KUADRAN
Jurnal Analitika Data dan Ilmu Komputer

Published by PT Yetti Afrida Center Bengkulu
e-ISSN: 3165-3677 | p-ISSN: 3165-3723
Vol. 1 No. 1 Oktober 2026 | Pages 15-28
https://yettiafridacenter.com/kuadran
This Article is licensed under a CC BY-SA 4.0 International License

Pengembangan Model Regresi dan Machine Learning untuk Meningkatkan Akurasi Prediksi pada Data Multivariat

Veggy Fikri*
Email: veggy.fikri@gmail.com

ABSTRACT

Purpose: *This article develops a comparative framework for linear regression, regularized regression, and machine-learning models to improve prediction in multivariate datasets while maintaining transparent assessment of uncertainty.*

Design/Methodology: *A methodological and analytical literature review examines ordinary least squares, ridge, lasso, elastic net, random forests, gradient boosting, and support vector regression. A reproducible experimental protocol is proposed using preprocessing pipelines, nested cross-validation, and MAE, RMSE, and R-squared; no unperformed empirical experiment is reported.*

Findings: *Prediction quality depends on variable relationships, sample size, collinearity, nonlinear patterns, and validation design. Regularization can stabilize estimates when predictors are correlated, whereas nonlinear learners may capture interactions but require tuning and controls against overfitting and data leakage.*

Practical Implications: *Researchers can apply the framework to select appropriate baselines, evaluate models on unseen data, assess subgroup performance, and report sensitivity before deployment.*

Originality/Value: *The contribution is an integrated evaluation framework linking statistical interpretability, predictive accuracy, feature stability, reproducibility, and domain-aware decisions.*

Keywords: *multivariate data; regression; machine learning; model validation; predictive accuracy; regularization.*

Received: 15/09/2026 Accepted: 22/09/2026 Published: 12/10/2026

A. INTRODUCTION (PENDAHULUAN)

Data multivariat telah menjadi bagian penting dari analisis di bidang pendidikan, ekonomi, kesehatan, manufaktur, pertanian, dan layanan publik. Sebuah hasil sering dipengaruhi oleh banyak variabel yang saling berhubungan, sehingga hubungan satu prediktor dengan variabel respons tidak dapat dipahami tanpa mempertimbangkan prediktor lainnya. Kemajuan perangkat lunak memungkinkan pengolahan data yang semakin kompleks, tetapi kemudahan membangun model tidak dengan sendirinya menjamin bahwa prediksi yang dihasilkan akurat atau dapat dipercaya.

Dalam praktik analisis multivariat, kerumitan sering berasal dari kombinasi perbedaan skala, jumlah variabel yang besar, kualitas pengukuran yang tidak seragam, dan tujuan prediksi yang belum dirumuskan secara tegas. Model yang terlihat unggul pada data pelatihan belum tentu mampu memprediksi pengamatan yang berasal dari waktu atau lingkungan berbeda. Oleh karena itu, pertanyaan utama tidak cukup berbunyi algoritma mana yang paling akurat, melainkan juga apakah pembandingnya adil, bagaimana kesalahan diukur, dan seberapa besar penurunan kinerja ketika kondisi berubah. Regresi konvensional menyediakan landasan untuk menjelaskan hubungan antarvariabel, sedangkan machine learning menawarkan fleksibilitas untuk menangkap pola interaksi dan ketidaklinearan. Penggabungan kedua perspektif memungkinkan peneliti mempertahankan keterjelasan penafsiran sambil menguji kemungkinan perbaikan prediksi.

Kajian ini menempatkan pengembangan model sebagai proses yang berurutan, dimulai dari perumusan target dan audit data, dilanjutkan pengembangan model dasar, regularisasi, serta metode nonlinear, kemudian ditutup dengan evaluasi pada data yang terpisah. Pemilihan algoritma bukan tujuan akhir penelitian, sebab nilai ilmiah sebuah prediksi bergantung pada

This Article is licensed under a CC BY-SA 4.0 International License

keterulangan proses, ketepatan penafsiran, dan kesesuaian dengan kebutuhan pengguna. Pembedaan antara klaim teoretis dan hasil pengujian empiris menjadi prinsip penting agar kesimpulan tidak melampaui bukti yang tersedia.

Regresi linear berganda merupakan titik awal yang kuat karena menyediakan persamaan yang dapat ditafsirkan, prosedur estimasi yang dikenal luas, serta cara eksplisit untuk menguji asumsi. Namun, ketika hubungan sebenarnya tidak linear, variabel sangat berkorelasi, atau pengaruh satu prediktor bergantung pada prediktor lain, model linear sederhana dapat menghasilkan kekeliruan sistematis. Algoritma machine learning menawarkan pilihan pemodelan yang lebih fleksibel; pilihan itu sekaligus membuka masalah baru berupa pemilihan hiperparameter, ketidakstabilan estimasi, dan risiko model hanya menyesuaikan diri dengan sampel pelatihan.

Tujuan artikel ini ialah membangun kerangka komparatif untuk memilih dan mengevaluasi regresi klasik serta beberapa pendekatan machine learning pada data multivariat. Analisis difokuskan pada hubungan antara sifat data, asumsi model, prosedur validasi, metrik kesalahan, penafsiran hasil, dan potensi penerapan. Pembahasan merupakan kajian konseptual-analitis yang menyusun protokol evaluasi yang dapat direplikasi; karena belum melaksanakan pengujian pada satu dataset empiris tertentu, artikel tidak mengklaim angka peningkatan akurasi yang belum diukur.

B. LITERATURE REVIEW (KAJIAN PUSTAKA)

Konsep Data Multivariat dan Ragam Hubungan Antarvariabel

Data multivariat memuat pengamatan beberapa atribut dalam satu unit analisis. Ketika respons bersifat kontinu, persoalan prediktif umumnya dirumuskan sebagai pendekatan fungsi yang menghubungkan matriks prediktor dengan nilai respons. Korelasi antarprediktor dapat mencerminkan informasi yang saling tumpang tindih atau mekanisme bersama, sedangkan interaksi berarti dampak suatu variabel berbeda pada tingkat variabel yang lain. Hubungan yang tampak sederhana dalam analisis bivariat dapat berubah setelah variabel pengganggu dimasukkan.

Dalam regresi linear, asumsi penting mencakup bentuk hubungan rerata bersyarat, perilaku residual, dan independensi yang sesuai rancangan pengamatan. Multikolinearitas tidak selalu mengurangi prediksi pada rentang data serupa, namun dapat membuat koefisien sangat sensitif terhadap perubahan sampel. Variabel yang berkorelasi kuat kadang justru bermanfaat untuk prediksi jika dikendalikan melalui regularisasi. Karena itu, penilaian kualitas model harus disesuaikan dengan tujuan: menjelaskan pengaruh variabel atau memprediksi observasi baru.

