

## ANALISIS BIBLIOMETRIK DAN SLR: TREN PENERAPAN JST DALAM PENILAIAN PEMBELAJARAN FISIKA BERBASIS *VIRTUAL LABORATORY*

**Delia Achadina Putri, Sri Endang Wahyuni\*, dan Popi Purwanti**

*Universitas Indraprasta PGRI, Indonesia*

*\*Email: sherie.endang.wahyuni@gmail.com .*

### **Abstrak**

Sistem penilaian pembelajaran fisika masih didominasi oleh pendekatan konvensional yang cenderung subjektif dan belum mampu mengakomodasi analisis data pembelajaran yang kompleks. Perkembangan teknologi, khususnya *virtual laboratory*, menghasilkan data interaksi belajar yang berpotensi dimanfaatkan untuk meningkatkan kualitas penilaian, namun pemanfaatannya masih terbatas. Penelitian ini bertujuan untuk mengidentifikasi tren penelitian, menganalisis perkembangan penerapan jaringan saraf tiruan (JST), serta memetakan peluang pengembangan sistem penilaian pembelajaran berbasis *virtual laboratory* pada periode 2016–2026. Metode yang digunakan adalah *Systematic Literature Review* (SLR) yang dikombinasikan dengan analisis bibliometrik. Data diperoleh dari database Scopus dengan 343 artikel yang teridentifikasi pada periode 2016–2026, kemudian disaring menjadi 143 artikel berdasarkan kriteria tertentu, dan diperoleh 20 artikel akhir untuk dianalisis lebih lanjut. Hasil analisis menunjukkan bahwa pendekatan *machine learning*, khususnya JST, banyak digunakan dalam prediksi dan evaluasi hasil belajar siswa dengan tingkat akurasi yang tinggi. Namun, sebagian besar penelitian masih berfokus pada hasil akhir pembelajaran dan belum mengintegrasikan data proses belajar secara menyeluruh. Selain itu, integrasi antara JST dan *virtual laboratory* dalam sistem penilaian pembelajaran masih sangat terbatas. Penelitian ini menegaskan bahwa pengembangan sistem penilaian berbasis JST yang terintegrasi dengan *virtual laboratory* berpotensi menghasilkan evaluasi pembelajaran yang lebih komprehensif, adaptif, dan berbasis data dalam pembelajaran fisika.

**Kata-kata kunci:** Jaringan Saraf Tiruan, *Virtual Laboratory*, Analisis Bibliometrik, *Systematic Literature Review*.

### **Abstract**

Physics learning assessment systems are still predominantly based on conventional approaches that tend to be subjective and are unable to accommodate the analysis of complex learning data. Technological developments, particularly virtual laboratories, generate learning interaction data that have the potential to improve the quality of assessment; however, their utilization remains limited. This study aims to identify research trends, analyze the development of artificial neural network (ANN) applications, and map opportunities for developing virtual laboratory-based learning assessment systems during the 2016–2026 period. The method employed is a Systematic Literature Review (SLR) combined with bibliometric analysis. Data were obtained from the Scopus database, yielding 343 articles identified during the 2016–2026 period. These articles were subsequently screened based on specific criteria, resulting in 143 articles, from which 20 final articles were selected for further analysis. The results indicate that machine learning approaches, particularly ANNs, are widely used for predicting and evaluating students' learning outcomes with high levels of accuracy. However, most studies still focus on final learning outcomes and have not comprehensively integrated learning process data. Furthermore, the integration of ANNs and virtual laboratories into learning assessment systems remains very limited. This study highlights that the development of ANN-based assessment systems

*integrated with virtual laboratories has the potential to provide more comprehensive, adaptive, and data-driven assessment in physics education.*

**Keywords :** *Artificial Neural Networks, Virtual Laboratory, Bibliometric Analysis, Systematic Literature Review.*

## I. PENDAHULUAN

Sistem penilaian pembelajaran fisika masih didominasi pendekatan konvensional yang bergantung pada interpretasi manual. Proses tersebut menghasilkan subjektivitas tinggi dan keterbatasan dalam mengolah data pembelajaran. Kondisi ini sejalan dengan temuan bahwa sistem penilaian berbasis manual memiliki keterbatasan dalam pengolahan data yang kompleks dan memerlukan pendekatan berbasis komputasi (Zola et al., 2018). Percepatan transformasi digital pasca pandemi mendorong sekolah mengoptimalkan pemanfaatan teknologi dalam pembelajaran. Perubahan ini tidak hanya memengaruhi media dan strategi mengajar, tetapi juga menuntut sistem penilaian yang objektif, efisien, dan mampu memantau proses belajar.

