



SKRIPSI

OPTIMASI DETEKSI SITUS PHISHING MENGUNAKAN METODE ENSEMBLE VOTING BERBASIS MODEL RANDOM FOREST, LOGISTIC REGRESSION, DAN SVM

MUHAMMAD IQBAL FIKRI ROBANI AMIN

NPM 21081010334

DOSEN PEMBIMBING

Dr. I Gede Susrama Mas Diyasa, ST., MT., IPU.

Achmad Junaidi, S.Kom., M.Kom

KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI INFORMATIKA

SURABAYA

2025



SKRIPSI

**OPTIMASI DETEKSI SITUS PHISHING
MENGUNAKAN METODE ENSEMBLE
VOTING BERBASIS MODEL RANDOM
FOREST, LOGISTIC REGRESSION, DAN SVM**

MUHAMMAD IQBAL FIKRI ROBANI AMIN
NPM 21081010334

DOSEN PEMBIMBING

Dr. I Gede Susrama Mas Diyasa, ST., MT.
Achmad Junaidi, S.Kom., M.Kom

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**

Lembar ini sengaja dikosongkan

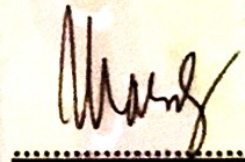
LEMBAR PENGESAHAN

OPTIMASI DETEKSI SITUS PHISHING MENGGUNAKAN METODE ENSEMBLE VOTING BERBASIS MODEL RANDOM FOREST, LOGISTIC REGRESSION, DAN SVM.

Oleh:
MUHAMMAD IQBAL FIKRI ROBANI AMIN
NPM. 21081010334

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Informatika
Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal
16 Mei 2025

Dr. Ir. I Gede Susrama Mas Divasa, ST, MT, IPU
NIP. 19700619 20212 1 1009



(Pembimbing I)

Achmad Junaidi, S.Kom., M.Kom
NPT. 3 7811 04 0199 1



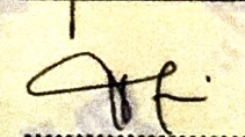
(Pembimbing II)

Yisti Vita Via, S.ST., M.Kom.
NIP. 19860425 202121 2 001



(Ketua Penguji)

Afina Lina Nuriaili, S.Kom., M.Kom.
NIP. 1993121 3202203 2 010



(Anggota Penguji)

Mengetahui,
Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

Lembar ini sengaja dikosongkan

LEMBAR PERSETUJUAN

**OPTIMASI DETEKSI SITUS PHISHING MENGGUNAKAN METODE
ENSEMBLE VOTING BERBASIS MODEL RANDOM FOREST,
LOGISTIC REGRESSION, DAN SVM.**

Oleh :

**MUHAMMAD IQBAL FIKRI ROBANI AMIN
NPM. 21081010334**



**Menyetujui,
Koordinator Program Studi Informatika
Fakultas Ilmu Komputer**

**Fetty Tri Anggraeny, S.Kom. M.Kom
NIP. 19820211 2021212 005**

Lembar ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertandatangan di bawah ini:

Nama Mahasiswa : Muhammad Iqbal Fikri Robani Amin
NPM : 21081010334
Program Studi : Sarjana (S1)
Program Studi : Informatika
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 16 Mei 2025
Yang Membuat Pernyataan,



MUHAMMAD IQBAL FIKRI R. A.
NPM. 21081010334

Lembar ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM	Muhamamd Iqbal Fikri Robani Amin / 21081010334
Judul Skripsi	OPTIMASI DETEKSI SITUS <i>PHISHING</i> MENGUNAKAN METODE <i>ENSEMBLE VOTING</i> BERBASIS MODEL <i>RANDOM FOREST</i> , <i>LOGISTIC REGRESSION</i> , DAN <i>SVM</i> .
Dosen Pembimbing	1. Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU 2. Achmad Junaidi, S.Kom. M.Kom

