I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol.16 No.1, Mei 2024
eISSN : 2580-3042
pISSN : 1979-0694

DETEKSI PENIPUAN PADA SOSIAL MEDIA TWITTER DENGAN METODE BIDIRECTIONAL LONG SHORT TERM MEMORY (BI-LSTM)

Hansen Dusenov[1], Antoni Wibowo[2]

[1,2,]Computer Science, Binus Graduate Program-Master of Computer Science
Universitas Bina Nusantara
[1,2,]Jl. Kebon Jeruk, Jakarta Barat, 11530, Indonesia
e-mail : [1]hansen.dusenov@binus.ac.id , [2]anwibowo@binus.edu

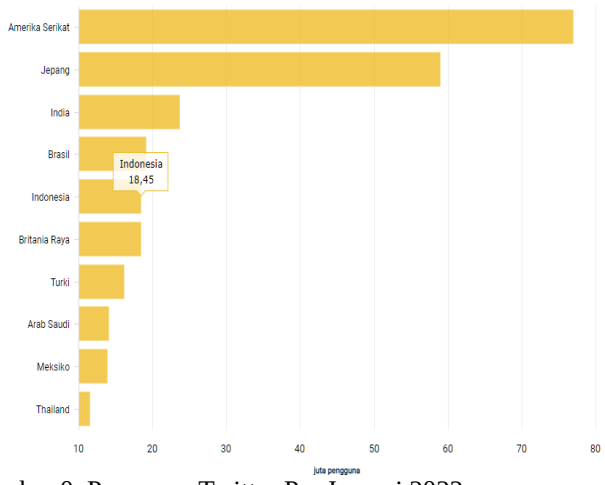
ABSTRAK

This research aims to address the issue of online fraud detection in Indonesia through the implementation of the Bidirectional Long Short Term Memory (BI-LSTM) method on the Twitter social media platform. Adopting a descriptive research approach, the study seeks to comprehend user behavior, interaction patterns, and sentiments expressed on Twitter without manipulating the studied variables. Data collection involves utilizing APIs and Web Crawlers to gather information regarding online behavior. The evaluation results indicate that the BI-LSTM model outperforms the LSTM model in detecting fraudulent and non-fraudulent transactions. The BI-LSTM model demonstrates higher precision, recall, and accuracy, showcasing its superior ability to identify genuine fraudulent transactions and avoid prediction errors. These evaluation outcomes are reinforced by training and validation graphs, illustrating that the model has reached its peak performance in learning from the available training data. The conclusion drawn from this research underscores the importance of understanding the common characteristics of online fraud, utilizing the Indonesian language, and employing relevant keywords during dataset collection to develop an effective deep learning model for online fraud detection. Furthermore, employing appropriate validation methods, periodic performance evaluations, hyperparameter tuning, and dataset adjustments are crucial steps in optimizing the outcomes of the developed model. The Early Stopping technique can also be utilized to halt training when the model no longer demonstrates significant performance improvements, thereby conserving computational resources and ensuring focus on the most optimal model.

Kata kunci *: Fraud Detection; BILSTM Model; Cyber Security; Machine Learning; Social Media Fraud*

1. PENDAHULUAN

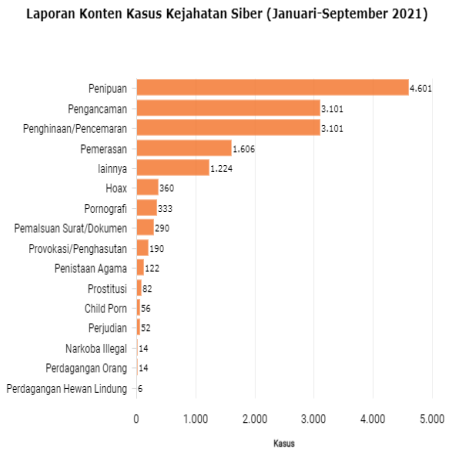
Sosial media sudah menjadi bagian dari kehidupan sehari – hari masyarakat modern saat ini. Banyak sekali pilihan platform sosial media diantaranya : Facebook, Instagram, Twitter dan Linkedin (Singh et al., 2019). Twitter merupakan salah satu media sosial yang cukup popular di Indonesia (Muzakir et al., 2022), merujuk pada Gambar 1.1 berdasarkan laporan per Januri 2022 dari statista Indonesia menduduki peringkat ke 5 sebagai negara pengguna Twitter terbanyak dengan jumlah 18,45 juta (Annur, 2022).



Gambar 0. Pengguna Twitter Per Januari 2022

I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol.16 No.1, Mei 2024
eISSN : 2580-3042
pISSN : 1979-0694

Dengan jumlah pengguna yang cukup banyak, twitter rentan akan tindak kejahatan siber. Menurut Girasa kejahatan siber adalah sebuah aksi yang dimana menggunakan teknologi sebagai komponen utama dalam melakukan tindak kejahatan (Girasa, 2020). Berdasarkan data Kata Data, Penipuan online merupakan kasus dengan nomor urut ke 1 paling banyak dilaporkan pada periode Januari – September 2021 (Pusparisa, 2020), seperti pada Gambar 1.2 yang menunjukan grafik pelaporan tindak kejahatan siber.