Pemanfaatan *virtual laboratory* telah digunakan sebagai alternatif dalam pembelajaran fisika. Media ini memungkinkan simulasi praktikum tanpa keterbatasan ruang dan alat. Penelitian menunjukkan bahwa *virtual laboratory* digunakan sebagai solusi keterbatasan fasilitas laboratorium dan meningkatkan efektivitas pembelajaran (Nurfadila & Arlita, 2025). Selain itu, penggunaan simulasi seperti PhET terbukti meningkatkan penguasaan konsep peserta didik dalam pembelajaran fisika (Musahir, 2022). Namun, sistem penilaian pada lingkungan tersebut masih berfokus pada hasil akhir dan belum mengakomodasi proses interaksi belajar secara menyeluruh. Meskipun demikian, metode evaluasi dalam pembelajaran berbasis *virtual laboratory* masih didominasi oleh pendekatan konvensional yang belum mampu mengakomodasi analisis proses belajar secara menyeluruh (Wang et al., 2024).

Penggunaan *virtual laboratory* menghasilkan data interaksi yang bersifat kompleks dan dinamis. Data tersebut mencakup aktivitas eksperimen, waktu respons, serta hasil simulasi peserta didik. Sebagian besar pendekatan penilaian belum memanfaatkan data tersebut secara optimal. Padahal, data digital tersebut memiliki potensi untuk dianalisis lebih lanjut dalam menentukan capaian pembelajaran secara lebih akurat. Hasil kajian literatur menunjukkan bahwa penggunaan *virtual laboratory* dalam pembelajaran fisika dapat meningkatkan berbagai keterampilan, termasuk *critical thinking*, *conceptual understanding*, dan *science process skills* (Darman et al., 2024).

Pendekatan berbasis kecerdasan buatan mulai digunakan untuk mengatasi keterbatasan tersebut. Salah satu metode yang banyak digunakan adalah Jaringan Syaraf Tiruan (JST). Metode ini memiliki kemampuan dalam mengenali pola dan melakukan prediksi berbasis data. Penelitian menunjukkan bahwa JST dapat digunakan untuk “memprediksi prestasi siswa” berdasarkan data pembelajaran (Zola et al., 2018). Pendekatan machine learning telah digunakan dalam pembelajaran fisika untuk memprediksi capaian belajar siswa serta melakukan penilaian otomatis terhadap hasil belajar berbasis data pembelajaran (Santoso et al., 2023). Selain itu, metode ini juga digunakan untuk mengolah data penilaian yang bersifat kuantitatif dan kualitatif agar mendekati penilaian pengajar. Hal ini didukung oleh penelitian yang menyatakan bahwa sistem berbasis JST dapat memanfaatkan data hasil belajar untuk diklasifikasikan dan diprediksi secara adaptif dalam proses penilaian siswa (N. Kapita et al., 2020). Dengan demikian, JST memiliki potensi untuk diterapkan dalam sistem penilaian pembelajaran.

Sejumlah penelitian telah mengkaji penggunaan JST dalam bidang pendidikan. Penelitian lain menunjukkan bahwa model JST mampu memprediksi capaian belajar siswa berdasarkan pola data pembelajaran dengan tingkat akurasi yang tinggi dalam analisis performa akademik (Lau et al., 2019; Yang, 2024). Studi lain juga menunjukkan bahwa metode backpropagation digunakan untuk memprediksi nilai akademik dan membantu pengambilan keputusan pendidikan (Handayani et al., 2021). Meskipun demikian, penelitian tersebut masih dilakukan secara parsial dan belum terintegrasi dengan sistem penilaian berbasis *virtual laboratory* dalam pembelajaran fisika.

Keterbatasan lain ditemukan pada belum adanya pemetaan penelitian yang terstruktur terkait topik ini. Informasi mengenai tren penelitian, perkembangan publikasi, serta keterkaitan antar konsep

masih tersebar pada berbagai sumber. Kondisi ini menyulitkan identifikasi arah penelitian yang potensial serta pengembangan model penilaian yang berbasis data.

Pendekatan *Systematic Literature Review* (SLR) digunakan untuk mengidentifikasi dan mensintesis temuan penelitian secara sistematis melalui seleksi literatur berbasis kriteria tertentu sehingga mengurangi bias. Analisis dilengkapi dengan *bibliometric analysis* untuk memetakan tren publikasi, pola sitasi, serta keterkaitan antar topik penelitian. Kebaruan penelitian terletak pada integrasi SLR dan analisis bibliometrik dalam mengkaji penerapan JST pada sistem penilaian pembelajaran fisika berbasis *virtual laboratory*, yang masih jarang diteliti secara terpadu. Penelitian ini bertujuan untuk mengidentifikasi tren publikasi, menganalisis perkembangan penerapan JST dalam sistem penilaian pembelajaran, serta memetakan peluang pengembangan penelitian pada lingkungan *virtual laboratory* fisika.

## II. METODE PENELITIAN

Penelitian ini menggunakan pendekatan *Systematic Literature Review* (SLR) yang dikombinasikan dengan analisis bibliometrik untuk mengidentifikasi, mengevaluasi, dan mensintesis penelitian terkait penerapan JST dalam sistem penilaian pembelajaran fisika berbasis *virtual laboratory*. Pendekatan SLR dipilih karena mampu menghasilkan kajian yang sistematis, transparan, dan terstruktur berdasarkan prosedur yang telah ditetapkan. Pelaksanaan SLR dalam penelitian ini mengacu pada pedoman PRISMA 2020 yang digunakan untuk memastikan proses identifikasi, seleksi, dan sintesis literatur dilakukan secara sistematis. Pedoman ini menyediakan kerangka kerja berupa tahapan identifikasi, penyaringan, kelayakan, dan inklusi dalam proses kajian literatur (Page et al., 2021).

