

**KLASIFIKASI CUSTOMER CHURN PADA PT. EDUCATION NUSANTARA
CONNECT MENGGUNAKAN METODE RANDOM
FOREST**

SKRIPSI

DISUSUN OLEH

ANISA JULIA RAHMAYANTI

NPM. 2109010079



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN
2025**

**KLASIFIKASI CUSTOMER CHURN PADA PT. EDUCATION NUSANTARA
CONNECT MENGGUNAKAN METODE RANDOM
FOREST**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
(S.Kom) dalam Program Studi Sistem Informasi pada Fakultas Ilmu Komputer
dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara**

**ANISA JULIA RAHMAYANTI
NPM. ISI 2109010079**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN
2025**

LEMBAR PENGESAHAN

Judul Skripsi : **KLASIFIKASI CUSTOMER CHURN PADA PT.
EDUCATION NUSANTARA CONNECT
MENGUNAKAN METODE RANDOM FOREST**

Nama Mahasiswa : **ANISA JULIA RAHMAYANTI**

NPM : **2109010079**

Program Studi : **SISTEM INFORMASI**

Menyetujui
Komisi Pembimbing



(MHD. BASRI, S.Kom., M.Kom)
NIDN. 0111078802

Ketua Program Studi



(Dr. Firahmi Rizky, S.Kom., M.Kom)
NIDN. 0116079201

Dekan



(Dr. Al-Khwarizmi, S.Kom., M.Kom)
NIDN. 0127099201

PERNYATAAN ORISINALITAS

KLASIFIKASI CUSTOMER CHURN PADA PT. EDUCATION NUSANTARA CONNECT MENGGUNAKAN METODE RANDOM FOREST

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, 16 Juli 2025

Yang membuat pernyataan



Anisa Julia Rahmayanti

NPM. 2109010079

**PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN
AKADEMIS**

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama	: Anisa Julia Rahmayanti
NPM	: 2109010079
Program Studi	: Sistem Informasi
Karya Ilmiah	: Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bebas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:
KLASIFIKASI CUSTOMER CHURN PADA PT. EDUCATION NUSANTARA CONNECT MENGGUNAKAN METODE RANDOM FOREST.

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, 16 Juli 2025

Yang membuat pernyataan



Anisa Julia Rahmayanti

NPM. 2109010079

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Anisa Julia Rahmayanti
Tempat dan Tanggal Lahir : Sukasari, 14 Juli 2003
Alamat Rumah : Sukasari dusun 4b
Telepon/Faks/HP : 081370773075
E-mail : anisajr140703@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD	: SDN 105375 Sukasari	TAMAT: 2015
SMP	: SMPN 2 Sei Rampah	TAMAT: 2018
SMA	: SMAN 1 Pegajahan	TAMAT: 2021

KATA PENGANTAR



Puji syukur saya panjatkan kepada Allah SWT, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan skripsi yang berjudul “ Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakna Metode Random Forest ”. Penulisan skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Program Studi Sistem Informasi Fakultas Ilmu Komputer Teknologi Informasi Universitas Muhammadiyah Sumatera Utara. Penulis tentunya berterima kasih kepada berbagai pihak dalam dukungan serta doa dalam penyelesaian skripsi. Penulis juga mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Agussani, M.AP., Rektor Universitas Muhammadiyah Sumatera Utara (UMSU)
2. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom. Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
3. Bapak Martiano S.Pd, S.Kom., M.Kom Ketua Program Studi Sistem Informasi
4. Ibu Yoshida Sary, SE, S.Kom., M.Kom. Sekretaris Program Studi Sistem Informasi
5. Pembimbing skripsi yaitu bapak MHD. Basri, S.Kom., M.Kom. Terima kasih telah memberikan arahan dan bimbingan dengan penuh kesabaran serta memudahkan penulis selama mengerjakan penelitian, sehingga penulis dapat menyelesaikan penelitian dan sampai ditahap ini.
6. Ibu Yohanni Syahra, S.Si., M.Kom, terima kasih sudah mempermudah penulis dalam menghadapi situasi skripsi ini. Atas kemudahan yang ibu berikan penulis dapat melanjutkan penelitian dan pada akhirnya sudah sampai tahap akhir menyelesaikan skripsi.
7. Cinta pertama penulis yaitu Alm.Bapak Suryono, S.pd. Terima kasih sudah menjadi garda terdepan dan memberikan banyak kebahagiaan, kasih

sayang untuk penulis, serta mengajarkan banyak hal tentang arti kehidupan. Bahagia disana ya pak, lihatlah anakmu sekarang sudah menyelesaikan Pendidikan S1 hasil dari keringat dan perjuanganmu. Setiap detik, rinduku padamu tak pernah pudar.

8. Surga dan pahlawan hidup penulis yaitu Ibu Sri Wulan, S.Pd. Terima kasih atas cinta dan kasih sayang dalam hidup ini. Betapa luar biasanya ibu dalam merawat dan membesarkan penulis sehingga penulis sampai dititik ini. Doa hari-harimu memberikan kemudahan bagi hari-hari penulis, semua bentuk kasih sayang ibu yang membuat penulis berhasil berdiri tegap diatas kaki penulis sendiri. Nasehat dan dukungan dari ibu yang menjadikan penulis kuat. Jadilah wanita yang lebih kuat, karena kekuatanmu adalah sumber energiku.
9. Adik-adik tersayang Hesti Duwi Chayani dan Widiya Tri Agustin yang telah memberikan dukungan dan semangat kepada penulis. Terima kasih sudah mendengarkan keluh kesah penulis dan siap sedia menolong penulis ketika penulis dalam kesulitan. Berkat dukungan kalian yang tak pernah putus, menjadi kekuatan bagi penulis dalam menyelesaikan skripsi ini. Teruslah saling merangkul, kita pasti sukses dan membahagiakan ibu.
10. Alm.Mbah Sukiye terncinta yang sudah membantu merawat dan membesarkan penulis dari bayi hingga saat ini dengan penuh kasih sayang. Terima kasih sudah membantu orang tua penulis disela-sela kesibukan mereka.
11. Penyemangat dalam hidup penulis yaitu Fishan Suryadinata. Terima kasih sudah sabar dalam menghadapi mood penulis. Terima kasih atas semua perhatian yang telah diberikan kepada penulis. Berkat cinta dan kasih sayang darimu penulis selalu semangat dalam menjalani kehidupan dan menyelesaikan skripsi ini. Semoga Allah mempermudah niat baikmu dan mempersatukan kita dalam ikatan yang halal.
12. Teman seperjuangan dalam perkuliahan Nurul Amanda dan Dea Arastika Purba, terimakasih atas segala keluh kesah kita selama dalam perkuliahan semoga kita tetap menjadi teman, sahabat, dan keluarga sampai kapanpun

dan semoga kita menjadi orang yang sukses dengan pekerjaan yang membanggakan.

13. Sabahat until Jannah Elsa Yunanda dan Fatimah Aulia. Terima kasih sudah memberikan waktu dan wadah untuk mendengarkan cerita dan keluh kesah penulis. Teruslah menjadi orang baik, jangan pernah berubah ya.
14. Terima kasih untuk diriku sendiri Anisa Julia Rahmayanti yang sudah berhasil menyelesaikan skripsi ini. Semua ketakutan sudah berhasil terlawan dan terlalui. Harus lebih kuat lagi, wujudkan semua mimpimu karena perjuangan sesungguhnya baru dimulai. Ingat Allah sebagai penolongmu dan jadikan Al-Quran sebagai petunjuk hidupmu.
15. Semua pihak yang terlibat langsung ataupun tidak langsung yang tidak dapat penulis ucapkan satu-persatu yang telah membantu penyelesaian skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna. Oleh karena itu penulis mengharapkan kritik dan saran yang membangun untuk kesempurnaan skripsi ini. Penulis berharap skripsi ini dapat bermanfaat bagi pihak-pihak terkait.

KLASIFIKASI CUSTOMER CHURN PADA PT. EDUCATION NUSANTARA CONNECT MENGGUNAKAN METODE RANDOM FOREST

ABSTRAK

Pertumbuhan pesat pengguna internet di Indonesia mendorong perusahaan penyedia layanan jaringan, seperti PT Education Nusantara Connect (PT ENC), untuk meningkatkan kualitas layanan dan mempertahankan pelanggan. Salah satu tantangan utama yang dihadapi adalah *customer churn*, yaitu fenomena di mana pelanggan berpindah ke layanan lain. Penelitian ini bertujuan untuk menerapkan metode *Random Forest* dalam mengklasifikasikan pelanggan yang berpotensi *churn*. Metode *Random Forest* dipilih karena kemampuannya dalam memberikan akurasi tinggi, menangani data dalam jumlah besar, serta efektif dalam mengelola data yang tidak lengkap. Dalam penelitian ini, model *Random Forest* dibangun menggunakan data pelanggan yang mencakup atribut seperti total transaksi, lama langganan, frekuensi login, dan rating terakhir. Hasil evaluasi model menunjukkan bahwa akurasi yang diperoleh adalah 38.1%, dengan *confusion matrix* yang menunjukkan 4 prediksi benar untuk pelanggan yang tidak *churn* dan 4 prediksi benar untuk pelanggan yang *churn*. Selain itu, analisis fitur menunjukkan bahwa total transaksi memiliki pengaruh terbesar terhadap prediksi *churn* dengan kontribusi sebesar 37.81%. Hasil dari penelitian ini diharapkan dapat membantu PT Education Nusantara Connect dalam menekan tingkat *churn* dan meningkatkan loyalitas pelanggan melalui strategi retensi yang lebih efektif.

Kata Kunci: *Customer churn*, *Random Forest*, klasifikasi, PT Education Nusantara Connect.

CUSTOMER CHURN CLASSIFICATION AT PT. EDUCATION NUSANTARA CONNECT USING THE RANDOM FOREST METHOD

ABSTRACT

The rapid growth of internet users in Indonesia has encouraged network service providers, such as PT Education Nusantara Connect (PT ENC), to improve service quality and retain customers. One of the main challenges faced is customer churn, a phenomenon where customers switch to other services. This study aims to apply the Random Forest method to classify customers who are likely to churn. The Random Forest method was chosen because of its ability to provide high accuracy, handle large amounts of data, and be effective in managing incomplete data. In this study, a Random Forest model was built using customer data that includes attributes such as total transactions, subscription length, login frequency, and last rating. The model evaluation results showed an accuracy of 38.1%, with a confusion matrix showing 4 correct predictions for non-churn customers and 4 correct predictions for churned customers. In addition, feature analysis showed that total transactions had the largest influence on churn prediction, contributing 37.81%. The results of this study are expected to help PT Education Nusantara Connect in reducing churn rates and increasing customer loyalty through more effective retention strategies.

Keywords: Customer churn, Random Forest, classification, PT Education Nusantara Connect.

DAFTAR ISI

LEMBAR PENGESAHAN	i
PERNYATAAN ORISINALITAS.....	ii
PERNYATAAN PERSETUJUAN PUBLIKASI.....	iii
RIWAYAT HIDUP	iv
KATA PENGANTAR.....	v
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
BAB II LANDASAN TEORI	7
2.1 Klasifikasi	7
2.2 Data Mining	7
2.3 Random <i>Forest</i>	10
2.3.1 Pengertian Random <i>Forest</i>	10
2.3.2 Cara Kerja Random <i>Forest</i>	11
2.3.3 Keunggulan Random <i>Forest</i>	11
2.3.4 Kelemahan Random Forest	14
2.3.5 Penerapan Random Forest dalam Customer Churn.....	15
2.4 Customer Churn.....	15
2.5 Prediksi	16
2.6 Algoritma.....	16
2.7 Web.....	17
2.7.1 HTML.....	18
2.7.2 PHP.....	18
2.8 Machine Learning	20
2.8.1 Pengertian Machine Learning.....	20
2.9 Jenis-jenis Machine Learning	21

2.9.1 Supervised Learning (Pembelajaran Terawasi)	21
2.10 UML (Unified Modeling Language)	27
2.11 Use Case Diagram	27
2.12 Flowchart	28
2.13 Activity Diagram	29
2.14 Studi Literatur	31
BAB III METODOLOGI PENELITIAN	33
3.1 Jenis Penelitian	33
3.2 Diagram Alur Penelitian	34
3.3 Rancangan.....	36
3.4 Class Diagram.....	40
3.5 Kerangka Berfikir	42
3.6 Activity Diagram	44
3.7 Rancangan Prediksi Hasil Random Forest	46
3.8 Tempat dan Waktu Penelitian.....	49
BAB IV HASIL DAN UJI COBA	51
4.1. Hasil.....	51
4.2. Pembahasan	56
BAB V PENUTUP.....	73
5.1 Kesimpulan.....	73
5.2. Saran	74
DAFTAR PUSTAKA.....	76

DAFTAR TABEL

Tabel 2.1 Penjelasan Entropy.....	13
Tabel 2.2 Use Case Diagram.....	28
Tabel 2.3 Activity Diagram.....	29
Tabel 2.4 Studi Literatur	31
Tabel 3.1 Rancangan Data Base.....	37
Tabel 3.2 Waktu Penelitian	49
Tabel 4.1. Data Churn Pelanggan	56
Tabel 4.2. Confusion Matrix	62
Tabel 4.3 <i>Blackbox Testing Form</i> Login.....	63
Tabel 4.4 <i>Blackbox Testing Form</i> Home	64
Tabel 4.5. <i>Blackbox Testing Form</i> Pelanggan	65
Tabel 4.6. <i>Blackbox Testing Form</i> Aktivitas Pelanggan.....	66
Tabel 4.7. <i>Blackbox Testing Form</i> Feedback Pelanggan	66
Tabel 4.8. <i>Blackbox Testing Form</i> Transaksi.....	67
Tabel 4.8. <i>Blackbox Testing Form</i> Transaksi.....	68
Tabel 4.9. <i>Blackbox Testing Form</i> Hasil Prediksi Churn.....	68
Tabel 4.10. <i>Blackbox Testing Form</i> Monitoring Churn.....	69

DAFTAR GAMBAR

Gambar 2.1 Proses Data Mining	9
Gambar 2.2 Proses Kerja Random Forest	12
Gambar 2.3 Proses Supervised Learning	22
Gambar 2.4 Proses Unsupervised Learning	24
Gambar 2.5 Proses <i>Reinforcement Learning</i>	26
Gambar 3.1 Diagram Alur Penelitian.....	34
Gambar 3.2 Rancangan Database	36
Gambar 3.3 Use Case Diagram PT Education Nusantara Connect.....	39
Gambar 3.4 Class Diagram	41
Gambar 3.5 Diagram Kerangka Berfikir.....	43
Gambar 3.6 Activity Diagram.....	45
Gambar 3.7 Rancangan Prediksi Hasil Random Forest	46
Gambar 3.8. Tampilan pilihan load data rapid miner	48
Gambar 3.9 Tampilan format coloum data produk	48
Gambar 3.10 Workflow	49
Gambar 4.1. <i>Form</i> Login	51
Gambar 4.2. Form Home	52
Gambar 4.3. Form Pelanggan.....	52
Gambar 4.3. Form Pelanggan.....	53
Gambar 4.5. Form Feedback Pelanggan	53
Gambar 4.6. Form Transaksi.....	54
Gambar 4.7. Form Dataset Model.....	54
Gambar 4.8. Form Hasil Prediksi.....	55
Gambar 4.9. Form Monitoring Hasil Prediksi	55
Gambar 4.10 Hasil Uji Coba Rapid Miner.....	71
Gambar 4.11 Deskripsi Hasil Uji Coba Rapid Miner	71

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Jaringan internet masuk pada berbagai sektor dan digunakan dalam aktifitas berbagai kegiatan, terlebih pada tren otomasi, industri 4.0. Dimana hampir semua aplikasi, mulai dari desktop, website, mobile (android dan ios) yang digunakan dalam berbagai bidang pendidikan, transportasi, bank, logistik, jasa tentunya menggunakan jaringan internet yang kompleksitas nya tinggi perlu untuk dilakukan analisa sehingga dapat diperoleh performa yang tinggi.(Rahman & Nurdin, 2020)

Pada saat ini penggunaan internet di Indonesia berkembang secara luar biasa cepatnya dan internet memegang peran penting di dalam aspek kehidupan manusia. Berdasarkan data dari Asosiasi Penyelenggara Jasa Internet Indonesia pada tahun 2018 pengguna internet mencapai angka 171,17 juta. Penggunaan internet yang menempati urutan pertama paling sering digunakan adalah layanan untuk berkomunikasi lewat pesan yang mencapai 24,7 % dan yang menempatkan urutan kedua yaitu sosial media yang mencapai 18,9 %, media sosial yang sering digunakan yaitu Facebook, Twitter, Instagram dan lain-lain(Ananda & Pristyanto, 2021)

Industri jaringan merupakan salah satu sektor strategis yang menopang perkembangan teknologi informasi dan komunikasi di era digital. Industri ini mencakup berbagai layanan seperti telekomunikasi, penyedia layanan internet (ISP), serta infrastruktur jaringan yang menghubungkan individu, bisnis, dan

sistem secara real-time. Seiring meningkatnya kebutuhan akan konektivitas yang cepat dan stabil, industri jaringan memegang peranan penting dalam mendukung transformasi digital di berbagai sektor, termasuk pendidikan, kesehatan, pemerintahan, dan perdagangan elektronik. Di Indonesia, pertumbuhan pengguna internet yang pesat mendorong perusahaan jaringan untuk terus berinovasi dalam meningkatkan kualitas layanan dan menjangkau wilayah-wilayah terpencil. Namun demikian, industri ini juga menghadapi berbagai tantangan seperti customer churn, keamanan data, dan pemerataan infrastruktur. Oleh karena itu, pemanfaatan teknologi seperti big data, machine learning, dan 5G menjadi kunci untuk mempertahankan daya saing dan menjawab kebutuhan pelanggan yang semakin kompleks. Indonesia merupakan negara yang memiliki jumlah penduduk yang banyak, tidak menutup kemungkinan banyak penduduk Indonesia yang menggunakan internet atau jaringan sebagai alat telekomunikasi. (Fadila et al., n.d.)

