



Analisis Literatur Terhadap Penerapan Machine Learning dalam Automasi Manajemen Basis Data

A. Akhmad Sultan^a, Kurnia Wahyu Prima^{b*}, Ayu Hasnining^c, Nur Azizah^d

^{a,b,c,d}Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Makassar, Makassar (90222), Indonesia

ARTICLE INFO

Accepted by the Editor: 10 June 2025
Final Revision: 12 June 2025
Published Online: 13 June 2025

KEYWORDS

automasi *database*; *machine learning*;
manajemen indeks; optimasi *query*;
tuning konfigurasi

CORRESPONDENCE*

E-mail:
kurnia.wahyu.prima@unm.ac.id

ABSTRACT

Perkembangan teknologi seperti *Internet of Things* dan *cloud computing*, membawa dampak pada pertumbuhan data yang sangat cepat dan menuntut adanya inovasi dalam pengelolaan *database*. Masalah klasik manual dan tradisional, upaya mengelola beban kerja serta dinamisnya pola data yang *shifting*, menjadi tantangan utama. Untuk itu, pendekatan cerdas yang dapat mengotomatisasi manajemen *database* sangat diperlukan. Artikel ini bertujuan untuk mengeksplorasi secara sistematis penggunaan *Machine Learning* (ML) dalam automasi manajemen *database*. Pada penelitian ini, penulis menerapkan *Systematic Literature Review* (SLR) pada beberapa artikel dalam lima tahun terakhir yang relevan. Dengan menjadikan makalah tersebut sebagai fokus utama, hasil *review* menyatakan bahwa dalam empat domain utama, yakni optimasi *query*, deteksi anomali, *Auto-ML*, dan analisis prediktif, ML terbukti sangat mendukung. Implementasi ML terbukti mengoptimalkan efisiensi dan responsif sistem, ancaman di dalam sistem terdeteksi secara *real-time*, dan pengambilan keputusan yang bergantung pada *database* dipercepat. Sementara demikian, perlu dicermati, meskipun sangat bermanfaat, sulitnya interpretabilitas serta kebutuhan komputasi tinggi menjadi tantangan untuk implementasi lebih luas ke depan.

1. Introduction

Sekarang lebih dari sebelumnya, mengelola basis data secara otomatis telah menjadi aspek krusial dari infrastruktur sistem informasi digital yang canggih. Rentang yang dibawa oleh *World Wide Web* bersamaan dengan sistem media sosial dan komputasi awan menuntut adanya sistem DBMS yang tidak hanya fungsional, tetapi juga efektif dalam hal penyimpanan dan pemrosesan big data. Sistem semacam itu juga harus beradaptasi secara proaktif terhadap perubahan pola beban kerja dari dataset yang berubah dinamis.

Semua bentuk perangkat lunak yang dioptimalkan dan terstruktur dapat disebut sebagai sistem DBMS [1]. Mengotomatisasi pengelolaan basis data secara langsung merujuk pada proses pengelolaan datanya. Di sini, pembelajaran mesin jelas merupakan teknologi terpenting untuk otomatisasi cerdas karena kemampuannya untuk

menganalisis data dan membuat keputusan di masa depan tanpa intervensi manusia memungkinkan pembelajaran dari data yang sudah tersedia [2].

Penerapan teknologi *Machine Learning* ke dalam sistem pengelolaan basis data telah maju pesat dalam lima tahun yang lalu. Menurut Gartner, sekitar 45% perusahaan besar di dunia telah menerapkan teknologi ML pada sistem *database* mereka, dan angka ini akan meloncat menjadi 70% di 2026 karena fitur penyetalan otomatis, deteksi anomali, dan optimasi kinerja. Juga, menurut Kraska et al, Pengantar Isi struktur indeks pembelajaran. Ini menunjukkan bahwa algoritma ML mampu menggantikan indeks tradisional dengan keefisienan yang lebih tinggi [3].

