

# Sistem Penilaian Mahasiswa Terhadap Fasilitas Kampus Politeknik Negeri Lhokseumawe Menggunakan *Natural Language Processing*

Muhammad Icshan<sup>1</sup>, Huzaeni<sup>2\*</sup>, Amirullah<sup>3</sup>

<sup>1,2,3</sup> Jurusan Teknologi Informasi dan Komputer Politeknik Negeri Lhokseumawe  
Jln. B.Aceh Medan Km.280 Buketrata 24301 INDONESIA

<sup>1</sup>muhammadiicshan7@gmail.com

<sup>2\*</sup>huzaeni@pnl.ac.id

<sup>1</sup>amir@pnl.ac.id

**Abstrak**— Fasilitas, sebagai elemen pendukung dalam pelaksanaan tugas dan kegiatan, mendefinisikan sarana serta prasarana yang esensial untuk menunjang aktivitas suatu institusi. Studi ini dilakukan di Politeknik Negeri Lhokseumawe (PNL), di mana mahasiswa partisipan diminta untuk mengevaluasi fasilitas kampus melalui kuesioner Google. Kuesioner ini menyajikan 15 pertanyaan yang dapat diakses melalui tautan formulir kuesioner, mencakup aspek-aspek seperti ruang kelas, laboratorium, perabot belajar, peralatan pengajaran, dan layanan perpustakaan hingga teknologi e-learning. Saat ini, belum ada metode formal untuk mengklasifikasikan pandangan mahasiswa terhadap fasilitas kampus. Oleh karena itu, diperkenalkan metode klasifikasi sentimen dengan menggunakan Natural Language Processing (NLP) untuk membersihkan dan mengolah data teks. Setelah itu, data tersebut dimasukkan ke dalam model klasifikasi menggunakan Multi-layer Perceptron, yang mampu memberikan prediksi akurasi sebesar 80% berdasarkan 78 responden yang berpartisipasi dalam penelitian ini. Skenario eksperimen melibatkan 1023 data latih dan 105 data uji, termasuk ekstraksi fitur opini, menunjukkan bahwa model klasifikasi ini efektif dalam mengelola data berukuran besar dan kompleks.

**Kata kunci**— Fasilitas, Klasifikasi, *Natural Language Processing* (NLP), Sentimen

**Abstract**— Facilities, as a supporting element in the implementation of tasks and activities, define the facilities and infrastructure that are essential to support the activities of an institution. This study was conducted at Politeknik Negeri Lhokseumawe (PNL), where participating students were asked to evaluate campus facilities through a Google questionnaire. The questionnaire presents 15 questions that can be accessed through the questionnaire form link, covering aspects such as classrooms, laboratories, learning furniture, teaching equipment, and library services to e-learning technology. Currently, there is no formal method to classify students' views on campus facilities. Therefore, a sentiment classification method is introduced by using Natural Language Processing (NLP) to clean and process the text data. Afterwards, the data was fed into a classification model using Multi-layer Perceptron, which was able to provide a prediction accuracy of 80% based on 78 respondents who participated in the study. Experimental scenarios involving 1023 training data and 105 test data, including opinion feature extraction, show that this classification model is effective in managing large and complex data.

**Keywords**— Facilities, Classification, *Natural Language Processing* (NLP), Sentiment

## I. PENDAHULUAN

Fasilitas merujuk pada segala hal yang dapat membantu mempermudah dan mempercepat pelaksanaan suatu tugas atau aktivitas, dan juga termasuk sebagai sarana atau prasarana yang dibutuhkan untuk menunjang suatu kegiatan dan memperlancar jalannya pelaksanaan kegiatan tersebut.

Fasilitas kampus mencakup berbagai sarana atau peralatan yang tersedia dalam lingkungan kampus untuk memudahkan dan mendukung pelaksanaan kegiatan belajar-mengajar dan kegiatan mahasiswa.

Meningkatkan kualitas sebuah lembaga pendidikan tinggi dapat didukung oleh berbagai alat, salah satunya adalah pendapat dari mahasiswa. Mahasiswa memiliki peran penting dalam memberikan kritik dan saran yang dapat membantu meningkatkan fasilitas lembaga pendidikan tinggi. Oleh karena itu, pendapat mahasiswa menjadi salah satu elemen yang diperlukan untuk keberhasilan suatu lembaga pendidikan tinggi dalam meningkatkan kualitasnya.

