

Ensemble Learning untuk Klasifikasi: Tinjauan Komprehensif Metode, Aplikasi, dan Perkembangan Terkini

Muhammad Ibnu Choldun Rachmatullah¹

¹ D3 Manajemen Informatika, Universitas Logistik dan Bisnis Internasional

¹muhammadibnucholdun@ulbi.ac.id

Abstrak— Ensemble learning telah menjadi salah satu paradigma paling efektif dalam pembelajaran mesin, khususnya untuk permasalahan klasifikasi yang menuntut akurasi, ketahanan, dan kemampuan generalisasi yang tinggi. Dengan mengombinasikan beberapa pengklasifikasi dasar, metode ensemble mampu menurunkan varians, mengurangi bias, serta meningkatkan stabilitas prediksi dibandingkan pendekatan model tunggal. Dalam beberapa tahun terakhir, ensemble learning semakin mendapat perhatian seiring dengan kemajuan pada metode gradient boosting, deep learning, penanganan data tidak seimbang, serta integrasinya dalam kerangka kerja automated machine learning (AutoML). Artikel ini menyajikan tinjauan komprehensif mengenai ensemble learning untuk klasifikasi, mencakup landasan teoretis, strategi utama seperti bagging, boosting, voting, dan stacking, serta perkembangan mutakhir termasuk deep ensemble, ensemble berbasis efisiensi, dan optimasi berbasis diversitas. Selain itu, strategi evaluasi, pertimbangan praktis, serta tantangan riset terbuka juga dibahas. Artikel ini diharapkan dapat memberikan wawasan teoretis sekaligus panduan praktis bagi peneliti dan praktisi dalam merancang sistem klasifikasi berbasis ensemble learning.

Kata kunci— Ensemble learning, klasifikasi, bagging, boosting, stacking

Abstract— Ensemble learning has become one of the most effective paradigms in machine learning, particularly for classification problems that demand high accuracy, robustness, and generalizability. By combining multiple base classifiers, ensemble methods can reduce variance, mitigate bias, and improve prediction stability compared to single-model approaches. In recent years, ensemble learning has received increasing attention due to advances in gradient boosting, deep learning, imbalanced data handling, and their integration into automated machine learning (AutoML) frameworks. This article presents a comprehensive overview of ensemble learning for classification, covering theoretical foundations, key strategies such as bagging, boosting, voting, and stacking, as well as recent developments including deep ensembles, efficiency-based ensembles, and diversity-based optimization. Evaluation strategies, practical considerations, and open research challenges are also discussed. This article is expected to provide theoretical insights and practical guidance for researchers and practitioners in designing ensemble learning-based classification systems..

Keywords— Ensemble learning, classification, bagging, boosting, stacking

I. PENDAHULUAN

Klasifikasi merupakan salah satu tugas fundamental dalam pembelajaran mesin yang memiliki aplikasi luas di berbagai bidang, seperti diagnosis medis, analisis risiko keuangan, keamanan siber, penginderaan jauh, serta pengolahan bahasa alami. Tujuan utama dari klasifikasi adalah memetakan data masukan ke dalam kelas-kelas tertentu dengan tingkat akurasi dan reliabilitas yang tinggi. Namun, dalam praktiknya, model klasifikasi tunggal sering menghadapi keterbatasan, terutama yang berkaitan dengan trade-off bias-varians, sensitivitas terhadap noise, serta keterbatasan dalam menangkap kompleksitas dan heterogenitas data.

Ensemble learning muncul sebagai solusi yang efektif untuk mengatasi keterbatasan tersebut dengan menggabungkan beberapa pengklasifikasi ke dalam satu kerangka prediksi terpadu. Gagasan utama ensemble learning adalah bahwa sekumpulan model yang cukup akurat dan saling beragam dapat menghasilkan kinerja yang lebih baik dibandingkan model individu mana pun. Konsep ini telah dibuktikan baik secara teoretis maupun empiris pada berbagai permasalahan klasifikasi, sehingga menjadikan ensemble learning sebagai paradigma dominan dalam sistem pembelajaran mesin modern [1].

