

SKRIPSI
ANALISIS KINERJA DENGAN *CLUSTERING* TERHADAP KEPUASAN
PELANGGAN PADA PERUSAHAAN AIR MINUM PERUMDA
TIRTANADI MENGGUNAKAN METODE *FUZZY K-MEANS*

DISUSUN OLEH
SULTAN ZAKY
2109010134



PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN
T.A 2024/2025

**ANALISIS KINERJA DENGAN *CLUSTERING* TERHADAP KEPUASAN
PELANGGAN PADA PERUSAHAAN AIR MINUM PERUMDA
TIRTANADI MENGGUNAKAN METODE *FUZZY K-MEANS***

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**

SULTAN ZAKY NPM. 2109010134

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

2025

LEMBAR PENGESAHAN

Judul Skripsi : ANALISIS KINERJA DENGAN *CLUSTERING*
TERHADAP KEPUASAN PELANGGAN PADA
PERUSAHAAN AIR MINUM PERUMDA
TIRTANADI MENGGUNAKAN METODE *FUZZY K-
MEANS*

Nama Mahasiswa : SULTAN ZAKY
NPM : 2109010134
Program Studi : SISTEM INFORMASI

Menyetujui
Komisi Pembimbing



(Indah Purnama Sari, S.T., M.Kom)
NIDN. 0116049001

Ketua Program Studi



(Dr. Firahmi Rizky, S.Kom., M.Kom)
NIDN. 3048770671230343



Dekan

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

PERNYATAAN ORISINALITAS

ANALISIS KINERJA DENGAN *CLUSTERING* TERHADAP KEPUASAN PELANGGAN PADA PERUSAHAAN AIR MINUM PERUMDA TIRTANADI MENGGUNAKAN METODE *FUZZY K-MEANS*

SKRIPSI

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

Medan, Agustus 2025

Yang membuat pernyataan



SULTAN ZAKY

NPM. 2109010134

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

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

Nama	: SULTAN ZAKY
NPM	2109010134
Program Studi	: SISTEM INFORMASI
Karya Ilmiah	: Skripsi

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

**ANALISIS KINERJA DENGAN *CLUSTERING* TERHADAP KEPUASAN
PELANGGAN PADA PERUSAHAAN AIR MINUM PERUMDA
TIRTANADI MENGGUNAKAN METODE *FUZZY K-MEANS***

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, Agustus 2025

Yang membuat pernyataan



SULTAN ZAKY

NPM. 2109010134

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap	: SULTAN ZAKY
Tempat dan Tanggal Lahir	: Kisaran, 12-01-2004
Alamat Rumah	: Prov.Sumatera Utara Kab.Batu Bara Dusun III Sidodadi Desa Binjai Baru
Telepon/Faks/HP	081362278570
E-mail	: sultanzaky45@gmail.com
Instansi Tempat Kerja	: -
Alamat Kantor	: -

DATA PENDIDIKAN

SD	:	TAMAT: 2015
SMP	:	TAMAT: 2018
SMA	:	TAMAT: 2021

KATA PENGANTAR



Penulis tentunya berterima kasih kepada berbagai pihak dalam dukungan serta doa dalam penyelesaian skripsi. Penulis juga mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Agussani, M.AP., sebagai Rektor Universitas Muhammadiyah Sumatera Utara (UMSU).
2. Bapak Assoc. Prof. Dr. Al-Khowarizmi, S.Kom., M.Kom sebagai Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
3. Bapak Halim Maulana, S.T, M.Kom sebagai Wakil Dekan I Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
4. Bapak Dr. Lutfi Basit, S.Sos, M.I.Kom sebagai Wakil Dekan III Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
5. Ibu Dr. Firahmi Rizky, S.Kom.,M.Kom sebagai Ketua Program Studi Sistem Informasi.
6. Bapak Mahardika Abdi Prawira Tanjung, S.Kom., M.Kom sebagai Sekretaris Program Studi Sistem Informasi
7. Ibu Indah Purnama Sari, S.T., M.Kom sebagai Dosen Pembimbing, penulis menyampaikan rasa terima kasih yang tulus atas segala bimbingan, arahan, serta masukan yang sangat berarti selama proses penyusunan skripsi ini. Dedikasi, kesabaran, dan waktu yang telah diberikan menjadi bekal berharga bagi penulis dalam menyelesaikan penelitian ini dengan sebaik-baiknya.
8. Ucapan terima kasih yang tak terhingga juga penulis sampaikan kepada kedua orang tua tercinta, Papa dan Mama, atas doa, kasih sayang, dukungan moral maupun materiil, serta semangat yang tiada henti. Segala pengorbanan dan cinta yang telah diberikan menjadi sumber kekuatan dan motivasi utama dalam menyelesaikan proses pendidikan dan penulisan skripsi ini.

9. Penulis juga menyampaikan rasa terima kasih yang sebesar-besarnya kepada seluruh teman-teman Kuliah Kerja Nyata (KKN) atas kebersamaan, kekompakan, dan semangat solidaritas yang senantiasa terjaga. Dukungan serta dorongan semangat yang saling diberikan menjadi pengalaman berharga yang turut membentuk semangat penulis dalam menyelesaikan skripsi ini.
10. Kepada seseorang yang pernah bersama penulis dan tidak bisa penulis sebutkan namanya. Terima kasih atas segala rasa yang pernah ada, termasuk luka yang diberikan saat proses penyusunan skripsi atau tugas akhir ini. Kepergianmu memang menyisakan kecewa, namun justru dari sana penulis mendapatkan motivasi untuk terus maju dan berproses menjadi pribadi yang lebih baik dan belajar tentang arti mencintai, dicintai, dan menerima arti kehilangan sebagai bentuk proses penempatan dalam menghadapi dinamika hidup. Tanpamu, mungkin penulis tak akan pernah benar-benar memahami bahwa cinta tak selalu harus dimiliki, dan perpisahan pun juga bisa menjadi awal dari proses memaknai hidup. Terima kasih telah menjadi bagian menyenangkan sekaligus menyakitkan dari pendewasaan ini. Pada akhirnya setiap orang ada masanya dan setiap masa ada orangnya.
11. Semua pihak yang terlibat langsung ataupun tidak langsung yang tidak dapat penulis ucapkan satu-persatu yang telah membantu penyelesaian skripsi ini.
12. Dan yang terakhir penulis berterima kasih kepada diri penulis sendiri, karena telah menunjukkan keteguhan luar biasa dalam melalui setiap proses yang tidak selalu mudah. Penulis bangga karena mampu melewati rasa lelah, ragu, dan berbagai tekanan tanpa benar-benar menyerah. Terima kasih karena terus memilih untuk maju, bahkan ketika jalan terasa berat dan sunyi. Penulis telah membuktikan bahwa diri penulis kuat, pantang menyerah, dan layak untuk dihargai. Perjalanan ini adalah bukti bahwa penulis mampu berdiri, tumbuh, dan menyelesaikan apa yang telah dimulai. Terima kasih telah menjadi pribadi yang tidak hanya bertahan, tetapi juga berkembang.

**ANALISIS KINERJA DENGAN *CLUSTERING* TERHADAP KEPUASAN
PELANGGAN PADA PERUSAHAAN AIR MINUM PERUMDA
TIRTANADI MENGGUNAKAN METODE *FUZZY K-MEANS***

ABSTRAK

Kepuasan pelanggan merupakan indikator penting dalam menilai keberhasilan perusahaan penyedia layanan, termasuk Perumda Tirtanadi sebagai penyedia air minum. Tingkat kepuasan pelanggan mencerminkan kemampuan perusahaan dalam memenuhi kebutuhan dasar masyarakat terhadap ketersediaan dan distribusi air bersih. Namun, Perumda Tirtanadi masih menghadapi tantangan berupa keluhan kualitas air, gangguan distribusi, dan respons terhadap pengaduan, yang menunjukkan adanya kesenjangan antara ekspektasi pelanggan dan layanan yang diberikan. Untuk mengatasi permasalahan tersebut, diperlukan pendekatan berbasis teknologi informasi dan analisis data guna memahami karakteristik serta preferensi pelanggan secara lebih mendalam. Penelitian ini menerapkan metode *Fuzzy K-Means* dalam proses klasterisasi data opini pelanggan terhadap layanan Perumda Tirtanadi. Metode ini dipilih karena mampu menangani data yang bersifat tidak pasti dengan memberikan derajat keanggotaan pada lebih dari satu klaster, sehingga hasil analisis menjadi lebih fleksibel dan akurat. Melalui penerapan metode ini, perusahaan dapat mengidentifikasi segmen pelanggan yang memerlukan perhatian khusus serta menyusun strategi peningkatan layanan secara lebih terarah. Hasil penelitian diharapkan mampu memberikan kontribusi terhadap peningkatan kualitas layanan, penguatan citra positif perusahaan, serta mendukung pengambilan keputusan dalam upaya meningkatkan kepuasan pelanggan secara berkelanjutan.

Kata Kunci : *Kepuasan Pelanggan, Kualitas Layanan, Perumda Tirtanadi, Fuzzy K-Means, Klasterisasi*

PERFORMANCE ANALYSIS OF CUSTOMER SATISFACTION AT PERUMDA TIRTANADI WATER UTILITY USING THE FUZZY K- MEANS CLUSTERING METHOD

ABSTRACT

Customer satisfaction is a key indicator in evaluating the success of service providers, including Perumda Tirtanadi as a regional water utility company. The level of customer satisfaction reflects the company's ability to fulfill the basic needs of the community regarding the availability and distribution of clean water. However, Perumda Tirtanadi still faces several challenges such as complaints about water quality, distribution disruptions, and responses to customer grievances, indicating a gap between customer expectations and the services provided. To address these issues, the application of information technology and data analysis is required to better understand customer characteristics and preferences. This study applies the *Fuzzy K-Means* method in clustering customer opinion data on Perumda Tirtanadi's services. The method is chosen due to its ability to handle uncertainty in data by assigning degrees of membership across multiple clusters, thus producing more flexible and accurate analysis results. Through this approach, the company can identify customer segments requiring special attention and design more targeted service improvement strategies. The findings of this study are expected to contribute to service quality enhancement, strengthen the company's positive image, and support decision-making processes in achieving sustainable customer satisfaction.

Keyword : *Customer Satisfaction, Service Quality, Perumda Tirtanadi, Fuzzy K-Means, Clustering*

DAFTAR ISI

LEMBAR PENGESAHAN	i
PERNYATAAN ORISINALITAS.....	ii
PERNYATAAN PERSETUJUAN PUBLIKASI.....	iii
RIWAYAT HIDUP	iv
KATA PENGANTAR	v
ABSTRAK.....	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
BAB I. PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah	3
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian.....	4
BAB II. LANDASAN TEORI	6
2.1. Kepuasan Pelanggan	6
2.1.1. Pengertian Kepuasan Pelanggan... ..	6
2.1.2. Faktor Faktor yang mempengaruhi kepuasan Pelanggan.....	7
2.2. Clustering.....	7
2.3. Metode <i>Fuzzy K-Means</i>	9
2.3.1. Rumus <i>Fuzzy K-Means</i>	10
2.3.2. Atribut <i>Fuzzy K-Means</i>	11
2.4. Phyton.....	12
2.5. Penelitian Terdahulu	14
BAB III. METODOLOGI PENELITIAN	16
3.1. Tempat dan Waktu Penelitian	16
3.2. Jenis Pendekatan penelitian	17
3.3. Subjek Penelitian.....	17
3.4. Prosedur Penelitian	18

3.5. Teknik Pengumpulan Data	18
3.6. Indikator Pencapaian	18
3.7. Flowchart Penelitian.....	19
3.8. Pembahasan Alur Penelitian	20
3.9. Alur Sistem Pengolahan Data	22
BAB IV. HASIL DAN PEMBAHASAN	27
4.1. Pengenalan Masalah.....	27
4.2. Pengumpulan Data	27
4.3. Pra-Pemrosesan Data	28
4.4. Penerapan.....	33
4.4.1. Input Data	35
4.4.2. Preprocessing Data.....	36
4.4.3. TF-IDF Vectorizer	37
4.4.4. Penerapan Fuzzy K-Means Clustering	37
4.5. Evaluasi	39
4.6. Distribusi Cluster (3 Sentimen : Negatif, Netral, Positif)	41
4.7. Visualisasi WordCloud Percluster	42
4.8. Visualisasi Keanggotaan Cluster.....	42
BAB V. PENUTUP.....	43
5.1. Saran.....	43
5.2. Kesimpulan	43
DAFTAR PUSTAKA	45

DAFTAR TABEL

Tabel 2.1	Tabel Penelitian Terdahulu	14
Tabel 3.1	Tabel Penelitian	16
Tabel 3.2	Data Hasil Kuesioner	24
Tabel 3.3	Data yang Sudah di klaster.....	25
Tabel 4.1	Hasil Kuisisioner Kepuasan Pelanggan.....	28
Tabel 4.2	Hasil <i>Case Folding</i>	28
Tabel 4.3	Hasil <i>Filtering</i>	29
Tabel 4.4	Hasil <i>Stemming</i>	31
Tabel 4.5	Hasil <i>Tokenizing</i>	32

DAFTAR GAMBAR

Gambar 2.1 Python	12
Gambar 3.1 Flowchart Penelitian	19
Gambar 3.2 Kuesioner Penelitian	21
Gambar 3.3 Google Colab.....	22
Gambar 4.1 TF-IDF Menggunakan Google Collab.....	33
Gambar 4.2 Library Fuzzy K-means yang digunakan	35
Gambar 4.3 Input Data Pada Google Collab	35
Gambar 4.4 Preprocessing Data Pada Google Collab	36
Gambar 4.5 Hasil Perhitungan TF-IDF Pada Google	37
Gambar 4.6 Penerapan Fuzzy K-Means Pada Google Collab.....	37
Gambar 4.7 Hasil Perhitungan Fuzzy K-Means Pada Google Collab	38
Gambar 4.8 Evaluasi dari Google Collab.....	39
Gambar 4.9 Hasil Evaluasi dari Google Collab	39
Gambar 4.10 Distribusi Cluster pada Google Collab	40
Gambar 4.11 Grafik Hasil Distribusi Cluster pada Google Collab.....	41
Gambar 4.12 WordCloud	41
Gambar 4.13 Visualisasi Keanggotaan Fuzzy Pada Google Collab	42
Gambar 4.14 Hasil Visualisasi Keanggotaan Fuzzy Pada Google Collab	42

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Kepuasan pelanggan merupakan indikator utama dalam menilai keberhasilan suatu perusahaan dalam memberikan layanan yang berkualitas bagi perusahaan penyedia air minum seperti Perumda Tirtanadi, tingkat kepuasan pelanggan menunjukkan kemampuan perusahaan dalam memenuhi kebutuhan dasar masyarakat terhadap ketersediaan air bersih yang layak serta pendistribusian yang lancar. Kepuasan yang tinggi dari pelanggan tidak hanya menjadi indikator keberhasilan operasional, tetapi juga berperan dalam membangun citra positif perusahaan di mata masyarakat.

Peningkatan kepuasan pelanggan menjadi salah satu fokus utama bagi perusahaan-perusahaan di berbagai sektor, termasuk dalam layanan air minum. Perusahaan Umum Daerah (Perumda) Tirtanadi, sebagai penyedia layanan air minum di daerah tertentu, menghadapi tantangan yang kompleks dalam menjaga dan meningkatkan kepuasan pelanggan. Berbagai faktor seperti kualitas pelayanan, ketersediaan sumber air, harga, serta responsivitas terhadap keluhan pelanggan dapat mempengaruhi tingkat kepuasan tersebut.

