

Analisis Ketidakseimbangan Tegangan Baterai dengan Pendekatan Random Forest, K-Nearest Neighbors untuk Prediksi Balancing Charger

Irwan Novianto^{1,*}, Septian Rico Hernawan²

¹ Fakultas Teknologi Informasi, Teknik Elektro, Universitas Nahdlatul Ulama Yogyakarta, Yogyakarta, Indonesia

² Fakultas Teknologi Informasi, Teknik Komputer, Universitas Nahdlatul Ulama Yogyakarta, Yogyakarta, Indonesia

Email: ^{1,*}irwannovianto@unu-jogja.ac.id, ²rico@unu-jogja.ac.id

Email Penulis Korespondensi: irwannovianto@unu-jogja.ac.id

Submitted: 27/07/2025; Accepted: 13/08/2025; Published: 31/08/2025

Abstrak—Ketidakseimbangan tegangan antarsel baterai menurunkan efisiensi, mempercepat degradasi, dan menaikkan risiko kegagalan sistem penyimpanan energi listrik. Penelitian ini bertujuan memodelkan serta memprediksi kondisi *balancing charger* menggunakan dua algoritma pembelajaran mesin, *Random Forest* (RF) dan *K-Nearest Neighbors* (KNN), pada kondisi sel baterai berisi 4, 8, 10, hingga 15 sel dengan lima skala dataset (1000, 5000, 10000, 15000, 20000 sampel). Data tegangan diperoleh melalui simulasi dan pengukuran laboratorium pada baterai lithium-ion dengan rentang 3,2–4,2 V, kemudian dinormalisasi dan dibagi menjadi data pelatihan serta pengujian. Kinerja dievaluasi memakai akurasi, *confusion matrix*, dan *feature importance*. Hasil menunjukkan RF mencapai akurasi 0,98 pada 4 sel dan tetap tinggi 0,93 pada 15 sel, sedangkan KNN hanya 0,94 dan 0,37 pada konfigurasi serupa. RF memperlihatkan prediksi yang terpusat pada diagonal *confusion matrix* dan bobot fitur merata, menandakan adaptabilitas terhadap kenaikan dimensi. Kontribusi penelitian ini adalah: (1) kerangka evaluasi untuk membandingkan algoritma klasifikasi pada skenario multi-sel; (2) bukti empiris skalabilitas RF untuk mendeteksi kondisi balancing dengan data tegangan tunggal; dan (3) implikasi praktis bagi BMS, yaitu pengambilan keputusan *balancing* yang lebih akurat, prioritas sel yang bermasalah, pengurangan siklus penyeimbangan sia-sia, dan potensi penghematan energi serta perpanjangan umur pakai. Temuan ini merekomendasikan RF sebagai algoritma inti pada sistem *balancing charger* berbasis *machine learning*, khususnya untuk implementasi nyata pada perangkat tepi berdaya terbatas.

Kata Kunci: Baterai; *Balancing charger*; *Random Forest*; *K-Nearest Neighbors*; *Machine Learning*; RF; KNN

Abstract—Inter-cell voltage imbalance degrades efficiency, accelerates aging, and increases failure risk in electrochemical energy storage systems. This study models and predicts balancing-charger conditions using two machine-learning algorithms Random Forest (RF) and K-Nearest Neighbors (KNN) across packs of 4, 8, 10, and 15 cells with five dataset scales (1,000; 5,000; 10,000; 15,000; and 20,000 samples). Voltage data were obtained through simulation and laboratory measurements on lithium-ion cells within 3.2–4.2 V, then normalized and split into training and testing sets. Performance was evaluated using accuracy, confusion matrices, and feature-importance analysis. Results show RF achieves 0.98 accuracy for 4-cell packs and remains high at 0.93 for 15-cell packs, whereas KNN attains only 0.94 and 0.37 on the same configurations. RF exhibits predictions concentrated along the confusion-matrix diagonal with well-distributed feature weights, indicating robustness to increasing dimensionality. The contributions are threefold: (1) an evaluation framework for comparing classifiers in multi-cell scenarios; (2) empirical evidence of RF's scalability for detecting balancing conditions from single-voltage inputs; and (3) practical implications for BMS operation more accurate balancing decisions, prioritization of problematic cells, reduced futile equalization cycles, and potential energy savings together with extended service life. These findings recommend RF as a core algorithm for machine-learning-based balancing chargers, particularly for real-world deployment on power-constrained edge devices.

Keywords: Battery; Balancing charger; Random Forest; K-Nearest Neighbors; Machine Learning; RF; KNN

1. PENDAHULUAN

Pertumbuhan kendaraan listrik di Indonesia saat ini mengalami peningkatan yang cukup signifikan, berdasarkan data dari Rencana Usaha Penyediaan Tenaga Listrik (RUPTL) tahun 2025 sampai dengan tahun 2034 yang disusun oleh PT PLN yang dipublikasikan pada Kementerian Energi dan Sumber Daya Mineral (KESDM) mencatat pertumbuhan jumlah mobil listrik di Indonesia khususnya kenaikan EV pada kendaraan roda 4 ini mencapai 54.856% dalam lima tahun terakhir [1]. Seiring dengan meningkatnya jumlah kendaraan listrik yang beroperasi di Indonesia, permintaan terhadap sistem baterai berkualitas tinggi khususnya baterai lithium-ion, menjadi semakin penting. Dimana salah satu bagian dari EV adalah baterai, yang berperan penting, sebagai sumber utama penyimpanan dan penyuplai energi bagi kendaraan. Keandalan baterai sangat menentukan performa, jangkauan, serta keamanan kendaraan listrik, sehingga tantangan dalam pengelolaan dan perawatan baterai pun semakin kompleks. Baterai memainkan peranan penting sebagai media penyimpan energi [2]. Salah satu tantangan dalam penggunaan baterai adalah memastikan keseimbangan tegangan antar sel baterai agar kinerja baterai tetap optimal [3], efisien, dan memiliki umur pakai yang panjang.

Balancing baterai adalah proses penyamaan tegangan antar sel baterai yang berbeda, dari akibat berbagai faktor seperti variasi produksi, perbedaan suhu operasional, dan usia baterai yang berbeda-beda [4]. Ketidakseimbangan tegangan antar sel baterai dapat menyebabkan kerusakan dini [5], menurunkan efisiensi sistem, dan bahkan menimbulkan risiko keamanan seperti overheating atau kegagalan sistem baterai secara menyeluruh [6]. Ketidakseimbangan tegangan baterai merupakan masalah yang signifikan dalam sistem

penyimpanan energi, terutama pada aplikasi kendaraan listrik dan perangkat lain berbasis baterai lainnya. Selain itu Teknik *balancing charger* menjadi solusi penting dalam mengatasi permasalahan ini, salah satu caranya dengan *machine learning* (ML) yang kini semakin banyak digunakan sebagai metode untuk memprediksi kondisi *balancing* baterai secara lebih presisi. Salah satu Pendekatan yang efektif untuk memprediksi keadaan ini dapat diimplementasikan menggunakan algoritma pembelajaran mesin atau *machine learning* (ML) seperti *Random Forest* (RF) dan *K-Nearest Neighbors* (KNN). Penelitian terbaru menunjukkan bahwa kedua metode ini memiliki potensi besar dalam mengelola dan memantau kesehatan serta kinerja baterai secara efisien.