Dalam beberapa tahun terakhir, ensemble learning mengalami perkembangan pesat. Metode gradient boosting berbasis pohon keputusan telah mencapai kinerja unggul pada klasifikasi data tabular, sementara pendekatan deep ensemble terbukti meningkatkan akurasi dan estimasi ketidakpastian pada jaringan saraf dalam [2]. Selain itu, ensemble learning banyak diterapkan untuk menangani tantangan klasifikasi yang kompleks, seperti data tidak seimbang [3], klasifikasi lalu lintas jaringan [4], dan analisis sinyal biomedis [5]. Integrasi ensemble learning ke dalam sistem AutoML semakin menegaskan peran sentral pendekatan ini dalam pipeline klasifikasi modern [6].

Meskipun demikian, ensemble learning juga menghadirkan tantangan baru, antara lain meningkatnya biaya komputasi, berkurangnya interpretabilitas model, serta risiko kebocoran data apabila prosedur validasi tidak dirancang dengan benar. Oleh karena itu, diperlukan sebuah tinjauan pustaka yang menyeluruh untuk merangkum landasan teoretis, perkembangan metodologis, serta pertimbangan praktis dalam penerapan ensemble learning untuk klasifikasi. Artikel ini bertujuan untuk memenuhi kebutuhan tersebut dengan menyajikan ulasan bergaya jurnal IEEE.

II. LANDASAN TEORETIS ENSEMBLE LEARNING

Cara Secara formal, ensemble learning dapat didefinisikan sebagai paradigma pembelajaran di mana beberapa pengklasifikasi dasar dilatih dan digabungkan untuk menghasilkan keputusan klasifikasi akhir. Misalkan $h_1(x), h_2(x), \dots, h_M(x)$ menyatakan sekumpulan model dasar, di mana masing-masing model menghasilkan label kelas atau distribusi probabilitas kelas. Prediksi ensemble diperoleh melalui suatu fungsi agregasi yang mengombinasikan keluaran model-model tersebut.

Efektivitas ensemble learning erat kaitannya dengan dekomposisi bias-varians dari galat klasifikasi. Metode ensemble berbasis bagging terutama bertujuan untuk menurunkan varians dengan merata-ratakan prediksi dari model-model yang dilatih pada variasi data yang berbeda. Sebaliknya, metode boosting dirancang untuk menurunkan bias dengan membangun model secara berurutan yang secara eksplisit berfokus pada contoh-contoh yang sulit diklasifikasikan. Pendekatan stacking berupaya menyeimbangkan bias dan varians dengan melatih sebuah meta-pengklasifikasi yang mempelajari cara optimal menggabungkan keluaran model dasar [1].

Konsep diversitas merupakan elemen kunci dalam keberhasilan ensemble learning. Diversitas mengacu pada sejauh mana model-model dalam ensemble menghasilkan kesalahan yang berbeda terhadap sampel data yang sama. Jika kesalahan model sangat berkorelasi, maka keuntungan ensemble menjadi minimal. Oleh karena itu, berbagai strategi untuk meningkatkan diversitas, seperti resampling data, pemilihan subset fitur, variasi arsitektur model, serta perbedaan hiperparameter, menjadi fokus utama dalam desain ensemble modern [2]. Penelitian terkini bahkan mengusulkan metrik diversitas khusus untuk klasifikasi data tidak seimbang yang terbukti mampu meningkatkan kinerja deteksi kelas minoritas [7].

A. Metode Ensemble Berbasis Bagging

Bagging atau bootstrap aggregating merupakan salah satu teknik ensemble paling awal yang dikembangkan untuk meningkatkan stabilitas model dengan varians tinggi. Pada bagging, beberapa dataset pelatihan dibentuk melalui proses sampling dengan pengembalian dari dataset asli. Setiap pengklasifikasi dasar dilatih secara independen pada

satu dataset bootstrap, dan keputusan akhir diperoleh melalui mekanisme voting mayoritas.