Gambar 2. Pelaporan Tindak Kejahatan Siber Januari - September 2021

Menurut Riset Nasional yang dilakukan terdapat 5 jenis penipuan yang paling banyak diterima responden (Finaka et al., 2022). Diantaranya seperti pada Tabel 1.1

Tabel 0. Jenis penipuan terbanyak

Jenis Penipuan	Persentase
Penipuan berkedok hadiah	91,2%
Pinjaman digital ilegal	74,8%
Tautan malware atau virus	65,2%
Berkedok krisis keluarga	59,8%
Investasi illegal	56%

Kerugian yang diakibatkan oleh penipuan online di Indonesia diperkirakan terdapat sebesar 18,7 triliun selama periode 2017 – 2021. Salah satu faktor yang menyebabkan terjadinya penipuan online adalah karena kurangnya literasi digital yang menyebabkan kurangnya kewaspadaan terhadap penggunaan sosial media (Prayuti, 2023).

Salah satu contoh penipuan online yang terjadi adalah adanya oknum yang berpura-pura sebagai akun call center bank palsu. Oknum tersebut membuat sebuah akun palsu dan menggunakan nama sedemikian rupa untuk menipu nasabah.

Berdasarkan penelitian yang sudah pernah dilakukan sebelumnya oleh Muharomah (2023) dengan menggunakan 9 label dan 4000 unlabel contoh data, didapat akurasi sebesar 87% untuk mendeteksi penipuan di twitter yang berada di new jersey (Muharomah & Ratnasari, 2023).

Penelitian lain yang menggunakan twitter sebagai sumber data pernah dilakukan oleh Abdulloh dan rekan dalam mendeteksi cyber bullying (Khairunnisa & Pithaloka, 2023).

Dalam era digital saat ini, pengolahan dan analisis data menjadi kunci untuk mendapatkan wawasan yang berharga. Salah satu bentuk data yang sering dihadapi adalah urutan data atau data temporal, seperti time series. Penerapan jaringan saraf rekurensi (RNN), khususnya arsitektur Long Short-Term Memory (LSTM), telah menjadi landasan yang efektif dalam memhami dan memproses urutan data yang panjang (Shrestha & Pradhanang, 2023).

Long Short-Term Memory (LSTM), sebagai modifikasi dari RNN hadir untuk mengatasi tantangan utama yang dihadapi oleh RNN, yaitu masalah hilangnya atau meledaknya gradien dan kesulitan pelatihan pada urutan data yang panjang. Meskipun Long Short-Term Memory (LSTM) telah membuktikan efektivitasnya dalam menangani masalah hilangnya atau meledaknya gradien pada jaringan saraf rekurensi (RNN), terdapat tantangan dalam pemodelan ketergantungan berurutan yang melibatkana informasi dari kedua arah dalam urutan data. Maka muncul Bi-Directional Long Short-Term Memory (Bi-LSTM) yang memperluas konsep LSTM dengan memasukkan koneksi rekuren pada kedua arah, yaitu maju dan mundur (Zhao et al., 2020).

Dengan menggunakan Bi-LSTM, diharapkan dapat diperoleh model yang lebih adapatif dan mampu mengidentifikasi pola penipuan yang kompleks dalam urutan data dari berbagai arah. Penelitian ini dapat membuka jalan untuk pengembangan solusi yang lebih canggih dan handal dalam mendeteksi penipuan pada platform media sosial yang dinamis dan kompleks seperti Twitter. Penelitian ini akan menggunakan dataset yang didapat dari Twitter pada rentang waktu Januari – Agustus 2023 dan dibatasi dengan wilayah yang ada di Indonesia.

2. METODE PENELITIAN

A. Jenis Penelitian

Penelitian ini mengusung pendekatan penelitian deskriptif dengan tujuan agar dapat memahami dengan mendalam tentang fenomena yang terjadi pada sosial medial Twitter (Creswell, 2017). Dengan menggunakan pendekatan yang deskriptif, penelitian ini dapat menggambarkan secara akurat perilaku dari para pengguna, pola interaksi dan sentimen yang muncul tanpa adanya melakukan manipulasi

I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol.16 No.1, Mei 2024
eISSN : 2580-3042
pISSN : 1979-0694

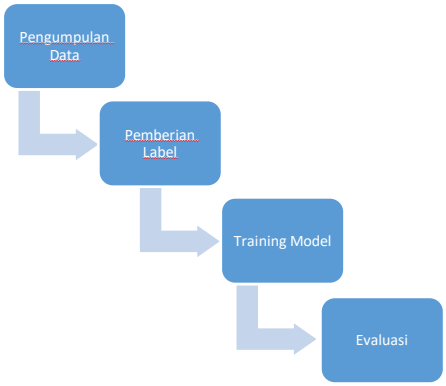
terhadapat variabel yang diteliti.