Untuk mengetahui penilaian mahasiswa terhadap fasilitas di Politeknik Negeri Lhokseumawe (PNL), umumnya dilakukan survei melalui google kuesioner yang diakses melalui tautan formulir. Google kuesioner ini berisi serangkaian pertanyaan yang dapat dijawab dengan mengisi kolom yang tersedia. Namun, saat ini belum dilakukan proses pengklasifikasian data sehingga sulit untuk mengevaluasi

penilaian mahasiswa terhadap fasilitas kampus. Oleh karena itu, metode survei ini dianggap kurang efisien untuk melakukan analisis terhadap penilaian fasilitas di PNL.

Dalam rangka mengatasi masalah tersebut, dibutuhkan metode klasifikasi sentimen pada data kuesioner yang diberikan oleh mahasiswa terkait fasilitas di PNL. Hal ini akan membantu mengetahui pandangan mahasiswa tentang fasilitas yang telah disediakan dan melakukan klasifikasi data menjadi dua kategori, yaitu positif dan negatif. Hasil klasifikasi sentimen ini kemudian akan divisualisasikan untuk membantu proses evaluasi dan pengambilan keputusan pihak yang berwenang dalam meningkatkan fasilitas kampus. Untuk membangun sistem yang mampu menentukan sentimen

tersebut, diperlukan metode analisis yang tepat untuk mengolah data kuesioner.

Untuk mengklasifikasi sentimen dari opini dalam bahasa Indonesia, dibutuhkan suatu metode klasifikasi yang dapat membagi opini menjadi dua kelas, yaitu sentimen positif dan negatif. Pada proyek akhir ini, menggunakan metode NLP untuk mengklasifikasikan opini dalam bahasa Indonesia. NLP atau pemrosesan bahasa alami digunakan untuk mengolah data atau teks sehingga dapat dimengerti oleh mesin, termasuk pra-pemrosesan atau pembersihan teks. Data atau teks kemudian dimasukkan ke dalam model mesin klasifikasi menggunakan model *Multi-layer perceptron* yang akan menghasilkan prediksi persentase akurasi.

Beberapa studi sebelumnya telah meneliti sentimen, seperti yang dilakukan oleh Astri Asroviana Putri dan rekannya pada tahun 2019, serta Ali Imron pada tahun yang sama. Kedua studi tersebut menggunakan algoritma naive bayes. Ada juga penelitian lain yang dilakukan oleh Sabrah Aliyya pada tahun 2020, yang menggunakan algoritma *Support Vector Machine* (SVM). Namun, penelitian-penelitian tersebut memiliki kelemahan karena menggunakan algoritma yang bersifat probabilistik, sangat sensitive dan sulit dipakai dalam problem berskala besar, sehingga sulit mendapatkan akurasi yang baik[1] (A. P. Wibawa et al. Berbeda dengan itu, algoritma NLP dapat mengolah data dalam jumlah besar dan kompleks, sehingga dapat menghasilkan akurasi yang lebih baik[2], [3]. Namun, beberapa proses algoritma tersebut dapat mempengaruhi hasil sentiment, seperti data yang mengandung kata-kata yang tidak diperlukan dianggap kurang efisien dalam menggambarkan sentimen. Maka akan dilakukan penambahan proses data *Part of Speech* (POS) Tagging menggunakan algoritma *Hidden Markov Model* pada algoritma NLP dan akan diekstraksi fitur pendapat, sehingga diharapkan dapat meningkatkan kecepatan dan kualitas proses klasifikasi sentimen.

Oleh karena itu, dalam penelitian ini akan menggunakan pendekatan NLP dan algoritma Jaringan Saraf Tiruan untuk mengklasifikasi sentimen dengan penambahan proses data menggunakan POS tagging. Hal ini diharapkan dapat mendukung keputusan petinggi kampus dalam meningkatkan fasilitas kampus ke depannya. Dengan adanya kemampuan pengguna untuk menginput teks sentimen, para petinggi kampus dapat memperoleh wawasan langsung dari pengguna terkait persepsi dan pendapat mereka tentang fasilitas kampus. Informasi ini dapat menjadi landasan yang berharga untuk pengambilan keputusan, memungkinkan mereka untuk fokus pada area yang memerlukan perbaikan atau pengembangan lebih lanjut.

#### A. Tinjauan Pustaka

##### 1. Fasilitas

Fasilitas adalah elemen fisik yang dibangun sebagai tempat pelaksanaan aktivitas manusia. Contohnya adalah bangunan gedung, jalan, jembatan, pabrik, dan lainnya yang dapat dibangun sebagai konstruksi baru atau direnovasi. Fasilitas merupakan bagian penting dari

investasi yang didesain dan dibangun untuk memenuhi kebutuhan bisnis dan sosial manusia[4].

