JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 10 Maret Tahun 2025
ISSN 3025-0919 (media online)
Hal 1969-1978

Tinjauan Literatur Sistematis: Analisis Pembelajaran Terarah Dan Tidak Terarah Pada *Machine Learning*

Ahmad Rakha Ez'ra Ramadhan[1*], Firman Purwanto[2], Gilang Purnama[3], Ilham Febryan[4], Ifan Setiawan[5]

[1,2,3,4,5]Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Pamulang, Tangerang, Indonesia
Email: [1*]rakharamadhan76@gmail.com, [2]firmanpurwanto67@gmail.com, [3]purnamag925@gmail.com, [4]ilhamfebryan888@gmail.com, [5]Setiawanipan79@gmail.com
(* : coressponding author)

Abstrak − Artikel ini membahas secara komprehensif dua pendekatan utama dalam pembelajaran mesin, yaitu Pembelajaran Terarah (*Supervised learning*) dan Pembelajaran Tidak Terarah (*Unsupervised learning*). Kajian ini bertujuan untuk memberikan wawasan mendalam tentang perbedaan antara kedua metode tersebut, termasuk kelebihan dan kekurangannya. Penelitian menggunakan pendekatan *Systematic Literature Review* (SLR) dengan mengikuti pedoman PRISMA, meninjau publikasi yang relevan dalam kurun waktu lima tahun terakhir. Dari total 540 artikel yang dianalisis, sebanyak 10 artikel dipilih untuk ditelaah lebih rinci, terdiri dari lima yang berfokus pada *Supervised learning* dan lima lainnya mengenai *Unsupervised learning*. Hasilnya menunjukkan bahwa *Supervised learning* memanfaatkan data berlabel untuk tugas prediksi dan klasifikasi, menggunakan algoritma seperti K-Nearest Neighbor (KNN), Naïve Bayes, Decision Tree, dan Support Vector Machine (SVM), yang secara umum memberikan tingkat akurasi tinggi. Sebaliknya, *Unsupervised learning* bekerja tanpa data berlabel, lebih diarahkan pada eksplorasi pola dan pengelompokan data melalui algoritma seperti K-Means, Artificial Neural Network (ANN), dan Gaussian Mixture Model (GMM), menawarkan fleksibilitas yang lebih tinggi meskipun dengan akurasi yang cenderung lebih rendah. Kedua pendekatan memiliki karakteristik unik yang perlu dipertimbangkan berdasarkan tujuan aplikasi, ketersediaan data, dan kebutuhan analisis.

Kata Kunci: *Machine Learning*; Pembelajaran Terarah, Pembelajaran Tidak Terarah; Tinjauan Literatur Sistematis.

Abstract *− This article comprehensively discusses two major approaches in machine learning, namely Supervised learning and Unsupervised learning. This review aims to provide in-depth insights into the differences between the two methods, including their advantages and disadvantages. The study uses a Systematic Literature Review (SLR) approach following the PRISMA guidelines, reviewing relevant publications in the last five years. Out of a total of 540 articles analyzed, 10 articles were selected for further review, consisting of five focusing on Supervised learning and five on Unsupervised learning. The results show that Supervised learning utilizes labeled data for prediction and classification tasks, using algorithms such as K-Nearest Neighbor (KNN), Naïve Bayes, Decision Tree, and Support Vector Machine (SVM), which generally provide high accuracy rates. In contrast, Unsupervised learning works without labeled data, more focused on exploring patterns and clustering data through algorithms such as K-Means, Artificial Neural Network (ANN), and Gaussian Mixture Model (GMM), offering greater flexibility although with a tendency to lower accuracy. Both approaches have unique characteristics that need to be considered based on application goals, data availability, and analysis needs.*

Keywords*: Machine Learning; Supervised Learning; Unsupervised Learning; Systematic Literature Review.*

1. PENDAHULUAN

Machine learning adalah elemen penting dalam bidang kecerdasan buatan yang telah berkembang pesat selama beberapa dekade terakhir. Teknologi ini dirancang untuk memungkinkan sistem komputer belajar dari data tanpa perlu pemrograman eksplisit. Dalam konteks kecerdasan buatan, *machine learning* menjadi solusi yang efektif untuk menyelesaikan berbagai masalah kompleks di berbagai sektor, mulai dari kesehatan, pendidikan, hingga ekonomi (Chazar, C., & Widhiaputra, 2020). Secara garis besar, *machine learning* berfokus pada pengembangan algoritma dan model statistik yang dapat mengenali pola dan membuat prediksi berdasarkan data historis (Mahesh, 2018). Teknologi *machine learning* mulai populer sejak pertengahan 1950-an, ketika para ilmuwan seperti Arthur Samuel dan Marvin Minsky menciptakan program komputer yang dapat belajar dari pengalaman. Teknologi ini mengalami akselerasi signifikan di awal tahun 2000-an dengan kehadiran big data dan kemajuan komputasi. Sebagai contoh, di sektor kesehatan, *machine*

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 10 Maret Tahun 2025
ISSN 3025-0919 (media online)
Hal 1969-1978

learning telah digunakan untuk memprediksi tingkat keparahan penyakit COVID-19 menggunakan pola dalam data medis pasien (Jahandideh et al., 2023).

