

## *Sentiment Analysis of Online Services at the Engineering and Vocational Faculty of Ganesha Education University Using Naïve Bayes and LSTM Algorithms*

### **Analisis Sentimen Pelayanan Daring di Fakultas Teknik dan Kejuruan Universitas Pendidikan Ganesha Menggunakan Algoritma Naïve Bayes dan LSTM**

Ni Kadek Tesya Ari Saputri<sup>1</sup>, I Gede Aris Gunadi<sup>2</sup>, I Made Gede Sunarya<sup>3\*</sup>

<sup>1,2,3</sup>Pascasarjana Ilmu Komputer, Universitas Pendidikan Ganesha, Indonesia

E-Mail: <sup>1</sup>tesya@undiksha.ac.id, <sup>2</sup>igede.arisgunadi@undiksha.ac.id, <sup>3</sup>sunarya@undiksha.ac.id

*Received Apr 2nd 2024; Revised Jun 23th 2024; Accepted Jul 5th 2024*  
*Corresponding Author: I Made Gede Sunarya*

#### **Abstract**

*One of the faculties at Ganesha Education University is the Technical and Vocational Faculty (FTK) which is implementing online services during the COVID-19 pandemic. Various comments emerged when this online service was carried out so that an analysis was needed. Students gave positive and negative opinions. To analyze student comments using the Naive Bayes method and long short-term memory (LSTM). The data used is information obtained from distributing questionnaires filled out by FTK students. Confusion matrix evaluation includes accuracy, precision, recall, and f-measure. This research aims to compare the two Naive Bayes and LSTM methods and detect words that frequently appear in FTK online services by comparing stopwords collected during questionnaire processing. Tested the accuracy of the Naive Bayes classification method and the result was 83.69%. The classification results of the LSTM method achieved an accuracy of 53.12%. The resulting LSTM accuracy value is very low. Perhaps the main reason is that only positive comments are read when trying the LSTM method. Another cause is the dataset which can also influence. By comparing the performance of Naive Bayes and LSTM, results were obtained showing that the Naive Bayes method was superior for analyzing FTK Undiksha student comments.*

*Keyword: COVID-19, LSTM, Naïve Bayes, Sentiment Analysis*

#### **Abstrak**

Salah satu fakultas di Universitas Pendidikan Ganesha yaitu Fakultas Teknik dan Kejuruan (FTK) yang menerapkan pelayanan daring di masa pandemi COVID-19. Berbagai komentar muncul saat dilakukan pelayanan daring ini sehingga perlu dilakukannya sebuah analisis. Mahasiswa memberikan pendapat positif dan negatif. Untuk menganalisis komentar mahasiswa menggunakan metode Naive Bayes dan Long Short-Term Memory (LSTM). Data yang digunakan merupakan informasi yang diperoleh dari penyebaran angket yang diisi oleh mahasiswa FTK. Evaluasi matriks konfusi termasuk akurasi, presisi, recall, dan f-measure. Penelitian ini bertujuan untuk membandingkan kedua metode Naive Bayes dan LSTM serta mendeteksi kata-kata yang sering muncul pada pelayanan daring FTK dengan membandingkan stopwords yang dikumpulkan pada saat pemrosesan kuesioner. Menguji keakuratan metode klasifikasi Naive Bayes dan hasilnya adalah 83,69%. Hasil klasifikasi metode LSTM mencapai akurasi sebesar 53,12%. Nilai akurasi LSTM yang dihasilkan sangat rendah. Mungkin alasan utamanya adalah hanya komentar positif yang terbaca saat mencoba metode LSTM. Penyebab lainnya yaitu dataset yang juga dapat mempengaruhi. Dengan membandingkan kinerja Naive Bayes dan LSTM, diperoleh hasil yang menunjukkan metode Naive Bayes lebih unggul untuk menganalisis komentar mahasiswa FTK Undiksha.

Kata Kunci: Analisis Sentimen, COVID-19, LSTM, Naive Bayes

#### **1. PENDAHULUAN**

Akibat merebaknya infeksi virus corona (COVID-19) di Indonesia, setiap daerah sudah beralih ke daring. Universitas Pendidikan Ganesha (Undiksha) merupakan salah satu perguruan tinggi yang menerapkan sistem online dalam melaksanakan kegiatan lingkungan hidup. Fakultas Teknik dan Kejuruan (FTK)