Dalam konteks klasifikasi, bagging sangat efektif ketika digunakan bersama pengklasifikasi yang sensitif terhadap perubahan data, seperti decision tree. Random Forest merupakan contoh paling terkenal dari pendekatan bagging, di mana selain sampling data, dilakukan pula pemilihan subset fitur secara acak pada setiap node pohon. Strategi ini meningkatkan diversitas antar pohon dan menghasilkan kinerja generalisasi yang kuat pada berbagai permasalahan klasifikasi data tabular.

Meskipun bagging mampu meningkatkan stabilitas dan ketahanan terhadap noise, pendekatan ini tidak secara langsung mengurangi bias model. Oleh karena itu, efektivitas bagging menjadi terbatas ketika pengklasifikasi dasar memiliki bias yang tinggi. Kendati demikian, bagging tetap menjadi metode ensemble yang andal dan sering digunakan sebagai baseline pada banyak studi klasifikasi [1].

B. Metode Ensemble Berbasis Boosting

Boosting merupakan pendekatan ensemble yang membangun pengklasifikasi secara berurutan, di mana setiap model baru difokuskan untuk memperbaiki kesalahan model sebelumnya. Gradient boosting memformalisasikan konsep ini sebagai proses optimasi yang meminimalkan fungsi kerugian melalui pendekatan gradien.

Gradient boosting berbasis pohon keputusan telah menjadi metode unggulan untuk klasifikasi data tabular karena kemampuannya dalam memodelkan hubungan nonlinier dan interaksi antar fitur. Berbagai studi empiris menunjukkan bahwa metode ini secara konsisten menghasilkan kinerja terbaik pada berbagai dataset dunia nyata, menjadikannya standar de facto dalam aplikasi klasifikasi modern [1], [4].

Namun, metode boosting memiliki sensitivitas yang lebih tinggi terhadap noise dan kesalahan label jika tidak disertai regularisasi yang memadai. Overfitting dapat terjadi apabila kedalaman pohon terlalu besar atau laju pembelajaran tidak dikendalikan dengan baik. Oleh karena itu, proses validasi silang dan penalaan hiperparameter menjadi aspek krusial dalam penerapan boosting untuk klasifikasi.

C. Ensemble Berbasis Voting dan Averaging

Ensemble berbasis voting menggabungkan prediksi dari beberapa pengklasifikasi melalui mekanisme suara mayoritas atau perataan probabilitas. Pada hard voting, setiap model memberikan satu label kelas, sedangkan pada soft voting, probabilitas kelas dirata-ratakan sebelum menentukan kelas akhir.

Pendekatan voting relatif sederhana dan efisien secara komputasi dibandingkan metode ensemble yang lebih kompleks seperti stacking. Voting sering digunakan untuk menggabungkan pengklasifikasi heterogen, misalnya regresi logistik, support vector machine, dan model berbasis pohon. Meskipun tidak secara eksplisit

mempelajari cara optimal menggabungkan model, voting dapat memberikan peningkatan kinerja yang signifikan ketika model-model dasar cukup beragam dan terkalibrasi dengan baik [8].

D. Metode Ensemble Berbasis Stacking

Stacking atau stacked generalization merupakan metode ensemble tingkat lanjut yang melibatkan pelatihan sebuah meta-pengklasifikasi untuk mengombinasikan keluaran dari beberapa pengklasifikasi dasar. Meta-pengklasifikasi ini mempelajari bobot dan pola kombinasi yang optimal berdasarkan performa model dasar.

Aspek krusial dalam stacking adalah penggunaan prediksi out-of-fold untuk melatih meta-pengklasifikasi. Pendekatan ini mencegah kebocoran data dan mengurangi risiko overfitting. Jika dirancang dengan benar, stacking sering kali menghasilkan kinerja yang lebih baik dibandingkan metode ensemble yang lebih sederhana, terutama pada permasalahan klasifikasi yang kompleks [9].

Penelitian terkini juga mengusulkan pendekatan stacking hibrida yang mengintegrasikan teknik optimasi seperti simulated annealing atau algoritma genetika untuk meningkatkan kinerja klasifikasi lebih lanjut [10].

