

Penerapan Model *Deep Learning* BiLSTM untuk Klasifikasi Multi-Kelas pada Data Aduan Masyarakat

Kautsa Adi Garda^{*1}, Galuh Wilujeng Saraswati², Erba Lutfina³

¹²³Sistem Informasi, Fakultas Sistem Informasi, Universitas Dian Nuswantoro
Kediri, Jawa Timur, Indonesia
Email: ¹612202400171@mhs.dinus.ac.id, ²galuhwilujeng@dsn.dinus.ac.id,
³erba.lutfina@dsn.dinus.ac.id

Abstrak

Layanan Pengaduan Masyarakat merupakan suatu kanal layanan yang dipergunakan masyarakat untuk melaporkan suatu kejadian pelanggaran dan masalah yang tidak sesuai dengan aturan tertulis maupun tidak tertulis. Pemerintah Kota Kediri melalui Dinas Komunikasi dan Informatika mempunyai suatu sistem layanan pengaduan online untuk menampung keluhan, saran, serta aspirasi masyarakat. Keberagaman dan tingginya volume aduan menuntut operator untuk meneruskan laporan kepada banyak pilihan instansi terkait sekaligus. Oleh karena itu, perlu adanya sistem yang mampu melakukan klasifikasi otomatis terhadap teks aduan masyarakat. Penelitian ini bertujuan untuk otomatisasi klasifikasi teks aduan menggunakan algoritma *Bidirectional Long Short-Term Memory* (BiLSTM) dengan skema *multiclass* sebanyak 21 kelas. Pembagian data dilakukan dengan rasio 80:20 untuk data *training* dan *testing*, kemudian 10% dari data pelatihan digunakan sebagai data *validation*. Penelitian ini menggunakan *Word2Vec* untuk pembobotan kata, *Random Oversampling* untuk *balancing* data serta membandingkan uji parameter *batch size* dan *optimizer*. Berdasarkan hasil dari 8 skenario pengujian, diperoleh skenario model dengan akurasi tertinggi sebesar 0,6977 (70%). Nilai tersebut tergolong cukup baik mengingat jumlah kelas yang relatif banyak dan variasi penggunaan bahasa daerah dan *slang* pada data aduan. Secara keseluruhan, BiLSTM menunjukkan kinerja yang memadai untuk klasifikasi teks *multiclass*.

Kata kunci: aduan, *BiLSTM*, klasifikasi, *multiclass*, *random oversampling*, *word2vec*.

Abstract

The Public Complaint Service is a service channel used by the public to report violations and issues that don't comply with written or unwritten rules. The Kediri City Government, through the Department of Communication and Information Technology, has an online complaint service system to receive complaints, suggestions, and public feedback. The diversity and high volume of complaints require operators to forward reports to multiple relevant agencies simultaneously. Therefore, a system capable of automatically classifying public complaint texts is necessary. This study aims to automate the classification of complaint texts using *Bidirectional Long Short-Term Memory* (BiLSTM) algorithm with a 21-class *multiclass* scheme. Data was split in an 80:20 ratio for training and testing, with 10% of the training data used as validation data. This study utilized *Word2Vec* for word weighting, *Random Oversampling* for data balancing, and compared tests of batch size and optimizer parameters. Based on the results of 8 testing scenarios, the model scenario with the highest accuracy of 0.6977 (70%) was obtained. This value is considered quite good given the relatively large number of classes and the variation in the use of regional languages and slang in the complaint data. Overall, BiLSTM demonstrates adequate performance for *multiclass* text classification.

Keywords: *BiLSTM*, complaint, classification, *multiclass*, *random oversampling*, *word2vec*.

This work is an open access article and licensed under a Creative Commons Attribution-NonCommercial ShareAlike 4.0 International (CC BY-NC-SA 4.0)



1. PENDAHULUAN

Penerapan teknologi informasi dalam tata kelola pemerintahan merupakan hal yang sangat penting, karena tidak hanya berfokus pada transformasi digital saja tetapi juga dalam optimalisasi pemanfaatan teknologi informasi terutama pada sektor pelayanan publik [1]. Dalam menjalankan fungsi pelayanan publik, pemerintah sering kali menghadapi tantangan berupa proses pelayanan yang belum efektif, cenderung lambat, dan mekanisme yang dinilai menyulitkan masyarakat. Oleh karena itu, diperlukan

penerapan *e-Government* dalam sistem pemerintahan untuk membentuk suatu transformasi birokrasi menuju sistem yang lebih responsif, cepat, serta berorientasi pada pelayanan masyarakat [1],[2].

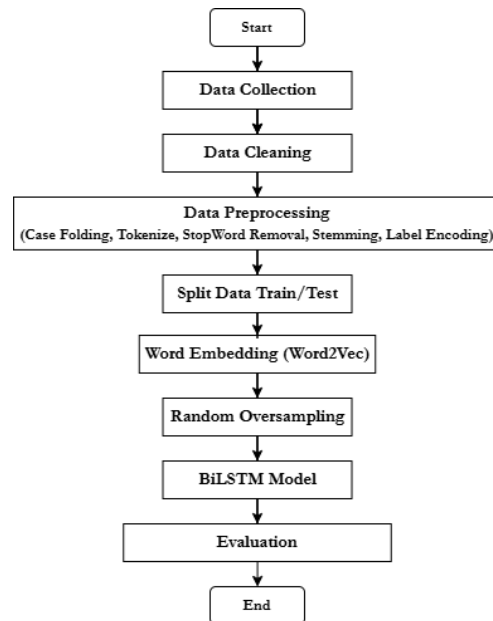
Sistem Pemerintahan Berbasis Elektronik (SPBE) adalah inisiatif Pemerintah Indonesia dalam mewujudkan transformasi digital birokrasi. Melalui Peraturan Presiden Nomor 95 Tahun 2018, ditegaskan bahwa seluruh instansi pemerintah, baik di tingkat pusat maupun daerah, wajib menerapkan SPBE dengan memanfaatkan Teknologi Informasi dan Komunikasi (TIK) dalam setiap penyelenggaraan layanan pemerintahan [3]. Ketentuan ini diperkuat melalui Peraturan Menteri PANRB Nomor 59 Tahun 2020 yang mengatur tentang mekanisme pemantauan dan evaluasi pelaksanaan SPBE.

Layanan aduan masyarakat merupakan salah satu upaya pemerintah dalam meningkatkan pelayanan publik berupa fasilitas tempat keluhan, saran, serta aspirasi warga terkait pelayanan publik, kebijakan, maupun penyelenggaraan pemerintahan [4]. Keberagaman aduan masyarakat yang disampaikan dengan bahasa non-formal atau tidak baku serta tingginya volume aduan masuk menyebabkan operator mengalami kesulitan dalam mengelompokkan dan menindaklanjuti laporan secara tepat. Untuk itu, sistem otomatisasi klasifikasi terhadap teks aduan masyarakat adalah solusi tepat untuk mengurangi resiko kesalahan dan kemudahan bagi operator dalam mengelompokkan aduan berdasarkan instansi yang berwenang [5],[6]. Model *Deep Learning Bidirectional Long Short-Term Memory* (BiLSTM) merupakan model yang cocok dalam penelitian ini karena memiliki kemampuan generalisasi yang baik terhadap data baru serta mampu mempercepat proses klasifikasi dibandingkan metode manual. Selain itu, BiLSTM mampu memahami konteks kalimat secara lebih menyeluruh karena memproses informasi dari dua arah, sekaligus mengatasi keterbatasan pada *Long Short-Term Memory* (LSTM) yang hanya mempertimbangkan aliran informasi satu arah [7].