Random Forest merupakan teknik pembelajaran mesin yang dapat digunakan untuk klasifikasi dan regresi dengan membangun banyak pohon keputusan dan menggabungkan hasilnya untuk meningkatkan akurasi dan mencegah *overfitting* [7]. Dengan menyusun berbagai pohon, *Random Forest* (RF) dapat menangkap hubungan non-linear dalam data yang berkaitan dengan tegangan dan keadaan baterai lainnya. Dalam konteks manajemen baterai, dengan pembelajaran mesin menawarkan metodologi yang lebih tepat untuk memprediksi keadaan pengisian (*state of charge*, SoC) serta keadaan kesehatan (*state of health*, SoH) dari baterai [8]. Sementara itu, algoritma *K-Nearest Neighbors* (KNN) bekerja dengan prinsip kedekatan jarak antar data, mengklasifikasikan data baru berdasarkan mayoritas kelas data terdekatnya, di mana *K-Nearest Neighbors* (KNN) adalah algoritma yang cukup sederhana namun efektif, yang mengandalkan kedekatan antara data untuk membuat prediksi. Ini dapat berfungsi sebagai metode pendukung dalam mengestimasi kondisi baterai [9]. *Random Forest* (RF) dan *K-Nearest Neighbors* (KNN), memiliki kelebihan masing-masing. *Random Forest* cenderung lebih robust dan efektif dalam mengolah dataset besar dengan banyak variabel, sementara *K-Nearest Neighbors* (KNN) mudah diterapkan dan dipahami, menjadikannya pilihan yang baik untuk aplikasi yang membutuhkan interpretasi cepat [10].

Dalam berbagai studi penelitian terdahulu, *Random Forest* (RF) telah terbukti efektif dalam memprediksi sifat-sifat baterai, seperti daya tahan dan kapasitas [11]. Dari segi inovasi di dalam pengelolaan manajemen baterai, *Random Forest* (RF) telah digunakan untuk memperkirakan masa dari umur pakai baterai dengan akurasi yang cukup tinggi, serta memungkinkan manajemen baterai untuk merencanakan perawatan dan penggantian lebih efisien [12]. Penelitian juga menunjukkan bahwa penggunaan *Random Forest* (RF) dalam prediksi biaya perawatan dan pengoperasian baterai dapat mengurangi secara signifikan kesalahan estimasi yang sering terjadi di lingkungan industri [13].

Penelitian lain juga menunjukkan penggunaan *Random Survival Forest* dalam memperkirakan degradasi baterai berdasarkan kondisi penggunaan aktual [14]. Dengan pola pembacaan untuk pengguna dan parameter lingkungan, dimana model RSF ini memungkinkan pemantauan terhadap perubahan tegangan yang jauh lebih akurat dan cukup dapat diandalkan. Selain itu ketersediaan data juga dapat memainkan peran kunci dalam efektivitas model *Random Forest* (RF). Variabel-variabel seperti kecepatan pengisian, temperatur, dan jenis penggunaan juga dapat memberikan hasil yang baik saat akan digunakan dalam analisis [11]. Penelitian tersebut menunjukkan bahwa mengoptimalkan pilihan pada variabel masukan tidak hanya meningkatkan akurasi akan tetapi juga memungkinkan informasi yang lebih mendalam tentang kondisi kesehatan baterai yang ada [15].

Dalam banyak kasus, data monitoring pada jangka panjang terbukti akan memberikan sinyal dini tentang masalah mendatang, serta memudahkan tindakan preventif untuk menjaga performa [16]. Sementara itu, penelitian dengan metode *K-Nearest Neighbors* (KNN) menjadi alat yang sangat baik untuk aplikasi yang memerlukan kecepatan dan respons rendah terhadap perubahan data. Dalam skenario di mana jumlah data terbatas atau variabilitas tidak terlalu kompleks, *K-Nearest Neighbors* (KNN) sering memberikan hasil yang dapat diterima tanpa proses pelatihan yang berat [17]. Namun, terdapat tantangan yang dihadapi dalam metode *K-Nearest Neighbors* (KNN) yaitu sensitivitasnya terhadap outlier, yang dapat menyebabkan hasil prediksi yang kurang terlalu akurat [18]. Pada penelitian lain, algoritma *Random Forest* (RF) juga telah terbukti unggul dibandingkan KNN dalam berbagai skenario klasifikasi *balancing* baterai, terutama ketika jumlah fitur meningkat secara signifikan [19]. Namun, eksplorasi mendalam tentang pengaruh variasi jumlah baterai dan dataset secara sistematis masih terbatas, khususnya pada implementasi *balancing* baterai berbasis ML dengan konfigurasi yang bervariasi secara kompleks.

Berdasarkan uraian kajian sebelumnya terdapat permasalahan utama yaitu belum adanya kajian mendalam terkait pengaruh variasi jumlah baterai terhadap kinerja algoritma *Random Forest* (RF) dan *K-Nearest Neighbors* (KNN) dalam prediksi *balancing charger*, serta kurangnya analisis adaptabilitas algoritma tersebut terhadap jumlah fitur dan ukuran dataset yang beragam. Dimana penelitian terbaru menunjukkan bahwa *Random Forest* (RF) unggul dibanding *K-Nearest Neighbors* (KNN) pada konfigurasi terbatas, namun belum banyak penelitian yang mengeksplorasi pengaruh variasi jumlah sel baterai secara sistematis dari jumlah kecil hingga besar secara komprehensif. Oleh karena itu, penelitian ini menawarkan eksplorasi yang lebih luas dan mendalam terhadap kedua algoritma tersebut.

Tujuan penelitian ini untuk mengevaluasi dan membandingkan secara komprehensif kinerja algoritma *Random Forest* dan *K-Nearest Neighbors* dalam berbagai konfigurasi jumlah baterai dan variasi dataset, serta merekomendasikan algoritma yang paling efektif dan adaptif untuk aplikasi *balancing charger* berbasis *machine learning*. Dengan penggabungan kedua metode ini *Random Forest* dan *K-Nearest Neighbors* dapat menciptakan sistem prediktif yang efektif untuk manajemen kesehatan baterai melalui pendekatan *ensemble*. Dalam praktiknya, prediksi ketidakseimbangan tegangan baterai hanya dapat dicapai dengan mengandalkan model-

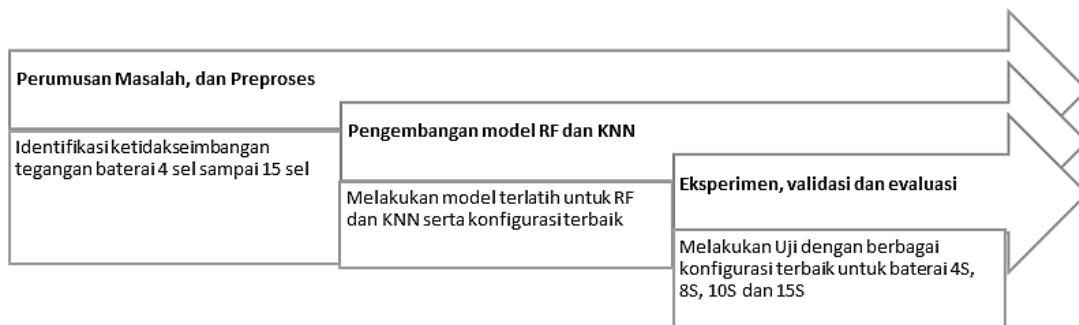
model yang mampu mengelola data dinamis, menciptakan database terpadu yang dapat mendeteksi pola dan tren, dan memungkinkan pembaruan model yang berkelanjutan seiring pertumbuhan data [10], [20].

