

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Untuk memperjelas posisi penelitian ini, maka disusun suatu tinjauan ulang terhadap penelitian-penelitian sejenis yang telah dilakukan sebelumnya berdasarkan pada pendekatan tinjauan dari kesamaan obyek bahasan maupun dengan kesamaan metode yang digunakan.

2.1.1 Penelitian yang Terkait dengan Prediksi Kinerja Akademik

Pada bagian ini akan dibahas mengenai kepustakaan penelitian terdahulu terkait dengan prediksi kinerja akademik. Penulis mencoba untuk mencari literatur tentang faktor-faktor yang dapat dijadikan variabel independen untuk klasifikasi prediksi kinerja dan keberhasilan mahasiswa yang dibahas dalam penelitian ini. Berikut adalah beberapa jurnal terkait dengan prediksi kinerja akademik.

Penelitian yang dilakukan oleh Sadiq Hussain, Neema Abdulaziz Dahan, Fadl Mutaher Ba-Alwib, Najoua Ribata (2018) berjudul *Educational Data Mining and Analysis of Students' Academic Performance Using WEKA*. Penelitian yang dilakukan peneliti dari Dibrugarh University, India tersebut membahas tentang bagaimana mengidentifikasi kinerja siswa dari tiga sekolah untuk meningkatkan kinerja siswa dan mencegah Drop Out. Penelitian ini menggunakan teknik klasifikasi data mining dengan berbagai macam algoritma. Data rill dari tiga Sekolah di Assam, India diambil untuk dilakukan filter sesuai dengan variable potensial yang diinginkan menggunakan *tool* WEKA. Data terdiri dari social ekonomi, demografi seperti informasi akademik siswa. Dataset informasi akademik siswa dikumpulkan, diuji dan diterapkan pada empat algoritma klasifikasi seperti J48, PART, *Random Forest* dan *Bayes Network Classifiers* dengan *tool* WEKA.

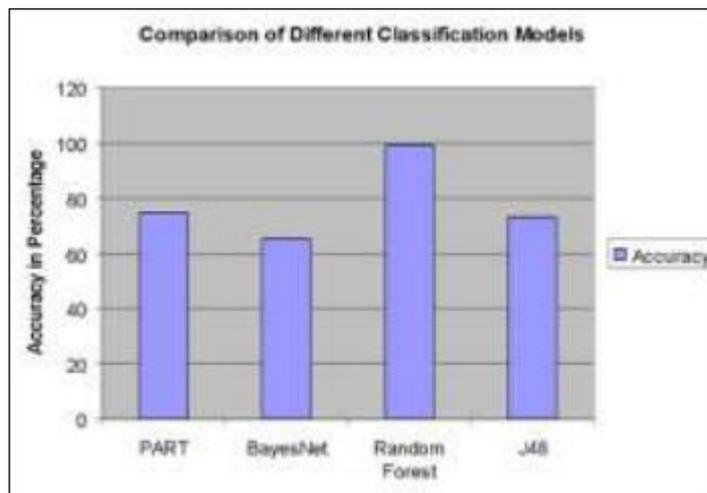
Sebanyak 300 data digunakan untuk analisis penelitian ini. Dalam penelitian ini, pemilihan variable potensial yang berpengaruh untuk klasifikasi menggunakan lima metode seleksi fitur seperti *correlation-based attribute evaluation*, *gain-ratio attribute evaluation*, *information-gain attribute evaluation*, *relief attribute*

evaluation, dan *symmetrical uncertainty attribute evaluation*. Setelah itu dilakukan evaluasi ranking/peringkat variabel sesuai dengan yang paling berpengaruh. Perangkat lunak WEKA memiliki menu Metode Ranker Search Technique untuk evaluasi ranking variabel tersebut. Atribut awal terdiri dari 24 atribut selanjutnya menjadi 22 atribut melalui proses pembersihan data sehingga dari proses seleksi fitur diperoleh sebelas atribut yang paling potensial. Berikut Tabel 2.1 merupakan hasil pemilihan variabel potensial yang sudah dievaluasi dengan metode *Ranker Search Technique*.

