



Tren dan Perkembangan *Supervised* versus *Unsupervised Learning*

Muhammad Naufal Ammr Dzakwan¹, Fadillah Dani Prawoto²,
Ahmad Nur Ihsan Purwanto^{3*}

¹⁻³Universitas Ary Ginanjar, Jakarta, Indonesia

Alamat Kampus: Menara 165, Cilandak No.kav 1 lt 2 & 18, Cilandak Tim., Ps. Minggu, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12560

Korespondensi penulis: ahmadnur.ihsan@esqbs.ac.id*

Abstract The rapid development of artificial intelligence (AI) in the last decade has driven the widespread adoption of machine learning in various sectors, from healthcare and finance to manufacturing and education. Two main approaches dominate: supervised and unsupervised learning. Supervised learning leverages labeled data to build predictive models with high accuracy and strong generalization capabilities, but its application is often hampered by the need for expensive, time-consuming, and error-prone labeling processes. In contrast, unsupervised learning operates on unlabeled data to identify latent patterns, clustering, and novel data representations. This approach excels in large-data scale and exploration efficiency, but faces challenges in interpreting results, validating performance, and implementing them in real-world systems. This article presents a recent literature review (2020–2024) based on searches in Scopus, IEEE Xplore, and Google Scholar using the keywords “supervised learning,” “unsupervised learning,” “machine learning trends,” and “hybrid approaches.” The results show a significant trend toward integrating hybrid approaches such as semi-supervised learning, self-supervised learning, and a combination of transfer learning approaches, aiming to overcome the limitations of each paradigm. Furthermore, research gaps were identified regarding the need for robust evaluation methods for unsupervised learning, the development of adaptive models for dynamic data, and increased transparency in the learning process. Critically, this article highlights the importance of an interdisciplinary approach in designing hybrid algorithms that are not only efficient and accurate, but also interpretive and reliable for data-driven decision-making. Recommendations for future research directions focus on the creation of a standardized evaluation framework, the integration of AI with edge computing technology, and the exploration of the potential of hybrid learning in the context of real-time data.

Keywords: Machine Learning, Self- Supervised Learning, Semi-Supervised Learning, Supervised Learning, Unsupervised Learning.

Abstrak. Perkembangan pesat kecerdasan buatan (AI) dalam dekade terakhir telah mendorong adopsi luas machine learning di berbagai sektor, mulai dari kesehatan, keuangan, manufaktur, hingga pendidikan. Dua pendekatan utama yang mendominasi adalah *supervised learning* dan *unsupervised learning*. *Supervised learning* memanfaatkan data berlabel untuk membangun model prediktif dengan akurasi tinggi dan kemampuan generalisasi yang kuat, namun penerapannya sering terkendala oleh kebutuhan proses *labeling* yang mahal, memakan waktu, serta rawan kesalahan manusia. Sebaliknya, *unsupervised learning* beroperasi pada data tidak berlabel untuk mengidentifikasi pola laten, klusterisasi, serta representasi data baru. Pendekatan ini unggul dalam skala data besar dan efisiensi eksplorasi, tetapi menghadapi tantangan dalam hal interpretasi hasil, validasi kinerja, dan penerapan pada sistem nyata. Artikel ini menyajikan tinjauan literatur terkini (2020–2024) berdasarkan pencarian pada basis data Scopus, IEEE Xplore, dan Google Scholar dengan kata kunci “supervised learning”, “unsupervised learning”, “machine learning trends”, dan “hybrid approaches”. Hasil kajian menunjukkan tren signifikan menuju integrasi pendekatan *hybrid* seperti *semi-supervised learning*, *self-supervised learning*, serta kombinasi *transfer learning* yang bertujuan mengatasi keterbatasan masing-masing paradigma. Selain itu, ditemukan adanya kesenjangan riset terkait kebutuhan metode evaluasi yang *robust* bagi *unsupervised learning*, pengembangan model adaptif terhadap data dinamis, serta peningkatan transparansi dalam proses pembelajaran. Secara kritis, artikel ini menyoroti pentingnya pendekatan interdisipliner dalam merancang algoritma *hybrid* yang tidak hanya efisien dan akurat, tetapi juga interpretatif dan dapat diandalkan untuk pengambilan keputusan berbasis data. Rekomendasi arah penelitian ke depan difokuskan pada penciptaan kerangka kerja evaluasi yang standar, integrasi AI dengan teknologi *edge computing*, serta eksplorasi potensi *hybrid learning* dalam konteks data real-time.

