

**A COMPARATIVE ANALYSIS OF HYPERPARAMETER-TUNED XGBOOST AND LIGHTGBM FOR MULTICLASS RAINFALL CLASSIFICATION IN JAKARTA**

**LAPORAN SCIENTIST**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



Disusun oleh :

**COKORDA GDE LANANG PRINGANDANA**

**22.11.4952**

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2025**

**A COMPARATIVE ANALYSIS OF HYPERPARAMETER-TUNED  
XGBOOST AND LIGHTGBM FOR MULTICLASS RAINFALL  
CLASSIFICATION IN JAKARTA**

**LAPORAN SCIENTIST**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



Disusun oleh :

**COKORDA GDE LANANG PRINGANDANA**

**22.11.4952**

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2025**

**HALAMAN PERSETUJUAN**

**JALUR SCIENTIST**

**A COMPARATIVE ANALYSIS OF HYPERPARAMETER-TUNED XGBOOST  
AND LIGHTGBM FOR MULTICLASS RAINFALL CLASSIFICATION IN  
JAKARTA**

yang disusun dan diajukan oleh  
**Cokorda Gde Lanang Pringandana**  
22.11.4952

telah disetujui oleh Dosen Pembimbing  
pada tanggal 24 Oktober 2025

Dosen Pembimbing,



**Kusnawi, S.Kom., M.Eng.**  
**NIK. 190302112**

HALAMAN PENGESAHAN

JALUR SCIENTIST

A COMPARATIVE ANALYSIS OF HYPERPARAMETER-TUNED XGBOOST  
AND LIGHTGBM FOR MULTICLASS RAINFALL CLASSIFICATION IN  
JAKARTA

yang disusun dan diajukan oleh

Cokorda Gde Lanang Pringandana  
22.11.4952

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 24 Oktober 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Haryoko, S.Kom., M.Cs.  
NIK. 190302286

Anna Balta, S.Kom., M.Kom.  
NIK. 190302290

Kusnawi, S.Kom., M.Eng.  
NIK. 190302112



Laporan ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 24 Oktober 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriji, M.Kom.  
NIK. 190302106

## HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Cokorda Gde Lanang Pringandana

NIM : 22.11.4952

Menyatakan bahwa Laporan dengan judul berikut:

**A Comparative Analysis Of Hyperparameter-Tuned Xgboost And Lightgbm For Multiclass Rainfall Classification In Jakarta**

Dosen Pembimbing : Kusnawi, S.Kom., M.Eng.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan kegiatan SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidak-benaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 24 Oktober 2025

Yang Menyatakan,



Cokorda Gde Lanang Pringandana

## HALAMAN PERSEMBAHAN

Puji syukur kehadiran Tuhan Yang Maha Esa. Dengan izin-Nya, penulis dapat menyelesaikan Tugas Akhir ini sebagai wujud tanggung jawab akademik sekaligus langkah akhir dalam menempuh pendidikan sarjana. Kehadiran karya ini menjadi bukti perjalanan panjang penuh doa, usaha, dan pengorbanan. Dengan penuh rasa syukur, penulis mempersembahkan karya ini kepada:

1. Dengan segenap rasa syukur, penulis haturkan puji kehadiran Tuhan Yang Maha Esa yang atas rahmat dan ridha-Nya telah memberikan kekuatan, kesehatan, dan kemudahan, sehingga penulis dapat menyelesaikan penelitian ini.
2. Bapak Ir. Tjokorda Gde Putra Tanaya (Alm) dan Ibu Anak Agung Sri Angga Sukaratni, S.Si, kedua orang tua tercinta, yang menjadi sumber kekuatan penulis. Atas segala limpahan kasih sayang yang tak hingga, doa yang tiada henti, dukungan moral maupun material yang tak pernah putus, serta pengorbanan dalam mengusahakan segala hal demi menunjang kehidupan dan studi penulis.
3. Saudara kandung penulis, Cokorda Istri Ambidika, S.E, yang selalu memberikan dukungan, doa, dan semangat dalam setiap langkah perjalanan penulis.
4. Bapak Kusnawi, S.Kom, M.Eng, selaku Dosen Pembimbing Kolaborasi Riset, atas segala bimbingan, arahan, semangat yang diberikan, serta kesabarannya dalam membimbing penulis hingga terselesaikannya penelitian ini.
5. Bapak Majid Rahardi, S.Kom, M.Eng, selaku Dosen Wali, atas perhatian, arahan, serta dukungan yang selalu diberikan kepada penulis selama menempuh perkuliahan.
6. Teman-teman seperjuangan dari SMA yang berkuliah ke Yogyakarta "Roses Cabang Jogja" Putu Nadia Astiasih dan Dewa Ngakan Gede Mahadewa,

