

**SKRIPSI**  
**ANALISIS SENTIMEN KEPUASAN PENGGUNA**  
**LABORATORIUM ICT UNIVERSITAS SULAWESI BARAT**  
**MENGGUNAKAN METODE (*K-NEAREST NEIGHBOR*)**



**SAMSANI**  
**D0218044**

**PROGRAM STUDI INFORMATIKA**  
**FAKULTAS TEKNIK**  
**UNIVERSITAS SULAWESI BARAT**  
**2023**

**HALAMAN PERSETUJUAN**  
**SKRIPSI**  
**ANALISIS SENTIMEN KEPUASAN PENGGUNA**  
**LABORATORIUM ICT UNIVERSITAS SULAWESI BARAT**  
**MENGGUNAKAN METODE (*K-NEAREST NEIGHBOR*)**

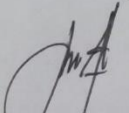
Diusulkan oleh

**SAMSANI**  
**D0218044**

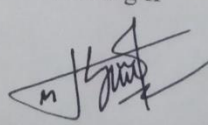
Telah disetujui

Pada tanggal 17 November 2023

Pembimbing I

  
**Indra. S.Kom..M.M**  
NIDN: 0031077904

Pembimbing II

  
**Muzaki. S.Kom..M.M**  
NIP: 198704302022031002

**LEMBAR PENGESAHAN**  
**ANALISIS SENTIMEN KEPUASAN PENGGUNA**  
**LABORATORIUM ICT UNIVERSITAS SULAWESI BARAT**  
**MENGGUNAKAN METODE (K-NEAREST NEIGHBOR)**

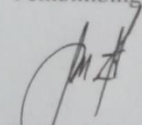
Telah dipersiapkan dan disusun oleh

**SAMSANI**  
**D0218044**

Telah dipertahankan di depan Tim Penguji  
Pada tanggal 8 November 2023

Susunan Tim Penguji

Pembimbing 1

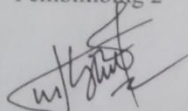


**Indra. S.Kom..M.M**  
NIDN: 0031077904

Penguji 1

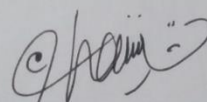
**Arnita Irianti. S.Si..M.Si**  
NIP:198708062018032001

Pembimbing 2



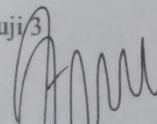
**Muzaki. S.Kom..M.M**  
NIP: 198704302022031002

Penguji 2



**Chairi Nur Insani. S.Kom..M.T**  
NIDN: 0027079404

Penguji 3



**Nurhikma Arifin. S.Kom..MT**  
NIP:199304252022032011DN

## **PERNYATAAN**

Dengan ini saya menyatakan bahwa penelitian ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan disuatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Majene, 16- Oktober - 2023

**SAMSANI**  
D0218044

## KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT atas segala rahmat-Nya, sehingga penulis dapat menyelesaikan yang berjudul “**analisis sentimen kepuasan pengguna laboratorium ICT universitas sulawesi barat**”.

Shalawat dan salam atas junjungan Nabi Muhammad SAW yang menjadi suritauladan bagi hidup penulis dan umat manusia, sehingga mendorong dan memberi motivasi kepada penulis untuk selalu menuntut ilmu sesuai denganuntutannya.

Dalam menyusun penelitian ini, walaupun banyak kesulitan dan hambatan yang penulis alami, namun berkat adanya dukungan, dorongan dan semangat dari orang terdekat di sekitar, sehingga penulis mampu menyelesaikan penelitian ini. Oleh karena itu penulis pada kesempatan mengucapkan terima kasih sedalam-dalamnya kepada:

1. Ibu, Ayah dan Kakak atas semua doa dan bantuan finansial yang senantiasa diberikan guna kelancaran penulis dalam menyelesaikan pendidikan serta seluruh keluarga yang senantiasa memberikan dukungan kepada penulis.
2. Ibu Dr.Ir. Hafsa Nirwana, M.T. Selaku Dekan Fakultas Teknik Universitas Sulawesi Barat.
3. Bapak Ir. Sugiarto Cokrowibowo, S.Si., M.T selaku wakil dekan Fakultas Teknik Universitas Sulawesi Barat

4. Bapak Muh. Fahmi Rustan, S. Kom., M.T selaku Ketua Program Studi dan seluruh dosen serta staf Teknik Informatika, Fakultas Teknik, Universitas Sulawesi Barat.
5. Bapak Indra, S.Kom.,M.M. selaku dosen pembimbing pertama saya, dengan arahan, tenaga, dan pikiran yang selalu membimbing penulis dalam menyelesaikan tulisan ini.
6. Bapak Muzaki, S.Kom.,M.M. selaku dosen pembimbing kedua saya, dengan arahan, tenaga, dan pikiran yang selalu membimbing penulis dalam menyelesaikan tulisan ini.
7. Teman-teman seperjuangan kelas Informatika E 2018 dan seluruh teman-teman angkatan 2018 yang telah banyak memberikan bantuan dan dukungan kepada penulis selama perkuliahan sampai pada tahap penyusunan ini.

Penulis menyadari bahwa penulisan ini masih jauh dari kesempurnaan. Oleh karena itu, kritik dan saran diharapkan penulis dalam penyempurnaan penulisan ini.

Majene, 01 Desember 2022

Penulis

SAMSANI  
D0218044

## DAFTAR ISI

KATA PENGANTAR .....	ii
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR .....	x
BAB I PENDAHULUAN.....	1
A. Latar Belakang.....	1
B. Rumusan Masalah.....	3
C. Batasan Masalah .....	3
D. Tujuan Penelitian.....	3
E. Manfaat Penelitian .....	3
BAB II TINJAUAN PUSTAKA.....	5
A. Analisis Sentimen.....	5
B. Text Mining .....	5
C. Laboratorium .....	6
D. <i>K-Nearest Neighbor</i> .....	7
E. Pembobotan TF-IDF .....	9
F. Evaluasi .....	10
G. Penelitian terkait.....	11
BAB III METODOLOGI PENELITIAN.....	15
A. Jenis Penelitian .....	15
B. Jadwal dan Lokasi Penelitian.....	15
C. Metode Pengumpulan Data.....	16
D. Teknik Analisis Data .....	16
E. Alur Penelitian .....	17
F. Alur Sistem.....	18
BAB IV HASIL DAN PEMBAHASAN .....	23
A. Implementasi .....	23

B. Pengumpulan Data .....	23
C. Teks <i>Preprocessing</i> .....	24
D. Pelabelan .....	26
E. Pembobotan TF-IDF .....	26
F. Pembagian Data .....	27
G. Klasifikasi Algoritma .....	28
G. <i>Confusion Matrix</i> .....	30
H. Hasil pengujian Akurasi.....	32
I. Hasil Sentimen .....	35
BAB V KESIMPULAN DAN SARAN.....	36
A. Kesimpulan .....	36
B. Saran.....	36
DAFTAR PUSTAKA .....	37



## DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i> .....	9
Tabel 2.2 Penelitian Terkait .....	11
Tabel 3.1 Jadwal dan Lokasi Penelitian .....	15
Tabel 4.1 <i>Preprocessing</i> .....	22
Tabel 4.2 Hasil Pelabelan .....	23
Tabel 4.3 Pembagian Data .....	25
Tabel 4.4 Perbandingan Akurasi .....	25
Tabel 4.5 Perbandingan Presisi .....	26
Tabel 4.6 Perbandingan Recall .....	27
Tabel 4.7 Perbandingan <i>F-1 Score</i> .....	28
Tabel 4.8 Kelas <i>Confusion Matrix</i> .....	30

## DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	17
Gambar 3.2 Alur Sistem.....	18
Gambar 4.1 Hasil TF-IDF.....	24
Gambar 4.2 <i>Confusion Matrix</i> .....	29
Gambar 4.3 Diagram Lingkar Analisis Sentimen.....	33

# **BAB I**

## **PENDAHULUAN**

### **A. Latar Belakang**

Universitas sebagai salah satu lembaga pendidikan memiliki peran penting dalam meningkatkan kualitas pendidikan dan ilmu pengetahuan. Universitas pada dasarnya memiliki tujuan untuk memberikan kesiapan kepada mahasiswa dalam berperan menghadapi lingkungan hidup yang selalu mengalami perubahan setiap tahunnya. Alasan yang mendukung dari pentingnya suatu lembaga pendidikan adalah karena banyak bagian dunia, lingkungan lembaga pendidikan, khususnya pendidikan tinggi yang beroperasi telah menjadi semakin kompetitif.