Kata kunci: Pembelajaran Mandiri, Pembelajaran Mesin, Pembelajaran Semi-Terawasi, Pembelajaran Tanpa Pengawasan, Pembelajaran Terawasi.

1. LATAR BELAKANG

Artificial Intelligence (AI) telah menjadi bagian integral dalam sistem informasi modern. AI memungkinkan sistem untuk melakukan analisis data secara otomatis, mengidentifikasi pola, dan memberikan rekomendasi berbasis kecerdasan buatan. Dengan kemampuannya yang terus berkembang, AI telah digunakan dalam berbagai sektor, termasuk bisnis, kesehatan, dan pemerintahan (Antares, Hasugian, & Satria, 2025). Machine Learning menggunakan algoritma untuk memahami pola dan hubungan yang kompleks dalam data dibandingkan mengandalkan pendekatan berbasis aturan (Rijcken et al., 2022). Untuk menggali pengetahuan dari data, dapat digunakan tiga paradigma utama: *Supervised Learning*, *Unsupervised Learning*, dan *Reinforcement Learning* (Andreanus & Kurniawan, n.d.).

Dalam *Supervised Learning*, metode ini menggunakan data yang sudah diklasifikasikan sebagai dasar untuk prediksi. Sedangkan dalam *Unsupervised Learning*, metode ini menggunakan data yang tidak memiliki label sehingga harus menemukan pola dari data sendiri (Boja, 2025). Kedua paradigma ini memiliki tujuan yang sama, yaitu menghasilkan model yang mampu mengenali pola dalam data, namun perbedaan mendasar dalam pendekatan mereka memiliki dampak signifikan terhadap hasil yang diperoleh (Gupta, Pathak, & Kumar, 2022). Hal ini memungkinkan pengguna *Machine Learning* untuk membuat keputusan dengan akurasi di atas rata-rata (Nur Ihsan Purwanto, et.al, 2024).

Berdasarkan hal tersebut, rumusan masalah dalam penelitian ini difokuskan pada tiga pertanyaan utama: bagaimana tren dan perkembangan *supervised learning* dan *unsupervised learning* dalam lima tahun terakhir, apa kelebihan, kekurangan, dan gap riset utama dari kedua pendekatan tersebut, serta bagaimana kontribusi *hybrid approaches* dalam menjawab tantangan yang ada. Penelitian ini bertujuan memberikan tinjauan kritis, mendalam, dan berbasis literatur ilmiah mutakhir guna menawarkan rekomendasi yang relevan dan aplikatif.

Perkembangan kecerdasan buatan (Artificial Intelligence/AI) yang semakin pesat telah menempatkan machine learning sebagai salah satu teknologi inti dalam transformasi digital berbagai sektor, mulai dari kesehatan, pendidikan, bisnis, hingga industri kreatif. Dua pendekatan utama yang banyak digunakan adalah supervised learning dan unsupervised learning. Supervised learning memungkinkan komputer belajar dari data berlabel untuk menghasilkan model prediktif yang akurat, sedangkan unsupervised learning bekerja dengan data tidak berlabel untuk menemukan pola tersembunyi atau klusterisasi dalam suatu himpunan data. Kedua pendekatan ini memiliki keunggulan sekaligus keterbatasan yang menjadikannya saling melengkapi dalam praktik.

