

Analisis Dampak SMOTE terhadap Feature Importance pada Klasifikasi Data Migraine menggunakan Random Forest dan Extra Trees

Henny Leidiyana^{1*}

^{1*}Informatika/Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika
e-mail: henny.hnl@nbsi.ac.id

DOI: <https://doi.org/10.31603/komtika.v10.i1.16684>

ABSTRACT

This study analyzes the impact of the Synthetic Minority Over-sampling Technique (SMOTE) on model performance and feature importance in the classification of migraine patients using Random Forest (RF) and Extra Trees (ET) algorithms. Evaluation was conducted based on recall and F1-Score for the minority class, as well as Permutation Importance analysis. The results indicate that ET, especially when combined with SMOTE (ET + SMOTE), delivers the best performance for the minority class. ET + SMOTE achieved an average F1-Score of 0.7000 and an average recall of 0.8041 using 11 optimal features, indicating better feature efficiency. The application of SMOTE significantly affected the ranking of important features. Although SMOTE improved detection for some minority classes, its impact was not always consistent and occasionally reduced performance on other minority classes. This study concludes that SMOTE alters feature contributions and model interpretability, as well as enhances performance on certain minority classes, particularly when combined with ET.

Keywords: SMOTE, Feature Importance, Random Forest, Extra Trees, Permutation Importance.

ABSTRAK

Penelitian ini menganalisis dampak Synthetic Minority Over-sampling Technique (SMOTE) terhadap kinerja model dan fitur penting (*feature importance*) dalam klasifikasi penderita migrain menggunakan algoritma Random Forest (RF) dan Extra Trees (ET). Evaluasi dilakukan berdasarkan recall dan F1-Score untuk kelas minoritas, serta analisis Permutation Importance. Hasil menunjukkan bahwa ET, khususnya dengan SMOTE (ET + SMOTE), memberikan kinerja terbaik untuk kelas minoritas. ET + SMOTE mencapai Avg_F1_Score 0.7000 dan Avg_Recall 0.8041 dengan 11 fitur optimal, mengindikasikan efisiensi fitur yang lebih baik. Penerapan SMOTE secara signifikan memengaruhi peringkat fitur penting. Meskipun SMOTE meningkatkan deteksi beberapa kelas minoritas, dampaknya tidak selalu konsisten dan terkadang dapat menurunkan kinerja pada kelas minoritas lain. Penelitian ini menyimpulkan bahwa SMOTE mengubah kontribusi fitur dan interpretasi model, serta meningkatkan kinerja pada kelas minoritas tertentu, terutama bila dikombinasikan dengan ET.

Kata Kunci: SMOTE, Feature Importance, Random Forest, Extra Trees, Permutation Importance.

PENDAHULUAN

Migrain merupakan gangguan sakit kepala primer yang pada sebagian besar kasus bersifat episodik, biasanya berlangsung selama 4–72 jam, serta disertai mual, muntah, dan/atau fotofobia serta fonofobia [1]. Prevalensi global migrain sangat tinggi, mempengaruhi sekitar 1 dari 7 orang dewasa, dengan dampak signifikan terhadap kualitas hidup, produktivitas kerja, dan beban ekonomi Kesehatan [2]. Oleh karena itu, klasifikasi yang akurat terhadap jenis migrain sangat krusial untuk diagnosis yang tepat dan penentuan strategi penanganan yang efektif.

Dalam upaya klasifikasi penderita migrain menggunakan metode pembelajaran mesin, seringkali dihadapkan pada tantangan ketidakseimbangan kelas (*class imbalance*) [3] [4]. Dataset medis umumnya menunjukkan bahwa jumlah sampel untuk kelas-kelas minoritas (misalnya, jenis migrain yang langka) jauh lebih sedikit dibandingkan dengan kelas mayoritas (misalnya, migrain tanpa aura). Ketidakseimbangan ini dapat menyebabkan model cenderung bias terhadap kelas mayoritas, menghasilkan kinerja yang buruk dalam memprediksi kelas minoritas yang justru seringkali merupakan kondisi klinis yang lebih krusial untuk diidentifikasi[5].