Fenomena churn, yaitu perilaku pelanggan yang beralih ke penyedia layanan telekomunikasi lain. Churn mengurangi keuntungan dan pendapatan perusahaan. Mempertahankan pelanggan yang ada menjadi tantangan utama, karena mendapatkan pelanggan baru lebih sulit dan biayanya lebih tinggi. Untuk mengatasi hal ini, perusahaan telekomunikasi perlu memprediksi kapan pelanggan akan churn. Prediksi customer *churn* dapat dilakukan dengan teknik Machine Learning. Dalam *Machine Learning*, diperlukan algoritma yang memiliki kemampuan untuk melakukan klasifikasi terhadap pelanggan apakah akan churn atau tidak. (Amirulhaq Iskandar & Latifa, 2023).

PT Education Nusantara Connect (PT ENC) sebagai salah satu perusahaan penyedia layanan koneksi jaringan internet, menghadapi tantangan besar dalam mempertahankan loyalitas pelanggannya. Persaingan yang ketat dan banyaknya alternatif layanan serupa membuat pelanggan memiliki banyak pilihan, sehingga potensi terjadinya *customer churn* atau kehilangan pelanggan menjadi sangat tinggi.

Perkembangan Teknologi Informasi telah membawa transformasi besar dalam cara perusahaan mengelola dan memanfaatkan data. TI berperan penting dalam menyediakan infrastruktur yang memungkinkan pengumpulan, penyimpanan, dan pemrosesan data dalam skala besar, termasuk data pelanggan yang sangat kompleks dan dinamis. Dalam memanfaatkan data tersebut, pendekatan data mining digunakan untuk menggali informasi yang tersembunyi dan menemukan pola yang tidak terlihat secara kasat mata. Data mining merupakan proses analisis data yang menggabungkan teknik statistik, pembelajaran mesin (*machine learning*), dan algoritma komputasi untuk menghasilkan pengetahuan baru dari data. Salah satu metode populer dalam data mining adalah algoritma Random Forest. Random Forest adalah metode klasifikasi dalam statistika yang berbasis komputasi.

Metode klasifikasi digunakan untuk pembelajaran fungsi-fungsi berbeda yang memetakan masing-masing data terpilih ke dalam salah satu dari kategori kelas yang telah ditetapkan. Memasuki era *big data*, penggunaan metode statistika berbasis komputasi sangat banyak digunakan. Metode Random Forest dipilih karena menghasilkan kesalahan yang lebih rendah, memberikan akurasi yang baik

dalam klasifikasi, dapat menangani data yang jumlahnya sangat besar, dan efektif untuk mengatasi data yang tidak lengkap. (Suci Amaliah et al., 2022)

Dengan dukungan teknologi informasi yang memadai, metode Random Forest dalam data mining dapat diimplementasikan secara efektif untuk memprediksi perilaku pelanggan, seperti kemungkinan terjadinya churn. Penerapan metode ini tidak hanya membantu perusahaan dalam mengidentifikasi pelanggan yang berpotensi churn, tetapi juga dalam menyusun strategi retensi yang lebih tepat sasaran. Melalui penerapan metode *Random Forest*, PT Education Nusantara Connect diharapkan dapat melakukan klasifikasi pelanggan yang berpotensi churn dengan lebih baik, sehingga keputusan bisnis dapat diambil berdasarkan analisis yang berbasis data. Dengan demikian, perusahaan tidak hanya dapat menekan tingkat churn, tetapi juga meningkatkan kepuasan dan loyalitas pelanggan secara berkelanjutan. Berdasarkan uraian latar belakang, maka diusulkan sebuah penelitian dengan judul “ **KLASIFIKASI *CUSTOMER CHURN* PADA PT. EDUCATION NUSANTARA CONNECT MENGGUNAKAN METODE *RANDOM FOREST*”.**

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut :

1. Bagaimana penerapan algoritma Random Forest dapat digunakan untuk mengklasifikasikan pelanggan yang berpotensi churn pada PT Education Nusantara Connect?
2. Apa saja faktor-faktor yang memengaruhi terjadinya customer churn pada PT Education Nusantara Connect?

3. Seberapa akurat metode Random Forest dalam memprediksi customer churn pada PT Education Nusantara Connect?
4. Bagaimana hasil evaluasi model berdasarkan *metrik akurasis, precision, recall*, dan *F1-Score*?

1.3 Batasan Masalah

Agar penelitian ini tetap fokus dan sesuai dengan tujuan yang telah ditetapkan, maka batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini hanya dilakukan pada data pelanggan PT Education Nusantara Connect yang tersedia dalam periode waktu tertentu, yaitu dari bulan Januari 2025 hingga Juni 2025.
2. Penelitian ini hanya membahas proses klasifikasi pelanggan yang berpotensi churn menggunakan data transaksi dibulan Mei 2025.
3. Evaluasi model hanya dilakukan berdasarkan *metrik akurasis, precision, recall*, dan *F1-Score*, tanpa menggunakan metrik lain.
4. Tidak membahas faktor eksternal seperti kondisi pasar, kompetitor, atau perubahan kebijakan perusahaan yang akan mempengaruhi keputusan pelanggan.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk membangun dan mengevaluasi model klasifikasi dalam memprediksi *customer churn* pada PT Education Nusantara Connect dengan menggunakan algoritma *Random Forest*. Adapun tujuan spesifik dari penelitian ini adalah sebagai berikut:

1. Memprediksi customer churn pada PT Education Nusantara Connect menggunakan algoritma Random Forest.

2. Mengetahui faktor apa saja yang dapat mempengaruhi terjadinya customer churn pada PT Education Nusantara Connect.
3. Mengetahui seberapa akurat metode Random Forest dalam memprediksi customer churn pada PT Education Nusantara Connect.
4. Mengetahui hasil evaluasi model berdasarkan *metrik akurasis, precision, recall*, dan *F1-Score*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis maupun praktis kepada Perusahaan maupun peneliti. Adapun manfaat dari penelitian ini sebagai berikut :

1. Memberikan kontribusi terhadap pengembangan ilmu di bidang data mining dan machine learning, khususnya dalam penerapan algoritma Random Forest untuk prediksi customer churn.
2. Memberikan manfaat langsung bagi PT Education Nusantara Connect dalam memprediksi pelanggan yang berisiko churn.
3. Menjadi dasar dalam perancangan strategi retensi pelanggan yang lebih efektif dan terfokus bagi Perusahaan.

BAB II

LANDASAN TEORI

2.1 Klasifikasi

Klasifikasi adalah salah satu teknik dalam data mining dan machine learning yang digunakan untuk memetakan objek atau data ke dalam kategori atau kelas tertentu. Teknik ini termasuk dalam kategori *supervised learning*, di mana model dilatih menggunakan dataset yang sudah diberi label atau kelas yang diketahui. Proses ini bertujuan untuk memprediksi label atau kategori untuk data yang belum diketahui kelasnya berdasarkan pola-pola yang ditemukan pada data pelatihan. klasifikasi merupakan pengelompokan objek kedalam kelas tertentu berdasarkan kelompoknya yang biasanya disebut dengan class. (Sidik et al., 2020).

Klasifikasi dapat diterapkan pada berbagai jenis data dan kasus, seperti prediksi apakah pelanggan akan berhenti menggunakan layanan (*customer churn*), identifikasi spam dalam email, klasifikasi jenis penyakit berdasarkan gejala, dan sebagainya. Algoritma klasifikasi yang umum digunakan termasuk *Decision Tree*, *Random Forest*, *Support Vector Machine (SVM)*, dan *K-Nearest Neighbors (KNN)*. Menurut Han, Kamber, dan Pei (2012), klasifikasi merupakan “proses untuk menemukan model atau fungsi yang dapat menjelaskan dan membedakan kelas-kelas data atau konsep-konsep tertentu.”

2.2 Data Mining

Data Mining merupakan suatu metode yang dijalankan untuk mencari pola, kejadian yang tidak biasa, dan kaitan di dalam kumpulan data yang besar dengan maksud untuk meramalkan hasil di masa yang akan datang. Langkah ini

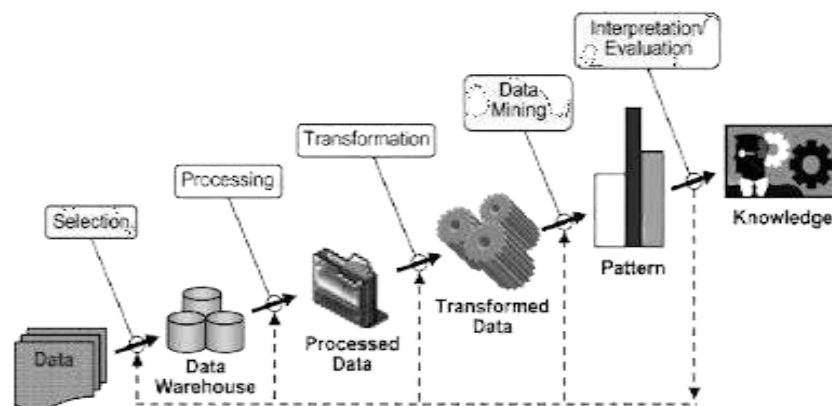
mencakup penggunaan metode analisis statistik, matematika, dan pembelajaran mesin untuk menarik informasi yang bermanfaat dari data yang belum diolah.

Data mining juga merupakan pendekatan untuk mengungkap informasi yang tidak terlihat di dalam basis data dan merupakan komponen dari proses *Knowledge Discovery in Database (KDD)* untuk menemukan informasi dan pola yang berguna dalam data (Dewi et al., 2022). Berikut adalah tahapan KDD:

- a. Data Selection. Sebelum tahap penggalian informasi dalam knowledge data discovery dimulai. pemilihan data dari sekumpulan data operasional harus dilakukan. Data yang dipilih untuk digunakan dalam proses penggalian informasi disimpan dalam suatu berkas yang berbeda dari basis data operasional. Sebelum proses penggalian informasi dapat dimulai, dalam proses ini juga dilakukan proses pengembangan, yaitu proses menambah data yang sudah ada dengan data atau informasi lain yang relevan.
- b. Pre-processing atau Cleaning Sebelum proses data mining dapat dimulai, data yang menjadi fokus KDD harus dibersihkan. Proses pembersihan meliputi penggantian data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal, dan penghapusan duplikat data.
- c. Transformation. Mentransformasikan atau menggabungkan data adalah proses menggabungkan berbagai kelompok data menjadi satu kelompok data. Tahap ini merubah data yang telah dipilih dengan tujuan membuatnya lebih sesuai untuk mempermudah proses pemrosesan data,

yaitu dengan mengubah jumlah data menjadi lebih tepat untuk melakukan pemrosesan data melalui peringkasan.

- d. Data Mining Proses data mining adalah mencari pola atau informasi yang menarik dalam sejumlah data. Ini dilakukan dengan menggunakan teknik, metode, atau algoritma tertentu yang disesuaikan dengan tujuan umum proses KDD.
- e. Interpretation / Evaluasi Interpretation / Evaluasi merupakan proses menerjemahkan pola-pola yang dihasilkan dari data mining dan mengevaluasi (menguji) apakah pola atau informasi yang ditemukan sesuai atau bertentangan dengan fakta atau hipotesa sebelumnya. Informasi yang diperoleh dari pola-pola ini dipresentasikan sebagai visualisasi. (Bahtiar, 2023)



Gambar 2.1 Proses Data Mining

Sumber gambar : <https://www.kajianpustaka.com/2017/09/data-mining.html>

1. Pembersihan data, bertujuan untuk melakukan *cleaning* pada data sehingga memberikan data *knowledge* yang sesuai dengan *study case* yang ada.
2. Integrasi data, merupakan tahapan yang bertujuan menggabungkan data yang telah dilakukan pembersihan sehingga menjadi data yang utuh.
3. Transformasi, bertujuan untuk memilah dan mengubah format data agar sesuai dengan teknik atau metode algoritma yang akan digunakan dalam pemrosesannya.
4. *Data Mining*, bertujuan untuk menggali data *knowledge* lebih dalam dengan implementasi dari pemodelan *machine learning*.
5. Evaluasi pola bertujuan untuk menguji hipotesis awal sehingga pemrosesan data sesuai dengan *study case*.

2.3 Random Forest

2.3.1 Pengertian Random Forest

Random Forest, merupakan sebuah metode yang dikembangkan dari metode CART (*Classification and Regression Trees*), yang juga merupakan metode atau algoritma dari teknik pohon keputusan. Yang membedakan metode *random forest* dari metode CART adalah *Random Forest* menerapkan metode *bootstrap aggregating (bagging)* dan juga seleksi fitur *random* atau bisa disebut *random feature selection*. *Random Forest* adalah kombinasi dari masing masing teknik pohon keputusan yang ada, lalu kemudian digabung dan dikombinasikan kedalam suatu model. Ada tiga poin utama dalam metode *Random Forest*, tiga poin utama tersebut yaitu (1) melakukan *bootstrap sampling* untuk membangun pohon prediksi; (2) masing-masing pohon keputusan memprediksi dengan

prediktor acak; (3) kemudian *Random Forest* melakukan prediksi dengan mengombinasikan hasil dari tiap tiap pohon keputusan dengan cara *majority vote* untuk klasifikasi atau rata-rata untuk regresi.(Adrian et al., 2021).

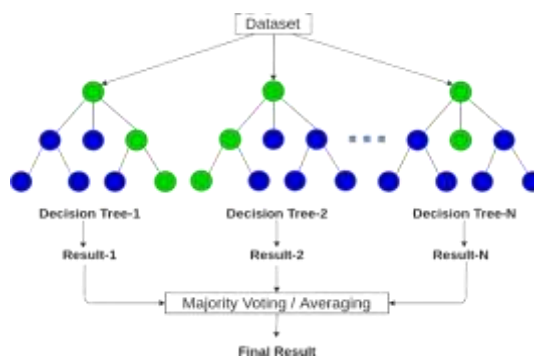
2.3.2 Cara Kerja Random Forest

1. *Bootstrap Sampling* : Data latih yang besar akan dibagi menjadi beberapa subset secara acak melalui metode *bootstrap sampling*. Setiap subset digunakan untuk melatih satu pohon keputusan.
2. Pembangunan Pohon Keputusan : Di setiap pohon, fitur yang digunakan untuk memecah node dipilih secara acak. Ini membuat setiap pohon lebih independen dan mengurangi kemungkinan model mengalami *overfitting*.
3. Voting atau Rata-Rata : Setiap pohon dalam *Random Forest* memberikan hasil prediksi (untuk klasifikasi, ini berupa kelas). Prediksi akhir dihitung dengan cara *majority voting* untuk klasifikasi (dari suara terbanyak antar pohon) atau rata-rata untuk regresi.
4. *Output* : Hasil prediksi yang dihasilkan oleh *Random Forest* adalah akumulasi dari semua pohon yang ada, yang menjadikan model lebih stabil dan lebih akurat daripada penggunaan satu pohon keputusan saja.

2.3.3 Keunggulan Random Forest

1. Akurasi Tinggi : *Random Forest* mampu menghasilkan model yang lebih akurat dibandingkan dengan algoritma model tunggal, seperti pohon keputusan.
2. Tahan Terhadap *Overfitting* : Dengan menggunakan banyak pohon dan memilih fitur secara acak, *Random Forest* mengurangi risiko *overfitting*, yang sering menjadi masalah dalam pohon keputusan tunggal.

3. *Feature Importance* : *Random Forest* dapat digunakan untuk mengidentifikasi fitur-fitur mana yang paling berpengaruh terhadap hasil prediksi, memberikan wawasan yang lebih dalam tentang data.
4. *Robust Terhadap Data yang Hilang* : Algoritma ini cukup kuat untuk menangani data yang hilang tanpa memerlukan *imputation* data secara berlebihan.



Gambar 2.2 Proses Kerja Random Forest

Sumber gambar : <https://www.tibco.com/reference-center/what-is-a-random-forest>

Pada gambar 2.3.3 merupakan alur proses dari algoritma *random forest* dengan detail penjelasan,

1. Algoritma akan memilih sampel data acak Algoritma akan membuat *decision tree* pada setiap sampel acak yang dipilih, sehingga mendapatkan nilai dari prediksi dari setiap *decision tree*.

2. Dilakukan pemilihan terhadap nilai yang telah didapatkan, dengan kondisi ketika terkait nilai klasifikasi akan menggunakan modus, sedangkan nilai regresi akan menggunakan mean.
3. Algoritma akan memilih nilai sesuai kondisi yang ada, untuk mendapatkan hasil terbaik.

Dalam menentukan nilai *decision tree* pada *random forest* akan dilakukan perhitungan nilai *entropy* untuk mengukur tingkat ketidakmurnian atribut dengan rumus sebagai berikut : .

$$Entropy(Y) = - \sum p(c|Y) \log_2 p(c|Y) \quad [1]$$

Tabel 2.1 Penjelasan Entropy

Notasi	Penjelasan
Y	Variabel target (label) yang ingin diprediksi. Contoh: "Apakah pelanggan churn atau tidak".
c	Kelas dari variabel target Y. Contoh: $c \in \{\text{churn}, \text{tidak churn}\}$.
$p(c Y)$	
$\log_2 p(c Y)$	
$\sum$	Simbol sigma berarti penjumlahan. Kamu menjumlahkan nilai untuk semua kelas c yang mungkin dari Y. Misalnya, jika hanya ada dua kelas (churn & tidak churn), maka jumlah untuk dua nilai.
Tanda minus (-)	Karena nilai $\log_2(p)$ untuk $p < 1$ akan menghasilkan angka negatif, maka diberikan tanda minus agar hasil entropi menjadi positif. Entropi selalu ≥ 0 .