Dalam beberapa tahun terakhir, tren riset menunjukkan adanya pergeseran dari penggunaan tunggal menuju integrasi keduanya dalam bentuk pendekatan hybrid seperti semi-supervised learning dan self-supervised learning. Hal ini mencerminkan kebutuhan akan metode yang lebih adaptif, efisien, dan mampu menangani data dalam skala besar sekaligus kompleks. Di sisi lain, tantangan juga muncul terkait keterbatasan dalam evaluasi hasil unsupervised learning serta kebutuhan standarisasi dalam penerapan supervised learning. Oleh karena itu, pembahasan mengenai perkembangan terkini kedua pendekatan ini menjadi penting untuk memberikan gambaran komprehensif mengenai arah penelitian, peluang, dan tantangan yang akan dihadapi dalam pengembangan machine learning di masa depan.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kajian literatur sebagai kerangka metodologis utama, dengan tujuan untuk memperoleh pemahaman yang mendalam dan komprehensif mengenai dampak kebijakan sertifikasi halal wajib terhadap daya saing Usaha Mikro, Kecil, dan Menengah (UMKM), khususnya di wilayah Jakarta. Pemilihan metode ini merujuk pada definisi yang disampaikan oleh Zulfikar dan Mayu (2020), yang menjelaskan bahwa kajian literatur merupakan proses sistematis yang mencakup tahapan identifikasi, evaluasi kritis, serta integrasi terhadap beragam temuan penelitian dan gagasan konseptual yang relevan. Seluruh tahapan tersebut dilakukan secara berulang, tersusun secara sistematis, dan disampaikan secara transparan sehingga memungkinkan replikasi dan validasi dalam kerangka ilmiah yang dapat dipertanggungjawabkan (At Thur Jasson Cassidy, et.al, 2024).

Data dalam penelitian ini diperoleh melalui analisis mendalam terhadap berbagai literatur yang berkaitan langsung dengan isu sertifikasi halal serta implikasinya terhadap daya saing UMKM. Dalam proses ini, peneliti melakukan kajian kritis terhadap publikasi yang diterbitkan oleh individu maupun institusi, dengan penekanan pada sumber-sumber yang bersumber dari situs resmi dan lembaga yang memiliki otoritas serta kredibilitas dalam bidang sertifikasi halal.

3. HASIL DAN PEMBAHASAN

Tren Terkini

Supervised learning tetap menjadi andalan untuk tugas prediksi dan klasifikasi berkat performa tinggi pada data berlabel, didukung oleh kemajuan deep learning dan transfer learning (Gupta et al., 2022). Unsupervised learning berkembang pesat dalam aplikasi *exploratory data analysis*, *anomaly detection*, dan *clustering*, terutama pada data besar dan heterogen (Abukmeil et al., 2021). *Hybrid approaches* seperti *semi-supervised* dan *self-supervised learning* menjadi

tren utama, memanfaatkan data berlabel dan tidak berlabel secara efisien untuk meningkatkan performa dan generalisasi model (Yang et al., 2021). Supervised learning: didominasi oleh CNN, RNN, dan *transfer learning*. Banyak dipakai untuk diagnosis medis (Aljuadi & Anwar, 2022) dan *cybersecurity* (Alif Auladi, et.al, 2025). Unsupervised learning: populer di *clustering* genomik, segmentasi pasien, hingga analisis *fraud detection*. Teknik K-Means, SOM, dan PCA mendominasi (Massi et al., 2020).

Tabel 1. Tabel Perbandingan Keunggulan & Kelemahan (Oye et al., n.d.)

Aspek	Supervised Learning	Unsupervised Learning
Data	Membutuhkan data berlabel besar (mahal)	Tidak memerlukan label
Akurasi	Umumnya lebih tinggi (pada prediksi)	Lebih eksploratif, akurasi bervariasi
Interpretasi	Lebih mudah diverifikasi	Sulit dijelaskan (<i>black-box effect</i>)
Biaya	Biaya labeling tinggi	Biaya labeling minimal
Aplikasi	Klasifikasi citra, diagnosis, prediksi harga	Clustering, segmentasi, <i>anomaly detection</i>

Gap dan Kritik

Supervised: ketergantungan tinggi pada data berlabel → mahal dan lambat (Boja, 2025). Unsupervised: sulit diinterpretasikan, rawan bias, hasil *clustering* kadang ambigu (Locatello et al., 2018). Tantangan *fairness* dan *explainability* menuntut pengembangan XAI (*explainable AI*) untuk meningkatkan kepercayaan publik (Grabowicz et al., 2022).