Urgensi penelitian ini penting untuk dilakukan mengingat meningkatnya kebutuhan terhadap sistem manajemen baterai yang handal dan efisien, terutama dengan pesatnya perkembangan kendaraan listrik dan energi baru terbarukan. Optimasi penggunaan algoritma ML dalam sistem *balancing* baterai sangat dibutuhkan agar dapat meningkatkan efisiensi operasional baterai, serta memperpanjang umur pakai baterai dan mengurangi risiko kerusakan sistem akibat ketidakseimbangan tegangan antar sel baterai. Secara keseluruhan, ketidakseimbangan tegangan baterai dapat difasilitasi dengan menggunakan metode pembelajaran mesin yang tangguh seperti *Random Forest* dan *K-Nearest Neighbors*, yang bersama-sama berkontribusi dalam meningkatkan manajemen dan pengoperasian dari sistem penyimpanan energi ini. Evolusi teknik dan metodologi dalam pembelajaran mesin menunjukkan harapan yang signifikan untuk sistem manajemen kesehatan baterai yang lebih cerdas dan efisien di masa depan.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Pada penelitian ini tahapan penelitian ini dirancang untuk mengevaluasi secara komprehensif kinerja *Random Forest* (RF), dan *K-Nearest Neighbors* (KNN) pada prediksi *balancing charger* dengan variasi jumlah sel baterai dan ukuran data. Adapun tahapan seperti pada gambar 1.



Gambar 1. Tahapan penelitian

1. Perumusan Masalah dan Preproses.

Langkah pertama dilakukan identifikasi terhadap ketidakseimbangan tegangan antarsel pada baterai *lithium-ion* pada berbagai konfigurasi paket. Dari temuan awal ini, peneliti merumuskan variabel penelitian, meliputi tegangan tiap sel, riwayat siklus, serta menetapkan hipotesis hubungan variabel-variabel tersebut dengan status “*balanced*” atau “*unbalanced*” untuk baterai tipe 4 sel, 8 sel, 10 sel dan 15 sel.

2. Pengembangan model RF dan KNN.

Pada tahap ini dilakukan Model Terlatih *Random Forest* (RF) dan *K-Nearest Neighbors* (KNN), fokus utama adalah mengembangkan model pembelajaran mesin berdasarkan dataset hasil praproses, serta mengoptimalkan konfigurasi parameter agar performa prediksi mencapai hasil terbaik. Tahapan ini dimulai dengan membagi data menjadi data latih dan data uji secara *stratified*, memastikan proporsi kelas “*balanced*” dan “*unbalanced*” tetap seimbang di setiap subset. Selanjutnya, dilakukan proses pelatihan (training) menggunakan dua algoritma utama: *Random Forest* (RF) dan *K-Nearest Neighbors* (KNN).

Hasil dari tahapan ini adalah dua model terlatih *Random Forest* (RF) dan *K-Nearest Neighbors* (KNN) yang telah diuji pada berbagai konfigurasi jumlah baterai dan skala dataset. Model-model ini kemudian digunakan dalam tahap berikutnya untuk pengujian performa, analisis sensitivitas, serta pemilihan metode yang paling adaptif untuk sistem prediksi *balancing* baterai secara otomatis berbasis *machine learning*. Tahapan ini krusial dalam memastikan sistem prediktif mampu bekerja secara akurat, efisien, dan andal di lingkungan nyata.

3. Eksperimen, validasi dan evaluasi.

Pada tahap ini dilakukan uji terhadap algoritma *Random Forest* (RF) dan *K-Nearest Neighbors* (KNN) diuji secara sistematis dalam simulasi *balancing charger* baterai dengan empat konfigurasi jumlah sel baterai, yaitu 4, 8, 10, dan 15 sel. Setiap konfigurasi diuji menggunakan variasi jumlah dataset yang berbeda, yaitu 1.000, 5.000, 10.000, dan 20.000 sampel. Tujuan dari pengujian ini adalah mengevaluasi kinerja kedua algoritma dalam mengenali pola ketidakseimbangan tegangan antarsel, serta mengukur tingkat adaptabilitas algoritma terhadap kompleksitas sistem dan skala data yang meningkat. Kemudian ditahap terakhir dilakukan evaluasi dan analisis terhadap data akurasi uji.

2.2 Random Forest (RF)

Random Forest (RF) merupakan salah satu metode *ensemble learning* yang digunakan untuk menyelesaikan masalah klasifikasi maupun regresi dengan menggabungkan prediksi dari beberapa pohon keputusan (*decision trees*) secara paralel. Algoritma ini diperkenalkan oleh Leo Breiman [21] sebagai pengembangan dari metode *bagging* (bootstrap aggregating) yang disertai dengan proses pemilihan fitur secara acak dalam setiap pemisahan cabang pohon. *Random Forest* (RF) bekerja dengan membangun sejumlah besar pohon keputusan dari sampel acak dataset (dengan pengembalian) dan subset acak dari fitur. Hasil prediksi akhir didapat dari mayoritas suara (*majority voting*) pada klasifikasi atau rata-rata (*mean*) hasil pada regresi. Persamaan *Random Forest* (RF) sebagai berikut:

$$\hat{y}(x) = \arg \max_{c \in C} \left\{ \frac{1}{B} \sum_{b=1}^B 1(T_b(x) = c) \right\} \quad (1)$$

Berdasarkan fungsi persamaan 1, Dimana untuk B adalah Jumlah pohon dalam hutan, kemudian untuk $T_b(x)$ adalah Prediksi pohon ke-b untuk input x, sedangkan untuk C adalah kelas ke-c.

Fungsi pada persamaan 1 digunakan untuk menghitung proporsi pohon yang memprediksi kelas tertentu c, dan memilih kelas dengan jumlah suara terbanyak. Kemudian proses *majority voting* yang membuat *Random Forest* (RF) lebih stabil dibanding satu pohon saja.

2.3 K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) merupakan salah satu algoritma pembelajaran mesin non-parametrik yang digunakan untuk tugas klasifikasi maupun regresi. *K-Nearest Neighbors* (KNN) bekerja berdasarkan prinsip data baru akan diklasifikasikan atau diprediksi berdasarkan kesamaan (kedekatan jarak) dengan data-data tetangga terdekat yang telah diberi label sebelumnya. Algoritma ini diperkenalkan oleh Cover dan Hart [22], dan sangat bergantung pada metrik jarak antar titik data dalam ruang fitur. Dalam konteks klasifikasi, *K-Nearest Neighbors* (KNN) mengidentifikasi k tetangga terdekat dari suatu data uji, lalu menentukan kelas mayoritas dari tetangga-tetangga tersebut sebagai hasil prediksi. Adapun formula dari *K-Nearest Neighbors* (KNN) sebagai berikut:

$$\hat{y}(x) = \arg \max_{c \in C} \left\{ \sum_{i \in N_k(x)} 1(y_i = c) \right\} \quad (2)$$

Berdasarkan fungsi persamaan 2, dimana untuk C adalah himpunan kelas, kemudian untuk $N_k(x)$ adalah himpunan k tetangga terdekat dari uji x, sedangkan untuk y_i adalah label kelas dari data latih ke-i.

Persamaan 2 merupakan formula untuk memprediksi dimana dilakukan dengan mengambil kelas mayoritas dari tetangga-tetangga tersebut. kelebihan dari *K-Nearest Neighbors* (KNN) ini Non-parametrik dimana tidak membuat asumsi awal terhadap distribusi data. Kemudian *Lazy learner* dimana tidak melakukan proses pelatihan eksplisit; perhitungan dilakukan saat inferensi. sehingga cocok untuk data kecil hingga menengah serta cocok digunakan sebagai baseline awal sebelum menggunakan model yang lebih kompleks.

3. HASIL DAN PEMBAHASAN

Penelitian ini di laksanakan untuk mengevaluasi performa algoritma *Random Forest* (RF) dan *K-Nearest Neighbors* (KNN) dalam mendeteksi kondisi *balancing charger* baterai *lithium-ion* berdasarkan tegangan antar sel baterai dengan konfigurasi 4S, 8S, 10S, dan 15S. Setiap konfigurasi diuji menggunakan variasi jumlah data pelatihan, yaitu 1000, 5000, 10000, 15000, dan 20000 sampel data seperti pada Tabel 1. Data tegangan listrik baterai diperoleh melalui simulasi dan pengukuran di laboratorium, dimana untuk sampel tegangan baterai yang digunakan adalah baterai jenis *lithium-ion* dengan rentang tegangan baterai untuk dataset dari 3,2 volt sampai dengan 4,2 volt. Pendekatan ini bertujuan untuk memahami bagaimana perubahan skala data dan kompleksitas jumlah sel baterai memengaruhi akurasi klasifikasi kedua algoritma, serta mengevaluasi sejauh mana mereka mampu beradaptasi terhadap peningkatan dimensi fitur dan volume data.