merupakan departemen di Undiksha telah menerapkan sistem pelayanan berbentuk daring. Daring merupakan kata yang berasal dari singkatan bahasa asing yaitu Bahasa Inggris untuk "online" atau "*on the network*". Online mengacu pada penggunaan WiFi dan paket data yang digunakan oleh pengguna ponsel. Dengan kata lain, pelayanan daring adalah layanan yang dapat dilakukan hanya dengan menggunakan telepon seluler, tanpa memerlukan kontak pribadi. Layanan FTK telah berubah dari offline menjadi online sehingga semakin mempersulit mahasiswa dalam berkomunikasi. Fakta ini diperoleh dari wawancara acak dengan beberapa mahasiswa FTK. Masalah komunikasi dapat disebabkan oleh berbagai sumber, termasuk Internet. Yang paling terkena dampaknya adalah mahasiswa yang tinggal jauh dari kampus. Permasalahan yang umum terjadi adalah perencanaan KRS yang salah. Sebelum adanya pandemi COVID-19, mahasiswa harus datang ke kampus jika terjadi kesalahan jadwal perkuliahan KRS, namun kini hanya bisa mengakses kampus secara online. Hal inilah yang menjadi salah satu kendala penerapan pelayanan daring.

Hambatannya antara lain seberapa cepat atau lambat staf FTK merespons pesan, dan apakah pesan tersampaikan. COVID-19 belum berakhir dan FTK masih menyediakan beberapa pelayanan daring. Analisis sentimen diperlukan untuk mengetahui berfungsinya pelayanan daring FTK pada masa wabah COVID-19. Terlepas dari kinerja pelayanan daring yang memuaskan atau adanya keluhan dari mahasiswa, tidak ada keraguan dari mahasiswa terhadap pelayanan daring FTK. Analisis sentimen adalah pemrosesan alami atas opini-opini tertentu. Analisis sentimen ini memungkinkan kami mengidentifikasi opini di bidang ini. Selain dari komentar tersebut, Anda juga dapat menyimpulkan dari cara seseorang mengutarakan pendapatnya [1]. Penelitian ini menggunakan teknik Naive Bayes dan LSTM untuk analisis sentimen. Pelayanan Pelayanan adalah suatu kegiatan yang dilakukan oleh seseorang atau sekelompok orang berdasarkan faktor-faktor penting dengan menggunakan sistem, prosedur, dan cara yang berupaya memuaskan kepentingan orang lain sesuai dengan haknya [2].

Naive Bayes adalah metode Bayesian untuk mengklasifikasikan probabilitas yang tunduk pada asumsi independensi yang naif (kuat). Naive Bayes merupakan algoritma yang sederhana dan oleh karena itu banyak digunakan untuk klasifikasi dokumen [3]. Naive Bayes sangat sederhana, sehingga banyak orang menggunakannya untuk analisis sentimen. LSTM adalah jenis jaringan saraf berulang (RNN). LSTM merupakan modifikasi dari metode RNN [4]. RNN memiliki kelemahan yaitu tidak mampu menyimpan informasi terlalu banyak sehingga digantikan oleh Long Short-Term Memory (LSTM). LSTM dapat memodelkan urutan kronologis dan ketergantungan jangka panjang. LSTM dapat mengingat semua informasi dalam waktu lama dan menghapus informasi yang tidak relevan.

Evaluasi model dilakukan untuk mengetahui cara kerja metode. Pengujian model dilakukan dengan memeriksa hasil dari nilai akurasi dengan matriks konfusi. Kemudian data uji diuji dan data latih, maka dibuatlah daftar dari data uji tersebut. Ini disebut prediksi kelas. Kelas yang diprediksi kemudian dibandingkan dengan kelas sebenarnya dari data tes yang disembunyikan sebelumnya. Ini memungkinkan Anda untuk memeriksa dan menghitung nilai akurasi, ketajaman, pemulihan, dan poin f1 [5].

Penelitian sebelumnya yang dilakukan oleh Khurotul Aeni dan M.Faisal bertujuan untuk mengevaluasi atau memprediksi kualitas dari layanan yang berbasis IT. Data yang didapat dari pengumpulan kuesioner. Hasilnya adalah Naive Bayes mampu menghitung 52 layanan Pendidikan berdasarkan 100 data training dan 17 data validation [6]. Muhammad Afrizal kembali melakukan penelitian berjudul "Analisis Sentimen Evaluasi Mengajar Dosen Di Perguruan Tinggi Menggunakan Metodologi LSTM". Bahan menggunakan 2.280 catatan, dan jumlah kata dalam teks ulasan bervariasi antara 3 hingga 50 kata. Langkah pertama adalah penelitian kepustakaan yaitu pencarian informasi dari berbagai sumber, dan langkah kedua adalah pengolahan data. Langkah ketiga dalam mempersiapkan data yang akan digunakan adalah membangun model dengan menggunakan metode LSTM, dan langkah terakhir adalah membangun model. Metode yang dipilih harus digunakan untuk menganalisis model secara efektif. Hasil penelitian ini hasil akurasi sangat baik mencapai 91,08% [7]. Kajian lainnya yang dilakukan oleh Auliya Rahman bertajuk "Analisis Perbandingan Algoritma LSTM dan Analisis Sentimen Naive Bayes". Hasil dari penelitian ini adalah recall sebesar 83,33% [8]. Dimana penelitian ini menggunakan Langkah processing kemudian pengujian.