Tabel 2. Variabel dengan Potensial Tinggi (Hussain et al., 2018)

Feature Selection Method	High Influence Attributes
Correlation-based Attribute Evaluation	arr, iap, tnp, as, twp, sh, me, fs, nf, atd, fo, fmi, fq, tt, ss
Gain-Ratio Attribute Evaluation	iap, ms, arr, tnp, twp, as, me, sh, atd, fmi, fq, nf, fo, mq, fs
Information-Gain Attribute Evaluation	iap, tnp, twp, arr, fmi, as, fq, me, atd, sh, fo, mq, nf, cst, tt
Relief Attribute Evaluation	iap, tnp, arr, tnp, nf, as, atd, me, fo, sh, fmi, fs, ls, ge, tt
Symmetrical Uncertainty Attribute	iap, tnp, twp, arr, as, me, fmi, atd, sh, fq, fo, mq, nf, fs, tt

Setelah melakukan analisis variabel yang berpotensi mempengaruhi proses klasifikasi, selanjutnya dilakukan teknik klasifikasi dengan berbagai macam algoritma yang sudah disebutkan sebelumnya. Hasil akhir yang didapatkan pada penelitian tersebut berupa tingkat akurasi prediksi berdasarkan semua algoritma klasifikasi dan perbandingannya untuk menemukan kinerja algoritma klasifikasi terbaik di antara semua. Penelitian ini menampilkan pentingnya Prediksi dan Klasifikasi berdasarkan algoritma data mining di bidang pendidikan dan juga menyajikan hasil terbaiknya dari beberapa klasifikasi yang dilakukan. Adapun hasil pengujian akurasi masing-masing algoritma ditampilkan dalam Gambar 2.1.



Gambar 2. 1 Perbandingan klasifikasi (Hussain et al., 2018)

Di antara semua data hasil uji akurasi, metode *Random Forest* memiliki tingkat akurasi tertinggi yakni 99%. Dengan demikian *Random Forest* terbukti menjadi algoritma classifier yang paling efektif dan efisien. Kemudian peneliti melakukan komparasi *Random Forest* dengan seleksi fitur dan tanpa seleksi fitur. Hasil klasifikasi menunjukkan *Random Forest* dengan dua belas atribut yang dipilih memiliki tingkat akurasi sebesar 99% sedangkan *Random Forest* tanpa seleksi fitur (menggunakan semua atribut) memiliki tingkat akurasi sebesar 84,33 %. Dengan demikian *Random Forest* dengan seleksi fitur terbukti memiliki akurasi yang terbaik. Tahap akhir dari penelitian ini, peneliti menerapkan algoritma apriori untuk mencari aturan yang terbaik dalam mengatasi kinerja akademik siswa dengan membuat kluaster yang sesuai dengan masing-masing karakteristik yang sama.

Penelitian yang dilakukan oleh Fadhilah Ahmad, Nur Hafieza Ismail dan Azwa Abdul Aziz (2015) berjudul *The Prediction of Students' Academic Performance Using Classification Data Mining Techniques*. Penelitian yang dilakukan oleh peneliti dari Universiti Sultan Zainal Abidin (UniSZA), Kuala Terengganu, Malaysia ini membahas tentang perbandingan metode untuk prediksi kinerja akademik mahasiswa. Pada awal semester untuk mahasiswa baru, seorang dosen kesulitan untuk mengetahui dan menganalisis kinerja mahasiswanya karena kurangnya informasi tentang latar belakang mahasiswa sebelumnya. Data yang

