

Bidirectional Long Short-Term Memory untuk Ekstraksi Informasi pada Struk Belanja

I Gede Widiantara^{1*}, Moh. Ali Romli²

^{1,2}Universitas Teknologi Yogyakarta

Jl. Siliwangi (Ringroad Utara), Jombor, Sleman, 55285, D. I. Yogyakarta, Indonesia

*Corresponding author e-mail : gdwidi13@gmail.com

ABSTRAK

Struk belanja merupakan dokumen bukti yang diperoleh setelah melakukan transaksi pembelian. Struk belanja mengandung informasi penting yang berguna untuk pelacakan pengeluaran dan pelaporan keuangan. Namun, proses pengambilan informasi dari struk secara manual cenderung rentan terhadap kesalahan dan memakan waktu. Penelitian ini bertujuan untuk menerapkan metode *Bidirectional Long-Short Term Memory* (Bi-LSTM) dalam ekstraksi informasi pada struk belanja. Bi-LSTM dipilih karena kemampuannya menangkap pola dalam teks tidak terstruktur melalui analisis sekuensial dari dua arah. Proses ekstraksi informasi dilakukan dalam beberapa tahapan, mulai dari pengumpulan data dari *dataset* CORD (*Consolidated Receipt Dataset for Post-OCR Parsing*) dan foto struk belanja yang diambil dengan ponsel. Tahapan berikutnya meliputi pemrosesan gambar (*image preprocessing*), ekstraksi teks menggunakan Optical Character Recognition (OCR), pemrosesan teks (*text preprocessing*), pelabelan, dan pembuatan model Bi-LSTM melalui tahap pelatihan, pengujian, serta evaluasi model dengan *confusion matrix*. Hasil penelitian menunjukkan bahwa model Bi-LSTM yang dikembangkan mencapai akurasi sebesar 95%, *precision* sebesar 95%, dan *recall* sebesar 95%. Penelitian ini diharapkan dapat berkontribusi dalam pengembangan teknologi otomatisasi yang lebih optimal untuk pengelolaan dan analisis data dari dokumen tidak terstruktur.

Kata kunci : Bi-LSTM, Ekstraksi Informasi, Struk Belanja, Literasi Keuangan.

ABSTRACT

Shopping receipt is a proof-of-purchase document received after a transaction. Receipts contain important information that is useful for tracking expenses and financial reporting. However, manually extracting information from receipts is prone to errors and time-consuming. This study aims to apply the Bidirectional Long-Short Term Memory (Bi-LSTM) method for information extraction from receipts. Bi-LSTM was chosen due to its ability to capture patterns in unstructured text through sequential analysis from both directions. The information extraction process involves several stages, starting with data collection from the CORD dataset (Consolidated Receipt Dataset for Post-OCR Parsing) and photos of receipts taken with a smartphone. The next steps include image preprocessing, text extraction using Optical Character Recognition (OCR), text preprocessing, labeling, and building the Bi-LSTM model through training, testing, and model evaluation using a confusion matrix. The results of the study show that the Bi-LSTM model developed achieved 95% accuracy, 95% precision, and 95% recall. This research is expected to contribute to the development of more optimal automation technology for the management and analysis of data from unstructured documents.

Keywords: Bi-LSTM, Information Extraction, Shopping Receipts, Financial Literacy.

I. PENDAHULUAN

Struk belanja adalah dokumen yang diberikan oleh penjual kepada pembeli sebagai bukti resmi transaksi pembelian barang atau jasa. Struk belanja memiliki informasi penting seperti tanggal transaksi, nama barang, jumlah barang, dan total harga. Informasi ini bermanfaat bagi individu dan

perusahaan dalam mengelola anggaran, melacak pengeluaran, dan membuat laporan keuangan [1].

Namun, pengambilan informasi dari struk belanja secara manual memakan banyak waktu dan rentan terhadap kesalahan. Selain itu, pengembangan sistem otomatis untuk mengekstraksi informasi dari struk belanja adalah tugas yang kompleks. Salah satu

tantangan utama dalam memproses struk belanja adalah variasi formatnya [2]. Meskipun teknologi *Optical Character Recognition* (OCR) dapat mengonversi gambar atau teks cetak menjadi data yang dapat dibaca mesin, teknologi ini masih menghadapi beberapa masalah [3]. Sistem OCR konvensional dirancang untuk dokumen yang terstruktur, dimana teks memiliki tata letak yang konsisten. Namun, struk belanja memiliki variasi gaya huruf, ukuran, dan spasi, yang dapat membingungkan sistem. Selain itu, logo, gambar, dan *barcode* dapat menambah kompleksitas, sehingga semakin sulit bagi OCR konvensional untuk mengekstrak informasi dengan akurat.