Beberapa penelitian telah menunjukkan arsitektur BiLSTM lebih baik dari metode lainnya dalam menangani kasus klasifikasi teks, khususnya pada aduan masyarakat, jika dibandingkan dengan metode klasifikasi lainnya. Penelitian [7],[8] melakukan klasifikasi teks aduan masyarakat sebagai bagian dari pengembangan layanan e-government. Hasil penelitian menunjukkan bahwa metode BiLSTM mampu mencapai akurasi 98% pada klasifikasi dengan 4 label, serta akurasi 0,884 pada klasifikasi dengan 7 kategori. Beberapa penelitian lain dengan topik yang berbeda seperti yang dilakukan oleh [9],[10],[11] dan [12], menunjukkan bahwa BiLSTM menghasilkan kinerja yang unggul dibandingkan dengan metode lain pada kasus sentimen analisis. Penelitian [10] menggunakan *Word2Vec* sebagai representasi kata dan memperoleh akurasi sebesar 95% dengan 3 label yaitu *positive*, *neutral*, *negative*. Sementara itu, penelitian [11] menunjukkan bahwa BiLSTM memiliki performa lebih baik dibandingkan *Naïve Bayes*, dengan akurasi mencapai 94,90% dengan menggunakan 2 label yaitu positif dan negatif. Selanjutnya, penelitian [12] melakukan perbandingan antara beberapa metode, yaitu LSTM, BiLSTM, GRU, CNN, dan RNN, dan hasilnya menunjukkan bahwa BiLSTM merupakan metode paling efektif, dengan akurasi tertinggi sebesar 98,45%, dengan menggunakan 2 kategori yaitu depresif dan tidak depresif. Adapun penelitian [13] yang melakukan prediksi terhadap harga saham Bank Rakyat Indonesia (Persero) membuktikan bahwa akurasi BiLSTM termasuk kategori sangat baik dengan topik dan model evaluasi yang berbeda. Hasil dari semua penelitian tersebut mempertegas bahwa arsitektur BiLSTM merupakan pendekatan yang relevan dan baik dalam menangani permasalahan klasifikasi [7],[12] maupun dengan topik penelitian yang berbeda sekalipun [13]. Namun, permasalahan utama terletak pada klasifikasi *multiclass* dengan jumlah label kelas yang lebih kompleks serta ketidakseimbangan distribusi data (*imbalanced data*), hal ini mempengaruhi penurunan kinerja model dan permasalahan tersebut masih belum sepenuhnya teratasi [14]. Selain itu, pada setiap kategori tidak selalu memiliki kata kunci yang spesifik, beberapa jenis aduan memiliki kemiripan konteks dan kompleksitas semantik dalam menangani *multiclass* data, sehingga menyulitkan model untuk membedakan kategori yang tepat [15]. *Imbalance data* juga menjadi faktor penyebab terjadinya bias pada model dan berpotensi mengurangi kinerja model klasifikasi [16]. Hal tersebut menjadi dasar bagi peneliti untuk mengeksplorasi sekaligus menerapkan model *deep learning* BiLSTM pada klasifikasi teks *multiclass* dari aduan masyarakat yang memiliki 21 label kelas dengan kondisi distribusi data yang tidak seimbang. Penelitian ini bertujuan tidak hanya melakukan otomatisasi klasifikasi terhadap teks aduan masyarakat, tetapi juga menganalisis performa BiLSTM terhadap *multiclass* dan *imbalance* data melalui pengujian menggunakan *Confusion Matrix*.

2. METODE PENELITIAN

Pada bagian ini disajikan alur penelitian yang digunakan dalam proses klasifikasi teks aduan masyarakat. Alur tersebut dirancang secara sistematis untuk memastikan setiap tahapan pengolahan data hingga evaluasi model berjalan terstruktur dan menghasilkan model yang optimal. Penelitian ini diimplementasikan menggunakan bahasa pemrograman *Python* versi 3 dengan memanfaatkan berbagai *tools* dan *library* seperti *Pandas* dan *NumPy* untuk analisis data. Pada tahap *preprocessing* teks memanfaatkan *library NLTK* dan *Sastrawi* untuk proses *tokenization*, *stopword removal*, dan *stemming*. Gambar 1. menyajikan alur penelitian klasifikasi teks aduan masyarakat menggunakan algoritma BiLSTM.



Gambar 1. Alur Penelitian Klasifikasi Teks Aduan Masyarakat Menggunakan Algoritma BiLSTM

Secara umum, data mentah yang diperoleh pada tahap awal masih mengandung berbagai *noise*, seperti duplikasi entri, data kosong, maupun data yang tidak lengkap, sehingga diperlukan proses pembersihan data (*cleaning*) agar data lebih siap untuk dianalisis. Setelah data dinyatakan bersih, tahap selanjutnya adalah *preprocessing* data yang meliputi beberapa langkah, yaitu *case folding* untuk mengubah seluruh teks menjadi huruf kecil, kemudian kata tersebut akan dipecah dengan menggunakan *tokenization*, setelah itu *stopword removal* akan menghapus kata-kata yang umum dan tidak memiliki makna, serta *stemming* untuk mengembalikan kata ke bentuk dasarnya. Selain itu, dilakukan pula *label encoding* untuk mengubah label kategorikal menjadi bentuk numerik yang dapat diproses oleh model. Selanjutnya, data dibagi menjadi data *training* dan *testing*, kemudian dilakukan pembuatan representasi kata menggunakan metode *word embedding (Word2Vec)*. Setelah itu, data *training* akan diseimbangkan menggunakan metode *Random Oversampling* terlebih dahulu sebelum digunakan pada tahap pemodelan klasifikasi *Bidirectional Long Short-Term Memory (BiLSTM)*. Tahap terakhir adalah evaluasi model menggunakan *confusion matrix* untuk mengukur performa model berdasarkan metrik evaluasi seperti akurasi, presisi, dan *recall*, sehingga dapat diketahui sejauh mana model BiLSTM mampu melakukan klasifikasi secara efektif dan akurat terhadap *multiclass* data yang telah diproses sebelumnya.

2.1. Data Collection

Data penelitian ini diperoleh dari hasil ekspor aplikasi layanan pengaduan masyarakat di Kota Kediri. Perolehan data dilakukan melalui prosedur resmi dengan mengajukan surat permohonan kepada instansi yang berwenang, yaitu Pemerintah Kota Kediri. Dataset berisi kumpulan teks aduan masyarakat yang ditujukan kepada pemerintah daerah dengan jumlah 5.714 data. Keberagaman jenis peristiwa

dalam aduan, data ini menjadi relevan untuk dianalisis menggunakan teknik text mining guna mengklasifikasikan tujuan OPD secara cepat dan efektif dengan algoritma BiLSTM.

2.2. *Cleaning*

Tahap *Cleaning* dilakukan sebelum memasuki proses preprocessing. Pada tahap ini, teks aduan dibersihkan untuk memastikan konsistensi dan keandalan dataset. Secara umum, proses *cleaning* meliputi beberapa langkah, antara lain mengekstraksi teks yang relevan, menghapus nilai kosong (*null*), mengurangi *noise*, menyaring data yang tidak diperlukan, serta menggabungkan data dari berbagai sumber agar dataset siap digunakan pada tahap pemrosesan berikutnya [17].

2.3. *Preprocessing*

Data yang bersih dan berkualitas tinggi sangat berperan dalam meningkatkan kinerja *machine learning*, terutama dalam hal akurasi. Hal ini juga membantu algoritma bekerja lebih cepat dan efisien, mengurangi kompleksitas data, serta menghindari kesalahan dalam interpretasi hasil akhir. *Preprocessing* merupakan serangkaian teknik untuk membersihkan dan mentransformasi data [18]. Proses yang dilakukan meliputi:

a. *Case Folding*

Tahap dimana semua isi aduan disamakan dan menjadi huruf kecil. Misalnya, kata “Lapor” diubah menjadi “lapor”.

b. *Tokenize*

Tahap memotong setiap kalimat menjadi bagian-bagian atau yang disebut dengan *token*. Berikut adalah contoh proses input dan *output tokenize*.

