

**ANALISIS BIAS MODEL HIBRIDA 1DCNN-XGBOOST
DALAM PINJAMAN HIPOTEK PERUMAHAN DENGAN
SHAP**

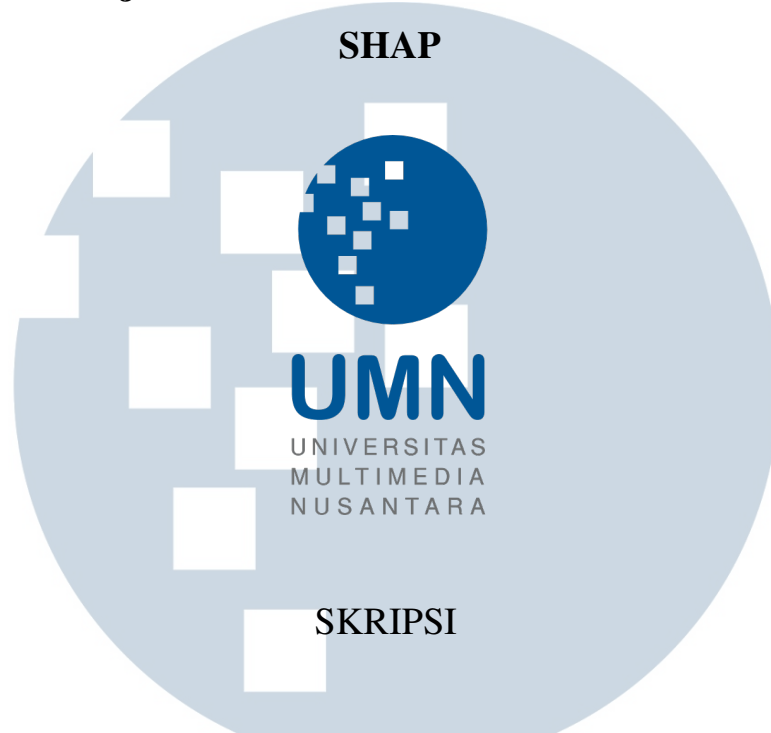


SKRIPSI

**BRYAN RICHIE IRAWAN
00000056044**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**ANALISIS BIAS MODEL HIBRIDA 1DCNN-XGBOOST
DALAM PINJAMAN HIPOTEK PERUMAHAN DENGAN
SHAP**



Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

BRYAN RICHIE IRAWAN
00000056044

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Bryan Richie Irawan

Nomor Induk Mahasiswa : 00000056044

Program Studi : Informatika

Skripsi dengan judul:

Analisis Bias Model Hibrida 1DCNN-XGBoost dalam Pinjaman Hipotek Perumahan dengan SHAP

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 4 Juli 2025



(Bryan Richie Irawan)

UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISIS BIAS MODEL HIBRIDA IDCNN-XGBOOST DALAM PINJAMAN HIPOTEK PERUMAHAN DENGAN SHAP

oleh

Nama : Bryan Richie Irawan
NIM : 00000056044
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

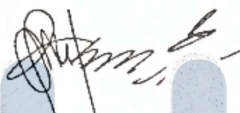
Telah diujikan pada hari Senin, 14 Juli 2025

Pukul 15.00 s/s 17.00 dan dinyatakan

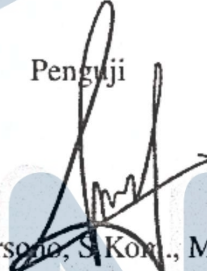
LULUS

Dengan susunan penguji sebagai berikut

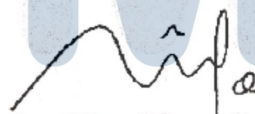
Ketua Sidang


(Adhi Kusnadi, S.T, M.Si.)
NIDN: 303037304

Penguji


(Eka Jaya Harseno, S.Kom., M.Eng.Sc.)
NIDN: 8343771672130333

Pembimbing


(Suwito Pomalingo, S.Kom., M.Kom.)
NIDN: 0911098201

Ketua Program Studi Informatika,


(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)
NIDN: 0315109103

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Bryan Richie Irawan
NIM : 00000056044
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Analisis Bias Model Hibrida 1DCNN-
XGBoost dalam Pinjaman Hipotek
Perumahan dengan SHAP

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 4 Juli 2025

Yang menyatakan

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Bryan Richie Irawan

**Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

HALAMAN PERSEMBAHAN / MOTTO

"A good name is to be more desired than great wealth, Favor is better than silver and gold."