Ekstraksi informasi pada struk belanja menggunakan *Bidirectional Long-Short Term Memory* (Bi-LSTM) dapat menjadi solusi yang tepat. Ekstraksi informasi merupakan proses pengambilan data faktual dari teks bahasa alami yang tidak terstruktur dan semi terstruktur, seperti entitas, atribut entitas, hubungan antar entitas, dan peristiwa [4]. Bi-LSTM memiliki keunggulan untuk mempertahankan informasi penting dalam data sekuensial dan menangkap pola dalam teks yang tidak terstruktur [5]. Data sekuensial adalah jenis data di mana urutan atau urutan elemen-elemen data sangat penting. Bi-LSTM dirancang untuk menangkap hubungan kontekstual antar kata dalam kedua arah (maju dan mundur), sehingga cocok untuk ekstraksi informasi dari teks sekuensial tersebut [6].

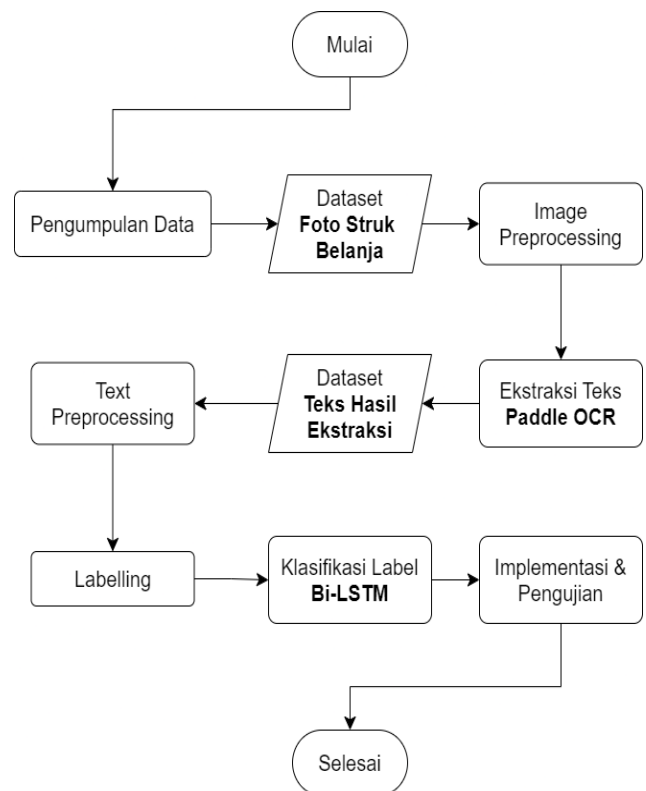
Untuk memaksimalkan kinerja Bi-LSTM dalam ekstraksi informasi, teknik *preprocessing* yang tepat sangat penting. Metode seperti normalisasi teks, deteksi tata letak, dan pemrosesan gambar dapat membantu meningkatkan akurasi model. Menghilangkan kata-kata yang tidak diperlukan (*stopword*) dan memecah teks menjadi unit-unit yang lebih kecil (tokenisasi) juga membantu membersihkan input sebelum diproses. Langkah-langkah ini memudahkan model Bi-LSTM mengenali pola dan meningkatkan akurasi ekstraksi informasi.

Penelitian ini bertujuan untuk mengaplikasikan Bi-LSTM dalam ekstraksi informasi pada struk belanja. Model Bi-LSTM yang dikembangkan bertujuan untuk memudahkan ekstraksi informasi dari struk belanja dengan variasi format dan penyajian data. Sehingga diharapkan dapat menjadi solusi dalam mengelola keuangan melalui otomatisasi pencatatan informasi dari struk belanja.

