

**ANALISIS DAN IMPLEMENTASI ALGORITMA LOGISTIK
REGRESI DALAM MENGLASIFIKASI CALON
KONSUMEN BARBERSHOP DI KOTA MEDAN SEBAGAI
IDE PENGEMBANGAN BISNIS APLIKASI PANGKASIN**

SKRIPSI

DISUSUN OLEH

FARHAN AL IKHSAN

2109010114



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

**ANALISIS DAN IMPLEMENTASI ALGORITMA LOGISTIK REGRESI
DALAM MENGLASIFIKASI CALON KONSUMEN BARBERSHOP DI
KOTA MEDAN SEBAGAI IDE PENGEMBANGAN BISNIS APLIKASI
PANGKASIN**

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**

FARHAN AL IKHSAN

2109010114

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

LEMBAR PENGESAHAN

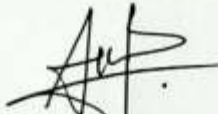
Judul Skripsi : ANALISIS DAN IMPLEMENTASI ALGORITMA
LOGISTIK REGRESI DALAM MENGLASIFIKASI
CALON KONSUMEN BARBERSHOP DI KOTA
MEDAN SEBAGAI IDE PENGEMBANGAN BISNIS
APLIKASI PANGKASIN

Nama Mahasiswa : FARHAN AL IKHSAN

NPM : 2109010114

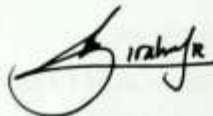
Program Studi : SISTEM INFORMASI

Menyetujui
Komisi Pembimbing



(Halim Maulana, S.T., M.Kom., MTA)
NIDN.

Ketua Program Studi



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

Dekan



(Dr. Al Khawarizmi, S.Kom., M.Kom.)
NIDN. 0127099201

PERNYATAAN ORISINALITAS

ANALISIS DAN IMPLEMENTASI ALGORITMA LOGISTIK REGRESI DALAM MENGLASIFIKASI CALON KONSUMEN BARBERSHOP DI KOTA MEDAN SEBAGAI IDE PENGEMBANGAN BISNIS APLIKASI PANGKASIN

SKRIPSI

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

Medan, 7 Oktober 2025

Yang membuat pernyataan



FARHAN AL IKHSAN

NPM. 2109010114

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

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

Nama	: FARHAN AL IKHSAN
NPM	: 2109010114
Program Studi	: SISTEM INFORMASI
Karya Ilmiah	: Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bedas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:

**ANALISIS DAN IMPLEMENTASI ALGORITMA LOGISTIK REGRESI
DALAM MENGLASIFIKASI CALON KONSUMEN BARBERSHOP DI
KOTA MEDAN SEBAGAI IDE PENGEMBANGAN BISNIS APLIKASI
PANGKASIN**

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, 7 Oktober 2025

Yang membuat pernyataan


FARHAN AL IKHSAN
NPM. 2109010114

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : FARHAN AL IKHSAN
Tempat dan Tanggal Lahir : Jakarta, 20 Juni 2001
Alamat Rumah : Jl. Bengkalis Komplek Macan Tutul No
AL 19, Medan Belawan, Kota Medan, Sumatera Utara
Telepon/Faks/HP : 082273371660
E-mail : farhansan47@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD	: 2007	TAMAT: 2013
SMP	: 2013	TAMAT: 2016
SMA	: 2016	TAMAT: 2019

KATA PENGANTAR



Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, karunia, serta petunjuk-Nya sehingga penulis dapat menyelesaikan tugas akhir yaitu penyusunan skripsi yang berjudul **“Analisis dan Implementasi Algoritma Logistik Regresi dalam Mengklasifikasikan Calon Konsumen Barbershop di Kota Medan sebagai Ide Pengembangan Bisnis Aplikasi Pangkasin”** dengan baik.

Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan studi pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara. Dalam penyusunan skripsi ini, penulis melakukan analisis terhadap metode Logistik Regresi yang diimplementasikan untuk memahami potensi calon konsumen barbershop di Kota Medan, yang menjadi dasar dari ide pengembangan bisnis digital dalam bentuk aplikasi bernama Pangkasin.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi perbaikan di masa mendatang. Segala hambatan yang penulis hadapi selama proses penelitian dan penulisan ini akhirnya dapat teratasi berkat dukungan dari banyak pihak yang telah dengan ikhlas membantu.

Pada kesempatan ini, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada pihak-pihak yang telah memberikan dukungan moril, bimbingan, serta motivasi, antara lain:

1. Bapak Prof. Dr. Agussani, M.AP., selaku Rektor Universitas Muhammadiyah Sumatera Utara.
2. Kedua orang tua tercinta penulis yang selalu mendoakan dan memberikan dukungan yang tak ternilai sepanjang perjalanan akademik penulis, Ibu Salbiah Yus dan Ayahanda Elfanda.
3. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom., selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi.

4. Bapak Halim Maulana, S.T., M.Kom., MTA., selaku Wakil Dekan I FIKTI, Serta selaku dosen pembimbing yang dengan baik memberikan arahan dan bimbingan selama proses penulisan skripsi ini.
5. Bapak Lutfi Basit, S.Sos., M.I.Kom., selaku Wakil Dekan III FIKTI.
6. Bapak Martiano, S.Pd., S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi.
7. Ibu Yoshida Sary, S.E., S.Kom., M.Kom., selaku Sekretaris Program Studi Sistem Informasi.
8. Rekan dan sahabat yang telah memberi semangat dan menemani proses penulis dalam menyelesaikan skripsi ini.
9. Partner terbaik dalam pengembangan startup Pangkasin, Farhan Febryan, atas kerja sama, semangat, dan kontribusi dalam merealisasikan ide bisnis digital ini.

Penulis berharap karya ini dapat memberikan manfaat, khususnya bagi penulis sebagai bentuk pengaplikasian ilmu selama menjalani studi, serta bagi pembaca yang ingin memahami potensi teknologi dalam pengembangan bisnis barbershop berbasis data.

Akhir kata, penulis memohon ampun atas segala kekurangan dalam penyusunan skripsi ini, dan semoga Allah SWT senantiasa membalas semua kebaikan dari pihak-pihak yang telah membantu.

Medan, 7 Oktober 2025

Penulis



(Farhan Al Ikhsan)

ANALISIS DAN IMPLEMENTASI ALGORITMA LOGISTIK REGRESI DALAM MENGLASIFIKASI CALON KONSUMEN BARBERSHOP DI KOTA MEDAN SEBAGAI IDE PENGEMBANGAN BISNIS APLIKASI PANGKASIN

ABSTRAK

Perkembangan teknologi digital mendorong transformasi di berbagai sektor, termasuk jasa pangkas rambut (barbershop). Startup Pangkasin hadir sebagai aplikasi yang menawarkan layanan reservasi dan pangkas rambut berbasis digital. Namun, tantangan muncul dalam mengidentifikasi konsumen potensial yang berminat menggunakan aplikasi ini. Penelitian ini bertujuan mengimplementasikan algoritma Logistik Regresi untuk mengklasifikasikan calon konsumen barbershop di Kota Medan berdasarkan data observasi yang meliputi umur, jenis kelamin, dan frekuensi cukur rambut. Penelitian ini menggunakan metode observasi langsung di delapan barbershop serta analisis data dengan Python. Data dibagi menjadi data latih (training) dan data uji (testing) dengan perbandingan 70:30. Hasil penelitian menunjukkan model Logistik Regresi yang dikembangkan mampu mengklasifikasikan konsumen potensial dengan tingkat akurasi 80%, recall 91,8%, presisi 85%, dan F1-Score 88%. Karakteristik pengguna potensial aplikasi Pangkasin cenderung berumur 20–35 tahun, pria, dan memiliki frekuensi cukur 2–3 kali per bulan. Hasil penelitian ini diharapkan dapat menjadi acuan dalam pengembangan strategi bisnis digital pada sektor jasa pangkas rambut.

Kata Kunci: Logistik Regresi, Klasifikasi, Barbershop, Pangkasin, Machine Learning, Python.

ANALYSIS AND IMPLEMENTATION OF LOGISTIC REGRESSION
ALGORITHM IN CLASSIFYING PROSPECTIVE BARBERSHOP CONSUMERS
IN MEDAN CITY AS A BUSINESS DEVELOPMENT IDEA FOR PANGKASIN
APPLICATION

ABSTRACT

The development of digital technology drives transformation in various sectors, including barbershop services. The Pangkasin startup is present as an application that offers digital-based hairdressing and reservation services. However, challenges arise in identifying potential consumers who are interested in using this application. This study aims to implement the Logistic Regression algorithm to classify barbershop consumers in Medan City based on observational data including age, prospective gender, and frequency of shaving. This study uses direct observation methods in eight barbershops and data analysis with Python. The data is divided into training data and testing data with a ratio of 70:30. The results of the study show that the developed Logistic Regression model is able to classify potential consumers with an accuracy level of 80%, recall 91,8%, precision 85%, and F1-Score 88%. The characteristics of potential users of the Pangkasin application tend to be 20–35 years old, male, and have a shaving frequency of 2–3 times per month. The results of this study are expected to be a reference in developing digital business strategies in the barbershop service sector.

Keywords: Logistic Regression, Classification, Consumer, Barbershop, Pangkasin, Machine Learning, Python.

DAFTAR ISI

HALAMAN SAMPUL.....	i
LEMBAR PENGESAHAN	Error! Bookmark not defined.
PERNYATAAN ORISINALITAS.....	Error! Bookmark not defined.
PERNYATAAN PERSETUJUAN PUBLIKASI...Error! Bookmark not defined.	Error! Bookmark not defined.
KARYA ILMIAH UNTUK KEPENTINGAN	Error! Bookmark not defined.
AKADEMIS.....	Error! Bookmark not defined.
RIWAYAT HIDUP	iii
KATA PENGANTAR.....	v
ABSTRAK	i
ABSTRACT	ii
DAFTAR ISI.....	iii
DAFTAR TABEL.....	iv
DAFTAR GAMBAR.....	v
DAFTAR LAMPIRAN	viii
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
BAB II	7
LANDASAN TEORI.....	7
2.1 Regresi Logistik	7
2.1.1 Logistik Regresi Biner.....	8
2.2 Klasifikasi	9
2.3 Teknik Analisis	11
2.4 Implementasi.....	11
2.5 Customer Relationship Management (CRM)	12
2.6 Aplikasi.....	12
2.7 Python	13
2.8 Penelitian Terdahulu	16
2.9 Kerangka Berpikir Konseptual	17
BAB III.....	19
METODOLOGI PENELITIAN	19

3.1 Lingkungan yang di pakai untuk mengembangkan Program	19
3.1.1 Perangkat Lunak yang di pakai	19
3.1.2 Sistem Operasi yang di pakai	20
3.1.3 Jenis Laptop yang digunakan	20
3.2 Strategi Pemecahan Masalah	20
3.3 Data Observasi	22
3.3.1 The Man Barbershop	26
3.3.2 Trees Barbershop	26
3.3.3 D'colonel's Barbershop	27
3.3.4 Premium Barbershop ZIE & ZKA	28
3.3.5 Barbershop	29
3.3.6 Ivon Pangkas	29
3.3.7 Mr. Klimis Marelان	30
3.3.8 Aa Barbershop	31
3.4 Jadwal Penelitian	32
3.5 Flowchart Penelitian	33
BAB IV	37
HASIL DAN PEMBAHASAN	37
4.1 Perhitungan Algoritma Logistik Regresi	37
4.1.1 Fungsi Linier	37
4.1.2 Fungsi Sigmoid	38
4.1.3 Data Bobot	38
4.2 Implementasi Algoritma	40
4.2.1 Code data training dan Hasil	41
4.2.2 Code data testing dan Hasil	44
4.2.3 Output code	51
4.3 Data Observasi	54
4.3.1 Code Exploratory Data Analysis	55
4.3.2 Output code Exploratory Data Analysis	58
4.4 Hasil Evaluasi Model	61
BAB V	64
KESIMPULAN DAN SARAN	64
5.1 Kesimpulan	64
5.2 Saran	64
DAFTAR PUSTAKA	66
LAMPIRAN	68

DAFTAR TABEL

Tabel 3. 1 Variabel Observasi pada Konsumen di Barbershop	21
Tabel 3. 2 Data Observasi Pada Barbershop	24
Tabel 3. 3 Tabel Jadwal Kegiatan	32
Tabel 4. 1 Tabel Kesimpulan Evaluasi Model	63

DAFTAR GAMBAR

Gambar 2. 1 Persamaan utama Regresi Logistik fungsi logit	7
---	---

Gambar 2. 2Model Regresi Logistik Biner	9
Gambar 2. 3Model Regresi Logistik Biner variabel bebas	9
Gambar 2. 4 Kerangka Berpikir Konseptual.....	17
Gambar 3. 1 Flowchart atau Diagram alur Penelitian.....	33
Gambar 4. 1 Rumus Fungsi Linier	37
Gambar 4. 2 Rumus Fungsi Sigmoid	38
Gambar 4. 3 Pseudocode perhitungan bias dan bobot fitur	41
Gambar 4. 4 Output bias dan bobot	43
Gambar 4. 5 Pseudocode 1.....	44
Gambar 4. 6 Pseudocode 2.....	46
Gambar 4. 7 Pseudocode 3.....	47
Gambar 4. 8 Pseudocode 4.....	48
Gambar 4. 9 Pseudocode 5.....	49
Gambar 4. 10 Pseudocode 6.....	50
Gambar 4. 11 EDA Metriks Performa Model.....	51
Gambar 4. 12 EDA Prediksi Data Baru	52
Gambar 4. 13 EDA Jumlah aktual per kelas	52
Gambar 4. 14 Rumus Precision	53
Gambar 4. 15 Rumus Recall	53
Gambar 4. 16 Rumus F1-Score.....	53
Gambar 4. 17 Grafik Confusion Matrix.....	54
Gambar 4. 18 Gambar Code Exploratory Data Analysis	55
Gambar 4. 19 Pseudocode Exploratory Data Analysis	57
Gambar 4. 20 Komposisi Umur	59

Gambar 4. 21 Distribusi Umur.....	60
Gambar 4. 22 Distribusi Frekuensi Cukur	61

DAFTAR LAMPIRAN

Lampiran 1 Data Set Excel 149 Data	68
Lampiran 2 Code training Penentuan Bias dan bobot.....	71
Lampiran 3 Out put dari code penentuan Bias dan Bobot	71
Lampiran 4 code program data testing Algoritma Logistik Regresi.....	72
Lampiran 5 code program data testing Algoritma Logistik Refresi 2	72
Lampiran 6 code program data testing Algoritma Logistik Regresi 3.....	73
Lampiran 7 code program data testing Algoritma Logistik Regresi 4.....	73
Lampiran 8 Output code data testing Algoritma Logistik Regresi	74
Lampiran 9 Code Program EDA (Exploratory Data Analysis)	74
Lampiran 10 Code Program EDA (Exploratory Data Analysis) 2	75
Lampiran 11 Output code EDA (Exploratory Data Analysis).....	75
Lampiran 12 narasumber bang Niko selaku Hairstylist di barbershop “The Man Barbershop”	76
Lampiran 13 Foto Observasi di Barbershop “Trees Barbershop” bersama narasumber Kak Zura, beliau tidak mau masuk kamera, peneliti foto dari atas ...	77
Lampiran 14 Foto Observasi bersama narasumber Bang Putra di Barbershop “D’COLONEL’S Barbershop”	77
Lampiran 15 Foto dengan Hairstylist Barbershop “Premium Barbershop Zie & Zka”	78
Lampiran 16 Foto dengan Hairstylist Pak Saprin di Barbershop “Barbershop 168”	78
Lampiran 17 Foto dengan Hairstylist Barbershop “Ivon Pangkasi” Nahazisokhi Buulolo dengan mengajukan pertanyaan observasi	79

Lampiran 18 Foto dengan Hairstylist Barbershop “Mr.Klimis” di Marelan bang Ilham dengan mengajukan pertanyaan.....	80
Lampiran 19 Foto dengan Hairstylist bang Aa di Barbershop “Aa Barbershop” di belawan dengan mengajukan pertanyaan.....	80

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan zaman dan teknologi yang pesat saat ini telah membawa perubahan yang signifikan dalam berbagai aspek kehidupan manusia, termasuk sektor jasa. Salah satu industri yang masih menghadapi tantangan dalam mengadopsi digitalisasi adalah gaya hidup, terutama jasa cukur rambut atau barbershop. Saat ini, sebagian besar barbershop masih mengandalkan sistem antrian manual tradisional, yang sering kali mengakibatkan waktu tunggu yang tidak pasti bagi pelanggan. Hal ini dapat menyebabkan ketidaknyamanan, penurunan tingkat kepuasan pelanggan, dan berkurangnya loyalitas terhadap layanan yang diberikan.

Menurut data dari jurnal "Pengaruh Kemudahan Penggunaan, Kepercayaan Konsumen & Kreativitas Iklan terhadap Minat Beli Konsumen Online-Shop" (Khotimah & Febriansyah, 2018), pengguna internet di Indonesia hingga Juni 2016 mencapai 88 juta jiwa dari total populasi 258 juta jiwa, dengan penetrasi sekitar 34%. Pada periode 2019–2020, menurut (Gunawan et al., 2021) penetrasi pengguna internet meningkat menjadi 73.7%. Hal ini menunjukkan peluang besar dalam penerapan teknologi digital di berbagai sektor, termasuk jasa pangkas rambut.

Melihat peluang ini, Startup Pangkasin hadir sebagai aplikasi jasa cukur rambut yang menawarkan layanan hairstylist ke rumah pelanggan atau penjadwalan di barbershop mitra. Namun, salah satu tantangan yang dihadapi oleh tim Pangkasin adalah kesulitan dalam mengidentifikasi dan mengelompokkan calon pengguna aplikasi potensial. Hingga saat ini, belum ada sistem yang mampu

mengklasifikasikan calon pengguna secara otomatis berdasarkan data karakteristik mereka.