Serangan phishing merupakan salah satu ancaman siber yang paling umum dan berbahaya, karena dapat menyebabkan pencurian informasi pribadi dan kerugian finansial. Penelitian ini bertujuan untuk membangun sistem deteksi situs phishing menggunakan metode *Ensemble Weighted Voting* yang menggabungkan tiga algoritma klasifikasi, yaitu *Logistic Regression*, *Random Forest*, dan *Support Vector Machine* (SVM). Setiap algoritma memberikan kontribusi berbeda berdasarkan keunggulan masing-masing dalam memproses data. Tahapan awal adalah *preprocessing* yang mencakup *data cleaning*, *standarisasi* dan *feature selection* menggunakan *Random Forest*, yang kemudian dilanjutkan dengan optimasi *hyperparameter* melalui teknik *GridSearchCV* untuk memperoleh parameter terbaik dari setiap model. Masing-masing model selanjutnya akan diuji melalui beberapa skenario pengujian yang melibatkan kombinasi *hyperparameter tuning* dan *feature selection*. Setiap skenario tersebut dievaluasi menggunakan metrik performa seperti *akurasi*, *presisi*, *recall*, *F1-score*, dan *ROC – AUC*. Hasil evaluasi menunjukkan bahwa model *Random Forest* memberikan *akurasi* tertinggi sebesar 99,33% setelah optimasi. Sementara itu, metode *Ensemble Weighted Voting* memberikan hasil terbaik secara keseluruhan, dengan *precision* sebesar 98,97%, *recall* 98,98%, *F1-score* 98,97%, dan nilai *ROC – AUC* mendekati sempurna.

Kata kunci : *Deteksi Phishing, Ensemble Voting, Random Forest, Logistic Regression, SVM, GridSearchCV*

Lembar ini sengaja dikosongkan

ABSTRACT

Student Name / NPM	Muhammad Iqbal Fikri Robani Amin / 21081010334
Thesis Title	OPTIMASI DETEKSI SITUS PHISHING MENGUNAKAN METODE ENSEMBLE VOTING BERBASIS MODEL RANDOM FOREST, LOGISTIC REGRESSION, DAN SVM.
Advisor	1. Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU 2. Achmad Junaidi, S.Kom. M.Kom

Phishing attacks are among the most common and dangerous cyber threats, as they can lead to the theft of personal information and financial losses. This study aims to develop a phishing website detection system using the Ensemble Weighted Voting method, which combines three classification algorithms: Logistic Regression, Random Forest, and Support Vector Machine (SVM). Each algorithm contributes uniquely based on its individual strengths in processing data. The initial stage involves preprocessing, including data cleaning, standardization, and feature selection using Random Forest, followed by hyperparameter optimization using the GridSearchCV technique to obtain the best parameters for each model. Each model is then evaluated through several testing scenarios involving combinations of hyperparameter tuning and feature selection. These scenarios are assessed using performance metrics such as accuracy, precision, recall, F1-score, and ROC – AUC. Evaluation results show that the Random Forest model achieved the highest accuracy of 99.33% after optimization. Meanwhile, the Ensemble Weighted Voting method provided the best overall performance, with a precision of 98.97%, recall of 98.98%, F1-score of 98.97%, and an almost perfect ROC – AUC value. This research demonstrates that the ensemble weighted voting method, when combining the right models, can significantly enhance the accuracy of phishing detection systems and has the potential to be applied in strengthening cybersecurity at both individual and organizational levels.