##### 2. Text classification

*Text classification* merupakan salah satu cabang ilmu dalam *Natural Language Processing* (NLP) yang bertujuan untuk mengkategorikan teks dengan target/label yang sudah ditentukan (*supervised learning*). Proses yang dilakukan dalam *text classification* secara berurutan adalah *text preprocessing*, *feature extraction*, *classifier algorithm selection* dan *evaluation*. *Text preprocessing* merupakan tahap membersihkan teks dari *noise*, *redundancy*, *stopwords* dan mengubah teks menjadi kata dasar. *Features extraction* mengubah data teks menjadi data numerik, hal ini dilakukan karena komputer hanya dapat memproses data numerik. *Classifier algorithm selection* melakukan pemilihan algoritma yang akan digunakan untuk pemodelan klasifikasi teks. *Evaluation* melakukan evaluasi untuk menilai kinerja model yang dihasilkan oleh algoritma yang digunakan[6].

##### 3. Natural Language Processing

*Natural Language Processing* (NLP) adalah sub-bidang Kecerdasan Buatan yang membantu komputer untuk memahami bahasa manusia yang alami. Karena mesin sekarang dapat menganalisis volume besar data tekstual dalam kecepatan dan cara yang tidak bias, kebutuhan untuk NLP semakin meningkat. Oleh karena itu, NLP membantu meningkatkan interpretasi mesin dengan menambahkan struktur numerik ke dalam data, baik itu teks atau ucapan, sehingga mempermudah analisis oleh berbagai teknik kecerdasan buatan. Algoritma pembelajaran seperti *Convolutional Neural Networks* (CNN) dan *Long Short-Term Memory* (LSTM) yang menggunakan teknik seperti penyematan kata, sekarang semakin banyak digunakan untuk membantu mesin mempelajari data bahasa manusia dalam bentuk teks dan ucapan yang tidak terstruktur[5]

##### 4. Confusion matrix

*Confusion matrix* merupakan sebuah tabel yang digunakan untuk mengukur performa algoritma klasifikasi. *Confusion Matrix* merupakan sebuah metode visualisasi dan ringkasan dari kinerja algoritma klasifikasi yang melibatkan beberapa istilah, antara lain *True Positive* (TP), *True Negative* (TN), *False Negative* (FN), dan *False Positive* (FP), sebagaimana terlihat pada Gambar 2.4. *True Positive* (TP) mengindikasikan bahwa algoritma berhasil memprediksi dengan benar kelas positif (prediksi kelas positif sesuai dengan kenyataan). *True Negative* (TN) berarti algoritma dengan tepat memprediksi data sebagai kelas negatif (prediksi kelas negatif sesuai dengan kenyataan). *False Positive* (FP) merujuk pada prediksi algoritma yang salah, mengelompokkan data sebagai kelas positif, padahal sebenarnya data tersebut adalah kelas negatif. *False Negative* (FN) artinya algoritma salah mengelompokkan data sebagai kelas negatif, padahal seharusnya data tersebut adalah kelas positif. Keempat parameter ini (TP, TN, FN, dan FP) berguna untuk mengevaluasi performa klasifikasi menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan F1 score[7]

II. METODOLOGI PENELITIAN

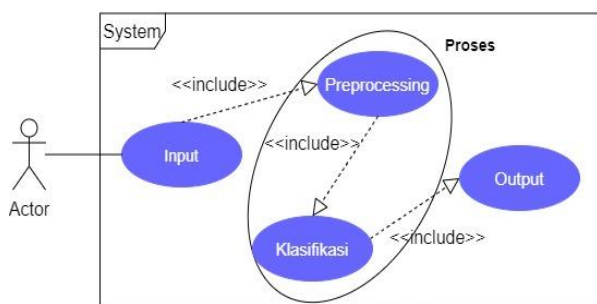
A. Analisis Kebutuhan Fungsional

Kebutuhan fungsional adalah kebutuhan yang dapat mempermudah suatu proses pengolahan data pada sistem. Adapun kebutuhan yang diperlukan dalam membuat perancangan sistem ini adalah sebagai berikut:

1. Kebutuhan Fungsional *User*  
Kebutuhan fungsional pengguna ini mencakup kemampuan bagi pengguna untuk memasukkan teks sentimen. Sentimen dalam konteks ini mengacu pada perasaan atau pendapat positif atau negatif, yang terkandung dalam suatu teks. Dengan kata lain, pengguna ingin dapat menyampaikan sentimen dari teks yang mereka masukkan.
2. Kebutuhan Fungsional *Admin*  
Beberapa kebutuhan fungsional admin:
  - a. Admin dapat melakukan login.
  - b. Admin dapat melihat dataset sentimen.
  - c. Admin dapat melihat grafik frekuensi kata sentiment positif dan negatif.
  - d. Admin dapat melihat grafik jumlah sentiment positif dan negatif
3. Pengumpulan Data  
Berikut adalah analisis kebutuhan data:
  - a. Data *User*  
User merupakan pengguna web aplikasi. Kebutuhan data user adalah data sentiment user.
  - b. Data *Training*  
Kebutuhan data untuk training yaitu data text, tag, dan label.
  - c. Data Uji  
Kebutuhan data untuk data uji yaitu data text dan label.

B. Rancangan Use Case Diagram

Use Case Diagram digunakan untuk menggambarkan bagaimana aktor dan sistem berinteraksi. Rancangan sistem pada aplikasi ini memiliki dua actor yakni *user* dan *admin*.



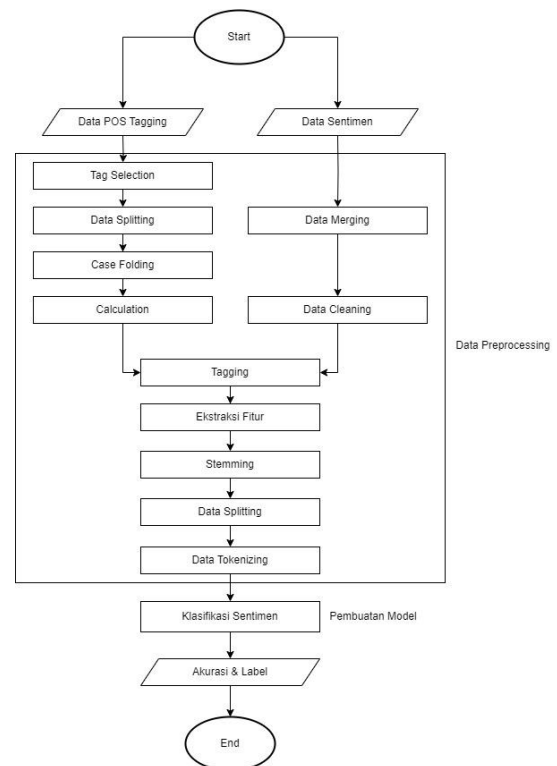
Gambar 1 Rancangan Use Case Diagram

Gambar 1 merupakan use case diagram sistem, Dalam proses input pendapat, mahasiswa memiliki kesempatan untuk secara rinci menjelaskan aktifitas yang dapat mereka lakukan. Untuk mendukung proses ini, telah disediakan sebuah form yang terdiri dari 15 pertanyaan yang harus dijawab oleh responden. Masing-masing pertanyaan dirancang untuk mencakup berbagai fasilitas yang tersedia. Pada preprocessing

menjelaskan tentang aktifitas system dalam melakukan proses pembersihan data, *case folding*, ekstraksi opini, dan klasifikasi data sentimen yang diinput oleh mahasiswa. Output yang dihasilkan pada penelitian ini adalah sentiment positif dan negatif. Hasil tersebut dimasukkan ke database dan dibuatkan grafik hasil pada halaman admin.

C. Perancangan Natural Language Processing

Rancangan flowchart *Natural Language Processing* ini bertujuan untuk merincikan langkah-langkah dalam mengimplementasi metode tersebut.



Gambar 2 Flowchart Natural Language Processing

Pada Gambar 2 diatas merupakan flowchart untuk metode *Natural Language Processing*. Dalam proses *POS Tagging*, tag selection pertama kali dilakukan dengan menghapus tag-tag yang tidak dipakai, seperti tag Z yang melambangkan punctuation dan tag SYM yang melambangkan simbol, keduanya dianggap tidak dibutuhkan dan kemudian dihapus. Setelah itu, dilakukan data splitting yang memisahkan data menjadi dua bagian, yaitu data text untuk perhitungan probabilitas emisi dan data tag untuk perhitungan probabilitas transisi oleh model. Tahapan selanjutnya adalah case folding, di mana karakter dalam data disamakan dengan mengonversi seluruh huruf menjadi huruf kecil ('a'-'z'). Terakhir, dalam fase calculation, dilakukan perhitungan probabilitas emisi dan probabilitas transisi untuk mendukung proses tagging. Dalam proses data sentimen, proses dimulai dengan data merging, yang melibatkan penggabungan data sentiment positif dan negatif dari output AI. Kemudian, dilakukan data cleaning dengan menghapus konten teks pada bagian sentimen dan mengonversi huruf menjadi huruf kecil. Proses Tagging dilakukan dengan