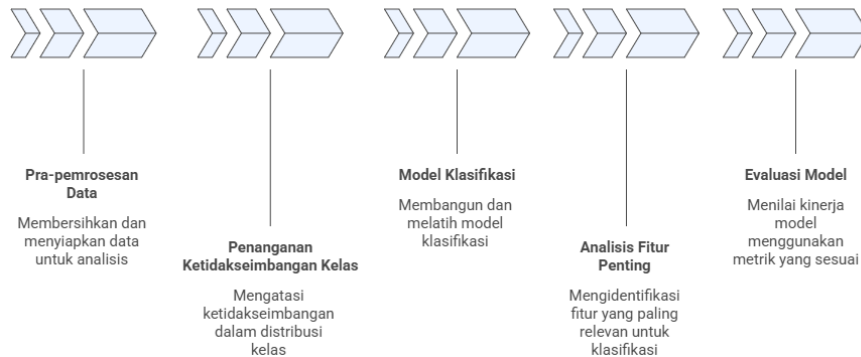
Salah satu metode populer untuk mengatasi ketidakseimbangan kelas adalah Synthetic Minority Over-sampling Technique (SMOTE), yang bekerja dengan menghasilkan sampel sintesis baru dari kelas minoritas untuk menyeimbangkan distribusi kelas dalam dataset pelatihan [6]. SMOTE telah terbukti efektif dalam meningkatkan kinerja model pada kelas minoritas dalam berbagai domain. Namun, meskipun SMOTE banyak digunakan, dampaknya terhadap fitur penting (*feature importance*) yang diidentifikasi oleh model klasifikasi, khususnya pada algoritma ensemble seperti Random Forest (RF) dan Extra Trees (ET), masih memerlukan eksplorasi lebih lanjut. Perubahan pada distribusi data setelah SMOTE dapat memengaruhi bagaimana model memandang kontribusi setiap fitur, yang pada akhirnya dapat memengaruhi interpretasi model dan seleksi fitur.

Feature importance menjadi aspek penting dalam analisis data karena mampu mengidentifikasi variabel yang paling berpengaruh terhadap hasil klasifikasi. Meskipun berbagai penelitian telah mengkaji penggunaan algoritma *machine learning* untuk klasifikasi penyakit, studi yang secara khusus mengintegrasikan analisis fitur dominan dan evaluasi kinerja model berbasis ensemble tree dalam konteks klasifikasi tipe migrain masih terbatas. Sejumlah penelitian cenderung berfokus pada peningkatan performa model, sementara aspek interpretabilitas seperti kontribusi masing-masing fitur belum banyak dieksplorasi secara mendalam [7][8]. Oleh karena itu, diperlukan penelitian yang tidak hanya mengevaluasi kinerja model, tetapi juga mengidentifikasi fitur-fitur dominan yang berperan dalam klasifikasi tipe migrain.

Penelitian ini bertujuan untuk menganalisis dampak penerapan SMOTE terhadap nilai penting fitur dalam tugas klasifikasi penderita migrain. Konsistensi fitur penting dianalisis berdasarkan perubahan peringkat feature importance pada model RF dan ET. Penelitian sebelumnya menunjukkan bahwa penerapan SMOTE dapat memengaruhi proses pembelajaran model dan hasil analisis variabel pada data tidak seimbang, terutama pada proses seleksi fitur dan identifikasi variabel yang berkontribusi terhadap klasifikasi [9][10]. Penulis akan membandingkan bagaimana SMOTE memengaruhi identifikasi fitur-fitur kunci oleh model Random Forest dan Extra Trees, baik sebelum maupun setelah teknik oversampling diterapkan. Penggunaan RF dan ET dipilih karena kedua algoritma berbasis pohon tersebut mampu menghasilkan nilai feature importance secara otomatis dari proses pembentukan model sehingga dapat digunakan untuk mengidentifikasi fitur-fitur yang paling berkontribusi terhadap hasil klasifikasi [11]. Analisis ini diharapkan dapat memberikan pemahaman mengenai interaksi antara teknik penanganan ketidakseimbangan kelas dan interpretasi model, serta memberikan kontribusi dalam optimasi seleksi fitur dalam diagnosis migrain yang lebih akurat.

METODE

Bagian ini menguraikan metodologi yang digunakan dalam penelitian ini, meliputi sumber data, tahap pra-pemrosesan, teknik penanganan ketidakseimbangan kelas, model klasifikasi yang diuji, serta metode analisis fitur penting dan evaluasi kinerja model seperti disajikan pada Gambar 1.



Gambar 1. Tahapan Penelitian

1. Sumber Data

Dataset yang digunakan dalam penelitian ini adalah 'migraine_data.csv' yang diperoleh dari dataset publik pada platform Kaggle, berisi fitur-fitur klinis dan demografis pasien migrain. Data ini terdiri dari 23 fitur prediktor dan satu variabel target, yaitu 'Type' migrain. Total sampel dalam dataset adalah 400. Pembagian dataset dilakukan secara stratifikasi untuk memastikan proporsi kelas yang sama pada set pelatihan dan pengujian.

2. Pra-pemrosesan Data

Tahap pra-pemrosesan data meliputi pemisahan fitur dan target, pembagian dataset data pelatihan dan data pengujian. Pembagian dilakukan secara stratifikasi, yaitu membagi data dengan tetap mempertahankan proporsi/jumlah setiap kelas target agar distribusinya tetap sama pada data latih dan data uji.