Keywords: *Phishing detection, Ensemble Voting, Random Forest, Logistic Regression, SVM, GridSearchCV*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Optimasi Deteksi Situs Phishing Menggunakan Metode Ensemble Voting Berbasis Model Random Forest, Logistic Regression, dan SVM”** dengan baik. Penyusunan skripsi ini tidak terlepas dari bantuan beberapa pihak sehingga penulis mengucapkan terimakasih kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT., selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur
3. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom. selaku Koordinator Program Studi Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Bapak Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU, selaku Dosen Pembimbing utama, yang telah memberikan banyak arahan, bimbingan dan bantuan kepada penulis.
5. Bapak Achmad Junaidi, S.Kom., M.Kom, selaku Dosen Pembimbing Kedua atas saran, bimbingan dan arahan yang telah diberikan kepada penulis.
6. Bapak/Ibu dosen informatika Universitas Pembangunan Nasioan “Veteran” Jawa timur, yang telah memberikan ilmu dan wawasan selama masa perkuliahan penulis.
7. Kedua orang tua saya Bapak Subianto dan Ibu Suci Rachmawati, yang senantiasa memberikan doa dan dukungan, baik materi maupun non materi yang tidak ternilai dalam menyelesaikan skripsi ini. Selalu mengusahakan anak tunggalnya dalam mencapai mimpi – mimpinya. Terakhir terimakasih karena telah menjadi panutan dalam hidup penulis.
8. Saudari Novita Dayanti yang selalu menemani di masa sulit penulis dan selalu menjadi patner terbaik bagi penulis.
9. Teman – teman kuliah Faizal, Firhan, Varell, Rafly, Ahill, Bagus, Rheza, Elang,

Napis, Blek, Adhim, Erik, Bagas, Arab, Yazid, Verdy, Bagjo, Bima, Roy, dan masih banyak lagi. See u on top rek.

Penulis menyadari bahwa penyusunan skripsi ini masih memiliki keterbatasan dan kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi penyempurnaan di masa yang akan datang.

Surabaya, 16 Mei 2025

Penulis,



Muhammad Iqbal Fikri Robani Amin

DAFTAR ISI

LEMBAR PENGESAHAN	iii
LEMBAR PERSETUJUAN	v
SURAT BEBAS PLAGIASI.....	vii
ABSTRAK	ix
ABSTRACT	xi
KATA PENGANTAR.....	xiii
DAFTAR ISI	xiv
DAFTAR GAMBAR	xix
DAFTAR TABEL	xxii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Tujuan Penelitian.....	3
1.4. Manfaat Penelitian	4
1.5. Batasan Masalah	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terdahulu	7
2.2 Landasan Teori.....	11
2.2.1 Situs Phishing.....	11
2.2.2 Machine Learning.....	13
2.2.3 Preprocessing Data	14
2.2.4 Hyperparameter Tuning	17
2.2.5 Logistic Regression	18
2.2.6 Support Vector Machine (SVM)	20
2.2.7 Random Forest	22

2.2.8	Ensemble Learning	24
2.2.9	Optimasi	25
2.2.10	Evaluasi	26
BAB III DESAIN DAN IMPLEMENTASI SISTEM.....		29
3.1.	Metodologi Penelitian.....	29
3.2.	Pengumpulan Dataset	30
3.3.	Preprocessing Data.....	34
3.3.1.	Data Cleaning	35
3.3.2.	Encoding	35
3.3.3.	Pembagian Dataset	35
3.3.4.	Standarisasi	37
3.3.5.	Feature Selection	38
3.4.	Implementasi Algoritma	39
3.4.1.	Algoritma Logistic Regression.....	39
3.4.2.	Algoritma Support Vector Machine	41
3.4.3.	Algoritma Random Forest	44
3.4.4.	Implementasi Algoritma Ensemble	47
3.5.	Hyperparameter Tuning.....	49
3.5.1	Implementasi Hyperparametertuning	50
3.6.	Model Development.....	51
3.7.	Evaluasi Model	53
3.8.	Skenario Pengujian	55
BAB IV PENGUJIAN DAN ANALISA		57
4.1.	Persiapan data	59
4.2.	Preprocessing Data	61
4.2.1.	Data Cleaning	61
4.2.2.	Encoding.....	62
4.2.3.	Memisahkan Fitur dan label	64
4.2.4.	Pembagian data	65