terima kasih atas kebersamaan, dukungan, dan semangat yang selalu diberikan dalam perjalanan hidup dan studi penulis.

7. Rekan di PC KMHDI Yogyakarta, terima kasih atas pengalaman, dukungan, serta kebersamaan dalam perjuangan organisasi yang turut membentuk perjalanan penulis.
8. Teman-teman di 22 IF 07, terima kasih atas kebersamaan, kerja sama, dan semangat yang selalu menguatkan selama masa perkuliahan.
9. Teman-teman tongkrongan di Cheers Lab, terima kasih untuk tawa, canda, dan kebersamaan yang selalu menghadirkan semangat baru di tengah kesibukan kuliah.
10. Cokorda Gde Lanang Pringandana, terima kasih atas segala usaha, perjuangan, dan keteguhan yang telah membawa karya ini hingga terwujud, menjadi bukti nyata perjalanan panjang penulis.

## KATA PENGANTAR

Segala puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga laporan kelulusan ini dapat terselesaikan dengan baik. Keberhasilan ini tidak terlepas dari bimbingan, dukungan, dan bantuan tulus berbagai pihak. Untuk itu, dengan penuh hormat penulis menyampaikan ucapan terima kasih yang mendalam kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Prof. Dr. Kusrini, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Eli Pujastuti, M.Kom., selaku Ketua Program Studi SI Informatika Universitas AMIKOM Yogyakarta.
4. Bapak Kusnawi, S.Kom., M.Eng., selaku Dosen Pembimbing, atas bimbingan, arahan, serta motivasi yang senantiasa diberikan hingga terselesaikannya penelitian ini.
5. Bapak Ir. Tjokorda Gde Putra Tanaya (Alm) dan Ibu Anak Agung Sri Angga Sukaratni, S.Si, selaku orang tua penulis, atas cinta, kasih sayang, doa, dukungan, dan pengorbanan yang tidak ternilai harganya.

Akhir kata, penulis menyadari bahwa laporan ini masih jauh dari kata sempurna dan memiliki banyak keterbatasan. Oleh karena itu, segala bentuk kritik dan saran yang membangun sangat diharapkan demi perbaikan di masa yang akan datang. Semoga karya ini dapat memberikan manfaat bagi pembaca dan pengembangan ilmu pengetahuan.

