

**KLASIFIKASI KEPUTUSAN PELANGGAN TERHADAP
PRODUK PT TELEKOMUNIKASI MENGGUNAKAN
*SUPERVISED LEARNING***

SKRIPSI

Oleh:

JHON KRISTIAN VIERI

201710225325



PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS BHAYANGKARA JAYA JAKARTA RAYA

2023

LEMBAR PERSETUJUAN PEMBIMBING

Judul Skripsi : Klasifikasi Keputusan Pelanggan Terhadap
Produk PT Telekomunikasi Menggunakan
Supervised Learning

Nama Mahasiswa : Jhon Kristian Vieri

Nomor Pokok Mahasiswa : 201710225325

Program Studi/Fakultas : Informatika / Ilmu Komputer



Pembimbing I

Pembimbing II

Dr. Tb. Ai Munandar.,S.Kom.,M.T

Dwi Budi Srisulistiwati.,S.Kom.,M.M

NIDN: 0413098403

NIDN: 0323057701

LEMBAR PENGESAHAN

Judul Skripsi : Klasifikasi Keputusan Pelanggan Terhadap Produk PT
Telekomunikasi Menggunakan Supervised Learning
Nama Mahasiswa : Jhon Kristian Vieri
Nomor Induk Mahasiswa : 201710225325
Program Studi / Fakultas : Informatika / Ilmu Komputer
Tanggal Lulus Ujian Skripsi : 03 Februari 2023

Bekasi, 14 Februari 2023

MENGESAHKAN,

Ketua Tim Penguji : Wowon Priatna, ST, M.Ti.

NIDN : 0429118007

Penguji II

: Ahmad Fathurrozi, S.E., M.M.S.I.

NIDN : 0327117402

Penguji III

: Dr. Tb. Ai Munandar, S.Kom., MT.

NIDN : 0413098403

MENGETAHUI,

Ketua Program Studi Informatika

Dekan Fakultas Ilmu Komputer

Ahmad Fathurrozi, S.E., M.M.S.I.

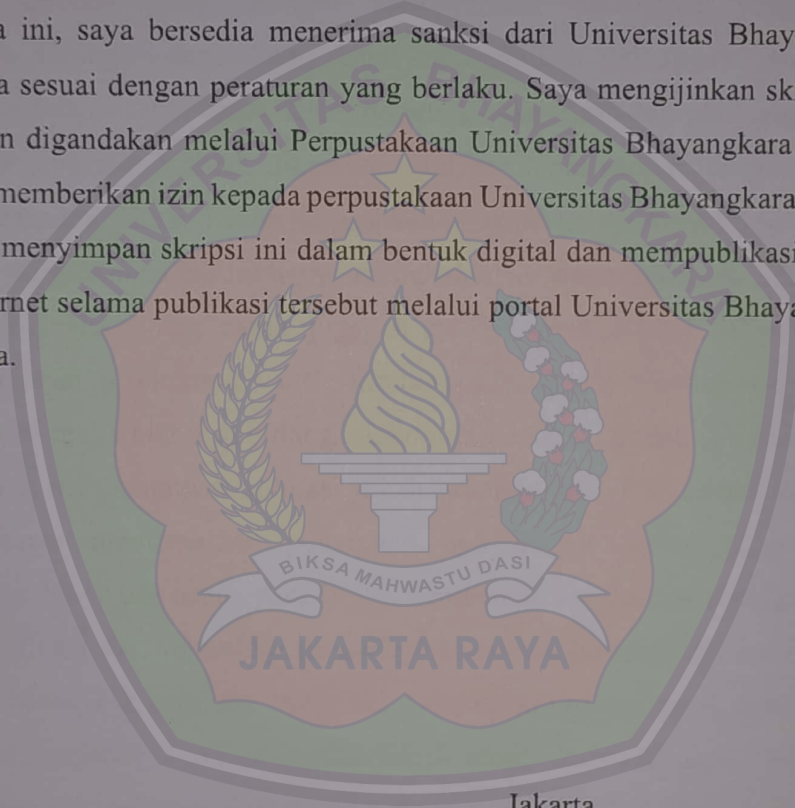
NIP. 2012486

Dr. Dra. Tyastuti Sri Lestari, MM.

