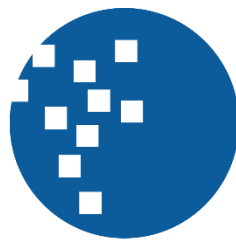


**PERBANDINGAN *GRID SEARCH* DAN *BAYESIAN*
OPTIMIZATION PADA KLASIFIKASI *COLLECTION*
INTENSITY SCORING DAN PREDIKSI *CHANNEL*
RECOMMENDATION (STUDI KASUS: *P2P LENDING*)**



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Kelly Mae

00000051428

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2024

PERBANDINGAN *GRID SEARCH* DAN *BAYESIAN OPTIMIZATION* PADA KLASIFIKASI *COLLECTION INTENSITY SCORING* DAN *PREDIKSI CHANNEL RECOMMENDATION* (STUDI KASUS: *P2P LENDING*)



Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Komputer (S.Kom.)

Kelly Mae
0000051428

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Kelly Mae
Nomor Induk Mahasiswa : 00000051428
Program studi : Sistem Informasi

Skripsi dengan judul:

“Perbandingan *Grid Search* dan *Bayesian Optimization* pada Klasifikasi *Collection Intensity Scoring* dan Prediksi *Channel Recommendation* (Studi Kasus: *P2P Lending*)”

merupakan hasil karya saya sendiri bukan plagiat dari karya 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 skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas Akhir yang telah saya tempuh.

Tangerang, 24 Mei 2024



Kelly Mae

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

HALAMAN PENGESAHAN

Skripsi dengan judul


“Perbandingan *Grid Search* dan *Bayesian Optimization*
pada Klasifikasi *Collection Intensity Scoring*
dan Prediksi *Channel Recommendation* (Studi Kasus: *P2P Lending*)”

Oleh

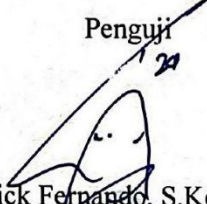
Nama : Kelly Mae
NIM : 00000051428
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Jumat, 24 Mei 2024
Pukul 13.00 s.d 15.00 dan dinyatakan
LULUS
Dengan susunan penguji sebagai berikut.

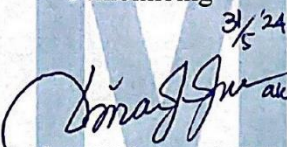
Ketua Sidang


31/05/24
Samuel Azy Sanjaya, S.T., M.T.
0305049402/075049

Penguji


24
Dr. Erick Fernando, S.Kom., M.S.I.
1029118501

Pembimbing


31/5/24
Dinar Ajeng Kristiyanti, S.Kom.,
M.Kom.
0330128801/079159

Ketua Program Studi Sistem Informasi


31/5/24
Ririn Ikana Desanti, S.Kom., M.Kom.

LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHAISWA

Yang bertanda tangan dibawah ini:

Nama : Kelly Mae
Nomor Induk Mahasiswa : 00000051428
Program Studi : Sistem Informasi
Jenjang : S2 / S1 / D3
Judul Karya Ilmiah :

**PERBANDINGAN *GRID SEARCH* DAN *BAYESIAN OPTIMIZATION*
PADA KLASIFIKASI *COLLECTION INTENSITY SCORING* DAN
PREDIKSI *CHANNEL RECOMMENDATION* (STUDI KASUS: *P2P*
LENDING)**

Menyatakan dengan sesungguhnya bahwa saya bersedia:

Memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.

Saya tidak bersedia, dikarenakan:

Dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)*.

Tangerang, 24 Mei 2024



Kelly Mae

* Jika tidak bisa membuktikan LoA jurnal/HKI selama 6 bulan kedepan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

KATA PENGANTAR

Puji syukur kepada Tuhan yang Maha Esa karena atas rahmat dan karunia-Nya, penulisan skripsi yang berjudul “Perbandingan *Grid Search* dan *Bayesian Optimization* pada Klasifikasi *Collection Intensity Scoring* dan Prediksi *Channel Recommendation* (Studi Kasus: *P2P Lending*)” dapat terselesaikan tepat pada waktunya. Pembuatan skripsi ini diajukan sebagai salah satu prasyarat kelulusan untuk menyelesaikan Program Strata 1 dan memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Teknik dan Informatika di Universitas Multimedia Nusantara.