Berdasarkan penjelasan tersebut, solusi yang dapat diberikan adalah analisis opini komentar mahasiswa FTK Undiksha. Analisis ini dilakukan untuk mengetahui kinerja metode naif Bayes dan LSTM dalam klasifikasi berdasarkan komentar mahasiswa FTK Undiksha serta mengukur kepuasan mahasiswa terhadap kinerja pelayanan daring FTK. Hasil dari penelitian ini adalah hasil persentase metode Naive Bayes dan LSTM serta word cloud dari seluruh komentar mahasiswa FTK Undiksha. Algoritma Naive Bayes dan LSTM digunakan untuk mengatasi masalah yang diteliti dalam konteks analisis sentimen pelayanan daring FTK Undiksha. Masalah ini terjadi karena ini adalah pelayanan daring. Untuk mengetahui lebih jauh mengenai pelayanan daring FTK Undiksha, peneliti menggunakan teknik Naive Bayes dan LSTM. Metode ini digunakan untuk menganalisis data komentar mahasiswa FTK. Ada beberapa rumusan masalah yang dapat dijelaskan. Analisis sentiment dilakukan untuk menganalisis tentang pelayanan daring FTK Undiksha. Karena dengan adanya pelayanan daring tersebut menimbulkan beberapa komentar mahasiswa. Dengan menggunakan metode Naive Bayes dan LSTM.

## 2. METODOLOGI PENELITIAN

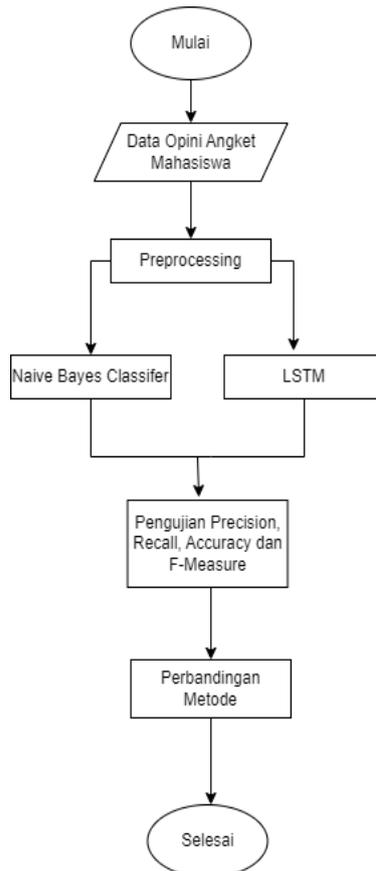
Tujuan penggunaan metode penelitian ini adalah agar proses penelitian menjadi lebih terstruktur, sistematis, terkontrol, dan terfokus. Lebih lanjut, metode penelitian bertujuan untuk melacak kemajuan dan keberhasilan penelitian serta mengidentifikasi hasil yang diharapkan dari setiap masukan yang berurutan. Tahapan penelitian ini ditunjukkan pada Gambar 1.

### 2.1. Studi Kepustakaan

Studi literatur adalah serangkaian kegiatan yang melibatkan proses mempertimbangkan topik yang akan dibahas [9].

### 2.2. Alur Penelitian

Alur penelitian ditunjukkan pada gambar 1.



**Gambar 1.** Alur Penelitian

Alur penelitian yang dilakukan adalah sebagai berikut.

#### 1. Pengumpulan Data

Pengumpulan data merupakan suatu proses yang menggambarkan metode yang digunakan peneliti untuk mengumpulkan data [10]. Proses pengumpulan data dilakukan dengan mengumpulkan data survei yang dikumpulkan menjadi dataset untuk memprediksi kepuasan server online yang dicapai melalui pengiriman Google Form. Dengan menyebarkan angket, mahasiswa akan diberikan pertanyaan-pertanyaan yang dapat mereka jawab dengan menggunakan teks yang ada di angket. Survei Mahasiswa FTK Undiksha angkatan 2020, 2021 dan 2022.

#### 2. *Preprocessing* Data

Pengolahan data adalah suatu proses pengolahan data yang digunakan menurut teknik yang ditentukan [11]. Proses ini terlebih dahulu membersihkan data agar mudah diolah menggunakan *text mining*. Informasi yang dikumpulkan dari survei diolah menjadi data bersih. Langkah-langkah prapemrosesan data ditunjukkan pada Gambar 2.

**Gambar 2. Preprocessing Data**

3. Model Klasifikasi

Langkah selanjutnya adalah pemodelan klasifikasi. Ini diimplementasikan menggunakan Naive Bayes dan LSTM. Nilai presisi, akurasi, recall, dan F1 tertinggi menunjukkan metode terbaik. Pada penelitian ini data latih digunakan untuk pemodelan klasifikasi dengan menggunakan 70% data latih dan 30% data uji.