II. METODE

Penelitian ini mencakup beberapa tahap penting untuk mencapai tujuan akhir, yaitu mengembangkan sistem yang mampu melakukan ekstraksi informasi dari struk belanja. Tahap pertama adalah pengumpulan data. Selanjutnya, pada tahap

kedua dilakukan pemrosesan gambar (*image preprocessing*). Pada tahap ketiga, ekstraksi teks dilakukan menggunakan *Optical Character Recognition* (OCR). Tahap keempat memproses teks yang telah diekstraksi melalui pemrosesan teks (*text preprocessing*). Kemudian, tahap kelima melibatkan pelabelan data. Pada tahap keenam, klasifikasi label diterapkan menggunakan metode *Bidirectional Long Short-Term Memory* (Bi-LSTM). Terakhir, kinerja model dievaluasi menggunakan Confusion matrix. Semua tahap ini ditunjukkan pada Gambar 1.



Gambar 1. Diagram Alur Penelitian

A. Pengumpulan Data

Penelitian ini memanfaatkan data struk belanja berbahasa Indonesia yang diperoleh melalui pengumpulan langsung dari berbagai gerai retail, dengan total sampel sebanyak 500 foto struk. Selain itu, penelitian ini juga menggunakan data tambahan dari sumber publik seperti *dataset CORD* (*Consolidated Receipt Dataset for Post-OCR Parsing*). Data yang dikumpulkan secara langsung oleh peneliti akan difungsikan sebagai data uji, sementara *dataset* publik dimanfaatkan sebagai data latih. Ilustrasi contoh citra struk belanja yang diperoleh menggunakan kamera ponsel dapat dilihat pada Gambar 2.



Gambar 2. Contoh Citra Struk Belanja

B. Pemrosesan Gambar

Pemrosesan gambar (*image preprocessing*) adalah strategi yang diterapkan pada citra mentah sebelum digunakan sebagai *input* pada jaringan saraf tiruan atau model pembelajaran mesin lainnya [7]. Berbagai teknik digunakan untuk melakukan *image processing*, seperti deteksi tepi menggunakan *Canny Edge Detection*. Algoritma *Canny Edge Detection* adalah pendekatan yang sederhana namun efisien untuk mendeteksi garis tepi sebuah objek dalam gambar [8]. Algoritma *Canny Edge Detection* digunakan karena memiliki kemampuan untuk menjaga tingkat kesalahan yang rendah, sekaligus mempertahankan informasi penting dalam gambar dengan menyaring gangguan [9]. *Canny Edge Detection* dipilih untuk mempermudah pemrosesan gambar selanjutnya. Teknik pemrosesan gambar berikutnya adalah pemotongan area relevan (*cropping*), perbaikan kemiringan (*deskewing*), penghilangan *noise*, serta konversi gambar ke dalam format biner. Pemrosesan gambar membantu menyederhanakan dan memperjelas fitur penting dalam gambar, sehingga memudahkan model dalam mengenali pola dan meningkatkan akurasi prediksi.

C. Ekstraksi Teks

Teks dari citra struk belanja yang telah melalui tahap pemrosesan gambar, diekstraksi menggunakan teknologi *Optical Character Recognition (OCR)*. OCR mengonversi teks visual, seperti hasil pemindaian gambar, dokumen PDF, atau foto, menjadi teks digital yang dapat diproses oleh komputer [10]. Penelitian ini menggunakan PaddleOCR karena akurasinya yang tinggi, kecepatannya, fleksibilitasnya, dan sifatnya yang

open-source. PaddleOCR menggunakan *Differentiable Binarization* untuk deteksi teks dan *Convolutional Recurrent Neural Network (CRNN)* untuk pengenalan teks. *Differentiable Binarization* adalah metode yang memungkinkan jaringan syaraf tiruan untuk belajar bagaimana melakukan binarisasi gambar secara adaptif dengan menjadikan proses *thresholding* bagian dari pelatihan, sehingga meningkatkan hasil deteksi [11]. Sedangkan CRNN dalam pengenalan sekuensial tanpa segmentasi adalah jaringan saraf yang mengintegrasikan ekstraksi fitur, pemodelan urutan, dan transkripsi [12]. Hal ini memungkinkan PaddleOCR memproses berbagai gaya teks dengan akurat dan adaptif di berbagai bahasa [13].