Dengan selesainya proses penulisan laporan skripsi ini, penulis ingin mengucapkan terima kasih kepada Universitas Multimedia Nusantara atas kesempatan yang diberikan kepada penulis untuk menimba ilmu. Di samping itu, penulis juga ingin menyampaikan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, M.A., 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. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Dinar Ajeng Kristiyanti, S.Kom., M.Kom., sebagai Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya laporan skripsi ini.
5. Keluarga penulis yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas laporan skripsi ini.
6. Geraldy Kumara, sebagai *support system* penulis yang selalu mendampingi sekaligus memberikan motivasi atas terselesainya laporan skripsi ini.
7. Teman-teman yang telah memberikan dukungan semangat atas terselesainya laporan skripsi ini.

Penulis juga mengucapkan terima kasih kepada seluruh sahabat, teman, rekan kerja, dan kepada seluruh pihak yang tidak dapat disebutkan satu-persatu yang telah mendukung kelangsungan pengerjaan hingga terealisasinya laporan skripsi ini.

Dengan demikian, penulis sudah berusaha sebaik mungkin dalam penyusunan skripsi ini, namun penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna. Oleh karena itu, penulis sangat terbuka terhadap kritik maupun saran dari berbagai pihak. Dengan adanya laporan skripsi ini, diharapkan dapat memberikan informasi yang bermanfaat dan menginspirasi para pembaca, serta memberikan kontribusi pada dunia riset dan pendidikan Indonesia.

Tangerang, 13 Mei 2024



Kelly Mae

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

PERBANDINGAN *GRID SEARCH* DAN *BAYESIAN OPTIMIZATION* PADA KLASIFIKASI *COLLECTION INTENSITY SCORING* DAN PREDIKSI *CHANNEL RECOMMENDATION* (STUDI KASUS: *P2P LENDING*)

Kelly Mae

ABSTRAK

Perkembangan teknologi yang pesat telah mendorong tumbuhnya *financial technology (fintech)*, mengubah metode keuangan tradisional menjadi digital. Salah satu solusi *fintech* yang seringkali digunakan adalah *Peer-to-Peer (P2P) Lending*, yang menyediakan pinjaman *online* namun menghadapi tantangan dalam pembayaran kembali akibat kondisi keuangan nasabah yang tidak stabil. Untuk mengatasi masalah ini, *P2P lending* berupaya meningkatkan produktivitas penagihan dan komunikasi dengan nasabah melalui penerapan model klasifikasi *Collection Intensity Scoring (CIS)* dan model prediksi *channel recommendation* untuk mengoptimalkan proses penagihan. Penelitian ini bertujuan membandingkan kinerja *Grid Search* dan *Bayesian Optimization* pada model klasifikasi *Random Forest* dan model prediksi *K-Nearest Neighbors Regressor* dengan mengikuti *CRISP-DM Framework*. Penelitian ini melibatkan pengumpulan dan pemrosesan data pribadi, pinjaman, dan interaksi nasabah dari perusahaan *P2P Lending* menggunakan metode *Extract, Transformation, dan Load (ETL)*. *Data pre-processing* dan *Exploratory Data Analysis* dilakukan untuk mempersiapkan data untuk pemodelan dengan *hyperparameter optimization*. Kinerja model dievaluasi berdasarkan *accuracy* dan *Mean Absolute Error (MAE)*. Hasil penelitian menunjukkan *Bayesian Optimization* mengungguli *Grid Search*, meningkatkan *accuracy* pada model *Random Forest* menjadi 98.34% dan menurunkan *MAE* pada model *K-Nearest Neighbors Regressor* menjadi 0.245308. Model dengan metode *Bayesian Optimization* mengklasifikasikan data intensitas penagihan dengan benar 98 dari 100 kali percobaan, dengan rata-rata kesalahan prediksi sekitar 0.245 unit dari nilai sebenarnya.