4. Evaluasi Model Klasifikasi

Komentar positif dan negatif ditentukan melalui pengujian oleh para ahli (ahli bahasa). Gunakan pengujian ahli untuk menentukan secara akurat komentar positif dan negatif. Pengujian ini menggunakan matriks konfusi [12] yang berisi nilai presisi, presisi, recall, dan nilai f berikut ini [13].

a. Accuracy

Accuracy merupakan prediksi yang dilakukan dengan data dengan data yang diprediksi benar.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} * 100\% \tag{1}$$

b. Precision

Precision merupakan prediksi yang dilakukan dengan data tidak sebenarnya positif dengan data yang diprediksi positif.

$$Precision = \frac{TP}{TP+FP} * 100\% \tag{2}$$

c. Recall

Recall merupakan prediksi yang dilakukan dengan data positif dengan data prediksi positif.

$$Recall = \frac{TP}{TP+FN} * 100\% \tag{3}$$

d. F-measure

Ketepatan dan perolehan rata-rata harmonik disebut pengukuran.

$$F-Measure = 2 * \frac{Presisi * Recall}{Presisi+Recall} * 100\% \tag{4}$$

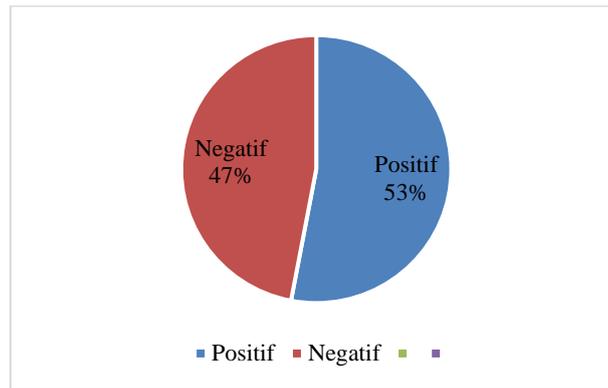
**3. HASIL DAN PEMBAHASAN**

**3.1. Pengumpulan Data**

Pengumpulan data merupakan suatu proses yang menggambarkan proses pengumpulan data [14]. Pengumpulan data dilakukan dengan membagikan Google Form kepada mahasiswa FTK 3 angkatan 2020, 2021, dan 2022. Data ini kemudian diolah untuk mendapatkan data yang bersih.

**Tabel 1. Pengumpulan Data**

Data Clean	Label
Harap jelas layanan mahasiswa	Negative
Banyak alas susah	Negative
Perlu ditingkatkan	Positive
Ramah mahasiswa tingkat	Positive
Lebih tanggap selesai masalah mahasiswa	Negative
Lebih tanggap selesai masalah siswa	Negative
Ramah tingkat	Positive
Lebih tingkat ramah perhati laku layanan	Negative
Tahan tingkat layanan	Positive
Tahan tingkat layanan prima	Positive
Tingkat cepat tepat waktu layanan	Positive
Jaga tahan layanan	Positive
Layanan daring lebih cepat tanggap	Negative



**Gambar 3.** Persentase Sentimen Komentar Mahasiswa

Pengumpulan data dilakukan dengan membagikan Google Form kepada mahasiswa FTK 3 angkatan 2020, 2021, dan 2022. Data ini kemudian diolah untuk mendapatkan data yang bersih. Dari data 425 dari hasil komentar mahasiswa FTK menunjukkan hasil presentase yaitu 53% positif dan 47% negatif.

### 3.2. Preprocessing Data

1. *Case folding* merupakan operasi yang mengubah huruf kecil dari data yang dibersihkan, disebut case folding [15].

**Tabel 2.** Case Folding

No	Komentar
1	terus meningkatkan kecepatan dan ketepatan waktu pelayanannya,
2	terus menjaga dn mempertahankan pelayanannya.
3	pelayanan daring bisa lebih cepat ditanggapi/
4	lebih ditingkatkan pelayanan di fakultas[
5	lebih memberikan perhatian yang bagus'

2. *Cleansing* merupakan proses yang dilakukan untuk memastikan bahwa data komentar tidak mengandung tanda baca, normalisasi unicode, dll.

**Tabel 3.** Cleansing

No	Komentar
1	terus meningkatkan kecepatan dan ketepatan waktu pelayanannya
2	terus menjaga dn mempertahankan pelayanannya
3	pelayanan daring bisa lebih cepat ditanggapi
4	lebih ditingkatkan pelayanan di fakultas
5	lebih memberikan perhatian yang bagus

3. Tokenisasi adalah proses pemisahan kata, simbol, frasa, dan elemen penting lainnya dari teks. Fase tokenisasi dilakukan menggunakan pustaka "re", yang merupakan komponen bawaan bahasa pemrograman *Python*.