Sebagai solusi, penelitian ini mengusulkan penggunaan algoritma Logistik Regresi untuk mengklasifikasikan calon pelanggan berdasarkan pola perilaku dan karakteristik layanan. Logistik Regresi adalah metode klasifikasi biner dalam machine learning yang digunakan untuk memprediksi probabilitas suatu entitas masuk ke dalam salah satu dari dua kategori (misalnya: potensial atau tidak potensial). Metode ini bekerja dengan mengubah kombinasi linear dari variabel input menjadi nilai probabilitas menggunakan fungsi sigmoid. Fungsi sigmoid mengkonversi nilai numerik menjadi rentang antara 0 hingga 1, memungkinkan untuk menginterpretasikan hasil model sebagai peluang seorang pelanggan termasuk dalam kategori tertentu.

Logistik Regresi sangat sesuai digunakan dalam penelitian ini karena mampu memetakan hubungan antar variabel seperti umur, jenis kelamin, dan frekuensi kunjungan dengan hasil prediksi berupa klasifikasi pelanggan. Dalam konteks penelitian ini, klasifikasi yang dilakukan terdiri dari dua kategori, yaitu pelanggan potensial dan pelanggan tidak potensial. Dengan menggunakan model ini, sistem dapat membantu tim Pangkasin untuk mengarahkan strategi promosi dan layanan secara lebih efektif.

Berdasarkan penelitian oleh (Tran et al., 2023) dalam jurnal "Customer Churn Prediction In The Banking Sector Using Machine Learning-Based Classification Models", Logistik Regresi mampu mencapai akurasi antara 86,26% hingga 87,57% dalam mengklasifikasikan pelanggan berdasarkan interaksi digital

mereka. Hal ini menunjukkan potensi Logistik Regresi untuk digunakan dalam skenario sektor jasa seperti barbershop.

Langkah-langkah yang akan dilakukan dalam penelitian ini meliputi:

1. Pengumpulan Data: Menggunakan data hasil observasi dari beberapa barbershop di Kota Medan yang relevan dengan pola penggunaan layanan pelanggan di barbershop.
2. Preprocessing Data: Melakukan pembersihan serta transformasi data agar dapat dianalisis dengan lebih akurat.
3. Implementasi Algoritma Logistic Regression: Membangun model klasifikasi pelanggan berdasarkan data yang telah diproses.
4. Evaluasi Model: Menggunakan metrik seperti F1 Score, Precision, dan Recall untuk mengukur kinerja model.
5. Analisis Hasil dan Penerapan: Memberikan rekomendasi strategis bagi Tim Pangkasin untuk tau strategi penyesuain ke pasar Cukur Rambut.

Hasil yang diharapkan dari penelitian ini adalah pengembangan sistem prediksi pelanggan yang dapat membantu tim Pangkasin memahami kebutuhan calon pelanggan serta cara meningkatkan loyalitas mereka. Dengan adanya sistem klasifikasi berbasis data, Pangkasin diharapkan dapat mengalami peningkatan retensi pelanggan ketika mulai beroperasi. Selain itu, penelitian ini juga berkontribusi dalam pengembangan teknologi bagi sektor jasa di Indonesia. Dengan memanfaatkan sistem berbasis kecerdasan buatan, industri pangkas rambut dapat beradaptasi dengan tren digitalisasi dan memberikan layanan yang lebih optimal bagi pelanggan.

1.2 Rumusan Masalah

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

1. Bagaimana implementasi algoritma Logistic Regression dalam mengklasifikasikan calon pelanggan pada aplikasi Pangkasin?
2. Seberapa akurat algoritma Logistic Regression dalam melakukan prediksi pelanggan berdasarkan data yang tersedia?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah Ruang lingkup data berasal dari hasil observasi langsung di beberapa Barbershop di Kota Medan. Fokus penelitian ini hanya pada klasifikasi Logistik Regresi pada data dari setiap barbershop untuk menjadi masukan pada Tim Pangkasin pada penerapan marketing dan strategi bisnis.

Untuk Fokus dan Keterbatasan ruang lingkup Penelitian, Penelitian ini dibatasi hanya pada beberapa aspek sebagai berikut:

1. Pengumpulan Data

Data yang digunakan dalam penelitian diperoleh melalui observasi langsung pada beberapa barbershop di sekitaran Kota Medan. Data yang di ambil berdasarkan pelanggan mencakup umur, jenis kelamin, frekuensi kunjungan, dan minat terhadap layanan digital.

2. Pra-pemrosesan Data

Data yang di kumpulkan akan masuk ke tahap pembersihan dan transformasi, agar siap di proses pada pemodelan logistik regresi. Proses

ini termasuk perubahan pada data Kategorikal menjadi Numerik dan penyesuaian format data.

3. Implementasi Algoritma

Penelitian ini terbatas pada penggunaan Algoritma Logistik Regresi sebagai klasifikasi utama. Algoritma ini dipilih karena sesuai untuk kasus klasifikasi biner dan mudah diinterpretasi.

4. Evaluasi Model

Kinerja model akan dievaluasi menggunakan metrik seperti F1 Score, Precision, dan Recall. Evaluasi dilaksanakan untuk mengukur kinerja model dalam indentifikasi calon pengguna Potensial.

5. Analisis Hasil dan Penerapan

Hasil klasifikasi akan di analisis dan digunakan sebagai data untuk menyusun strategi pemasaran bagi tim Pangkasin dalam menyasar calon pengguna aplikasi secara terarah dan efisien.

1.4 Tujuan Penelitian

Adapun tujuan penelitian ini adalah sebagai berikut:

1. Mengkaji potensi penerapan Algoritma Logistik Regresi dalam pengklasifikasian calon pengguna aplikasi di kota Medan.
2. Mengembangkan model klasifikasi yang dapat membedakan Konsumen potensial dan tidak potensial, berdasarkan data Observasi, seperti usia, jenis kelamin, jenis layanan, dan frekuensi kunjungan.

3. Menilai tingkat akurasi dan efektivitas model yang di bangun melalui pengukuran metrik evaluasi klasifikasi, yakni Precision, Recall, dan F1 Score.
4. Memberikan rekomendasi ide untuk perkembangan strategi bisnis dan marketing pada Tim Pangkasin untuk optimalisasi sistem logistik regresi.

1.5 Manfaat Penelitian

Menambah wawasan dan literatur terkait Logistik Regresi pada klasifikasi pelanggan, menjadi referensi ide untuk penelitian selanjut nya untuk metode klasifikasi calon pelanggan, dan membantu dalam memahami karakteristik calon pengguna aplikasi Pangkasin derta memberikan rekomendasi strategi pemasaran yang lebih efektif dengan klasifikasi menggunakan Logistik Regresi.

BAB II

LANDASAN TEORI

2.1 Regresi Logistik

Regresi logistik adalah suatu analisis regresi yang digunakan untuk menggambarkan hubungan antara variabel respon dengan sekumpulan variabel bebas, dimana variabel dependennya biner. Regresi yang berarti peramalan, penaksiran, atau pendugaan pertama kali diperkenalkan oleh Sir Francis Galton pada tahun 1877 sehubungan dengan penelitiannya terhadap tinggi manusia. Persamaan regresi adalah persamaan matematik yang memungkinkan untuk meramalkan nilai-nilai suatu variabel tak bebas dari nilai nilai satu atau lebih variabel bebas (Ndangi et al., 2019).

Menurut Hosmer et al. (2013) dalam buku *Applied Logistic Regression*, regresi logistik menyusun probabilitas berdasarkan fungsi logit (log-odds), yang dituliskan sebagai:

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

Gambar 2. 1 Persamaan utama Regresi Logistik fungsi logit
(sumber Hosmer et al., 2013)

Manfaat Logistik Regresi adalah dapat digunakan dengan data nonlinier. Dengan regresi linier, memprediksi bagaimana suatu variabel meningkat atau menurun berdasarkan perubahan variabel penjelas. Dengan regresi logistik, dapat memodelkan hubungan kompleks yang tidak bergantung pada linearitas, termasuk model klasifikasi, yang sangat penting dalam pembelajaran mesin. Regresi logistik juga dapat memproses data dengan kecepatan tinggi sekaligus tetap fleksibel terhadap beberapa jenis pertanyaan penelitian. (Coursera, 2025).

2.1.1 Logistik Regresi Biner

Regresi logistik biner adalah metode statistik yang digunakan untuk menganalisis hubungan antara variabel independen dan variabel dependen biner. Dalam konteks ini, variabel dependen adalah niat beli ulang pelanggan (biner: ya/tidak), variabel independen adalah karakteristik produk. Metode ini cocok untuk situasi di mana hasil yang diprediksi hanya memiliki dua kemungkinan, dan memberikan estimasi probabilitas bahwa pelanggan akan melakukan pembelian ulang berdasarkan karakteristik produk (Faisal et al., 2024).

Selain itu Metode Logistic Regression dipilih di penelitian ini karena sederhana namun efektif untuk klasifikasi biner. Algoritma ini termasuk yang paling luas digunakan dalam tugas klasifikasi biner karena kemudahannya dalam implementasi dan interpretasi. Model logistik regresi juga menghasilkan output probabilitas yang dapat membantu memahami seberapa yakin model terhadap prediksinya – hal ini penting untuk pengambilan keputusan bisnis.

Model regresi logistik biner ini dimanfaatkan agar menguji hubungan dari satu variabel dependen dengan beberapa variabel yang independen. Apabila variabel dependen menghasilkan dua kategori dengan nilai 0 dan 1, digunakan model regresi logistik biner, model ini mengikuti distribusi Bernoulli, yaitu sebagai berikut:

$$f(y_1) = \pi_i^{y_i} (1 - \pi_i)^{1 - y_i}$$

dengan

π_i : probabilitas kejadian ke-i

y_i : pengubah acak ke-i, mencakup atas 0 dan 1

Model dari regresi logistik satu variabel bebas memiliki bentuk persamaan yaitu (Agresti, 2002) :

$$\pi(x_i) = \frac{e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}}$$

Gambar 2. 2 Model Regresi Logistik Biner

Sumber: (Alfayulanda & Murni, 2024)

Agar mempermudah pendugaan parameter regresi, transformasikan $\pi(x)$ pada persamaan (2) yang hasilnya merupakan bentuk dari logit dari regresi logistik:

$$g(x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \beta_0 + \beta_1 x$$

Gambar 2. 3 Model Regresi Logistik Biner variabel bebas

Sumber: (Alfayulanda & Murni, 2024)

2.2 Klasifikasi

Klasifikasi merupakan cara pengelompokkan benda berdasarkan ciri – ciri yang dimiliki oleh objek klasifikasi. Dalam prosesnya, klasifikasi dapat dilakukan dengan banyak cara baik secara manual ataupun dengan bantuan teknologi. Klasifikasi yang dilakukan secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa adanya bantuan dari algoritma cerdas komputer. klasifikasi yang dilakukan dengan bantuan teknologi, memiliki beberapa algoritma, diantaranya Naïve Bayes, Support Vector Machine, Decision Tree, Fuzzy dan Jaringan Saraf Tiruan (Wibawa et al., 2018).

Selain itu Menurut jurnal dari (Hendrian, 2018) Klasifikasi adalah tipe analisis data yang dapat membantu orang menentukan kelas label dari sampel yang ingin di klasifikasi. Klasifikasi merupakan Metode supervised learning, metode yang mencoba menemukan hubungan antara atribut masukan dan atribut

target. Tujuan klasifikasi untuk meningkatkan kehandalan hasil yang diperoleh dari data.

Menurut (Irma Devi Lestari, 2016) Klasifikasi merupakan kata serapan dari bahasa Belanda, *classificatie*, yang sendirinya berasal dari bahasa Prancis *classification*. Istilah ini menunjuk kepada sebuah metode untuk menyusun data secara sistematis atau menurut beberapa aturan atau kaidah yang telah ditetapkan. Secara harafiah bisa pula dikatakan bahwa klasifikasi adalah pembagian sesuatu menurut kelas-kelas.

Menurut Ilmu Pengetahuan, Klasifikasi adalah Proses pengelompokkan benda berdasarkan ciri-ciri persamaan dan perbedaan.(Wikipedia). Ini juga dapat mengatakan bahwa Klasifikasi merupakan proses-proses dalam hal : x mendefinisikan kelas; x menentukan hubungan antara kelas (seperti hubungan hierarki, antara lain), yaitu membuat sebuah sistem klasifikasi; dan x menetapkan elemen (dalam LIS, dokumen) untuk sebuah class dalam suatu sistem klasifikasi. Ini sama dengan proses saling tergantung dari : x mendefinisikan konsep-konsep (lihat Hjørland, 2009); x menentukan hubungan semantik antara konsep-konsep (lihat Hjørland, jumlah b); dan x menentukan elemen-elemen yang jatuh di bawah konsep yang diberikan (untuk menetapkan sebuah "hal" untuk sebuah konsep).

Contoh : Untuk mengatakan bahwa konsep "Mu'ller-Lyer illusion khayalan" adalah sejenis "khayalan optis", yang merupakan jenis "fenomena persepsi", yang merupakan jenis " fenomena SVLNRORJLV DGDODK VDPD VHSHUWL NODVLILNDVLµ 0XéOOHU-Lyer illusion khayalan" dalam kelas disebut "ilusi optik", yang merupakan bagian dari kelas lebih luas dari " fenomena persepsi", yang merupakan bagian dari kelas lebih luas dari "fenomena psikologis". Ingetraut

Dahlberg, menyatakan bahwa "unsur-unsur" dari skema klasifikasi adalah "konsep atau representasi konsep-konsep".

Satu demikian dapat disimpulkan bahwa ketika sebuah konsep yang dipilih untuk sebuah kelas, konsep pertanyaan dalam merujuk kepada sesuatu yang harus dipakai bersama oleh semua dokumen-dokumen dari kelas itu, dan bahwa konsep ini diperlukan oleh dokumen-dokumen. Namun, mengambil label "011" dalam DDC sebagai contoh, ia merujuk ke kelas dokumen yang memiliki fitur umum yang mereka bibliografi dan tidak tentang bibliografi.

2.3 Teknik Analisis

Kata analysis berasal dari bahasa Greek (Yunani), terdiri dari kata "ana" dan "lysis". Ana artinya atas (above), lysis artinya memecahkan atau menghancurkan. Secara definitif ialah: "Analysis is a process of resolving data into its constituent components to reveal its characteristic elements and structure" Ian Dey (1995: 30). Agar data bisa dianalisis data tersebut harus dipecah dahulu menjadi bagian-bagian kecil (menurut element atau struktur), kemudian mengaduknya bersama untuk memperoleh pemahaman yang baru (Kurniasih et al., 2021).

2.4 Implementasi

Secara umum Implementasi dalam kamus besar Indonesia berarti pelaksanaan atau penerapan. Istilah suatu implementasi biasanya dikaitkan dengan suatu kegiatan yang dilaksanakan untuk mencapai tujuan tertentu. Implementasi merupakan sebuah penempatan ide, konsep, kebijakan, atau inovasi dalam suatu tindakan praktis memberikan dampak, baik berupa perubahan pengetahuan, keterampilan maupun nilai dan sikap. Implementasi merupakan aspek penting dalam keseluruhan proses kebijakan dan merupakan suatu upaya untuk mencapai

tujuan tertentu dengan sarana dan prasarana tertentu dan dalam urutan waktu tertentu. Pada dasarnya implementasi kebijakan adalah upaya untuk mencapai tujuan yang sudah ditentukan melalui program-program agar dapat terpenuhi pelaksanaan kebijakan itu (Badan Pengembangan dan Pembinaan Bahasa, 2016).

2.5 Customer Relationship Management (CRM)

Setelah era digital yang selalu menggunakan “e-“ didepan semua kata seperti e-commerce, e-book, esales, e-mail, dst, Customer Relationship Management atau CRM merupakan akronim atau singkatan yang paling populer dikalangan orang-orang sales dan marketing. Kalau di-bahasa Indonesia-kan kira-kira adalah Manajemen Hubungan Pelanggan (MHP). Telaah per katanya adalah sebagai berikut; Pelanggan atau customer, kata pertama, di kamus bahasa Inggris artinya adalah seseorang yang berulang kali atau teratur melakukan pembelian kepada seorang pedagang. Jadi pelanggan adalah orangnya (dalam definisi ini tidak disinggung tentang kepuasan, mahal-murahnya pembelian, dll). Hubungan atau relationship, kata kedua, adalah bentuk komunikasi dua arah antara pembeli dan penjual. Manajemen, kata terakhir, artinya pengelolaan (secara luas tanpa perlu menjabarkan detail bagaimana mengelola sesuatu). Jadi definisi diatas kalau digabungkan kira-kira menjadi pengelolaan hubungan dua arah antara suatu perusahaan dengan orang yang menjadi pelanggan di perusahaan tersebut (Danardatu, 2019).

2.6 Aplikasi

Mengutip dari jurnal Aplikasi Jasa Tukang Berbasis Website (Alwin & Aditiawan, 2022) Aplikasi Adalah program siap pakai yang digunakan untuk menjalankan perintah-perintah dari pengguna aplikasi tersebut dengan tujuan

pembuatan aplikasi tersebut, aplikasi mempunyai arti yaitu pemecahan masalah yang menggunakan salah satu teknik pemrosesan data aplikasi yang biasanya berpacu pada sebuah komputasi yang diinginkan atau diharapkan maupun pemrosesan data yang diharapkan.

Menurut Telkomuniversity (2024) Aplikasi juga adalah sebuah program atau perangkat lunak yang dirancang untuk menjalankan fungsi tertentu pada perangkat elektronik, seperti komputer, ponsel pintar, atau tablet. Istilah ini berasal dari bahasa Inggris “application software,” yang berarti perangkat lunak aplikasi. Dalam konteks teknologi saat ini, aplikasi sangat beragam jenisnya, mencakup berbagai sektor dan kebutuhan.