Di sisi lain, kontrol manual atau semi otomatis pada pengelolaan basis data masih memiliki banyak masalah, terutama dengan tingginya level kompleksitas konfigurasi sistem *tuning*, abnormalitas

dalam melakukan akses data, dan masalah dalam mengenali limit adaptasi terhadap perubahan pola *cosine*. Masalah ini menambah waktu pengambilan keputusan dan menambah resiko keamanan data. Dalam bisnis yang sangat kompetitif, keterlambatan dalam data *capture* dan analisis dapat mempengaruhi kinerja organisasi secara keseluruhan.

Pembelajaran dapat dibagi menjadi tiga komponen utama: pembelajaran terawasi, pembelajaran tidak terawasi, dan intensitas yang dipelajari [2]. Pembelajaran terawasi berusaha untuk membuat prediksi menggunakan data historis, seperti peramalan beban kerja atau estimasi *query* timeline. Kompromi tidak terawasi dapat berguna untuk pembentukan cluster dan deteksi anomali outlier yang tidak dilabeli, dan sangat membantu dalam mengenali pola akses data yang anomali. Sistem pembelajaran memungkinkan sistem untuk belajar dari interaksi dengan lingkungan melalui mekanisme hadiah dan umpan balik, yang digunakan dalam proses pengambilan keputusan yang canggih. Misalnya, kasus antrean umum dan alokasi sumber daya dinamis.

Ketiga pendekatan ini telah dipelajari dan diterapkan secara luas pada beberapa tugas penting dalam manajemen basis data seperti optimasi *query*, penyetelan konfigurasi, dan analisis beban kerja [4], [5]. Misalnya, konsolidasi pembelajaran mendalam telah diterapkan untuk mengoptimalkan rencana implementasi *Query* untuk membuat pesanan umum dalam waktu sesingkat mungkin [4]. Selain itu, tautan yang dipelajari telah memberikan pendekatan yang dapat diprediksi untuk mendefinisikan strategi penggabungan tabel dan menyesuaikan konfigurasi untuk basis data konfigurasi *autoclass* sistem untuk memungkinkan penyetelan basis data secara *real-time*.

Tidak hanya terbatas pada proses internal, penggunaan *ML* eksternal dan internal sebagai komponen dalam sistem basis data otonom juga telah banyak dibahas dalam dokumen akhir [10]. Sistem ini memiliki kemampuan untuk melakukan fungsi-fungsi penting seperti pemantauan kinerja, menyesuaikan konfigurasi, untuk secara otomatis pulih setelah kegagalan, tanpa intervensi manusia.

Namun, penerapan pembelajaran otomatis dalam konteks *SGD* juga memiliki tantangan besar. Salah satunya adalah permintaan untuk sumber daya komputer yang tinggi, terutama pelatihan tentang model besar dan kompleks. Tantangan lain adalah kemampuan untuk menafsirkan atau kemampuan

untuk menjelaskan hasil model *ML*, yang sangat penting dalam konteks audit sistem dan mematuhi peraturan. Integrasi model *ML* dengan infrastruktur basis data yang ada juga tidak berfungsi dengan baik, dengan kompatibilitas arsitektur dan sumber daya yang terbatas di banyak organisasi.

Selain itu, sebagian besar studi saat ini masih fokus pada solusi teknis dan puing-puing tertentu. Beberapa penelitian telah menyajikan pendekatan yang komprehensif dan sistematis untuk pengembangan *SGBD* berdasarkan *ML* dari awal hingga akhir. Misalnya, bagaimana pipa pembelajaran mesin diintegrasikan ke dalam pengelolaan indeks, perencanaan *Query* dan manajemen beban kerja dalam sistem terintegrasi [7], [8].

Oleh karena itu, artikel ini bertujuan untuk memberikan penilaian dokumen yang sistematis dan lengkap dari otomatisasi manajemen basis data dengan menggunakan pembelajaran otomatis. Tujuan artikel termasuk aplikasi *ML* yang berbeda seperti optimasi *query*, manajemen indeks, penyesuaian konfigurasi otomatis, serta deteksi abnormal dan analisis prediktif. Dengan pengetahuan tentang penelitian terbaru, artikel ini akan menjadi referensi awal yang bermanfaat bagi para peneliti dan siswa dalam mengembangkan sistem manajemen basis data yang cerdas, lebih efektif dan mudah beradaptasi di masa depan. Dengan pemahaman penuh tentang konteks teknologi ini, para pihak yang terlibat dalam sistem informasi dapat membuat keputusan strategis yang tepat di era digital *database*.