Kontribusi Pemikiran

Integrasi *active learning* dan *transfer learning* dapat mempercepat *labeling* pada supervised learning. Pengembangan metrik evaluasi baru untuk unsupervised learning, misal berbasis konsistensi atau validasi eksternal, sangat dibutuhkan (Zailani Iman, et.al, 2025). Kolaborasi antara ahli domain dan *data scientist* penting untuk meningkatkan interpretasi hasil unsupervised dan *hybrid approaches*. Tantangan *fairness* dan *explainability* menuntut pengembangan XAI (*explainable AI*) untuk meningkatkan kepercayaan publik (Grabowicz et al., 2022).

Keterbatasan Studi

kajian ini sepenuhnya didasarkan pada literatur sekunder yang bersumber dari jurnal, prosiding konferensi, dan preprint yang tersedia secara daring. Pendekatan ini berpotensi menimbulkan bias publikasi, di mana penelitian yang tidak dipublikasikan atau menghasilkan temuan negatif mungkin tidak terliput (Syarof Azzaky, et.al, 2024). Selain itu, keterbatasan akses terhadap artikel berbayar dapat menyebabkan adanya celah informasi pada pembahasan. Cakupan waktu penelitian dibatasi hanya pada publikasi lima tahun terakhir, sehingga pembahasan mengenai perkembangan awal atau konsep dasar yang masih relevan menjadi terbatas. Studi ini juga tidak melakukan pengujian eksperimental terhadap metode yang direview, sehingga kesimpulan yang diperoleh bersifat teoretis dan sangat bergantung pada

kredibilitas serta kualitas sumber yang digunakan. Akibatnya, temuan yang disajikan lebih bersifat indikatif dan memerlukan pembuktian lebih lanjut melalui riset empiris..

4. KESIMPULAN DAN SARAN

Supervised dan unsupervised learning memiliki peran saling melengkapi dalam ekosistem machine learning modern. Supervised learning unggul dalam akurasi prediksi pada data berlabel, namun terbatas oleh biaya labeling. Unsupervised learning efisien dalam eksplorasi data besar tanpa label, meski interpretasinya menantang. Hybrid approaches seperti semi-supervised dan self-supervised learning menjadi tren utama untuk mengatasi gap tersebut, namun masih menghadapi tantangan interpretabilitas dan efisiensi. Ke depan, riset perlu difokuskan pada pengembangan metode evaluasi robust untuk unsupervised learning, model adaptif terhadap data dinamis, serta integrasi explainable AI untuk meningkatkan kepercayaan dan adopsi di industri.

Untuk mendukung perkembangan tersebut, kolaborasi antara akademisi, industri, dan pembuat kebijakan perlu ditingkatkan guna menciptakan standar evaluasi yang terukur, dataset benchmark yang representatif, serta regulasi yang mendorong transparansi dan akuntabilitas AI. Selain itu, pelatihan sumber daya manusia di bidang machine learning dan XAI harus diperluas agar inovasi yang dihasilkan tidak hanya canggih secara teknis, tetapi juga relevan dan dapat diterima secara sosial.

DAFTAR REFERENSI

- Abukmeil, M., Ferrari, S., Genovese, A., Piuri, V., & Scotti, F. (2021). A survey of unsupervised generative models for exploratory data analysis and representation learning. *ACM Computing Surveys*, 54(5), Article 99. <https://doi.org/10.1145/3450963>
- Alif Auladi, G., & Muwahid, F. (2025). Cybersecurity and moral responsibility: A philosophical-Islamic approach to digital trust. *TechComp Innovations: Journal of Computer Science and Technology*, 2(1), 53–65. <https://doi.org/10.70063/techcompinnovations.v2i1.93>
- Aljuadi, A., & Anwar, M. (2022). Survey of supervised learning for medical image processing. *SN Computer Science*, 3, Article 392. <https://doi.org/10.1007/s42979-022-01166-1>
- Andreanus, J., & Kurniawan, A. (n.d.). Sejarah, teori dasar dan penerapan reinforcement learning: Sebuah tinjauan pustaka. *Jurnal Telematika*, 12(2). Institut Teknologi Harapan Bangsa, Bandung. <https://doi.org/10.61769/telematika.v12i2.193>
- Antares, J., Hasugian, B. S., & Satria, W. (2025). Integrasi artificial intelligence dalam sistem informasi: Meningkatkan efisiensi dan otomatisasi. *Device: Journal of Information System, Computer Science and Information Technology*, 6(1), 256–265. <https://doi.org/10.46576/device.v6i1.6627>