menggunakan *Hidden Markov Model* (HMM) untuk memberikan label pada setiap unit dalam suatu rangkaian peristiwa. Setelah itu, dilakukan Ekstraksi Fitur dengan mengekstraksi jenis kata yang menggambarkan opini. Proses *Stemming* melibatkan pemetaan dan penguraian kata menjadi bentuk kata dasarnya. Data *Splitting* membagi data sentimen menjadi data pelatihan dan data pengujian dengan proporsi 80% dan 20% secara berturut-turut. Data *Tokenization* mengubah setiap kata menjadi token numerik, diikuti oleh proses padding untuk mendapatkan input dengan panjang yang seragam. Klasifikasi sentimen menggunakan struktur *Multi Layer Perceptron* (MLP) dengan proses *feed forward* dan *back forward*. Evaluasi performa dilakukan melalui proses menghitung akurasi, presisi, *recall*, dan *F1 score*. Hasil akhir proyek mencakup tabel data, grafik frekuensi kata positif dan negatif, serta grafik jumlah sentiment positif dan negatif.

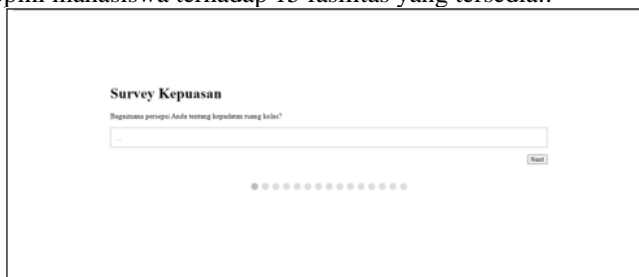
Terdapat pilihan menu tabel untuk dapat dipilih menu yang ingin ditampilkan informasi. Informasi yang ditampilkan adalah tabel, grafik frekuensi kata positif dan negative, dan grafik jumlah sentiment positif dan negatif. Tabel dapat diunduh dengan format excel.

### III. HASIL DAN PEMBAHASAN

#### A. Implementasi Sistem

##### 1. Halaman Survey

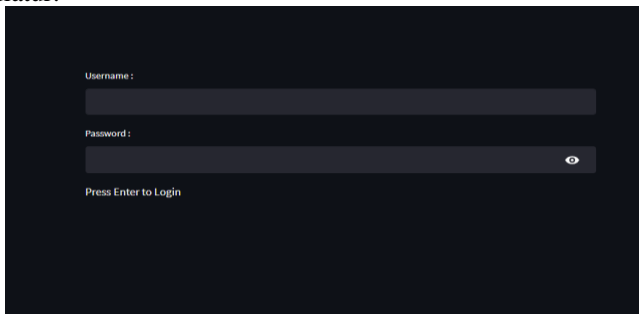
Mahasiswa akan mengisi formulir dengan sebanyak 15 halaman, disertai 15 pertanyaan yang relevan. Setelah seluruh formulir terisi, langkah selanjutnya adalah menekan tombol "submit". Proses pengisian formulir ini mencakup penilaian opini mahasiswa terhadap 15 fasilitas yang tersedia..



Gambar 2 Halaman Survey

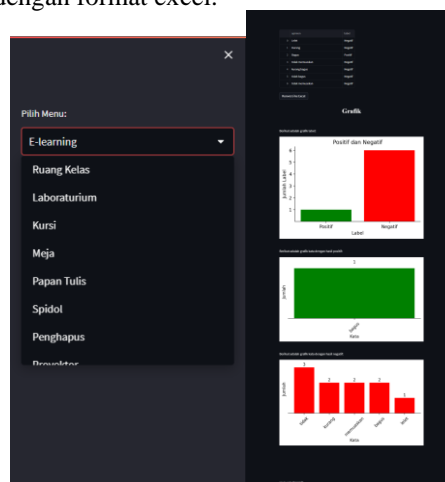
##### 2. Halaman login admin

Halaman login admin merupakan antarmuka yang dipergunakan untuk masuk ke dalam area administratif atau kendali khusus pada suatu situs web, aplikasi, atau sistem. Halaman login digunakan oleh admin untuk masuk kedalam halaman informasi dengan *username* dan *password* yang telah diatur.