Desain penelitian menggunakan metode SLR dengan tahapan utama meliputi: (1) perumusan pertanyaan penelitian, (2) pencarian literatur, (3) seleksi studi, (4) ekstraksi data, dan (5) sintesis hasil. Metode ini memungkinkan identifikasi penelitian yang relevan secara sistematis dan meminimalkan bias dalam pemilihan literatur (Tedja et al., 2024).

Pencarian literatur dilakukan menggunakan dua database utama, yaitu Scopus dan Google Scholar. Database Scopus digunakan untuk memperoleh artikel bereputasi terkait kecerdasan buatan dan sistem penilaian, sedangkan Google Scholar digunakan untuk melengkapi konteks pembelajaran fisika dan *virtual laboratory*. Pencarian dilakukan menggunakan kombinasi kata kunci berbasis operator Boolean sebagai berikut:

("neural network" OR "machine learning") AND ("student assessment" OR "learning evaluation") AND ("prediction" OR "classification")

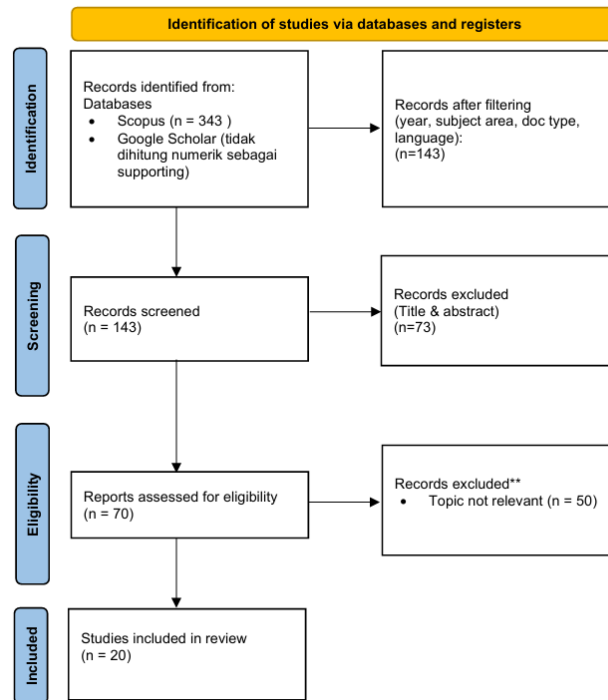
("virtual laboratory" OR "virtual lab") AND ("physics education") AND ("assessment" OR "evaluation")

Hasil pencarian pada database Scopus kemudian dilakukan penyaringan berdasarkan tahun publikasi (2016–2026), bidang kajian (computer science, engineering, dan social science), jenis dokumen (article dan conference paper), serta bahasa (English). Analisis data dilakukan melalui proses ekstraksi informasi penting dari setiap artikel, seperti metode yang digunakan, fokus serta hasil penelitian. Selanjutnya, dilakukan sintesis untuk mengidentifikasi tren penelitian, pola penggunaan metode, serta kesenjangan penelitian. Analisis bibliometrik digunakan untuk memetakan perkembangan publikasi, pola sitasi, dan keterkaitan antar topik penelitian.

## III. HASIL DAN PEMBAHASAN

Proses seleksi literatur dilakukan secara sistematis mengacu pada pedoman PRISMA 2020 yang meliputi tahap identifikasi dan penyaringan. Pada tahap identifikasi, pencarian melalui database Scopus menghasilkan sebanyak 343 artikel. Selanjutnya dilakukan penyaringan awal berdasarkan kriteria yang telah ditentukan, meliputi tahun publikasi (2016–2026), bidang kajian, jenis dokumen, dan bahasa, sehingga jumlah artikel berkurang menjadi 143 artikel. Tahap berikutnya adalah screening melalui analisis judul dan abstrak. Pada tahap ini, sebanyak 73 artikel dieliminasi karena tidak relevan dengan fokus penelitian, sehingga tersisa 70 artikel untuk ditelaah lebih lanjut.