Tabel 1. *Input dan Output Tokenize*

No	Input	Output
1	Jembatan Masuk Kelurahan Rejomulyo	["Jembatan", "Masuk", "Kelurahan", "Rejomulyo"]
2	banyak warga yang buang sampah	["banyak", "warga", "yang", "buang", "sampah"]
3	lihat sekolah yang menerima dana	["lihat", "sekolah", "menerima", "dana"]
4	tertibkan bangunan liar	["tertibkan", "bangunan", "liar"]
5	Pelayanan obat di rs gambiran kurang memuaskan	["Pelayanan", "obat", "di", "rs", "gambiran", "kurang", "memuaskan"]

c. *Stopword Removal*

Tahap ini melibatkan penghapusan kata-kata yang tidak memiliki makna signifikan atau dianggap kurang penting dengan menggunakan *library sastrawi*. Contohnya adalah kata-kata seperti “lah”, “kah”, “pun”, “tah”, “ku”, “mu”, dan “nya” yang termasuk dalam kategori *stopword* [19].

d. *Stemming*

Tahap ini merupakan proses mengubah kata yang memiliki imbuhan atau bentuk turunan menjadi bentuk dasarnya [20]. Tujuan *stemming* adalah untuk menyederhanakan kata-kata tersebut agar lebih mudah dikenali dan dianalisis, karena kata-kata dengan makna yang mirip akan direpresentasikan dalam satu bentuk dasar yang sama. Misalnya, kata “menerima” menjadi “terima” atau kata “membuang” menjadi “buang”. Pada penelitian ini menggunakan *NLTK* dan kamus *sastrawi* sebagai dukungan untuk melakukan proses *stemming*.

e. *Label Encoding*

Salah satu teknik *preprocessing* data untuk mengubah teks target menjadi bentuk numerik seperti 0,1,2,3 yang diperlukan pada proses penelitian selanjutnya. Kategori label dalam penelitian ini

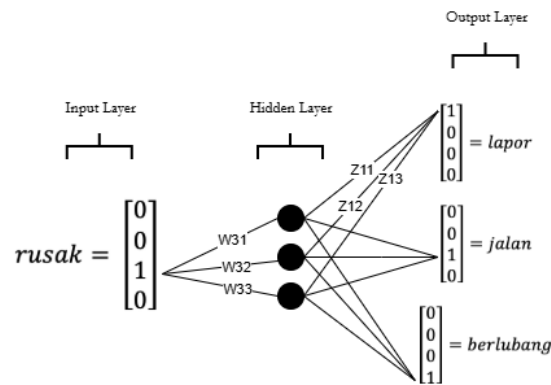
merepresentasikan unit kerja atau OPD dan memiliki kompleksitas tinggi yang bersifat *multiclass* atau memiliki lebih dari 10 label kelas.

2.4. Split Data

Data yang telah melalui proses pembobotan selanjutnya dibagi menjadi dua bagian, yaitu data pelatihan (*training*) dan data pengujian (*testing*). Data *training* digunakan untuk melatih data dan membangun model analisis terbaik, sedangkan data *testing* digunakan untuk mengevaluasi kinerja model [21]. Pada penelitian ini, proporsi pembagian data yang digunakan adalah 80% untuk data *training* dan 20% untuk data *testing*.

2.5. Word2Vec

Salah satu teknik pembobotan kata yang banyak digunakan adalah *Word2Vec*, yang mampu merepresentasikan hubungan semantik antar kata dengan baik melalui pengelompokan kata-kata serupa ke dalam vektor dengan nilai yang berdekatan [22]. Pada penelitian ini digunakan arsitektur Skip-gram karena efektif dalam memprediksi kata target berdasarkan konteks kata di sekitarnya [23]. Gambar 2. menunjukkan Arsitektur *Skip-gram* pada penelitian ini.



Gambar 2. Arsitektur *Skip-gram*

Perhitungan matematis dari pada *Skip-gram*, $X_{(wk)}$ merupakan Vektor *one-hot*, W_k dan X_i adalah Indeks posisi kata ke- i , Matriks bobot pada input *hidden layer* ($W_{V \times N}$), jumlah kata konteks yang dipilih (C) dan kolom ke- i dari matriks bobot $W'_{N \times V}$ adalah V'_{wi} terdapat persamaan matematis (1),(2),(3) sebagai berikut:

Input Layer:

$$X_{(wk)} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_v \end{bmatrix}; X_k = 1 \text{ dan } 0 \text{ Lainnya} \quad (1)$$

Hidden Layer (*Skip-gram*):

$$h = V_{(wk)}^T W_{V \times N} = V_{(wk)} \quad (2)$$

Output Layer:

$$U_{c,i} = H v'_{wi} \quad (3)$$

untuk menghitung probabilitas kemunculan suatu kata $W_{c,i}$ menjadi konteks dari kata *input* W_k . Adapun persamaan (4) sebagai berikut:

$$p(W_{c,i}|W_k) = Y_{ci} = \frac{e^{u_{c,i}}}{\sum_{t'=1}^V e^{u_{c,i'}}} \quad (4)$$

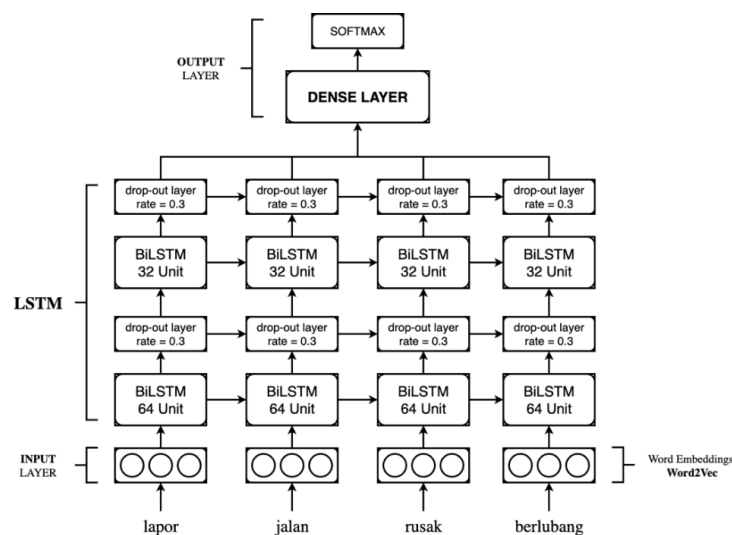
Pada penelitian ini, telah dilakukan percobaan hingga diperoleh parameter *vector_size* = 100, *window* = 5, *min_count* = 2, *workers* = 4, dan *sg* = 1 yang menunjukkan performa yang lebih baik. Metrik yang digunakan dalam menentukan konfigurasi terbaik adalah dengan melihat *validation accuracy* dan kestabilan nilai *validation loss*.

2.6. Random Oversampling

Setelah melalui tahap pembobotan serta pembagian data *training* dan *testing*, data yang dihasilkan tidak seimbang. Ketimpangan data yang terjadi mengakibatkan proses dalam mengenali kelas minoritas menjadi sulit karena bias terhadap kelas mayoritas. Maka perlu dilakukan penyeimbangan data menggunakan *Random Oversampling*. Untuk mencegah terjadinya data *leakage*, proses penyeimbangan data dilakukan hanya pada data *training*. Metode ini efektif digunakan untuk mengatasi ketidakseimbangan kelas, khususnya pada kelas minoritas yang jumlah datanya lebih sedikit [24].

2.7. Bidirectional Long Short-Term Memory (BiLSTM)

Schuster dan Paliwal mengusulkan suatu pengembangan dari algoritma *Long Short-Term Memory* (LSTM) dengan tujuan agar algoritma tersebut lebih memahami makna kata secara komprehensif. Pengembangan pada unit dari LSTM yang awalnya berjalan pada satu arah menjadi dua arah yaitu maju (*forward*) dan mundur (*backward*). Kombinasi arah dari LSTM merupakan konsep utama algoritma *Bidirectional Long Short-Term Memory* (BiLSTM) [9]. Berikut adalah Diagram Arsitektur BiLSTM:



Gambar 3. Diagram Arsitektur BiLSTM [25]

Penelitian ini menggunakan konfigurasi pelatihan berupa jumlah epoch sebanyak 20, *learning rate* sebesar 0,001. Pada tahap ini, juga diterapkan *dropout layer* dengan tujuan untuk mencegah terjadinya *overfitting* dengan rate 0,3 dan *batch size* 64 bit, serta menerapkan *class weight balancing* untuk mengatasi ketidakseimbangan data dan *early stopping* dengan *patience* 4 guna mencegah *overfitting* serta dalam pengujian menggunakan *Optimizer Adam* dan *Adagrad*. Model BiLSTM merupakan hasil pengembangan terakhir dari awal mulai *Recurrent Neural Network* (RNN) kemudian disempurnakan oleh *Long Short-Term Memory* (LSTM). Semua algoritma tersebut adalah turunan dari *Neural Network*. Adapun tiga layer utama antara lain *Input Layer*, *Bidirectional Layer* dan *Output Layer*.

a. Input Layer

Tahap awal dimulai dengan pemrosesan setiap kata yang telah melalui proses pembobotan kata dan pembagian data. Proses ini memastikan bahwa setiap kata direpresentasikan dalam bentuk vektor numerik yang menggambarkan makna dari teks aslinya.

b. Bidirectional Layer

Setiap kata yang telah melalui proses pembobotan kemudian diproses pada layer BiLSTM dengan 64 unit neuron. Secara konsep, BiLSTM bekerja dengan prinsip yang sama seperti LSTM yang memiliki tiga komponen utama, yaitu *Forget Gate* untuk menentukan informasi dari tahap sebelumnya yang perlu dihapus, dilanjutkan *Input Gate* untuk mengatur informasi baru yang akan disimpan ke memori, dan *Output Gate* untuk menentukan informasi yang akan diteruskan ke tahap selanjutnya [25]. Jumlah unit neuron menunjukkan dimensi vektor yang dihasilkan oleh layer tersebut. Penentuan jumlah unit bergantung pada ukuran dataset biasanya 16–32 unit untuk dataset kecil, 64 unit untuk dataset menengah, dan 128–256 unit untuk dataset besar.

