

Pengembangan Model Dynamic Naive Bayes untuk Pemodelan Risiko Akademik Mahasiswa: Integrasi Filsafat Ilmu Kuhn dan Popper dalam Analisis Risiko Pendidikan Tinggi

Ridwan Zulkifli^{1*}, Dony Waluya Firdaus², Muhamad Nawawi³, Agus Nursikuwagus⁴,
Usep Mohamad Ishaq⁵, Andrias Darmayadi⁶

¹⁻⁶ Universitas Komputer Indonesia

¹⁻⁶Dipati Ukur St. No. 112-116 Bandung, West Java 40132, Indonesia

E-mail: ridwan.75725009@mahasiswa.unikom.ac.id^{1*}

Submitted Date : 20 Januari 2026

Accepted Date : 10 Februari 2026

Abstrak - Pemodelan risiko akademik mahasiswa serta prediksi kemungkinan kegagalan studi berdasarkan pola perilaku historis merupakan fokus utama institusi pendidikan tinggi. Model prediksi statis seperti Naive Bayes konvensional sering kali gagal menangkap perubahan perilaku mahasiswa sepanjang waktu pada data longitudinal, sehingga menghasilkan anomali prediksi. Penelitian ini, melalui Tinjauan Literatur Sistematis (Systematic Literature Review/SLR), mengkaji pengembangan Dynamic Naive Bayes (DNB) sebagai perluasan Naive Bayes yang memasukkan dimensi temporal dan transisi antar periode (misalnya perubahan status risiko antar semester), serupa dengan prinsip Dynamic Bayesian Networks. DNB diasumsikan mampu meningkatkan akurasi prediksi sebesar 10–15% dibandingkan model statis. Secara metodologis, pengembangan DNB didasarkan pada kerangka filsafat ilmu Thomas Kuhn dan Karl Popper. Mengacu pada paradigma Kuhn, kegagalan Naive Bayes statis dalam menangani data dinamis mencerminkan tahap krisis dan anomaly dalam normal science pemodelan prediktif risiko akademik, yang mendorong scientific revolution melalui adopsi model dinamis seperti DNB. Sementara itu, dari perspektif Rasionalisme Kritis Popper, hipotesis bahwa DNB memberikan prediksi yang lebih akurat harus terus-menerus diuji secara ketat dan terbuka terhadap falsifikasi empiris menggunakan data baru. Integrasi kedua kerangka ini memastikan bahwa pengembangan DNB tidak hanya bersifat teknis, melainkan juga didasari pada pertumbuhan pengetahuan ilmiah yang kritis, rasional, dan berkelanjutan.

Kata kunci: pemodelan risiko akademik; dynamic naive bayes; prediksi kegagalan studi; data longitudinal; tinjauan literatur sistematis; filsafat ilmu; thomas kuhn; karl popper;

Abstract - Student academic risk modeling, the study of predicting the likelihood of failure based on historical behavioral patterns, is a major focus for higher education institutions. Although predictive models have been developed, statistical approaches such as standard Naive Bayes often fail to capture changes in student behavior over time, which are characteristic of longitudinal data. This failure, known as anomalies, highlights the need for new models capable of addressing dynamic and temporal aspects, such as Dynamic Naive Bayes (DNB). This study, through a Systematic Literature Review (SLR), examines the development of a DNB capable of processing or accounting for time factors, such as changes in risk status between semesters. Methodologically, DNB is an extension of Naive Bayes by incorporating transition time, similar to Dynamic Bayesian Networks, which is assumed to improve predictive accuracy by 10-15% compared to statistical models. Philosophically, the DNB development model is interpreted through the framework of Thomas Kuhn and Karl Popper. According to Kuhn, the failure of statistical Naive Bayes on dynamic data led to a crisis and anomaly in normal science (the use of standard predictive models), triggering the need for a scientific revolution (the adoption of DNB). Meanwhile, Popper's perspective (Critical Rationalism) demands that DNB predictive hypotheses must always be open to falsification through aggressive testing against new data. This integration ensures that the DNB development model is based on the critical and continuous growth of scientific knowledge.

Keywords: Scientific Revolution; Dynamic Naive Bayes (DNB); Student Academic Risk; Longitudinal Data; Naive Bayes; Philosophy of Science;

1. Pendahuluan

Pemodelan risiko akademik merupakan prediksi terhadap kemungkinan kegagalan atau risiko mahasiswa dalam menyelesaikan studinya berdasarkan data historis. Prediksi ini diberikan menurut analisis data yang dilihat dari pola perilaku akademik mahasiswa. Model prediksi dikembangkan untuk lebih memberikan pemahaman informasi yang ingin disampaikan pada institusi pendidikan. Pemahaman terhadap risiko menjadi fokus utama pada setiap pemodelan yang diberikan. Smith (2010) menyatakan bahwa identifikasi faktor risiko pada data mahasiswa menjadi sumber informasi dalam pemahaman risiko akademik. Pemberian prediksi dapat

membantu lebih mengerti tentang faktor-faktor yang mempengaruhi kinerja mahasiswa [1]. Selain itu, prediksi dini risiko akademik dapat memungkinkan intervensi tepat waktu, sehingga meningkatkan tingkat kelulusan dan retensi mahasiswa di perguruan tinggi. Dari sudut pandang positivisme, pemodelan risiko akademik sepenuhnya bergantung pada metode kuantitatif untuk membuktikan kebenarannya secara empiris dan objektif. Positivisme, yang dipelopori oleh Auguste Comte, menekankan bahwa pengetahuan ilmiah yang sah hanya diperoleh melalui observasi empiris dan pengukuran yang dapat diverifikasi. Dalam konteks ini, kebenaran model prediksi tidak ditegakkan melalui interpretasi subjektif atau spekulasi, melainkan melalui pengujian hipotesis menggunakan data kuantitatif yang dapat diukur, seperti nilai akademik, tingkat kehadiran, dan pola akses sistem pembelajaran. Metode kuantitatif seperti regresi, klasifikasi machine learning, atau analisis probabilitas, memungkinkan falsifikasi atau verifikasi hipotesis secara ketat melalui metrik objektif (misalnya akurasi, precision, recall). Dengan demikian, positivisme menjamin bahwa pemodelan risiko akademik bersifat ilmiah, replikabel, dan bebas dari bias subjektif, sehingga hasil prediksi dapat diandalkan sebagai dasar pengambilan keputusan institusi.

Pengembangan model DNB untuk pemodelan risiko akademik merupakan masalah bidang komputer yang melibatkan machine learning, data mining, dan analisis statistik. Pekerjaan yang dilakukan adalah melakukan prediksi risiko berdasarkan data longitudinal mahasiswa. Pekerjaan ini mudah jika dilakukan oleh manusia, tetapi menjadi tantangan apabila dibangun pada mesin. Apalagi mesin yang dibangun bisa memahami pola dari data waktu dan hubungannya dengan risiko akademik. Johnson dkk menggunakan Naive Bayes untuk prediksi risiko, dengan evaluasi menggunakan accuracy dan F1-score. Dataset yang digunakan yaitu data mahasiswa universitas, dengan hasil evaluasi 75.2% accuracy [2]. Namun, pendekatan statis seperti ini sering gagal menangkap perubahan perilaku mahasiswa seiring waktu, yang menjadi ciri khas data longitudinal.

