



**SISTEM PREDIKSI *LEAD CONVERSION* MENGGUNAKAN
MACHINE LEARNING MODEL PADA PT. XYZ**

SKRIPSI

RADYAN NUGROHO

1910512046

UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”

JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI SISTEM INFORMASI

2023



**UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”
JAKARTA**

**SISTEM PREDIKSI *LEAD CONVERSION* MENGGUNAKAN
MACHINE LEARNING MODEL PADA PT. XYZ**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana
komputer**

RADYAN NUGROHO

1910512046

**UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”
JAKARTA**

**FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
2023**

LEMBAR PERSETUJUAN SIDANG

LEMBAR PERSETUJUAN SIDANG

Dengan ini dinyatakan bahwa Tugas Akhir berikut:

Nama : Radyan Nugroho
NIM : 1910512046
Program Studi : S1 Sistem Informasi
Judul : Sistem Prediksi *Lead Conversion* Menggunakan *Machine Learning Model* pada PT. XYZ

Telah disetujui untuk diujikan oleh Tim Penguji pada ujian sidang Tugas Akhir sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S-1 Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta.

Menyetujui,



Dr. Bambang Saras Yulistiawan, S.T., M.Kom.
Ketua Program Studi

Mengetahui,



Helena Nurramdhani Irmada, S.Pd., M.Kom.
Ketua Program Studi

Ditetapkan di : Jakarta
Tanggal Persetujuan : 5 Juni 2023

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

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

Nama : Radyan Nugroho

NIM : 1910512046

Program Studi : S1 Sistem Informasi

Judul : Sistem Prediksi *Lead Conversion* Menggunakan *Machine Learning Model* pada PT. XYZ

Tanggal : 5 Juni 2023

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

Jakarta, 5 Juni 2023

Yang menyatakan,

A handwritten signature in black ink is written over a yellow and black 1000 Rupiah stamp. The stamp features the Garuda Pancasila emblem and the text 'METERAI TEMPEL' and 'D2BAKX435094920'.

(Radyan Nugroho)

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Radyan Nugroho
NIM : 1910512046
Fakultas : Ilmu Komputer
Program Studi : S1 Sistem Informasi

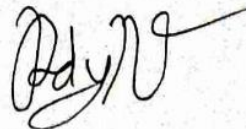
Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non eksklusif (*Non – exclusive Royalty Free Right*) atas skripsi saya yang berjudul:

Sistem Prediksi *Lead Conversion* Menggunakan *Machine Learning Model* pada PT. XYZ

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

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 5 Juni 2023
Yang menyatakan,



(Radyan Nugroho)

LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa Skripsi berikut:

Nama : Radyan Nugroho

NIM : 1910512046


Program Studi : S1 - Sistem Informasi

Judul : Sistem Prediksi *Lead Conversion* Menggunakan Model *Machine Learning* Pada PT. XYZ

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta.


(Dr. Widva Cholil, S.Kom., M.I.T)

Penguji I


(Helena Nurramdhani Irmanda, S.Pd.,

M.Kom.)

Penguji II


(Dr. Bambang Saras Yulistiawan, S.T.,

M.Kom.)

Pembimbing




(Helena Nurramdhani Irmanda, S.Pd., M.Kom.)

Kepala Program Studi

Ditetapkan di : Jakarta

Tanggal Persetujuan : 27 Juni 2023



SISTEM PREDIKSI *LEAD CONVERSION* MENGGUNAKAN *MACHINE LEARNING MODEL* PADA PT. XYZ

Radyan Nugroho

ABSTRAK

Tantangan dari waktu ke waktu bagi para pelaku bisnis salah satunya adalah menyaring pelanggan yang berpotensi dari pasar mereka. Namun yang menjadi masalah adalah saat ini tim *marketing* diterjang banyak ketidakpastian kondisi pasar. Sampai saat ini tercatat banyak pengguna aplikasi yang masih belum bertransaksi sama sekali. Penelitian ini dilakukan untuk membantu tim *marketing* bagaimana melihat potensi calon pelanggannya berdasarkan data yang dimiliki saat ini dari aspek demografis dan aktivitas terbaru dari pengguna aplikasi. Penerapan algoritma *machine learning model* pada penelitian ini akan digunakan untuk membantu perusahaan dalam pengklasifikasian pelanggan yang berpotensi menggunakan data perusahaan yang tersedia. Algoritma *machine learning model* yang digunakan pada penelitian ini adalah *tree based learning* yang di antaranya peneliti mengusung tiga *classifier* yaitu *decision tree*, *random forest* dan *gradient boosting*. Sebelum melakukan perbandingan, data akan melewati proses pemodelan dengan diawali dengan data *preprocessing* hingga di akhir melakukan pengukuran performa algoritma model menggunakan *confusion matrix*. Terdapat 9,976 baris data yang telah bersih dan siap dianalisis, kemudian dipisah menjadi dua bagian menjadi data *training* dan *test* dengan rasio 8:2. Hasil evaluasi dari ke tiga model ditemukan pada algoritma *random forest* mampu mendapatkan skor akurasi sebesar 76%, *recall* 76%, presisi 78%, F-1 skor 74% dan nilai AUC sebesar 0.836. Algoritma *random forest* kemudian diimplementasikan ke dalam sistem berbasis *website* untuk tahap *deployment* yang diharapkan dapat membantu tim *marketing* dalam melihat ciri-ciri pengguna aplikasi yang berpotensi untuk konversi menjadi pelanggan.