Pendidikan pada dasarnya tidak terlepas dari proses mengamati, membaca, meniru, mencoba sesuatu, mendengar dan mengikuti arahan tertentu. Hal ini menyebabkan sistem pendidikan yang diterapkan di lingkungan perguruan tinggi lebih mengedepankan dari kesadaran masing-masing mahasiswa dalam menyerap ilmu yang diberikan. Namun dari perguruan tinggi sendiri telah menyediakan berbagai fasilitas penunjang untuk memudahkan mahasiswa dalam memahami ilmu pengetahuan.

Fasilitas yang harus disediakan oleh perguruan tinggi diantaranya yaitu sarana dan prasarana. Sarana pendidikan mencakup semua peralatan dan perlengkapan yang secara langsung menunjang pendidikan sedangkan prasarana pendidikan tidak secara langsung. Apabila sebuah perguruan tinggi telah memiliki sarana dan prasarana yang memadai seperti ruang pembelajaran dan laboratorium, maka kualitas pendidikan

yang diterima oleh mahasiswa akan lebih optimal (Hidayat L, Mulyana M, & Effendy M (2018).

Laboratorium ICT (*Information, Communication, Technologi*) yang merupakan salah satu fasilitas universitas Sulawesi barat, menjadi salah satu objek penilaian terkait kepuasan pengguna laboratorium kepada mahasiswa. Hal ini karena laboratorium merupakan tempat mahasiswa berlatih secara langsung dalam menerapkan ilmu-ilmu serta memahami konsep-konsep dengan melakukan percobaan dan pengamatan. Untuk mencapai tujuan kepuasan yang berkualitas, Laboratorium dituntut untuk memenuhi kebutuhan penggunanya (Rikin Tiyas Permana, 2020).

Seiring dengan upaya peningkatan yang dilakukan, masih terdapat permasalahan yang ditemui di lapangan, diakui bahwa masih terdapat keluhan dari mahasiswa. Tentu hal ini tidak boleh dibiarkan begitu saja, diperlukan perhatian dan penanganan yang efektif agar kepuasan mahasiswa dalam penggunaan Laboratorium dapat diciptakan.

Tujuan dari penelitian ini adalah untuk mengetahui kepuasan mahasiswa dalam penggunaan laboratorium, kepuasan mahasiswa yang berisi komentar positif ataupun negatif. Hasil dari kepuasan mahasiswa tersebut juga dapat menjadi tolak ukur sebuah laboratorium dalam penanganan yang lebih efektif kedepannya.

## **B. Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan maka penulis merumuskan suatu masalah yaitu kepuasan mahasiswa dalam penggunaan laboratorium untuk mengetahui analisis sentiment positif dan negatif dengan menggunakan metode KNN (*K-Nearest Neighbor*)?

## **C. Batasan Masalah**

Dalam merancang dan membangun serta menulis penelitian ini diperlukan batasan-batasan masalah agar lebih mengarahkan tujuan.

Adapun batasan masalah dalam penelitian ini yaitu:

1. Data yang digunakan dalam penelitian ini diambil dari responden kuesioner secara online terhadap mahasiswa pengguna Laboratorium ICT.
2. Klasifikasi data menggunakan algoritma KNN (*K-Nearest Neighbor*).
3. Output dari penelitian ini berupa hasil klasifikasi positif dan negatif.

## **D. Tujuan Penelitian**

Tujuan dari penelitian ini adalah menerapkan algoritma KNN (*K-Nearest Neighbor*) untuk mengetahui analisis sentimen positif dan negatif kepuasan mahasiswa dalam penggunaan laboratorium.

## **E. Manfaat Penelitian**

Adapun manfaat dari penelitian ini yang dibagi menjadi dua yaitu manfaat bagi peneliti dan manfaat bagi pihak lain :

1. Manfaat bagi peneliti

- a. Untuk mengetahui kinerja dari algoritma KNN (*K-Nearest Neighbor*) dalam mengklasifikasikan tingkat kepuasan mahasiswa dalam penggunaan laboratorium ICT.
  - b. Untuk mengetahui tingkat sentiment positif dan negatif mahasiswa dalam penggunaan laboratorium ICT.
  - c. Hasil penelitian ini merupakan salah satu syarat dalam menyelesaikan Pendidikan Strata Satu (S1).
2. Manfaat bagi pihak lain
- a. Untuk menjadi acuan bagi peneliti yang ingin melakukan penelitian sejenis.
  - b. Dasar peningkatan pelayanan dalam pengguna laboratorium ICT

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **A. Analisis Sentimen**

Analisis sentimen adalah bidang yang menarik dikembangkan dalam dunia digital sebab masyarakat saat ini lebih bebas bersuara dalam dunia maya untuk mengutarakan pendapat dan hasil pemikiran mereka dengan bahasa teks yang dimana para pembaca terkadang memiliki kesalah pahaman dalam menerjemahkan sentimen yang terdapat didalamnya (Balya, 2019). Sementara menurut (Sari & Wibowo, 2019) Analisis sentimen merupakan sebuah metode digunakan untuk mengekstrak data opini, memahami serta mengolah tekstuan data secara otomatis untuk mengetahui sentimen yang terdapat dalam sebuah opini.

Analisis Sentimen adalah sebuah teknik atau cara yang digunakan untuk mengidentifikasi bagaimana sebuah sentimen diekspresikan menggunakan teks dan bagaimana sentimen tersebut bisa dikategorikan sebagai sentimen negatif maupun sentimen positif. Singkatnya Analisis sentimen merupakan sebuah metode yang digunakan untuk mengekstrak data opini, memahami serta mengolah tekstual data secara otomatis untuk melihat sentimen yang terkandung dalam sebuah opini.

#### **B. Text Mining**

*Text mining* adalah suatu proses untuk mengekstrak informasi menarik, pengetahuan atau pola dari dokumen tidak terstruktur yang berasal dari sumber yang berbeda. *Text mining* dapat dikatakan gabungan dari *text analysis* dan data

mining karena *text mining* bertujuan untuk mendapatkan representasi dari suatu dokumen atau teks dengan teknik aspek-aspek yang ada di data mining. Sebelum melakukan *text mining*, seluruh dokumen akan dilakukan praproses teks terlebih dahulu agar dokumen yang didapatkan lebih mudah diklasifikasi (Ernawati & Ika Nurlaili, 2020).

### **C. Laboratorium**

Proses meningkatkan sumber daya manusia yang memiliki intelektual dan unggul, mampu bersaing dalam bidangnya, tidak terlepas dari peranan Perguruan Tinggi sebagai institusi pendidikan. Perguruan Tinggi bertugas menghasilkan lulusan-lulusan yang berkualitas. Proses belajar dan mengajar saja tidak cukup tetapi perlu didukung dengan fasilitas-fasilitas penunjang lainnya seperti Laboratorium Komputer.

Laboratorium komputer merupakan salah satu sarana pelayanan yang dipergunakan sebagai penunjang proses pembelajaran. Peningkatan laboratorium komputer perlu terus diupayakan guna memenuhi kebutuhan dan kepuasan mahasiswa sebagai pengguna serta menunjang keberhasilan mahasiswa dalam proses pembelajaran (Syukhri 2018).