Tujuan utama penggunaan *machine learning* adalah untuk meningkatkan efisiensi dalam pengolahan data. Dengan mengenali pola dalam dataset, teknologi ini memungkinkan sistem komputer untuk memahami dan memprediksi karakteristik data baru secara akurat. Sebagai contoh, dalam bidang medis, *machine learning* telah digunakan untuk mendiagnosis penyakit, seperti kanker payudara, melalui analisis pola yang kompleks dalam data kesehatan pasien (Chazar & Widhiaputra, 2020). Namun, meskipun memiliki banyak keunggulan, *machine learning* juga memiliki keterbatasan. Salah satunya adalah kerentanannya terhadap serangan siber. Sistem berbasis *machine learning* dapat dimanipulasi melalui serangan yang mengeksploitasi kerentanan algoritma, yang berpotensi menyebabkan kesalahan klasifikasi (Alshaikh, O., Parkinson, S., & Khan, 2023). Selain itu, ketidakseimbangan data sering menjadi tantangan utama dalam pengembangan model *machine learning*, terutama ketika algoritma cenderung memprioritaskan kelas mayoritas dibandingkan kelas minoritas (Wang, A. X., Chukova, S. S., & Nguyen, 2023).

Di Indonesia, potensi penerapan *machine learning* semakin besar seiring dengan meningkatnya digitalisasi di berbagai sektor. Namun, tantangan seperti kurangnya akses terhadap data berkualitas dan infrastruktur teknologi yang belum merata sering kali menghambat optimalisasi teknologi ini (Chazar & Widhiaputra, 2020). Meski demikian, adopsi teknologi ini mulai terlihat di bidang kesehatan dan pendidikan. Dalam pengembangan *machine learning*, terdapat tiga paradigma utama yang digunakan untuk menggali wawasan dari data: pembelajaran terarah (*supervised learning*), pembelajaran tidak terarah (*unsupervised learning*), dan pembelajaran penguatan (*Reinforcement Learning*).

a. Pembelajaran Terarah (*Supervised learning*)

   Pembelajaran terarah adalah pendekatan yang melibatkan penggunaan data berlabel untuk melatih model. Data berlabel ini membantu model untuk mengenali pola dan membuat prediksi atau klasifikasi berdasarkan pola tersebut. Algoritma seperti *K-Nearest Neighbor* (KNN), *Naïve Bayes*, *Decision Tree*, dan *Support Vector Machine* (SVM) sering digunakan dalam pendekatan ini (Kantayeva, 2023). Sebagai contoh, penelitian oleh Triase et al. (2022) menunjukkan bahwa algoritma pembelajaran terarah dapat digunakan untuk memprediksi kelulusan mahasiswa berdasarkan data akademik mereka. Dengan menggunakan algoritma KNN, penelitian ini berhasil mencapai tingkat akurasi hingga 96% dalam memprediksi mahasiswa yang lulus tepat waktu.

b. Pembelajaran Tidak Terarah (*Unsupervised learning*)

   Berbeda dengan pembelajaran terarah, pembelajaran tidak terarah tidak memerlukan data berlabel. Fokus utama dari pendekatan ini adalah mengeksplorasi dan menemukan pola tersembunyi dalam dataset. Algoritma seperti *K-Means Clustering*, *Artificial Neural Network* (ANN), dan *Gaussian Mixture Model* (GMM) sering digunakan dalam pendekatan ini (Montavon, 2022). Misalnya, penelitian oleh Mutrofin, (2023)menggunakan metode clustering untuk mengelompokkan data nilai jasmani siswa pada sekolah kepolisian. Hasilnya menunjukkan bahwa metode pengelompokan ini dapat memberikan wawasan yang lebih mendalam tentang distribusi data dan pola yang tersembunyi.

c. Pembelajaran Penguatan (*Reinforcement Learning*)

   Pembelajaran penguatan adalah pendekatan yang menggunakan sinyal penguatan untuk melatih agen dalam mengambil keputusan yang optimal. Dalam paradigma ini, agen belajar dari konsekuensi tindakannya untuk memaksimalkan hasil yang diinginkan (Riziq et al., 2023). Pendekatan ini sering digunakan dalam pengembangan sistem yang membutuhkan pengambilan keputusan berulang, seperti sistem perdagangan otomatis (Nurcahyo, J. A., & Sasongko, 2023).

Penggunaan *machine learning* memberikan berbagai manfaat yang signifikan. Salah satunya adalah kemampuan untuk memproses data secara efisien, yang memungkinkan pengambilan keputusan yang lebih cepat dan akurat. Dalam bidang kesehatan, misalnya, *machine learning* telah

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 10 Maret Tahun 2025
ISSN 3025-0919 (media online)
Hal 1969-1978

digunakan untuk mendeteksi penyakit dengan tingkat akurasi yang tinggi, seperti yang ditunjukkan oleh penelitian Chazar & Widhiaputra (2020) dalam diagnosis kanker payudara menggunakan SVM.