Contoh Kasus:

Misalnya, dalam dataset terdapat:

- 60 pelanggan tidak churn $\rightarrow p(\text{tidak churn}|Y) = 0.6$
- 40 pelanggan churn $\rightarrow p(\text{churn}|Y) = 0.4$

Maka

Entropy(Y) adalah :

$$= - [0.6 \log_2 0.6 + 0.4 \log_2 0.4]$$

$$= - [0.6 \times (-0.7369) + 0.4 \times (-1.3219)]$$

$$= - [-0.4421 - 0.5288]$$

$$= -0.9709$$

2.3.4 Kelemahan Random Forest

1. Kompleksitas Model : Karena menggunakan banyak pohon, Random Forest bisa lebih lambat dalam hal pelatihan dan prediksi, terutama ketika data yang digunakan sangat besar.
2. Sulit untuk Interpretasi : Dibandingkan dengan pohon keputusan tunggal, Random Forest lebih sulit diinterpretasikan karena banyaknya pohon yang terlibat. Hal ini bisa menjadi tantangan ketika ingin memahami bagaimana model membuat prediksi.
3. Konsumsi Memori yang Tinggi : Random Forest memerlukan lebih banyak sumber daya komputasi dan memori dibandingkan dengan model lain yang lebih sederhana.

2.3.5 Penerapan Random Forest dalam Customer Churn

Dalam konteks *customer churn*, *Random Forest* dapat digunakan untuk memprediksi pelanggan yang berisiko untuk berhenti berlangganan berdasarkan berbagai atribut pelanggan, seperti lama berlangganan, frekuensi penggunaan layanan, keluhan pelanggan, dan status pembayaran. Dengan kemampuan untuk menangani variabel yang banyak dan mengidentifikasi pola yang kompleks, *Random Forest* dapat memberikan hasil prediksi yang lebih akurat dan dapat diandalkan untuk membantu perusahaan mengurangi *churn* dan meningkatkan retensi pelanggan.

2.4 Customer Churn

Customer churn merujuk pada keadaan di mana pelanggan berhenti menggunakan layanan atau produk dari suatu perusahaan atau berpindah ke pesaing. Hal ini bisa disebabkan oleh berbagai faktor, seperti ketidakpuasan pelanggan, penawaran yang lebih baik dari pesaing, atau perubahan kebutuhan pelanggan. Memahami dan memprediksi customer churn menjadi penting bagi perusahaan untuk mengambil tindakan yang tepat guna mempertahankan pelanggan yang ada. (Amirulhaq Iskandar & Latifa, 2023)

Pelanggan yang mengalami churn bisa disebabkan oleh berbagai faktor, seperti ketidakpuasan terhadap layanan, harga yang tidak kompetitif, kurangnya fitur yang diinginkan, atau karena berpindah ke pesaing. Dalam dunia analisis data dan *machine learning*, *customer churn* biasanya menjadi masalah klasifikasi *biner*, di mana pelanggan dikategorikan menjadi dua kelas: *churn* (berhenti) dan *non-churn* (tetap). Contoh Sederhana:

- Pelanggan yang membatalkan langganan layanan internet → *churn*
- Pelanggan yang masih aktif membayar dan menggunakan layanan → *non-churn*

2.5 Prediksi

Menurut Muhammad, dkk (2020) prediksi merupakan suatu tindakan untuk memperkirakan keadaan pada masa mendatang berdasarkan data masa lampau. Salah satu metode untuk melakukan prediksi yang dapat digunakan adalah metode regresi linier. Metode regresi linier sendiri terdiri dari 2 jenis yaitu metode regresi linier sederhana dan metode regresi linier berganda. Pada penelitian ini menggunakan metode regresi linier sederhana dengan satu variabel bebas. (Almumtazah et al., 2021)

2.6 Algoritma

Algoritma merupakan urutan logis pengambilan keputusan untuk memecahkan sebuah masalah matematika. Kata logis itu merupakan suatu langkah-langkah yang nilai nya harus logis kebenarannya harus logis dan dapat ditentukan nilai nya itu benar atau salah. Suatu langkah-langkah yang tidak benar dapat memberikan suatu yang salah. Programan adalah suatu kumpulan intruksi-intruksi tersendiri yang bisa disebut dengan source code yang dibuat oleh programan.

Algoritma adalah urutan dari langkah-langkah logis penyelesaian masalah yang disusun secara sistematis. Kata logis yang berasal dari kata kunci dari sebuah algoritma. Langkah langkah tersebut harus dapat ditentukan. benar atau salah langkah-langkah didalam algoritma tersebut harus dapat ditentukan.

Algoritma juga merupakan suatu jantung ilmu computer atau informatika. Banyak cabang ilmu Komputer yang diacu dalam teknologi dalam terminology algoritma. Namun jangan beranggapan algoritma selalu identic dengan ilmu komputer saja. Komputer hanyalah salah satu dari pemproses agar dilaksanakan oleh computer, algoritma harus ditulis dalam notasi Bahasa perograman sehingga dinamakan program. Program adalah perwujudan atau implementasi teknis algoritma yang ditululis dalam Bahasa programan tertentu sehingga dapat dilangsanakan oleh komputer. (Rambe, 2021)

2.7 Web

Web juga dikenal sebagai situs, adalah halaman web yang menyajikan informasi kepada user oleh pemilik situs. Halaman pertama setiap situs web dikenal sebagai halaman depan atau halaman rumah. Web, juga dikenal sebagai *World Wide Web* merupakan sebuah sistem yang terintegrasi dalam dokumen *hypertext* yang berisi berbagai jenis informasi, termasuk tulisan, gambar, suara, video, dan informasi multimedia lainnya.

Semua informasi ini dapat diakses melalui browser web, Dengan menggunakan web browser dan web client, halaman web yang disimpan di web server akan diterjemahkan ke dalam bentuk dokumen yang dapat dipahami. melalui protokol yang biasa disebut *http* atau *Hypertext Tansfer Protocol* (Wahyuni & Irawan, 2020).

2.7.1 HTML

HTML atau Hypertext Markup Language adalah bahasa markup umum yang digunakan untuk membuat halaman web. HTML adalah pemrograman web yang mengajarkan peramban web (web browser) bagaimana menyusun dan menampilkan konten di halaman web..(Sudaria et al., 2021)

2.7.2 PHP

PHP merupakan singkatan dari Hypertext Preprocessor yang digunakan sebagai bahasa script server-side dalam pengembangan web yang disisipkan pada dokumen HTML. PHP merupakan software open source yang disebar dan dilisensikan secara gratis serta dapat didownload secara bebas dari situs resminya <http://www.php.net>. Kelebihan dari PHP, yaitu:

- a. PHP adalah bahasa script yang tidak melakukan kompilasi oleh pengguna.
- b. Web server yang mendukung PHP termasuk apache, IIS, Lighttpd, nginx, dan Xitami, dengan konfigurasi yang lebih mudah. Dalam hal pengembangan, ada banyak developer yang siap membantu. Dari perspektif pemahaman, PHP adalah bahasa scripting yang paling mudah digunakan karena memiliki banyak referensi. PHP adalah bahasa open source yang dapat digunakan di berbagai mesin, seperti Linux, Unix, Macintosh, dan Windows. Ini juga dapat dijalankan secara runtime melalui console dan juga memiliki kemampuan untuk menjalankan perintah sistem.(Sahi, 2020).

c. CSS

CSS atau Cascading Style Sheet digunakan untuk mengubah gaya tampilan website Anda, seperti tata letak halaman, warna, dan font. CSS bertujuan untuk membuat website terlihat lebih menarik. Cascading Style Sheet (CSS) merupakan aturan untuk mengatur berbagai bagian dari situs web sehingga lebih terstruktur. (Permana et al., 2022).

d. JavaScript

JavaScript adalah bahasa pemrograman yang digunakan oleh developer untuk menghasilkan halaman web yang menarik. JavaScript merupakan bahasa pemrograman tingkat tinggi yang pertama kali dirilis pada tahun 1995 untuk lingkungan pengembangan aplikasi berbasis web dinamis di sisi client, yang memungkinkan pengembang untuk mengembangkan situs web yang memiliki tampilan yang menarik. Kemampuan JavaScript telah berkembang pesat, termasuk dukungan untuk pemrograman berorientasi objek, pemrograman fungsional, struktural, prosedural, event-driven, prototyping, dan kemudahan pengembangan aplikasi di sisi server. (Eka Pratama, 2020).

e. MySQL

MySQL adalah perangkat lunak sistem manajemen basis data SQL atau DBMS yang multithread dan multiuser yang mendukung bahasa database SQL sebagai bahasa interaktif untuk mengelola data. MySQL juga merupakan database engine atau server database. (Sitanggang Rianto et al., 2022). MySQL adalah Relational Database Management System (RDBMS) yang didistribusikan secara gratis dibawah lisensi GPL (General

Public License). Setiap orang dapat menggunakannya secara bebas, tetapi tidak boleh dibuat menjadi produk turunan komersial atau dari sumber tertutup. SQL (Structur Query Language) adalah konsep pengoperasian database, terutama untuk pemilihan, seleksi, dan pemasukan data, yang memungkinkan pengoperasian data dilakukan dengan mudah secara otomatis. MySQL sebenarnya merupakan turunan dari konsep ini sejak lama. Dalam hal database server, MySQL dapat dikatakan lebih unggul dalam query data dibandingkan database server lainnya.

f. Notepad++

Notepad++ merupakan sebuah program untuk mengedit teks dan kode sumber yang bekerja dengan Windows. Komponen Scintilla memberikan kemampuan Notepad++ untuk menampilkan dan menyunting teks serta berkas kode sumber dari sejumlah bahasa pemrograman. Notepad ++ adalah perangkat lunak bebas. Salah satu fitur yang sangat kecil adalah Notepad, yang biasanya digunakan untuk mencatat nomor HP, mengingat password, dan fitur lainnya. Notepad, aplikasi yang sudah terintegrasi dengan Windows sejak awal, sering digunakan untuk mengedit source code programmer berbasis web.(Ridarmin et al., 2020).

2.8 Machine Learning

2.8.1 Pengertian Machine Learning

Machine Learning (ML) adalah salah satu aplikasi dari *Artificial Intilligent* (AI) yang fokus kepada pengembangan sebuah sistem yang mampu belajar sendiri tanpa harus diprogram berulang kali. ML membutuhkan sebuah

data (*data training*) sebagai proses *learning* sebelum menghasilkan sebuah hasil. Jadi, secara sederhana dapat dijelaskan bahwa *Machine Learning* adalah pemrograman komputer untuk mencapai kriteria/performa tertentu dengan menggunakan sekumpulan data *training* atau pengalaman di masa lalu (*past experience*). (Chazar & Erawan, 2020)

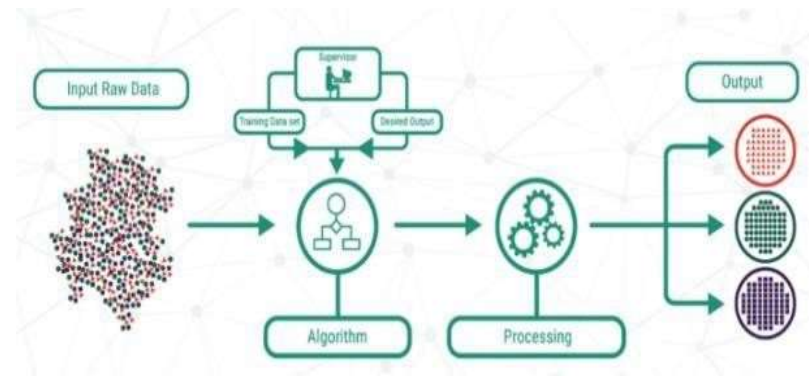
2.9 Jenis-jenis Machine Learning

2.9.1 Supervised Learning (Pembelajaran Terawasi)

Supervised learning merupakan sebuah metode dari *machine learning*, pada *supervised learning* ini sudah memiliki nilai dari label data sebagai *training* dari algoritma dan juga nilai keluaran data yang dapat diamati, (Shah et al., 2019). Teknik *Supervised machine learning*, terbagi menjadi tiga kategori yaitu : (1) *similarity-based*, (2) *model-based* dan (3) *probabilistic-approaches*. (Tomasevic et al., 2020),

Tujuan dari pemodelan algoritma *supervised learning* adalah melatih algoritma untuk berjalan secara otomatis memetakan data *input* ke data *output*, (Shah et al., 2019). Pemodelan *supervised learning* memiliki tahapan dan prosesnya sebagai berikut :

- a) Model dilatih dengan data berlabel (ada input dan output yang diketahui).
- b) Contoh: klasifikasi churn pelanggan (churn/tidak churn).
- c) Algoritma: Decision Tree, Random Forest, KNN, SVM, dll.



Gambar 2.3 Proses Supervised Learning

Sumber gambar : <https://www.techfor.id/konsep-supervised-learning-dalam-membangun-model-machine-learning/>

Pada gambar diatas proses *supervised learning* memiliki beberapa tahapan dengan detail penjelasan, (TechforId, 2021).

1. *Input data raw*, pada tahap ini melakukan input data mentah yang akan diproses dan memastikan bahwa dataset memiliki label pada datanya.
2. *Algoritma*, pada tahap ini melakukan pemilihan algoritma yang akan dipakai, menyesuaikan dengan kasus yang akan dipecahkan, serta melakukan *split data* training dan test data.
3. *Processing*, pada tahap ini melakukan pemrosesan data dengan algoritma yang telah dipilih sesuai dengan pemecahan kasus yang ada.
4. *Output*, merupakan tahap terakhir yaitu hasil dari evaluasi penilaian yang didapatkan dengan pemecahan algoritma yang digunakan.

Pemodelan *supervised learning* memiliki beberapa algoritma dalam penerapannya. Selain itu pada setiap algoritma memiliki keunggulan dan fungsinya masing – masing, sebagai berikut, (Tavasoli, 2023):

1. Regresi linear, bertujuan untuk mengidentifikasi hubungan antara

variabel dependen dengan satu atau lebih variabel independen dan digunakan untuk memprediksi nilai di masa mendatang.

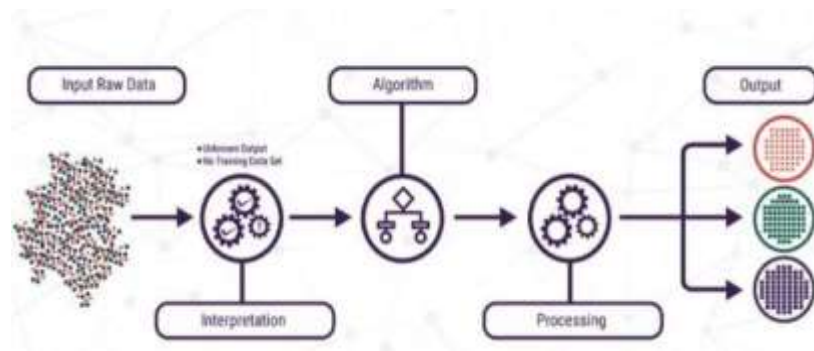
2. *K-Nearest Neighbours*, bertujuan untuk mengelompokkan nilai setiap data dengan menggunakan algoritma non- parametrik yang mengelompokkan nilai poin data berdasarkan kedekatan dan juga asosiasi dengan data yang ada.
3. *Support Vector Machine*, bertujuan untuk melakukan klasifikasi dan juga dapat dilakukan untuk regresi, sehingga algoritma ini berfokus pada penyelesaian klasifikasi serta pembangunan *hyperlane* atau batas keputusan yang memisahkan kelas – kelas titik data.
4. *naïve Bayes*, bertujuan untuk memberikan nilai probabilitas dalam algoritma ini tidak terpengaruh oleh komponen lain dalam hasilnya.
5. *Random Forest*, bertujuan untuk melakukan klasifikasi maupun regresi.
6. *Neural Network* digunakan dalam pembelajaran algoritma yang mendalam metode ini melatih data dengan meniru *interkonektivitas* otak manusia yang saling terhubung melalui lapisan setiap *node*.

2.8.2.2 Unsupervised Learning (Pembelajaran Tak Terawasi)

Unsupervised learning merupakan pemodelan dari *machine learning*, pada *unsupervised learning* ini tidak memiliki sebuah label data dalam pengamatan data-nya, tujuan dari pemodelan ini untuk melatih algoritma menemukan sebuah pola, korelasi atau kluster dalam sebuah dataset, (Shah et al., 2019).

Dalam *Unsupervised learning*, pengamatan atau data digunakan tanpa label, kelas, atau keputusan yang telah ditentukan. Metode pembelajaran tanpa

pengawasan sangat umum digunakan untuk berbagai tugas statistik yang penting, dengan aplikasi yang luas dalam *sains* dan industri. *Unsupervised machine learning* semakin umum dalam memprediksi cacat perangkat lunak. Pendekatan ini dapat bermanfaat bagi praktisi perangkat lunak karena mengurangi ketergantungan pada data pelatihan yang telah dilabeli. Pada *Unsupervised learning* menilai hasil dari metode pembelajaran tanpa pengawasan atau menganalisis data menjadi sulit karena tidak ada mekanisme yang secara umum diterima untuk melakukan validasi silang atau memverifikasi hasil pada set data yang independen. (Nurhalizah et al., 2024)



Gambar 2.4 Proses Unsupervised Learning

Sumber gambar : <https://www.techfor.id/konsep-unsupervised-learning-jenis-dan-contohnya-dalam-machine-learning/>

Pada gambar diatas, unsupervised learning memiliki beberapa tahapan dengan detail penjelasan (M. Sharma, 2019):

1. *Input data raw* pada tahap ini melakukan input data mentah yang akan diproses dan memastikan bahwa *dataset* tidak memiliki label dalam datanya.

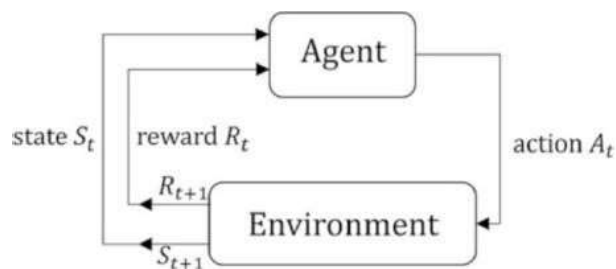
2. *Interpretation* melakukan pengolahan data *knowledge* pada data yang akan diproses.
3. Algoritma pada tahap ini melakukan pemilihan algoritma yang sesuai dengan *study case* untuk mendapatkan hasil yang optimal.
4. *Processing* pada tahap ini algoritma akan melakukan proses algoritma dengan hasil yang optimal.
5. *Output* melakukan evaluasi data yang telah didapatkan apakah data sudah sesuai atau tidaknya.