Proverbs 22:1 (NASB)



KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan Skripsi dengan judul **Analisis Bias Model Hibrida 1DCNN-XGBoost dalam Pinjaman Hipotek Perumahan dengan SHAP** yang dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai dengan pada penyusunan Skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan Skripsi ini. Saya berharap hasil dari Skripsi ini dapat bernilai baik, bermanfaat, dan memberi kontribusi positif.

Mengucapkan terima kasih

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Suwito Pomalingo, S.Kom., M.Kom., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga Skripsi ini dapat memberikan kontribusi positif bagi pengembangan teknologi kecerdasan buatan yang adil dan bertanggung jawab, khususnya terhadap keputusan otomatis pada praktik keuangan.

Tangerang, 4 Juli 2025



Bryan Richie Irawan

ANALISIS BIAS MODEL HIBRIDA 1DCNN-XGBOOST DALAM PINJAMAN HIPOTEK PERUMAHAN DENGAN SHAP

Bryan Richie Irawan

ABSTRAK

Penelitian ini membahas penerapan model hibrida 1D Convolutional Neural Network (1DCNN) dan Extreme Gradient Boosting (XGBoost) dalam skenario persetujuan pinjaman hipotek perumahan, serta analisis bias model menggunakan metode Explainable AI, yaitu SHapley Additive Explanations (SHAP). Model hibrida digunakan untuk meningkatkan performa klasifikasi dan diuji menggunakan metrik evaluasi untuk klasifikasi biner, misalnya *accuracy*, *recall*, hingga *h-score*. Hasilnya menunjukkan bahwa model hibrida mampu mendapat performa yang tinggi dengan nilai *recall* misalnya, sebesar 0,9983 dan *accuracy* 0,9648, namun tidak secara signifikan lebih baik dibandingkan kedua model tunggal. Selain itu, analisis SHAP mengungkapkan bahwa meskipun model lebih banyak bergantung pada fitur keuangan, masih terdapat indikasi penggunaan fitur sensitif ras dan jenis kelamin dalam pengambilan keputusan, baik secara lokal maupun global. Hal ini menunjukkan bahwa model hibrida belum sepenuhnya bebas dari bias fitur sensitif. Penelitian ini menegaskan pentingnya penerapan XAI untuk mengevaluasi keadilan dalam sistem pengambilan keputusan otomatis di sektor keuangan.

Kata kunci: bias, explainable AI, pinjaman hipotek, SHAP, 1DCNN-XGBoost



ANALYSIS OF BIAS IN THE 1DCNN-XGBOOST HYBRID MODEL FOR MORTGAGE LOANS USING SHAP

Bryan Richie Irawan

ABSTRACT

This research examines the application of a hybrid 1D Convolutional Neural Network (1DCNN) and Extreme Gradient Boosting (XGBoost) model for mortgage loan approval scenarios, coupled with bias analysis using Explainable AI (XAI), specifically SHapley Additive Explanations (SHAP). The hybrid model was designed to improve classification performance and was evaluated using binary classification metrics such as accuracy, recall, h-score, etc. Results indicate that the hybrid model achieved high performance, with a recall of 0.9983 and an accuracy of 0.9648, though it did not significantly outperform the individual standalone models. Additionally, SHAP analysis revealed that while the model primarily relied on financial features, there was evidence of sensitive features such as race and gender influencing decision making at both local and global levels, suggesting that the hybrid model is not entirely free from bias related to protected attributes. This study emphasizes the critical role of XAI in evaluating fairness within automated decision-making systems in the financial sector.