2.7 Python

Python adalah sebuah bahasa pemrograman yang digunakan untuk menyederhanakan perhitungan matematis, termasuk pertidaksamaan nilai mutlak. Menurut penelitian Amri Hamzah et al. (2023), penerapan Python dalam perhitungan pertidaksamaan telah menunjukkan efektivitas yang signifikan dalam mengurangi kesalahan perhitungan manual, yang sering kali dialami oleh mahasiswa. Python juga dapat digunakan untuk visualisasi penyelesaian, yang membantu mahasiswa memahami struktur permasalahan secara lebih jelas dan terorganisir. (Ameliya et al., 2024)

Selain itu menurut buku An Introduction To Python (Van Rossum & Drake, 2006) bahwa Python adalah bahasa pemrograman yang mudah dipelajari dan canggi. Bahasa ini memiliki struktur data tingkat tinggi yang efisien dan pendekatan yang sederhana namun efektif untuk pemrograman berorientasi objek. Sintaksis Python yang elegan dan pengetikan dinamis, bersama dengan sifatnya

yang ditafsirkan, menjadikannya bahasa yang ideal untuk pembuatan skrip dan pengembangan aplikasi cepat di banyak area di sebagian besar platform. Interpreter Python dan pustaka standar yang luas tersedia secara gratis dalam bentuk sumber atau biner untuk semua platform utama dari situs Web Python, <http://www.python.org/>, dan dapat didistribusikan secara bebas.

Situs yang sama juga berisi distribusi dan petunjuk ke banyak modul, program, dan alat Python pihak ketiga gratis, serta dokumentasi tambahan. Interpreter Python mudah diperluas dengan fungsi dan tipe data baru yang diimplementasikan dalam C atau C++ (atau bahasa lain yang dapat dipanggil dari C). Python juga cocok sebagai bahasa ekstensi untuk aplikasi yang dapat disesuaikan.

Tutorial ini memperkenalkan pembaca secara informal pada konsep dasar dan fitur bahasa dan sistem Python. Akan sangat membantu jika memiliki interpreter Python yang praktis untuk pengalaman langsung, tetapi semua contoh bersifat mandiri, tutorial ini juga dapat dibaca secara offline. Untuk deskripsi objek dan modul standar, lihat dokumen Referensi Pustaka Python. Manual Referensi Python memberikan definisi bahasa yang lebih formal. Untuk menulis ekstensi dalam C atau C++, baca Memperluas dan Menyematkan Interpreter Python dan Referensi API Python/C. Ada juga beberapa buku yang membahas Python secara mendalam. Tutorial ini tidak berupaya untuk menjadi komprehensif dan mencakup setiap fitur, atau bahkan setiap fitur yang umum digunakan.

Sebaliknya, tutorial ini memperkenalkan banyak fitur Python yang paling penting, dan akan memberi gambaran yang baik tentang cita rasa dan gaya bahasa tersebut. Setelah membacanya akan dapat membaca dan menulis modul dan

program Python, dan akan siap untuk mempelajari lebih lanjut tentang berbagai modul pustaka Python yang dijelaskan dalam Referensi Pustaka Python.

Python juga Bahasa pemrograman paling cocok sebagai bahasa pertama yang harus dipelajari oleh programmer pemula, karena bahasa ini memiliki perangkat yang canggih yang mencerminkan cara orang berpikir dan cara mereka mengimplementasikan kode.

[1] Selain itu, bahasa ini meminimalkan kata kunci tambahan yang diperlukan untuk menulis program yang benar secara sintaksis. Pendekatan ini tampaknya lebih produktif daripada mengajarkan bahasa C++ atau Java, yang memiliki banyak istilah dan elemen yang terkait dengan spesifikasi bahasa daripada realisasi algoritma. Selain itu, instruktur di lebih dari selusin universitas, seperti MIT, UC Berkeley, UC Davis, Sonoma State University, University of Washington, University of Waterloo, Luther College, dan Swarthmore College, telah menggunakannya untuk mengajarkan kursus pemrograman pengantar kepada mahasiswa jurusan ilmu komputer.

[2] Saat ini, bagi seorang ilmuwan komputer, penting untuk mempelajari setidaknya satu bahasa pemrograman, karena semua inovasi dan teknologi didasarkan pada pemahaman menyeluruh tentang komputer, sistem operasi, API perangkat lunak, atau beberapa periferal perangkat keras. Semuanya diciptakan oleh programmer yang menggunakan cara berpikir tertentu. Dan untuk memperoleh cara berpikir tersebut, seseorang harus terbiasa dengan salah satu bahasa pemrograman dan menjadi ahli dalam pengembangan perangkat lunak.

[6] Bagi siapa pun yang mulai mempelajari pemrograman, penting untuk berkonsentrasi pada konsep pemrograman daripada pada spesifikasi bahasa,

karena konsep tersebut mungkin berbeda untuk berbagai bahasa pemrograman. Namun, Python menyediakan tingkat pemrograman tertinggi. Jadi, siswa tidak perlu memikirkan manajemen memori, yang tidak dapat dihindari dalam C++, atau hierarki kelas, yang tidak dapat dihindari dalam Java, atau tipe dan deklarasi variabel, yang ada di hampir setiap bahasa pemrograman. (Bogdanchikov et al., 2013).

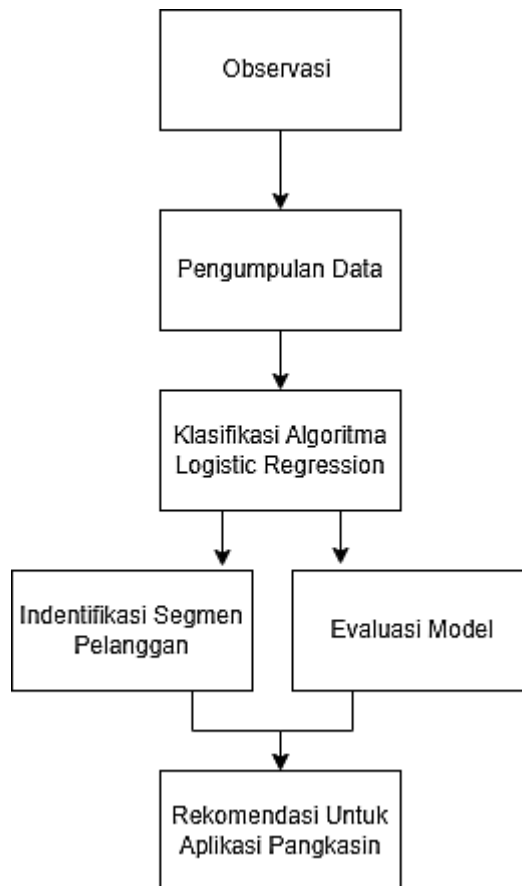
2.8 Penelitian Terdahulu

Penelitian yang dilakukan oleh (Naïve et al., 2024) membahas tentang Klasifikasi untuk membuktikan status penderita hipertensi atau tidak hipertensi. Metode klasifikasi yang digunakan adalah Binary Logistic Regression dan Naïve Bayes Classifier. Faktor yang digunakan pada penelitian ini adalah jenis kelamin, usia, tinggi badan, dan berat badan. Akurasi hasil klasifikasi paling besar terdapat pada proporsi 90%:10%. Akurasi klasifikasi yang dihasilkan oleh metode Binary Logistic Regression menghasilkan sensitivitas sebesar 93,33%. Akurasi klasifikasi yang diperoleh oleh Naïve Bayes Classifier dengan sensitivitas sebesar 63,64%. Hal ini menunjukkan bahwa metode Regresi Logistik Biner mempunyai nilai sensitivitas yang lebih baik.

Selain itu penelitian yang dilakukan oleh (Ndangi et al., 2019) Penelitian ini membahas tentang klasifikasi konsumen berdasarkan tempat berbelanja di pasar modern atau pasar tradisional. Metode klasifikasi yang digunakan adalah Analisis Diskriminan dan Regresi Logistik. Faktor yang digunakan dalam penelitian ini meliputi variabel promosi dan pendidikan. Hasil penelitian menunjukkan bahwa Regresi Logistik memiliki tingkat ketepatan klasifikasi yang lebih tinggi dibandingkan dengan Analisis Diskriminan, dengan akurasi 88,0% untuk data

learning dan 80,0% untuk data testing. itu, Analisis Diskriminan hanya mencapai 86,7% untuk data learning dan 76,0% untuk data testing. Hal ini menunjukkan bahwa Regresi Logistik lebih efektif dalam mengelompokkan konsumen berdasarkan tempat berbelanja.

2.9 Kerangka Berpikir Konseptual



Gambar 2. 4 Kerangka Berpikir Konseptual

Keterangan:

1. Observasi Klasifikasi

Melakukan Observasi pada beberapa Barbershop di kota Medan, dengan pertanyaan untuk mengetahui mengenai Identifikasi Profil konsumen

barbershop, pola konsumsi dan preferensi pelanggan, serta Persepsi barbershop terhadap digitalisasi dengan adanya Platform Pangkasin nantinya.

2. Pengumpulan Data

Pengumpulan data hasil dari observasi ke beberapa barbershop di kota Medan.

3. Klasifikasi Algoritma Logistic Regression

Menggunakan Algoritma Logistic Regression untuk dapat mengklasifikasi pengunjung barbershop untuk mendapatkan data sebagai ide preferensi calon pengguna aplikasi Pangkasin nantinya.

4. Identifikasi segmen Pelanggan dan Evaluasi Model

Proses Identifikasi segmentasi pelanggan dari hasil klasifikasi dan penyesuaian model Algoritma atau metode yang relevan dari Logistic Regression.

5. Rekomendasi untuk Tim Pangkasin

Setelah mendapatkan data dari klasifikasi, Identifikasi segmen pelanggan, dan evaluasi model penulis memberikan kesimpulan dan masukkan ke Tim Pangkasin untuk dapat referensi untuk perkembangan bisnis.

BAB III METODOLOGI PENELITIAN

Pada Bab 3 ini Peneliti menjelaskan tahapan-tahapan dalam penelitian ini berawal dari melakukan observasi untuk mendapatkan data dari setiap beberapa barbershop di kota Medan, Penelitian ini termasuk jenis penelitian Kuantitatif di bidang data Mining. Metode yang digunakan adalah Observasi Lapangan dan analisis Data Sekunder. Data Utama di dapatkan pada melalui Observasi di 8 barbershop di Kota Medan dengan Wilayah berbeda-beda, yang telah setuju menjadi sampel penelitian. ketika sudah mendapatkan data peneliti melakukan klasifikasi dari data yang di dapat kan hasil nya akan mendapatkan kesimpulan untuk memberikan masukan pada tim Pangkasin agar dapat melakukan pengembangan bisnis dari hasil preferensi konsumen.

3.1 Lingkungan yang di pakai untuk mengembangkan Program

Lingkungan yang di pakai untuk mengembangkan program meliputi jenis komputer dan perangkat lunak apa saja yang di pakai, serta sistem operasi apa yang di gunakan untuk klasifikasi.

3.1.1 Perangkat Lunak yang di pakai

Perangkat lunak yang di pakai untuk menjalankan Algoritma klasifikasi

Logistik Regresi:

A. Python Language versi 3.13.2, dengan Library utama Pandas untuk pengolahan data, Scikit-Learn untuk model Machine Learning.

B. Visual Studio Code, untuk environment coding agar dokumentasi proses analisis terstruktur.

3.1.2 Sistem Operasi yang di pakai

Sistem operasi yang di pakai untuk menjalankan Algoritma klasifikasi Logistik Regresi pada pengunjung barbershop di kota Medan untuk referensi ide terhadap tim Pangkasin adalah Windows.

3.1.3 Jenis Laptop yang digunakan

Jenis komputer yang digunakan untuk menjalankan Algoritma klasifikasi Logistik Regresi pada pengunjung barbershop di Kota Medan untuk referensi ide terhadap tim Pangkasin adalah laptop dengan spesifikasi yaitu:

- A. Processor AMD Ryzen 3 2200U with Radeon Vega Mobile Gfx 2.50 GHz.
- B. RAM 12 GB.
- C. Sistem Operasi Windows 10 Pro.
- D. 64-bit.

3.2 Strategi Pemecahan Masalah

Untuk menyelesaikan permasalahan dalam penelitian ini, peneliti merancang beberapa langkah strategis sebagai berikut:

1. Identifikasi Masalah dan Kebutuhan Pengguna
2. Dilakukan dengan observasi langsung ke beberapa barbershop di Kota Medan untuk memahami karakteristik pelanggan, kebutuhan layanan, dan potensi digitalisasi sistem pemesanan.
3. Pengumpulan dan Penyusunan Data Hasil Observasi
4. Karena aplikasi Pangkasin belum memiliki mitra aktif dan data real, digunakan data hasil observasi dari beberapa barbershop di Kota Medan berdasarkan hasil observasi, seperti data usia pelanggan, jenis kelamin, frekuensi kunjungan, dan minat terhadap layanan.

No	Variabel	Deskripsi	Tipe Data
1	Usia	Umur Konsumen (Tahun)	Numerik (Continue)
2	Jenis Kelamin	Gender Konsumen (Pria/Wanita)	Kategorikal
3	Pekerjaan	Pekerjaan (Mahasiswa, karyawan,dll)	Kategorikal
4	Frekuensi Kunjungan	Rata-rata Kunjungan ke Barbershop	Numerik (Diskrit)
5	Preferensi Layanan	Layanan Utama yang Diminati (Cukur/Styling)	Kategorikal
6	Respon Terhadap Promosi	Minat pada Promo (Mis. Tertarik/tidak)	Kategorikal (Biner)
7	Label: Potensial/Tidak	Kelas Output – Apakah Konsumen Potensial	Kategorikal (Biner)

Tabel 3. 1 Variabel Observasi pada Konsumen di Barbershop

5. Penerapan Algoritma Logistic Regression
6. Logistic Regression digunakan karena kemampuannya dalam mengklasifikasikan variabel dependen biner (dalam hal ini: minat atau tidak minat terhadap layanan digital). Model dilatih menggunakan data dummy yang mencerminkan kondisi lapangan.
7. Evaluasi Model
8. Model yang telah dibangun dievaluasi menggunakan metrik akurasi, precision, recall, dan F1-score untuk menilai performanya dalam mengklasifikasikan calon pengguna potensial aplikasi Pangkasin.
9. Analisis dan Penarikan Kesimpulan
10. Hasil klasifikasi digunakan untuk memberikan masukan strategi pemasaran dan pengembangan fitur aplikasi berdasarkan segmentasi pelanggan. Hal ini bertujuan untuk meningkatkan efektivitas pemasaran dan adopsi aplikasi Pangkasin di masa depan.

3.3 Data Observasi

Data beberapa Barbershop di Kota Medan:

No	Nama Barbershop	Rata-rata Pelanggan	Layanan Favorit	Jam Ramai	Sistem Reservasi	Minat Digitalisasi
1	The Man Barbershop	10/Hari	Premium (Rp.80.000)	Sabtu dan Minggu sore	WhatsApp	Setuju
2	Tress Barbershop	20/Hari	Basic (Rp.75.000)	Sabtu dan	WhatsApp	Setuju

				Minggu Pagi/sore		
3	D'Colonel's Barbershop	10/Hari	D'Colonel's Cut (Rp.50.000)	Sabtu dan Minggu Sore ke malam	WhatsApp	Setuju
4	Premium Barbershop ZIE & ZKA	5-10/Hari	Basic (Rp.40.000)	Sabtu dan Minggu Sore ke Malam	WhatsApp	Setuju
5	168 Barbershop	5-10/Hari	Basic (Rp.30.000)	Sabtu dan Minggu	Manual	Setuju
6	Ivon Pangkas	5-16/Hari	Basic (Rp.20.000)	Sabtu dan Minggu	WhatsApp	Setuju
7	Mr.Klimis Marelan	15- 20/Hari	Basic (Rp.40.000)	Sabtu dan Minggu Sore ke Malam	WhatsApp	Setuju
8	Aa	15-	Basic	Sabtu	WhatsApp	Belum/Tidak

	Barbershop	25/Hari	(Rp.40.000)	dan Minggu Sore ke Malam		
--	------------	---------	-------------	-----------------------------------	--	--

Tabel 3. 2 Data Observasi Pada Barbershop

Hasil Observasi di atas berdasarkan format pertanyaan sebagai berikut:

A. Identifikasi Profil Konsumen Barbershop

1. Berapa rata-rata jumlah pelanggan perhari di barbershop ini?
2. Bagaimana komposisi umur pelanggan (misalnya: di bawah 18 tahun, 18-25 tahun, 26-35 tahun, di atas 35 tahun)?
3. Apakah lebih banyak pelanggan pria atau ada juga pelanggan wanita?
4. Jenis layanan apa yang paling sering digunakan oleh pelanggan? (misal: potong rambut, cukur janggut, hair coloring, dll.)

B. Pola Konsumsi dan Preferensi Pelanggan

1. Apakah pelanggan lebih sering datang langsung (walk-in) atau melakukan reservasi terlebih dahulu?
2. Jika reservasi sudah diterapkan, bagaimana mekanisme yang digunakan (WhatsApp, aplikasi lain, atau datang langsung)?
3. Apakah pelanggan lebih suka memilih tukang cukur tertentu atau menerima siapa saja yang tersedia?
4. Seberapa besar pengaruh promosi terhadap peningkatan jumlah pelanggan? (misalnya diskon, promo loyalitas)

5. Apakah ada pola pelanggan berdasarkan profesi? (misalnya mahasiswa, pekerja kantoran, dll.)

C. Persepsi Barbershop terhadap Digitalisasi dan Pangkasin

1. Apakah barbershop pernah menggunakan sistem pencatatan pelanggan atau masih manual?
2. Jika pernah menggunakan sistem digital, seberapa efektifkah sistem tersebut dalam mengelola pelanggan?
3. Bagaimana pendapat barbershop tentang konsep aplikasi seperti Pangkasin yang memungkinkan pelanggan melakukan booking online?
4. Apa tantangan utama dalam mengelola antrian pelanggan di barbershop ini?
5. Apakah barbershop bersedia bergabung dengan Pangkasin? Jika tidak, apa alasannya?

D. Evaluasi Kinerja dan Efektivitas Layanan

1. Berapa rata-rata waktu yang dibutuhkan untuk melayani satu pelanggan?
2. Apakah ada perbedaan waktu layanan berdasarkan jenis potongan rambut atau layanan lainnya?
3. Bagaimana tingkat kepuasan pelanggan terhadap layanan yang diberikan?
4. Apakah pelanggan sering memberikan umpan balik atau keluhan mengenai layanan? Jika iya, keluhan apa yang paling sering muncul?