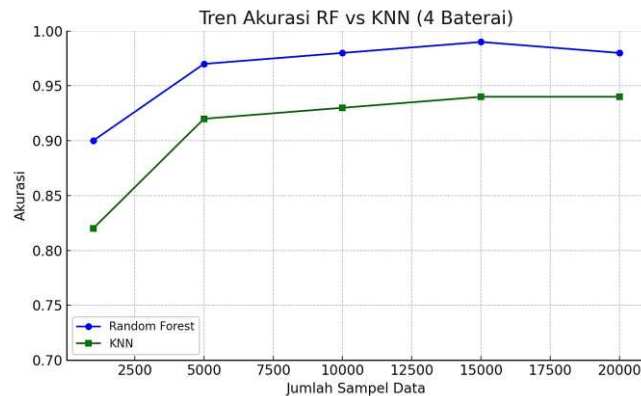
Tabel 1. Distribusi dataset penelitian

Skenario	Sel Baterai	Uji Sampel dataset
S1	4	1.000, 5.000, 10.000, dan 20.000
S2	8	1.000, 5.000, 10.000, dan 20.000
S3	10	1.000, 5.000, 10.000, dan 20.000
S4	15	1.000, 5.000, 10.000, dan 20.000

3.1 Pengujian konfigurasi baterai 4S.

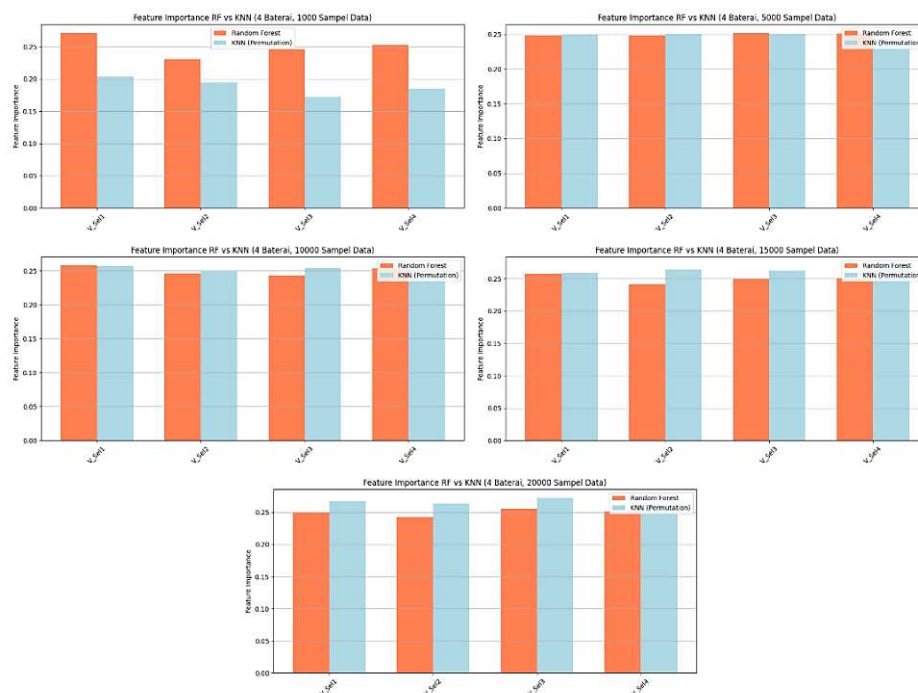
Hasil pengujian untuk konfigurasi baterai 4S, algoritma *Random Forest* (RF) menunjukkan performa superior pada semua ukuran dataset. Dengan dataset kecil berisi 1000 sampel, *Random Forest* (RF) memperoleh akurasi sebesar 0,90 sedangkan algoritma *K-Nearest Neighbors* (KNN) hanya mencapai 0,82. Peningkatan jumlah data berdampak positif bagi keduanya, namun *Random Forest* (RF) tetap mempertahankan keunggulannya. Pada 20000 sampel, *Random Forest* (RF) mencatat akurasi 0,98 sementara *K-Nearest Neighbors* (KNN) mencapai

0,94. Perbedaan akurasi ini menunjukkan keunggulan *Random Forest* (RF) dalam mengenali pola-pola yang kompleks, bahkan dalam dataset sederhana. Hasil pengujian di tunjukkan pada gambar 2.



Gambar 2. Tren akurasi *Random Forest* dan *K-Nearest Neighbors* pada dataset simulasi balancing 4 baterai.

Pada *confusion matrix* menunjukkan pola prediksi yang konsisten. *Confusion matrix Random Forest* (RF) didominasi oleh nilai diagonal, yang menunjukkan prediksi yang tepat pada masing-masing kelas target balancing, dengan sedikit kesalahan prediksi. Sebaliknya, *confusion matrix K-Nearest Neighbors* (KNN) menunjukkan jumlah kesalahan prediksi yang relatif lebih tinggi, terutama pada dataset kecil, meskipun jumlah kesalahan ini berkurang pada dataset yang lebih besar.



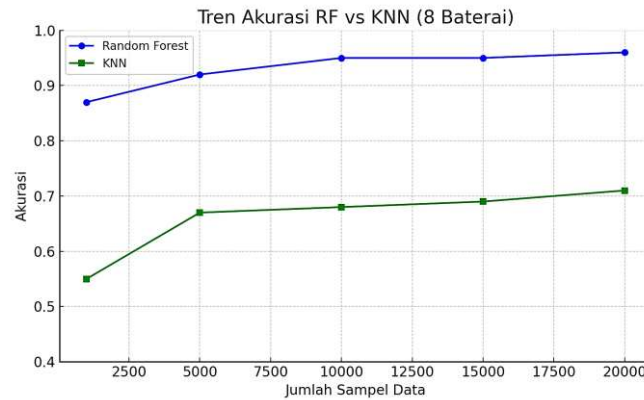
Gambar 3. *Feature importance* algoritma *Random Forest* dan *K-Nearest Neighbors* untuk konfigurasi 4 baterai (4S).

Selain itu, analisis *feature importance* pada gambar 3 memberikan gambaran mengenai perilaku kedua algoritma dalam memanfaatkan informasi dari fitur input. Pada dataset kecil, *Random Forest* (RF) mengindikasikan dominasi fitur V_Sel1 sebagai faktor penentu utama, sedangkan fitur lainnya berkontribusi lebih rendah. Namun, seiring bertambahnya jumlah data, distribusi *feature importance* pada *Random Forest* (RF) menjadi semakin merata, yang menunjukkan adaptabilitas *Random Forest* (RF) dalam memahami seluruh fitur secara proporsional. Sementara itu, *K-Nearest Neighbors* (KNN) menunjukkan pola *feature importance* yang cenderung rata pada semua skenario, yang sejalan dengan sifat algoritmanya yang berbasis jarak dan tidak memprioritaskan fitur tertentu.