D. Pemrosesan Teks

Dalam pemrosesan teks (*text preprocessing*), beberapa teknik dapat meningkatkan akurasi deteksi model. *Case folding* mengubah semua karakter menjadi huruf kecil. Teknik ini mencegah perbedaan interpretasi yang disebabkan oleh kapitalisasi dan membuat teks lebih konsisten. Tokenisasi memecah kalimat menjadi unit-unit kecil atau kata, sehingga model lebih mudah mengenali pola. Penghapusan *stopword* menghilangkan kata-kata umum yang tidak signifikan, seperti "dan" atau "di". Teknik ini meninggalkan hanya kata-kata penting untuk diproses. Penghapusan karakter *non-alphanumeric* membersihkan teks dari simbol-simbol yang tidak relevan yang dapat mengganggu analisis. *Padding* memastikan panjang input seragam, terutama untuk model berbasis sekuens seperti Bi-LSTM[14]. Teknik ini membantu model menangani teks dengan panjang yang bervariasi secara lebih efektif. Secara keseluruhan, teknik-teknik pemrosesan teks mengurangi *noise*, menyederhanakan data, dan memastikan input yang lebih relevan, yang pada akhirnya meningkatkan akurasi deteksi model.

E. Labelling

Teks yang telah dibersihkan akan melalui proses pelabelan (*sequence labeling*). Setiap kata dalam teks akan diberi label sesuai dengan jenis informasinya, seperti "O" untuk kata tidak penting, "ADDRESS" untuk alamat, "DATE" untuk tanggal, "ITEM_NAME" untuk nama barang, dan "PRICE" untuk harga.

F. Klasifikasi Bi-LSTM

Pada tahap klasifikasi, model Bidirectional Long Short-Term Memory (Bi-LSTM) digunakan sebagai metode deep learning. Model Bi-LSTM menangkap informasi konteks dari dua arah (maju dan mundur) dalam urutan teks [15]. Penggunaan Bi-

LSTM dalam klasifikasi teks memungkinkan model mengenali pola yang lebih kompleks dan memahami hubungan antar kata dalam kalimat. Hal ini meningkatkan akurasi dalam pelabelan setiap kata atau token [16]. Arsitektur yang digunakan melibatkan dua lapisan Bi-LSTM, yang berfungsi memperkuat kemampuan model dalam memahami pola yang lebih rumit dan menyeluruh.

G. Implementasi dan Pengujian

Setelah tahap klasifikasi label pada teks dilakukan, langkah berikutnya adalah melakukan evaluasi kinerja model. Salah satu metode yang umum digunakan untuk mengevaluasi performa klasifikasi adalah *confusion matrix*. *Confusion matrix* merupakan representasi tabular yang menggambarkan hasil klasifikasi dari data yang diuji [17]. Melalui *confusion matrix*, dapat dilakukan penghitungan metrik performa seperti akurasi, presisi, *recall*, serta F1-score, yang memberikan gambaran tentang kualitas model klasifikasi [18]. *Confusion matrix* dipilih karena kemampuannya untuk menganalisis kesalahan klasifikasi secara mendetail. *Confusion matrix* juga dapat menangani klasifikasi *multi-label* dengan efektif, sehingga memberikan gambaran menyeluruh tentang kinerja model [19]. Masing-masing metrik tersebut dihitung dengan menggunakan rumus-rumus tertentu, yaitu:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - Score = \frac{2 * Recall + Precision}{Recall + Precision} \quad (4)$$

III. HASIL DAN PEMBAHASAN

Berikut adalah hasil penelitian berdasarkan tahapan yang ada pada metode penelitian ini, seperti dataset yang didapatkan, hasil dataset yang telah dilakukan pre-processing, hasil dari ekstraksi teks dengan Optical Character Recognition, hasil akurasi dari klasifikasi menggunakan *Bidirectional Long Short-Term Memory* (Bi-LSTM), dan hasil perhitungan evaluasi dengan *confussion matrix*. Beberapa tabel yang ditampilkan merupakan hasil dari beberapa tahapan penelitian. Adapun penjelasannya sebagai berikut:

A. Deskripsi Dataset

Pada penelitian yang dilakukan, dataset yang digunakan adalah struk belanja berbahasa Indonesia