Kata kunci: *Bayesian Optimization, Channel Recommendation, Collection Intensity Scoring, Grid Search, K-Nearest Neighbors, Random Forest*

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

**COMPARISONS OF GRID SEARCH AND BAYESIAN
OPTIMIZATION ON COLLECTION INTENSITY SCORING
CLASSIFICATION AND CHANNEL RECOMMENDATION
PREDICTION (CASE STUDY: P2P LENDING)**

Kelly Mae

ABSTRACT (English)

The rapid advancement of technology has spurred the expansion of financial technology (fintech), which has revolutionized conventional financial practices into digital formats. Peer-to-Peer (P2P) Lending stands out as a popular fintech solution, offering online loans. However, it encounters repayment hurdles due to the inconsistent financial circumstances of borrowers. To address this challenge, P2P lending is employing Collection Intensity Scoring (CIS) classification model and channel recommendation prediction model to enhance collection productivity and customer communication. These models aim to optimize the collection process. This research evaluates the performance of Grid Search and Bayesian Optimization techniques on Random Forest classification and K-Nearest Neighbors Regressor prediction models within the CRISP-DM Framework. Through the Extract, Transformation, and Load (ETL) method, personal, loan, and interaction data from a P2P Lending firm are gathered and processed. Data pre-processing and Exploratory Data Analysis precede modeling with hyperparameter optimization. Bayesian Optimization is found to be superior, enhancing the Random Forest model's accuracy to 98.34% and reducing the K-Nearest Neighbors Regressor model's Mean Absolute Error (MAE) to 0.245308. With Bayesian Optimization, the CIS classification model accurately predicts collection intensity correctly 98 out of 100 times, with an average prediction error of approximately 0.245 units from the actual value.

Keywords: *Bayesian Optimization, Channel Recommendation, Collection Intensity Scoring, Grid Search, K-Nearest Neighbors, Random Forest*

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 ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PENGESAHAN	iii
LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHAISWA	iv
KATA PENGANTAR	v
ABSTRAK	vii
<i>ABSTRACT (English)</i>	viii
DAFTAR ISI	ix
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR RUMUS	xvii
DAFTAR LAMPIRAN	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	8
1.3 Batasan Masalah	9
1.4 Tujuan dan Manfaat Penelitian	10
1.4.1 Tujuan Penelitian	10
1.4.2 Manfaat Penelitian	10
1.5 Sistematika Penulisan	11
BAB II LANDASAN TEORI	12
2.1 Penelitian Terdahulu	12
2.2 Tinjauan Teori	16
2.2.1 <i>Financial Technology (Fintech)</i>	16
2.2.1.1 <i>Peer-to-Peer Lending (P2P Lending)</i>	16
2.2.1.2 <i>Credit Scoring</i>	17
2.2.2 <i>Recommendation System</i>	18
2.2.3 <i>Extract, Transformation, Load (ETL)</i>	18
2.2.4 <i>Grid Search</i>	19
2.2.5 <i>Bayesian Optimization</i>	19
2.2.6 <i>Evaluation Metrics</i>	22
2.2.6.1 <i>Accuracy</i>	22
2.2.6.2 <i>Precision</i>	23

2.2.6.3	<i>Recall</i>	23
2.2.6.4	<i>F1-Score</i>	23
2.2.6.5	<i>Mean Absolute Error</i>	24
2.2.6.6	<i>Mean Square Error</i>	24
2.2.6.7	<i>Root Mean Square Error</i>	25
2.2.7	<i>Statistical Tests</i>	25
2.2.7.1	<i>Shapiro-Wilk Test</i>	25
2.2.7.2	<i>T-Test</i>	26
2.3	<i>Framework dan Algoritma</i>	28
2.3.1	<i>CRISP-DM Framework</i>	28
2.3.2	<i>Random Forest</i>	30
2.3.3	<i>K-Nearest Neighbors</i>	31
2.4	<i>Tools</i>	32
2.4.1	<i>PostgreSQL</i>	32
2.4.2	<i>DBeaver</i>	33
2.4.3	<i>Python</i>	33
2.4.4	<i>Jupyter Notebook</i>	33
2.4.5	<i>R Programming Language</i>	34
2.4.6	<i>RStudio</i>	34
BAB III METODOLOGI PENELITIAN		35
3.1	Gambaran Umum Objek Penelitian	35
3.2	Metode Penelitian	35
3.2.1	Alur Penelitian	36
3.2.1.1	<i>Business Understanding</i>	38
3.2.1.2	<i>Data Understanding</i>	38
3.2.1.3	<i>Data Preparation</i>	39
3.2.1.4	<i>Modeling</i>	41
3.2.1.5	<i>Evaluation</i>	42
3.2.2	Metode Data Mining	44
3.3	Teknik Pengumpulan Data	45
3.3.1	Populasi dan Sampel	45
3.3.2	Periode Pengambilan Data	46
3.4	Variabel Penelitian	46