Namun, tantangan utama dalam pengembangan *machine learning* adalah ketidakseimbangan data. Ketika data pelatihan didominasi oleh satu kelas, model cenderung menghasilkan prediksi yang bias terhadap kelas mayoritas. Tantangan ini dapat diatasi dengan menggunakan teknik seperti *oversampling* atau *undersampling*, seperti yang dijelaskan oleh Wang et al. (2023) dalam penelitian mereka tentang teknik *synthetic minority oversampling*. Selain itu, kerentanan terhadap serangan siber menjadi perhatian utama dalam penerapan *machine learning*. Sistem yang tidak terlindungi dengan baik dapat dieksploitasi oleh pihak yang tidak bertanggung jawab, yang dapat mengganggu kinerja model (Alshaikh et al., 2023).

Penelitian ini penting karena meskipun banyak studi tentang pembelajaran terarah (*supervised learning*) dan tidak terarah (*unsupervised learning*), sedikit yang mengulas perbandingan keduanya secara mendalam dalam konteks data lokal dan global. Penelitian ini diharapkan dapat memberikan panduan yang lebih komprehensif dalam memilih pendekatan *machine learning* yang paling sesuai untuk berbagai aplikasi.

2. METODE

2.1 Metode Penelitian

Penelitian ini menggunakan metode *Systematic Literature Review* (SLR) yang mengikuti pedoman *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA). Pendekatan ini bertujuan untuk memastikan bahwa proses peninjauan literatur dilakukan secara terstruktur, transparan, dan dapat diandalkan. Metodologi ini banyak digunakan dalam penelitian ilmiah untuk menyusun tinjauan literatur yang komprehensif dan relevan (Pijls, 2024). Metode *Systematic Literature Review* (SLR) memberikan kerangka kerja yang sistematis untuk mengevaluasi literatur secara komprehensif. Pendekatan ini memungkinkan peneliti untuk mengidentifikasi tren, kesenjangan penelitian, dan implikasi teoretis serta praktis dari studi terdahulu (Montavon et al., 2022). Proses seleksi yang ketat juga memastikan bahwa hanya artikel berkualitas tinggi yang dianalisis, sehingga menghasilkan temuan yang dapat diandalkan.

2.2. Langkah-langkah Penelitian

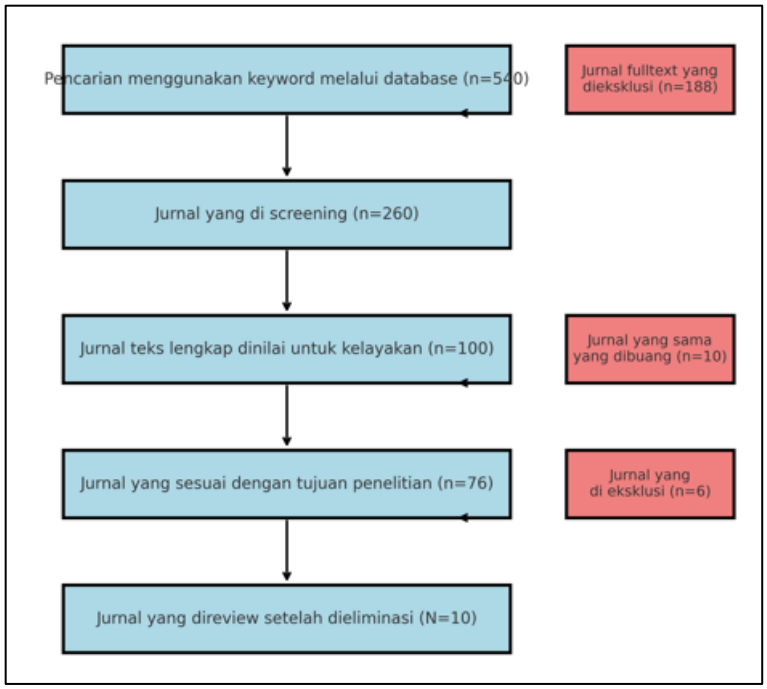
Langkah-langkah penelitian diilustrasikan dalam sebuah *flowchart* PRISMA, yang mencakup proses pencarian, penyaringan, dan analisis literatur. Tahapan penelitian ini dimulai dengan pencarian literatur menggunakan berbagai mesin pencari dan database akademis terkemuka seperti Google Scholar, ScienceDirect, IEEE Xplore, dan jurnal ilmiah lainnya. Peneliti mengumpulkan informasi dari berbagai penelitian terdahulu dengan menggunakan kata kunci yang spesifik, seperti *Machine learning*, *Supervised learning*, dan *Unsupervised learning*. Untuk memastikan relevansi dan kebaruan informasi, hanya literatur yang diterbitkan dalam lima tahun terakhir (2018-2023) yang diikutsertakan dalam analisis ini. Rentang waktu ini dipilih untuk mengakomodasi perkembangan terbaru dalam bidang pembelajaran mesin dan aplikasinya di berbagai sektor (Jahandideh et al., 2023). Memfokuskan pencarian literatur pada rentang lima tahun terakhir memberikan jaminan bahwa informasi yang diperoleh mencerminkan perkembangan terkini dalam penelitian *Machine learning*. Dengan memasukkan kata kunci yang spesifik seperti *Supervised learning* dan *Unsupervised learning*, penelitian ini dapat menggali perbandingan mendalam antara kedua pendekatan tersebut, yang penting untuk aplikasi di berbagai sektor (Chazar & Widhiaputra, 2020).