Pemodelan *unsupervised learning* memiliki beberapa algoritma dalam penerapannya, sebagai berikut : (Behnke, 2003; Ghahramani, 2004):

1. K-Means merupakan algoritma *clustering* dengan metode yang iteratif pada situasi dimana semua variabel dari tipe kuantitatif dan kuadrat jarak *euclidean*. Dan *K-means* melakukan pengelompokan meminimalkan ukuran distorsi daripada menyesuaikan model probabilistik.
2. SVD, digunakan untuk mengurangi nilai matriks
3. PCA atau dikenal dengan *principal components analysis* merupakan analisis komponen utama sebagai faktor pembatas yang penting untuk analisis.
4. Apriori merupakan algoritma dengan tingkatan *advanced* pada *data mining*.
5. FP-Growth atau *frequent pattern growth* merupakan sebuah pengembangan model dari algoritma *apriori*, digunakan sebagai alternatif untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) pada sebuah kumpulan data.

2.1.1.1 Reinforcement Learning

Reinforcement learning merupakan sebuah metode dari *machine learning* berbasis umpan balik di mana agen belajar berperilaku di lingkungan dengan melakukan tindakan dan melihat hasil tindakannya. Dalam kasus ini agen belajar secara otomatis menggunakan umpan balik tanpa data yang berlabel. Maka algoritma yang akan diimplementasikan berdasarkan pengalamannya saja atau berdasarkan umpan balik yang didapatkan, (Trivusi, 2022b). Tujuan dari *reinforcement learning* adalah untuk pembelajaran agen seperti melakukan interaksi yang baik dengan lingkungannya sehingga agen akan mendapatkan nilai yang baik dalam performa algoritmanya. Berikut merupakan alur dari sebuah *reinforcement learning*, (Zhang & Yu, 2020):



Gambar 2.5 Proses Reinforcement Learning

Sumber gambar : <https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292>

2.10 UML (Unified Modeling Language)

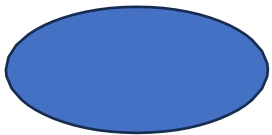
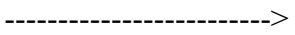
UML (Unified Modeling Language) adalah sistem penjualan dan penyewaan. Contohnya adalah system informasi untuk pengelolaan data penyewaan alat musik. Implementasi UML dapat membuat sistem lebih terstruktur, efisien dan user-friendly. Dengan menggunakan UML saat merencanakan sistem persewaan alat musik berbasis web, diharapkan dapat menyederhanakan proses persewaan dan pengelolaan inventaris, serta meningkatkan efisiensi pengelolaan data dan informasi. (Elis & Voutama, 2023)

2.11 Use Case Diagram

Use Case merupakan salah satu *tools* yang digunakan untuk membuat pemodelan interaksi user dengan sistem. Penelitian ini bertujuan untuk membuat pemodelan interaksi user dengan sistem pada sistem informasi penjualan dengan menggunakan use case. (Setiyani, 2021).

Menurut (Tabrani & Rezqy Aghniya, 2020) menyimpulkan bahwa, “use case adalah rangkaian atau uraian sekelompok yang saling terkait dan membentuk sistem secara teratur yang dilakukan atau diawasi oleh sebuah aktor”. Simbol-simbol yang terdapat pada use case diagram.

Tabel 2.2 Use Case Diagram

No.	Simbol	Keterangan
1.	<i>Actor</i> 	Orang, proses, atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat di luar sistem informasi. Tapi aktor belum tentu merupakan orang, biasanya dinyatakan menggunakan kata bendadi awal frasa.
2.	<i>Use Case</i> 	Fungsionalitas yang disediakan system sebagai unit – unit yang saling bertukaran pesan antar unit atau aktor, biasanya digunakan dengan menggunakan kata kerja diawal frasa.
3.	<i>Asosiasi / association</i> 	Komunikasi antara actor dan use case yang berpartisipasi pada usecase atau use case memiliki interaksi denganactor.
4.	<i>Eksistensi / extend</i> 	Relasi use case ditambahkan ke sebuah usecase di mana use case yang ditambahkan dapat berdiri sendiri walaupun tanpa use case tambahan itu.
5.	<i>Generalisasi / generalization</i> 	Hubungan generalisasi dan spesialisasi (umum–khusus) antara dua buah use case di mana fungsi yang satu adalah fungsi yang lebih umum dari lainnya.

2.12 Flowchart

Flowchart adalah penggambaran secara grafik dari langkah-langkah dan urutan-urutan prosedur dari suatu program. *Flowchart* menolong analis dan programmer untuk memecahkan masalah kedalam segmen-segmen yang lebih

kecil dan menolong dalam menganalisis alternative-alternatif lain dalam pengoperasian. *Flowchart* biasanya mempermudah penyelesaian suatu masalah khususnya masalah yang perlu dipelajari dan di evaluasi lebih lanjut. Perangkat lunak pembelajaran implementasi *flowchart* ini dapat membantu dalam memahami algoritma-algoritma yang ada dalam menampilkan proses dan tahapan-tahapan dalam pembuatan *flowchart*. Perangkat lunak pembelajaran implementasi *flowchart* ini mampu merancang suatu *flowchart* berbasis objek tanpa perlu mengisi kode program isi instruksi dari masing-masing objek *flowchart*. Perangkat lunak pembelajaran implementasi *flowchart* ini juga menyediakan objek-objek yang dibutuhkan dalam pembuatan suatu algoritma *flowchart* sehingga mempermudah pengguna dalam pembuatan suatu diagram alir (*flowchart*). (Zalukhu et al., 2023)

2.13 Activity Diagram

Diagram aktivitas memodelkan tingkah laku dalam sebuah proses. Oleh karena itu, diagram aktivitas tidak secara tepat menggambarkan cara kerja bagian dalam sistem, tetapi umumnya menggambarkan proses dan jalur aktivitas dari tingkat atas. (Handie Pramana Putra et al., 2022)

Tabel 2.3 Activity Diagram

No	Simbol	Keterangan
1.	Status Awal 	Status awal aktifitas sistem, semua diagram aktivitas memiliki sebuah status awal.
2.	Aktivitas 	Aktivitas-aktivitas yang dilakukan sistem, biasanya diawali dengan kata

		kerja.
3.	Swimlane Nama Swimlane	Memisahkan organisasi bisnis yang bertanggung jawab terhadap aktivitas yang terjadi.
4.	Status Akhir 	Status akhir yang dilakukan oleh sistem, semua diagram aktivitas memiliki sebuah status akhir.

2.14 Rapid Minner

Rapid Miner merupakan perangkat lunak yang dibuat oleh Dr. Markus Hofmann dari Insitute of Technologi Blachardstown dan Ralf Klinkenberg dari rapidi.com dengan tampilan GUI (Graphical User Interface) yang dapat mempermudah pengguna dalam menggunakan perangkat lunak ini. Perangkat lunak ini bersifat open source dan dibuat dengan menggunakan program java dibawah lisensi GNU Public Licence dan Rapid Miner dapat dijalankan pada sistem operasi manapun. Dengan menggunakan rapid miner, tidak dibutuhkan kemampuan koding khusus, karena semua fasilitas sudah tersedia. Rapid miner dikhususkan untuk penggunaan data mining, model yang telah disediakan juga cukup banyak dan lengkap, seperti Model Bayesian, Modeling, Tree Induction, Neural Network dan lain-lain. (Suherman et al., 2024)

RapidMiner dapat digunakan untuk melakukan proses analisis pada data mining, text mining, dan analisis prediksi. RapidMiner menggunakan berbagai cara dan teknik deskriptif serta prediksi dalam pembuatan model yang dapat digunakan dalam pengambilan keputusan (Chisholm, 2013). Penggunaan RapidMiner dalam pembuatan model tidak memerlukan program dan semua yang digunakan dalam pembuatan model sudah tersedia dalam bentuk operator. Pembuatan model menggunakan berbagai operator yang sesuai dan saling dikaitkan dalam membentuk suatu model. Proses pembuatan data mining dengan

RapidMiner sudah dilakukan beberapa peneliti dengan berbagai datasheet dan model serta algoritma yang digunakan. Penelitian tersebut di antaranya dilakukan oleh Chisholm (2013), Prasetyo et al. (2021), dan Sudarsono et al. (2021).(Sudarsono et al., 2021)

2.15 Studi Literatur

Tabel 2.4 Studi Literatur

No	Judul	Metode dan Proses	Hasil	Kelebihan & Kekurangan
1	<i>Comparing to techniques used in customer churn analysis (Osmanoglu, 2019)</i>	Metode, churn analytics dengan algoritma (deep learning, random forest, (SVM). Proses, dilakukan pemilahan dataset, lalu split data dan implementasi machine Learning.	Implementasi nilai Churn dapat Dilakukan oleh Beragam machine Learning tetapi Diantaranya deep learning memiliki nilai yang tinggi dibandingkan dengan yang lainnya.	Kelebihan, Deep learning memiliki nilai yang cukup baik dalam penentuan churn dengan 77,8% dan disusul oleh random forest 73,8%. Kekurangan, Pada implementasi deep learning memiliki kompleksitas algoritma yang cukup sulit.
2	<i>Customer Churn Prediction in influencer Commerce : An Application of decision tree (Kim & Lee, 2021)</i>	Metode, decision tree Proses, penarikan data, split data (training dan testing data) dilakukan. Proses, machine learning decision tree dan akan dilakukan penilaian berdasarkan confusion matrix.	Implementasi Decision tree memiliki nilai confusion matrix yang bagus dengan rerata 82% dalam memprediksi nilai churn.	Kelebihan, Terdapat implementasi perbandingan nilai <i>accuracy</i> tiga algoritma, sehingga mengetahui nilai algoritma yang optimal Kekurangan, Tidak adanya implementasi data mining pada dataset. Penggunaan Decision tree Memiliki nilai akurasi yang tinggi dan algoritma yang tidak kompleks.
3	<i>A comparative study of employee churn prediction model. (Alamsyah & Salma, 2018)</i>	Metode, <i>naïve bayes</i> , <i>decision tree</i> dan <i>random forest</i> Proses, <i>data preprocessing</i> , <i>model construction</i> dan dilakukan perbandingan nilai model.	Implementasi <i>random forest</i> memiliki nilai yang tinggi dengan nilai <i>accuracy</i> 97.5% dibanding algoritma <i>naïve bayes</i> dan <i>decision tree</i> .	Kekurangan, Tidak adanya penerapan implementasi <i>data mining</i> lebih dalam.

4	<i>Dynamic churn prediction using machine learning algorithms—predict your customer through customer behavior</i> (Vln & Deeplakshmi, 2021)	Metode, prediksi nilai <i>churn</i> menggunakan algoritma SVM Proses, preprocessing data, split data(training dan testing), implementasi model SVM.	Konsep pembahasan penerapan <i>machine learning</i> pada studi kasus pencarian nilai <i>churn</i> .	Kelebihan Implementasi detail terkait konsep dan pembahasan dari nilai <i>churn</i> pada konsumen. Kekurangan, Tidak adanya hasil pembuktian nilai penggunaan algoritma SVM.
---	---	---	---	--

BAB III

METODOLOGI PENELITIAN

3.1 Jenis Penelitian

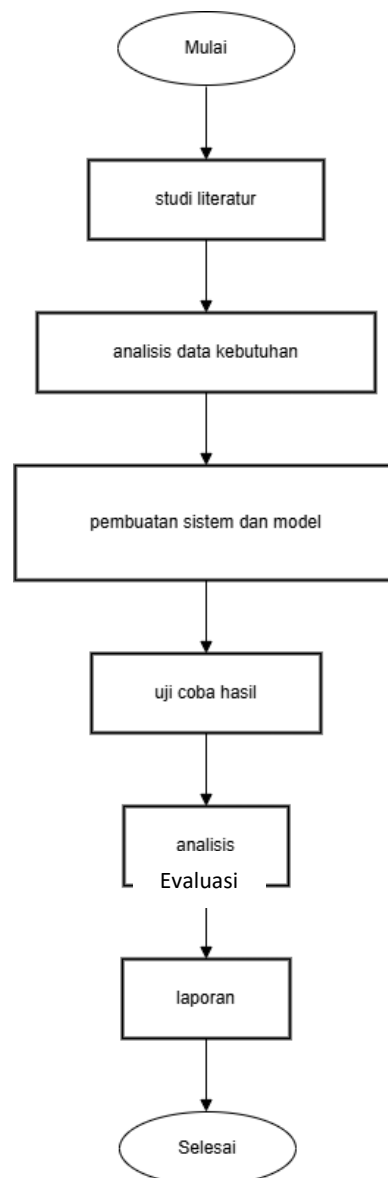
Pada pengumpulan data yang diperlukan dalam penelitian ini, penulis membuat *dataset* dalam merancang model RFM dan nilai *churn* konsumen. Tak lepas dari bahan bacaan berupa jurnal skala nasional maupun internasional, *website* dan buku untuk dijadikan referensi sebagai pengetahuan tambahan penulis dalam melakukan pengembangan metode yang akan diimplementasikan pada penelitian ini untuk mencari pemodelan yang optimal.

Model penelitian ini termasuk dalam kategori penelitian kuantitatif, Penelitian kuantitatif bersifat deskriptif dan cenderung menggunakan analisis pendekatan induktif. Penelitian kuantitatif lebih menonjol disusun dalam bentuk narasi yang bersifat kreatif dan mendalam serta menunjukkan ciri-ciri naturalistic yang penuh dengan nilai-nilai otentik. Penelitian kuantitatif adalah penelitian yang menitikberatkan pada pengukuran dan analisis hubungan sebab-akibat antara bermacam macam variabel, bukan prosesnya, penyelidikan dipandang berada dalam kerangka bebas nilai. Penelitian kuantitatif menghasilkan informasi yang lebih terukur. Hal ini karena ada data yang dijadikan landasan untuk menghasilkan informasi yang lebih terukur. Penelitian kuantitatif tidak mempermasalahkan hubungan antara peneliti dengan subyek penelitian karena hasil penelitian lebih banyak tergantung dengan instrumen yang digunakan dan

terukur variabel yang digunakan, dari pada intim dan keterlibatan emosi antara peneliti dengan subyek yang diteliti.

3.2 Diagram Alur Penelitian

Berikut merupakan flowchart system yang digunakan pada penelitian customer churn pada PT Education Nusantara Connect.



Gambar 3.1 Diagram Alur Penelitian

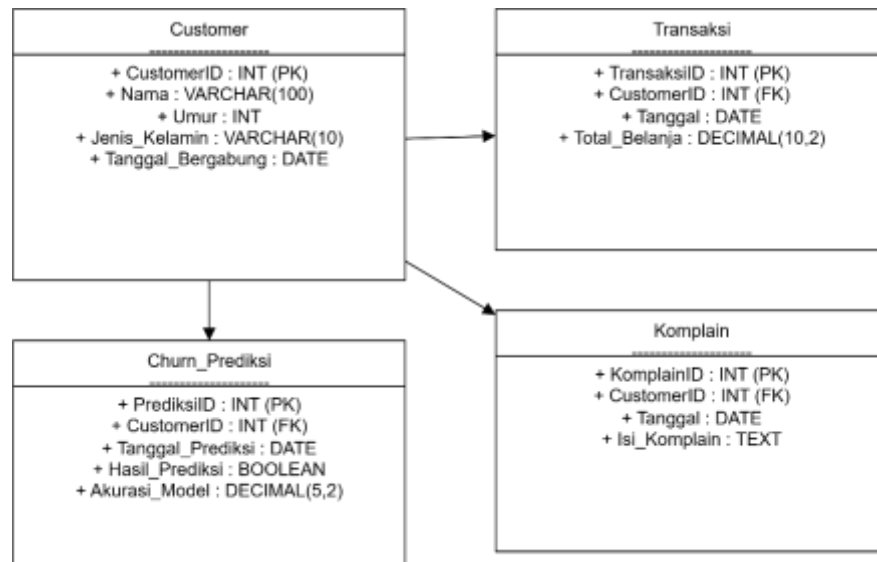
Penelitian ini diawali dengan tahap **Mulai**, yang menandai dimulainya keseluruhan proses penelitian. Selanjutnya, dilakukan **Studi Literatur** guna mengumpulkan teori-teori, metode, dan penelitian terdahulu yang relevan, khususnya terkait customer churn, data mining, dan algoritma Random Forest. Setelah dasar teori dipahami, masuk ke tahap **Analisis Data & Kebutuhan**, yang mencakup identifikasi data pelanggan dari PT Education Nusantara Connect, pemahaman atribut penting, serta penentuan kebutuhan sistem untuk proses klasifikasi.

Tahap berikutnya adalah **Pembuatan Model Prediksi**, di mana dilakukan pengolahan data, pelatihan model menggunakan algoritma Random Forest, serta pengembangan sistem klasifikasi churn. Setelah sistem selesai dibangun, dilakukan **Uji Coba Model** untuk mengetahui tingkat akurasi model dan performanya terhadap data uji menggunakan rapidminer. Kemudian dianalisis pada tahap **Evaluasi**, yang meliputi evaluasi performa model, interpretasi hasil klasifikasi, serta identifikasi faktor-faktor yang memengaruhi churn.

Setelah analisis dilakukan secara menyeluruh, hasil penelitian dirangkum dalam **Laporan**, yang mendokumentasikan seluruh proses dan temuan. Akhirnya, penelitian dinyatakan **Selesai** setelah seluruh tahapan berhasil diselesaikan dan laporan akhir siap dipresentasikan atau dipublikasikan.

3.3 Rancangan

3.3.1 Rancangan Database



Gambar 3.2 Rancangan Database

Diagram di atas merupakan model Entity Relationship Diagram (ERD) yang menggambarkan struktur basis data untuk sistem prediksi *customer churn*. Diagram ini terdiri dari empat entitas utama, yaitu Customer, Transaksi, Komplain, dan Churn_Prediksi yang saling berelasi satu sama lain untuk mendukung proses analisis perilaku pelanggan dan prediksi potensi churn.