Di beberapa laboratorium komputer digunakan untuk simulasi atau analisis data yang dikumpulkan di tempat lain, seperti yang biasa digunakan oleh ilmuwan komputer. Ilmuwan di bidang lain akan menggunakan jenis laboratorium lainnya. Insinyur menggunakan laboratorium untuk merancang, membangun, dan juga menguji perangkat teknologi.



#### **D. *K-Nearest Neighbor***

*K-Nearest Neighbor* merupakan sebuah algoritma yang sering digunakan untuk klasifikasi teks dan data (Ernawati S & Wati R (2018)). Tujuan dari algoritma ini adalah mengklasifikasikan obyek berdasarkan atribut dan training sample. Classifier tidak menggunakan apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik query, akan ditemukan sejumlah k obyek atau (titik training) yang paling dekat dengan titik query. Klasifikasi menggunakan voting terbanyak diantara klasifikasi dari k obyek. Algoritma *K-Nearest Neighbor* (K-NN) menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari query instance yang baru (Ernawati S & Wati R (2018)). Tahapan klasifikasi KNN dapat dilihat pada Gambar 3.3 berikut:

K-NN akan dihitung memakai jarak *Euclidean* dengan persamaan sebagai berikut:

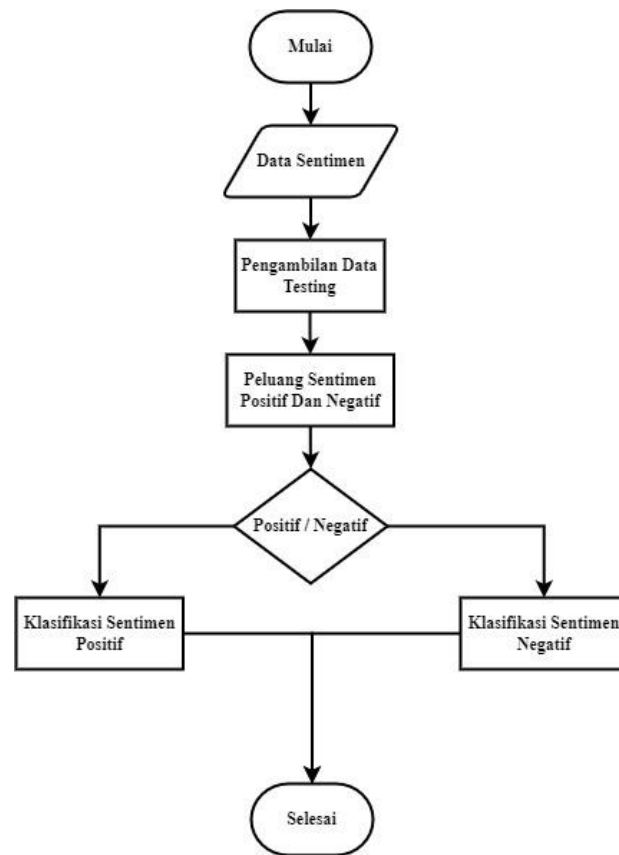
$$\begin{aligned} d(x,y) &= \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2 \dots + (x^n - y^n)^2} & 2 \\ &= \sqrt{\sum_{ki=1} (xi - yi)^2} \end{aligned}$$

Keterangan :

$d(x,y)$  = Sebagai jarak antar variabel  $x_1$  dan  $x_2$

$x$  = Sebagai variabel

$k$  = Sebagai jumlah variabel



Gambar 2.1 Tahapan klasifikasi k nearest neighbor

Dimulai dengan menginput data kemudian pengambilan data testing, kemudian mencari peluang kemunculan sentimen positif dan negatif. Setelah diketahui peluang sentimen positif dimana sentiment positif diberi label 1.0. Sentiment negatif diberi label 0.0 Hasil dari pelabelan adalah data yang disimpan dalam bentuk csv.

### E. Pembobotan TF-IDF

Pembobotan TF-IDF (*Term Frequency-Inverse Document Frequency*) merupakan metode yang melakukan perhitungan pada tiap bobot kata yang terdapat didalam data dokumen. Data dokumen akan diubah menjadi vektor sebanyak kata (term) yang akan digunakan untuk klasifikasi (Furqan M., Sriani S. & Sari S M. (2022). TF-IDF digunakan agar bisa melakukan analisis dengan memakai algoritma *K-Nearest Neighbor* (K-NN).

Adapun perhitungan pembobotan TF-IDF ialah sebagai berikut :

$$idf = \log \frac{N}{df}$$

$$w(k,d) = tf(k,d) * idf$$

$$w(k,d) = tf(k,d) * \log \frac{N}{df}$$

Keterangan :

- W(k,d) = Bobot kata (*term*) yang tersedia dalam dokumen
- tf (k,d) = Jumlah frekuensi munculnya kata didalam dokumen
- N = Jumlah seluruh dokumen yang tersedia didalam database
- df = Jumlah dokumen yang mengandung *term*.

## F. Evaluasi

Teknik yang digunakan untuk mengevaluasi hasil yaitu *confusion matrix* banyak digunakan dalam penelitian untuk mengevaluasi hasil dan mengukur kinerja suatu metode klasifikasi dan juga digunakan untuk menghitung serta bisa menarik kesimpulan dari hasil proses penelitian yang sudah dilakukan. Dalam *confusion matrix* akan dihitung *accuracy*, *precision*, *f1-score* dan *recall* yang dirumuskan pada persamaan.

Table 2.1 *Confusion matrix*

		Nilai Sebenarnya	
		True	False
Nilai Prediksi	<i>True</i>	TP ( <i>True Positive</i> )	FP ( <i>True Positive</i> )
	<i>False</i>	FN ( <i>False Negatif</i> )	TN ( <i>False Negatif</i> )

Keterangan :

FN = jika data aktualnya positif dan diidentifikasi negatif oleh sistem

TP = jika data aktualnya positif dan diidentifikasi positif oleh sistem

TN = jika data aktualnya negatif dan diidentifikasi negatif oleh sistem

FP = jika data aktualnya negatif dan diidentifikasi positif oleh sistem

Nilai *accuracy* merupakan rasio prediksi Benar (positif dan negatif) dengan keseluruhan data. Nilai *accuracy confusion matrix* tersebut :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

Nilai *precision* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Nilai *precision confusion matrix* tersebut :

$$Precision = \frac{TP}{FP+TP} \times 100\%$$

Nilai *recall* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Nilai *recall confusion matrix* tersebut :

$$Recall = \frac{TP}{FN+TP} \times 100\%$$

### G. Penelitian terkait

Beberapa penelitian yang berkaitan dengan analisis sentimen positif dan negatif kepuasan mahasiswa dalam penggunaan laboratorium ICT.

Tabel 2.2 Penelitian terkait

No	Judul	Nama dan tahun penelitian	Hasil Penelitian	Keterkaitan
1	Analisis kepuasan pengguna laboratorium:	(Rikin Tiyas Permana ,2020)	Kuesioner yang dihasilkan telah dapat mempresentasikan	Menggunakan objek yang sama dengan studi kasus

	studi kasus di laboratorium jurusan teknik industri universitas islam indonesia		<p>karakteristik dari keempat laboratorium dan menjadikan standarisasi dalam prosedur penilaian laboratorium sehingga hasil penilaian dan pengolahan dapat digunakan sebagai perbandingan tingkat kepuasan antar laboratorium dengan tujuan agar semua laboratorium di jurusan Teknik Industri memiliki kualitas pelayanan yang sama dan merata dengan memperhatikan laboratorium yang memiliki kualitas pelayanan paling baik sebagai acuan.</p>	yang berbeda.
2	Analisa kualitas pelayanan terhadap kepuasan mahasiswa menggunakan laboratorium komputer universitas simalungun	(Rosita Manawari Girsang & Rosita Manawari Girsang,2019)	<p>Hasil penelitian menunjukkan bahwa <i>variable</i> kehandalan (<i>reliability</i>) berpengaruh positif dan signifikan bagi kepuasan mahasiswa, berbeda halnya dengan keinginan pegawai untuk membantu mahasiswa dan memberikan pelayanan dengan baik (<i>Responsiveness</i>), serta meliputi kemudahan dalam</p>	Menggunakan objek yang hampir sama dengan sumber data yang beda