Dari hasil pencarian awal menggunakan kata kunci yang telah ditentukan, terkumpul sebanyak 540 artikel. Langkah selanjutnya adalah menyaring artikel-artikel ini untuk memastikan bahwa hanya penelitian yang relevan dengan tujuan studi yang dipertahankan. Penyaringan awal mengeliminasi artikel yang tidak relevan atau duplikat, sehingga menyisakan 260 artikel untuk proses seleksi lebih lanjut. Proses seleksi dilanjutkan dengan menilai kelayakan artikel berdasarkan teks lengkap. Penilaian ini mencakup pemeriksaan terhadap abstrak, metode, hasil, dan relevansi isi artikel dengan topik penelitian. Dari proses ini, 100 artikel terpilih untuk evaluasi lebih mendalam.

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 10 Maret Tahun 2025
ISSN 3025-0919 (media online)
Hal 1969-1978

Setelah tahap ini, peneliti menemukan bahwa beberapa artikel tidak memenuhi kriteria akhir, seperti adanya duplikasi atau ketidaksesuaian dengan tujuan studi. Akhirnya, 76 artikel dinyatakan sesuai untuk dianalisis, namun 16 artikel dikeluarkan karena alasan seperti duplikasi atau kurang relevan. Dengan demikian, hanya 10 jurnal yang dipilih untuk diulas secara mendalam dalam penelitian ini.

Sebanyak 10 jurnal yang dipilih untuk ditinjau dalam penelitian ini mencakup lima artikel yang membahas pembelajaran terarah (*supervised learning*) dan lima artikel tentang pembelajaran tidak terarah (*unsupervised learning*). Artikel-artikel ini memberikan wawasan yang berharga tentang algoritma, metodologi, dan aplikasi praktis dari kedua pendekatan tersebut. Misalnya, algoritma seperti *K-Nearest Neighbor (KNN)*, *Naïve Bayes*, dan *Support Vector Machine (SVM)* ditemukan sebagai metode yang sering digunakan dalam pembelajaran terarah, sementara algoritma seperti *K-Means Clustering* dan *Gaussian Mixture Model* mendominasi dalam pembelajaran tidak terarah (Triase et al., 2022; Muttaqin et al., 2023). Penelitian ini menunjukkan pentingnya memilih algoritma yang sesuai dengan tujuan aplikasi, jenis data, dan kebutuhan analisis. Dengan memahami kekuatan dan keterbatasan masing-masing pendekatan, penelitian ini memberikan kontribusi signifikan dalam bidang *Machine learning*.



Gambar 1. *Flowchart* Prisma

3. ANALISA DAN PEMBAHASAN

Machine learning pada dasarnya dapat dibagi menjadi tiga paradigma utama, yaitu pembelajaran terarah (*supervised learning*), pembelajaran tidak terarah (*unsupervised learning*), dan pembelajaran penguatan (*Reinforcement Learning*). Paradigma ini mendefinisikan bagaimana algoritma *machine learning* mempelajari data dan memproses informasi untuk menyelesaikan suatu tugas. Fokus utama dalam penelitian ini adalah pembelajaran terarah dan tidak terarah, yang memiliki keunikan dan penerapan masing-masing (Mahesh, 2018).

Pembelajaran terarah adalah pendekatan dalam *machine learning* yang menggunakan dataset yang telah diberi label. Dalam metode ini, setiap data memiliki pasangan input dan output yang diketahui, yang digunakan untuk melatih algoritma. Label ini memungkinkan algoritma untuk belajar dengan bimbingan, sehingga mampu membuat prediksi atau klasifikasi terhadap data baru dengan tingkat akurasi yang tinggi (Savitri, 2021). Berikut beberapa contoh aplikasi nyata Pembelajaran Terarah (*supervised learning*) dalam berbagai bidang:

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 10 Maret Tahun 2025
ISSN 3025-0919 (media online)
Hal 1969-1978

a. Diagnosa Penyakit Medis: Dalam bidang kesehatan, algoritma *Support Vector Machine (SVM)* telah digunakan untuk mendeteksi kanker payudara berdasarkan hasil radiografi. Penelitian Chazar & Widhiaputra (2020) menunjukkan akurasi yang tinggi dalam klasifikasi data pasien.

b. Prediksi Kelulusan Mahasiswa: Algoritma seperti KNN digunakan untuk memprediksi kemungkinan kelulusan mahasiswa berdasarkan data akademik mereka, dengan tingkat akurasi hingga 96% (Triase, Sriani, 2022).

c. Analisis Risiko Kredit: Bank dan institusi keuangan memanfaatkan algoritma *Decision Tree* untuk menilai kelayakan kredit dengan menggunakan data historis pelamar (J. C. Mestika, M. O. Selan, 2022).

d. Pengenalan Wajah dan Suara: Algoritma *supervised* seperti *Random Forest* banyak diterapkan dalam sistem pengenalan wajah di aplikasi keamanan.