Penelitian deksriptif juga sesuai untuk mengungkapkan pola-pola yang muncul dalam sebuah data sosial media yang kompleks, seperti variasi dari gaya penulisan, perbedaan budaya, dan reaksi yang bervariasi terhadap sebuah topik. Dalam hal penipuan penilitian deskriptif, memungkinkan dapat memberikan pemahaman yang lebih baik terhadap konteks dan variabel yang berpotensi menjadi indikator penipuan.

B. Tahapan Penelitian

Penelitian ini terdapat serangkaian tahapan esensial yang membimbing keseluruhan proses analisis. Pertama-tama, penelitian ini diawali dengan proses pengumpulan data yang dimana informasi terkait perilaku online diperoleh melalui pemanfaatan API dan Web Crawler. Selanjutnya, tahap pemberian label menjadi langkah penting, dimana algoritma klasifikasi otomatis digunakan untuk membedakan antara data yang merepresentasikan kegiatan normal dan yang mencurigakan.

Proses selanjutnya adalah tahap pelatihan model, dimana model akan dilatih menggunakan arsitektur LSTM dan BiLSTM, disesuaikan dengan data yang telah diberi label. Pada Tahap ini penyesuaian bertujuan untuk mengidentifikasi pola dan tren terkait potensi penipuan (Gao et al., 2019). Terakhir penelitian ini mencapai tahapan evaluasi, dimana model di uji menggunakan data testing. Evaluasi ini melibatkan pengukuran akurasi, presisi, recall, dan F1-score untuk menilai sejauh mana kemampuan model dalam mendeteksi penipuan, serta untuk menjamin hasil dari analisis yang handal dan relevan. Dengan adanya tahapan – tahapan yang dilakukan, penelitian ini menghadirkan metodologi yang kuat untuk mendeteksi permasalahan penipuan dengan pendekatan yang holistik dan terstruktur. Alur tahapan ini dapat dilihat pada Gambar 3.



Gambar 0. Alur Penelitian

C. Pengumpulan Data

Pengumpulan data diperoleh dari situs sosial media Twitter, dengan menggunakan Teknik web crawler. Teknik memanfaatkan fitur pencarian yang ada pada twitter lalu diambil menggunakan python dengan mengkombinasikan keyword yang berbeda pada setiap proses pengambilan data. Dari tahapan tersebut data yang digunakan terkumpul dari Januari 2023 – Agustus 2023 dengan total 6655 row data. Contoh data yang dikumpulkan dapat dilihat pada Gambar 4

	id_str	full_text
0	1,69368E+18	Angka Result Pasaran KENTUCKYMID Hari Ini : Se...
1	1,69368E+18	Angka Result Pasaran KENTUCKYMID Hari Ini : Se...
2	1,69368E+18	Hasil Pengeluaran Pasaran KENTUCKYMID Hari Sel...
3	1,69368E+18	Angka Result Pasaran KENTUCKYMID Hari Ini : S...
4	1,69368E+18	@JUJA_KOT @Acoco_ke Mimi hujiambiaga jackpot n...
...	...	...
1438	1,69472E+18	Watched #GunsAndGulaabs and it's mind boggling...
1439	1,69472E+18	This no one cant tipu. Exo got the package. ht...
1440	1,69472E+18	jelas, lepas dari orgil tukang tipu bahagia ba...
1441	1,69472E+18	🔥 BRCLife Giveaway ! 🔥 🏆 Rewards Pools: $99.96 ...
1442	1,69472E+18	Malam jumat Thailand open vcs bugil ini real y...

[1443 rows x 2 columns]

Gambar 4 Data yang dikumpulkan Pemberian Label

Pada tahap ini, dilakukan pemberian label otomatis pada dataset menggunakan metode auto-labeling berdasarkan analisis sentimen menggunakan polarity. Proses ini bertujuan untuk mengkategorikan setiap entri yang digunakan sebagai “penipuan” dan “non penipuan” berdasarkan sentiment yang terdeteksi. Tahapan yang dilakukan adalah menerapkan analisis sentiment menggunakan metode analisis. sentiment. polarity pada teks entri data set. Metode ini mengukur polaritas sentiment dari teks, dengan nilai positif menunjukan sentimen positif, nilai negatif menunjukan sentimen negatif, dan nilai nol menunjukkan sentimen netral.

Selanjutnya menentukan ambang batas polarity yang digunakan untuk memisahkan entri yang dianggap sebagai “penipuan” dan “non penipuan”. Misalnya entri dengan nilai polarity di atas ambang positif dianggap sebagai “non penipuan”, sedangkan nilai di bawah ambang negative dianggap sebagai “penipuan”. Berdasarkan polarity, setiap entri dataset diberi label otomatis sebagai “penipuan” atau “non penipuan”. Proses ini dilakukan secara otomatis tanpa intervensi manusia. Berikut adalah data yang sudah diberikan label secara otomatis pada gambar 5