3.4.1	Variabel Dependen	46
3.4.2	Variabel Independen	47
3.5	Teknik Analisis Data	49
BAB IV	ANALISIS DAN HASIL PENELITIAN	52
4.1	<i>Business Understanding</i>	52
4.2	<i>Data Understanding</i>	53
4.3	<i>Data Preparation</i>	56
4.3.1	<i>Extract, Transformation, Load (ETL)</i>	56
4.3.1.1	<i>Extract</i>	56
4.3.1.2	<i>Transformation</i>	57
4.3.1.3	<i>Load</i>	71
4.3.2	<i>Data Pre-Processing and Exploratory Data Analysis (EDA)</i>	73
4.3.2.1	<i>Data Pre-Processing for General</i>	73
4.3.2.2	<i>Data Pre-Processing for Classification</i>	79
4.3.2.3	<i>Data Pre-Processing for Prediction</i>	83
4.3.2.4	<i>Exploratory Data Analysis (EDA)</i>	95
4.3.2.5	<i>Data Splitting for Classification</i>	106
4.3.2.6	<i>Data Splitting for Prediction</i>	109
4.4	<i>Modeling</i>	110
4.4.1	<i>Classification</i>	110
4.4.1.1	<i>Random Forest</i>	111
4.4.1.2	<i>Hyperparameter Optimization (Grid Search & Bayesian Optimization)</i>	113
4.4.2	<i>Prediction</i>	118
4.4.2.1	<i>K-Nearest Neighbors Regressor</i>	119
4.4.2.2	<i>Hyperparameter Optimization (Grid Search & Bayesian Optimization)</i>	121
4.4.3	<i>Model Testing</i>	125
4.4.3.1	<i>Classification</i>	125
4.4.3.2	<i>Prediction</i>	128
4.5	<i>Evaluation</i>	131
4.5.1	<i>Random Forest</i>	132
4.5.2	<i>K-Nearest Neighbors Regressor</i>	138

4.5.3	<i>Statistical Tests</i>	142
4.5.3.1	<i>Shapiro-Wilk Test</i>	142
4.5.3.2	<i>T-Test</i>	143
4.6	<i>Result and Discussion</i>	145
BAB V SIMPULAN DAN SARAN		150
5.1	Simpulan	150
5.2	Saran	152
DAFTAR PUSTAKA		154
LAMPIRAN		164

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	12
Tabel 3.1 Rincian Data Splitting	41
Tabel 3.2 Perbandingan Framework KDD, SEMMA, dan CRISP-DM	44
Tabel 3.3 Struktur Data pada Variabel Dependen	47
Tabel 3.4 Struktur Data pada Variabel Independen untuk Klasifikasi Collection Intensity Scoring	47
Tabel 3.5 Struktur Data pada Variabel Independen untuk Prediksi Channel Recommendation.....	48
Tabel 3.6 Komparasi Algoritma pada Penelitian Terdahulu terkait Credit Scoring	50
Tabel 3.7 Komparasi Algoritma pada Penelitian Terdahulu terkait Recommendation System.....	51
Tabel 4.1 Tabel customer_personal_information	53
Tabel 4.2 Tabel loan_information.....	54
Tabel 4.3 Tabel email_interaction.....	54
Tabel 4.4 Tabel robocall_interaction	54
Tabel 4.5 Tabel sms_interaction	55
Tabel 4.6 Tabel telephony_interaction.....	55
Tabel 4.7 Robocall Intention.....	65
Tabel 4.8 Telephony Call Type.....	66
Tabel 4.9 Telephony Action Result	67
Tabel 4.10 Hasil Load Data	72
Tabel 4.11 Perbandingan Kinerja Model Random Forest.....	136
Tabel 4.12 Hasil Evaluasi Performa Accuracy pada Model Random Forest.....	138
Tabel 4.13 Perbandingan Kinerja Model K-Nearest Neighbors Regressor	139
Tabel 4.14 Hasil Evaluasi Performa Mean Absolute Error (MAE) pada Model K-Nearest Neighbors Regressor.....	141
Tabel 4.15 Hasil Shapiro-Wilk Test berdasarkan Metrik Accuracy dari Model Random Forest dan Mean Absolute Error dari Model K-Nearest Neighbors Regressor.....	143
Tabel 4.16 Hasil T-Test berdasarkan Metrik Accuracy dari Model Random Forest x Bayesian Optimization dan Mean Absolute Error dari Model K-Nearest Neighbors Regressor x Bayesian Optimization dengan Model Lainnya	144
Tabel 4.17 Perbandingan Performa Model Random Forest dalam Penelitian Terdahulu dengan Penelitian Collection Intensity Scoring (CIS) dan Channel Recommendation.....	146
Tabel 4.18 Perbandingan Performa pada Model K-Nearest Neighbors dalam Penelitian Terdahulu dengan Penelitian Collection Intensity Scoring (CIS) dan Channel Recommendation	147