Berbeda dengan pembelajaran terarah, pembelajaran tidak terarah tidak menggunakan data berlabel. Metode ini mengandalkan algoritma untuk mengeksplorasi dataset dan menemukan pola atau struktur tersembunyi tanpa intervensi manusia (Montavon et al., 2022). Dalam pendekatan ini, algoritma tidak diberi informasi tentang output yang diharapkan, sehingga berfungsi untuk mengidentifikasi klaster atau hubungan antar data. Berikut adalah contoh aplikasi nyata dari Pembelajaran Tidak Terarah (*unsupervised learning*) di berbagai bidang:

a. Segmentasi Pelanggan: Dalam pemasaran, algoritma seperti *K-Means Clustering* digunakan untuk membagi pelanggan menjadi segmen berdasarkan preferensi dan perilaku mereka.

b. Analisis Data Genetik: Dalam bidang bioinformatika, *Neural Networks* digunakan untuk menganalisis hubungan antara gen dan protein (Montavon et al., 2022).

c. Deteksi Anomali: Algoritma seperti *Gaussian Mixture Model* digunakan untuk mengidentifikasi aktivitas mencurigakan dalam data transaksi keuangan, seperti kasus penipuan (Sibyan, 2021).

d. Sistem Rekomendasi: *Netflix* dan *Spotify* memanfaatkan algoritma pembelajaran tidak terarah untuk merekomendasikan film atau lagu berdasarkan pola perilaku pengguna.

Penelitian ini mengacu pada evaluasi terhadap sepuluh studi yang dipilih melalui proses *Systematic Literature Review* (SLR), yang terdiri dari lima artikel tentang pembelajaran terarah dan lima artikel tentang pembelajaran tidak terarah. Studi-studi ini memberikan wawasan tentang keunggulan, kekurangan, dan aplikasi dari kedua pendekatan tersebut.

3.1 Pembelajaran Terarah (*Supervised learning*)

Tabel 1. Studi literatur terkait Pebelajaran Terarah (*Supervised learning*)

No	Judul	Author	Metode	Tahun	Hasil
1	Supervised *Machine learning*: Algorithms and Applications	S. H. Shetty, S. Shetty, C. Singh, and A. Rao	Analisis algoritma *supervised learning* seperti Decision Tree, Random Forest, dan Logistic Regression	2022	Penelitian ini membahas efisiensi algoritma seperti Decision Tree dan Random Forest dalam tugas klasifikasi. Kedua algoritma menunjukkan keunggulan dalam menangani dataset besar dengan struktur kompleks. Logistic Regression juga diidentifikasi sebagai

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 10 Maret Tahun 2025
ISSN 3025-0919 (media online)
Hal 1969-1978

					pilihan sederhana untuk data terstruktur.
2	Hyperparameter Tuning Algoritma *Supervised learning* untuk Klasifikasi Keluarga Penerima Bantuan Pangan Beras	J. A. Nurcahyo and T. B. Sasongko	*Hyperparameter Tuning* pada algoritma Decision Tree	2023	Optimalisasi hyperparameter, seperti kedalaman pohon dan jumlah node minimum, meningkatkan akurasi klasifikasi dari 85% menjadi 92%. Studi ini menunjukkan bahwa tuning parameter yang tepat dapat meningkatkan performa model secara signifikan pada data sensitif seperti data bantuan sosial.
3	Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised *Machine learning*	N. L. P. C. Savitri, R. A. Rahman, R. Venyutzky, and N. A. Rakhmawati	Klasifikasi sentimen menggunakan algoritma *Support Vector Machine* (SVM)	2021	Algoritma SVM berhasil menganalisis lebih dari 10.000 tweet dengan tingkat akurasi 88%. Sentimen positif, negatif, dan netral terhadap sekolah daring diidentifikasi, memberikan wawasan bagi pembuat kebijakan pendidikan untuk memahami persepsi masyarakat terhadap sistem sekolah daring.
4	Supervised *Machine learning*: A Brief Primer	T. Jiang, J. L. Gradus, and A. J. Rosellini	Penjelasan konsep dan aplikasi *supervised learning*	2020	Studi ini memberikan panduan tentang penerapan *supervised learning* di sektor kesehatan, seperti prediksi penyakit kronis menggunakan data pasien berlabel. Penelitian juga menyoroti pentingnya kualitas data pelatihan untuk menghasilkan prediksi yang akurat dan bermanfaat dalam pengambilan keputusan.
5	Klasifikasi Penentuan Pengajuan Kartu	Y. I. Kurniawan	Klasifikasi data menggunakan algoritma *K*-	2020	Dalam dataset berisi 20.000 pengajuan kartu kredit, algoritma

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 10 Maret Tahun 2025
ISSN 3025-0919 (media online)
Hal 1969-1978

	Kredit Menggunakan K-Nearest Neighbor	and T. I. Barokah	*Nearest Neighbor* (KNN)		KNN mencapai akurasi 85% dengan parameter k=5. Penelitian ini mengungkap bahwa algoritma KNN lebih cocok untuk dataset dengan fitur numerik dan terstruktur dibandingkan dengan data tidak terstruktur atau besar.