Kata Kunci: *Lead Conversion, Sistem Prediksi, Predictive Analytics, Machine Learning Algorithm, Data Mining*

SISTEM PREDIKSI *LEAD CONVERSION* MENGGUNAKAN *MACHINE LEARNING MODEL* PADA PT. XYZ

Radyan Nugroho

ABSTRACT

The challenge over time for business practitioners is to filter potential customers from their market. However, the problem lies in the current uncertainties faced by marketing teams in market conditions. Until now, there have been many application users who have not made any transactions at all. This research is conducted to assist the marketing team in identifying the potential customer base based on the available data, specifically focusing on demographic aspects and recent activities of application users. The application of machine learning algorithms in this research aims to help the company classify potential customers using the available company data. The tree-based learning algorithms, namely decision tree, random forest, and gradient boosting, are employed as classifiers in this study. Prior to the comparison, the data undergoes modeling processes, starting from data preprocessing, and culminating in the measurement of algorithm performance using a confusion matrix. A total of 9,976 clean and analyzable data rows are obtained, which are then divided into two parts for training and testing, with an 8:2 ratio. The evaluation results of the three models reveal that the random forest algorithm achieves an accuracy score of 76%, recall of 76%, precision of 78%, F-1 score of 74%, and an AUC value of 0.836. Subsequently, the random forest algorithm is implemented into a website-based system for deployment, with the aim of assisting the marketing team in identifying characteristics of application users who have potential for conversion into customers.

Keywords: *Lead Conversion, Prediction System, Predictive Analytics, Machine Learning Algorithm, Data Mining*

KATA PENGANTAR

Puji dan syukur kita panjatkan kehadirat Allah SWT Tuhan Yang Maha Esa ,atas berkat dan karunia-Nya peneliti dapat menyelesaikan Tugas Akhir (Skripsi) dengan baik. Tugas Akhir ini dilakukan sebagai bentuk ikhtiar dalam proses penyelesaian perkuliahan dan juga karena merupakan prasyarat untuk lulus.

Dalam penyelesaian proposal ini, peneliti tidak berdiri sendiri , karena pembuatan proposal ini tidak lepas dari bantuan banyak pihak yang telah memberikan bantuan dan dukungan. Untuk itu peneliti mengucapkan banyak terima kasih kepada:

1. Allah SWT.
2. Kedua orang tua yang telah memberikan dukungan baik secara moral maupun materil.
3. Dr. Ermatita, M.Kom., selaku dekan Fakultas Ilmu Komputer.
4. Ibu Helena Nurramadhan Irmada, S.Pd., M.Kom selaku Ketua Program Studi Sarjana Jurusan Sistem Informasi.
5. Bapak Dr. Bambang Saras Yulistiawan, S.T., M.Kom selaku Dosen Pembimbing.
6. Tim *marketing*, P&T, dan AI PT. XYZ.
7. Bang Reza Kurniawan selaku *Performance Marketing Lead* di PT. XYZ.
8. Kak Dayu dan Bapak Lucky selaku *Business Intelligence Associate* dan *Business Intelligence Senior Manager* PT. XYZ yang telah membantu dalam pengumpulan data.
9. Teman-teman seperjuangan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jakarta angkatan 2019.
10. Serta semua pihak yang tidak dapat peneliti sebutkan satu persatu tanpa mengurangi rasa hormat.

Disadari bahwa masih banyaknya kekurangan dari Tugas Akhir ini, baik dari materi maupun teknik penelitian, mengingat kurangnya pengetahuan dan pengalaman peneliti. Oleh karena itu, kritik dan saran yang membangun akan sangat berarti bagi peneliti.