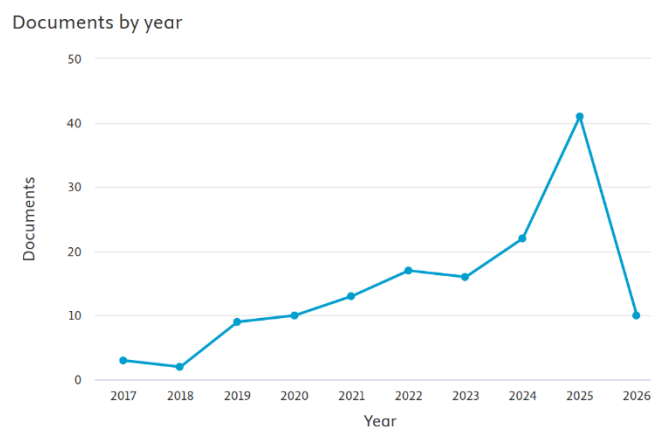
Selanjutnya dilakukan seleksi lanjutan berdasarkan kesesuaian topik penelitian, khususnya pada penerapan kecerdasan buatan dalam sistem penilaian pembelajaran. Hasilnya, sebanyak 50 artikel dieliminasi, sehingga diperoleh 20 artikel akhir yang dianalisis dalam penelitian ini. Selain itu, artikel dari Google Scholar digunakan sebagai sumber pelengkap untuk memperkaya konteks pembelajaran fisika berbasis *virtual laboratory*, namun tidak dimasukkan dalam perhitungan kuantitatif utama. Diagram PRISMA pada penelitian ini disajikan dalam Gambar 1.



**Gambar 1.** Diagram PRISMA Penelitian

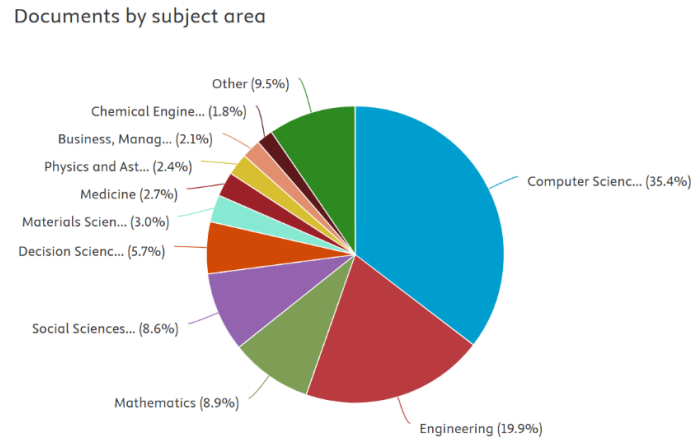
Setelah proses seleksi literatur dilakukan, langkah selanjutnya adalah melakukan analisis bibliometrik terhadap artikel yang diperoleh untuk mengidentifikasi tren publikasi, distribusi bidang kajian, serta pola penelitian dalam topik yang dikaji. Analisis ini bertujuan untuk memberikan gambaran umum mengenai perkembangan penelitian terkait penerapan kecerdasan buatan dalam sistem penilaian pembelajaran.

Hasil analisis bibliometrik berdasarkan data Scopus disajikan dalam beberapa visualisasi, meliputi distribusi publikasi berdasarkan tahun, negara, dan bidang kajian. Grafik publikasi berdasarkan tahun ditunjukkan pada Gambar 2, yang menggambarkan perkembangan jumlah artikel dalam rentang waktu penelitian.



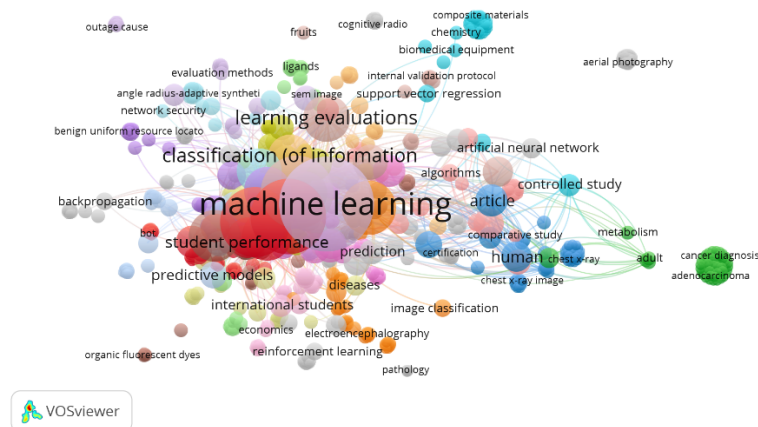
**Gambar 2.** Grafik Publikasi Scopus dengan Topik Terkait Berdasarkan Tahun

Selanjutnya, distribusi bidang kajian (*subject area*) ditampilkan pada Gambar 3 untuk mengidentifikasi bidang ilmu yang mendominasi penelitian pada topik ini.



**Gambar 3.** Grafik Publikasi Scopus dengan Topik Terkait Berdasarkan Subject Area

Untuk melengkapi analisis bibliometrik, dilakukan pemetaan hubungan antar kata kunci menggunakan perangkat lunak VOSviewer. Analisis ini bertujuan untuk mengidentifikasi keterkaitan topik penelitian serta pola kemunculan kata kunci dalam studi yang dianalisis. Visualisasi jaringan kata kunci (*network visualization*) disajikan pada Gambar 4.