			<p>hubungan, komunikasi yang baik, perhatian yang baik (<i>Emphaty</i>) memiliki hubungan terbalik dengan kepuasan mahasiswa, hal ini disebabkan karena mahasiswa lebih mengutamakan fasilitas dan jaminan pengetahuan dalam pengolahan data serta sertifikat, serta instruktur yang handal.</p>	
3	<p>Analisis sentimen terhadap pelayanan PT PLN di Jakarta pada <i>Twitter</i> dengan algoritma <i>K-Nearest Neighbor</i> (K-NN)</p>	<p>(M. Suyudi Alrajak, Iin Ernawati &amp; Ika Nurlaili, 2020)</p>	<p>Hasil analisis sentimen mendapatkan 426 laporan (sentimen netral), 71 keluhan (sentimen negatif), dan 3 dukungan (sentimen positif). Dari 500 tweets, hanya terdapat 71 keluhan atau sekitar 14,2% dari total seluruh opini. Maka, dapat disimpulkan pelayanan PLN di Jakarta pada bulan Februari 2020 sudah cukup baik.</p>	<p>Penelitian ini menggunakan analisis sentimen positif, negatif dan netral, sedangkan penelitian yang akan dilakukan hanya menggunakan positif dan negatif.</p>
4	<p>Analisis Sentimen pada <i>Twitter</i> terhadap Pelayanan Pemerintah Kota Makassar</p>	<p>(Rosdiana, Eddy Tungadi, Zawiyah Saharuna, MuhammadNur Yasir Utomo, 2019)</p>	<p>Hasil dari analisis sentimen yang di simpan pada Elasticsearch berhasil ditampilkan pada aplikasi web</p>	<p>Pada penelitian ini tidak menggunakan pembobotan tf-idf, sedangkan pada penelitian yang akan</p>

			dengan bantuan REST API. Aplikasi web juga dilengkapi dengan fitur pencarian yang dapat mencari data hasil analisis sentimen berdasarkan keyword dengan rentang waktu tertentu.	dilakukan menggunakan pembobotan tf-idf.
5	Analisis Sentimen Kepuasan Pelanggan Pada Penyedia Layanan Pengantaran Makanan	(Andrew,Erick Dazki, Richardus Eko Indrajit3, 2022)	Hasil yang didapatkan ketiga model dapat dikatakan cukup baik dengan nilai akurasi dan F yang berada di atas 78%. Dengan menggunakan TF-IDF untuk mengekstraksi fitur dapat dikatakan berhasil meningkatkan hasil evaluasi dari ketiga model yang digunakan.	Keterkaitan penelitian ini adalah untuk melihat keterkaitan performa tiga algoritma yaitu <i>Stochastic Gradient Descent</i> , <i>Support Vector Machine</i> , dan <i>K-Nearest Neighbor</i> .



## **BAB III**

### **METODOLOGI PENELITIAN**

#### **A. Jenis Penelitian**

Dari judul yang telah ada maka jenis penelitian yang di gunakan yaitu kuantitatif, Jenis penelitian ini adalah penelitian kuantitatif yaitu penelitian ilmiah yang sistematis terhadap bagian-bagian fenomena serta hubungan-hubungan yang digunakan untuk menguji hipotesis yang telah ditetapkan melalui perhitungan ilmiah yang berasal dari sampel konsumen yang diminta menjawab atas sejumlah pertanyaan tentang survey untuk melakukan frekuensi dan presentase tanggapan yang mereka berikan.

Jenis penelitian ini merupakan penelitian yang banyak menggunakan angka, mulai dari pengumpulan data, penafsiran terhadap data tersebut, serta penampilan hasilnya. Dalam hal ini diperlukan data responden kuesioner kepuasan mahasiswa dalam penggunaan laboratorium ICT. Adapun data isian pada kuesioner berupa saran, pendapat atau keluhan mahasiswa dalam penggunaan laboratorium.

#### **B. Jadwal dan Lokasi Penelitian**

Penelitian ini dilaksanakan di Laboratorium ICT Universitas Sulawesi Barat. Adapaun rencana waktu pelaksanaan akan dilaksanakan dalam waktu terhitung mulai dari bulan Januari hingga Maret 2023.

Tabel 3.1 Jadwal dan Lokasi Penelitian

No	Kegiatan Penelitian	Bulan														
		Januari					Februari					Maret				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	Identifikasi masalah															
2	Pengambilan data															
4	Evaluasi															
5	Pengujian															

### C. Metode Pengumpulan Data

Pengumpulan data yang dilakukan pada penelitian ini yaitu dengan dilakukannya pengumpulan data responden kepuasan mahasiswa dengan menggunakan kuesioner, Adapun data isian pada kuesioner berupa Nama, Nomor induk mahasiswa (NIM), Fakultas, prodi, tahun angkatan, saran, pendapat atau keluhan mahasiswa dalam penggunaan laboratorium. Setelah mendapatkan data yang diinginkan maka dilanjutkan ketahap labeling. Labeling yaitu proses memberikan label data yang sudah didapatkan sebelumnya dengan memberikan label apakah data tersebut bernilai positif atau negatif.

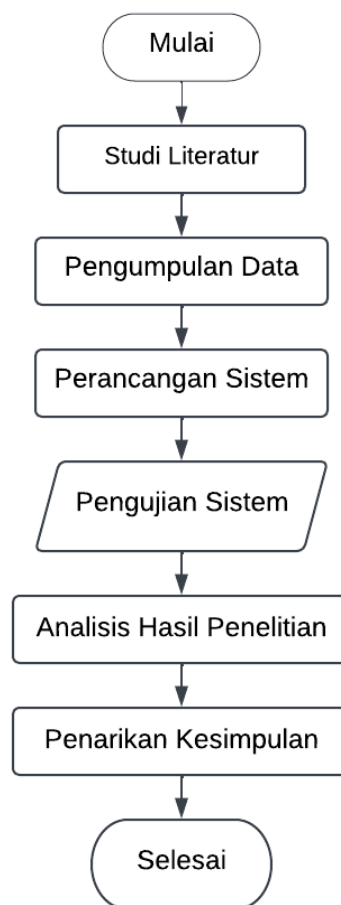
### D. Teknik Analisis Data

Pada penelitian ini terdapat beberapa metode analisis data yang digunakan. Berikut merupakan metode analisis sentimen digunakan untuk melakukan pelabelan data ke dalam kelas sentimen positif dan negatif. Metode *K-Nearest*

*Neighbors* digunakan sebagai suatu metode yang menunjang ketika melakukan proses klasifikasi ulasan, baik ulasan positif maupun ulasan negatif.

