



UNIVERSITAS DIPONEGORO

**IMPLEMENTASI *MACHINE LEARNING* UNTUK
MEMPREDIKSI KELAYAKAN KREDIT PENGGUNA GAWAI
TELEPON GENGAM DI INDONESIA PADA MASA
PANDEMI COVID-19**

TESIS

Diajukan sebagai salah satu syarat untuk memperoleh gelar Magister Teknik

RADEN RARA KARTIKA KUSUMA WINAHYU

21060120420020

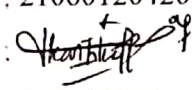
**FAKULTAS TEKNIK
DEPARTEMEN TEKNIK ELEKTRO
PROGRAM STUDI MAGISTER**

SEMARANG

JUNI 2022

HALAMAN PERNYATAAN ORISINALITAS

**Tesis ini adalah hasil karya saya sendiri,
dan semua sumber baik yang dikutip maupun yang dirujuk
telah saya nyatakan dengan benar.**

NAMA : Raden rara Kartika Kusuma
NIM : 21060120420020
Tanda Tangan : 
Tanggal : 2 Juni 2022

HALAMAN PENGESAHAN

Tesis ini diajukan oleh:

NAMA : Raden rara Kartika Kusuma Winahyu
NIM : 21060120420020
Departemen/Program Studi : Teknik Elektro / Magister (S2)
Judul Tesis : Implementasi *Machine Learning* untuk
Memprediksi Kelayakan Kredit Pengguna Gawai
Telepon Genggam di Indonesia pada Masa Pandemi
COVID-19

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Teknik pada Program Studi Magister, Departemen Teknik Elektro, Fakultas Teknik, Universitas Diponegoro.

TIM PENGUJI

Pembimbing 1 : Dr. Maman Somantri, S.T., M.T. (.....)
Pembimbing 2 : Dr. Oky Dwi Nurhayati, S.T., MT. (.....)
Penguji 1 : Dr. Iwan Setiawan, S.T., M.T. (.....)
Penguji 2 : Dr. Eng. Wahyul Amien Syafei, S.T., M.T. (.....)
Penguji 3 : Mochammad Facta, S.T., M.T., Ph.D (.....)

Semarang, 2 Juni 2022

Ketua Departemen Teknik Elektro



Aghus Sofwan, S.T., M.T., Ph.D

NIP. 197302041997021001

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TESIS UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademika Universitas Diponegoro, saya yang bertanda tangan di bawah ini:

Nama : RADEN RARA KARTIKA KUSUMA WINAHYU
NIM : 21060120420020
Program Studi : MAGISTER (S2)
Departemen : TEKNIK ELEKTRO
Fakultas : TEKNIK

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Diponegoro Hak Bebas Royalti Noneksklusif (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul :

IMPLEMENTASI *MACHINE LEARNING* UNTUK MEMPREDIKSI
KELAYAKAN KREDIT PENGGUNA GAWAI TELEPON GENGAM DI
INDONESIA PADA MASA PANDEMI COVID-19

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Diponegoro berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Semarang
Pada Tanggal : 2 Juni 2022

Yang menyatakan,



(RADEN RARA KARTIKA
KUSUMA WINAHYU)

ABSTRAK

Penelitian ini bertujuan untuk memprediksi kelayakan kredit pengguna gawai telepon genggam di Indonesia pada masa pandemi COVID-19 dengan menggunakan algoritma machine learning. Algoritma PCA dan K-means digunakan untuk mengelompokkan kelayakan kredit berdasarkan dataset yang terdiri dari duabelas pertanyaan yang diajukan kepada 1050 pengguna gawai telepon genggam di Indonesia selama masa pandemi COVID-19 (Juli-Agustus 2021). Selanjutnya, empat algoritma klasifikasi yang berbeda yaitu : Logistic Regression, Support Vector Machine (SVM), Decision Tree, dan Naïve Bayes (NB) digunakan untuk memprediksi kelayakan kredit pengguna gawai telepon di Indonesia dan membandingkan kinerja keempat algoritma klasifikasi tersebut dengan melakukan pengujian akurasi, presisi, recall, F1-score, dan Area Under Curve Receiver Operating Characteristics (AUCROC). Algoritma Logistik Regression keluar sebagai algoritma dengan performa yang sempurna pada penelitian ini, sedangkan algoritma Naïve Bayes (NB) menjadi algoritma dengan performa yang paling rendah. Hasil penelitian ini juga memberikan pengetahuan baru tentang variabel yang paling berpengaruh dan tidak berpengaruh bagi pengguna gawai telepon genggam di Indonesia pada masa pandemi COVID-19 yang dapat membantu lembaga keuangan di Indonesia khususnya perbankan dalam menilai kelayakan kredit calon nasabah di masa pandemi COVID-19.

