

Kajian Performa Metode Class Weight Random Forest pada Klasifikasi Imbalance Data Kelas Curah Hujan

Hafiz Akbar

Program Studi Meteorologi
Sekolah Tinggi Meteorologi
Klimatologi dan Geofisika
Tangerang Selatan, Indonesia
hhafizzaakbarr2305@gmail.com

Wisnu Karya Sanjaya

Stasiun Meteorologi Soekarno-Hatta
Tangerang, Yogyakarta
me.wisnu@gmail.com

Abstraksi—Data hujan merupakan aspek penting dalam bidang hidrometeorologi dan penerbangan. Namun, data curah hujan yang dicatat oleh sistem cuaca otomatis yaitu Automatic Weather Station (AWS) selama satu tahun cenderung tidak seimbang, dengan kelas yang tidak hujan mendominasi. Hal ini mengakibatkan bias dalam klasifikasi, karena ketidakseimbangan antara kelas yang dominan dan yang tidak dominan. Random forest merupakan salah satu model machine learning yang dapat digunakan dalam proses klasifikasi dan dalam menangani masalah ketidakseimbangan data digunakan penambahan metode class weight pada model random forest dengan beberapa kombinasi nilai. Percobaan untuk pengaturan class weight dilakukan dengan berbagai pengaturan nilai, dan hasil kombinasi terbaik ditemukan pada pengaturan class weight {0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5} pada data latihan dan pada data uji, model dengan pengaturan class weight yang sama mencapai akurasi sebesar 73%.

Kata Kunci—hujan, klasifikasi, random forest, class weight

I. PENDAHULUAN

Wilayah Indonesia merupakan wilayah kepulauan yang terletak di wilayah ekuatorial dimana intensitas dari penyinaran matahari yang didapatkan lebih banyak, intensitas penyinaran matahari yang besar dapat menjadi faktor pemicu pertumbuhan awan-awan konvektif yang menjadi penyebab peningkatan curah hujan. Bencana hidrometeorologi dapat berkaitan dengan curah hujan [1]. Komponen penting dari sistem iklim dan siklus air global adalah presipitasi. Uap air mengembun menjadi awan, dimana presipitasi dapat turun, sebagai bagian dari siklus air. Proses ini membawa air dari atmosfer ke permukaan dan berfungsi sebagai sumber utama air tawar, yang penting bagi kehidupan di bumi. Kesejahteraan manusia dipengaruhi pada pemantauan dan pengukuran curah hujan, kurangnya curah hujan mengakibatkan kekeringan yang mempengaruhi pertanian dan dapat menyebabkan kelaparan, sementara curah hujan yang berlebihan membahayakan keselamatan dan harta benda [2].

Pengukuran curah hujan dapat dilakukan menggunakan berbagai alat, dari yang bersifat manual maupun otomatis. Sebuah stasiun cuaca yang menggunakan sensor otomatis disebut juga Automatic Weather Station (AWS) yang memantau dan mencatat beberapa parameter cuaca meteorologi yaitu curah hujan, kecepatan angin, suhu udara

maksimum, suhu udara minimum, suhu udara rata-rata, rata-rata kelembapan udara, tekanan udara, evaporasi. Sensor ini berfungsi sebagai alat ukur untuk mendeteksi fluktuasi perubahan parameter cuaca. Data logger akan mengolah data keluaran sensor sebelum dikirim ke stasiun cuaca BMKG, yang dapat diakses oleh pengguna atau masyarakat umum yang membutuhkan data stasiun cuaca [3]. Data akan disimpan di data logger selain dikirim langsung ke web server sebagai metode backup [4].

Data AWS yang mencatat beragam parameter cuaca dapat digunakan sebagai bahan analisis untuk klasifikasi. Klasifikasi diperlukan untuk mengidentifikasi konsep dan pola kelas yang saling terkait antar satu kelas dengan kelas yang lainnya [5]. Teknik klasifikasi data memiliki banyak metode seiring perkembangan zaman, salah satunya dengan memanfaatkan metode pendekatan statistik seperti machine learning (ML). ML adalah suatu pemrograman komputer untuk mengoptimalkan kinerja kriteria dengan menggunakan contoh data yang diberikan dengan menerapkan teori statistik dalam membangun model matematika, sehingga memberikan kesimpulan berdasarkan pembelajaran data dan metode pembelajaran yang dilakukan oleh ML salah satunya adalah supervised learning dimana pembelajaran yang sifatnya terawasi dengan tujuan mempelajari pemetaan dari input ke output yang bernilai benar[6]. Supervised learning pada data AWS data digunakan dengan memanfaatkan parameter cuaca untuk pelatihan model ML pada klasifikasi intensitas curah hujan. Salah satu algoritma ML yang termasuk kedalam supervised learning adalah Random Forest (RF). RF adalah kombinasi prediktor pohon sedemikian rupa sehingga setiap pohon bergantung pada nilai vektor acak yang disampel secara independen dan dengan distribusi yang sama untuk semua pohon di dalam hutan[7]. Secara statistis, RF memiliki beberapa kelebihan yaitu banyak fitur tambahan yang bisa dikonfigurasi seperti tingkat variabel, pembobotan kelas, mengatasi missing value dan visualisasi [8].

Beberapa penelitian terkait RF dalam klasifikasi imbalance data sudah pernah dilakukan. Perbandingan metode RF dan XGBOOST seperti di [9] menggunakan metode SMOTE pada RF dan XGBOOST untuk klasifikasi tingkat penyakit hepatitis C, hasil yang didapatkan bahwa penambahan metode SMOTE meningkatkan nilai akurasi pada RF dan XGBOOST dengan model RF memiliki nilai akurasi yang sedikit lebih tinggi dibandingkan XGBOOST.