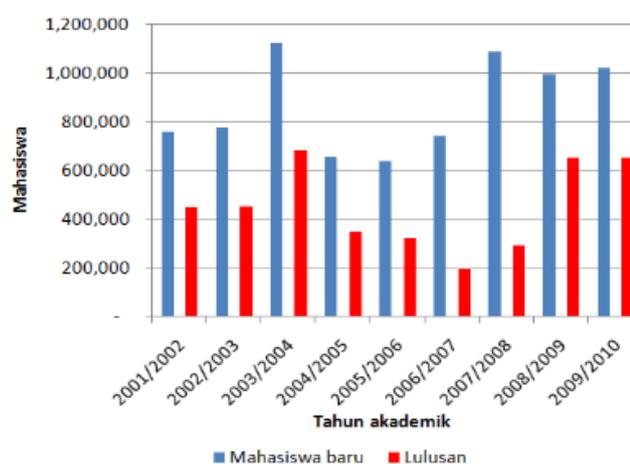
digunakan yaitu data dari Juli 2006/2007 sampai Juli 2013/2014 yang berisi informasi demografi siswa, catatan akademik sebelumnya, dan informasi latar belakang keluarga. Pada penelitian ini, IPK mahasiswa dipilih sebagai parameter dependen. Parameter Independen terdiri dari Sembilan atribut seperti *gender*, *race*, *hometown (ht)*, *GPA*, *family income (fi)*, *university mode entry (ume)*, and *SPM grades in three subjects; Malay Language (bm)*, *English (bi)*, dan *Mathematics (math)*. Dataset yang digunakan hanya 399 dari 497 setelah melalui proses pembersihan data. Selanjutnya dilakukan klasifikasi dengan tiga algoritma seperti *Decision Tree*, *Naïve Bayes*, dan *Rule Based classification*. Hasil eksperimen menunjukkan *Rule Based classification* merupakan model terbaik daripada teknik klasifikasi lainnya dengan nilai akurasi tertinggi sebesar 71,3 %. Model prediksi ini akan digunakan untuk identifikasi dan profil mahasiswa untuk menentukan tingkat keberhasilan mahasiswa di semester pertama.

Penelitian yang dilakukan oleh Satria Abadi, Kamarul Shukri Mat, Badlihisam Mohd Nasir dkk (2018) berjudul *Application model of k-means clustering: insights into promotion strategy of vocational high school* yang dipublikasikan di *International Journal of Engineering & Technology*, membahas strategi promosi untuk mencapai target pada proses penerimaan mahasiswa baru di SMK Muhammadiyah 1 Pringsewu, Lampung, Indonesia. Penelitian ini dilakukan guna meminimalisasi biaya dalam proses pemasaran melalui penentuan strategi promosi yang tepat. Metode yang digunakan yaitu *clustering K-Means* dengan alat bantu WEKA. Atribut yang digunakan meliputi alamat rumah, sekolah asal, transportasi, dan alasan memilih sekolah. Hasil penelitian ini ditemukan bahwa ada dua strategi promosi yang dapat diimplementasikan oleh SMK Muhammadiyah 1 Prongsewu yaitu: 1) melakukan strategi promosi mengundang teman-teman sekolah dengan memberikan insentif kepada masyarakat yang mampu mengajak calon siswa untuk bersekolah di SMK Muhammadiyah 1 Pringsewu; 2) melakukan strategi promosi dengan sosialisasi langsung ke sekolah,

Penelitian yang dilakukan oleh Nur Aminudin dkk (2018) berjudul *Higher education selection using simple additive weighting* yang dipublikasikan di *International Journal of Engineering & Technology*, membahas Sistem pendukung

keputusan untuk menentukan universitas terbaik. Penelitian ini dilakukan karena banyaknya pilihan kriteria perguruan tinggi dan sebagian besar masyarakat bingung memilih perguruan tinggi yang sesuai untuk diri mereka sendiri dan tuntutan pekerjaan. Metode yang digunakan yaitu Simple Additive Weighting (SAW). Kriteria yang digunakan seperti *Academic achievement Master Lecturer Extracurricular*, dan *Accreditation Status Facilities Scholarship*. Hasil dari penelitian ini berupa model sistem pendukung keputusan dengan SAW untuk menentukan universitas terbaik berdasarkan kriteria yang ditentukan.