Kata Kunci: Kelayakan kredit, Telepon Genggam, Machine learning, Pandemi

ABSTRACT

This study aimed to predict the creditworthiness of smartphone users in Indonesia during the COVID-19 pandemic using machine learning algorithms. The PCA and K-means algorithms were used to group the creditworthiness based on the dataset consisting of twelve questions posed to 1050 smartphone users in Indonesia during the COVID-19 pandemic (July-August 2021). Furthermore, four different classification algorithms, namely: Logistic Regression, Support Vector Machine (SVM), Decision Tree, and Naïve Bayes (NB) were used to predict the creditworthiness of smartphone users in Indonesia and compare their performance by testing the accuracy, precision, recall, F1-score, and Area Under Curve Receiver Operating Characteristics (AUCROC). The logistic Regression algorithm came out as the algorithm with the perfect performance in this study, while the Naïve Bayes (NB) algorithm became the algorithm with the lowest performance. The results of this study also provide new knowledge about the most influential and non-influential variables for smartphone users in Indonesia during the COVID-19 pandemic that can help financial institutions in Indonesia, especially banks, in assessing the creditworthiness of prospective customers during the COVID-19 pandemic.

Keywords: Creditworthiness, Smartphone, Machine learning, Pandemic.

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Allah Subhanahu wa Ta'ala atas rahmat, karunia, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan tesis dan penyusunan laporan ini.

Tesis dengan judul “Implementasi *Machine Learning* untuk Memprediksi Kelayakan Kredit Pengguna Gawai Telepon Genggam di Indonesia pada Masa Pandemi COVID-19” ini diajukan untuk memenuhi syarat akhir dalam menyelesaikan pendidikan Program Studi Magister Departemen Teknik Elektro, Fakultas Teknik, Universitas Diponegoro, Semarang.

Rasa terima kasih yang tulus penulis ucapkan kepada semua pihak yang telah membantu selama pelaksanaan tesis ini.

1. Bapak Aghus Sofwan, S.T., M.T., Ph.D selaku Ketua Departemen Teknik Elektro Fakultas Teknik Universitas Diponegoro Semarang.
2. Bapak Mochammad Facta, S.T., M.T., Ph.D selaku Ketua Program Studi Magister Teknik Elektro.
3. Bapak Dr. Maman Somantri, S.T., M.T. selaku Dosen Pembimbing I.
4. Ibu Dr. Oky Dwi Nurhayati, S.T., M.T. selaku Dosen Pembimbing II.
5. Bapak/Ibu dosen Departemen Teknik Elektro Fakultas Teknik Universitas Diponegoro Semarang yang membimbing penulis hingga menyelesaikan semua mata kuliah.
6. Bapak/Ibu karyawan tata usaha Departemen Teknik Elektro Fakultas Teknik Universitas Diponegoro Semarang yang telah membantu dalam proses menuju sidang akhir.
7. Kedua orangtua, kakak, adik, dan suami yang telah menjadi sosok penyemangat yang selalu bisa memberikan inspirasi, dukungan, dan motivasi bagi penulis dalam keadaan apa pun.

Semarang, 2 Juni 2022

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI.....	iv
ABSTRAK	v
ABSTRACT.....	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xv
BAB I.....	1
1.1. Latar Belakang.....	1
1.2. Tujuan Penelitian	7
1.3. Perumusan Masalah	7
1.4. Batasan Masalah	8
1.5. Sistematika Penulisan	8
BAB II.....	10
2.1. Kajian Hasil Penelitian Terdahulu	10
2.2. Sistem Skor untuk Kredit.....	15
2.3. <i>Machine Learning</i>	6
2.3.1. <i>Logistic Regression</i>	7
2.3.2. <i>Decision Tree Classifier</i>	18
2.3.3. <i>Support Vector Machine</i>	19
2.3.4. <i>Naïve Bayes Classifier</i>	20
2.3.5. <i>Principal Component Analysis</i>	20
2.3.6. <i>K-means Clustering</i>	21
2.4. Tahapan dalam Proses <i>Machine Learning</i>	22
2.4.1 <i>Pre-processing / Data cleaning</i>	22
2.4.2 Analisa dan Persiapan Data	22
2.4.3 Pembangunan Model dan Pengujian Performa.....	24

2.5.	Kebaruan yang Diusulkan.....	28
BAB III		
3.1	Perlengkapan yang Digunakan	29
3.2	Tahap-Tahap Penelitian	29
3.3.	Metode Pengumpulan Data.....	31
3.3.1.	Responden.....	31
3.3.2.	Kuesioner	32
3.3.3.	<i>Flowchart Diagram</i>	35
3.3.4.	Langkah Mengimplementasikan Algoritma ML	36
3.3.5.	Cara Mengukur Evaluasi Performa.....	40
BAB IV		
4.1.	Persiapan Data	44
4.1.1.	Penyusunan dan Pengambilan Data Kuesioner.....	44
4.2.	Pengolahan Data dengan <i>Machine Learning</i>	46
4.2.1	<i>Pre-processing / Data cleaning</i>	46
4.2.2	Analisa dan Persiapan Data	48
4.2.2.1	Teknik Menghilangkan <i>Outlier</i>	48
4.2.2.2	Analisa Dataset menggunakan <i>Seaborn / Heatmap</i>	54
4.2.2.3	Proses Reduksi Dimensi dengan PCA	55
4.2.2.4	Proses <i>Clustering</i> dengan K-means	58
4.2.2.5	Analisa Data ANOVA, Bivariate Analysis, VIF Test	61
4.2.2.6	Pengujian Performa Algoritma <i>Clustering</i>	66
4.2.3.	Pembangunan Model Algoritma Klasifikasi <i>Machine Learning</i>	67
4.2.3.1.	Implementasi <i>Logistic Regression</i> dan Pengujian Performa.....	67
4.2.3.2.	Implementasi <i>Decision Tree</i> dan Pengujian Performa.....	70
4.2.3.3.	Implementasi SVM dan Pengujian Performa	72
4.2.3.4.	Implementasi <i>Naïve Bayes</i> dan Pengujian Performa	74
4.2.3.5.	Ringkasan Hasil Pengujian Performa dan Analisis	76
4.2.3.6.	<i>Assesment</i> Akurasi Algoritma Klasifikasi	77
4.2.3.6.1	<i>Assesment</i> Akurasi <i>Logistic Regression</i>	77
4.2.3.6.2	<i>Assesment</i> Akurasi <i>Decision Tree</i>	78

4.2.3.6.3 <i>Assesment</i> Akurasi SVM.....	79
4.2.3.6.4 <i>Assesment</i> Akurasi <i>Naïve Bayes</i>	79
4.2.3.6.5 Ringkasan Hasil <i>Assesment</i> Akurasi dan Analisis	80
4.3. Perbandingan Antar Kelas dan Hasil Analisa	80
BAB V.....	82
5.1. Kesimpulan	82
5.2. Saran	83
DAFTAR PUSTAKA	85
BIODATA	
LAMPIRAN	

DAFTAR GAMBAR

Gambar 2. 1 Pendekatan <i>Machine Learning</i>	17
Gambar 2. 2 Struktur Algoritma <i>Decision Tree</i>	19
Gambar 2. 3 Contoh <i>Decision Tree</i>	19
Gambar 2. 4 Matriks Kesalahan (<i>Confusion Matrix</i>).....	5
Gambar 2. 5 Contoh Kurva AUCROC	7
Gambar 3. 1 Grafik Distribusi Usia Responden.....	32
Gambar 3. 2 <i>Diagram Pie</i> Distribusi Pekerjaan Responden.....	32
Gambar 3. 3 Data Hasil Kuesioner	34
Gambar 3. 4 <i>Diagram Flowchart</i> Penelitian.....	35
Gambar 3. 5 Contoh Matriks Kesalahan Multi Kelas (3 kelas)	41
Gambar 4.1 Pengecekan Nilai <i>Null</i>	46
Gambar 4.2 Pemisahan Jenis Data	47
Gambar 4.3 Evaluasi Data Tiap Kategori	47
Gambar 4.4 Manipulasi Data Kategoris menjadi Numerik	48
Gambar 4.5 <i>Box Plot</i> Usia.....	49
Gambar 4.6 <i>Box Plot</i> Pekerjaan	50
Gambar 4.7 <i>Box Plot</i> Merk	50
Gambar 4.8 <i>Box Plot</i> Frekuensi Ganti	50
Gambar 4.9 <i>Box Plot</i> Alasan Ganti	51
Gambar 4.10 <i>Box Plot</i> Biaya <i>Smartphone</i>	51
Gambar 4.11 <i>Box Plot</i> Penggunaan Data Media Sosial.....	51
Gambar 4.12 <i>Box Plot</i> Sisa Uang.....	52
Gambar 4.13 <i>Box Plot</i> Penggunaan <i>Data Game</i>	53
Gambar 4.14 <i>Box Plot</i> Biaya Internet	53
Gambar 4.15 <i>Box Plot</i> Menabung.....	54
Gambar 4.16 <i>Box Plot</i> Bayar Sendiri.....	54
Gambar 4.17 <i>Seaborn Heatmap</i> Korelasi Antar Variabel	55
Gambar 4.18 Proses Standarisasi, Normalisasi Data, dan Proses PCA	56

Gambar 4.19 Grafik Kumulatif Menjabarkan Varians tiap Komponen.....	56
Gambar 4.20 Proses PCA dengan memilih nilai n komponen sejumlah 2	56
Gambar 4.21 <i>Scatter</i> diagram hasil proses PCA	57
Gambar 4.22 Proses Mengetahui 2 Variabel Komponen Teratas	57
Gambar 4.23 Plot SSE (<i>sum of square error</i>) untuk Tiap n-Cluster	58
Gambar 4.24 <i>Scatter Diagram Clustering</i> dengan Algoritma K-means.....	59
Gambar 4.25 Hasil Visualisasi <i>Clustering</i> dengan Algoritma K-means	59
Gambar 4.26 Proses Mengetahui Titik Pusat <i>Cluster</i>	60
Gambar 4.27 Distribusi Jumlah Masing-masing <i>Cluster</i>	60
Gambar 4.28 Presentase <i>Cluster</i> Berdasarkan Distribusinya.....	61
Gambar 4.29 Hasil Tes ANOVA	61
Gambar 4.30 <i>Box Plot</i> Biaya <i>Smartphone</i> per <i>Cluster</i>	63
Gambar 4.31 <i>Box Plot</i> Biaya Internet per <i>Cluster</i>	63
Gambar 4.32 <i>Box Plot</i> Menabung per <i>Cluster</i>	64
Gambar 4.33 <i>Box Plot</i> Sisa Uang per <i>Cluster</i>	64
Gambar 4.34 <i>Box Plot</i> Penggunaan Data Media Sosial.....	65
Gambar 4.35 Hasil VIF Test	66
Gambar 4.36 Nilai Skor <i>Silhouette</i>	66
Gambar 4.37 Proses Pemisahan Data.....	67
Gambar 4.38 Pemodelan Algoritma <i>Logistic Regression</i>	68
Gambar 4.39 Hasil Prediksi Model Algoritma <i>Logistic Regression</i>	68
Gambar 4.40 Matriks Kesalahan Algoritma <i>Logistic Regression</i>	68
Gambar 4.41 Skor presisi dan akurasi Algoritma <i>Logistic Regression</i>	69
Gambar 4.42 Laporan Klasifikasi Algoritma <i>Logistic Regression</i>	69
Gambar 4.43 Pemodelan Algoritma <i>Decision Tree</i>	70
Gambar 4.44 Hasil Prediksi Model Algoritma <i>Decision Tree</i>	70
Gambar 4.45 Matriks Kesalahan (<i>Confusion Matrix</i>) Algoritma <i>Decision Tree</i> .	71
Gambar 4.46 Skor presisi dan akurasi Algoritma <i>Decision Tree</i>	71
Gambar 4.47 Laporan Klasifikasi Algoritma <i>Decision Tree</i>	71
Gambar 4.48 Pemodelan Algoritma <i>Support Vector Machine</i> (SVM).....	72
Gambar 4.49 Hasil Prediksi Model Algoritma <i>Support Vector Machine</i> (SVM). 72	