2. Methodology

Penelitian ini menggunakan pendekatan *Systematic Literature Review* (*SLR*). *SLR* adalah metode yang sangat terstruktur dan ketat untuk meninjau literatur, yang menghasilkan ringkasan bukti yang andal dan komprehensif tentang suatu topik. Metode *SLR* ini digunakan untuk memperoleh pemahaman yang komprehensif mengenai penerapan *Machine Learning* (*ML*) dalam automasi manajemen *database*.

Metode ini dipilih karena memiliki keunggulan dalam menyusun sintesis pengetahuan dari beragam studi yang telah dipublikasikan sebelumnya secara sistematis, transparan, dan bisa direplikasi. Tujuan utama dari penggunaan *SLR* yaitu untuk mengidentifikasi tren, aplikasi utama, serta tantangan implementatif yang telah dibahas dalam literatur ilmiah lima tahun terakhir, khususnya dalam konteks

sistem manajemen *database* yang berbasis kecerdasan buatan. Urutan dari metode yang digunakan diwakilkan oleh Gambar 1.

Penentuan Topik Masalah

Pada tahap ini, peneliti memperjelas dan membatasi ruang lingkup topik dengan fokus pada Manajemen basis data secara manual (seperti *tuning*, penjadwalan *backup*, dan optimasi *query*) memakan waktu, rentan terhadap kesalahan, dan membutuhkan keahlian tinggi. Penerapan teknik *machine learning* untuk mengotomatisasi tugas-tugas tersebut. Dengan tujuan analisis untuk mengidentifikasi, mengevaluasi, dan memetakan bagaimana *machine learning* telah digunakan dalam automasi manajemen basis data berdasarkan literatur ilmiah yang sudah ada.



Gambar 1. Tahapan Penelitian

Penetapan Kata Kunci (*Keyword*)

Berdasarkan topik yang dipilih, penentuan kata kunci dilakukan untuk mencari artikel ilmiah. Kata kunci relevan yang digunakan sebagai kata kunci konsep utama adalah *Machine Learning*, *Database Automation*, *Database Management*. Sedangkan istilah terkait yang digunakan seperti *Autonomous*

Database, *Self-Driving Database*, *AI in Databases*, *Database Administration*, serta teknik spesifik kata kunci yang digunakan seperti *Reinforcement Learning for tuning*, *Query Optimization with ML*, *Automated Indexing*.

Pencarian Artikel dari Database Ilmiah

Dengan menggunakan kata kunci yang telah ditetapkan, peneliti melakukan pencarian literatur secara sistematis pada berbagai *database* ilmiah terkemuka seperti: *IEEE Xplore*, *ACM Digital Library*, *SpringerLink*, *ScienceDirect* dan *Google Scholar* dengan tujuan untuk mengumpulkan sebanyak mungkin artikel ilmiah berupa jurnal, makalah konferensi yang relevan dan diterbitkan antara tahun 2019 hingga 2024.

Seleksi Artikel Berdasarkan Abstrak dan Relevansi

Kumpulan artikel yang didapat dari tahap sebelumnya masih sangat luas. Oleh karena itu, dilakukan proses penyaringan yang ketat dengan membaca judul dan abstrak dari setiap artikel untuk secara cepat menentukan apakah artikel tersebut benar-benar membahas penerapan *machine learning* untuk automasi basis data.

Menetapkan kriteria inklusi atau eksklusi yang tidak relevan, seperti yang hanya membahas *machine learning* secara umum atau manajemen basis data tanpa automasi. Artikel seperti itu akan disingkirkan dan hanya artikel yang paling relevan yang akan disimpan untuk dianalisis lebih lanjut.

