkat dalam penelitian ini menggunakan tema Lulusan Tepat Waktu telah dilakukan sesuai dengan Tabel 1. Beberapa tahapan selanjutnya yang akan dilakukan dalam penelitian ini disajikan dalam proses diagram alir penelitian yang tersaji pada Gambar 1.



Gambar 1. Diagram Alir Penelitian

Tahapan yang terdapat pada Gambar 1, dilaksanakan dalam penelitian ini. Untuk tahapan pertama pernyataan tentang penelitian dilaksanakan seperti Tabel 1.

**Pengumpulan Data**

Penelitian tentang lulusan tepat waktu menggunakan data yang dikumpulkan dari Universitas Multi Data Palembang. Data yang dikumpulkan adalah mahasiswa dengan tahun masuk 2015 - 2019. Penelitian ini terdapat 26 atribut yang digunakan. Beberapa atribut tersebut yaitu tahun\_masuk merupakan data yang diambil dari data mahasiswa ketika memasuki perkuliahan,

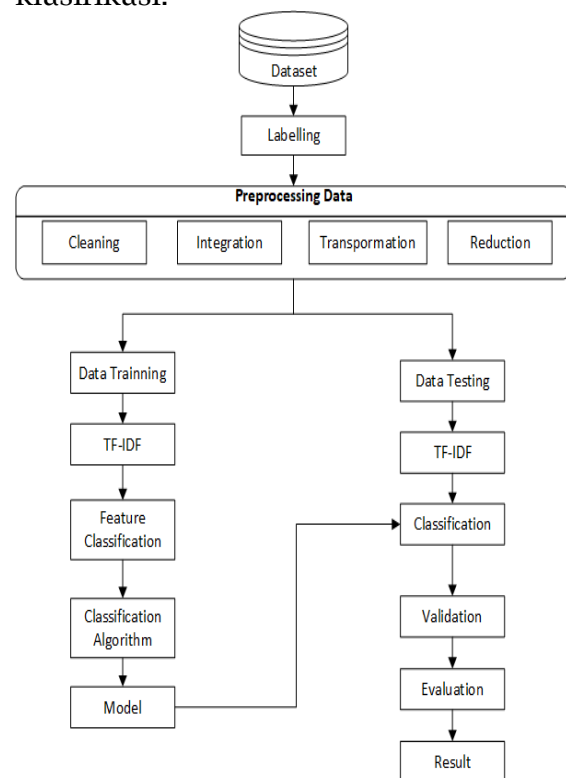
waktu\_kuliah merupakan data yang diambil berupa pembagian kuliah dengan jadwal pagi atau sore, jenis\_kelamin merupakan data yang dibedakan atas jenis kelamin laki-laki atau perempuan, jenis\_sekolah merupakan data yang diambil dari jenis sekolah yang ada di Indonesia yaitu SMA, SMK atau MA. Jurusan merupakan data yang tersedia di lokasi data penelitian yaitu terdiri atas program studi Sistem Informasi, Teknik Informatika, Teknik Elektro, Manajemen dan Akuntansi, IPS 1 - 10 merupakan data Indeks Prestasi setiap semester dari mahasiswa yang digunakan sebagai objek penelitian, SKS 1 -10 merupakan data yang diambil dari jumlah mata kuliah yang dapat diambil dengan total maksimal 24 SKS. Data ini diambil dari data konversi nilai IPS yang diperoleh setiap semester. Atribut terakhir yang digunakan yaitu status. Atribut ini digunakan sebagai label untuk menentukan status mahasiswa apakah lulus atau tidak lulus.

Pra Pemrosesan dan Pengumpulan Data

Tahapan pengumpulan data dilakukan dengan mengumpulkan data dari bidang akademik melalui aplikasi pendukung proses pembelajaran. Total data yang terkumpul yaitu sebanyak 1307 baris. Proses selanjutnya yang dilakukan adalah pembersihan terhadap data yang telah dikumpulkan. Salah satu proses yang dilakukan yaitu membuang data yang semua data terisi hampir semuanya kosong (tidak terisi sama sekali), kemudian mengisi data yang hanya bagian tertentu saja tidak terisi dengan benar. Proses yang dilakukan yaitu dengan mengambil nilai rata-rata dari data yang dimiliki yang telah dikumpulkan dalam proses pengumpulan data.