Pembeda mendasar antara pemodelan inferensial dan prediktif sering diabaikan. Inferensi berupaya menafsirkan parameter di bawah asumsi tertentu, sementara prediksi menekankan kesalahan di luar sampel. Sebuah model dapat mempunyai koefisien signifikan tetapi performa prediksi rendah, atau sebaliknya dapat memprediksi dengan baik namun sulit menjelaskan mekanisme pengaruh. Literatur pembelajaran statistik mengingatkan bahwa kompleksitas model perlu ditimbang terhadap bias dan varians prediksi (Hastie et al., 2009; James et al., 2023).

Kesulitan lain berkaitan dengan cara informasi disimpan. Dataset multivariat kerap menggabungkan data transaksi, survei, sensor, dan catatan administratif yang memiliki ketelitian berbeda. Tanpa penyamaan definisi variabel, model mungkin menganggap perbedaan praktik pencatatan sebagai sinyal prediktif. Kondisi tersebut menyulitkan penafsiran pengaruh

This Article is licensed under a CC BY-SA 4.0 International License

Vol. 1 No. 1, Oktober 2026, hal. 15–28 e-ISSN: 3165-3677 | p-ISSN: 3165-3723

variabel dan berpotensi menimbulkan perubahan performa ketika data baru dikumpulkan menggunakan prosedur berbeda. Karena itu, kamus data, riwayat perubahan variabel, dan aturan pemeriksaan konsistensi perlu diperlakukan sebagai bagian metodologi, bukan pekerjaan teknis tambahan.

Keputusan pemilihan variabel juga harus mempertimbangkan kapan informasi tersedia. Misalnya, indikator yang baru tercatat setelah suatu kejadian terjadi tidak layak digunakan untuk memprediksi kejadian sebelum berlangsung. Pengujian dengan memasukkan indikator tersebut akan menghasilkan kesan ketepatan yang tinggi, tetapi nilai praktisnya rendah. Kerangka pemodelan yang baik sejak awal membedakan prediktor yang tersedia pada waktu pengambilan keputusan dari informasi yang hanya diketahui sesudahnya.

Regresi Linear dan Regularisasi

Regresi kuadrat terkecil mengestimasi koefisien dengan meminimalkan jumlah kuadrat selisih antara respons dan nilai prediksi. Keunggulannya meliputi transparansi, kemudahan komunikasi, dan kesesuaian dengan hubungan yang mendekati linear. Akan tetapi, estimasi dapat berfluktuasi besar ketika jumlah prediktor tinggi dibanding ukuran sampel. Dalam kondisi tersebut, ridge regression menambahkan penalti kuadrat koefisien sehingga estimasi menyusut dan menjadi lebih stabil.

Lasso menggunakan penalti nilai absolut koefisien, sehingga beberapa koefisien dapat tepat bernilai nol. Sifat ini membuat lasso berguna sebagai penyaring fitur, tetapi pada kelompok prediktor yang sangat berkorelasi ia mungkin memilih satu fitur secara tidak stabil. Elastic net mengombinasikan kedua penalti dan memberi kompromi antara penyusutan dan pemilihan variabel. Nilai penalti tidak boleh dipilih dengan melihat performa pada data pengujian final; pemilihannya perlu berlangsung pada proses validasi di dalam data pelatihan (Tibshirani, 1996; Zou & Hastie, 2005).

Penyusutan menimbulkan bias terkontrol namun dapat mengurangi varians prediksi. Ini menjelaskan mengapa parameter dengan interpretasi kurang tepat pada sampel pelatihan kadang menghasilkan prediksi lebih baik pada sampel baru. Keberhasilan tidak bersifat otomatis: jika hubungan amat sederhana dan dataset cukup besar, kuadrat terkecil bisa menjadi pilihan yang lebih efisien. Pembandingan yang adil perlu menggunakan pemisahan data dan ukuran evaluasi sama untuk setiap kandidat.

Pemahaman mengenai regularisasi juga menegaskan bahwa ketepatan prediksi dan pemilihan fitur tidak identik. Dua model dapat mempunyai kesalahan uji yang hampir sama, tetapi menghasilkan daftar prediktor berbeda akibat korelasi yang kuat. Ketika hasil digunakan untuk mendukung keputusan, perbedaan tersebut harus diperiksa melalui kestabilan pemilihan fitur serta pengetahuan domain. Koefisien menyusut bukan bukti bahwa variabel bersangkutan tidak mempunyai relevansi substantif.

Ridge, lasso, dan elastic net sebaiknya tidak dipandang sebagai prosedur yang saling menggantikan dalam seluruh keadaan. Ridge mempertahankan seluruh prediktor melalui penyusutan yang kontinu, sedangkan lasso dapat menyingkirkan sebagian koefisien sekaligus melakukan pemilihan fitur. Elastic net menambahkan ruang kompromi ketika kelompok fitur saling berkorelasi. Perbandingan ketiganya menuntut standardisasi fitur yang dihitung hanya pada data pelatihan, terutama bila skala numerik berbeda jauh. Tanpa perlakuan tersebut, besarnya penalti akan bergantung pada satuan variabel dan bukan semata kekuatan hubungan prediktif.

This Article is licensed under a CC BY-SA 4.0 International License

Vol. 1 No. 1, Oktober 2026, hal. 15–28 e-ISSN: 3165-3677 | p-ISSN: 3165-3723

Perlu pula dibedakan kemampuan memilih fitur dan kemampuan mengidentifikasi faktor penyebab. Koefisien yang bernilai nol setelah penalti tidak membuktikan suatu variabel tidak berpengaruh secara substantif, sebab efeknya dapat dibawa oleh fitur lain yang berkorelasi. Karena itu, daftar variabel terpilih sebaiknya dibandingkan pada beberapa pembagian data. Jika hasilnya berubah-ubah, peneliti perlu melaporkan ketidakstabilan tersebut dan menghindari kesimpulan yang terlalu tegas.

Algoritma Machine Learning untuk Respons Kontinu

Random forest membangun banyak pohon keputusan menggunakan variasi sampel serta subset prediktor. Prediksi kemudian dirata-ratakan agar ketidakstabilan sebuah pohon berkurang. Pendekatan ini mampu menangkap hubungan nonlinear dan interaksi tanpa menentukan seluruh bentuk fungsi sejak awal. Namun, prediksi model pohon di luar rentang nilai pelatihan dapat kurang andal, dan ukuran pentingnya fitur tertentu dapat bias apabila tidak dihitung dengan prosedur yang sesuai (Breiman, 2001).