Forward LSTM:

$$ht^{\rightarrow} = LSTM(xt, ht^{\rightarrow 1}) \quad (5)$$

Backward LSTM:

$$ht^{\rightarrow} = LSTM(xt, ht^{\rightarrow} + 1) \quad (6)$$

Output BiLSTM:

$$ht = (ht^{\rightarrow} ht) \quad (7)$$

c. *Output Layer*

Merupakan tahap akhir dari arsitektur BiLSTM yang berfungsi menghasilkan prediksi akhir berdasarkan hasil pemrosesan teks dari layer sebelumnya. Layer ini menghasilkan output berupa OPD. Proses prediksi menggunakan fungsi aktivasi *Softmax*, yang menghitung probabilitas untuk setiap label kelas. Kelas dengan nilai probabilitas tertinggi dianggap sebagai hasil prediksi yang paling relevan. *Optimizer* yang digunakan dalam penelitian ini adalah *Adam* dan *Adagrad*.

2.8. Evaluasi Model

Confusion Matriks menjadi metode pilihan untuk mengevaluasi performa dari hasil klasifikasi jenis aduan masyarakat menggunakan algoritma BiLSTM berdasarkan data sebelumnya (data *testing*). *Confusion Matrix* memiliki struktur kelas yaitu positif dan negatif dengan beberapa komponen seperti nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN), sehingga mampu membantu untuk mengidentifikasi masalah dalam pemodelan [26]. Berikut adalah tabel dan metrik yang umum digunakan.

Tabel 2. *Confusion Matrix*

	Prediksi Positif	Prediksi Negatif
Aktual Positif	TP	FN
Aktual Negatif	FP	TN

Dikarenakan penelitian ini menggunakan klasifikasi *multiclass*, maka pencarian *accuracy*, *precision*, *recall* menggunakan pendekatan *macro average* dengan persamaan sebagai berikut:

$$Macro Accuracy = \frac{1}{K} \sum_{i=1}^K \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \quad (8)$$

$$Macro Precision = \frac{1}{K} \sum_{i=1}^K \frac{TP_i}{TP_i + FP_i} \quad (9)$$

$$Macro Recall = \frac{1}{K} \sum_{i=1}^K \frac{TP_i}{TP_i + FN_i} \quad (10)$$

$$Macro\ F1 = \frac{1}{K} \sum_{i=1}^K F1_i \quad (11)$$

3. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil pengujian dari beberapa skenario dan tahapan yang telah dirancang dalam kerangka kerja awal untuk membangun model yang optimal. Penelitian ini diimplementasikan menggunakan bahasa pemrograman *Python* beserta berbagai *Library* pendukung proses analisis data.

3.1 Dataset

Data merupakan kumpulan teks aduan masyarakat yang diambil dari suatu aplikasi layanan pengaduan masyarakat di Pemerintah Kota Kediri. Data berbentuk *file excel* dengan 5.765 baris data dan 2 kolom. Berikut ini adalah tabel sampel data yang digunakan dalam penelitian ini.

Tabel 3. Dataset

No	Isi	Nama Departemen
1	selamat siang dengan ini kami mau melaporkan bahwa tutup saluran yang ada di jalan patiunus bagian sisi barat banyak yang sudah jebol atau lobang jadi sangat membahayakan bagi pengguna jalan mohon dengan sangat perhatiannya mengingat sebentar lagi musim penghujan akan datang terima kasih atas perhatiannya	Dinas Pekerjaan Umum dan Penataan Ruang
2	mohon utk bagian terkait dklh maaf nuhunsewu untuk membersihkan pohon di pinggir jalan penanggungan secara rutin pohon di depan rumah saya banyak semut angkrang sehingga merusak masuk ke dalam rumah selain itu rontokan daun dan dahannya mengotori tanaman sy d depan rmh pohon blimbing dan daun salam menjadi penyakitn gara2 kena rontokan daun pohon trsebut...	Dinas Lingkungan Hidup, Kebersihan dan Pertamanan

Adapun distribusi awal data untuk setiap kelas disajikan pada Gambar 4. Distribusi jumlah data per kelas (OPD). Gambar Diagram ini menunjukkan sebaran jumlah data pada masing-masing kelas dengan *threshol* = 10 .



Gambar 4. Distribusi jumlah data per kelas (OPD)

3.2 Exploratory Data Analysis (EDA)

Setelah data dikumpulkan, tahap selanjutnya adalah analisis awal terhadap dataset. Tahap ini bertujuan untuk memahami karakteristik isi data sehingga dapat ditentukan data yang relevan untuk dipertahankan maupun yang perlu dieliminasi. Hasil dari tahap ini menjadi dasar dalam penentuan fitur yang akan digunakan pada proses analisis selanjutnya. Berikut ini adalah karakteristik data teks aduan dalam bentuk *word cloud* dapat dilihat pada Gambar 5.



Gambar 5. *Word Cloud*

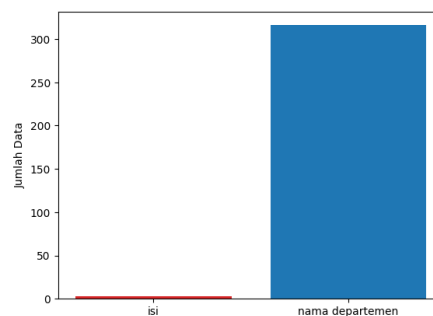
3.3 Persiapan Data

3.3.1 Tahap *Cleaning*

Tahap ini merupakan proses pembersihan data (*cleaning*) yang meliputi penghapusan data duplikat, penanganan data kosong (*null*), penghilangan spasi berlebih, serta penghapusan karakter atau simbol yang tidak diperlukan. Berikut ini merupakan proses penghapusan terhadap data null, ditunjukkan pada Tabel 4.

Tabel 4. Proses Menghapus Data Kosong (*null*)

Atribut Data	Jumlah Data		Jumlah / Persentase Data (Null)
	Sebelum	Sesudah	
Isi	5.711	5.395	3 (0.05%)
nama_departemen	5.398	5.395	316 (5.85%)

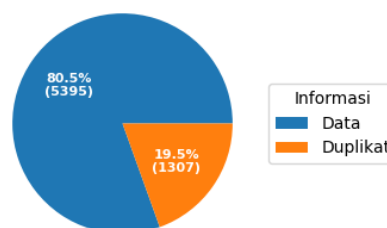


Gambar 6. Diagram *Bar Chart* Jumlah Data Kosong (*Null*)

Setelah melakukan proses penghapusan data kosong dilanjutkan pada proses penghapusan data duplikat yang ditunjukkan pada Tabel 5.

Tabel 5. Proses Menghapus Data Duplikat

Jumlah Data		Jumlah Data
Sebelum	Sesudah	Duplikat
5.395	4.090	1.307



Gambar 7. Diagram *Pie Chart* Persentase Data Duplikat

3.3.2 Tahap *Preprocessing*

Setelah data dibersihkan, tahap selanjutnya adalah data preprocessing untuk mempersiapkan data sebelum analisis lanjutan. Proses ini meliputi *case folding*, *tokenization*, *normalization* data, penghapusan kata yang termasuk kategori *stopword*, serta *stemming*. Hasil dari proses preprocessing ditunjukkan pada Tabel 6.