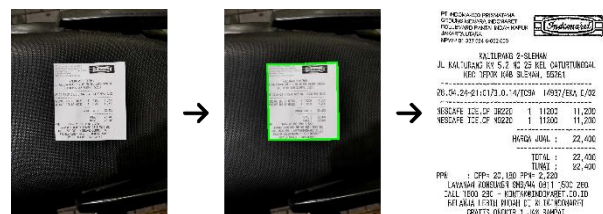
yang diperoleh melalui pengumpulan langsung dari berbagai gerai retail. Data primer tersebut berjumlah sebanyak 500 foto struk. Selain itu, penelitian ini juga mengintegrasikan data tambahan dari sumber publik seperti *dataset CORD* (*Consolidated Receipt Dataset for Post-OCR Parsing*). Berikut adalah tabel deskripsi sumber data penelitian:

Tabel 1. Sumber Data Penelitian

No	Sumber Data	Keterangan
1	Dataset publik dari Hugging Face, pranala: https://huggingface.co/datasets/naver-clova-ix/cord-v2	Data latih berjumlah 2000
2	Data primer	Data uji berjumlah 500

B. Pemrosesan Gambar

Berbagai teknik digunakan untuk melakukan pemrosesan gambar (*image preprocessing*) seperti deteksi tepi menggunakan *Canny Edge Detection*. Teknik ini memperjelas batas tepi citra struk belanja, sehingga memudahkan langkah-langkah selanjutnya seperti pemotongan (*cropping*), perbaikan kemiringan (*deskewing*), dan penghilangan *noise*. Teknik pemrosesan gambar berikutnya adalah konversi gambar ke dalam format biner menggunakan *adaptive thresholding*. Metode ini berguna untuk menangani kondisi pencahayaan yang tidak merata dengan menyesuaikan *threshold* secara dinamis di berbagai area. Tujuan penerapan teknik-teknik ini adalah untuk memastikan data citra yang dihasilkan lebih akurat dan siap untuk tahap selanjutnya.



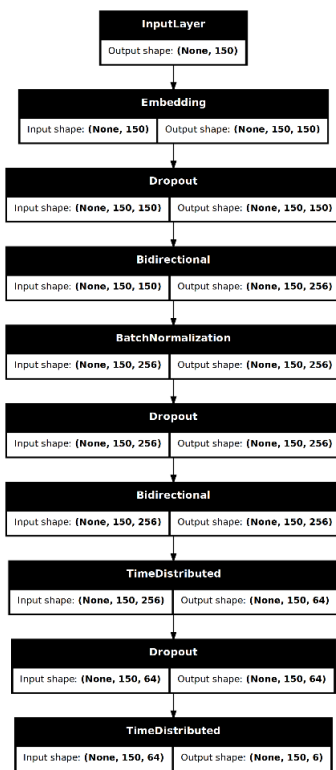
Gambar 3. Hasil Pemrosesan Gambar

C. Ekstraksi Teks

Dalam penelitian ini PaddleOCR digunakan karena akurasinya yang tinggi, kecepatan, fleksibilitas, dan *open-source*. PaddleOCR akan mendeteksi area dalam citra yang kemungkinan besar berisi teks, kemudian mengenali karakter-karakter yang ada di dalam area tersebut menggunakan model pengenalan teks yang telah dilatih pada *dataset* dengan jumlah yang besar. Adapun hasil dari ekstraksi teks pada Gambar 3.

F. Pembuatan Model Bi-LSTM

Model Bi-LSTM dilatih dengan menggunakan fungsi aktivasi ReLU pada lapisan dense dan dioptimalkan menggunakan algoritma Adam, dengan tingkat pembelajaran (learning rate) yang diatur sebesar 0,001. Model ini dirancang dengan arsitektur yang mencakup dua lapisan Bi-LSTM yang masing-masing menggunakan regularisasi kernel untuk mencegah overfitting, serta tiga lapisan dropout untuk menambah robustitas jaringan dalam proses pelatihan. Selain itu, model ini menggunakan lapisan dense yang dibungkus dalam layer time distributed, yang bertujuan untuk menangani data sekuensial secara lebih efektif. Pelatihan dilakukan selama maksimal 10 *epoch* dengan ukuran *batch* 32. Rincian arsitektur lengkap dari model yang dibuat dapat dilihat pada Gambar 5.