**Gambar 4.** Visualisasi Jaringan Kata Kunci dengan VOSviewer

Berdasarkan hasil analisis tersebut, selanjutnya dilakukan kajian lebih mendalam terhadap artikel yang memiliki tingkat relevansi tertinggi. Dari total 20 artikel yang dianalisis, dipilih 10 artikel utama yang memiliki tingkat relevansi tertinggi terhadap fokus penelitian. Artikel-artikel ini merepresentasikan perkembangan penelitian terkait penerapan kecerdasan buatan dalam sistem penilaian pembelajaran yang disajikan dalam Tabel 1.

**Tabel 1.** Artikel Inti Hasil Screening

| No | Penulis, Tahun                   | Judul Artikel                                                                                                                                                  | Fokus Penelitian                          | Metode                                                            | Temuan                                                           |
|----|----------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------|-------------------------------------------------------------------|------------------------------------------------------------------|
| 1  | (Cheng, 2026)                    | Prediction and Classification of Student Performance Through an Ensemble Machine Learning Approach                                                             | Prediksi performa siswa                   | Ensemble ML (RF, SVM, AdaBoost, GBM) + optimasi genetic algorithm | Akurasi meningkat signifikan (>94%) dibanding model tunggal      |
| 2  | (You et al., 2025)               | Machine Learning Approaches for Predicting U.S. Students' Scientific Literacy: An Analysis of Key Factors Across Performance Levels and Socioeconomic Statuses | Prediksi literasi sains                   | Random Forest & XGBoost dibanding regresi tradisional             | ML lebih akurat; faktor utama: grade, enjoyment, self-perception |
| 3  | (Kaur et al., 2025)              | AI-Powered and Mobile-Integrated Assessment Models Using Random Forest: Redefining Examination and Grading                                                     | Penilaian otomatis berbasis AI            | Random Forest dengan integrasi mobile learning                    | Sistem mampu memberikan penilaian real-time dan lebih adaptif    |
| 4  | (Javed et al., 2025)             | An Intelligent Model for Parametric Cognitive Assessment of E-Learning-Based Students                                                                          | Penilaian kognitif berbasis teks          | NLP + RNN untuk klasifikasi level Bloom's Taxonomy                | Akurasi klasifikasi kognitif mencapai >90%                       |
| 5  | (Hariharasitaraman et al., 2025) | Interpretable and ethical learning assessment transformer (IELAT): an                                                                                          | Assessment berbasis AI yang interpretable | Transformer + SHAP (explainable AI)                               | Akurasi sangat tinggi (~99%) dengan transparansi model           |

| No | Penulis, Tahun          | Judul Artikel                                                                                                                            | Fokus Penelitian              | Metode                                             | Temuan                                                 |
|----|-------------------------|------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------|----------------------------------------------------|--------------------------------------------------------|
|    |                         | explainable transformer model for personalized student assessments                                                                       |                               |                                                    |                                                        |
| 6  | (Xuan Lam et al., 2024) | Enhancing Educational Evaluation through Predictive Student Assessment Modeling                                                          | Prediksi nilai akademik       | Perbandingan ML (RF, SVM, NB, KNN) + Stacking      | Stacking menghasilkan prediksi paling akurat           |
| 7  | (Tin Tin et al., 2024)  | Educational Big Data Mining : Comparison of Multiple Machine Learning Algorithms in Predictive Modelling of Student Academic Performance | Prediksi performa skala besar | Gradient Boosting, RF, SVM pada dataset PISA besar | Gradient Boosting paling unggul (akurasi ~74%)         |
| 8  | (Çınar et al., 2020)    | Machine Learning Algorithm for Grading Open-Ended Physics Questions in Turkish                                                           | Penilaian fisika otomatis     | SVM, KNN, AdaBoost untuk short-answer grading      | AdaBoost terbaik; penilaian otomatis akurat & objektif |
| 9  | (Öz et al., 2024)       | Stacking: An Ensemble Learning Approach to Predict Student Performance in PISA 2022                                                      | Prediksi performa global      | Stacking vs boosting & blending                    | Stacking memberikan error terendah (MAE, MSE)          |

| No | Penulis, Tahun            | Judul Artikel                                                                                                                                          | Fokus Penelitian                                                                            | Metode                                                                              | Temuan                                                                                                                                                                                                |
|----|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 10 | (Thavasi & Revathi, 2025) | A personalized machine learning-based system to evaluate students' skillset and analyze the gap between academia and industry for engineering students | Evaluasi skill, prediksi karier, dan identifikasi gap kompetensi berbasis machine learning. | Perbandingan ML (NB, KNN, SVM, ANN) untuk prediksi job role berbasis data akademik. | ANN memberikan performa terbaik akurasi prediksi mencapai 93,33% precision mencapai 98% sistem mampu mengidentifikasi kekuatan & kelemahan skill mahasiswa memberikan rekomendasi pengembangan karier |

### ***Tren Penggunaan Machine Learning dalam Penilaian Pembelajaran***

Berdasarkan hasil analisis bibliometrik, terlihat bahwa penelitian terkait penerapan machine learning dalam sistem penilaian pembelajaran didominasi oleh topik prediksi, klasifikasi, dan evaluasi hasil belajar. Visualisasi jaringan kata kunci juga menunjukkan bahwa machine learning menjadi pusat utama penelitian dengan keterkaitan yang kuat terhadap student performance dan prediction. Temuan ini selanjutnya dianalisis lebih mendalam melalui kajian terhadap artikel terpilih.