INFORMATIKA
Jurnal Informatika, Manajemen dan Komputer, Vol.16 No.1, Mei 2024
eISSN : 2580-3042
pISSN : 1979-0694

	full_text	label
0	Angka Result Pasaran KENTUCKYMID Hari Ini : Se...	penipuan
1	Angka Result Pasaran KENTUCKYMID Hari Ini : Se...	penipuan
2	Hasil Pengeluaran Pasaran KENTUCKYMID Hari Sel...	penipuan
3	Angka Result Pasaran KENTUCKYMID Hari Ini : S...	penipuan
4	@JUJA_KOT @Acoco_ke Mimi hujiambiaga jackpot n...	penipuan
...	...	...
1438	Watched #GunsAndGulaabs and it's mind boggling...	non penipuan
1439	This no one cant tipu. Exo got the package. ht...	non penipuan
1440	jelas, lepas dari orgil tukang tipu bahagia ba...	non penipuan
1441	🔥 BRCLife Giveaway ! 🔥 🏆 Rewards Pools: $99.96 ...	non penipuan
1442	Malam jumat Thailand open vcs bugil ini real y...	non penipuan

[1443 rows x 2 columns]

Gambar 5 Pemberian Label

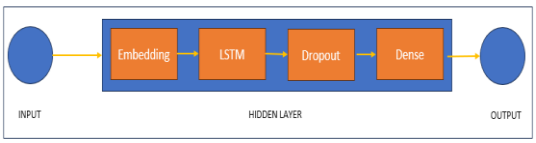
Proses pemberian label ini menggunakan Pustaka yang ada pada python bernama TextBlob. TextBlob adalah sebuah Pustaka di python yang digunakan untuk Natural Language Processing (NLP). Pustaka ini memungkinkan untuk melakukan berbagai hal seperti tokenisasi, part-of-speech tagging, parsing, analisis sentiment, penerjemahan dan lainnya. Cara kerja untuk polarity dalam TextBlob menggunakan beberapa tahapan, diantaranya :

1. Tokenisasi : membagi teks menjadi frasa-frasa yang relevan
2. Analisis Kata : setiap kata dalam teks dianalisis untuk menentukan makna dan kontribusi terhadap sentiment secara keseluruhan
3. Penghitungan polarity : Berdasarkan analisis kata, TextBlob menghitung nilai polaritas keseluruhan dari teks tersebut.
4. Penyimpulan sentiment : Dengan menggunakan polaritas yang telah dihitung maka dapat disimpulkan bersifat penipuan atau non penipuan

D. Pembuatan Model

1. Arsitektur Model LSTM

Arsitektur model LSTM yang digunakan dalam penelitian ini terdiri dari beberapa lapisan yang masing-masing memiliki peran dan fungsi yang penting dalam pemrosesan data berurutan. Berikut tampilan arsitektur yang digunakan pada Gambar 6



Gambar 6 Arsitektur LSTM

Lapisan pertama dari model LSTM adalah lapisan embedding. Lapisan ini bertanggung jawab untuk mengubah setiap kata dalam teks menjadi representasi vector yang padat, dalam penelitian ini lapisan embedding ditetapkan sebesar 128. Setelah melalui lapisan embedding data diproses melalui lapisan LSTM, lapisan ini memiliki 128 unit LSTM yang memungkinkan untuk memahami pola pola temporal dalam data berurutan dengan lebih baik. Selanjutnya Dropout layar untuk mengurangi overfitting, lapisan dropout diterapkan setelah lapisan LSTM, dengan tingkat 0.2 digunakan untuk secara acak mengabaikan sebagian unit selama proses pelatihan. Terakhir dense layer melewati satu neuron menggunakan fungsi aktivasi sigmoid. Hal ini digunakan untuk menghasilkan output akhir dalam bentuk probabilitas kelas positif

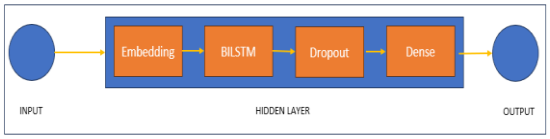
Untuk meningkatkan kinerja model, beberapa perubahan pada hyperparameter yang dilakukan dalam penilitian ini adalah pada tabel 2:

Tabel 2. Perbedaan Hyperparameter untuk LSTM

Parameter	Sebelum	Sesudah
Unit LSTM	64	128
Embedding	64	128
Dropout	0.5	0.2

2. Arsitektur Model BILSTM

Arsitektur model BI-LSTM yang digunakan dalam penilitian ini terdiri dari beberapa lapisan yang sama seperti arsitektur LSTM sebelumnya, seperti pada Gambar 7



Gambar 7. Arsitektur BILSTM

Lapisan-Lapisan yang diterapkan pada BILSTM sama seperti pada lapisan yang digunakan pada arsitektur LSTM sebelumnya. Serta untuk meningkatkan kinerja model menggunakan beberapa perubahan pada hyperparemeter sebagai berikut pada Tabel 3

Tabel 0. Perbedaan Hyperparameter untuk Bi-LSTM

Parameter	Sebelum	Sesudah
Unit BI-LSTM	64	128
Embedding	64	128
Dropout	0.5	0.2

3. Training Model

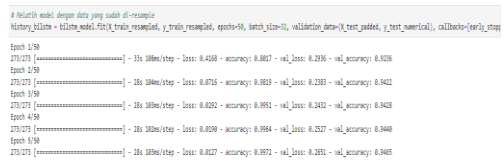
Pertama-tama, dataset yang digunakan diimpor dari file CSV menggunakan pustaka Pandas. Data