Namun, dalam praktiknya, Perumda Tirtanadi masih menghadapi berbagai tantangan dalam memenuhi harapan pelanggan. Keluhan terkait kualitas air, gangguan distribusi, dan respons terhadap pengaduan masih sering muncul. Hal ini menunjukkan adanya kesenjangan antara ekspektasi pelanggan dan layanan yang diberikan, yang dapat berdampak pada menurunnya tingkat kepuasan dan loyalitas pelanggan.

Dalam era digital saat ini, dengan adanya teknologi informasi dan analisis data, perusahaan memiliki akses untuk lebih memahami perilaku dan preferensi pelanggan. Salah satu metode yang dapat digunakan untuk menganalisis tingkat kepuasan pelanggan adalah melalui teknik klastering. Teknik ini memungkinkan perusahaan untuk mengelompokkan pelanggan berdasarkan karakteristik dan tingkat kepuasan mereka. Analisis klastering memiliki peran signifikan dalam mengukur tingkat kepuasan pelanggan, yang secara langsung berdampak pada

kualitas produk dan layanan perusahaan. Penerapan metode klastering dapat memberikan gambaran yang lebih terstruktur mengenai segmen pelanggan, sehingga memungkinkan perusahaan untuk lebih fokus dalam meningkatkan kualitas produk dan layanan guna meningkatkan kepuasan pelanggan. Evaluasi terhadap kualitas layanan dan produk menjadi hal yang krusial dalam upaya membangun citra positif perusahaan di mata pelanggan

Salah satu metode yang umum digunakan dalam pengelompokan data adalah *Fuzzy K-Means*. Metode ini memiliki keunggulan dalam menangani data yang bersifat tidak pasti dan ambigu. Pendekatan fuzzy logic yang diterapkan dalam metode ini memungkinkan adanya tingkat keanggotaan data dalam lebih dari satu kelompok, sehingga memberikan representasi yang lebih fleksibel dibandingkan dengan metode klastering konvensional. Hasil evaluasi menunjukkan bahwa metode ini mampu mengelompokkan data secara otomatis dengan tingkat akurasi dan interpretabilitas yang tinggi, (Budiono et al., 2024).

Selain itu, metode *Fuzzy K-Means* juga memiliki keunggulan dalam menentukan keanggotaan klaster dengan lebih akurat. Tidak seperti metode klastering "hard" seperti K-Means, yang mengharuskan setiap data masuk ke dalam satu klaster secara tegas, *Fuzzy K-Means* memungkinkan suatu data memiliki derajat keanggotaan dalam beberapa klaster secara bersamaan, dengan rentang nilai antara 0 hingga 1. Hal ini mencerminkan kondisi nyata bahwa pelanggan tidak selalu dapat dikategorikan ke dalam satu segmen secara. Selain itu, metode ini memiliki tingkat akurasi yang tinggi serta efisiensi waktu komputasi yang lebih baik

Melalui penelitian ini, diharapkan Perumda Tirtanadi dapat mengidentifikasi segmen pelanggan yang memerlukan perhatian khusus, meningkatkan efisiensi layanan, serta merumuskan strategi yang lebih efektif untuk meningkatkan kepuasan pelanggan secara keseluruhan. Penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam pengembangan strategi pelayanan yang lebih baik, sehingga Perumda Tirtanadi dapat memenuhi ekspektasi dan kebutuhan pelanggan dengan lebih optimal. Dengan demikian, hasil penelitian ini dapat dijadikan dasar pengambilan keputusan untuk perbaikan layanan di masa mendatang.

1.2. Rumusan Masalah

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

- 1) Bagaimana tingkat kepuasan pelanggan terhadap layanan yang diberikan oleh Perusahaan Umum Daerah (Perumda) Tirtanadi saat ini?
- 2) Apa saja permasalahan yang dihadapi pelanggan terkait dengan kualitas layanan air minum yang disediakan oleh Perumda Tirtanadi?
- 3) Bagaimana metode *Fuzzy K-Means* dapat digunakan untuk mengelompokkan pelanggan berdasarkan tingkat kepuasan mereka?
- 4) Bagaimana hasil pengelompokan pelanggan menggunakan metode *Fuzzy K-Means* dapat memberikan gambaran yang lebih terarah dalam mengevaluasi kinerja layanan Perumda Tirtanadi?
- 5) Bagaimana pemanfaatan hasil analisis klastering dapat mendukung strategi peningkatan mutu pelayanan dan kepuasan pelanggan di Perumda Tirtanadi?

1.3. Batasan Masalah

Berdasarkan Rumusan Masalah yang telah diuraikan di atas, maka batasan masalah dalam penelitian ini adalah sebagai berikut

- 1) Penelitian hanya difokuskan pada pelanggan Perusahaan Umum Daerah (Perumda) Tirtanadi.
- 2) Penelitian ini menggunakan algoritma *Fuzzy K-Means* sebagai pendekatan dalam proses pengelompokan data berdasarkan tingkat keanggotaan yang bersifat tidak tegas
- 3) Penelitian ini berfokus pada analisis tingkat kepuasan yang dirasakan oleh pelanggan terhadap layanan yang diterima
- 4) Bahasa pemrograman yang digunakan dalam proses analisis untuk menentukan tingkat kepuasan pelanggan dalam penelitian ini adalah Python.
- 5) Data yang digunakan dalam penelitian ini dibatasi pada data yang diperoleh melalui kuesioner, wawancara mendalam, dan observasi partisipatif selama masa penelitian berlangsung.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan sebelumnya, peneliti dapat merangkai tujuan penelitian diantaranya sebagai berikut:

- 1) Melakukan identifikasi terhadap berbagai faktor yang memengaruhi tingkat kepuasan pelanggan terhadap layanan yang diberikan oleh Perumda Tirtanadi.
- 2) Melakukan pengelompokan pelanggan berdasarkan karakteristik serta tingkat kepuasan yang mereka rasakan terhadap layanan yang disediakan oleh Perumda Tirtanadi.
- 3) Menganalisis hubungan atau pola interaksi antara karakteristik pelanggan dengan tingkat kepuasan yang mereka alami terhadap layanan yang diberikan oleh Perumda Tirtanadi.
- 4) Mengeksplorasi pengaruh peningkatan tingkat kepuasan terhadap loyalitas pelanggan yang dimiliki terhadap Perumda Tirtanadi.

1.5. Manfaat Penelitian

Adapun manfaat yang diharapkan atas dilaksanakannya penelitian ini ialah sebagai berikut;;

- 1) Bagi Mahasiswa
Penelitian ini memberikan tambahan wawasan dan pengetahuan mengenai analisis kepuasan pelanggan dengan metode klastering, serta melatih kemampuan dalam melakukan pengolahan data dan penyusunan penelitian ilmiah. Selain itu, penelitian ini menjadi salah satu syarat untuk memperoleh gelar sarjana di Fakultas Ilmu Komputer Dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara.
- 2) Bagi Kampus
Penelitian ini menambah koleksi referensi ilmiah di kampus, yang dapat digunakan sebagai bahan rujukan untuk penelitian selanjutnya di bidang manajemen pelayanan dan teknologi informasi.
- 3) Bagi Perusahaan
Penelitian ini memberikan informasi yang bermanfaat dalam meningkatkan kualitas layanan dengan cara memahami kelompok

pelanggan yang perlu mendapat perhatian khusus, sehingga perusahaan dapat menyusun strategi pelayanan yang lebih efektif.

4) Bagi Masyarakat

Penelitian ini diharapkan berdampak pada meningkatnya kualitas layanan air bersih yang diterima oleh masyarakat, sehingga kebutuhan masyarakat terhadap air bersih dapat terpenuhi secara optimal.

BAB II

LANDASAN TEORI

2.1. Kepuasan Pelanggan

2.1.1. Pengertian Kepuasan Pelanggan

Dalam menghasilkan suatu produk, perusahaan perlu memastikan bahwa produk tersebut memenuhi harapan dan kebutuhan pelanggan. Tingkat kepuasan pelanggan menjadi salah satu faktor penentu keberhasilan perusahaan dalam memasarkan produknya serta membangun citra dan kepercayaan di masyarakat. Kepuasan adalah perasaan yang timbul ketika seseorang membandingkan kualitas layanan yang diterima dengan harapannya. Jika layanan yang diberikan sesuai atau melebihi harapan, maka kepuasan akan tercapai

“kepuasan pelanggan adalah perasaan senang atau kecewa yang muncul setelah membandingkan kinerja produk atau jasa yang dirasakan dengan harapan pelanggan,(Dr. Meithiana Indrasari, 2019). Jika kinerja produk memenuhi atau melebihi harapan, pelanggan akan merasa puas; sebaliknya, jika kinerja produk tidak memenuhi harapan, pelanggan akan merasa kecewa, (H Jasin, 2014) . “kepuasan pelanggan adalah pengukuran atau indikator sejauh mana pelanggan atau pengguna produk perusahaan atau jasa sangat senang dengan produk-produk atau jasa yang diterima, kepuasan pelanggan adalah perbandingan antara harapan terhadap persepsi pengalaman (dirasakan/diterima)”, (Dr.didin Fatihudin, se., 2019).

Berdasarkan definisi tersebut, dapat disimpulkan bahwa Kepuasan pelanggan adalah tingkat kepuasan yang dirasakan pelanggan setelah membandingkan harapan mereka dengan pengalaman nyata saat menggunakan suatu produk atau layanan. Jika kualitas yang diterima lebih rendah dari harapan, maka kepuasan pelanggan akan menurun. Sebaliknya, apabila produk atau layanan yang diberikan mampu memenuhi atau bahkan melampaui harapan, maka kepuasan pelanggan akan meningkat. Oleh karena itu, perusahaan perlu memahami kebutuhan dan ekspektasi pelanggan secara mendalam serta secara berkelanjutan mengevaluasi kualitas produk dan layanan yang ditawarkan. Dengan demikian, perusahaan dapat meningkatkan pengalaman pelanggan serta membangun loyalitas dalam jangka

panjang

2.1.2. Faktor Faktor yang mempengaruhi kepuasan Pelanggan

a) Kualitas produk

Pelanggan akan merasa puas apabila produk yang mereka gunakan memiliki kualitas yang baik dan sesuai dengan harapan.

b) Kualitas Pelayanan

Dalam sektor jasa, kualitas layanan menjadi aspek krusial dalam menentukan kepuasan pelanggan. Konsumen cenderung merasa puas apabila mereka mendapatkan pelayanan yang sesuai dengan harapan, baik dari segi kecepatan, keramahan, maupun efektivitas dalam menangani permasalahan

c) Aspek emosional

Kepuasan pelanggan juga dapat dipengaruhi oleh faktor emosional, terutama terkait dengan citra merek. Pelanggan sering kali merasa lebih puas ketika menggunakan produk dari merek tertentu yang memberikan kebanggaan serta pengakuan sosial. Dalam hal ini, kepuasan yang diperoleh bukan hanya berasal dari kualitas produk, tetapi juga dari nilai sosial yang melekat pada merek tersebut

d) Harga

Harga menjadi salah satu pertimbangan utama dalam kepuasan pelanggan. Produk yang memiliki kualitas serupa tetapi ditawarkan dengan harga lebih terjangkau akan memberikan nilai lebih bagi pelanggan, sehingga meningkatkan kepuasan mereka

e) Pengeluaran Tambahan

Pelanggan cenderung lebih puas apabila mereka tidak perlu mengeluarkan biaya tambahan atau menghabiskan banyak waktu untuk memperoleh produk atau layanan. Kemudahan akses serta efisiensi dalam biaya dan waktu menjadi faktor yang dapat meningkatkan kepuasan pelanggan

2.2. Clustering

Clustering atau klusterisasi merupakan suatu teknik atau metode untuk

mengelompokkan data. Clustering adalah salah satu alat yang penting dalam pengolahan data statistik untuk melakukan analisis data,(Nur Afidah, 2023). Clustering merupakan proses pengelompokan objek dalam suatu kumpulan data ke dalam beberapa klaster yang memiliki karakteristik serupa, sehingga setiap klaster yang terbentuk bersifat homogen. Tujuan utama dari metode klaster adalah pengelompokkan sejumlah data/obyek kedalam klaster (grup) sehingga dalam setiap klaster akan berisi data yang semirip mungkin,(Bahauddin et al., 2021).

Para ahli telah mengembangkan berbagai metode clustering, di mana setiap metode memiliki karakteristik unik, serta kelebihan dan keterbatasan masing-masing. Proses clustering sendiri dapat dibedakan berdasarkan tingkat kesamaan data dalam satu klaster, keanggotaan setiap klaster, serta struktur klaster yang terbentuk. Perbedaan ini berperan penting dalam menentukan efektivitas suatu metode dalam mengelompokkan data, sehingga pemilihan metode yang tepat sangat bergantung pada karakteristik data yang dianalisis dan tujuan penelitian yang ingin dicapai.

Berdasarkan strukturnya, metode clustering dikategorikan menjadi dua jenis utama, yaitu hierarchical clustering dan partition clustering, (Alfina et al., 2012) . Hierarchical clustering menerapkan prinsip di mana setiap kelompok dapat terdiri dari satu data tunggal, kemudian kelompok-kelompok kecil dapat bergabung membentuk kelompok yang lebih besar secara bertahap hingga seluruh data tergabung dalam satu klaster utama,(Dan & Tik, 2023). Sementara itu, partition clustering membagi kumpulan data ke dalam beberapa kelompok yang bersifat eksklusif, sehingga tidak ada tumpang tindih antar kelompok, dan setiap data hanya dapat menjadi bagian dari satu klaster. Metode yang termasuk dalam kategori ini antara lain K-Means dan *Fuzzy K-Means*.

Menurut keanggotaan, metode clustering dikategorikan menjadi dua jenis utama, yaitu kategori eksklusif dan kategori tumpang tindih, (Yulita Pratiwi, 2023). Kategori eksklusif, atau yang sering disebut hard clustering, mengelompokkan setiap objek data ke dalam satu klaster tertentu tanpa adanya tumpang tindih. Artinya, setiap data hanya dapat menjadi anggota dari satu kelompok dan tidak dapat berada dalam lebih dari satu klaster pada waktu yang bersamaan. Metode seperti K-Means Clustering merupakan contoh dari

pendekatan ini

Sementara itu, kategori tumpang tindih, yang dikenal sebagai fuzzy clustering, memungkinkan setiap data memiliki derajat keanggotaan dalam beberapa klaster dengan nilai keanggotaan antara 0 hingga 1. Dengan metode ini, suatu data tidak sepenuhnya dikategorikan dalam satu klaster tertentu, melainkan memiliki kemungkinan untuk termasuk dalam beberapa klaster berdasarkan tingkat kedekatannya. Salah satu metode yang sering digunakan dalam kategori ini adalah *Fuzzy K-Means*, yang memberikan fleksibilitas dalam pengelompokan data yang memiliki karakteristik saling tumpang tindih.