Perbandingan metode oversampling, undersampling dan SMOTE seperti di [10] untuk klasifikasi penerima bidikmisi sejawa timur tahun 2017 menggunakan SVM dan RF, hasil yang didapatkan adalah akurasi yang didapatkan tanpa handling pada data input memiliki akurasi sebesar 96% namun 0 pada variabel yang mengindikasikan bahwa model ini masih kurang bagus, setelah dilakukan oversampling dan SMOTE pada data hasil SVM dan RF memiliki akurasi yang mirip dengan adanya peningkatan nilai recall pada kedua model tersebut. Penerapan metode Random Over-under sampling dan RF pada klasifikasi imbalance data pernah dilakukan [11] untuk dataset german credit, hasil yang didapatkan adalah dengan membandingkan dengan metode SMOTE pada data yang menunjukkan bahwa metode Random Over-under sampling memiliki nilai akurasi yang lebih tinggi dibandingkan SMOTE. Selain menggunakan teknik resampling seperti oversampling dan undersampling terdapat metode yang memanfaatkan class weight seperti di [12] dan [13]. Pada kasus multiclass classification [14] RF memiliki performa klasifikasi yang lebih baik. Penelitian ini bertujuan melihat performa algoritma random forest dengan memanfaatkan metode pembobotan pada klasifikasi kelas curah hujan berdasarkan kriteria BMKG Tahun 2020 menggunakan data AWS pada stasiun meteorologi Soekarno-Hatta tahun 2022. Sehingga berdasarkan penelitian yang telah dilakukan sebelumnya bahwa RF dapat digunakan membantu proses klasifikasi.

II. KAJIAN TEORI

A. Hujan

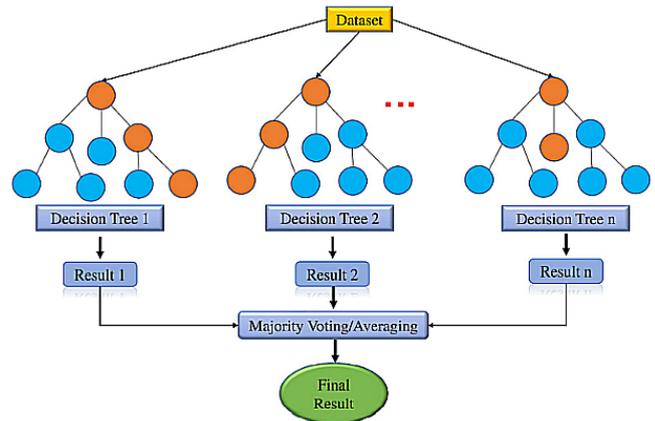
Hujan merujuk pada zat yang terdiri dari tetesan air cair dengan diameter setidaknya 0,5 mm, drizzle merujuk pada tetesan air dengan diameter kurang dari 0,5 mm yang terbentuk dari awan stratus atau hujan yang tidak sepenuhnya tergenang di udara sebelum jatuh ke bumi, sedangkan virga merujuk pada hujan yang terpantulkan sebelum mencapai permukaan bumi karena udara yang terlalu kering[15]. Hujan merupakan hal yang penting bagi kehidupan dan lingkungan di bumi dan ketika intensitas curah hujan mengalami peningkatan intensitas dapat menyebabkan banjir, sementara curah hujan yang rendah dapat menyebabkan kekeringan. Faktor posisi lintang wilayah, pola pergerakan angin, bentuk permukaan bumi dan sebaran wilayah laut dan daratan memberikan dampak terhadap perbedaan curah hujan yang terdapat di Indonesia[16]. Tabel 1 merupakan pembagian kelas intensitas curah hujan BMKG [17].

TABEL 1. KLASIFIKASI UNTUK INTENSITAS CURAH HUJAN

Kategori	Intensitas Curah Hujan	
	Setiap Jam (mm/jam)	Setiap Hari (mm/hari)
Tidak Hujan	-	-
Hujan Ringan	0.5 - 5.0	5.0 - 20.0
Hujan Sedang	5.0 - 10.0	20.0 - 50.0
Hujan Lebat	10.0 - 20.0	50.0 - 100.0
Hujan Sangat Lebat	>20.0	>100.0

B. Random Forest

Random forest (RF) adalah sistem klasifikasi yang tersusun atas kumpulan pohon klasifikasi yang terstruktur dimana adalah vektor acak independen yang terdistribusi identik dan setiap pohon memberikan suara unit untuk kelas yang paling populer pada input x [7]. Algoritma RF memiliki beberapa parameter yang dapat dikonfigurasi, di antaranya adalah jumlah decision tree yang akan dibuat, kriteria yang digunakan untuk mengevaluasi pemisahan dalam setiap node, batas maksimal dari kedalaman decision tree, lihat Gambar 1.



Gambar 1. Visualisasi Random Forest [18].

1) *Jumlah Decision Tree*: Jumlah pohon keputusan (decision tree) adalah parameter dalam algoritma RF yang menentukan jumlah pohon yang terdapat pada model RF dalam membantu mengambil keputusan dan prediksi klasifikasi. Banyaknya jumlah pohon dalam RF akan berpengaruh terhadap proses dan akurasi perhitungan. Peningkatan performa pada model ini dapat dipengaruhi oleh banyaknya jumlah pohon keputusan [19].

2) *Kriteria Pemisahan Data*: Kriteria pemisahan data atau criterion (C) adalah nilai yang besar mengindikasikan kecocokan model yang buruk (regresi) atau klasifikasi node yang salah untuk setiap masing-masing output berdasarkan kriteria pemisahan, dimana mengukur kualitas kecocokan model node (regresi) atau kemurnian node (klasifikasi) [8].

3) *Max Depth*: Max depth adalah parameter yang terdapat pada algoritma RF yang menentukan kedalaman maksimum dari pohon dalam hutan. Setiap tingkat kerumitan perhitungan pada pohon keputusan dibatasi oleh kedalaman pohon maksimum, dan seiring dengan bertambahnya kedalaman pohon, begitu juga dengan biaya komputasi (waktu berjalan) [20].