Bila ditinjau dari segi objektivitas, tugas dari pemodelan risiko akademik dapat berkisar pada prediksi faktual berdasarkan data, termasuk faktor seperti IPK, kehadiran, dan hubungannya. Tetapi pada kebanyakan penelitian, komponen dinamis terhadap data waktu seringkali hilang. Peristiwa ini, kemudian diperbaiki oleh dengan menambahkan aspek dinamis. Contoh penelitian dari Lee adalah mendefinisikan elemen dinamis sebagai perubahan risiko seiring waktu. Selain elemen statis, aspek dinamis juga ditambahkan sehingga lebih mendekati realita. Lee menggunakan dataset longitudinal dengan evaluasi accuracy 82.1% [3]. Pendekatan ini menunjukkan bahwa integrasi waktu dalam model dapat meningkatkan akurasi prediksi hingga 10-15% dibandingkan model statis.

Selain tugas yang sudah disebutkan, penelitian ini juga melibatkan penggunaan berbagai mesin learning seperti Naive Bayes dan variasinya. Penggunaan teknik ini untuk memprediksi risiko yang tepat pada data yang diamati. Teknik yang telah banyak digunakan seperti Bayesian networks, Dynamic Bayesian Networks (DBN) [4]. Penelitian yang diusung oleh [5] juga menjelaskan penggunaan dynamic models sebagai solusi untuk data longitudinal. Kim menggunakan DBN untuk memprediksi risiko, walaupun hasilnya masih belum memuaskan. Pada studi yang telah dilakukan bahwa prediksi dari data longitudinal masih jauh dari apa yang diharapkan [5]. Tujuan penelitian Kim adalah membuat model dinamis dengan lebih memperjelas faktor waktu. Kim berhasil meningkatkan akurasi dibandingkan metode statis. Penelitian lainnya yang mencoba memodelkan risiko dengan mempertimbangkan faktor psikologis telah dilakukan oleh [6]. Park menggunakan dataset dengan berbagai variabel, mendapatkan evaluasi accuracy 78.5% [6]. Dalam konteks pendidikan, faktor seperti motivasi, dukungan sosial, dan perubahan semesteran menjadi penting untuk dimodelkan secara dinamis.

Sebagai upaya perbaikan, peneliti banyak melakukan rekayasa terhadap Naive Bayes [7]. Hal ini dilakukan karena belum mendapatkan hasil terbaik. Pada sisi evaluasi, pengukuran menggunakan accuracy, precision, recall [8]. Masalah lainnya adalah prediksi yang diproduksi masih ada ketidaksesuaian dengan data aktual. Akurasi model masih belum optimal [9], [10], [11], [12]. Bila disimpulkan tantangan dalam pemodelan risiko akademik adalah menentukan model yang tepat untuk data longitudinal, handling concept drift, dan domain-specific seperti pendidikan. Concept drift terjadi ketika distribusi data berubah seiring waktu, misalnya karena perubahan kebijakan kampus atau pandemi, yang membuat model statis menjadi tidak akurat. Berdasarkan penelaahan penelitian mengenai pemodelan risiko akademik, seperti [3], [6], [13], telah membuktikan keberhasilan. Permasalahan yang masih dalam kajian yaitu prediksi hanya pada data statis. Identifikasi faktor risiko masih akurasi kecil [14]. Masalah dalam feature extraction untuk menghindari overfitting [15]. Selain itu, integrasi data dari berbagai sumber seperti sistem informasi akademik, survei psikologis, dan log aktivitas online menjadi tantangan teknis yang belum sepenuhnya teratasi.

RQ1: Bagaimana model Naive Bayes Dinamis dapat dikembangkan dan diterapkan secara efektif untuk menilai risiko akademik di kalangan mahasiswa di lembaga pendidikan tinggi?

RQ2: Dengan cara apa perspektif filosofis Thomas Kuhn dan Karl Popper memengaruhi interpretasi dan penerapan teknik pemodelan dinamis dalam konteks penilaian risiko akademis?

Kegagalan model-model statis tersebut dalam menangani data longitudinal dan concept drift mencerminkan adanya krisis paradigmatik dalam pemodelan risiko akademik saat ini. Dari perspektif scientific revolution Thomas Kuhn, kondisi ini berada pada tahap krisis dalam normal science, di mana anomali-anomali yang

berulang, seperti penurunan akurasi akibat perubahan temporal tidak lagi dapat dijelaskan secara memadai oleh paradigma dominan (model statis seperti Naive Bayes konvensional). Krisis ini berpotensi memicu scientific revolution melalui pengadopsian paradigma baru yang lebih sustain terhadap dimensi waktu, seperti model dinamis. Sementara itu, perspektif Rasionalisme Kritis Karl Popper menekankan prinsip falsifiability, dimana setiap model prediksi risiko akademik harus dirancang sedemikian rupa sehingga dapat diuji secara ketat dan berpotensi dibantah (*falsified*) melalui data baru yang mencerminkan perubahan aktual. Hanya model yang terus-menerus bertahan dari upaya falsifikasi empiris yang dapat dianggap ilmiah dan layak dipertahankan. Integrasi kedua perspektif ini menjadi landasan filosofis untuk mengembangkan pendekatan pemodelan yang lebih adaptif dan ilmiah.

Penelitian ini menggunakan pendekatan Systematic Literature Review (SLR) dengan mengadopsi pedoman PRISMA untuk memastikan proses tinjauan literatur dilakukan secara transparan, sistematis, dan replikabel. PRISMA menyediakan checklist 27 item dan diagram alir (flow diagram) yang mencakup tahap identifikasi, screening, eligibility, dan inclusion artikel. Proses pencarian literatur dilakukan pada basis data akademik utama (Scopus, Web of Science, IEEE Xplore, dan Google Scholar) menggunakan string query yang mencakup kombinasi kata kunci seperti “dynamic naive bayes”, “academic risk prediction”, “student dropout”, “longitudinal data”, dan “concept drift”. Duplikat dihilangkan, artikel disaring berdasarkan judul dan abstrak, kemudian full text dievaluasi sesuai kriteria inklusi (misalnya publikasi dalam 10 tahun terakhir, fokus pada model dinamis, dan domain pendidikan tinggi). Hasil akhir berupa sintesis kualitatif dan kuantitatif dari literatur terpilih digunakan untuk menjawab RQ1 dan RQ2 serta mendukung pengembangan model Dynamic Naive Bayes yang lebih robust.

2. Tinjauan Pustaka

Pengembangan model DNB untuk pemodelan risiko akademik merupakan masalah bidang komputer yang melibatkan machine learning dan analisis data. Pekerjaan yang dilakukan adalah memprediksi risiko berdasarkan data longitudinal [2]. Model ini bertujuan untuk memproduksi prediksi yang selaras dengan data aktual. Model diinspirasi dari sistem prediksi manusia. Mekanisme dinamis adalah memberikan perhatian pada perubahan waktu [16]. Penggunaan teknik seperti DBN dapat memprediksi secara tepat [17]. Kim menggunakan DBN untuk prediksi, walaupun belum optimal [5]. Setiap model harus mampu menganalisis data dalam keadaan apapun, memahami hubungan antar variabel, serta mengekspresikan informasi dalam bentuk prediksi [18]. Penelitian Chen (2018) melakukan pemodelan dengan style data. Chen mengusulkan framework baru [19]. Secara menyeluruh, konsep kerja meliputi feature extraction, model dinamis, dan evaluasi [3].