Keywords: *bias, explainable AI, mortgage loans, SHAP, 1DCNN-XGBoost*



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR KODE	xiii
DAFTAR RUMUS	xv
DAFTAR LAMPIRAN	xvi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Permasalahan	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB 2 LANDASAN TEORI	5
2.1 Home Mortgage Disclosure Act	5
2.2 Min-Max Scaling	6
2.3 Standard Scaling	6
2.4 Recursive Feature Elimination	7
2.5 Information Value	7
2.6 Extreme Gradient Boosting	8
2.7 1D Convolutional Neural Network	9
2.8 Model Hibrida 1DCNN-XGBoost	12
2.9 Metrik Evaluasi Model	13
2.9.1 Precision	13
2.9.2 Recall	14
2.9.3 F1-score	14
2.9.4 Area Under the Curve	15
2.9.5 Macro Average	15
2.9.6 Weighted Average	15
2.9.7 H-score	16
2.10 SHapley Additive exPlanations	17
BAB 3 METODOLOGI PENELITIAN	19
3.1 Alur Penelitian	19
3.1.1 Studi Literatur	20
3.1.2 Pengumpulan Data	21
3.1.3 Praproses Statis	21
3.1.4 Praproses Dinamis	23
3.1.5 Implementasi Model	26
3.1.6 Pengujian dan Evaluasi Model	30
3.1.7 Analisis Kontribusi Fitur	30

BAB 4	HASIL DAN DISKUSI	31
4.1	Spesifikasi Sistem	31
4.2	Implementasi Kode	31
4.2.1	Impor Library	32
4.2.2	Pemuatan Data	33
4.2.3	Eliminasi Fitur	34
4.2.4	Perbaikan Tipe Data Tiap Fitur	35
4.2.5	Memperbaiki Nilai Incomprehensible Menjadi NaN	37
4.2.6	Menghapus Observasi dengan Fitur Krusial yang Tidak Lengkap	40
4.2.7	Reduksi Dimensi Fitur census_tract	41
4.2.8	Penyaringan Label	42
4.2.9	Pendefinisian Metode Encoding dan Normalisasi	43
4.2.10	Fungsi Inisialisasi Model 1DCNN	47
4.2.11	Penerapan K-Fold Validation	49
4.2.12	Penyiapan Data dengan Fitur Terpilih dalam Tiap Fold	50
4.2.13	Pencarian Parameter XGBoost Terbaik dalam Tiap Fold	63
4.2.14	Perekaman Parameter dan Fitur dalam Tiap Fold	68
4.2.15	Seleksi Parameter dan Fitur Terbaik dari Kelima Fold	69
4.2.16	Penyiapan Data dengan Fitur Terbaik	71
4.2.17	Pelatihan dan Pengujian Model Akhir	71
4.2.18	Evaluasi Model	75
4.2.19	Analisis Kontribusi Fitur dengan SHAP	76
4.3	Analisis Performa Model Hibrida 1DCNN-XGBoost	79
4.3.1	Perbandingan Model Hibrida dalam Pinjaman Hipotek dengan Pinjaman Peer-to-peer	80
4.3.2	Perbandingan Model Hibrida dengan Model Tunggal dalam Pinjaman Hipotek	81
4.4	Analisis Bias Ras dan Jenis Kelamin dalam Model Hibrida 1DCNN-XGBoost	82
BAB 5	SIMPULAN DAN SARAN	84
5.1	Simpulan	84
5.2	Saran	84
DAFTAR PUSTAKA		86

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR TABEL

Tabel 2.1	Nilai ambang batas untuk skor IV.	8
Tabel 4.1	Perbandingan performa model hibrida dalam dua skenario berbeda	80
Tabel 4.2	Perbandingan performa ketiga model	81
Tabel 5.1	Perbandingan performa dengan rasio pemisahan data 80:20 dan 70:30	105
Tabel 5.2	Parameter XGBoost terbaik	107
Tabel 5.3	Data individu yang dipakai dalam <i>waterfall plot</i>	108