Hasil kajian menunjukkan bahwa pendekatan machine learning menjadi metode dominan dalam pengembangan sistem penilaian pembelajaran. Berbagai algoritma seperti Random Forest, Support Vector Machine, serta metode ensemble banyak digunakan untuk meningkatkan akurasi prediksi. Penggunaan metode ensemble terbukti mampu menggabungkan keunggulan beberapa model sehingga menghasilkan performa yang lebih stabil dan akurat.

Hal ini didukung oleh penelitian Prediction and Classification of Student Performance Through an Ensemble Machine Learning Approach (Cheng, 2026) yang menunjukkan bahwa pendekatan ensemble mampu meningkatkan akurasi prediksi secara signifikan hingga lebih dari 94% ("the constructed model... achieved 94.42% accuracy"). Temuan ini menegaskan bahwa pendekatan berbasis kombinasi model lebih unggul dibandingkan model tunggal dalam konteks penilaian pembelajaran.

### ***Peran Jaringan Syaraf Tiruan dalam Penilaian***

Di antara berbagai pendekatan machine learning, jaringan syaraf tiruan (JST) memiliki keunggulan dalam memodelkan hubungan non-linear dan mengenali pola kompleks pada data pembelajaran. JST banyak diimplementasikan dalam bentuk Artificial Neural Network (ANN), Recurrent Neural Network (RNN), serta model berbasis transformer.

Penelitian An Intelligent Model for Parametric Cognitive Assessment of E-Learning-Based Students (Javed et al., 2025) menunjukkan bahwa model berbasis RNN mampu mencapai akurasi klasifikasi hingga 98% dalam mengidentifikasi level kognitif siswa ("the cognitive-based questions pool classification achieved by RNN results in 98% accuracy"). Selain itu, integrasi JST dengan metode lain seperti Random Forest juga mampu meningkatkan akurasi evaluasi hingga lebih dari 92% ("student assessments... increased to an accuracy ratio of 92.16%").

Perkembangan lebih lanjut terlihat pada penggunaan model transformer dalam sistem penilaian. Penelitian Interpretable and ethical learning assessment transformer (IELAT): an explainable transformer model for personalized student assessments menunjukkan bahwa model ini mampu mencapai akurasi sangat tinggi sekaligus memberikan transparansi dalam proses penilaian ("achieving 99.81% accuracy... ensuring transparency and trust"). Transformer merupakan pengembangan arsitektur jaringan syaraf tiruan berbasis mekanisme self-attention yang mampu memproses data

sekuensial secara paralel, sehingga lebih efisien dibandingkan model sebelumnya seperti Recurrent Neural Network (RNN) dan Long Short-Term Memory (LSTM) dalam menangani dependensi jangka panjang (Vaswani, 2017). Hal ini menunjukkan bahwa JST tidak hanya unggul dalam akurasi, tetapi juga berkembang dalam aspek interpretabilitas.

#### ***Fokus Penelitian: Prediksi Hasil Belajar***

Sebagian besar penelitian yang dianalisis masih berfokus pada prediksi hasil belajar siswa, seperti nilai akademik dan tingkat literasi sains. Pendekatan ini memanfaatkan data historis siswa untuk membangun model prediktif yang akurat.

Sebagai contoh, penelitian Machine Learning Approaches for Predicting U.S. Students' Scientific Literacy: An Analysis of Key Factors Across Performance Levels and Socioeconomic Statuses (You et al., 2025) menunjukkan bahwa faktor seperti tingkat kelas, minat belajar, dan persepsi diri menjadi prediktor utama dalam menentukan literasi sains siswa ("grade, enjoyment, and self-perceived scientific literacy were the most important predictors"). Hal ini menunjukkan bahwa machine learning efektif dalam mengidentifikasi faktor-faktor yang mempengaruhi hasil belajar, namun masih terbatas pada aspek output.

#### ***Keterbatasan Penilaian Berbasis AI***

Meskipun teknologi AI menunjukkan potensi besar dalam meningkatkan kualitas penilaian, terdapat beberapa keterbatasan yang masih ditemukan. Sebagian besar penelitian masih menggunakan data berbasis nilai atau jawaban akhir siswa, sehingga belum mampu merepresentasikan proses berpikir secara menyeluruh.

Penelitian Machine Learning Algorithm for Grading Open-Ended Physics Questions in Turkish (Çınar et al., 2020) menunjukkan bahwa algoritma machine learning mampu melakukan penilaian otomatis terhadap jawaban terbuka dengan tingkat akurasi yang tinggi ("AdaBoost.M1 technique had the best performance... accuracy, precision and F1-score"), namun masih berfokus pada hasil akhir jawaban, bukan proses penyelesaiannya.