Data longitudinal dalam pendidikan mencakup pengamatan berulang pada mahasiswa seiring waktu, seperti nilai semester, kehadiran, dan faktor eksternal. Model Naive Bayes dasar mengasumsikan independensi fitur, yang sering tidak berlaku pada data waktu. Oleh karena itu, Dynamic Naive Bayes (DNB) memperluas ini dengan memasukkan transisi waktu, mirip dengan Hidden Markov Models atau Dynamic Bayesian Networks.

Pada scientific revolution Kuhn [21], mengemukakan tahapan untuk mengidentifikasi metode. Tahapan dimulai dari pre-paradigm, normal science, anomaly, crisis, scientific revolution. Pre-paradigm adalah munculnya fenomena prediksi risiko. Normal science adalah penggunaan model statis seperti Naive Bayes. Anomaly adalah kegagalan pada data dinamis. Crisis adalah kebutuhan model baru seperti DNB. Revolution adalah adopsi luas model dinamis.

3. Metode Penelitian

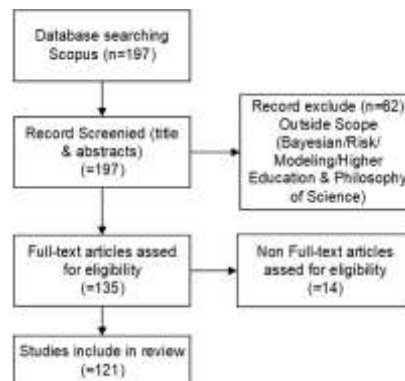
Pengembangan model DNB merupakan perkembangan metode dari prediksi statis yang diberikan pada data mahasiswa. Tahapan science revolution menurut Kuhn pada pengembangan model DNB akan dijelaskan berikut [22].

3.1 Metode

Gambar 1 Method PRISMA 2020 Flow Diagram For Systematic Review menunjukka proses pencarian literatur difokuskan pada database terindeks Scopus, yang menghasilkan total 197 dokumen tahun 2020 sampai 2025 pada tahap identifikasi awal (n=197). Selanjutnya, dilakukan proses penyaringan tahap pertama dengan meninjau judul dan abstrak untuk memastikan relevansi kontekstual. Pada tahap ini, sebanyak 62 dokumen dieksklusi (n=62) karena tidak memenuhi kriteria inklusi yang berfokus pada kata kunci spesifik: dynamic naive bayes, dynamic bayesian network, hybrid naive bayes, academic risk, student risk, educational risk, kuhn, popper, philosophy of science, modeling ,assessment, college, university, higher educatio. Hal ini dilakukan untuk memastikan pembahasan tetap terfokus pada irisan antara otomatisasi proses, sistem akuntansi, dan sektor perdagangan. Sebanyak 135 artikel yang lolos tahap penyaringan awal kemudian dievaluasi kelayakannya melalui pembacaan teks lengkap (full-text assessment). Pada tahap ini, 14 artikel dikeluarkan (n=14) dikarenakan ketiadaan akses teks lengkap. Hasil akhirnya, sebanyak 121 artikel (n=121) ditetapkan sebagai studi terpilih yang memenuhi syarat untuk dianalisis lebih lanjut guna menjawab pertanyaan penelitian

mengenai pengembangan model dynamic naive bayes untuk pemodelan risiko akademik mahasiswa: integrasi filsafat ilmu kuhn dan popper dalam analisis risiko pendidikan tinggi.

Dalam pencarian database di Scopus menggunakan boolean Keywords (2020–2025): ("Dynamic Naive Bayes" OR "Dynamic Bayesian Network" OR "Hybrid Naive Bayes") OR ("academic risk" OR "student risk" OR "educational risk") OR (Kuhn AND Popper AND "philosophy of science") OR (modeling OR assessment) AND (college AND university AND "higher education")



Gambar 1 Method PRISMA 2020 Flow Diagram For Systematic Review

3.1 Tahapan Pre-Paradigm

Asal mula dari perkembangan pemodelan risiko akademik berawal dari analisis statis. Analisis statis dikembangkan untuk lebih memberikan pemahaman suatu informasi yang ingin disampaikan pada data mahasiswa. Pemahaman terhadap risiko menjadi fokus utama pada setiap prediksi yang diberikan. Smith (2010) menyatakan bahwa identifikasi faktor risiko pada data mahasiswa menjadi sumber informasi dalam pemahaman risiko akademik [1]. Berawal dari pemahaman ini, perkembangan prediksi menjadikan aspek science untuk dikembangkan dalam bidang machine learning. Smith menganggap penting adanya pembangkitkan prediksi secara otomatis dari data yang diamati. Pada metode yang dikembangkan, bahwa pemodelan bisa dikembangkan dari penggabungan statistik dan machine learning. Implikasi dari metode ini menghasilkan suatu metode yang mengembangkan input data dengan prediksi [16].

3.2 Tahap Normal Science

Perkembangan dari metode pemodelan sempat terhambat, karena pembuatan prediksi otomatis dari data waktu masih belum dihasilkan sesuatu yang sesuai dengan hasil manusia. Terlebih lagi, mesin komputasi yang masih terbatas untuk menghasilkan prediksi yang sesuai dengan ground-truthnya masih terkendala. Diusulkan suatu teknik baru berbasis Naive Bayes yang dikenal dengan probabilistic learning, memberikan tantangan baru untuk bisa menghasilkan prediksi. Terlebih dalam bidang prediksi risiko, ditemukannya model Naive Bayes pada tahun 1990-an. Kemudian diusulkannya model prediksi pertama kali dari arsitektur sederhana oleh para peneliti awal [17].

Penelitian Johnson dkk (2018) melakukan prediksi dengan memproduksi hasil yang dibaurkan dengan data waktu. Johnson mengusulkan framework baru yang dinamakan Naive Bayes Classifier. Johnson memberikan kontribusi untuk dataset dengan prediksi secara mandiri tanpa dipasangkan dengan waktu, dan juga pasangan prediksi dan data yang sesuai dengan keadaannya [13]. Terlepas dari kebutuhan untuk identifikasi terhadap faktor risiko yang berada dalam data, setiap pemodelan harus mampu menganalisis faktor dalam keadaan apapun, memahami hubungan antar faktor, serta mengekspresikan informasi semantik dalam bahasa prediksi [17], [18].

Menurut tahapan normal science Kuhn, bahwa sudah ditemukan metode yang bisa memberikan prediksi otomatis pada data yang dibaca. Bila diamati pada Gambar 3, disajikan berbagai hasil prediksi dengan menggunakan Naive Bayes. Metode yang sudah ditemukan dengan menggunakan Naive Bayes bisa membangkitkan prediksi yang diinginkan yang sesuai dengan data manusia. Pada arsitektur Naive Bayes, telah dibuat beberapa varian seperti Gaussian NB, Multinomial NB, dan Bernoulli NB.