3.3.1 The Man Barbershop

Lokasi: Jl. Halat No.39A, Kotamatsum III, Kec. Medan Kota, Kota Medan, Sumatera Utara 20132

Dengan Narasumber bernama Bang Niko (Hairstylist)

Dengan Data Observasi sebagai berikut:

1. Jumlah pelanggan per hari : Sekitar 10 orang
2. Umur rata-rata pelanggan: 25 tahun keatas
3. Pekerjaan pelanggan rata-rata: Pekerja atau pebisnis
4. Pria/Wanita: Pria
5. Jenis layanan favorit : Premium (Rp.80.000)
6. Hari dan jam kunjungan tertinggi : Sabtu dan Minggu sore
7. Sistem reservasi : Melalui WhatsApp
8. Minat bergabung Pangkasin : Menyatakan setuju

Hasil observasi di The Man Barbershop menunjukkan rata-rata pelanggan harian mencapai 10 orang, dengan layanan paling diminati adalah paket premium. Waktu kunjungan tertinggi terjadi pada akhir pekan di sore hari. Meskipun saat ini sistem antrian dilakukan secara manual, barbershop ini terbuka terhadap penggunaan aplikasi seperti Pangkasin dan menggunakan WhatsApp sebagai media reservasi.

3.3.2 Trees Barbershop

Lokasi: Komplek J City. Jl. Karya Wisata No.12 Blok A, Pangkalan Masyhur, Kec. Medan Johor, Kota Medan, Sumatera Utara 20146.

Dengan Narasumber bernama Kak Zura (Kasir)

Dengan Data Observasi sebagai berikut:

1. Jumlah pelanggan per hari : Sekitar 20 orang
2. Umur rata-rata pelanggan: 20 tahun keatas
3. Pekerjaan rata-rata pelanggan: Pelajar dan Pekerja
4. Pria/Wanita: Kebanyakan Pria, ada Wanita itu pun seperti Tomboy
5. Jenis layanan favorit : Premium (Rp.75.000)
6. Hari dan jam kunjungan tertinggi : Sabtu dan Minggu pagi hingga sore
7. Sistem reservasi : Melalui WhatsApp chat/telpon
8. Minat bergabung Pangkasin : Setuju dan terbuka bergabung

Trees Barbershop memiliki volume pelanggan yang cukup tinggi, terutama pada akhir pekan. Mayoritas pelanggan memilih layanan basic. Mereka telah menggunakan WhatsApp sebagai media reservasi dan menyatakan kesediaannya untuk bergabung dengan Pangkasin jika dinilai dapat membantu pengelolaan layanan.

3.3.3 D'colonel's Barbershop

Lokasi: Jalan Alfalah Raya No.18 Glugur Darat 1 Medan Timur, Kota Medan, Indonesia 20236

Dengan Narasumber Bang Putra (Hairstylist)

Dengan Observasi Observasi sebagai berikut:

1. Jumlah pelanggan per hari : Sekitar 10 orang
2. Umur rata-rata Pelanggan: 16 tahun keatas
3. Pekerjaan rata-rata pelanggan: Mahasiswa, Pelajar, dan Pekerja
4. Pria/Wanita: Pria
5. Jenis layanan favorit : Colonel's Cut (Rp.50.000)
6. Hari dan jam kunjungan tertinggi : Sabtu dan Minggu sore - malam

7. Sistem reservasi : Melalui WhatsApp

8. Minat bergabung Pangkasin : Setuju selama membantu

Barbershop ini memiliki pelanggan reguler dengan waktu kunjungan dominan di akhir pekan sore. Meskipun pencatatan masih dilakukan secara manual, D'Colonel's menunjukkan sikap positif terhadap konsep digitalisasi. Reservasi dilakukan fleksibel melalui WhatsApp.

3.3.4 Premium Barbershop ZIE & ZKA

Lokasi: Jalan Bilal No.62, Pulo Brayan Darat 1 Medan Timur, Kota Medan, Indonesia 20239

Dengan Narasumber Bang Zais

Dengan Data Observasi sebagai berikut:

1. Jumlah pelanggan per hari : Sekitar 5-10 orang kadang 25 orang
2. Umur rata-rata Pengunjung: 20 tahun keatas
3. Pekerjaan rata-rata pelanggan: Mahasiswa dan pekerja
4. Pria/Wanita: Pria
5. Jenis layanan favorit : Basic (Rp.40.000)
6. Hari dan jam kunjungan tertinggi : Sabtu dan Minggu sore - malam
7. Sistem reservasi : Melalui WhatsApp
8. Minat bergabung Pangkasin : Setuju

Premium Barbershop Zie & Zka melayani sekitar 5–10 pelanggan per hari, meningkat saat ramai. Mereka telah terbiasa dengan sistem semi-digital dan menyambut baik konsep aplikasi Pangkasin yang diharapkan bisa meningkatkan jumlah pelanggan serta efisiensi reservasi.

3.3.5 Barbershop

Lokasi: Jalan Gunung Krakatau, Glugur Darat 1 Medan Timur, Kota Medan, Indonesia 20238

Dengan Narasumber Pak Saprin

Dengan Data Observasi sebagai berikut:

1. Jumlah pelanggan per hari : Sekitar 5-10 orang
2. Umur Rata-rata pengunjung: 25 tahun keatas
3. Pekerjaan Pengunjung: Pekerja dan Pengusaha
4. Pria/Wanita: Pria
5. Jenis layanan favorit : Basic (Rp.30.000)
6. Hari dan jam kunjungan tertinggi : Sabtu dan Minggu
7. Sistem reservasi : Langsung datang
8. Minat bergabung Pangkasin : Belum mau

168 Barbershop tergolong kecil dengan jumlah pelanggan sedikit karena bekerja sendiri. Semua pencatatan dilakukan secara manual dan tidak ada sistem reservasi aktif. Namun, pemilik menyambut baik ide digitalisasi tetapi masih belum mau karena bekerja sendiri takut menyulitkan.

3.3.6 Ivon Pangkas

Lokasi: Jalan Rahayu No.117, Bandar Klippa, Kec. Percut Sei Tuan, Kabupaten Deli Serdang, Indonesia 20371

Dengan Narasumber Nahazisokhi Buulolo (Freelance Hairstylist)

Dengan Data Observasi sebagai berikut:

1. Jumlah pelanggan per hari : Sekitar 5-16 orang
2. Umur rata-rata Pengunjung: 18 tahun ke atas

3. Pekerjaan pengunjung: Pekerja dan Wiraswasta
4. Pria/Wanita: Khusus Pria
5. Jenis layanan favorit : Basic (Rp.20.000)
6. Hari dan jam kunjungan tertinggi : Sabtu dan Minggu
7. Sistem reservasi : Melalui WhatsApp
8. Minat bergabung Pangkasin : Setuju meningkatkan pelanggan

Ivon Pangkas merupakan barbershop dengan jumlah pelanggan harian bervariasi, tergantung hari. Layanan yang ditawarkan tergolong sangat terjangkau. Meskipun sebagian besar pelanggan datang langsung, beberapa juga memesan via WhatsApp. Pemilik menyatakan dukungan terhadap konsep Pangkasin karena dianggap dapat membantu memperluas jangkauan pelanggan.

3.3.7 Mr. Klimis Marelan

Lokasi: Jalan Marelan Raya Pasar 2, Rengas Pulau, Kec. Medan Marelan, Kota Medan, Sumatera Utara 20255

Dengan Narasumber Bang Ilhsam (Hairstylist)

Dengan Data Observasi sebagai berikut:

1. Jumlah pelanggan per hari : Sekitar 15-20 orang
2. Umur rata-rata Pelanggan: 15 Tahunan keatas
3. Pekerjaan Pelanggan: Mahasiswa, Pelajar, Pekerja, dan Pengusaha
4. Pria/Wanita: Pria ada, wanita juga ada, biasa potong pendek
5. Jenis layanan favorit : Basic (Rp.40.000)
6. Hari dan jam kunjungan tertinggi : Sabtu dan Minggu sore-malam
7. Sistem reservasi : Melalui WhatsApp
8. Minat bergabung Pangkasin : Setuju tapi sistem nya harus bagus

Mr. Klimis Marelan memiliki volume pelanggan yang cukup tinggi, terutama pada akhir pekan. Sebagian pelanggan sudah terbiasa melakukan pemesanan lewat WhatsApp. Pemilik menyambut baik ide digitalisasi, tetapi mengingatkan agar sistem yang ditawarkan tetap praktis dan tidak merepotkan dari sisi operasional.

3.3.8 Aa Barbershop

Lokasi: Jl. Veteran No.21, Belawan 1, Medan Kota Belawan, Kota Medan, Sumatera Utara 20411

Dengan Narasumber Bang Aa (Hairstylist)

Dengan Data Observasi sebagai berikut:

1. Jumlah pelanggan per hari : Sekitar 15-25 orang
2. Umur rata-rata pengunjung: 22 Tahun keatas
3. Pekerjaan pelanggan rata-rata: Pekerja dan Pengusaha
4. Pria/Wanita: Pria
5. Jenis layanan favorit : Basic (Rp.40.000)
6. Hari dan jam kunjungan tertinggi : Sabtu dan Minggu sore-malam
7. Sistem reservasi : Melalui WhatsApp
8. Minat bergabung Pangkasin : Belum mau, karena lihat dulu

Aa Barbershop melayani pelanggan dalam jumlah cukup besar dengan waktu kunjungan puncak di akhir pekan sore. Layanan basic menjadi favorit pelanggan. Mereka sudah familiar dengan sistem pemesanan lewat WhatsApp dan tapi belum bersedia bekerja sama dengan Pangkasin karena mau melihat perkembangan nya dulu.

3.4 Jadwal Penelitian

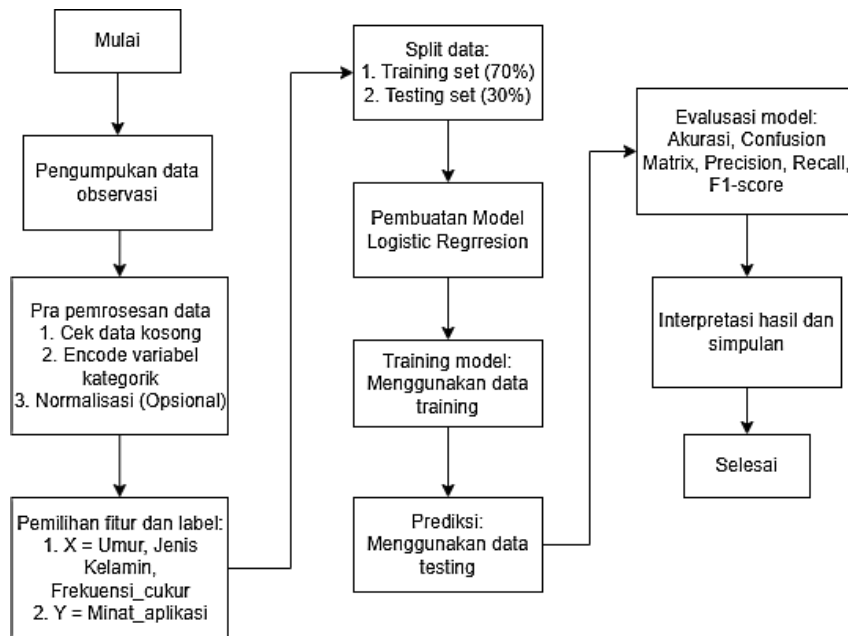
Untuk Jadwal penelitian Peneliti hanya pada saat melakukan coding pada Python di laptop peneliti mulai dari bulan Februari-April tahun 2025, karena saat observasi sebagian sudah ada data nya sebelum nya, jadi peneliti hanya meminta data hasil observasi untuk diolah datanya menggunakan Logistik Regresi, Tabel waktu kegiatan sebagai berikut:

Kegiatan	Waktu Kegiatan											
	Februari 2025				Maret 2025				April 2025			
	Minggu ke				Minggu ke				Minggu ke			
	1	2	3	4	1	2	3	4	1	2	3	4
Observasi												
Pengumpulan Data												
Pra-pemrosesan & Analisis Awal Data												
Pembuatan & Pelatihan Model Logistic Regression												
Evaluasi Model & Pelatihan Model Logistic Regression												

Tabel 3. 3 Tabel Jadwal Kegiatan

3.5 Flowchart Penelitian

Untuk menggambarkan alur proses penelitian ini secara sistematis, digunakan sebuah flowchart. Flowchart ini menjelaskan tahapan-tahapan penelitian mulai dari pengumpulan data hingga proses klasifikasi menggunakan algoritma Logistic Regression. Adapun tahapan dalam flowchart terdiri atas:



Gambar 3. 1 Flowchart atau Diagram alur Penelitian

Flowchart penelitian yang digunakan ini menggambarkan tahapan-tahapan dalam proses klasifikasi calon konsumen barbershop menggunakan algoritma Logistic Regression. Tujuannya adalah untuk memberikan gambaran mengenai alur logika penelitian, mulai dari pengumpulan data hingga evaluasi model. Berikut ini adalah penjelasan tiap-tiap tahapan dalam flowchart tersebut:

1. **Mulai**
2. Proses dimulai dengan penetapan tujuan penelitian, yaitu untuk mengklasifikasikan konsumen potensial barbershop berdasarkan data observasi.

3. Pengumpulan Data Observasi

4. Data dikumpulkan secara manual melalui observasi langsung ke delapan barbershop di kota Medan. Data tersebut meliputi atribut seperti umur konsumen, jenis kelamin, frekuensi cukur, dan minat terhadap penggunaan aplikasi digital pemesanan waktu dan tempat di barbershop (Pangkasin).

5. Pra-pemrosesan Data (Preprocessing)

6. Data mentah yang telah dikumpulkan selanjutnya diproses dengan beberapa langkah seperti:
 - A. Membersihkan data (menghapus nilai kosong/duplikat).
 - B. Mengubah variabel kategorik menjadi numerik (contohnya: pria = 0, wanita = 1).
 - C. Menyiapkan struktur data agar siap untuk dilatih oleh model.

7. Pemisahan Data (Data Splitting)

8. Setelah data bersih, dilakukan pemisahan dataset menjadi dua bagian, yaitu:

Data latih (training data): digunakan untuk melatih model.

Data uji (testing data): digunakan untuk mengukur performa model terhadap data baru.

9. Pembangunan Model Logistic Regression

10. Selanjutnya Pada tahap ini, model Logistic Regression dibentuk menggunakan data latih. Model akan belajar mengenali pola hubungan antara fitur input (umur, jenis kelamin, frekuensi cukur) dan target output (minat terhadap aplikasi).

11. Pelatihan Model (Training)

12. Model dijalankan dan dioptimalkan menggunakan data latih untuk memperoleh koefisien (bobot) yang paling sesuai dalam memprediksi probabilitas klasifikasi biner.

13. Prediksi (Prediction)

14. Setelah model terlatih, data uji diberikan untuk menghasilkan prediksi berupa dua kelas:

1: Konsumen potensial (berminat menggunakan aplikasi).

0: Konsumen tidak potensial.

15. Evaluasi Model (Evaluation)

16. Kinerja model dinilai menggunakan metrik klasifikasi seperti:

A. Accuracy (akurasi).

B. Precision (presisi).

C. Recall (daya ingat).

D. F1-Score.

Evaluasi dilakukan untuk menilai sejauh mana model mampu mengklasifikasikan data secara benar.

17. Interpretasi Hasil

18. Setelah model dievaluasi, dilakukan analisis terhadap hasil klasifikasi untuk mengetahui variabel mana yang paling berpengaruh serta potensi penerapan model ke sistem nyata.

19. Selesai

20. Tahapan akhir ditutup dengan penyusunan kesimpulan berdasarkan hasil klasifikasi serta saran pengembangan untuk penelitian lanjutan.

Perlu diketahui bahwa teknik analisis data yang digunakan adalah Observasi. Jadi para pengembang memantau jalannya algoritma, apakah berjalan dengan baik atau tidak, jika tidak algoritma akan dijalankan ulang.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Perhitungan Algoritma Logistik Regresi

Pada Penelitian ini peneliti telah melakukan observasi dan mendapatkan data, setelah itu mencari tau rumus dari Algoritma Logistik Regresi dari beberapa jurnal informasi, serta tahapan-tahapan terkait Logistik Regresi dan penerapannya, peneliti dapat melakukan penyesuaian dengan objek penelitian mencari korelasi antar objek penelitian dan Algoritma Logistik Regresi itu sendiri, didapatkan tahapan-tahapan perhitungan sebagai berikut:

4.1.1 Fungsi Linier

Untuk mengetahui setiap data Pelanggan dari beberapa Barbershop di Kota Medan, memiliki beberapa Variabel yaitu Umur (X1), Jenis Kelamin (X2), dan Frekuensi cukur (X3). Dikalikan dengan Bobot (w) dengan setiap Variabel tersebut dan di jumlahkan dengan bias/intersep (b), membentuk persamaan linier sebagai nilai tambahan untuk membantu nilai Model untuk memprediksi apakah benar atau salah, tetapi jika disesuaikan dengan penelitian Minat atau Tidak Minat, dengan rumus nya sebagai berikut:

$$z = b + w_1x_1 + w_2x_2 + w_3x_3$$

Gambar 4. 1 Rumus Fungsi Linier

Dengan rumus di atas, diketahui b: Bias/Intersep = 2, w_1: bobot umur = -0.08, w_2: jenis kelamin = 0.62, w_3: frekuensi cukur = 0.92.

Dan dengan data nya misal ambil satu umur = 27 Tahun, jenis kelamin = 0 (0 = pria dan 1 = wanita), frekuensi cukur = 3 kali.