Penelitian yang dilakukan oleh Hermansyah Nur Ahmad, Vincent Suhartono, dan Ika Novita Dewi (2017) berjudul Penentuan Tingkat Kelulusan Tepat Waktu Mahasiswa Stmik Subang Menggunakan Algoritma C4.5. Penelitian ini membahas prediksi tingkat kelulusan tepat waktu dengan menentukan faktor-faktor yang mempengaruhi tingkat kelulusan mahasiswa tepat waktu dengan hasil prediksi yang lebih akurat menggunakan algoritma C4.5. Berdasarkan data dari Pusat Statistik Pendidikan Badan Penelitian dan Pengembangan Departemen Pendidikan Nasional Republik Indonesia pada tahun akademik 2001/2002 sampai dengan 2009/2010 menunjukkan bahwa perguruan tinggi menerima rata-rata sebanyak 868.050 mahasiswa baru dan meluluskan rata-rata 451.168 mahasiswa setiap tahunnya (seperti Gambar 2.2). Ketidakjelasan status tersebut bisa jadi karena mahasiswa menempuh studi tidak tepat waktu, memiliki status non aktif (*mangkir*) atau bahkan *drop out*.



Gambar 2. 2 Grafik Perbandingan Jumlah Mahasiswa Baru dan Lulusan Perguruan Tinggi di Indonesia

(Sumber: Data dari pusat statistik Pendidikan Badan Penelitian dan Pengembangan Departemen Pendidikan Nasional Republik Indonesia tahun akademik 2001/2002 sampai dengan 2009/2010 via (H. N. Ahmad et al., 2017))

Pada penelitian ini, peneliti menggunakan data mahasiswa Tahun Akademik 2013 STMIK Subang jurusan teknik informatika, data yang diperoleh memiliki 200 record dengan field yang ada pada data tersebut sebanyak 14 atribut data. Atribut Data yang digunakan adalah: Nama, Jenis_Kelamin, Usia, Status_Mahasiswa, Status_Pengambilan_TA, IPS_Smt1, IPS_Smt2, IPS_Smt3, IPS_Smt4, IPS_Smt5, IPS_Smt6, IPS_Smt7, IPS_Smt8, IPK, Status_Kelulusan. Berdasarkan hasil eksperimen diperoleh tingkat akurasi dari penerapan C.45 sebesar 95%, Presisi sebesar 91,89 % dan Recall sebesar 93,00%.

Penelitian yang dilakukan oleh V.L. Miguéisa, Ana Freitasb, Paulo J.V. Garciab, dan André Silva (2018) berjudul *Early segmentation of students according to their academic performance: A predictive modelling approach*. Penelitian ini dipublikasikan di *Decision Support Systems journal*, peneliti membahas segmentasi mahasiswa berdasarkan bukti kegagalan atau kinerja tinggi saat memasuki program sarjana dengan mengkomparasikan 6 algoritma prediksi seperti *random forests*, *decision trees*, *support vector machines*, *naive Bayes*, *bagged trees* dan *boosted trees*). Penelitian ini menggunakan Data set sebanyak 2459 mahasiswa dari tahun 2003 – 2015, dari *European Engineering School of a public research University*. Atribut yang digunakan terdiri dari 19 atribut (seperti Gambar 2.2), yang diambil dari latar belakang SMA, Faktor Psikometrik, faktor social dan demografi. Hasil penelitian menunjukkan bahwa Random Forest merupakan algoritma terbaik untuk model prediksi kinerja akademik mahasiswa di perguruan tinggi.

Tabel 2. 1 Atribut berpengaruh (Miguéis et al, 2018)

Explanatory variables	Type of variable	Normalized weight
Enrolment average grade	Enrolment process	1.00
Enrolment exams average grade	Enrolment process	0.79
Average grade in the first semester	First year assessment	0.76
Average grade in the second semester	First year performance	0.73
Number of ECTS completed in the second semester	First year performance	0.62
Number of ECTS completed in the first semester	First year performance	0.43
Average # of exams attended to conclude the courses in the first academic year	First year performance	0.40
High school average grade	High school background	0.39
Degree	Enrolment process	0.29
Mother's education level	Socio-economic status	0.26
Father's education level	Socio-economic status	0.26
Mother's job	Socio-economic status	0.23
Father's job	Socio-economic status	0.21
Academic year of enrolment	Enrolment process	0.17
Enrolment option	Enrolment process	0.17
Marital status	Socio-demographic	0.09
School type	High school background	0.06
Gender	Socio-demographic	0.05
Enrolment stage	Enrolment process	0.00

Dari *review* jurnal di atas, untuk membantu mencapai tujuan dalam penelitian ini maka diperoleh atribut-atribut yang dapat menjadi referensi untuk prediksi performansi mahasiswa. Adapun prediksi performansi dapat diproses dalam berbagai macam metode dengan hasil akurasi yang berbeda-beda.