Gambar 3 Halaman login admin

##### 3. Halaman Informasi tabel dan grafik



Gambar 4 Halaman Informasi

#### B. Hasil Perancangan Metode *Natural Language Processing*

##### 1. *Preprocessing*

###### a. *Case Folding*

*Case Folding* dilakukan untuk mengasimilasi atau menyederhanakan teks sehingga kapitalisasi tidak lagi menjadi faktor yang berpengaruh dalam analisis teks. Proses *case folding* dapat dilihat pada Tabel 1.

Sebelum dikonversi menjadi huruf kecil	Sesudah dikonversi menjadi huruf kecil
Ruang kelas, sangat nyaman	ruang kelas, sangat nyaman
Ruang kelas, suram dan mengganggu	ruang kelas, suram dan mengganggu
Pencahayaannya kurang memadai.	pencahayaannya kurang memadai.

Tabel 1 Proses *Case Folding*

###### b. Penghapusan tanda baca

Penghapusan tanda baca bertujuan untuk menghasilkan teks yang lebih "bersih" dan memudahkan ekstraksi informasi atau fitur dari teks..

Tabel 2 Proses Penghapusan Tanda Baca

Sebelum dihapus tanda baca	Sesudah dihapus tanda baca
ruang kelas, sangat nyaman	ruang kelas sangat nyaman
ruang kelas, suram dan mengganggu	ruang kelas suram dan mengganggu
pencahayaannya kurang memadai.	pencahayaannya kurang memadai

###### c. *Split Kata*

*Split* kata adalah proses pemisahan kata menjadi entitas tunggal. Proses ini bertujuan agar setiap kata dapat diproses dalam proses tagging.

Tabel 3 Proses *Split Kata*

Sebelum split kata	Sesudah split kata
ruang kelas sangat nyaman	'ruang;', 'kelas', 'sangat', 'nyaman'
ruang kelas suram dan mengganggu	'ruang', 'kelas', 'suram', 'dan', 'mengganggu'
pencahayaan kurang memadai	'pencahayaan', 'kurang', 'memadai'

d. *Tagging*

Proses *tagging* merujuk pada tindakan memberikan label atau penanda pada suatu objek, data, atau informasi. Proses *tagging* dapat dilihat pada Tabel 4.

Tabel 4 Proses *tagging*

Kata sebelum di tagging	Kata sesudah di tagging
'ruang;', 'kelas', 'sangat', 'nyaman'	('NN', 'ruang'), ('NN', 'kelas'), ('RB', 'sangat'), ('JJ', 'nyaman')
'ruang', 'kelas', 'suram', 'dan', 'mengganggu'	('NN', 'ruang'), ('NN', 'kelas'), ('JJ', 'suram'), ('CC', 'dan'), ('VB', 'mengganggu')
'pencahayaan', 'kurang', 'memadai'	('NN', 'pencahayaan'), ('RB', 'kurang'), ('VB', 'memadai')

e. Ekstraksi data

Proses ekstraksi data yaitu mengekstraksi jenis kata yang paling menggambarkan mengandung opini. Tag yang dipilih adalah kata sifat (JJ), kata keterangan (RB), kata benda (NN) dan kata kerja (VB)[8].

Tabel 5 Proses Ekstraksi Data

Data sudah diekstraksi	Kata digabung kembali menjadi kalimat
('NN', 'ruang'), ('NN', 'kelas'), ('RB', 'sangat'), ('JJ', 'nyaman')	'ruang kelas sangat nyaman',
('NN', 'ruang'), ('NN', 'kelas'), ('JJ', 'suram'), ('CC', 'dan'), ('VB', 'mengganggu')	'ruang kelas suram mengganggu',
('VB', 'senang'), ('NN', 'kepadatan'), ('VB', 'ada'), ('NN', 'ruang'), ('NN', 'kelas')	'pencahayaan kurang memadai'

f. Proses *Stemming*

*Stemming* adalah proses pemetaan dan penguraian bentuk dari suatu kata menjadi bentuk kata dasarnya

Tabel 6 Proses *Stemming*

Proses sebelum stemming	Hasil sesudah stemming
'ruang kelas sangat nyaman',	'ruang kelas sangat nyaman'

'ruang kelas suram mengganggu',	'ruang kelas suram ganggu'
'pencahayaan kurang memadai'	'cahaya kurang pada'

g. Tokenisasi Data

Tokenisasi data merupakan suatu metode yang digunakan untuk memisahkan setiap sentimen data menjadi kata-kata individual yang dikenal sebagai token, mengingat bahwa model hanya dapat memproses data dalam bentuk angka. Setiap token akan diubah menjadi representasi numerik yang mewakili katakata dalam data yang telah diberikan tokenisasi.