2.3. Metode *Fuzzy K-Means*

Metode *Fuzzy K-Means* merupakan salah satu varian dari algoritma K-Means yang dikembangkan untuk mengatasi keterbatasan metode klasterisasi konvensional. *Fuzzy K-Means* adalah teknik untuk kelompok data yang non-hierarkis dimana mencoba membagi data yang ada menjadi satu atau dalam beberapa kelompok- kelompok, (Lesmana, 2023). Algoritma K-Means sendiri pertama kali diperkenalkan oleh James MacQueen pada tahun 1967 sebagai metode klasterisasi berbasis partisi yang membagi data ke dalam kelompok-kelompok berdasarkan kedekatan titik data dengan pusat klaster. Namun, pendekatan ini memiliki kelemahan karena setiap data hanya dapat masuk ke satu klaster secara eksklusif. Untuk mengatasi keterbatasan tersebut, James C. Bezdek pada tahun 1981 memperkenalkan konsep *Fuzzy K-Means* sebagai bagian dari pengembangannya terhadap metode Fuzzy K- Means ,yang dijelaskan dalam bukunya “*Pattern Recognition with Fuzzy Objective Function Algorithms.*”

Fuzzy K-Means memungkinkan setiap data memiliki derajat keanggotaan terhadap pengelompokan data yang lebih dari satu klaster, sehingga memberikan fleksibilitas dalam pengelompokan data yang lebih kompleks. Pendekatan ini sangat berguna dalam berbagai bidang seperti pengolahan citra, segmentasi data, dan sistem pengenalan pola, di mana batas antar kelompok sering kali tidak tegas atau saling tumpang tindih, (Desy Purnamasari, 2015).

Dalam metode *Fuzzy K-Means*, setiap titik data dalam dataset akan memiliki probabilitas keanggotaan dalam beberapa cluster, dan tidak hanya

menjadi anggota dari satu cluster secara mutlak. Dengan demikian, sebuah data akan memiliki nilai derajat keanggotaan terbesar pada cluster yang paling sesuai, dibandingkan dengan cluster lainnya.

2.3.1. Rumus *Fuzzy K-Means*

1) Fungsi Tujuan (objective Function)

Fungsi tujuan dari *Fuzzy K-Means* bertujuan untuk meminimalkan jarak antar data dan pusat kluster dengan mempertimbangkan tingkat keanggotaan setiap data dalam kluster, (Siregar, 2016). Fungsi tujuan *Fuzzy K-Means* dapat dituliskan sebagai berikut :

$$J(U, V) = \sum_{i=1}^N \sum_{k=1}^C u_{ik}^m \cdot \|x_i - v_k\|^2$$

di mana :

- a) u_{ik} adalah derajat keanggotaan titik data x_i pada kluster k ,
- b) m adalah parameter fuzziness (biasanya $m > 1$)
- c) x_i adalah pusat kluster ke- i
- d) v_k adalah pusat kluster ke- k
- e) C adalah jumlah kluster, dan
- f) N adalah jumlah titik data

2) Derajat keanggotaan (membership degree)

Derajat keanggotaan u_{ik} untuk data x_i pada kluster k dihitung dengan rumus berikut:

$$u_{ik} = \frac{1}{\sum_{j=1}^C \left(\frac{\|x_i - v_k\|}{\|x_i - v_{kj}\|} \right)^{2/(m-1)}}$$

di mana :

- a) u_{ik} adalah derajat keanggotaan titik data x_i dalam kluster k
- b) m adalah parameter fuzziness yang mengontrol seberapa "fuzzy"

pengelompokan tersebut

- c) v_k adalah pusat klaster k
- d) x_i adalah titik data ke- i ,
- e) C adalah jumlah klaster, dan
- f) j adalah indeks klaster yang berbeda

3) Pusat Klaster (Cluster Center)

Pusat klaster v_k dihitung sebagai rata-rata tertimbang dari titik data berdasarkan derajat keanggotaan:

$$v_k = \frac{\sum_{i=1}^N u_{ik}^m \cdot x_i}{\sum_{i=1}^N u_{ik}^m}$$

di mana :

- a) v_k adalah pusat klaster k
- b) u_{ik} adalah derajat keanggotaan titik data x_i dalam klaster k
- c) m adalah parameter fuzziness, dan
- d) N adalah jumlah titik data

2.3.2. Atribut Fuzzy K-Means

1) Derajat Keanggotaan (Membership Degree)

Setiap data memiliki derajat keanggotaan dalam satu atau lebih klaster, yang menunjukkan sejauh mana data tersebut berafiliasi dengan masing-masing klaster.

2) Parameter Fuzziness (m)

Parameter m mengontrol tingkat keaburan pengelompokan. Semakin besar nilai m , semakin fuzzy atau kabur batas antar klaster.

3) Pusat Klaster (Cluster Centers)

Pusat klaster dihitung sebagai nilai rata-rata tertimbang dari data berdasarkan derajat keanggotaan.

4) Jarak (Distance)

Pengukuran jarak antara data dan pusat klaster yang digunakan untuk menentukan seberapa besar keanggotaan data pada suatu klaster.

2.4. Python

Python adalah bahasa pemrograman yang menggunakan interpreter untuk menjalankan kode programnya, (Rahman et al., 2023). Bahasa ini telah menjadi salah satu pilihan utama dalam bidang ilmu data dan pembelajaran mesin (machine learning) karena fleksibilitasnya dalam menangani berbagai kebutuhan analisis dan pengolahan data Python juga didukung oleh beragam pustaka (library) seperti NumPy, Pandas, Matplotlib, hingga Scikit-learn dan Scikit-fuzzy yang mendukung penerapan algoritma klusterisasi termasuk *Fuzzy K-Means*.



Gambar 2.1 . Python

Fuzzy K-Means merupakan algoritma klusterisasi yang mengadopsi pendekatan fuzzy logic, di mana setiap data tidak diklasifikasikan secara tegas ke dalam satu kluster, melainkan memiliki derajat keanggotaan dalam beberapa kluster sekaligus. Hal ini menjadikan metode ini sangat cocok digunakan dalam situasi di mana batas antar kelompok data tidaklah jelas atau tumpang tindih. Oleh karena itu, algoritma ini sering digunakan dalam sistem rekomendasi, pengolahan citra digital, serta segmentasi pelanggan.

Dalam implementasinya, Python menyediakan pustaka seperti *scikit-fuzzy* yang mempermudah pengguna dalam menerapkan algoritma *Fuzzy K-Means* melalui fungsi *cmeans*. Fungsi ini memungkinkan pengguna menentukan parameter seperti jumlah kluster, batas toleransi error, serta nilai derajat fuzzifikasi (m), yang kemudian menghasilkan output berupa pusat kluster, matriks keanggotaan, dan nilai performa klusterisasi. Melalui Python, proses klusterisasi tidak hanya menjadi lebih

mudah, tetapi juga dapat dilakukan dengan efisiensi tinggi.

Secara matematis, algoritma *Fuzzy K-Means* bertujuan untuk meminimalkan fungsi objektif berikut:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2$$

Di mana J_m adalah fungsi objektif, u_{ij} adalah derajat keanggotaan data ke-1 terhadap kluster ke-j, x_i adalah data ke-I c_j adalah pusat kluster ke-j, dan m adalah parameter fuzzifikasi (biasanya $m > 1$). Tujuan dari algoritma ini adalah meminimalkan fungsi objektif tersebut untuk mendapatkan pembagian kluster yang paling optimal.

Beberapa studi menunjukkan bahwa integrasi antara Python dan metode *Fuzzy K-Means* menghasilkan efisiensi analisis yang tinggi serta akurasi pengelompokan yang baik, terutama untuk data yang memiliki kecenderungan tumpang tindih. Oleh karena itu, pendekatan ini sangat relevan untuk diterapkan dalam konteks penelitian kepuasan pelanggan, karena pelanggan tidak selalu dapat diklasifikasikan secara kaku dalam satu kelompok tertentu. Dengan memanfaatkan Python, peneliti memiliki fleksibilitas dalam mengolah data serta mengembangkan model analisis yang adaptif terhadap kompleksitas data di dunia nyata.

2.5. Penelitian Terdahulu

Tabel 2.1 . Tabel Penelitian terdahulu

No	Judul Penelitian	Penulis	Tahun	Research Gap
1	Analisis Cluster Kepuasan Pengguna Terhadap Layanan Shopee Menggunakan Algoritma K-Means (Patimah et al., 2021)	Patimah, Endah Ermatita, Ermatita Chamidah, Nurul	2021	Pada penelitian ini, algoritma K-Means digunakan untuk mengelompokkan kepuasan pengguna Shopee, namun belum menerapkan <i>Fuzzy K-Means</i> yang dapat memberikan informasi lebih rinci terkait keanggotaan pengguna dalam setiap cluster kepuasan.
2	Cluster Analysis Menggunakan Algoritma <i>Fuzzy K-Means</i> Untuk Tingkat Pengangguran Di Provinsi Riau (Sarbaini et al., 2022)	Sarbaini Saputri, Windylia Nazaruddin Muttakin, Fitriani	2022	Pada penelitian ini, <i>Fuzzy K-Means</i> digunakan untuk mengelompokkan tingkat pengangguran kabupaten/kota, namun belum digunakan pada analisis kepuasan pelanggan perusahaan air minum untuk strategi pelayanan.
3	Pengukuran Tingkat Kepuasan Mahasiswa Terhadap Pelayanan di Kantin Kampus Menggunakan Algoritma K-means Clustering (Carelsa et al., 2023)	Carelsa, H V Malik, R A Putra, D J	2023	Pada penelitian ini, algoritma K-Means diterapkan untuk mengukur kepuasan mahasiswa terhadap pelayanan kantin, namun belum menggunakan <i>Fuzzy K-Means</i> yang dapat menunjukkan keanggotaan majemuk pada tiap cluster kepuasan.
4	Segmentasi Pelanggan Menggunakan <i>Fuzzy k-means</i> Clustering berdasarkan RFM Model pada E-Commerce (Studi Kasus: E-Commerce XYZ) (Wicaksono et al., 2021)	Wicaksono, Ariefia Bachtiar, Fitra A Setiawan, Nanang Yudi	2021	Pada penelitian ini, metode <i>Fuzzy k-means</i> dan RFM digunakan untuk segmentasi pelanggan e-commerce, namun belum diterapkan pada data kepuasan dan konsumsi pelanggan perusahaan air minum.
5	Penerapan Metode <i>Fuzzy K-Means</i> Clustering untuk Pengelompokan Konten Halaman Web secara Otomatis. (Budiono et al., 2024)	Budiono, Wahyu Nugroho, Indarmawan Santoso, Gunawan,	2024	Pada penelitian ini, metode <i>Fuzzy K-Means</i> digunakan untuk mengelompokkan konten halaman web, namun belum diterapkan untuk segmentasi kepuasan dan konsumsi air pelanggan perusahaan air minum.
6	Penerapan Algoritma <i>Fuzzy k-means</i> Menggunakan Model Rfm Dalam Klasterisasi Pelanggan Pada Toko	Moh. Azlan Shah Putra, Julhandri, Ihsan Khoiru, Siti Monalisa	2020	Pada penelitian ini, <i>Fuzzy k-means</i> diterapkan dengan model RFM untuk klasterisasi pelanggan toko kue, namun

	Kue Feandra Cake(Shah Putra , 2020)			belum digunakan pada segmentasi kepuasan dan konsumsi pelanggan perusahaan air minum.
7	Implementasi Data Mining Pengelompokan Tingkat Kepuasan Pelanggan Pada Pt. Bisa Group Menggunakan <i>Fuzzy k-means</i> . (Maria Yostin Br Tarigan, 2023)	Maria Yostin Br Tarigan,Mita Aprila Silpa Br.Simanjuntak, Musa Andrew Loyd Sitanggang, Etriska Prananta S., Delima Sitanggang	2023	Pada penelitian ini, <i>Fuzzy k-means</i> digunakan untuk mengelompokkan kepuasan pelanggan restoran PT. Bisa Group, namun belum menggunakan <i>Fuzzy K-Means</i> pada pelanggan perusahaan air minum untuk mendukung kebijakan pelayanan
8	Implementasi K-Means Untuk Clustering Kepuasan Mahasiswa Teknik Informatika Terhadap Layanan Akademik. (JASMINE, 2024)	Achmad Maulidin, Rudiman, Ahmad Ridhani Mubaraq Mohamad Febriansyah AlAkbar, Muhammad Farid Azis	2024	Pada penelitian ini, metode K-Means digunakan untuk mengelompokkan kepuasan mahasiswa terhadap layanan akademik, namun belum menggunakan <i>Fuzzy K-Means</i> untuk mengetahui derajat keanggotaan mahasiswa dalam setiap cluster sehingga informasi segmentasi lebih detail.
9	Penerapan Text Mining Dan <i>Fuzzy k-means</i> Clustering Untuk Identifikasi Keluhan Utama Pelanggan Pdam Tirta Moedal Kota Semarang. (Aulia ., 2023)	Genisia Pramestiloka Aulia , Tatik Widiharih , Iut Tri Utami	2023	Pada penelitian ini, <i>Fuzzy k-means</i> digunakan untuk mengelompokkan keluhan pelanggan PDAM, namun belum menggabungkan analisis kepuasan dan konsumsi air pelanggan dengan metode <i>Fuzzy K-Means</i> .
10	Pendekatan Metode Serqual Dan Klaster <i>Fuzzy K-Means</i> Untuk Menganalisis Indeks Kepuasan Nasabah Bank X(Erliana ., 2024)	Dewi Erliana , Mustafid , Abdul Hoyyi	2024	Pada penelitian ini, metode SERVQUAL dan <i>Fuzzy K-Means</i> digunakan untuk menganalisis indeks kepuasan nasabah bank, namun belum diterapkan pada pelanggan perusahaan air minum dengan variabel konsumsi air dan kepuasan layanan secara bersamaan.

BAB III

METODOLOGI PENELITIAN

3.1. Tempat dan Waktu Penelitian

Penelitian ini dilaksanakan di PDAM Tirtanadi yang berlokasi di Kota Medan, Sumatera Utara. Waktu pelaksanaan penelitian dimulai pada bulan Juni sampai dengan Agustus. Lokasi ini dipilih secara purposif karena PDAM Tirtanadi merupakan badan usaha milik daerah yang memiliki cakupan layanan luas dan pelanggan dari berbagai segmen masyarakat, sehingga sesuai untuk dilakukan eksplorasi terhadap persepsi dan pengalaman pelanggan.

Tabel 3.1 . Tabel Penelitian

No	Kegiatan	Bulan									
		Des 24	Jan 25	Feb 25	Mar 25	Ap 25	Mei 25	Jun 25	Jul 25	Agt 25	Sep 25
1	ACC Judul Skripsi										
2	Observasi										
3	Bimbingan Dengan dosen										
4	Penyusunan Proposal Skripsi										
5	Pengajuan sempro										
6	Pengambilan Data										
7	Pengolahan Data										
8	Penyusunan Skripsi										

9	Pengumpulan Skripsi										
10	Sidang Skripsi										

3.2. Jenis dan Pendekatan Penelitian

Jenis penelitian ini adalah penelitian kualitatif dengan pendekatan studi kasus deskriptif. Penelitian kualitatif dipilih karena bertujuan untuk memahami secara mendalam persepsi, pengalaman, dan kepuasan pelanggan terhadap layanan PDAM Tirtanadi. Pendekatan studi kasus digunakan untuk menggali informasi secara intensif pada satu unit analisis, yaitu pelanggan perusahaan air minum tersebut. Menurut pendekatan ini memungkinkan peneliti untuk memahami konteks fenomena yang kompleks secara holistik dan mendalam. Penelitian ini juga mengadopsi prinsip naturalistik, yakni data dikumpulkan dalam kondisi alamiah tanpa manipulasi.

3.3. Subjek Penelitian

Pemilihan subjek dalam penelitian ini menggunakan teknik purposive sampling, yaitu memilih individu yang dianggap memiliki pengetahuan dan pengalaman yang relevan dengan fokus penelitian. Dalam hal ini, subjek penelitian adalah pelanggan Perumda Tirtanadi yang telah menggunakan layanan perusahaan selama minimal enam bulan. Purposive sampling memungkinkan peneliti untuk memilih informan yang dapat memberikan informasi yang mendalam dan tepat sesuai dengan tujuan penelitian. Teknik ini dianggap penting dalam penelitian kualitatif karena dapat memastikan bahwa informan yang terpilih.