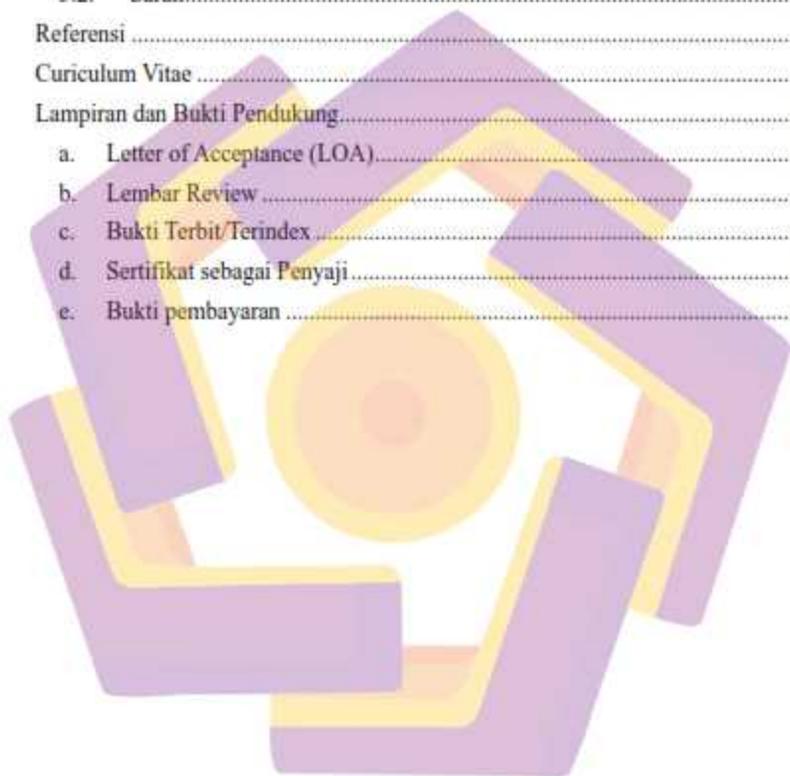
Yogyakarta, 5 September 2025

Penulis

## DAFTAR ISI

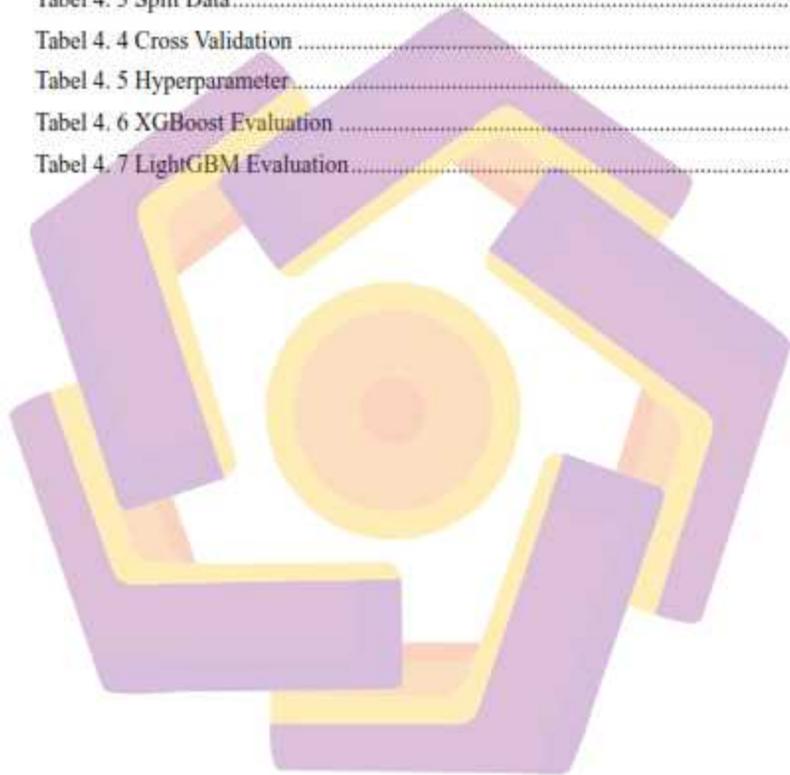
Halaman Judul.....	i
Halaman Persetujuan.....	ii
Halaman Pengesahan .....	iii
Halaman Pernyataan Keaslian Karya.....	iv
Halaman Persembahan .....	v
Kata Pengantar .....	vii
Daftar Isi.....	viii
Daftar Tabel.....	x
Daftar Gambar.....	xi
Daftar Lambang dan Singkatan.....	xii
Daftar Istilah.....	xiii
Intisari .....	xiv
<i>Abstract</i> .....	xv
<b>Bab I</b> Pendahuluan .....	1
1.1.    Gambaran Umum.....	1
1.2.    Rumusan Masalah.....	2
1.3.    Batasan Masalah.....	2
1.4.    Tujuan.....	3
<b>Bab II</b> Tinjauan Pustaka .....	4
2.1.    Studi Literatur .....	4
2.2.    Landasan Teori.....	5
<b>BAB III</b> Metode Penelitian.....	6
3.1.    Metode.....	6
3.1.1.    Data Colaction.....	7
3.1.2.    Data Preparation.....	7
3.1.3.    Data Transformation .....	8
3.1.4.    Modeling .....	9
3.1.5.    Evaluation .....	11
<b>BAB IV</b> Pembahasan .....	14
4.1.    Sub Pembahasan .....	14