Tabel 6. Hasil Proses *Preprocessing*

Teks Awal	Hasil <i>Preprocessing</i>
assalamualaikum wr wb sebagai saran masukan lahan dbelakang water torn sdptnya dbenahi dfungsikn mungkin sbg taman kota sarana tempat main dsb sdptnya agar tidak dsalah gunakn hal2 negatif tindak kejahatn lbh2 pada mlm hr serta selokan jalan masuk dekat kantor kecamatn sering dgunakan orang2 untuk buang hajat besar smg masukn sumbang saran ini mnjadikn prhatian dtindak lanjuti mohon mfm trm ksh slm hrmt syng km sklga	assallamualaikum wr wb saran masukan lahan dibelakang water torn sedapatnya dibanahi difungsikan taman kota sarana main dan bagai sedapatnya tidak disalah gunakan negatif tindak kejahatan malam selokan jalan masuk kantor kecamatan digunakan orang buang hajat moga masukan sumbang saran menjadikan perhatian ditindak lanjut mohon maaf terima kasih salam hormat sayang sekeluarga
saya sudah isi form pindah dll trus karena syarat upload dokumen saya tunda dlu karena hrus persiapan ap aj yang hrus di aplot trus pkl 10 30 saya mau aplot gbs daftar lg jg gbs gimana nge cek berkas melengkapi berkas yang di aplot saya cek di kelengkapan jg yang kluar cm dftar invoice nama dn nik	isi form pindah dan lain lain syarat unggah dokumen tunda siap unggah unggah gabisa daftar gabisa nge cek berkas lengkap berkas unggah cek lengkap daftar tagih nama nik

Selanjutnya dilakukan proses pemberian label pada label kelas menggunakan *Label Encoding* yang ditunjukkan pada Tabel 7.

Tabel 7. Hasil Proses *Label Encoding*

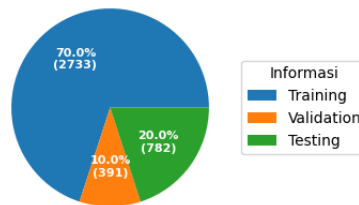
OPD / Unit Kerja	Label
Badan Daerah	0
Bagian Sekretariat Daerah	1
Dinas Kebudayaan, Pariwisata, Pemuda dan Olahraga	2
Dinas Kependudukan dan Pencatatan Sipil	3

3.4 Pembagian Data Training, Validation dan Testing

Pembagian ini menghasilkan proporsi 80:20. Selanjutnya 80% dari data *training* dibagi lagi untuk data *validation* sebanyak 10%, Hasil pembagian data dapat dilihat melalui Tabel 8.

Tabel 8. Hasil Pembagian Data

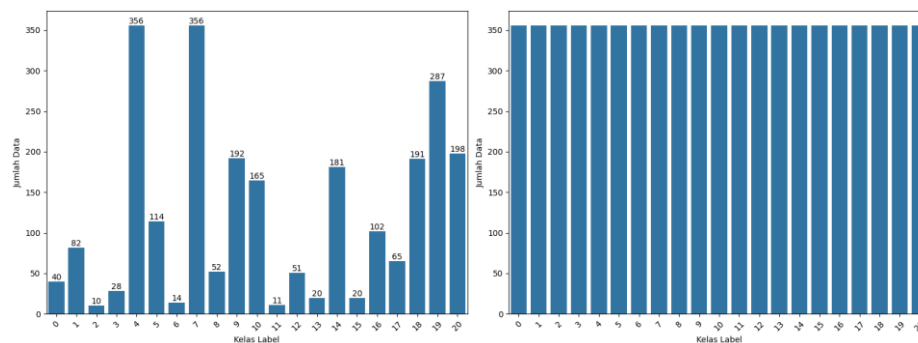
Data	Jumlah
<i>Training</i>	2.733
<i>Validation</i>	391
<i>Testing</i>	782



Gambar 8. Diagram *Pie Chart* Persentase Pembagian Data

3.5 Penanganan Data Tidak Seimbang

Berdasarkan analisis sebelumnya, teridentifikasi adanya ketidakseimbangan kelas pada dataset. Untuk mengatasinya, dilakukan penyeimbangan data menggunakan teknik *Random Oversampling* dengan konfigurasi umum yaitu *sampling_strategy* = 'auto' dan *random_state* = 42. Proses ini diterapkan hanya pada data pelatihan (*X_train* dan *y_train*). Pada Gambar 9. menunjukkan penanganan data tidak seimbang.



Gambar 9. Hasil Penanganan Data Tidak Seimbang

Setelah penyeimbangan, ukuran *X_train* menjadi (12.978, 100) dengan distribusi yang seimbang, di mana setiap kelas memiliki 618 data.

3.6 Pemodelan

Pemodelan data yang digunakan dalam penelitian ini adalah BiLSTM. Tabel 9. menunjukkan arsitektur model Bi-LSTM.

Tabel 9. Arsitektur Model *Deep Learning* (BiLSTM)

Layer (type)	Output Shape	Param #
<i>embedding (Embedding)</i>	(None, 100, 100)	360,500
<i>bidirectional (Bidirectional)</i>	(None, 128)	84,480
<i>dropout (Dropout)</i>	(None, 128)	0
<i>dense (Dense)</i>	(None, 21)	2,709

Terdapat 4 layer yaitu *embedding*, *bidirectional*, *dropout* dan *dense*. Total params 447,689 dan Trainable Params 447,689. Arsitektur ini menggunakan 64 dan 128 unit neuron pada dataset berukuran menengah. Berdasarkan penelitian [27], perlu adanya pengujian yang bervariasi dengan menggunakan jumlah unit neuron rentang 32, 64 dan 128 untuk menemukan nilai yang optimal. Selain dari hasil *trial and error*, Pemilihan jumlah 64 unit neuron didasarkan pada penelitian [28] yang melakukan klasifikasi dengan 12 label kelas dan model *single stacked*, berhasil memperoleh akurasi tinggi yaitu 80,25% dengan menggunakan *batch size* 64 unit neuron. Dari hasil penelitian tersebut menunjukkan konsistensi akurasi lebih tinggi dari model LSTM dan *hyperparameter* yang digunakan mampu memberikan kinerja model yang baik dalam proses klasifikasi.

3.7 Evaluasi

Beberapa kendala ditemukan yang mengakibatkan pengujian tidak maksimal. Kata *slang* dan variasi bahasa merupakan masalah utama pada penelitian ini, dibuktikan pada Tabel 10 hasil sesudah *preprocessing* tidak maksimal karena banyak kata *slang* yang masih lolos contohnya ‘kerjoo’, ‘ki’, ‘lo’.

Tabel 10. Hasil *Preprocessing* kurang maksimal

Tahap <i>Preprocessing</i>	
Sebelum	Sesudah
Butuh kerjoo boss, kerjoan malah mbo tutupi ki lo yo pieeee, ojol e yo sepi no akeh wong seng butuh ki Pak/ibu ngapunten nggeh.kulo bade tangklet nopo sebabu bantuan bpnt kulo koq mboten medal kados tiang lintu	butuh kerjoo kerjoan tutup ki lo ya ojol elektronik ya sepi nomor akeh seng butuh ki bapak ibu ngapunten kulo bade tangklet nopo babe bantu bpnt kulo kok mboten medal kados tiang lintu

Hal ini menunjukkan adanya keterbatasan pada tahap *preprocessing* yang belum mampu menyaring seluruh kata *slang* melalui kamus *stopword* dan proses *stemming* standar. Untuk mengatasinya, penelitian ini menambahkan detail kata yang terlewat pada kamus normalisasi khusus guna mengidentifikasi dan memperbaiki kata *slang* yang masih lolos pada tahap *preprocessing*. Adapun kendala lain yang dialami seperti kemiripan kosakata pada konteks kalimat pada label kelas yang berbeda yang ditunjukkan pada Tabel 11. Hal ini mempengaruhi pada perfoma kinerja model pada saat proses pembelajaran.