3. Penanganan Ketidakseimbangan Kelas

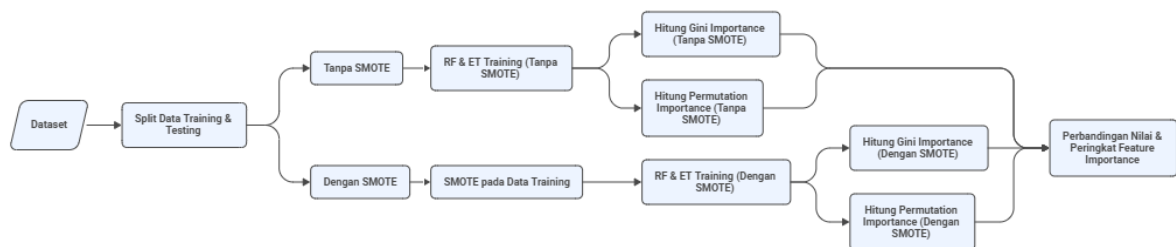
Untuk mengatasi masalah ketidakseimbangan kelas pada dataset migrain, teknik Synthetic Minority Over-sampling Technique (SMOTE) diterapkan pada data pelatihan. SMOTE bekerja dengan menghasilkan sampel sintesis dari kelas minoritas, sehingga menyeimbangkan distribusi kelas dan mencegah model bias terhadap kelas mayoritas.

4. Model Klasifikasi

Dua algoritma ensemble berbasis pohon digunakan dalam penelitian ini Random Forest (RF) dan Extra Trees (ET). RF membangun beberapa pohon keputusan secara independen dan menggabungkan prediksinya. RF dikenal handal dalam menangani fitur berdimensi tinggi dan mengurangi overfitting [12]. ET mirip dengan RF tetapi menambahkan randomisasi lebih lanjut dalam proses pembangunan pohon dengan memilih ambang pemisahan fitur secara acak yang membuat pohon-pohon dalam ensemble menjadi lebih bervariasi dan seringkali mengurangi variansi model [13]. Kedua model ini diuji dalam dua skenario, yaitu tanpa SMOTE (*baseline*) dan dengan penerapan SMOTE pada data pelatihan.

5. Analisis Fitur Penting (*Feature Importance*)

Analisis penting fitur seperti pada Gambar 2 penting dilakukan untuk mengidentifikasi fitur-fitur yang paling berpengaruh dalam klasifikasi migrain, baik sebelum maupun setelah penerapan SMOTE. Dua metode utama digunakan yaitu *Gini Importance* (MDI) dan *Permutation Importance*. *Gini importance* merupakan metrik yang sudah tersedia secara otomatis di dalam algoritma berbasis pohon seperti Random Forest dan Extra Trees untuk mengukur seberapa besar penurunan ketidakmurnian (*impurity*) yang disebabkan oleh fitur tersebut [14]. Metode *permutation importance* mengukur fitur penting dengan mengacak nilai fitur tertentu pada data pengujian dan mengamati seberapa besar penurunan kinerja model. Keunggulan *permutation importance* adalah kemampuannya untuk mengukur fitur penting secara global pada model yang sudah terlatih dan tidak bias terhadap fitur dengan banyak nilai unik yang berbeda [15]. Perbandingan dilakukan antara fitur penting yang diidentifikasi oleh RF dan ET, baik tanpa SMOTE maupun setelah SMOTE, untuk melihat pergeseran atau perubahan relatif dalam kontribusi fitur.



Gambar 2. Proses Analisa Fitur Penting

Tahapan analisis dilakukan dengan terlebih dahulu melatih model Random Forest dan Extra Trees menggunakan dataset asli sebelum penerapan SMOTE. Setelah model terbentuk, nilai feature importance dihitung menggunakan metode Gini Importance dan Permutation Importance untuk memperoleh fitur-fitur yang paling berkontribusi terhadap proses klasifikasi migrain. Selanjutnya, SMOTE diterapkan pada data pelatihan untuk menyeimbangkan distribusi kelas melalui pembentukan sampel sintesis pada kelas minoritas. Dataset hasil oversampling kemudian digunakan kembali untuk melatih model Random Forest dan Extra Trees. Setelah proses pelatihan ulang dilakukan, nilai feature importance dihitung kembali menggunakan metode yang sama. Hasil sebelum dan sesudah penerapan SMOTE kemudian dibandingkan untuk menganalisis perubahan nilai maupun peringkat fitur penting pada masing-masing model. Analisis ini dilakukan untuk mengetahui bagaimana perubahan distribusi data akibat SMOTE memengaruhi identifikasi fitur-fitur kunci dalam klasifikasi migrain.