Contoh perhitungan nya sebagai berikut:

$$Z = 2 + (-0.08 \times 27) + (0.62 \times 0) + (0.92 \times 3)$$

$$= 2 + (-2.16) + 0 + 2.76 = 2.6$$

4.1.2 Fungsi Sigmoid

Masuk ke bagian Fungsi Sigmoid merupakan upaya atau rumus untuk merubah output linier (Z) ke bentuk probabilitas 0 atau 1, jadi seperti jika $0.44 = 44\%$, contoh nya seperti itu, atau fungsi sigmoid dapat juga merubah nilai yang dalam bentuk negatif atau positif ke dalam probabilitas. Dengan rumus nya sebagai berikut:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Gambar 4. 2 Rumus Fungsi Sigmoid

Dengan rumus di atas diketahui fungsi sigmoid sebagai mengubah nilai z (bisa positif dan negatif) menjadi probabilitas (Nilai 0 - 1), dan e sebagai bilangan euler (2.718). Dengan perhitungan nya sebagai berikut:

$$Z = 2,6, \sigma(2.6) = 1/1 + 2,718^{-2,6}$$

$$2.718^{-2.6} = 0.0507$$

$$\sigma(2.6) = 1/1+0.0507 = 1/1.0507 = 0.93$$

Dengan hasil perdata nya itu 0.93, yang berarti 93% jika dalam bentuk persentase. Dengan kelas nya itu P (Probabilitas) = 1 (Minat), jika $p > 0.5$ dan 0 (tidak Minat), jika $p < 0.50$. dari data di atas menunjukkan bahwa umur 27 tahun, Jenis Kelamin Pria, dan fengan frekuensi cukur 3 kali dalam sebulan, dengan data tersebut menunjukkan bahwa Indikasi minat terhadap ada nya inovasi Aplikasi Pangkasin.

4.1.3 Data Bobot

Dalam Algoritma Logistik Regresi Data Bobot atau disebut juga dengan Optimasi Bobot pada data latih (training) itu sangat penting untuk dapat mencari

nilai terbaik dari setiap parameter model, pada Algoritma penelitian terdapat beberapa bobot yaitu w_1 (umur), w_2 (jenis kelamin), w_3 (frekuensi cukur), dan bias (b). Tujuan nya ialah untuk model dapat memprediksi secara akurat terhadap data yang dimiliki. Pada saat memulai menjalankan algoritma dengan data training, bobot dan bias nya dijalankan dengan acak atau dengan nilai kecil.

Selanjut nya model melakukan iterasi berulang, yang disebut juga namanya model Gradient descent, dalam setiap iterasi nya model melakukan prediksi ke semua data training dengan rumus linier dan rumus sigmoid, mengukur model ketika mendapatkan data kesalahan prediksi (error), dan melakukan update nilai bobot dan bias secara bertahap untuk meminimalkan data salah atau error.

Proses tersebut akan dilakukan terus nilai error atau loss nya mencapai titik minimum atau sudah tidak dapat dirubah kembali secara signifikan, dari itu diperolehnya kombinasi bobot dan bias yang dimana hasil nya atau akurasi nya mendekati data asli. Pada praktik nya secara manual, menghitung bobot dan bias ya itu sangat rumit dan proses yang sangat panjang karena memiliki banyak perhitungan, itu bisa pada bagian fokus penelitian yang lain, oleh karena itu pada bagian penelitian ini peneliti menggunakan nilai bobot w_1 , w_2 , w_3 , dan bias (b), di ambil dari hasil training model library scikit_learning secara otomatis, tersebut telah menerapkan library yang secara otomatis proses optimasi secara efisien dan praktis melalui program Python, dapat mengambil hasilnya untuk perhitungan secara manual.

Catatan: Nilai bias (b) dan bobot, dapat berubah-ubah tergantung pada data latih yang digunakan peneliti, ketika terjadi modifikasi atau perubahan data bobot dan bias nya akan berubah. Intinya optimasi bobot ini adalah proses belajar pada

data, memperbaiki diri setiap kali terjadi nya loss atau error, hingga akhirnya mendapatkan bobot dan bias paling optimal untuk memprediksi minat atau tidak seorang pelanggan terhadap adanya Aplikasi Pangkasin.

Dari Perhitungan dan Code manual dari Algoritma Logistik Regresi ini untuk peneliti dapat memahami alur program Algoritma Logistik Regresi ini jika tanpa Library atau model bawaan dari Bahasa Pemrograman Python, Peneliti dan Pembaca dapat memahami alur manual dari perhitungan dan penelitian ini, Library juga untuk memudahkan program dalam membaca data hasil Observasi dari data Excel untuk dapat membaca keseluruhan data observasi untuk dapat di olah dalam Algoritma Logistik Regresi, menghasilkan data untuk bahan pengembangan Market bisnis Aplikasi Pangkasin.

4.2 Implementasi Algoritma

Algoritma Logistik regresi ini peneliti rancang hanya untuk melakukan pengklasifikasi untuk memilah konsumen potensial dan konsumen tidak potensial untuk pada akhirnya data nya di olah sendiri sama tim Pangkasin. Data ini juga dapat dimanfaatkan untuk ide konten Digital Marketing untuk dapat mendapatkan audiens yang sesuai dengan harapan tim Pangkasin. Ketika melakukan promosi bisa secara terarah langsung ke target atau konsumen potensial dari hasil data klasifikasi dari Algoritma Logistik Regresi.

Program Algoritma Logistik Regresi ini peneliti kembangkan dengan menggunakan Bahasa Pemrograman Python sebagai Bahasa pengembangan untuk menjalankan program nya serta Visual Studio Code sebagai IDE (Integrated Development Environment) yaitu software untuk menjalankan program Programmer menulis, tes, dan debug code pada satu aplikasi.

4.2.1 Code data training dan Hasil

Kode Model Logistik Regresi dibangun dengan 2 model yaitu model perhitungan bias (b) dan bobot setiap fitur otomatis menggunakan library untuk dapat membaca hasil observasi pada data Excel untuk mendapatkan nilai bias dan bobot pada setiap fitur dan model perhitungan secara manual. Dengan gambar pseudocode penentuan bias dan bobot sebagai berikut:

```
BEGIN
// 1. Baca dataset
FILE_PATH ← "C:/Users/HP/Documents/Kuliah/Skripsi/coding/Data_Observasi_149.xlsx"
DATA ← Baca file excel dari FILE_PATH

// 2. Tentukan variabel independen (X) dan variabel dependen (y)
X ← kolom [ 'umur', 'jenis_kelamin', 'frekuensi_cukur' ] dari DATA
y ← kolom [ 'minat_aplikasi' ] dari DATA

// 3. Bagi dataset menjadi data training dan data testing
(X_train, X_test, y_train, y_test) ← Bagi data X dan y
    dengan porsi 70% training dan 30% testing
    stratifikasi berdasarkan y agar seimbang

// 4. Buat model Logistic Regression
MODEL ← LogisticRegression()

// 5. Latih model dengan data training
Latih MODEL menggunakan (X_train, y_train)

// 6. Ambil parameter hasil training
INTERCEPT ← bias (w0) dari MODEL
W1 ← bobot untuk fitur 'umur'
W2 ← bobot untuk fitur 'jenis_kelamin'
W3 ← bobot untuk fitur 'frekuensi_cukur'

// 7. Tampilkan hasil bobot dan bias
CETAK "Bias (intercept) = ", INTERCEPT
CETAK "Bobot umur (w1) = ", W1
CETAK "Bobot jenis kelamin (w2) = ", W2
CETAK "Bobot frekuensi cukur (w3) = ", W3
END
```

Gambar 4. 3 Pseudocode perhitungan bias dan bobot fitur

Pada bagian pertama code terdapat langkah untuk mengambil data dari Data_Observasi_149.xlsx untuk dapat di baca dan mengolah data pada dataframe di Python dan disimpan ke dalam variabel. Data set di atas berisi umur,

jenis_kelamin, frekuensi_cukur, dan minat_aplikasi. terdapat juga untuk pembagian dataset menjadi dua bagian yaitu data training dan data testing, serta terdapat form sklearn linier model dan import Logistik Regresi nya untuk proses pengembangan model logistik regresi nya, untuk mengklasifikasi data berdasarkan fitur yang diberikan.

Dibaris code kedua selanjutnya terdapat perintah pembagian variabel untuk dapat memisahkan x = faktor yang mempengaruhi minat pengguna (umur, jenis kelamin, frekuensi cukur) dan y = minat atau tidak terhadap aplikasi (1 = minat dan 0 = tidak minat) yang digunakan untuk melatih program dan uji program nya dan terdapat `df = pd.read_excel` untuk membaca file_path nya yang dalam bentuk file Excel yang dalam bentuk df atau data frame bentuk tabel.

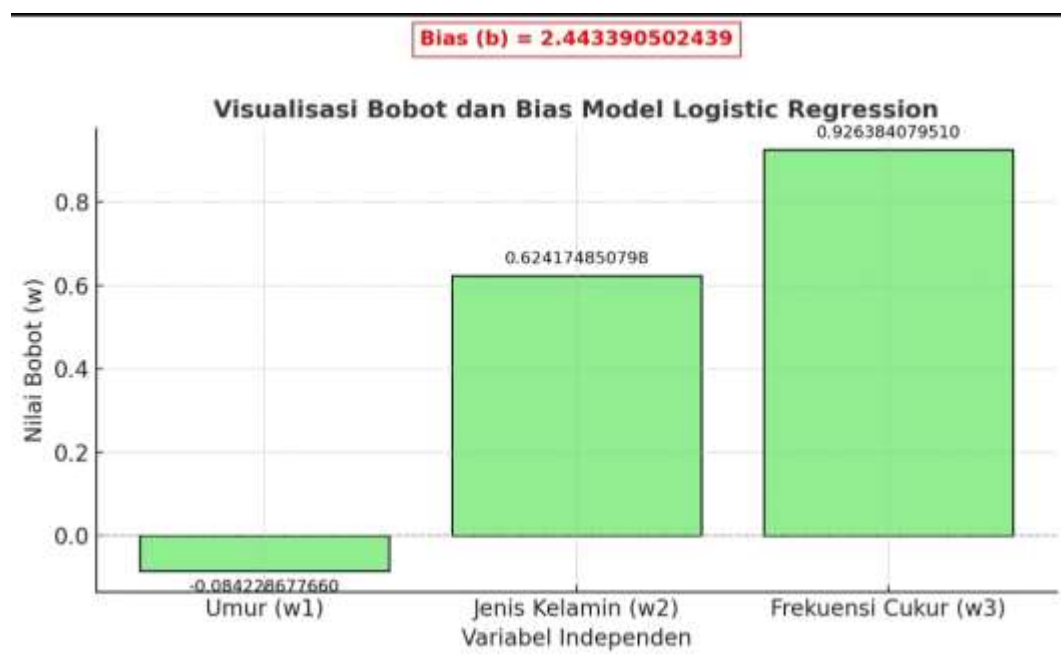
Selanjutnya pada baris code ke tiga terdapat bagian pisahkan bagian x (umur, jenis kelamin, dan frekuensi cukur) dan y (minat aplikasi, 0 = tidak minat dan 1 = minat). Serta pada baris code ke empat yaitu pembagian data training dan data testing, dari perintah `train_test_split` dengan x adalah fitur yang digunakan model untuk pelatihan model dan y sebagai target yang di prediksi model, dengan ukuran split nya itu 70% untuk data training dan 30% untuk data testing nya, dengan menggunakan `random_state = 42` sebagai penetapan pada angka acak yang terus konsisten ketika program di jalankan, serta menggunakan `stratify = y` untuk memastikan pembagian data merata pada distribusi kelas y .

Selanjutnya pada baris code ke empat yaitu membangun model logistik regresi nya untuk melatih data dan membuat prediksi, pada baris code ke enam proses latih model pada data training dengan `model.fit(x_train {fitur melatih model}, y_train {label melatih model})`.

Pada baris code ke lima yaitu data latih untuk ambil bias (intercept) dan bobot nilai data nya dihitung pada saat pelatihan model. Dengan bisa sebagai nilai kostanta yang digunakan dalam meyesuaikan prediksi, dengan w_1 = umur, w_2 = jenis kelamin, dan w_3 = frekuensi cukur.

Pada baris ke 6 bagian code ke delapan yaitu perintah print untuk menampilkan bias, bobot umur, bobot jenis kelamin, dan bobot frekuensi cukur, hasil nya itu dari proses training, dengan hasil nya untuk melengkapi rumus Algoritma Logistik Regresi tersebut.

Dan pada barik ke 7 adalah perintah menampilkan hasil program yaitu sebagai Berikut adalah hasil atau output dari program penentuan bias dan bobot pada setiap fitur, dengan gambar dibawah ini:



Gambar 4. 4 Output bias dan bobot

Dari gambar hasil code program model training diatas, didapatkan nilai bias (b) = 2.44 (2), bobot umur = -0.08, bobot jenis kelamin = 0.62, dan bobot frekuensi cukur = 0.92.

4.2.2 Code data testing dan Hasil

Pada bagian ini peneliti menjelaskan soal Pseudocode data testing dengan menggunakan pethitungan secara manual dengan hanya menggunakan library pandas dan math, tanpa menggunakan sklearn library, dengan gambar sebagai berikut:

```
#Deklarasi variabel:
#   b ← 2           # Bias model
#   w1 ← -0.08      # Bobot untuk umur
#   w2 ← 0.62       # Bobot untuk jenis kelamin
#   w3 ← 0.92       # Bobot untuk frekuensi cukur

#Fungsi sigmoid(z):
#   #return 1 / (1 + e^(-z))

#Fungsi calculate_z(umur, jenis_kelamin, frekuensi_cukur):
#   # z ← b + (w1 * umur) + (w2 * jenis_kelamin) + (w3 * frekuensi_cukur)
#   #return z

#Fungsi prediksi_minat(umur, jenis_kelamin, frekuensi_cukur):
#   # z ← calculate_z(umur, jenis_kelamin, frekuensi_cukur)
#   # prob ← sigmoid(z)
#   # jika prob ≥ 0.5 maka
#   #   pred ← 1           # Prediksi: Berminat
#   # selain itu
#   #   pred ← 0           # Prediksi: Tidak berminat
#   #return (pred, prob)

#Fungsi logistic_regression(data_input):
#   # hasil ← list kosong
#   # untuk setiap baris i dalam data_input lakukan:
#   #   umur ← data_input[i]["umur"]
#   #   jenis_kelamin ← data_input[i]["jenis_kelamin"]
#   #   frekuensi_cukur ← data_input[i]["frekuensi_cukur"]
#   #   (pred, prob) ← prediksi_minat(umur, jenis_kelamin, frekuensi_cukur)
#   #   tambahkan (pred, prob) ke dalam hasil
#   # return hasil
```

Gambar 4. 5 Pseudocode 1

Dengan penjelasan pada baris pertama yaitu import math untuk dapat menghitung rumus Algoritma Logistik Regresi, serta import pandas untuk

membaca data set tabel excel nya dan di proses. Selanjutnya pada baris code kedua menunjukkan hasil dari proses training seblum nya yang menunjukkan bias $(b) = 2$ (untuk membuat data lebih pesimis terhadap hasil prediksi dapat di rubah ke -2), selanjut nya menunjukkan nilai bobot umur = -0.08, bobot jenis kelamin = 0.62, dan bobot frekuensi cukur = 0.92.

Selanjut nya pada baris ke dua dan tiga menjelaskan soal rumus sigmoid untuk merubah hasil linier ke bentuk probabilitas 0 – 1. Selanjut nya pada baris ke empat yaitu menghitung nilai z dengan menggunakan def calculate _z, untuk menghitung nilai z dengan menggunakan rumus Algoritma Logistik Regresi.

Pada baris ke empat yaitu fungsi dengan menggunakan rumus linier untuk prediksi satu data dengan fungsi def prediksi_minat, dengan menginput data umur, jenis kelamin, dan frekuensi cukur, calculate_(z) proses menghitung berdasarkan fitur-fitur tersebut, hasil nya dirubah menggunakan rumus sigmoid yang mana keluaran nya dalam bentuk probabilitas, jika $p > 0.50$ minat tetapi jika $p < 0.50$ tidak minat.

```

#Fungsi logistic_regression(data_input):

#   Buat list kosong bernama hasil

#   Untuk setiap indeks i dari 0 sampai jumlah baris data_input - 1 lakukan:
#       Ambil nilai umur dari data_input baris ke-i
#       Ambil nilai jenis_kelamin dari data_input baris ke-i
#       Ambil nilai frekuensi_cukur dari data_input baris ke-i

#       Panggil fungsi prediksi_minat(umur, jenis_kelamin, frekuensi_cukur)
#       → hasilkan prediksi (pred) dan probabilitas (diabaikan)

#       #Tambahkan nilai pred ke dalam list hasil

#Tambahkan kolom baru ke data_input dengan nama "pred_minat"
#   dan isi dengan nilai-nilai dari list hasil

# Kembalikan (return) data_input yang telah memiliki kolom prediksi

```

Gambar 4. 6 Pseudocode 2

Serta pada baris ke 6 yaitu fungsi memprediksi seluruh dataframe (input) testing yaitu seluruh data testing di iterasi berulang atau looping pada setiap baris data, perintah ambil nilai untuk mengambil nilai setiap fitur untuk setiap data, kemudian di panggil untuk mendapatkan prediksi setiap data, dan hasil proses di simpan dalam hasil[] yang di tambahkan pada Dataframe sebagai kolom pred_minat.

Dan di baris 7 dan 8 terakhir terdapat perintah ["pred_minat"] = hasil, untuk menginput secara otomatis dari model Algoritma Logistik Regresi dalam uji testing model seberapa akurat, karena Machine Learning itu tidak pasti jawaban dikarenakan pada bias dan bobot bisa dirubah tergantung situasi yang di inginkan dari peneliti untuk penyesuaian pada tim Pangkasin.

```

#Fungsi evaluate_metrics(data, kolom_target = "minat_aplikasi", kolom_prediksi = "pred_minat"):

  # Hitung jumlah prediksi berdasarkan kondisi benar dan salah
# TP ← jumlah data di mana kolom_target = 1 dan kolom_prediksi = 1 # True Positive
# TN ← jumlah data di mana kolom_target = 0 dan kolom_prediksi = 0 # True Negative
# FP ← jumlah data di mana kolom_target = 0 dan kolom_prediksi = 1 # False Positive
#FN ← jumlah data di mana kolom_target = 1 dan kolom_prediksi = 0 # False Negative

  # Hitung metrik performa model
# accuracy ← (TP + TN) / (TP + TN + FP + FN)
#jika (TP + FP) ≠ 0 maka
| # precision ← TP / (TP + FP)
#selain itu
| # precision ← 0

# jika (TP + FN) ≠ 0 maka
# recall ← TP / (TP + FN)
# selain itu
# recall ← 0

#jika (precision + recall) ≠ 0 maka
| # f1 ← (2 * precision * recall) / (precision + recall)
#elain itu
| # f1 ← 0

```

Gambar 4. 7 Pseudocode 3

Pada baris code 9 di atas perintah dari `evaluate_metrics (data)` yaitu program evaluasi performa model prediksi Algoritma Logistik Regresi, dengan menghitung metrik evaluasi Akurasi, Presisi, Recall, dan F1-Score. Dengan tujuan mengukur sebaik apa performa dalam memprediksi minat terhadap aplikasi, dengan cara membandingkan antara nilai sebenarnya (`minat_aplikasi`) dan hasil prediksi (`pred_minat`).