NIP. 1408206

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa: Skripsi yang berjudul “Klasifikasi Keputusan Pelanggan Terhadap Produk PT Telekomunikasi Menggunakan Supervised Learning”. Ini adalah benar-benar merupakan hasil karya saya sendiri dan tidak mengandung materi yang ditulis oleh orang lain kecuali pengutipan sebagai referensi yang sumbernya telah di tuliskan secara jelas sesuai dengan kaidah penulisan karya ilmiah. Apabila di kemudian hari di temukan adanya kecurangan dalam karya ini, saya bersedia menerima sanksi dari Universitas Bhayangkara Jakarta Raya sesuai dengan peraturan yang berlaku. Saya mengizinkan skripsi ini dipinjam dan digandakan melalui Perpustakaan Universitas Bhayangkara Jakarta Raya. Saya memberikan izin kepada perpustakaan Universitas Bhayangkara Jakarta Raya untuk menyimpan skripsi ini dalam bentuk digital dan mempublikasikannya melalui Internet selama publikasi tersebut melalui portal Universitas Bhayangkara Jakarta Raya.



Jakarta,

Yang Membuat Pernyataan,


John Kristian Vieri

201710225325

ABSTRAK

John Kristian Vieri, 201710225325. Pelanggan adalah tujuan utama dari semua bidang usaha, tanpa adanya pelanggan perusahaan tidak akan mampu melanjutkan atau bersaing dalam bidang bisnis yang dijalani nya meskipun perusahaan tersebut memiliki produk yang cemerlang jika tidak memiliki peningkatan jumlah pelanggan usaha tersebut tidak akan dapat berkembang atau bahkan dapat mengalami kebangkrutan. Maka dari itu dibutuhkan observasi dan pembuatan aplikasi yang mampu memprediksi pelanggan yang akan berlangganan sehingga perusahaan dapat memprediksi pelanggan yang akan berlangganan dengan tepat tanpa harus menunggu kepastian dari pelanggan yang kemungkinan nya masih tidak diketahui. Hal ini dapat sangat berguna untuk perusahaan manapun dikarenakan perusahaan tidak perlu lagi mencari pelanggan secara acak yang dimana hal itu hanya menghabiskan waktu untuk mendapatkan pelanggan. Masalah yang sama juga di hadapi oleh PT. Telekomunikasi Indonesia dengan produknya yaitu (Indihome) yang sedang berjuang untuk bersaing dalam dunia bisnis di bidang telekomunikasi dan internet. Maka dari itu penelitian dan pembangunan aplikasi ini di lakukan agar PT. Telekomunikasi indonesia dapat mendapatkan pelanggan nya dengan cepat tanpa harus menghabiskan biaya dan usaha yang besar. Pembuatan aplikasi ini menggunakan metode klasifikasi dari teknologi pembelajaran mesin berdasarkan data histori pelanggan. Metode klasifikasi memiliki banyak algoritma yang kuat untuk memprediksi variabel yang memiliki label lebih daripada satu. Beberapa algoritma yang digunakan adalah *Logistic Regression*, *Random Forest Classifier*, *Support Vector Machine* dan *Decision Tree* Yang di sediakan oleh modul pada bahasa pemograman *python* yaitu *Sk-Learn*. Keempat algoritma tersebut akan di uji dengan data yang diseimbangkan menggunakan metode *Oversampling* dari algoritma *Smote* untuk mendapatkan hasil yang optimal dalam memprediksi pelanggan secara otomatis.

Kata Kunci : Klasifikasi, *Logistic Regression*, *Random Forest*, *Support Vector Machine*, *Decision Tree Classifier*, *Smote*, *Sk-Learn*.

ABSTRACT

John Kristian Vieri, 201710225325. *Customers are the main goal of all business fields, without customers the company will not be able to continue or compete in the business field it is in, even though the company has brilliant products. Therefore, it is necessary to make observations and make applications that are able to predict customers who will subscribe so that companies can predict customers who will subscribe correctly without having to wait for confirmation from customers whose possibilities are still unknown. This can be very useful for any company because companies no longer need to look for random customers where it only takes time to find customers. PT. Telekomunikasi Indonesia with its product (Indihome), which is struggling to compete in the business world in the telecommunications and internet sector. Therefore research and development of this application are carried out so that PT. Indonesian telecommunications can get its customers quickly without having to spend a lot of money and effort. Making this application uses a Classification method from Machine Learning technology based on customer historical data. The Classification method has many strong algorithms to predict variables with more than 1 label. Some of the algorithms used are Logistic Regression, Random Forest Classifier, Support Vector Machine and Decision Tree. The four algorithms will be tested with balanced data to get optimal results in predicting customers automatically.*

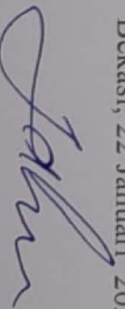
Key Word : Classification, Logistic Regression, Random Forest Classifier, Support Vector Machine, Decision Tree Classifier, Smote, Sk-Learn.