Tabel 7 Tokenisasi Data

key	value	key	value
1	x	6	suram
2	ruang	7	ganggu
3	kelas	8	cahaya
4	sangat	9	kurang
5	nyaman	10	pada

h. Proses *Sekuens and Padding*

*Sekuens* adalah sebuah larik yang berisi kumpulan token sesuai dengan setiap kata pada sebuah kalimat dalam teks. Proses *padding* adalah proses untuk membuat setiap kalimat pada teks memiliki panjang yang seragam.

Tabel 8 *Sekuens dan Padding*

Sebelum	Diubah Menjadi Sekuens	Diubah Menjadi Padding
ruang kelas sangat nyaman	[2, 3, 4, 5]	[2, 3, 4, 5, 0, 0, 0,...]
ruang kelas suram ganggu	[2, 3, 6, 7]	[2, 3, 6, 7, 0,0,0,...]
cahaya kurang pada	[8, 9, 10]	[8, 9, 10, 0,0,0,...]

2. Klasifikasi sentimen

Klasifikasi sentimen adalah tugas dalam analisis sentimen untuk menentukan polaritas emosional suatu teks. dibangun model jaringan saraf tiruan menggunakan *tensorflow* dengan *layer embedding*, *GlobalAveragePooling1D*, 1 *hidden layer* dengan 1000 *node* dan 1 *layer output*. *Optimizer* yang digunakan adalah *optimizer gradien descent* dengan *learning\_rate=0.2*. kemudian compile model dengan *loss binary\_crossentropy* dan metrik akurasi. Kemudian melatih model dengan *epoch* 1000.

Tabel 9 Hasil Klasifikasi

Kalimat opini	Nilai prediksi	Label
ruang kelas sangat nyaman	0.86257213	Positif

ruang kelas suram ganggu	0.07318411	Negatif
cahaya kurang pada	0.01645908	Negatif

Jika hasil prediksi dibawah angka 0,5 maka kalimat dinyatakan negatif dan jika hasil prediksi diatas 0,5 maka kalimat dinyatakan positif.

C. Pengujian Sistem

Pengujian yang dilakukan pada aplikasi ini yakni pengujian *black box* dan *whitebox*. Pengujian *black box* dilakukan untuk menguji *user interface* dari aplikasi, sedangkan pengujian *whitebox* dilakukan untuk menguji hasil prediksi sentimen menggunakan *confussion matrix*.

1. Pengujian *Black Box*

Dalam pengujian *black box*, dua kategori yang dianalisis adalah kelulusan dan ketidakkelulusan dari *test cases* yang telah dilakukan. *Test case* yang lulus menunjukkan bahwa sistem berhasil menjalankan fungsi-fungsi yang diuji dengan sukses, sementara *test case* yang tidak lulus mengindikasikan adanya masalah atau kesalahan dalam fungsionalitas sistem.

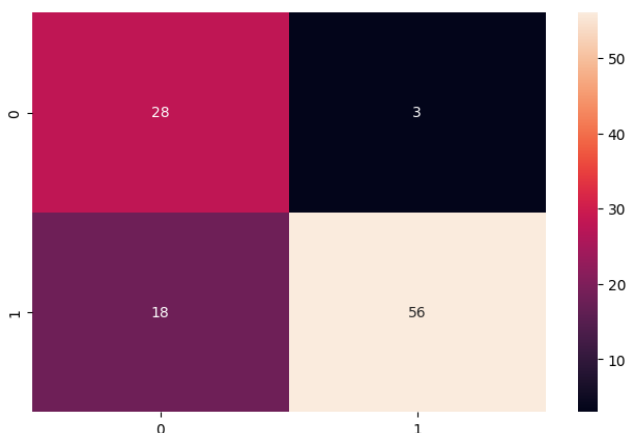
Dalam kasus ini, dari total 12 test case, sebanyak 10 test case lulus dan 2 test case tidak lulus. Maka perhitungannya adalah sebagai berikut:

$$\begin{aligned} \text{Keberhasilan (\%)} &= (\text{jumlah test case lulus} / \text{total test case}) \times 100\% \\ &= (10 / 12) \times 100\% \\ &= 83,33\% \end{aligned}$$

$$\begin{aligned} \text{Tidak Berhasil} &= (\text{jumlah test case gagal} / \text{total test case}) \times 100\% \\ &= (2 / 12) \times 100\% \\ &= 16,66\% \end{aligned}$$

2. *Confussion Matrix*

*Confussion Matrix* dapat memberikan informasi tentang berapa banyak data yang diklasifikasikan dengan benar atau salah untuk kategori sentimen yang diuji.