Dalam konteks penelitian ini, penggunaan kuesioner memungkinkan pengumpulan data yang terstruktur dari pelanggan yang memiliki pengalaman yang relevan dengan layanan Perumda Tirtanadi. Kuesioner yang disusun akan berfokus pada aspek-aspek yang berhubungan langsung dengan persepsi dan pengalaman pelanggan selama menggunakan layanan tersebut, dengan harapan dapat memberikan wawasan yang lebih mendalam mengenai kepuasan dan kualitas layanan yang diterima oleh pelanggan.

3.4. Prosedur Penelitian

Prosedur penelitian meliputi tahap persiapan (penentuan fokus dan pengembangan instrumen wawancara), tahap pengumpulan data (wawancara mendalam dan observasi), tahap analisis data, dan tahap penyusunan laporan. Data dikumpulkan secara kualitatif, dianalisis secara induktif melalui pengkodean tematik, dan disajikan dalam bentuk narasi deskriptif. Setiap proses dilakukan dengan mempertimbangkan etika penelitian seperti persetujuan informan dan jaminan kerahasiaan data.

3.5. Teknik Pengumpulan Data

Teknik pengumpulan data dilakukan melalui:

1. Kuesioner

Selain wawancara mendalam dan observasi partisipatif, penelitian ini juga menggunakan kuesioner sebagai teknik pengumpulan data. Kuesioner disusun dengan pertanyaan terstruktur yang dirancang untuk mengukur tingkat kepuasan pelanggan Perumda Tirtanadi terhadap layanan yang diberikan. Pertanyaan-pertanyaan dalam kuesioner mencakup aspek kualitas layanan, kecepatan pelayanan, kejelasan informasi, dan kendala yang pernah dialami pelanggan selama menggunakan layanan Perumda Tirtanadi. Kuesioner disebarkan kepada pelanggan yang menjadi responden penelitian dengan tujuan untuk memperoleh data kuantitatif yang dapat mendukung hasil wawancara dan observasi, sehingga penelitian ini memiliki data yang lebih luas, objektif, dan dapat digunakan untuk mengidentifikasi tingkat kepuasan pelanggan secara lebih terukur dan komprehensif.

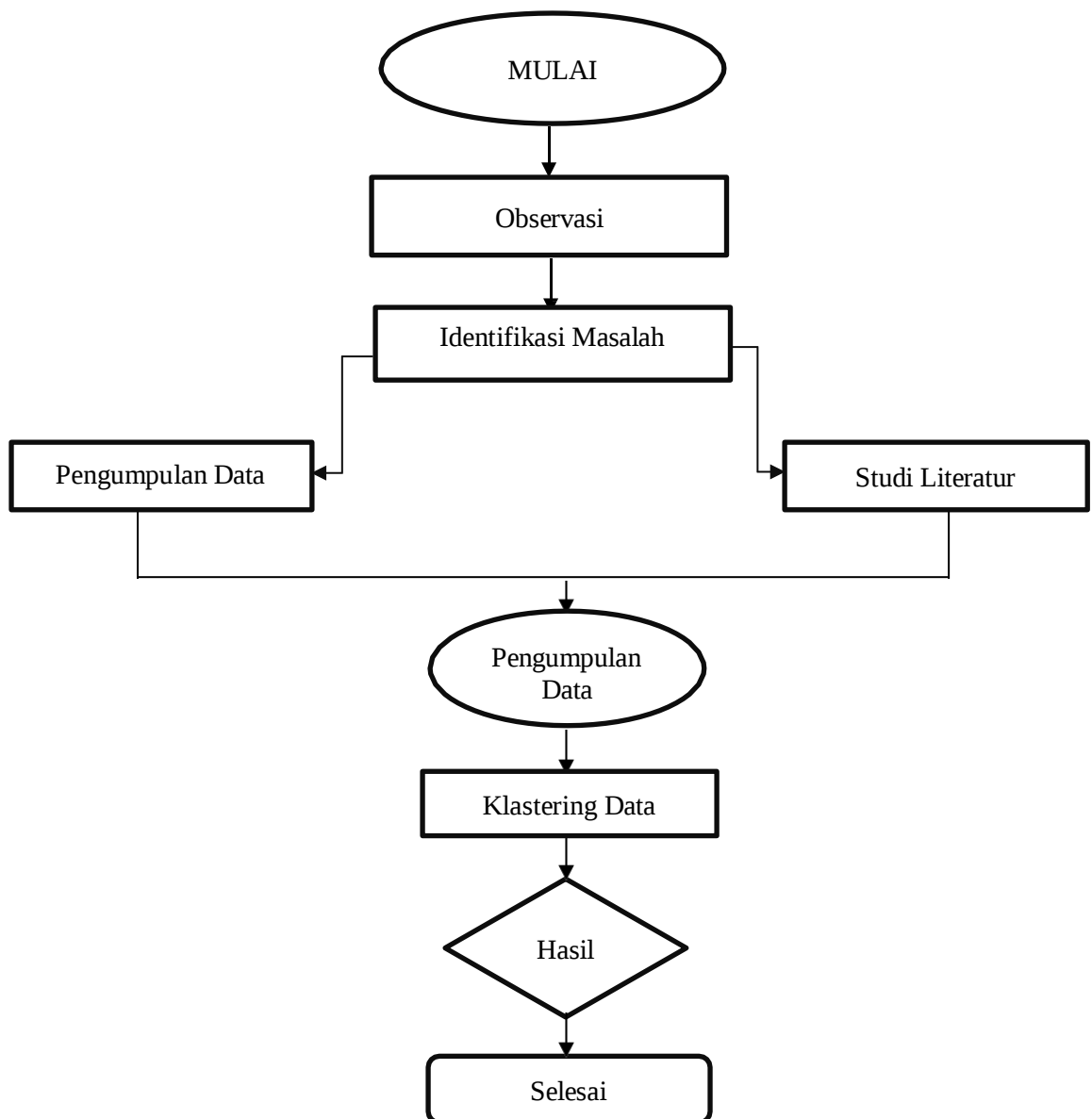
3.6. Indikator Pencapaian

Indikator Keberhasilan Penelitian ini adalah :

- 1) Tersusunnya kategori dan tema-tema utama yang mencerminkan faktor-faktor kepuasan dan ketidakpuasan pelanggan
- 2) Diperolehnya pemahaman yang mendalam tentang persepsi pelanggan terhadap kualitas layanan Perumda Tirtanadi

- 3) Tersusunnya rekomendasi strategis berbasis pengalaman pelanggan untuk peningkatan mutu layanan
- 4) Terpenuhinya validitas kualitatif yang baik melalui proses triangulasi dan verifikasi data.

3.7. Flowchart Penelitian



Gambar 3.1. Flowchart Penelitian

3.8. Pembahasan Alur Penelitian

Tahap awal pada sebuah penelitian merupakan langkah pertama yang memiliki peran penting dalam keseluruhan sebuah proses penelitian. Pada tahap ini, saya memulainya dengan mengidentifikasi dan merumuskan masalah yang ingin diteliti berdasarkan fenomena atau kondisi yang ada. Setelah itu saya mulai melakukan studi literatur untuk mengumpulkan informasi dan referensi terkait topik penelitian saya guna memperkuat dasar teori dan memahami penelitian terdahulu, pada tahap ini juga dilakukan penentuan tujuan penelitian dan batasan penelitian agar fokus penelitian menjadi lebih jelas dan terarah.

1. Tahap Observasi

Pada tahap observasi saya melakukan pengamatan secara langsung terhadap objek atau fenomena yang menjadi fokus penelitian saya yaitu soal kepuasan pelanggan di perusahaan umum daerah Tirtanadi Sumatera Utara. Tujuan utama dari tahap ini adalah mengumpulkan data awal yang akurat dan relevan, sehingga dapat memberikan gambaran nyata tentang kondisi atau perilaku yang terjadi. Data yang diperoleh dari observasi ini nantinya akan menjadi dasar dalam merumuskan masalah, menyusun hipotesis, serta menentukan langkah-langkah penelitian saya selanjutnya.

2. Tahap Identifikasi Masalah

Pada tahap identifikasi masalah, saya mengamati dan menganalisis fenomena atau kondisi yang terjadi pada (perumda) Tirtanadi untuk menemukan permasalahan yang relevan dan signifikan untuk sebuah bahan penelitian. Tahap ini melibatkan pengumpulan informasi awal yang bertujuan untuk mengungkap gap atau kekurangan dalam pengetahuan yang ada. Dengan mengidentifikasi masalah secara jelas dan spesifik sehingga saya dapat merumuskan fokus penelitian yang tepat dan saya berharap hasil dari penelitian saya nantinya dapat memberikan kontribusi yang berarti bagi (Perumda) Tirtanadi.

3. Studi Literatur

Pada tahap ini, saya melakukan studi literatur untuk memahami konsep dasar, teori-teori yang relevan, dari hasil penelitian sebelumnya yang berkaitan dengan topik yang saya angkat. Studi literatur dilakukan dengan menelusuri

berbagai sumber terpercaya seperti jurnal ilmiah, buku, artikel, dan laporan penelitian.

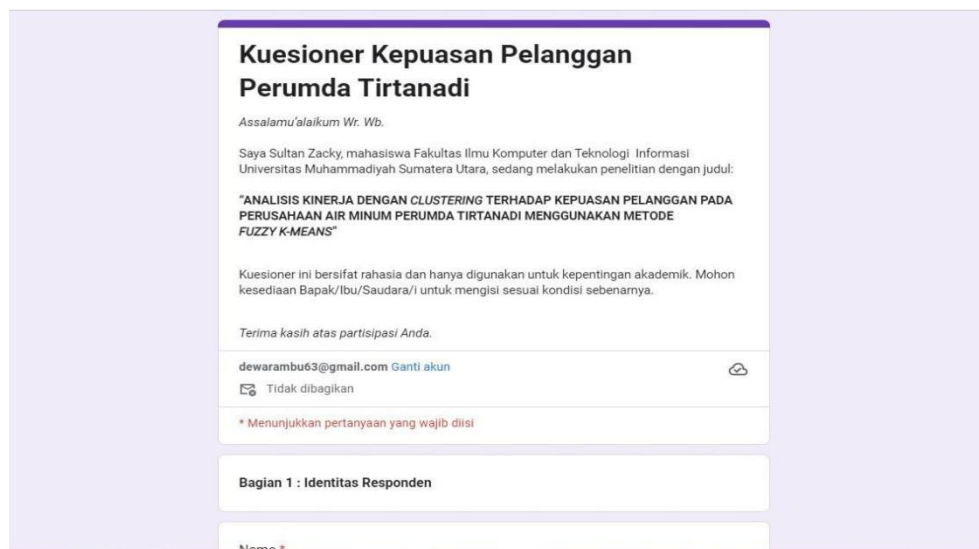
Tujuan dari studi literatur adalah untuk:

- a) Mengidentifikasi kesenjangan penelitian (research gap),
- b) Memperkuat landasan teori,
- c) Menyusun kerangka berpikir,
- d) Serta merumuskan hipotesis atau arah penelitian secara lebih sistematis.

Dengan mempelajari penelitian terdahulu, saya dapat memastikan bahwa penelitian yang saya lakukan memiliki kontribusi ilmiah yang baik dan tidak mengulang kajian yang telah ada sebelumnya.

4. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan melalui kuesioner. Kuesioner digunakan sebagai metode pengumpulan data kuantitatif dengan pertanyaan terstruktur yang dirancang untuk mengukur tingkat kepuasan pelanggan terhadap layanan yang diberikan, sehingga data yang diperoleh bersifat luas, objektif, dan mendukung hasil wawancara serta observasi untuk memberikan gambaran yang komprehensif mengenai tingkat kepuasan pelanggan Perumda Tirtanadi.

The image shows a digital questionnaire form titled "Kuesioner Kepuasan Pelanggan Perumda Tirtanadi". The form is white with a purple header and footer. It contains the following text: "Assalamu'alaikum Wr. Wb.", "Saya Sultan Zacky, mahasiswa Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara, sedang melakukan penelitian dengan judul: 'ANALISIS KINERJA DENGAN CLUSTERING TERHADAP KEPUASAN PELANGGAN PADA PERUSAHAAN AIR MINUM PERUMDA TIRTANADI MENGGUNAKAN METODE FUZZY K-MEANS'", "Kuesioner ini bersifat rahasia dan hanya digunakan untuk kepentingan akademik. Mohon kesediaan Bapak/Ibu/Saudara/i untuk mengisi sesuai kondisi sebenarnya.", "Terima kasih atas partisipasi Anda.", a contact email "dewarambu63@gmail.com" with a "Ganti akun" link, a "Tidak dibagikan" (Not shared) icon, and a red asterisk note "* Menunjukkan pertanyaan yang wajib diisi". Below this is a section titled "Bagian 1 : Identitas Responden" with a text input field for "Nama *".

Gambar 3.2. Kuesioner Penelitian

5. Klastering Data

Setelah data dikumpulkan dan melalui tahap praproses, penelitian dilanjutkan ke tahap klastering data. Klastering merupakan proses

pengelompokan data ke dalam beberapa kelompok (cluster) berdasarkan kemiripan karakteristik atau pola tertentu. Pada tahap ini, algoritma klustering seperti *Fuzzy K-Means*, digunakan untuk menemukan struktur tersembunyi dalam data. Data yang memiliki karakteristik serupa akan dikelompokkan dalam satu klaster, sehingga memudahkan dalam analisis lebih lanjut.

Tujuan dari tahap ini adalah untuk:

- a) Mengidentifikasi pola atau segmentasi dalam data,
 - b) Menyederhanakan kompleksitas data,
 - c) Dan mendukung pengambilan keputusan berdasarkan hasil pengelompokan
- Hasil dari klustering ini selanjutnya dianalisis dan diinterpretasikan untuk memberikan pemahaman yang lebih mendalam terhadap objek penelitian

3.9. Alur Sistem Pengolahan Data

Alur sistem pengolahan data dalam penelitian ini merupakan tahapan terstruktur yang mencakup seluruh proses dari pengumpulan data mentah hingga analisis hasil pengelompokan menggunakan metode *Fuzzy K-Means*. Seluruh proses dirancang dan diimplementasikan menggunakan bahasa pemrograman Python pada platform *Google Colaboratory (Colab)*.



Gambar 3.3. Google Colab

Google Colaboratory (Google Colab) merupakan platform komputasi berbasis *cloud* yang dikembangkan oleh Google untuk mendukung implementasi dan eksekusi kode pemrograman secara daring. Dalam konteks penelitian ini, Google Colab berfungsi sebagai sarana utama dalam mengimplementasikan algoritma *Clustering Fuzzy K-Means*, karena menyediakan lingkungan pemrograman yang interaktif serta terintegrasi dengan berbagai pustaka Python yang relevan, seperti *NumPy*, *Pandas*, dan *Scikit-learn*.

Pemilihan platform ini didasarkan pada kemampuannya dalam menyediakan lingkungan pemrograman berbasis cloud yang efisien, mendukung eksekusi kode Python tanpa perlu konfigurasi lokal, serta memiliki integrasi langsung dengan pustaka-pustaka analisis data dan machine learning yang relevan, ketersediaan sumber daya komputasi berbasis CPU maupun GPU pada Google Colab memungkinkan proses pengolahan data dan perhitungan iteratif algoritma dilakukan secara lebih efisien tanpa memerlukan perangkat keras dengan spesifikasi tinggi. Dengan demikian, Google Colab memberikan kontribusi signifikan terhadap efektivitas dan efisiensi analisis dalam penelitian ini.