### E. Alur Penelitian

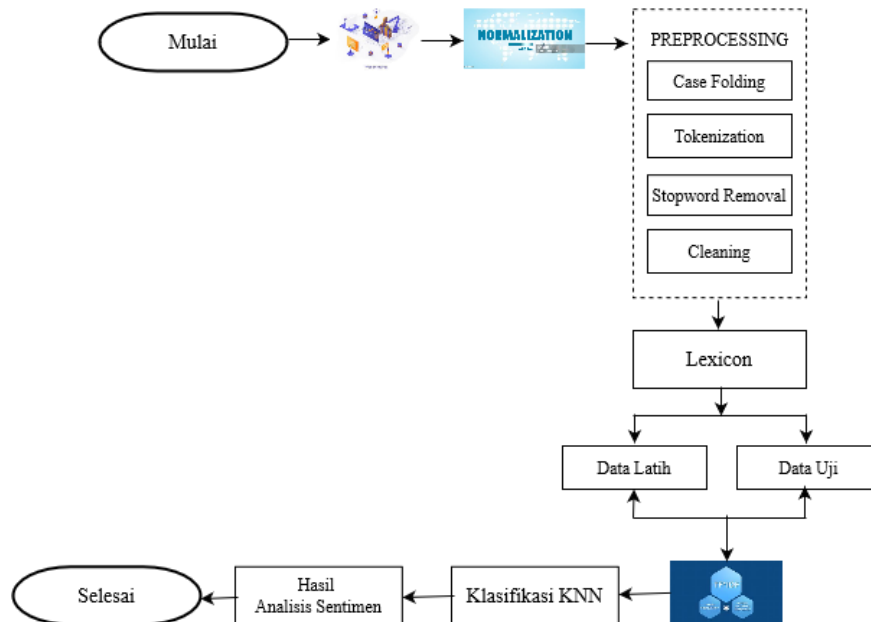
Berikut merupakan alur penelitian yang dapat dilihat pada gambar 3.1:



Gambar 3.1 Alur Penelitian

## F. Alur Sistem

Berikut merupakan gambaran alur sistem yang dapat dilihat pada gambar 3.2



Gambar 3.2 Alur Sistem

### 1. Pengumpulan data

Proses pengumpulan data ini dilakukan dengan cara mengumpulkan data informasi baik, teks atau angka menggunakan scraping. scraping adalah sebuah cara pengambilan suatu data atau informasi tertentu dengan jumlah besar. Dibandingkan dengan melakukan survei secara manual. Scraping bisa mengambil data secara praktis dan waktu yang singkat. Data diambil dari responden kuesioner.

### 2. Normalisasi

Menormalisasikan kata yaitu sebuah proses untuk mengubah dan memperbaiki kata yang disingkat ke dalam kata yang memiliki arti sama yang berdasarkan KBBI agar menjadikannya informasi yang dapat diproses dan memudahkan dalam proses selanjutnya.

<b>Komentar</b>	<b>Case Folding</b>
sdah cukup baik tpi mungkin bisa melakukan pengecekan komputer untuk bisa memastikan tdk adanya kendala apabila sdah dgunakn oleh mahasiswa	sudah cukup baik tapi mungkin bisa melakukan pengecekan komputer untuk bisa memastikan tidak adanya kendala apabila sudah digunakan oleh mahasiswa

### 3. *Preprocessing*

Merupakan tahapan pemrosesan data agar menjadi data yang siap untuk dianalisis. Setelah data terstruktur maka dapat diolah lebih lanjut. Beberapa proses yang dilakukan pada tahap ini adalah sebagai berikut :

#### a. *Case folding*

Proses perubahan semua huruf menjadi huruf kecil. Hanya huruf a sampai z yang diproses. Karakter selain huruf akan dibiarkan.

Tabel 3.2 Case folding

<b>Komentar</b>	<b>Case Folding</b>
Sudah Cukup Baik Tapi Mungkin Bisa Melakukan Pengecekan Komputer Untuk Bisa Memastikan Tidak Adanya Kendala Apabila Sudah Digunakan Oleh Mahasiswa	sudah cukup baik tapi mungkin bisa melakukan pengecekan komputer untuk bisa memastikan tidak adanya kendala apabila sudah digunakan oleh mahasiswa

b. *Tokenization*

*Tokenizing* adalah proses untuk memecah kalimat yang terdapat dalam dokumen text dengan cara memisahkan kata berdasarkan kata per-spasi. Potongan kata atau kata yang sudah terpisahkan disebut sebagai token. Pemisahan kata ini akan dipisahkan dengan tanda koma (,) untuk setiap katanya.

Tabel 3.3 Tokenization

<b>Komentar</b>	<b><i>Tokenization</i></b>
sudah cukup baik tapi mungkin bisa melakukan pengecekan komputer untuk bisa memastikan tidak adanya kendala apabila sudah digunakan oleh mahasiswa	['sudah'], ['cukup'], ['baik'], ['tapi'], ['mungkin'], ['bisa'], ['melakukan'], ['pengecekan'], ['komputer'], ['untuk'], ['bisa'], ['memastikan'], ['tidak'], ['adanya'], ['kendala'], ['apabila'], ['sudah'], ['digunakan'], ['oleh'], ['mahasiswa'].

c. *Stopword removal*

*Filtering* bertujuan untuk menghapuskan kata yang tidak mengandung makna ataupun kurang penting pada data. kata-kata yang biasa dipakai tetapi tidak jelas dan dapat dibuang seperti "dari", "yang", "untuk", "dan", "di" dan sebagainya.

Tabel 3.4 Stopword removal

<b>Komentar</b>	<b><i>Stopword Removal</i></b>
sudah cukup baik tapi mungkin bisa melakukan pengecekan komputer <b>untuk</b> bisa memastikan <b>tidak</b>	sudah cukup baik tapi mungkin bisa melakukan pengecekan komputer untuk bisa memastikan tidak adanya

adanya kendala apabila sudah digunakan oleh mahasiswa	kendala apabila sudah digunakan oleh mahasiswa
---	--

#### d. *Cleaning*

*Cleaning* yaitu untuk membersihkan atau menghilangkan karakter yang tidak diperlukan pada data seperti tanda baca, numeric, url, username.

Tabel 3.5 Tokenization

<b>Komentar</b>	<b><i>Cleaning</i></b>
sudah cukup baik, tapi mungkin bisa melakukan pengecekan komputer, untuk bisa memastikan tidak adanya kendala, apabila sudah digunakan oleh mahasiswa,	sudah cukup baik tapi mungkin bisa melakukan pengecekan komputer untuk bisa memastikan tidak adanya kendala apabila sudah digunakan oleh mahasiswa

#### 4. *Lexicon*

Proses ini dilakukan dengan *Lexicon Based*, yaitu pendekatan yang menggunakan suatu kamus sentimen berisi kata positif dan kata negatif yang dibandingkan dan dicocokkan dengan kata pada kalimat untuk diketahui tingkat polaritasnya dan skor kata opini negatif pada tiap kalimat. Data yang akan digunakan dibagi menjadi dua bagian yaitu data latih dan data uji.

#### 5. Pembobotan TF-IDF

*Term Frequency -Inverse Document Frequency* atau TF-IDF adalah suatu metode algoritma yang berguna untuk menghitung bobot setiap kata

yang umum digunakan. Pada proses pembobotan untuk mengevaluasi seberapa penting sebuah kata dalam dokumen

#### 6. Klasifikasi *K-Nearest Neighbor*

*K-Nearest Neighbor* merupakan sebuah algoritma yang sering digunakan untuk klasifikasi teks dan data (Samuel, Delima, & Rachmat, 2014). Tujuan dari algoritma ini adalah mengklasifikasikan obyek berdasarkan atribut dan training sample.

#### 7. Hasil Analisis Sentimen

Hasil proses klasifikasi sentimen positif dan negatif dengan menggunakan algoritma *K-Nearest Neighbor* diperlukan *confusion matrix* dengan hasil *accuracy*, *precision* dan *recall*.



## **BAB IV**

### **HASIL DAN PEMBAHASAN**

Dalam bab ini akan menjelaskan hasil dan pembahasan tentang penelitian yang telah dilakukan berdasarkan tahapan-tahapan yang dilakukan dalam penelitian ini.

#### **A. Implementasi**

Dalam penelitian ini dilakukan analisis yang dimulai dari tahapan pengumpulan data yang dikumpulkan dari kuesioner, melakukan, proses *preprocessing*, pelabelan dengan *lexicon*, menghitung bobot perkataanya, selanjutnya pengklasifikasian sehingga mendapatkan output dalam penelitian ini.

Penelitian analisis sentiment positif dan negatif *twitter* diimplementasikan menggunakan *software jupyter notebook* dengan spesifikasi *processor* AMD Athlon Silver 3050U *with* Radeon Graphics (2 CPUs). ~2.3GHz dan *memory* 8192 MB RAM dengan *Operating System* Windows 11.