**Tabel 4.** Tokenisasi

No	Komentar	Tokenisasi
1	terus meningkatkan kecepatan dan ketepatan waktu pelayanannya	[terus' meningkatkan' kecepatan 'dan' ketepatan' waktu 'pelayanannya]
2	terus menjaga dn mempertahankan pelayanannya	[terus' menjaga' dn 'mempertahankan' pelayanannya]
3	pelayanan daring bisa lebih cepat ditanggapi	[elayanan' daring' bisa' lebih 'cepat 'ditanggapi]
4	lebih ditingkatkan pelayanan di fakultas	[lebih 'ditingkatkan' pelayanan 'di 'fakultas]
5	lebih memberikan perhatian yang bagus	[lebih' memberikan' perhatian' yang' bagus]

4. Perumusan dan penerjemahan merupakan perubahan ke kata baku. Contoh kata yang tidak baku seperti sngkatan yang akan diubah menjadi kata standar.

**Tabel 5.** Formalasi dan Translasi

No	Komentar	Tokenisasi
1	terus meningkatkan kecepatan dan ketepatan waktu pelayanannya	tingkat cepat pelayanannya
2	terus menjaga dn mempertahankan pelayanannya	jaga tahan layan
3	pelayanan daring bisa lebih cepat ditanggapi	layan daring cepat tanggap
4	lebih ditingkatkan pelayanan di fakultas lebih memberikan perhatian yang bagus	tingkat layan fakultas perhati bagus

- Library Sastrawi bahasa pemrograman Python digunakan untuk melakukan *stemming*. *Stemming* ini bertujuan untuk menghilangkan imbuhan. Perpustakaan literatur sumber utama yaitu mengkodekannya untuk mendapatkan data dari data *database stopwords*. Setelah data master disimpan, data tersebut disimpan dengan file baru.

**Tabel 6.** Stemming

No	Komentar	Stemming
1	tingkat cepat pelayanannya	tingkat cepat layan
2	jaga tahan layan	jaga tahan layan
3	layan daring cepat tanggap	layan daring cepat tanggap
4	tingkat layan fakultas	tingkat layan fakultas
5	perhati bagus	perhati bagus

### 3.3. Ekstraksi Dataset

Untuk mengekstrak teks ke dalam array vektor, kita perlu mengekstrak kumpulan datanya. Vektor ini digunakan untuk memberi masukan pada model LSTM dan Naive Bayes. Teknik *Term Frekuensi-Inverse Document* (TF-IDF) digunakan dalam penelitian ini. Proses TF-IDF mengintegrasikan frekuensi *term* (tf) menjadi frekuensi dokumen *invers* (idf) [16].

**Tabel 7.** Data Sampel

No	Data Sampel
1	ramah tingkat
2	layan mahasiswa banyak

Langkah-langkah TF-IDF yaitu [17]:

- Perhitungan tf dilakukan untuk pada setiap dokumen ditentukan jumlahnya.

**Tabel 8.** Perhitungan tf

Term	D1	D2
ramah	1	0
tingkat	1	0
mahasiswa	1	0
layan	0	1
banyak	0	1

- Menghitung df adalah kata ke-I dihitung jumlah dokumennya.

**Tabel 9.** Menghitung df

Term	df
ramah	1
tingkat	1
mahasiswa	1
layan	1
banyak	1

- Hitung idf dengan rumus  $\log(df / \text{jumlah dokumen})$ .

**Tabel 10.** Hitung idf

Term	df
ramah	0,3
tingkat	0,3
mahasiswa	0,3
layan	0,3
banyak	0,3

## 4. Menghitung Tf-Idf

**Tabel 11.** Menghitung Tf-Idf

Term	D1	D2
ramah	0,3	0
tingkat	0,3	0
mahasiswa	0,3	0
layan	0	0,3
banyak	0	0,3

## 5. Tabulasi nilai pada Tf-Idf

**Tabel 12.** Hasil Tabulasi Nilai pada Tf-Idf

Term	D1	D2	TF (1)	TF (2)	IDF
ramah	1	0	00.30	0	00.30
tingkat	1	0	00.30	0	00.30
mahasiswa	1	0	00.30	0	00.30
layan	0	1	0	00.30	00.30
banyak	0	1	0	00.30	00.30

**Tabel 13.** Hasil Tabulasi Nilai pada Tf-Idf

Term	D1	D2
Total	0,9	0,6

**3.4. Pembagian Dataset Untuk Model Klasifikasi**

Pemisahan dataset merupakan pemisahan yang dilakukan untuk menentukan data pelatihan dan pengujian [18]. Pelabelan komentar dilakukan secara manual sebelum menentukan data pengujian dan pelatihan. Selanjutnya menggunakan proses partisi untuk membedakan kumpulan data. Membedakan kumpulan data menggunakan nilai dari `dataset_type` yang menentukan apakah dataset dimasukkan sebagai data pelatihan atau pengujian. Jika proses memisahkan kumpulan data berdasarkan rasio atau persentase, gunakan opsi Stratify untuk menyamakan proporsi sesuai dengan presentase yang digunakan [19]. Untuk presentase datanya digunakan 70% data training dan 30% data validation.