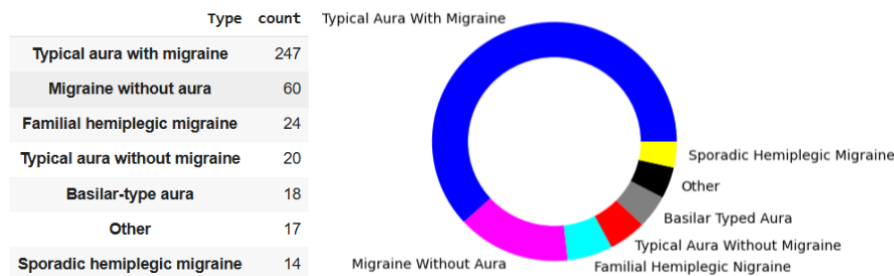
6. Evaluasi Model

Kinerja model dievaluasi menggunakan Classification Report, Confusion matrix, F1-Score untuk membandingkan kinerja model secara keseluruhan dalam mendeteksi kelas yang kurang terwakili, dan Optimal K-Features Analysis untuk menentukan jumlah fitur optimal

(K) yang memberikan F1-Score dan Recall rata-rata tertinggi untuk kelas minoritas berdasarkan urutan Permutation Importance dari masing-masing model.

HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil dari eksperimen yang telah dilakukan dan pembahasannya, fokus pada dampak penerapan SMOTE terhadap kinerja model klasifikasi (Random Forest dan Extra Trees) serta perubahan pada identifikasi fitur penting. Sebagai tahap awal, dilakukan analisis terhadap distribusi kelas target untuk mengidentifikasi kondisi ketidakseimbangan data yang berpotensi memengaruhi performa model klasifikasi seperti disajikan pada Gambar 3.



Gambar 3. Distribusi Kelas Target

Visualisasi dalam bentuk diagram pada Gambar 3 menjelaskan jumlah data pada kelas *Typical Aura with Migraine* jauh lebih banyak dibandingkan kelas lainnya. Hal ini menunjukkan bahwa dataset tidak seimbang dan memerlukan penanganan khusus, seperti teknik resampling atau penyesuaian parameter model, agar hasil klasifikasi menjadi lebih adil dan representatif. Evaluasi awal terhadap model Random Forest (RF) dan Extra Trees (ET) tanpa penerapan SMOTE ditunjukkan pada Tabel 1.

Tabel 1. Kinerja Model Klasifikasi sebelum SMOTE

RF Baseline					ET Baseline				
	precision	recall	f1-score	support		precision	recall	f1-score	support
Basilar-type aura	0.67	0.50	0.57	4	Basilar-type aura	1.00	0.50	0.67	4
Familial hemiplegic migraine	0.50	0.60	0.55	5	Familial hemiplegic migraine	0.50	0.80	0.62	5
Migraine without aura	0.92	1.00	0.96	12	Migraine without aura	0.92	1.00	0.96	12
Other	1.00	0.67	0.80	3	Other	1.00	0.67	0.80	3
Sporadic hemiplegic migraine	1.00	0.67	0.80	3	Sporadic hemiplegic migraine	1.00	0.67	0.80	3
Typical aura with migraine	0.94	0.96	0.95	49	Typical aura with migraine	0.96	0.96	0.96	49
Typical aura without migraine	1.00	1.00	1.00	4	Typical aura without migraine	1.00	1.00	1.00	4
accuracy			0.90	80	accuracy			0.91	80
macro avg	0.86	0.77	0.80	80	macro avg	0.91	0.80	0.83	80
weighted avg	0.90	0.90	0.90	80	weighted avg	0.93	0.91	0.91	80

Kedua model baseline menunjukkan akurasi yang tinggi secara keseluruhan. Namun, pada kelas minoritas seperti 'Basilar-type aura', 'Familial hemiplegic migraine', 'Other', dan 'Sporadic hemiplegic migraine', nilai recall dan f1-score cenderung lebih rendah dibandingkan kelas mayoritas. Extra Trees Baseline menunjukkan sedikit keunggulan pada metrik macro average F1-score (0.83 vs 0.80) dan recall (0.80 vs 0.77) dibandingkan Random Forest

Baseline. Setelah penerapan SMOTE pada data pelatihan, kinerja model berubah seperti yang ditunjukkan pada Tabel 2.

Tabel 2. Kinerja Model Klasifikasi Setelah SMOTE

RF+SMOTE					ET+SMOTE				
	precision	recall	f1-score	support		precision	recall	f1-score	support
Basilar-type aura	0.40	0.50	0.44	4	Basilar-type aura	0.50	0.75	0.60	4
Familial hemiplegic migraine	0.40	0.80	0.53	5	Familial hemiplegic migraine	0.40	0.80	0.53	5
Migraine without aura	0.86	1.00	0.92	12	Migraine without aura	0.86	1.00	0.92	12
Other	1.00	0.33	0.50	3	Other	1.00	0.33	0.50	3
Sporadic hemiplegic migraine	1.00	0.33	0.50	3	Sporadic hemiplegic migraine	1.00	0.33	0.50	3
Typical aura with migraine	0.98	0.90	0.94	49	Typical aura with migraine	0.98	0.88	0.92	49
Typical aura without migraine	1.00	1.00	1.00	4	Typical aura without migraine	1.00	1.00	1.00	4
accuracy			0.85	80	accuracy			0.85	80
macro avg	0.80	0.69	0.69	80	macro avg	0.82	0.73	0.71	80
weighted avg	0.90	0.85	0.85	80	weighted avg	0.90	0.85	0.86	80