#### **B. Pengumpulan Data**

Pada penelitian ini, proses pengambilan data dilakukan melalui penyebaran kuesioner berupa link google form. Hasil kuesioner kemudian akan disimpan ke dalam file CSV untuk kemudahan pengolahan dan analisis data.

Jumlah data yang diambil dalam penelitian ini sebanyak 1055 data untuk sentimen. Pengambilan jumlah data yang relatif besar ini bertujuan untuk memperoleh sampel data yang cukup representatif untuk menganalisis sentimen publik terkait Laboratorium ICT di Universitas Sulawesi Barat. Selain itu, jumlah

data yang cukup besar juga dapat meningkatkan keakuratan hasil analisis dan membuat kesimpulan yang lebih solid.

### C. Teks *Preprocessing*

Untuk melanjutkan ke proses Langkah awal yang akan dilakukan dengan import library. Berikut adalah penjelasan singkat tentang library yang digunakan:

1. *numpy* adalah library Python yang digunakan untuk melakukan operasi matematika dan statistika pada data.
2. *matplotlib.pyplot* adalah library Python yang digunakan untuk membuat grafik dan visualisasi data.
3. *sklearn.feature\_extraction.text* adalah library Python yang digunakan untuk melakukan ekstraksi fitur teks pada data.
4. *string* adalah library Python yang berisi kumpulan fungsi untuk memanipulasi string, seperti penghapusan karakter atau penggantian karakter tertentu pada string.
5. *re* adalah library Python yang digunakan untuk melakukan operasi pemrosesan string dengan pola tertentu (*regular expression*).
6. *StopWordRemoverFactory* dan *StemmerFactory* adalah library Python dari Sastrawi yang digunakan untuk melakukan penghapusan kata stopwords dan stemming pada teks Bahasa Indonesia.
7. *nlk* adalah library Python yang digunakan untuk memproses teks Bahasa Inggris, termasuk tokenisasi dan penghapusan kata *stopword*.
8. *csv* adalah library Python yang digunakan untuk membaca dan menulis data dalam format CSV.

9. *nlk.download('punkt')* dan *nlk.download('stopwords')* adalah kode untuk mendownload *resource* untuk pemrosesan teks Bahasa Inggris menggunakan *library nltk*.

Tujuan dilakukan *preprocessing* dalam penelitian ini untuk menyeleksi data teks agar lebih tersruktur dan menghilangkan kata yang tidak diperlukan.

Adapun tahapan dalam *preprocessing* ini yaitu :

Tabel 4.1 *Preprocessing*

<b>Sebelum <i>Preprocessing</i></b>	<b>Setelah <i>Preprocessing</i></b>
Lab ICT sudah lumayan bagus cuma spesifikasi komputernya perlu di tambah	LAB ICT lumayan bagus spesifikasi komputernya
Bagus sesuai fungsinya banyak membantu praktikum mahasiswa	Bagus fungsinya banyak mbantu praktikum mahasiswa
Fasilitasnya sudah lumayan memadai, nyaman ketika kita belajar didalam, stafnya juga sangat ramah dan memperhatikan keadaan Laboratorium ketika ada mahasiswa yang sedang menggunakan ruangan tersebut.	Fasilitasnya lumayan memadai nyaman belajar didalam stafnya ramah memperhatikan keadan laboratorium mahasiswa menggunakan ruangan
Lab ICT sekarang sudah bagus, hanya masih kurang pada tingkat pemeliharaan sarananya saja	LAB ICT bagus pemeliharaan sarananya

### D. Pelabelan

Data yang dihasilkan dalam *preprocessing* akan dilabeli secara menggunakan proses *lexicon*, pelabelan teks diubah kedalam bentuk numerik agar sesuai dengan kebutuhan sistem yang dimana sentiment positif diberi label 1.0. Sentiment negatif diberi label 0.0 Hasil dari pelabelan adalah data yang disimpan dalam bentuk csv.

Tabel 4. 2 Contoh Hasil Pelabelan

<b>Positif</b>	Fasilitasnya memadai nyaman belajar di dalam, stafnya ramah memperhatikan keadaan laboratorium mahasiswa menggunakan ruangan
<b>Negatif</b>	Lumayan nyaman saat di ruangan lab berdesakan terbilang sempit

### E. Pembobotan TF-IDF

Pembobotan TF-IDF (*term Frequency-Inverse Document Frequency*) untuk melihat berapa nilai frekuensi kemunculan terhadap suatu kata. Hasil yang didapatkan dalam pembobotan TF-IDF merupakan hasil perkalian dari TF dikalikan dengan IDF.

Gambar 4. 1 Hasil TF-IDF

```
ranking_br=ranking.sort_values('TF-IDF', ascending=False)
print(ranking_br)
```

	term	TF-IDF
130	memuaskan	49.001825
10	bagus	43.638648
216	tik	43.260305
110	lumayan	36.974675
104	laboratorium	36.464463
..	...	...
201	stafnya	0.352167
32	didalam	0.352167
84	keluhan	0.000000
66	inovasi	0.000000
57	frustasi	0.000000

[231 rows x 2 columns]

Keterangan :

*Term* : kalimat

TF-IDF : hasil pembobotan

Hasil pembobotan TF-IDF merupakan data numerik yang dimana nantinya data tersebut digunakan dalam proses perhitungan klasifikasi *KNN*.

## F. Pembagian Data

Pada tahap proses klasifikasi KNN pada penelitian ini, dibangun sebuah *machine learning* yang menggunakan data sebanyak 1055 dengan perbandingan data seperti tabel berikut:

Tabel 4. 3 Pembagian data

Berikut pembagian Data latih dan data uji sebagai berikut

Perbandingan	Data Latih	Data Uji
70:30	739	316
80:20	844	211
90:10	950	105

### G. Klasifikasi Algoritma

Pada penelitian ini yang digunakan adalah algoritma K-Nearest Neighbors dengan metode perhitungan jarak menggunakan *Euclidean Distance*, dengan menguji beberapa nilai K dari K=1, K=3, dan K=5 untuk melihat akurasi tertinggi yang dihasilkan dari pengujian beberapa nilai K tersebut.

Tabel 4. 4 Perbandingan Akurasi

Nilai K	Perbandingan %		
	70:30	80:20	90:10
K1	94.21	98.28	93.75
K3	90.53	92.13	89.06
K5	89.47	92.13	89.06

Tabel 4.6 menunjukkan perbandingan akurasi antara tiga nilai K (K1, K3, dan K5) dengan tiga perbandingan data uji dan data latih (70:30, 80:20, dan 90:10). Dari tabel tersebut, dapat dilihat bahwa nilai K yang memberikan akurasi tertinggi adalah **K1** dengan **98.28%** pada perbandingan data 80:20. Nilai K yang memberikan akurasi terendah adalah **K5** dengan **89.47%** pada perbandingan data 70:30. Secara umum, dapat disimpulkan bahwa nilai K yang lebih kecil cenderung memberikan akurasi yang lebih tinggi daripada nilai K yang lebih besar. Hal ini mungkin disebabkan oleh fakta bahwa nilai K yang lebih kecil lebih sensitif

terhadap variasi data dan lebih mampu mengenali pola-pola yang spesifik. Namun, nilai **K** yang terlalu kecil juga berisiko menyebabkan overfitting, yaitu ketika model terlalu menyesuaikan diri dengan data latih dan kurang mampu memprediksi data uji yang baru. Oleh karena itu, pemilihan nilai **K** yang optimal harus mempertimbangkan keseimbangan antara akurasi dan generalisasi model.