4.1.1.	Data Preparation Results.....	14
4.1.2.	Modeling Results .....	16
4.1.3.	Evaluation Results.....	19
BAB V	Kesimpulan .....	24
5.1.	Kesimpulan .....	24
5.2.	Saran.....	25
Referensi	.....	26
Curriculum Vitae	.....	31
Lampiran dan Bukti Pendukung	.....	32
a.	Letter of Acceptance (LOA).....	32
b.	Lembar Review .....	33
c.	Bukti Terbit/Terindex .....	40
d.	Sertifikat sebagai Penyaji.....	41
e.	Bukti pembayaran .....	42



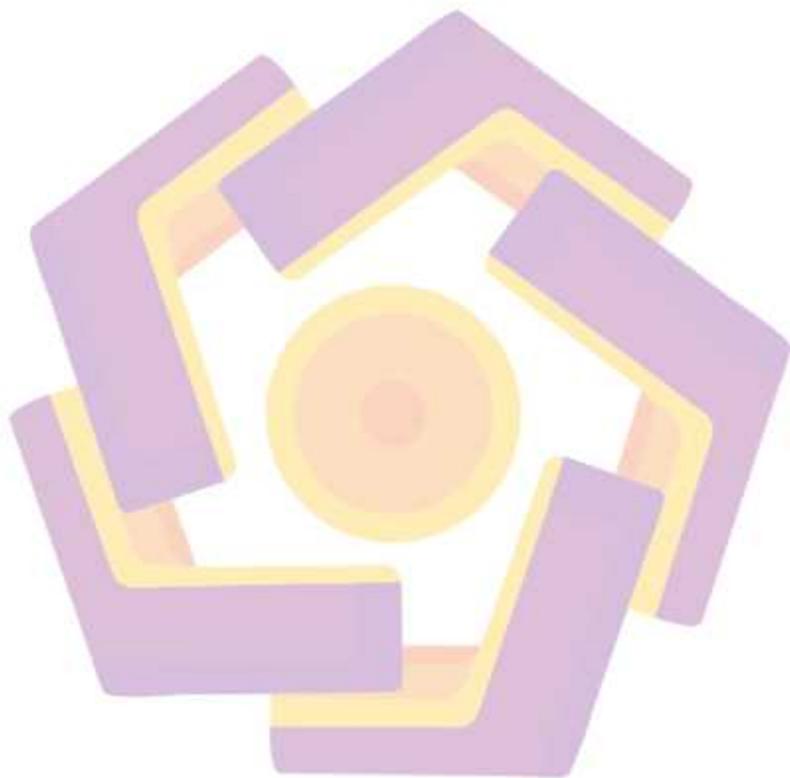
## DAFTAR TABEL

Tabel 3. 1 Dataset .....	7
Tabel 3. 2 Evaluation Matrix .....	12
Tabel 4. 1 Missing Value Imputation .....	14
Tabel 4. 2 mean imputation .....	15
Tabel 4. 3 Split Data .....	17
Tabel 4. 4 Cross Validation .....	17
Tabel 4. 5 Hyperparameter .....	19
Tabel 4. 6 XGBoost Evaluation .....	19
Tabel 4. 7 LightGBM Evaluation .....	20