Penerapan SMOTE secara umum sedikit menurunkan akurasi keseluruhan, tetapi dampaknya pada kelas minoritas bervariasi. Untuk RF + SMOTE, recall untuk 'Familial hemiplegic migraine' meningkat dari 0.60 menjadi 0.80, namun 'Other' dan 'Sporadic hemiplegic migraine' mengalami penurunan recall yang signifikan. Untuk ET + SMOTE, recall untuk 'Basilar-type aura' meningkat dari 0.50 menjadi 0.75 dan 'Familial hemiplegic migraine' tetap 0.80. Namun, sama seperti RF + SMOTE, recall untuk 'Other' dan 'Sporadic hemiplegic migraine' juga menurun.

Dapat disimpulkan bahwa SMOTE tidak secara konsisten meningkatkan recall untuk semua kelas minoritas, dan dalam beberapa kasus justru menurunkan kinerja. Hal ini mungkin terjadi karena pembuatan sampel sintetis dapat memperkenalkan noise atau over-generalisasi pada kelas yang sangat kecil. Ringkasan fitur penting berdasarkan *Gini Importance* dan *Permutation Importance* untuk model RF dan ET sebelum SMOTE ditunjukkan pada Tabel 3

Tabel 3. Analisis Feature Importance Sebelum Penerapan SMOTE

Fitur	RF Importance	Permutation Importance (RF)	ET Importance	Permutation Importance (ET)
Visual	0.119952	0.14250	0.117407	0.12375
Intensity	0.098695	0.02250	0.100042	0.04875
Vertigo	0.054717	0.00750	0.060430	0.03625
Sensory	0.036932	0.02000	0.038738	0.02625
Tinnitus	0.026449	0.03375	0.027819	0.02375
Dysphasia	0.034878	0.02500	0.044698	0.02125
DPF	0.073125	0.02750	0.077862	0.01750
Character	0.097719	0.01625	0.084858	0.01625
Hypoacusis	0.025090	0.01250	0.033095	0.01500
Age	0.118714	0.02375	0.064883	0.01375
Location	0.073300	0.01000	0.088355	0.01250
Defect	0.027154	0.01125	0.028901	0.01125
Frequency	0.057502	0.00125	0.056771	0.01000
Vomit	0.022455	0.00125	0.025828	0.00500
Phonophobia	0.027752	0.00000	0.038596	0.00250
Duration	0.029343	-0.01375	0.029624	-0.00125
Photophobia	0.032803	0.00000	0.030682	0.00000
Conscience	0.018604	0.00000	0.019896	0.00000
Nausea	0.014970	0.00000	0.016785	0.00000
Paresthesia	0.004021	0.00000	0.006674	0.00000
Diplopia	0.003702	0.00000	0.006065	0.00000
Dysarthria	0.002123	0.00000	0.001992	0.00000
Ataxia	0.000000	0.00000	0.000000	0.00000

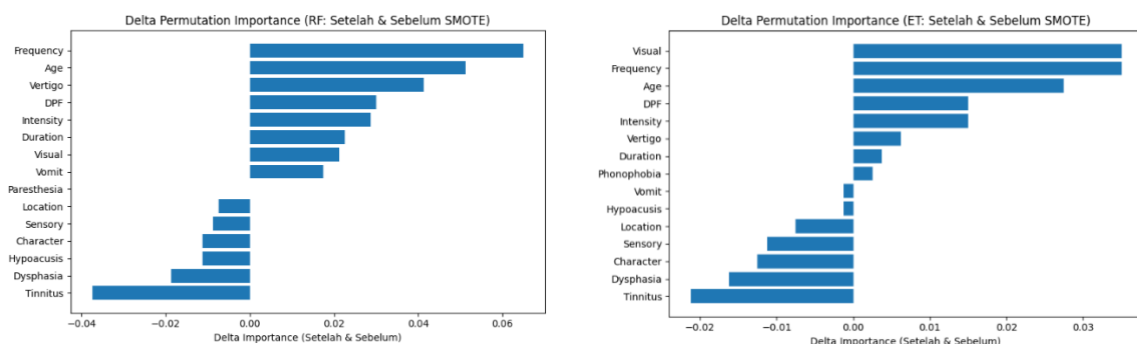
Sebelum SMOTE, fitur 'Visual' secara konsisten menjadi fitur paling penting menurut Permutation Importance untuk RF (0.14250) dan ET (0.12375). Fitur 'Tinnitus', 'DPF', dan 'Dysphasia' juga menunjukkan nilai penting (*importance*) yang signifikan untuk RF,

sedangkan 'Intensity', 'Vertigo', dan 'Sensory' penting untuk ET. Perbedaan antara Gini Importance dan Permutation Importance terlihat, dengan Permutation Importance cenderung lebih mampu mengidentifikasi fitur-fitur yang benar-benar memengaruhi kinerja model pada data testing. Ringkasan fitur penting setelah SMOTE diterapkan ditunjukkan pada Tabel 4.