Analisis Mendalam Artikel

Artikel-artikel yang lolos seleksi akan dibaca secara penuh dan mendalam. Pada tahap ini, peneliti mengekstrak informasi kunci dari setiap artikel yang mencakup metode seperti teknik atau algoritma *machine learning* yang digunakan seperti *Neural Networks*, *Reinforcement Learning*, atau *Decision Trees* serta bagaimana arsitektur sistem yang diusulkan. Informasi hasil meliputi pencapaian metode tersebut apakah berhasil meningkatkan performa, mengurangi latensi, atau menghemat biaya serta seberapa signifikan peningkatannya.

Sintesis Temuan dan Penarikan Kesimpulan

Tahapan ini adalah tahap akhir analisis di mana peneliti tidak hanya merangkum tetapi juga menghubungkan informasi dari semua artikel yang telah dianalisis. Pada tahap sintesis temuan, peneliti mengidentifikasi pola, tren, dan celah dalam literatur.

Contohnya adalah temuan bahwa teknik reinforcement learning paling banyak digunakan untuk query tuning, atau bahwa masih sedikit penelitian yang fokus pada automasi keamanan basis data menggunakan *machine learning*.

Peneliti melakukan penarikan kesimpulan berdasarkan sintesis tersebut. Kesimpulan ini merangkum kondisi terkini penerapan machine learning di bidang yang dikaji, menyoroti metode yang paling efektif, serta memberikan rekomendasi untuk arah penelitian di masa depan.

3. Results and Discussion

Berdasarkan tahapan pada metode penelitian yang digunakan dalam melakukan analisis beberapa artikel ilmiah, ditemukan bahwa penerapan *Machine Learning* dalam automasi manajemen *database* mengarah pada empat kategori utama.

Optimasi Query dan Pengelolaan Sumber Daya

Salah satu area penerapan *Machine Learning* (ML) yang paling dominan dalam automasi manajemen *database* yaitu optimasi *query* dan pengelolaan sumber daya sistem. Proses eksekusi *query* dalam sistem manajemen *database* (DBMS) melibatkan pemilihan strategi rencana eksekusi yang optimal untuk meminimalkan waktu dan konsumsi sumber daya. Strategi ini secara tradisional bergantung pada heuristik dan statistik yang dibangun secara manual, yang seringkali tidak mampu mengikuti dinamika beban kerja modern yang sangat berubah-ubah dan tidak terduga. Dalam konteks inilah ML berperan, dengan kemampuannya mempelajari pola dari histori eksekusi *query* dan memprediksi performa dari berbagai rencana eksekusi yang mungkin.

Pendekatan yang banyak digunakan dalam kategori ini melibatkan *supervised learning*, khususnya algoritma misalnya *decision tree*, *random forest*, dan *gradient boosting*, yang dilatih menggunakan fitur-fitur misalnya jumlah baris, jenis *join*, indeks yang digunakan, dan estimasi biaya dari *query optimizer*. Model-model ini digunakan untuk memprediksi waktu eksekusi atau estimasi performa, dan hasil prediksinya digunakan untuk memilih rencana eksekusi terbaik secara otomatis. Di samping itu, *deep reinforcement learning* telah diperkenalkan dalam studi-studi terkini untuk enumerasi *join order*, di mana agen pembelajar mencoba berbagai kombinasi *join* dan menbisakan reward berdasarkan efisiensi eksekusi aktual.

Hasil dari 12 artikel yang dianalisis menunjukkan bahwa penerapan ML secara signifikan bisa meningkatkan kecepatan eksekusi *query* serta efisiensi pemanfaatan CPU dan memori. Konsep *learned index structures* yang mampu menggantikan indeks *B-tree* tradisional dengan model *neural network*, menghasilkan efisiensi akses data yang lebih tinggi di beberapa skenario. Studi lain menunjukkan bahwa sistem berbasis ML mampu menyesuaikan strategi eksekusi *query* terhadap pola data dan *workload* secara *real-time*, tanpa perlu intervensi manual berulang.

Walaupun hasilnya menjanjikan, implementasi solusi berbasis ML ini memerlukan pelatihan model yang cukup mahal dari segi komputasi dan data pelatihan yang representatif. Selain itu, tantangan interpretabilitas model juga perlu diperhatikan, terutama dalam konteks *debugging* dan audit. Namun secara umum, integrasi ML dalam optimasi *query* dan pengelolaan sumber daya terbukti memberikan kontribusi besar terhadap peningkatan performa dan skalabilitas sistem *database* modern, khususnya dalam lingkungan *cloud* dan *big data* yang sangat dinamis.