**3.5. Hasil Evaluasi Model Klasifikasi**

Evaluasi model melibatkan pengujian model yang dilakukan sesuai dengan model yang diuji [20]. Membandingkan rata-rata kinerja akurasi dan hasil eksperimen dari total perbandingan 70:30, kinerja pertama percobaan (yaitu 85,93), kinerja kedua percobaan (yaitu 82,35), dan kinerja ketiga percobaan (yaitu 82,81). Tidak ada perubahan signifikan pada ketiga percobaan yang dilakukan atau hasil percobaan pertama, kedua dan ketiga tidak berbeda nyata. Hasil rata-ratanya sebesar 83,69. Nilai akurasi yang tinggi menunjukkan tingginya proporsi nilai prediksi yang benar terhadap total data, atau tingkat kemiripan yang tinggi antara data positif dan negatif.

**Tabel 14.** Hasil Perhitungan Naïve Bayes

Presentase	Percobaan	Accuracy	Precision	Recall	F-Measure
70%-30%	1	85,93	86,02	85,93	85,89
	2	82,35	84,46	82,35	81,9
	3	82,81	83,47	82,81	82,61
Rata-Rata		83,69	84,65	83,69	83,46

Hasil yang didapat pada nilai LSTM berdasarkan dari jumlah *epoch* dan neuronnya. *Training loss* terbawah pada 256 neuron dan 50 *epoch* hasil 69.27, dan *training loss* teratas pada 128 neuron dan 250 *epoch* nilai 69.68. *Validation loss* terbawah pada 64 neuron dan 250 *epoch* dengan nilai 69.14, dan



### 3.7. Diskusi

Penelitian ini menggunakan Google Forms untuk mengumpulkan data dari FTK. Data yang diperoleh dari survei Google Forms yang akan dijadikan sampel pada analisis berdasarkan populasi mahasiswa FTK pada ketiga angkatan yaitu 2020, 2021, dan 2022. Pada pengumpulan data tahap awal dilakukan purposive sampling dengan mempertimbangkan tiga kelas yang mempunyai dampak paling besar terhadap pelayanan daring FTK. Gaya-gaya yang digunakan ditentukan dan diperoleh tiga sampel gaya dengan menggunakan rumus Slovin. Pengambilan sampel dilakukan dengan persamaan Slovin karena ini paling cocok. Banyaknya jumlah sampel data yang menggunakan persamaan Slovin berjumlah 314. Banyaknya menggunakan data adalah 425 dimana sudah lebih dari data minimum. Setelah sampel diolah, kemudian praproses untuk menghilangkan huruf kecil, tanda baca, bahasa gaul, dan singkatan dari data tersebut. Preprocessing dilakukan untuk memperoleh data yang lebih bersih dan memudahkan analisis komentar yang terdiri dari kata baku. Sekarang setelah kita memproses data sebelumnya, kita dapat mengekstrak kumpulan data dan mengonversi kata menjadi vektor. Dilakukan ekstraksi dataset sedemikian rupa sehingga system dapat membaca metode yang digunakan. Betapa pentingnya sebuah kata dalam kumpulan dokumen. Ukuran statistik ini digunakan sebagai faktor pembobotan untuk pengambilan informasi, teks, dan pemodelan pengguna. TF-IDF dilakukan untuk mengetahui frekuensi kata dalam sebuah kalimat. Analisis kemudian dilakukan dengan menggunakan metode Naive Bayes dan LSTM. Teorema Bayes adalah dasar klasifikasi Naive Bayes dan digunakan dalam metode probabilistik dan statistik untuk memprediksi peluang di masa depan.