Entitas Customer merupakan pusat dari sistem, yang merepresentasikan data pelanggan, termasuk atribut seperti *id_customer*, nama, email, dan informasi lainnya. Setiap pelanggan dapat melakukan beberapa transaksi yang dicatat dalam entitas Transaksi. Entitas Transaksi menyimpan data seperti *id_transaksi*, tanggal_transaksi, dan total_pembayaran, serta dihubungkan ke entitas Customer melalui atribut *id_customer*. Hubungan ini menunjukkan bahwa satu pelanggan

bisa melakukan banyak transaksi, tetapi satu transaksi hanya dimiliki oleh satu pelanggan.

Selain transaksi, pelanggan juga dapat melakukan komplain terhadap layanan. Hal ini direpresentasikan oleh entitas Komplain, yang menyimpan data seperti `id_komplain`, `tanggal_komplain`, dan `isi_komplain`. Relasi antara Customer dan Komplain juga bersifat satu ke banyak, di mana satu pelanggan dapat memiliki beberapa komplain.

Untuk tujuan utama sistem ini, yaitu melakukan prediksi churn, digunakan entitas Churn_Prediksi. Entitas ini mencatat hasil prediksi apakah seorang pelanggan berpotensi untuk berhenti menggunakan layanan (churn) atau tidak. Atribut seperti `id_prediksi`, `hasil_prediksi`, `probabilitas`, dan `tanggal_prediksi` disediakan untuk menyimpan data hasil analisis. Setiap entri dalam entitas ini terkait dengan satu pelanggan melalui `id_customer`.

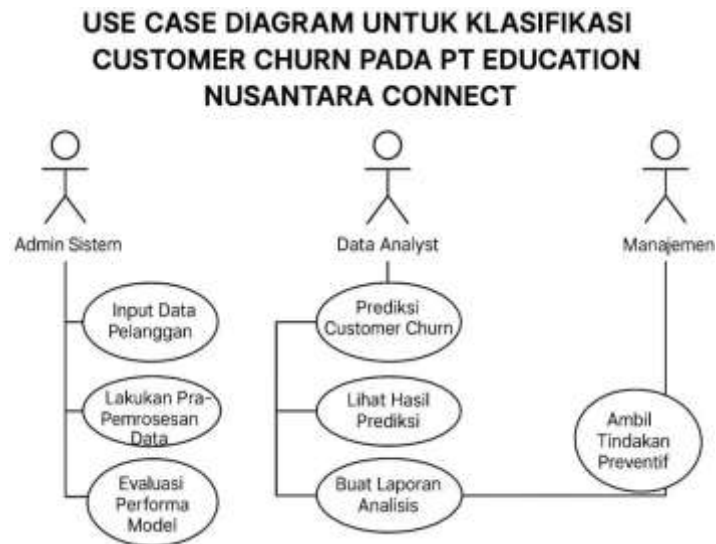
Secara keseluruhan, model ERD ini mendukung proses prediksi churn dengan cara mengintegrasikan data historis pelanggan (melalui transaksi dan komplain) serta mengaitkannya dengan hasil prediksi. Dengan struktur ini, sistem mampu menganalisis perilaku pelanggan secara menyeluruh dan menghasilkan prediksi churn yang dapat dimanfaatkan perusahaan untuk mengambil langkah strategis dalam mempertahankan pelanggan.

Tabel 3.1 Rancangan Data Base

Nama Kolom	Tipe Data	Keterangan
<code>id_pelanggan</code>	INT (Primary Key)	ID unik untuk setiap pelanggan

Nama Kolom	Tipe Data	Keterangan
nama	VARCHAR	Nama pelanggan (opsional untuk privasi)
jenis_kelamin	VARCHAR(10)	Laki-laki / Perempuan
usia	INT	Usia pelanggan
lokasi	VARCHAR	Lokasi/kota domisili pelanggan
status_pembayaran	VARCHAR	Tepat waktu / Menunggak
jumlah_keluhan	INT	Total keluhan yang pernah diajukan
lama_berlangganan	INT	Dalam bulan (misal: 12 bulan, 6 bulan)
frekuensi_login	INT	Jumlah login rata-rata per bulan
aktivitas_terakhir	DATE	Tanggal terakhir pelanggan aktif
jenis_paket	VARCHAR	Basic / Premium / Trial / dll.
churn	BOOLEAN / INT	1 = churn, 0 = tidak churn (target klasifikasi)

3.3.2 Rancangan Use Case Diagram



Gambar 3.3 Use Case Diagram PT Education Nusantara Connect

Pertama, Admin Sistem memiliki tanggung jawab awal dalam alur sistem. Aktivitasnya dimulai dengan Input Data Pelanggan, yaitu memasukkan data historis pelanggan ke dalam sistem. Setelah data masuk, admin melakukan Pra-pemrosesan Data, yaitu tahap pembersihan dan transformasi data agar siap digunakan oleh model klasifikasi. Selanjutnya, admin juga bertanggung jawab dalam Evaluasi Performa Model, yakni melakukan penilaian terhadap akurasi atau metrik performa dari model klasifikasi yang digunakan.

Kemudian, Data Analyst menjalankan proses utama dalam pengambilan keputusan berbasis data. Mereka melakukan Prediksi Customer Churn menggunakan model yang sudah dilatih. Setelah prediksi dilakukan, Data Analyst dapat Melihat Hasil Prediksi, yakni menginterpretasikan pelanggan mana saja yang berpotensi churn. Berdasarkan hasil tersebut, mereka juga membuat Laporan

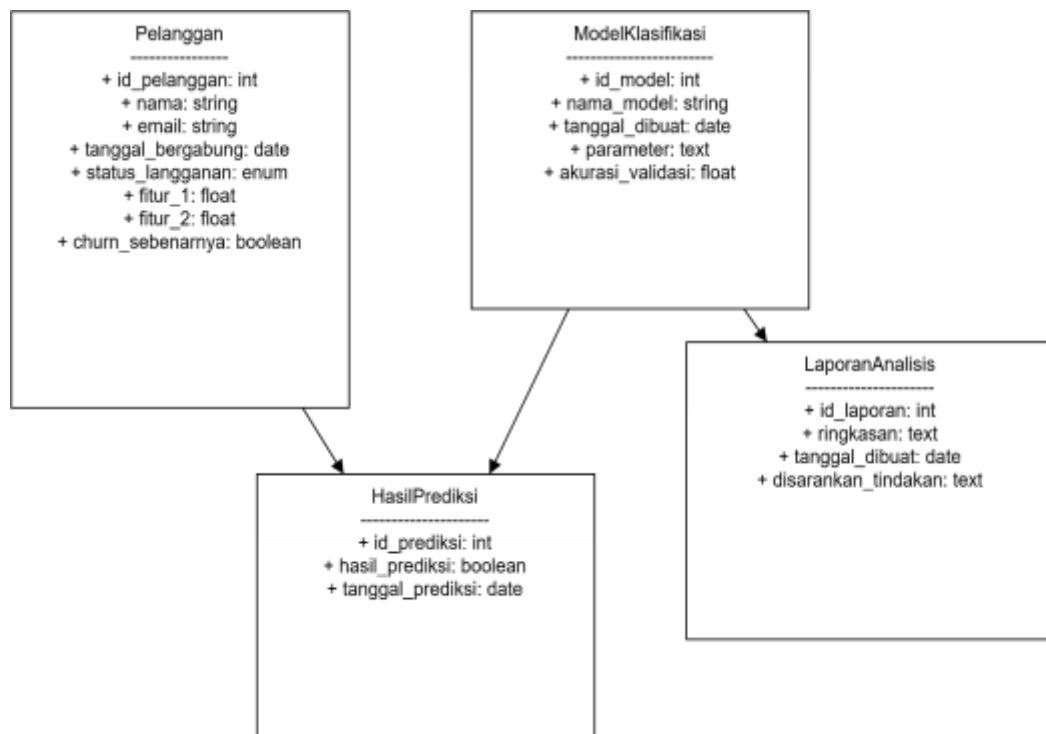
Analisis yang mendokumentasikan hasil dan insight yang diperoleh dari proses klasifikasi.

Terakhir, pihak Manajemen menggunakan informasi dari laporan tersebut untuk melakukan Tindakan Preventif, yaitu merancang strategi atau program retensi guna mencegah pelanggan yang diprediksi akan churn agar tetap bertahan menggunakan layanan PT Education Nusantara Connect.

Dengan alur kerja ini, use case diagram menunjukkan kolaborasi antara berbagai peran untuk membentuk sistem klasifikasi churn yang terstruktur dan berbasis data, sehingga mendukung pengambilan keputusan yang lebih tepat sasaran dalam upaya menjaga loyalitas pelanggan.

3.4 Class Diagram

Class diagram adalah salah satu jenis diagram dalam UML (Unified Modeling Language) yang digunakan untuk memodelkan struktur sistem perangkat lunak berbasis objek. Diagram ini menggambarkan kelas-kelas dalam sistem, atribut dan metode yang dimiliki masing-masing kelas, serta hubungan antar kelas. Class diagram ini menggambarkan alur data yang sistematis dalam sistem klasifikasi customer churn. Data pelanggan diproses oleh model klasifikasi, menghasilkan output berupa hasil prediksi churn. Hasil tersebut kemudian dikompilasi dalam laporan analisis yang menjadi acuan bagi pengambilan keputusan strategis oleh pihak manajemen.



Gambar 3.4 Class Diagram

Diagram di atas menggambarkan rancangan sistem prediksi berbasis model klasifikasi, seperti Random Forest, yang terdiri dari beberapa entitas utama, yaitu **Pelanggan**, **Model Klasifikasi**, **Hasil Prediksi**, dan **Laporan Analisis**. Entitas *Pelanggan* menyimpan informasi identitas dan karakteristik pengguna layanan, yang menjadi objek utama dalam proses prediksi. Setiap pelanggan dapat memiliki satu atau lebih hasil prediksi yang tersimpan dalam entitas *HasilPrediksi*. Entitas ini mencatat informasi seperti hasil prediksi (misalnya churn atau tidak), nilai probabilitas prediksi, serta waktu pelaksanaan prediksi.

Proses prediksi tersebut dilakukan menggunakan model yang tersimpan pada entitas Model Klasifikasi. Setiap model memiliki identitas unik dan dapat digunakan untuk menghasilkan berbagai prediksi untuk banyak pelanggan. Selain

itu, model juga dievaluasi melalui entitas Laporan Analisis yang berfungsi untuk menyimpan metrik evaluasi seperti akurasi, precision, recall, dan f1-score. Laporan ini berguna untuk menilai performa dari model yang digunakan sehingga dapat dibandingkan antar model jika dilakukan eksperimen lebih dari satu jenis algoritma klasifikasi.

Dengan rancangan ini, sistem mampu melakukan prediksi perilaku pelanggan secara otomatis dan menyimpan riwayat prediksi serta performa modelnya. Struktur ini juga memungkinkan pengembangan sistem analisis churn yang terukur, terdokumentasi, dan dapat diandalkan untuk pengambilan keputusan bisnis.

3.5 Kerangka Berfikir

Kerangka berpikir adalah suatu rancangan yang digunakan untuk membantu peneliti dalam menyelesaikan penelitian yang sudah dibuatnya. Kerangka berpikir dibuat dalam bagian-bagian penting yang harus dikerjakan terlebih dahulu. Kerangka berpikir merupakan dasar pemikiran dari penulisan ataupun penelitian yang disusun dari faktafakta, observasi, serta kajian kepustakaan. Sehingga itulah sebabnya pada saat melakukan proses penulisan atau penelitian, seorang penulis atau peneliti harus menyiapkan kerangka berpikir.(Ekayanti H., Makkasau, Fitriani., 2023). Berikut merupakan gambar dari kerangka berfikir dari penelitian yang saya buat.



Gambar 3.5 Diagram Kerangka Berfikir

Penelitian ini dimulai dari permasalahan tingginya tingkat *customer churn* atau pelanggan yang berhenti berlangganan di PT Education Nusantara Connect (PT ENC). Fenomena ini berdampak negatif terhadap keberlangsungan bisnis dan menuntut perusahaan untuk segera mencari solusi prediktif guna mempertahankan pelanggan. Oleh karena itu, langkah awal yang dilakukan adalah mengidentifikasi kebutuhan prediksi churn sebagai upaya untuk mengantisipasi potensi kehilangan pelanggan di masa depan. Salah satu pendekatan yang digunakan adalah pemanfaatan data historis pelanggan yang telah dikumpulkan oleh PT ENC.

Data pelanggan tersebut kemudian dijadikan sumber utama dalam proses eksplorasi dan analisis data. Proses eksplorasi dilakukan untuk memahami karakteristik data, membersihkan data dari noise atau nilai yang tidak valid, serta mengidentifikasi variabel-variabel yang berpengaruh terhadap kemungkinan terjadinya churn. Setelah data siap digunakan, langkah selanjutnya adalah membangun model klasifikasi menggunakan algoritma Random Forest. Algoritma ini dipilih karena kemampuannya yang baik dalam menangani data dengan dimensi tinggi dan menghasilkan performa prediksi yang akurat.

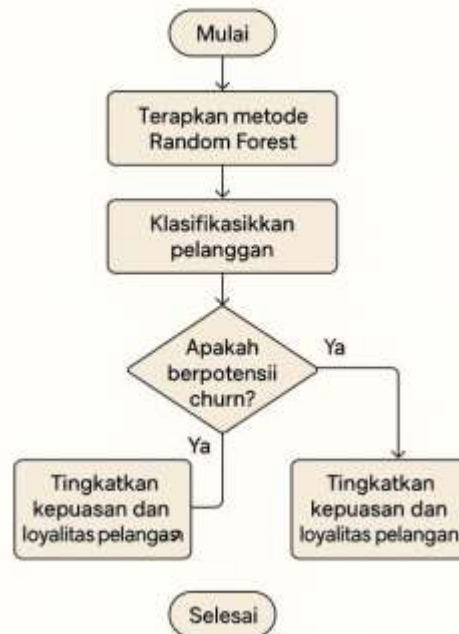
Model yang telah dibangun dievaluasi menggunakan metrik-metrik seperti accuracy, precision, recall, dan F1-score guna mengukur sejauh mana model mampu memprediksi churn dengan baik. Hasil evaluasi ini menjadi dasar untuk menilai apakah model layak digunakan pada data pelanggan yang sebenarnya. Setelah model tervalidasi, dilakukan proses prediksi churn pelanggan, di mana pelanggan yang berisiko tinggi untuk berhenti dapat terdeteksi lebih awal. Akhir dari tahapan ini adalah menghasilkan rekomendasi strategi retensi, yang bertujuan untuk mempertahankan pelanggan berdasarkan hasil prediksi yang telah diperoleh. Strategi ini diharapkan dapat membantu perusahaan dalam mengambil keputusan yang lebih tepat sasaran dalam mengurangi churn.

3.6 Activity Diagram

Berikut merupakan alur dari penelitian mengenai customer churn untuk mendapatkan hasil churn atau non-churnnya customer PT Education Nusantara Connect.

Klasifikasi Customer Churn Menggunakan Metode Random Forest

PT Education Nusantara Connect



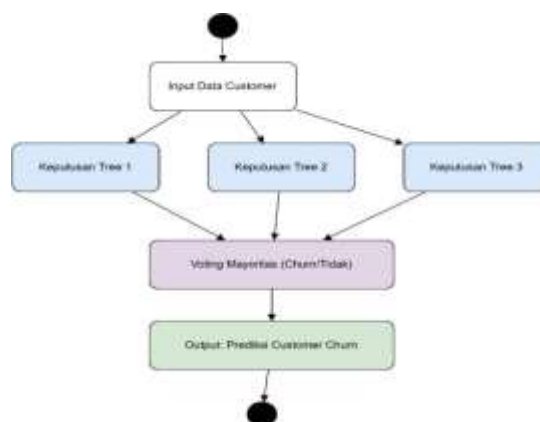
Gambar 3.6 Activity Diagram

Diagram alir di atas menggambarkan proses klasifikasi *customer churn* menggunakan metode Random Forest di PT Education Nusantara Connect. Proses dimulai dari tahap awal, yaitu penerapan metode Random Forest pada data pelanggan yang tersedia. Metode ini dipilih karena kemampuannya yang baik dalam mengklasifikasikan data dengan tingkat akurasi tinggi dan resistensi terhadap overfitting. Setelah model diterapkan, pelanggan diklasifikasikan ke dalam dua kelompok utama: pelanggan yang berpotensi churn dan pelanggan yang tidak.

Langkah berikutnya adalah evaluasi hasil klasifikasi untuk mengetahui apakah seorang pelanggan tergolong dalam kelompok yang berisiko churn. Jika pelanggan teridentifikasi memiliki potensi churn, maka perusahaan perlu segera melakukan tindakan strategis untuk meningkatkan kepuasan dan loyalitas pelanggan tersebut, misalnya dengan memberikan penawaran khusus, meningkatkan layanan, atau mengadakan komunikasi yang lebih personal. Demikian pula, jika pelanggan tidak termasuk dalam kelompok berisiko churn, perusahaan tetap dianjurkan untuk menjaga dan meningkatkan kepuasan serta loyalitas agar pelanggan tidak berpindah di masa mendatang. Seluruh rangkaian proses ini bertujuan untuk menekan tingkat churn dan mempertahankan basis pelanggan yang sudah ada, sehingga berkontribusi terhadap pertumbuhan dan stabilitas bisnis perusahaan.

3.7 Rancangan Prediksi Hasil Random Forest

Rancangan prediksi adalah suatu kerangka kerja atau langkah-langkah sistematis yang dirancang untuk membangun model prediksi berdasarkan data yang ada, dengan tujuan untuk memprediksi hasil. Berikut merupakan gambar dari rancangan prediksi hasil dari penelitian yang saya buat.