Tabel 4.5 Perbandingan Presisi

Nilai <b>K</b>	Perbandingan %		
	70:30	80:20	90:10
K1	91	92	94
K3	86	87	86
K5	89	87	86

Tabel 4.5 menunjukkan perbandingan presisi antara tiga nilai **K** (**K1**, **K3**, dan **K5**) dengan tiga perbandingan data uji dan data latih (70:30, 80:20, dan 90:10). Presisi adalah ukuran yang menunjukkan seberapa akurat model dalam mengklasifikasikan data positif. Dari tabel tersebut, dapat dilihat bahwa nilai **K** yang memberikan presisi tertinggi adalah **K1** dengan **94%** pada perbandingan data 90:10. Nilai **K** yang memberikan presisi terendah adalah **K3** dengan **86%** pada perbandingan data 70:30 dan 90:10. Secara umum, dapat disimpulkan bahwa nilai **K** yang lebih kecil cenderung memberikan presisi yang lebih tinggi daripada nilai **K** yang lebih besar. Hal ini sesuai dengan hasil analisis akurasi pada tabel sebelumnya. Nilai **K** yang lebih kecil lebih mampu mengidentifikasi data positif dengan tepat, tetapi juga berpotensi menimbulkan kesalahan klasifikasi pada data negatif.

Tabel 4.6 Perbandingan Recall

Nilai K	Perbandingan %		
	70:30	80:20	90:10
K1	91	92	90
K3	84	87	85
K5	77	87	85

Tabel 4.6 menunjukkan perbandingan recall antara tiga nilai K (K1, K3, dan K5) dengan tiga perbandingan data uji dan data latih (70:30, 80:20, dan 90:10). Recall adalah ukuran yang menunjukkan seberapa lengkap model dalam mengklasifikasikan data positif. Dari tabel tersebut, dapat dilihat bahwa nilai K yang memberikan recall tertinggi adalah **K1** dengan **92%** pada perbandingan data 80:20. Nilai K yang memberikan recall terendah adalah **K5** dengan **77%** pada perbandingan data 70:30. Secara umum, dapat disimpulkan bahwa nilai K yang lebih kecil cenderung memberikan recall yang lebih tinggi daripada nilai K yang lebih besar. Hal ini sesuai dengan hasil analisis akurasi dan presisi pada tabel sebelumnya. Nilai K yang lebih kecil lebih mampu mengklasifikasikan data positif secara lengkap, tetapi juga berpotensi mengabaikan data negatif.

Tabel 4.7 Perbandingan F1-Score

Nilai K	Perbandingan %		
	70:30	80:20	90:10
K1	91	92	92
K3	85	87	86
K5	81	87	86

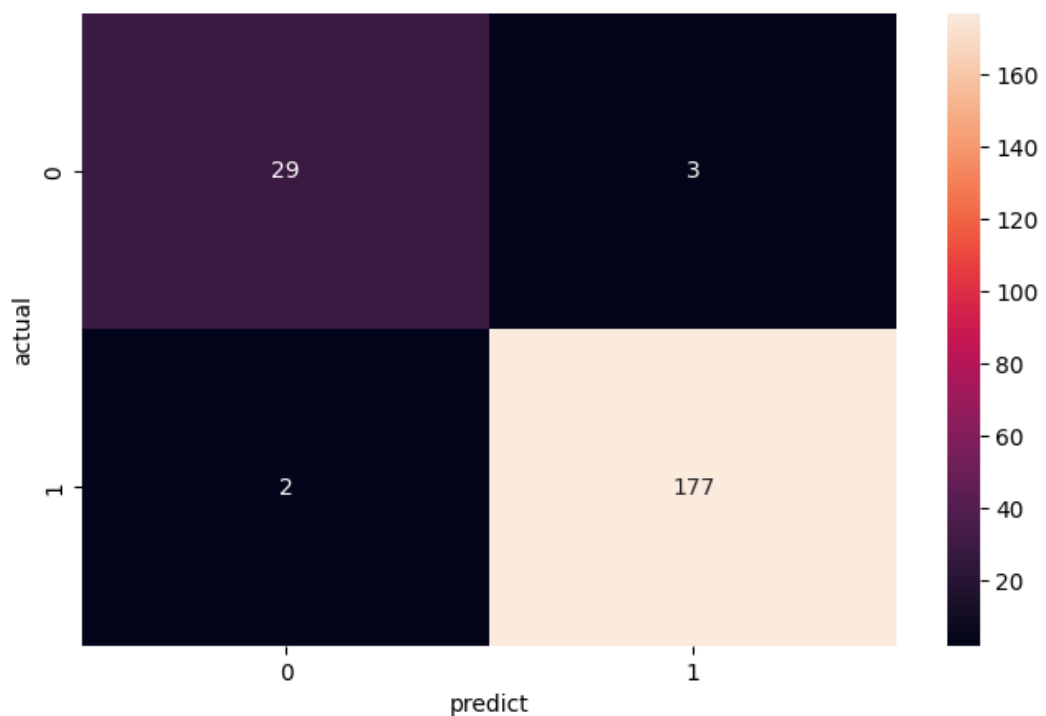
Tabel di atas merupakan hasil dari metode K-Nearest Neighbors (KNN) yang telah dijalankan dengan tiga set perbandingan nilai K yang berbeda: 70:30, 80:20, dan 90:10. Setiap baris dalam tabel mewakili hasil klasifikasi berdasarkan metode KNN untuk beberapa nilai K yang berbeda pada beberapa kelompok data



yang berbeda (dalam hal ini, K1, K3, dan K5). Berdasarkan hasil pengujian tersebut di dapat bahwa nilai perbandingan terbaik berada di 80:20 dengan nilai tetangga terdekat K1.

### G. *Confusion Matrix*

*Confusion Matrix* adalah alat yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan prediksi model dengan label sebenarnya. Dalam kasus ini, *Confusion Matrix* memiliki bentuk sebagai berikut:



Gambar 4.2 *Confusion Matrix*

Dalam *Confusion Matrix* ini, terdapat 2 kelas: kelas positif dan kelas negatif. Di sini, kita bisa mengidentifikasi matriks sebagai berikut:

1. *True Positive* (TP): 177
2. *True Negative* (TN): 2
3. *False Positive* (FP): 3
4. *False Negative* (FN): 29

## H. Hasil pengujian Akurasi

*Confusion matrix* adalah sebuah tabel yang menunjukkan kinerja sebuah model klasifikasi pada data uji yang memiliki label sebenarnya. *Confusion matrix* dapat digunakan untuk menghitung berbagai metrik evaluasi, seperti akurasi, presisi, *recall*, dan *F1-score*.

Untuk menghitung *confusion matrix*, Anda perlu melakukan langkah-langkah berikut:

1. Memerlukan dataset uji atau dataset validasi dengan nilai hasil yang diharapkan.
2. Buat prediksi untuk setiap baris di dataset uji.
  - a. Dari hasil yang diharapkan dan prediksi, hitung:
  - b. Jumlah prediksi yang benar untuk setiap kelas.

Jumlah prediksi yang salah untuk setiap kelas, disusun berdasarkan kelas yang diprediksi.

Dalam kasus ini, *confusion matrix* yang diberikan memiliki dua kelas: 0 dan 1. Dapat menginterpretasikan *confusion matrix* sebagai berikut:

Tabel 4.8 Kelas *Confusion Matrix*

	<b>Prediksi 0</b>	<b>Prediksi 1</b>
Kelas 0	29 (TN)	3 (FP)
Kelas 1	2 (FN)	177 (TP)

Di mana:

1. TN (*True Negative*) adalah jumlah prediksi yang benar untuk kelas 0.
2. FP (*False Positive*) adalah jumlah prediksi yang salah untuk kelas 0, tetapi diprediksi sebagai kelas 1.
3. FN (*False Negative*) adalah jumlah prediksi yang salah untuk kelas 1, tetapi diprediksi sebagai kelas 0.
4. TP (*True Positive*) adalah jumlah prediksi yang benar untuk kelas 1.