Perancangan Sistem

Tahap ini yang dilakukan adalah dimulai dengan melakukan proses penilaian dan pengambilan bobot dari suatu atribut yang akan digunakan dalam penelitian. Tahapan proses dalam perancangan sistem yang dilakukan dalam penelitian ini tersaji pada Gambar 2. Hasil yang akan diperoleh yaitu berupa model yang akan digunakan dalam pengujian data yang diperoleh selanjutnya dalam penelitian untuk melakukan sebuah klasifikasi.



Gambar 2. Perancangan Sistem

Validasi dan Evaluasi

Tahap ini merupakan proses validasi dan evaluasi. Hasil awal diperoleh sebuah confusion matrix yang terdiri atas nilai true positive, false positive, true negative dan false negative. Hasil yang diperoleh selanjutnya dalam penelitian ini yaitu berupa Accuracy (Acc), Precision (Pr) dan Recall (Rc). Setelah hasil tersebut diperoleh maka dilakukan perhitungan F1-Score (F1). Tabel untuk confusion matrix tersaji pada Tabel 2.

Tabel 2. Confusion Matrix

		Predicted Values	
		Positive	Negative
Actual Values	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Keterangan:

TP : Nilai yang berisi positif

FP : Nilai yang berisi positif tapi hasil salah

FN : Nilai yang berisi negatif, tetapi nilai yang diberikan salah

TN : Nilai yang berisi tidak benar

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision+Recall}$$

### Hasil dan Pembahasan

Proses pengujian data yang telah dikumpulkan sebelumnya dalam penelitian ini dibagi menjadi dua bagian, yaitu 80% data sebagai data latih dan 20% data digunakan sebagai data Uji. Menggunakan metode klasifikasi yang terpilih yaitu DT, NB, LR, KNN dan RF, peneliti memperoleh hasil seperti tersaji pada Tabel 3.

Terdapat lima kolom hasil yang diperoleh dari masing-masing algoritma yang digunakan yaitu accuracy, precision, recall, AUC (Area Under Curve) dan F1-Score.

accuracy: 89.12%

	true 0	true 1	class precision
pred. 0	132	1	99.25%
pred. 1	30	122	80.26%
class recall	81.48%	99.19%	

Gambar 3. Accuracy Algoritma DT

Tabel 3. Hasil Pengujian

Klasifikasi	Acc	Pr	Rc	AUC	F1
DT	0.8912	0.8026	0.9919	0.951	0.887
NB	0.7965	0.9114	0.5854	0.927	0.713
LR	0.8947	0.8908	0.8618	0.951	0.876
KNN	0.8772	0.8607	0.8573	0.957	0.859
RF	0.9088	0.8593	0.9431	0.972	0.899

Hasil yang disajikan dari Gambar 3 merupakan hasil pengujian yang diperoleh berupa accuracy dari algoritma decision tree. Hasil yang diperoleh yaitu sebesar 89.12%

precision: 80.26% (positive class: 1)

	true 0	true 1	class precision
pred. 0	132	1	99.25%
pred. 1	30	122	80.26%
class recall	81.48%	99.19%	

Gambar 4. Precision Algoritma DT

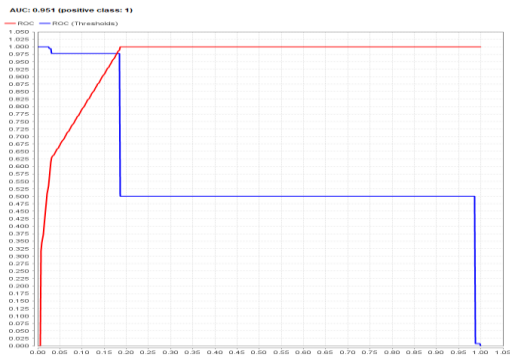
Hasil pengujian selanjutnya yang disajikan pada Gambar 4, merupakan hasil pengujian yang dilakukan menggunakan algoritma decision tree untuk mendapatkan hasil precision. Dari pengujian data yang dikumpulkan, diperoleh nilai precision yaitu sebesar 80.26%.

recall: 99.19% (positive class: 1)

	true 0	true 1	class precision
pred. 0	132	1	99.25%
pred. 1	30	122	80.26%
class recall	81.48%	99.19%	

Gambar 5. Recall Algoritma DT

Untuk hasil dari Gambar 5 merupakan hasil pengujian dari algoritma decision tree untuk pencarian recall. Hasil yang diperoleh yaitu sebesar 99.19%.



Gambar 6. ROC Curve Algoritma DT

Hasil yang ditunjukkan oleh Gambar 6 yaitu merupakan hasil pengujian untuk mendapatkan AUC (area Under Curve) dengan menggunakan algoritma decision tree.