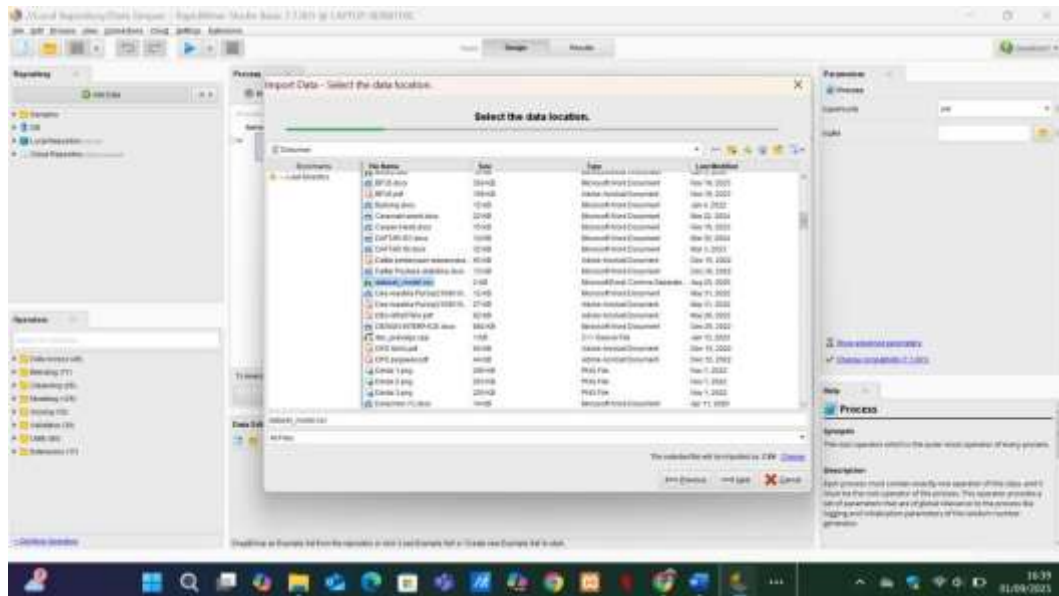
Gambar 3.7 Rancangan Prediksi Hasil Random Forest

Diagram tersebut menjelaskan mekanisme kerja algoritma Random Forest dalam memprediksi *customer churn* berdasarkan data pelanggan. Proses dimulai dengan melakukan input data customer, yaitu data historis atau profil pelanggan yang mencakup berbagai variabel seperti durasi langganan, frekuensi penggunaan, dan keluhan pelanggan. Setelah data dimasukkan, model Random Forest akan meneruskannya ke beberapa decision tree (pohon keputusan) yang bekerja secara paralel. Pada diagram ini ditunjukkan tiga pohon keputusan, yakni *Tree 1*, *Tree 2*, dan *Tree 3*—meskipun dalam implementasi nyata bisa saja terdapat puluhan atau ratusan tree.

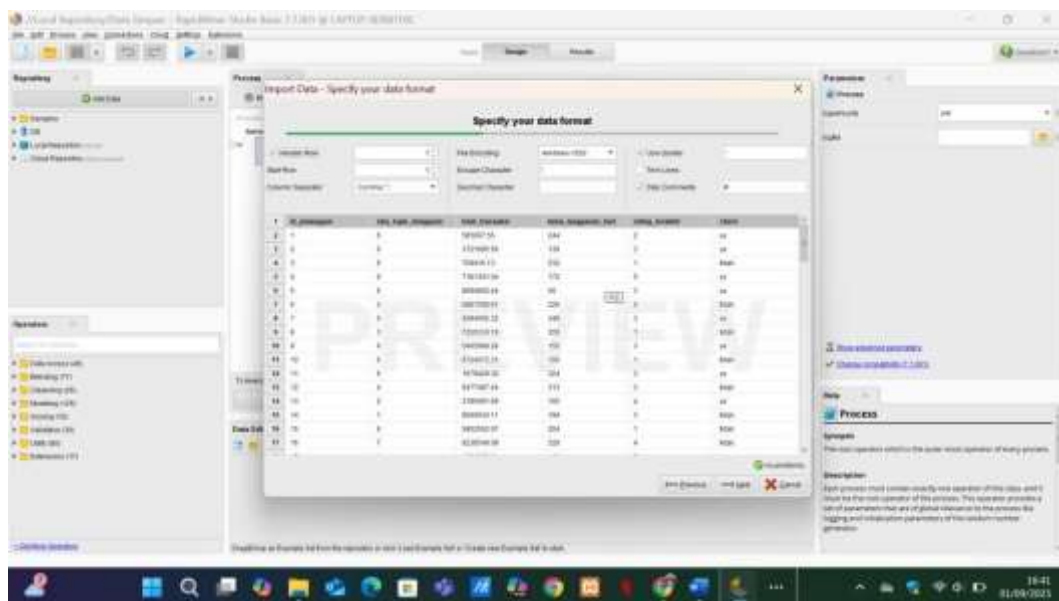
Masing-masing pohon akan memberikan hasil prediksi independen terhadap apakah pelanggan tersebut berpotensi churn atau tidak. Kemudian, semua hasil dari pohon-pohon tersebut dikumpulkan dalam proses voting mayoritas, di mana keputusan akhir diambil berdasarkan suara terbanyak dari semua tree. Jika mayoritas pohon memprediksi bahwa pelanggan akan churn, maka hasil akhirnya adalah “churn”, dan sebaliknya. Tahap akhir dari proses ini adalah output berupa prediksi customer churn, yaitu kesimpulan apakah pelanggan tertentu termasuk kelompok yang berisiko churn atau tidak. Pendekatan ini memungkinkan hasil prediksi yang lebih akurat dan stabil dibandingkan penggunaan satu pohon keputusan saja, karena mengurangi risiko kesalahan akibat bias atau overfitting dari satu model tunggal.

Pembangunan model klasifikasi dilakukan menggunakan perangkat lunak RapidMiner dengan algoritma **Random Forest**. Tahapan dimulai dari proses pengambilan dataset melalui operator *Retrieve* yang telah disimpan dalam repository. Selanjutnya, atribut target ditentukan sebagai label menggunakan operator *Set Role*. Untuk membangun dan menguji model, digunakan operator *Cross Validation* yang membagi data menjadi data latih dan data uji. Pada bagian pelatihan, data diolah dengan algoritma Random Forest, sedangkan pada bagian pengujian hasil prediksi dievaluasi menggunakan operator *Performance (Classification)*. Parameter utama pada Random Forest, seperti jumlah pohon (*number of trees*) dan kedalaman maksimum (*maximal depth*), diatur sesuai

kebutuhan. Hasil akhir dari proses ini berupa model klasifikasi serta nilai evaluasi performa seperti akurasi, precision, recall, F-measure, dan confusion matrix.



Gambar 3.8. Tampilan pilihan *load* data rapid miner



Gambar 3.9 Tampilan format coloum data produk

3.8 Tempat dan Waktu Penelitian

Tempat penelitian ini dilakukan di PT Education Nusantara Connect yang merupakan salah satu perusahaan penyedia jasa layanan jaringan internet yang beralamat di Cikampak, Labuhan Batu Utara, Sumatera Utara.

Tabel 3.2 Waktu Penelitian

[illegible]

BAB IV

HASIL DAN UJI COBA

4.1. Hasil

Hasil dari Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dilihat sebagai berikut:

1. Form Login

Form Login dari Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dilihat pada Gambar 4.1.



The image shows a web application interface for PT. Education Nusantara Connect. At the top left is the logo 'ENC' with the text 'Klasifikasi Customer Churn PT. Education Nusantara Connect'. At the top right is a header bar with 'Customer Churn' and a phone number '+62 812-3456-7890'. The main content area is a light blue box titled 'Masukan Login'. It contains two input fields: 'USERNAME' and 'PASSWORD', and a 'SUBMIT' button. Below the input fields, there is a small text block: 'Klasifikasi adalah salah satu teknik dalam data mining dan machine learning yang digunakan untuk menentukan objek atau data ke dalam kategori atau kelas tertentu'.

Gambar 4.1. Form Login

2. Form Home

Form Home dari Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dilihat pada Gambar 4.2.



Gambar 4.2. Form Home

3. Form Pelanggan

Form Pelanggan dari Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dilihat pada Gambar 4.3.

	No HP	Nama	Email	Tanggal Daftar	Status
Tambah	081234567890	Ahmad	ahmad@gmail.com	2025-07-01	AKTIF
Tambah	08987654321	Budi	budi@gmail.com	2025-07-16	AKTIF
Tambah	091234567890	Ahmad Nurharyo	ahmadnurharyo@gmail.com	2023-09-30	AKTIF
Tambah	082345678901	Suci	sucikaa34@gmail.com	2023-09-16	AKTIF
Tambah	095643207651	Supardi	supardi78@gmail.com	2023-06-25	AKTIF
Tambah	081234567890	Suciati	suciati42@gmail.com	2024-05-30	AKTIF
Tambah	081367542901	Rini Cahya	rinicahya43@gmail.com	2024-06-17	AKTIF
Tambah	0813627724589	Dina Puspita	dina54@gmail.com	2023-04-26	AKTIF
Tambah	082176528763	Rendi Shopping	rendishopping1@gmail.com	2022-09-30	AKTIF
Tambah	081376568754	Rudians	rudians00@gmail.com	2023-07-24	AKTIF

Gambar 4.3. Form Pelanggan

4. Form Aktivitas Pelanggan

Form Aktivitas Pelanggan dari Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dilihat pada Gambar 4.4.

The screenshot shows the 'Data Aktivitas Pelanggan' form. At the top, there is a header with the ENC logo and the text 'Klasifikasi Customer Churn PT. Education Nusantara Connect'. To the right of the header, there is a 'Customer Churn' button and a phone number '+62 812-3456-7890'. Below the header is a navigation bar with buttons: HOME, DATA +, MODEL +, PREDIKSI +, MONITORING CHURN, and EXIT. The main title of the form is 'Data Aktivitas Pelanggan'. Below the title is a table with the following columns: ID Pelanggan, Tanggal, Aktivitas, and Durasi (menit). The table is currently empty. Below the table, there is a text box with the following text: 'Klasifikasi adalah salah satu teknik dalam data mining dan machine learning yang digunakan untuk memetakan objek atau data ke dalam kategori atau kelas tertentu'.

Gambar 4.4. Form Aktivitas Pelanggan

5. Form Feedback Pelanggan

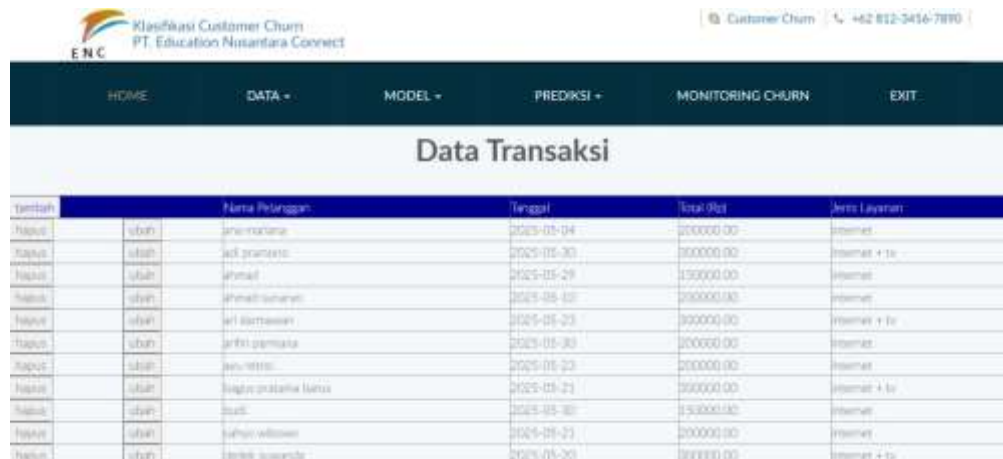
Form Feedback Pelanggan dari Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dilihat pada Gambar 4.5.

The screenshot shows the 'Data Feedback Pelanggan' form. At the top, there is a header with the ENC logo and the text 'Klasifikasi Customer Churn PT. Education Nusantara Connect'. To the right of the header, there is a 'Customer Churn' button and a phone number '+62 812-3456-7890'. Below the header is a navigation bar with buttons: HOME, DATA +, MODEL +, PREDIKSI +, MONITORING CHURN, and EXIT. The main title of the form is 'Data Feedback Pelanggan'. Below the title is a table with the following columns: Nama Pelanggan, Tanggal, Rating, and Komentar. The table is currently empty. Below the table, there is a text box with the following text: 'Klasifikasi adalah salah satu teknik dalam data mining dan machine learning yang digunakan untuk memetakan objek atau data ke dalam kategori atau kelas tertentu'.

Gambar 4.5. Form Feedback Pelanggan

6. Form Transaksi

Form Transaksi dari Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dilihat pada Gambar 4.6.

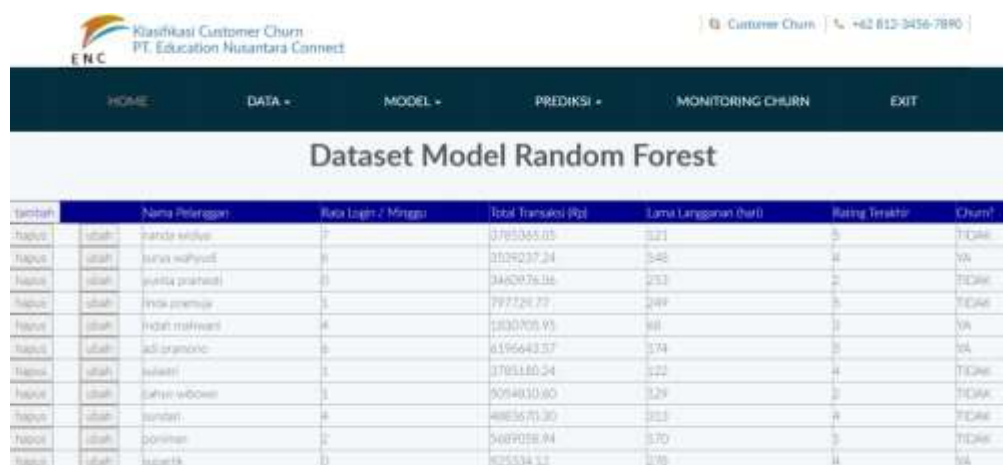


Tambahan	statis	Nama Pelanggan	Tanggal	Total Rp	Jenis Layanan
Hasbi	statis	arif notomane	2025-05-04	200000.00	Internet
Hasbi	statis	adi pramono	2025-05-30	200000.00	Internet + tv
Hasbi	statis	ahmad	2025-05-29	150000.00	Internet
Hasbi	statis	ahmad sudarso	2025-05-02	200000.00	Internet
Hasbi	statis	ari darmawan	2025-05-23	200000.00	Internet + tv
Hasbi	statis	arif pambata	2025-05-30	200000.00	Internet
Hasbi	statis	awidhi	2025-05-23	200000.00	Internet
Hasbi	statis	bagus pratomo harso	2025-05-21	200000.00	Internet + tv
Hasbi	statis	isak	2025-05-30	150000.00	Internet
Hasbi	statis	rahya widowati	2025-05-21	200000.00	Internet
Hasbi	statis	devi suwandi	2025-05-20	200000.00	Internet + tv

Gambar 4.6. Form Transaksi

7. Form Dataset Model

Form Dataset Model dari Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dilihat pada Gambar 4.7.

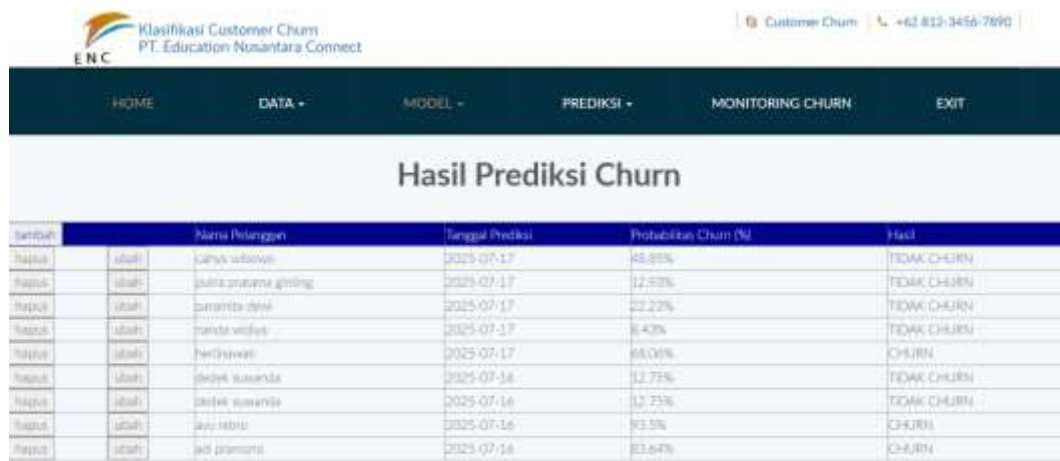


Tambahan	statis	Nama Pelanggan	Rata Login / Minggu	Total Transaksi (Rp)	Jumlah Layanan (hari)	Rating Transaksi	Churn?
Hasbi	statis	franda widya	7	3785365.03	321	5	TIKAL
Hasbi	statis	harva wahyudi	9	3124237.24	348	4	YA
Hasbi	statis	yusita pratika	20	3462076.06	233	2	TIDAK
Hasbi	statis	irisa purnama	5	797729.72	249	5	TIDAK
Hasbi	statis	hidat mahyadi	4	2830705.93	88	3	YA
Hasbi	statis	adi pramono	6	6196443.57	374	3	YA
Hasbi	statis	iswani	1	3783180.54	222	4	TIDAK
Hasbi	statis	rahya widowati	3	5054830.60	329	3	TIDAK
Hasbi	statis	sumbari	4	4885671.30	213	4	TIDAK
Hasbi	statis	poniman	2	5689058.04	370	3	TIDAK
Hasbi	statis	iswari	0	405534.32	278	4	YA

Gambar 4.7. Form Dataset Model

8. Form Hasil Prediksi

Form Hasil Prediksi dari Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dilihat pada Gambar 4.8.