Pada baris ke 10 dan 11 code menjelaskan dan menyusun penjelasan istilah seperti TP, TN, FP, dan FN. Selanjutnya perhitungan akurasi, presisi, recall, dan F1-Score dari data yang di ambil.

```

# Hitung support (jumlah data pada masing-masing kelas)
# support_class_1 ← TP + FN      # Jumlah data aktual kelas 1 (Minat)
# support_class_0 ← TN + FP      # Jumlah data aktual kelas 0 (Tidak Minat)

# Tampilkan hasil evaluasi model secara manual
#Cetak "Evaluasi Model Manual"
#Cetak "Akurasi      : ", accuracy (4 angka di belakang koma)
#Cetak "Precision    : ", precision (4 angka di belakang koma)
#Cetak "Recall       : ", recall (4 angka di belakang koma)
#Cetak "F1-Score     : ", f1 (4 angka di belakang koma)
#Cetak "Support Kelas 1 (Minat)      : ", support_class_1
#Cetak "Support Kelas 0 (Tidak Minat): ", support_class_0
#Cetak "TP = ", TP, ", TN = ", TN, ", FP = ", FP, ", FN = ", FN

```

Gambar 4. 8 Pseudocode 4

Pada Baris ke 12 dan 13 yaitu proses hitung support, dengan perintah support class fungsi ini memprediksi semua data yang sebenarnya minat dan tidak minat dari hasil model prediksi. Selanjutnya adalah menunjukkan evaluasi model (secara menyeluruh) yaitu menunjukkan hasil perhitungan dari akurasi, presisi, recall, F1-Score, dan Support. Akurasi yaitu seberapa banyak prediksi benar secara keseluruhan, presisi (seberapa banyak prediksi benar dari semua prediksi positif), recall (seberapa banyak prediksi yang benar dari seluruh data yang harusnya positif), dan F1-Score (rata-rata harmonis dari precision dan recall), dan support (setiap masing-masing kelas dihitung sebagai jumlah data dalam kelas minat dan tidak minat). Fungsi tersebut di simbolkan dengan TP (True Positive), TN (True Negative), FP (False Positive), dan FN (False Negative) yang di ambil dari Confusion matrix.

```

# --- Tahap 1: Acak dan Split Data ---

1. Acak seluruh data dalam variabel df secara random
| - Gunakan random_state = 42 agar hasil acakan tetap konsisten
| - Reset index agar urutan baris kembali berurutan

2. Hitung ukuran data training
| - train_size = 70% dari total jumlah data

3. Bagi dataset menjadi dua bagian:
| - train_df = data dari indeks 0 sampai train_size
| - test_df = data dari indeks train_size sampai akhir

# --- Tahap 2: Prediksi dan Evaluasi Model ---

4. Jalankan fungsi logistic_regression() pada data uji (test_df)
| - Simpan hasil prediksi ke dalam test_df

5. Jalankan fungsi evaluate_metrics() dengan input test_df
| - Fungsi ini akan menghitung akurasi, precision, recall, F1-score, dan support
| - Tampilkan hasil evaluasi performa model di layar

```

Gambar 4. 9 Pseudocode 5

Pada code diatas berisi 2 tahapan yaitu acak dan split data, pada baris pertama seluruh baris pada df diacak agar susunan data tidak mengikuti pola tertentu misal nya urutan berdasarkan nilai umur, dengan menggunakan perintah `random_state = 42`, ini supaya hasil dapat sama terus dengan data yang sama, dengan harapan urutan acak ya tetap sama juga.

Berikutnya hitung set pelatihan data nya sebesar 70% dari jumlah baris data Excel nya, dengan komposisi 70:30 dengan perintah `train_df` di ambil dari indeks 0 hingga sebelum `train_size`, intinya pengacakan memastikan representativitas acak, dan pemotongan 70/30 menyediakan data untuk melatih/evaluasi tanpa terganggu dengan data uji.

Tahap selanjut nya algoritma logistik regresi yang di jalan kan di atas data uji (`test_df`) dengan hasil nya nanti label prediksi apakah minat atau tidak minat, dan probabilitas keyakinan model, dalam konteks Pseudocode fungsi ini untuk

menerapkan bobot bias pada tiap baris, hasil nya Pred prob tersebut di masukkan pada test_df, sehingga data uji memiliki nilai aktual (minat aplikasi) berdampingan dengan hasil prediksi dan probabilitas.

Terakhir pada baris terakhir fungsi evaluate_metrics() di panggil dengan dengan fungsi test_df. Dimana fungsi membandingkan kolom target aktual (minat_aplikasi) dengan kolom prediksi untuk menghitung TP, TN, FP, dan FN. Dari keempat variabel itu di hitung untuk menghasilkan performa model data uji.

```
# --- Tahap 3: Prediksi Data Baru (User Baru) ---

6. Tampilkan teks "Prediksi Data Baru" di layar

7. Tentukan nilai input pengguna baru:
   - umur_baru = 30
   - jenis_kelamin_baru = 1      (misal 0 = laki-laki, 1 = perempuan)
   - frekuensi_cukur_baru = 2    (misal frekuensi cukur per minggu)

8. Jalankan fungsi prediksi_minat() dengan input:
   (umur_baru, jenis_kelamin_baru, frekuensi_cukur_baru)

9. Simpan hasilnya sebagai:
   - pred → hasil prediksi minat (1 = berminat, 0 = tidak berminat)
   - prob → nilai probabilitas hasil prediksi

10. Tampilkan hasil prediksi ke layar:
    - Cetak: "Prediksi Minat: [pred] (Probabilitas: [prob])"
```

Gambar 4. 10 Pseudocode 6

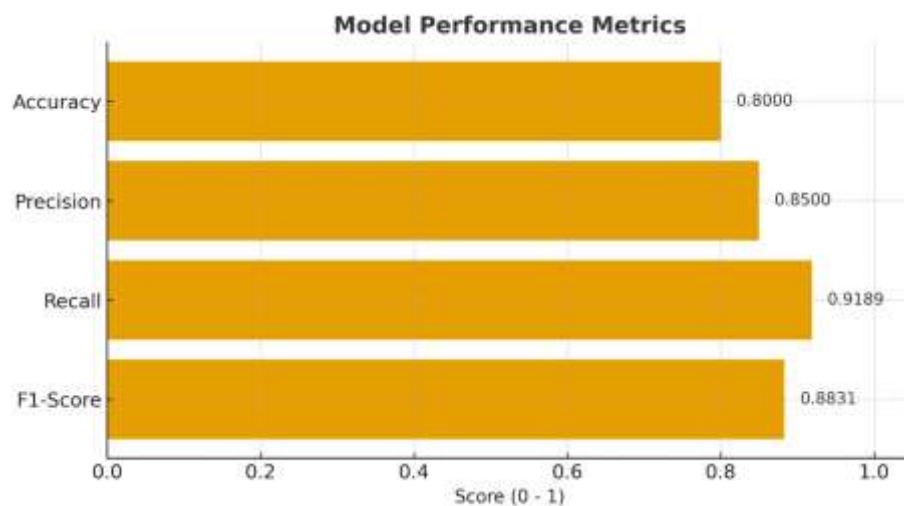
Pada Tahap 3 di atas yaitu prediksi data baru menjelaskan bagaimana model yang sudah di latih digunakan untuk menginferensikan minat seorang pengguna yang belum terdata, selanjut nya di atas menggunakan model pada inputan baru, ditetapkan lah satu orang pengguna dengan umur 30 tahun, jenis kelamin laki-laki, dan frekuensi cukur rambut 2 kali dalam sebulan, ini sebagai form baru uji untuk data baru saja.

Selanjut nya ketiga nilai tersebut di kirim ke fungsi prediksi, dengan di balik nya menggunakan skor linier, lalu merubah nya ke fungsi sigmoid, yaitu probabilitas jika di atas $>50\%$ maka minat (1) dan $<50\%$ tidak minat (0). Fungsi mengembalikan dua hal sekaligus yaitu prediksi label kelas dan probabilitas (Keyakinan model terhadap keputusan).

Terakhir hasil nya itu disimpan pada variabel dan di tampilkan dalam bentuk nol koma dirubah ke bentuk desimal sehingga mudah di baca dan di pahami, tahapan ini diperluas untuk pengguna (batch) asalkan setiap input baru melewati pra-pemrosesan data latih yang sama. Sehingga setiap prediksi tetap valid dan konsisten.

4.2.3 Output code

Hasil dari pemrograman sebelum nya menghasilkan output dengan menunjukkan Evaluasi Model bekerja, Berikut adalah gambar dari output Code hasil dari program Logistik regresi:



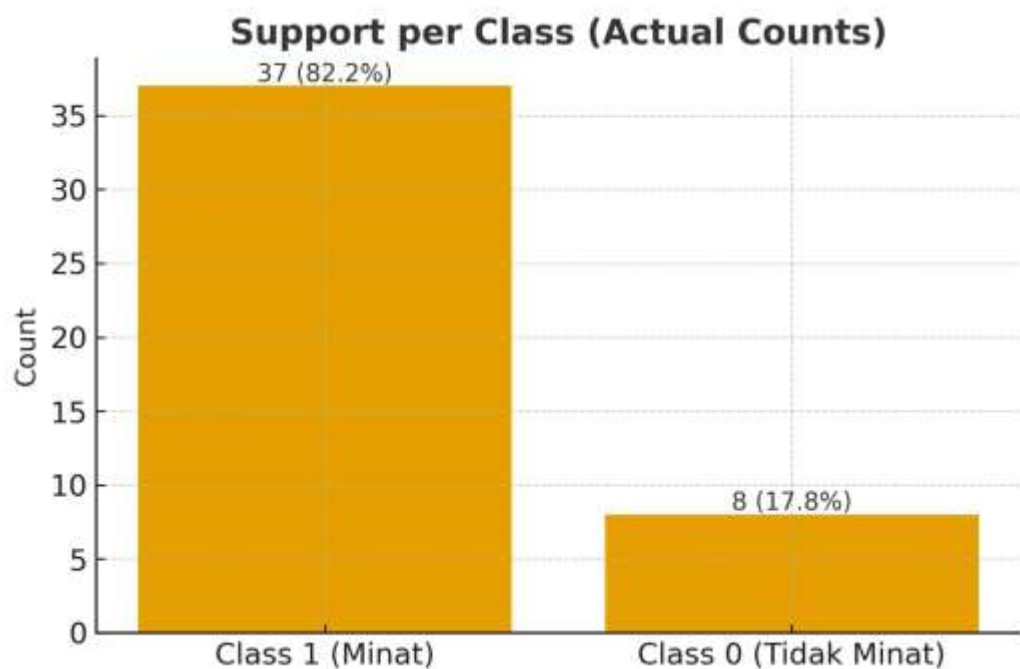
Gambar 4. 11 EDA Metriks Performa Model

New Data Prediction

Predicted Interest (class): 1
Probability: 0.8085

TP=34, TN=2, FP=6, FN=3

Gambar 4. 12 EDA Prediksi Data Baru



Gambar 4. 13 EDA Jumlah aktual per kelas

Dari gambar diatas hasil nya menunjukkan Evaluasi Model Manual keseluruhan data testing nya, untuk akurasi : 0.80, yang berarti memprediksi benar 80% dari data yang diuji, dengan baik minat dan tidak minat), untuk precission 0.86, yang berarti banyak prediksi positif yang benar (minat) dibandingkan

dengan total prediksi positif yang dilakukan oleh model, jika model tidak banyak menghasilkan minat yang banyak, dengan rumus sebagai berikut:

$$Precision = \frac{TP}{TP + FP}$$

Gambar 4. 14 Rumus Precision

Serta dengan Recall = 0.91, recall mengukur seberapa banyak prediksi positif yang benar (minat) dibandingkan dengan total dengan data yang benar-benar positif (minat). Jika recall tinggi itu dikarenakan model tidak dapat membuang data yang seharusnya positif yang seharusnya minat. Dengan rumus sebagai berikut:

$$Recall = \frac{TP}{TP + FN}$$

Gambar 4. 15 Rumus Recall

Selanjutnya adalah F1-Score = 0.88, yaitu nilai seimbang dengan Precision dan Recall, dengan rumus sebagai berikut:

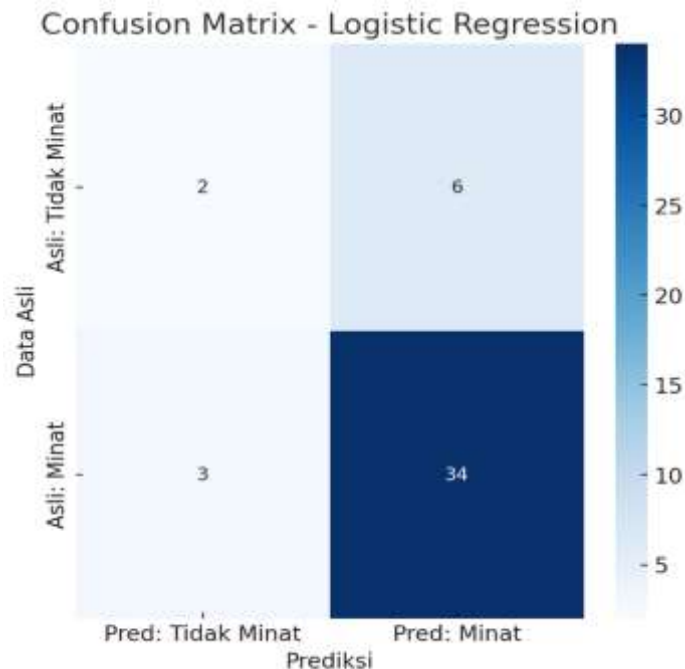
$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Gambar 4. 16 Rumus F1-Score

Serta dengan Support, terdapat 37 yang data minat dan 8 data yang tidak minat setelah melakukan uji dari program.

Dengan Confusion Matrix TP = 31, yaitu data yang benar-benar minat dan diprediksi minat, TN = 7, yaitu data yang benar-benar tidak minat dan di prediksi tidak minat, FP = 6, yaitu data yang seharusnya tidak minat tetapi di prediksi model minat, dan FN = 6, yaitu data yang harus nya minat tetapi di prediksi tidak minat.

Dengan Gambar Grafik Confusion Matrix sebagai berikut:



Gambar 4. 17 Grafik Confusion Matrix

Selanjutnya pada prediksi data baru, diprediksi minat, karena dengan inputan umur = 30, jenis kelamin = 1 (wanita), dan frekuensi cukur = 2 kali, di klasifikasikan model dalam minat, dengan probabilitas 0.88 / 88%.

4.3 Data Observasi

Data berikut adalah hasil dari wawancara dengan beberapa Barbershop di kota Medan dengan meminta Data set referensi pengunjung barbershop tersebut. Dengan data yang di perlukan untuk Algoritma Logistik Regresi ini adalah Umur (Tahun), Jenis kelamin (0 = Pria dan 1 = Wanita), frekuensi cukur (1 kali, 2 kali, 3 kali, 4 kali dalam waktu bulan), Serta minat Aplikasi (1 Setuju dan 0 Tidak Setuju). Total data set nya adalah sebanyak 149 data dari yang peneliti dapatkan dari 8 Barbershop di sekitaran Kota Medan serta beberapa data yang diperoleh secara mandiri dengan melakukan wawancara langsung ke beberapa konsumen barbershop.

Data berikut dalam file Excel, yang kemudian di integrasi kan dengan Visual Studio Code nya, membaca program setelah itu di eksekusi menjadi data olah menghasilkan data yang di perlukan tim Pangkasin untuk kebutuhan internal mereka.

4.3.1 Code Exploratory Data Analysis

Data observasi juga di integrasikan dengan EDA (Exploratory Data Analysis) dengan menggunakan Tools Visual Studio Code, serta menggunakan Bahasa Pemrograman Python untuk dapat menjelaskan secara visual data observasi secara keseluruhan, dengan Pseudocode sebagai berikut:

```
#PROGRAM AnalisisDataObservasi

# // --- 1) Persiapan ---
# DECLARE df                // tabel data
#DECLARE rata_umur          // float
#DECLARE persentase_pria    // float (dalam %)
#DECLARE rata_frekuensi_cukur // float
#DECLARE persentase_minat   // float (dalam %)
#DECLARE grouped            // tabel ringkasan

// --- 2) Input Data ---
#df ← READ_EXCEL("C:/Users/HP/Documents/Kuliah/Skripsi/coding/Data_Observasi_149.xlsx")

// --- 3) Statistik Deskriptif ---
#rata_umur ← MEAN( df["umur"] )

// asumsikan: jenis_kelamin = 0 (pria), 1 (wanita)
#persentase_pria ← MEAN( df["jenis_kelamin"] == 0 ) * 100

#rata_frekuensi_cukur ← MEAN( df["frekuensi_cukur"] )

// asumsikan: minat_aplikasi = 1 (minat), 0 (tidak)
#persentase_minat ← MEAN( df["minat_aplikasi"] == 1 ) * 100

#PRINT "Rata-rata Umur:", rata_umur
#PRINT "Persentase Pria (%)": persentase_pria
#PRINT "Rata-rata Frekuensi Cukur:", rata_frekuensi_cukur
#PRINT "Persentase Minat Aplikasi (%)": persentase_minat
```

Gambar 4. 18 Gambar Code Exploratory Data Analysis

Pada baris Pseudocode pertama terdapat gambar merangkum alur program Analisis_observasi dengan tugas nya membaca data survei setelah itu membuat

statistik dasar dan perbandingan status minat terhadap aplikasi, bagian awal terdapat deklarasi variabel: df sebagai tabel data (dataset utama), kemudian 4 variabel float untuk tampung hasil perhitungan rata-rata umur, jenis kelamin yang sering, frekuensi cukur nya, dan persentase minat. Terdapat juga grouped sebagai tabel ringkasan hasil pengelompokkan. setelah itu masuk ke bagian input data, program membaca data excel pada path yang di tentukkan dan menyimpannya pada path.