LSTM adalah suatu teknik untuk mengolah, memprediksi dan mengklasifikasikan informasi berdasarkan time series. LSTM merupakan salah satu jenis arsitektur jaringan saraf yang sangat efektif dalam memproses data deret waktu, seperti data teks. Data deret waktu ditampilkan sebagai deret waktu di mana setiap kata atau token direpresentasikan sebagai suatu titik waktu. Metode Naive Bayes dilakukan dalam beberapa percobaan dengan rasio variasi persentase 60: 40, 70: 30, dan 80: 20. Dari lima percobaan yang dilakukan pada ketiga perbandingan tersebut, ditentukan nilai rata-rata akurasi pada ketiga perbandingan tersebut. Perbandingan 60:40 mempunyai rata-rata skor akurasi 83.52, Perbandingan 70:30 mempunyai rata-rata skor akurasi 83.96, Perbandingan 80:20 mempunyai rata-rata skor akurasi 82.35. Melihat hasil nilai rata-rata presisi dari ketiga percobaan yang dilakukan, diperoleh nilai terbaik yaitu rasio 70: 30 dan nilai presisi 83,96, sehingga perbandingan tersebut kami gunakan untuk melakukan percobaan lainnya. Dari lima epoch yang dilakukan, skor tertinggi berhasil diraih. Epoch 50 mencapai nilai akurasi tertinggi sebesar 53.12 dan nilai loss sebesar 69.22 untuk Neuron 64. Pada epoch 100, nilai terbaik dari nilai akurasi 53.12 dan nilai loss 69.18 dicapai dengan jumlah neuron sebanyak 256. Epoch 150 mencapai nilai akurasi tertinggi Neuron 128 sebesar 53.12 dan nilai loss sebesar 69.17, Epoch 200 mencapai nilai akurasi tertinggi Neuron 64 sebesar 53.12 dan nilai loss sebesar 69.19. Epoch 50 mencapai nilai akurasi tertinggi Neuron 64 sebesar 53.12 dan nilai loss sebesar 69.14. Lima nilai teratas tidak berbeda nyata, artinya nilai LSTM tidak terpengaruh oleh perubahan epoch atau neuron. Nilai terbaik diperoleh dari lima kali percobaan. Artinya, hasil terbaik dengan 250 epoch dan 64 neuron yaitu hasil akurasi 53,12 dan hasil loss 69,14. Oleh karena itu, nilai matriks konfusi untuk LSTM adalah presisi 53,12, presisi 82,82, recall 53,12, dan f-value 36,65. Hasil nilai akurasi kedua metode menunjukkan bahwa metode Naive Bayes lebih cocok untuk melakukan analisis sentimen terhadap pelayanan daring FTK Undiksha.

## 4. KESIMPULAN

Berdasarkan penjelasan di atas dapat disimpulkan bahwa metode Naive Bayes lebih baik dibandingkan metode LSTM untuk pelayanan daring FTK Undiksha. Metode ini menghasilkan skor akurasi sebesar 83,69, presisi sebesar 84,65, recall sebesar 83,69, dan skor f-measure sebesar 83,46 dengan persentase rasio 70:30. Nilai diperoleh dari beberapa percobaan dengan perbandingan 60: 40, 70: 30, dan 80: 20. Keunggulan Bayes antara lain proses komputasi yang cepat, algoritma yang sederhana, dan akurasi yang tinggi. Kekurangan Bayes adalah membutuhkan waktu untuk memilih banyak atribut. Dengan menggunakan metode LSTM diperoleh skor akurasi sebesar 53,12, presisi sebesar 82,82, skor recall sebesar 53,12, dan f-measure sebesar 36,65 dengan persentase rasio sebesar 70: 30. Nilai-nilai yang diperoleh ditentukan sebagai nilai terbaik karena diperoleh melalui eksperimen dengan beberapa neuron. LSTM memiliki beberapa keunggulan. Artinya pola pembelajaran proses pelatihan arsitektur lebih stabil, namun kekurangannya adalah pelatihan yang membutuhkan waktu lama.

Hasil perbandingan kedua metode tersebut jelas menunjukkan bahwa metode Naive Bayes mengungguli metode LSTM dengan nilai akurasi yang berbeda secara signifikan. Skor akurasi Naive Bayes sebesar 83,69, sedangkan skor akurasi LSTM sebesar 53,12. Jika membaca semua komentar saat pengujian, nilai 0 akan menjadi mayoritas, sehingga LSTM akan memiliki nilai kecil. Akibatnya, LSTM tidak bisa membedakan komentar positif dan negatif. Karena nilai 0 dominan, semua komentar dari LSTM dibaca positif. Hal ini meningkatkan nilai akurasi LSTM karena akurasinya membandingkan data positif aktual dengan data prediksi positif. Berbeda dengan presisi, recall rendah komntar dilihat positif. Hasil f-measure kecil karena selisih nilai presisi dan recall besar. Selain itu, karakteristik data itu sendiri yaitu perbedaan panjang data juga mempengaruhi nilai LSTM.