Proses yang sama dilakukan untuk memperoleh hasil seperti Tabel 3 dengan cara mengubah algoritma yang digunakan, agar memperoleh hasil yang berbeda antar algoritma.

Proses selanjutnya setelah didapat hasil berupa accuracy, precision, recall dan auc yaitu melakukan perhitungan terhadap F1-Score. Proses ini dilakukan terhadap 5 algoritma yang berbeda.

$$F1\ Score = 2 \times \frac{0.8026 \times 0.9919}{0.8026 + 0.9919}$$

$$F1\ Score = 2 \times \frac{0.7961}{1.7945}$$

$$F1\ Score = 0.8873$$

Perhitungan F1-Score dari algoritma decision tree, diperoleh hasil seperti perhitungan diatas. Proses yang dilakukan yaitu hasil yang diperoleh sebesar 0.8873 atau sebesar 88.73%

$$F1\ Score = 2 \times \frac{0.9114 \times 0.5854}{0.9114 + 0.5854}$$

$$F1\ Score = 2 \times \frac{0.5335}{1.4968}$$

$$F1\ Score = 0.713$$

Proses perhitungan diatas dilakukan untuk memperoleh hasil dari pengukuran nilai dengan menggunakan algoritma naive bayes. Hasil yang diperoleh dari perhitungan menggunakan algoritma tersebut adalah sebesar 0.713 atau sebesar 71.3%.

$$F1\ Score = 2 \times \frac{0.8908 \times 0.8618}{0.8908 + 0.8618}$$

$$F1\ Score = 2 \times \frac{0.7677}{1.7526}$$

$$F1\ Score = 0.876$$

Pengujian selanjutnya yaitu menggunakan algoritma linier regression. Proses perhitungan yang dilakukan untuk mendapatkan F1-Score disajikan pada perhitungan diatas. Hasil yang diperoleh yaitu sebesar 0.876 atau 87.6%.

$$F1\ Score = 2 \times \frac{0.8607 \times 0.8573}{0.8607 + 0.8573}$$

$$F1\ Score = 2 \times \frac{0.7379}{1.718}$$

$$F1\ Score = 0.859$$

Perhitungan F1-Score untuk algoritma KNN, disajikan pada perhitungan diatas. Hasil yang diperoleh dengan menggunakan algoritma ini yaitu nilai F1-Score sebesar 0.859 atau 85.9%.

$$F1\ Score = 2 \times \frac{0.8593 \times 0.9431}{0.8593 + 0.9431}$$

$$F1\ Score = 2 \times \frac{0.8104}{1.8024}$$

$$F1\ Score = 0.899$$

Perhitungan dengan algoritma yang terakhir yaitu algoritma random forest untuk memperoleh nilai F1-Score disajikan pada perhitungan diatas. Hasil yang diperoleh yaitu sebesar 0.899 atau sebesar 89.9%.

## Kesimpulan

Menurut hasil pengujian yang telah dilakukan yaitu algoritma random forest memiliki tingkat akurasi yang paling tinggi yaitu sebesar 90.88% dengan hasil dari AUC yang diperoleh yaitu sebesar 97.2% dan perhitungan F1-Score dari hasil nilai precision dan recall diperoleh sebesar 89.9%, tertinggi dari empat algoritma lainnya. Sedang untuk algoritma decision tree dan logistic regression memiliki nilai akurasi masing-masing yaitu sebesar 89.12% dan 89.47%. Nilai dari logistic regression lebih tinggi untuk akurasi, akan tetapi untuk nilai F1-Score decision tree lebih baik dari logistic regression yaitu 88.7% berbanding 87.6%.

Untuk penelitian selanjutnya dapat menggunakan algoritma yang berbeda yang tersedia baik menggunakan data mining atau menggunakan pembelajaran mesin.