2.2 Posisi Penelitian

Tujuan dari *review* artikel dan tinjauan pustaka terhadap penelitian terdahulu adalah agar penulis dapat mengetahui posisi penelitian saat ini. Pada penelitian ini yang dijadikan rujukan adalah atribut-atribut yang mempengaruhi prediksi kinerja mahasiswa dan algoritma atau metode yang digunakan. Adapun lingkup penelitian dan metodologi yang digunakan pada penelitian terdahulu seperti yang terlihat pada Tabel 2.3.

Berdasarkan Tabel 2.2, yang membedakan dan memposisikan penelitian ini terhadap penelitian terdahulu diantaranya:

- a. Melakukan komparasi metode klasifikasi *Naïve Bayes Classification* dengan *K-Nearest Neighbors* untuk mendapatkan algoritma untuk model prediksi kelulusan mahasiswa dengan hasil akurasi terbaik.
- b. Objek penelitian dilakukan untuk klasifikasi prediksi ketepatan waktu lulus mahasiswa sebagai data training dan data testing, yang kemudian

nantinya dapat diaplikasikan sebagai salah satu pendukung kebijakan pengelola program studi D3 Teknik Laboratorium Medik dalam memantau kinerja mahasiswa agar lulus tepat waktu.

- c. Atribut-atribut yang digunakan berdasarkan dari hasil litatur studi dan FGD dari pengelola program studi dan bagian admisi atau penerimaan mahasiswa baru. Atribut yang digunakan sebanyak 15 yaitu jenis kelamin, asal jurusan, asal daerah, Indeks Prestasi Semester (IPS) 1, IPS 2, IPS 3, IPS 4, Nilai Mata kuliah wajib semester 1 (TLM 1), TLM 2, TLM 3, TLM 4, Nilai Mata Kuliah Muatan Lokal semester 1 (MLK 1), MLK 2, MLK 3 dan MLK 4.

Tabel 2. 2 Tinjauan pustaka

Nama, Tahun	Permasalahan	Data	Fokus Penelitian	Metode/Algoritma	Variabel / atribut	Kontribusi
(H. N. Ahmad et al., 2017)	jumlah mahasiswa yang masuk jauh lebih besar dibandingkan dengan jumlah mahasiswa yang telah menyelesaikan masa studinya.	data mahasiswa TA 2013 jurusan teknik informatika STMIK Subang, data yang diperoleh memiliki 200 record	prediksi tingkat kelulusan mahasiswa	C45	- 14 atribut : Nama, Jenis_Kelamin, Usia, Status_Mahasiswa, Status_Pengambilan_TA, IPS_Smt1, IPS_Smt2, IPS_Smt3, IPS_Smt4, IPS_Smt5, IPS_Smt6, IPS_Smt7, IPS_Smt8, IPK, Status_Kelulusan	melakukan prediksi tingkat kelulusan tepat waktu dengan menentukan faktor-faktor yang mempengaruhi tingkat kelulusan mahasiswa tepat waktu. Output: Lulus tepat waktu & Lulus terlambat Nilai akurasi algoritma C4.5 sebesar 95,00%, artinya sangat akurat untuk diterapkan.
(Hussain et al., 2018)	membandingkan masing-masing algoritma berdasarkan akurasi untuk memilih algoritma dengan kinerja terbaik	Data akademik dan personal mahasiswa. 300 data dari Tiga Perguruan Tinggi yang berbeda	Perbandingan metode untuk prediksi kinerja akademik mahasiswa	J48, PART, BayesNetwork dan Random Forest	24 atribut:	metode Random Forest adalah algoritma yang paling cocok untuk prediksi kinerja akademik mahasiswa.
(F. Ahmad et al., 2015)	Pada awal semester untuk mahasiswa baru, seorang dosen kesulitan untuk mengetahui dan menganalisis kinerja mahasiswanya karena kurangnya informasi tentang latar belakang mahasiswa sebelumnya.	Data dari Juli 2006/2007 sampai Juli 2013/2014 yang berisi informasi demografi siswa, catatan akademik sebelumnya, dan informasi latar belakang keluarga.	Perbandingan metode untuk prediksi kinerja akademik mahasiswa	Decision Tree, Naïve Bayes, dan Rule Based classification techniques	demografi siswa, catatan akademik sebelumnya, dan informasi latar belakang keluarga	Model dapat berfungsi sebagai alat bantu bagi dosen untuk merencanakan bahan ajar guna meningkatkan kinerja mahasiswa, dan untuk menurunkan angka kegagalan pada mata kuliah ilmu komputer
(Abadi et al., 2018)	Input mahasiswa berasal dari berbagai daerah, maka diperlukan strategi	Data siswa, 100 record	Strategi Promosi	K-Means	alamat rumah, sekolah asal, transportasi, dan alasan memilih sekolah	Metode K-means dapat diterapkan untuk mendukung strategi