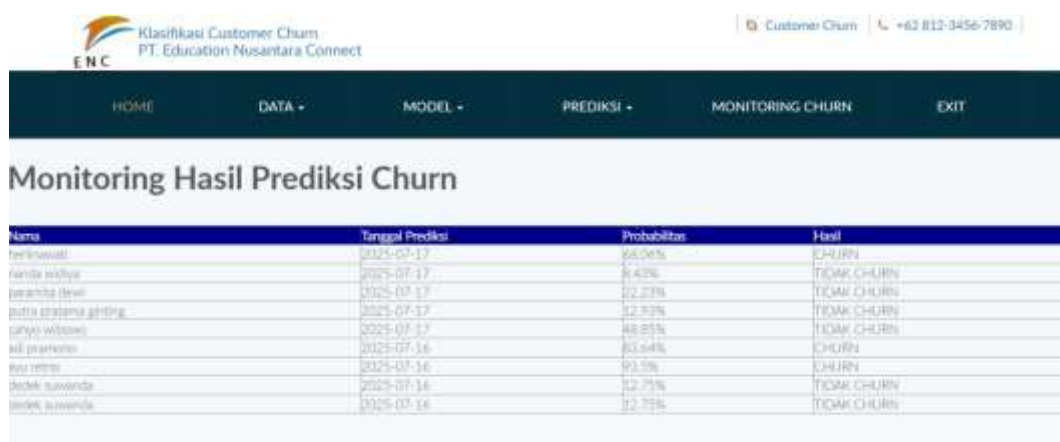


Jenis Kelamin	Status	Nama Pelanggan	Tanggal Prediksi	Probabilitas Churn (%)	Hasil
laki-laki	tidak	carly wibisono	2025-07-17	48.85%	TIDAK CHURN
laki-laki	tidak	putra pratama giring	2025-07-17	12.93%	TIDAK CHURN
laki-laki	tidak	parantha dewi	2025-07-17	22.23%	TIDAK CHURN
laki-laki	tidak	nanda wibisono	2025-07-17	8.43%	TIDAK CHURN
laki-laki	tidak	herlinawati	2025-07-17	66.06%	CHURN
laki-laki	tidak	dedek suwanda	2025-07-16	12.75%	TIDAK CHURN
laki-laki	tidak	dedek suwanda	2025-07-16	12.75%	TIDAK CHURN
laki-laki	tidak	ayu retno	2025-07-16	93.5%	CHURN
laki-laki	tidak	adi pramono	2025-07-16	83.64%	CHURN

Gambar 4.8. Form Hasil Prediksi

9. Form Monitoring Hasil Prediksi

Form Monitoring Hasil Prediksi dari Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dilihat pada Gambar 4.9.



Nama	Tanggal Prediksi	Probabilitas	Hasil
herlinawati	2025-07-17	66.06%	CHURN
nanda wibisono	2025-07-17	8.43%	TIDAK CHURN
parantha dewi	2025-07-17	22.23%	TIDAK CHURN
putra pratama giring	2025-07-17	12.93%	TIDAK CHURN
carly wibisono	2025-07-17	48.85%	TIDAK CHURN
adi pramono	2025-07-16	83.64%	CHURN
ayu retno	2025-07-16	93.5%	CHURN
dedek suwanda	2025-07-16	12.75%	TIDAK CHURN
dedek suwanda	2025-07-16	12.75%	TIDAK CHURN

Gambar 4.9. Form Monitoring Hasil Prediksi

4.2 Pembahasan

Pebahasan meliputi kebutuhan perangkat, hasil yang digunakan dan pengujian pada penelitian ini.

1. Kebutuhan Perangkat

Kebutuhan perangkat keras dan perangkat lunak untuk membuat aplikasi adalah sebagai berikut:

a. Satu unit laptop dengan spesifikasi sebagai berikut:

- 1) *Processor* minimal *Core 2 Duo*
- 2) RAM minimal 1 Gb
- 3) *Hardisk* minimal 80 Gb

b. Perangkat Lunak dengan spesifikasi sebagai berikut:

- 1) Sistem Operasi *Windows*
- 2) Notepad++
- 3) Appserv

2. Metode Random Forest

Penelitian ini menggunakan Metode Random Forest yang digunakan sebagai proses pencarian hasil. Berikut adalah tahapan Metode Random Forest dan data yang digunakan:

Tabel 4.1. Data Churn Pelanggan

Rata-Rata Login/Minggu	Total Transaksi	Lama Langganan	Rating Teakhir	Churn
4	2500000.00	180	4	YA

7	8744482.73	59	4	TIDAK
2	7892178.59	35	4	YA
4	4514710.02	200	5	TIDAK
4	1591007.63	130	5	YA
4	8530188.96	43	3	YA
1	8018466.97	104	2	YA
1	5011246.13	84	3	TIDAK
6	8587885.35	285	3	TIDAK
7	6639983.69	221	1	TIDAK
0	5729327.61	196	2	TIDAK
5	3382935.06	115	2	TIDAK
4	8829671.26	148	1	TIDAK
2	3522354.92	296	2	TIDAK
3	5906638.34	340	4	YA
5	3408593.27	86	2	YA
6	8222874.28	348	5	YA
6	8646042.80	334	4	YA
2	8530619.96	61	5	TIDAK
4	2417885.99	202	1	TIDAK
1	4280608.02	120	1	TIDAK
4	9042842.40	296	2	YA
1	845391.54	147	4	YA

0	2858201.01	244	4	TIDAK
7	7842239.12	70	5	YA
5	5363183.60	315	1	YA
4	4753663.72	203	3	YA
3	6821494.91	136	1	TIDAK
1	5963364.34	365	1	YA
5	7290132.31	200	5	TIDAK
2	8978796.52	185	3	TIDAK
1	8182338.89	160	4	YA
7	4706028.65	199	5	YA
3	7271719.08	362	1	YA
1	4159929.53	268	4	TIDAK
0	4638061.25	324	2	TIDAK
0	4540357.72	141	3	TIDAK
1	3380463.31	119	2	YA
2	1925894.82	189	4	YA
3	9949792.24	328	3	YA
2	8410928.57	200	5	TIDAK
5	6835956.18	55	4	YA
4	311913.47	180	1	TIDAK
0	2638777.03	208	3	TIDAK
6	426403.98	120	4	TIDAK

2	6730797.69	325	5	YA
4	6293603.37	266	1	TIDAK
4	9946771.90	120	2	YA
5	5686698.62	202	5	TIDAK
0	8830058.56	348	4	TIDAK
2	6668635.37	343	5	TIDAK
6	5829357.62	126	5	TIDAK
1	9433876.21	292	5	YA
1	885228.46	92	1	TIDAK
3	244948.48	274	4	YA
4	7172661.00	358	2	TIDAK
3	9562359.90	243	5	TIDAK
7	9580298.58	252	3	YA
5	6893320.59	268	3	TIDAK
2	8178902.63	326	4	YA
4	9759779.67	237	2	YA
7	4160220.02	79	4	YA
2	4897566.24	170	5	TIDAK
7	1969238.29	333	1	TIDAK
1	676339.40	356	4	TIDAK
3	5417809.93	312	3	TIDAK
2	5956902.60	211	2	TIDAK

1	9613972.71	256	1	YA
3	8706329.38	307	4	YA
1	954294.22	346	5	TIDAK
0	3734869.46	242	5	TIDAK
5	4984483.43	219	5	TIDAK
1	3577406.39	240	3	YA
2	866331.27	54	1	TIDAK
0	6776642.74	259	2	TIDAK
1	9579951.95	43	5	TIDAK
7	7840609.55	173	3	YA
0	9435734.87	199	3	YA
4	2696221.03	77	2	YA
7	5063423.65	81	2	YA
4	3220908.26	117	4	TIDAK
3	5033772.74	338	2	TIDAK
0	4555074.07	315	1	TIDAK
5	7988675.94	204	3	TIDAK
5	2625929.75	169	4	YA
1	9396265.75	66	3	TIDAK
6	2442987.15	200	5	YA
3	7860967.99	141	1	YA
1	8962107.65	106	5	YA

4	4121332.08	313	5	TIDAK
7	6110767.76	78	5	YA
5	9027197.91	325	1	TIDAK
3	6434733.70	144	5	TIDAK
3	5018956.78	363	2	YA
0	4609093.96	164	5	YA
5	3496074.72	156	5	YA
0	3147071.52	189	1	YA
1	3445960.45	147	3	TIDAK
6	6862948.76	125	1	YA
4	3119237.48	169	3	YA
4	2500000.00	180	4	TIDAK

1. Training Model

a. Ringkasan Data

Data Latih: 80 baris

Data Uji: 21 baris

Akurasi Model: 38.1%

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{4 + 4}{4 + 4 + 6 + 7} = \frac{8}{21} \approx 38.1\%$$

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{4}{4 + 6} = \frac{4}{10} = 0.40 \text{ (40\%)}$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{4}{4 + 7} = \frac{4}{11} \approx 0.3636 \text{ (36.36\%)}$$

$$F1 = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} = 2 \times \frac{(0.4 \times 0.3636)}{0.4 + 0.3636} \approx 2 \times \frac{0.1454}{0.7636} \approx 0.381 \text{ (38.1\%)}$$

b. Confusion Matrix

Tabel 4.2. Confusion Matrix

	Pred: Tidak	Pred: Churn
Aktual: Tidak	4	6
Aktual: Churn	7	4

c. Pengaruh setiap fitur (Feature importance)

- 1) **total_transaksi**: 37.81%
- 2) **lama_langganan_hari**: 30.85%
- 3) **rating_terakhir**: 15.79%
- 4) **rata_login_mingguan**: 15.55%

2. Data pelanggan yang diuji

Nama Pelanggan: Ahmad

- a. Rata-rata Login Mingguan: 6 kali/minggu

(Didapat langsung dari sistem login pelanggan)

- b. Total Transaksi: Rp 585.008,00,-,

(Hasil penjumlahan seluruh pembayaran transaksi pelanggan)

- c. Lama Langganan: 244 hari

(Selisih antara hari ini dan tanggal daftar pelanggan)

d. Rating Terakhir: 2 dari 5

(Rating terakhir yang diberikan pelanggan di sistem ulasan)

3. Hasil Prediksi Model:

Probabilitas Churn: 35%

Status: TIDAK CHURN

Penjelasan:

Model Machine Learning (Random Forest) memprediksi kemungkinan pelanggan akan **berlanjut** menggunakan layanan.

Jika Probabilitas > 0.5 , maka pelanggan dianggap CHURN.

4. Uji Coba Program

Uji coba terhadap sistem bertujuan untuk memastikan bahwa sistem sudah berada pada kondisi siap pakai. Instrumen yang digunakan untuk melakukan pengujian ini yaitu dengan menggunakan *Blackbox Testing*:

Tabel 4.3 Blackbox Testing Form Login

No	Form Login	Keterangan	Validitas
1.	Jika pengguna mengisi username dan password dengan benar kemudian melakukan Klik Tombol Submit	Aplikasi menampilkan form Home	Valid
2	Jika pengguna mengisi username dan password dengan salah kemudian melakukan Klik	Aplikasi menampilkan pesan kesalahan	Valid

	Tombol Submit		
--	---------------	--	--

Tabel 4.4 *Blackbox Testing Form Home*

No	Form Home	Keterangan	Validitas
1.	Klik Tombol Home	Aplikasi menampilkan form Home	Valid
2.	Klik Tombol Pelanggan	Aplikasi menampilkan form Pelanggan	Valid
3.	Klik Tombol Aktivitas Pelanggan	Aplikasi menampilkan form Aktivitas Pelanggan	Valid
4.	Klik Tombol Feedback Pelanggan	Aplikasi menampilkan form Feedback Pelanggan	Valid
5.	Klik Tombol Transaksi	Aplikasi menampilkan form Transaksi	Valid
6.	Klik Tombol Dataset Model	Aplikasi menampilkan	Valid

		form Dataset Model	
6.	Klik Tombol Hasil Prediksi	Aplikasi menampilkan form Hasil Prediksi	Valid
7.	Klik Tombol Monitoring Churn	Aplikasi menampilkan form Monitoring Churn	Valid

Tabel 4.5. Blackbox Testing Form Pelanggan

No	Form Pelanggan	Keterangan	Validitas
1.	Klik Tombol Simpan	Aplikasi menyimpan seluruh data di textbox ke dalam table database	Valid
2.	Klik Tombol Ubah	Aplikasi mengubah isi di table database sesuai data yang diubah	Valid

3.	Klik Tombol Hapus	Aplikasi menghapus isi data di database	
----	-------------------	---	--

Tabel 4.6. *Blackbox Testing Form* Aktivitas Pelanggan

No	Form Aktivitas Pelanggan	Keterangan	Validitas
1.	Klik Tombol Simpan	Aplikasi menyimpan seluruh data di textbox ke dalam table database	Valid
2.	Klik Tombol Ubah	Aplikasi mengubah isi di table database sesuai data yang diubah	Valid
3.	Klik Tombol Hapus	Aplikasi menghapus isi data di database	

Tabel 4.7. *Blackbox Testing Form* Feedback Pelanggan

No	Form Feedback Pelanggan	Keterangan	Validitas
1.	Klik Tombol Simpan	Aplikasi	Valid

		menyimpan seluruh data di textbox ke dalam table database	
2.	Klik Tombol Ubah	Aplikasi mengubah isi di table database sesuai data yang diubah	Valid
3.	Klik Tombol Hapus	Aplikasi menghapus isi data di database	

Tabel 4.8. *Blackbox Testing Form Transaksi*

No	Form Transaksi	Keterangan	Validitas
1.	Klik Tombol Simpan	Aplikasi menyimpan seluruh data di textbox ke dalam table database	Valid
2.	Klik Tombol Ubah	Aplikasi mengubah isi di table database	Valid

		sesuai data yang diubah	
3.	Klik Tombol Hapus	Aplikasi menghapus isi data di database	

Tabel 4.9. *Blackbox Testing Form Dataset Model*

No	Form Dataset Model	Keterangan	Validitas
1.	Klik Tombol Simpan	Aplikasi menyimpan seluruh data di textbox ke dalam table database	Valid
2.	Klik Tombol Ubah	Aplikasi mengubah isi di table database sesuai data yang diubah	Valid
3.	Klik Tombol Hapus	Aplikasi menghapus isi data di database	

Tabel 4.9. *Blackbox Testing Form Hasil Prediksi Churn*

No	Form Hasil Prediksi Churn	Keterangan	Validitas
1.	Klik Tombol Simpan	Aplikasi menyimpan seluruh data di textbox ke dalam table database	Valid
2.	Klik Tombol Ubah	Aplikasi mengubah isi di table database sesuai data yang diubah	Valid
3.	Klik Tombol Hapus	Aplikasi menghapus isi data di database	

Tabel 4.10. *Blackbox Testing Form Monitoring Churn*

No	Form Monitoring Churn	Keterangan	Validitas
1.	Klik Tombol Simpan	Aplikasi menyimpan seluruh data di textbox ke dalam table database	Valid

2.	Klik Tombol Ubah	Aplikasi mengubah isi di table database sesuai data yang diubah	Valid
3.	Klik Tombol Hapus	Aplikasi menghapus isi data di database	

4.3 Hasil Uji Coba

Setelah melakukan uji coba terhadap aplikasi, maka dapat disimpulkan hasil yang didapatkan yaitu:

1. *Interface* rancangan telah sesuai dengan *Interface* hasil.
2. Metode Random Forest telah diterapkan pada aplikasi yang dibuat.
3. *Interface* aplikasi bersifat *user friendly* sehingga pengguna dapat menggunakannya dengan mudah.
4. Aplikasi yang telah dibuat berjalan dengan baik.
5. Aplikasi yang telah dibuat tidak memiliki kesalahan logika.

4.4 Kekurangan Aplikasi

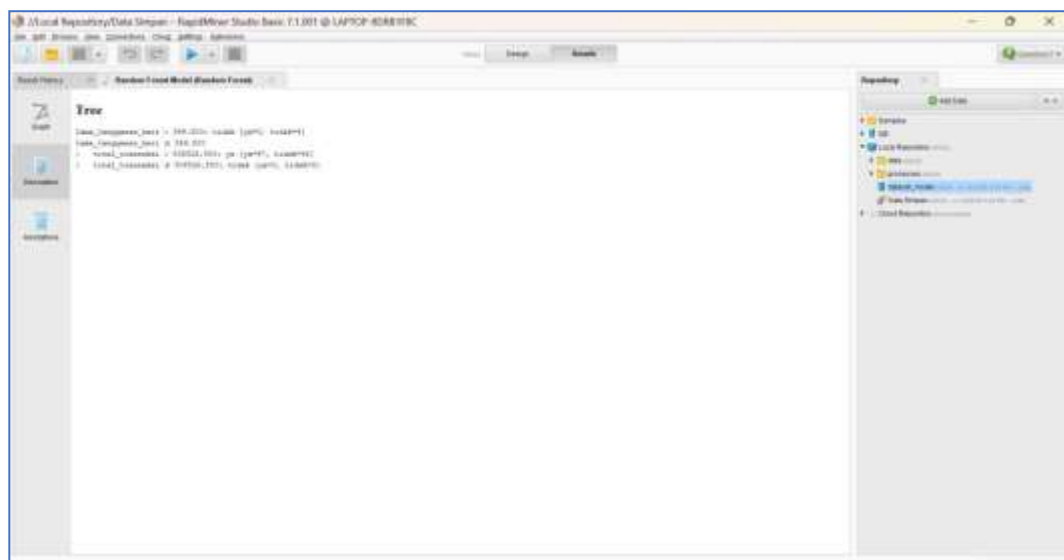
Kekurangan aplikasi pada penelitian ini diantaranya:

1. Aplikasi yang telah dibuat tidak memiliki petunjuk penggunaan.
2. Aplikasi yang telah dibuat terlalu banyak pilihan menu.
3. Aplikasi pada bagian admin tidak menggunakan pemrograman visual.

4.5 Uji Coba Rapid Miner



Gambar 4.10 Hasil Uji Coba Rapid Miner



Gambar 4.11 Deskripsi Hasil Uji Coba Rapid Miner

Tampilan Pohon Keputusan

Walaupun Random Forest sebenarnya terdiri dari banyak pohon, RapidMiner biasanya menampilkan salah satu pohon yang terbentuk untuk interpretasi.