C. Imbalance Data

Imbalance data atau data yang tidak seimbang adalah kondisi di mana terdapat ketidakseimbangan jumlah antara kelas atau label dalam suatu dataset. Hal ini dapat menyebabkan masalah dalam pemodelan machine learning karena model cenderung lebih baik memprediksi kelas yang lebih banyak muncul dibandingkan dengan kelas yang lebih sedikit. Tugas-tugas pengolahan dan analisis data seperti klasifikasi, clustering, prediksi, dan lainnya menjadi tidak mudah apabila satu kelas memiliki jumlah data yang jauh

lebih tinggi atau lebih rendah dibandingkan dengan kelas lain, sehingga dapat menyebabkan imbalance data [21]. Class weight adalah teknik yang dapat digunakan dalam pembelajaran mesin untuk mengatasi kelas yang tidak seimbang dengan cara memberikan bobot yang lebih tinggi pada sampel dari kelas minoritas. Permasalahan dapat timbul dalam menilai classifier secara adil karena ketidakseimbangan distribusi kelas pada dataset yang tidak seimbang, sehingga menyebabkan terjadinya bias prediksi pada model terhadap kelas mayoritas [22].

III. METODE

Penelitian ini menggunakan metode yang bersifat eksperimental dengan tujuan untuk menguji dengan mengontrol variabel-variabel yang dapat digunakan sebagai data input untuk klasifikasi kelas curah hujan antara lain kecepatan angin rata-rata, permukaan dan maksimum, suhu udara minimum, rata-rata dan maksimum, intensitas penyinaran rata-rata matahari dan evaporasi. Data yang digunakan yaitu data observasi permukaan dari AWS yang terdapat pada stasiun meteorologi Cengkareng. Data AWS pada penelitian ini merupakan data hujan dengan resolusi temporal 1 jam yang terkumpul dari tanggal 1 Januari 2022 hingga 31 Maret 2023. Data ini diperoleh dari website awscenter BMKG yang menyediakan akses ke data hujan dari AWS stasiun meteorologi Cengkareng dan dapat diakses melalui link <https://awscenter.bmkg.go.id/>. Variabel curah hujan yang terdapat pada data AWS digunakan sebagai data dependen dan variabel lainnya seperti kecepatan angin, tekanan udara, penguapan, kelembaban relative dan penyinaran matahari menjadi data imput latih dan pada model prediksi yang akan dibangun dalam penelitian ini.

A. Teknik Pengolahan Data

Berikut adalah beberapa proses pengolahan data:

1) *Mengatasi Missing Data*: Mengatasi missing data atau handling missing data dilakukan sebelum pembangunan model dimulai dan memastikan bahwa data yang digunakan untuk analisis atau pemodelan merupakan data yang lengkap dan akurat. Data tersebut kemudian dibagi menjadi lima kelas intensitas hujan berdasarkan kategori klasifikasi intensitas curah hujan [17]. Namun, dalam proses pengumpulan data terdapat nilai-nilai yang hilang atau tidak lengkap. Oleh karena itu, dilakukan pembersihan data dengan cara menghapus setiap baris yang memiliki nilai kosong sebelum data tersebut diproses lebih lanjut dan dimasukkan ke dalam model. Informasi untuk data mentah ada pada Tabel 2 dan 3.

TABEL 2. JUMLAH DATA SEBELUM HANDLING

Data	Jumlah	Tipe data
rr	8416	float64
ws_avg	8322	float64
ws_max	8322	float64
tt_air_max	8374	float64
tt_air_avg	8374	float64
tt_air_min	8374	float64
rh_avg	8374	float64

pp_air	8376	float64
sr_avg	8377	float64
tt_sea	8376	float64
ws_50cm	8322	float64
ev_pan	8190	float64

Adapun jumlah data setelah dilakukan proses handling missing data terlihat pada Tabel 4.

Data curah hujan yang sudah dikumpulkan kemudian dikelompokkan berdasarkan kategori dari BMKG menjadi 5 kelas intensitas hujan pada data training dan data testing.

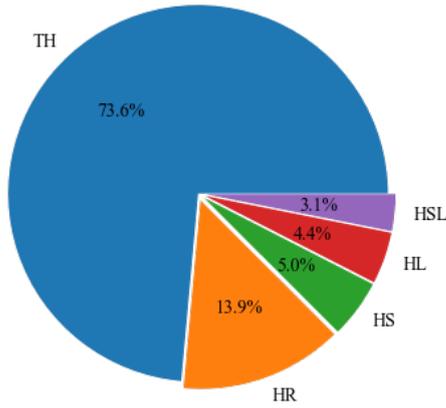
TABEL 3. JUMLAH MISSING DATA

Data	Jumlah
rr	306
ws_avg	400
ws_max	400
tt_air_max	348
tt_air_avg	348
tt_air_min	348
rh_avg	348
pp_air	346
sr_avg	345
tt_sea	346
ws_50cm	400
ev_pan	532

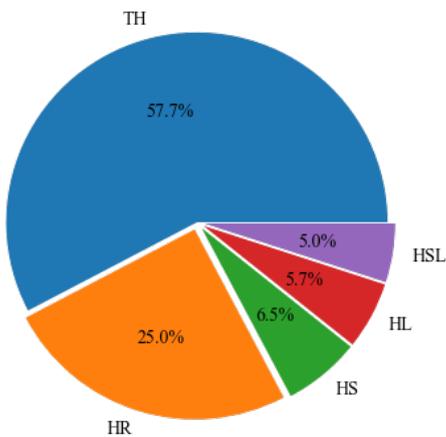
TABEL 4. JUMLAH DATA SETELAH HANDLING

Data	Jumlah	Tipe data
rr	8114	float64
ws_avg	8114	float64
ws_max	8114	float64
tt_air_max	8114	float64
tt_air_avg	8114	float64
tt_air_min	8114	float64
rh_avg	8114	float64
pp_air	8114	float64
sr_avg	8114	float64
tt_sea	8114	float64
ws_50cm	8114	float64
ev_pan	8114	float64

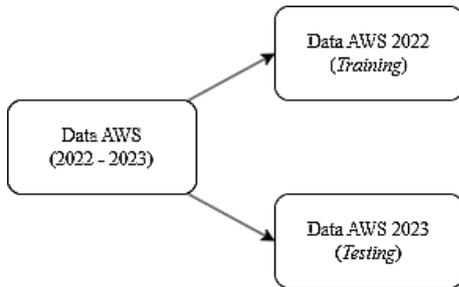
2) *Pembagian dan Normalisasi Data*: Pembangunan model pada data yang telah dikumpulkan dibagi menjadi dua bagian, yaitu data latih (training) dan data uji (testing) pada data AWS tahun 2022 digunakan sebagai data training Sedangkan untuk data AWS pada periode 1 Januari 2023 hingga 31 Maret 2023 digunakan sebagai data testing dan evaluasi dari hasil pelatihan model terbaik. Lihat Gambar 2, 3, dan 4.