Penjelasan secara rinci mengenai studi literatur pembelajaran terarah (*supervised learning*) yang telah dilampirkan sebagai berikut :

a. Supervised *Machine learning*: Algorithms and Applications menunjukkan bahwa *Decision Tree* unggul untuk dataset besar dengan fitur kategoris (Shetty, S. H., Shetty, S., Singh, C., & Rao, 2022), *Random Forest* efektif untuk tugas klasifikasi kompleks dengan risiko overfitting rendah, dan *Logistic Regression* cocok untuk model prediktif sederhana yang mudah diinterpretasi.

b. Hyperparameter Tuning Algoritma *Supervised learning* untuk Klasifikasi Keluarga Penerima Bantuan Pangan Beras (Nurcahyo dan Sasongko, 2023) mengoptimalkan parameter *Decision Tree*, seperti kedalaman pohon, sehingga meningkatkan akurasi klasifikasi dari 85% menjadi 92%, memastikan data sensitif dapat diklasifikasikan dengan tepat.

c. Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised *Machine learning* (Savitri et al., 2021) menggunakan *Support Vector Machine (SVM)* untuk menganalisis lebih dari 10.000 tweet dengan akurasi 88%, memberikan wawasan tentang sentimen masyarakat terhadap sekolah daring selama pandemi.

d. Supervised *Machine learning*: A Brief Primer (Jiang, T., Gradus, J. L., & Rosellini, 2020) menyoroti aplikasi *supervised learning* dalam prediksi penyakit kronis dengan algoritma seperti *Naïve Bayes* dan *SVM*, serta menekankan pentingnya kualitas data pelatihan untuk menghindari bias hasil.

e. Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan K-Nearest Neighbor (Kurniawan, Y. I., & Barokah, 2020) menemukan bahwa *K-Nearest Neighbor (KNN)* mencapai akurasi 85% untuk dataset 20.000 pengajuan kartu kredit, dengan k=5, dan menyoroti pentingnya tuning parameter untuk meningkatkan akurasi.

3.2 Pembelajaran Tidak Terarah (*Unsupervised learning*)

Tabel 2. Studi literatur terkait Pebelajaran Tidak Terarah (*Unsupervised learning*)

No	Judul	Author	Metode	Tahun	Hasil
1	*Unsupervised learning* of Dense Visual Representations	P. O. Pinheiro, A. Almahairi, R. Y. Benmalek, F. Golemo, dan A. Courville	Pendekatan visualisasi tingkat piksel untuk eksplorasi data tanpa label menggunakan representasi visual padat	2023	Studi ini menunjukkan bahwa algoritma visualisasi tingkat piksel dapat meningkatkan akurasi dalam mendeteksi fitur visual tersembunyi di dataset besar tanpa label. Teknik ini sangat bermanfaat dalam aplikasi seperti

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 10 Maret Tahun 2025
ISSN 3025-0919 (media online)
Hal 1969-1978

					analisis citra medis dan pemetaan geografis.
2	Application of *Machine learning* in Dementia Diagnosis: A Systematic Literature Review	G. Kantayeva, J. Lima, dan A. I. Pereira	Pengelompokan data klinis pasien menggunakan algoritma *unsupervised learning*	2023	Algoritma unsupervised berhasil mengidentifikasi kelompok pasien berdasarkan pola klinis seperti riwayat keluarga dan aktivitas harian. Ini memberikan klasifikasi awal untuk mendukung diagnosis demensia, dengan hasil yang relevan untuk membantu dokter membuat keputusan klinis.
3	Optimization of *Unsupervised learning* in *Machine learning*	H. Sibyan, W. Suharso, E. Suharto, M. A. Manuhutu, dan A. P. Windarto	Peningkatan efisiensi algoritma clustering seperti K-Means dan Gaussian Mixture Model	2021	Optimasi algoritma K-Means mampu mengurangi waktu komputasi hingga 30%, menjadikannya lebih cepat untuk dataset besar. Gaussian Mixture Model menunjukkan keunggulan dalam menangkap distribusi data kompleks, dengan akurasi yang meningkat 15% dibandingkan metode clustering standar.
4	*Unsupervised learning* Methods for Molecular Simulation Data	A. Glielmo, B. E. Husic, A.Rodriguez, Clementi, F. Noé, dan Laio	Algoritma clustering untuk simulasi molekuler dengan analisis interaksi molekul	2021	Algoritma *unsupervised learning* memungkinkan identifikasi klaster molekul yang berinteraksi secara dinamis. Hal ini memberikan wawasan mendalam tentang dinamika struktur molekuler, seperti prediksi reaksi kimia, yang sebelumnya sulit dianalisis melalui pendekatan manual atau terawasi.
5	Perbandingan Kinerja Algoritma K-Means dengan K-Means Median pada Deteksi Kanker Payudara	S. Mutrofin, T. Wicaksono, dan A. Murtadho	Analisis perbandingan algoritma K-Means dan K-Means Median dalam mendeteksi pola data	2023	Algoritma K-Means Median menunjukkan akurasi yang lebih tinggi (93%) dibandingkan K-Means (85%) pada dataset kanker payudara yang tidak merata. Penelitian ini