Nama, Tahun	Permasalahan	Data	Fokus Penelitian	Metode/Algoritma	Variabel / atribut	Kontribusi
	khusus oleh departemen pemasaran dalam melakukan pemasaran untuk mencari calon siswa agar promosi dilakukan dengan lebih efektif dan efisien.					promosi dalam menjangkau calon mahasiswa baru.
(Purba et al., 2018)	Salah satu faktor siswa gagal adalah putus sekolah. Siswa yang berpotensi drop out karena siswa kurang bersemangat dalam belajar, orang tua yang tidak didukung, siswa yang malu-malu dan waktu berperilaku siswa yang kurang.	Data akademik, 36 record	Prediksi mahasiswa Drop Out	K-Means	data penilaian mahasiswa, kualitas dan IPK.	Model ini dapat digunakan dalam membantu perguruan tinggi untuk mengetahui situasi kemahasiswaan dan menjadi pendidikan awal dalam proses pengambilan keputusan preventif dalam mengantisipasi mahasiswa drop out,
(Aminudin et al., 2018)	Banyaknya pilihan kriteria perguruan tinggi dan sebagian besar masyarakat bingung memilih perguruan tinggi yang sesuai untuk diri mereka sendiri dan tuntutan pekerjaan.	Data dari studi literatur	Sistem pendukung keputusan untuk menentukan universitas terbaik	Simple Additive Weighting (SAW)	<i>Academic achievement Master Lecturer Extracurricular Accreditation Status Facilities Scholarship</i>	model ini berupa sistem pendukung keputusan dengan SAW untuk menentukan universitas terbaik berdasarkan kriteria yang ditentukan
(Alharbi, Cornford, Dolder, & Iglesia, n.d.)		Data dari 984 mahasiswa sarjana yang memperoleh penghargaan akademik selama tahun 2005 hingga 2013	Perbandingan metode untuk prediksi Kinerja buruk mahasiswa	logistic regression; Neural Network; Decision List; Bayesian Network; C5, C&R Tree, Quest dan CHAID.	Demografi siswa dan kinerja umum, modul siswa	mengidentifikasi siswa yang lemah sedini mungkin, yaitu siswa yang akan berakhir dengan hasil yang buruk
(Negara & Doni, 2020)		Data pengolahan kas	Perbandingan Performa Algoritma Data Mining pada	decision tree C4.5 dan metode K-means	Hajj savings program Magnitude of Savings Total Savings Id_member	menentukan strategi mencari santri yang