DAFTAR GAMBAR

Gambar 2.1 Confusion Matrix	22
Gambar 2.2 Kerangka Kerja CRISP-DM.....	28
Gambar 3.1 Alur Penelitian.....	37
Gambar 3.2 Data Preparation.....	39
Gambar 3.3 Modeling	41
Gambar 3.4 Evaluation.....	42
Gambar 4.1 Konfigurasi Import Data di DBeaver	56
Gambar 4.2 Hasil Extract.....	57
Gambar 4.3 Query customer_loan_information	58
Gambar 4.4 Query telephony_interaction	58
Gambar 4.5 Query mst_customer_information.....	59
Gambar 4.6 Hasil user_id dalam Tabel mst_customer_information	60
Gambar 4.7 Query mst_loan_information	60
Gambar 4.8 Query mst_email_interaction	61
Gambar 4.9 Query mst_robocall_interaction.....	61
Gambar 4.10 Query mst_sms_interaction.....	62
Gambar 4.11 Query mst_telephony_interaction	63
Gambar 4.12 Query mst_email_interaction_filtered	64
Gambar 4.13 Query mst_robocall_interaction_filtered	64
Gambar 4.14 Query mst_sms_interaction_filtered	65
Gambar 4.15 Query mst_telephony_interaction_filtered.....	66
Gambar 4.16 Query interaction_history.....	68
Gambar 4.17 Query nearest_interaction_to_payment.....	69
Gambar 4.18 Query persona_cis_with_comm_channels.....	70
Gambar 4.19 Hasil Transformation.....	71
Gambar 4.20 Import Data	73
Gambar 4.21 View Data Info	74
Gambar 4.22 Display Data Head	74
Gambar 4.23 Check Null Data	75
Gambar 4.24 Replace Null Data	76
Gambar 4.25 Change Data Types	76
Gambar 4.26 View Updated Data Info	77
Gambar 4.27 Hasil Data Cleaning	78
Gambar 4.28 Data Standardization	78
Gambar 4.29 Hasil Data Standardization.....	78
Gambar 4.30 Scoring Variables Initialization.....	79
Gambar 4.31 Intensity Scoring Calculation.....	80
Gambar 4.32 Apply Intensity Score Calculation	81
Gambar 4.33 Categorizing Intensity Scoring.....	81
Gambar 4.34 Apply Categorizing Intensity Score	82
Gambar 4.35 Hasil Feature Engineering untuk Pendekatan Classification	82
Gambar 4.36 Data Encoding untuk Pendekatan Classification	83