Tabel 4. Analisis Feature Importance Setelah Penerapan SMOTE

Fitur	RF+SMOTE Importance	ET+SMOTE Importance	Permutation Importance (RF+SMOTE)	Permutation Importance (ET+SMOTE)
Visual	0.112658	0.112584	0.16375	0.15875
Intensity	0.122157	0.135408	0.05125	0.06375
Frequency	0.066301	0.061500	0.06625	0.04500
Vertigo	0.080639	0.083927	0.04875	0.04250
Age	0.151736	0.114400	0.07500	0.04125
DPF	0.093646	0.088152	0.05750	0.03250
Sensory	0.022757	0.026499	0.01125	0.01500
Hypoacusis	0.009428	0.012318	0.00125	0.01375
Defect	0.003216	0.004027	0.01250	0.01250
Phonophobia	0.061775	0.069314	0.00625	0.00500
Dysphasia	0.015475	0.022114	0.00625	0.00500
Location	0.056387	0.062309	0.00250	0.00500
Vomit	0.013606	0.016743	0.01875	0.00375
Character	0.069043	0.073062	0.00500	0.00375
Duration	0.021204	0.020987	0.00875	0.00250
Tinnitus	0.013342	0.016611	-0.00375	0.00250
Photophobia	0.056713	0.046576	0.00000	0.00000
Nausea	0.027099	0.029477	0.00000	0.00000
Conscience	0.002067	0.002597	0.00000	0.00000
Diplopia	0.000551	0.001048	0.00000	0.00000
Paresthesia	0.000159	0.000254	0.00000	0.00000
Dysarthria	0.000042	0.000094	0.00000	0.00000
Ataxia	0.000000	0.000000	0.00000	0.00000

Setelah SMOTE, Permutation Importance menunjukkan bahwa 'Visual' masih dominan untuk RF+SMOTE (0.16375) dan ET+SMOTE (0.15875). Namun, ada perubahan signifikan pada peringkat fitur lainnya. Untuk RF+SMOTE, fitur 'Age', 'Frequency', 'DPF', 'Intensity', dan 'Vertigo' mengalami peningkatan nilai penting yang substansial. Ini menunjukkan bahwa SMOTE mengubah distribusi data sedemikian rupa sehingga model RF lebih mengandalkan fitur-fitur tersebut untuk klasifikasi kelas minoritas yang telah disintesis. Analisis delta Permutation Importance menunjukkan fitur-fitur yang mengalami perubahan paling besar setelah SMOTE ditunjukkan pada Gambar 4.



Gambar 4. Pergeseran Fitur penting (*Delta Permutation Importance*)

Tabel 5. Pergeseran Pergeseran terbesar

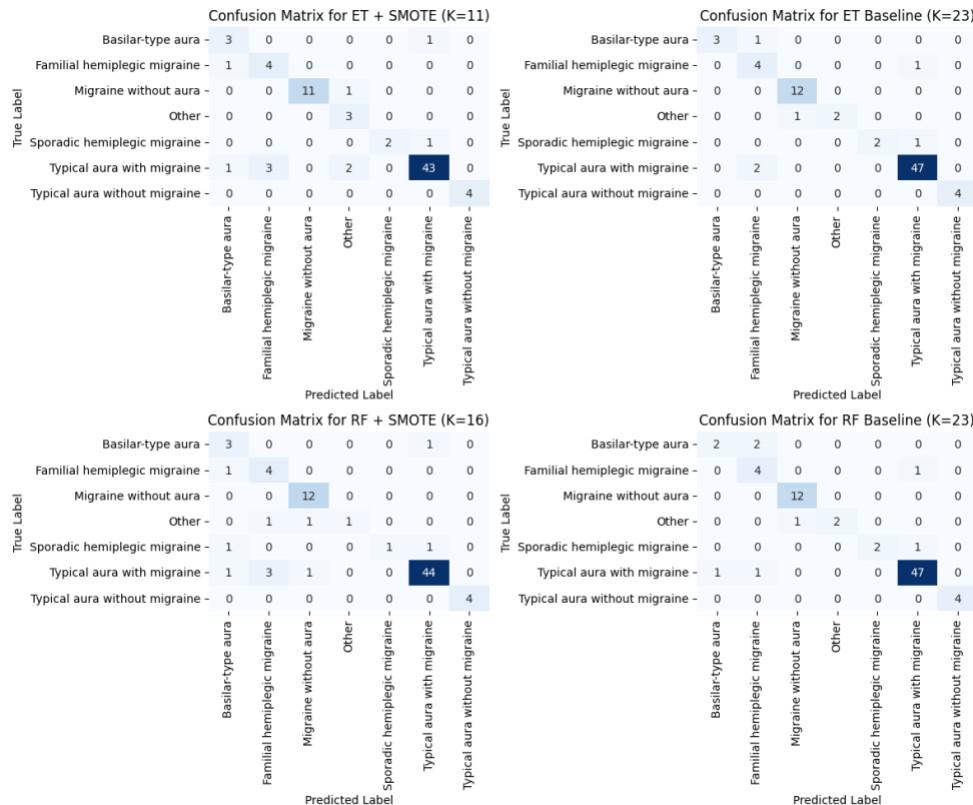
RF		ET	
Frequency	0.06500	Frequency	0.03500
Age	0.05125	Age	0.02750
DPF	0.03000	Tinnitus	0.02125
Intensity	0.02875	Dysphasia	0.01625
Tinnitus	0.03750	Intensity	0.01500
Vertigo	0.04125	DPF	0.01500
Visual	0.02125	Visual	0.03500
Duration	0.02250	Vertigo	0.00625
Dysphasia	0.01875	Location	0.00750
Hypoacusis	0.01125	Sensory	0.01125