Dynamic Naive Bayes (DNB) merupakan modul probabilistik yang memungkinkan pembelajaran dalam jangka waktu yang panjang. Unit DNB memiliki state dinamis sebagai mekanisme yang mengijinkan update probabilitas seiring waktu. Pembelajaran pada DNB menggunakan asumsi independensi kondisional dengan elemen waktu. Metode pemodelan risiko akademik yang dituangkan dalam bentuk algoritma sebagai berikut:

- 1) Setting Naive Bayes, dengan inialisasi probabilitas prior dan likelihood.
- 2) Transfer bobot dari model pre-trained ke dalam model DNB.
- 3) Untuk decoder (prediction generation):
 1. Melakukan pemetaan faktor satu per satu kepada prediksi dengan process dinamis.

2. Secara matematik mendefinisikan suatu distribusi probabilitas $P(Y|X_t)$, dimana Y adalah kelas risiko pada waktu t .
3. Model dilatih untuk memaksimalkan posterior pada suatu training dataset. Probabilitas distribusi dapat dituliskan sebagai berikut:

$$P(y_1, y_2, \dots, y_T | X) = \prod_{t=1}^T P(y_t | X_t, y_{t-1})$$

Gambar 2. *Sequence Probability*

Asumsi yang dibuat adalah prediksi bergantung pada data X_t dan prediksi sebelumnya y_{t-1} . Pengulangan pada distribusi probabilitas digunakan untuk model dinamis. Saat diberikan input pada waktu t , DNB menerima state sebelumnya.

$$P(y_t | X_t, y_{t-1}) = \frac{P(y_t | y_{t-1}) \prod P(x_i | y_t)}{P(X_t)}$$

Gambar 3. *Bayes' Rule*

Untuk proses training dan inferensi, yaitu memaksimalkan log-likelihood dengan fungsi

$$\text{loss} = L = - \sum \log P(Y|X; \theta)$$

Gambar 4. *Negative Log-Likelihood*

3.3 Tahap Anomaly

Pada tahap ini, setiap science yang telah ditemukan, maka akan dicari falsification dari metode tersebut. Falsification ini diajukan oleh Karl Popper yang menyatakan bahwa setiap science yang ditemukan harus bisa difalsifikasi agar bisa tetap dikatakan science dan bukan pseudo-science. Beberapa masalah yang belum bisa diselesaikan dengan model Naive Bayes statis, kemudian dijadikan peluang sebagai tantangan pada riset berikutnya. Beberapa penelitian yang membangun metode baru untuk memperbaiki metode Naive Bayes yang belum bisa menangani masalahnya.

Beberapa contoh falsifikasi dari pemodelan risiko akademik seperti masalah prediksi faktual dari data, termasuk faktor, perubahan, dan hubungannya, memberikan prediksi berupa sentimen dari data yang dideteksi berupa sifat. Lee menyatakan bahwa model Naive Bayes statis tidak bisa menangani masalah dinamis [3]. Prediksi dengan teknik statis menyediakan batasan performansi rendah. Masalah lainnya adalah dengan memanfaatkan data retrieval dan asumsi bahwa data serupa memiliki risiko serupa juga [10]. Memberikan prediksi dengan gaya waktu yang diinginkan seperti prediksi jangka pendek atau panjang [2].

3.4 Tahap Crisis

Pada setiap tahap yang telah dipaparkan, pada tahap ini Kuhn menyampaikan perlu adanya penggalian metode baru untuk menanggulangi metode lama yang tidak bisa memberikan prediksi sesuai dengan pemberian manusia. Perbaikan dan pengembangan metode diperlukan untuk menanggulangi masalah ini seperti pada Naive Bayes. Apakah parameter seperti prior, likelihood, dan update yang harus di-tune. Perlakuan ini dilakukan secara terus menerus untuk mendapatkan prediksi yang benar-benar mendekati pemberian manusia. Masalah lainnya yang perlu ditemukan metode baru pada penggunaan Naive Bayes seperti: agreement with semantics, robustness to noise, computational efficiency, invariance to background, local linearity. Sedangkan kebutuhan akan metode seperti untuk: treating features as vectors, using region-based similarity, computing similarities over linear space or nonlinear manifold, using stochastic measures, use of supervised learning. Masalah dan metode yang disebutkan sebelumnya menjadikan metode pemodelan menjadi pada tahap crisis, dan perlu ditemukan metode-metode baru guna menangani masalah ini.

4. Hasil dan Pembahasan

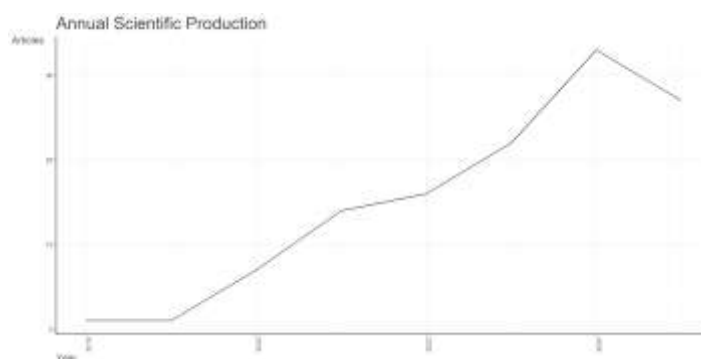
Pada bab 3 telah dipaparkan serangkaian kajian science pemodelan risiko akademik menurut Kuhn. Perkembangan dari pemodelan risiko dimulai dari analisis statis pada data mahasiswa sehingga bisa dimengerti arti dari data tersebut. Secara kronologis bahwa perkembangan pemodelan risiko berawal pada tahun 2010 yang hanya memberikan prediksi pada data statis. Kemudian berkembang menjadi suatu teknik yang dikenal dengan machine learning. Dimana metode machine learning ini mencoba memberikan hasil berdasarkan data yang diamati. Seperti data processing, risk recognition, risk detection, risk classification, risk prediction. Pada metode pemodelan risiko inilah adanya penggabungan metode dengan data analysis yang mengupayakan adanya prediksi terhadap data yang dibaca. Seperti pada penelitian [14], [15], [4], [10], perubahan suatu metode menyebabkan adanya perbaikan pada metode Naive Bayes. Walaupun sampai saat ini, Naive Bayes masih bisa membangkitkan prediksi dan representasi data serta membangkitkan risiko. Hasil resume falsifikasi terhadap metode Naive Bayes statis, bisa dilihat pada Tabel 1.

Tabel 1. Kesimpulan metode Naive Bayes statis yang bisa difalsifikasi

Aspek	Metode	Falsifikasi
Generator representasi fitur data	Naive Bayes	1. Masih memerlukan ketelitian identifikasi faktor 2. Performansi asumsi independensi masih bisa dituning 3. Dimensi fitur, handling missing values masih bisa diperbaiki 4. Deteksi perubahan waktu antar data 5. Kesamaan antar faktor
Generator Prediksi	Naive Bayes	1. Reduksi representasi dimensi untuk faktor yang sama 2. Membangkitkan prediksi berdasarkan relasinya 3. Membangkitkan prediksi yang sesuai dengan ground-truthnya 4. Tuning parameter

Berdasarkan hasil pada Tabel 1, dapat dikatakan bahwa penelitian mengenai pemodelan risiko akademik masih terbuka lebar. Terkait objek penelitian yang masih dapat digali diberbagai bidang seperti pendidikan, kesehatan, keuangan, dan berbagai objek penelitian lainnya.