4.2.5.	Standarisasi data	67
4.2.6.	Feature Selection	68
4.3.	Model Random Forest.	70
4.4.	Model SVM	73
4.5.	Model Logistic Regression.....	75
4.6.	Hyperparameter tuning.....	78
4.6.1	Model Random Forest Hyperparameter Tuning	78
4.6.2	Model SVM Hyperparameter Tuning.....	80
4.6.3	Model Logistic Regression Hyperparameter Tuning	83
4.6.4	Model Ensemble Hyperparameter Tuning.....	85
4.7.	Skenario Pengujian	88
4.7.1.	Skenario Pengujian Logistic Regression.....	88
4.7.1.1.	Skenario Pengujian dengan Feature Selection dan GridSeaercCV	88
4.7.1.2.	Skenario Pengujian tanpa Feature Selection, dengan GridSerchCV	91
4.7.1.3.	Skenario Pengujian dengan Feature Selection, tanpa GridSearchCV	94
4.7.1.4.	Skenario Pengujian tanpa Feature Selection, dengan GridSearchCV	96
4.7.2.	Skenario Pengujian Support Vector Machine.....	100
4.7.2.1.	Skenario Pengujian dengan Feature Selection dan GridSearchCV	100
4.7.2.2.	Skenario Pengujian tanpa Feature Selection, dengan GridSearchCV	103
4.7.2.3.	Skenario Pengujian dengan Feature Selection, tanpa GridSearchCV	105
4.7.2.4.	Skenario Pengujian tanpa Feature Selection dan GridSearchCV	108
4.7.3.	Skenario Pengujian Random Forest	111
4.7.3.1.	Skenario Pengujian dengan Feature Selection dan GridSearchCV	111
4.7.3.2.	Skenario Pengujian tanpa Feature Selection, dengan GridSearchCV	114
4.7.3.3.	Skenario Pengujian dengan Feature Selection, tanpa GridSearchCV	117
4.7.3.4.	Skenario Pengujian tanpa Feature Selection dan GridSearchCV	119
4.7.4.	Skenario Pengujian Ensemble Model Bobot berbeda	123
4.7.4.1.	Skenario Pengujian dengan Feature Selection dan GridSearchCV	123
4.7.4.2.	Skenario Pengujian tanpa Feature Selection dengan GridSearchCV	126
4.7.4.3.	Skenario Pengujian dengan Feature Selection, tanpa GridSearchCV	129

4.7.4.4.	Skenario Pengujian tanpa Feature Selection dan GridSearchCV	131
4.7.5.	Skenario Pengujian Ensemble Model Bobot Sama	134
4.7.5.1.	Skenario Pengujian dengan Feature Selection dan GridSearchCV	135
4.7.5.2.	Skenario Pengujian tanpa Feature Selection, dengan GridSearchCV	137
4.7.5.3.	Skenario Pengujian dengan Feature Selection, tanpa GridSearchCV	140
4.7.5.4.	Skenario Pengujian tanpa Feature Selectio dan GridSearch CV	143
4.7.6.	Skenario Pengujian Semua Data Model Ensemble.	149
4.7.6.1.	Skenario Pengujian Ensemble bobot Berbeda Terbaik.	149
4.7.6.2.	Skenario Pengujian Ensemble Bobot Sama Terbaik.	153
4.8.	Konfigurasi Model Terbaik	157
4.9.	Evaluasi Model Terbaik	158
4.9.1	Evaluasi Grafik Model Terbaik	158
4.9.2	Evaluasi Confusion Matrix Model Terbaik	160
4.9.3	Evaluasi ROC AUC Model Terbaik	161
4.10.	Hasil Klasifikasi	162
BAB V PENUTUP		166
5.1.	Kesimpulan	167
5.2.	Saran	169
DAFTAR PUSTAKA		170