Gambar 4.37 Hasil Data Encoding untuk Pendekatan Classification	83
Gambar 4.38 Hasil Data Cleaning untuk raw_data.....	84
Gambar 4.39 Data Cleaning untuk interaction_data	85
Gambar 4.40 Calculate Sparsity.....	85
Gambar 4.41 Compressed Sparse Row Matrix Computation	86
Gambar 4.42 Hasil Data Transformation pada csr.....	86
Gambar 4.43 Change Dataframe to Long Format.....	87
Gambar 4.44 Hasil Data Transformation pada melted.....	88
Gambar 4.45 K-Nearest Neighbors Model Initialization and Fitting	88
Gambar 4.46 Channel Recommendation Based on Similarity.....	89
Gambar 4.47 Apply Channel Recommendation for All Users	90
Gambar 4.48 Hasil Feature Engineering untuk Pendekatan Prediction	91
Gambar 4.49 Data Integration.....	92
Gambar 4.50 Hasil Data Integration	92
Gambar 4.51 Data Encoding untuk Pendekatan Prediction	93
Gambar 4.52 Hasil Data Encoding untuk Pendekatan Prediction.....	94
Gambar 4.53 Data Standardization untuk Pendekatan Prediction	94
Gambar 4.54 Hasil Data Standardization untuk Pendekatan Prediction.....	95
Gambar 4.55 Copy and Merge Data for Exploratory Data Analysis	95
Gambar 4.56 Descriptive Statistics	96
Gambar 4.57 Gender Distribution.....	97
Gambar 4.58 Age Distribution.....	98
Gambar 4.59 Email Distribution	98
Gambar 4.60 Robocall Distribution	99
Gambar 4.61 SMS Distribution	100
Gambar 4.62 Telephony Distribution	101
Gambar 4.63 Total Loan Distribution.....	101
Gambar 4.64 Paid Loans Distribution.....	102
Gambar 4.65 Average Interacted to Paid Distribution.....	103
Gambar 4.66 Intensity Category Distribution.....	104
Gambar 4.67 Recommended Channel Distribution	104
Gambar 4.68 Correlation Plot	105
Gambar 4.69 Data Splitting untuk Pendekatan Classification	107
Gambar 4.70 Hasil Data Splitting untuk Pendekatan Classification.....	107
Gambar 4.71 SMOTE	108
Gambar 4.72 Perbandingan Distribusi Sebelum dan Sesudah SMOTE	108
Gambar 4.73 Hasil SMOTE.....	109
Gambar 4.74 Data Splitting untuk Pendekatan Prediction.....	109
Gambar 4.75 Hasil Data Splitting untuk Pendekatan Prediction	110
Gambar 4.76 Model Random Forest.....	111
Gambar 4.77 Feature Importance dari Model Random Forest	112
Gambar 4.78 Prediksi Model Random Forest pada Testing Data.....	113
Gambar 4.79 Grid Search pada Model Random Forest	114
Gambar 4.80 Hasil Grid Search pada Model Random Forest.....	115

Gambar 4.81 Prediksi Grid Search pada Model Random Forest	115
Gambar 4.82 Bayesian Optimization pada Model Random Forest.....	116
Gambar 4.83 Hasil Bayesian Optimization pada Model Random Forest	117
Gambar 4.84 Prediksi Bayesian Optimization pada Model Random Forest	118
Gambar 4.85 Model K-Nearest Neighbors Regressor	119
Gambar 4.86 Prediksi K-Nearest Neighbors Regressor pada Testing Data.....	120
Gambar 4.87 Grid Search pada Model K-Nearest Neighbors Regressor	121
Gambar 4.88 Hasil Grid Search pada Model K-Nearest Neighbors Regressor ..	122
Gambar 4.89 Prediksi Grid Search pada Model K-Nearest Neighbors Regressor	122
Gambar 4.90 Bayesian Optimization pada Model K-Nearest Neighbors Regressor	123
Gambar 4.91 Hasil Bayesian Optimization pada Model K-Nearest Neighbors Regressor.....	124
Gambar 4.92 Prediksi Bayesian Optimization pada Model K-Nearest Neighbors Regressor.....	124
Gambar 4.93 Dummy Data	125
Gambar 4.94 Prediksi dari Model Random Forest x Bayesian Optimization pada Dummy Data	126
Gambar 4.95 Tabel Hasil Prediksi dari Klasifikasi Intensitas Penagihan pada Dummy Data	127
Gambar 4.96 Predicted Intensity Category Distribution.....	127
Gambar 4.97 Predicted Intensity Category per User	128
Gambar 4.98 Prediksi dari Model K-Nearest Neighbors Regressor x Bayesian Optimization pada Dummy Data	129
Gambar 4.99 Tabel Hasil Prediksi Channel Recommendation pada Dummy Data	130
Gambar 4.100 Recommended Channel Distribution	130
Gambar 4.101 Predicted Interaction Count per User	131
Gambar 4.102 Confusion Matrix dari Model Random Forest.....	132
Gambar 4.103 Confusion Matrix dari Model Random Forest dengan Grid Search	133
Gambar 4.104 Confusion Matrix dari Model Random Forest dengan Bayesian Optimization.....	134
Gambar 4.105 Random Forest Evaluation	135
Gambar 4.106 Random Forest Accuracy Comparison Plot.....	136
Gambar 4.107 Evaluasi Performa Accuracy pada Model Random Forest	137
Gambar 4.108 K-Nearest Neighbors Regressor Evaluation	139
Gambar 4.109 Mean Absolute Error Comparison Plot.....	140
Gambar 4.110 Evaluasi Performa Mean Absolute Error (MAE) pada Model K-Nearest Neighbors Regressor	141