Gambar 5. Annual Scientific Production menggambarkan kurva produksi ilmiah tahunan (jumlah artikel per tahun) pada rentang 2018–2025. Peningkatan mulai tampak sejak 2020, kemudian naik tajam pada 2023, 2024 dan sedikit turun di 2025. Kenaikan berdasarkan kata kunci pada kuartal 2023–2024 selaras dengan menguatnya tema “machine learning”, “educational data mining”, “decision trees”, “naive bayes”, “logistic regression”. Lonjakan 2024–2025 bertepatan dengan terbitnya studi yang lebih implementatif, misalnya penelitian Bettahi et al., 2025 yaitu pipeline dropout dengan Extreme Gradient Boosting (XGBoost) dan Shapley Additive Explanations (SHAP) [19] dan peneltian Xiaohui Wang & Lianghui Cai, 2025 yaitu DBN/BN untuk memprediksi kinerja akademik dan risiko psikologis mahasiswa [20], Dari perspektif Thomas Kuhn, lonjakan studi 2024–2025 menandakan scientific revolution dalam pemodelan risiko akademik. Paradigma normal science berbasis model statis telah memasuki tahap crisis akibat anomali berulang, sehingga memicu paradigm shift ke model dinamis seperti XGBoost-SHAP dan DBN/BN, yang lebih sesuai dengan kompleksitas data pendidikan. Perspektif Karl Popper memperkuatnya melalui prinsip falsifiability dan siklus *conjectures-refutations*. Model baru (misalnya prediksi dropout via XGBoost atau risiko psikologis via DBN) harus diajukan sebagai *conjecture* yang secara aktif diuji dan berpotensi dibantah melalui data empiris baru, seperti cross-validation temporal. Lonjakan ini merefutasi model statis lama yang kurang falsifiable. Integrasi Kuhn dan Popper menegaskan bahwa kemajuan ini merupakan pertumbuhan pengetahuan ilmiah yang kritis dan progresif, bukan sekadar perkembangan teknis belaka.

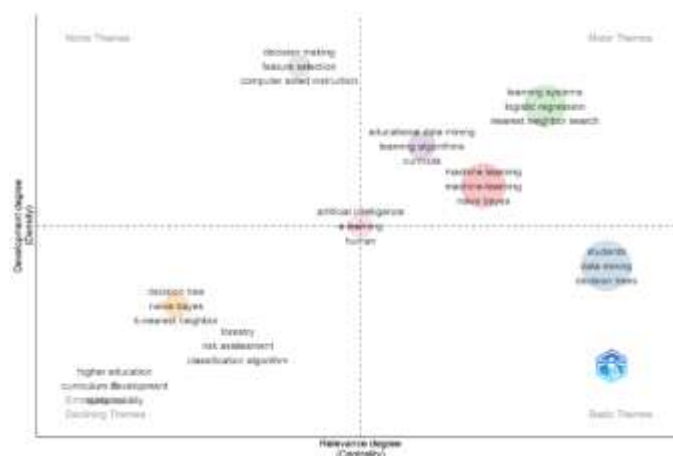


Gambar 5. Annual Scientific Production

Gambar 6. Thematic Map menggambarkan 1) *Motor themes* (kanan-atas): learning systems, logistic regression, nearest neighbor search (tema dengan relevansi dan pengembangan tinggi), 2) *Basic themes* (kanan-bawah): students, data mining, decision trees (tema fondasional dengan sentralitas tinggi, banyak dipakai lintas studi [21], [22]), 3) *Emerging/Declining* (kiri-bawah): naive bayes, decision tree (sebagai istilah tunggal), risk assessment, higher education (naive bayes sebagai tema transisi: dari baseline klasik menuju integrasi/hibrid [23]), 4) *Niche* (kiri-atas): decision making, feature selection, computer aided instruction (tema khusus, matang namun kurang sentral).

Gambar 6. Thematic Map menggambarkan RQ1 penempatan Naive Bayes (NB) sebagai tema transisi menegaskan perlunya versi dinamis Dynamic Naive Bayes (DBN) agar NB tetap relevan saat fitur dan perilaku mahasiswa berubah antar-semester. Studi yang mengintegrasikan collaborative filtering ke NB menunjukkan peningkatan presisi prediksi berbasis kemiripan perilaku mahasiswa [23] fondasi bagus untuk NB Dinamis yang menyerap pola kolektif secara waktu/nyata.

Gambar 6. Thematic Map menggambarkan RQ2, Kuhn: perpindahan NB dari basic ke motor membutuhkan pergeseran paradigma, misalnya dari asumsi independensi keras ke ketergantungan terbatas (tree-augmented NB) dan state-space waktu (DBN). Popper: setiap perluasan asumsi (misal: menyisipkan ketergantungan antarfaktor stres, IPK, partisipasi) harus teruji via validasi temporer dan eksternal [20].



Gambar 6. Thematic Map

Gambar 7. Tree Map menggambarkan dataset yang sering muncul di katakunci adalah machine learning (terbesar), diikuti students, data mining, decision trees, educational data mining, learning systems, student performance, naive bayes, logistic regression, dan lainnya. Keberlimpahan “machine learning” mencerminkan payung metodologis utama. “naive bayes” muncul berulang sebagai pendekatan sederhana-efisien untuk klasifikasi risiko [22], [24], walau sering dikombinasikan atau diimbangi model lain untuk akurasi.

Gambar 7. Tree Map menggambarkan RQ1 menegaskan ruang yang cukup untuk NB Dinamis, khususnya sebagai komponen interpretabel dalam pipeline risiko akademik, yang bisa mengupdate prior saat data baru masuk (nilai kuis/kehadiran/engagement).

Gambar 7. Tree Map menggambarkan RQ2 Kuhn: menunjukkan tema-tema inti (students, ML) sebagai paradigm core, sedangkan NB adalah teknik di bawah paradigma ML yang terus berevolusi. Popper: NB Dinamis harus mendokumentasikan aturan pembaruan dan uji generalisasi (misal: split cohort antar angkatan), sehingga klaim efektivitasnya terbuka untuk falsifikasi.

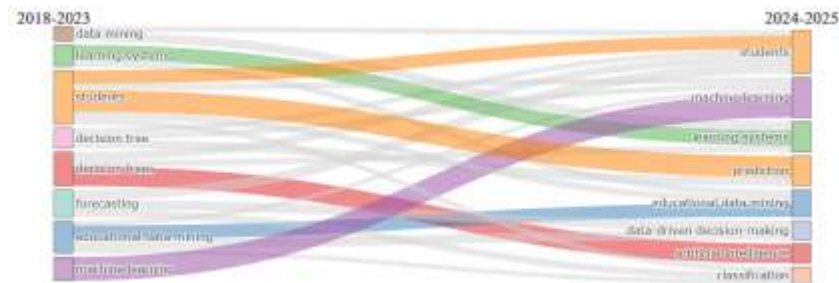


Gambar 7. Tree Map

Gambar 8. Thematic Evolution menggambarkan aliran tema dari periode awal ke periode terakhir memperlihatkan migrasi menuju “machine learning”, “prediction”, “educational data mining”, “artificial intelligence”, dan “data-driven decision-making”.