DAFTAR GAMBAR

Gambar 2. 1. Alur Situs Phishing. [8].	11
Gambar 2. 2. Model Arsitektur Logistic Regression [13].	18
Gambar 2. 3. Model Arsitektur Support Vector Machine [14]	21
Gambar 2. 4. Model Arsitektur Random Forest[7].	23
Gambar 2. 5. Confusion Matrix[1]	26
Gambar 2. 6. ROC Curve False Positive Rate[21]	28
Gambar 3. 1. Alur Penelitian.	29
Gambar 3. 2. Dataset Kaggle.	31
Gambar 3. 3. Alur Preprocessing Data	34
Gambar 3. 4. Algoritma Logistic Regression.	39
Gambar 3. 5. Algoritma SVM	41
Gambar 3. 6. Algoritma Random Forest.	44
Gambar 3. 7. Alur Implementasi Ensemble	48
Gambar 3. 8 Algoritma Hyperparameter tuning.	50
Gambar 3. 9 Alur Hyperparamater Tuning	50
Gambar 3. 10 Alur Data Modelling	52
Gambar 4. 1. Mount Drive.	60
Gambar 4. 2. Dataset berhasil dimuat	60
Gambar 4. 3. Hasil Missing Values	62
Gambar 4. 4. Hasil Pengecekan Baris Duplikat	62
Gambar 4. 5. Hasil penghapusan fitur unamed0	62
Gambar 4. 6. Hasil Encoding	63
Gambar 4. 7. Hasil Pemisahann Fitur dan Label	64
Gambar 4. 8. Hasil Pembagian Data	66
Gambar 4. 9. Visualisasi Hasil Pembagian Data.	66
Gambar 4. 10. Hasil Standarisasi	68
Gambar 4. 11. Hasil Feature Selection	70
Gambar 4. 12. Hasil Feature Selection	70
Gambar 4. 13. Training Loss Model Random Forest	72
Gambar 4. 14. Training Loss SVM	75
Gambar 4. 15. Training Loss Logistic Regression	77
Gambar 4. 16. Parameter Terbaik Random Forest	80
Gambar 4. 17. Parameter Terbaik SVM	82
Gambar 4. 18. Parameter Terbaik Logistic Regression.	84
Gambar 4. 19. Bobot Model	86
Gambar 4. 20. Hasil Ensemble bobot sama.	87
Gambar 4. 21. Hasil Evaluasi Skenario 1 Model Logistic Regression.	89
Gambar 4. 22. Classification Report Logistic Regression Skenario 1	90
Gambar 4. 23. Confusion Matrix Logistic Regression Skenario 1	90
Gambar 4. 24. Hasil Evaluasi Matrix Skenario 2 Model Logistic Regression.	92
Gambar 4. 25. Classification Report Skenario 2 Logistic Regression	93
Gambar 4. 26. Confusion Matrix Skenario 2 Logistic Regression.	93
Gambar 4. 27. Hasil Evaluasi Matrix Model Logistic Regression	95
Gambar 4. 28. Classification Report Skenario 3 Logistic Regression	95
Gambar 4. 29. Confusin Matrix skenario 3 Logistic Regression	96
Gambar 4. 30. Hasil Evaluasi Matrix Skenario 4 Model Logistic Regression	98
Gambar 4. 31. Classification Report Skenario pengujian 4 logistic regression	98
Gambar 4. 32. Confusion Matrix Skenario Pengujian 4 Logistic Regression.	99