Deteksi Anomali dan Keamanan Data

Dalam konteks manajemen *database*, keamanan menjadi aspek krusial yang tidak hanya mencakup perlindungan terhadap serangan *eksternal*, tetapi juga kemampuan sistem dalam mendeteksi aktivitas tidak normal yang bisa mengindikasikan penyalahgunaan, *data breach*, atau bahkan korupsi data. Dalam hal ini, *Machine Learning* memberikan pendekatan yang adaptif dan proaktif melalui kemampuan deteksi anomali berdasarkan pembelajaran pola akses data historis. Teknik ML memungkinkan sistem mengenali penyimpangan dari pola normal tanpa perlu eksplisit diberi tahu misalnya apa bentuk anomali tersebut.

Pendekatan deteksi anomali umumnya menggunakan metode *unsupervised learning* misalnya *clustering* (misalnya *DBSCAN*, *K-means*) dan model statistik misalnya *Gaussian Mixture Model* atau *Isolation Forest*. Beberapa studi juga menggunakan *autoencoder* atau model pembelajaran representasi lain berbasis *deep learning* untuk membangun representasi kompak dari aktivitas normal, sehingga penyimpangan terhadap representasi ini bisa diidentifikasi sebagai anomali. Integrasi *autoencoder* dalam log audit *database* mampu meningkatkan sensitivitas deteksi terhadap serangan *SQL injection* maupun *privilege escalation*.

Dari delapan studi yang dianalisis dalam kategori ini, seluruhnya menunjukkan bahwa sistem yang dilengkapi dengan komponen deteksi anomali berbasis ML memiliki kemampuan untuk mendeteksi pola akses tidak biasa secara *real-time* dan memberikan peringatan dini terhadap kemungkinan insiden keamanan. Studi-studi ini juga menyebutkan bahwa penggunaan ML mampu mengurangi tingkat *false positive* secara signifikan dibanding pendekatan berbasis aturan statis. Selain itu, sistem bisa belajar dan beradaptasi terhadap pola akses pengguna yang berubah seiring waktu, sesuatu yang sulit dicapai oleh pendekatan tradisional.

Namun, tantangan utama dalam implementasi sistem deteksi anomali berbasis ML yaitu kebutuhan akan data pelatihan yang memadai dan representatif terhadap kondisi normal. Dalam beberapa kasus, data pelatihan yang bias atau tidak lengkap bisa menyebabkan kegagalan deteksi atau sebaliknya, menghasilkan terlalu banyak *false alarm*. Selain itu, interpretabilitas hasil deteksi juga masih menjadi isu, terutama dalam organisasi yang membutuhkan dokumentasi yang kuat untuk alasan audit atau kepatuhan regulasi.

Secara keseluruhan, deteksi anomali berbasis *Machine Learning* telah menjadi pilar penting dalam transformasi keamanan *database*. Pendekatan ini tidak hanya meningkatkan keamanan sistem secara teknis, tetapi juga membuka peluang untuk pengembangan arsitektur *database* yang lebih tanggap terhadap risiko dan lebih resilien terhadap serangan atau gangguan operasional.

Automated Machine Learning (Auto-ML)

Automated Machine Learning (Auto-ML) merupakan pendekatan yang bertujuan untuk mengotomatiskan seluruh proses pengembangan model pembelajaran mesin, mulai dari pemilihan algoritma, ekstraksi fitur, hingga proses *tuning* parameter dan validasi model. Dalam konteks manajemen *database*, *Auto-ML* memberikan peluang besar untuk mengurangi ketergantungan terhadap keahlian teknis tinggi, sekaligus mempercepat penerapan solusi cerdas dalam pengelolaan data secara adaptif. Keunggulan utama *Auto-ML* terletak pada kemampuannya untuk menghasilkan model-model prediktif yang optimal tanpa intervensi manual secara mendalam, sehingga memungkinkan integrasi *Machine Learning* ke dalam sistem *database* yang bersifat otonom.