I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol.16 No.1, Mei 2024
eISSN : 2580-3042
pISSN : 1979-0694

kemudian dibagi menjadi set pelatihan dan pengujian menggunakan fungsi "train_test_split" dengan pembagian training 75% dan testing 25%, Selanjutnya label yang sudah diberikan dikonversi dari text menjadi numerik menggunakan pemetaan yang ditentukan sebelumnya, di mana "penipuan" direpresentasikan sebagai 1 dan "non penipuan" sebagai 0. Setelah data dibagi dan diberikan label secara numerik, maka data akan dihitung frekuensinya dengan menggunakan Tokenizer lalu setiap kata diubah menjadi indeks dengan teks to sequence. Karena data yang dimasukkan cukup beragam maka akan dilakukan padding. Seperti pada gambar 8berikut.

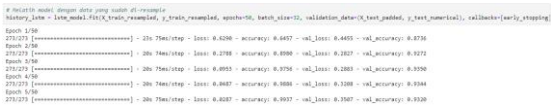
Gambar 8. Inisialisasi Model

Training menggunakan 2 model LSTM dan BI-LSTM yang terdiri dari lapisan Embedding, LSTM, Dropout dan Dense. Lapisan Embedding digunakan untuk mengonversi kata-kata menjadi vector numeric, sedangkan LSTM berperan dalam pemahaman konteks dan urutan teks. Dropout diterapkan untuk mencegah overfitting, dan Dense memberikan output biner untuk klasifikasi. Sebagai pembanding untuk mendapatkan akurasi. Setiap algoritma menggunakan 50 epoch dengan setiap iterasinya akan melakukan training dan validity terhadap model yang terbentuk serta menerapkan fungsi callback yang akan berhenti jika tidak ada peningkatan akurasi. Selanjutnya model yang dibuat juga menggunakan Learning Rate 0,0001 yang ada pada library keras optimizer Adam.
Dataset yang digunakan pada penelitian terdapat imbalance, maka perlu dilakukan oversampling menggunakan Random OveSampler, Berikut adalah gambaran yang didapat saat proses training dengan metode LSTM pada gambar 9.



Gambar 9. Proses Training Model LSTM

Berikut adalah gambar yang didapat saat proses training dengan metode BI-LSTM pada



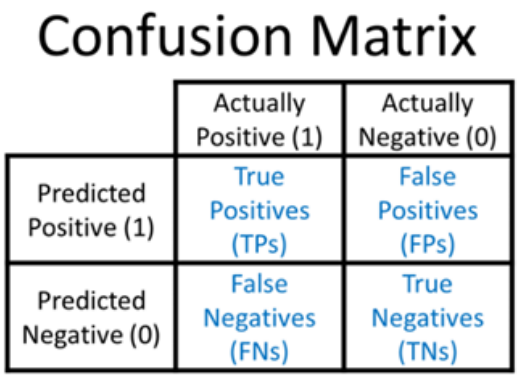
Gambar 10 Proses Training Model BI-LSTM

E. Pengujian

Setelah model berhasil dibuat maka akan dilakukan pengujian performa model LSTM yang telah dilatih pada subset pengujian yang terpisah. Pada tahapan ini akan dihitung metrik kinerja seperti akurasi, presisi, recal, dan F1-score untuk melakukan evaluasi kemampuan model dalam melakuakn analisis sentiment dan deteksi penipuan atau scam. Analisis juga akan dilakukan lebih lanjut terhadap hasil pengujian seperti memeriksa contoh prediksi yang salah atau pengaruh hyperparameter terhadap kinerja model. Jika diperlukan maka akan mengulangi tahap desain dan implementasi untuk memperbaiki model LSTM dan hyperparameter.

F. Evaluasi

Evaluasi akan dilakukan terhadap hasil implementasi mengenai akurasi dan performa dalam klasifikasi. Pada evaluasi ini akan digunakan confusion matrix untuk deteksi penipuan menggunakan label yang sudah di definisikan sebelumnya seperti pada gambar 3.9.



Gambar 11. Confusion Matrix (Draelos, 2019)

Confusion matrix adalah suatu metode yang digunakan dalam evaluasi performa model klasifikasi yang terdiri dari 4 komponen utama : True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) (Heydarian et al., 2022).

Confusion matrix digunakan untuk mengevaluasi kemampuan model dalam memprediksi apakah suatu tweet mengandung informasi penipuan atau tidak. Berikut adalah penjelasan komponen-komponen dalam confusion matrix (Raj et al., 2020):

True Positive (TP): Jumlah tweet yang benar-benar terdeteksi sebagai penipuan oleh model. Ini berarti model memprediksi tweet sebagai penipuan, dan memang benar adanya.

I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol.16 No.1, Mei 2024
eISSN : 2580-3042
pISSN : 1979-0694

True Negative (TN): Jumlah tweet yang benar-benar terdeteksi sebagai non-penipuan oleh model. Ini berarti model memprediksi tweet sebagai non-penipuan, dan memang benar adanya.