Gradient boosting menyusun model lemah secara bertahap untuk memperbaiki kesalahan model sebelumnya. Pengaturan kedalaman pohon, laju pembelajaran, dan jumlah iterasi sangat menentukan kemampuan generalisasi. Model yang terlalu rumit dapat mengejar fluktuasi acak; validasi silang dan penghentian dini pada partisi yang tepat membantu mengendalikan risiko ini. Algoritma boosting modern menawarkan efisiensi komputasi yang tinggi tetapi tetap memerlukan pelaporan rancangan eksperimen.

Support vector regression berupaya menemukan fungsi yang menjaga kesalahan dalam toleransi tertentu sembari mengendalikan kompleksitas. Dengan kernel nonlinear, model dapat merepresentasikan hubungan yang sulit ditangkap regresi linear. Akan tetapi, pemilihan kernel, parameter regularisasi, dan lebar kernel harus didasarkan pada validasi. Tidak ada satu algoritma yang mendominasi seluruh dataset: pendekatan paling tepat bergantung pada ukuran sampel, sinyal yang tersedia, tingkat kebisingan, serta kebutuhan interpretasi.

Model nonlinear harus diberi baseline yang masuk akal agar nilai tambahnya terlihat. Sebuah random forest yang tampak kuat dibanding model linear tanpa prapemrosesan yang memadai belum membuktikan bahwa fleksibilitas algoritma menjadi penyebab perbedaan performa. Semua kandidat seyogianya menerima perlakuan yang konsisten terhadap partisi, definisi target, dan keterbatasan fitur. Jika perbaikan prediksi kecil tetapi biaya komputasi jauh lebih tinggi, kelebihan praktisnya patut dipertanyakan.

Model berbasis pohon berguna ketika hubungan antara prediktor dan respons mempunyai perubahan tajam pada rentang tertentu, tetapi tetap perlu dievaluasi terhadap data yang mewakili kondisi penerapan. Random forest cenderung mengurangi varians melalui penggabungan banyak pohon, sedangkan boosting memperbaiki kesalahan secara bertahap dan berpotensi menyesuaikan pola lebih rinci. Pada kedua pendekatan, nilai pentingnya fitur perlu dibaca secara hati-hati karena hasilnya bergantung pada metode pengukuran dan struktur korelasi antarprediktor.

Support vector regression menambahkan pilihan lain melalui pengaturan margin toleransi dan fungsi kernel. Pada data dengan ribuan atribut atau sampel yang terbatas, biaya pencarian hiperparameter serta kebutuhan standardisasi dapat menjadi pertimbangan penting. Pendekatan nonlinear yang kompleks tidak selalu memberikan keuntungan praktis apabila kesalahan uji hampir sama dengan model regresi sederhana. Oleh sebab itu, evaluasi juga mencatat waktu pelatihan, kemampuan menjelaskan keluaran, dan kemudahan pembaruan model ketika data baru tersedia.

This Article is licensed under a CC BY-SA 4.0 International License

Vol. 1 No. 1, Oktober 2026, hal. 15–28 e-ISSN: 3165-3677 | p-ISSN: 3165-3723

C. METHOD (METODE)

Desain Kajian dan Pertanyaan Analitis

Artikel disusun melalui kajian metodologis terhadap konsep regresi statistik dan pembelajaran mesin untuk data multivariat. Objek perbandingan meliputi regresi kuadrat terkecil, ridge, lasso, elastic net, random forest, gradient boosting, dan support vector regression. Literatur metodologi digunakan untuk mengidentifikasi asumsi, keunggulan, keterbatasan, indikator akurasi, serta prinsip evaluasi yang dapat dipertanggungjawabkan. Penelitian ini tidak mengklaim penggunaan data responden, data laboratorium, atau hasil pelatihan model yang belum dilaksanakan.

Kerangka analisis terdiri atas lima tahap. Pertama, peneliti mendefinisikan target prediksi, populasi sasaran, dan unit pengamatan. Kedua, dilakukan audit kualitas data meliputi nilai hilang, variabel kategorikal, pencilan, dan potensi pergeseran distribusi. Ketiga, data dibagi menurut rancangan validasi yang menjaga independensi pengamatan. Keempat, kandidat model dilatih melalui pipeline yang memuat seluruh transformasi data. Kelima, hasil dievaluasi memakai ukuran kesalahan, interval ketidakpastian, analisis subkelompok, dan kebutuhan penjelasan.

Kriteria evaluasi bukan hanya ketepatan rata-rata. Peneliti perlu mencatat stabilitas prediksi terhadap perubahan partisi, kompleksitas pemrosesan, kebutuhan komputasi, dan potensi konsekuensi dari prediksi salah. Jika data memiliki urutan waktu, pembagian acak dapat menyebabkan informasi masa depan masuk ke pelatihan. Jika beberapa pengamatan berasal dari individu yang sama, pembagian berdasarkan individu lebih tepat daripada memisahkan baris secara acak.

Dengan rancangan yang jelas, analisis dapat dipisahkan antara aktivitas deskriptif, estimasi hubungan, serta prediksi prospektif. Ketiganya dapat memakai data yang sama, tetapi memerlukan kriteria keberhasilan yang berlainan. Rancangan dokumentasi sejak awal memudahkan pembaca membedakan hasil konseptual dari eksperimen yang benar-benar telah dilakukan, termasuk menjelaskan mengapa sejumlah metode dipilih dan metode lain tidak dicakup dalam evaluasi.

Tahap audit data dilakukan sebelum pelatihan agar karakteristik observasi dapat dipahami secara sistematis. Pemeriksaan mencakup jumlah unit yang benar-benar independen, proporsi nilai hilang pada tiap variabel, jenis pengkodean kategori, pola pencilan, dan kemungkinan nilai respons mengalami kesalahan pencatatan. Strategi imputasi harus disesuaikan dengan mekanisme hilangnya informasi; pengisian rerata tidak selalu tepat apabila kehilangan data berkaitan dengan karakteristik individu atau institusi. Pencilan perlu dianalisis untuk menentukan apakah ia kesalahan pencatatan atau keadaan sah yang memang relevan bagi prediksi.

Pada tahap perancangan, seluruh kandidat model harus memanfaatkan target, partisi evaluasi, serta definisi keberhasilan yang sama. Jumlah fitur, strategi pemilihan variabel, dan ruang pencarian hiperparameter perlu didokumentasikan. Kesetaraan prosedur membantu memastikan bahwa perbedaan hasil benar-benar berkaitan dengan kemampuan model, bukan dengan akses informasi yang lebih besar atau keuntungan dari pemilihan partisi yang berbeda.

Rancangan Validasi dan Pencegahan Kebocoran Data

Data leakage terjadi ketika informasi yang pada saat penerapan belum tersedia digunakan pada tahap pelatihan atau pemilihan model. Kebocoran dapat muncul dari normalisasi yang

This Article is licensed under a CC BY-SA 4.0 International License

dihitung pada keseluruhan data sebelum pemisahan, imputasi berdasarkan data pengujian, variabel yang merupakan turunan langsung dari target, atau pengulangan entitas dalam partisi pelatihan dan pengujian. Kesalahan semacam itu dapat menghasilkan skor yang tampak sangat baik tetapi gagal ketika sistem diterapkan.

Rancangan yang disarankan ialah menyisihkan data uji final satu kali, kemudian menggunakan validasi silang pada bagian pelatihan untuk memilih hiperparameter. Untuk membandingkan banyak model dan mengestimasi kinerja prosedur pemilihan secara lebih tidak bias, nested cross-validation menyediakan validasi luar dan pencarian hiperparameter di bagian dalam. Seluruh tahapan imputasi, standardisasi, encoding, serta seleksi fitur harus dipelajari hanya dari fold pelatihan yang bersangkutan.

Penilaian memakai mean absolute error (MAE) untuk rata-rata selisih absolut, root mean squared error (RMSE) untuk memberi penalti lebih besar pada kesalahan ekstrem, serta koefisien determinasi R^2 untuk membandingkan variasi yang dapat dijelaskan relatif terhadap prediktor konstan. Nilai R^2 negatif pada data uji dimungkinkan, sehingga penafsiran wajib membedakan hasil pelatihan dan hasil pengujian. Ukuran yang dipilih harus mempertimbangkan satuan target dan biaya substantif dari kesalahan.

Ketika selisih kesalahan dua kandidat kecil, peneliti perlu menilai rentang variasinya, bukan langsung menyatakan model tertentu unggul. Bootstrap yang menghormati struktur data atau pengulangan validasi dapat memberi gambaran ketidakpastian. Hasil tersebut sebaiknya dilaporkan bersama jumlah observasi, pembagian data, rentang hiperparameter, dan seed acak.

Penentuan satuan analisis merupakan keputusan penting yang sering kurang mendapat perhatian. Pada data sekolah, sejumlah siswa mungkin berasal dari sekolah yang sama; pada data kesehatan, banyak kunjungan dapat dilakukan oleh pasien yang sama; dan pada data manufaktur, catatan berulang dapat berasal dari mesin yang sama. Apabila ketergantungan ini diabaikan, pemisahan acak per baris dapat menilai performa secara terlalu optimistis. Prinsip pemisahan berkelompok dan temporal memberikan evaluasi yang lebih mendekati penggunaan di lingkungan baru.

Pemilihan skema validasi harus mengikuti cara model nantinya dipakai. Untuk data lintas individu yang independen, validasi silang acak dapat memberikan estimasi awal yang berguna. Sebaliknya, untuk data yang dikumpulkan secara bertahap, evaluasi temporal lebih mendekati situasi ketika model dilatih dari masa lalu dan diminta memprediksi masa depan. Apabila terdapat beberapa catatan dari unit yang sama, seluruh catatan satu unit sebaiknya ditempatkan dalam kelompok yang sama pada satu partisi. Pilihan ini mencegah model hanya mengenali karakteristik unit yang sebelumnya telah muncul saat pelatihan.

Selain angka MAE, RMSE, dan R^2, laporan dapat menyajikan distribusi residual untuk mengidentifikasi pola kesalahan menurut besaran target. Kinerja agregat yang baik bisa menyembunyikan kesalahan ekstrem pada rentang respons tertentu. Dengan demikian, evaluasi perlu menyertakan perbandingan terhadap baseline sederhana, misalnya prediksi rerata data pelatihan, agar manfaat tambahan algoritma dapat dibuktikan secara proporsional.

D. RESULT AND DISCUSSION (HASIL DAN PEMBAHASAN)

Persamaan Dasar dan Interpretasi Koefisien

Pada regresi linear berganda, hubungan target y dengan prediktor x_1 hingga x_p direpresentasikan sebagai $y = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p + \varepsilon$. Koefisien β_j mengukur perbedaan rerata respons yang berasosiasi dengan perubahan satu unit x_j ketika prediktor lain dikendalikan,

This Article is licensed under a CC BY-SA 4.0 International License

sepanjang model dan asumsi penafsiran sesuai. Persamaan tidak membuktikan pengaruh kausal semata-mata karena koefisiennya bernilai besar atau signifikan.

Dalam ridge, objektif kuadrat terkecil ditambah penalti $\lambda\Sigma\beta_j^2$, sedangkan lasso menggunakan $\lambda\Sigma|\beta_j|$. Semakin besar λ, semakin kuat penyusutan; untuk λ terlalu besar model berisiko kurang menangkap pola yang benar. Penalti umumnya tidak diterapkan pada intersep dan variabel numerik perlu distandardisasi dalam fold pelatihan agar penalti sebanding. Elastic net menimbang kombinasi kedua penalti sehingga memberi fleksibilitas untuk fitur yang berkorelasi.

Hubungan rumus dengan penilaian operasional penting dijelaskan. Model dengan jumlah parameter banyak tidak otomatis memiliki kemampuan prediksi lebih tinggi. Kesalahan yang rendah pada data pelatihan hanyalah bukti kecocokan terhadap data yang telah dilihat. Kemampuan generalisasi perlu diuji pada pengamatan baru yang tidak ikut memengaruhi pelatihan, penyesuaian hiperparameter, atau pemilihan fitur.

Bila variabel memiliki hubungan nonlinear yang dapat dijelaskan, transformasi atau basis spline juga perlu dipertimbangkan sebelum langsung memilih algoritma yang lebih rumit. Pilihan seperti ini menjaga keterbacaan model sambil memperluas fleksibilitas. Keseimbangan antara kemudahan interpretasi dan kemampuan menangkap struktur data merupakan pertimbangan utama.

Dalam komunikasi ilmiah, simbol matematika perlu selalu disertai penjelasan yang dapat dibaca pembaca lintas disiplin. Penambahan penalti pada suatu fungsi objektif bukan semata prosedur komputasi, melainkan pilihan untuk mengurangi ketidakstabilan yang muncul dari data terbatas. Nilai parameter yang optimal ditentukan berdasarkan bukti pada sampel validasi, bukan preferensi algoritma. Pendekatan demikian menempatkan rumus sebagai alat penalaran dan pengambilan keputusan metodologis.

Interpretasi koefisien regresi linear perlu selalu memperhatikan konteks data. Perubahan prediksi yang berkaitan dengan satu fitur dihitung dengan mempertahankan fitur lain tetap pada nilai tertentu, tetapi keadaan demikian belum tentu dapat diwujudkan jika sejumlah variabel memiliki hubungan yang kuat. Karena itu, koefisien dapat menjadi alat deskripsi bersyarat tanpa harus diterjemahkan sebagai pengaruh sebab-akibat. Dalam kajian prediktif, fokus tambahan diberikan pada residual yang konsisten, perubahan kesalahan sepanjang rentang fitur, dan performa ketika pengamatan baru berada dekat batas data pelatihan.

Regularisasi membantu mengendalikan ketidakstabilan koefisien, namun tidak memperbaiki kesalahan target atau membenarkan penggunaan prediktor yang tersedia terlambat. Pengembangan model semestinya berlangsung dari baseline linear menuju variasi yang semakin fleksibel secara bertahap. Setiap tambahan kompleksitas harus dinilai berdasarkan besarnya manfaat pada pengujian luar sampel dan apakah tambahan tersebut memberi informasi bermakna bagi pengguna.

Perbandingan Karakteristik Model

Model regresi berganda biasanya cocok sebagai baseline karena memberikan tolok ukur transparan yang sulit dilewati tanpa alasan substantif. Ridge dapat diprioritaskan ketika prediktor saling berkorelasi, lasso ketika seleksi variabel dianggap penting, dan elastic net ketika keduanya muncul bersamaan. Random forest serta gradient boosting lebih menarik ketika data menampilkan interaksi dan ambang nonlinear, sedangkan support vector regression dapat menjadi kandidat untuk dataset berukuran sedang yang telah diskalakan dengan benar.

This Article is licensed under a CC BY-SA 4.0 International License

Dalam praktik, pemilihan model perlu memerhatikan ukuran sampel efektif. Data yang memiliki ribuan baris belum tentu menyediakan ribuan unit independen apabila semua pengamatan diambil berulang dari beberapa entitas. Pada jumlah pengamatan independen yang kecil, model fleksibel lebih rentan terhadap ketidakstabilan. Sebaliknya, pada data yang beragam dan cukup besar, pola nonlinear dapat membantu memperkecil kesalahan, asalkan pembagian uji menjaga kesamaan kondisi penerapan.

Pembandingan tidak seharusnya mengandalkan satu peringkat berdasarkan satu angka. MAE dapat menunjukkan performa yang baik pada kasus biasa, tetapi RMSE membuka kemungkinan kesalahan besar pada sebagian pengamatan. Jika nilai target sangat menceng, evaluasi dalam satuan alami perlu dipertahankan sekalipun respons ditransformasi sewaktu pelatihan. Metode yang dipilih juga harus mempunyai strategi penanganan nilai hilang, outlier, dan penyimpangan data operasional.

Kesimpulan analitis dari literatur ialah bahwa pengembangan akurasi merupakan proses rekayasa evaluasi, bukan akibat otomatis dari label machine learning. Tanpa validasi yang setara, klaim superioritas algoritma tidak dapat dibedakan dari keuntungan akibat proses pengujian yang tidak adil. Oleh karena itu, pelaporan sebaiknya memprioritaskan prosedur, rentang kesalahan, serta bukti replikasi.

Perbandingan yang bermakna juga perlu memerhatikan biaya kesalahan yang tidak simetris. Dalam beberapa aplikasi, salah memprediksi terlalu rendah dapat lebih merugikan daripada terlalu tinggi, atau sebaliknya. MAE dan RMSE tidak selalu menangkap konsekuensi tersebut. Peneliti dapat menambahkan ukuran berbobot atau fungsi kerugian yang sesuai dengan kebutuhan operasional, tetapi pilihan tersebut harus ditetapkan secara terbuka agar tidak dipakai sekadar untuk memenangkan suatu algoritma.

Tabel evaluasi yang ideal menempatkan model-model pembanding pada baris yang sama dengan ukuran kesalahan yang konsisten. Selain peringkat, peneliti dapat menambahkan rentang variasi antarpartisi, jumlah fitur yang dipakai, dan waktu komputasi. Pemilihan model tidak harus mengikuti selisih numerik terkecil jika selisih itu berada dalam variasi yang mungkin timbul karena perubahan sampel. Model yang lebih sederhana dapat lebih menguntungkan ketika memerlukan pelatihan ulang yang sering atau harus dijelaskan kepada pemangku kepentingan nonteknis.

Pada aplikasi praktis, konsekuensi kesalahan perlu didefinisikan sejak rancangan penelitian. Prediksi permintaan yang terlalu rendah mungkin menyebabkan kekurangan pasokan, sedangkan prediksi terlalu tinggi dapat menimbulkan pemborosan. Perbedaan ini dapat menuntut penilaian tambahan selain rata-rata kesalahan absolut atau kuadrat. Meski demikian, metrik khusus tidak boleh dipilih sesudah melihat skor akhir karena prosedur itu dapat menghasilkan penilaian yang memihak satu kandidat.

Pengaruh Multikolinearitas dan Seleksi Fitur

Multikolinearitas terjadi saat sejumlah prediktor menyediakan informasi serupa. Dalam model kuadrat terkecil, kondisi ini dapat memperbesar varians koefisien dan menimbulkan perubahan tanda yang tidak stabil ketika satu variabel dihapus. Namun, menyingkirkan semua prediktor berkorelasi tanpa mempertimbangkan tujuan prediksi bisa mengurangi informasi. Pilihan penanganan meliputi pemeriksaan korelasi, faktor inflasi varians dalam konteks yang tepat, regularisasi, atau reduksi dimensi.

This Article is licensed under a CC BY-SA 4.0 International License

Vol. 1 No. 1, Oktober 2026, hal. 15–28 e-ISSN: 3165-3677 | p-ISSN: 3165-3723

Penyaringan fitur menggunakan target harus dilakukan di dalam setiap fold pelatihan. Jika peneliti lebih dahulu memilih prediktor berdasarkan korelasi dengan target pada seluruh dataset, maka data uji secara tidak langsung memengaruhi model. Ini meningkatkan peluang optimisme dalam estimasi performa. Seleksi yang stabil juga memerlukan pemeriksaan seberapa sering suatu fitur terpilih dalam partisi berbeda, bukan hanya daftar fitur dari satu putaran.

Ketika fitur memiliki makna substantif yang kuat, pilihan yang paling akurat belum tentu paling berguna. Misalnya, dalam perencanaan layanan, prediktor yang tersedia setelah keputusan dibuat tidak dapat dipakai untuk prediksi prospektif. Fitur juga bisa bertindak sebagai proksi yang tidak diinginkan untuk atribut sensitif. Audit kesiapan fitur pada saat penggunaan menjadi bagian dari kualitas ilmiah model.

Penjelasan model perlu dibedakan antara interpretasi global dan lokal. Koefisien regresi memberi arah perubahan bersyarat di bawah spesifikasi linear, sedangkan importance pada model pohon menunjukkan kontribusi prediktif menurut prosedur tertentu. Kedua ukuran tidak dapat disamakan dengan dampak kausal. Peneliti harus menjelaskan batas tersebut sebelum menghasilkan rekomendasi kebijakan.

Selain seleksi fitur yang dilakukan pada pelatihan, kualitas label target perlu mendapat perhatian. Ketika nilai respons berasal dari pengukuran yang tidak konsisten, model mungkin hanya mempelajari praktik pencatatan daripada fenomena yang ingin diprediksi. Dokumentasi definisi operasional, perubahan instrumen pengukuran, dan proses penggabungan beberapa sumber data membantu mengurangi kemungkinan kesalahan. Tanpa pemeriksaan tersebut, penggunaan teknik yang lebih canggih belum tentu memberi peningkatan kualitas informasi.

Hubungan antarvariabel juga dapat berubah ketika data berasal dari kelompok yang berbeda. Dua fitur mungkin berkorelasi kuat pada keseluruhan data tetapi menunjukkan pola berbeda ketika diamati secara terpisah menurut lokasi atau periode. Pemeriksaan struktur korelasi sebaiknya dilakukan pada data pelatihan dan disertai pertimbangan tentang apakah keterkaitan tersebut mempunyai dasar substantif. Pengurangan dimensi seperti principal component analysis dapat membantu dalam situasi tertentu, tetapi komponennya perlu ditafsirkan dengan hati-hati karena merupakan kombinasi variabel asli.

Kestabilan fitur terpilih dapat dievaluasi dengan menghitung seberapa sering suatu prediktor muncul pada model dari pembagian data yang berbeda. Namun, frekuensi ini pun tidak identik dengan bukti hubungan kausal. Ketika tujuan penelitian menuntut penjelasan kebijakan, hasil seleksi perlu dikombinasikan dengan rancangan inferensial yang sesuai dan pengetahuan tentang mekanisme domain. Pemisahan tujuan ini menjaga agar rekomendasi tidak didasarkan pada korelasi prediktif semata.

Pengendalian Overfitting dan Tuning Hiperparameter

Overfitting muncul ketika model mempelajari pola khusus sampel yang tidak bertahan pada pengamatan baru. Gejalanya berupa kesalahan pelatihan sangat kecil tetapi kesalahan validasi relatif besar. Risiko ini meningkat ketika ruang pencarian hiperparameter luas, jumlah kandidat dicoba berkali-kali, atau keputusan analisis diubah berdasarkan data uji final. Model sederhana sekalipun dapat mengalami overfitting melalui seleksi berulang dan interaksi yang tidak direncanakan.

Random forest memiliki parameter seperti jumlah pohon, jumlah fitur pada tiap split, dan ukuran minimum daun. Pada boosting, kedalaman pohon, laju pembelajaran, jumlah iterasi, serta regularisasi memengaruhi bias dan varians. Support vector regression membutuhkan pengaturan parameter penalti, toleransi kesalahan, dan kernel. Tidak tersedia satu kombinasi

This Article is licensed under a CC BY-SA 4.0 International License

universal yang selalu optimal; pencarian harus dibatasi secara rasional dan dinilai menggunakan validasi yang tepat.

Pemilihan pemenang sebaiknya mempertimbangkan variasi kinerja antar-fold. Bila dua model mempunyai kesalahan yang hampir sama, pendekatan lebih sederhana, mudah diaudit, dan murah dikelola mungkin lebih layak dipilih. Keputusan tersebut bukan menolak inovasi algoritmik, melainkan menghindari kompleksitas yang tidak dibayar dengan keuntungan prediksi yang bermakna.

Kinerja setelah penerapan harus dipantau karena data berubah. Pergeseran prevalensi, perubahan hubungan prediktor dengan respons, atau modifikasi proses pencatatan dapat membuat model lama menurun akurasinya. Rancangan evaluasi awal idealnya menetapkan indikator pemantauan, ambang investigasi, mekanisme pembaruan, dan kewajiban dokumentasi perubahan versi.

Upaya mencegah overfitting dapat dilakukan sejak perumusan hipotesis prediksi. Jumlah variasi model yang dicoba sebaiknya sebanding dengan informasi yang tersedia, sementara keputusan penting seperti pilihan metrik dan rentang parameter disimpan dalam catatan penelitian. Kedisiplinan ini membantu membedakan hasil yang stabil dari keberuntungan pada satu partisi data. Dengan demikian, evaluasi tidak sekadar memilih nilai skor tertinggi, tetapi menyajikan bukti yang dapat diperiksa kembali.

Overfitting tidak selalu tampak dari jumlah parameter yang besar. Proses mencoba banyak rancangan sampai memperoleh hasil uji terbaik juga dapat secara tidak langsung menyesuaikan penelitian terhadap data uji. Karena itu, jumlah eksperimen, alasan perubahan rancangan, dan aturan penghentian pencarian perlu dicatat. Ketika hasil eksperimen akan dilaporkan sebagai perbandingan ilmiah, satu dataset uji akhir yang tidak pernah digunakan untuk memilih model mempunyai peran khusus sebagai pemeriksaan independen.

Penting pula membedakan kebutuhan tuning antaralgoritma. Sebagian metode mempunyai parameter dasar yang relatif stabil pada banyak persoalan, sedangkan algoritma lain peka terhadap rentang pencarian dan skala data. Perbandingan yang adil tidak selalu berarti seluruh model diberi jumlah percobaan yang identik, melainkan sumber daya pencarian dan alasan pemilihannya dilaporkan secara terbuka. Dengan transparansi tersebut, pembaca dapat menilai apakah peningkatan performa cukup kuat untuk membenarkan kerumitan tambahan.

Validitas Eksternal, Ketidakpastian, dan Keadilan Prediksi

Validitas eksternal menanyakan apakah model bekerja pada populasi, periode, dan lokasi yang berbeda dari data pelatihan. Skor rata-rata yang baik tidak menjamin performa konsisten untuk tiap kelompok pengguna. Kesalahan dapat meningkat pada wilayah yang kurang terwakili, rentang umur tertentu, atau observasi dengan nilai prediktor ekstrem. Audit subkelompok perlu dirancang hati-hati agar tidak menyimpulkan ketidakadilan dari sampel yang terlalu kecil.

Ketidakpastian prediksi dapat berasal dari kebisingan data, variasi sampel, dan ketidakpastian spesifikasi model. Interval prediksi tidak sama dengan interval keyakinan rerata: interval prediksi bertujuan menggambarkan kemungkinan respons individual dan biasanya lebih lebar. Metode yang digunakan untuk membangun interval perlu dinyatakan bersama asumsi cakupannya. Pelaporan titik prediksi tunggal tanpa rentang risiko bisa memberi kesan kepastian yang keliru.

This Article is licensed under a CC BY-SA 4.0 International License

Pengujian pada data lintas waktu menjadi sangat penting bila kondisi dunia nyata berubah. Model yang dilatih sebelum perubahan aturan, teknologi, atau pola konsumsi mungkin perlu divalidasi ulang. Evaluasi hanya menggunakan partisi acak dapat menyembunyikan kegagalan yang akan muncul saat model dioperasikan pada periode berikutnya. Pembagian temporal harus disertai informasi ketersediaan fitur pada waktu prediksi.

Implikasi etis tidak terpisahkan dari mutu statistik. Dalam aplikasi berisiko tinggi, model dengan penurunan MAE kecil tidak otomatis layak digunakan jika sulit diaudit atau menimbulkan disparitas kesalahan yang besar. Transparansi asumsi, mekanisme koreksi, dan pengawasan manusia diperlukan untuk menempatkan prediksi sebagai alat bantu keputusan yang proporsional.

Perbedaan kinerja antarwilayah atau antarperiode tidak selalu berarti algoritma bermasalah; penyebabnya dapat berupa perubahan cara data dikumpulkan, berkurangnya representasi kelompok tertentu, atau munculnya hubungan baru yang tidak ada dalam pelatihan. Prosedur pemantauan perlu mencari sumber perubahan sebelum model dilatih ulang. Penambahan data baru juga perlu diaudit agar tidak memperkuat bias pengamatan yang sudah ada. Hal ini membuat pemeliharaan model menjadi proses ilmiah berkelanjutan.

Pergeseran distribusi data dapat terjadi ketika komposisi subjek, cara pengukuran, atau hubungan antarvariabel berubah. Pergeseran pada sebaran prediktor belum tentu langsung menyebabkan penurunan akurasi, tetapi memberi alasan untuk melakukan pemeriksaan tambahan. Jika hubungan antara prediktor dan target berubah, model yang sebelumnya terkalibrasi mungkin perlu diperbarui. Karena itu, pengujian periodik perlu disertai batas kinerja yang jelas dan aturan kapan model harus ditinjau ulang, bukan otomatis dilatih ulang setiap kali skor berfluktuasi.

Penilaian kelompok juga tidak cukup hanya membandingkan satu nilai kesalahan. Jumlah observasi, rentang nilai target, serta ketidakpastian perhitungan harus dipertimbangkan agar perbedaan antarkelompok tidak ditafsirkan berlebihan. Pada penerapan dengan dampak penting bagi manusia, dokumentasi tentang siapa yang dapat meninjau prediksi dan bagaimana koreksi dilakukan menjadi pelengkap evaluasi statistik.

Protokol Pelaporan dan Reproduksibilitas

Sebuah laporan pemodelan yang dapat direproduksi sebaiknya mencatat sumber dan periode data, kriteria inklusi, kamus variabel, cara menangani nilai hilang, distribusi target, serta struktur pengulangan pengamatan. Rincian pembagian data, pipeline transformasi, daftar hiperparameter, ukuran evaluasi, dan perangkat lunak harus dapat ditelusuri. Jika data tidak boleh dibuka, peneliti dapat membagikan kode dan contoh sintetis yang jelas dilabeli agar prosedur tetap dapat diperiksa.

Hasil eksperimen hendaknya memuat tabel komparatif kesalahan data uji beserta ukuran variasi antar-validasi. Grafik residual dapat menunjukkan ketidakseragaman varians atau hubungan yang belum tertangkap, sedangkan plot prediksi terhadap nilai aktual membantu mengidentifikasi sistematisasi kesalahan. Namun, grafik hanya merupakan pelengkap, bukan pengganti pengujian pada data yang benar-benar tidak dipakai dalam proses pemilihan.

Transparansi juga mencakup pelaporan percobaan yang gagal atau tidak unggul. Pemilihan satu hasil terbaik dari banyak percobaan tanpa memberi gambaran rentang eksperimen berpotensi membuat estimasi akurasi tampak terlalu optimistis. Peneliti dapat merancang protokol sebelum pemodelan sehingga pemilihan metrik, pembagian data, dan batas pencarian ditentukan terlebih dahulu.

This Article is licensed under a CC BY-SA 4.0 International License

Dalam konteks penelitian lintas bidang, keterlibatan ahli domain diperlukan untuk mengidentifikasi definisi target dan risiko interpretasi. Model yang secara statistik konsisten mungkin tidak relevan apabila target hanya proksi lemah dari fenomena yang ingin dipahami. Kerangka yang diajukan dengan demikian menghubungkan validasi algoritmik, kualitas pengukuran, dan tanggung jawab penggunaan.

Bagi pengelola jurnal dan pembaca akademik, keterlacakan proses merupakan indikator penting mutu laporan penelitian. Artikel yang hanya menyatakan sebuah algoritma menghasilkan akurasi terbaik tanpa menjelaskan data, validasi, dan konfigurasi tidak menyediakan dasar yang memadai untuk reproduksi. Pada kajian konseptual, batas klaim perlu lebih tegas lagi: manfaat suatu metode dijelaskan berdasarkan teori dan bukti literatur, sedangkan hasil eksperimen baru harus dinyatakan sebagai agenda penelitian berikutnya.

Reproduksibilitas memerlukan pemisahan yang jelas antara rancangan analisis, data yang digunakan, dan hasil perhitungan. Berkas kode dapat dibagikan bersama petunjuk lingkungan komputasi, versi pustaka, pengaturan acak, dan urutan transformasi. Bila data utama tidak tersedia secara terbuka, contoh sintetis dapat membantu memeriksa apakah alur komputasi berjalan, tetapi tidak dapat digunakan untuk mengklaim replikasi hasil empiris pada data asli. Penjelasan tentang pembatasan data membuat ruang verifikasi menjadi lebih transparan.

Laporan yang baik juga memberi tahu pembaca prosedur apa yang tidak dilakukan. Jika suatu model belum diuji pada lokasi atau periode lain, keterbatasan tersebut perlu dinyatakan dalam pembahasan. Jika ketidakpastian prediksi belum dihitung, angka hasil sebaiknya tidak dibahas seolah-olah memiliki ketepatan mutlak. Disiplin pelaporan membantu menjembatani pengembangan algoritma dengan kebutuhan penilaian ilmiah yang dapat dipertanggungjawabkan.

E. CONCLUSION (KESIMPULAN)

Pembahasan menunjukkan bahwa pengembangan prediksi pada data multivariat harus dimulai dari pemahaman karakteristik data dan tujuan penggunaan. Regresi linear berganda memberikan baseline interpretatif yang kuat, sementara ridge, lasso, serta elastic net menawarkan penyusutan koefisien ketika korelasi antarprediktor atau kompleksitas fitur menjadi persoalan. Random forest, gradient boosting, dan support vector regression dapat menambah fleksibilitas untuk pola nonlinear, tetapi memerlukan validasi dan pengendalian kompleksitas yang konsisten.

Kontribusi utama kajian ini ialah kerangka evaluasi yang memperlakukan akurasi sebagai hasil dari keseluruhan proses: perumusan target, audit kualitas data, pipeline prapemrosesan, pemilihan model, validasi, pengukuran ketidakpastian, dan pemeriksaan penerapan. Kajian ini tidak memuat hasil eksperimen empiris, sehingga tidak dapat menyatakan algoritma tertentu meningkatkan akurasi sebesar persentase khusus. Pengujian langsung pada dataset nyata menjadi langkah penting sebelum memberi rekomendasi operasional.

Penelitian lanjutan perlu menguji protokol menggunakan beberapa dataset lintas bidang yang mempunyai karakter berbeda, mencatat waktu komputasi, dan menilai stabilitas performa di luar sampel. Analisis fairness, pergeseran distribusi, serta interpretasi model juga perlu dijadikan bagian dari evaluasi yang setara. Dengan prinsip tersebut, regresi dan machine learning dapat dipadukan sebagai pendekatan yang saling melengkapi untuk menghasilkan prediksi yang tidak hanya lebih tepat, tetapi juga lebih dapat dijelaskan dan dipertanggungjawabkan.

This Article is licensed under a CC BY-SA 4.0 International License

Secara lebih luas, hasil kajian menegaskan bahwa literasi statistika tidak kehilangan arti ketika algoritma pembelajaran mesin semakin mudah digunakan. Pemahaman asumsi, kesalahan pengukuran, generalisasi, dan ketidakpastian justru semakin dibutuhkan untuk menilai keluaran perangkat lunak. Kolaborasi antara pendekatan regresi dan pemodelan nonlinear akan memberikan manfaat terbesar jika berorientasi pada pertanyaan yang jelas, kualitas data yang baik, serta evaluasi yang dapat diulang.

Secara konseptual, kontribusi pengembangan model terletak pada keterpaduan antara prinsip statistika dan praktik komputasi modern. Statistik menyediakan alat untuk memahami asumsi, galat, dan ketidakpastian, sedangkan pembelajaran mesin menambah ruang pemodelan untuk interaksi atau ketidaklinearan yang sulit dirumuskan sebelumnya. Keduanya seharusnya saling menguatkan: model yang mudah ditafsirkan digunakan sebagai pembanding, sedangkan model fleksibel diuji untuk menunjukkan manfaat nyata melalui evaluasi yang setara.

Batas utama kajian ini adalah belum dilakukannya estimasi langsung pada suatu himpunan data empiris. Oleh karena itu, kesimpulan tidak menunjukkan bahwa satu algoritma pasti unggul pada seluruh persoalan. Penelitian berikutnya disarankan menguji kerangka evaluasi pada data dengan tingkat korelasi, ukuran sampel, dan pola nonlinear berbeda, serta melaporkan variasi hasil dan konsekuensi operasionalnya. Prinsip tersebut mendukung pengembangan prediksi multivariat yang lebih akurat sekaligus dapat diperiksa.

Implikasi metodologis dari pembandingan model juga perlu dibaca melalui konteks pengumpulan data. Dalam penelitian dengan pengamatan berulang, ukuran sampel efektif dapat lebih kecil daripada jumlah baris data. Karena itu, pembagian data yang mengikuti kelompok atau urutan waktu perlu diprioritaskan ketika hasil akhirnya dipakai untuk meramalkan entitas dan periode yang belum diamati. Kesalahan validasi yang tampak rendah tidak dapat menjadi dasar generalisasi jika informasi yang sama telah terbagi antara pelatihan dan pengujian. Pemeriksaan kesesuaian asumsi tersebut merupakan persyaratan sebelum mengambil keputusan tentang model paling unggul.

Keterbatasan kajian ini terletak pada tidak tersedianya pengujian empiris lintas dataset. Oleh sebab itu, kelebihan setiap algoritma tetap bersifat kondisional, tergantung struktur hubungan, jumlah prediktor, besar sampel, tingkat kebisingan, dan tujuan penggunaan. Penelitian berikutnya disarankan membangun perbandingan dengan data terbuka yang memiliki profil berbeda, misalnya data dengan multikolinearitas tinggi, pola nonlinear, serta perubahan distribusi antarkurun waktu. Seluruh eksperimen perlu menggunakan partisi yang konsisten dan melaporkan ketidakpastian skor agar perbedaan kecil antarmodel tidak diberi makna berlebihan.

Secara operasional, setiap pengembangan model perlu dilengkapi catatan definisi target, asal dan waktu tersedianya prediktor, keputusan prapemrosesan, serta versi perangkat lunak. Kriteria model terbaik dapat diperluas dari akurasi menjadi kombinasi tingkat kesalahan, stabilitas, waktu komputasi, kemudahan penjelasan, dan kesiapan pemeliharaan. Pendekatan demikian membantu peneliti memilih metode yang sepadan dengan kebutuhan dan sumber daya institusi. Dalam penelitian berisiko tinggi, keluaran prediksi juga memerlukan penelaahan manusia dan mekanisme koreksi apabila muncul ketidaksesuaian pada penggunaan nyata.

F. BIBLIOGRAPHY (DAFTAR PUSTAKA)

Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32. https://doi.org/10.1023/A:1010933404324

This Article is licensed under a CC BY-SA 4.0 International License

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of KDD, 785–794. https://doi.org/10.1145/2939672.2939785

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression evaluation. PeerJ Computer Science, 7, e623. https://doi.org/10.7717/peerj-cs.623

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. The Annals of Statistics, 29(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow (3rd ed.). O'Reilly Media.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). An introduction to statistical learning: With applications in Python. Springer. https://doi.org/10.1007/978-3-031-38747-0

Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. Springer. https://doi.org/10.1007/978-1-4614-6849-3

Kuhn, M., & Silge, J. (2022). Tidy modeling with R. O'Reilly Media. https://www.tmwr.org/

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.

Molnar, C. (2022). Interpretable machine learning (2nd ed.). https://christophm.github.io/interpretable-ml-book/

Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

Probst, P., Wright, M. N., & Boulesteix, A.-L. (2019). Hyperparameters and tuning strategies for random forest. WIREs Data Mining and Knowledge Discovery, 9(3), e1301. https://doi.org/10.1002/widm.1301

Roberts, D. R., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913–929. https://doi.org/10.1111/ecog.02881

Scikit-learn developers. (2025). Cross-validation: Evaluating estimator performance. https://scikit-learn.org/stable/modules/cross_validation.html

Scikit-learn developers. (2025). Common pitfalls and recommended practices. https://scikit-learn.org/stable/common_pitfalls.html

Steyerberg, E. W. (2019). Clinical prediction models (2nd ed.). Springer. https://doi.org/10.1007/978-3-030-16399-0

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B, 58(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Varoquaux, G., & Cheplygina, V. (2022). Machine learning for medical imaging: Methodological failures and recommendations for the future. NPJ Digital Medicine, 5, 48. https://doi.org/10.1038/s41746-022-00592-y

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B, 67(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

This Article is licensed under a CC BY-SA 4.0 International License