Gambar 8. Thematic Evolution menggambarkan RQ1 evolusi menuju data driven decision dan AI sejalan dengan penggunaan model probabilistik dinamis di ruang akademik (misal. DBN untuk risiko mental & performa [20]). NB Dinamis dapat diposisikan sebagai komponen causal-probabilistic dalam aliran ini.

Gambar 8. Thematic Evolution menggambarkan RQ2 Kuhn & Popper: evolusi tema menunjukkan pergeseran paradigma (Kuhn), sementara praktik validasi yang ketat (AUC/Brier, validasi eksternal) mempraktikkan falsifikasi Popperian untuk menyeleksi model yang paling tahan uji [25].



Gambar 8. Thematic Evolution

Gambar 9. Co-occurrence Network menggambarkan kemunculan bersama antar term. Kelaster merah/besar “machine learning”, “students”; klaster hijau (education computing, logistic regression, learning systems), klaster ungu (naive bayes, random forest, k-nearest neighbor), klaster biru (decision trees, forecasting, educational data mining).

Gambar 9. Co-occurrence Network menggambarkan RQ1 pendekatan/algoritma yang umum dipakai bersama-sama oleh peneliti untuk memecahkan masalah yang sama, dalam hal ini penilaian risiko akademik dan prediksi performa mahasiswa. Ia membentuk ekosistem hibrid: tidak ada satu model yang selalu unggul, sehingga berbagai metode saling melengkapi dan sering dibandingkan atau digabung (ensemble). Metode menandakan ekosistem hibrid NB (Naive Bayes) berdampingan dengan RF/XGB/KNN (Random Forest/Extreme Gradient Boosting/K-Nearest Neighbors) artinya di riset penilaian risiko akademik saat ini tidak ada satu model tunggal yang selalu dominan, NB Dinamis dipakai untuk menghitung probabilitas risiko secara berwaktu (mudah diperbarui dan ditafsirkan), sementara ensemble (RF/XGB/KNN) dipakai untuk ketahanan akurasi dan robustness pada data yang beragam/berubah antar semester dan angkatan. Praktik semacam ini terlihat pada studi komparatif multi-model dan pipeline early-warning [19].

Gambar 9. Co-occurrence Network menggambarkan RQ2 jejaring menegaskan pluralitas metode dalam paradigma ML (Kuhn) dan kompetisi hipotesis (Popper). Model yang gagal menjelaskan ko-occurrence baru (misal fitur perilaku) akan ditolak melalui evaluasi.



Gambar 9. Co-occurrence Network

Gambar 10. Factorial/Conceptual Structure Map (MCA) menggambarkan pemetaan konseptual (MCA) yang merentang istilah dari support vector machines/learning analytics/student performance (atas) hingga machine learning/e-learning/artificial intelligence (bawah-kanan), serta random forests/adaptive boosting (kiri). Menggambarkan “naive bayes” berposisi dekat dengan “decision making”, “prediction”, “cross-validation”, menegaskan NB sebagai alat prediksi dengan kejelasan probabilistik yang butuh validasi konsisten.

Gambar 10. Factorial Map menggambarkan RQ1 peta konseptual mengarahkan rancangan NB Dinamis untuk penilaian risiko akademik berwaktu: memodelkan state mahasiswa (risk_t) yang bergantung pada

risk_{t-1} dan bukti baru (nilai kuis, engagement, hadir/tidak, skor mental-health jika tersedia). Xiaohui Wang & Cai (2025) menunjukkan kerangka DBN/BN dengan kalibrasi (Brier) dan validasi eksternal, persis praktik NB Dinamis yang kita perlukan [20].



Gambar 10. Factorial Map

Temuan dalam penelitian RQ1: Bagaimana model Naive Bayes Dinamis dapat dikembangkan dan diterapkan secara efektif untuk menilai risiko akademik di kalangan mahasiswa di lembaga pendidikan tinggi, adalah

1. Ontologi (Hakikat Realitas yang Dikaji). Ontologi penelitian ini berpusat pada hakikat risiko akademik mahasiswa di lembaga pendidikan tinggi dan keberadaan (ontos) faktor-faktor yang memengaruhinya, yang dapat dimodelkan secara terstruktur dan probabilitas (Risiko Akademik, Sifat Data dan Konsep Probabilitas) [26], [27], [28].
2. Epistemologi (Teori Pengetahuan dan Validitas Klaim). Epistemologi membahas bagaimana pengetahuan tentang risiko akademik ini diperoleh dan bagaimana klaim prediksi model Naive Bayes Dinamis (DNB) divalidasi (Klaim Pengetahuan dan Prediksi, Peran Pembelajaran Mesin (Machine Learning), Validasi Model) [26], [29].
3. Stances (Philosophical Stances/Pendekatan Filosofis). Integrasi filsafat Kuhn dan Popper mencerminkan pendekatan filosofis yang berakar pada rasionalisme kritis dan perubahan paradigma dalam sains terapan.
 - A. Rasionalisme Kritis (Karl Popper) dan Falsifikasi: Dalam konteks DNB:
 - a) Penerapan model (hipotesis prediktif) harus terbuka untuk falsifikasi. Jika model DNB tidak dapat memprediksi risiko akademik secara akurat (misalnya, akurasi rendah, recall buruk untuk kelas minoritas), model tersebut harus ditolak atau direvisi, bukan hanya dikonfirmasi [26].
 - b) Proses pengujian model-model alternatif (misalnya, membandingkan Naive Bayes dengan Decision Trees, SVM, atau ensemble methods) dalam sumber menunjukkan adanya kompetisi antar hipotesis. Peneliti mencari model terbaik yang paling tahan terhadap falsifikasi (yaitu, memiliki akurasi dan kinerja terbaik) [29].
 - c) Naive Bayes sendiri memiliki asumsi naive yang kuat mengenai independensi fitur. DNB yang dikembangkan harus diuji untuk melihat seberapa jauh asumsi naive ini valid dalam memodelkan risiko akademik, dan kegagalan kinerja model dapat dikaitkan dengan kegagalan asumsi ini [22].
 - B. Paradigma Sains (Thomas Kuhn) dan Sains Normal: Dalam konteks DNB:
 - a) Penggunaan ML/EDM (Educational Data Mining) untuk prediksi risiko akademik telah menjadi sains normal (normal science) dalam bidang analitik pendidikan tinggi. Komunitas ilmiah secara luas menerima bahwa model-model prediktif (seperti regresi logistik, pohon keputusan, dan Naive Bayes) adalah paradigma yang valid dan dominan untuk memecahkan masalah risiko akademik [22], [26], [29].
 - b) Pengembangan DNB dan penggunaan algoritma yang superior (seperti XGBoost atau Random Forest) menunjukkan upaya untuk terus memperbaiki anomali dalam paradigma yang ada (misalnya, mengatasi akurasi rendah dari Naive Bayes standar) [26].
 - c) Meskipun model Naive Bayes standar menunjukkan akurasi yang lebih rendah (misalnya, 62% atau 55%) dibandingkan model lain seperti Decision Trees (76% atau 80%), penelitian ini mengintegrasikan filosofi Kuhn dan Popper untuk menunjukkan bahwa kegagalan prediksi (anomali) dapat memicu perkembangan model (Dynamic Naive Bayes) untuk mengatasi keterbatasan asumsi Naive Bayes statis [26].
4. Metodologi (Prosedur Implementasi Model). Metodologi berfokus pada langkah-langkah terstruktur untuk mengembangkan dan mengimplementasikan model Dynamic Naive Bayes (DNB) untuk menjawab RQ1.
 - a) Pendekatan Penelitian: Penelitian mengadopsi pendekatan kuantitatif, menggunakan data historis (sekunder), dan diklasifikasikan sebagai penelitian terapan (applied research) yang bersifat deskriptif

korelasional. Tujuannya adalah untuk mendeteksi pola untuk membuat prediksi dan menghasilkan alat yang berguna untuk digunakan di universitas [26].