False Positive (FP): Jumlah tweet yang salah terdeteksi sebagai penipuan oleh model. Ini berarti model memprediksi tweet sebagai penipuan, tetapi sebenarnya tweet tersebut tidak mengandung informasi penipuan (false alarm).

False Negative (FN): Jumlah tweet yang salah terdeteksi sebagai non-penipuan oleh model. Ini berarti model memprediksi tweet sebagai non-penipuan, tetapi sebenarnya tweet tersebut mengandung informasi penipuan (missed detection) (Alhashmi et al., 2021).

3. HASIL DAN PEMBAHASAN

A. Evaluasi Model LSTM

Evaluasi dari model menggunakan LSTM yang telah berhasil dibuat, mulai dari evaluasi akurasi dan confusion matrix secara keseluruhan oleh model. Confusion matrix berguna untuk menunjukan jumlah prediksi yang benar dan salah dari setiap label (Krstinic et al., 2023). Metrik yang dapat dilihat dari confusion matrix adalah akurasi, presisi, recall dan F1-score pada Tabel 4.

Tabel 4. Confusion Matrix Model

TP	1449
TN	92
FP	105
FN	16

Lalu evaluasi untuk label sebagai berikut pada tabel 5.

Tabel 5. Confusion Matrix Label LSTM

TP	1449
TN	92
FP	105
FN	16

Precision (Presisi): 0.9324 Dari semua transaksi yang diprediksi sebagai penipuan oleh model LSTM, sekitar 93.24% di antaranya memang benar-benar penipuan. Ini menunjukkan bahwa model cenderung cukup hati-hati dalam mengklasifikasikan transaksi sebagai penipuan.

Recall (Sensitivitas): 0.9891 Model LSTM mampu mendeteksi sekitar 98.91% dari semua transaksi penipuan yang sebenarnya. Hal ini menunjukkan bahwa model cukup efektif dalam mendeteksi transaksi penipuan yang sebenarnya.

F1-Score: 0.9599 Rata-rata dari presisi dan sensitivitas sekitar 95.99%, menunjukkan keseimbangan antara kedua metrik tersebut. Ini menunjukkan bahwa model LSTM memiliki kinerja yang seimbang antara presisi dan recall.

Akurasi: 0.9272 Secara keseluruhan, model LSTM berhasil memprediksi sekitar 92.72% transaksi dengan benar. Meskipun nilai akurasi yang cukup tinggi, perlu diperhatikan bahwa akurasi dapat menyesatkan jika ada ketidakseimbangan kelas.

B. Evaluasi Model BI-LSTM

Evaluasi dari model menggunakan BI-LSTM adalah sebagai berikut pada tabel 6.

Tabel 6. Confusion Matrix Model BI-LSTM

TP	1446
TN	120
FP	77
FN	19

Selanjutnya evaluasi untuk label sebagai berikut pada tabel 7.

Tabel 7. Confusion Matrix Label LSTM

Precision	0.9494
Recall	0.9870
F1-Score	0.9679
Accuracy	0.9422

Precision (Presisi): 0.9494 Dari semua transaksi yang diprediksi sebagai penipuan oleh model BI-LSTM, sekitar 94.94% di antaranya memang benar-benar penipuan. Ini menunjukkan bahwa model BI-LSTM memiliki kehati-hatian yang sedikit lebih tinggi dalam mengklasifikasikan transaksi sebagai penipuan.

Recall (Sensitivitas): 0.9870 Model BI-LSTM mampu mendeteksi sekitar 98.70% dari semua transaksi penipuan yang sebenarnya. Hal ini menunjukkan bahwa model BI-LSTM memiliki kemampuan yang sangat baik dalam mendeteksi transaksi penipuan yang sebenarnya.

F1-Score:0.9679 Rata-rata dari presisi dan sensitivitas sekitar 96.79%, menunjukkan keseimbangan antara kedua metrik tersebut. Ini menunjukkan bahwa model BI-LSTM memiliki kinerja yang seimbang antara presisi dan recall, dengan nilai yang sedikit lebih tinggi daripada LSTM.

Akurasi: 0.9422 Secara keseluruhan, model BI-LSTM berhasil memprediksi sekitar 94.22% transaksi dengan benar. Akurasi yang lebih tinggi dibandingkan dengan LSTM menunjukkan bahwa model BI-LSTM memiliki kinerja yang lebih baik secara keseluruhan.

C. Perbandingan Model LSTM dan BI-LSTM

Model BI-LSTM menunjukan kinerja yang lebih baik daripada model LSTM dengan nilai akurasi, presisi, dan F1-score yang lebih tinggi seperti pad tabel 8. (Hong & Oh, 2021).