Tahap awal dalam alur sistem ini adalah pengumpulan data (data acquisition) yang dilakukan melalui beberapa metode, yaitu wawancara, observasi partisipatif, dan kuesioner. Pengumpulan data dilakukan melalui wawancara langsung kepada pelanggan Perumda Tirtanadi untuk memperoleh pendapat, pengalaman, serta penilaian mereka secara mendalam mengenai kualitas layanan. Observasi partisipatif dilakukan dengan peneliti terlibat langsung dalam pelayanan di Perumda Tirtanadi untuk mengamati prosedur, interaksi, dan dinamika secara akurat. Di samping itu, kuesioner juga disebarakan kepada pelanggan untuk memperoleh data kuantitatif maupun kualitatif berupa jawaban tertutup dan opini terbuka. Contoh data mentah yang diperoleh dapat berupa hasil wawancara langsung terkait pengalaman pelanggan menggunakan layanan air bersih, jawaban kuesioner mengenai kepuasan harga dan kualitas air, serta komentar positif maupun negatif terhadap layanan Perumda Tirtanadi.

Tabel. 3.2. Data Hasil Kuesioner

No	Pertanyaan	Jawaban
1	Bagaimana kepuasan Anda terhadap harga air PDAM Tirta Nadi Medan?	Sangat puas
2	Mohon berikan alasan Anda atas jawaban tersebut	Karena harga yg diberikan masih terjangkau
3	Bagaimana kepuasan Anda terhadap harga air PDAM Tirta Nadi Medan?	Puas
4	Mohon berikan alasan Anda atas jawaban tersebut	Karna sesuai pemakaian
5	Bagaimana kepuasan Anda terhadap harga air PDAM Tirta Nadi Medan?	Tidak puas
6	Mohon berikan alasan Anda atas jawaban tersebut	Karena harga mahal
7	Bagaimana kepuasan Anda terhadap harga air PDAM Tirta Nadi Medan?	Sangat puas
8	Mohon berikan alasan Anda atas jawaban tersebut	Harga sesuai kualitas
9	Bagaimana kepuasan Anda terhadap harga air PDAM Tirta Nadi Medan?	Tidak puas
10	Mohon berikan alasan Anda atas jawaban tersebut	Di tempat kami air PAM hidup hanya di malam hari

Tahap selanjutnya adalah pra-pemrosesan data (data preprocessing). Data ulasan yang bersifat teks bebas umumnya mengandung elemen-elemen tidak terstruktur yang perlu dibersihkan agar dapat digunakan dalam proses analisis. Pada tahap ini dilakukan proses *case folding*, *filtering*, *stemming*, dan *tokenizing* yang akan menghasilkan teks bersih (*cleaned text*) untuk digunakan dalam proses TF-IDF (*Term Frequency-Inverse Document Frequency*), yang berfungsi untuk mengukur tingkat kepentingan kata dalam sebuah dokumen relatif terhadap seluruh korpus.

Tahap utama dari sistem ini adalah proses clustering menggunakan metode *Fuzzy K-Means*, yang merupakan salah satu teknik dalam *unsupervised learning* untuk mengelompokkan data ke dalam sejumlah klaster berdasarkan tingkat kemiripan fitur. Berbeda dengan metode K-Means konvensional yang bersifat “crisp” (tegas), *Fuzzy K-Means* memperbolehkan satu data untuk memiliki derajat keanggotaan pada lebih dari satu klaster secara bersamaan. Dalam konteks

penelitian ini, metode *Fuzzy K-Means* digunakan untuk mengelompokkan data ulasan pelanggan ke dalam beberapa kategori kepuasan, seperti “sangat puas”, “netral”, dan “tidak puas”. Dengan contoh data yang di Klastering sebagai berikut

Setelah proses clustering selesai, sistem akan memasuki tahap visualisasi dan interpretasi hasil analisis.

Tabel. 3.3. Data yang Sudah di klaster

NO	Cluster Positif	Cluster Negatif	Cluster Overlapping
1	Harga air PDAM terjangkau sesuai dengan kualitasnya.	Harga air PDAM mahal dan tidak sesuai pemakaian.	Harga air PDAM sebenarnya terjangkau, tapi airnya sering mati.
2	Harga air PDAM cukup murah dan sesuai pemakaian.	Harga mahal tapi air sering mati.	Harga murah tapi air kadang keruh saat hujan.
3	Harga stabil dan pelayanan memuaskan.	Harga air PDAM mahal, kualitasnya jelek.	Harga sesuai pemakaian, cuma airnya sering mati di siang hari.
4	Air bersih, harga sesuai pemakaian.	Harga mahal tidak sesuai dengan pelayanan.	Harga air PDAM terjangkau, hanya saja airnya kecil kalau malam
5	Harga normal dan airnya lancar.	Harga air mahal dan airnya sering keruh.	Harga air tidak terlalu mahal, tapi kualitas airnya kurang bersih.
6	Harga air standar dan terjangkau.	Harga mahal tapi air kecil.	Harga sesuai , tetapi respon pelayanan lambat
7	Harga air sesuai dengan kebutuhan sehari-hari	Harga mahal tidak sesuai kualitas air.	. Air nya bersih, tapi pelayanan buruk
8	Harga ekonomis dan kualitas air bersih.	Harga PDAM terlalu mahal.	Pelayanan bagus, air kotor
9	Air lancar dan harga terjangkau.	Harga air mahal dan pelayanan buruk.	Air kotor, pelayanan kurang memuaskan
10	Harga tidak mahal dan air bersih.	Harga mahal dan air sering mati di siang hari.	Harga mahal, tetapi air bersih
11	Harga air PDAM sesuai kubik meter yang digunakan.	Harga air mahal dan kualitas tidak bagus.	Air bersih , pelayanan kurang baik
12	Harga PDAM tergolong murah dan pelayanan baik.	Harga mahal dan tidak sesuai dengan kubik pemakaian.	Harga terjangkau, Tapi Air keruh

Hasil pengelompokan divisualisasikan melalui berbagai bentuk grafik, seperti diagram batang dan diagram lingkaran untuk menunjukkan distribusi jumlah data pada setiap klaster. Analisis lanjutan juga dilakukan dengan menampilkan contoh ulasan dari tiap klaster untuk memperkuat interpretasi hasil. Data hasil clustering beserta label klasternya kemudian disimpan dalam format CSV untuk keperluan dokumentasi dan analisis lebih lanjut.

Dengan demikian, alur sistem yang dirancang ini diharapkan mampu menyediakan suatu mekanisme yang efisien dan berbasis data dalam mengevaluasi tingkat kepuasan pelanggan terhadap layanan Perumda Tirtanadi. Hasil analisis ini diharapkan dapat digunakan sebagai dasar pengambilan keputusan strategis dalam peningkatan kualitas layanan perusahaan.

BAB IV

HASIL DAN PEMBAHASAN

4.1. Pengenalan Masalah

PDAM Tirtanadi adalah perusahaan daerah yang bertanggung jawab dalam pengelolaan serta pengembangan sistem penyediaan air bersih bagi masyarakat Kota Tirtanadi. Perusahaan ini berlokasi di Jl. Sisingamangaraja No.1, Ps. Baru, Kec. Medan Kota, Kota Medan, Sumatera Utara 20212. Sebagai penyedia layanan publik, PDAM Tirtanadi masih menghadapi kendala dalam memenuhi kepuasan para pelanggannya secara optimal. Beberapa permasalahan yang masih dirasakan oleh pelanggan meliputi aspek kualitas air, jumlah pasokan air, dan komunikasi antara perusahaan dengan pelanggan. Oleh sebab itu, penting bagi PDAM Tirtanadi untuk mengetahui pandangan dan opini masyarakat terkait pelayanan mereka, sehingga perbaikan dapat dilakukan berdasarkan kebutuhan dan keluhan nyata yang dialami masyarakat.

4.2. Pengumpulan Data

Berdasarkan hasil kuesioner terkait penilaian kinerja PDAM Tirtanadi, data yang terkumpul dapat diolah dengan memetakan beberapa variabel kepuasan pelanggan. Variabel tersebut meliputi kepuasan terhadap harga layanan yang direpresentasikan oleh *kepuasan harga*, yang mengukur tingkat kepuasan pelanggan terhadap harga layanan yang diberikan oleh PDAM Tirtanadi. Selanjutnya, *kepuasan pelayanan air* digunakan untuk menilai kepuasan pelanggan terhadap kualitas pelayanan air secara keseluruhan, sedangkan *kepuasan aliran air* merepresentasikan penilaian pelanggan terhadap kelancaran aliran air yang diterima. Selain itu, variabel *kepuasan daya tanggap* digunakan untuk mengukur tingkat kepuasan pelanggan terkait respon dan kecepatan petugas dalam menangani keluhan, dan *kepuasan kualitas air* menilai sejauh mana kualitas air yang disediakan dapat memenuhi kebutuhan pelanggan.. Seluruh variabel tersebut diukur menggunakan skala Likert guna memperoleh gambaran tingkat kepuasan masyarakat terhadap PDAM Tirtanadi. Secara menyuluruh, hasil penilaian ini akan digunakan untuk evaluasi kinerja dan perbaikan pelayanan ke depannya. Rincian

data observasi dapat dilihat pada lampiran, berikut disajikan data kepuasan pelanggan terkait harga layanan.

Tabel. 4.1. Hasil Kuisisioner kepuasan pelanggan

No	Klaster	Banyaknya Responden
1	Positif	89
2	Negatif	5
3	Netral	389

4.3. Pra-Pemrosesan Data

Sebelum dilakukan proses klasterisasi menggunakan metode Fuzzy K-Means, data naratif yang diperoleh dari opini terbuka responden harus melalui beberapa tahap pra-pemrosesan teks agar dapat dianalisis secara komputasional. Tahapan ini bertujuan untuk membersihkan dan menyiapkan data teks mentah menjadi bentuk yang dapat digunakan dalam representasi numerik melalui metode seperti *TF-IDF* (*Term Frequency–Inverse Document Frequency*).

a) Case Folding

Tahap *Case Folding* dilakukan dengan mengubah semua huruf kapital dalam data opini menjadi huruf kecil. Tujuannya adalah untuk menghindari perbedaan pengenalan kata akibat penggunaan huruf besar dan kecil, sehingga proses identifikasi kata menjadi lebih konsisten dan akurat. Berikut ini merupakan contoh data yang telah melalui proses *Case Folding*.

Tabel 4.2. Hasil *Case Folding*

1	Ai	karena, harga yg di berikan masih terjangkau
2	Rahmat	karena pdam memberikan harga yg standar
3	Rini	karena masih terjangkau
4	Deni	sebab harga pdam sibolga masih terjangkau seluruh lapisan masyarakatkau oleh lapisan masyarakat khususnya masyarakat menengah ke bawah

5	Arif	sangat membantu
6	Agung	karna sesuai pemakaian
7	Yuli	harga terjangkau dikalangan masyarakat menengah kebawah
8	Eni	karna harga masih terjangkau dan bisa buat kebutuhan tsb
9	Hadi	sangat puas
10	Wawan	karena menurut saya sesuai dengan yang dipakai
11	Andri	di tempat kami air pam hidup hanya di malam hari
12	Iis	karena terkadang harga tidak sesuai dengan pemakaian
13	Indah	karna belum adanya program air gratis
14	yusufsyahputra	karna menurut saya kemahalan, tidak sesuai dgn yg saya pakai
15	Abdul	suka melonjak tidak tentu

b) Filtering

Tahapan berikutnya adalah *Filtering*, yaitu proses penghapusan seluruh simbol atau karakter non-alfabet dari dataset. Langkah ini dilakukan untuk mengurangi gangguan (*noise*) dalam data dan meningkatkan akurasi analisis yang akan dilakukan. Berikut adalah contoh data yang telah diproses melalui tahap *Filtering*.

Tabel 4.3. Hasil *Filtering*

1	Ai	karena harga yg beri masih terjangkau
2	Rahmat	karen, pdam beri harga yg standar
3	Rini	karena masih terjangkau

4	Deni	sebab harga pdam sibolga masih terjangkau seluruh lapisan masyarakatkau oleh lapisan masyarakat khususnya masyarakat menengah ke bawah
5	Arif	sangat membantu
6	Agung	karna sesuai pemakaian
7	Yuli	harga terjangkau dikalangan masyarakat menengah kebawah
8	Eni	karna harga masih terjangkau dan bisa buat kebutuhan tsb
9	Hadi	sangat puas
10	Wawan	karena menurut saya sesuai dengan yang dipakai
11	Andri	di tempat kami air pam hidup hanya di malam hari
12	Iis	karena terkadang harga tidak sesuai dengan pemakaian
13	Indah	karna belum adanya program air gratis
14	yusufsyahputra	karna menurut saya kemahalan, tidak sesuai dgn yg saya pakai
15	Abdul	suka melonjak tidak tentu

c) Stemming

Pada tahap *Stemming*, setiap teks dalam dataset akan dianalisis untuk menghapus semua imbuhan yang melekat pada kata. Proses ini bertujuan untuk mengembalikan kata ke bentuk dasarnya, sehingga memudahkan dalam memahami makna inti dari kata tersebut serta mengurangi variasi kata yang memiliki arti serupa. Dengan demikian, analisis data menjadi lebih efisien dan terfokus pada inti informasi yang terkandung.

Tabel 4.4. Hasil *Stemming*

1	Ai	karena harga yg beri masih terjangkau
2	Rahmat	karen, pdam beri harga yg standar
3	Rini	karena masih terjangkau
4	Deni	sebab harga pdam sibolga masih terjangkau seluruh lapisan masyarakatgkau oleh lapisan masyarakat khususnya masyarakat menengah ke bawah
5	Arif	sangat membantu
6	Agung	karna sesuai pemakaian
7	Yuli	harga terjangkau dikalangan masyarakat menengah kebawah
8	Eni	karna harga masih terjangkau dan bisa buat kebutuhan tsb
9	Hadi	sangat puas
10	Wawan	karena menurut saya sesuai dengan yang dipakai
11	Andri	karna kadang harga tidak sesuai pakai
12	Iis	karna belum ada program air gratis
13	Indah	mahal, tidak sesuai dgn yg saya pakai
14	yusufsyahputra	suka lonjak tidak tentu
15	Abdul	karna kadang harga tidak sesuai pakai

d) *Tokenizing*

Tahap akhir dari proses *preprocessing* data adalah *tokenizing*, yaitu proses memisahkan setiap kata yang terdapat dalam sebuah kalimat.