- Atribut yang dipakai untuk split terlihat adalah "**lama_bergeser_bulan**".
- Dari pohon yang ditampilkan, decision rule-nya adalah:

- Jika **lama_bergeser_bulan** ≤ 346.54 , maka prediksi = "**tidak**" (tidak churn).
- Jika **lama_bergeser_bulan** > 346.54 , maka lanjut ke percabangan berikutnya.
- Split berikutnya menggunakan nilai **246.65**, kemudian menghasilkan prediksi "ya" (churn) atau "tidak".

Label Kelas

- "**ya**" → menunjukkan pelanggan churn.
- "**tidak**" → menunjukkan pelanggan tidak churn.
- Warna biru dan merah pada node menunjukkan distribusi data (misalnya berapa banyak data training yang jatuh pada kelas tersebut).

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan pada PT. Edducation Nusantara Connect mengenai penerapan algoritma Random Forest pada judul penelitian skripsi “ Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest “. Maka dapat disimpulkan penelitian tersebut sebagai berikut ;

1. Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest dapat dengan mudah mengklasifikasi pelanggan yang berpotensi churn (berhenti menggunakan layanan) sangat efektif karena kemampuannya dalam menangani data kompleks dan memberikan prediksi yang akurat.
2. Faktor yang mempengaruhi terjadinya customer churn pada penelitian ini adalah total transaksi, lama langganan hari, rating terakhir, rata login mingguan. Dengan menggunakan data Pelanggan, Aktivitas Pelanggan, dan Feedback Pelanggan, dan lain sebagainya maka dapat menerapkan metode Random Forest untuk Klasifikasi Customer Churn.
3. Berdasarkan data historis seperti rata-rata login per minggu, total transaksi, lama langganan, dan rating terakhir. Model dilatih dengan 80 data latih dan diuji pada 21 data uji. akurasi sebesar 38,1% dengan precision 40%, recall 36,36%, dan F1-score 38,1%.
4. Model Random Forest yang digunakan belum optimal karena seluruh metrik berada di bawah 50%. Hal ini menunjukkan model belum dapat

membedakan pelanggan churn dan tidak churn secara akurat. penyebabnya antara lain Jumlah data latih terlalu sedikit, Data tidak seimbang (jumlah churn vs tidak churn), Perlu hyperparameter tuning, Fitur yang digunakan belum cukup representative. Berdasarkan confusion matrix, model memprediksi churn dengan presisi yang kurang baik, menunjukkan bahwa model masih memiliki keterbatasan dalam membedakan pelanggan churn dan tidak churn secara akurat.

5.2. Saran

Berdasarkan hasil penelitian yang telah dilakukan dan evaluasi terhadap performa model Random Forest dalam mengklasifikasikan pelanggan berpotensi churn, maka penulis memberikan beberapa saran untuk pengembangan lebih lanjut, baik dari sisi teknis maupun metodologis dari judul penelitian skripsi yang berjudul “Klasifikasi Customer Churn Pada PT. Education Nusantara Connect Menggunakan Metode Random Forest” dapat dilihat sebagai berikut:

1. Penambahan Jumlah Data Latih. Dataset yang digunakan dalam penelitian ini masih terbatas, yaitu hanya terdiri dari 80 data latih dan 21 data uji. Untuk mendapatkan hasil prediksi yang lebih akurat dan generalisasi yang baik, disarankan untuk menggunakan jumlah data yang lebih besar. Semakin banyak data yang dimiliki, semakin baik algoritma dalam memahami pola dan variasi dalam perilaku pelanggan.
2. Tuning Parameter Model Random Forest. Penelitian ini belum mengoptimalkan hyperparameter Random Forest. Untuk hasil yang lebih optimal, disarankan melakukan tuning parameter seperti jumlah pohon

(`n_estimators`), kedalaman pohon (`max_depth`), dan minimal sampel split menggunakan teknik Grid Search atau Random Search.

3. Aplikasi yang telah dibuat sebaiknya memiliki petunjuk penggunaan.

DAFTAR PUSTAKA

- Adrian, M. R., Putra, M. P., Rafialdy, M. H., & Rakhmawati, N. A. (2021). Perbandingan Metode Klasifikasi Random Forest dan SVM Pada Analisis Sentimen PSBB. *Jurnal Informatika Upgris*, 7(1), 36–40. <https://doi.org/10.26877/jiu.v7i1.7099>
- Almumtazah, N., Azizah, N., Putri, Y. L., & Novitasari, D. C. R. (2021). Prediksi Jumlah Mahasiswa Baru Menggunakan Metode Regresi Linier Sederhana. *Jurnal Ilmiah Matematika Dan Terapan*, 18(1), 31–40. <https://doi.org/10.22487/2540766x.2021.v18.i1.15465>
- Amirulhaq Iskandar, M., & Latifa, U. (2023). Website Prediksi Customer Churn Untuk Mempertahankan Pelanggan Pada Perusahaan Telekomunikasi. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(2), 1308–1316. <https://doi.org/10.36040/jati.v7i2.6639>
- Ananda, F. D., & Pristyanto, Y. (2021). Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine. *MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 20(2), 407–416. <https://doi.org/10.30812/matrik.v20i2.1130>
- Chazar, C., & Erawan, B. (2020). Machine Learning Diagnosis Kanker Payudara Menggunakan Algoritma Support Vector Machine. *INFORMASI (Jurnal Informatika Dan Sistem Informasi)*, 12(1), 67–80. <https://doi.org/10.37424/informasi.v12i1.48>
- Ekayanti H., Makkasau, Fitriani., et all. (2023). *Metodologi Penelitian Kesehatan*.
- Elis, E., & Voutama, A. (2023). Pemanfaatan Uml (Unified Modeling Language) Dalam Perencanaan Sistem Penyewaan Baju Adat Berbasis Website. *INFORMATIKA*, 14(2), 26. <https://doi.org/10.36723/juri.v14i2.445>
- Fadila, W. A., Aini, Q., Wahyudi, F. A., & Fisika, P. (n.d.). *Perkembangan teknologi pemanfaatan fiber optik dalam industri telekomunikasi untuk koneksi jaringan*. 8(2), 309–320.
- Handie Pramana Putra, Sugiarto, Hendra Maulana, Evi Triandini, & Praja Firdaus Nuryananda. (2022). Relasional Desain Activy Diagram Sistem Informasi Agen Travel. *Prosiding Seminar Nasional Teknologi Dan Sistem Informasi*, 2(1), 238–241. <https://doi.org/10.33005/sitasi.v2i1.303>
- Nurhalizah, R. S., Ardianto, R., & Purwono, P. (2024). Analisis Supervised dan Unsupervised Learning pada Machine Learning: Systematic Literature Review. *Jurnal Ilmu Komputer Dan Informatika*, 4(1), 61–72. <https://doi.org/10.54082/jiki.168>
- Rahman, T., & Nurdin, H. (2020). Abdul Hamid No.77, RT.8/RW.4, Cawang, Kramat Jati, Jakarta Timur 13630 1 Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri. *Jl. Raya Jatiwaringin*, 5(1), 23.
- Rambe, P. P. (2021). Teori Atau Konsep Algoritma Pemrograman. *Ilmu Tarbiyah Dan Keguruan*, 1(1), 6–11.
- Setiyani, L. (2021). Desain Sistem : Use Case Diagram Pendahuluan. *Prosiding Seminar Nasional : Inovasi & Adopsi Teknologi 2021, September*, 246–260. <https://journal.uui.ac.id/AUTOMATA/article/view/19517>
- Sidik, A. D. W. M., Himawan Kusumah, I., Suryana, A., Edwinanto, Artiyasa, M.,

- & Pradiftha Junfithrana, A. (2020). Gambaran Umum Metode Klasifikasi Data Mining. *FIDELITY: Jurnal Teknik Elektro*, 2(2), 34–38. <https://doi.org/10.52005/fidelity.v2i2.111>
- Suci Amaliah, Nusrang, M., & Aswi, A. (2022). Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijiwa Bantaeng. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 4(3), 121–127. <https://doi.org/10.35580/variansiunm31>
- Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner. *JBASE - Journal of Business and Audit Information Systems*, 4(1), 13–21. <https://doi.org/10.30813/jbase.v4i1.2729>
- Suherman, Arsita, A. D., & Afriantoro, I. (2024). Analisis Pola Penjualan di Toko Retail Menggunakan Algoritma Apriori dengan Rapid Minner. *Jurnal Teknologi Informatika Dan Komputer*, 10(2), 734–753. <https://doi.org/10.37012/jtik.v10i2.2407>
- Zalukhu, A., Swingly, P., & Darma, D. (2023). Perangkat Lunak Aplikasi Pembelajaran Flowchart. *Jurnal Teknologi, Informasi Dan Industri*, 4(1), 61–70. <https://ejurnal.istp.ac.id/index.php/jtii/article/view/351>

Lampiran



MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi A Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 8819/KUBAN-PT/Akred/PT/06/2019
 Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6651003
 Email: (061)umsu.ac.id hty: 061@umsu.ac.id f: umsumedan i: umsumedan t: umsumedan u: umsumedan

**PENETAPAN DOSEN PEMBIMBING
 PROPOSAL/SKRIPSI MAHASISWA
 NOMOR : 290/IL3-AU/UMSU-09/F/2025**

Assalamu'alaikum Warahmatullahi Wabarakatuh

Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara, berdasarkan Persetujuan permohonan judul penelitian Proposal / Skripsi dari Ketua / Sekretaris.

Program Studi : Sistem Informasi
Pada tanggal : 05 Februari 2025

Dengan ini menetapkan Dosen Pembimbing Proposal / Skripsi Mahasiswa.

Nama : Anisa Julia Rahmayanti
NPM : 2109010079
Semester : VII (Tujuh)
Program studi : Sistem Informasi
Judul Proposal / Skripsi : Klasifikasi Customer Churn pada PT Education Nusantara Connect Menggunakan Metode Random Forest

Dosen Pembimbing : Mhd. Basri, S.Si. M.Kom

Dengan demikian di izinkan menulis Proposal / Skripsi dengan ketentuan

1. Penulisan berpedoman pada buku panduan penulisan Proposal / Skripsi Fakultas Ilmu Komputer dan Teknologi Informasi UMSU
2. Pelaksanaan Sidang Skripsi harus berjarak 3 bulan setelah dikeluarkannya Surat Penetapan Dosen Pembimbing Skripsi.
3. **Proyek Proposal / Skripsi dinyatakan " BATAL " bila tidak selesai sebelum Masa Kadaluarsa tanggal : 05 Februari 2026**
4. Revisi judul.....

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Ditetapkan di : Medan

Pada Tanggal : 06 Sya'ban 1446 H

05 Februari 2025 M

a.n.Dekan
 Wakil Dekan I



Haidar Maulana, S.T., M.Kom.
 NIDN : 0121119102







Cc. File

**PT. EDUKASI NUSANTARA CONNECT**

Jl. Lintas Sumatera Cikampak - Ds. Cikampak 1A-kef. Aek Batu
Kec. Tonggamba-Labuhan Batu Selatan-Sumatera Utara - Phone :061361229981

SURAT KETERANGAN SELESAI PENELITIAN

Yang bertanda tangan dibawah ini :

Nama : JURI
Jabatan : Direktur Utama
Unit Kerja : PT. Edukasi Nusantara Connect

Dengan ini menerangkan bahwa mahasiswa berikut :

Nama : ANISA JULIA RAHMAYANTI
NIM : 2109010079
Prodi : Sistem Informasi
Asal PT : Universitas Muhammadiyah Sumatera Utara

Telah melaksanakan penelitian untuk memenuhi tugas skripsi, terhitung tanggal 27 Januari – 7 juni 2025 untuk memperoleh data dalam rangka penyusunan skripsi yang berjudul "Klasifikasi Customer Churn Pada PT.Edukasi Nusantara Connect Menggunakan Metode Random Forest".

Demikian surat ini dibuat untuk dipergunakan sebagaimana mestinya.

Cikampak , 7 juni 2025

Direktur Utama

JURI



UMSU

Stepad | Cerdas | Berprestasi

MAJLIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi A Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 88/TK/BDAN-PT/Akred/PT/182019

Pusat Administrasi: Jalan Mukhter Basri No. 3 Medan 20236 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003

Mail: info@umsu.ac.id

Telp: 662@umsu.ac.id

umsu.medan

umsu.medan

umsu.medan

umsu.medan

Berita Acara Pembimbingan Skripsi

Nama Mahasiswa : Anisa Julia Rahmayanti
NPM : 21090100
Nama Dosen Pembimbing : Mhd. Basri, S.Kom., M.Kom
Program Studi : Sistem Informasi
Konsentrasi :
Judul Penelitian : Klasifikasi Customer Churn Pada PT. Education Nusantara
Connect Menggunakan Metode Random Forest

Item	Hasil Evaluasi	Tanggal	Paraf Dosen
	Bab 1 - Latar Belakang Kurang Komplex - Tujuan masalah ditulis dengan benar	12 Mei 2025	[Signature]
	Bab 2 - Perbaikan Penulisan 1 Paragraf	22 Mei 2025	[Signature]
	- Flowchart diganti dengan diagram alir	28 Mei 2025	[Signature]
	- Diagram alir penulisan kurang sesuai dengan format yang ditentukan	3 Juni 2025	[Signature]
	- Gantikan diagram masih kurang sesuai dengan alir penulisan	6 Juni 2025	[Signature]
	- Tabel Studi Literatur bagian akhir tidak lengkap	9 Juni 2025	[Signature]
	- Input data transaksi ke dalam aplikasi - Promotivitas hasil belum konsisten	14 Juni 2025	[Signature]
	ACC Berhang	16 Juni 2025	[Signature]

Medan, 16 Juli 2025

Diketahui oleh :

Ketua Program Studi
Sistem Informasi

(Martiano S.Pd, S.Kom., M.Kom)

Disetujui oleh :

Dosen Pembimbing

[Signature]

(Mhd. Basri, S.Kom., M.Kom)





UMSU
Unggul | Cerdas | Berprestasi

MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PEMIPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi A Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 88/SK/BAN-PT/Akred/PT/001/2019
Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224587 Fax. (061) 6625474 - 6631093
<https://id.umsu.ac.id> info@umsu.ac.id [umsuamedan](https://www.facebook.com/umsuamedan) [umsuamedan](https://www.instagram.com/umsuamedan) [umsuamedan](https://www.youtube.com/umsuamedan) [umsuamedan](https://www.tiktok.com/umsuamedan)

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ
BERITA ACARA SEMINAR PROPOSAL
TAHUN AJARAN 2024/2025

Hari/Tanggal: SUMBER, 13-05-2025

Nama Mahasiswa: ANISA JULIA RAHMAYANTI
NPM: 2109010079
Program Studi: SISTEM INFORMASI
Nama Dosen Penanggap: YOKHANI RAHMAT, S.P., M.Kom
Judul Proposal: KLASIFIKASI CUSTOMER CHURN PADA PT. BENCOTAN
NUBUNETA CONNECT MELAYUKAN MATEMATIKA
PORDOM FOREST.

Materi/Point yang Diperbaiki: 1. Perbaikan Rumus dan Materi
2. Perbaikan Rumus Perhitungan
3. Perbaikan Rumus Logika dan Rumus
4. Ti Logika Rumus Matematika 19.

Dosen Penanggap

(Yokhani Rahmat, S.P., M.Kom)

Mahasiswa

(Anisa Julia Rahmayanti)



UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
Nomor : 662/IL3-AU/UMSU-09/E/2025

UNDANGAN SEMINAR PROPOSAL

Fakultas : Ilmu Komputer dan Teknologi Informasi
 Program Studi : Sistem Informasi
 Hari/Tanggal : Juni'at, 13 Juni 2025
 Waktu /Tempat : 09.00/G
 Pemimpin Seminar : **Martiano,S.Pd., S.Kom., M.Kom**

No.	SPN	SENAMUAMADIASISWA	Dosen Pembimbing	Dokter Pembimbing	Klasifikasi	Customer Churn pada PT Education Nusantara Connect menggunakan metode Random Forest	Klasifikasi Customer Inten Dalam Mengetahui Tingkat Kepuasan Pelanggan Pada PT Education Nusantara Connect Dengan Menggunakan Metode Exboost	Analisis Kepuasan Pengguna Terhadap Kualitas Layanan Aplikasi Livin By Mandiri Menggunakan Metode E-Servqual Dan Importance Performance Analysis (IPA)
9	2109010079	Anisa Julia Rahmasyanti	Mhd. Basri, S.Si. M.Kom	Yohani Syahra, S.Si. M.Kom	Martiano, S. Pd. S. Kom, M. Kom.			
10	2109010081	Nurul Amanda	Rizalddy Khair, M.Kom					
11	2109010068	Anis Badriah Harahap	Mhd. Basri, S.Si. M.Kom	Halim Maulana, S.T., M.Kom.				



NB: - Laki-laki berbukuca hitam putih dan memakai dari
 Perempuan berbukuca muslimah hitam putih

Medan, 15 Dzulhijjah 1446 H
 11 Juni 2025 M



Dekan
 Dr. H. Djalal Sholahudin, M.Kom
 NIDN : 127059201





UMSU UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

KARTU KEHADIRAN AUDIENCE SEMINAR PROPOSAL SKRIPSI

NAMA : ANISA JULIA RAHMAYANTI
NPM : 2109010079
PRODI : SISTEM INFORMASI

No	Jadwal Seminar	Pemateri	Judul Seminar Proposal	Paraf Ketua Sidang
1	13 Juni 2024	Nurul Azzahra	Klasifikasi Customer Intern dalam meng tahu tingkat kepuasan pelanggan pada PT. Education Nusantara Connect menggunakan Metode ESBBoost	
2	— " —	Andi Rahmat Lili	Analisis Kepuasan pengguna terhadap tampilan layanan Aplikasi Ulin Ey Maslin menggunakan Metode E-Survey dan dengan Importance Performance Analysis (IPA)	
3	— " —	Aditya Asyraf Firda	Analisis persediaan obat dengan menggunakan FIFO dan Average di Klinik Asrama Arinmut Binjai	

catatan :

1. Kartu ini di bawa pada saat pelaksanaan seminar proposal dan disahkan oleh ketua sidang seminar proposal skripsi
2. Kehadiran dan keaktifan minimal 3 kali
3. Kartu ini syarat untuk mengikuti Seminar Proposal