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah melimpahkan karunia-Nya sehingga penulis dapat melaksanakan penelitian dan dapat menyelesaikan proposal skripsi dengan lancar yang berjudul “Klasifikasi Keputusan Pelanggan Terhadap Produk PT Telekomunikasi Menggunakan Supervised Learning”. Dalam pelaksanaan penelitian dan penyusunan skripsi, penulis mendapat banyak sekali bantuan, dukungan dan bimbingan dari berbagai pihak. Oleh karena itu dalam kesempatan ini saya ingin menyampaikan terimakasih kepada:

1. Kepada Tuhan Yang Maha Esa, yang sudah memberikan perlindungan serta pertolongan dalam menjalani penelitian dan penyusunan skripsi.
2. Bapak Dr. Drs. H. Bambang Karsono, S.H.,M.M selaku Rektor Universitas Bhayangkara Jakarta Raya.
3. Ibu Dr. Dra. Tyastuti Sri Lestari., S.Si., M.M selaku Dekan fakultas ilmu komputer Universitas Bhayangkara Jakarta Raya.
4. Bapak Ahmad Fathurrozi., S.Kom., M.M.S.I selaku ketua program studi informatika.
5. Bapak Ahmad Noeman, S.kom ., M.Kom selaku dosen pembimbing akademik
6. Bapak Dr.Tb.Ai Munandar., S.Kom., M.T selaku dosen pembimbing pertama dalam penelitian ini yang selalu mengarahkan dan membimbing dengan sepenuh hati.
7. Ibu Dwi Budi Srisulitiowati,S.Kom.,M.M selaku dosen pembimbing kedua dalam penelitian ini yang selalu mengarahkan dan membimbing sepenuh hati.

Bekasi, 22 Januari 2023



Jhon Kristian Vieri

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR.....	x
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	2
1.3 Rumusan Masalah	3
1.4 Batasan Masalah.....	3
1.5 Tujuan Penelitian	3
1.6 Manfaat Penelitian	3
1.7 Tempat dan waktu pelaksanaan	4
BAB II LANDASAN TEORI	5
2.1 Tinjauan Pustaka	5
2.2 <i>Machine Learning</i>	7
2.3 Analisis Data	9
2.4 <i>Knowledge Discovery in Databases (KDD)</i>	10
2.5 <i>Decision Tree</i>	11
2.6 <i>Random Forest</i>	12
2.7 Statistika Deskriptif.....	13
2.8 <i>Scikit-Learn</i>	13

2.9 <i>Business Intelligence</i>	14
2.10 Smote	14
2.11 <i>Confusion Matrix</i>	16
2.12 <i>Logistic Regression</i>	19
2.13 <i>Support Vector Machine</i>	20
2.12 <i>Flow Chart</i>	21
BAB III METODOLOGI PENELITIAN	23
3.1 Desain Penelitian.....	23
3.1.1 Alur Penelitian	23
3.1.2 Alat Penelitian.....	24
3.1.3 Persiapan Data.....	24
3.1.5 Analisis dan Visualisasi Data.....	24
3.1.6 <i>Preprocessing</i>	24
3.1.7 Normalisasi Data.....	24
3.1.8 Modelling	24
3.1.9 Evaluasi	24
3.1.10 Penetapan	24
3.1.11 <i>Deployment</i>	25
3.2 Pengumpulan data	25
3.2.1 Observasi.....	25
3.2.2 Wawancara.....	25
BAB IV HASIL DAN PEMBAHASAN	26
4.1 Hasil Penelitian	26
4.1.1 Persiapan Data.....	26
4.1.2 Analisis dan Visualisasi Data.....	28
4.1.3 <i>Preprocessing</i>	37

4.1.4 Transformasi Data.....	40
4.1.5 Modelling	42
4.1.6 <i>Deployment</i>	45
4.2 Pembahasan.....	46
4.2.1 Seleksi Fitur	47
4.2.1 Evaluasi pada model <i>Logistic Regression</i>	48
4.2.2 Evaluasi pada model <i>Support Vector Machine</i>	48
4.2.3 Evaluasi pada model <i>Decision Tree</i>	49
4.2.4 Evaluasi pada model <i>Random Forest</i>	50
4.2.5 Penetapan Model.....	50
BAB V PENUTUP	52
5.1 Kesimpulan	52
5.2 Saran.....	52
DAFTAR PUSTAKA	54