Gambar 7 *Confussion Matrix*

Tabel *confussion matrix* pada gambar 7 memiliki dua kolom dan dua baris dimana dua baris merupakan nilai aktual atau data dengan label yang sesuai (*true label*) sedangkan dua kolom adalah nilai prediksi atau data dengan label hasil prediksi

(*predicted label*). Tabel ini menyimpan empat nilai dimana baris pertama (dari kiri ke kanan) menyimpan nilai TN (*true negative*) dan FP (*false positive*) sedangkan baris kedua (dari kiri ke kanan) menyimpan nilai FN (*false negative*) dan TP (*true positive*). Dari hasil klasifikasi model yang diterapkan, didapat bahwa nilai TN berjumlah 28, nilai FP berjumlah 3, nilai FN berjumlah 18 dan nilai TP berjumlah 56. Dari nilai-nilai tersebut dapat dihitung akurasi, presisi, *recall* dan *F1 score* sebagai evaluasi dari kinerja model yang dibuat.

$$\begin{aligned} \text{accuracy} &= ((\text{tp} + \text{tn})) / ((\text{tp} + \text{tn} + \text{fp} + \text{fn})) \\ \text{accuracy} &= ((56 + 28)) / ((56 + 28 + 18 + 3)) = 0,8 \\ \text{precision} &= (\text{tp}) / ((\text{tp} + \text{fp})) \\ \text{precision} &= (56) / ((56 + 18)) = 0,75 \\ \text{recall} &= (\text{tp}) / ((\text{tp} + \text{fn})) \\ \text{recall} &= (56) / ((56 + 3)) = 0,94 \\ \text{F1 Score} &= 2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall}) \\ \text{F1 Score} &= 2 \times (0,75 \times 0,94) / (0,75 + 0,94) = 0,84 \end{aligned}$$

Dari perhitungan diperoleh skor neural network dalam melakukan klasifikasi yaitu akurasi sebesar 80%, *precision* sebesar 75%, *recall* sebesar 94%, dan *f-score* sebesar 84%.

IV. KESIMPULAN

- Berdasarkan hasil *confussion matrix*, sistem dapat memberikan prediksi dengan diperoleh skor neural network dalam melakukan klasifikasi yaitu akurasi sebesar 80%, *precision* sebesar 75%, *recall* sebesar 94%, dan *f-score* sebesar 84%.
- Berdasarkan pengujian *black box*, dapat disimpulkan bahwa 83,33% fungsi dari sistem berjalan dengan baik.

REFERENSI

- [1] A. P. Wibawa, M. Guntur, A. Purnama, M. Fathony Akbar, dan F. A. Dwiyanto, "Metode-metode Klasifikasi," Prosiding Seminar Ilmu Komputer dan Teknologi Informasi, vol. 3, no. 1, 2018.
- [2] E. H. Muktafin, K. Kusriani, dan E. T. Luthfi, "Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing," Jurnal Eksplora Informatika, vol. 10, no. 1, hlm. 32–42, Sep 2020, doi: 10.30864/eksplora.v10i1.390.
- [3] A. N. Rohman, E. Utami, dan S. Raharjo, "Deteksi Kondisi Emosi pada Media Sosial Menggunakan Pendekatan Leksikon dan Natural Language Processing," Eksplora Informatika, vol. 9, no. 1, hlm. 70–76, Sep 2019, doi: 10.30864/eksplora.v9i1.277.
- [4] M. Labombang, "Manajemen Pemeliharaan Fasilitas Dalam Pengelolaan Gedung," Majalah Ilmiah Mektek, Tahun X, No.1, Januari 2008.
- [5] S. Verma, A. Paul, S. S. Kariyannavar, dan R. Katarya, "Understanding the Applications of Natural Language Processing on COVID-19 Data," dalam Proceedings of the 4th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2020, Institute of Electrical and Electronics Engineers Inc., Nov 2020, hlm. 1157–1162. doi: 10.1109/ICECA49313.2020.9297490
- [6] K. Kowsari, K. J. Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, dan D. Brown, "Text classification algorithms: A survey,"

Information (Switzerland), vol. 10, no. 4. MDPI AG, 2019. doi:  
10.3390/info10040150.

- [7] P. Dellia dan A. Tjahyanto, "Tax Complaints Classification on Twitter Using Text Mining," *J Sci*, vol. 2, no. 1, hlm. 2337–8530, 2017.
- [8] G. Asrofi Buntoro, T. Bharata Adji, dan A. Erna Purnamasari, "Sentiment Analysis Twitter dengan Kombinasi Lexicon Based dan Double Propagation," Yogyakarta, Okt 2014.