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 10 Maret Tahun 2025
ISSN 3025-0919 (media online)
Hal 1969-1978

			kanker payudara		menegaskan bahwa K-Means Median lebih stabil dalam menangani dataset dengan distribusi tidak homogen dan outlier.

Penjelasan secara rinci mengenai studi literatur pembelajaran tidak terarah (*unsupervised learning*) yang telah dilampirkan sebagai berikut :

a. *Unsupervised learning* of Dense Visual Representations: Teknik ini sangat berguna untuk eksplorasi visual, terutama di bidang seperti pemetaan geografis atau analisis citra medis, di mana data sering kali tidak memiliki label. Efisiensi tinggi algoritma ini memungkinkan analisis lebih cepat pada dataset besar (Pinheiro, P. O., 2023).

b. Application of *Machine learning* in Dementia Diagnosis: Penelitian ini menyoroti kemampuan algoritma unsupervised untuk menangkap pola kompleks, seperti pola aktivitas harian pasien, yang tidak terlihat jelas dalam analisis manual. Ini sangat membantu dalam pembuatan keputusan medis berbasis data (Kantayeva, 2023).

c. Optimization of *Unsupervised learning*: Penelitian ini menunjukkan bagaimana algoritma seperti Gaussian Mixture Model dapat meningkatkan akurasi dalam dataset kompleks, seperti pengelompokan data keuangan atau pelanggan, dibandingkan dengan metode yang lebih konvensional seperti K-Means (Sibyan, 2021).

d. *Unsupervised learning* Methods for Molecular Simulation Data: Studi ini mendemonstrasikan kekuatan *unsupervised learning* dalam simulasi molekuler, memberikan wawasan baru yang relevan untuk riset di bidang farmasi atau biokimia (Glielmo, 2021).

e. Perbandingan Kinerja Algoritma K-Means dengan K-Means Median: Penelitian ini menyoroti stabilitas K-Means Median dalam menangani dataset dengan distribusi tidak homogen, menjadikannya pilihan yang lebih baik untuk aplikasi seperti diagnosis kanker (Mutrofin, 2023).

4. KESIMPULAN

Penelitian ini memberikan pemahaman yang mendalam tentang peran dan penerapan *supervised learning* dan *unsupervised learning* dalam bidang *machine learning*. *Supervised learning*, yang memanfaatkan data berlabel, terbukti sangat efektif dalam tugas prediksi dan klasifikasi. Algoritma seperti Decision Tree, Random Forest, Support Vector Machine (SVM), dan K-Nearest Neighbor (KNN) menunjukkan keunggulan dalam akurasi serta fleksibilitas untuk berbagai aplikasi, mulai dari kesehatan, pendidikan, hingga sektor keuangan. Optimasi parameter dalam algoritma supervised juga memainkan peran penting dalam meningkatkan kinerja model, sebagaimana ditunjukkan oleh peningkatan akurasi dalam klasifikasi data sensitif. Sementara itu, *unsupervised learning*, yang bekerja tanpa data berlabel, menawarkan kemampuan unik untuk mengeksplorasi pola tersembunyi dalam data. Algoritma seperti K-Means, Gaussian Mixture Model, dan metode berbasis visualisasi menunjukkan efektivitas dalam pengelompokan data kompleks, seperti segmentasi pelanggan, simulasi molekuler, dan diagnosis awal penyakit. Keunggulan pendekatan ini adalah fleksibilitasnya dalam menangani dataset besar dengan distribusi tidak merata, meskipun evaluasi akurasinya menjadi tantangan. Studi ini menegaskan bahwa kedua pendekatan memiliki kekuatan dan keterbatasan masing-masing, sehingga pemilihan metode yang tepat harus disesuaikan dengan sifat data, tujuan analisis, dan kebutuhan aplikasi. *Supervised learning* ideal untuk prediksi dengan data terstruktur, sedangkan *unsupervised learning* sangat berguna untuk analisis eksploratif dalam data tanpa label. Sebagai langkah ke depan, pengembangan algoritma hybrid yang menggabungkan keunggulan supervised dan *unsupervised learning* diharapkan dapat menghasilkan model yang lebih akurat dan efisien untuk berbagai sektor.