Auto-ML bisa diintegrasikan dengan sistem *database* untuk menyusun *pipeline* otomatis yang mampu mengidentifikasi konfigurasi optimal berdasarkan histori *workload* [6]. *Auto-ML* secara dinamis menyesuaikan parameter misalnya ukuran *buffer*, jumlah *thread*, dan strategi *indexing*, sehingga menghasilkan performa sistem yang lebih stabil dan responsif terhadap variasi beban kerja. Hal ini sangat relevan terutama dalam lingkungan *database cloud* yang menghadapi lonjakan beban tidak terduga dan variabilitas sumber daya.

Dalam praktiknya, *Auto-ML* bekerja dengan menguji kombinasi berbagai algoritma dan parameter menggunakan teknik pencarian misalnya *grid search*, *random search*, atau bahkan pendekatan berbasis *Bayesian optimization*. Model terbaik kemudian dipilih berdasarkan metrik kinerja tertentu, misalnya akurasi prediksi waktu eksekusi *query*, *throughput* sistem, atau efisiensi penggunaan sumber daya. Selain itu, beberapa platform *Auto-ML* modern sudah mendukung integrasi dengan *pipeline DevOps* dan sistem manajemen *database* misalnya *PostgreSQL* atau *MySQL* melalui API.

Dari enam artikel yang dianalisis, seluruhnya melaporkan bahwa penggunaan *Auto-ML* berhasil menurunkan waktu konfigurasi sistem, mengurangi beban kerja administrator, serta meningkatkan stabilitas performa sistem secara keseluruhan. Beberapa studi bahkan menunjukkan bahwa performa model yang dihasilkan *Auto-ML* sebanding atau melebihi model yang dibangun secara manual oleh pakar domain.

Namun demikian, implementasi *Auto-ML* dalam *database* masih menghadapi kendala, misalnya waktu pelatihan yang lama dan kebutuhan komputasi tinggi, terutama saat melakukan eksplorasi banyak kombinasi parameter. Selain itu, dalam lingkungan produksi, *Auto-ML* harus dijalankan secara efisien tanpa mengganggu kinerja operasional utama dari sistem *database*.

Secara keseluruhan, *Auto-ML* menjanjikan sebuah lompatan besar dalam automasi manajemen *database* dengan menggabungkan kecepatan, efisiensi, dan fleksibilitas dalam pengembangan model *Machine Learning*. Teknologi ini membuka jalan bagi terciptanya *autonomous database systems* yang benar-benar mampu beradaptasi dan belajar secara mandiri dari data dan lingkungan operasionalnya, sekaligus menurunkan hambatan adopsi teknologi cerdas dalam skala industri.

Analisis Prediktif dan Pengambilan Keputusan

Penerapan *Machine Learning* (ML) dalam manajemen *database* tidak hanya terbatas pada pengoptimalan teknis misalnya eksekusi *query* atau pengaturan konfigurasi, tetapi juga telah berkembang ke arah yang lebih strategis, yaitu mendukung analisis prediktif dan pengambilan keputusan otomatis. Analisis prediktif memungkinkan sistem untuk mengantisipasi kebutuhan di masa mendatang berdasarkan pola historis yang terbiasa dalam data operasional. Dalam konteks *database*, hal ini mencakup prediksi terhadap beban sistem, estimasi kapasitas penyimpanan yang akan dibutuhkan, hingga rekomendasi penyesuaian konfigurasi untuk mencegah degradasi performa.

Beberapa pendekatan prediktif yang umum digunakan antara lain yaitu *time series forecasting*, *regresi multivariat*, dan *recurrent neural networks* (RNN), terutama jenis *long short-term memory* (LSTM) yang unggul dalam memodelkan data sekuensial. Model-model ini dilatih untuk mengenali pola temporal dalam aktivitas sistem misalnya jumlah transaksi, waktu akses, dan pola *query*. Dengan kemampuan ini, sistem bisa secara otomatis menyesuaikan sumber daya atau mengirimkan peringatan dini sebelum terjadi *bottleneck* atau gangguan performa.

Integrasi model prediktif dalam manajemen *database* memberikan dampak signifikan terhadap efisiensi jangka panjang sistem. Misalnya, prediksi beban kerja memungkinkan penjadwalan ulang tugas-tugas berat ke waktu-waktu yang lebih ringan, sehingga menghindari konflik akses dan menjaga kestabilan sistem. Selain itu, prediksi kebutuhan penyimpanan atau pertumbuhan data juga membantu dalam perencanaan kapasitas, terutama untuk organisasi yang bergantung pada penyimpanan *cloud* berbasis langganan.

Dari empat artikel yang dikaji dalam kategori ini, semuanya menekankan pentingnya *predictive analytics* dalam meningkatkan responsivitas sistem terhadap perubahan lingkungan kerja. Rekomendasi konfigurasi berbasis prediksi juga mengurangi waktu yang dibutuhkan administrator dalam menganalisis *log* dan menentukan tindakan, karena sistem bisa memberikan saran berdasarkan evaluasi historis dan tren data.

Meski demikian, tantangan yang dihadapi mencakup akurasi model prediktif dalam lingkungan dengan tingkat entropi tinggi, serta integrasi hasil

prediksi ke dalam sistem pengambilan keputusan secara *real-time*. Dalam beberapa kasus, prediksi yang meleset bisa menyebabkan tindakan otomatis yang tidak tepat, sehingga dibutuhkan mekanisme validasi tambahan atau intervensi semi-otomatis.

Secara keseluruhan, analisis prediktif dalam sistem manajemen *database* menawarkan dimensi baru dalam automasi, dengan potensi besar untuk meningkatkan efisiensi, mengurangi risiko kegagalan sistem, dan mendukung keputusan yang lebih berdata. Kombinasi pendekatan ini dengan *Auto-ML* dan pembelajaran berkelanjutan akan memperkuat arah perkembangan sistem *database* menuju platform yang benar-benar otonom dan adaptif terhadap perubahan.

Tabel 1 menunjukkan rangkuman jumlah studi dan manfaat utama dari masing-masing kategori.

Tabel 1. Ringkasan Temuan Utama

No.	Kategori Aplikasi ML	Jumlah Studi	Manfaat Utama
1	Optimasi Query dan Sumber Daya	12	Pemilihan strategi eksekusi query secara otomatis, Alokasi sumber daya yang efisien, Percepatan proses query melalui supervised & deep reinforcement learning
2	Deteksi Anomali dan Keamanan	8	Deteksi akses mencurigakan secara real-time, Pencegahan risiko keamanan, Perlindungan terhadap integritas data
3	Automated Machine Learning (Auto-ML)	6	Otomatisasi pemilihan algoritma dan tuning parameter, Penyusunan pipeline ML, Mengurangi ketergantungan terhadap tenaga ahli

4	Analisis Prediktif dan Keputusan	4	Prediksi beban sistem dan kebutuhan kapasitas, Rekomendasi konfigurasi otomatis, Mendukung pengambilan keputusan berbasis data
---	----------------------------------	---	--

Temuan ini menunjukkan bahwa integrasi ML memberikan kontribusi signifikan terhadap efisiensi, keamanan, dan skalabilitas sistem *database* modern. Selain itu, pendekatan misalnya *Auto-ML* dan *reinforcement learning* membuka peluang baru untuk pengembangan sistem *database* otonom yang adaptif terhadap beban kerja yang dinamis. Namun, tantangan misalnya interpretabilitas model dan beban komputasi tetap menjadi kendala yang perlu ditangani secara khusus.

Secara keseluruhan, hasil temuan dari tinjauan literatur ini sangat relevan dengan fokus utama artikel, yaitu menelusuri tren automasi manajemen *database* menggunakan *Machine Learning*. Keempat kategori temuan utama memperlihatkan bagaimana ML berperan penting dalam mengotomatisasi tugas-tugas inti manajemen *database*, mulai dari optimasi *query* hingga pengambilan keputusan.

Tren penggunaan *Auto-ML* tidak hanya mengurangi kebutuhan akan tenaga ahli, tetapi juga mempercepat siklus pengembangan model yang adaptif terhadap data baru. Ini selaras dengan misi automasi, di mana sistem mampu beroperasi mandiri dalam kondisi dinamis. Sementara itu, pendekatan misalnya *deep reinforcement learning* untuk enumerasi urutan *join query* menunjukkan bahwa ML mampu mengatasi masalah yang sebelumnya membutuhkan optimasi kompleks berbasis heuristik. Hal ini merupakan bentuk nyata dari *transformasi* manajemen *database* menuju sistem yang lebih cerdas dan efisien.

Selain efisiensi, dimensi keamanan dan pengambilan keputusan otomatis juga menjadi bagian penting dari tren ini. Integrasi ML dalam pengawasan aktivitas *database* bisa mengidentifikasi anomali dengan tingkat akurasi yang tinggi, meningkatkan keamanan sistem secara signifikan. bagaimana prediksi *database* memungkinkan perencanaan kapasitas dan konfigurasi *database* yang lebih responsif.

4. Conclusions

Berdasarkan tinjauan literatur terhadap 30 artikel ilmiah, dapat disimpulkan bahwa integrasi machine learning telah menjadi tren utama dalam manajemen database untuk mengatasi skala dan kompleksitas data modern. Aplikasi utamanya mencakup optimasi *query*, deteksi anomali, AutoML, dan analisis prediktif, yang secara kolektif mendorong pergeseran paradigma dari sistem yang dikelola manusia menuju sistem *database* otonom yang cerdas dan adaptif. Meskipun teknik seperti *deep reinforcement learning* dan *supervised learning* terbukti sangat efektif, implementasinya masih menghadapi tantangan signifikan terkait kebutuhan komputasi yang tinggi, interpretabilitas model, dan kesulitan integrasi dengan sistem lawas.

Untuk penelitian selanjutnya, disarankan agar fokus dialihkan pada pengujian empiris melalui implementasi langsung di lingkungan *database* nyata untuk memperoleh bukti kuantitatif mengenai manfaat dan keterbatasan praktisnya. Selain itu, eksplorasi integrasi teknologi baru seperti *Edge AI* dan *federated learning* perlu dilakukan untuk mengatasi isu efisiensi energi, privasi data, dan skalabilitas, sehingga pengembangan sistem database masa depan dapat menjawab tuntutan era big data secara efektif dan berkelanjutan.

Reference

- [1] R. Elmasri and S. B. Navathe, *Fundamentals of Database Systems*, 7th ed., Boston: Pearson, 2016.
- [2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge: MIT Press, 2016.
- [3] T. Kraska, A. Beutel, E. H. Chi, J. Dean, and N. Polyzotis, "The Case for Learned Index Structures," in Proc. of the 2018 ACM SIGMOD, pp. 489–504, 2018.
- [4] R. Marcus and O. Papaemmanouil, "Deep Reinforcement Learning for Join Order Enumeration," arXiv preprint arXiv:1803.00055, 2018.
- [5] M. Sharma and M. Kaushik, "A Review on the Use of *Machine Learning* in *Database* Management," *Journal of Big Data*, vol. 9, no. 1, pp. 1–21, 2022.
- [6] D. Van Aken, A. Pavlo, G. Gordon, and B. Zhang, "An Inquiry into *Machine Learning*-based Automatic Configuration Tuning Services for *Databases*," in Proc. of the 2021 ACM SIGMOD, pp. 1241–1253, 2021.
- [7] G. Roberts, "The Use of *Machine Learning* in *Database* Systems," University of Virginia Technical Report, 2020.
- [8] Z. Yan, V. Uotila, and J. Lu, "Join Order Selection with Deep Reinforcement Learning," *PVLDB*, vol. 16, no. 12, pp. 3882–3894, 2023.
- [9] I. Sabek and M. Hassan, "The Case for Learned In-Memory Joins," *PVLDB*, vol. 16, no. 7, pp. 1749–1761, 2023.
- [10] A. Pavlo, "External vs. Internal: An Essay on *Machine Learning* Agents for Autonomous *Database*