Dengan menggunakan *confusion matrix*, Dapat menghitung metrik evaluasi lainnya, seperti:

1. Akurasi: proporsi prediksi yang benar dari total prediksi. Rumusnya adalah:

$$= \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

2. Presisi: proporsi prediksi yang benar untuk kelas 1 dari total prediksi untuk kelas 1. Rumusnya adalah:

$$= \frac{TP}{FP + TP} \times 100\%$$

3. *Recall*: proporsi prediksi yang benar untuk kelas 1 dari total kelas yang sebenarnya adalah kelas 1. Rumusnya adalah:

$$= \frac{TP}{FN + TP} \times 100\%$$

4. *F1-score*: rata-rata harmonik dari presisi dan *recall*. Rumusnya adalah:

$$= \frac{2 \times \text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \times 100\%$$

Dalam kasus ini, metrik evaluasi yang dapat dihitung adalah:

1. Akurasi:

$$\begin{aligned} &= \frac{29 + 177}{29 + 177 + 3 + 2} \times 100\% \\ &= 97\% \end{aligned}$$

2. Presisi:

$$\begin{aligned} &= \frac{177}{177 + 3} \times 100\% \\ &= 98\% \end{aligned}$$

3. *Recall*:

$$\begin{aligned} &= \frac{177}{177 + 2} \times 100\% \\ &= 99\% \end{aligned}$$

4. *F1-score*:

$$\begin{aligned} &= \frac{2 \times 0.98 \times 0.99177}{0.98 + 0.99} \times 100\% \\ &= 99\% \end{aligned}$$

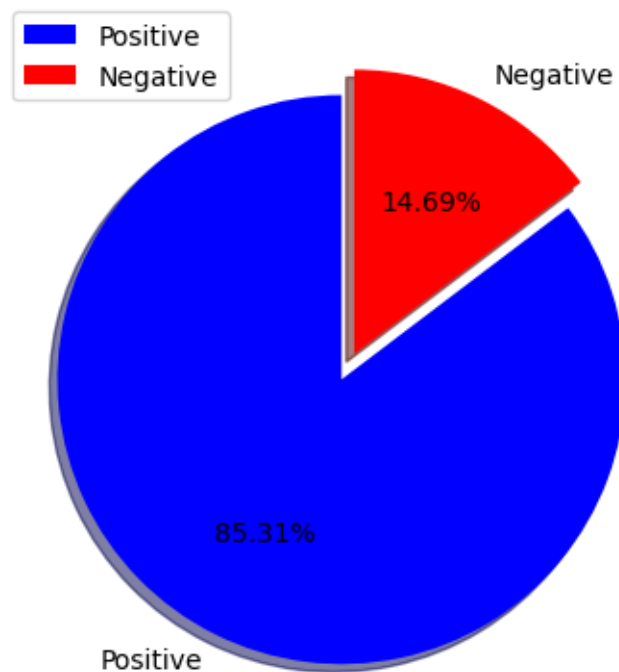
Dengan demikian, berdasarkan *Confusion Matrix* yang diberikan, model memiliki akurasi sekitar 97 %, presisi sekitar 98 %, *recall* sekitar 99 %, dan F1-

Skor sekitar 99 %. Ini menunjukkan bahwa model memiliki kinerja yang baik dalam mengklasifikasikan data.

### I. Hasil Sentimen

Berdasarkan gambar 4.3 dapat diketahui bahwa sebanyak 85.31% data memiliki sentimen "*Positive*" dan 14.69% data memiliki sentimen "*Negative*". Hal ini mengindikasikan bahwa mayoritas data dalam dataset tersebut memiliki sentimen positif. Namun, tanpa mengetahui detail tentang dataset dan metode analisis yang digunakan, sulit untuk memberikan penjelasan lebih lanjut tentang hasil analisis sentimen tersebut.

Diagram Lingkaran Data Hasil Prediksi Klasifikasi KNN



Gambar 4. 3 Diagram lingkaran Analisis sentimen

## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **A. Kesimpulan**

Kesimpulan dari hasil analisis sentimen dapat diambil dari dua bagian informasi yang diberikan:

1. Dalam Analisis Sentimen Data: Berdasarkan hasil analisis sentimen, mayoritas data dalam dataset memiliki sentimen positif (85.31%), sementara minoritas memiliki sentimen negatif (14.69%). Ini menunjukkan adanya kecenderungan yang kuat menuju sentimen positif dalam data tersebut.
  
2. Dalam Pengujian Kepuasan Pengguna Laboratorium ICT: Berdasarkan hasil pengujian, mayoritas mahasiswa mengungkapkan sentimen positif terkait kepuasan pengguna terhadap Laboratorium ICT di Universitas Sulawesi Barat. Ini dapat diartikan bahwa mahasiswa umumnya merasa puas dengan layanan dan fasilitas yang diberikan oleh laboratorium tersebut.

#### **B. Saran**

1. Memeriksa algoritma atau metode yang digunakan dalam analisis sentimen untuk memastikan hasil yang akurat.
  
2. Penelitian selanjutnya dapat memanfaatkan kamus kata selain kamus *lexicon* untuk mendapatkan performa akurasi yang lebih baik.

## DAFTAR PUSTAKA

- Alrajak, M. S., Ernawati, I., & Nurlaili, I. (2020). Analisis Sentimen Terhadap Pelayanan PT. PLN Di Jakarta Pada Twitter Dengan Algoritma *K-Nearest Neighbor* (K-NN). *Senamika*, 1(2), 110-122.
- Balya. (2019). Analisis Sentimen Pengguna Youtube Di Indonesia Pada Review Smartphone Menggunakan *Naïve Bayes*. Skripsi Universitas Sumatera Utara, 4–16.
- Cahyawati, D. (2010). Analisis Faktor-Faktor yang Mempengaruhi Kepuasan Mahasiswa Terhadap Pelayanan Laboratorium Komputer Jurusan Matematika Fakultas MIPA. *Jurnal Penelitian Sains*, 13(3).
- Deviyanto, A., & Wahyudi, M. D. R. (2018). Penerapan analisis sentimen pada pengguna twitter menggunakan metode K-Nearest Neighbor. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 3(1), 1-13.
- Ernawati, S., & Wati, R. (2018). Penerapan Algoritma *K-Nearest Neighbors* Pada Analisis Sentimen Review Agen Travel. *jurnal khatulistiwa informatika*, 6(1).
- Furqan, M., Sriani, S., & Sari, S. M. (2022). Analisis Sentimen Menggunakan *K-Nearest Neighbor* Terhadap New Normal Masa Covid-19 Di Indonesia. *Techno. Com*, 21(1), 51-60.
- Girsang, R. M., & Saragih, L. (2019). Analisa Kualitas Pelayanan Terhadap Kepuasan Mahasiswa Menggunakan Laboratorium Komputer Universitas Simalungun. *Jesya (Jurnal Ekonomi dan Ekonomi Syariah)*, 2(1), 136-144.
- Hidayat, L., Mulyana, M., & Effendy, M. (2018). Membangun Kepuasan Mahasiswa Pengguna Laboratorium Komputer. *JAS-PT (Jurnal Analisis Sistem Pendidikan Tinggi Indonesia)*, 1(2), 93-101.
- Permana, R. T. (2020). Analisis Kepuasan Pengguna Laboratoriumorato: Studi Kasus di Laboratorium Jurusan Teknik Industri Universitas Islam Indonesia.
- Sari, F. V., & Wibowo, A. (2019). Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi. *Jurnal SIMETRIS*, 10(2), 681–686.
- Syukhri, S. (2018). Analisis Kepuasan Mahasiswa Terhadap Pelayanan Laboratoriumoratorium Jaringan Menggunakan Pendekatan Importance-Performance Analysis. *INVOTEK: Jurnal Inovasi Vokasional dan Teknologi*, 18(2), 109-114.

Wijaya, A. (2022). Analisis Sentimen Kepuasan Pelanggan pada Penyedia Layanan Pengantaran Makanan. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 9(3), 1699-1708.