DAFTAR RUMUS

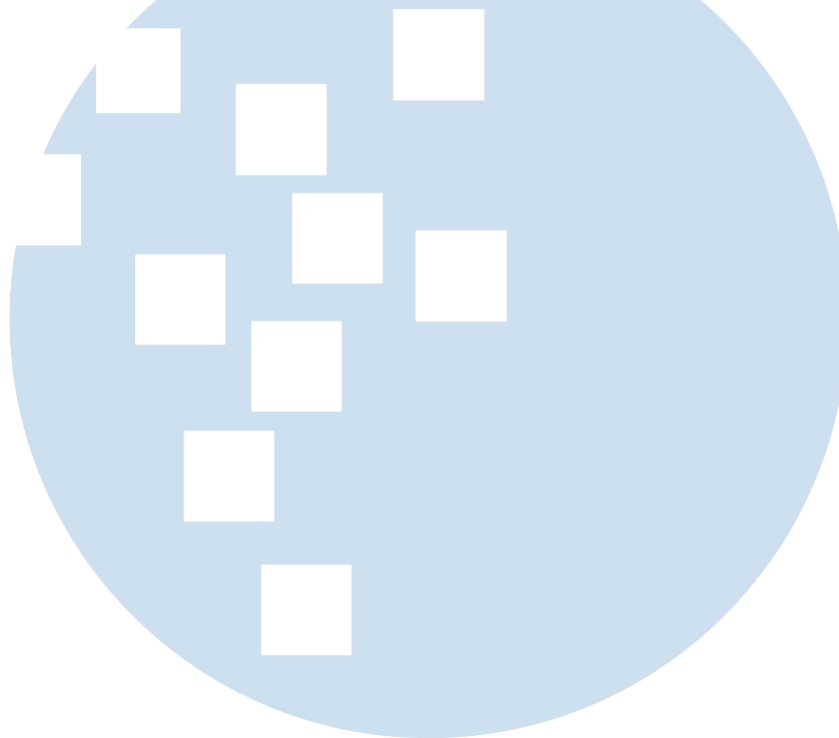
Rumus 2.1 Prior Distribution	20
Rumus 2.2 Posterior Probability Distribution	20
Rumus 2.3 Posterior Mean	21
Rumus 2.4 Posterior Variance	21
Rumus 2.5 Expected Improvement	21
Rumus 2.6 Accuracy	22
Rumus 2.7 Precision	23
Rumus 2.8 Recall	23
Rumus 2.9 F1-Score	24
Rumus 2.10 Mean Absolute Error	24
Rumus 2.11 Mean Square Error	25
Rumus 2.12 Root Mean Square Error	25
Rumus 2.13 Shapiro-Wilk Test	26
Rumus 2.14 Pairwise Difference	27
Rumus 2.15 Mean of Difference	27
Rumus 2.16 Standard Deviation of Difference	27
Rumus 2.17 t-Statistic	27
Rumus 2.18 Margin Function	31
Rumus 2.19 Generalization Error	31
Rumus 2.20 Cosine Similarity	32

UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran A Formulir Konsultasi Skripsi	164
Lampiran B Lampiran Pengecekan Hasil Turnitin.....	166



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA