



**PENERAPAN KLASIFIKASI *BACKPROPAGATION*
TERHADAP DATA *TELCO CUSTOMER CHURN* DENGAN
SELEKSI FITUR *FORWARD SELECTION***

SKRIPSI

DEBY RIANASARI

1810511024

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN

JAKARTA

2022



**PENERAPAN KLASIFIKASI *BACKPROPAGATION*
TERHADAP DATA *TELCO CUSTOMER CHURN* DENGAN
SELEKSI FITUR *FORWARD SELECTION***

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar
Sarjana Komputer**

DEBY RIANASARI

1810511024

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN

JAKARTA

2022

PERNYATAAN ORISINALITAS

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

Nama : Deby Rianasari

NIM : 1810511024

Tanggal : 8 Juni 2022

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 8 Juni 2022

Yang Menyatakan,



(Deby Rianasari)

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademika Universitas Pembangunan Nasional Veteran Jakarta, saya bertanda tangan di bawah ini:

Nama : Deby Rianasari
NIM : 1810511024
Fakultas : Ilmu Komputer
Program Studi : S1 Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

PENERAPAN KLASIFIKASI *BACKPROPAGATION* TERHADAP DATA *TELCO CUSTOMER CHURN* DENGAN SELEKSI FITUR *FORWARD SELECTION*

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 27 Juli 2022

Yang Menyatakan,



(Deby Rianasari)

LEMBAR PENGESAHAN

Skripsi diajukan oleh :

Nama : Deby Rianasari
NIM : 1810511024
Program Studi : SI Informatika
Judul Tugas Akhir : Penerapan Klasifikasi *Backpropagation* Terhadap Data
Telco Customer Churn Dengan Seleksi Fitur *Forward Selection*

Telah diuji dan dipertahankan di depan Tim Penguji pada Juli 2022 dan telah dinyatakan LULUS sebagai salah satu persyaratan untuk memperoleh gelar Sarjana Komputer (S. Kom.) pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Yuni Widlastiwi, S.Kom., M.Si.

Penguji I



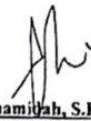
Mayanda Mega Santoni, S.Kom., M.Kom.

Penguji II



Dr. Djidit Widlyanto, S.Kom., M.Si.

Pembimbing I



Nurul Chamidah, S.Kom., M.Kom.

Pembimbing II




Dr. Ermatita, M.Kom.

Dekan



Desta Sandya Prasvita, S.Kom., M.Kom.

Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 13 Juli 2022



Dipindai dengan CamScanner

Dipindai dengan CamScanner

ABSTRAK

Penelitian ini di latar belakang oleh dampak pandemi *Covid 19* yang mendorong orang untuk melakukan kegiatan dari rumah atau *Work From Home* (WFH). Sehingga pemakaian internet yang semakin intensif akan mengakibatkan dinamisnya *customer churn* untuk berpindah layanan dari operator penyedia jasa sambungan internet satu ke lainnya yang lebih menguntungkan. Tujuan dari penelitian ini adalah melakukan klasifikasi terhadap data *telco customer churn*. Dataset yang digunakan berasal dari *Kaggle.com*. Penelitian ini menggunakan algoritma *backpropagation* untuk klasifikasi dan *confusion matrix* untuk evaluasi. Pada penelitian ini juga menerapkan *forward selection* untuk mendapatkan fitur yang relevan. Pada penelitian ini dihasilkan akurasi sebesar 78.79%, presisi sebesar 82.51%, dan *recall* sebesar 81.10% dari 15 fitur yang relevan.

Kata Kunci : *Backpropagation, Confusion Matrix, Forward Selection, Telco Customer Churn*

ABSTRACT

This research is motivated by the impact of the Covid 19 pandemic which encourages people to carry out activities from home or Work From Home (WFH). So that increasingly intensive internet usage will result in dynamic customer churn to switch services from one internet connection service provider to another that is more profitable. The purpose of this research is to classify the telco customer churn data. The dataset used is from Kaggle.com. This research uses backpropagation algorithm for classification and confusion matrix for evaluation. This research also applies forward selection to get the relevant features. In this study, an accuracy of 78.79%, precision of 82.51%, and recall of 81.10% of 15 relevant features resulted.

Keywords: *Backpropagation, Confusion Matrix, Forward Selection, Telco Customer Churn*

KATA PENGANTAR

Puji syukur dan ucapan terima kasih yang tidak henti - hentinya penulis panjatkan kepada kehadiran Allah SWT yang mana telah memberikan kesehatan, kekuatan dan segala karunia-Nya kepada penulis sehingga skripsi ini berhasil diselesaikan. Tidak lupa juga penulis mengucapkan shalawat dan salam kepada junjungan Nabi Besar Muhammad SAW. Skripsi ini disusun oleh penulis guna memenuhi salah satu syarat untuk menyelesaikan pendidikan Sarjana Komputer pada Program Studi S1 Informatika di Fakultas Ilmu Komputer (FIK) Universitas *Pembangunan Nasional Veteran Jakarta*. *Judul yang dipilih pada penelitian ini adalah Penerapan Klasifikasi Backpropagation Terhadap Data Telco Customer Churn*. Penulis berharap, dengan adanya skripsi ini dapat menambah referensi para pembaca. Penyelesaian tulisan ini tidak terlepas bantuan dari berbagai pihak, terutama dan teristimewa penulis ucapkan Terima kasih juga kepada:

1. Ayah dan Ibu penulis, serta seluruh keluarga yang selalu memberikan penulis dukungan, semangat dan doa hingga penulis dapat menyelesaikan skripsi ini.
2. Pak Dr. Didit Widiyanto, S.Kom., M.Si. dan Bu Nurul Chamidah, S.Kom., M.Kom. selaku dosen pembimbing satu dan pembimbing dua penulis yang telah memberikan saran dan arahan dalam proses penyusunan skripsi penulis.
3. Bapak/Ibu dosen Universitas Pembangunan Nasional Veteran Jakarta, terkhusus Fakultas Ilmu Komputer yang telah memberikan ilmu yang bermanfaat sehingga penulis dapat menyelesaikan skripsi ini berkat seluruh ilmu yang penulis dapatkan selama ini.
4. Teman-teman penulis, dan seluruh pihak yang telah berkontribusi baik langsung dan tidak langsung, yang tidak dapat penulis sebutkan satu - persatu dalam menyelesaikan skripsi ini.

Penulis sadar bahwa skripsi ini masih memiliki banyak kekurangan. Semua kesalahan dan kekurangan yang ada dalam skripsi ini berasal dari penulis. Penulis

harap dengan semua kekurangan tersebut, skripsi ini dapat tetap memberikan manfaat.

Jakarta, 8 Juni 2022

Penulis

Deby Rianasari

DAFTAR ISI

PERNYATAAN ORISINALITAS	ii
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS	iii
LEMBAR PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR SIMBOL	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Ruang Lingkup	2
1.4 Tujuan	2
1.5 Manfaat	2
1.6 Luaran yang diharapkan	3
1.7 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Churn	5
2.2 Machine Learning	5
2.3 Artificial Neural Network	6
2.4 Normalisasi Min-Max	6
2.5 Synthetic Minority Oversampling Technique	7
2.6 Forward Selection	8
2.7 Backpropagation	9
2.4.1 Fungsi Aktivasi Sigmoid	12
2.8 Confusion Matrix	12
2.9 Penelitian Terkait	13
2.9.1 Penelitian Backpropagation	13
2.9.2 Permasalahan Detail	14

BAB III METODOLOGI PENELITIAN	16
3.1 Kerangka Pikir	16
3.1.1 Identifikasi Masalah	17
3.1.2 Studi Literatur.....	17
3.1.3 Pengumpulan Data.....	17
3.1.4 <i>Exploratory Data Analysis</i>	18
3.1.5 <i>Preprocessing Data</i>	18
3.1.6 Seleksi Fitur	18
3.1.7 <i>Klasifikasi Backpropagation</i>	19
3.1.8 Evaluasi.....	20
3.2 Perangkat Penelitian.....	20
3.3 Jadwal Penelitian	21
BAB IV PEMBAHASAN	22
4.1 Data	22
4.2 <i>Exploratory Data Analysis</i>	24
4.2.1 <i>Descriptive Statistics</i>	24
4.2.2 <i>Univariate Analysis</i>	25
4.2.3 <i>Multivariate Analysis</i>	26
4.3 <i>Preprocessing Data</i>	26
4.3.1 Mengecek Dataset	27
4.3.2 Mengubah Tipe Data	27
4.3.3 Memisahkan Data Kategorik dan Data Numerik	28
4.3.4 Mengecek <i>Missing Value</i>	28
4.3.5 Menangani <i>Missing Value</i>	29
4.3.6 Mengganti Nilai “ <i>No Internet Service</i> ” dan “ <i>No Phone Service</i> ” menjadi “ <i>No</i> ”	30
4.3.7 Mengubah Nilai Kategorik menjadi Numerik	31
4.3.8 Melakukan Normalisasi Data	31
4.3.9 Menangani Ketidakseimbangan Data	32
4.4 Seleksi Fitur	33
4.5 Pembagian Data	33
4.6 <i>Klasifikasi Backpropagation</i>	34
4.6.1 Menentukan Parameter Jaringan	34
4.6.2 Parameter Jaringan Terbaik	38
4.7 Evaluasi.....	39

BAB V PENUTUP	41
5.1 Simpulan	41
5.2 Saran	41
DAFTAR PUSTAKA	42
RIWAYAT HIDUP	45

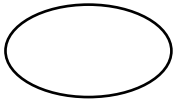

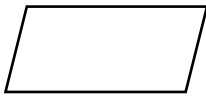



DAFTAR TABEL

Tabel 2.1 Confusion Matrix.....	12
Tabel 2.2 Resume Literatur Review.....	14
Tabel 3.1 Perangkat Penelitian.....	20
Tabel 3.2 Jadwal Penelitian.....	21
Tabel 4.1 Deskripsi Variabel.....	22
Tabel 4.2 Dataset Telco Customer Churn.....	23
Tabel 4.3 Ringkasan Statistika Pada Data Numerik.....	24
Tabel 4.4 Ringkasan Statistika Pada Data Kategorik.....	24
Tabel 4.5 Informasi Data.....	27
Tabel 4.6 Perubahan Tipe Data Pada Variabel TotalCharges.....	28
Tabel 4.7 Hasil Pemisahan Data Kategorik dan Data Numerik.....	28
Tabel 4.8 Jumlah Missing Value.....	28
Tabel 4.9 Sampel Data yang Terdapat Missing Value.....	29
Tabel 4.10 Sampel Data yang Sudah Diganti Dengan Nilai Mean.....	30
Tabel 4.11 Label Encode.....	31
Tabel 4.12 Inisialisasi Parameter.....	34
Tabel 4.13 Inisialisasi Parameter Konstan.....	35
Tabel 4.14 Perbandingan Hasil Variasi Hidden Neuron.....	36
Tabel 4.15 Perbandingan Hasil Variasi Learning Rate.....	37
Tabel 4.16 Perbandingan Hasil Variasi Epoch.....	38
Tabel 4.17 Inisialisasi Parameter Terbaik.....	38
Tabel 4.18 Hasil Pengujian.....	40

DAFTAR GAMBAR

Gambar 2.1 Arsitektur Jaringan Syaraf Tiruan.....	6
Gambar 2.2 Flowchart Forward Selection.....	8
Gambar 2.3 Fungsi Sigmoid.....	12
Gambar 3.1 Flowchart Metodologi Penelitian.....	16
Gambar 4.1 Visualisasi Data Numerik menggunakan Box Plot.....	25
Gambar 4.2 Visualisasi Data Kategorik menggunakan Count Plot.....	25
Gambar 4.3 Visualisasi Korelasi antar Variabel Pada Data Numerik menggunakan Heatmap.....	26
Gambar 4.4 Visualisasi Pada Variabel <i>OnlineSecurity</i> , <i>OnlineBackup</i> , <i>DeviceProtection</i> , <i>TechSupport</i> , <i>SteamingTV</i> , <i>StreamingMovies</i> dan <i>MultipleLines</i>	30
Gambar 4.5 Hasil Normalisasi Data.....	32
Gambar 4.6 Distribusi Class Churn tanpa SMOTE.....	32
Gambar 4.7 Distribusi Class Churn dengan SMOTE.....	32
Gambar 4.8 Plot Peforma Training Pada Parameter Jaringan Terbaik.....	39
Gambar 4.9 Evaluasi menggunakan Confusion Matrix.....	39

DAFTAR SIMBOL

Simbol	Keterangan
	Terminator berfungsi sebagai permulaan / akhir program
	Proses berfungsi sebagai proses perhitungan atau pengolahan data
	Input / Output Data berfungsi sebagai proses input/output data, parameter, informasi.
	Dokumen berfungsi untuk menyatakan dokumen tunggal
	Dokumen berfungsi untuk menyatakan dokumen jamak atau lebih dari satu dokumen
	Garis Penghubung berfungsi untuk menghubungkan simbol