Selanjutnya tahap statistik deskriptif, disini program menghitung rata-rata umur dari kolom umur untuk mengetahui proporsi pria dan wanita, lalu mengambil rata-rata nya frekuensi cukur, serta menghitung minat aplikasi == 1 dikalikan 100. Keempat nilai ini nantinya ditampilkan sebagai ringkasan sempel, setelah statistik dasar program dapat membuat ringkasan per kelompok minat: data di kelompokkan berdasarkan minat_aplikasi lalu dihitung rata-rata untuk kolom umur, jenis-kelamin, dan frekuensi_cukur pada setiap kelompok. Karena rata-rata setiap jenis_kelamin disini berupa proporsi, nilai nya nantinya di kalikan 100, agar terbaca sebagai persentase dalam ringkasan.

Tabel ringkasan itu di cetak untuk membandingkan karakteristik responden yang minat dan tidak minat (misal nya apakah kelompok yang umur nya muda lebih sering cukur rambut), secara keseluruhan Pseudo code ini menunjukkan struktur yang rapi deklarasi – baca data – hitung statistik kunci – kelompokan dan konversi proporsi – tampilkan hasil, tanpa ubah data asli.

```

// --- 4) Ringkasan per status Minat ---
#grouped <- GROUP_BY(df, key="minat_aplikasi")
#      .AGGREGATE( MEAN, columns=["umur","jenis_kelamin","frekuensi_cukur"] )

// ubah kolom jenis_kelamin pada ringkasan menjadi persen (0-100)
#FOR EACH row IN grouped DO
#   row["jenis_kelamin"] <- row["jenis_kelamin"] * 100
#END FOR

#PRINT "\nRingkasan Minat vs Tidak Minat:"
#PRINT grouped

// --- 5) Visualisasi: Histogram Umur ---
#FIGURE_NEW()
#HISTOGRAM(
#   data=df["umur"],
#   bins=10,
#   bar_color="skyblue",
#   edge_color="black"
# )
#TITLE("Distribusi Umur")
#XLABEL("Umur")
#YLABEL("Jumlah")
#GRID(ON)
#SHOW_FIGURE()

#END PROGRAM

```

Gambar 4. 19 Pseudocode Exploratory Data Analysis

Pada baris code diatas program melanjutkan tahap ringkasan untuk status minat dan visualisasi. Pertama data utama (df) dikelompokkan berdasarkan kolom minat_aplikasi, dari tiap kelompok (minat = 1 dan tidak minat = 0) di hitung rata-rata untuk ketiga kolom: umur, jenis kelamin, dan frekuensi cukur. Hasil nya nantinya di simpan pada ringkasan grouped. Karena kolom jenis kelamin dikodekan biner (0/1).

Nilai rata-rata presentasikan proporsi; supaya lebih informatif, setiap grouped kemudian diiterasikan dan nilai jenis kelamin x 100 sehingga bentukan nya (0x100%). Setelah di konversi program mencetak hasil “ringkasan minat vs tidak minat” diikuti tabel grouped agar berbeda karakteristik nya.

Bagian berikut nya adalah visualisasi distribusi umur. Program membuat EDA (Exploratory Data Analysis) bentuk tabel histogram dari kolom umur

dengan 10 bin, batang histogram berwarna sky blue dan garis tepi nya hitam. Dengan judul distribusi umur, lalu sumbu x umur, sumbu y jumlah, serta grid di aktifkan agar skala frekuensi mudah di baca. Terakhir grafik atau EDA nya langsung muncul setelah program nya jalan.

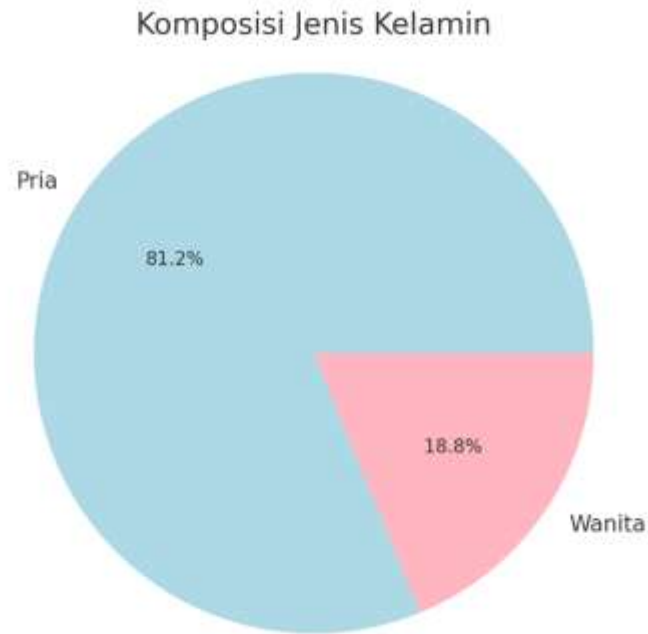
4.3.2 Output code Exploratory Data Analysis

Rata-rata umur pada data observasi sekitar 31 tahun, dengan persentase pria 81%, dengan rata-rata frekuensi cukur 2 kali dalam sebulan, dan dengan persentase minat aplikasi sebesar 83%.

Dengan ringkasan tidak minat, dengan umur 36 tahun, jenis kelamin (Pria dan Wanita) 12% yang tidak minat dan frekuensi cukur 1,64 kali dalam sebulan. ringkasan minat, dengan umur 29 tahun, Jenis kelamin (pria dan wanita) 20.16%, dengan frekuensi cukur 2.31 kali perbulan, dengan Kesimpulan yang berminat terhadap aplikasi lebih muda , lebih sering cukur, dan mendominasi pria, yang tidak minat cenderung lebih tua, jarang cukur , dan Sebagian besar Wanita. Itu merupakan hasil dari analisi program EDA Python.

Berikut adalah tampilan dari visualisasi grafik hasil dari data set observasi Excel:

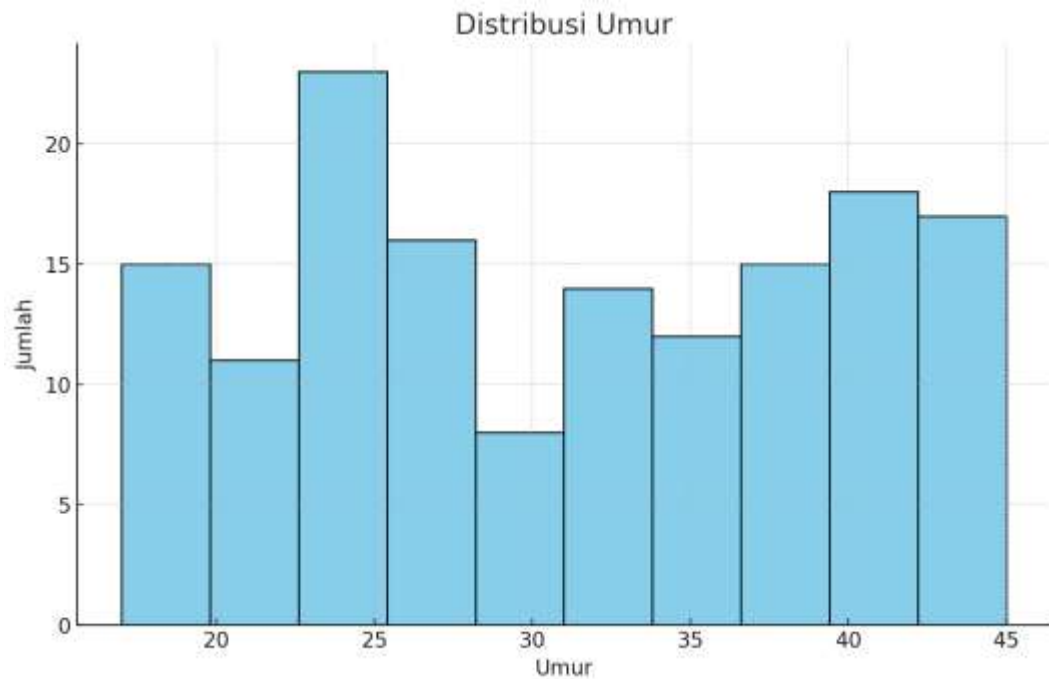
1. Gambar Komposisi Umur, dengan bentuk lingkaran Pie Chart, Pria 81.2% mendominasi dari pada Wanita sebesar 18%, dengan chart sebagai berikut:



Gambar 4. 20 Komposisi Umur

Dari visualisasi ini menunjukkan bahwa pria yang warna biru muda mendominasi dari pada Wanita dengan warna merah muda. memang untuk bisnis barbershop memang dibuktikan konsumen potensial nya adalah pria.

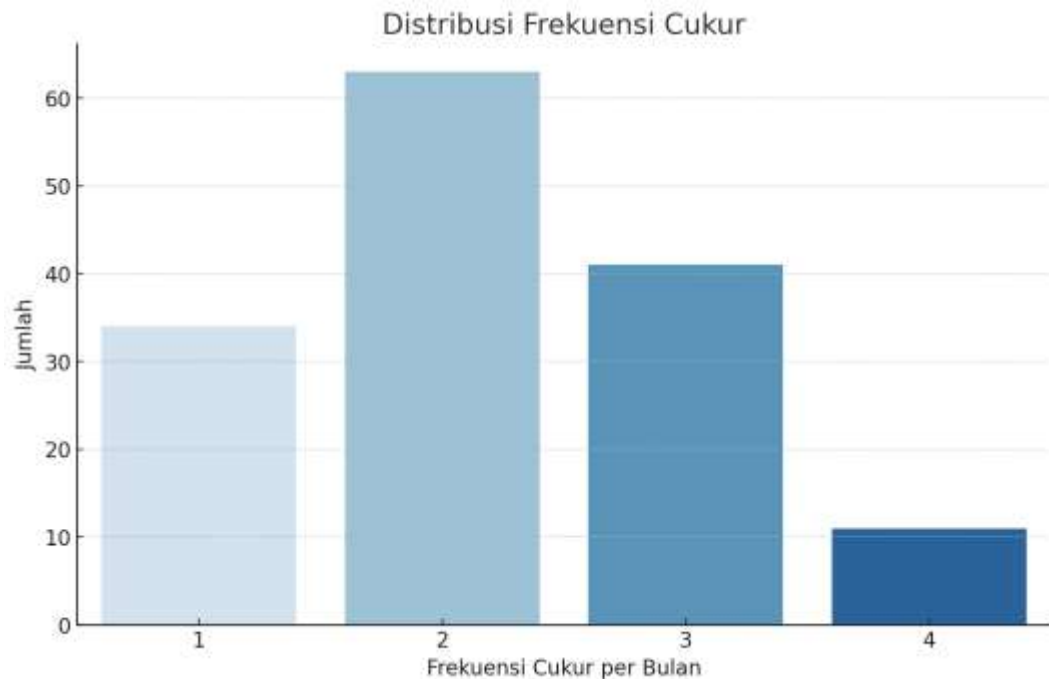
2. Gambar Distribusi Umur, Data ini berbentuk anak tangga Histogram untuk menunjukkan variabel umur mana yang lebih tinggi, dengan chart sebagai berikut:



Gambar 4. 21 Distribusi Umur

Dari gambar di atas menunjukkan bahwa umur 20 - 25 tahun untuk pengunjung barbershop terbanyak sesuai dengan data set Excel dengan angka rata-rata nya 20 orang, serta di umur terendah nya terdapat pada 30 tahun yang sedikit ke barbershop untuk cukur rambut dengan jumlah sekitar 5 orang. Dari data distribusi umur yang sering di temukan umur-umur di atas tapi tidak menutup kemungkinan ada umur di atas dari angka 45 tahun.

3. Gambar distribusi frekuensi cukur, yang dalam bentuk Bar Chart, karena hasil observasi hanya 4 kali yang ditemukan cukur rambut dalam sebulan, dengan grafik sebagai berikut:



Gambar 4. 22 Distribusi Frekuensi Cukur

Dari Visualisasi data Bar Chart di atas menunjukkan mendominasi 2 kali sebulan pangkas dengan 60 orang lebih di temukan dari pada yang data terendahnya 4 kali dalam sebulan yang hanya 10 orang, serta yang 1 kali sebulan 30 orang lebih dan 3 kali sebulan 40 orang lebih.

Kesimpulan pada perbedaan hasil Algoritma Logistik Regresi ini dengan EDA (Exploratory Data Analysis) ialah pada Algoritma Logistik Regresi dibagi data nya pada data training dan data testing yaitu 70% dan 30% berbeda dengan data EDA yang memvisualisasi keseluruhan tanpa ada nilai akurasi, recall, precision, dan support nya.

4.4 Hasil Evaluasi Model

Hasil dari pengembangan Model Algoritma Logistik Regresi peneliti mendapatkan hasil yang dimana Klasifikasi Logistik Regresi ini dapat mengidentifikasi calon pengguna potensial Aplikasi Pangkasin nantinya dengan tingkat akurasi mencapai 80%. Informasi ini juga memberikan gambaran

karakteristik pengguna yang cenderung minat dengan adanya Aplikasi Pangkasin ini, yang mana hasil nya menunjukkan umur 20 – 35 tahun, Pria, dengan frekuensi cukur rambut antara 2 – 3 kali dalam sebulan, dapat menjadi target potensial yang akan menggunakan Aplikasi Pangkasin ini, hasil dari evaluasi Akurasi, Recall, dan F1-Score nya dapat bekerja dengan baik dalam klasifikasi pengguna yang berminat.

Dengan penjelasan kesimpulan akhir dari data testing dalam Tabel sebagai berikut:

Metrik Evaluasi	Nilai	Penjelasan
Akurasi	0.80	Persentase yang di baca benar oleh model terhadap seluruh data pengujian.
Recall	0.91	Persentase model dalam menemukan data yang benar-benar minat.
Precision	0.85	Persentase Proporsi prediksi “minat” yang benar-benar akurat secara prediksi.
F1-Score	0.88	Rata-rata antara precision dan recall, jika persentase semakin tinggi semakin

		seimbang data nya.
Support	45	Jumlah data uji keseluruhan menyesuaikan kelas minat dan tidak minat terhadap aplikasi dari data uji.

Tabel 4. 1 Tabel Kesimpulan Evaluasi Model

Dari tabel di atas hasil dari pemrograman Python Algoritma Logistik Regresi, ini dapat berubah-ubah sewaktu ada melakukan observasi kembali, dan perubahan pada bias dan bobot pada setiap fitur nya.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

1. Logistik Regresi terbukti efektif dan bermanfaat dalam mengklasifikasi calon pengguna Aplikasi Pangkasin nantinya dari hasil analisis data pelanggan barbershop di sekitaran Kota Medan, terkhusus dilihat dari segi Umur, Frekuensi cukur rambut, dan Jenis kelamin pelanggan, yang dapat berpotensi menjadi calon pengguna aplikasi Pangkasin.
2. Model Algoritma Logistik di bangun dengan menggunakan hanya 3 variabel yaitu Umur, Frekuensi cukur, dan jenis kelamin, ini memiliki korelasi terhadap minat pengguna Aplikasi nantinya.
3. Dari 149 data set observasi yang mencapai 84%, menunjukkan kinerja Algoritma cukup baik yang membuktikan dapat di pakai dalam Prediksi Pasar.
4. Penggunaan Logistik Regresi ini dapat menjadi Indikator dalam penentuan Marketing suatu bisnis terkhusus Tim Pangkasin dalam membaca Pasar serta aksi Tim dalam melakukan tindakan bisnis nya terhadap perkembangan bisnis Startup Pangkasin.

5.2 Saran

1. Dalam pengembangan nya selanjut nya penambahan Variabel mungkin akan lebih baik, seperti penambahan Pekerjaan, Lokasi Geografis, dan Penghasilan Konsumen, untuk meningkatkan akurasi, serta menggambarkan keinginan pasar yang lebih spesifik.

2. Penambahan kuantitas dalam observasi mungkin di perluas lagi, seperti ke Deli Serdang, Tanjung Morawa, dan sebagai nya, untuk mendapatkan data yang lebih luas dan Komprehensif.
3. Implementasi Algoritma Logistik Regresi mungkin bisa di masukkan pada Backend Aplikasi Pangkasin bisa langsung dapat data dari pengguna aplikasi, tanpa harus melakukan secara manual observasi nya.
4. Dan untuk penelitian selanjutnya mungkin dapat memberikan rekomendasi layanan berdasarkan klasifikasi minat pelanggan, untuk meningkatkan loyalitas pengguna Aplikasi nantinya.