Jakarta, 29 Oktober 2022

Penulis

DAFTAR ISI

LEMBAR PERSETUJUAN SIDANG	iii
PERNYATAAN ORISINALITAS	iv
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS	v
Lembar Pengesahan	vi
Abstrak	vii
<i>ABSTRACT</i>	viii
Kata pengantar	ix
DAFTAR ISI	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Ruang Lingkup Penelitian	3
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian	4
1.6. Luaran yang Diharapkan	5
1.7. Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1. <i>Digital Marketing</i>	7
2.1.1. Pengertian Digital Marketing	7
2.1.2. Manfaat Digital Marketing	7
2.2. PT. XYZ	8
2.3. <i>Lead Conversion & Predictive Lead Conversion</i>	8
2.3.1. Pengertian <i>Lead Conversion</i>	8
2.3.2. <i>Predictive Lead Conversion</i>	8
2.4. <i>Data Mining</i>	9
2.4.1. Fungsi <i>Predictive Analytic</i>	9
2.4.2. <i>Data Preprocessing</i>	9
2.4.3. <i>Data Splitting</i>	11
2.5. Pemodelan Data	11
2.5.1. Klasifikasi	11
2.5.2. Algoritma <i>Decision Tree</i>	12

2.5.3.	Algoritma <i>Random Forest</i>	13
2.5.4.	Algoritma <i>Gradient Boost</i>	14
2.5.5.	<i>Imbalance Data</i>	15
2.5.6.	<i>K Fold Cross-Validation</i>	16
2.6.	Evaluasi Model.....	16
2.7.	Pembangunan Sistem Prediksi	19
2.7.1.	Sistem	19
2.7.2.	<i>SDLC Waterfall</i>	19
2.8.	Pengertian UML.....	20
2.9.	<i>Website</i>	21
2.9.1.	<i>Low Fidelity Wireframe</i>	21
2.9.2.	HTML, CSS & JavaScript	22
2.9.3.	<i>Bootstrap Framework</i>	22
2.9.4.	<i>Microframework Flask</i>	23
2.10.	Bahasa Pemrograman <i>Python</i>	23
2.11.	<i>Blackbox Testing</i>	24
2.12.	<i>Review Penelitian Terdahulu</i>	25
BAB III METODOLOGI PENELITIAN		31
3.1.	Alur Penelitian.....	31
3.2.	Identifikasi Masalah	32
3.3.	Studi Literatur	32
3.4.	Deskripsi Dataset.....	32
3.5.	Analisis Kebutuhan Sistem	34
3.6.	Desain Sistem & Data <i>Pre-Processing</i>	34
3.6.1.	<i>Cleansing</i> (Pembersihan).....	35
3.6.2.	<i>Splitting</i> (Pemisahan).....	35
3.6.3.	<i>Transformation</i> (Perubahan).....	35
3.6.4.	<i>Normalization</i> (Normalisasi)	35
3.6.5.	<i>Label Encoding</i> (Pengkodean Label).....	36
3.6.6.	<i>Selection</i> (Seleksi)	36
3.7.	Evaluasi & Pemilihan Model	36
3.8.	Implementasi Pengkodean Sistem.....	37
3.9.	Pengujian Sistem & Evaluasi Sistem	38
3.10.	Alat dan Bahan	38
3.11.	Waktu dan Tempat Penelitian.....	39
3.12.	Jadwal Penelitian	39

BAB IV ANALISIS DAN PEMBAHASAN.....	40
4.1. Persiapan Dataset	40
4.2. Perancangan Sistem Prediksi.....	41
4.3. Desain Sistem & <i>Pre-processing</i> Data.....	41
4.3.1. Desain Sistem	41
4.3.2. Data <i>Cleansing</i> (Pembersihan)	45
4.3.3. Data <i>Split</i> (Pemisahan)	47
4.3.4. Data <i>Transformation</i> (Perubahan)	48
4.3.5. Data <i>Normalization</i> (Normalisasi).....	48
4.3.6. Data <i>Selection</i> (Seleksi).....	50
4.4. Implementasi Model <i>Machine Learning</i>	52
4.5. Evaluasi Model.....	53
4.5.1. <i>Decision Tree</i>	53
4.5.2. <i>Random Forest</i>	55
4.5.3. <i>Gradient Boosting</i>	57
4.5.4. Pemilihan Model.....	59
4.6. Implementasi Sistem Prediksi (Pengkodean).....	60
4.6.1. Menyimpan Model.....	61
4.6.2. Membangun <i>Front-end</i>	61
4.6.3. Membangun <i>Back-end Flask</i>	62
4.6.4. Menjalankan Aplikasi	63
4.7. Evaluasi Sistem Prediksi	67
BAB V PENUTUP	70
5.1. Kesimpulan.....	70
5.2. Saran.....	70
DAFTAR PUSTAKA	72
LAMPIRAN.....	77
Lampiran 1. Surat Pengajuan Izin Penelitian.....	78
Lampiran 2. Bukti Persetujuan Penelitian.....	79
Lampiran 3. Wawancara Analisa Kebutuhan Sistem & Validasi Data	80
Lampiran 4 Dataset Awal.	84
Lampiran 5. Dataset Setelah <i>Preprocessing</i>	85
Lampiran 6. Eksplorasi Data (EDA).....	85
Lampiran 7. Kode <i>routing</i> Aplikasi (app.py).....	90
Lampiran 8. Hasil Turnitin	91