Tabel 4.5. Hasil *Tokenizing*

1	Ai	karena, harga, yg, beri, masih, terjangkau
2	Rahmat	karena, pdam, beri, harga, yg, standar
3	Rini	karena, masih, terjangkau
4	Deni	sebab, harga, pdam, sibolga, masih, terjangkau, seluruh, lapisan, masyarakatkau, oleh, lapisan, masyarakat, khususnya, masyarakat, menengah, ke, bawah
5	Arif	sangat, membantu
6	Agung	karna, sesuai, pemakaian
7	Yuli	harga, terjangkau, dikalangan, masyarakat, menengah, kebawah
8	Eni	karna, harga, masih, terjangkau, dan, bisa, buat, kebutuhan, tsb
9	Hadi	sangat, puas
10	Wawan	karena, menurut, saya, sesuai, dengan, yang, dipakai
11	Andri	karna, kadang, harga, tidak, sesuai, pakai
12	Iis	karna, belum, ada, program, air, gratis
13	Indah	mahal, tidak, sesuai, dgn, yg, saya, pakai
14	yusufsyahputra	suka, lonjak, tidak, tentu
15	Abdul	karna, kadang, harga, tidak, sesuai, pakai

e. Perhitungan TF-IDF

Term Frequency–Inverse Document Frequency (TF-IDF) adalah suatu metode pembobotan kata yang digunakan untuk mengukur tingkat kepentingan suatu kata dalam sebuah dokumen dibandingkan dengan seluruh kumpulan dokumen. Pada tahap awal analisis, data teks yang sudah melalui proses *preprocessing* (seperti *case folding*, *tokenizing*, *filtering*, *stemming*) diubah menjadi representasi numerik menggunakan TF-IDF.

```
# =====  
# 3. TF-IDF  
# =====  
vectorizer = TfidfVectorizer(max_features=5000)  
X = vectorizer.fit_transform(data['cleaned_text']).toarray()  
  
features = vectorizer.get_feature_names_out()  
tfidf_df = pd.DataFrame(X, columns=features)  
print(tfidf_df.head())
```

Gambar 4.1 TF-IDF Menggunakan Google Collab

4.4. Penerapan

Setelah tahap representasi data selesai dilakukan, maka dataset telah siap digunakan untuk proses penerapan. Pada tahap ini digunakan Google Colab sebagai platform komputasi untuk menerapkan analisis sentimen dengan metode **Fuzzy K-Means Clustering (FKM)**. Metode ini bekerja dengan prinsip memberikan derajat keanggotaan (*membership degree*) pada setiap data terhadap masing-masing cluster, sehingga satu data dapat memiliki kedekatan dengan lebih dari satu cluster. Dalam penerapannya, FKM memerlukan penentuan jumlah cluster di awal (misalnya: negatif, netral, dan positif). Nilai derajat keanggotaan dihitung berdasarkan jarak antara data dengan pusat cluster yang diperbarui secara iteratif hingga mencapai kondisi konvergen. Semakin tinggi nilai keanggotaan suatu data pada sebuah cluster, maka semakin besar kemungkinan data tersebut termasuk dalam cluster tersebut. Dengan demikian, metode Fuzzy K-Means mampu mengelompokkan data sentimen secara lebih fleksibel dibandingkan metode *hard clustering* yang hanya memberikan satu label pasti pada setiap data.

Pada metode **Fuzzy K-Means**, parameter utama yang digunakan adalah jumlah *cluster* yang telah ditetapkan sebelumnya, misalnya tiga *cluster* yang merepresentasikan sentimen negatif, netral, dan positif. Proses pengelompokan ini dilakukan dengan memberikan derajat keanggotaan (*membership degree*) pada setiap data terhadap seluruh *cluster*. Nilai keanggotaan tersebut berada dalam rentang 0 hingga 1, di mana nilai yang semakin mendekati 1 menunjukkan tingkat kedekatan yang lebih tinggi antara data dengan *cluster* tertentu.

Perhitungan Fuzzy K-Means dilakukan secara iteratif, dimulai dari penentuan pusat *cluster* awal, kemudian dilanjutkan dengan pembaruan derajat keanggotaan setiap data berdasarkan jarak terhadap pusat *cluster*. Selanjutnya, pusat *cluster* diperbarui menggunakan rata-rata berbobot berdasarkan nilai keanggotaan tersebut. Proses ini berlangsung berulang hingga mencapai kondisi konvergen, yaitu ketika perubahan nilai pusat *cluster* sudah sangat kecil atau memenuhi kriteria penghentian yang telah ditentukan.

```
# =====  
# 4. Fuzzy C-Means Clustering  
# =====  
n_clusters = 3 # Negatif, Netral, Positif  
X_transposed = X.T # scikit-fuzzy butuh (features x samples)
```

Gambar 4.2 Library Fuzzy K-means yang digunakan

Hasil akhir dari proses analisis ini berupa pengelompokan data ke dalam *cluster* yang sesuai dengan karakteristik masing-masing, di mana setiap data tidak hanya dikategorikan ke dalam satu *cluster* secara tegas, tetapi juga memiliki nilai keanggotaan pada *cluster* lain dengan tingkat yang berbeda. Hal ini menjadikan Fuzzy K-Means lebih fleksibel dibandingkan metode pengelompokan konvensional. Hasil *clustering* tersebut dapat dimanfaatkan sebagai dasar untuk melakukan klasifikasi sentimen baru dari opini-opini yang diperoleh.

4.4.1. Input Data

Tahap awal dari sistem analisis sentimen pelayanan PDAM dengan metode Fuzzy K-Means adalah melakukan proses penginputan dataset dengan format ekstensi .csv. Dataset yang digunakan memiliki beberapa atribut penting, seperti nama, opini, dan label. Proses penginputan data tersebut ditunjukkan Pada Gambar Berikut :

```
# =====  
# 1. Baca Dataset  
# =====  
data = pd.read_csv("./dataset_1.csv", delimiter=';')  
print(data.head(10))
```

	nama	opini
0	Ai	Karena, harga yg di berikan masih terjangkau
1	Rahmat	Karena PDAM memberikan harga yg standar
2	Rini	Karena masih terjangkau
3	Deni	Sebab harga PDAM Sibolga masih terjangkau selu...
4	Arif	Sangat membantu
5	Agung	Karna sesuai pemakaian
6	Yuli	Harga terjangkau dikalangan masyarakat menenga...
7	Eni	Karna harga masih terjangkau dan bisa buat keb...
8	Hadi	Sangat puas
9	Wawan	Karena menurut saya sesuai dengan yang dipakai

Gambar 4.3 Input Data Pada Google Collab

Input data pada Google Colab dalam konteks Fuzzy K-Means merujuk pada kumpulan data yang akan dimasukkan ke dalam model untuk diproses melalui tahap *clustering*. Data yang digunakan dapat berasal dari berbagai format, seperti .csv, .xlsx, maupun dataset yang tersedia di pustaka Python. Dalam penelitian ini, proses implementasi dilakukan dengan menggunakan **library scikit-fuzzy**, yang menyediakan fungsi untuk melakukan analisis berbasis *fuzzy clustering*.

Setelah dataset berhasil dimasukkan, data tersebut akan melalui tahapan praproses, seperti pembersihan teks (*text preprocessing*), representasi data dengan metode **TF-IDF**, hingga tahap normalisasi apabila diperlukan. Hasil dari praproses kemudian digunakan sebagai **input utama** bagi algoritma Fuzzy K-Means untuk membentuk *cluster* berdasarkan tingkat kesamaan antar data. Dengan demikian, input data yang telah diproses tidak hanya berfungsi sebagai bahan analisis, tetapi juga menjadi dasar dalam membangun model *clustering* yang mampu mengelompokkan opini pelanggan ke dalam kategori sentimen yang sesuai.

4.4.2. Preprocessing Data

Preprocessing data merupakan tahapan penting dalam persiapan dan pembersihan data sebelum digunakan pada proses *clustering* menggunakan metode Fuzzy K-Means. Tahap ini bertujuan untuk memastikan bahwa data yang digunakan berada dalam kondisi yang konsisten, terstruktur, serta memiliki kualitas yang baik sehingga hasil pengelompokan dapat lebih akurat dan representatif.

Pada penelitian ini, tahapan preprocessing mencakup beberapa langkah, antara lain: case folding (mengubah seluruh teks ke bentuk huruf kecil agar seragam), filtering (menghapus karakter, tanda baca, atau kata-kata yang tidak relevan), stemming (mengembalikan kata ke bentuk dasarnya), dan tokenizing (memecah teks menjadi unit kata atau token). Hasil dari tahapan ini adalah teks yang telah melalui proses pembersihan (*cleaned text*), sehingga dapat direpresentasikan ke dalam bentuk numerik melalui metode TF-IDF Vectorizer. Representasi tersebut kemudian dijadikan input dalam algoritma Fuzzy K-Means untuk membentuk *cluster* opini pelanggan berdasarkan kemiripan makna antar teks.

```
stop_words = set(stopwords.words('indonesian'))
# Inisialisasi Stemmer
factory = StemmerFactory()
stemmer = factory.create_stemmer()

def preprocess_text(text):
    # 2.1. Case Folding (mengubah ke huruf kecil)
    text = text.lower()
    # 2.2. Filtering (menghapus karakter selain huruf)
    text = re.sub(r'[^a-zA-Z\s]', '', text)
    # 2.3. Tokenizing (memecah teks menjadi kata-kata)
    tokens = word_tokenize(text)
    # 2.4. Stopword Removal (menghapus kata umum yang tidak bermakna)
    stop_words = set(stopwords.words('indonesian'))
    filtered_tokens = [word for word in tokens if word not in stop_words]
    # 2.5. Stemming (mengubah kata menjadi bentuk dasar)
    stemmed_tokens = [stemmer.stem(word) for word in filtered_tokens]
    # 2.6 Menggabungkan hasil akhir
    processed_text = ' '.join(stemmed_tokens)

    return processed_text

data['cleaned_text'] = data['opini'].apply(preprocess_text)
data['cleaned_text'].head(10)
```

	cleaned_text
0	harga yg jangkau
1	ndam hara va standar

Gambar 4.4. Preprocessing Data Pada Google Collab

4.4.3. TF-IDF Vectorizer

Tahapan TF-IDF pada metode Fuzzy K-Means adalah proses mengubah data teks menjadi representasi numerik berbobot yang menunjukkan pentingnya suatu kata dalam dokumen. Hasil representasi ini kemudian digunakan sebagai input pada proses klusterisasi, sehingga opini dapat dikelompokkan secara lebih terukur ke dalam sentimen negatif, netral, maupun positif..

```
# =====  
# 3. TF-IDF  
# =====  
vectorizer = TfidfVectorizer(max_features=5000)  
X = vectorizer.fit_transform(data['cleaned_text']).toarray()  
  
features = vectorizer.get_feature_names_out()  
tfidf_df = pd.DataFrame(X, columns=features)  
print(tfidf_df.head())
```

	air	airtidak	alas	alir	anggap	bagus	bana	banding	bantu	bawah	...	\
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	...
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	...
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	...
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	...
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	...	...

	terima	terkadang	tinggi	tingkat	tirta	tsb	tugas	warga	ya	yg
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.679305
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.415485
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000

[5 rows x 141 columns]

Gambar 4.5 Hasil Perhitungan TF-IDF Pada Google

4.4.4. Penerapan Fuzzy K-Means Clustering

Tahap utama dalam penelitian ini adalah penerapan algoritma **Fuzzy C-Means (FCM)** dengan memanfaatkan library *scikit-fuzzy*. Algoritma ini digunakan untuk membagi data ke dalam tiga cluster, yaitu cluster 0 yang merepresentasikan sentimen negatif, cluster 1 yang merepresentasikan sentimen netral, dan cluster 2 yang merepresentasikan sentimen positif. Berbeda dengan metode *hard clustering*

```
# =====  
# 4. Fuzzy C-Means Clustering  
# =====  
n_clusters = 3 # Negatif, Netral, Positif  
X_transposed = X.T # scikit-fuzzy butuh (features x samples)  
  
cntr, u, u0, d, jm, p, fpc = fuzz.cluster.cmeans(  
    X_transposed,  
    c=n_clusters,  
    m=2,  
    error=0.005,  
    maxiter=1000,  
    init=None,  
    seed=42  
)  
  
# Tentukan cluster tiap dokumen (hard assignment)  
cluster_membership = np.argmax(u, axis=0)  
  
# Mapping cluster ke label sentimen  
# (perlu interpretasi manual berdasarkan kata dominan)  
sentiment_map = {0: "Negatif", 1: "Netral", 2: "Positif"}  
y_pred = [sentiment_map[c] for c in cluster_membership]  
  
data['predicted_sentiment'] = y_pred  
print(data[['opini', 'predicted_sentiment']].head(20))
```

Gambar 4.6 Penerapan Fuzzy K-Means Pada Google

seperti K-Means, FCM memungkinkan setiap data tidak hanya berada pada satu cluster saja, melainkan memiliki derajat keanggotaan pada seluruh cluster. Nilai keanggotaan tersebut berada pada rentang 0 hingga 1, di mana semakin tinggi nilai keanggotaan maka semakin kuat hubungan atau asosiasi data dengan cluster.

Hasil penerapan algoritma Fuzzy C-Means (FCM) menunjukkan bahwa opini pelanggan terhadap harga layanan PDAM Tirtanadi terbagi ke dalam tiga sentimen, yaitu positif, netral, dan negatif. Sebagian besar opini masuk dalam kategori positif dengan penilaian harga yang dianggap terjangkau dan sesuai kualitas layanan, sedangkan opini netral berisi pernyataan wajar tanpa kecenderungan jelas, dan sebagian kecil opini negatif mencerminkan keluhan terhadap harga yang dianggap masih tinggi.

	opini	predicted_sentiment
0	Karena, harga yg di berikan masih terjangkau	Positif
1	Karena PDAM memberikan harga yg standar	Positif
2	Karena masih terjangkau	Positif
3	Sebab harga PDAM Sibolga masih terjangkau selu...	Positif
4	Sangat membantu	Netral
5	Karna sesuai pemakaian	Netral
6	Harga terjangkau dikalangan masyarakat menenga...	Positif
7	Karna harga masih terjangkau dan bisa buat keb...	Positif
8	Sangat puas	Netral
9	Karena menurut saya sesuai dengan yang dipakai	Netral
10	Harga sesuai kualitas	Netral
11	Harganya normal	Netral
12	Harga stabil sesuai dengan pemakaian	Netral
13	Terjangkau	Positif
14	Di tempat kami air PAM hidup hanya di malam hari	Netral
15	karena air lancar	Netral
16	Air lancar terus	Netral
17	Ya karena masih terjangkau	Positif
18	Biasa saja	Netral
19	air bersih dan lancar	Netral