DAFTAR GAMBAR

Gambar 1. 1 Pengertian Machine Learning, Sumber: Dokumentasi Pribadi Diolah dari berbagai sumber	7
Gambar 2. 2 Contoh SVM dalam Menentukan Hyperplane, Sumber: https://medium.com/@samsudiney/penjelasan-sederhana-tentang-apa-itu-svm-149fec72bd02	20
Gambar 3. 1 Desain Penelitian, Sumber: Dokumentasi Pribadi	23
Gambar 4. 1 Distribusi Variabel Numerikal, Sumber: Hasil Penelitian	28
Gambar 4. 2 Visualisasi Korelasi Numerikal, Sumber: Hasil Penelitian.....	29
Gambar 4. 3 Visualisasi Variabel Target, Sumber: Hasil Penelitian	30
Gambar 4. 4 Visualisasi Variabel Bantuan Teknisi, Sumber: Hasil Penelitian	30
Gambar 4. 5 Visualisasi Variabel Backup, Sumber: Hasil Penelitian	31
Gambar 4. 6 Visualisasi Variabel Keamanan Jaringan, Sumber: Hasil Penelitian	31
Gambar 4. 7 Visualisasi Variabel Layanan Internet, Sumber: Hasil Penelitian ...	32
Gambar 4. 8 Visualisasi Variabel Layanan Telepon, Sumber: Hasil Penelitian...	32
Gambar 4. 9 Visualisasi Variabel Paket Langganan, Sumber: Hasil Penelitian.....	33
Gambar 4. 10 Visualisasi Variabel Partner, Sumber: Hasil Penelitian	33
Gambar 4. 11 Visualisasi Variabel Tagihan Tanpa Surat, Sumber: Hasil Penelitian	34
Gambar 4. 12 Visualisasi Variabel Tanggungan, Sumber: Hasil Penelitian.....	34
Gambar 4. 13 Visualisasi Variabel Jenis Kelamin, Sumber: Hasil Penelitian.....	35
Gambar 4. 14 Visualisasi Variabel Kontrak, Sumber: Hasil Penelitian	35
Gambar 4. 15 Visualisasi Variabel Metode Pembayaran, Sumber: Hasil Penelitian	36
Gambar 4. 16 Visualisasi Variabel Penduduk Lama, Sumber: Hasil Penelitian ..	36
Gambar 4. 17 Visualisasi Variabel Streaming Film, Sumber: Hasil penelitian....	37
Gambar 4. 18 Visualisasi Variabel Streaming TV, Sumber: Hasil Penelitian.....	37
Gambar 4. 19 Hasil Identifikasi Nilai Ekstrim, Sumber: Hasil Penelitian.....	38

Gambar 4. 20 Penulisan Kode dan Hasil Identifikasi Missing Value, Sumber Hasil Penelitian.....	39
Gambar 4. 21 Visualisasi hasil Oversampling Menggunakan SMOTE, Sumber: Hasil Penelitian	40
Gambar 4. 22 Matriks Hasil Encoding Menggunakan OrdinalEncoder, Sumber: Hasil Penelitian	41
Gambar 4. 23 Matriks Hasil Scalling Menggunakan MinMaxScaller, Sumber: Hasil Penelitian.....	41
Gambar 4. 24 Hasil Klasifikasi Logistic Regression, Sumber: Hasil Penelitian ..	42
Gambar 4. 25 Hasil Klasifikasi Support Vector Machine, Sumber: Hasil Penelitian	43
Gambar 4. 26 Hasil Klasifikasi Decision Tree, Sumber: Hasil Penelitian.....	44
Gambar 4. 27 Hasil Klasifikasi Random Forest, Sumber: Hasil Penelitian	44
Gambar 4. 28 Halaman Atas Aplikasi, Sumber: Hasil Penelitian.....	45
Gambar 4. 29 Halaman Tengah Aplikasi, Sumber: Hasil Penelitian.....	46
Gambar 4. 30 Halaman Bawah Aplikasi, Sumber: Hasil Penelitian.....	46
Gambar 4. 31 Hasil Uji Korelasi Terhadap Variabel Kategorikal	47
Gambar 4. 32 Hasil Uji Korelasi Terhadap Variabel Numerik.....	47
Gambar 4. 33 Hasil Confusion Matrix Logistic Regression, Sumber: Hasil Penelitian.....	48
Gambar 4. 34 Hasil Confusion Matrix SVM, Sumber: Hasil Penelitian	49
Gambar 4. 35 Hasil Confusion Matrix Decision Tree, Sumber: Hasil Penelitian	49
Gambar 4. 36 Hasil Confusion Matrix Random Forest, Sumber: Hasil Penelitian	50

DAFTAR TABEL

Tabel 2. 1 Tinjauan Pustaka	5
Tabel 2. 2 Tabel Contoh Confusion Matrix	17
Tabel 2. 3 Tabel Flow Chart	21
Tabel 4. 1 Keterangan Data.....	26
Tabel 4. 2 Tabel Nilai Evaluasi Model	51



LAMPIRAN

Lampiran 1.1 Biodata Penulis.....	56
Lampiran 1.2 Kartu Konsultasi.....	57
Lampiran 1.3 Pendaftaran Ujian Skripsi.....	60
Lampiran 1.4 Rekomendasi Ujian Sidang.....	61
Lampiran 1.5 Persentasi Plagiarisme.....	62
Lampiran 1.6 Prototype.....	80