Tabel 11. Contoh data yang kurang tepat

Aduan	OPD	OPD yang benar
Ternak burung puyuh di lingkungan padat penduduk,hingga terjadinya bau kotoran yang mengganggu.bahkan rt nya pun ikut memelihara	Dinas Komunikasi dan Informatika	Dinas Lingkungan Hidup, Kebersihan dan Pertamanan
lampu penerangan jalan di depan rumah saya telah lama sekali tidak mnyala alias mati mohon perhatiannya	Dinas Perumahan dan Kawasan Pemukiman	Dinas Perhubungan

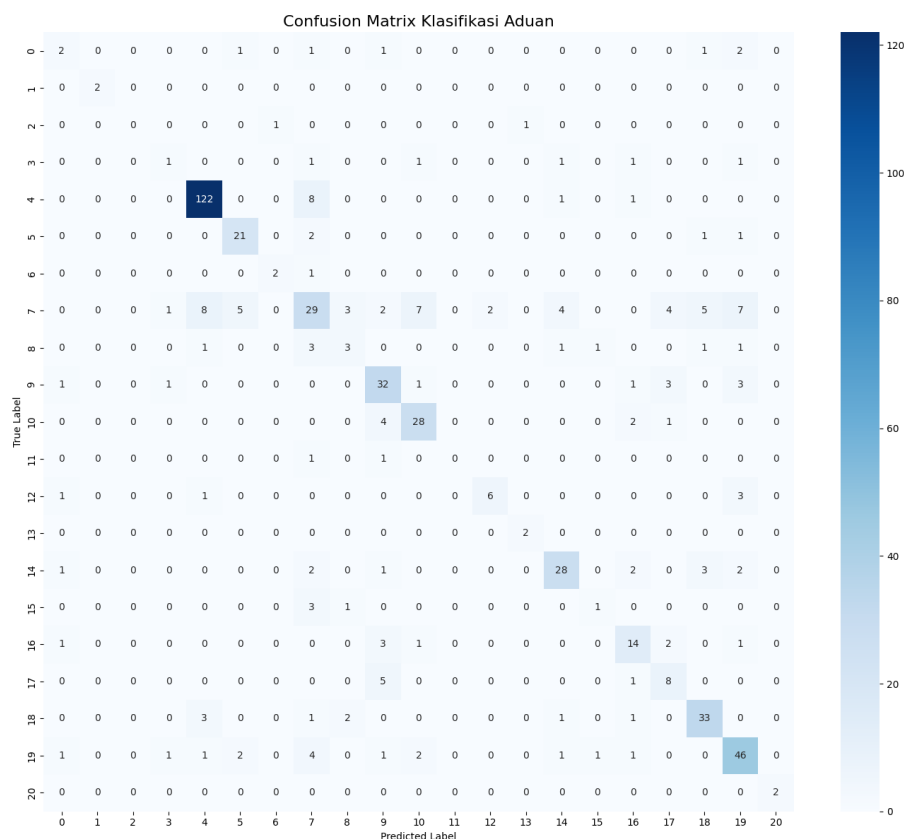
Pengujian model dilakukan sebanyak 8 kali, ditunjukkan pada Tabel 10. Hasil Evaluasi Model BiLSTM. Hasil akurasi pada Trial#5 memperoleh hasil akurasi tertinggi yaitu sebesar 0,6977 (70%) dengan menggunakan *random oversampling* dan *adam optimizer* di *batch size* 64 bit.

Tabel 12. Hasil Evaluasi Model BiLSTM

Trial	Teknik Imbalance	Optimize	Batch Size	Test Accuracy
<i>Trial #1</i>	-	Adam	64	0.6871
<i>Trial #2</i>	-	Adagrad	64	0.6359
<i>Trial #3</i>	-	Adam	128	0.6270
<i>Trial #4</i>	-	Adagrad	128	0.6409
<i>Trial #5</i>	<i>Random oversampling</i>	Adam	64	0.6977
<i>Trial #6</i>	<i>Random overampling</i>	Adagrad	64	0.6910
<i>Trial #7</i>	<i>Random oversampling</i>	Adam	128	0.6359
<i>Trial #8</i>	<i>Random oversampling</i>	Adagrad	128	0.6018

Berdasarkan pada Tabel 12, menunjukkan bahwa penggunaan Adam lebih baik daripada Adagrad. Pada penggunaan *random oversampling* dan *batch size* 64 bit, Adam berhasil memperoleh 06977 (*Trial#5*) sedangkan Adagrad lebih rendah yaitu 0.6910 (*Trial#6*). Di kondisi yang sama tetapi dengan *batch size* berbeda yaitu 128 bit, Adam tetap unggul daripada Adagrad dengan selisih 0,0341. Adagrad (*Adaptive Gradient*) merupakan pengembangan dari metode gradient descent yang menyesuaikan learning rate secara otomatis untuk setiap parameter. Sementara itu, Adam merupakan kombinasi dari

Adagrad dan *RMSProp* yang tidak hanya menyesuaikan learning rate tiap parameter, tetapi juga menjaga kestabilan selama proses pelatihan [29]. Hal ini membuktikan bahwa performa optimizer *Adam* lebih baik dibandingkan *Adagrad*. Didukung adanya penelitian [30] terkait dengan perbandingan antara *Adam*, *RmsProp*, *SGD*, *Adagrad* dan *Adadelata*. *Adam optimizer* menunjukkan kinerja terbaik dengan akurasi 86.74%. Berikut ini merupakan gambar *Confusion Matrix* dari klasifikasi teks aduan masyarakat.



Gambar 10. Confusion Matrix

Dari Gambar 10. Confusion Matrix menunjukkan bahwa model melakukan klasifikasi terhadap 21 label kelas. Beberapa kelas menunjukkan kinerja yang sangat baik, misalnya label 4 dengan 122 prediksi benar, yang menunjukkan bahwa model mampu mengenali pola aduan pada kelas tersebut dengan sangat baik. Selain itu, label 19 juga menunjukkan performa yang baik dengan 46 prediksi benar. Namun, terdapat pula kelas yang sering mengalami kesalahan klasifikasi, seperti label 2 yang hampir tidak memiliki prediksi benar, label 3 yang hanya memiliki sedikit prediksi benar, serta label 8 yang menunjukkan cukup banyak kesalahan klasifikasi ke beberapa kelas lain. Berdasarkan hasil evaluasi, kelas minoritas masih menunjukkan nilai *recall* dan *f1-score* rendah meskipun telah diseimbangkan. Hal ini menunjukkan bahwa *random oversampling* masih belum mampu meningkatkan kemampuan model dalam mengenali pola khususnya pada kelas minoritas. Alternatif lain dapat digunakan seperti *Synthetic Minority Over-sampling technique* (SMOTE) untuk menambah variasi data, *Adaptive Synthetic Minority* (ADASYN) yaitu pengembangan dari smote fokus pada sampel minoritas yang lebih sulit diklasifikasikan [31] atau *class weight* dengan tidak menambah data tetapi memberi bobot lebih besar pada kelas minoritas [32].

Tabel 13. Hasil Pengujian Setiap Label Kelas

No	OPD	Label	Precision	Recall	F1-Score
1	Badan Daerah	0	0.33	0.25	0.29
2	Bagian Sekretariat Daerah	1	0.67	1.00	0.80
3	Dinas Kearsipan dan Perpustakaan	2	0.00	0.00	0.00

No	OPD	Label	Precision	Recall	F1-Score
4	Dinas Kebudayaan, Pariwisata, Pemuda & Olahraga	3	0.50	0.33	0.40
5	Dinas Kependudukan dan Pencatatan Sipil	4	0.94	0.91	0.92
6	Dinas Kesehatan	5	0.71	0.80	0.75
7	Dinas Ketahanan Pangan dan Pertanian	6	0.50	0.33	0.40
8	Dinas Komunikasi dan Informatika	7	0.48	0.36	0.41
9	Dinas Koperasi, Usaha Mikro dan Tenaga Kerja	8	0.40	0.18	0.25
10	Dinas Lingkungan Hidup, Kebersihan dan Pertamanan	9	0.56	0.83	0.67
11	Dinas Pekerjaan Umum dan Penataan Ruang	10	0.74	0.71	0.72
12	Dinas Pemberdayaan Perempuan, Perlindungan Anak, Pengendalian Penduduk & Keluarga Berencana	11	0.50	0.50	0.50
13	Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu	12	0.70	0.64	0.67
14	Dinas Pendapatan	13	0.67	1.00	0.80
15	Dinas Pendidikan	14	0.71	0.64	0.68
16	Dinas Perdagangan dan Perindustrian	15	0.67	0.40	0.50
17	Dinas Perhubungan	16	0.59	0.59	0.59
18	Dinas Perumahan dan Kawasan Pemukiman	17	0.47	0.64	0.55
19	Dinas Sosial	18	0.65	0.85	0.74
20	Instansi Lainnya	19	0.73	0.72	0.73
21	Kelurahan dan Kecamatan	20	0.67	1.00	0.80
Macro Avg			0.58	0.60	0.58
Weighted Avg			0.69	0.70	0.69