Gambar 2. Frekuensi Kelas Hujan Training Data



Gambar 3. Frekuensi Kelas Hujan Testing Data



Gambar 4. Pembagian Dataset Training dan Testing

Proses selanjutnya adalah normalisasi data dengan tujuan mengubah ukuran data menjadi lebih kecil daripada data asli, namun tetap menjaga karakteristik data yang sama sehingga rentang nilai data dari setiap parameter akan menjadi sama. Selain itu, normalisasi data juga bertujuan untuk mengurangi waktu komputasi yang dibutuhkan dalam proses pemodelan. Teknik min and max scalling dilakukan pada proses normalisasi data penelitian ini yang merujuk pada penelitian [23].

$$X_{norm} = (X - X_{min}) / (X_{max} - X_{min})$$

Keterangan:

X_{norm} = data normalisasi

X = nilai *input* parameter yang akan dinormalisasi

X_{max} = nilai terbesar dalam seluruh data *input*

X_{min} = nilai terkecil dalam seluruh data *input*

3) *Penyusunan Konfigurasi Nilai Hyperparameter:*

Hyperparameter adalah parameter dalam algoritma ML yang tidak diatur oleh model itu sendiri, tetapi harus diatur oleh pengguna sebelum proses pelatihan dimulai. Hyperparameter sangat penting dalam algoritma ML karena dapat mempengaruhi kinerja model dan hasil yang dihasilkan. Pada penelitian ini variabel parameter yang akan dikonfigurasi terdapat pada Tabel 5.

TABEL 5. HYPERPARAMETER MODEL RF

Model	Hyperparameter	Nilai
Random Forest	Number of Decision Tree	100, 200, 300, 400, 500
	Criterion	Gini
		Entropy
	Max Depth	5, 10, 15, 20
Class Weight	Balanced, {0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1}, {0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5}	

Hyperparameter class weight adalah parameter yang digunakan untuk menyeimbangkan bobot kelas dalam proses pembelajaran model saat melakukan proses training yang memungkinkan pemberian bobot yang berbeda kepada setiap kelas dalam dataset dengan frekuensi kelas yang rendah. Bobot kelas ini mempengaruhi perhitungan kerugian (loss) dalam proses pembelajaran dan memberikan penalti yang lebih besar untuk kesalahan klasifikasi pada kelas-kelas minoritas. Nilai dari parameter class weight balanced untuk setiap kelasnya dihitung dari banyak sample dibagi dengan kelas yang dikalikan dengan banyak jumlah kelasnya sedangkan untuk nilai {0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1}, {0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5} merupakan bobot yang kita tentukan sendiri untuk kelasnya, dimana dalam klasifikasi terdapat 5 kelas hujan dan angka 0 hingga 4 merupakan perwakilan dari kelas tidak hujan (TH), Hujan Ringan (HR), Hujan Sedang (HS), Hujan Lebat (HL) dan Hujan Sangat Lebat (HSL). Beberapa parameter yang ditentukan pada model kemudian dilakukan proses hyperparameter tuning menggunakan metode random search.

4) *K-Fold Cross Validation:* K-Fold cross validation adalah teknik untuk mengevaluasi performa model dalam dengan memisahkan menjadi beberapa bagian yang besarnya sama untuk dilakukan prediksi secara acak dan bergantian. K-Fold cross validation bekerja dengan membagi data menjadi k subset (k = jumlah fold yang ditentukan). Dengan cara ini, K-Fold Cross Validation dapat

memberikan estimasi performa model yang lebih akurat dan mengurangi kemungkinan terjadinya overfitting atau underfitting. Jumlah pembagian data pada penelitian ini merujuk pada penelitian [24] dan [25] yang membagi data menjadi 5 bagian untuk proses K-Fold cross validation.

B. Verifikasi

Pada penelitian ini verifikasi dilakukan menggunakan confusion matrix, nilai yang didapat dari confusion matrix digunakan untuk menghitung beberapa parameter untuk melihat akurasi performa model dalam klasifikasi, lihat Tabel 6.

TABEL 6. CONFUSION MATRIX

	<i>Predicted Negative</i>	<i>Predicted Positive</i>
<i>Actual Negative</i>	True Negative (TN)	False Positive (FP)
<i>Actual Positive</i>	False Negative (FN)	True Positive (TP)

Dengan menggunakan confusion matrix, Nilai yang dapat diperoleh dari True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) untuk mengukur performa suatu model:

1) *Akurasi*: Akurasi (Accuracy) adalah proporsi dari instance yang terprediksi dengan benar oleh model. Akurasi dapat dihitung dengan rumus:

$$\text{Akurasi} = (TP + TN) / (TP + TN + FP + FN)$$

2) *Presisi*: Presisi (Precision) adalah proporsi dari instance yang terprediksi sebagai positif yang sebenarnya positif. Presisi dapat dihitung dengan rumus:

$$\text{Presisi} = (TP) / (TP + FP)$$

3) *Sensitivitas*: Sensitivitas (Recall) adalah proporsi dari instance yang sebenarnya positif yang terprediksi sebagai positif oleh model. Sensitivitas dapat dihitung dengan rumus:

$$\text{Sensitivitas} = (TP) / (TP + FN)$$

4) *F1 Score*: Nilai F1 Score adalah rata-rata harmoni dari presisi dan sensitivitas. F1 Score dapat dihitung dengan rumus:

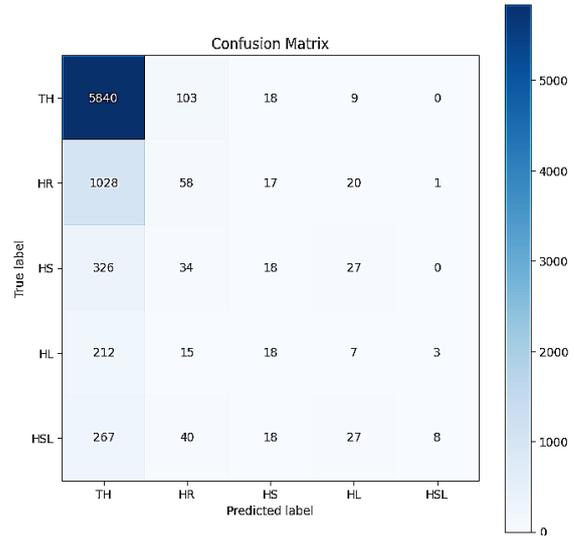
$$\text{F1 score} = 2 * (\text{presisi} * \text{sensitivitas}) / (\text{presisi} + \text{sensitivitas})$$

5) *PR Curve*: Precision-Recall curve (PR curve) menampilkan Precision, yaitu proporsi observasi dengan nilai prediksi positif yang benar-benar positif, dibandingkan dengan Recall, yang mengukur proporsi contoh dengan label positif yang mendapatkan hasil prediksi positif [26]. PR memberikan informasi tentang seberapa baik model dapat mengklasifikasikan kelas minoritas (kelas positif) dengan benar.

IV. HASIL DAN PEMBAHASAN

A. Hasil Performa Model Pada Klasifikasi Data Training

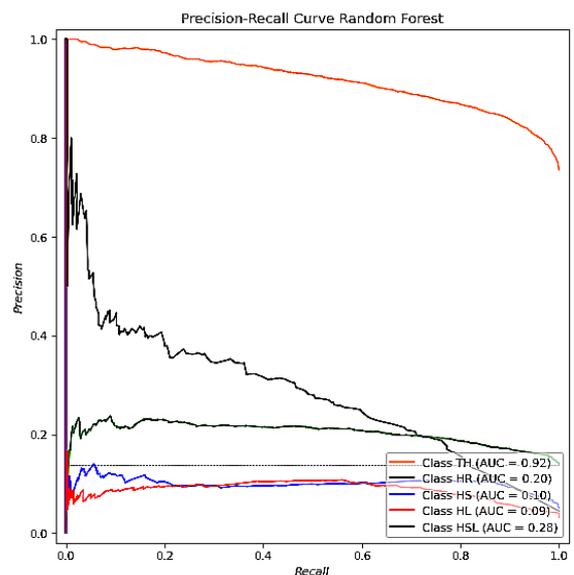
1) *Class Weight (Balance)*: Berdasarkan confusion matrix model RF pada Gambar 5 dengan menggunakan metode class weight balance pada kelas hujan, nilai metrics yang diperoleh terlihat pada Tabel 7.



Gambar 5. Confusion Matrix Skenario 1 (Balance)

TABEL 7. NILAI METRICS UNTUK SKENARIO 1 (BALANCE)

	Precision	Recall	F1 score
TH	0.76	0.98	0.86
HR	0.23	0.05	0.08
HS	0.20	0.04	0.07
HL	0.08	0.03	0.04
HSL	0.67	0.02	0.04

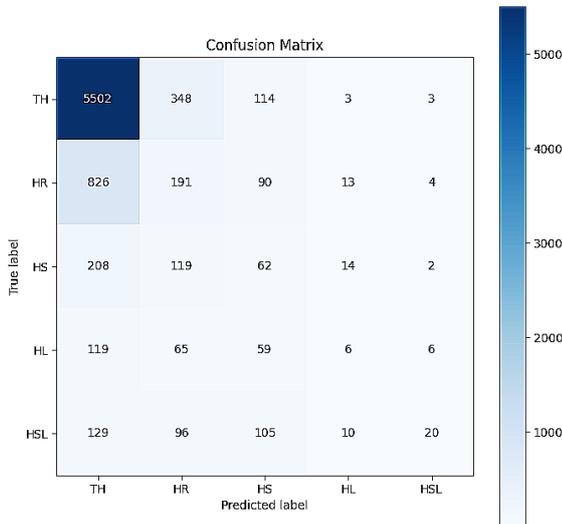


Gambar 6. PR Curve Skenario 1 (Balance)

Dari Tabel 7, terlihat akurasi yang didapatkan yaitu 73% secara keseluruhan dengan nilai precision kelas TH memiliki nilai yang paling tinggi yaitu 0.76 mengindikasikan model dapat memprediksi kelas TH dengan baik. Sedangkan kelas HR, HS, dan HL memiliki nilai precision yang rendah 0.23, 0.2 dan 0.08 mengindikasikan model belum mampu dalam memprediksi kelas ini dan kelas HSL memiliki nilai precision yang cukup tinggi yaitu 0.67, namun tetap di bawah kelas TH. Nilai recall kelas TH memiliki nilai tertinggi yaitu 0.98, mengindikasikan model sangat baik dalam mengklasifikasikan kelas TH secara benar. Namun pada kelas HR, HS, HL, dan HSL memiliki nilai recall yang sangat rendah menunjukkan model belum mampu dalam mengklasifikasikan kelas-kelas tersebut secara tepat dan benar. Sedangkan untuk metrics F1 score kelas TH memiliki nilai yang paling tinggi yaitu 0.86 menunjukkan model dapat memprediksi dan mengklasifikasikan kelas TH dengan baik. Sedangkan kelas HR, HS, HL, dan HSL memiliki nilai F1 score yang sangat rendah menunjukkan model tidak efektif dalam memprediksi dan mengklasifikasikan kelas-kelas tersebut.

Berdasarkan PR Curve pada Gambar 6, dapat dilihat bahwa model dengan class weight balance hanya dominan memprediksi kelas TH sedangkan pada kelas HR, HS, HL dan HSL kurang baik, hal ini ditandai dengan nilai auc kelas TH yang tinggi namun kelas lainnya sangat rendah.