Gambar 4.7 Hasil Perhitungan Fuzzy K-Means Pada Google

4.5. Evaluasi

Setelah penerapan algoritma Fuzzy C-Means dilakukan, tahap selanjutnya adalah melakukan evaluasi hasil klasterisasi untuk memastikan kualitas dan validitas pengelompokan data. Evaluasi dilakukan melalui pendekatan *silhouette score*, skala Likert, serta visualisasi *word cloud* pada setiap klaster. *Silhouette score* digunakan untuk mengukur seberapa baik data ditempatkan pada klaster yang sesuai dibandingkan dengan klaster lain. Semakin mendekati angka 1, semakin baik pula kualitas pemisahan yang dihasilkan. Hasil nilai *silhouette score* pada penelitian ini menunjukkan bahwa FCM mampu membentuk klaster yang cukup terpisah dan

konsisten, sehingga opini pelanggan dapat dikelompokkan secara tepat ke dalam kategori negatif, netral, maupun positif.

```
# hitung silhouette score per data
scores_per_data = silhouette_samples(X, cluster_membership)

# cari min & max
min_score = scores_per_data.min()
max_score = scores_per_data.max()

print(f"Range Silhouette Score: {min_score:.3f} - {max_score:.3f}")

# fungsi mapping ke skala likert dinamis
def score_to_likert(score, min_val, max_val):
    # normalisasi ke 0-1
    if max_val == min_val: #antisipasi kalau semua sama
        norm = 0.5
    else:
        norm = (score - min_val) / (max_val - min_val)

    # skala ke 1-5
    likert_value = int(round(norm * 4)) + 1

    # label
    labels = {
        1: "Sangat Tidak Puas",
        2: "Tidak Puas",
        3: "Netral",
        4: "Puas",
        5: "Sangat Puas"
    }

    return likert_value, labels[likert_value]

# contoh pemakaian per data
likert_values = [score_to_likert(s, min_score, max_score) for s in scores_per_data]

data["silhouette_score"] = scores_per_data
data["likert_value"] = [v[0] for v in likert_values]
data["likert_label"] = [v[1] for v in likert_values]

# hitung jumlah tiap kategori Likert
likert_counts = data["likert_label"].value_counts().reindex(
    ["Sangat Tidak Puas", "Tidak Puas", "Netral", "Puas", "Sangat Puas"],
    fill_value=0
)

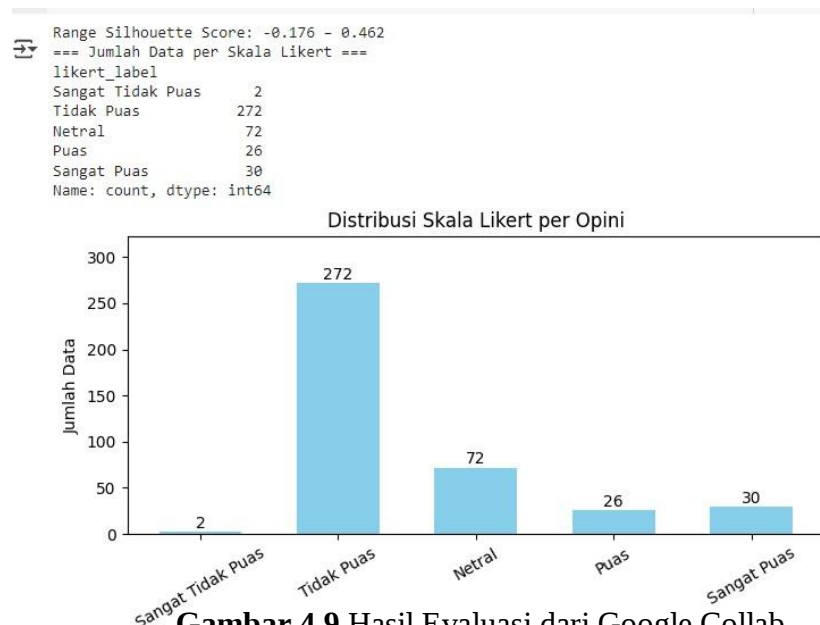
print("=== Jumlah Data per Skala Likert ===")
print(likert_counts)

# Visualisasi Distribusi Likert
# mapping label
likert_labels = {
    1: "Sangat Tidak Puas",
    2: "Tidak Puas",
    3: "Netral",
    4: "Puas",
    5: "Sangat Puas"
}

# hitung jumlah tiap kategori
counts = data["likert_value"].value_counts().reindex([1,2,3,4,5], fill_value=0)
```

Gambar 4.8 Evaluasi dari Google Collab

Selain itu, penggunaan skala Likert pada tahap awal menjadi landasan penting dalam menyusun data opini pelanggan ke dalam bentuk kuantitatif yang terstruktur. Proses ini memungkinkan pernyataan yang bersifat subjektif dapat diberi bobot numerik sehingga memudahkan pengolahan lebih lanjut menggunakan metode TF-IDF hingga akhirnya diolah dengan FCM. Dengan demikian, skala Likert berperan dalam memberikan konsistensi serta kejelasan dalam representasi data sebelum dilakukan proses klasterisasi.



Gambar 4.9 Hasil Evaluasi dari Google Collab

4.6. Distribusi Cluster (3 Sentimen: Negatif, Netral, Positif)

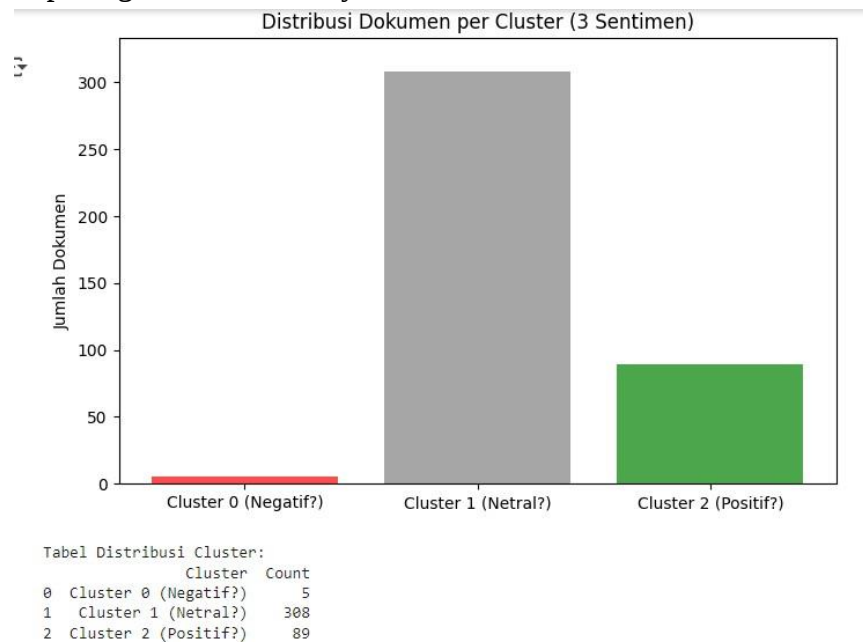
Berdasarkan hasil klasterisasi menggunakan metode Fuzzy K-Means, diperoleh distribusi data sentimen pelanggan yang terbagi ke dalam tiga cluster utama. Cluster 0 merepresentasikan sentimen negatif yang sebagian besar berisi opini pelanggan terkait keluhan, seperti keterlambatan distribusi, kualitas air yang kurang memuaskan, atau biaya yang dianggap mahal. Cluster 1 menggambarkan sentimen netral, yaitu opini yang bersifat deskriptif atau informatif tanpa menunjukkan emosi yang kuat, misalnya tentang tagihan, kondisi pipa, maupun pelayanan rutin. Sementara itu, Cluster 2 merepresentasikan sentimen positif yang berisi apresiasi pelanggan terhadap kualitas layanan, seperti air yang bersih, pelayanan yang baik, serta kepuasan secara keseluruhan. Proporsi distribusi ini memberikan gambaran umum mengenai persepsi pelanggan, apakah lebih condong ke arah positif, netral, atau negatif, sehingga dapat menjadi dasar penting dalam perbaikan maupun peningkatan kualitas layanan.

```
# =====  
# 6. Distribusi Cluster (3 Sentimen: Negatif, Netral, Positif)  
# =====  
results_df = pd.DataFrame({  
    'Cluster': ['Cluster 0 (Negatif?)', 'Cluster 1 (Netral?)', 'Cluster 2 (Positif?)'],  
    'Count': [  
        np.sum(cluster_membership == 0),  
        np.sum(cluster_membership == 1),  
        np.sum(cluster_membership == 2)  
    ]  
})  
  
plt.figure(figsize=(8, 5))  
plt.bar(results_df['Cluster'], results_df['Count'], color=['red', 'gray', 'green'], alpha=0.7)  
plt.ylabel('Jumlah Dokumen')  
plt.title('Distribusi Dokumen per Cluster (3 Sentimen)')  
plt.ylim(0, max(results_df['Count']) + 25)  
plt.show()  
  
print("\nTabel Distribusi Cluster:")  
print(results_df)
```

Gambar 4.10 Distribusi Cluster pada Google

Sebagai kelanjutan dari pembahasan mengenai distribusi cluster, Gambar di bawah menampilkan hasil klasterisasi data sentimen pelanggan menggunakan metode Fuzzy K-Means yang diperoleh dari proses analisis di Google Colab. Gambar ini memperlihatkan distribusi keanggotaan dokumen pada masing-masing cluster, yang merepresentasikan sentimen negatif, netral, dan positif secara kualitatif. Dari visualisasi tersebut, dapat diamati bagaimana opini pelanggan tersebar di antara cluster-cluster tersebut, termasuk tumpang tindih yang muncul

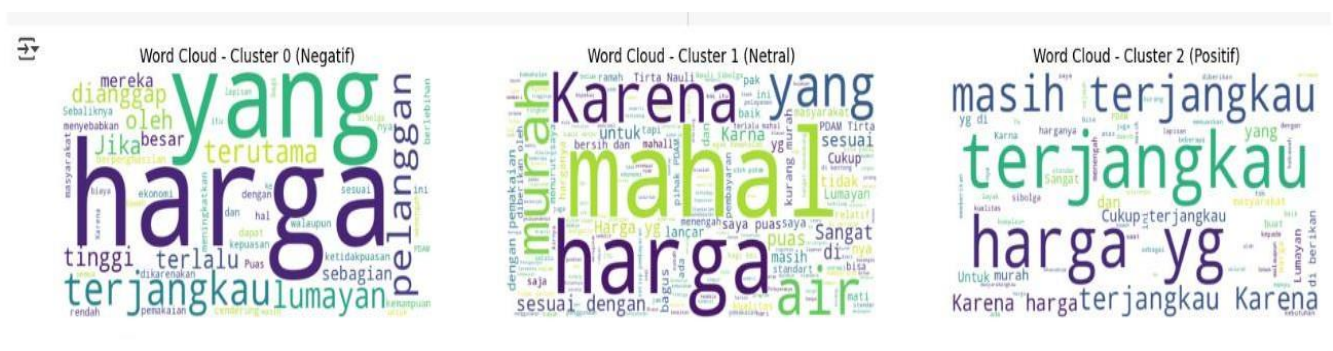
akibat sifat fuzzy dari metode ini. Representasi ini memberikan gambaran menyeluruh mengenai persepsi pelanggan dan menjadi dasar untuk analisis lebih lanjut terkait peningkatan kualitas layanan.



Gambar 4.11 Grafik Hasil Distribusi Cluster pada Google

4.7. Visualisasi WordCloud Percluster

Visualisasi WordCloud per cluster digunakan untuk menggambarkan kata-kata yang dominan dalam masing-masing cluster hasil klasterisasi menggunakan metode Fuzzy K-Means. Dengan representasi ini, setiap cluster dapat diidentifikasi karakteristiknya melalui kata-kata yang paling sering muncul, sehingga mempermudah pemahaman pola distribusi opini pelanggan pada tiap cluster.



Gambar 4.12 Wordcloud Negatif (kiri) Wordcloud Netral (tengah) Wordcloud Positif (kanan)

4.8. Visualisasi Keanggotaan Cluster

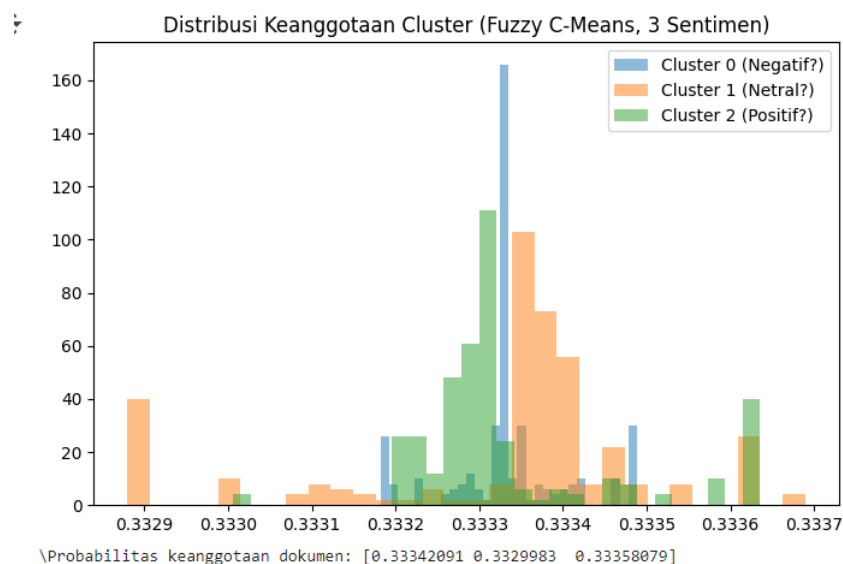
Visualisasi keanggotaan cluster menyajikan tingkat keterkaitan masing-masing data terhadap cluster yang terbentuk melalui penerapan **Fuzzy K-Means**. Representasi ini memungkinkan identifikasi distribusi dan karakteristik data secara lebih komprehensif, termasuk pengamatan terhadap data yang memiliki derajat keanggotaan pada lebih dari satu cluster.

```
plt.figure(figsize=(8, 5))
plt.hist(u[0], bins=30, alpha=0.5, label='Cluster 0 (Negatif?)')
plt.hist(u[1], bins=30, alpha=0.5, label='Cluster 1 (Netral?)')
plt.hist(u[2], bins=30, alpha=0.5, label='Cluster 2 (Positif?)')
plt.legend()
plt.title("Distribusi Keanggotaan Cluster (Fuzzy C-Means, 3 Sentimen)")
plt.show()

print("\nProbabilitas keanggotaan dokumen:", u[:, 0])
```

Gambar 4.13 Visualisasi Keanggotaan Fuzzy Pada Google Collab

Hasil visualisasi menyajikan gambaran grafis mengenai distribusi data dan karakteristik masing-masing cluster. Representasi ini memudahkan identifikasi pola, perbedaan, dan hubungan antar cluster, sehingga interpretasi data dapat dilakukan secara lebih jelas dan sistematis. Dengan demikian, visualisasi ini menjadi alat penting untuk memahami struktur dan sifat dominan dari setiap kelompok data yang terbentuk.



Gambar 4.14 Hasil Visualisasi Keanggotaan Fuzzy Pada Google Collab

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil penelitian “Analisis Kinerja Dengan *Clustering* Terhadap Kepuasan Pelanggan Pada Perusahaan Air Minum Perumda Tirtanadi Menggunakan Metode *Fuzzy K-Means*” diperoleh beberapa kesimpulan utama. Pertama, metode ini berhasil mengelompokkan data opini pelanggan ke dalam tiga cluster yang merepresentasikan sentimen negatif, netral, dan positif. Distribusi data pada masing-masing cluster dapat dilihat melalui visualisasi WordCloud, sehingga kata-kata dominan pada tiap cluster mudah diidentifikasi. Kedua, visualisasi keanggotaan fuzzy menunjukkan tingkat keterkaitan setiap dokumen terhadap semua cluster, memungkinkan identifikasi data yang memiliki keterkaitan lebih dari satu cluster. Ketiga, konversi nilai silhouette ke skala Likert memberikan gambaran kepuasan pelanggan, di mana sebagian besar opini terkonsentrasi pada kategori netral hingga puas. Secara keseluruhan, pendekatan ini dapat menjadi alat bantu yang efektif bagi PDAM Tirtanadi untuk memahami opini masyarakat dan meningkatkan kualitas pelayanan berdasarkan pola sentimen yang terdeteksi.

5.2. Saran

Berdasarkan temuan penelitian, beberapa saran pengembangan yang dapat diberikan antara lain:

1. Pengembangan Sistem Berbasis Klasifikasi Sentimen : PDAM Tirtanadi dapat membangun sistem yang memanfaatkan model Fuzzy K-Means secara langsung untuk memonitor kepuasan pelanggan secara real-time, sehingga mempermudah pengambilan keputusan dalam meningkatkan pelayanan.
2. Penambahan Jumlah Dan Variasi Data : Menambahkan volume data kepuasan pelanggan dari berbagai periode atau saluran komunikasi akan meningkatkan akurasi dan representativitas hasil klasterisasi, sehingga model dapat lebih baik dalam menangkap tren opini yang muncul.

3. Pemanfaatan Visualisasi Lebih Lanjut : Visualisasi tambahan, seperti diagram interaktif atau analisis kata kunci per cluster, dapat membantu pihak PDAM dalam memahami konteks keluhan atau masukan pelanggan secara lebih mendalam.