**REFERENSI**

- [1] Bhatia, S., Sharma, M., & Bhatia, K. K. (2018). Sentiment Analysis and Mining of Opinions. *Studies in Big Data*, 30(May), 503–523. [https://doi.org/10.1007/978-3-319-60435-0\\_20](https://doi.org/10.1007/978-3-319-60435-0_20).
- [2] Fernandes, Y. D., & Marlius, D. (2018). Peranan Customer Service Dalam Meningkatkan Pelayanan Kepada Nasabah Pada PT. Bank Pembangunan Daerah Sumatera Barat Cabang Utama Padang.
- [3] Cahyo, A. D. (2023). METODE NAIVE BAYES UNTUK KLASIFIKASI MASA STUDI SARJANA. *Jurnal Teknologi Pintar*, 3(4).
- [4] Iman, F. N., & Wulandari, D. (2023). Prediksi Harga Saham Menggunakan Metode Long Short Term Memory. *LOGIC: Jurnal Ilmu Komputer Dan Pendidikan*, 1(3), 601-616.
- [5] Muhammadin, A., & Sobari, I. A. (2021). Analisis Sentimen Pada Ulasan Aplikasi Kredivo Dengan Algoritma Svm Dan Nbc. *Reputasi: Jurnal Rekayasa Perangkat Lunak*, 2(2), 85-91.
- [6] Aeni, K. A. (2020). Prediksi Kepuasan Layanan Akademik Menggunakan Algoritma Naïve Bayes. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 7(3), 601–609. <https://doi.org/10.35957/jatisi.v7i3.603>.
- [7] Amrustian, M. A., Widayat, W., & Wirawan, A. M. (2022). Analisis Sentimen Evaluasi Terhadap Pengajaran Dosen di Perguruan Tinggi Menggunakan Metode LSTM. *Jurnal Media Informatika Budidarma*, 6(1), 535. <https://doi.org/10.30865/mib.v6i1.3527>.
- [8] Isnain, A. R., Sulistiani, H., Hurohman, B. M., Nurkholis, A., & Styawati, S. (2022). Analisis Perbandingan Algoritma LSTM dan Naive Bayes untuk Analisis Sentimen. *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, 8(2), 299-303.
- [9] Hadna, N. M. S., Santosa, P. I., & Winarno, W. W. (2016). Studi literatur tentang perbandingan metode untuk proses analisis sentimen di Twitter. *Semin. Nas. Teknol. Inf. dan Komun*, 2016, 57-64.
- [10] Makbul, M. (2021). Metode pengumpulan data dan instrumen penelitian.
- [11] Harmayani, H., Abdilah, D., Mapilindo, M., Oktopanda, O., & Hutahaeon, J. (2021). Aplikasi Komputer. Yayasan DPI, 1-89.
- [12] Syukri Mustafa, M., Rizky Ramadhan, M., & Thenata, A. P. (2017). Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Citec Journal*, 4(2), 151–162.
- [13] Rahman, M. Z., Sari, Y. A., & Yudistira, N. (2021). Analisis Sentimen Tweet COVID-19 menggunakan Word Embedding dan Metode Long Short-Term Memory (LSTM). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(11), 5120–5127. Retrieved from <http://j-ptiik.uib.ac.id>.
- [14] -, S., Fadlil, A., & -, S. (2018). Analisis Sentimen Menggunakan Metode Naïve Bayes Classifier Pada Angket Mahasiswa. *Saintekbu*, 10(2), 1–9. <https://doi.org/10.32764/saintekbu.v10i2.190>.
- [15] Amrustian, M. A., Widayat, W., & Wirawan, A. M. (2022). Analisis Sentimen Evaluasi Terhadap Pengajaran Dosen di Perguruan Tinggi Menggunakan Metode LSTM. *Jurnal Media Informatika Budidarma*, 6(1), 535. <https://doi.org/10.30865/mib.v6i1.3527>.
- [16] Wijaya, A. P., & Santoso, H. A. (2016). Naive Bayes Classification pada Klasifikasi Dokumen Untuk Identifikasi Konten E-Government Naive Bayes Classification on Document Classification to Identify E-Government Content. *Journal of Applied Intelligent System*, 1(1), 48–55.
- [17] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [18] Gunawan, B., Pratiwi, H. S., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 4(2), 113. <https://doi.org/10.26418/jp.v4i2.27526>.
- [19] Aldi, M. W. P., Jondri, & Aditsania, A. (2018). Analisis dan Implementasi Long Short-Term Memory Neural Network untuk Prediksi Harga Bitcoin. *Jurnal Informatika*, 5, No (2), 3548. Retrieved from <http://openlibrarypublications.telkomniversity.ac.id>.
- [20] Syafrizal, S., Afdal, M., & Novita, R. (2024). Analisis Sentimen Ulasan Aplikasi PLN Mobile Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor: Sentiment Analysis of PLN Mobile Application Review Using Naïve Bayes Classifier and K-Nearest Neighbor Algorithm. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(1), 10-19.
- [21] Nasution, M. R. A., & Hayaty, M. (2019). Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter. *Jurnal Informatika*, 6(2), 226–235. <https://doi.org/10.31311/ji.v6i2.5129>.
- [22] Muhammadin, A., & Sobari, I. A. (2021). Analisis Sentimen Pada Ulasan Aplikasi Kredivo Dengan Algoritma Svm Dan Nbc. *Reputasi: Jurnal Rekayasa Perangkat Lunak*, 2(2), 85-91.