I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol.16 No.1, Mei 2024
eISSN : 2580-3042
pISSN : 1979-0694

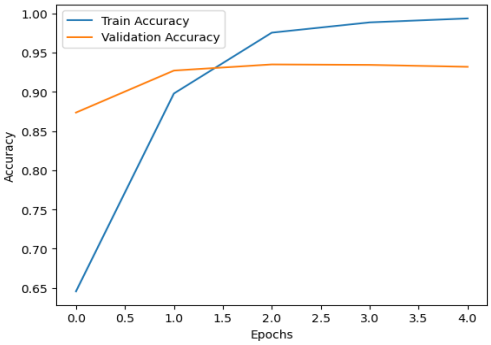
Tabel 8. Rangkuman hasil perbandingan performa

	LSTM	BILSTM
Accuracy	92.72%	94.22%
Precision	93.24%	94.94%
Recall	98.91%	98.70%
F1-Score	95.99%	96.79%

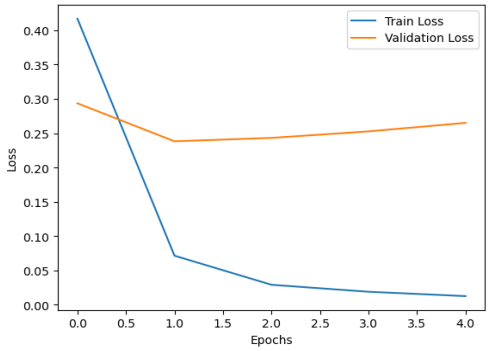
Dari analisis di atas, dapat disimpulkan bahwa model BI-LSTM memiliki kinerja yang sedikit lebih baik dibandingkan dengan model LSTM dalam mendeteksi transaksi penipuan dan non-penipuan. BI-LSTM memiliki presisi yang sedikit lebih tinggi, serta recall dan akurasi yang lebih tinggi, menunjukkan kemampuan yang lebih baik dalam mendeteksi transaksi penipuan yang sebenarnya dan menghindari kesalahan prediksi. Oleh karena itu, dalam kasus deteksi penipuan dan non-penipuan ini, penggunaan arsitektur BI-LSTM mungkin menjadi pilihan yang lebih baik.

D. Grafik Pelatihan dan Validasi

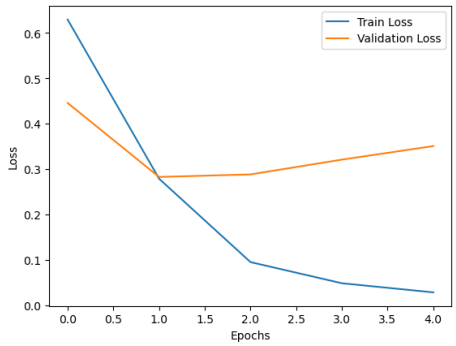
Visualisasi grafik pelatihan dan validasi untuk kedua model dapat dilihat pada Gambar (12, 13, 14, dan 15) grafik dibawah ini :



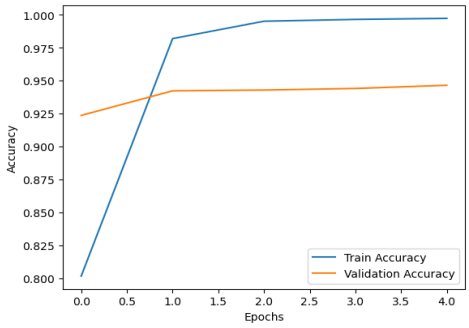
Gambar 12. Grafik Akurasi LSTM



Gambar 13. Grafik Loss LSTM



Gambar 14. Grafik Akurasi BI-LSTM



Gambar 15. Grafik Loss BI-LSTM

Setelah lima epoch pelatihan, model telah diawasi dengan cermat melalui callbacks Early Stopping. Meskipun diharapkan bahwa setiap epoch akan membawa peningkatan kinerja, pada titik ini, metrik yang dimonitor tidak menunjukkan perbaikan yang signifikan dalam performa model (Carr et al., 2021). Ini dapat diartikan sebagai sinyal bahwa model telah mencapai puncak kemampuannya dalam mempelajari pola dari data pelatihan yang tersedia. Dalam rangka untuk mencegah overfitting dan meminimalkan waktu komputasi yang tidak perlu, pelatihan dihentikan lebih awal pada epoch ke-5.

Dengan menggunakan teknik Early Stopping, dapat dipastikan bahwa model diberhentikan saat tidak lagi menghasilkan peningkatan yang signifikan dalam kinerja, sehingga menghemat sumber daya komputasi dan memastikan fokus pada model yang paling optimal (Vilares Ferro et al., 2023). Dalam hal ini, keputusan untuk menghentikan pelatihan pada epoch ke-5 didasarkan pada analisis yang cermat terhadap kinerja model dan tujuan untuk mencapai keseimbangan yang optimal antara generalisasi dan akurasi.

4. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, maka dapat diambil kesimpulan yang menjawab sesuai dengan rumusan masalah yang diangkat, yaitu:

Untuk membuat model deep learning yang dapat digunakan untuk deteksi penipuan online di Indonesia perlu memahami karakteristik penipuan online yang umum terjadi, harus dapat menangani bahasa Indonesia dan menggunakan keyword yang tepat untuk mengumpulkan dataset.