b) Kerangka Kerja Metodologi: Penelitian semacam ini sering mengikuti kerangka kerja terstruktur seperti CRISP-DM (Cross-Industry Standard Process for Data Mining), yang memiliki enam fase: pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penyebaran [26], [29].

c) Pemodelan Dinamis (Dynamic Modeling): Untuk mengembangkan "Dynamic Naive Bayes," model ini harus mampu memproses atau memperhitungkan faktor waktu (misalnya, perubahan status risiko mahasiswa antar semester atau tahun akademik). Data perlu dikumpulkan dalam bentuk yang mencerminkan perkembangan temporal atau dianalisis pada berbagai titik waktu kritis (misalnya, memprediksi risiko pada minggu ke-3, ke-6, dan ke-9 semester) [29].

d) Teknik Inti (Naive Bayes Dinamis): Model Naive Bayes (NB) adalah pengklasifikasi probabilistik yang didasarkan pada teorema Bayes dan asumsi independensi antar variabel prediktor [26], [27], [29].

Temuan dalam penelitian RQ2: Dengan cara apa perspektif filosofis Thomas Kuhn dan Karl Popper memengaruhi interpretasi dan penerapan teknik pemodelan dinamis dalam konteks penilaian risiko akademis adalah integrasi perspektif Thomas Kuhn dan Karl Popper secara mendasar memengaruhi bagaimana model Dynamic Naive Bayes diinterpretasikan, diuji, dan diterapkan dalam konteks penilaian risiko akademis:

1. Pengaruh Karl Popper (Falsifikasi dan Rasionalisme Kritis)

a) Interpretasi Model: Perspektif Popper menuntut agar model Dynamic Naive Bayes (sebagai suatu "teori" atau "dugaan" tentang risiko) harus selalu dianggap fallible (dapat salah) dan bersifat sementara. Interpretasi hasil model tidak boleh dogmatis, melainkan harus objektif, tidak bergantung pada opini komunitas tertentu [30], [31].

b) Penerapan Teknik Dinamis: Model ini harus diterapkan dengan fokus pada pengujian yang agresif terhadap hipotesisnya. Alih-alih mencari bukti yang memastikan kebenaran model (verifikasi), penerapannya harus secara aktif mencari data atau situasi yang dapat memfalsifikasi model tersebut. Dalam konteks risiko akademis, ini berarti mencari kasus di mana model memprediksi kegagalan (risiko tinggi) tetapi mahasiswa tersebut berhasil, atau sebaliknya, untuk mengidentifikasi batasan dan kesalahan model [25], [30].

c) Revisi Model: Gagalnya model Dynamic Naive Bayes dalam pengujian kritis (falsifikasi) akan memicu methodological revisionism (merevisi metodologi) atau pengembangan model yang lebih baik (trial and error). Popper meyakini bahwa pengetahuan ilmiah bertumbuh secara revolusioner, menggantikan teori lama dengan teori baru yang lebih progresif, bukan hanya sekadar melengkapi [30], [31].

2. Pengaruh Thomas Kuhn (Paradigma dan Ilmu Normal).

a) Interpretasi Model: Perspektif Kuhn menempatkan model Dynamic Naive Bayes dalam bingkai paradigma yang sudah ada dalam pendidikan tinggi (misalnya, definisi yang diterima mengenai "risiko akademik," "indikator keberhasilan," dan "data yang relevan"). Model ini berfungsi sebagai alat puzzle-solving (pemecahan teka-teki) yang bertujuan untuk mengatasi masalah yang ditentukan oleh paradigma yang berlaku [25], [32].

b) Penerapan Teknik Dinamis: Selama model bekerja dengan baik (ilmu normal), fokusnya adalah pada peningkatan presisi dan cakupan penerapannya dalam batas-batas yang ditetapkan oleh paradigma (aturan, metode, dan konsep yang disepakati). Jika model menghadapi anomali yang signifikan, kegagalan prediksi yang tidak dapat dijelaskan dengan penyesuaian kecil, maka penerapan teknik dinamis mungkin memasuki masa krisis [25], [32].

c) Pergeseran Paradigma: Keputusan untuk meninggalkan model Dynamic Naive Bayes dan mengadopsi pendekatan pemodelan risiko yang sama sekali baru (misalnya, beralih ke struktur pemodelan risiko yang berbeda secara fundamental) akan menjadi Pergeseran Paradigma (Paradigm Shift). Pergeseran ini mungkin dipengaruhi oleh nilai-nilai dan konsensus sosial komunitas peneliti, dan bukan hanya oleh bukti empiris murni. Kuhn juga menunjukkan bahwa pergeseran ini dapat menyebabkan incommensurability (inkomensurabilitas), di mana konsep risiko dalam model baru mungkin tidak dapat dibandingkan secara langsung dengan model Dynamic Naive Bayes yang lama [25], [32].

5. Kesimpulan

Pengembangan Model Dynamic Naive Bayes (DNB) merupakan upaya metodologis untuk mengatasi keterbatasan model statis seperti Naive Bayes dalam memprediksi risiko akademik mahasiswa pada data longitudinal. DNB memungkinkan pembaruan probabilitas secara dinamis seiring waktu, sehingga menghasilkan prediksi yang lebih akurat dan selaras dengan perubahan perilaku mahasiswa. Dari perspektif Thomas Kuhn, model statis berada pada fase normal science yang mengalami krisis akibat anomali berulang dalam menangani data temporal, sehingga memerlukan paradigm shift menuju DNB sebagai solusi puzzle solving dalam Educational Data Mining. Sementara Karl Popper melalui Rasionalisme Kritis menekankan falsifiability, di mana hipotesis DNB harus terus diuji secara agresif dan berpotensi dibantah melalui data baru, mendorong proses trial and error yang berkelanjutan. Integrasi kedua perspektif ini selaras dengan tujuan

penelitian, yaitu tidak hanya mengembangkan model dinamis yang lebih robust, tetapi juga memastikan pertumbuhan pengetahuan ilmiah yang kritis dan adaptif dalam pemodelan risiko akademik. Ke depan, penelitian lanjutan dapat mengkaji aspek filsafat ilmu yang belum tergali mendalam, seperti integrasi perspektif Imre Lakatos (research programmes) untuk mengevaluasi apakah DNB merupakan bagian dari progressive atau degenerating problemshift dalam domain Educational Data Mining, serta perspektif Paul Feyerabend (anarkisme metodologis) yang memungkinkan eksplorasi pendekatan non-ortodoks atau hibrida di luar kerangka Bayesian. Kajian ini akan memperkaya diskusi epistemologis sekaligus membuka peluang inovasi metodologis yang lebih pluralistik dalam prediksi risiko akademik berbasis data dinamis.