Berdasarkan Tabel 13, kelas dengan performa terendah adalah Dinas Kearsipan dan Perpustakaan yang memiliki nilai *precision*, *recall*, dan *f1-score* sebesar 0,00, yang menunjukkan bahwa model tidak berhasil mengklasifikasikan data pada kelas tersebut karena hanya memiliki 10 sampel *training* dari total 15 sampel setelah melalui proses pembagian data *training* yang mana hasil tersebut termasuk kelas minoritas. Selain itu, tidak ada kata kunci spesifik seperti “arsip”, penggunaan istilah yang bersifat umum seperti “layanan publik”, serta singkatan seperti “perpusda” yang belum teridentifikasi pada tahap *preprocessing*, menyebabkan model kesulitan mengenali pola pada kelas tersebut. Kondisi ini berdampak pada rendahnya performa, bahkan menghasilkan nilai *F1-score* sebesar 0,00. Kelas dengan performa menengah adalah Dinas Pekerjaan Umum dan Penataan Ruang dengan nilai *precision* 0,74, *recall* 0,71, dan *f1-score* 0,72, yang menunjukkan bahwa model cukup mampu mengidentifikasi aduan pada kategori tersebut. Sementara itu, performa tertinggi diperoleh pada kelas Dinas Kependudukan dan Pencatatan Sipil dengan nilai *precision* 0,94, *recall* 0,91, dan *f1-score* 0,92. Hal ini dikarenakan distribusi data pada kelas tersebut relatif lebih banyak sehingga model dapat mempelajari pola teks aduan dengan sangat baik. Selain itu, Bagian Sekretariat Daerah, Dinas Pendapatan, Kelurahan, dan Kecamatan memperoleh nilai *recall* 1,00, yang menunjukkan bahwa seluruh data pada kelas tersebut berhasil terdeteksi oleh model, meskipun nilai *precision* belum mencapai nilai maksimum. Metode BiLSTM dinilai mampu melakukan klasifikasi aduan masyarakat dengan 21 label kelas dan mencapai akurasi sebesar 0,6977, yang menunjukkan performa yang cukup baik. Hanya saja, Model ini masih memiliki keterbatasan dalam menangani penggunaan bahasa daerah dan kata *slang*, mengingat masyarakat di wilayah Jawa Timur cenderung menggunakan campuran bahasa daerah dan bahasa tidak baku dalam menyampaikan aduan. Kondisi tersebut dapat memengaruhi kualitas representasi teks serta berdampak pada penurunan kinerja model dalam proses klasifikasi.

Pada penelitian ini memiliki nilai akurasi yang lebih rendah dibandingkan penelitian terdahulu yang dilakukan oleh peneliti [7] yang menggunakan 7 kelas dengan hasil akurasi 98% serta peneliti [11] yang menggunakan 2 kelas menghasilkan akurasi sebesar 94,90%. Hal ini dikarenakan penelitian ini mengimplementasikan klasifikasi *multiclass* dengan 21 kelas sehingga menghasilkan akurasi lebih rendah yakni 70% mengingat tantangan yang dihadapi jauh lebih kompleks dibandingkan dengan pendekatan klasifikasi pada penelitian sebelumnya.

4. KESIMPULAN

Secara keseluruhan, penerapan model BiLSTM pada penelitian ini dapat dikategorikan memadai sebagai *baseline* model dengan memperoleh hasil akurasi sebesar 0,6977 (70%). Model terbukti mampu melakukan klasifikasi dengan baik pada kondisi ketika data pelatihan tersedia dalam jumlah yang cukup. Sebaliknya, performa model menurun secara signifikan pada kelas dengan jumlah data terbatas dan dalam kondisi data tidak seimbang (*imbalance data*). Adapun saran untuk peneliti selanjutnya adalah mempertimbangkan penggunaan model berbasis *transformer*, seperti IndoBERT yang lebih mampu untuk menangani kalimat *noisy* misalnya typo, slang, singkatan, atau struktur tidak baku dengan cara memecah menjadi *subword*, atau mengkombinasikannya dengan pendekatan yang telah digunakan pada penelitian ini. Pendekatan tersebut berpotensi meningkatkan kualitas representasi konteks bahasa sehingga kinerja klasifikasi pada kasus *multiclass* dengan kompleksitas tinggi dapat lebih optimal.

5. DAFTAR PUSTAKA

- [1] D. Prihatono, "Penerapan Pelayanan Publik Berbasis E-Governance Pada Era Revolusi Industri 4.0," *Jurnal Ilmu Administrasi*, vol. 14, no. 2, pp. 192–201, Jun. 2023.
- [2] P. A. Lestari, A. Tasyah, A. Syofira, C. A. Rahmayani, R. D. Cahyani, and N. Tresiana, "Inovasi Pelayanan Publik Berbasis Digital (E-Government) Di Era Pandemi COVID-19," *Jurnal Ilmu Administrasi*, vol. 18, no. 2, pp. 212–224, Dec. 2021.
- [3] R. I. Arti, F. Dewi, and W. A. Nurtrisha, "Analisis Dan Perancangan Enterprise Architecture Pada Badan Pendapatan Daerah DKI Jakarta Untuk Mendukung Implementasi SPBE," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 3, pp. 2550–2560, Sep. 2025, doi: 10.29100/jupi.v10i3.6620.
- [4] P. Parisca, K. Khairul, and Z. Syahputra, "Platform Layanan Pengaduan Masyarakat Berbasis Web Pada Kantor Desa Helvetia," *Jurnal Minfo Polgan*, vol. 14, no. 2, pp. 1550–1557, Jul. 2025, doi: 10.33395/jmp.v14i2.15065.
- [5] M. Azis Suprayogi and R. Adrianti Supono, "Klasifikasi Helpdesk Menggunakan Metode K-Nearest Neighbor dan TF-ABS," *Techno.COM*, vol. 20, no. 4, pp. 508–517, Nov. 2021, doi: 10.33633/tc.v20i4.5094.
- [6] M. A. Rasyida and M. F. Rizal, "Optimalisasi Klasifikasi Disposisi Pengaduan Masyarakat melalui Kanal LAPOR Menggunakan Algoritma Naïve Bayes dan Integrasi Ekstensi Chrome," *Integrative Perspectives of Social and Science Journal*, vol. 2, no. 3, pp. 4115–4121, Jun. 2025.
- [7] P. Pan and Y. Chen, "Automatic Subject Classification of Public Messages in E-government Affairs," *Data Inf. Manag.*, vol. 5, no. 3, pp. 336–347, Sep. 2021, doi: 10.2478/dim-2021-0004.
- [8] A. Oktarino, S. Defit, and Yuhandri, "Improving Accuracy Performance on CNN and BiLSTM Algorithms for Classification of Public Complaints Against Police Through WhatsApp," in *2024 8th International Conference on Information Technology, Information Systems and Electrical Engineering, ICITISEE 2024*, Institute of Electrical and Electronics Engineers Inc., Aug. 2024, pp. 399–404. doi: 10.1109/ICITISEE63424.2024.10730082.
- [9] A. Fatimah, B. Badieah, and S. F. C. Haviana, "Implementasi Bidirectional Long Short-Term Memory untuk Identifikasi Entitas Saham," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 11, no. 1, pp. 78–90, Apr. 2025, doi: 10.28932/jutisi.v11i1.9775.
- [10] S. Mutmainah, D. H. Fudholi, and S. Hidayat, "Analisis Sentimen dan Pemodelan Topik Aplikasi Telemedicine Pada Google Play Menggunakan BiLSTM dan LDA," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 1, p. 312, Jan. 2023, doi: 10.30865/mib.v7i1.5486.
- [11] H. Ma', A. Yudo Husodo, and B. Irmawati, "PERFORMANCE COMPARISON OF NAIVE BAYES AND BIDIRECTIONAL LSTM ALGORITHMS IN BSI MOBILE REVIEW SENTIMENT ANALYSIS," *Jurnal Teknik Informatika (JUTIF)*, vol. 6, no. 1, pp. 159–172, Feb. 2025, doi: 10.52436/1.jutif.2025.6.1.4178.
- [12] A. Muhammad Huda, G. F. Shidik, and V. Praskatama, "COMPARATIVE ANALYSIS OF LSTM, BiLSTM, GRU, CNN, AND RNN FOR DEPRESSION DETECTION IN SOCIAL MEDIA," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 6, pp. 1723–1735, Dec. 2024, doi: 10.52436/1.jutif.2024.5.6.4060.
- [13] R. Darja, M. Irwan, F. Sains dan Teknologi, U. Islam Negeri Alauddin Makassar Gowa, and S. Selatan, "Memprediksikan Harga Saham Bank Rakyat Indonesia (Persero) Tbk Menggunakan