DAFTAR GAMBAR

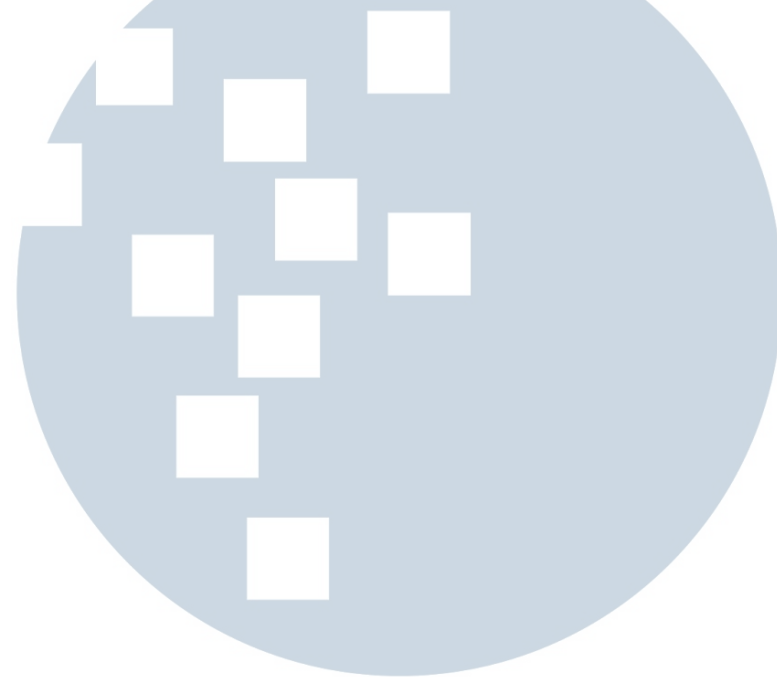
Gambar 3.1	Alur penelitian	19
Gambar 3.2	Arsitektur model tunggal 1DCNN	27
Gambar 3.3	Arsitektur model hibrida 1DCNN-XGBoost	29
Gambar 4.1	<i>Waterfall plot</i> untuk sampel data pertama	82
Gambar 4.2	<i>Beeswarm plot</i> untuk penjelasan global	83
Gambar 5.1	<i>Fitur terpilih melalui stability selection</i>	106