Untuk mengoptimalkan hasil dari model yang dibuat harus menggunakan metode validasi yang tepat untuk menghindari overfitting, melakukan evaluasi kinerja secara berkala, menerapkan hyperparameter dan menyesuaikan dengan dataset yang digunakan. Callback dapat digunakan untuk tahap mencari hyperparameter yang dapat diterapkan

5. REFERENSI

Alhashmi, S. M., Khedr, A. M., Arif, I., & El Bannany, M. (2021). Using a Hybrid-Classification Method to Analyze Twitter Data during Critical Events. *IEEE Access*. https://doi.org/10.1109/ACCESS.2021.3119063

Annur, C. M. (2022). *Pengguna Twitter Indonesia Masuk Daftar Terbanyak di Dunia, Urutan Berapa?* Databoks. https://databoks.katadata.co.id/datapublish/2022/03/23/pengguna-twitter-indonesia-masuk-daftar-terbanyak-di-dunia-urutan-berapa

Carr, S. J. A., Chen, W., Fondran, J., Friel, H., Sanchez-Gonzalez, J., Zhang, J., & Tatsuoka, C. (2021). Early Stopping in Experimentation With Real-Time Functional Magnetic Resonance Imaging Using a Modified Sequential Probability Ratio Test. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2021.643740

Creswell, J. (2017). *Pendekatan Metode Kualitatif, Kuantittatif dan Campuran*. Pustaka Pelajar.

Finaka, A. W., Oktari, R., & Devina, C. (2022). *Maraknya Penipuan Digital di Indonesia*. Indonesiabaik.Id. https://indonesiabaik.id/infografis/maraknya-penipuan-digital-di-indonesia

Gao, C., Yan, J., Zhou, S., Varshney, P. K., & Liu, H. (2019). Long short-term memory-based deep recurrent neural networks for target tracking. *Information Sciences*. https://doi.org/10.1016/j.ins.2019.06.039

Girasa, R. (2020). Artificial intelligence as a disruptive technology: Economic transformation and government regulation. In *Artificial Intelligence as a Disruptive Technology: Economic Transformation and Government Regulation*. https://doi.org/10.1007/978-3-030-35975-1

Heydarian, M., Doyle, T. E., & Samavi, R. (2022). MLCM: Multi-Label Confusion Matrix. *IEEE Access*. https://doi.org/10.1109/ACCESS.2022.3151048

Hong, C. S., & Oh, T. G. (2021). TPR-TNR plot for confusion matrix. *Communications for Statistical Applications and Methods*. https://doi.org/10.29220/CSAM.2021.28.2.161

Khairunnisa, K., & Pithaloka, D. (2023). Use of Twitter Account Autobase @ JPFBASE as A Communication Media For Japanese Pop-Culture Viewers in Pekanbaru. *AICCON 1, August*, 30–31.

Krstinic, D., Seric, L., & Slapnicar, I. (2023). Comments on "MLCM: Multi-Label Confusion Matrix." *IEEE Access*. https://doi.org/10.1109/ACCESS.2023.3267672

Muharomah, S., & Ratnasari, C. I. (2023). Latent Dirichlet Allocation for Uncovering Fraud Cases on Twitter. *Jurnal Riset Informatika*. https://doi.org/10.34288/jri.v5i3.551

Muzakir, A., Syaputra, H., & Panjaitan, F. (2022). A Comparative Analysis of Classification Algorithms for Cyberbullying Crime Detection: An Experimental Study of Twitter Social Media in Indonesia. *Scientific Journal of Informatics*. https://doi.org/10.15294/sji.v9i2.35149

Prayuti, Y. (2023). The Gift Cards Fraud : Challenges and Strategies for Consumer Protection in the Digital Era. *Law Development Journa*, *5*(225), 484–495.

Pusparisa, Y. (2020). *Ribuan Penipuan Online Dilaporkan dalam Lima Tahun Terakhir*. Katadata Media Network. https://databoks.katadata.co.id/datapublish/2020/09/11/ribuan-penipuan-online-dilaporkan-tiap-tahun

Raj, R. J. R., Srinivasulu, S., & Ashutosh, A. (2020). A multi-classifier framework for detecting spam and fake spam messages in Twitter. *Proceedings - 2020 IEEE 9th International Conference on Communication Systems and Network Technologies, CSNT 2020*. https://doi.org/10.1109/CSNT48778.2020.9115796

Shrestha, S. G., & Pradhanang, S. M. (2023). Performance of LSTM over SWAT in Rainfall-Runoff Modeling in a Small, Forested Watershed: A Case Study of Cork Brook, RI. *Water (Switzerland)*.

https://doi.org/10.3390/w15234194

Singh, A., Halgamuge, M. N., & Moses, B. (2019). An Analysis of Demographic and Behavior Trends Using Social Media: Facebook, Twitter, and Instagram. In *Social Network Analytics*. https://doi.org/10.1016/b978-0-12-815458-8.00005-0

Vilares Ferro, M., Doval Mosquera, Y., Ribadas Pena, F. J., & Darriba Bilbao, V. M. (2023). Early stopping by correlating online indicators in neural networks. *Neural Networks*. https://doi.org/10.1016/j.neunet.2022.11.035

Zhao, J., Huang, F., Lv, J., Duan, Y., Qin, Z., Li, G., & Tian, G. (2020). Do RNN and LSTM have long memory? *37th International Conference on Machine Learning, ICML 2020*.