#### ***Kesenjangan Penelitian: Integrasi dengan Virtual Laboratory***

Temuan ini sejalan dengan hasil analisis bibliometrik yang menunjukkan bahwa kata kunci yang berkaitan dengan *virtual laboratory* dan pembelajaran fisika masih belum dominan dalam jaringan penelitian. Di sisi lain, *virtual laboratory* sebagai lingkungan pembelajaran digital memiliki potensi besar dalam menyediakan data proses belajar yang lebih kaya, seperti interaksi eksperimen, waktu pengerjaan, dan langkah penyelesaian masalah. Namun, hasil kajian menunjukkan bahwa integrasi antara JST dan *virtual laboratory* dalam sistem penilaian pembelajaran masih sangat terbatas. Sebagian besar penelitian mengkaji sistem penilaian berbasis AI dan penggunaan *virtual laboratory* secara terpisah, sehingga belum memanfaatkan potensi data proses belajar secara optimal. Padahal, data yang dihasilkan dari aktivitas eksperimen virtual sangat relevan untuk dianalisis menggunakan JST guna mengembangkan sistem penilaian yang lebih komprehensif.

#### ***Implikasi dan Arah Penelitian Selanjutnya***

Berdasarkan hasil analisis, integrasi jaringan syaraf tiruan dalam sistem penilaian berbasis *virtual laboratory* memiliki potensi besar untuk dikembangkan. Pendekatan ini memungkinkan analisis data pembelajaran yang lebih kompleks dan dinamis, sehingga dapat menghasilkan sistem penilaian yang tidak hanya berfokus pada hasil akhir, tetapi juga proses belajar siswa. Dengan demikian, pengembangan sistem penilaian berbasis JST dalam lingkungan *virtual laboratory* dapat menjadi solusi untuk meningkatkan kualitas evaluasi pembelajaran fisika yang lebih adaptif, objektif, dan berbasis data.

## **IV. KESIMPULAN**

Berdasarkan hasil Systematic Literature Review yang dikombinasikan dengan analisis bibliometrik, dapat disimpulkan bahwa penerapan machine learning, khususnya jaringan syaraf tiruan (JST), memiliki potensi yang signifikan dalam pengembangan sistem penilaian pembelajaran. Tren penelitian menunjukkan peningkatan penggunaan pendekatan berbasis data untuk memprediksi dan

mengevaluasi performa belajar siswa. Namun, sebagian besar penelitian masih berfokus pada hasil akhir pembelajaran dan belum mengakomodasi analisis proses belajar secara menyeluruh.

Hasil analisis juga menunjukkan bahwa integrasi antara JST dan lingkungan pembelajaran berbasis *virtual laboratory* masih sangat terbatas, meskipun lingkungan tersebut menghasilkan data interaksi yang kaya dan berpotensi untuk dianalisis secara lebih mendalam. Oleh karena itu, penelitian ini menegaskan bahwa pengembangan sistem penilaian berbasis JST yang terintegrasi dengan *virtual laboratory* merupakan peluang penting untuk menghasilkan evaluasi pembelajaran yang lebih komprehensif, adaptif, dan berbasis data, khususnya dalam pembelajaran fisika.