DAFTAR KODE

Kode 4.1	Python <i>script</i> untuk impor <i>library</i>	32
Kode 4.2	Python <i>script</i> untuk memuat data	33
Kode 4.3	Python <i>script</i> eliminasi fitur	34
Kode 4.4	Python <i>script</i> untuk perbaikan tipe data	35
Kode 4.5	Python <i>script</i> untuk mengganti nilai <i>incomprehensible</i> menjadi <i>NaN</i>	38
Kode 4.6	Python <i>script</i> untuk eliminasi observasi dengan fitur krusial yang tidak lengkap	40
Kode 4.7	Python <i>script</i> reduksi kategori <i>census</i> menjadi kategori <i>county</i> . .	41
Kode 4.8	Python <i>script</i> untuk penyaringan label	42
Kode 4.9	Python <i>script</i> untuk pengelompokan fitur numerik	43
Kode 4.10	Python <i>script</i> untuk pengelompokan fitur	45
Kode 4.11	Python <i>script</i> untuk inisialisasi model 1DCNN dalam model hibrida	47
Kode 4.12	Python <i>script</i> untuk inisialisasi model tunggal 1DCNN	48
Kode 4.13	Python <i>script</i> penerapan <i>k-fold validation</i>	49
Kode 4.14	Python <i>script</i> penyiapan data dalam <i>k-fold validation</i>	51
Kode 4.15	Python <i>script</i> untuk keseluruhan pembersihan data dalam model tunggal 1DCNN dan model hibrida 1DCNN-XGBoost	52
Kode 4.16	Python <i>script</i> untuk keseluruhan pembersihan data dalam model tunggal XGBoost	54
Kode 4.17	Python <i>script</i> pembuatan <i>pipeline</i> pengisian data yang hilang . .	56
Kode 4.18	Python <i>script</i> penerapan <i>pipeline</i> penanganan data yang hilang .	56
Kode 4.19	Python <i>script</i> pembuatan <i>pipeline encoding</i> dan normalisasi . . .	58
Kode 4.20	Python <i>script</i> penerapan <i>pipeline encoding</i> dan normalisasi . . .	60
Kode 4.21	Python <i>script</i> seleksi fitur	61
Kode 4.22	Python <i>script</i> untuk perhitungan nilai IV tiap fitur	62
Kode 4.23	Python <i>script</i> deklarasi parameter untuk optimasi hiperparameter XGBoost	64
Kode 4.24	Python <i>script</i> pencarian parameter dan fitur terbaik untuk model tunggal XGBoost	64
Kode 4.25	Python <i>script</i> pencarian parameter dan fitur terbaik untuk model hibrida 1DCNN-XGBoost	65
Kode 4.26	Python <i>script</i> penyimpanan fitur terbaik untuk model tunggal XGBoost dan model hibrida 1DCNN-XGBoost	68
Kode 4.27	Python <i>script</i> pencarian parameter dan fitur terbaik untuk model tunggal XGBoost	69
Kode 4.28	Python <i>script</i> fungsi pengambilan fitur dengan <i>stability selection</i> .	69
Kode 4.29	Python <i>script</i> penerapan <i>stability selection</i> dan penentuan parameter XGBoost	70
Kode 4.30	Python <i>script</i> pembagian data menjadi data latih dan data uji untuk model akhir	71
Kode 4.31	Python <i>script</i> pelatihan model akhir XGBoost	72
Kode 4.32	Python <i>script</i> pelatihan model akhir XGBoost	73
Kode 4.33	Python <i>script</i> pelatihan model akhir 1DCNN-XGBoost	74
Kode 4.34	Python <i>script</i> evaluasi model	75

Kode 4.35	Python <i>script</i> fungsi <i>custom</i> model hibrida untuk dianalisis dengan SHAP	76
Kode 4.36	Python <i>script</i> pelacakan kontribusi fitur dengan SHAP	77
Kode 4.37	Python <i>script</i> untuk menampilkan <i>waterfall plot</i>	78
Kode 4.38	Python <i>script</i> untuk menampilkan <i>beeswarm plot</i>	79



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR RUMUS

Rumus 2.1	<i>Min-max scaling</i>	6
Rumus 2.2	<i>Standard scaling</i>	7
Rumus 2.3	<i>Information value</i>	8
Rumus 2.4	Prediksi agregat XGBoost	9
Rumus 2.5	Fungsi objektif XGBoost	9
Rumus 2.6	Lapisan <i>convolutional 1D</i>	10
Rumus 2.7	Lapisan <i>pooling</i>	11
Rumus 2.8	Lapisan <i>dense</i>	11
Rumus 2.9	Fungsi aktivasi <i>sigmoid</i> lapisan <i>output</i>	12
Rumus 2.10	Fungsi <i>binary cross-entropy</i> lapisan <i>output</i>	12
Rumus 2.11	<i>Precision</i>	14
Rumus 2.12	<i>Recall</i>	14
Rumus 2.13	<i>F1-score</i>	15
Rumus 2.14	<i>Area under the curve</i>	15
Rumus 2.15	<i>Macro average</i>	15
Rumus 2.16	<i>Weighted average</i>	16
Rumus 2.17	<i>H-score</i>	17
Rumus 2.18	Nilai <i>shapley</i>	18



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	90
Lampiran 2	Formulir Bimbingan	102
Lampiran 3	Perbandingan Performa Model 1DCNN-XGBoost dengan Rasio Data 80:20 dan 70:30	105
Lampiran 4	Daftar fitur akhir yang Terpilih melalui Stability Selection . . .	106
Lampiran 5	Parameter XGBoost Terbaik	107
Lampiran 6	Sampel Data Individu yang Dipakai untuk Analisis Kontribusi Fitur Waterfall Plot	108