2) *Class Weight* ({0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1}): Berdasarkan gambar confusion matrix model RF (Gambar 7) dengan menggunakan metode class weight {0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1} pada kelas hujan memiliki nilai metrics seperti pada Tabel 8.

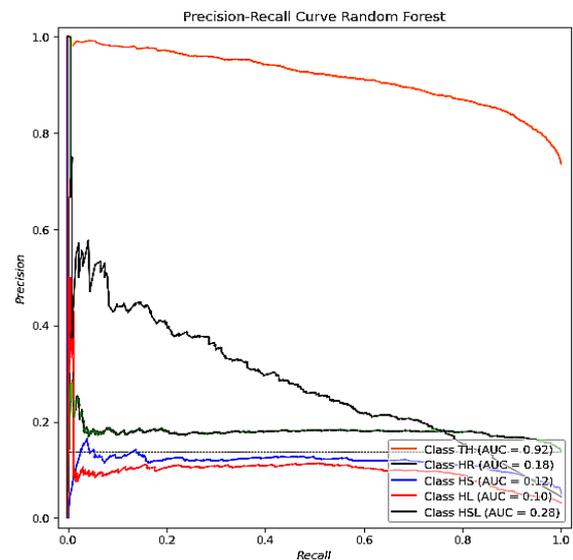


Gambar 7. Confusion Matrix Skenario 2

TABEL 8. NILAI METRICS UNTUK SKENARIO 2

	Precision	Recall	F1-score
TH	0.81	0.92	0.86
HR	0.23	0.17	0.20
HS	0.14	0.15	0.15
HL	0.13	0.02	0.04
HSL	0.57	0.06	0.10

Akurasi yang didapatkan yaitu 71%, sedikit lebih rendah dibandingkan class weight {0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1}, namun dengan nilai precision kelas TH memiliki nilai yang lebih tinggi yaitu 0.81. Sedangkan kelas HR, HS, dan HL memiliki nilai precision yang lebih baik dibandingkan sebelumnya yaitu 0.23, 0.14 dan 0.13 namun hal ini mengindikasikan model masih belum mampu dalam memprediksi kelas ini dan kelas HSL memiliki nilai precision yang menurun yaitu 0.57. Nilai recall kelas TH sedikit mengalami penurunan menjadi 0.92, mengindikasikan model dapat mengklasifikasikan kelas TH secara benar. Sedangkan pada kelas HR, HS, HL, dan HSL memiliki nilai recall yang sedikit meningkat. Sedangkan untuk metrics F1 score kelas TH memiliki nilai yang tidak berubah dari metode sebelumnya yaitu 0.86. Sedangkan kelas HR, HS, dan HSL memiliki nilai F1 score yang lebih tinggi mengindikasikan bahwa dengan menggunakan nilai parameter ini memiliki kemampuan klasifikasi lebih baik dari parameter balance, namun pada kelas HL tetap memiliki nilai yang sama. Secara keseluruhan terdapat beberapa kenaikan dalam nilai precision, recall dan F1-score pada kelas HR, HS, HL dan HSL



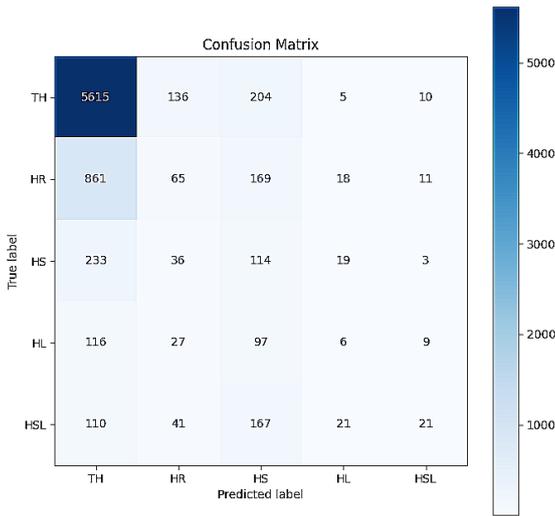
Gambar 8. PR Curve Skenario 2

Berdasarkan PR Curve pada Gambar 8, model dengan class weight {0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1} hanya dominan memprediksi kelas TH sedangkan pada kelas HR, HS, HL dan HSL kurang baik meskipun mengalami kenaikan pada metrics precision dan recall pada beberapa kelas

3) *Class Weight* ({0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5}): Berdasarkan gambar confusion matrix model RF (Gambar 9) dengan menggunakan metode class weight {0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5} pada kelas hujan memiliki nilai metrics seperti pada Tabel 9.

Akurasi yang didapatkan yaitu 72%, sedikit lebih rendah dibandingkan class weight balance, dengan nilai precision pada setiap kelas kecuali TH meningkat namun sedikit menurun pada kelas HR, HL dan HSL untuk class weight {0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1} menjadi 0.21, 0.09 dan 0.39 sedangkan untuk kelas HS mengalami kenaikan nilai sebesar 0.01. Pada metrics recall didapatkan nilai yang lebih tinggi dari kelas balance dan sedikit lebih rendah dari class weight

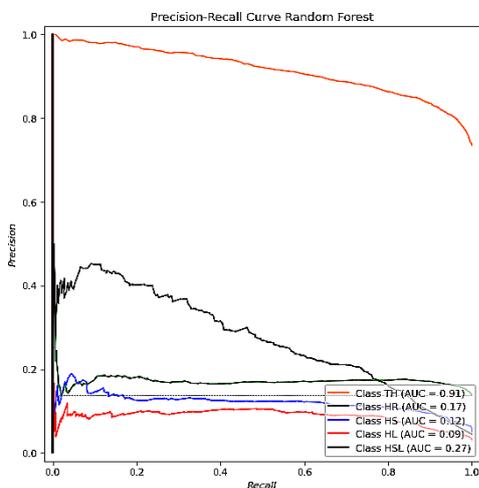
{0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1} pada kelas HR dan HL namun pada kelas TH dan HSL mengalami kenaikan dan kelas HL tetap menghasilkan nilai recall yang sama. Pada metrics F1 score mendapatkan nilai yang lebih tinggi dari class weight balance dan nilai yang lebih rendah pada kelas HR dari class weight {0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1} sedangkan pada kelas TH dan HS mendapatkan nilai yang lebih tinggi. Secara umum pada class weight {0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5} mendapatkan nilai yang lebih rendah dibandingkan dengan class weight {0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1} namun lebih baik dibandingkan dengan class weight balance, hal ini dapat dilihat dari tinggi nilai pada satu kelas dan rendah pada kelas-kelas lain untuk parameter recall dan precision.



Gambar 9. Confusion Matrix Skenario 3

TABEL 9. NILAI METRICS UNTUK SKENARIO 3

	Precision	Recall	F1-score
TH	0.81	0.94	0.87
HR	0.21	0.06	0.09
HS	0.15	0.28	0.20
HL	0.09	0.02	0.04
HSL	0.39	0.06	0.10

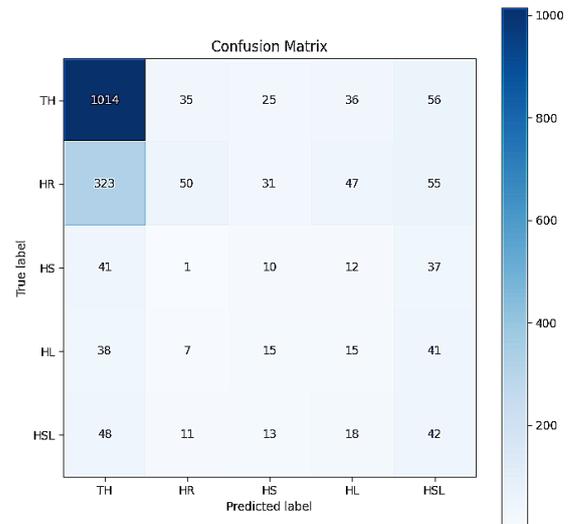


Gambar 10. PR curve Skenario 3

Berdasarkan PR Curve Gambar 10, model dengan class weight {0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5} hanya dominan memprediksi kelas TH sedangkan pada kelas HR, HS, HL dan HSL kurang baik. Pada model ini nilai auc pada beberapa kelas mengalami penurunan dibandingkan 2 model sebelumnya. Perbandingan nilai metrics yang didapatkan dari 3 model sebelumnya yang menggunakan metode class weight, {0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5} memiliki hasil yang paling baik dibandingkan 2 percobaan lainnya.

B. Penerapan Performa Model Terbaik pada Klasifikasi Data Testing

Confusion matrix model RF (Gambar 11) pada data testing dengan menggunakan metode class weight {0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5} pada kelas hujan memiliki nilai metrics yang terdapat pada Tabel 10.



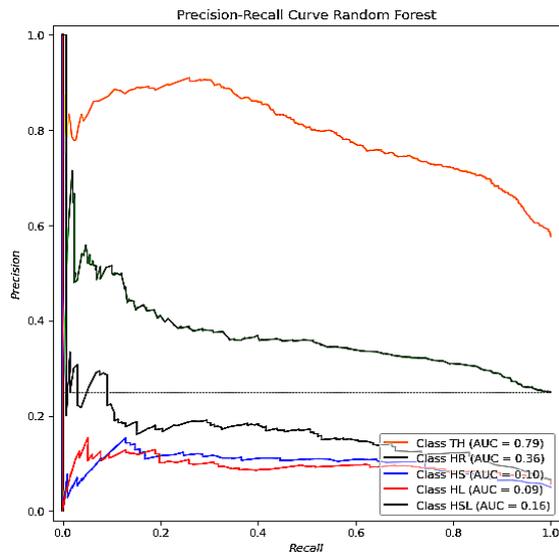
Gambar 11. Confusion Matrix Data Testing

TABEL 10. NILAI METRICS UNTUK DATA TESTING

	Precision	Recall	F1-score
TH	0.76	0.98	0.86
HR	0.23	0.05	0.08
HS	0.20	0.04	0.07
HL	0.08	0.03	0.04
HSL	0.67	0.02	0.04

Kelas TH memperoleh hasil yang cukup baik dengan nilai precision sebesar 0.76 dan recall sebesar 0.98, serta F1 score sebesar 0.86, hal ini mengindikasikan bahwa model dapat mengklasifikasikan data ke dalam kelas TH dengan cukup baik. Namun, pada hasil kelas HR, HS, HL, dan HSL menunjukkan performa yang buruk. Kelas-kelas tersebut memiliki nilai precision rendah di bawah 0.25, menunjukkan model cenderung mengklasifikasikan data ke dalam kelas-kelas tersebut sebagai false positive. Sementara itu, nilai Recall yang rendah, yaitu di bawah 0.05, menunjukkan model memiliki kesulitan untuk mengidentifikasi data yang sebenarnya masuk ke dalam kelas-kelas tersebut.. secara keseluruhan mode tetap hanya baik untuk klasifikasi kelas tidak hujan namun buruk pada kelas lainnya, hal ini juga

dapat dilihat pada nilai auc dari PR curve (Gambar 12), kelas TH memiliki nilai yang paling tinggi diikuti dengan kelas HR, HS, HSL dan HR.



Gambar 12. PR curve data testing

V. KESIMPULAN

Perbandingan nilai metrics yang dilakukan pada model dengan pengaturan metode class weight yang berbeda yaitu Balanced, {0: 0.5, 1: 1, 2: 1, 3: 1, 4: 1}, {0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5} dimana nilai class weight {0: 0.5, 1: 1.5, 2: 1.5, 3: 1.5, 4: 1.5} memiliki hasil yang paling baik pada training data dan pada testing data memiliki akurasi 73%, namun belum mampu dengan baik mengatasi ketidakseimbangan data berdasarkan nilai AUC pada PR curve.

DAFTAR PUSTAKA

[1] N. H. Muzaki, I. N. Wijaya, M. Agritami, and R. F. Ramdani, "Perbandingan Produk Estimasi Curah Hujan Global Precipitation Measurement (GPM) Dan Global Satellite Mapping of Precipitation (GSMaP) Di Wilayah Kalimantan Tahun 2020," vol. 1, pp. 19–31, 2022.