Gambar 4. 33.	Hasil Evaluasi Matrix skenario 1 Model SVM.....	101
Gambar 4. 34.	Classification Report skenario pengujian 1 Model SVM	102
Gambar 4. 35.	Confusion Matrix Skenario 1 SVM	102
Gambar 4. 36.	Hasil Evaluasi Matrix Skenario 2 Model SVM.....	104
Gambar 4. 37.	Classification Report Skenario 2 SVM.....	104
Gambar 4. 38.	Confusion Matrix Skenario 2 SVM	105
Gambar 4. 39.	Hasil Evaluasi Matrix Skenario 3 Model SVM	107
Gambar 4. 40.	Classification Report Skenario Pengujian 3 SVM.....	107
Gambar 4. 41.	Confusion Matrix Skenario Pengujian 3 SVM	107
Gambar 4. 42.	Hasil Evaluasi Matrix Skenario 4 Model SVM	109
Gambar 4. 43.	Classification Skenario Pengujian 4 SVM.....	110
Gambar 4. 44.	Confusion Matrix Skenario pengujian 4 SVM	110
Gambar 4. 45.	Hasil Evaluasi Matrix Skenario 1 Model Random Forest	113
Gambar 4. 46.	Classification Report Skenario 1 Random Forest.....	113
Gambar 4. 47.	Confusion Matrix Skenario 1 Random Forest.	114
Gambar 4. 48.	Hasil Evaluasi Matrix Skenario 2 Model Random Forest	115
Gambar 4. 49.	Classification Report Skenario 2 Random Forest.....	116
Gambar 4. 50.	Confusion Matrix Skenario 2 Random Forest	116
Gambar 4. 51.	Hasil Evaluasi Matrix Skenario 3 Random Forest.....	118
Gambar 4. 52.	Classification Report skenario 3 Random Forest.....	118
Gambar 4. 53.	Confusion Matrix Skenario 3 Random Forest	119
Gambar 4. 54.	Hasil Evaluasi Matrix Skenario 4 Random Forest.....	121
Gambar 4. 55.	Classification Report Skenario 4 Random Forest.....	121
Gambar 4. 56.	Confusion Matrix Skenario 4 Random Forest	122
Gambar 4. 57.	Hasil Evaluasi Matrix Skenario 1 Model Ensemble bobot berbeda	124
Gambar 4. 58.	Classification Report Skenario 1 Ensemble Bobot Berbeda.....	125
Gambar 4. 59.	Confusion Matrix Skenario 1 Ensemble Bobot Berbeda	125
Gambar 4. 60.	Hasil Evaluasi Matrix Skenario 2 Model Ensemble bobot Berbeda.....	127
Gambar 4. 61.	Classification Report Skenario 2 Ensemble Bobot Berbeda.....	128
Gambar 4. 62.	Confusion Matrix Skenario 2 Ensemble Bobot Berbeda	128
Gambar 4. 63.	Hasil Evaluasi Matrix Skenario 3 Model Ensemble Bobot Berbeda	130
Gambar 4. 64.	Classification Report Skenario 3 Ensemble Bobot Berbeda.....	130
Gambar 4. 65.	Confusion Matrix Skenario 3 Ensemble Berbeda.....	131
Gambar 4. 66.	Hasil Evaluasi Skenario 4 Model Ensemble Bobot Berbeda.....	133
Gambar 4. 67.	Classification Report Skenario 4 Ensemble Bobot Berbeda.....	133
Gambar 4. 68.	Confusion Matrix Skenario 4 Ensemble Bobot Berbeda	134
Gambar 4. 69.	Hasil Evaluasi Matrix Skenario 1 Model Ensemble Bobot sama	136
Gambar 4. 70.	Classification Report Skenario 1 Model Ensemble Bobot Sama.....	136
Gambar 4. 71.	Confusion Matrix Skenario 1 Model Ensemble Bobot Sama	137
Gambar 4. 72.	Hasil Evaluasi Matrix Skenario 2 Model Ensemble Bobot Sama.....	139
Gambar 4. 73.	Classification Report Skenario 2 Model Ensemble Bobot Sama.....	139
Gambar 4. 74.	Confusion Matrix Skenario 2 Model Ensemble Bobot Sama	140
Gambar 4. 75.	Hasil Evaluasi Matrix Skenario 3 Model Ensemble Bobot Sama.....	142
Gambar 4. 76.	Classification Report Skenario 3 Ensemble Bobot Sama.	142
Gambar 4. 77.	Confusion Matrix Skenario 3 Model Ensemble Bobot Sama	143
Gambar 4. 78.	Hasil Evaluasi Matrix Skenario 4 Model Ensemble bobot sama.....	145
Gambar 4. 79.	Classification Report skenario 4 Model Ensemble	145
Gambar 4. 80.	Confusion Matris Skenario 4 Model Ensemble Bobot Sama	146
Gambar 4. 81.	Hasil Evaluasi Matrix Semua data model Ensemble bobot Berbeda.....	151
Gambar 4. 82.	Classification Report Semua data model Ensemble bobot berbeda.....	152