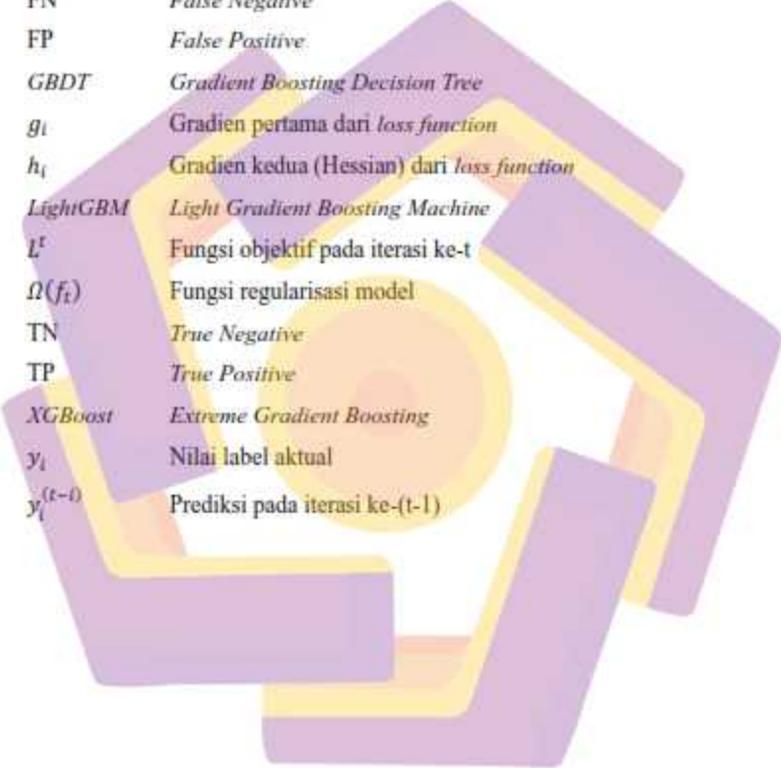


## DAFTAR GAMBAR

Gambar 3. 1 Model Overview.....	6
Gambar 4. 1 Confusion matrix of classification results.....	21
Gambar 4. 2 Comparative performance of XGBoost and LightGBM.....	22



## DAFTAR LAMBANG DAN SINGKATAN



BMKG	Badan Meteorologi, Klimatologi, dan Geofisika
$f_t(x_i)$	Fungsi model pada iterasi ke-t
<i>F1-score</i>	Rata-rata harmonik precision & recall
FN	<i>False Negative</i>
FP	<i>False Positive</i>
<i>GBDT</i>	<i>Gradient Boosting Decision Tree</i>
$g_t$	Gradien pertama dari <i>loss function</i>
$h_t$	Gradien kedua (Hessian) dari <i>loss function</i>
<i>LightGBM</i>	<i>Light Gradient Boosting Machine</i>
$L^t$	Fungsi objektif pada iterasi ke-t
$\Omega(f_t)$	Fungsi regularisasi model
TN	<i>True Negative</i>
TP	<i>True Positive</i>
<i>XGBoost</i>	<i>Extreme Gradient Boosting</i>
$y_i$	Nilai label aktual
$y_i^{(t-1)}$	Prediksi pada iterasi ke-(t-1)

## DAFTAR ISTILAH

Vektor	besaran yang mempunyai arah
<i>Eigen Value</i>	akar-akar persamaan
<i>Overfitting</i>	Model terlalu sesuai data latih.
<i>Underfitting</i>	Model terlalu sederhana.
<i>Cross Validation</i>	Evaluasi model dengan lipatan data.
<i>Feature Engineering</i>	Modifikasi fitur agar lebih informatif.
<i>Hyperparameter Tuning</i>	Penyesuaian parameter model.
<i>Confusion Matrix</i>	Matriks prediksi benar/salah.
Imputasi Data	Mengisi nilai hilang.
<i>Class Imbalance</i>	Distribusi kelas tidak seimbang.
<i>One-Hot Encoding</i>	Ubah kategori ke format biner.
<i>Ensemble Learning</i>	Gabungan beberapa model.
Regularisasi	Pencegah <i>overfitting</i> .
<i>Evaluation Metrics</i>	Ukuran performa model.
<i>Specificity</i>	Kemampuan deteksi negatif benar.
<i>Precision</i>	Proporsi positif yang benar.
<i>Recall</i>	Proporsi positif terdeteksi.
<i>Oversampling</i>	Menambah data kelas minoritas.
<i>RandomizedSearchCV</i>	Pencarian hyperparameter acak.
Dataset	Kumpulan data analisis.