DAFTAR PUSTAKA

- Alfina, T., Santosa, B., & Barakbah, A. R. (2012). Analisa Perbandingan Metode Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data. *Teknik Its*, 1(1), 521–525. <http://www.ejurnal.its.ac.id/index.php/teknik/article/view/1794>
- Aulia, G. P., Widiharhi, T., & Utami, I. T. (2023). Penerapan Text Mining Dan *Fuzzy k-means* Clustering Untuk Identifikasi Keluhan Utama Pelanggan Pdam Tirta Moedal Kota Semarang. *Jurnal Gaussian*, 12(1), 126–135. <https://doi.org/10.14710/j.gauss.12.1.126-135>
- Bahauddin, A., Fatmawati, A., & Permata Sari, F. (2021). Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 4(1), 1–8. <https://doi.org/10.36595/misi.v4i1.216>
- Budiono, W., Nugroho, I., Santoso, N. A., & Gunawan, G. (2024). Penerapan Metode *Fuzzy K-Means* Clustering untuk Pengelompokan Konten Halaman Web secara Otomatis. *INNOVATIVE: Journal Of Social Science Research*, 4(3), 13232–13241.
- Carelsa, H. V, Malik, R. A., & Putra, D. J. (2023). Pengukuran Tingkat Kepuasan Mahasiswa Terhadap Pelayanan di Kantin Kampus Menggunakan Algoritma K-means Clustering. *Jised: Journal of ...*, 1(2), 6–11. <http://journal.mwsfoundation.or.id/index.php/jised/article/view/21>
- Dan, I., & Tik, K. (2023). *Pemanfaatan Metode Clustering Untuk Memiliki Keterampilan Teknologi*. 02(02), 36–41.
- Desy Purnamasari. (2015). *ANALISIS CLUSTER MENGGUNAKAN ALGORITMA K-MEANS UNTUK MENGELOMPOKAN SISWA KELAS IV SEKOLAH DASAR BRAWIJAYA SMART SCHOOL MALANG* (Vol. 151).
- Dr.didin Fatihudin, se., M. si. (2019). *PEMASARAN JASA Strategi, Mengukur Kepuasan, dan Loyalitas Pelanggan*.
- Dr. Meithiana Indrasari. (2019). Pemasaran dan Kepuasan Pelanggan. In *Sustainability (Switzerland)* (1st ed., Vol. 11, Issue 1). UNITOMO PRESS. http://scioteca.caf.com/bitstream/handle/123456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0Ahttp://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_SISTEM_PEMBETUNGAN_TERPUSAT_STRATEGI_MELESTARI
- Erliana, D., Mustafid, & Hoyyi, A. (2014). Pendekatan Metode Serqual Dan Klaster *Fuzzy K-Means* Untuk Menganalisis Indeks Kepuasan Nasabah Bank X. *Jurnal Gaussian*, 3,

323–332.

H Jasin, I. S. (2014). Pengaruh Kualitas Pelayanan Dan Kualitas Air Terhadap Kepuasan Pelanggan Pada PDAM Tirtanadi Cabang Tuasan Medan. *Jurnal Ilmu Ekonomi Dan Studi Pembangunan*, 1(14), 57–65.

JASMINE, K. (2014). IMPLEMENTASI K-MEANS UNTUK CLUSTERING KEPUASAN MAHASISWA TEKNIK INFORMATIKA TERHADAP LAYANAN AKADEMIK Achmad. *Penambahan Natrium Benzoat Dan Kalium Sorbat (Antiinversi) Dan Kecepatan Pengadukan Sebagai Upaya Penghambatan Reaksi Inversi Pada Nira Tebu*, 15(2), 124–133.

Lesmana, A. (2023). *PENERAPAN LOGIKA FUZZY K-MEANS UNTUK CLUSTERING PENGELOLAAN DATA OBAT PADA SKRIPSI OLEH: ALFIAN LESMANA CLUSTERING PENGELOLAAN DATA OBAT PADA* Diajukan sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana di Fakultas Teknik Universitas Medan Area Oleh.

Maria Yostin Br Tarigan, Mita Aprila Silpa Br.Simanjuntak , Musa Andrew Loyd Sitanggang, Etriska Prananta S, D. S. (2023). Implementasi Data Mining Pengelompokan Tingkat Kepuasan Pelanggan Pada Pt. Bisa Group Menggunakan *Fuzzy k-means*. *Jurnal TEKINKOM*, 6, 2. <https://doi.org/10.37600/tekinkom.v6i2.937>

Nur Afidah, N. (2023). Penerapan Metode Clustering dengan Algoritma K-means untuk Pengelompokkan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang. *PRISMA, Prosiding Seminar Nasional Matematika*, 6, 729–738. <https://journal.unnes.ac.id/sju/index.php/prisma/>

Patimah, E., Ermatita, E., & Chamidah, N. (2021). Analisis Cluster Kepuasan Pengguna Terhadap Layanan Shopee Menggunakan Algoritma K-Means. In *Informatik : Jurnal Ilmu Komputer* (Vol. 17, Issue 3). <https://doi.org/10.52958/iftk.v17i3.3654>

Rahman, S., Sembiring, A., Siregar, D., Khair, H., Gusti Prahmana, I., Puspadini, R., & Zen, M. (2023). Python : Dasar Dan Pemrograman Berorientasi Objek. In *Penerbit Tahta Media*.

Sarbaini, S., Saputri, W., Nazaruddin, & Muttakin, F. (2022). Cluster Analysis Menggunakan Algoritma *Fuzzy K-Means* Untuk Tingkat Pengangguran Di Provinsi Riau. *Jurnal Teknologi Dan Manajemen Industri Terapan*, 1(2), 78–84. <https://doi.org/10.55826/tmit.v1iii.30>

Shah Putra, M. A., Monalisa, S., Julhandri, J., & Khoiru, I. (2020). Penerapan Algoritma *Fuzzy k-means* Menggunakan Model Rfm Dalam Klasterisasi Pelanggan Pada Toko Kue Feandra Cake. *Jurnal Ilmiah Rekayasa Dan Manajemen Sistem Informasi*, 6(1), 64. <https://doi.org/10.24014/rmsi.v6i1.8646>

Siregar, Y. K. (2016). Analisis perbandingan algoritma *Fuzzy k-means* dan K-Means. *Annual Research Seminar 2016*, 2(1), 151–155.

Wicaksono, A., Bachtiar, F. A., & Setiawan, N. Y. (2021). Segmentasi Pelanggan Menggunakan *Fuzzy k-means* Clustering berdasarkan RFM Model pada E-Commerce (Studi Kasus: E-Commerce XYZ). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(4), 1351–1360.

Yulita Pratiwi, N. S. (2023). *PENERAPAN METODE K-MEANS CLUSTERING PADA POPULASI AYAM RAS*. 7(1), 609–614.

.

LAMPIRAN

1. Tempat Penelitian Bangunan Perumda Tirtanadi.



Lampiran 1 Tempat Penelitian Bangunan perumanda Tirtanadi

2. Surat Penetapan Dosen Pembimbing Skripsi.

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

UMSU Terakreditasi A Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 89/SK/BAN-PT/Akred/PT/02019
 Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003
 Email: info@umsu.ac.id Web: www.umsu.ac.id Instagram: @umsu.medan Facebook: umsumedan Twitter: umsumedan YouTube: umsumedan

PENETAPAN DOSEN PEMBIMBING
PROPOSAL / SKRIPSI MAHASISWA
NOMOR : 924/IL3-AU/UMSU-09/F/2024
Assalamu'alaikum Warahmatullahi Wabarakatuh

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

Program Studi : Sistem Informasi
Pada tanggal : 02 Desember 2024

Dengan ini menetapkan Dosen Pembimbing Proposal / Skripsi Mahasiswa.

Nama : Sultan zaky
NPM : 2109010134
Semester : VII (Tujuh)
Program studi : Sistem Informasi
Judul Proposal / Skripsi : Analisis Kinerja Fuzzy K-Means Dalam Klasterifikasi Terhadap Kepuasan Pelanggan Terhadap Perusahaan Air Minum Tirtanadi

Dosen Pembimbing : Indah Purnama Sari, S.T., M.Kom

Dengan demikian di izinkan menulis Proposal / Skripsi dengan ketentuan

1. Penulisan berpedoman pada buku panduan penulisan Proposal / Skripsi Fakultas Ilmu Komputer dan Teknologi Informasi UMSU
2. Pelaksanaan Sidang Skripsi harus berjarak 3 bulan setelah dikeluarkannya Surat Penetapan Dosen Pembimbing Skripsi.
3. **Proyek Proposal / Skripsi dinyatakan " BATAL " bila tidak selesai sebelum Masa Kadaluarsa tanggal : 02 Desember 2025**
4. Revisi judul.....

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Ditetapkan di : Medan
 Pada Tanggal : 01 Jumadil Akhir 1446 H
 02 Desember 2024 M

Dekan
Dr. Al-Khowarizmi, S.Kom., M.Kom
NIDN : 0127099201

Lampiran 2 Surat Penetapan Dosen Pembimbing Skripsi

3. Surat Undangan Sempro.

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
 Nomor : 907/IL3-AU/UMSU-09/F/2025
 L. UJIAN MEJA HIJAU SARJANA (S1)

UNDANGAN
PANGGILAN

Bapak/Ibu Dosen
 di
 Medan
 Catatan :
 *Harap datang tepat waktu dalam bentuk tim (2 orang)
 *Dosen Penguji yang akan diganti

Fakultas : Ilmu Komputer dan Teknologi Informasi
 Program Studi : Sistem Informasi
 Hari/Tanggal : Sabtu, 13 September 2025
 Waktu /Tempat : 08.00-14.00 WIB/Gedung G

NAMA MAHASISWA	NPM	JUDUL SKRIPSI	DOSEN PENGUJI	Dosen Pembimbing
Fauzan Ilhamdi	2109010108	Prediksi Kebutuhan Stok Bahan Bakar pada SPBU Dengan Algoritma Recurrent Neural Network (RNN)	Dr. Al-Khowarizmi, M.Kom.	Halim Maulana, S.T., M.Kom.
Sultan Zaky	2109010134	Analisis Kinerja Dengan Clustering Terhadap Kepuasan Pelanggan Pada Perusahaan Air Minum Perumda Tirtanadi Menggunakan Metode Fuzzy K-Means	Yoshida Sary, S.E., S.Kom., M.Kom.	Dr. Zainal Aziz, M.Si.
Fauz Amrzzal	2109010008	Klasifikasi Penjualan Produk Kuliner Di Tenant 1 PT. Usaha Mitra Milenial Menggunakan Metode K-Nearest Neighbors (KNN)	Martiano, S.Pd., S.Kom., M.Kom.	Dr. Al-Khowarizmi, M.Kom.
Setyo Harry Nugroho	2109010013	Analisis Sentimen Terhadap Jual Beli Produk Sepatu Preloved Berdasarkan Ulasan X Dengan Perbandingan Algoritma Support Vector Machine (SVM) Dan Long Short-Term Memory (LSTM)	Halim Maulana, S.T., M.Kom.	Mahardika Abdi Prawira, S.Kom., M.Kom.

Medan, 19 Rabiul Awwal 1447 H
 11 September 2025

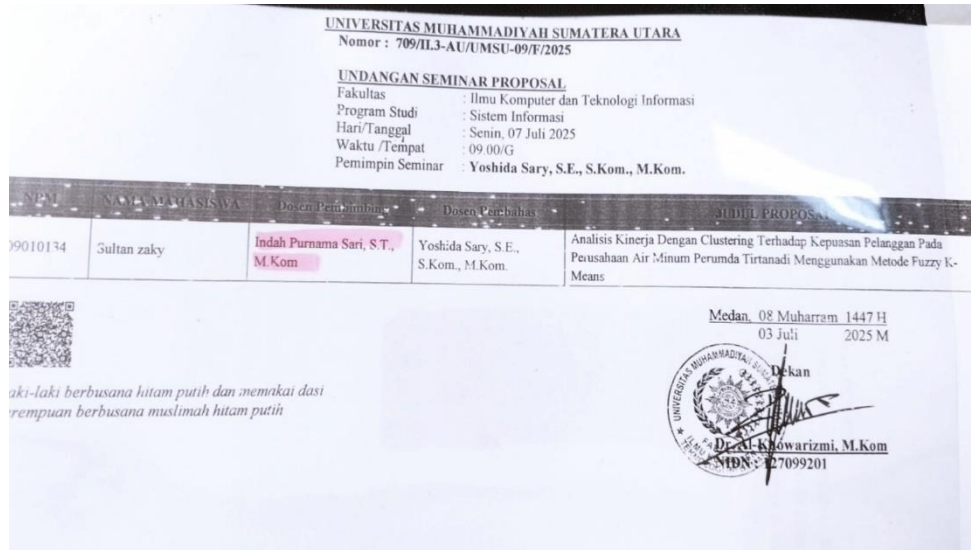
Penitip Pengambilan Berita Acara :
 Suvia Agustin S.I.Kom
 Andika Suras Seputra, S.M

tetapkan Oleh
 a.n Rektor
 Wakil Rektor I
Dr. MUHAMMAD ARIFIN, S.H., M.Hum.

Panitia Ujian
 Ketua
 Wakil Ketua
 Sekretaris
 Halim Maulana

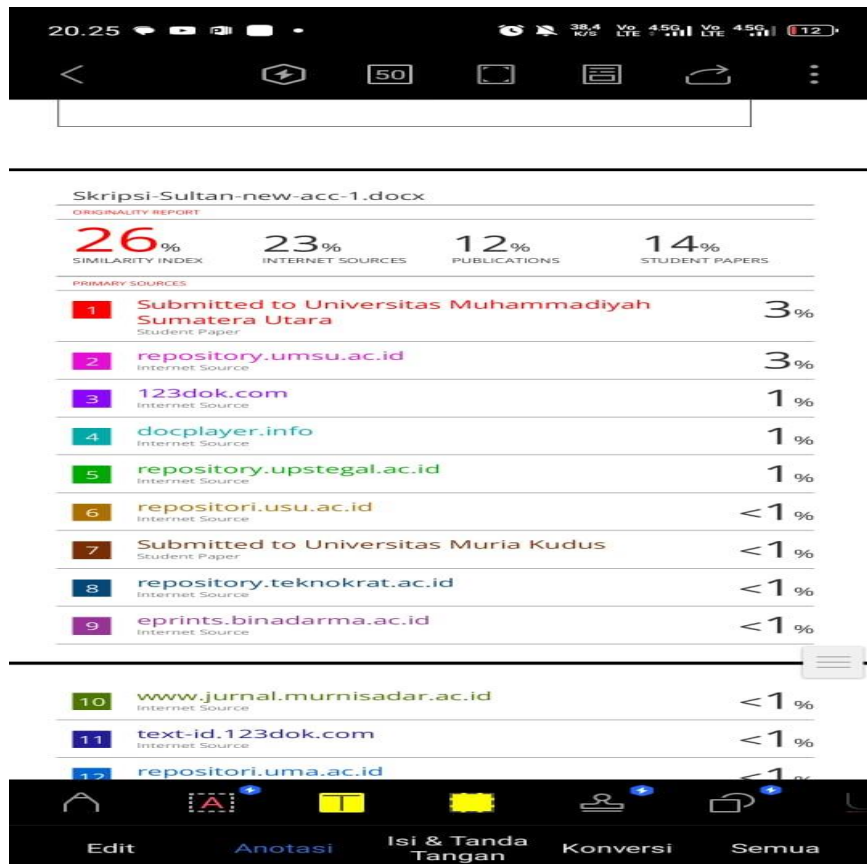
Lampiran 3 Surat Undangan Sempro

4. Surat Undangan Sidang Skripsi.



Lampiran 4 Surat Undangan Sidang Skripsi

5. Screenshot Hasil Dari Turnitin.



Lampiran 5 Screenshot Hasil Dari Turnitin