- At Thur Jasson Cassidy, A., Fuad, A., & Ulil Abshor As Shofy, M. (2024). Emerging trends and challenges in digital crime: A study of cyber criminal tactics and countermeasures. *TechComp Innovations: Journal of Computer Science and Technology*, 1(1), 38–45. <https://doi.org/10.70063/techcompinnovations.v1i1.25>
- Boja, R. R. (2025). A comparative study of supervised and unsupervised learning approaches. *International Multidisciplinary Research Journal Reviews (IMRJR)*, 2(4). <https://doi.org/10.17148/IMRJR.2025.020411>
- Grabowicz, P., Perello, N., & Mishra, A. (2022). Marrying fairness and explainability in supervised learning. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 165–174). ACM. <https://doi.org/10.1145/3531146.3533236>
- Gupta, J., Pathak, S., & Kumar, G. (2022). Deep learning (CNN) and transfer learning: A review. *Journal of Physics: Conference Series*, 2273(1), 012029. <https://doi.org/10.1088/1742-6596/2273/1/012029>
- Locatello, F., Bauer, S., Lucic, M., Rätsch, G., Gelly, S., Schölkopf, B., & Bachem, O. (2018). Challenging common assumptions in the unsupervised learning of disentangled representations. *arXiv preprint arXiv:1811.12359*. <https://arxiv.org/abs/1811.12359>
- Massi, M. C., Ieva, F., & Lettieri, E. (2020). Data mining application to healthcare fraud detection: A two-step unsupervised clustering method for outlier detection with administrative databases. *BMC Medical Informatics and Decision Making*, 20, Article 160. <https://doi.org/10.1186/s12911-020-01143-9>
- Nur Ihsan Purwanto, A., & Abdan Hanif, A. (2024). Strategic synergy: Integrating business management with computer science for competitive advantage. *TechComp Innovations: Journal of Computer Science and Technology*, 1(1), 10–18. <https://doi.org/10.70063/techcompinnovations.v1i1.23>
- Ouali, Y., Hudelot, C., & Tami, M. (2020). An overview of deep semi-supervised learning. *arXiv preprint arXiv:2006.05278*. <https://doi.org/10.48550/arXiv.2006.05278>
- Oye, E., Frank, E., & Owen, J. (n.d.). Unsupervised and supervised models in machine learning: Unsupervised vs. supervised learning. Retrieved from <https://www.researchgate.net/publication/387136555>
- Rijcken, E., Zervanou, K., Mosteiro, P., Scheepers, F., Spruit, M., & Kaymak, U. (2022, December 2). Machine learning vs. rule-based methods for document classification of electronic health records within psychiatry: A systematic literature review. *Research Square*. <https://doi.org/10.21203/rs.3.rs-2320804/v1>
- Syarof Azzaky, N., Salimah, A., & Risma Saputri, C. (2024). Revolutionizing business: The role of AI in driving Industry 4.0. *TechComp Innovations: Journal of Computer Science and Technology*, 1(1), 28–37. <https://doi.org/10.70063/techcompinnovations.v1i1.24>

- Yang, X., Song, Z., King, I., & Xu, Z. (2021). A survey on deep semi-supervised learning. *arXiv preprint arXiv:2103.00550*. <https://arxiv.org/abs/2103.00550>
- Zailani Iman, M., Monic Veronica, A., & Airlangga Asis, A. (2025). Managing machine learning integration in English language curriculum: Challenges and innovations in teacher training. *TechComp Innovations: Journal of Computer Science and Technology*, 2(1), 16–27. <https://doi.org/10.70063/techcompinnovations.v2i1.91>