Gambar 4. 83.	Confusion Matrix Semua data Model Ensemble bobot berbeda	153
Gambar 4. 84.	Hasil Evaluasi Semua Data Ensemble Model sama	156
Gambar 4. 85.	Classification Report Semua data Ensemble model sama	156
Gambar 4. 86.	Confusion Matrix Semua data Ensemble Model bobot Sama	157
Gambar 4. 87.	Training Loss Model Terbaik	159
Gambar 4. 88.	Kurva Akurai Model Terbaik	159
Gambar 4. 89.	Confusion Matrix Model Terbaik	160
Gambar 4. 90.	Evaluasi Best model dengan 4 Matrix	161
Gambar 4. 91.	ROC - AUC untuk Semua Model.....	161

DAFTAR TABEL

Tabel 2. 1. Penelitian Terdahulu.....	7
Tabel 3. 1. Dataset phishing.....	31
Tabel 3. 2. Daftar Fitur Dataset	32
Tabel 3. 3. Pembagian Dataset	36
Tabel 3. 4. Sample Data Standarisasi	37
Tabel 3. 5 Data Hasil Standarisasi.....	38
Tabel 3. 6. Tabel Sample Data Untuk perhitungan Logistic Regression.....	40
Tabel 3. 7. Koefisien Perhitungan Logistic Regression	40
Tabel 3. 8. Hasil Perhitungan Manual Logistic Regression	41
Tabel 3. 9 Sampel Data SVM.....	42
Tabel 3. 10. Hasil Perhitungan Manual SVM.....	43
Tabel 3. 11. Sample Data Random Forest	44
Tabel 3. 12. Bootstrap Sampling Pohon 1	45
Tabel 3. 13 Nilai Gini	46
Tabel 3. 14. Hasil Klasifikasi Random Forest.....	47
Tabel 3. 15. Sample Data Ensemble.....	48
Tabel 3. 16. Sample data Evaluasi.....	54
Tabel 3. 17. Tabel Parameter Model	55
Tabel 3. 18. Skema Skenario Pengujian	56
Tabel 4. 1 Skenario Pengujian 1 Logistic Regression.....	88
Tabel 4. 2. Skenario Pengujian 2 Logistic Regression	91
Tabel 4. 3. Skenario Pengujian 3 Logistic Regression	94
Tabel 4. 4. Skenario Pengujian 4 Logistic Regression	96
Tabel 4. 5. Hasil pengujian Logistic Regression	99
Tabel 4. 6. Skenario 1 pengujian SVM	100
Tabel 4. 7. Skenario 2 pengujian SVM	103
Tabel 4. 8. Skenario 3 pengujian SVM	105
Tabel 4. 9. Skenario 4 pengujian SVM	108
Tabel 4. 10. Hasil Skenario Pengujian SVM.....	111
Tabel 4. 11. Skenario 1 pengujian Random Forest.....	111
Tabel 4. 12. Skenario 2 pengujian Random Forest.....	114
Tabel 4. 13. Skenario 3 pengujian Random Forest.....	117
Tabel 4. 14. Skenario 4 pengujian Random Forest.....	119
Tabel 4. 15. Hasil Skenario Pengujian Random Forest.....	122
Tabel 4. 16. Skenario 1 pengujian Ensemble	123
Tabel 4. 17. Skenario 2 pengujian Ensemble	126
Tabel 4. 18. Skenario 3 pengujian Ensemble	129
Tabel 4. 19. Skenario 4 pengujian Ensemble	131
Tabel 4. 20. Skenario Pengujian 1 Ensemble Bobot Sama.....	135
Tabel 4. 21. Skenario Pengujian 2 Ensemble bobot sama.....	137
Tabel 4. 22. Skenario Pengujian 3 Ensemble bobot sama.....	140
Tabel 4. 23. Skenario Pengujian 4 Ensemble	143
Tabel 4. 24. Hasil Skenario Pengujian Ensemble.....	146
Tabel 4. 25. Hasil Skenario Pengujian	147
Tabel 4. 26. Skenario Pengujian Ensemble Semua Data.....	149
Tabel 4. 27. Skenario Pengujian Semua data Model Ensemble Bobot Sama.....	154
Tabel 4. 28. Konfigurasi Model Terbaik	158
Tabel 4. 30. Hasil Klasifikasi	162