



**KLASIFIKASI PENERAPAN PENGOLAHAN DATA MINING
DALAM BIDANG AKADEMIK**

TUGAS AKHIR

Hendri Setiawan
41517010069

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2021**



**KLASIFIKASI PENERAPAN PENGOLAHAN DATA MINING
DALAM BIDANG AKADEMIK**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:
Hendri Setiawan
41517010069

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2021

MERCU BUANA

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41517010069

Nama : Hendri Setiawan

Judul Tugas Akhir : Klasifikasi Penerapan Pengolahan Data Mining Dalam Bidang Akademik

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.



UNIVERSITAS
MERCU BUANA

Jakarta, 3 Agustus 2021



Hendri Setiawan

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Hendri Setiawan
NIM : 41517010069
Judul Tugas Akhir : Klasifikasi Penerapan Pengolahan Data Mining
Dalam Bidang Akademik

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 3 Agustus 2021



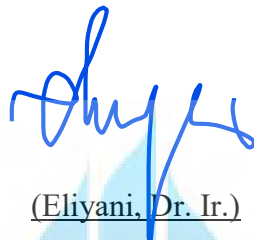
Hendri Setiawan

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010069
Nama : Hendri Setiawan
Judul Tugas Akhir : Klasifikasi Penerapan Pengolahan Data Mining
Dalam Bidang Akademik

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 3 Agustus 2021



(Eliyani, Dr. Ir.)


UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010069
Nama : Hendri Setiawan
Judul Tugas Akhir : Klasifikasi Penerapan Pengolahan Data Mining
Dalam Bidang Akademik

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 3 Agustus 2021



(Achmad Kodar, Drs. MT)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010069
Nama : Hendri Setiawan
Judul Tugas Akhir : Klasifikasi Penerapan Pengolahan Data Mining
Dalam Bidang Akademik

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 3 Agustus 2021



(Wawan Gunawan, S.Kom., MT)

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

NIM : 41517010069

Nama : Hendri Setiawan

Judul Tugas Akhir : Klasifikasi Penerapan Pengolahan Data Mining Dalam Bidang Akademik

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 3 Agustus 2021

Menyetujui



(Dr. Leonard Goeirmanto, ST, M.Sc)
Dosen Pembimbing

Mengetahui,

UNIVERSITAS
MERCU BUANA



(WaWawan Gulawan, S.Kom., M.Koord.)
Tugas Akhir Teknik Informatika



(Ir. Emil R. Kaburuan, Ph.D., IPM.)
Ka. Prodi Teknik Informatika

ABSTRAK

Nama : Hendri Setiawan
NIM : 41517010069
Pembimbing TA : Dr. Leonard Goeirmento, ST, M.Sc
Judul : Klasifikasi Penerapan Pengolahan Data Mining
Dalam Bidang Akademik

Data mining dalam dunia pendidikan merupakan teknik yang dapat digunakan untuk menganalisis data yang diperoleh dari proses dalam dunia pendidikan untuk memperoleh informasi baru dari data yang dimiliki sehingga berguna untuk meningkatkan kualitas pembelajaran. Salah satu cara yang dapat dilakukan untuk meningkatkan kualitas pembelajaran adalah dengan memprediksi prestasi akademik dari pelajar atau mahasiswa. Klasifikasi akademik berdasarkan potensi akademik dapat menjadi strategi untuk meningkatkan kelulusan pelajar, meningkatkan prestasi belajar, dan juga untuk pengelolaan data akademik yang lebih baik. Hingga saat ini banyak Perguruan Tinggi yang membantu pemerintah dalam percepatan peningkatan mutu pendidikan sehingga terciptanya lingkungan yang kompetitif. Banyaknya data yang terdapat di Perguruan Tinggi dapat dimanfaatkan dengan baik sesuai dengan kebutuhan dan diolah menjadi informasi yang bermanfaat sehingga dapat untuk mengetahui hubungan antara atribut-atribut data di dalamnya dapat di analisis dan keluaran yang diharapkan berupa kinerja siswa yang berkaitan dengan masa studi yaitu dapat dikategorikan sesuai atau terlambat dalam masa studi yang diantisipasi. Data mining dapat digunakan untuk institusi atau institusi pendidikan dan sering disebut dengan Educational Data Mining (EDM).

Kata kunci:

Klasifikasi, Clustering, Data Mining, Akademik, Ilmu Komputer, Universitas Mercu Buana

ABSTRACT

Name : Hendri Setiawan
Student Number : 41517010069
Counsellor : Dr. Leonard Goeirmanto, ST, M.Sc
Title : Klasifikasi Penerapan Pengolahan Data Mining
Dalam Bidang Akademik

Data mining in education is a technique that can be used to analyze data obtained from processes in the world of education to obtain new information from the data held so that it is useful for improving the quality of learning. One way that can be done to improve the quality of learning is to predict the academic achievement of students or students. Academic classification based on academic potential can be a strategy to increase student graduation, improve learning achievement, and also for better management of academic data. Until now, many universities have assisted the government in accelerating the improvement of the quality of education so as to create a competitive environment. The amount of data contained in tertiary institutions can be used properly according to needs and processed into useful information so that it is possible to find out the relationship between the data attributes in it can be analyzed and the expected output in the form of student performance related to the study period can be categorized appropriate or late in the anticipated study period. Data mining can be used for educational institutions or institutions and is often referred to as Educational Data Mining (EDM).

Kata kunci:

Clustering, Data Mining, akademik, Ilmu Komputer, Universitas Mercu Buana

KATA PENGANTAR

Puji syukur kita panjatkan kepada ALLAH SWT yang telah memberikan nikmat dan rahmatnya sehingga penulis dapat menyelesaikan laporan tugas akhir ini yang berjudul “Klasifikasi Penerapan Pengolahan Data Mining Dalam Bidang Akademik” Tugas akhir ini disusun untuk memenuhi syarat memperoleh gelar sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Penulis menyadari bahwa tanpa bantuan dan bimbingan dari dosen pembimbing, orang tua serta teman-teman tidak akan terselesaikan dengan baik tugas akhir ini. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Hery Derajad Wijaya, S.Kom, MM selaku Kepala Prodi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana dan selaku dosen pembimbing tugas akhir yang telah memberikan masukan saat bimbingan dan meluangkan waktu sebagian besarnya untuk melakukan bimbingan dalam penyusunan tugas akhir ini hingga selesai.
2. Bapak Wawan Gunawan, S.Kom., MT Selaku Koordinator Tugas Akhir Teknik Informatika di Universitas Mercu Buana.
3. Bapak Leonard Goeirmanto, DR., ST, M.Sc selaku Dosen Pembimbing Akademik yang selalu membimbing dan menasehati pencapaian akademik saya selama kuliah.
4. Orang tua yang selalu memberikan dukungan secara penuh dan doa sehingga dapat menyelesaikan semua laporan tugas akhir dengan lancar.

Akhir kata, penulis berharap tugas akhir ini dapat menjadi bermanfaat bagi pembaca dan menambah wawasan pengetahuan semua pihak.

Jakarta, 3 Agustus 2021



Hendri Setiawan

I. TOPIK / BIDANG ILMU

Berisi topik / bidang ilmu yang akan dibahas, disesuaikan dengan topik / bidang ilmu yang ada pada Tugas Akhir, yaitu

- Kecerdasan Buatan, Data Mining, Machine Learning, Pengolahan Citra, dan topik sejenis lainnya
- Pemrograman Web / Mobile
- Jaringan Komputer dan atau Keamanan Jaringan

II. DAFTAR JURNAL (MINIMAL 20)

Bagian ini berisi **daftar judul artikel** yang akan di review. Dimana terdiri dari*

1. Minimal 5 Jurnal dari jurnal terakreditasi Nasional Sinta pada web <http://sinta.ristekbrin.go.id/>, bisa dicari di <http://garuda.ristekbrin.go.id/>
2. Minimal 5 Jurnal yang terindeks Scopus atau science direct dan jurnal bereputasi lainnya
3. Maximal 5 jurnal yang tidak termasuk dalam point 1,2

No	Judul Jurnal	Kategori (diisi dengan 1/2/3)*
1.	Using Data Mining Techniques to Predict Student Performance to Support Decision Making in University Admission Systems	3
2.	Students' Academic Performance Prediction using Data Mining	3
3.	Predicting Students' Academic Procrastination in Blended Learning Course Using Homework Submission Data	3
4.	A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining	3

5.	Combining University Student Self Regulated Learning Indicators and Engagement with Online Learning Events to Predict Academic Performance	3
6.	Data Mining Mid-Semester Exam Data Clustering Using K-Means Algorithm	1
7.	An Educational Data Mining For Student Academic Prediction Using K-Means Clustering And Naïve Bayes Classifier	1
8.	The uses of educational data mining in academic performance analysis at higher education institutions (case study at UNJANI)	1
9.	Classification Analysis Using C4.5 Algorithm To Predict The Level of Graduation of Nurul Falah Pekanbaru High School Students	1
10.	The Decision Tree c5.0 Classification Algortima For Prediction Student Academic Performancse	1
11.	Application of Data Mining Classification Method for Student Graduation Prediction Using K-Nearest Neighbor (K-NN) Algorithm	1
12.	Nave Bayes Algorithm for Classification of Student Major Specialization	1
13.	Classification and Prediction of Student GPA Using K-Means Clustering Algorithm to Assist Student Admission Process	1
14.	Implementation Naïve Bayes Algorithm	1

	for Student Classification Based on Graduation Status	
--	--	--



15.	Text Clasification using Naïve Baye Updateable Algorithm In SBMPTN Test Questions	1
16.	Determinating Student Interactions in a Virtual Learning Environment Using Data Mining	2
17.	The Third Information Systems International Conference A Review on Predicting Student's Performance using Data Mining Techniques	2
18.	An Intelligent Prediction System for Educational Data Mining Based on Ensemble and Filtering approaches	2
19.	Using Data Mining Techniques to Guide Academic Programs Design and Assessment	2
20.	Mining authentic Student feedback for Faculty Using Naïve Bayes Classifier	2

UNIVERSITAS
MERCU BUANA

III. TABEL REVIEW

No	1
Judul Artikel	Using Data Mining Techniques to Predict Student
Topik	Data Mining
Data	Dua Semester Pertama, Koefisien Korelasi, Koefisien Determinasi.
Metode / Algoritma	Artificial Neural Network (ANN), Decision Tree, Support Vector Machine (SVM), dan Naive Bayes
Abstrak	<p>Sistem penerimaan yang didasarkan pada kriteria penerimaan yang valid dan dapat diandalkan sangat penting untuk menyeleksi kandidat yang memiliki prestasi akademik baik di institusi pendidikan tinggi. Studi ini berfokus pada cara-cara untuk mendukung universitas dalam pengambilan keputusan penerimaan menggunakan teknik data mining untuk memprediksi kinerja akademik pelamar di universitas. Kumpulan data dari 2.039 siswa yang terdaftar di Computer Science and Information College di universitas negeri Saudi dari 2016 hingga 2019 digunakan untuk memvalidasi metodologi yang diusulkan. Hasilnya menunjukkan bahwa kinerja universitas awal pelamar dapat diprediksi sebelum masuk berdasarkan kriteria pra-penerimaan tertentu (rata-rata kelas sekolah menengah, skor Tes Prestasi Scholastic, dan skor Tes Bakat Umum). Hasilnya juga menunjukkan bahwa skor Tes Masuk Prestasi Sekolah adalah kriteria pra-masuk yang paling akurat memprediksi kinerja siswa di masa depan. Oleh karena itu, skor ini harus diberi bobot lebih dalam sistem penerimaan. Kami juga</p>

	<p>menemukan bahwa teknik Jaringan Syaraf Tiruan memiliki tingkat akurasi di atas 79%, membuatnya lebih unggul daripada teknik klasifikasi lainnya yang dipertimbangkan (Decision Trees, Support Vector Machines, and Naïve Bayes).[1]</p>
<p>Hasil</p>	<p>Berdasarkan hasil dan rekomendasi yang diberikan, penelitian PNU yang mengaku dan Pendaftaran Deanship memutuskan bahwa sistem penerimaan saat ini harus diubah untuk memodifikasi bobot kriteria penerimaan untuk memastikan kinerja tinggirst- siswa tahun. Dekan memutuskan untuk menambah bobot kriteria SAAT dan mengubah bobot tiga kriteria penerimaan (HSGA, SAAT, dan GAT) masing - masing menjadi 30%, 40%, dan 30%, untuk tahun ajaran 2018-2019. Bobot yang sesuai sebelumnya adalah 60%, 20%, 20%, masing-masing.</p>
<p>Kesimpulan</p>	<p>Studi ini mengkonfirmasi keefektifan pemodelan prediksi di institusi pendidikan tinggi di mana pengambil keputusan dapat menggunakan model ini dalam merencanakan dan mengoptimalkan alokasi sumber daya institusi yang terbatas. Selain itu, hasil penelitian menunjukkan bahwa model kinerja tinggi untuk memprediksi kinerja awal siswa dapat dikembangkan berdasarkan informasi pra-masuk. Sebagai contoh, dalam studi khusus ini, model JST mampu mencapai tingkat akurasi kinerja sekitar 79,22%. Studi ini menyimpulkan bahwa skor SAAT adalah kriteria penerimaan yang paling akurat memprediksi kinerja akademik di kemudian hari dan karena itu bobot lebih harus diberikan padanya. Berdasarkan rekomendasi ini, pengambil</p>

	keputusan di PNU, tempat penelitian ini dilakukan, meninjau sistem penerimaan dan memutuskan untuk mengubah bobot dari tiga mereka yang kriteria penerimaandisebutkan dalam penelitian ini dengan memberikan bobot lebih pada skor SAAT.
Penulis	Hanana Abdulah Mengash
Nama jurnal, Volume, Nomor, Tahun	<i>IEEE Access</i> , vol. 8, pp. 55462–55470, 2020
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian yang memfokuskan pada cara-cara untuk mendukung universitas dalam pengambilan keputusan penerimaan menggunakan teknik data mining untuk memprediksi kinerja akademik pelamar di universitas. Penulis menggunakan Kumpulan data yang terdiri dari 2.039 siswa yang terdaftar di Computer Science dan Information College di universitas negeri Saudi dari 2016 hingga 2019 digunakan untuk memvalidasi metodologi yang diusulkan. Pada penelitian ini penulis menggunakan algoritma (Decision Trees, Support Vector Machines, and Naïve Bayes). Untuk melihat klasifikasi mana yang lebih unggul dari pada teknik klasifikasi lainnya yang dipertimbangkan. Penulis memaparkan bahwa penelitian ini menunjukkan bahwa model kinerja tinggi untuk memprediksi kinerja awal siswa dapat dikembangkan berdasarkan informasi pra-masuk. Sebagai contoh, dalam studi khusus ini, model JST mampu mencapai tingkat akurasi kinerja sekitar 79,22%. Studi ini menyimpulkan bahwa skor SAAT adalah kriteria penerimaan yang paling akurat memprediksi kinerja akademik di kemudian hari dan karena itu bobot lebih harus diberikan padanya.</p>

Link URL Jurnal	https://ieeexplore.ieee.org/document/9042216
-----------------	---

No	2
Judul Artikel	Students' Academic Performance Prediction using
Topik	Data Mining
Data	kuesioner yang berisi tentang demografi siswa, IPK sebelumnya
Metode / Algoritma	Decision Tree, Random Forest
Abstrak	<p>Data Mining Pendidikan telah menjadi topik yang muncul saat ini karena pertumbuhan data pendidikan. Bidang ini memungkinkan untuk mengembangkan metode untuk menemukan tersembunyi pola dari data pendidikan. Metode yang diambil dari disiplin Educational Data Mining kemudian digunakan untuk memahami siswa termasuk perilaku belajar mereka serta untuk memprediksi kinerja akademik mereka. Penelitian ini mengusulkan model untuk memprediksi kinerja akademik mahasiswa Ilmu Komputer menggunakan teknik Data Mining. Pengumpulan data dilakukan dengan menggunakan kuesioner yang berisi tentang demografi siswa, IPK sebelumnya, dan informasi latar belakang keluarga. Dua model Data Mining (Decision Tree and Random Forest) diterapkan pada data siswa untuk membuat model prediksi kinerja akademik siswa terbaik. Hasil penelitian menunjukkan bahwa Decision Tree merupakan model terbaik dibandingkan dengan Random Forest dengan nilai akurasi tertinggi yaitu 66,9%. Studi ini menunjukkan bahwa ada beberapa fitur relevan yang mempengaruhi kinerja akademik siswa. [2]</p>

<p>Hasil</p>	<p>Pengklasifikasi Decision Tree menghasilkan hasil yang lebih baik dibandingkan dengan pengklasifikasi Random Forest. Ini memberikan akurasi 66,85% yang menunjukkan bahwa ada 117 siswa yang diklasifikasikan dengan benar ke kelas yang tepat (Buruk, Sedang, dan Baik) dari 175 siswa dalam dataset pelatihan. Selain itu, kinerja prediktif dari model pelatihan yang dibangun menghasilkan kinerja yang stabil masing-masing sebesar 0,622, 0,669, dan 0,632 nilai presisi, recall, dan F-measure. Hal ini membuat Decision Tree adalah model terbaik untuk model prediksi kami.</p> <p>Selanjutnya, untuk memvalidasi model prediksi kami, dengan menggunakan dataset pengujian kami melakukan pengujian independen menggunakan algoritma Decision Tree hanya karena menunjukkan hasil yang lebih baik daripada Random Forest. Seperti yang terlihat pada Tabel 4, model kami menghasilkan akurasi 66,2%, nilai presisi 0,594, recall 0,635, dan F-measure 0,614. Ini menunjukkan bahwa model kami memberikan hasil yang adil. Kinerja presisi dan recall menunjukkan kinerja yang seimbang yang menunjukkan bahwa tidak ada overfitting dari dataset pelatihan ke dataset pengujian.</p>
<p>Kesimpulan</p>	<p>Penelitian ini melakukan analisis perbandingan dua teknik klasifikasi antara Decision Tree dan Random Forest menggunakan WEKA. Hasil penelitian menunjukkan bahwa Decision Tree memiliki kinerja klasifikasi yang lebih baik daripada Random Forest. Ada pengetahuan menarik yang disediakan di bagian sebelumnya dari makalah ini. Secara keseluruhan, hasil penelitian menunjukkan faktor-faktor yang mempengaruhi</p>

	<p>perilaku belajar siswa. Misalnya, siswa tahun pertama yang menghabiskan terlalu banyak waktu untuk pergi keluar dengan teman-temannya cenderung menunjukkan prestasi akademik yang buruk serta jika mereka menghabiskan sedikit waktu dengan keluarga mereka. Pengetahuan yang diperoleh dari kamipercobaan dapat menjadi bahan pertimbangan bagi siswa sebagai evaluasi perilaku belajar mereka, untuk dosen sebagai referensi latar belakang siswa serta perilaku belajar siswa, dan untuk orang tua juga.</p>
Penulis	Fergie Joanda Kaunang, Reymon Rotikan
Nama jurnal, Volume, Nomor, Tahun	Conf. Informatics Comput. ICIC 2018, pp. 1–5, 2018
Ulasan artikel	<p>Penelitian ini dilakukan untuk mengusulkan model untuk memprediksi kinerja akademik mahasiswa Ilmu Komputer menggunakan teknik Data Mining. Penulis menggunakan Pengumpulan data dilakukan dengan menggunakan kuesioner yang berisi tentang demografi siswa, IPK sebelumnya, dan informasi latar belakang keluarga. Pada penelitian ini penulis menggunakan algoritma (Decision Tree and Random Forest) yang diterapkan pada data siswa untuk membuat model prediksi kinerja akademik siswa terbaik, yaitu penelitian menunjukkan bahwa Decision Tree merupakan model terbaik dibandingkan dengan Random Forest dengan nilai akurasi tertinggi yaitu 66,9%. Studi ini menunjukkan bahwa ada beberapa fitur relevan yang mempengaruhi kinerja akademik siswa.</p>
Link URL Jurnal	https://ieeexplore.ieee.org/document/8780547

No	3
Judul Artikel	Predicting Students' Academic Procrastination
Topik	Data Mining
Data	Log Course SCHOLAT homework : Datestart, Dateend, Dateupload
Metode / Algoritma	Algoritma K-means Clustering
Abstrak	<p>Penundaan akademik telah dilaporkan mempengaruhi kinerja siswa di lingkungan belajar yang didukung komputer. Studi telah menunjukkan bahwa siswa yang menunjukkan kecenderungan penundaan yang lebih tinggi mencapai kurang dari siswa dengan kecenderungan penundaan yang lebih rendah. Penting bagi seorang guru untuk menyadari perilaku siswa terutama kecenderungan menunda-nunda mereka. Teknik EDM dapat digunakan untuk menganalisis data yang dikumpulkan melalui lingkungan belajar yang didukung komputer dan untuk memprediksi perilaku siswa. Dalam makalah ini, kami menyajikan sebuah algoritma yang disebut peningkatan kinerja akademik siswa melalui pendeteksian keterlambatan pekerjaan rumah/non-submission (SAPE) untuk memprediksi kinerja akademik siswa. Algoritma ini dirancang untuk memprediksi siswa dengan kesulitan belajar melalui perilaku penyerahan pekerjaan rumah mereka. Pertama, siswa diberi label sebagai prokrastinator atau non-penunda menggunakan algoritma k-means clustering. Kemudian, metode klasifikasi yang berbeda digunakan untuk mengklasifikasikan siswa menggunakan vektor fitur pengiriman pekerjaan rumah. Kami menggunakan sepuluh metode klasifikasi, yaitu ZeroR, OneR, ID3, J48, random</p>

	<p>forest, decision stump, JRip, PART, NBTree, dan Prism. Sebuah analisis rinci disajikan mengenai kinerja metode klasifikasi yang berbeda untuk jumlah kelas yang berbeda. Analisis mengungkapkan bahwa secara umum akurasi prediksi semua metode menurun dengan bertambahnya jumlah kelas. Namun, metode yang berbeda melakukan yang terbaik atau terburuk untuk jumlah kelas yang berbeda. [3]</p>
<p>Hasil</p>	<p>Sepuluh metode klasifikasi, yaitu Zero, OneR, ID3, J48, random forest, decision stump, JRip, PART, NBTree dan Prism digunakan untuk mengklasifikasikan data. Hasil klasifikasi dibandingkan dengan jumlah kelas yang berbeda. Empat Evaluasi ukuran, persentase contoh yang diklasifikasikan dengan benar, contoh yang persentase salah diklasifikasikan, statistik kappa dan kesalahan kuadrat rata-rata akar digunakan untuk mengevaluasi kinerja metode klasifikasi dengan jumlah berbeda kelas. Gambar 3 menunjukkan persentase diklasifikasikan dengan benar contoh yang. Terbukti bahwa persentase diklasifikasikan dengan benar contoh yang menurun seiring dengan meningkatnya jumlah kelas untuk metode yang sama. Namun, Zero R dan OneR menunjukkan terburuk kinerja, menurun dari 80,7 dan 90,8 menjadi 51,4 dan 65,1 masing-masing ketika jumlah kelas ditingkatkan dari 2 menjadi 5. Gambaran serupa terlihat pada Gambar 4, di mana ZeroR dan OneR kembali menjadi metode terburuk dalam kinerja. perbandingan, masing-masing 19,3 dan 9,2 menjadi 48,6 dan 34,9 ketika jumlah kelas dinaikkan dari 2 menjadi 5. Gambar 5 dan 6</p>

	<p>menunjukkan statistik kappa dan nilai RMSE untuk metode yang berbeda ketika jumlah kelas ditingkatkan dari 2 menjadi 5 secara berurutan. Tabel 2 menunjukkan standar deviasi dari metode yang berbeda untuk empat ukuran yang berbeda. Tabel Menunjukkan bagaimana empat langkah-langkah ini, persentase kasus benar dan salah diklasifikasikan, statistik kappa dan akar berarti perubahan error kuadrat sebagai jumlah kelas berubah dari 2 sampai 5. Untuk contoh diklasifikasikan dengan benar, ZeroR dan Oner menunjukkan nilai-nilai deviasi standar yang lebih tinggi wrt lainnya metode, jumlah diklasifikasikan dengan benar contoh yang sangat bervariasi untuk kedua metode ini. Tren serupa terlihat dalam kasus contoh yang salah diklasifikasikan. Dalam hal statistik kappa, NBTree menunjukkan kinerja paling konsisten dengan standar deviasi terendah, 0,017. Random forest menjadi classier terbaik kedua karena standar deviasi untuk metode ini terendah kedua, 0,023. Demikian pula, NBTree juga menunjukkan kinerja yang konsisten dalam nilai RMSE dengan standar deviasi 0,01. Metode lain dengan standar deviasi terendah adalah Random forest dan J48 dengan nilai standar deviasi masing-masing 0,013 dan 0,017.</p>
Kesimpulan	<p>Tiga algoritma disajikan dalam pekerjaan ini. Algoritme pertama, yaitu, algoritma 1 adalah cara baru untuk membangun vektor fitur pengiriman pekerjaan rumah siswa. Vektor fitur ini dapat digunakan untuk merepresentasikan perilaku penyerahan pekerjaan rumah siswa dalam satu semester. Algoritma 2 merinci langkah-langkah yang diambil untuk</p>

	<p>memverifikasi kebenaran vektor fitur. Algoritma 3 menyajikan proses penerapan metode clustering dan klasifikasi untuk memprediksi perilaku prokrastinasi dan non prokrastinasi siswa. Algoritma digeneralisasi dan fleksibel untuk sepenuhnya mengotomatisasi proses mengidentifikasi siswa dengan kesulitan belajar.</p>
Penulis	<p>Afyab Akram, Chengzhou Fu, Yuyao li, Muhamad Yaqoob Javed, Ronghua Lin, yungcheng, Yong Tang</p>
Nama jurnal, Volume, Nomor, Tahun	<p><i>IEEE Access</i>, vol. 7, pp. 102487–102498, 2019</p>
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian terkait menyajikan sebuah algoritma yang disebut peningkatan kinerja akademik siswa melalui pendeteksian keterlambatan pekerjaan rumah/non-submission (SAPE) untuk memprediksi kinerja akademik siswa penundaan akademik kinerja siswa di lingkungan belajar yang didukung komputer. Dimana studi telah menunjukkan bahwa siswa yang menunjukkan kecenderungan penundaan yang lebih tinggi mencapai kurang dari siswa dengan kecenderungan penundaan yang lebih rendah. Penulis menggunakan metode Teknik EDM dapat digunakan untuk menganalisis data yang dikumpulkan melalui lingkungan belajar yang didukung komputer dan untuk memprediksi perilaku siswa. Penulis menggunakan sepuluh metode klasifikasi, yaitu ZeroR, OneR, ID3, J48, random forest, decision stump, JRip, PART, NBTree, dan Prism. Sebuah analisis rinci disajikan mengenai kinerja metode klasifikasi yang berbeda untuk jumlah kelas yang berbeda. Analisis mengungkapkan bahwa secara umum akurasi prediksi semua metode menurun dengan</p>

	<p>bertambahnya jumlah kelas. Namun, metode yang berbeda melakukan yang terbaik atau terburuk untuk jumlah kelas yang berbeda. diklasifikasikan dengan benar, ZeroR dan Oner menunjukkan nilai-nilai deviasi standar yang lebih tinggi wrt lainnya metode, jumlah diklasifikasikan dengan benar contoh yang sangat bervariasi untuk kedua metode ini. Tren serupa terlihat dalam kasus contoh yang salah diklasifikasikan. Dalam hal statistik kappa, NBTree menunjukkan kinerja paling konsisten dengan standar deviasi terendah, 0,017. Random forest menjadi classier terbaik kedua karena standar deviasi untuk metode ini terendah kedua, 0,023. Demikian pula, NBTree juga menunjukkan kinerja yang konsisten dalam nilai RMSE dengan standar deviasi 0,01. Metode lain dengan standar deviasi terendah adalah Random forest dan J48 dengan nilai standar deviasi masing-masing 0,013 dan 0,017.</p>
Link URL Jurnal	https://ieeexplore.ieee.org/document/8778644

MERCU BUANA

No	4
Judul Artikel	A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining
Topik	Data Mining
Data	Data penerimaan siswa dikumpulkan setiap periode akademik
Metode / Algoritma	Decision Trees, Naïve Bayes
Abstrak	Makalah ini menyajikan hasil penerapan pendekatan data mining pendidikan untuk model atrisi akademik (kehilangan status akademik) di Universidad Nacional de Colombia. Dua model data mining didefinisikan untuk menganalisis

	<p>data akademik dan nonakademik; model menggunakan dua teknik klasifikasi, Nave Bayes dan pengklasifikasi Decision Three, untuk memperoleh pemahaman yang lebih baik tentang gesekan selama pendaftaran pertama dan untuk menilai kualitas data untuk tugas klasifikasi, yang dapat dipahami sebagai prediksi hilangnya status akademik karena prestasi akademik yang rendah. Model bertujuan untuk memprediksi gesekan dalam empat pendaftaran pertama siswa. Pertama, mempertimbangkan salah satu dari periode ini, dan kemudian, pada pendaftaran tertentu. Catatan akademik historis dan data dari proses penerimaan digunakan untuk melatih model yang dievaluasi menggunakan validasi silang dan catatan yang sebelumnya tidak terlihat dari periode akademik penuh. Hasil eksperimen menunjukkan bahwa prediksi hilangnya status akademik meningkat ketika data akademik ditambahkan. [4]</p>
<p>Hasil</p>	<p>Prediksi hilangnya status akademik di salah satu dari empat periode akademik pertama menggunakan data awal, tetapi hasilnya tidak memuaskan; akurasi seimbang berkisar 51-52% di Decision Trees dan antara 54-57% di Naïve Bayes</p>
<p>Kesimpulan</p>	<p>Dua algoritma pembelajaran, nave Bayes dan Decision Trees, digunakan untuk membangun model klasifikasi untuk memprediksi hilangnya status akademik karena rendah kinerja akademik pada dua program sarjana teknik di Universidad Nacional de Colombia. Model dievaluasi menggunakan skenario yang berbeda tergantung pada informasi yang tersedia. Skenario ini termasuk prediksi hilangnya status akademik</p>

	<p>karena kinerja akademik yang rendah setiap saat dalam dua tahun pertama, pada pendaftaran tertentu hanya menggunakan data penerimaan dan kemudian termasuk informasi akademik, yaitu nilai dan kredit yang terdaftar. Model diuji dengan catatan tak terlihat yang sesuai dengan periode akademik penuh (2012-II).</p> <p>Kinerja pengklasifikasi Bayes meningkat ketika data akademik dari pendaftaran pertama ditambahkan; Namun, kinerjanya menurun setelah penambahan data akademik dari pendaftaran kedua. Hal ini mungkin disebabkan oleh asumsi independensi yang dibutuhkan oleh algoritma. Hasil Naïve Bayes lebih baik pada set tes; namun, ada perbedaan antara data pelatihan dan data uji. Hasil Decision Trees lebih konsisten mengenai subjek itu sehingga lebih dapat diandalkan saat menguji data baru.</p>
Penulis	Camilo E. López G., Elizabeth León Guzmán and Fabio A. González
Nama jurnal, Volume, Nomor, Tahun	<i>Rev. Iberoam. Tecnol. del Aprendiz.</i> , vol. 10, no. 3, pp. 119–125, 2015
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian terkait penerapan pendekatan data mining pendidikan dengan model atrisi akademik (kehilangan status akademik) di Universidad Nacional de Colombia. Model bertujuan untuk memprediksi gesekan dalam empat pendaftaran pertama siswa. Pertama, mempertimbangkan salah satu dari periode ini, dan kemudian, pada pendaftaran tertentu. Catatan akademik historis dan data dari proses penerimaan digunakan untuk melatih model yang dievaluasi menggunakan validasi silang dan catatan yang sebelumnya tidak terlihat dari periode akademik</p>

	<p>penuh. Hasil eksperimen menunjukkan bahwa prediksi hilangnya status akademik meningkat ketika data akademik ditambahkan. Penulis menggunakan algoritma pengklasifikasi Decision Tree, Naive Bayes untuk memperoleh pemahaman yang lebih baik tentang gesekan selama pendaftaran pertama dan untuk menilai kualitas data untuk tugas klasifikasi. Dari penelitian yang dilakukan Prediksi hilangnya status akademik di salah satu dari empat periode akademik pertama menggunakan data awal, tetapi hasilnya tidak memuaskan; akurasi seimbang berkisar 51-52% di Decision Trees dan antara 54-57% di Naive Bayes. Dimana Hasil Naive Bayes lebih baik pada set tes; namun, ada perbedaan antara data pelatihan dan data uji. Hasil Decision Trees lebih konsisten mengenai subjek itu sehingga lebih dapat diandalkan saat menguji data baru.</p>
Link URL Jurnal	https://ieeexplore.ieee.org/document/7156098

UNIVERSITAS
MERCU BUANA

No	5
Judul Artikel	Combining University Student Self Regulated Learning Indicators and Engagement with Online Learning Events to Predict Academic Performance
Topik	Data Mining
Data	Data penelitian ini sepenuhnya sukarela dan persetujuan tertulis mereka atas kuesioner, nilai akhir mereka, dan jejak digital mereka dalam Sistem Manajemen Pembelajaran.
Metode / Algoritma	Metode Ward
Abstrak	Teori belajar mandiri digunakan untuk memahami alasan perbedaan tingkat kinerja

	<p>akademik mahasiswa. Demikian pula, penelitian analitik pembelajaran mengusulkan kombinasi jejak data terperinci yang berasal dari tugas-tugas yang dimediasi teknologi dengan berbagai algoritma untuk memprediksi kinerja akademik siswa. Pendekatan pertama dirancang untuk memberikan panduan pedagogis yang bermakna, sedangkan yang kedua dirancang untuk mengidentifikasi pola dan hubungan peristiwa yang dapat diterjemahkan ke dalam perbaikan yang dapat ditindaklanjuti. Manfaat dari kedua pendekatan telah memotivasi penelitian ini untuk menyelidiki apakah kombinasi dari data laporan diri dan data yang timbul dari pengamatan keterlibatan siswa dengan peristiwa pembelajaran online menawarkan pemahaman yang lebih dalam dan penjelasan tentang mengapa beberapa siswa mencapai tingkat yang relatif lebih tinggi. prestasi akademik. Dalam makalah ini kami mengeksplorasi bagaimana menggabungkan data tentang keterampilan belajar mandiri dengan ukuran aktivitas online yang dapat diamati dalam kursus pembelajaran campuran untuk meningkatkan kemampuan prediktif kinerja akademik siswa untuk tujuan menginformasikan pengajaran dan desain tugas. Sebuah studi kasus di sebuah kursus dengan 145 siswa menunjukkan bahwa variasi nilai akhir siswa untuk kursus mereka lebih baik dijelaskan ketika faktor-faktor dari kedua pendekatan dipertimbangkan. [5]</p>
<p>Hasil</p>	<p>Hasil penelitian ini menawarkan sejumlah ide untuk mengajar siswa dalam kursus campuran. Hubungan antara variabel dalam analisis korelasi dan cluster menunjukkan bahwa</p>

	<p>pengalaman siswa yang positif tentang efikasi diri, tes, motivasi, regulasi diri <i>dan</i> interaksi positif dengan banyak acara online, terutama yang menawarkan umpan balik (seperti dasbor dan pertanyaan pilihan ganda) dan refleksi dan penalaran (seperti urutan latihan pemecahan masalah) akan berkorelasi dengan pencapaian akademik yang relatif lebih tinggi dibandingkan dengan pengalaman negatif dari aspek-aspek tersebut. Hasil ini menawarkan strategi bagi guru untuk mengeksplorasi desain tugas dan pendekatan yang melibatkan pengungkapan kepada siswa apa yang melibatkan pengaturan diri positif untuk pembelajaran, bagaimana berinteraksi secara efektif dengan acara pembelajaran online, atau mendesain ulang materi instruksional yang menyertai kegiatan ini.</p>
Kesimpulan	<p>Studi yang dijelaskan dalam makalah ini menyoroti pentingnya menganalisis pengalaman belajar yang menggabungkan wawasan yang diperoleh dengan data yang dilaporkan sendiri berdasarkan kerangka teoretis yang mapan seperti pengaturan diri, dengan yang diperoleh dengan metode seperti merekam interaksi antara siswa dan acara kursus di sebuah platform daring. Analisis faktor yang dilaporkan sendiri tidak diragukan lagi menjelaskan bagaimana siswa mendekati aspek pengalaman belajar mereka dan mengidentifikasi faktor-faktor yang berkontribusi pada hasil akademik yang lebih baik. Tetapi ketika dilengkapi dengan perilaku yang diamati dalam lingkungan online, ada porsi signifikan dari konteks spesifik yang dibawa ke dalam analisis dan model menawarkan deskripsi yang lebih akurat. Studi ini menggunakan kursus</p>

	<p>blended learning di mana siswa diminta untuk mempersiapkan dua dari tiga sesi mingguan menggunakan materi online. Relevansi skenario ini sangat penting, karena menempatkan kombinasi tugas tatap muka dan online sebagai salah satu fitur yang menentukan. Meskipun siswa melaporkan preferensi mereka sehubungan dengan aspek pengaturan diri dari pendekatan pembelajaran mereka, kombinasi dengan indikator keterlibatan online menawarkan model linier yang menjelaskan 32 persen varians kinerja akademik.</p>
Penulis	Abelardo Pardo, Feifei Han, and Robert A. Ellis
Nama jurnal, Volume, Nomor, Tahun	<i>IEEE Trans. Learn. Technol.</i> , vol. 10, no. 1, pp. 82–92, 2017
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian terhadap bagaimana menggabungkan data tentang keterampilan belajar mandiri dengan ukuran aktivitas online yang dapat diamati dalam kursus pembelajaran campuran untuk meningkatkan kemampuan prediktif kinerja akademik siswa untuk tujuan menginformasikan pengajaran dan desain tugas. Penulis menggunakan Data penelitian yang sepenuhnya sukarela dan persetujuan tertulis mereka atas kuesioner, nilai akhir mereka, dan jejak digital mereka dalam Sistem Manajemen Pembelajaran. Dalam penelitian yang dilakukan oleh penulis Hasil penelitian ini menawarkan sejumlah ide untuk mengajar siswa dalam kursus campuran. Hubungan antara variabel dalam analisis korelasi dan cluster menunjukkan bahwa pengalaman siswa yang positif tentang efikasi diri, tes, motivasi, regulasi diri dan interaksi positif dengan banyak acara online, terutama</p>

	yang menawarkan umpan balik (seperti dasbor dan pertanyaan pilihan ganda) dan refleksi dan penalaran (seperti urutan latihan pemecahan masalah) akan berkorelasi dengan pencapaian akademik yang relatif lebih tinggi dibandingkan dengan pengalaman negatif dari aspek-aspek tersebut. Hasil ini menawarkan strategi bagi guru untuk mengeksplorasi desain tugas dan pendekatan yang melibatkan pengungkapan kepada siswa apa yang melibatkan pengaturan diri positif untuk pembelajaran, bagaimana berinteraksi secara efektif dengan acara pembelajaran online, atau mendesain ulang materi instruksional yang menyertai kegiatan ini.
Link URL Jurnal	https://ieeexplore.ieee.org/document/7782787

No	6
Judul Artikel	Data Mining Mid-Semester Exam Data Clustering Using K-Means Algorithm
Topik	Data Mining
Data	Nilai UTS Mahasiswa Semester IV
Metode / Algoritma	K-Means
Abstrak	Ilmu pengetahuan dan teknologi akan mempermudah pekerjaan manusia. Namun, disisi lain akan meningkatkan persaingan. Dalam menghadapi ketatnya persaingan, perlu dipersiapkan sumber daya manusia yang kompeten. Mahasiswa diharapkan siap secara akademik, berupa kesiapan pengetahuan maupun keterampilan untuk menghadapi persaingan yang semakin ketat. Salah satu cara untuk melihat kompetensi mahasiswa adalah dengan melihat hasil belajar yang dapat direpresentasikan dengan nilai ujian yang ditempuh. Ujian tengah semester (UTS)

	<p>merupakan salah satu bentuk ujian yang menjadi komponen penilaian. Dengan mengetahui nilai UTS, dosen mengetahui sebaran mahasiswa dalam hal kompetensi akademik. Untuk itulah, diperlukan pengelompokan (<i>clustering</i>) menggunakan algoritma <i>k-means</i> sebagai pertimbangan dosen dalam membentuk kelompok belajar mahasiswa berdasarkan <i>cluster</i> nilai UTS. [6]</p>
<p>Hasil</p>	<p>Dari hasil pengolahan terhadap data nilai UTS mahasiswa dengan algoritma <i>k-means</i> untuk <i>clustering</i> menggunakan aplikasi WEKA 3.6, diperoleh dua hasil <i>clustering</i>. Hasil <i>clustering</i> pertama untuk kelas Teknik Informatika-A dari <i>output</i> aplikasi <i>weka 3.6</i> menunjukkan sebanyak 5 (10%) nilai UTS termasuk dalam <i>cluster</i> nilai rendah, sebanyak 17 (33%) nilai UTS mahasiswa termasuk dalam <i>cluster</i> nilai sedang, dan sebanyak 30 (58%) nilai UTS mahasiswa termasuk dalam <i>cluster</i> nilai tinggi.</p> <p>Hasil <i>clustering</i> kedua untuk kelas Teknik Informatika-B dari <i>output weka 3.6</i> menunjukkan sebanyak 12 (24%) nilai UTS termasuk dalam <i>cluster</i> nilai rendah, sebanyak 24 (49%) nilai UTS mahasiswa termasuk dalam <i>cluster</i> nilai sedang, dan sebanyak 13 (27%) nilai UTS termasuk dalam <i>cluster</i> nilai tinggi</p>
<p>Kesimpulan</p>	<p>Berdasarkan hasil dan pembahasan <i>output</i> dari aplikasi <i>weka</i>, maka dapat diambil dua kesimpulan. Pertama, <i>clustering</i> yang terbentuk adalah tiga <i>cluster</i>, yakni <i>cluster</i> nilai UTS tinggi, <i>cluster</i> nilai UTS sedang, dan <i>cluster</i> nilai UTS rendah. Kedua, dari <i>clustering</i> yang terbentuk dari masing-masing kelas, dapat dijabarkan bahwa untuk kelas Teknik</p>

	<p>Informatika-A, sebanyak 5 (10%) nilai UTS termasuk dalam cluster nilai rendah, sebanyak 17 (33%) nilai UTS mahasiswa termasuk dalam cluster nilai sedang, dan sebanyak 30 (58%) nilai UTS mahasiswa termasuk dalam cluster nilai tinggi. Sedangkan untuk Teknik Informatika-B, sebanyak 12 (24%) nilai UTS termasuk dalam cluster nilai rendah, sebanyak 24 (49%) nilai UTS mahasiswa termasuk dalam cluster nilai sedang, dan sebanyak 13 (27%) nilai UTS termasuk dalam cluster nilai tinggi.</p>
Penulis	Ginanjari Abdurrahman
Nama jurnal, Volume, Nomor, Tahun	JUSTINDO, Vol 1, No 2 (2016): JUSTINDO
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian terkait bagaimana melihat hasil belajar yang dapat direpresentasikan dengan nilai ujian yang ditempuh. Ujian tengah semester (UTS) merupakan salah satu bentuk ujian yang menjadi komponen penilaian. Penulis menggunakan algoritma <i>k-means</i> sebagai pertimbangan dosen dalam membentuk kelompok belajar mahasiswa berdasarkan <i>cluster</i> nilai UTS. Data yang digunakan oleh penulis adalah data yang berasal dari nilai UTS mahasiswa semester IV. Dari penelitian yang dilakukan oleh penulis diperoleh dua hasil <i>clustering</i>. Hasil <i>clustering</i> pertama untuk kelas Teknik Informatika-A dari <i>output</i> aplikasi <i>weka 3.6</i> menunjukkan sebanyak 5 (10%) nilai UTS termasuk dalam <i>cluster</i> nilai rendah, sebanyak 17 (33%) nilai UTS mahasiswa termasuk dalam <i>cluster</i> nilai sedang, dan sebanyak 30 (58%) nilai UTS mahasiswa termasuk dalam <i>cluster</i> nilai tinggi.</p> <p>Hasil <i>clustering</i> kedua untuk kelas Teknik Informatika-B dari <i>output weka 3.6</i></p>

	menunjukkan sebanyak 12 (24%) nilai UTS termasuk dalam <i>cluster</i> nilai rendah, sebanyak 24 (49%) nilai UTS mahasiswa termasuk dalam <i>cluster</i> nilai sedang, dan sebanyak 13 (27%) nilai UTS termasuk dalam <i>cluster</i> nilai tinggi
Link URL Jurnal	https://garuda.ristekbrin.go.id/documents/detail/1286509

No	7
Judul Artikel	An Educational Data Mining For Student Academic Prediction Using K-Means Clustering And Naïve Bayes Classifier
Topik	Data Mining
Data	Dataset Kinerja Akademik Mahasiswa
Metode / Algoritma	K-Means Clusterin, Naïve Bayes Classifier
Abstrak	<p>Data mining dalam dunia pendidikan merupakan teknik yang dapat digunakan untuk menganalisis data yang diperoleh dari proses dalam dunia pendidikan untuk memperoleh informasi baru dari data yang dimiliki sehingga berguna untuk meningkatkan kualitas pembelajaran. Salah satu cara yang dapat dilakukan untuk meningkatkan kualitas pembelajaran adalah dengan memprediksi prestasi akademik siswa. Klasifikasi akademik siswa berdasarkan potensi akademik dapat menjadi strategi untuk meningkatkan kelulusan mahasiswa, meningkatkan prestasi belajar, dan juga untuk pengelolaan data akademik siswa yang lebih baik. Penelitian ini mengusulkan penggabungan metode K-Means clustering data mining dan Naïve Bayes classifier (K-Means Bayes) untuk hasil yang lebih baik dalam pengolahan data kinerja akademik siswa. Data diambil dari dataset Kinerja Akademik Mahasiswa yang</p>

	<p>digunakan sebagai test case. Jumlah data yang digunakan dalam penelitian ini adalah 131 data dan 21 atribut. Akurasi hasil yang diperoleh dari kombinasi metode yang diusulkan adalah 97,44%. Meskipun penentuan centroid awal pada metode K-Means dilakukan secara acak, namun dampaknya dapat dikurangi dengan menambahkan metode Naïve Bayes Classifier yang menghasilkan nilai akurasi yang lebih baik, sehingga meningkatkan akurasi metode yang diusulkan. Dibandingkan dengan metode K-Means dan Naïve Bayes, metode yang diusulkan meningkatkan akurasi sekitar 27% dari algoritma Naïve Bayes dan sekitar 23% dari algoritma K-Means. Dengan hasil yang diperoleh, dapat disimpulkan bahwa metode yang diusulkan dapat meningkatkan prediksi data prestasi akademik siswa. [7]</p>
<p>Hasil</p>	<p>Dari hasil penelitian hasil akurasi yang diperoleh dari kombinasi metode yang diusulkan adalah 97,44%. Meskipun penentuan centroid awal pada metode K-Means dilakukan secara acak, namun dampaknya dapat dikurangi dengan menambahkan metode Naïve Bayes Classifier yang menghasilkan nilai akurasi yang lebih baik, sehingga meningkatkan akurasi metode yang diusulkan. Dibandingkan dengan metode K-Means dan Naïve Bayes, metode yang diusulkan meningkatkan akurasi sekitar 27% dari algoritma Naïve Bayes dan sekitar 23% dari algoritma K-Means.</p>
<p>Kesimpulan</p>	<p>Penelitian ini mengusulkan kombinasi teknik data mining klasterisasi K-Means dan klasifikasi Naive Bayes (K-Means Bayes) pada data</p>

	<p>prestasi akademik siswa untuk menghasilkan akurasi data yang lebih tinggi.</p> <p>Akurasi kombinasi metode yang diusulkan adalah 97,44%. Perhitungan dengan algoritma K-Means dan perhitungan dengan algoritma Naive Bayes, metode yang diusulkan memberikan hasil yang lebih baik. Walaupun penentuan centroid awal pada metode K-Means dilakukan secara acak, namun dampak tersebut dapat dikurangi dengan penambahan metode Naive Bayes Classifier sehingga menghasilkan akurasi yang lebih baik dan meningkatkan akurasi dari metode yang ada.</p> <p>Dibandingkan dengan K-Means dan Naive Bayes, metode yang diusulkan meningkatkan akurasi sekitar 27% dari algoritma Naive Bayes dan sekitar 23% dari algoritma K-Means. Dengan hasil yang diperoleh, dapat disimpulkan bahwa metode yang diusulkan dapat meningkatkan prediksi data kinerja akademik siswa. Penentuan centroid awal pada metode K-Means mempengaruhi bahwa kualitas akurasi pengelompokan bergantung pada centroid awal.</p>
<p>Penulis</p>	<p>Dewi Ayu Nur Wulandari, Riski Annisa, Lestari Yusuf, Titin Prihatin</p>
<p>Nama jurnal, Volume, Nomor, Tahun</p>	<p>Jurnal Pilar Nusa Mandiri Vol 16 No 2 (2020): Publishing Period for September 2020</p>
<p>Ulasan artikel</p>	<p>Pada penelitian ini penulis melakukan penelitian terkait bagaimana meningkatkan kualitas pembelajaran dengan memprediksi prestasi akademik siswa. Penulis menggunakan 2 algoritma klasifikasi yaitu K-Means dan Naive Bayes. Data yang digunakan oleh penulis diambil dari dataset kinerja akademik mahasiswa yang digunakan sebagai test case. Jumlah data</p>

	<p>yang digunakan dalam penelitian ini adalah 131 data dan 21 atribut. Dari penelitian yang dilakukan oleh penulis didapat Akurasi kombinasi metode yang diusulkan adalah 97,44%. Perhitungan dengan algoritma K-Means dan perhitungan dengan algoritma Naive Bayes, metode yang diusulkan memberikan hasil yang lebih baik. Walaupun penentuan centroid awal pada metode K-Means dilakukan secara acak, namun dampak tersebut dapat dikurangi dengan penambahan metode Naive Bayes Classifier sehingga menghasilkan akurasi yang lebih baik dan meningkatkan akurasi dari metode yang ada. Dibandingkan dengan K-Means dan Naive Bayes, metode yang diusulkan meningkatkan akurasi sekitar 27% dari algoritma Naive Bayes dan sekitar 23% dari algoritma K-Means. Dengan hasil yang diperoleh, dapat disimpulkan bahwa metode yang diusulkan dapat meningkatkan prediksi data kinerja akademik siswa. Penentuan centroid awal pada metode K-Means mempengaruhi bahwa kualitas akurasi pengelompokan bergantung pada centroid awal.</p>
Link URL Jurnal	https://garuda.ristekbrin.go.id/documents/detail/2019339

No	8
Judul Artikel	The Uses of Educational Data Mining in Academic Performance Analysis at Higher Education Institutions (Case Study at UNJANI)
Topik	Data Mining
Data	Data akademik mahasiswa jurusan Teknik Informatika
Metode / Algoritma	k-Means, k-Medoids

<p>Abstrak</p>	<p>Pendidikan merupakan suatu hal yang penting dalam kehidupan seseorang, karena dengan memiliki pendidikan yang memadai maka kehidupan seseorang akan lebih baik. Pendidikan dapat diperoleh secara formal melalui lembaga formal yang secara konstruktif memberikan kemampuan seseorang secara akademis. Penelitian ini bertujuan untuk mengetahui kinerja mahasiswa ditinjau dari ranah akademik dan non akademik pada waktu tertentu selama menempuh pendidikan dengan menggunakan teknik data mining (DM) yang diarahkan pada analisis data akademik. Kinerja akademik disampaikan melalui model data mining terintegrasi Educational Data Mining (EDM), di mana teknik yang digunakan meliputi klasifikasi (ID3, SVM), clustering (k-Means, k-Medoids), aturan asosiasi (Apriori) dan deteksi anomali (DBSCAN). Dataset yang digunakan adalah data akademik berupa hasil belajar selama periode waktu tertentu. Hasil EDM dapat digunakan untuk analisis terkait kinerja akademik yang dapat digunakan untuk pengambilan keputusan strategis dalam manajemen akademik di perguruan tinggi. Hasil penelitian ini menunjukkan bahwa penggunaan beberapa teknik dalam data mining secara bersama-sama dapat memaksimalkan kemampuan menganalisis kinerja akademik dengan sumber data yang sama dan menghasilkan pola analisis yang berbeda. [8]</p>
<p>Hasil</p>	<p>Dari hasil penelitian ini proses EDM dapat dilakukan dalam memperoleh pengetahuan. Proses clustering menggunakan teknik k-Means dengan nilai k ideal = 2, setelah terlebih dahulu</p>

	<p>dilakukan pengukuran SSE (sum square of error) = 18.1388, hasil clustering terdapat 2 (dua) kelompok pola akademik mahasiswa. Algoritma k-Means memberikan hasil yang lebih baik jika dibandingkan dengan algoritma k-Medoids. Teknik lain yang digunakan adalah teknik klasifikasi dengan algoritma pohon keputusan, dimana hasil dari teknik ini adalah terbentuknya pola aturan bersyarat (if-then-else) dengan menentukan atribut sebagai akar awal untuk menghitung nilai entropi. Dengan membandingkan kedua algoritma pada klasifikasi (ID3 dan SVM) dapat diketahui bahwa algoritma ID3 memiliki performansi yang lebih baik, hal ini dapat dipengaruhi oleh dataset yang digunakan. Untuk menentukan anomali data dari data akademik menggunakan algoritma DBSCAN, algoritma ini dapat mendeteksi anomali data akademik dengan nilai epsilon 0,6 dan nilai min poin 10.</p>
Kesimpulan	<p>Educational Data Mining (EDM) telah menjadi alat yang memadai untuk digunakan sebagai pendukung pengambilan keputusan strategis, terutama di institusi pendidikan tinggi. EDM merupakan konsep yang relatif baru digunakan untuk akuisisi pengetahuan di bidang pendidikan melalui teknik dalam data mining. Teknik yang digunakan dalam penelitian ini meliputi clusterization, klasifikasi, aturan asosiasi dan deteksi anomali. Teknik dalam EDM dapat digunakan sebagai cara untuk menganalisis pola yang terbentuk menjadi pengetahuan baru dalam pengelolaan pendidikan di perguruan tinggi dengan menggunakan kumpulan data akademik yang dihasilkan melalui berbagai sistem</p>

	informasi yang tersedia di institusi pendidikan tinggi.
Penulis	Yulison Herry Chrisnanto , Gunawan Abdullah
Nama jurnal, Volume, Nomor, Tahun	Matrix : Jurnal Manajemen Teknologi dan Informatika Vol 11 No 1 (2021): MATRIX - Jurnal Manajemen Teknologi dan Informatika
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian terkait mengetahui kinerja mahasiswa ditinjau dari ranah akademik dan non akademik pada waktu tertentu selama menempuh pendidikan. Penulis menggunakan model data mining terintegrasi Educational Data Mining (EDM), di mana teknik yang digunakan meliputi klasifikasi (ID3, SVM), clustering (K-Means, K-Medoids), aturan asosiasi (Apriori) dan deteksi anomali (DBSCAN). Penulis menggunakan Dataset akademik berupa hasil belajar selama periode waktu tertentu. Dalam penelitian yang dilakukan didapat proses clustering menggunakan teknik k-Means dengan nilai k ideal = 2, setelah terlebih dahulu dilakukan pengukuran SSE (sum square of error) = 18.1388, hasil clustering terdapat 2 (dua) kelompok pola akademik mahasiswa. Algoritma k-Means memberikan hasil yang lebih baik jika dibandingkan dengan algoritma k-Medoids. Teknik lain yang digunakan adalah teknik klasifikasi dengan algoritma pohon keputusan, dimana hasil dari teknik ini adalah terbentuknya pola aturan bersyarat (if-then-else) dengan menentukan atribut sebagai akar awal untuk menghitung nilai entropi, dimana teknik dalam EDM dapat digunakan sebagai cara untuk menganalisis pola yang terbentuk menjadi</p>

	pengetahuan baru dalam pengelolaan pendidikan di perguruan tinggi dengan menggunakan kumpulan data akademik yang dihasilkan melalui berbagai sistem informasi yang tersedia di institusi pendidikan tinggi.
Link URL Jurnal	https://garuda.ristekbrin.go.id/documents/detail/2027066

No	9
Judul Artikel	Classification Analysis Using C4.5 Algorithm To Predict The Level of Graduation of Nurul Falah Pekanbaru High School Students
Topik	Data Mining
Data	Data siswa 2016/2017 terdiri dari dua jurusan yaitu IPA, IPS dan USBN
Metode / Algoritma	C4.5 algorithm.
Abstrak	<p>Sekolah sebagai salah satu proses penyelenggaraan pendidikan formal dituntut untuk melaksanakan proses pembelajaran secara optimal untuk menghasilkan siswa yang berkualitas. Mengenai proses penelitian yang dilakukan untuk memprediksi tingkat kelulusan siswa SMA Nurul Falah dengan menggunakan metode Decision Tree. Data yang digunakan dalam penelitian ini adalah data mahasiswa dengan menggunakan kriteria nama mahasiswa, jurusan, rata-rata raport semester satu (I), dua (II), tiga (III), empat (IV), lima (V), dan nilai rata-rata Ujian Sekolah Standar Nasional (USBN). Data tersebut kemudian dikelola menggunakan software Rapidminer 5.3 untuk memudahkan dalam memprediksi tingkat kelulusan siswa. Aplikasi data mining digunakan untuk memprediksi tingkat kelulusan dengan menggunakan metode pohon keputusan dan</p>

	<p>algoritma C4.5 sebagai pendukung serta untuk mengetahui informasi tingkat kelulusan siswa SMA Nurul Falah. Penelitian ini bertujuan untuk memprediksi tingkat kelulusan siswa agar mendapatkan informasi yang bermanfaat dan pihak sekolah dapat membuat kebijakan di tahun mendatang.</p>
<p>Hasil</p>	<p>Dari hasil dapat diketahui bahwa atribut gain tertinggi adalah semester 4 yaitu sebesar 0,31128. Ada dua nilai atribut dari semester 4 yaitu B dan C. Dari kedua nilai atribut nilai atribut C sudah mengklasifikasi kasus yaitu keputusan “Tidak Lulus”, tetapi untuk nilai atribut B masih perlu dihitung lagi. Dengan demikian 4 semester bisa menjadi simpul cabang ke-2 nilai atribut semester B. Berikutnya bawah adalah keseluruhan formasi decision tree. [9]</p>
<p>Kesimpulan</p>	<p>Melalui penelitian ini simpulkan bahwa:</p> <ol style="list-style-type: none"> 1. Pengolahan data mining menggunakan metode klasifikasi dan algoritma C4.5 dapat memprediksi tingkat kelulusan mahasiswa dengan dua kategori yaitu lulus dan tidak lulus. lulus, dan atribut yang paling berpengaruh dalam hasil prediksi adalah semester tiga. 2. Analisis data siswa untuk memprediksi kelulusan siswa tersebut telah berhasil dilakukan dengan menggunakan metode klasifikasi dan algoritma C4.5. 3. Dengan demikian hal-hal yang dapat dijadikan masukan berdasarkan hasil penelitian ini adalah sebaiknya menambahkan bentuk pohon secara visual tentang atribut nilai Gain tertinggi sampai proses perhitungan selesai sehingga dapat

	diketahui atribut apa saja yang menempati root, branch, leaf posisi..
Penulis	Fana Wizal, Bayu Febriadi1
Nama jurnal, Volume, Nomor, Tahun	IJISTECH Vol 2, No 2 (2019): May
Ulasan artikel	<p>Pada peneletian ini penulis melakukan peneletian mengenai memprediksi tingkat kelulusan siswa SMA Nurul Falah dengan menggunakan metode Decision Three. Penulis menggunakan metode Decision Three dan algoritma C4.5 sebagai pendukung serta untuk mengetahui informasi tingkat kelulusan siswa SMA Nurul Falah. Data yang digunakan oleh penulis ini adalah data mahasiswa dengan menggunakan kriteria nama mahasiswa, jurusan, rata-rata raport semester satu (I), dua (II), tiga (III), empat (IV), lima (V), dan nilai rata-rata Ujian Sekolah Standar Nasional (USBN). Data tersebut kemudian dikelola menggunakan software Rapidminer 5.3 untuk memudahkan dalam memprediksi tingkat kelulusan siswa.</p> <p>Dari penelitian yang dilakukan oleh penulis didapat bahwa atribut gain tertinggi adalah semester 4 yaitu sebesar 0,31128. Ada dua nilai atribut dari semester 4 yaitu B dan C. Dari kedua nilai atribut nilai atribut C sudah mengklasifikasi kasus yaitu keputusan “Tidak Lulus”, tetapi untuk nilai atribut B masih perlu dihitung lagi. Dengan demikian 4 semester bisa menjadi simpul cabang ke-2 nilai atribut semester B. Berikutnya bawah adalah keseluruhan formasi decision tree.</p>
Link URL Jurnal	https://garuda.ristekbrin.go.id/documents/detail/1905075

No	10
Judul Artikel	The Decision Tree c5.0 Classification Algoritma For Prediction Student Academic Performance
Topik	Data Mining
Data	-
Metode / Algoritma	Algoritma C5.0
Abstrak	<p>Kinerja siswa sering digunakan sebagai tolak ukur dan keaktifan siswa sering digunakan sebagai kriteria seberapa baik kinerja siswa secara akademik di sekolah. Dimana dalam penelitian ini akan berusaha mencari tahu apakah keaktifan seorang siswa dapat memprediksi kinerja akademiknya. Data yang digunakan adalah dataset pendidikan yang dikumpulkan menggunakan learning management system (LMS), yang merupakan alat pelacak aktivitas pelajar yang terhubung oleh internet. Data ini memiliki variabel numerik dan kategorikal, sehingga diperlukan algoritma yang tepat untuk mengklasifikasikan data secara akurat dan memastikan validitas data. Dalam penelitian ini, algoritma C.50 digunakan untuk menguji data, di mana data dibagi menjadi data pelatihan sebesar 75% dan pengujian data sebesar 25%. Dan hasil dari data yang diuji, akurasi 71,667% diperoleh. [10]</p>
Hasil	<p>Dari hasil penelitian ini akurasi yang dicapai oleh algoritma C5.0 ini mencapai 71.667% untuk prediksi performa akademik siswa dari faktor keaktifan siswa di dalam kelas maupun di luar kelas yang mencakup tingkat keseringan dalam mengangkat tangan, keseringan berpartisipasi dalam diskusi, insiatif siswa untuk belajar di luar</p>

	sekolah, dan absensi dari siswa itu sendiri pada data testing.
Kesimpulan	<p>Kesimpulan yang dapat ditarik dari penelitian ini adalah penggunaan algoritma <i>Decision Tree</i> C5.0 ini cukuplah efektif untuk memprediksi performa akademik siswa dimana tercapai akurasi sebesar 71.667%.</p> <p>Kesimpulan lainnya yang dapat ditarik dari <i>decision tree</i> yang telah ditunjukkan adalah seorang siswa memang diperlukan untuk aktif dan selalu datang ke sekolah agar tidak tertinggal pelajaran untuk mendapatkan nilai yang memuaskan. Akan tetapi, memang ada saja siswa yang tidak aktif namun mempunyai nilai yang memuaskan ataupun sebaliknya yang sangat aktif namun mempunyai nilai yang kurang memuaskan. Namun, penelitian ini telah membuktikan bahwa sebagian besar yang aktif memang dapat mendapatkan nilai yang memuaskan.</p>
Penulis	Natanael Benediktus, Raymond Sunardi Oetama
Nama jurnal, Volume, Nomor, Tahun	Jurnal ULTIMATICS Vol 12 No 1 (2020): ULTIMATICS
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian mengenai bagaimana mencari tahu apakah keaktifan seorang siswa dapat memprediksi kinerja akademiknya. Dalam penelitian ini penulis menggunakan algoritma C.50 yang digunakan untuk menguji data, di mana data dibagi menjadi data pelatihan sebesar 75% dan pengujian data sebesar 25%. Dan hasil dari data yang diuji, akurasi 71,667% diperoleh. Data yang digunakan oleh penulis adalah dataset pendidikan yang dikumpulkan menggunakan</p>

	<p>learning management system (LMS), yang merupakan alat pelacak aktivitas pelajar yang terhubung oleh internet. Data ini memiliki variabel numerik dan kategorikal, sehingga diperlukan algoritma yang tepat untuk mengklasifikasikan data secara akurat dan memastikan validitas data. Dari penelitian yang dilakukan oleh penulis di dapat akurasi yang dicapai oleh algoritma C5.0 ini mencapai 71.667% untuk prediksi performa akademik siswa dari faktor keaktifan siswa di dalam kelas maupun di luar kelas yang mencakup tingkat keseringan dalam mengangkat tangan, keseringan berpartisipasi dalam diskusi, inisiatif siswa untuk belajar di luar sekolah, dan absensi dari siswa itu sendiri pada data testing. Kesimpulan lainnya yang dapat ditarik dari <i>decision tree</i> yang telah ditunjukkan adalah seorang siswa memang diperlukan untuk aktif dan selalu datang ke sekolah agar tidak tertinggal pelajaran untuk mendapatkan nilai yang memuaskan.</p>
Link URL Jurnal	https://garuda.ristekbrin.go.id/documents/detail/1729407

No	11
Judul Artikel	Application of Data Mining Classification Method for Student Graduation Prediction Using K-Nearest Neighbor (K-NN) Algorithm
Topik	Data mining
Data	Data kelulusan mahasiswa tahun 2010-2012
Metode / Algoritma	K-Nearest Neighbor (K-NN) Algorithm
Abstrak	<p>Tingkat kelulusan mahasiswa merupakan salah satu indikator untuk meningkatkan akreditasi suatu program studi. Diperlukan pemantauan dan evaluasi terhadap kecenderungan kelulusan mahasiswa, tepat waktu atau tidak. Salah satunya adalah memprediksi tingkat kelulusan dengan memanfaatkan teknik data mining. Metode Klasifikasi Data Mining yang digunakan adalah algoritma K-Nearest Neighbor (K-NN). Data yang digunakan berasal dari data siswa, data nilai siswa, dan data kelulusan siswa tahun ajaran 2010-2012 sebanyak 2.189 record. Atribut yang digunakan adalah jenis kelamin, asal sekolah, program studi IP Semester 1-6. Hasil penelitian menunjukkan bahwa metode K-NN menghasilkan akurasi yang tinggi sebesar 89,04%. [11]</p>
Hasil	<p>Dari hasil penelitian yang telah dilakukan menggunakan dataset kelulusan yang terdiri dari data training dan data testing, dengan menerapkan metode K-optimal pada 2 jenis dataset menghasilkan perolehan akurasi nilai pada data latih sebesar 88,16% pada pengujian dengan nilai k=1 sedangkan pengujian data pengujian dengan nilai k=1 menghasilkan akurasi sebesar 89,04%. Dapat juga dianalisis bahwa nilai k mempengaruhi nilai akurasi dari setiap dataset test. Hal ini dibuktikan dengan</p>

	<p>nilai akurasi yang dihasilkan pada pengujian pengujian ke 1 sampai 3, dimana setiap percobaan menggunakan nilai k yang berbeda yaitu 1, 3, dan 5. Pada pengujian pengujian data, training menghasilkan nilai akurasi tertinggi pada implementasi nilai $K = 1$ yaitu 88,16% sedangkan nilai akurasi terendah dihasilkan pada pengujian dengan nilai $k = 5$, maka dapat disimpulkan bahwa nilai akurasi tertinggi kemungkinan implementasi nilai K terkecil.</p>
Kesimpulan	<p>Kesimpulan dari penelitian yang dilakukan adalah, total gabungan data asli berjumlah 2.189 Record, kemudian dilakukan pembersihan dan Integrasi data diperoleh sebanyak 1.736 Record. Dari 1.736 Record yang terbagi menjadi 2 dataset yaitu data training dan data testing, data training berjumlah 667 Record dari mahasiswa angkatan 2011-2011 yang lulus baik tepat waktu maupun tidak tepat waktu, dan data testing berjumlah 867 Record mahasiswa angkatan 2010-2012 yang lulus. belum lulus.</p>
Penulis	Mohammad Imron, Satia Angga Kusumah
Nama jurnal, Volume, Nomor, Tahun	International Journal of Informatics and Informations Systems Vol 1, No 1: September 2018
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian mengenai memprediksi tingkat kelulusan dengan memanfaatkan teknik data mining. Metode Klasifikasi Data Mining yang digunakan oleh penulis adalah algoritma K-Nearest Neighbor (K-NN). Data set yang digunakan oleh penulis berasal dari data siswa, data nilai siswa, dan data kelulusan siswa tahun ajaran 2010-2012 sebanyak 2.189 record. Atribut yang digunakan adalah jenis kelamin, asal sekolah, program studi</p>

	<p>IP Semester 1-6. Dari penelitian yang dilakukan di dapat dataset kelulusan yang terdiri dari data training dan data testing, dengan menerapkan metode K-optimal pada 2 jenis dataset menghasilkan perolehan akurasi nilai pada data latih sebesar 88,16% pada pengujian dengan nilai k=1 sedangkan pengujian data pengujian dengan nilai k=1 menghasilkan akurasi sebesar 89,04%. Dapat juga dianalisis bahwa nilai k mempengaruhi nilai akurasi dari setiap dataset test. Hal ini dibuktikan dengan nilai akurasi yang dihasilkan pada pengujian ke 1 sampai 3, dimana setiap percobaan menggunakan nilai k yang berbeda yaitu 1, 3, dan 5. Pada pengujian pengujian data, training menghasilkan nilai akurasi tertinggi pada implementasi.</p>
Link URL Jurnal	https://garuda.ristekbrin.go.id/documents/detail/1849889

No	12
Judul Artikel	Naive Bayes Algorithm for Classification of Student Majors Specialization
Topik	Data Mining
Data	-
Metode / Algoritma	Naive Bayes Algorithm
Abstrak	<p>Jurusan penting dalam menentukan peminatan mahasiswa. Jika terjadi kesalahan dalam pengarahan siswa tentu akan mempengaruhi pendidikan siswa selanjutnya. Di SMA Negeri 1 Kampar Timur terdapat dua jurusan yaitu IPA dan IPS. Untuk menentukan jurusan tersebut perlu acuan nilai rata-rata nilai mahasiswa semester 3 sampai semester 5 yang meliputi nilai rata-rata Pendidikan Agama Islam, Bahasa Indonesia, Pendidikan Kewarganegaraan,</p>

	<p>Bahasa Inggris, IPA, IPS, dan Matematika. Algoritma Naive Bayes merupakan algoritma yang dapat digunakan dalam mengklasifikasikan jurusan yang terdapat di SMA Negeri 1 Kampar Timur. Untuk menentukan klasifikasi jurusan di SMA Negeri 1 Kampar Timur digunakan data latih dan data tes masing-masing sebesar 70% dan 30%. Data ini akan diuji akurasi dengan menggunakan confusion matrix, dan menghasilkan akurasi yang cukup tinggi yaitu 96,19%. Dengan akurasi yang tinggi tersebut maka algoritma Naive Bayes sangat cocok digunakan dalam menentukan arah siswa di SMA Negeri 1 Kampar Timur. [12]</p>
<p>Hasil</p>	<p>Dari proses perhitungan data mining menggunakan algoritma Naive Bayes dan tingkat keakuratannya, dihasilkan informasi baru yaitu perhitungan data mining berdasarkan data uji siswa SMA Negeri 1 Kampar Timur, menunjukkan siswa yang memilih IPS sebanyak 51 orang, dengan ketepatan nilai 100%, dan sebanyak 50 siswa yang memilih masuk IPA, dan ditambah 4 orang dari siswa yang memilih IPS diprediksi masuk IPA dengan presisi kelas 92,59%. Dari hasil ini diperoleh akurasi sebesar 96,19%, dengan nilai true recall kelas IPS 92,73% dan true IPA class recall 100%.</p>
<p>Kesimpulan</p>	<p>Berdasarkan percobaan dan analisis yang kami lakukan, diperoleh kesimpulan bahwa peminatan jurusan menggunakan algoritma Naive Bayes menghasilkan tingkat akurasi yang tinggi dalam pemilihan jurusan sebagai perencanaan untuk menentukan kursus berikutnya. Dari proses perhitungan data mining menggunakan algoritma Naive Bayes dan tingkat akurasi,</p>

	<p>dihasilkan informasi perhitungan data mining dengan presisi kelas 100%, dan sebanyak 50 siswa yang memilih untuk masuk IPA, dan ditambah 4 siswa yang memilih untuk masuk IPA. memilih IPS diprediksi masuk IPA dengan kelas presisi 92,59%. Dari hasil ini diperoleh akurasi sebesar 96,19%, dengan nilai true recall kelas IPS 92,73% dan true IPA class recall 100%. dapat disimpulkan bahwa metode pemilihan jurusan menggunakan algoritma ini sudah tepat. Oleh karena itu, penelitian ini dapat digunakan untuk keperluan pemilihan jurusan pada tahun-tahun berikutnya, karena model yang dihasilkan mampu diprediksi secara akurat dengan.</p>
Penulis	Astia Weni Syaputri, Erno Irwandi, Mustakim
Nama jurnal, Volume, Nomor, Tahun	Journal of Intelligent Computing and Health Informatics (JICHI) Vol 1, No 1 (2020)
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian mengenai menentukan peminatan atau jurusan mahasiswa. Penulis menggunakan metode Algoritma Naive Bayes yang dapat digunakan dalam mengklasifikasikan jurusan tersebut. Dari peneletian yang dilakukan oleh penulis didapat hasil informasi baru yaitu perhitungan data mining berdasarkan data uji siswa SMA Negeri 1 Kampar Timur, menunjukkan siswa yang memilih IPS sebanyak 51 orang, dengan ketepatan nilai 100%, dan sebanyak 50 siswa yang memilih masuk IPA, dan ditambah 4 orang dari siswa yang memilih IPS diprediksi masuk IPA dengan presisi kelas 92,59%. Dari hasil ini diperoleh akurasi sebesar 96,19%, dengan nilai true recall kelas IPS 92,73% dan true IPA class recall 100%, sehigga diperoleh kesimpulan</p>

	bahwa peminatan jurusan menggunakan algoritma Naive Bayes menghasilkan tingkat akurasi yang tinggi dalam pemilihan jurusan sebagai perencanaan untuk menentukan kursus berikutnya.
Link URL Jurnal	https://garuda.ristekbrin.go.id/documents/detail/1837643

No	13
Judul Artikel	Classification and Prediction of Student GPA Using K-Means Clustering Algorithm to Assist Student Admission Process
Topik	Data Mining
Data	Data siswa angkatan 2008-2017
Metode / Algoritma	K-Means Algorithm
Abstrak	<p>Penerimaan mahasiswa di perguruan tinggi bertujuan untuk menyeleksi kandidat terbaik yang akan berprestasi dan menyelesaikan studinya tepat waktu. Banyak faktor yang harus diperhatikan dalam penerimaan mahasiswa baru. Untuk membantu proses tersebut, diperlukan model cerdas untuk melihat siswa yang berpotensi berprestasi tinggi, serta mengidentifikasi siswa yang berpotensi bermasalah sedini mungkin. Penelitian ini menggunakan K-means clustering untuk memprediksi nilai rata-rata (IPK) siswa berdasarkan profil siswa, seperti status sekolah menengah dan lokasi, nilai tes masuk universitas dan kompetensi bahasa Inggris. Data siswa angkatan 2008-2017 digunakan untuk membuat dua cluster menggunakan algoritma clustering K-means. Dua centroid dari cluster digunakan untuk mengklasifikasikan semua data menjadi dua kelompok: IPK tinggi dan IPK rendah. Kami</p>

	menggunakan data dari angkatan 2018 sebagai data uji. Performa prediksi diukur dengan menggunakan akurasi, presisi dan recall. [13]
Hasil	Berdasarkan hasil analisis dari penelitian yang dilakukan, metode pengelompokan K-means adalah 78,59% akurat di antara siswa penerimaan berbasis prestasi dan 94,627% di antara siswa penerimaan reguler.
Kesimpulan	Prediksi yang melibatkan mahasiswa masuk berbasis merit memiliki nilai akurasi prediksi yang lebih rendah dibandingkan dengan yang melibatkan mahasiswa reguler karena model clustering untuk data penerimaan berbasis merit adalah $K = 3$, tetapi untuk prediksi, asumsinya adalah $K = 2$.
Penulis	Raden Gunawan Santosa, Yuan Lukito, Antonius Rachmat Chrismanto
Nama jurnal, Volume, Nomor, Tahun	Journal of Information Systems Engineering and Business Intelligence Vol 7, No 1 (2021): April
Ulasan artikel	Pada penelitian ini penulis melakukan penelitian mengenai Klasifikasi dan Prediksi IPK mahasiswa untuk membantu proses penerimaan mahasiswa baru. Penulis menggunakan algoritma K-means clustering untuk memprediksi nilai rata-rata (IPK) siswa berdasarkan profil siswa, seperti status sekolah menengah dan lokasi, nilai tes masuk universitas dan kompetensi bahasa Inggris. Dan data yang digunakan adalah Data siswa angkatan 2008-2017 digunakan untuk membuat dua cluster menggunakan algoritma clustering K-means. Dua centroid dari cluster digunakan untuk mengklasifikasikan semua data menjadi dua kelompok: IPK tinggi dan IPK rendah. Kami menggunakan data dari angkatan 2018 sebagai

	<p>data uji. Performa prediksi diukur dengan menggunakan akurasi, presisi dan recall. Dari penelitian yang dilakukan didapat metode pengelompokan K-means adalah 78,59% akurat di antara siswa penerimaan berbasis prestasi dan 94,627% di antara siswa penerimaan reguler. Prediksi yang melibatkan mahasiswa masuk berbasis merit memiliki nilai akurasi prediksi yang lebih rendah dibandingkan dengan yang melibatkan mahasiswa reguler karena model clustering untuk data penerimaan berbasis merit adalah $K = 3$, tetapi untuk prediksi, asumsinya adalah $K = 2$.</p>
Link URL Jurnal	https://garuda.ristekbrin.go.id/documents/detail/2008492

No	14
Judul Artikel	Implementation Naïve Bayes Algorithm for Student Classification Based on Graduation Status
Topik	Data Mining
Data	Data Registrasi Mahasiswa Jurusan Statistika
Metode / Algoritma	Naive Bayes Algorithm
Abstrak	<p>Lama studi di perguruan tinggi adalah waktu yang dibutuhkan seorang mahasiswa untuk menyelesaikan studi di perguruan tinggi. Dalam mencapai gelar sarjana normal itu membutuhkan waktu selama empat tahun, namun masih ada mahasiswa yang menyelesaikan studinya di luar batas normal (di atas empat tahun). Hal ini antara lain berpengaruh pada nilai akreditasi institusi. Dalam makalah ini kami menggunakan lima variabel: Indeks Prestasi Kumulatif (IPK), Konsentrasi di Sekolah Menengah Atas, Jenis Kelamin, Partisipasi dalam Pendampingan dan Kota Tempat Tinggal, yang diklasifikasikan</p>

	berdasarkan Status Kelulusan siswa di atas empat tahun dan kurang dari sama dengan empat tahun. Metode yang digunakan untuk klasifikasi waktu belajar siswa adalah algoritma Naive Bayes. [14]
Hasil	Dari hasil penelitian yang dilakukan yaitu menyelidiki klasifikasi mahasiswa berdasarkan Status Kelulusan pada Jurusan Statistika Universitas Islam Indonesia. Dari hasil klasifikasi algoritma Naive Bayes cukup baik dengan nilai akurasi untuk Naive Bayes sebesar 81,18%.
Kesimpulan	Dalam makalah ini, Naive Bayes Classifier telah dibahas sebagai classifier terbaik dalam masalah ini. Dengan demikian, tingkat akurasi pada model klasifikasi algoritma Naive Bayes 81,18%.
Penulis	Ayundyah Kesumawati, Din Waikabu
Nama jurnal, Volume, Nomor, Tahun	International Journal of Applied Business and Information Systems Vol. 1 No. 2 (2017)
Ulasan artikel	Pada penelitian ini penulis melakukan penelitian mengenai klasifikasi mahasiswa berdasarkan status kelulusan. Penulis menggunakan algoritma Naive Bayes klasifikasi. Data yang digunakan oleh penulis adalah data Data Registrasi Mahasiswa Jurusan Statistika yang mengacu pada atribut Indeks Prestasi Kumulatif (IPK), Konsentrasi di Sekolah Menengah Atas, Jenis Kelamin, Partisipasi dalam Pendampingan dan Kota Tempat Tinggal, yang kemudian diklasifikasikan berdasarkan Status Kelulusan siswa di atas empat tahun dan kurang dari sama dengan empat tahun. Dari penelitian yang dilakukan oleh penulis didapat algoritma Naive Bayes Classifier telah dibahas sebagai classifier

	terbaik dalam masalah ini. Dengan demikian, tingkat akurasi pada model klasifikasi algoritma Naive Bayes 81,18%.
Link URL Jurnal	https://garuda.ristekbrin.go.id/documents/detail/1687583

No	15
Judul Artikel	Text Clasification using Naïve Baye Updateable Algorithm In SBMPTN Test Questions
Topik	Data Mining
Data	Data yang digunakan bersumber dari buku kumpulan soal-soal ujian SBMPTN
Metode / Algoritma	Algoritma Naïve Bayes
Abstrak	<p>Klasifikasi dokumen merupakan minat dalam penelitian text mining. Klasifikasi dapat dilakukan berdasarkan topik, bahasa, dan sebagainya. Penelitian ini dilakukan untuk mengetahui bagaimana Naive Bayes Updateable melakukan dalam mengklasifikasikan soal ujian SBMPTN berdasarkan tema. Model kenaikan satu algoritma klasifikasi sering digunakan dalam klasifikasi teks klasifikasi memiliki kemampuan untuk belajar dari data baru memperkenalkan dengan sistem bahkan setelah classifier telah diproduksi dengan data yang ada. Klasifikasi mengklasifikasikan pertanyaan ujian berdasarkan tema bidang studi dengan menganalisis kata kunci yang muncul di soal ujian. Salah satu metode seleksi fitur DF-Thresholding diimplementasikan untuk meningkatkan kinerja klasifikasi. Evaluasi klasifikasi dengan algoritma classifier Naive Bayes menghasilkan akurasi 84,61%. [15]</p>
Hasil	Proses pelatihan data tekstual dilakukan dengan data sebanyak 299 data soal ujian SBMPTN.

	<p>Data yang dibutuhkan untuk proses training ini adalah kombinasi teks soal ujian dengan temanya masing-masing. Teks soal ujian yang telah di-preprocessing untuk setiap kata pada satu kelas dihitung nilai mean dan variansinya. Perhitungan ini dilakukan untuk setiap fitur kata pada satu kelas. Kombinasi nilai mean dan varian inilah yang merupakan hasil dari training data.</p>
Kesimpulan	<p>Pengklasifikasian soal ujian SBMPTN dengan algoritma <i>Naive Bayes</i> yang dilakukan pada penelitian ini dapat berjalan dengan baik, sebagian besar data soal ujian yang diklasifikasikan sesuai dengan pengklasifikasian secara manual dengan nilai akurasi 84,61%. Soal ujian yang gagal diklasifikasikan dengan benar ini disebabkan oleh adanya kata kunci baru yang memiliki kemungkinan sebagai kata kunci penting tetapi tidak dapat digunakan karena belum terdaftar pada data training, selain itu kegagalan klasifikasi juga dapat disebabkan oleh keterbatasan data <i>training</i> yang dimiliki, seperti dapat dilihat pada proses <i>memorizing</i> terdapat beberapa soal ujian yang tidak mampu diingat oleh sistem dikarenakan <i>keyword</i> yang dimiliki kelas sebenarnya masih kurang beragam maupun kurang frekuensinya sehingga mempengaruhi derajat konsentrasi, derajat persebaran dan derajat kontribusinya.</p>
Penulis	Ristu Saptono, Meiyanto Eko Sulistyono, Nur Shobriana Trihabsari
Nama jurnal, Volume, Nomor, Tahun	Telematika : Jurnal Informatika dan Teknologi Informasi Vol 13, No 2 (2016): Edisi Juli 2016
Ulasan artikel	Pada penelitian ini penulis melakukan penelitian mengenai klasifikasi teks pada soal tes

	<p>SBMPTN. Penulis menggunakan algoritma Naive Bayes Updateable selanjutnya melakukan dalam mengklasifikasikan soal ujian SBMPTN berdasarkan tema. Model kenaikan satu algoritma klasifikasi sering digunakan dalam klasifikasi teks klasifikasi memiliki kemampuan untuk belajar dari data baru memperkenalkan dengan sistem bahkan setelah classifier telah diproduksi dengan data yang ada.</p> <p>Penulis menggunakan data yang digunakan bersumber dari buku kumpulan soal-soal ujian SBMPTN. Dari penelitian yang dilakukan didapat pengklasifikasian soal ujian SBMPTN dengan algoritma Naive Bayes yang dilakukan pada penelitian ini dapat berjalan dengan baik, sebagian besar data soal ujian yang diklasifikasikan sesuai dengan pengklasifikasian secara manual dengan nilai akurasi 84,61%. Soal ujian yang gagal diklasifikasikan dengan benar ini disebabkan oleh adanya kata kunci baru yang memiliki kemungkinan sebagai kata kunci penting tetapi tidak dapat digunakan karena belum terdaftar pada data training, selain itu kegagalan klasifikasi juga dapat disebabkan oleh keterbatasan data <i>training</i> yang dimiliki, seperti dapat dilihat pada proses <i>memorizing</i> terdapat beberapa soal ujian yang tidak mampu diingat oleh sistem dikarenakan <i>keyword</i> yang dimiliki kelas sebenarnya masih kurang beragam maupun kurang frekuensinya sehingga mempengaruhi derajat konsentrasi, derajat persebaran dan derajat kontribusinya.</p>
Link URL Jurnal	https://garuda.ristekbrin.go.id/documents/detail/1752584

No	16
Judul Artikel	Determinating Student Interactions in a Virtual Learning Environment Using Data Mining
Topik	Data Mining
Data	Data interaksi siswa dalam Kursus Bahasa Inggris
Metode / Algoritma	Decision Tree Algorithms
Abstrak	<p>Artikel ini berfokus pada penentuan interaksi siswa dalam Kursus Bahasa Inggris Virtual dengan Model Pendidikan Jarak Jauh (DEM) di Universitas Mumbai, India. Untuk tujuan ini, analisis dilakukan pada database siswa selama periode akademik 2015 - 2018 untuk memilih atribut yang diperlukan yang memungkinkan untuk menghasilkan model data mining. Analisis metode penambangan kemudian dilakukan membandingkan masing-masing untuk memilih salah satu yang membantu pengembangan proyek, memilih metode Crisp-dm karena mengandung beberapa fase yang menunjukkan setiap aktivitas yang harus diselesaikan, sehingga menjadi praktis panduan. Selain itu, analisis komparati dikembangkan dengan mempertimbangkan fitur alat penambangan data di mana RapidMiner dipilih untuk melakukan proses menggunakan beberapa algoritma bersama dengan data siswa yang dibagi menjadi dua set untuk pelatihan dan validasi, menghasilkan Decisin Three sebagai algoritma terbaik untuk tujuan tersebut karena mengklasifikasikan instance dengan benar dengan margin kesalahan minimum. [16]</p>
Hasil	<p>Dari hasil penelitian ini untuk menentukan interaksi siswa, hasil terbaik dari algoritma dipertimbangkan, yang dianalisis dalam evaluasi model (Tabel 1) mendapatkan algoritma Pohon Keputusan sebagai hasil terbaik, menyajikan klasifikasi yang baik</p>

	dengan 92,9% dan margin kesalahan yang lebih rendah dalam validasi model dengan 7,1%.
Kesimpulan	<p>Melalui hasil, kesimpulan berikut diperoleh: Data mining sangat penting dalam bidang pendidikan karena membantu mengekstrak informasi yang tersembunyi dalam data sedemikian rupa sehingga memungkinkan analisis dan generasi pengetahuan baru untuk menentukan tingkat dari interaksi siswa. RapidMiner adalah alat penambangan data yang kuat karena berisi add-on yang memungkinkan untuk menggunakan berbagai algoritma dalam alat ini dan alat lainnya dengan operator yang membantu dalam pengembangan proses pembuatan model yang berlaku untuk analisis data.</p> <p>Untuk menentukan tingkat interaksi dalam kursus bahasa Inggris, algoritma klasifikasi yang berbeda diterapkan, menyajikan hasil terbaik di Pohon Keputusan karena algoritma ini memperoleh margin kesalahan paling sedikit selama klasifikasi data dari interaksi dalam kursus (tugas, tes, sumber daya), dan data pribadi, kelembagaan dan sosial ekonomi. Melalui model data mining ini dapat diketahui bahwa interaksi siswa di English Virtual Course menghasilkan tingkat rata-rata dengan persentase 69%, dan faktor yang paling mempengaruhi model adalah interaksi siswa dalam ujian, tugas, sumber daya, status perkawinan, dan status pekerjaan.</p>
Penulis	Amelec Viloría, Jorge Rodríguez López, Karen Payaresc, Carlos Vargas-Mercado, Sonia Ethel Duran, Hugo Hernández-Palma, Mónica Arrozola David
Nama jurnal, Volume, Nomor, Tahun	Procedia Computer Science Volume 155, 2019, Pages 587-592
Ulasan artikel	Pada penelitian ini penulis melakukan penelitian mengenai penentuan interaksi siswa dalam

lingkungan pembelajaran virtual menggunakan data mining. Data yang digunakan penulis adalah data interaksi siswa dalam Kursus Bahasa Inggris. Analisis metode penambangan yang dilakukan oleh penulis adalah membandingkan masing-masing untuk kemudian memilih salah satu yang membantu pengembangan proyek, memilih metode Crisp-dm karena mengandung beberapa fase yang menunjukkan setiap aktivitas yang harus diselesaikan, sehingga menjadi praktis panduan. Selain itu, analisis komparati dikembangkan dengan mempertimbangkan fitur alat penambangan data di mana RapidMiner dipilih untuk melakukan proses menggunakan beberapa algoritma bersama dengan data siswa yang dibagi menjadi dua set untuk pelatihan dan validasi, menghasilkan Decisin Three sebagai algoritma terbaik untuk tujuan tersebut karena mengklasifikasikan instance dengan benar dengan margin kesalahan minimum. Dari penelitian yang dilakukan oleh penulis didapat hasil terbaik dari algoritma dipertimbangkan algoritma Decision Three sebagai hasil terbaik, menyajikan klasifikasi yang baik dengan 92,9% dan margin kesalahan yang lebih rendah dalam validasi model dengan 7,1%. Dan untuk menentukan tingkat interaksi dalam kursus bahasa Inggris, algoritma klasifikasi yang berbeda diterapkan, menyajikan hasil terbaik di Decision Three karena algoritma ini memperoleh margin kesalahan paling sedikit selama klasifikasi data dari interaksi dalam kursus (tugas, tes, sumber daya), dan data pribadi, kelembagaan dan sosial ekonomi. Melalui model data mining ini dapat diketahui bahwa interaksi siswa di English Virtual Course menghasilkan tingkat rata-rata dengan persentase 69%, dan faktor yang paling

	mempengaruhi model adalah interaksi siswa dalam ujian, tugas, sumber daya, status perkawinan, dan status pekerjaan.
Link URL Jurnal	https://www.sciencedirect.com/science/article/pii/S1877050919309962

No	17
Judul Artikel	The Third Information Systems International Conference A Review on Predicting Student's Performance using Data Mining Techniques
Topik	Data Mining
Data	-
Metode / Algoritma	Decision Tree, Neural Network , Support Vector Machine, K-Nearest Neighbor
Abstrak	<p>Memprediksi kinerja siswa menjadi lebih menantang karena volume data yang besar dalam database pendidikan. Saat ini di Malaysia, kurangnya sistem yang ada untuk menganalisis dan memantau kemajuan dan kinerja siswa tidak ditangani. Ada dua alasan utama mengapa ini terjadi. Pertama, studi tentang metode prediksi yang ada masih belum cukup untuk mengidentifikasi metode yang paling cocok untuk memprediksi kinerja mahasiswa di institusi Malaysia. Kedua adalah karena kurangnya penyelidikan pada faktor-faktor yang mempengaruhi prestasi siswa dalam kursus tertentu dalam konteks Malaysia. Oleh karena itu, tinjauan pustaka sistematis untuk memprediksi kinerja siswa dengan menggunakan teknik data mining diusulkan untuk meningkatkan prestasi siswa. Tujuan utama dari makalah ini adalah untuk memberikan gambaran tentang teknik data mining yang telah digunakan untuk memprediksi kinerja siswa. Makalah ini juga berfokus pada bagaimana algoritma prediksi dapat digunakan untuk mengidentifikasi atribut yang paling penting</p>

	<p>dalam data siswa. Kami sebenarnya dapat meningkatkan prestasi dan keberhasilan siswa secara lebih efektif dengan cara yang efisien menggunakan teknik data mining pendidikan. Dapat membawa manfaat dan dampak bagi siswa, pendidik dan institusi akademik. [17]</p>
Hasil	<p>Dari hasil analisis terhadap penelitian yang dilakukan Neural Network memiliki akurasi prediksi tertinggi sebesar (98%) diikuti oleh Decision Tree sebesar (91%). Selanjutnya Support Vector Machine dan K-Nearest Neighbor memberikan akurasi yang sama, yaitu (83%). Terakhir, metode yang memiliki akurasi prediksi lebih rendah adalah metode Naive Bayes sebesar (76%). Hasil akurasi prediksi tergantung pada atribut atau fitur yang digunakan selama proses prediksi. Metode Neural Network memberikan akurasi prediksi tertinggi karena pengaruh dari atribut utama. Atribut ini merupakan hibridisasi dari dua fitur, yaitu penilaian internal dan eksternal. Dengan penggunaan satu variabel saja, yaitu penilaian eksternal, akurasinya berkurang (1%). Variabel ketiga yang paling banyak digunakan adalah penilaian internal yang memberikan hasil akurasi (81%). Hal ini menunjukkan bahwa penilaian eksternal, yaitu nilai yang diperoleh dalam ujian akhir, memainkan peran penting dalam memprediksi kinerja siswa. Sedangkan variabel signifikan.</p>
Kesimpulan	<p>Memprediksi kinerja siswa sebagian besar berguna untuk membantu pendidik dan peserta didik meningkatkan proses belajar dan mengajar mereka. Makalah ini telah mengulas penelitian sebelumnya tentang memprediksi kinerja siswa dengan berbagai metode analisis. Sebagian besar peneliti telah</p>

	<p>menggunakan rata-rata nilai kumulatif (CGPA) dan penilaian internal sebagai kumpulan data. Sedangkan untuk teknik prediksi, metode klasifikasi sering digunakan dalam bidang data mining pendidikan. Di bawah teknik klasifikasi, Neural Network dan Decision Tree adalah dua metode yang sangat digunakan oleh para peneliti untuk memprediksi kinerja siswa. Kesimpulannya, meta-analisis dalam memprediksi kinerja siswa telah memotivasi kami untuk melakukan penelitian lebih lanjut untuk diterapkan di lingkungan kami. Ini akan membantu sistem pendidikan untuk memantau kinerja siswa secara sistematis.</p>
Penulis	Amirah Mohamed Shahiria, Wahidah Husaina, Nur'aini Abdul Rashida.
Nama jurnal, Volume, Nomor, Tahun	Procedia Computer Science Volume 72, 2015, Pages 414-422
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitaian mengenai memprediksi kinerja siswa menggunakan teknik data mining. Penulis menggunakan algoritma Decision Tree, Neural Network , Support Vector Machine, K-Nearest Neighbor sebagai analisis dan perbandingan prediksi yang digunakan untuk mengidentifikasi atribut yang paling penting dalam data siswa. Kami sebenarnya dapat meningkatkan prestasi dan keberhasilan siswa secara lebih efektif dengan cara yang efisien menggunakan teknik data mining pendidikan. Dapat membawa manfaat dan dampak bagi siswa, pendidik dan institusi akademik. Dari penelitian yang dilakukan oleh penulis didapat analisis Neural Network memiliki akurasi prediksi tertinggi sebesar (98%) diikuti oleh Decision Tree sebesar (91%). Selanjutnya Support Vector Machine dan K-Nearest Neighbor memberikan akurasi yang sama, yaitu (83%). Terakhir, metode yang memiliki</p>

	<p>akurasi prediksi lebih rendah adalah metode Naive Bayes sebesar (76%). Hasil akurasi prediksi tergantung pada atribut atau fitur yang digunakan selama proses prediksi. Metode Neural Network memberikan akurasi prediksi tertinggi karena pengaruh dari atribut utama. Atribut ini merupakan hibridisasi dari dua fitur, yaitu penilaian internal dan eksternal. Dengan penggunaan satu variabel saja, yaitu penilaian eksternal, akurasinya berkurang (1%). Variabel ketiga yang paling banyak digunakan adalah penilaian internal yang memberikan hasil akurasi (81%). Hal ini menunjukkan bahwa penilaian eksternal, yaitu nilai yang diperoleh dalam ujian akhir, memainkan peran penting dalam memprediksi kinerja siswa. Sedangkan variabel signifikan.</p>
Link URL Jurnal	https://www.sciencedirect.com/science/article/pii/S1877050915036182

No	18
Judul Artikel	An Intelligent Prediction System for Educational Data Mining Based on Ensemble and Filtering approaches
Topik	Data Mining
Data	Dataset pedagogis dari University of Kashmir
Metode / Algoritma	Decision Tree, Naïve Bayes, K-Nearest Neighbor
Abstrak	<p>Pendekatan ensemble seperti boosting didasarkan pada sistem heuristik untuk mengembangkan paradigma prediksi. Teknik pembelajaran ensemble biasanya lebih akurat daripada pengklasifikasi individu untuk menghasilkan prediksi. Untuk alasan ini, terutama dalam penelitian ini beberapa teknik ensemble telah dibahas untuk mendapatkan pengetahuan yang komprehensif tentang metode kunci. Di antara berbagai pendekatan ensemble, peneliti telah mempraktekkan mekanisme boosting</p>

	<p>untuk memprediksi kinerja siswa. Karena penerapan metode ensemble dianggap sebagai fenomena signifikan dalam prosedur klasifikasi dan prediksi, oleh karena itu para peneliti memanfaatkan teknik boosting untuk mengembangkan model pedagogis prediksi yang akurat, mengingat sifat yang diucapkan dan kebaruan metode yang diusulkan dalam penambahan data pendidikan. Pengklasifikasi dasar termasuk Decision Tree, j48, knn dan nave bayes telah dievaluasi pada sistem validasi silang 10 kali lipat. Selain itu, prosedur penyaringan seperti oversampling (SMOTE) under-sampling (Spread subsampling), telah dimanfaatkan untuk memeriksa lebih lanjut setiap perubahan signifikan dalam hasil di antara pengklasifikasi meta dan dasar. Kedua pendekatan ensemble dan penyaringan telah menunjukkan peningkatan substansial dalam memprediksi kinerja siswa daripada penerapan pengklasifikasi konvensional. Selanjutnya, berdasarkan peningkatan hasil, dua model prediksi baru telah diajukan setelah melakukan analisis kinerja pada setiap pendekatan. [18]</p>
<p>Hasil</p>	<p>Decision Tree, K-Nearest Neighbor dan Naive Bayes dengan teknik boosting telah menunjukkan hasil yang luar biasa di semua pengklasifikasi. Pengklasifikasi pembelajaran dasar yaitu. j48 dan nave bayes dengan sistem boosting, yang diilustrasikan pada tabel 5 telah menunjukkan pencapaian penting dalam hasil prediksi saat mengalami teknik pengambilan sampel berlebih. Kinerja kedua classifier pembelajaran meningkat dari 95,32% menjadi 96,44% (j48 dengan boosting) dan 95,04% menjadi 96,06% (nave bayes dengan boosting). Namun, di antara semua ansambel</p>

	<p>pembelajaran dari masing-masing pengklasifikasi dasar yaitu. pohon acak telah menunjukkan akurasi prediksi yang paling penting atas semua pengklasifikasi dengan teknik oversampling dari 89,35% menjadi 92,03%, dan knn dari 93,59% menjadi 94,49%.</p>
Kesimpulan	<p>Tujuan utama dari penyelidikan penelitian ini adalah untuk menguatkan apakah pendekatan penyaringan atau metode ensemble memiliki dampak besar pada akurasi prediksi pengklasifikasi pembelajaran. Dan, akibatnya berdasarkan hasil yang dicapai, mengusulkan sistem prediksi yang lebih baik pada dataset pedagogis. Dengan demikian, dua model prediksi yaitu. pembelajaran tingkat dasar dan pembelajaran tingkat meta telah dikemukakan untuk mewujudkan hasil yang signifikan. Sebelum penerapan metode penyaringan atau ensemble, di antara semua pengklasifikasi pembelajaran individu, naive bayes memperoleh akurasi prediksi yang luar biasa sebesar 95,50%. Karena dataset dalam hal ini tidak seimbang yang akan menghasilkan hasil yang salah dan bias. Oleh karena itu, setelah menggunakan pendekatan penyaringan, para peneliti memeriksa bahwa setiap pengklasifikasi pembelajaran mengalami peningkatan yang signifikan, dan diantara setiap pengklasifikasi dasar, naive bayes mencapai akurasi prediksi yang patut dicatat sebesar 97,15%, dengan teknik over-sampling. Pengklasifikasi knn menunjukkan akurasi prediksi yang cukup besar dengan teknik under-sampling dari pada pengklasifikasi pembelajaran lainnya. Namun, saat memeriksa pendekatan ansambel dan menguatkan metode mana yang memberikan hasil yang lebih baik, para peneliti menemukan bahwa peningkatan</p>

	<p>tanpa dikenakan pendekatan penyaringan apapun, yaitu. prosedur oversampling dan under-sampling menggambarkan akurasi prediksi yang menonjol daripada pengklasifikasi individu. Selain itu, ketika pendekatan ensemble diverifikasi dengan praktik under-sampling dan under-sampling, dalam kedua kasus akurasi prediksi diperkuat dengan jumlah yang signifikan daripada pengklasifikasi individu.</p>
Penulis	Mudasir AshrafMajid Zaman, Muheet Ahmed
Nama jurnal, Volume, Nomor, Tahun	Procedia Computer Science Volume 167, 2020, Pages 1471-1483
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian mengenai sistem prediksi cerdas untuk pendidikan berdasarkan pendekatan ensemble dan filtering. Di antara berbagai pendekatan ensemble, penulis telah mempraktekkan mekanisme boosting untuk memprediksi kinerja siswa. Karena penerapan metode ensemble dianggap sebagai fenomena signifikan dalam prosedur klasifikasi dan prediksi, oleh karena itu para peneliti memanfaatkan teknik boosting untuk mengembangkan model pedagogis prediksi yang akurat, mengingat sifat yang diucapkan dan kebaruan metode yang diusulkan dalam penambahan data pendidikan. Penulis menggunakan algoritma Decision Tree, Naïve Bayes, K-Nearest Neighbor dan Dataset pedagogis dari University of Kashmir. Dari penelitian yang dilakukan oleh penulis didapat Decision Tree, K-Nearest Neighbor dan Naive Bayes dengan teknik boosting telah menunjukkan hasil yang luar biasa di semua pengklasifikasi. Pengklasifikasi pembelajaran dasar yaitu. j48 dan nave bayes dengan sistem boosting, yang diilustrasikan pada tabel 5 telah menunjukkan pencapaian penting dalam hasil prediksi saat mengalami teknik</p>

	<p>pengambilan sampel berlebih. Kinerja kedua classifier pembelajaran meningkat dari 95,32% menjadi 96,44% (j48 dengan boosting) dan 95,04% menjadi 96,06% (nave bayes dengan boosting). Para peneliti memeriksa bahwa setiap pengklasifikasi pembelajaran mengalami peningkatan yang signifikan, dan diantara setiap pengklasifikasi dasar, naive bayes mencapai akurasi prediksi yang patut dicatat sebesar 97,15%, dengan teknik over-sampling. Pengklasifikasi knn menunjukkan akurasi prediksi yang cukup besar dengan teknik under-sampling dari pada pengklasifikasi pembelajaran lainnya. Namun, saat memeriksa pendekatan ansambel dan menguatkan metode mana yang memberikan hasil yang lebih baik, para peneliti menemukan bahwa peningkatan.</p>
Link URL Jurnal	https://www.sciencedirect.com/science/article/pii/S1877050920308243



No	19
Judul Artikel	Using Data Mining Techniques to Guide Academic Programs Design and Assessment
Topik	Data Mining
Data	Data pemetaan antara PEO mereka dan ABET-EAC SO telah diekstraksi.
Metode / Algoritma	Algoritma Apriori
Abstrak	<p>Specifically, itu mengusulkan menggunakan teknik association rule mining untuk menemukan satu set aturan yang mengatur hubungan antara dua komponen inti dari program akademik, tujuan program pendidikan (PEO) dan hasil siswa (SOS). Sebagai studi kasus, makalah ini menunjukkan bagaimana teknik penambangan aturan asosiasi diterapkan pada aturan pemetaan tambang antara PEO dan seperangkat SO yang telah ditentukan sebelumnya yang diadopsi oleh American Board for Engineering and Technology-Engineering Accreditation Commission (ABET-EAC) untuk program rekayasa. Untuk tujuan ini, 152 laporan studi mandiri program teknik terakreditasi ABET-EAC telah dikumpulkan dan data pemetaan antara PEO mereka dan ABET-EAC SO telah diekstraksi. Dataset ini telah diproses sebelumnya dan diubah menjadi representasi yang cocok untuk menerapkan teknik penambangan aturan asosiasi. Ini melibatkan mengidentifikasi satu set label PEO, membubuhi keterangan contoh data dengan label PEO, dan memproyeksikan setiap contoh data multi-label ke dalam satu set contoh label tunggal. Algoritma Apriori kemudian diterapkan untuk menemukan aturan yang mengatur pemetaan antara PEO dan ABET-EAC SO. Aturan yang ditemukan sangat penting untuk memandu desain dan penilaian akademik teknik program. Selain itu,</p>

	<p>aturan yang ditemukan mengungkap sejumlah korelasi menarik antara PEO dan ABET-EAC SO yang perlu diselidiki lebih lanjut oleh para pedagogis. [19]</p>
Hasil	<p>Kumpulan aturan untuk setiap PEO telah diurutkan dalam urutan menurun sesuai dengan nilai kepercayaannya dan 10 aturan teratas telah digunakan untuk menghasilkan representasi yang lebih kompak dari aturan yang ditemukan seperti yang ditunjukkan pada Tabel 6. Dapat diamati bahwa PEO Ethical Conduct (EC) dan Social and Community (SC) memiliki rata-rata keyakinan tertinggi, sedangkan Kompetensi Pengetahuan (KC) dan Studi Pascasarjana (GS) memiliki terendah. Dalam aturan ini representasi PEO, Social and Community (SC) dan Knowledge Competency (KC), tidak bergantung pada keberadaan SO. Mereka terutama bergantung pada tidak adanya kombinasi SO yang berbeda. Pembelajaran Seumur Hidup (LL), Komunikasi (C), Perilaku Etis (EC), dan Profesionalisme (P) PEO bergantung pada keberadaan SO tunggal. Ini menunjukkan pemetaan satu-ke-satu antara atribut lulusan PEO dan keterampilan yang sesuai dari SO.</p> <p>Akhirnya, aturan pemetaan PEOs-SOs yang direkomendasikan memberikan wawasan yang sangat berguna bagi para pengambil keputusan program akademik teknik selama proses perbaikan berkelanjutan. Mereka merupakan pedoman panduan yang dapat digunakan untuk membantu akademisi ketika merancang program akademik baru atau meninjau yang sudah ada.</p>
Kesimpulan	<p>Teknik data mining aturan asosiasi diusulkan dan diterapkan untuk menemukan seperangkat aturan yang mengatur pemetaan antara PEO dan SO</p>

	<p>program akademik. Ini diterapkan pada kumpulan data yang diekstraksi dari SSR dari 152 program teknik terakreditasi ABET. Aturan yang ditemukan berguna untuk desain dan penilaian program akademik. Selain itu, aturan yang ditemukan mengungkapkan sejumlah korelasi menarik antara PEO dan ABET SO yang informatif untuk perancangan dan penilaian program serta layak untuk diselidiki lebih lanjut.</p>
Penulis	Anwar Ali Yahya, Addin Osman
Nama jurnal, Volume, Nomor, Tahun	Procedia Computer Science Volume 163, 2019, Pages 472-481
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian terkait memandu desain dan penilaian program Akademik. Dataset yang digunakan oleh penulis ini telah diproses sebelumnya dan diubah menjadi representasi yang cocok untuk menerapkan teknik penambangan aturan asosiasi. Ini melibatkan mengidentifikasi satu set label PEO, membubuhi keterangan contoh data dengan label PEO, dan memproyeksikan setiap contoh data multi-label ke dalam satu set contoh label tunggal. Algoritma Apriopi kemudian diterapkan untuk menemukan aturan yang mengatur pemetaan antara PEO dan ABET-EAC SO. Dari penelitian yang dilakukan oleh penulis didapat aturan pemetaan PEOs-SOs yang direkomendasikan memberikan wawasan yang sangat berguna bagi para pengambil keputusan program akademik teknik selama proses perbaikan berkelanjutan. Mereka merupakan pedoman panduan yang dapat digunakan untuk membantu akademisi ketika merancang program akademik baru atau meninjau yang sudah ada.</p>
Link URL Jurnal	https://www.sciencedirect.com/science/article/pii/S1877050919321696

No	20
Judul Artikel	Mining authentic Student feedback for Faculty Using Naïve Bayes Classifier
Topik	Data Mining
Data	Data sampel 1000 siswa Institusi Pendidikan Tinggi India
Metode / Algoritma	Naïve Bayes Classifier
Abstrak	<p>Keluaran dari analisis tradisional umpan balik siswa untuk penyampaian ruang kelas fakultas mengalami ketidakakuratan karena tidak mempertimbangkan pengaruh berbagai fitur kualitas langsung dan tidak langsung yang terkait dengan siswa seperti keteraturan kehadiran di kelas, upaya, latar belakang akademik, hasil kursus yang dicapai dan sikap positif pada ukuran umpan balik. Akibatnya, output dari analisis umpan balik fakultas tradisional bukanlah indikator yang benar dari efektivitas fakultas dalam proses belajar mengajar. Makalah ini menyajikan model analisis umpan balik fakultas proaktif dan berbasis hasil yang menggunakan Naïve Bayes Classifier untuk menyisihkan dan mengklasifikasikan umpan balik yang diberikan oleh setiap siswa ke dalam kategori yang valid atau tidak valid berdasarkan efek relatif dari fitur kualitas yang disebutkan di atas pada ukuran umpan balik. Fitur kualitas di atas digunakan untuk menyempurnakan ukuran umpan balik. Metode ini mencoba mengatasi ketidaktepatan untuk mengatasi keterbatasan model tradisional. Akibatnya, output dari analisis umpan balik fakultas menghasilkan Indeks Efektivitas Fakultas yang lebih halus dan akurat. Indeks Efektivitas Fakultas dihitung sebagai rata-rata tertimbang dari hanya ukuran umpan balik yang valid dengan validitas umpan balik yang diambil sebagai bobot terkait. Pengklasifikasi</p>

	<p>mempertimbangkan kontribusi independen dari masing-masing fitur serta beberapa bukti kemunculannya dalam umpan balik yang diberikan oleh setiap siswa. Metode ini juga menyarankan bentuk umpan balik yang komprehensif yang terdiri dari dua bagian yaitu bagian umpan balik subjektif untuk memperoleh umpan balik dengan cara tradisional dan bagian umpan balik berbasis hasil untuk mengumpulkan informasi tentang fitur kualitas tersebut di atas terkait dengan siswa yang mempengaruhi ukuran umpan balik. [20]</p>
<p>Hasil</p>	<p>Dari hasil penelitian ini pengklasifikasi yang dikembangkan menggunakan bahasa pemrograman R diterapkan pada kumpulan data sampel 1000 siswa dari Institusi Pendidikan Tinggi India yang berafiliasi dengan universitas negeri menunjukkan hasil yang memuaskan. Uji Z rata-rata dua yang menerapkan umpan balik subjektif atau umpan balik tradisional dan umpan balik yang valid dari siswa masing-masing untuk fakultas pada pengiriman ruang kelas memberikan hasil sebagai berikut: Hasil Uji untuk FEI. Hasil Uji Z Dua sampel untuk sarana yang diterapkan pada Indeks Efektivitas Fakultas yang dihitung pada umpan balik siswa diperoleh melalui pendekatan tradisional dan umpan balik siswa berdasarkan hasil yang divalidasi masing-masing.</p>
<p>Kesimpulan</p>	<p>Penelitian ini telah menerapkan pengklasifikasi pembelajaran mesin Naïve Bayes untuk membangun Model Validasi Umpan Balik dalam penambangan data Pendidikan. Pendekatan model probabilistik dapat secara efektif menggali alasan kemajuan belajar siswa dalam proses belajar mengajar secara individual. Ini juga menentukan sejauh mana masing-masing alasan mempengaruhi</p>

kemajuan belajar setiap siswa. Ini menyediakan mekanisme umpan maju untuk perbaikan tepat waktu penyimpangan dengan intervensi seperti konseling, kelas tambahan dll sesuai refleksi menengah. Umpan balik yang dipengaruhi oleh berbagai faktor dipisahkan dari pengaruh yang tidak produktif yang dapat merusak keaslian dan memproyeksikan umpan balik yang salah atau palsu. Umpan balik yang benar lebih lanjut dievaluasi. Umpan balik seperti itu tanpa pengaruh yang tidak produktif akan terbukti menjadi ukuran yang baik dari umpan balik siswa dan akibatnya kinerja fakultas juga. Umpan balik yang terotentikasi berfungsi sebagai alat untuk membantu menilai efektivitas pembelajaran dengan cara yang lebih proaktif dan objektif. Ini juga berfokus pada keterampilan dan hasil yang dicapai dan membantu meningkatkan hasil. Model Validasi Umpan Balik harus menjadi alat yang efektif untuk manajemen mutu proses belajar mengajar di perguruan tinggi. Ada banyak ruang untuk pekerjaan di masa depan menggunakan model ini. Anonimitas umpan balik siswa dapat dipertahankan dengan mengembangkan program pembuat kode identitas berdasarkan angka acak. Seorang siswa dapat menghasilkan identitas acak yang dapat digunakan untuk menghubungkan umpan balik subjektifnya dengan objektif rekan. Varian yang berbeda dari pengklasifikasi Nave Bayes hibrida dapat dieksplorasi untuk otentikasi umpan balik dan akurasinya dibandingkan Dalam hal pengurangan ketidakpastian dalam analisis data umpan balik. Pengklasifikasi Naïve Bayes dapat digunakan untuk mengeksplorasi area lain dari penambahan data pendidikan seperti analisis penempatan.

Penulis	Sandhya Maitra, Sushila Madan, Rekha Kandwal, Prerna Mahajan
Nama jurnal, Volume, Nomor, Tahun	Procedia Computer Science Volume 132, 2018, Pages 1171-1183
Ulasan artikel	<p>Pada penelitian ini penulis melakukan penelitian mengenai analisis tradisional umpan balik siswa untuk penyampaian ruang kelas fakultas. Penulis menggunakan Data sampel 1000 siswa Institusi Pendidikan Tinggi India. Penulis menggunakan metode algoritma yang menggunakan Naïve Bayes Classifier untuk menyisihkan dan mengklasifikasikan umpan balik yang diberikan oleh setiap siswa ke dalam kategori yang valid atau tidak valid berdasarkan efek relatif dari fitur kualitas yang disebutkan di atas pada ukuran umpan balik. Fitur kualitas di atas digunakan untuk menyempurnakan ukuran umpan balik. Dimana Metode ini mencoba mengatasi ketidak tepatan untuk mengatasi keterbatasan model tradisional. Akibatnya, output dari analisis umpan balik fakultas menghasilkan Indeks Efektivitas Fakultas yang lebih halus dan akurat. Dari penelitian yang dilakukan oleh penulis didapat pengklasifikasi yang dikembangkan menggunakan bahasa pemrograman R diterapkan pada kumpulan data sampel 1000 siswa dari Institusi Pendidikan Tinggi India yang berafiliasi dengan universitas negeri menunjukkan hasil yang memuaskan. Uji Z rata-rata dua yang menerapkan umpan balik subjektif atau umpan balik tradisional dan umpan balik yang valid dari siswa masing-masing untuk fakultas pada pengiriman ruang kelas memberikan hasil sebagai berikut: Hasil Uji untuk FEI.</p> <p>Hasil Uji Z Dua sampel untuk sarana yang diterapkan pada Indeks Efektivitas Fakultas yang</p>

	dihitung pada umpan balik siswa diperoleh melalui pendekatan tradisional dan umpan balik siswa berdasarkan hasil yang divalidasi masing-masing
Link URL Jurnal	https://www.sciencedirect.com/science/article/pii/S1877050918307646

Tabel ini terdiri dari minimal 20 Jurnal

IV. ULASAN

Berisi ulasan dari jurnal yang sudah direview, menjadi sebuah usulan penelitian yang baru. Gunakanlah hasil paraphrase yang telah dibuat pada setiap artikel dan dirangkai menjadi usulan penelitian baru.

No	Judul Jurnal	Ulasan
1.	Using Data Mining Techniques to Predict Student Performance to Support Decision Making in University Admission Systems	Pada penelitian ini penulis melakukan penelitian yang memfokuskan pada cara-cara untuk mendukung universitas dalam pengambilan keputusan penerimaan menggunakan teknik data mining untuk memprediksi kinerja akademik pelamar di universitas. Penulis menggunakan Kumpulan data yang terdiri dari 2.039 siswa yang terdaftar di Computer Science dan Information College di universitas negeri Saudi dari 2016 hingga 2019 digunakan untuk memvalidasi metodologi yang diusulkan. Pada penelitian ini penulis menggunakan algoritma (Decision Trees, Support Vector Machines, and Naïve Bayes). Untuk melihat klasifikasi mana yang lebih unggul dari pada teknik klasifikasi lainnya yang dipertimbangkan. Penulis memaparkan bahwa penelitian ini

		<p>menunjukkan bahwa model kinerja tinggi untuk memprediksi kinerja awal siswa dapat dikembangkan berdasarkan informasi pra-masuk. Sebagai contoh, dalam studi khusus ini, model JST mampu mencapai tingkat akurasi kinerja sekitar 79,22%. Studi ini menyimpulkan bahwa skor SAAT adalah kriteria penerimaan yang paling akurat memprediksi kinerja akademik di kemudian hari dan karena itu bobot lebih harus diberikan padanya.</p>
2.	<p>Students' Academic Performance Prediction using Data Mining</p> 	<p>Penelitian ini dilakukan untuk mengusulkan model untuk memprediksi kinerja akademik mahasiswa Ilmu Komputer menggunakan teknik Data Mining. Penulis menggunakan Pengumpulan data dilakukan dengan menggunakan kuesioner yang berisi tentang demografi siswa, IPK sebelumnya, dan informasi latar belakang keluarga. Pada penelitian ini penulis menggunakan algoritma (Decision Tree and Random Forest) yang diterapkan pada data siswa untuk membuat model prediksi kinerja akademik siswa terbaik, yaitu penelitian menunjukkan bahwa Decision Tree merupakan model terbaik dibandingkan dengan Random Forest dengan nilai akurasi tertinggi yaitu 66,9%. Studi ini menunjukkan bahwa ada beberapa</p>

		fitur relevan yang mempengaruhi kinerja akademik siswa
3.	Predicting Students' Academic Procrastination in Blended Learning Course Using Homework Submission Data	<p>Pada penelitian ini penulis melakukan penelitian terkait menyajikan sebuah algoritma yang disebut peningkatan kinerja akademik siswa melalui pendeteksian keterlambatan pekerjaan rumah/non-submission (SAPE) untuk memprediksi kinerja akademik siswa penundaan akademik kinerja siswa di lingkungan belajar yang didukung komputer. Dimana studi telah menunjukkan bahwa siswa yang menunjukkan kecenderungan penundaan yang lebih tinggi mencapai kurang dari siswa dengan kecenderungan penundaan yang lebih rendah. Penulis menggunakan metode Teknik EDM dapat digunakan untuk menganalisis data yang dikumpulkan melalui lingkungan belajar yang didukung komputer dan untuk memprediksi perilaku siswa. Penulis menggunakan sepuluh metode klasifikasi, yaitu ZeroR, OneR, ID3, J48, random forest, decision stump, JRip, PART, NBTree, dan Prism. Sebuah analisis rinci disajikan mengenai kinerja metode klasifikasi yang berbeda untuk jumlah kelas yang berbeda. Analisis mengungkapkan bahwa secara umum akurasi prediksi semua metode</p>

		<p>menurun dengan bertambahnya jumlah kelas. Namun, metode yang berbeda melakukan yang terbaik atau terburuk untuk jumlah kelas yang berbeda. diklasifikasikan dengan benar, ZeroR dan Oner menunjukkan nilai-nilai deviasi standar yang lebih tinggi wrt lainnya metode, jumlah diklasifikasikan dengan benar contoh yang sangat bervariasi untuk kedua metode ini. Tren serupa terlihat dalam kasus contoh yang salah diklasifikasikan. Dalam hal statistik kappa, NBTree menunjukkan kinerja paling konsisten dengan standar deviasi terendah, 0,017. Random forest menjadi classier terbaik kedua karena standar deviasi untuk metode ini terendah kedua, 0,023. Demikian pula, NBTree juga menunjukkan kinerja yang konsisten dalam nilai RMSE dengan standar deviasi 0,01. Metode lain dengan standar deviasi terendah adalah Random forest dan J48 dengan nilai standar deviasi masing-masing 0,013 dan 0,017.</p>
4.	A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining	<p>Pada penelitian ini penulis melakukan penelitian terkait penerapan pendekatan data mining pendidikan dengan model atrisi akademik (kehilangan status akademik) di Universidad Nacional de Colombia. Model bertujuan untuk memprediksi gesekan dalam empat pendaftaran pertama siswa. Pertama,</p>

		<p>mempertimbangkan salah satu dari periode ini, dan kemudian, pada pendaftaran tertentu. Catatan akademik historis dan data dari proses penerimaan digunakan untuk melatih model yang dievaluasi menggunakan validasi silang dan catatan yang sebelumnya tidak terlihat dari periode akademik penuh. Hasil eksperimen menunjukkan bahwa prediksi hilangnya status akademik meningkat ketika data akademik ditambahkan. Penulis menggunakan algoritma pengklasifikasi Decision Tree, Naive Bayes untuk memperoleh pemahaman yang lebih baik tentang gesekan selama pendaftaran pertama dan untuk menilai kualitas data untuk tugas klasifikasi. Dari penelitian yang dilakukan Prediksi hilangnya status akademik di salah satu dari empat periode akademik pertama menggunakan data awal, tetapi hasilnya tidak memuaskan; akurasi seimbang berkisar 51-52% di Decision Trees dan antara 54-57% di Naive Bayes. Dimana Hasil Naive Bayes lebih baik pada set tes; namun, ada perbedaan antara data pelatihan dan data uji. Hasil Decision Trees lebih konsisten mengenai subjek itu sehingga lebih dapat diandalkan saat menguji data baru.</p>
--	---	--

5.	<p>Combining University Student Self Regulated Learning Indicators and Engagement with Online Learning Events to Predict Academic Performance</p>	<p>Pada penelitian ini penulis melakukan penelitian terhadap bagaimana menggabungkan data tentang keterampilan belajar mandiri dengan ukuran aktivitas online yang dapat diamati dalam kursus pembelajaran campuran untuk meningkatkan kemampuan prediktif kinerja akademik siswa untuk tujuan menginformasikan pengajaran dan desain tugas. Penulis menggunakan Data penelitian yang sepenuhnya sukarela dan persetujuan tertulis mereka atas kuesioner, nilai akhir mereka, dan jejak digital mereka dalam Sistem Manajemen Pembelajaran. Dalam penelitian yang dilakukan oleh penulis Hasil penelitian ini menawarkan sejumlah ide untuk mengajar siswa dalam kursus campuran. Hubungan antara variabel dalam analisis korelasi dan cluster menunjukkan bahwa pengalaman siswa yang positif tentang efikasi diri, tes, motivasi, regulasi diri <i>dan</i> interaksi positif dengan banyak acara online, terutama yang menawarkan umpan balik (seperti dasbor dan pertanyaan pilihan ganda) dan refleksi dan penalaran (seperti urutan latihan pemecahan masalah) akan berkorelasi dengan pencapaian akademik yang relatif lebih tinggi dibandingkan dengan pengalaman</p>
----	---	--

		<p>negatif dari aspek-aspek tersebut. Hasil ini menawarkan strategi bagi guru untuk mengeksplorasi desain tugas dan pendekatan yang melibatkan pengungkapan kepada siswa apa yang melibatkan pengaturan diri positif untuk pembelajaran, bagaimana berinteraksi secara efektif dengan acara pembelajaran online, atau mendesain ulang materi instruksional yang menyertai kegiatan ini.</p>
6.	<p>Data Mining Mid-Semester Exam Data Clustering Using K-Means Algorithm</p> 	<p>Pada penelitian ini penulis melakukan penelitian terkait bagaimana melihat hasil belajar yang dapat direpresentasikan dengan nilai ujian yang ditempuh. Ujian tengah semester (UTS) merupakan salah satu bentuk ujian yang menjadi komponen penilaian. Penulis menggunakan algoritma <i>k-means</i> sebagai pertimbangan dosen dalam membentuk kelompok belajar mahasiswa berdasarkan <i>cluster</i> nilai UTS. Data yang digunakan oleh penulis adalah data yang berasal dari nilai UTS mahasiswa semester IV. Dari penelitian yang dilakukan oleh penulis diperoleh dua hasil clustering. Hasil <i>clustering</i> pertama untuk kelas Teknik Informatika-A dari <i>output</i> aplikasi <i>weka 3.6</i> menunjukkan sebanyak 5 (10%) nilai UTS termasuk dalam <i>cluster</i> nilai</p>

		rendah, sebanyak 17 (33%) nilai UTS mahasiswa termasuk dalam <i>cluster</i> nilai sedang, dan sebanyak 30 (58%) nilai UTS mahasiswa termasuk dalam <i>cluster</i> nilai tinggi
7.	An Educational Data Mining For Student Academic Prediction Using K-Means Clustering And Naïve Bayes Classifier.	<p>Pada penelitian ini penulis melakukan penelitian terkait bagaimana meningkatkan kualitas pembelajaran dengan memprediksi prestasi akademik siswa. Penulis menggunakan 2 algoritma klasifikasi yaitu K-Means dan Naïve Bayes. Data yang digunakan oleh penulis diambil dari dataset kinerja akademik mahasiswa yang digunakan sebagai test case. Jumlah data yang digunakan dalam penelitian ini adalah 131 data dan 21 atribut. Dari penelitian yang dilakukan oleh penulis didapat Akurasi kombinasi metode yang diusulkan adalah 97,44%. Perhitungan dengan algoritma K-Means dan perhitungan dengan algoritma Naive Bayes, metode yang diusulkan memberikan hasil yang lebih baik. Walaupun penentuan centroid awal pada metode K-Means dilakukan secara acak, namun dampak tersebut dapat dikurangi dengan penambahan metode Naive Bayes Classifier sehingga menghasilkan akurasi yang lebih baik dan meningkatkan akurasi dari metode yang ada.</p>

		<p>Dibandingkan dengan K-Means dan Naïve Bayes, metode yang diusulkan meningkatkan akurasi sekitar 27% dari algoritma Naïve Bayes dan sekitar 23% dari algoritma K-Means. Dengan hasil yang diperoleh, dapat disimpulkan bahwa metode yang diusulkan dapat meningkatkan prediksi data kinerja akademik siswa. Penentuan centroid awal pada metode K-Means mempengaruhi bahwa kualitas akurasi pengelompokan bergantung pada centroid awal.</p>
8.	<p>The Uses of Educational Data Mining in Academic Performance Analysis at Higher Education Institutions (Case Study at UNJANI)</p> 	<p>Pada penelitian ini penulis melakukan penelitian terkait mengetahui kinerja mahasiswa ditinjau dari ranah akademik dan non akademik pada waktu tertentu selama menempuh pendidikan. Penulis menggunakan model data mining terintegrasi Educational Data Mining (EDM), di mana teknik yang digunakan meliputi klasifikasi (ID3, SVM), clustering (K-Means, K-Medoids), aturan asosiasi (Apriori) dan deteksi anomali (DBSCAN). Penulis menggunakan Dataset akademik berupa hasil belajar selama periode waktu tertentu. Dalam penelitian yang dilakukan didapat proses clustering menggunakan teknik k-Means dengan nilai k ideal = 2, setelah terlebih dahulu dilakukan pengukuran SSE (sum square of error) = 18.1388, hasil clustering</p>

		<p>terdapat 2 (dua) kelompok pola akademik mahasiswa. Algoritma k-Means memberikan hasil yang lebih baik jika dibandingkan dengan algoritma k-Medoids. Teknik lain yang digunakan adalah teknik klasifikasi dengan algoritma pohon keputusan, dimana hasil dari teknik ini adalah terbentuknya pola aturan bersyarat (if-then-else) dengan menentukan atribut sebagai akar awal untuk menghitung nilai entropi, dimana teknik dalam EDM dapat digunakan sebagai cara untuk menganalisis pola yang terbentuk menjadi pengetahuan baru dalam pengelolaan pendidikan di perguruan tinggi dengan menggunakan kumpulan data akademik yang dihasilkan melalui berbagai sistem informasi yang tersedia di institusi pendidikan tinggi.</p>
9.	<p>Classification Analysis Using C4.5 Algorithm To Predict The Level of Graduation of Nurul Falah Pekanbaru High School Students</p>	<p>Pada penelitian ini penulis melakukan penelitian mengenai memprediksi tingkat kelulusan siswa SMA Nurul Falah dengan menggunakan metode Decision Three. Penulis menggunakan metode Decision Three dan algoritma C4.5 sebagai pendukung serta untuk mengetahui informasi tingkat kelulusan siswa SMA Nurul Falah. Data yang digunakan oleh penulis ini adalah data mahasiswa dengan menggunakan kriteria nama</p>

		<p>mahasiswa, jurusan, rata-rata raport semester satu (I), dua (II), tiga (III), empat (IV), lima (V), dan nilai rata-rata Ujian Sekolah Standar Nasional (USBN). Data tersebut kemudian dikelola menggunakan software Rapidminer 5.3 untuk memudahkan dalam memprediksi tingkat kelulusan siswa. Dari penelitian yang dilakukan oleh penulis didapat bahwa atribut gain tertinggi adalah semester 4 yaitu sebesar 0,31128. Ada dua nilai atribut dari semester 4 yaitu B dan C. Dari kedua nilai atribut nilai atribut C sudah mengklasifikasi kasus yaitu keputusan “Tidak Lulus”, tetapi untuk nilai atribut B masih perlu dihitung lagi. Dengan demikian 4 semester bisa menjadi simpul cabang ke-2 nilai atribut semester B. Berikutnya bawah adalah keseluruhan formasi decision tree.</p>
10.	<p>The Decision Tree c5.0 Classification Algoritma For Prediction Student Academic Performance</p>	<p>Pada penelitian ini penulis melakukan penelitian mengenai bagaimana mencari tahu apakah keaktifan seorang siswa dapat memprediksi kinerja akademiknya. Dalam penelitian ini penulis menggunakan algoritma C.50 yang digunakan untuk menguji data, di mana data dibagi menjadi data pelatihan sebesar 75% dan pengujian data sebesar 25%. Dan hasil dari data yang diuji, akurasi 71,667% diperoleh. Data yang digunakan oleh</p>

		<p>penulis adalah dataset pendidikan yang dikumpulkan menggunakan learning management system (LMS), yang merupakan alat pelacak aktivitas pelajar yang terhubung oleh internet. Data ini memiliki variabel numerik dan kategorikal, sehingga diperlukan algoritma yang tepat untuk mengklasifikasikan data secara akurat dan memastikan validitas data. Dari penelitian yang dilakukan oleh penulis di dapat akurasi yang dicapai oleh algoritma C5.0 ini mencapai 71.667% untuk prediksi performa akademik siswa dari faktor keaktifan siswa di dalam kelas maupun di luar kelas yang mencakup tingkat keseringan dalam mengangkat tangan, keseringan berpartisipasi dalam diskusi, inisiatif siswa untuk belajar di luar sekolah, dan absensi dari siswa itu sendiri pada data testing. Kesimpulan lainnya yang dapat ditarik dari <i>decision tree</i> yang telah ditunjukkan adalah seorang siswa memang diperlukan untuk aktif dan selalu datang ke sekolah agar tidak tertinggal pelajaran untuk mendapatkan nilai yang memuaskan.</p>
11.	<p>Application of Data Mining Classification Method for Student Graduation Prediction Using K-Nearest Neighbor (K-NN) Algorithm</p>	<p>Pada penelitian ini penulis melakukan penelitian mengenai memprediksi tingkat kelulusan dengan memanfaatkan teknik data mining. Metode Klasifikasi Data</p>

		<p>Mining yang digunakan oleh penulis adalah algoritma K-Nearest Neighbor (K-NN). Data set yang digunakan oleh penulis berasal dari data siswa, data nilai siswa, dan data kelulusan siswa tahun ajaran 2010-2012 sebanyak 2.189 record. Atribut yang digunakan adalah jenis kelamin, asal sekolah, program studi IP Semester 1-6. Dari penelitian yang dilakukan di dapat dataset kelulusan yang terdiri dari data training dan data testing, dengan menerapkan metode K-optimal pada 2 jenis dataset menghasilkan perolehan akurasi nilai pada data latih sebesar 88,16% pada pengujian dengan nilai k=1 sedangkan pengujian data pengujian dengan nilai k=1 menghasilkan akurasi sebesar 89,04%. Dapat juga dianalisis bahwa nilai k mempengaruhi nilai akurasi dari setiap dataset test. Hal ini dibuktikan dengan nilai akurasi yang dihasilkan pada pengujian ke 1 sampai 3, dimana setiap percobaan menggunakan nilai k yang berbeda yaitu 1, 3, dan 5. Pada pengujian pengujian data, training menghasilkan nilai akurasi tertinggi pada implementasi.</p>
12.	Naive Bayes Algorithm for Classification of Student Majors Specialization	<p>Pada penelitian ini penulis melakukan penelitian mengenai menentukan peminatan atau jurusan mahasiswa. Penulis menggunakan</p>

		<p>metode Algoritma Naive Bayes yang dapat digunakan dalam mengklasifikasikan jurusan tersebut. Dari penelitian yang dilakukan oleh penulis didapat hasil informasi baru yaitu perhitungan data mining berdasarkan data uji siswa SMA Negeri 1 Kampar Timur, menunjukkan siswa yang memilih IPS sebanyak 51 orang, dengan ketepatan nilai 100%, dan sebanyak 50 siswa yang memilih masuk IPA, dan ditambah 4 orang dari siswa yang memilih IPS diprediksi masuk IPA dengan presisi kelas 92,59%. Dari hasil ini diperoleh akurasi sebesar 96,19%, dengan nilai true recall kelas IPS 92,73% dan true IPA class recall 100%, sehingga diperoleh kesimpulan bahwa peminatan jurusan menggunakan algoritma Naive Bayes menghasilkan tingkat akurasi yang tinggi dalam pemilihan jurusan sebagai perencanaan untuk menentukan kursus berikutnya.</p>
13.	<p>Classification and Prediction of Student GPA Using K-Means Clustering Algorithm to Assist Student Admission Process</p>	<p>Pada penelitian ini penulis melakukan penelitian mengenai Klasifikasi dan Prediksi IPK mahasiswa untuk membantu proses penerimaan mahasiswa baru. Penulis menggunakan algoritma K-means clustering untuk memprediksi nilai rata-rata (IPK) siswa berdasarkan profil siswa, seperti status sekolah menengah dan lokasi, nilai tes masuk</p>