Fitur-fitur seperti 'Frequency', 'Age', 'DPF', 'Intensity', 'Tinnitus', 'Vertigo', 'Visual', 'Duration', 'Dysphasia', dan 'Hypoacusis' menunjukkan pergeseran pentingnya (*importance*) yang paling signifikan untuk RF. Sementara untuk ET, fitur 'Frequency', 'Age', 'Tinnitus', 'Dysphasia', 'Intensity', 'DPF', 'Visual', 'Vertigo', 'Location', dan 'Sensory' mengalami perubahan paling besar. Perubahan ini mengindikasikan bahwa SMOTE tidak hanya menyeimbangkan kelas, tetapi juga mengubah bagaimana model menginterpretasikan hubungan antara fitur dan target, menyoroti fitur-fitur tertentu yang mungkin menjadi lebih diskriminatif setelah oversampling. Analisis untuk menemukan jumlah fitur ('K') optimal yang memaksimalkan rata-rata F1-Score dan Recall untuk kelas minoritas memperoleh hasil yang diperlihatkan pada Tabel 6.

Tabel 6. Hasil Analisa Optimal K-Features

K_Features	Model	Avg_Recall	Avg_F1_Score
11	ET + SMOTE	0.804167	0.700000
23	ET Baseline	0.720833	0.780952
16	RF + SMOTE	0.554167	0.553846
23	RF Baseline	0.658333	0.709524

ET + SMOTE mencapai rata-rata F1-Score tertinggi untuk kelas minoritas (0.7000) dengan hanya 11 fitur dan juga mencapai rata-rata recall tertinggi (0.8041). ET Baseline mencapai rata-rata F1-Score tertinggi secara keseluruhan (0.7809) dengan 23 fitur . RF + SMOTE mencapai rata-rata F1-Score (0.5538) dan Recall (0.5541) yang lebih rendah dengan 16 fitur. RF Baseline memiliki F1-Score 0.7095 dan Recall 0.6583 dengan 23 fitur. Untuk kelas minoritas, ET + SMOTE cenderung mencapai kinerja optimal dengan jumlah fitur yang lebih sedikit, yang mengindikasikan efisiensi dan kemampuan model untuk fokus pada fitur-fitur paling relevan setelah oversampling. Visualisasi confusion matrix untuk setiap model pada jumlah fitur optimalnya pada Gambar 4 memberikan gambaran detail tentang kinerja klasifikasi.



Gambar 5. Confusion Matrix untuk Model Optimal

Dari confusion matrices, terlihat bahwa model ET + SMOTE dengan 11 fitur cenderung memiliki angka True Positives yang lebih baik untuk beberapa kelas minoritas dibandingkan model lainnya, terutama 'Basilar-type aura' dan 'Familial hemiplegic migraine', yang berkorelasi dengan nilai recall yang lebih tinggi. Namun, ada juga peningkatan False Positives untuk beberapa kelas, yang menyebabkan penurunan presisi seperti yang terlihat pada laporan klasifikasi.

Tabel 7. Recall per Class

Model	ET + SMOTE	ET Baseline	RF + SMOTE	RF Baseline
Class				
Basilar-type aura	0.750000	0.750000	0.750000	0.500000
Familial hemiplegic migraine	0.800000	0.800000	0.800000	0.800000
Other	1.000000	0.666667	0.333333	0.666667
Sporadic hemiplegic migraine	0.666667	0.666667	0.333333	0.666667

Berdasarkan Tabel 7, ET + SMOTE menunjukkan recall yang sangat kuat untuk 'Basilar-type aura' (0.75) dan 'Familial hemiplegic migraine' (0.80), bahkan mencapai recall sempurna (1.00) untuk kelas 'Other' meskipun dengan presisi yang mungkin bervariasi. Hal ini menunjukkan bahwa SMOTE, terutama dikombinasikan dengan Extra Trees, dapat secara efektif meningkatkan kemampuan model untuk mendeteksi instansi positif dari kelas-kelas

minoritas tertentu, meskipun ini bisa datang dengan *trade-off* pada presisi atau kinerja pada kelas minoritas lainnya.

KESIMPULAN

Penelitian ini bertujuan untuk menganalisis dampak penerapan SMOTE terhadap nilai *feature importance* pada klasifikasi penderita migrain menggunakan Random Forest (RF) dan Extra Trees (ET). Hasil penelitian menunjukkan bahwa model ET, baik sebelum maupun sesudah SMOTE, memiliki performa yang lebih baik dibandingkan RF, terutama dalam mendeteksi kelas minoritas berdasarkan nilai *recall* dan *F1-Score*. Penerapan SMOTE juga memengaruhi nilai dan peringkat fitur penting, di mana fitur *Visual* tetap menjadi fitur paling dominan, sementara fitur seperti *Age*, *Frequency*, *DPF*, *Intensity*, dan *Vertigo* menunjukkan peningkatan kontribusi setelah oversampling. Temuan ini menunjukkan bahwa SMOTE tidak hanya membantu menyeimbangkan distribusi kelas, tetapi juga memengaruhi identifikasi fitur-fitur yang berkontribusi terhadap proses klasifikasi migrain.

DAFTAR PUSTAKA

- [1] "Migraine and other headache disorders," WHO. [Online]. Available: <https://www.who.int/en/news-room/fact-sheets/detail/headache-disorders>
- [2] "Global, regional, and national burden of disorders affecting the nervous system, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021.," *Lancet. Neurol.*, vol. 23, no. 4, pp. 344–381, Apr. 2024, doi: 10.1016/S1474-4422(24)00038-3.
- [3] A. X. Wang, V.-T. Le, H. N. Trung, and B. P. Nguyen, "Addressing imbalance in health data: Synthetic minority oversampling using deep learning," *Comput. Biol. Med.*, vol. 188, p. 109830, 2025, doi: <https://doi.org/10.1016/j.compbiomed.2025.109830>.
- [4] M. Salmi, D. Atif, D. Oliva, A. Abraham, and S. Ventura, *Handling imbalanced medical datasets : review of a decade of research*, vol. 57, no. 10. Springer Netherlands, 2024. doi: 10.1007/s10462-024-10884-2.
- [5] W. W. Y. Ng, Z. Liu, J. Zhang, and W. Pedrycz, "Maximizing minority accuracy for imbalanced pattern classification problems using cost-sensitive Localized Generalization Error Model," *Appl. Soft Comput.*, vol. 104, p. 107178, 2021, doi: <https://doi.org/10.1016/j.asoc.2021.107178>.
- [6] G. Pradipta, R. Wardoyo, A. Musdholifah, H. Sanjaya, and M. Ismail, *SMOTE for Handling Imbalanced Data Problem : A Review*. 2021. doi: 10.1109/ICIC54025.2021.9632912.
- [7] A. M. Musolf, E. R. Holzinger, J. D. Malley, and J. E. Bailey, "What makes a good prediction? Feature importance and beginning to open the black box of machine learning in genetics," *Hum. Genet.*, vol. 141, no. 9, pp. 1515–1528, 2022, doi: 10.1007/s00439-021-02402-z.
- [8] M. Saarela and S. Jauhiainen, "Comparison of feature importance measures as explanations for classification models," *SN Appl. Sci.*, vol. 3, no. 2, pp. 1–12, 2021, doi: 10.1007/s42452-021-04148-9.
- [9] Ö. Akarçay, N. Y. Pehlivan, G. W. Weber, and N. Y. Pehlivan, "A comparative analysis based on feature selection methods and machine learning algorithms for evaluating heart disease prediction performance," vol. 1, pp. 253–269, 2026, doi: 10.17341/gazimmfd.1513919.
- [10] H. Ghinaya, R. Herteno, M. R. Faisal, A. Farmadi, and F. Indriani, "Analysis of Important Features in Software Defect Prediction using Synthetic Minority Oversampling Techniques (SMOTE), Recursive Feature Elimination (RFE) and

- Random Forest,” vol. 6, no. 2, pp. 276–288, 2024.
- [11] J. Liu and M. Shi, “A Hybrid Feature Selection and Ensemble Approach to Identify Depressed Users in Online Social Media,” vol. 12, no. January, pp. 1–13, 2022, doi: 10.3389/fpsyg.2021.802821.
 - [12] L. Breiman, “Random Forests,” vol. 45, no. 1, 2001, doi: 10.1007/978-3-030-62008-0_35.
 - [13] P. Geurts, D. Ernst, and L. Wehenkel, “Extremely randomized trees,” *Mach. Learn.*, vol. 63, no. 1, pp. 3–42, 2006, doi: 10.1007/s10994-006-6226-1.
 - [14] E. Scornet, “Trees , forests , and impurity-based variable importance in regression,” pp. 1–40.
 - [15] A. I. Adler and A. Painsky, “Cross-Validation Feature Selection,” 2021.



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)