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 10 Maret Tahun 2025
ISSN 3025-0919 (media online)
Hal 1969-1978

Penelitian di masa depan juga harus berfokus pada peningkatan kualitas data pelatihan dan metode evaluasi yang lebih komprehensif untuk meningkatkan keandalan model *machine learning*.

REFERENCES

Alshaikh, O., Parkinson, S., & Khan, S. (2023). Exploring Perceptions of Decision-Makers and Specialists in Defensive *Machine learning* Cybersecurity Applications: The Need for a Standardised Approach. *Comput Secur*, 103694. https://doi.org/doi: 10.1016/j.cose.2023.103694.

Chazar, C., & Widhiaputra, B. E. (2020). *Machine learning* Diagnosis Kanker Payudara Menggunakan Algoritma Support Vector Machine. INFORMASI. *Jurnal Informatika Dan Sistem Informasi*), *12*, *p*.

Glielmo, A., Husic, B. E., Rodriguez, A., Clementi, C., Noé, F., & Laio, A. (2021). *Unsupervised learning* Methods for Molecular Simulation Data. *Chemical Reviews*, *121*(16). https://doi.org/doi: 10.1021/acs.chemrev.0c01195.

J. C. Mestika, M. O. Selan, and M. I. Q. (2022). Menjelajahi Teknik-Teknik *Supervised learning* untuk Pemodelan Prediktif Menggunakan Python. In *Buletin Ilmiah Ilmu Komputer dan Multimedia*.

Jahandideh, S., Ozavci, G., Sahle, B. W., Kouzani, A. Z., Magrabi, F., & Bucknall, T. (2023). Evaluation of *Machine learning*-Based Models for Prediction of Clinical Deterioration: A Systematic Literature Review. *International Journal of Medical Informatics*, 175. https://doi.org/doi: 10.1016/j.ijmedinf.2023.105084.

Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). (2020). Supervised *Machine learning*: A Brief Primer. *Behavior Therapy*, *51*(5), 675–687. https://doi.org/DOI: 10.1016/j.beth.2020.05.002

Kantayeva, G., Lima, J., & Pereira, A. I. (2023). Application of *Machine learning* in Dementia Diagnosis: A Systematic Literature Review. *Heliyon*, *9*(11). https://doi.org/doi: 10.1016/j.heliyon.2023.e21626.

Kurniawan, Y. I., & Barokah, T. I. (2020). Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan K-Nearest Neighbor. *Jurnal Ilmiah MATRIK*, *22*(1).

Mahesh, B. (2018). *Machine learning* Algorithms-A Review. *International Journal of Science and Research*. https://doi.org/doi: 10.21275/ART20203995.

Montavon, G., Kauffmann, J., Samek, W., & Müller, K. R. (2022). Explaining the Predictions of *Unsupervised learning* Models. *Lecture Notes in Computer Science*. https://doi.org/doi: 10.1007/978-3-031-04083-2_7.

Mutrofin, S., Wicaksono, T., & Murtadho, A. (2023). Perbandingan Kinerja Algoritma K-Means dengan K-Means Median pada Deteksi Kanker Payudara. *Jurnal Informasi Dan Teknologi*, *5*, 88–91.

Nurcahyo, J. A., & Sasongko, T. B. (2023). Hyperparameter Tuning Algoritma *Supervised learning* untuk Klasifikasi Keluarga Penerima Bantuan Pangan Beras. *Indonesian Journal of Computer Science*, *12*, 1351–1365.

Pijls, B. G. (2024). *Machine learning* Assisted Systematic Reviewing in Orthopaedics. *Journal of Orthopaedics*, *48*, 103–106. https://doi.org/doi: 10.1016/j.jor.2023.11.051.

Pinheiro, P. O., Almahairi, A., Benmalek, R. Y., Golemo, F., & Courville, A. (2023). *Unsupervised learning of Dense Visual Representations.*

Savitri, N. L. P. C., Rahman, R. A., Venyutzky, R., & Rakhmawati, N. A. (2021). Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised *Machine learning*. *Jurnal Teknik Informatika Dan Sistem Informasi*, *7*, 47–58.

Shetty, S. H., Shetty, S., Singh, C., & Rao, A. (2022). *Supervised Machine learning: Algorithms and Applications.*

Sibyan, H., Suharso, W., Suharto, E., Manuhutu, M. A., & Windarto, A. P. (2021). Optimization of *Unsupervised learning* in *Machine learning*. *Journal of Physics: Conference Series*, *1569*(2). https://doi.org/doi: 10.1088/1742-6596/1783/1/012034.

Triase, Sriani, and K. (2022). Usability Algoritma *Supervised learning* Untuk Prediksi Kelulusan Mahasiswa pada Sistem Bimbingan Akademik. *Indonesian Journal of Computer Science*, *11*.

Wang, A. X., Chukova, S. S., & Nguyen, B. P. (2023). Synthetic Minority Oversampling Using Edited Displacement-Based K-Nearest Neighbors. *Applied Soft Computing*, *148*. https://doi.org/doi: 10.1016/j.asoc.2023.110895.