Nama, Tahun	Permasalahan	Data	Fokus Penelitian	Metode/Algoritma	Variabel / atribut	Kontribusi	
			Dataset Tabungan Mahasiswa			berpotensi menabung tabungan haji.	
(Rohmawan, 2018)		data mahasiswa Universitas Dehasen Bengkulu Program Studi Teknik Informatika tahun akademik 2009, 2010 dan 2011. Data yang terkumpul sebanyak 600 data mahasiwa.	Perbandingan algoritma untuk prediksi kelulusan mhs	decision tree dan artificial neural network.	NIM, jenis kelamin, asal sekolah, jalur masuk, nilai ujian nasional, gaji orang tua, IP semester 1-4, IPK semester 1-4 dan keterangan lulus.	mengetahui mahasiswa lebih dini nantinya bisa lulus tepat waktu atau tidak. metode Artificial Neural Network memiliki tingkat accuracy yang lebih tinggi jika dibandingkan dengan metode decision tree yaitu 79,74%.	
(Miguéis et al., 2018)	Segmentasi mahasiswa berdasarkan bukti kegagalan atau kinerja tinggi saat memasuki program sarjana.	Data set sebanyak 2459 mahasiswa dari tahun 2003 – 2015, dari European Engineering School of a public research University	Perbandingan algoritma untuk prediksi	<i>random forests, decision trees, support vector machines, naive Bayes, bagged trees and boosted trees)</i>	19 atribut/variable Variabel independen terdiri dari latar belakang SMA, Faktor Psikometrik, faktor social dan demografi.	Segmentasi mahasiswa berdasarkan bukti kegagalan atau kinerja tinggi saat memasuki program sarjana. Metode dengan hasil terbaik yaitu Random Forest.	
(Amra & Maghari, 2017)	Banyak siswa yang tidak sukses	Data sekolah SD tahun 2015	Prediksi siswa	Kinerja	KNN dan Naïve Bayes Classification	13 atribut	Prediksi sejak dini kinerja siswa untuk meningkatkan kualitas pendidikan SD di Gaza
(Francis & Babu, 2019)	Belum diperoleh pola pada data pendidikan yang semakin bertambah. Deteksi dini nilai ujian akhir siswa.		Prediksi Akademik	Kinerja	Hibrid Klustering dan klasifikasi. SVM, DT, Neural Network dan Naïve Bayes dan K-Means.	Faktor demografi, faktor akademik, faktor perilaku, faktor tambahan	Membantu pendidik untuk mengenali kelemahan siswa sehingga dapat meningkatkan proses pembelajaran.
(Pambudi, Supianto, & Setiawan, 2019)	Adanya ketidakseimbangan jumlah rata-rata mahasiswa masuk dengan jumlah rata-rata kelulusan mahasiswa	Data mahasiswa sebanyak 1354 dari tahun 2011-2016.	Prediksi kelulusan mahasiswa berdasarka akdemik	kinerja	Naïve bayes	28 Atribut	Mengetahui pola – pola berdasarkan probabilitas pada tiap atribut yang dapat digunakan untuk menentukan apakah

Nama, Tahun	Permasalahan	Data	Fokus Penelitian	Metode/Algoritma	Variabel / atribut	Kontribusi
Penelitian yang diusulkan	Banyaknya mahasiswa yang lulus tidak tepat waktu	Data akademik tahun 2015/2016-2018/2019	Perbandingan metode untuk prediksi ketepatan waktu lulus mahasiswa	Algoritma Naïve Bayes Classification dan K-Nearest Neighbors	Variabel yang akan digunakan berdasarkan dari hasil validasi oleh pengelola program studi.dan bagian admisi atau PMB. Atribut yang digunakan sebanyak 15 yaitu jenis kelamin, asal jurusan, asal daerah, Indeks Prestasi Semester (IPS) 1, IPS 2, IPS 3, IPS 4, Nilai Mata kuliah wajib semester 1 (TLM 1), TLM 2, TLM 3, TLM 4, Nilai Mata Kuliah Muatan Lokal semester 1 (MLK 1), MLK 2, MLK 3 dan MLK 4.	mahasiswa lulus tepat waktu atau tidak. Membantu pengelola program studi untuk mengambil keputusan lebih dini terhadap mahasiswa D3 Teknologi Laboratorium Medik STIKES Guna Bangsa Yogyakarta yang diprediksi lulus terlambat. Algoritma KNN memiliki tingkat akurasi yang lebih baik dari pada NBC