- Metode Bidirectional Long Short Term Memory (Bi-LSTM),” vol. 10, no. 2, 2025, [Online]. Available: <https://journal.uin-alauddin.ac.id/index.php/instek>
- [14] O. M. Alyasiri and Y. N. Cheah, “Multi-Class Text Classification using Machine Learning Techniques,” *Engineering, Technology and Applied Science Research*, vol. 15, no. 3, pp. 22598–22604, Jun. 2025, doi: 10.48084/etasr.9994.
- [15] A. S. Rizky and E. Y. Hidayat, “Klasifikasi Emosi Pada Teks Bahasa Indonesia Menggunakan IndoBERT,” *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 4, pp. 2963–2970, Nov. 2025, doi: 10.29100/jipi.v10i4.6602.
- [16] D. A. Sani and M. Z. Sarwani, “A Random Oversampling and BERT-based Model Approach for Handling Imbalanced Data in Essay Answer Correction,” *INFOTEL*, vol. 16, no. 4, pp. 729–739, Nov. 2024, doi: 10.20895/INFOTEL.V16I3.1224.
- [17] M. Wahyu, A. Pramana, D. Purnami, S. Putri, I. Ketut, and A. Purnawan, “Comparison of IndoBERT and Bi-LSTM Models for Indonesian Law Violation Text Classification,” *Jurnal Informatika: Jurnal pengembangan IT*, vol. 10, no. 4, pp. 2477–5126, 2025, doi: 10.30591/jpit.v10i4.8795.
- [18] P. Koukaras and C. Tjortjis, “Data Preprocessing and Feature Engineering for Data Mining: Techniques, Tools, and Best Practices,” *AI*, vol. 6, no. 10, pp. 1–40, Oct. 2025, doi: 10.3390/ai6100257.
- [19] M. Alkaff, A. Baskara Rizky, and I. Maulani, “Klasifikasi Laporan Keluhan Pelayanan Publik Berdasarkan Instansi Menggunakan Metode LDA-SVM,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 6, pp. 1265–1276, Dec. 2021.
- [20] J. Saputra, L. Maryani, Rahmadden, D. Wulandari, and W. Eka, “Analisis Performa Naive Bayes dan SVM terhadap Sentimen Teks Media Sosial dengan Word2Vec dan SMOTE,” *Informatika Sains dan Teknologi (INSTEK)*, vol. 10, no. 1, pp. 143–155, Apr. 2025, [Online]. Available: <https://journal.uin-alauddin.ac.id/index.php/instek>
- [21] M. Rafly Gusmansyah, H. Hendrawan, Rahmadden, and Rohid, “Peningkatan Kinerja Analisis Sentimen pada Ulasan Aplikasi Identitas Kependudukan Digital (IKD) di Indonesia Menggunakan Algoritma Support Vector Machine (SVM) dan SMOTE,” *Informatika Sains dan Teknologi (INSTEK)*, vol. 10, no. 1, pp. 185–196, Apr. 2025, [Online]. Available: <https://journal.uin-alauddin.ac.id/index.php/instek>
- [22] D. Intan Af'idah, S. Febbi Handayani, R. Wijayatun Pratiwi, and Dairoh, “Pengaruh Parameter Word2Vec terhadap Performa Deep Learning pada Klasifikasi Sentimen,” *Jurnal Pengembangan IT (JPIT)*, vol. 6, no. 3, pp. 156–161, Nov. 2021.
- [23] I. Yanti and E. Utami, “SENTIMENT ANALYSIS OF INDONESIA’S CAPITAL RELOCATION USING WORD2VEC AND LONG SHORT-TERM MEMORY METHOD,” *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 1, pp. 149–158, Sep. 2024, doi: 10.52436/1.jutif.2025.6.1.2712.
- [24] H. Argretya and K. Kusnawi, “Impelementasi Random Oversampling pada Algoritma Random Forest untuk Klasifikasi Indeks Standar Udara di DKI Jakarta 2024,” *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 10, no. 4, pp. 1189–1198, Sep. 2025, doi: 10.30591/jpit.v10i4.9466.
- [25] P. Zahwa, S. Agustian, F. Yanto, and Novriyanto, “IMPLEMENTASI BI-DIRECTIONAL LONG SHORT TERM MEMORY TERHADAP KLASIFIKASI SENTIMEN DI TWITTER PADA DATASET TERBATAS,” *Jurnal Sistem Informasi (ZONAsi)*, vol. 7, no. 1, pp. 11–5, Jan. 2025.
- [26] D. V. Setiawan and U. D. Iswavigra, “Sentiment Analysis to Evaluate Public Service Perception among Surakarta City Residents Using the BiLSTM Model,” *Journal Of Informatics and Telecommuniacion Engineering (JITE)*, vol. 9, no. 1, pp. 229–239, Jul. 2025, doi: 10.31289/jite.v9i1.15498.
- [27] N. H. A. Rahman, Q. P. Yin, and H. S. Zulkafli, “OPTIMISING LSTM AND BILSTM MODELS FOR TIME SERIES FORECASTING THROUGH HYPERPARAMETER TUNING,” *Journal of Quality Measurement and Analysis*, vol. 21, no. 3, pp. 191–205, Sep. 2025, doi: 10.17576/jqma.2103.2025.12.
- [28] E. Aurora Az Zahra, Y. Sibaroni, and S. Suryani Prasetyowati, “Classification of Multi-Label of Hate Speech on Twitter Indonesia using LSTM and BiLSTM Method,” *JINAV: Journal of*

- Information and Visualization*, vol. 4, no. 2, pp. 170–178, Jul. 2023, doi: 10.35877/454ri.jinav1864.
- [29] M. K. Najib, A. Irawan, F. N. Salsabilla, and S. Nurdianti, “Performance Comparison of Gradient-based Optimizer for Classification of Movie Genres,” *Indonesian Journal of Mathematics and Applications*, vol. 3, no. 1, pp. 1–18, Apr. 2025, doi: 10.21776/ub.ijma.2025.003.01.1.
- [30] M. R. Raharjo, Indra Riyana Rahadjeng, M. N. H. Siregar, and P. Alkhairi, “Optimasi Fungsi Aktivasi pada Artificial Neural Network untuk Prediksi Gagal Jantung Secara Akurat,” *Explorer (Hayward)*, vol. 5, no. 1, pp. 42–51, Jan. 2025, doi: 10.47065/explorer.v5i1.1840.
- [31] A. H. Putra and A. Salam, “A Comparative Performance of SMOTE, ADASYN and Random Oversampling in Machine Learning Models on Prostate Cancer Dataset,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 3, pp. 603–610, Jun. 2025, [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [32] J. Homepage *et al.*, “Implementation of the XGBoost Classifier Algorithm for Classifying Crowds of Hajj and Umrah Pilgrims,” *Indonesian Journal of Machine Learning and Computer Science (MALCOM)*, vol. 6, no. 1, pp. 106–115, Jan. 2026, doi: 10.57152/malcom.v6i1.2245.