## Referensi

- BAN-PT. (n.d.). *Badan Akreditasi Nasional - Perguruan Tinggi*. Retrieved October 20, 2022, from <https://www.banpt.or.id/>
- BAN-PT. (2018). *Pedoman Penilaian IAPT 3.0*.
- BAN-PT. (2019). *Matriks Penilaian Laporan Evaluasi Diri dan Laporan Kinerja Program Studi*.
- Batubara, D. N., Windarto, A. P., & Irawan, E. (2022). Analisis Prediksi Keterlambatan Pembayaran Listrik Menggunakan Komparasi Metode Klasifikasi Decision Tree dan Support Vector Machine. *JURIKOM (Jurnal Riset Komputer)*, 9(1), 102. <https://doi.org/10.30865/jurikom.v9i1.3833>
- Endang Etriyanti. (2021). Perbandingan Tingkat Akurasi Metode Knn Dan Decision Tree Dalam Memprediksi Lama Studi Mahasiswa. *Jurnal Ilmiah Binary STMIK Bina Nusantara Jaya Lubuklinggau*, 3(1), 6-14. <https://doi.org/10.52303/jb.v3i1.40>
- Gunawan, M., Zarlis, M., & Roslina, R. (2021). Analisis Komparasi Algoritma Naïve Bayes dan K-Nearest Neighbor Untuk Memprediksi Kelulusan Mahasiswa Tepat Waktu. *Jurnal Media Informatika Budidarma*, 5(2), 513. <https://doi.org/10.30865/mib.v5i2.2925>
- Hasibuan, T. H., & Mahdiana, D. (2023). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Algoritma C4.5 Pada Uin Syarif Hidayatullah Jakarta. *Skanika: Sistem Komputer Dan Teknik Informatika*, 6(1), 61-74. <https://doi.org/10.36080/skanika.v6i1.2976>
- Jananto, A., Sulastris, S., Nur Wahyudi, E., & Sunardi, S. (2021). Data Induk Mahasiswa sebagai Prediktor Ketepatan Waktu Lulus Menggunakan Algoritma CART Klasifikasi Data Mining. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 10(1), 71-78. <https://doi.org/10.32736/sisfokom.v10i1.991>
- Manullang, R. A., Sianturi, F. A., Penerapan, [, Neighbor, A. K.-N., Memprediksi, U., & Mahasiswa, K. (2021). Penerapan Algoritma K-Nearest Neighbour Untuk Memprediksi Kelulusan Mahasiswa. *JIKOMSI*, 4(2), 42-50.
- Nurhikam, A. S., Syaputra, R., Rohman, S., & ... (2023). Deteksi Berita Palsu Pada Pemilu 2024 Dengan Menggunakan Algoritma Random Forest. *DoubleClick: Journal of Computer and Information Technology*, 7(1), 41-50. <http://e-journal.unipma.ac.id/index.php/doubleclick/article/view/15456%0A> <http://e-journal.unipma.ac.id/index.php/doubleclick/article/download/15456/5393>
- Puspa, S. (2021). Ahmad Yusuf 3) 1), 3)

- Sistem Informasi, Fakultas Sains dan Teknologi, UIN Sunan Ampel, Surabaya 2) Matematika, Fakultas Sains dan Teknologi. *Jurnal Ilmiah NERO*, 6(1), 39-47.
- Qisthiano, M. R., Prayesy, P. A., & Ruswita, I. (2023). Penerapan Algoritma Decision Tree dalam Klasifikasi Data Prediksi Kelulusan Mahasiswa. *G-Tech: Jurnal Teknologi Terapan*, 7(1), 21-28.  
<https://doi.org/10.33379/gtech.v7i1.1850>
- Rahayu, T. M., Ningsi, B. A., Isnurani, & Arofah, I. (2021). Klasifikasi Ketepatan Waktu Kelulusan Mahasiswa Dengan Metode Naive Bayes. *Jurnal Binawakya*, 16(2), 6427-6438.
- Rahmayanti, A., Rusdiana, L., & Suratno, S. (2022). Perbandingan Metode Algoritma C4.5 Dan Naïve Bayes Untuk Memprediksi Kelulusan Mahasiswa. *Walisongo Journal of Information Technology*, 4(1), 11-22.  
<https://doi.org/10.21580/wjit.2022.4.1.9654>
- Situmorang, N., & Sirait, G. (2020). Implementasi Data Mining untuk Memprediksi Kelulusan Siswa dengan Metode Naive Bayes. *Comasie*, 6(2), 107-118.
- Veluchamy, A., Nguyen, H., Diop, M. L., & Iqbal, R. (2018). Comparative Study of Sentiment Analysis with Product Reviews Using Machine Learning and Lexicon-Based Approaches. *SMU Data Science Review*, 1(4), 1-22.  
<https://scholar.smu.edu/cgi/viewcontent.cgi?article=1051&context=datasciencereview>
- Yatimah, M. N. (2021). Implementasi Data Mining untuk Prediksi Kelulusan Tepat Waktu Mahasiswa STIMIK ESQ Menggunakan Decision Tree C4.5. *JUMANJI (Jurnal Masyarakat Informatika Unjani)*, 5(2), 89.  
<https://doi.org/10.26874/jumanji.v5i2.95>