[2] C. Kidd and G. Huffman, "Global precipitation measurement," *Meteorol. Appl.*, vol. 18, no. 3, pp. 334–353, 2011, doi: 10.1002/met.284.

[3] Z. H. Salindri, Darjat, and A. M. Riyadi, "Rancang Bangun Mini Weather Station menggunakan Web Berbasis Arduino ATmega 2560," *Transient*, vol. 4, no. 4, pp. 1079–1086, 2015.

[4] H. N. Isnianto and E. Puspitaningrum, "Stasiun Cuaca Secara Nirkabel Berbasis," *Semin. Nas. Inform.*, vol. 2012, no. semnasIF, pp. 51–58, 2012.

[5] L. M. Chan, "Classification , Present and Future Classification , Present and Future," vol. 9374, no. 1996, pp. 37–41, 2009, doi: 10.1300/J104v21n02.

[6] E. Alpaydin, *Introduction to Machine Learning Second Edition*. MIT press, 2020.

[7] L. Breiman, "Random Forests," 2001 Kluwer Acad. Publ. Manuf. Netherlands., 2001, doi: 10.1109/ICCECE51280.2021.9342376.

[8] A. Cutler and D. R. Cutler, "Ensemble Machine Learning," *Ensemble Mach. Learn.*, no. January, 2012, doi: 10.1007/978-1-4419-9326-7.

[9] M. Syukron, R. Santoso, and T. Widiari, "Perbandingan Metode Smote Random Forest Dan Smote Xgboost Untuk Klasifikasi Tingkat Penyakit Hepatitis C Pada Imbalance Class Data," *J. Gaussian*, vol. 9, no. 3, pp. 227–236, 2020, doi: 10.14710/j.gauss.v9i3.28915.

[10] L. Qadrini, H. Hikmah, and M. Megasari, "Oversampling, Undersampling, Smote SVM dan Random Forest pada Klasifikasi Penerima Bidikmisi Sejava Timur Tahun 2017," *J. Comput. Syst. Informatics*, vol. 3, no. 4, pp. 386–391, 2022, doi: 10.47065/josyc.v3i4.2154.

[11] A. Syukron and A. Subekti, "Penerapan Metode Random Over-Under Sampling dan Random Forest Untuk Klasifikasi Penilaian Kredit," *J. Inform.*, vol. 5, no. 2, pp. 175–185, 2018, doi: 10.31311/ji.v5i2.4158.

[12] M. Zhu et al., "Class weights random forest algorithm for processing class imbalanced medical data," *IEEE Access*, vol. 6, pp. 4641–4652, 2018, doi: 10.1109/ACCESS.2018.2789428.

[13] R. Prakash, "(PDF) Class Weight technique for Handling Class Imbalance," no. July, 2022.

[14] M. Liu, M. Wang, J. Wang, and D. Li, "Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar," *Sensors Actuators, B Chem.*, vol. 177, pp. 970–980, 2013, doi: 10.1016/j.snb.2012.11.071.

[15] C. D. Ahrens, *Essentials of Meteorology AN INVITATION TO THE ATMOSPHERE*, Sixth Edit. Cengage Learning, 2011.

[16] Tukidi, "Karakter Curah Hujan Di Indonesia," *J. Geogr.*, vol. 7, no. 2, pp. 136–145, 2010, [Online]. Available: <http://journal.unnes.ac.id/nju/index.php/JG/article/view/84>

[17] BMKG, "Probabilistik Curah Hujan 24 Jam [Online].," 2020. <https://www.bmkg.go.id/cuaca/probabilistik-curah-hujan.bmkg> (accessed Jan. 07, 2023).

[18] T. A. Pham and V. Q. Tran, "Developing random forest hybridization models for estimating the axial bearing capacity of pile," *PLoS One*, vol. 17, no. 3 March, 2022, doi: 10.1371/journal.pone.0265747.

[19] P. Probst, M. N. Wright, and A. L. Boulesteix, "Hyperparameters and tuning strategies for random forest," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 9, no. 3, pp. 1–19, 2019, doi: 10.1002/widm.1301.

[20] C. H. B. Liu, B. P. Chamberlain, D. A. Little, and A. Cardoso, "Generalising Random Forest Parameter Optimisation to Include Stability and Cost," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10536 LNAI, pp. 102–113, 2017, doi: 10.1007/978-3-319-71273-4_9.

[21] A. Indrawati, L. Ilmu, P. Indonesia, and P. I. Diabetes, "Penerapan Teknik Kombinasi Oversampling Dan Undersampling Hybrid Oversampling and Undersampling Techniques To Handling Imbalanced Dataset," vol. 4, no. 1, pp. 38–43, 2021, doi: 10.33387/jiko.

[22] A. Gupta, N. Tatbul, R. Marcus, S. Zhou, I. Lee, and J. Gottschlich, "Class-Weighted Evaluation Metrics for Imbalanced Data Classification," pp. 1–11, 2020, [Online]. Available: <http://arxiv.org/abs/2010.05995>

[23] S. Zahroh, Y. Hidayat, R. S. Pontoh, A. Santoso, Sukono, and A. T. Bon, "Modeling and forecasting daily temperature in bandung," *Proc. Int. Conf. Ind. Eng. Oper. Manag.*, no. November, pp. 406–412, 2019.

[24] H. Azis, F. Fattah, and P. Putri, "Performa Klasifikasi K-NN dan Cross-validation Pada Data Pasien Pengidap Penyakit Jantung," vol. 12, no. 2, pp. 81–86, 2020.

[25] G. D. Krisnawan, *Rancang Model Prediksi Probabilistik Intensitas Siklon Tropis Berbasis Pembelajaran Jaringan Saraf Tiruan*. Tangerang Selatan: STMKG, 2020.

[26] J. Miao and W. Zhu, "Precision-recall curve (PRC) classification trees," *Evol. Intell.*, vol. 15, no. 3, pp. 1545–1569, 2022, doi: 10.1007/s12065-021-00565-2.