DAFTAR PUSTAKA

- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.
- Telkomuniversity.ac.id. (2024, Oktober 28). Memahami Apa Itu Aplikasi dan Fungsinya di Era Digital. Diakses pada 7 April 2025, dari <https://dif.telkomuniversity.ac.id/memahami-apa-itu-aplikasi-dan-fungsinya-di-era-digital/>
- Alfayulanda, P., & Murni, D. (2024). Penerapan Analisis Regresi Logistik Biner Pada Faktor-Faktor Yang Mempengaruhi Keputusan Mahasiswa Berbelanja Online Di Tiktok. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 5(1), 178–189. <https://doi.org/10.46306/lb.v5i1.547>
- Alwin, M. I., & Aditiawan, F. P. (2022). Aplikasi Layanan Jasa Tukang Berbasis Website. *INTEGER: Journal of Information Technology*, 7(2), 93–99. <https://doi.org/10.31284/j.integer.2022.v7i2.3298>
- Ameliya, A., Nainggolan, D. O., & Pinem, H. G. (2024). Analisis Kemampuan Mahasiswa Matematika FMIPA Unimed dalam Menyelesaikan Pertidaksamaan Nilai Mutlak dengan Berbantuan Python. 2(5).
- Badan Pengembangan dan Pembinaan Bahasa, K. P. dan K. R. I. (2016). *Implementasi*. <https://kbbi.kemdikbud.go.id/entri/Implementasi>
- Bogdanchikov, A., Zhaparov, M., & Suliyev, R. (2013). Python to learn programming. *Journal of Physics: Conference Series*, 423(1). <https://doi.org/10.1088/1742-6596/423/1/012027>
- Danardatu, A. H. (2019). Pengenalan Customer Relationship Management (CRM). *Jurnal Ilmiah Teknologi Informasi Asia*, 1–5.
- Faisal, F., Sari, I. R., Saraswati, I., & Heikal, J. (2024). Analisis Pengaruh Karakteristik Produk terhadap Niat Beli Ulang Pelanggan menggunakan Metode Regresi Logistik Biner. *AKADEMIK: Jurnal Mahasiswa Humanis*, 4(3), 1182–1190. <https://doi.org/10.37481/jmh.v4i3.1064>
- Gunawan, R., Aulia, S., Supeno, H., Wijanarko, A., Uwiringiyimana, J. P., & Mahayana, D. (2021). Adiksi Media Sosial dan Gadget bagi Pengguna Internet di Indonesia. *Techno-Socio Ekonomika*, 14(1), 1–14. <https://doi.org/10.32897/techno.2021.14.1.544>
- Hendrian, S. (2018). Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan. *Faktor Exacta*, 11(3), 266–274. <https://doi.org/10.30998/faktorexacta.v11i3.2777>
- Irma Devi Lestari. (2016). 196927-ID-klasifikasi-online-dan-google. *Iqra'*, 83–

94.

- Khotimah, K., & Febriansyah, F. (2018). Pengaruh kemudahan penggunaan, kepercayaan konsumen dan kreativitas iklan terhadap minat beli konsumen online-shop. *Jurnal Manajemen Strategi Dan Aplikasi Bisnis*, 1(1), 19–26. <https://doi.org/10.36407/jmsab.v1i1.16>
- Kurniasih, D., Rusfiana, Y., Subagyo, A., & Nuradhawati, R. (2021). Teknik Analisa. *Alfabeta Bandung*, 1–119. www.cvalfabeta.com
- Naïve, A., Classifier, B., & Penyakit, P. (2024). Analisis klasifikasi menggunakan regresi logistik biner dan algoritma naïve bayes classifier pada penyakit hipertensi 1,2,3. 13(2007), 319–327. <https://doi.org/10.14710/j.gauss.13.2.319-327>
- Ndangi, W. R. A., Resmawan, R., & Djakaria, I. (2019). Perbandingan Analisis Diskriminan dan Regresi Logistik Multinomial. *Jambura Journal of Mathematics*, 1(2), 54–63. <https://doi.org/10.34312/jjom.v1i2.2100>
- Tran, H., Le, N., & Nguyen, V. H. (2023). Customer Churn Prediction in the Banking Sector Using Machine Learning-Based Classification Models. *Interdisciplinary Journal of Information, Knowledge, and Management*, 18(May), 87–105. <https://doi.org/10.28945/5086>
- Van Rossum, G., & Drake, F. L. . (2006). *An introduction to Python, release 2.5*.
- Wibawa, A. P., Guntur, M., Purnama, A., Fathony Akbar, M., & Dwiyanto, F. A. (2018). Metode-metode Klasifikasi. *Prosiding Seminar Ilmu Komputer Dan Teknologi Informasi*, 3(1), 134–138.

LAMPIRAN

Lampiran 1 Data Set Excel 149 Data

umur	jenis_kelamin	frekuensi_cukur	minat_aplikasi
23	0	2	1
27	0	3	1
20	0	3	1
39	1	1	0
18	0	2	1
24	0	2	1
45	0	2	0
37	0	3	1
23	0	1	1
25	1	2	1
35	0	2	1
29	0	2	0
27	0	3	1
27	0	3	1
40	0	4	1
37	1	2	0
20	0	2	1
24	0	2	1
40	0	4	1
19	1	1	1
38	0	2	1
37	0	1	1
18	1	2	1
40	0	2	1
28	1	2	1
22	0	2	1
18	0	3	1
44	1	2	1
37	0	1	0
17	0	2	1
31	0	2	1
42	0	2	0
38	0	2	1
45	0	2	1
28	0	2	1
41	0	3	1
33	0	1	1
43	1	3	1
43	0	2	1
26	0	2	1
44	0	2	0
36	1	2	0
32	0	3	1
31	0	1	1
31	0	1	0
35	0	3	1

28	0	1	1
39	0	1	0
36	0	2	1
41	0	1	0
19	1	3	1
21	0	1	1
35	0	2	0
23	0	2	1
37	0	3	1
25	0	2	1
23	0	1	1
34	0	4	1
20	0	3	1
41	0	3	1
44	0	2	1
30	0	1	1
34	0	2	1
42	0	2	0
25	0	2	0
42	1	3	1
37	0	2	0
18	0	4	1
36	0	2	1
44	1	3	1
31	0	1	1
44	0	2	1
23	0	3	1
28	0	1	1
45	0	1	1
24	0	4	1
31	0	3	1
19	0	1	1
30	1	2	1
33	0	2	0
20	1	3	1
34	0	2	1
24	1	1	1
20	1	1	1
18	0	2	1
22	0	3	1
38	1	1	1
26	0	2	1
20	0	2	1
38	0	2	1
45	1	1	1
34	1	1	1
42	0	4	1

28	0	3	1
18	1	2	1
26	0	3	0
20	0	2	1
30	1	3	1
32	0	1	1
31	0	2	1
24	0	1	1
30	0	1	1
39	1	2	1
44	0	1	0
41	0	1	0
24	1	2	1
37	0	4	1
32	0	1	1
29	0	1	0
34	0	3	1
31	0	3	1
37	1	2	1
40	0	2	1
42	0	3	1
41	0	1	0
44	0	4	1
44	0	3	1
44	0	2	1
29	1	1	1
25	0	4	1
45	0	2	1
31	0	2	1
29	0	3	1
17	0	3	1
41	0	4	1
23	1	4	1
25	0	3	1
40	0	3	1
17	0	3	1
28	0	3	1
24	0	2	1
40	0	1	0
27	0	3	1
35	0	3	0
33	0	3	1
24	0	3	1
19	0	3	1
19	1	2	1
17	0	2	1
43	0	2	1

21	0	2	1
26	0	1	0
29	0	3	1
42	0	2	0
25	0	3	1
26	1	2	1
29	0	3	1
25	0	2	1
28	0	2	1

Lampiran 2 Code training Penentuan Bias dan bobot

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression

# Baca data
file_path = 'C:/Users/HP/Documents/Kuliah/Skripsi/coding/Data_Observasi_149.xlsx'
df = pd.read_excel(file_path)

# Pisahkan fitur (X) dan label (y)
X = df[['umur', 'jenis_kelamin', 'frekuensi_cukur']]
y = df['minat_aplikasi']

# Bagi data menjadi training dan testing (70% untuk training)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42, stratify=y)

# Bangun model Logistic Regression
model = LogisticRegression()

# Latih model pada data training
model.fit(X_train, y_train)

# Ambil bobot dan bias hasil training
b = model.intercept_[0] # Bias (intercept)
w1 = model.coef_[0][0] # Bobot untuk umur
w2 = model.coef_[0][1] # Bobot untuk jenis_kelamin
w3 = model.coef_[0][2] # Bobot untuk frekuensi_cukur

# Tampilkan bobot dan bias yang dihasilkan
print(f"Bias (b): {b}")
print(f"Bobot untuk umur (w1): {w1}")
print(f"Bobot untuk jenis kelamin (w2): {w2}")
print(f"Bobot untuk frekuensi cukur (w3): {w3}")
```

Lampiran 3 Out put dari code penentuan Bias dan Bobot

```
PS C:\Users\HP\Documents\Latihan Coding> & C:/Users/HP/AppData/Local/Programs/Python/Python313/python.exe "C:/Users/HP/Documents/
Kuliah/Skripsi/coding/Program terdapat/Penentuan bias bobot logregresi.py"
Bias (b): 2.4433905024389224
Bobot untuk umur (w1): -0.88422867766829572
Bobot untuk jenis kelamin (w2): 0.6241748587977134
Bobot untuk frekuensi cukur (w3): 0.9263844079510141
```

Lampiran 4 code program data testing Algoritma Logistik Regresi

```
import math
import pandas as pd

# Bobot model hasil training manual
b = 2 # Bisa coba ganti dengan -4 untuk lebih pesimis
w1 = -0.08
w2 = 0.62
w3 = 0.92

# Fungsi sigmoid
def sigmoid(z):
    return 1 / (1 + math.exp(-z))

# Hitung nilai z dari input dan rumus algoritma logistik regresi
def calculate_z(umur, jenis_kelamin, frekuensi_cukur):
    return b + w1 * umur + w2 * jenis_kelamin + w3 * frekuensi_cukur

# Fungsi prediksi untuk satu orang
def prediksi_minat(umur, jenis_kelamin, frekuensi_cukur):
    z = calculate_z(umur, jenis_kelamin, frekuensi_cukur)
    prob = sigmoid(z)
    pred = 1 if prob >= 0.5 else 0
    return pred, prob

# Fungsi prediksi untuk seluruh DataFrame (dalam testing)
def logistic_regression(df_input):
    hasil = []
    for i in range(len(df_input)):
        umur = df_input.iloc[i]["umur"]
        jenis_kelamin = df_input.iloc[i]["jenis_kelamin"]
        frekuensi_cukur = df_input.iloc[i]["frekuensi_cukur"]

        pred, prob = prediksi_minat(umur, jenis_kelamin, frekuensi_cukur)
        hasil.append(pred)

    df_input["pred_minat"] = hasil
    return df_input
```

Lampiran 5 code program data testing Algoritma Logistik Refresi 2

```
# Fungsi prediksi untuk seluruh DataFrame (dalam testing)
def logistic_regression(df_input):
    hasil = []
    for i in range(len(df_input)):
        umur = df_input.iloc[i]["umur"]
        jenis_kelamin = df_input.iloc[i]["jenis_kelamin"]
        frekuensi_cukur = df_input.iloc[i]["frekuensi_cukur"]

        pred, _ = prediksi_minat(umur, jenis_kelamin, frekuensi_cukur)
        hasil.append(pred)

    df_input["pred_minat"] = hasil
    return df_input
```

Lampiran 6 code program data testing Algoritma Logistik Regresi 3

```
# Evaluasi performa model
def evaluate_metrics(df, target_col="minat_aplikasi", pred_col="pred_minat"):
    TP = ((df[target_col] == 1) & (df[pred_col] == 1)).sum() # True Positive
    TN = ((df[target_col] == 0) & (df[pred_col] == 0)).sum() # True Negative
    FP = ((df[target_col] == 0) & (df[pred_col] == 1)).sum() # False Positive
    FN = ((df[target_col] == 1) & (df[pred_col] == 0)).sum() # False Negative

    accuracy = (TP + TN) / (TP + TN + FP + FN)
    precision = TP / (TP + FP) if (TP + FP) != 0 else 0
    recall = TP / (TP + FN) if (TP + FN) != 0 else 0
    f1 = 2 * precision * recall / (precision + recall) if (precision + recall) != 0 else 0

    # Menghitung Support
    support_class_1 = TP + FN # Jumlah data asli kelas 1 (Minat)
    support_class_0 = TN + FP # Jumlah data asli kelas 0 (Tidak Minat)

    print("\n--- Evaluasi Model Manual ---")
    print(f"Akurasi : {accuracy:.4f}")
    print(f"Precision : {precision:.4f}")
    print(f"Recall : {recall:.4f}")
    print(f"F1-Score : {f1:.4f}")
    print(f"Support untuk kelas 1 (Minat): {support_class_1}")
    print(f"Support untuk kelas 0 (Tidak Minat): {support_class_0}")
    print(f"TP={TP}, TN={TN}, FP={FP}, FN={FN}")

# Baca file
df = pd.read_excel("C:/Users/HP/Documents/Kuliah/Skripsi/coding/Data_Observasi_149.xlsx")
```

Lampiran 7 code program data testing Algoritma Logistik Regresi 4

```
# Acak dan split data
df = df.sample(frac=1, random_state=42).reset_index(drop=True)
train_size = int(0.7 * len(df))
train_df = df[:train_size]
test_df = df[train_size:]

# Prediksi & evaluasi di test data
test_df = logistic_regression(test_df)
evaluate_metrics(test_df)

# Contoh prediksi user baru
print("\n--- Prediksi Data Baru ---")
umur_baru = 30
jenis_kelamin_baru = 0
frekuensi_cukur_baru = 2

pred, prob = prediksi_minat(umur_baru, jenis_kelamin_baru, frekuensi_cukur_baru)
print(f"Prediksi Minat : {pred} (Probabilitas: {prob:.4f})")
```

Lampiran 8 Output code data testing Algoritma Logistik Regresi

```
PS C:\Users\HP> & C:\Users\HP\AppData\Local\Programs\Python\Python311\python.exe "C:\Users\HP\Documents\Kuliah\Skripsi\coding\Program\Tertentu\Algoritma Logistik Regresi.py"
C:\Users\HP\Documents\Kuliah\Skripsi\coding\Program\Tertentu\Algoritma Logistik Regresi.py(38): Settingptiticcappwarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df_input["pred_minat"] = hasil

--- Evaluasi Model Manual ---
Akurasi : 0.8889
Precision : 0.8500
Recall : 0.8889
F1-Score : 0.8889
Support untuk kelas 1 (Minat): 27
Support untuk kelas 0 (Tidak Minat): 8
Thresh, Thresh, F1-Score, etc.

--- Prediksi Data Baru ---
Prediksi Minat : 1 (Probabilitas: 0.8889)
```

Lampiran 9 Code Program EDA (Exploratory Data Analysis)

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Baca file Excel
df = pd.read_excel("C:/Users/HP/Documents/Kuliah/Skripsi/coding/Data_Observasi_149.xlsx")

# Statistik Deskriptif
print("Rata-rata Umur:", df["umur"].mean())
print("Persentase Pria (%):", (df["jenis_kelamin"] == 0).mean() * 100)
print("Rata-rata Frekuensi Cukur:", df["frekuensi_cukur"].mean())
print("Persentase Minat Aplikasi (%):", (df["minat_aplikasi"] == 1).mean() * 100)

# Ringkasan Minat vs Tidak Minat
grouped = df.groupby("minat_aplikasi")[["umur", "jenis_kelamin", "frekuensi_cukur"]].mean()
grouped["jenis_kelamin"] *= 100
print("\nRingkasan Minat vs Tidak Minat:")
print(grouped)

# Histogram Umur
plt.figure()
plt.hist(df["umur"], bins=10, color='skyblue', edgecolor='black')
plt.title("Distribusi Umur")
plt.xlabel("Umur")
plt.ylabel("Jumlah")
plt.grid(True)
plt.show()
```

Lampiran 10 Code Program EDA (Exploratory Data Analysis) 2

```
# Pie Chart Jenis Kelamin
plt.figure()
labels = ["Pria", "Wanita"]
sizes = [(df["jenis_kelamin"] == 0).sum(), (df["jenis_kelamin"] == 1).sum()]
plt.pie(sizes, labels=labels, autopct="%1.1f%%", colors=["lightblue", "lightpink"])
plt.title("Komposisi Jenis Kelamin")
plt.axis("equal")
plt.show()

# Bar Chart Frekuensi Cukur
plt.figure()
sns.countplot(x="frekuensi_cukur", data=df, palette="Blues")
plt.title("Distribusi Frekuensi Cukur")
plt.xlabel("Frekuensi Cukur per Bulan")
plt.ylabel("Jumlah")
plt.grid(True, axis='y')
plt.show()
```

Lampiran 11 Output code EDA (Exploratory Data Analysis)

```
P5 C:\Users\HP\Documents\Latihan Coding> & C:/Users/HP/AppData/Local/Programs/Python/Python313/python.exe "c:/Users/HP/Documents/
Kuliah/Skripsi/coding/Program tarbaru/EDA.py"
Rata-rata Umur: 31.98724832214765
Persentase Pria (%): 81.28885369127517
Rata-rata Frekuensi Cukur: 2.1946308724832213
Persentase Minat Aplikasi (%): 83.22147651006712

Ringkasan Minat vs Tidak Minat:
      umur  jenis_kelamin  frekuensi_cukur
minat_aplikasi
0          36.640000      12.00000      1.640000
1          29.967742      20.16129      2.386452
c:\Users\HP\Documents\Kuliah\Skripsi\coding\Program tarbaru\EDA.py:40: FutureWarning:
Passing 'palette' without assigning 'hue' is deprecated and will be removed in v0.14.0. Assign the 'x' variable to 'hue' and set
'legend=False' for the same effect.
sns.countplot(x="frekuensi_cukur", data=df, palette="Blues")
```

Lampiran 12 narasumber bang Niko selaku Hairstylist di barbershop “The Man Barbershop”



Lampiran 13 Foto Observasi di Barbershop “Trees Barbershop” bersama narasumber Kak Zura, beliau tidak mau masuk kamera, peneliti foto dari atas



Lampiran 14 Foto Observasi bersama narasumber Bang Putra di Barbershop “D’COLONEL’S Barbershop”



Lampiran 15 Foto dengan Hairstylist Barbershop “Premium Barbershop Zie & Zka”



Lampiran 16 Foto dengan Hairstylist Pak Saprin di Barbershop “Barbershop 168”



**Lampiran 17 Foto dengan Hairstylist Barbershop “Ivon Pangkasi”
Nahazisokhi Buulolo dengan mengajukan pertanyaan observasi**



Lampiran 18 Foto dengan Hairstylist Barbershop “Mr.Klimis” di Marelan bang Ilham dengan mengajukan pertanyaan



Lampiran 19 Foto dengan Hairstylist bang Aa di Barbershop “Aa Barbershop” di belawan dengan mengajukan pertanyaan



Hasil Turnitin