Daftar Pustaka

- [1] R. K. Smith, "Risk Factors in Education." pp. 1–32, 2010.
- [2] J. Johnson and others, "An Empirical Study of Naive Bayes for Risk Prediction," in *IEEE International Conference on Machine Learning*, 2018.
- [3] Y. U and T. H. Lee, "Dynamic Models with Sentiment Terms via Longitudinal Dataset," in *British Machine Learning Conference*, 2019, p. 53.1-53.1.
- [4] J. Aneja and A. G. Schwing, "Dynamic Bayesian Networks."
- [5] B. Dai, "Contrastive Learning for Risk Prediction," in *Advances in Neural Information Processing Systems Conference*, 2017, pp. 898–907.
- [6] K. Shuster, S. Humeau, H. Hu, A. Bordes, and J. Weston, "Engaging Image Captioning via Personality," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 12516–12526. doi: 10.1109/CVPR.2019.01280.
- [7] B. Dai and S. Fidler, "A Neural Compositional Paradigm for Risk Prediction." 2018.
- [8] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "BLEU: a Method for Automatic Evaluation of Machine Translation," in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2002, pp. 311–318. doi: 10.3115/1073083.1073135.
- [9] J. Devlin and others, "Machine Learning Models for Risk Prediction: The Quirks and What Works," *Comput Lang*, pp. 100–105, 2015.
- [10] V. Kougia, J. Pavlopoulos, and I. Androutsopoulos, "A Survey on Biomedical Image Captioning," *Proceedings of the Second Workshop on Shortcomings in Vision and Language*. pp. 26–36, 2019. doi: 10.18653/v1/W19-1803.
- [11] X. Li, X. Song, L. Herranz, Y. Zhu, and S. Jiang, "Image Captioning with both Object and Scene Information," in *Proceedings of the 24th ACM International Conference on Multimedia*, 2016, pp. 1107–1110. doi: 10.1145/2964284.2984069.
- [12] D. Shin and I. Kim, "Deep Image Understanding Using Multilayered Contexts," *Math. Probl. Eng.*, vol. 2018, pp. 1–11, 2018, doi: 10.1155/2018/5847460.
- [13] Z. Gan et al., "Semantic Compositional Networks for Visual Captioning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1141–1150. doi: 10.1109/CVPR.2017.127.
- [14] A. Karpathy and L. Fei-Fei, "Deep Visual-Semantic Alignments for Generating Image Descriptions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 664–676, 2017, doi: 10.1109/TPAMI.2016.2598339.
- [15] X. He, B. Shi, X. Bai, G. Xia, and Z. Zhang, "Risk Prediction with Part of Speech Guidance," *Pattern Recognit. Lett.*, vol. 0, pp. 1–9, 2017.
- [16] D. Mahayana, *Filsafat Ilmu Pengetahuan*. Bandung, Indonesia: ITB Press, 2018.
- [17] F. Serafino, G. Pio, and M. Ceci, "Ensemble Learning for Multi-Type Classification in Heterogeneous Networks," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 12, pp. 2326–2339, 2018, doi: 10.1109/TKDE.2018.2822307.
- [18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, 2012, pp. 1–9. [Online]. Available: <https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>
- [19] A. Bettahi, F.-Z. Belouadha, and H. Harroud, "A Modular and Explainable Machine Learning Pipeline for Student Dropout Prediction in Higher Education," *Algorithms*, vol. 18, no. 10, p. 662, Oct. 2025, doi: 10.3390/a18100662.
- [20] W. Xiaohui and C. Lianghui, "Using the Bayesian Network Method to Evaluate the Effectiveness of College Students' Mental Health Intervention Strategies and Their Impact on Academic Performance," *J. Psychoeduc. Assess.*, p. 07342829251393575, Oct. 2025, doi: 10.1177/07342829251393575.

- [21] A. Santoso, H. Retnawati, Kartianom, E. Apino, I. Rafi, and M. N. Rosyada, "Predicting Time to Graduation of Open University Students: An Educational Data Mining Study," *Open Educ. Stud.*, vol. 6, no. 1, p. 20220220, Feb. 2024, doi: 10.1515/edu-2022-0220.
- [22] D. Khairy, N. Alharbi, M. A. Amasha, M. F. Areed, S. Alkhalaf, and R. A. Abougalala, "Prediction of student exam performance using data mining classification algorithms," *Educ. Inf. Technol.*, vol. 29, no. 16, pp. 21621–21645, Nov. 2024, doi: 10.1007/s10639-024-12619-w.
- [23] V. Nakhipova, Y. Kerimbekov, Z. Umarova, H. I. Bulbul, L. Suleimenova, and E. Adylbekova, "Integration of Collaborative Filtering Into Naive Bayes Method to Enhance Student Performance Prediction:," *Int. J. Inf. Commun. Technol. Educ.*, vol. 20, no. 1, pp. 1–18, Aug. 2024, doi: 10.4018/IJICTE.352512.
- [24] A. Harif and M. A. Kassimi, "Predictive Modeling of Student Performance Using RFECV-RF for Feature Selection and Machine Learning Techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 7, 2024, doi: 10.14569/IJACSA.2024.0150723.
- [25] C. R. Maboloc, "On the Scientific Methods of Kuhn and Popper: Implications of Paradigm-Shifts to Development Models," *Philosophia*, vol. 46, no. 2, pp. 387–399, June 2018, doi: 10.1007/s11406-017-9891-3.
- [26] Quispe, "Early Prediction of University Student Dropout Using Machine Learning Models," *Nanotechnol. Percept.*, vol. 20, no. S5, June 2024, doi: 10.62441/nano-ntp.v20iS5.62.
- [27] S. A. Abdul Rahim, F. Sidi, L. S. Affendey, I. Ishak, and A. Y. Nurlankyzy, "Leveraging Data Lake Architecture for Predicting Academic Student Performance," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 14, no. 6, pp. 2121–2129, Dec. 2024, doi: 10.18517/ijaseit.14.6.12408.
- [28] M. H. Afif, "Exploring the quality of the higher educational institution website using data mining techniques," *Decis. Sci. Lett.*, vol. 12, no. 2, pp. 279–290, 2023, doi: 10.5267/j.dsl.2023.1.007.
- [29] H. A. Abdelhafez and H. Elmannai, "Developing and Comparing Data Mining Algorithms That Work Best for Predicting Student Performance:," *Int. J. Inf. Commun. Technol. Educ.*, vol. 18, no. 1, pp. 1–14, Feb. 2022, doi: 10.4018/IJICTE.293235.
- [30] O. Pylypchuk and O. Strelko, "Karl Popper's 'Critical Rationalism': The way to freedom and democracy," *Hist. Sci. Technol.*, vol. 14, no. 2, pp. 419–435, Dec. 2024, doi: 10.32703/2415-7422-2024-14-2-419-435.
- [31] F. Wall, "Management Control from a Critical Rationalist Perspective: A Discussion on Results Controls," *Philos. Manag.*, vol. 24, no. 2, pp. 183–206, June 2025, doi: 10.1007/s40926-025-00336-4.
- [32] M. R. Matthews, "Thomas Kuhn and Science Education: Learning from the Past and the Importance of History and Philosophy of Science," *Sci. Educ.*, vol. 33, no. 3, pp. 609–678, June 2024, doi: 10.1007/s11191-022-00408-1.