## DAFTAR PUSTAKA

- Cheng, W. (2026). Prediction and Classification of Student Performance Through an Ensemble Machine Learning Approach. *Journal of The Institution of Engineers (India): Series B*, 107(B), 297–307. <https://doi.org/https://doi.org/10.1007/s40031-025-01240-1>
- Çınar, A., Ince, E., Gezer, M., & Yılmaz, Ö. (2020). Machine learning algorithm for grading open-ended physics questions in Turkish. *Education and Information Technologies*, 25, 3821–3844. <https://doi.org/https://doi.org/10.1007/s10639-020-10128-0>
- Darman, D. R., Suhandi, A., Kaniawati, I., Samsudin, A., & Wibowo, F. C. (2024). Virtual laboratory in physics education: A systematic review. *ELECTRONIC PHYSICS INFORMATICS INTERNATIONAL CONFERENCE (EPIIC)*. <https://doi.org/https://doi.org/10.1063/5.0210640>
- Handayani, K. S., Lumbanbatu, K., & Simanjuntak, M. (2021). Jaringan Syaraf Tiruan dengan Algoritma Backpropagation untuk Memprediksi Nilai Ujian Kompetensi Siswa (Studi Kasus Smks Jabal Rahmah Stabat). *Jurnal Ilmiah Abdi Ilmu*, 14(1), 73–80.
- Hariharasitaraman, S., Jayavel, A., Gnanasekaran, S., Sohail, S. S., & Madsen, D. Ø. (2025). Interpretable and ethical learning assessment transformer (IELAT): an explainable transformer model for personalized student assessments. *Cogent Education*, 12(1). <https://doi.org/10.1080/2331186X.2025.2485508>
- Javed, M. S., Aslam, M., & Khurshid, S. K. (2025). An Intelligent Model for Parametric Cognitive Assessment of E-Learning-Based Students. *Information*, 16(2), 1–35. <https://doi.org/https://doi.org/10.3390/info16020093>
- Kaur, S., Goel, D., Ramachandaran, D. S., Devi, S., Yadav, B., & Goyal, A. (2025). AI-Powered and Mobile-Integrated Assessment Models Using Random Forest: Redefining Examinations and Grading. *International Journal of Interactive Mobile Technologies*, 19(14), 57–69.
- Lau, E. T., Sun, L., & Yang, Q. (2019). Modelling, prediction and classification of student academic performance using artificial neural networks. *SN Applied Sciences*, 1(9), 1–10. <https://doi.org/10.1007/s42452-019-0884-7>
- Musahir, M. (2022). Penggunaan media pembelajaran fisika dengan virtual laboratory berbasis “PhET Interactive Simulations” untuk meningkatkan kemampuan penguasaan konsep kognitif rangkaian resistor. *Compton: Jurnal Ilmiah Pendidikan Fisika*, 9(2), 122–130. <https://doi.org/10.30738/cjipf.v9i2.13664>
- N. Kapita, S., Mahdi, S., & Tempola, F. (2020). Penilaian Pengetahuan Siswa Dengan Jaringan Syaraf Tiruan Algoritma Perceptron. *Techno: Jurnal Penelitian*, 9(1), 372. <https://doi.org/10.33387/tjp.v9i1.1712>
- Nurfadila, Z., & Arlita, W. (2025). Penerapan Laboratorium Virtual dalam Pembelajaran Fisika di Era Digital. *Jurnal Inovasi Pendidikan Dan Pengajaran (JIPP)*, 4(2), 33–36. <https://doi.org/10.31571/jipp.v4i2.9317>
- Öz, E., Bulut, O., Cellat, Z. F., & Yürekli, H. (2024). Stacking: An Ensemble Learning Approach to Predict Student Performance in PISA 2022. *Education and Information Technologies*, 30, 7753–7779. <https://doi.org/https://doi.org/10.1007/s10639-024-13110-2>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Thebmj*, 372. <https://doi.org/10.1136/bmj.n71>
- Santoso, P. H., Santosa, H. S., Istiyono, E., Haryanto, H., & Retnawati, H. (2023). Predicting Physics

- Students' Achievement Using In-Class Assessment Data: A Comparison of Two Machine Learning Models. *Physics Education Research Journal*, 5(2), 97–106. <https://doi.org/10.21580/perj.2023.5.2.14217>
- Tedja, B., Al Musadieq, M., Kusumawati, A., & Yulianto, E. (2024). Systematic literature review using PRISMA: exploring the influence of service quality and perceived value on satisfaction and intention to continue relationship. *Future Business Journal*, 10(1), 1–9. <https://doi.org/10.1186/s43093-024-00326-4>
- Thavasi, S., & Revathi, T. (2025). A personalized machine learning-based system to evaluate students' skillset and analyze the gap between academia and industry for engineering students. *Kybernetes: The International Journal of Cybernetics, Systems, and Management Sciences*, 54(9), 4850–4865. <https://doi.org/https://doi.org/10.1108/K-01-2024-0053>
- Tin Tin, T., Hock, L. S., & Ikumapayi, O. M. (2024). Educational Big Data Mining : Comparison of Multiple Machine Learning Algorithms in Predictive Modelling of Student Academic Performance. (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, 15(6), 633–645. <https://doi.org/https://dx.doi.org/10.14569/IJACSA.2024.0150664>
- Vaswani, A. (2017). Attention Is All You Need. *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Nips.
- Wang, Y., Zhang, L., & Pang, M. (2024). Virtual experiments in physics education: a systematic literature review. *Research in Science & Technological Education*, 43(2), 633–655. <https://doi.org/https://doi.org/10.1080/02635143.2024.2327995>
- Xuan Lam, P., Quoc Hung Mai, P., Hung Nguyen, Q., & Pham, T. (2024). Enhancing educational evaluation through predictive student assessment modeling. *Computers and Education: Artificial Intelligence*, 6. <https://doi.org/10.1016/j.caeai.2024.100244>
- Yang, K. (2024). Predicting Student Performance Using Artificial Neural Network Analysis. *Journal of Arts, Society, and Education*, 42(5), 516–517. <https://doi.org/10.1111/j.1365-2923.2008.03052.x>
- You, H., Hong, M., Zhu, L., & Zhenhan, F. (2025). Machine Learning Approaches for Predicting U.S. Students' Scientific Literacy: An Analysis of Key Factors Across Performance Levels and Socioeconomic Statuses. *International Journal of Science and Mathematics Education*, 23, 2755–2783. <https://doi.org/10.1007/s10763-025-10545-y>
- Zola, F., Nurcahyo, G. W., & Santony, J. (2018). Jaringan Syaraf Tiruan Menggunakan Algoritma Backpropagation Untuk Memprediksi Prestasi Siswa. *Jurnal Teknologi Dan Open Source*, 1(1), 58–72.