Gambar 4.50 Matriks Kesalahan Algoritma <i>Support Vector Machine</i> (SVM).....	73
Gambar 4.51 Skor Presisi dan Akurasi Algoritma SVM.....	73
Gambar 4.52 Laporan Klasifikasi Algoritma SVM.....	74
Gambar 4.53 Pemodelan Algoritma <i>Naïve Bayes</i>	74
Gambar 4.54 Hasil Prediksi Model Algoritma <i>Naïve Bayes</i>	75
Gambar 4.55 Matriks Kesalahan (<i>Confusion Matrix</i>) Algoritma <i>Naïve Bayes</i>	75
Gambar 4.56 Skor Presisi dan Akurasi Algoritma <i>Naïve Bayes</i>	76
Gambar 4.57 Laporan Klasifikasi Algoritma <i>Naïve Bayes</i>	76
Gambar 4.58 Hasil <i>Assessment</i> Akurasi Algoritma <i>Logistic Regression</i>	78
Gambar 4.59 Hasil <i>Assessment</i> Akurasi Algoritma <i>Decision Tree</i>	78
Gambar 4.60 Hasil <i>Assessment</i> Akurasi Algoritma SVM	79
Gambar 4.61 Hasil <i>Assessment</i> Akurasi Algoritma <i>Naïve Bayes</i>	79

DAFTAR TABEL

Tabel 1. 1 Perbandingan Penelitian Terdahulu dan Penelitian Penulis.....	3
Tabel 2. 1 Variabel Kualitatif dari Data <i>Survey</i>	13
Tabel 2. 2 Variabel Kuantitatif dari Data <i>Survey</i>	14
Tabel 2. 3 Variabel yang Berpengaruh Hasil Perhitungan Algoritma MCFS.....	14
Tabel 2. 4 Kriteria Subyektif Kualitas Pengelompokkan Berdasarkan Koefisien <i>Silhouette</i>	24
Tabel 2. 5 Nilai AUC dan Keterangannya	27
Tabel 3. 1 Variabel Kualitatif pada Pertanyaan Survey	33
Tabel 3. 2 Variabel Kuantitatif pada Pertanyaan Survey	33
Tabel 4. 1 Implementasi Parameter Kuesioner ke Variabel Data untuk Diolah dengan <i>Machine Learning</i>	45
Tabel 4. 2 Hasil Parameter Pengukuran Algoritma <i>Logistic Regression</i>	69
Tabel 4. 3 Hasil Parameter Pengukuran Algoritma <i>Decision Tree</i>	72
Tabel 4. 4 Hasil Parameter Pengukuran Algoritma SVM.....	74
Tabel 4. 5 Hasil Parameter Pengukuran Algoritma Naïve Bayes	76
Tabel 4. 6 Perbandingan Algoritma Berdasarkan Hasil Pengecekan Performa....	76
Tabel 4. 7 Perbandingan Pengujian AUCROC antar Algoritma.	80
Tabel 4. 8 Perbandingan Variabel Antar Kelas.....	80