E. Deep Ensemble Learning

Stacking atau stacked generalization merupakan metode ensemble tingkat lanjut yang melibatkan pelatihan sebuah meta-pengklasifikasi untuk mengombinasikan keluaran dari beberapa pengklasifikasi dasar. Meta-pengklasifikasi ini mempelajari bobot dan pola kombinasi yang optimal berdasarkan performa model dasar

III. ENSEMBLE LEARNING UNTUK KLASIFIKASI DATA TIDAK SEIMBANG

Klasifikasi data tidak seimbang, di mana satu atau lebih kelas memiliki jumlah sampel yang jauh lebih sedikit dibandingkan kelas lainnya, merupakan tantangan besar dalam banyak aplikasi nyata. Dalam kondisi ini, model klasifikasi cenderung bias terhadap kelas mayoritas, sehingga metrik akurasi menjadi kurang representatif.

Ensemble learning banyak digunakan untuk mengatasi permasalahan ini melalui penggabungan dengan teknik resampling, pembobotan kelas, serta desain ensemble yang mempertimbangkan diversitas. Survei terbaru menunjukkan bahwa pendekatan berbasis ensemble merupakan salah satu solusi paling efektif untuk klasifikasi data tidak seimbang, terutama ketika dikombinasikan dengan metrik evaluasi yang sesuai seperti macro-F1 dan precision-recall AUC [3]. Pendekatan ensemble yang memanfaatkan metrik diversitas khusus juga terbukti meningkatkan deteksi kelas minoritas secara signifikan [7].

IV. KESIMPULAN

Ensemble learning merupakan paradigma kunci dalam pengembangan sistem klasifikasi modern yang akurat dan andal. Melalui pendekatan bagging, boosting, voting, dan stacking, ensemble learning mampu mengatasi berbagai keterbatasan model tunggal. Perkembangan terbaru dalam

deep ensemble, klasifikasi data tidak seimbang, dan AutoML semakin memperkuat relevansi ensemble learning dalam penelitian dan aplikasi praktis. Dengan terus berkembangnya riset terkait efisiensi dan interpretabilitas, ensemble learning diperkirakan akan tetap menjadi fondasi utama dalam sistem klasifikasi berbasis pembelajaran mesin.

REFERENSI

- [1] I. D. Mienye dan Y. Sun, "A survey of ensemble learning: Concepts, algorithms, applications, and prospects," *IEEE Access*, vol. 10, hlm. 99129–99149, 2022.
- [2] A. Mohammed dan R. Kora, "A comprehensive review on ensemble deep learning," *Journal of King Saud University – Computer and Information Sciences*, vol. 35, no. 2, hlm. 757–774, 2023.
- [3] W. Chen, Y. Liu, dan Y. Zhang, "A survey on imbalanced learning," *Artificial Intelligence Review*, 2024.
- [4] G. Aceto dkk., "AI-powered internet traffic classification," *IEEE Communications Magazine*, 2024.
- [5] A. Karim, S. Ryu, dan I. C. Jeong, "Ensemble learning for biomedical signal classification," *Scientific Reports*, vol. 15, 2025.
- [6] E. LeDell dan S. Poirier, "H2O AutoML," *Proceedings of the AutoML Workshop (ICML)*, 2020.
- [7] Y. Pang dkk., "Imbalanced ensemble learning leveraging data-driven diversity metrics," *Pattern Recognition*, 2025.
- [8] T.-M. Chiu dkk., "Deep ensemble learning for multiclass skin lesion classification," *Bioengineering*, vol. 12, no. 9, 2025.
- [9] L. Liu, X. Zhang, dan J. Chen, "Hybrid stacking ensemble algorithm," *Expert Systems with Applications*, 2024.
- [10] A. Azedou dkk., "Genetic algorithm optimization of ensemble learning," *Knowledge-Based Systems*, 2025.
- [11] U.-J. Baek dkk., "Deep ensemble learning with differential model selection," *Sensors*, vol. 25, no. 9, 2025.