		<p>universitas dan kompetensi bahasa Inggris. Dan data yang digunakan adalah Data siswa angkatan 2008-2017 digunakan untuk membuat dua cluster menggunakan algoritma clustering K-means. Dua centroid dari cluster digunakan untuk mengklasifikasikan semua data menjadi dua kelompok: IPK tinggi dan IPK rendah. Kami menggunakan data dari angkatan 2018 sebagai data uji. Performa prediksi diukur dengan menggunakan akurasi, presisi dan recall. Dari penelitian yang dilakukan didapat metode pengelompokan K-means adalah 78,59% akurat di antara siswa penerimaan berbasis prestasi dan 94,627% di antara siswa penerimaan reguler. Prediksi yang melibatkan mahasiswa masuk berbasis merit memiliki nilai akurasi prediksi yang lebih rendah dibandingkan dengan yang melibatkan mahasiswa reguler karena model clustering untuk data penerimaan berbasis merit adalah $K = 3$, tetapi untuk prediksi, asumsinya adalah $K = 2$.</p>
14.	Implementation Naïve Bayes Algorithm for Student Classification Based on Graduation Status	<p>Pada penelitian ini penulis melakukan penelitian mengenai klasifikasi mahasiswa berdasarkan status kelulusan. Penulis menggunakan algoritma Naive Bayes klasifikasi. Data yang digunakan oleh penulis adalah data</p>

		<p>Data Registrasi Mahasiswa Jurusan Statistika yang mengacu pada atribut Indeks Prestasi Kumulatif (IPK), Konsentrasi di Sekolah Menengah Atas, Jenis Kelamin, Partisipasi dalam Pendampingan dan Kota Tempat Tinggal, yang kemudian diklasifikasikan berdasarkan Status Kelulusan siswa di atas empat tahun dan kurang dari sama dengan empat tahun. Dari penelitian yang dilakukan oleh penulis didapat algoritma Naïve Bayes Classifier telah dibahas sebagai classifier terbaik dalam masalah ini. Dengan demikian, tingkat akurasi pada model klasifikasi algoritma Naive Bayes 81,18%.</p>
15.	 <p>Text Clasification using Naïve Baye Updateable Algorithm In SBMPTN Test Questions</p>	<p>Pada penelitian ini penulis melakukan penelitian mengenai klasifikasi teks pada soal tes SBMPTN. Penulis menggunakan algoritma Naive Bayes Updateable selanjutnya melakukan dalam mengklasifikasikan soal ujian SBMPTN berdasarkan tema. Model kenaikan satu algoritma klasifikasi sering digunakan dalam klasifikasi teks klasifikasi memiliki kemampuan untuk belajar dari data baru memperkenalkan dengan sistem bahkan setelah classifier telah diproduksi dengan data yang ada. Penulis menggunakan data yang digunakan bersumber dari buku kumpulan soal-soal ujian SBMPTN.</p>

		<p>Dari penelitian yang dilakukan didapat pengklasifikasian soal ujian SBMPTN dengan algoritma Naive Bayes yang dilakukan pada penelitian ini dapat berjalan dengan baik, sebagian besar data soal ujian yang diklasifikasikan sesuai dengan pengklasifikasian secara manual dengan nilai akurasi 84,61%. Soal ujian yang gagal diklasifikasikan dengan benar ini disebabkan oleh adanya kata kunci baru yang memiliki kemungkinan sebagai kata kunci penting tetapi tidak dapat digunakan karena belum terdaftar pada data training, selain itu kegagalan klasifikasi juga dapat disebabkan oleh keterbatasan data <i>training</i> yang dimiliki, seperti dapat dilihat pada proses <i>memorizing</i> terdapat beberapa soal ujian yang tidak mampu diingat oleh sistem dikarenakan <i>keyword</i> yang dimiliki kelas sebenarnya masih kurang beragam maupun kurang frekuensinya sehingga mempengaruhi derajat konsentrasi, derajat persebaran dan derajat kontribusinya.</p>
16.	<p>Determinating Student Interactions in a Virtual Learning Environment Using Data Mining</p>	<p>Pada penelitian ini penulis melakukan penelitian mengenai penentuan interaksi siswa dalam lingkungan pembelajaran virtual menggunakan data mining. Data yang digunakan penulis adalah data</p>

		<p>interaksi siswa dalam Kursus Bahasa Inggris. Analisis metode penambangan yang dilakukan oleh penulis adalah membandingkan masing-masing untuk kemudian memilih salah satu yang membantu pengembangan proyek, memilih metode Crisp-dm karena mengandung beberapa fase yang menunjukkan setiap aktivitas yang harus diselesaikan, sehingga menjadi praktis panduan. Selain itu, analisis komparati dikembangkan dengan mempertimbangkan fitur alat penambangan data di mana RapidMiner dipilih untuk melakukan proses menggunakan beberapa algoritma bersama dengan data siswa yang dibagi menjadi dua set untuk pelatihan dan validasi, menghasilkan Decisin Three sebagai algoritma terbaik untuk tujuan tersebut karena mengklasifikasikan instance dengan benar dengan margin kesalahan minimum. Dari penelitian yang dilakukan oleh penulis didapat hasil terbaik dari algoritma dipertimbangkan algoritma Decision Three sebagai hasil terbaik, menyajikan klasifikasi yang baik dengan 92,9% dan margin kesalahan yang lebih rendah dalam validasi model dengan 7,1%. Dan untuk menentukan tingkat interaksi dalam kursus bahasa Inggris, algoritma</p>
--	---	---

		<p>klasifikasi yang berbeda diterapkan, menyajikan hasil terbaik di Decision Three karena algoritma ini memperoleh margin kesalahan paling sedikit selama klasifikasi data dari interaksi dalam kursus (tugas, tes, sumber daya), dan data pribadi, kelembagaan dan sosial ekonomi. Melalui model data mining ini dapat diketahui bahwa interaksi siswa di English Virtual Course menghasilkan tingkat rata-rata dengan persentase 69%, dan faktor yang paling mempengaruhi model adalah interaksi siswa dalam ujian, tugas, sumber daya, status perkawinan, dan status pekerjaan.</p>
17.	 <p>The Third Information Systems International Conference A Review on Predicting Student's Performance using Data Mining Techniques</p>	<p>Pada penelitian ini penulis melakukan penelitaian mengenai memprediksi kinerja siswa menggunakan teknik data mining. Penulis menggunakan algoritma Decision Tree, Neural Network , Support Vector Machine, K-Nearest Neighbor sebagai analisis dan perbandingan prediksi yang digunakan untuk mengidentifikasi atribut yang paling penting dalam data siswa. Kami sebenarnya dapat meningkatkan prestasi dan keberhasilan siswa secara lebih efektif dengan cara yang efisien menggunakan teknik data mining pendidikan. Dapat membawa manfaat dan dampak bagi siswa,</p>

		<p>pendidik dan institusi akademik. Dari penelitian yang dilakukan oleh penulis didapat analisis Neural Network memiliki akurasi prediksi tertinggi sebesar (98%) diikuti oleh Decision Tree sebesar (91%). Selanjutnya Support Vector Machine dan K-Nearest Neighbor memberikan akurasi yang sama, yaitu (83%). Terakhir, metode yang memiliki akurasi prediksi lebih rendah adalah metode Naive Bayes sebesar (76%). Hasil akurasi prediksi tergantung pada atribut atau fitur yang digunakan selama proses prediksi. Metode Neural Network memberikan akurasi prediksi tertinggi karena pengaruh dari atribut utama. Atribut ini merupakan hibridisasi dari dua fitur, yaitu penilaian internal dan eksternal. Dengan penggunaan satu variabel saja, yaitu penilaian eksternal, akurasinya berkurang (1%). Variabel ketiga yang paling banyak digunakan adalah penilaian internal yang memberikan hasil akurasi (81%). Hal ini menunjukkan bahwa penilaian eksternal, yaitu nilai yang diperoleh dalam ujian akhir, memainkan peran penting dalam memprediksi kinerja siswa. Sedangkan variabel signifikan.</p>
18.	An Intelligent Prediction System for Educational Data Mining Based on Ensemble and Filtering approaches	<p>Pada penelitian ini penulis melakukan penelitian mengenai sistem prediksi cerdas untuk pendidikan berdasarkan pendekatan</p>

		<p>ensemble dan filtering. Di antara berbagai pendekatan ensemble, penulis telah mempraktekkan mekanisme boosting untuk memprediksi kinerja siswa. Karena penerapan metode ensemble dianggap sebagai fenomena signifikan dalam prosedur klasifikasi dan prediksi, oleh karena itu para peneliti memanfaatkan teknik boosting untuk mengembangkan model pedagogis prediksi yang akurat, mengingat sifat yang diucapkan dan kebaruan metode yang diusulkan dalam penambahan data pendidikan. Penulis menggunakan algoritma Decision Tree, Naïve Bayes, K-Nearest Neighbor dan Dataset pedagogis dari University of Kashmir. Dari penelitian yang dilakukan oleh penulis didapat Decision Tree, K-Nearest Neighbor dan Naive Bayes dengan teknik boosting telah menunjukkan hasil yang luar biasa di semua pengklasifikasi. Pengklasifikasi pembelajaran dasar yaitu. j48 dan nave bayes dengan sistem boosting, yang diilustrasikan pada tabel 5 telah menunjukkan pencapaian penting dalam hasil prediksi saat mengalami teknik pengambilan sampel berlebih. Kinerja kedua classifier pembelajaran meningkat dari 95,32% menjadi 96,44% (j48 dengan</p>
--	---	--

		<p>boosting) dan 95,04% menjadi 96,06% (nave bayes dengan boosting). para peneliti memeriksa bahwa setiap pengklasifikasi pembelajaran mengalami peningkatan yang signifikan, dan diantara setiap pengklasifikasi dasar, naive bayes mencapai akurasi prediksi yang patut dicatat sebesar 97,15%, dengan teknik over-sampling. Pengklasifikasi knn menunjukkan akurasi prediksi yang cukup besar dengan teknik under-sampling dari pada pengklasifikasi pembelajaran lainnya. Namun, saat memeriksa pendekatan ansambel dan menguatkan metode mana yang memberikan hasil yang lebih baik, para peneliti menemukan bahwa peningkatan.</p>
19.	<p>UNIVERSITAS MERCU BUANA</p> <p>Using Data Mining Techniques to Guide Academic Programs Design and Assessment</p>	<p>Pada penelitian ini penulis melakukan penelitian terkait memandu desain dan penilaian program Akademik. Dataset yang digunakan oleh penulis ini telah diproses sebelumnya dan diubah menjadi representasi yang cocok untuk menerapkan teknik penambangan aturan asosiasi. Ini melibatkan mengidentifikasi satu set label PEO, membubuhi keterangan contoh data dengan label PEO, dan memproyeksikan setiap contoh data multi-label ke dalam satu set contoh label tunggal. Algoritma Apriopi</p>

		<p>kemudian diterapkan untuk menemukan aturan yang mengatur pemetaan antara PEO dan ABET-EAC SO. Dari penelitian yang dilakukan oleh penulis didapat aturan pemetaan PEOs-SOs yang direkomendasikan memberikan wawasan yang sangat berguna bagi para pengambil keputusan program akademik teknik selama proses perbaikan berkelanjutan. Mereka merupakan pedoman panduan yang dapat digunakan untuk membantu akademisi ketika merancang program akademik baru atau meninjau yang sudah ada.</p>
20.	 <p>Mining authentic Student feedback for Faculty Using Naïve Bayes Classifier</p>	<p>Pada penelitian ini penulis melakukan penelitian mengenai analisis tradisional umpan balik siswa untuk penyampaian ruang kelas fakultas. Penulis menggunakan Data sampel 1000 siswa Institusi Pendidikan Tinggi India. Penulis menggunakan metode algoritma yang menggunakan Naïve Bayes Classifier untuk menyisihkan dan mengklasifikasikan umpan balik yang diberikan oleh setiap siswa ke dalam kategori yang valid atau tidak valid berdasarkan efek relatif dari fitur kualitas yang disebutkan di atas pada ukuran umpan balik. Fitur kualitas di atas digunakan untuk menyempurnakan ukuran umpan balik. Dimana Metode ini mencoba</p>

		<p>mengatasi ketidak tepatan untuk mengatasi keterbatasan model tradisional. Akibatnya, output dari analisis umpan balik fakultas menghasilkan Indeks Efektivitas Fakultas yang lebih halus dan akurat. Dari penelitian yang dilakukan oleh penulis didapat pengklasifikasi yang dikembangkan menggunakan bahasa pemrograman R diterapkan pada kumpulan data sampel 1000 siswa dari Institusi Pendidikan Tinggi India yang berafiliasi dengan universitas negeri menunjukkan hasil yang memuaskan. Uji Z rata-rata dua yang menerapkan umpan balik subjektif atau umpan balik tradisional dan umpan balik yang valid dari siswa masing-masing untuk fakultas pada pengiriman ruang kelas memberikan hasil sebagai berikut: Hasil Uji untuk FEI.</p> <p>Hasil Uji Z Dua sampel untuk sarana yang diterapkan pada Indeks Efektivitas Fakultas yang dihitung pada umpan balik siswa diperoleh melalui pendekatan tradisional dan umpan balik siswa berdasarkan hasil yang divalidasi masing-masing</p>
--	---	---

DAFTAR PUSTAKA

- [1] H. A. Mengash, "Using data mining techniques to predict student performance to support decision making in university admission systems," *IEEE Access*, vol. 8, pp. 55462–55470, 2020, doi: 10.1109/ACCESS.2020.2981905.
- [2] F. J. Kaunang and R. Rotikan, "Students' academic performance prediction using data mining," *Proc. 3rd Int. Conf. Informatics Comput. ICIC 2018*, pp. 1–5, 2018, doi: 10.1109/IAC.2018.8780547.
- [3] A. Akram *et al.*, "Predicting Students' Academic Procrastination in Blended Learning Course Using Homework Submission Data," *IEEE Access*, vol. 7, pp. 102487–102498, 2019, doi: 10.1109/ACCESS.2019.2930867.
- [4] C. E. Lopez Guarin, E. L. Guzman, and F. A. Gonzalez, "A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining," *Rev. Iberoam. Tecnol. del Aprendiz.*, vol. 10, no. 3, pp. 119–125, 2015, doi: 10.1109/RITA.2015.2452632.
- [5] A. Pardo, F. Han, and R. A. Ellis, "Combining University student self-regulated learning indicators and engagement with online learning events to Predict Academic Performance," *IEEE Trans. Learn. Technol.*, vol. 10, no. 1, pp. 82–92, 2017, doi: 10.1109/TLT.2016.2639508.
- [6] G. Abdurrahman, "Clustering Data Ujian Tengah Semester (UTS) Data Mining Menggunakan Algoritma K-Means," *J. Sist. dan Teknol. Inf. Indones.*, vol. 1, no. 2, pp. 71–79, 2016, [Online]. Available: <http://jurnal.unmuhjember.ac.id/index.php/JUSTINDO/article/view/566>.
- [7] D. A. N. Wulandari, R. Annisa, and L. Yusuf, "an Educational Data Mining for Student Academic Prediction Using K-Means Clustering and Naïve Bayes Classifier," *J. Pilar Nusa Mandiri*, pp. 155–160, 2020.
- [8] Y. H. Chrisnanto and G. Abdullah, "The uses of educational data mining in academic performance analysis at higher education institutions (case study at UNJANI)," *Matrix J. Manaj. Teknol. dan Inform.*, vol. 11, no. 1, pp. 26–35, 2021, doi: 10.31940/matrix.v11i1.2330.
- [9] F. Wiza and B. Febriadi, "Classification Analysis Using C4.5 Algorithm To Predict The Level of Graduation of Nurul Falah Pekanbaru High School Students," *IJISTECH (International J. Inf. Syst. Technol.*, vol. 2, no. 2, p. 43, 2019, doi: 10.30645/ijistech.v2i2.21.
- [10] N. Benediktus and R. S. Oetama, "Algoritma Klasifikasi Decision Tree C5 . 0 untuk Memprediksi Performa Akademik Siswa," *Ultimatics*, vol. XII, no. 1, pp. 14–19, 2020.
- [11] M. Imron and S. A. Kusumah, "Application of Data Mining Classification Method for Student Graduation Prediction Using K-Nearest Neighbor (K-NN) Algorithm," *IJIS Int. J. Informatics*

- Inf. Syst.*, vol. 1, no. 1, pp. 1–8, 2018, doi: 10.47738/ijiis.v1i1.17.
- [12] A. W. Syaputri, E. Irwandi, and M. Mustakim, “Naïve Bayes Algorithm for Classification of Student Major’s Specialization,” *J. Intell. Comput. Heal. Informatics*, vol. 1, no. 1, p. 17, 2020, doi: 10.26714/jichi.v1i1.5570.
- [13] R. G. Santosa, Y. Lukito, and A. R. Chrismanto, “Classification and Prediction of Students’ GPA Using K-Means Clustering Algorithm to Assist Student Admission Process,” *J. Inf. Syst. Eng. Bus. Intell.*, vol. 7, no. 1, p. 1, 2021, doi: 10.20473/jisebi.7.1.1-10.
- [14] A. Kesumawati and D. Waikabu, “Implementation Naïve Bayes Algorithm for Student Classification Based on Graduation Status,” *Int. J. Appl. Bus. Inf. Syst.*, vol. 1, no. 2, pp. 6–12, 2018, doi: 10.31763/ijabis.v1i2.3.
- [15] R. Saptono, M. E. Sulistyono, and N. S. Trihabsari, “Text Classification Using Naive Bayes Updateable Algorithm in Sbmptn Test Questions,” *Telematika*, vol. 13, no. 2, p. 123, 2017, doi: 10.31315/telematika.v13i2.1728.
- [16] A. Vilorio *et al.*, “Determinating student interactions in a virtual learning environment using data mining,” *Procedia Comput. Sci.*, vol. 155, no. 2018, pp. 587–592, 2019, doi: 10.1016/j.procs.2019.08.082.
- [17] A. M. Shahiri, W. Husain, and N. A. Rashid, “A Review on Predicting Student’s Performance Using Data Mining Techniques,” *Procedia Comput. Sci.*, vol. 72, pp. 414–422, 2015, doi: 10.1016/j.procs.2015.12.157.
- [18] M. Ashraf, M. Zaman, and M. Ahmed, “An Intelligent Prediction System for Educational Data Mining Based on Ensemble and Filtering approaches,” *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 1471–1483, 2020, doi: 10.1016/j.procs.2020.03.358.
- [19] A. A. Yahya and A. Osman, “Using Data Mining Techniques to Guide Academic Programs Design and Assessment,” *Procedia Comput. Sci.*, vol. 163, pp. 472–481, 2019, doi: 10.1016/j.procs.2019.12.130.
- [20] S. Maitra, S. Madan, R. Kandwal, and P. Mahajan, “Mining authentic student feedback for faculty using Naïve Bayes classifier,” *Procedia Comput. Sci.*, vol. 132, pp. 1171–1183, 2018, doi: 10.1016/j.procs.2018.05.032.