3.2 Pengujian konfigurasi baterai 8S.

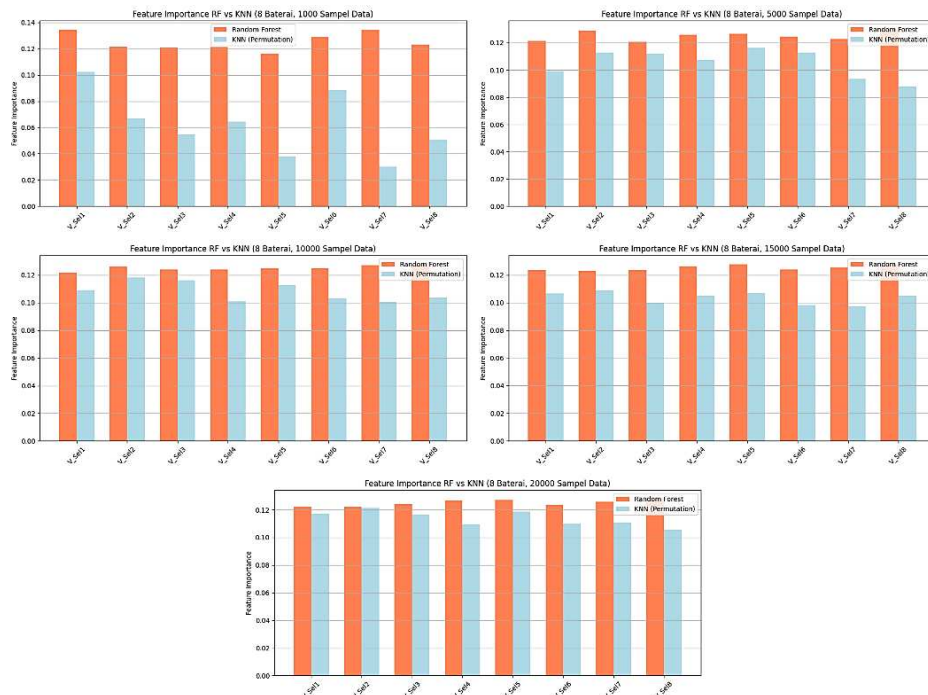
Hasil pengujian untuk konfigurasi baterai 8S, kompleksitas data bertambah dua kali lipat, menyebabkan tantangan lebih besar bagi algoritma. Pada algoritma *Random Forest* (RF) menunjukkan performa yang

konsisten dan signifikan lebih baik dibandingkan *K-Nearest Neighbors* (KNN). Pada dataset 1000 sampel, akurasi *Random Forest* (RF) tercatat 0,87 sedangkan *K-Nearest Neighbors* (KNN) hanya mencapai 0,55. Ketika dataset bertambah besar, akurasi *Random Forest* (RF) meningkat menjadi 0,96 pada 20000 sampel, sedangkan *K-Nearest Neighbors* (KNN) hanya mencapai 0,71. Hal ini menunjukkan kemampuan *Random Forest* (RF) dalam mempelajari pola distribusi tegangan antar sel baterai meskipun jumlah fitur meningkat, sedangkan *K-Nearest Neighbors* (KNN) mengalami degradasi performa akibat sensitivitas terhadap dimensionalitas tinggi. Hasil pengujian di tunjukkan pada gambar 4.



Gambar 4. Tren akurasi *Random Forest* dan *K-Nearest Neighbors* pada dataset simulasi balancing 8 baterai.

Pada *confusion matrix* memperlihatkan pola distribusi prediksi yang mendukung hasil akurasi. *Random Forest* (RF) menunjukkan dominasi prediksi yang benar pada diagonal matriks, sementara *K-Nearest Neighbors* (KNN) menunjukkan distribusi prediksi yang tersebar di luar diagonal, terutama pada dataset kecil, yang mengindikasikan jumlah *misclassifications* yang cukup besar.



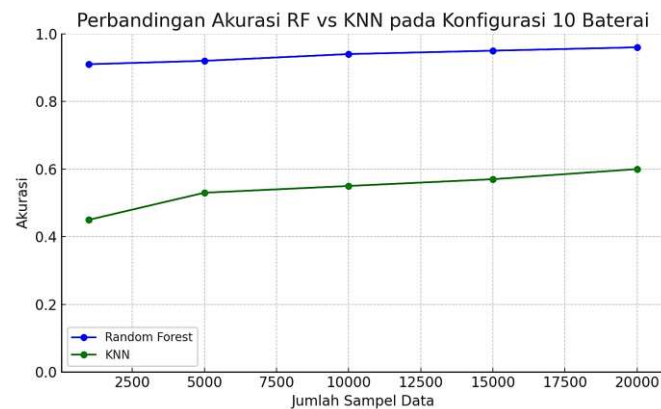
Gambar 5. Feature importance algoritma *Random Forest* dan *K-Nearest Neighbors* untuk konfigurasi 8 baterai (8S).

Selain itu untuk analisis *feature importance* pada gambar 5 memperlihatkan pola yang menarik. Pada dataset kecil (1000 sampel), *feature importance Random Forest* (RF) menunjukkan variasi antar fitur, dengan fitur V_Sel1 sebagai fitur dominan (0,13) sementara fitur lain memiliki nilai yang sedikit lebih rendah. Namun, ketika dataset bertambah besar, distribusi *feature importance Random Forest* (RF) menjadi semakin merata (0,12–0,13), menunjukkan adaptabilitas *Random Forest* (RF) dalam memanfaatkan seluruh fitur secara optimal seiring pertambahan data. Sebaliknya, *feature importance K-Nearest Neighbors* (KNN) menunjukkan distribusi yang jauh lebih rendah (0,03–0,10) dan cenderung tidak variatif pada dataset kecil. Meski distribusi ini

menjadi sedikit lebih merata pada dataset besar, nilai keseluruhan *feature importance* *K-Nearest Neighbors* (KNN) tetap lebih rendah dibandingkan RF.

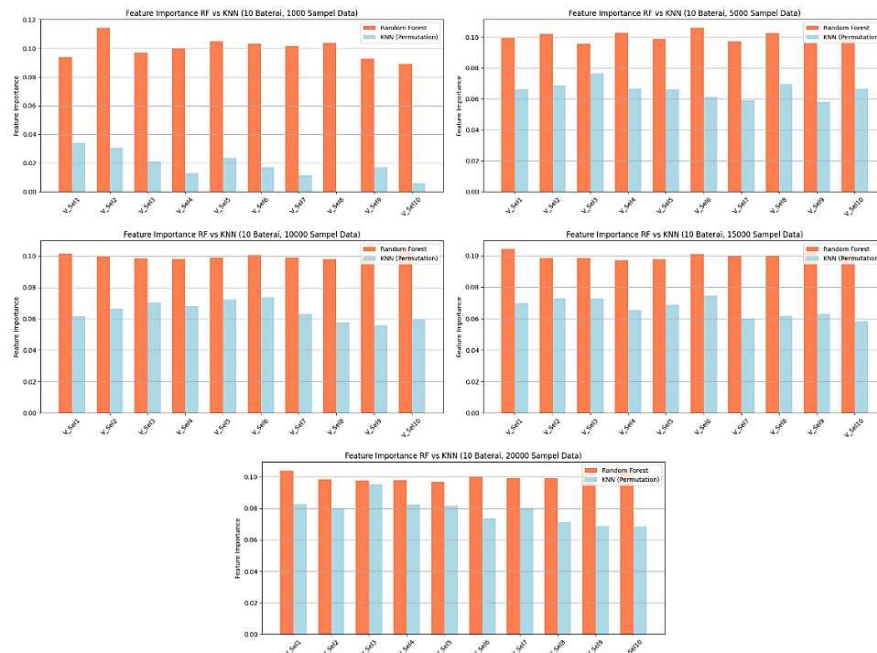
3.3 Pengujian konfigurasi baterai 10S.

Hasil pengujian untuk konfigurasi baterai 10S, untuk algoritma *Random Forest* (RF) menunjukkan performa yang konsisten dan signifikan lebih baik dibandingkan *K-Nearest Neighbors* (KNN). Pada dataset 1000 sampel, akurasi *Random Forest* (RF) tercatat 0,91 sedangkan *K-Nearest Neighbors* (KNN) hanya mencapai 0,45. Ketika dataset bertambah besar menjadi 20000 sampel, akurasi *Random Forest* (RF) meningkat menjadi 0,96 sedangkan *K-Nearest Neighbors* (KNN) hanya mencapai 0,60. Hal ini menunjukkan kemampuan *Random Forest* (RF) dalam mempelajari pola distribusi tegangan antar sel baterai meskipun jumlah fitur meningkat, sedangkan *K-Nearest Neighbors* (KNN) mengalami degradasi performa akibat sensitivitas terhadap dimensionalitas tinggi. Hasil pengujian di tunjukkan pada gambar 6.



Gambar 6. Tren akurasi *Random Forest* dan *K-Nearest Neighbors* pada dataset simulasi balancing 10 baterai.

Pada confusion matrix memperlihatkan pola distribusi prediksi yang mendukung hasil akurasi. *Random Forest* (RF) menunjukkan dominasi prediksi yang benar pada diagonal matriks, sementara *K-Nearest Neighbors* (KNN) menunjukkan distribusi prediksi yang tersebar di luar diagonal, terutama pada dataset kecil, yang mengindikasikan jumlah misclassifications yang cukup besar.



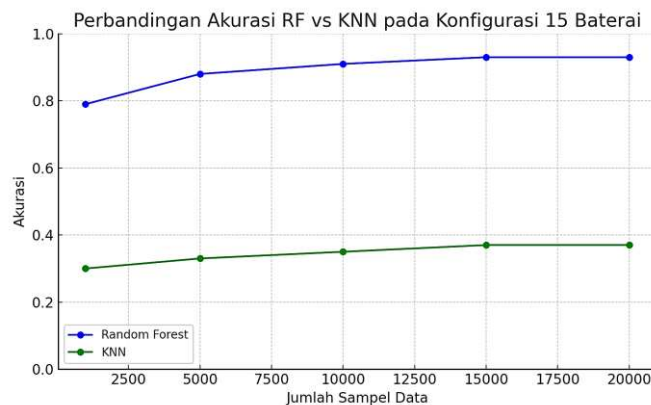
Gambar 7. *Feature importance* algoritma *Random Forest* dan *K-Nearest Neighbors* untuk konfigurasi 10 baterai (10S).

Selain itu untuk analisis *feature importance* pada gambar 7 memperlihatkan bahwa *Random Forest* (RF) menunjukkan variasi antar fitur yang semakin merata seiring bertambahnya jumlah sampel data, menunjukkan adaptabilitas *Random Forest* (RF) dalam memanfaatkan seluruh fitur secara optimal. Sebaliknya, *feature*

importance pada *K-Nearest Neighbors* (KNN) menunjukkan nilai yang relatif lebih rendah dan distribusi yang tidak variatif, sesuai karakteristik *K-Nearest Neighbors* (KNN) yang berbasis jarak.

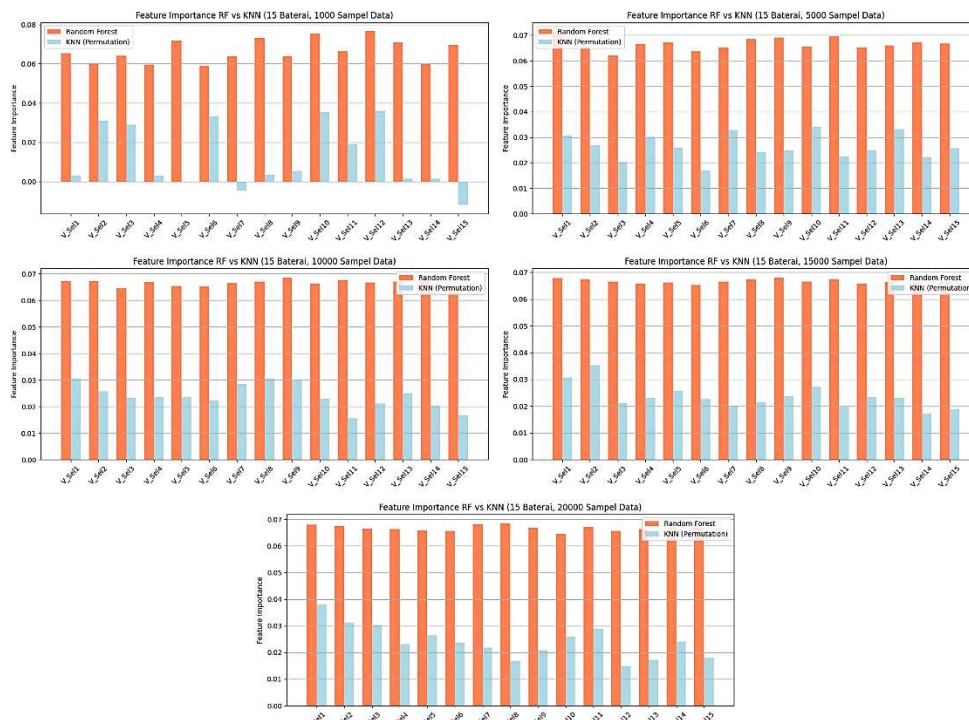
3.4 Pengujian konfigurasi baterai 15S.

Hasil pengujian untuk konfigurasi baterai 15S, untuk algoritma *Random Forest* (RF) menunjukkan performa yang konsisten dan lebih baik dibandingkan *K-Nearest Neighbors* (KNN). Pada dataset 1000 sampel, akurasi *Random Forest* (RF) tercatat 0,79 sedangkan *K-Nearest Neighbors* (KNN) hanya mencapai 0,30. Ketika dataset bertambah menjadi 20000 sampel, akurasi *Random Forest* (RF) meningkat hingga 0,93 sedangkan *K-Nearest Neighbors* (KNN) hanya mencapai 0,37. Hasil ini menunjukkan keunggulan *Random Forest* (RF) dalam mempelajari pola distribusi tegangan antar sel meskipun jumlah fitur meningkat, sementara *K-Nearest Neighbors* (KNN) mengalami degradasi performa akibat sensitivitas terhadap dimensionalitas tinggi. Hasil pengujian di tunjukkan pada gambar 8.



Gambar 8. Tren akurasi *Random Forest* dan *K-Nearest Neighbors* pada dataset simulasi balancing 15 baterai.

Pada *confusion matrix* memperlihatkan *Random Forest* (RF) memiliki prediksi yang dominan pada diagonal matriks, menandakan klasifikasi yang presisi. Sebaliknya, *K-Nearest Neighbors* (KNN) menunjukkan distribusi prediksi yang menyebar, menandakan tingginya tingkat kesalahan klasifikasi.



Gambar 9. Feature importance algoritma *Random Forest* dan *K-Nearest Neighbors* untuk konfigurasi 15 baterai (15S).

Selain itu untuk analisis *feature importance* gambar 9 juga memperlihatkan pola yang menarik. *Random Forest* (RF) menunjukkan distribusi yang relatif merata pada semua fitur (6-7%), menunjukkan kemampuannya

memanfaatkan semua informasi. Sementara itu, *K-Nearest Neighbors* (KNN) menunjukkan feature importance yang rendah (2-3%) dan tidak stabil, mengindikasikan keterbatasan dalam menangani dataset berdimensi tinggi.

3.5 Pembahasan

Dari hasil simulasi keempat konfigurasi baterai, yaitu 4S, 8S, 10S, dan 15S, dapat disimpulkan bahwa secara umum algoritma *Random Forest* (RF) selalu menunjukkan kinerja yang superior dibandingkan dengan *K-Nearest Neighbors* (KNN) dalam berbagai skenario pengujian, baik dalam konteks jumlah sel baterai maupun variasi skala dataset. Performa *Random Forest* (RF) yang konsisten ini menandakan bahwa algoritma tersebut lebih *robust* terhadap peningkatan kompleksitas data, baik dari segi dimensi fitur maupun jumlah sampel pelatihan..

Pada konfigurasi 4 baterai, *Random Forest* (RF) memperoleh akurasi awal sebesar 0,90 pada dataset kecil berukuran 1000 sampel, dan meningkat hingga 0,98 pada dataset terbesar (20000 sampel). Sementara itu, *K-Nearest Neighbors* (KNN) hanya mampu mencatatkan akurasi awal 0,82 yang kemudian meningkat secara bertahap hingga 0,94. Walaupun selisih akurasi tidak terlalu signifikan pada konfigurasi ini, *Random Forest* (RF) tetap menunjukkan keunggulan dari segi konsistensi dan kestabilan akurasi dalam berbagai ukuran data. Hal ini menunjukkan bahwa *Random Forest* (RF) memiliki kemampuan generalisasi yang lebih baik bahkan dalam kasus dengan kompleksitas rendah.

Pada konfigurasi 8 baterai, performa kedua algoritma mulai menunjukkan perbedaan yang mencolok. *Random Forest* (RF) mencatat akurasi awal sebesar 0,87 dan meningkat menjadi 0,96 seiring bertambahnya jumlah sampel. Di sisi lain, *K-Nearest Neighbors* (KNN) justru mengalami kesulitan, hanya memperoleh akurasi 0,55 pada dataset kecil dan maksimal 0,71 meskipun jumlah data telah diperbesar hingga 20000 sampel. Penurunan performa *K-Nearest Neighbors* (KNN) pada konfigurasi ini menunjukkan keterbatasannya dalam menangani data berdimensi lebih tinggi, di mana kompleksitas relasi antar fitur menjadi lebih kompleks dan tidak dapat sepenuhnya dijelaskan oleh metrik jarak semata.

Tren yang sama semakin terlihat pada konfigurasi 10 baterai. *Random Forest* (RF) masih menunjukkan keunggulannya dengan akurasi 0,91 pada dataset kecil dan meningkat menjadi 0,96 pada dataset besar. Sementara *K-Nearest Neighbors* (KNN) justru menurun performanya, hanya mampu mencapai akurasi 0,45 hingga 0,60. Semakin tinggi dimensi data menyebabkan jarak antar titik menjadi semakin homogen (fenomena *curse of dimensionality*), yang berdampak langsung pada efektivitas *K-Nearest Neighbors* (KNN) dalam membedakan kelas-kelas data.

Puncak dari perbandingan performa terjadi pada konfigurasi baterai 15S, yang memiliki kompleksitas tertinggi. *Random Forest* (RF) berhasil mencatat akurasi sebesar 0,79 pada dataset kecil dan meningkat signifikan hingga 0,93 pada dataset terbesar. Sebaliknya, *K-Nearest Neighbors* (KNN) mengalami kesulitan yang sangat besar, hanya mencatatkan akurasi 0,30 hingga 0,37. Ini menunjukkan bahwa *Random Forest* (RF) memiliki kapasitas pembelajaran yang jauh lebih baik terhadap pola distribusi tegangan antar sel baterai, bahkan dalam kondisi yang sangat kompleks. Sementara *K-Nearest Neighbors* (KNN) gagal menangkap karakteristik data yang semakin kompleks, yang menunjukkan keterbatasan algoritma tersebut dalam domain klasifikasi dengan dimensi tinggi.

Analisis lebih lanjut pada confusion matrix memperkuat temuan ini. *Random Forest* (RF) secara konsisten menunjukkan distribusi prediksi yang presisi, dengan nilai-nilai tertinggi terkonsentrasi pada diagonal utama matriks. Ini menandakan bahwa sebagian besar prediksi model berada pada kelas yang benar. Tingkat kesalahan (*misclassification*) sangat rendah dan semakin menurun seiring bertambahnya data. Di sisi lain, *K-Nearest Neighbors* (KNN) menunjukkan distribusi prediksi yang tersebar, terutama pada dataset kecil dan konfigurasi baterai tinggi. Banyaknya nilai prediksi yang salah menunjukkan bahwa *K-Nearest Neighbors* (KNN) tidak mampu membedakan secara akurat antara kelas yang berbeda dalam kondisi kompleks.

Selain evaluasi akurasi dan confusion matrix, analisis feature importance memberikan wawasan tambahan terkait bagaimana kedua algoritma memanfaatkan fitur input. Pada dataset kecil, *Random Forest* (RF) cenderung mengandalkan fitur dominan seperti V_Sel1, namun seiring bertambahnya jumlah data dan fitur, distribusi importance menjadi lebih merata. Ini mengindikasikan bahwa *Random Forest* (RF) mampu belajar secara adaptif terhadap kontribusi relatif dari semua fitur. Pola ini sangat penting karena mencerminkan bahwa *Random Forest* (RF) tidak hanya mengandalkan satu atau dua fitur saja, tetapi mampu memahami hubungan multivariat antar fitur tegangan sel baterai secara menyeluruh.

Berbeda dengan *Random Forest* (RF), *K-Nearest Neighbors* (KNN) menunjukkan pola feature importance yang rendah dan kurang bervariasi. Ini konsisten dengan sifat dasar *K-Nearest Neighbors* (KNN) yang tidak melakukan proses pelatihan eksplisit seperti *Random Forest* (RF), melainkan hanya mengandalkan perhitungan jarak pada saat inferensi. Oleh karena itu, *K-Nearest Neighbors* (KNN) tidak memiliki kemampuan untuk mengestimasi kontribusi relatif dari setiap fitur terhadap hasil klasifikasi, yang menjadi kelemahan signifikan dalam konteks data berdimensi tinggi.

Visualisasi tren akurasi mempertegas keunggulan *Random Forest* (RF). Pada seluruh konfigurasi, kurva akurasi *Random Forest* (RF) selalu berada di atas *K-Nearest Neighbors* (KNN), dan menunjukkan tren yang semakin meningkat secara linier maupun eksponensial dengan penambahan jumlah data. Ini menunjukkan

skalabilitas dan stabilitas *Random Forest* (RF), serta kemampuannya dalam menyerap informasi secara bertahap dan efektif dari data yang tersedia.

Secara keseluruhan, pembahasan ini memperjelas bahwa *Random Forest* adalah algoritma yang lebih andal dan efisien dalam aplikasi klasifikasi kondisi balancing baterai, terutama untuk konfigurasi multi-sel dengan jumlah fitur yang besar. Kemampuan *Random Forest* (RF) dalam mengatasi noise, menangani dataset berdimensi tinggi, serta memberikan interpretasi melalui feature importance menjadikannya pilihan ideal untuk diintegrasikan ke dalam sistem manajemen baterai berbasis kecerdasan buatan. Sebaliknya, meskipun *K-Nearest Neighbors* (KNN) masih relevan untuk aplikasi sederhana dan dataset kecil, performanya tidak cukup kompetitif untuk digunakan dalam sistem real-time yang kompleks dan menuntut akurasi tinggi.

4. KESIMPULAN

Berdasarkan hasil pembahasan yang telah dilakukan untuk seluruh konfigurasi baterai (4S, 8S, 10S, dan 15S baterai), dengan variasi dataset sebanyak 1000, 5000, 10000, 15000, 20000 sampel data, dapat disimpulkan bahwa algoritma *Random Forest* (RF) secara konsisten lebih unggul dibandingkan algoritma *K-Nearest Neighbors* (KNN) dalam simulasi prediksi balancing charger baterai. *Random Forest* mampu mencapai akurasi tinggi secara konsisten pada seluruh konfigurasi baterai 4S, 8S, 10S, dan 15S, bahkan pada dataset kecil, dan terus meningkat seiring bertambahnya ukuran dataset. Algoritma *Random Forest* ini juga menunjukkan tingkat kesalahan klasifikasi yang rendah dengan pola prediksi yang terpusat di diagonal *confusion matrix*, serta adaptif dalam memanfaatkan fitur secara optimal dengan seiring bertambahnya jumlah dataset serta kompleksitas konfigurasi baterai. Sebaliknya, untuk *K-Nearest Neighbors* menunjukkan keterbatasan signifikan terutama dalam menghadapi data berdimensi tinggi. KNN membutuhkan dataset yang sangat besar untuk mencapai akurasi yang mendekati RF, dan tetap menunjukkan tingkat kesalahan prediksi yang tinggi serta distribusi *feature importance* yang rendah dan tidak stabil. Oleh karena itu, berdasarkan keseluruhan analisis, penelitian ini merekomendasikan algoritma *Random Forest* sebagai pilihan utama dalam implementasi sistem *balancing charger* baterai berbasis *machine learning*, khususnya untuk konfigurasi baterai dengan jumlah sel yang bervariasi dari 4, 8, 10 dan 15 baterai. Algoritma ini memberikan performa yang stabil, akurasi tinggi, adaptabilitas optimal terhadap peningkatan jumlah fitur dan dataset, serta menghasilkan prediksi yang lebih presisi dibandingkan *K-Nearest Neighbors*.

REFERENCES

- [1] PT PLN (PERSERO), "Rencana Usaha Penyediaan Tenaga Listrik (RUPTL) 2025-2034," Jakarta, May 2025.
- [2] J. Langer, J. Quist, and K. Blok, "Review of Renewable Energy Potentials in Indonesia and Their Contribution to a 100% Renewable Electricity System," *Energies (Basel)*, vol. 14, no. 21, p. 7033, Oct. 2021, doi: 10.3390/en14217033.
- [3] M. Latif, R. Nazir, and H. Reza, "Analisa Proses Charging Akumulator Pada Prototipe Turbin Angin Sumbu Horizontal Di Pantai Purus Padang," *Jurnal Nasional Teknik Elektro*, vol. 2, no. 1, pp. 1–8, Apr. 2013, doi: 10.20449/jnte.v2i1.91.
- [4] M. R. Faqih, N. A. Windarko, and E. Wahjono, "Sistem Baterai Cell Balancing Pasif Menggunakan Kontrol Logika Fuzzy Tipe Mamdani untuk Baterai Pack Lithium," *J-Innovation*, vol. 10, no. 2, pp. 34–43, Dec. 2021, doi: 10.55600/jipa.v10i2.111.
- [5] O. Zebua, A. H. Setiawan, N. Soedjarwanto, J. Anggara, and A. Haris, "Rancang Bangun Alat Monitoring Ketidakseimbangan Beban Pada Jaringan Tegangan Menengah," *JURNAL NASIONAL TEKNIK ELEKTRO*, vol. 5, no. 3, p. 405, Nov. 2016, doi: 10.25077/jnte.v5n3.325.2016.
- [6] R. N. Ahmad, H. Suryatomo, and D. C. Riawan, "Rancang Bangun Pengisi Daya Untuk Baterai Lithium-Polymer Dengan Mempertimbangkan Kompensasi Resistansi," *Transmisi: Jurnal Ilmiah Teknik Elektro*, vol. 25, no. 2, pp. 48–57, Jun. 2023, doi: 10.14710/transmisi.25.2.48-57.
- [7] S. Voronov, E. Frisk, and M. Krysander, "Data-Driven Battery Lifetime Prediction and Confidence Estimation for Heavy-Duty Trucks," *IEEE Trans Reliab*, vol. 67, no. 2, pp. 623–639, Jun. 2018, doi: 10.1109/TR.2018.2803798.
- [8] H. Yedla, L. R. Koppada, and R. S. Bodala, "Advanced Battery Management: Forecasting Health, State of Charge & Maintenance Needs Using AI & ML Models (LSTM, Gradient Boosting, SVR, Random Forest)," *Asian Journal of Research in Computer Science*, vol. 17, no. 8, pp. 46–57, Jul. 2024, doi: 10.9734/ajrcos/2024/v17i7489.
- [9] T. Jim Hassan *et al.*, "UAS-Guided Analysis of Electric and Magnetic Field Distribution in High-Voltage Transmission Lines (Tx) and Multi-Stage Hybrid Machine Learning Models for Battery Drain Estimation," *IEEE Access*, vol. 12, pp. 82289–82317, 2024, doi: 10.1109/ACCESS.2024.3394532.
- [10] S. Rout, S. K. Samal, and S. R. Mahapatro, "Estimation of SOC for Li-ion battery-powered three-wheeled electric vehicle using machine learning methods," *Engineering Research Express*, vol. 6, no. 4, p. 045308, Dec. 2024, doi: 10.1088/2631-8695/ad8063.

- [11] K. S. R. Mawonou, A. Eddahech, D. Dumur, D. Beauvois, and E. Godoy, "State-of-health estimators coupled to a random forest approach for lithium-ion battery aging factor ranking," *J Power Sources*, vol. 484, p. 229154, Feb. 2021, doi: 10.1016/j.jpowsour.2020.229154.
- [12] G. Wang, Z. Lyu, and X. Li, "An Optimized Random Forest Regression Model for Li-Ion Battery Prognostics and Health Management," *Batteries*, vol. 9, no. 6, p. 332, Jun. 2023, doi: 10.3390/batteries9060332.
- [13] L. Shao, L. Zhao, H. Liu, D. Zhang, J. Li, and C. Li, "Research on the Remaining Useful Life Prediction Method of Energy Storage Battery Based on Multimodel Integration," *ACS Omega*, vol. 9, no. 39, pp. 40496–40510, Oct. 2024, doi: 10.1021/acsomega.4c03524.
- [14] U. Khan, M. Tariq, and A. I. Sarwat, "Adaptive Remaining Capacity Estimator of Lithium-Ion Battery Using Genetic Algorithm-Tuned Random Forest Regressor Under Dynamic Thermal and Operational Environments," *Energies (Basel)*, vol. 17, no. 22, p. 5582, Nov. 2024, doi: 10.3390/en17225582.
- [15] G. Crocioni, D. Pau, J.-M. Delorme, and G. Gruosso, "Li-Ion Batteries Parameter Estimation With Tiny Neural Networks Embedded on Intelligent IoT Microcontrollers," *IEEE Access*, vol. 8, pp. 122135–122146, 2020, doi: 10.1109/ACCESS.2020.3007046.
- [16] Dr. M. S. Dr. S. D. Akansh Garg and Mr. G. D. Ms. A. Dr. Abhishek Tripathi, "Machine Learning Based Risk Management of Credit Sales in Small and Mid-Size Business," *Journal of Informatics Education and Research*, vol. 4, no. 2, May 2024, doi: 10.52783/jier.v4i2.842.
- [17] I. Andik, F. Y. Arslan, and A. Uysal, "Comparison of prediction performance of lithium titanate oxide battery discharge capacity with machine learning methods," *Electrical Engineering*, vol. 107, no. 5, pp. 6721–6734, May 2025, doi: 10.1007/s00202-024-02503-8.
- [18] C. Chang, Z. Feng, and Z. Liu, "A Study of the Distribution of Forest Density in Inner Mongolia Based on Environmental Factors," *Forests*, vol. 13, no. 2, p. 313, Feb. 2022, doi: 10.3390/f13020313.
- [19] Z. B. Omariba, L. Zhang, and D. Sun, "Review of Battery Cell Balancing Methodologies for Optimizing Battery Pack Performance in Electric Vehicles," *IEEE Access*, vol. 7, pp. 129335–129352, 2019, doi: 10.1109/ACCESS.2019.2940090.
- [20] X. Wang, "Stock Price Prediction: A Comparative Study of Random Forest and LSTM Models," *Highlights in Science, Engineering and Technology*, vol. 107, pp. 117–123, Aug. 2024, doi: 10.54097/70a8b947.
- [21] L. Breiman, "Random Forests," *Mach Learn*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
- [22] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans Inf Theory*, vol. 13, no. 1, pp. 21–27, Jan. 1967, doi: 10.1109/TIT.1967.1053964.