DAFTAR GAMBAR

Gambar 1 Ilustrasi Split Dataset	11
Gambar 2 Ilustrasi Algoritma Gradient Boost	15
Gambar 3 Ilustrasi Teknik Random Undersampling	16
Gambar 4 SDLC Waterfall Model	19
Gambar 5 Alur Penelitian	31
Gambar 6 Informasi Dataset	40
Gambar 7 Nilai 0 pada Fitur Numerik	40
Gambar 8 Use Case Diagram Sistem Prediksi Lead Conversion	42
Gambar 9 Mockup Halaman Home Sistem Prediksi	43
Gambar 10 Mockup Halaman Login Sistem Prediksi	43
Gambar 11 Mockup Halaman About Sistem Prediksi	44
Gambar 12 Mockup Halaman Predict Sistem Prediksi	44
Gambar 13 Mockup Halaman Bulk Predict Sistem Prediksi	45
Gambar 14 Deskripsi Fitur Numerik Sebelum Cleansing	45
Gambar 15 Deskripsi Fitur Numerik Setelah Cleansing	46
Gambar 16 Komposisi Kelas dari Dataset	47
Gambar 17 One-hot encoding pada data training	48
Gambar 18 Normalisasi Kategori Numerik	49
Gambar 19 Jumlah Kolom Data Train Setelah di Transformasi	50
Gambar 20 Confusion Matrix untuk Algoritma Decision Tree	54
Gambar 21 Kurva AUC untuk Algoritma Decision Tree	55
Gambar 22 Confusion Matrix untuk Algoritma Random Forest	56
Gambar 23 Kurva AUC untuk Algoritma Random Forest	57
Gambar 24 Confusion Matrix untuk Algoritma Gradient Boosting	58
Gambar 25 Kurva AUC untuk Algoritma Gradient Boosting	59
Gambar 26 Menjalankan Sistem Prediksi Melalui Terminal	63
Gambar 27 Halaman Utama Sistem Prediksi	64
Gambar 28 Halaman About Sistem Prediksi	64
Gambar 29 Halaman Login Sistem Prediksi	65
Gambar 30 Halaman Utama Setelah Login Sistem Prediksi	65
Gambar 31 Tes Memasukkan Nilai pada Form Prediksi	66
Gambar 32 Tes Output Hasil Sistem Prediksi	66
Gambar 33 Halaman Bulk Predict Sistem Prediksi	67
Gambar 34 Output Hasil Bulk Predict Sistem Prediksi	67

DAFTAR TABEL

Tabel 1 Pengukuran Performa Model	17
Tabel 2 Penelitian Terdahulu	25
Tabel 3 Rasio Pemisahan	47
Tabel 4 Konfigurasi Parameter	52
Tabel 5 Implementasi Training Data kedalam Algoritma ML	53
Tabel 6 Perhitungan Confusion Matrix Pada Algoritma Decision Tree.....	54
Tabel 7 Rata-rata dan Pembobotan Metrik Pada Algoritma Decision Tree	54
Tabel 8 Perhitungan Confusion Matrix Pada Algoritma Random Forest.....	56
Tabel 9 Rata-rata dan Pembobotan Metrik Pada Algoritma Random Forest	56
Tabel 10 Perhitungan Confusion Matrix Pada Algoritma Gradient Boosting	58
Tabel 11 Rata-rata dan Pembobotan Metrik Pada Algoritma Gradient Boosting	58
Tabel 12 Hasil Evaluasi Model Terhadap Data Test	60
Tabel 13 Hasil Evaluasi Model Terhadap Data Training	60
Tabel 14 Blackbox Testing	68