## INTISARI

Peningkatan frekuensi kejadian cuaca ekstrem di Jakarta telah mengganggu kehidupan sehari-hari dan infrastruktur kritis, sehingga menyoroti kebutuhan mendesak akan model prediksi curah hujan yang akurat untuk mendukung mitigasi bencana dan sistem peringatan dini. Penelitian ini bertujuan untuk mengevaluasi dan membandingkan kinerja dua algoritma *machine learning*, yaitu *Extreme Gradient Boosting (XGBoost)* dan *Light Gradient Boosting Machine (LightGBM)*, untuk klasifikasi curah hujan multikelas menggunakan data meteorologi historis. Dataset yang digunakan mencakup fitur-fitur seperti suhu, kelembapan, kecepatan angin, dan curah hujan, yang diproses melalui imputasi rata-rata, *oversampling* untuk mengatasi ketidakseimbangan kelas, *one-hot encoding*, dan rekayasa fitur. Kedua model dilatih dan *tuning* menggunakan *RandomizedSearchCV* serta dievaluasi melalui *cross-validation* dan pengujian independen. Hasil penelitian menunjukkan bahwa *XGBoost* secara konsisten mengungguli *LightGBM*, dengan akurasi 94% dibandingkan 91%. Selain itu, *XGBoost* menunjukkan nilai presisi, *recall*, *F1-score*, dan spesifisitas yang lebih tinggi pada semua kategori curah hujan, sehingga menghasilkan lebih sedikit kesalahan klasifikasi dan prediksi yang lebih stabil. *Confusion matrix* mengonfirmasi kemampuan *XGBoost* yang lebih baik dalam membedakan kondisi cuaca yang serupa, seperti kelas berawan dan hujan. Temuan ini menunjukkan bahwa *XGBoost* lebih efektif dalam menangkap interaksi nonlinier antar fitur cuaca dan oleh karena itu lebih sesuai digunakan pada iklim tropis yang kompleks. Penelitian ini menyimpulkan bahwa *XGBoost* adalah model yang lebih andal dan merekomendasikan integrasinya ke dalam sistem peringatan dini *real-time* untuk meningkatkan ketahanan iklim dan kesiapsiagaan bencana di wilayah perkotaan seperti Jakarta yang semakin terdampak oleh variabilitas iklim.

**Kata kunci:** Klasifikasi, Hyperparameter-Tuning, LightGBM, Machine-Learning, XGBoost

## ABSTRACT

The increasing frequency of extreme weather events in Jakarta has disrupted daily life and critical infrastructure, highlighting the urgent need for accurate rainfall prediction models to support disaster mitigation and early warning systems. This study aims to evaluate and compare the performance of two machine learning algorithms Extreme Gradient Boosting (XGBoost) and Light Gradient Boosting Machine (LightGBM) for multiclass rainfall classification using historical meteorological data. The dataset, which includes features such as temperature, humidity, wind speed, and rainfall, was preprocessed through mean imputation, oversampling to address class imbalance, one-hot encoding, and feature engineering. Both models were trained and tuned using RandomizedSearchCV and assessed through cross-validation and independent testing. The results show that XGBoost consistently outperformed LightGBM, achieving 94% accuracy compared to 91%. Furthermore, XGBoost demonstrated higher precision, recall, F1-score, and specificity across all rainfall categories, resulting in fewer misclassifications and more stable predictions. Confusion matrices confirmed its superior ability to distinguish between similar weather conditions such as cloudy and rainy classes. These findings indicate that XGBoost is more effective in capturing nonlinear interactions between weather features and is therefore better suited for use in complex tropical climates. The study concludes that XGBoost is the more reliable model and recommends its integration into real-time early warning systems to improve climate resilience and disaster preparedness in urban areas like Jakarta that are increasingly affected by climate variability.

**Keyword:** Classification, Hyperparameter-Tuning, LightGBM, Machine-Learning, XGBoost