Gambar 5. Arsitektur Model

G. Evaluasi

Metrik evaluasi untuk model disajikan dalam Tabel 4 dan Tabel 5. Tabel 4 menunjukkan *confusion matrix*, sementara Tabel 5 merangkum nilai *precision*, *recall*, dan *F1-score* untuk setiap label, serta rata-ratanya. Akurasi yang diperoleh selama pelatihan mencapai 94,52%, sedangkan rata-rata *precision*, *recall*, dan *F1-score* adalah 95%. *Precision* adalah proporsi prediksi positif yang benar dari semua prediksi positif. Sebagai contoh, model mencapai *precision* 100% untuk label *address*, karena memprediksi 2.787 *instance* dengan hanya satu *false*

positive (FP), sedangkan untuk label *price*, *precision* lebih rendah, yaitu 73% karena lebih banyak *false positive*. *Recall* mengukur seberapa baik model mengidentifikasi instance positif yang sebenarnya. Untuk *address*, *recall* adalah 100%, artinya semua instance teridentifikasi dengan benar, tetapi untuk *price*, *recall* adalah 79%, menunjukkan bahwa beberapa *instance* terlewatkan. *F1-score* menyeimbangkan *precision* dan *recall* dengan menghitung rata-rata harmoniknya. Untuk *address*, *F1-score* adalah 100%, mencerminkan *precision* dan *recall* yang tinggi, sedangkan untuk *price*, *F1-score* lebih rendah, yaitu 76%. Secara keseluruhan, model bekerja dengan baik, dengan *precision*, *recall*, dan *F1-score* rata-rata mencapai 95%. Model ini sangat efektif untuk label seperti *address* dan O, meskipun ada ruang untuk perbaikan dalam prediksi label *price*.

Tabel 4. Hasil Evaluasi Proses Pelatihan dan *Confusion Matrix*

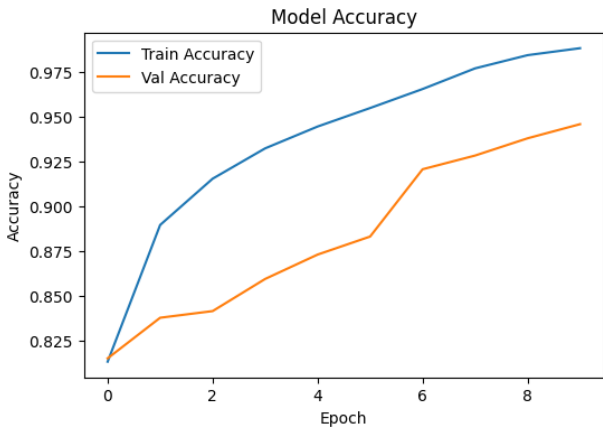
	Akurasi					94,52
Label	address	date	item_name	O	price	
address	2787	0	0	1	0	
date	0	656	68	24	0	
Item_name	0	0	11407	135	1594	
O	0	5	182	50840	438	
price	0	0	1121	346	5396	

Tabel 5. Hasil Evaluasi Proses Pengujian

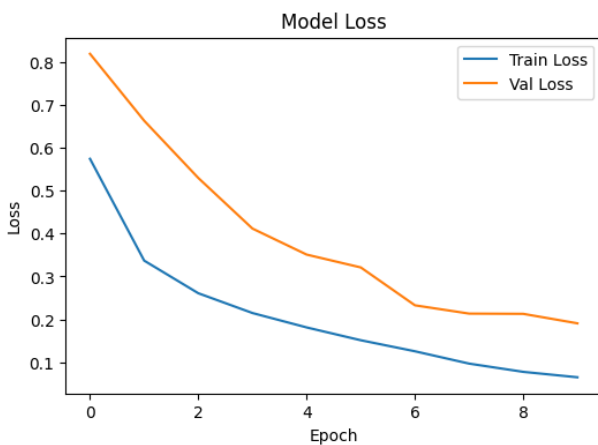
Label	Precision	Recall	F1-Score	Support
address	100	100	100	2788
date	99	88	93	748
item_name	89	87	88	13136
O	99	99	99	51465
price	73	79	76	6863
Avg	95	95	95	75000

Label *price* memiliki nilai *precision*, *recall*, dan *F1-score* yang lebih rendah dibandingkan label lainnya. Hal ini disebabkan oleh karakteristik dataset pelatihan yang digunakan, yaitu struk dalam bahasa Indonesia. Banyak struk di Indonesia yang tidak mencantumkan simbol mata uang (Rp) saat menampilkan harga item. Kondisi ini berbeda dengan struk dalam bahasa lain, seperti yang menggunakan tanda dolar (\$). Sehingga harga item dapat ditandai secara jelas. Ketiadaan penanda mata uang pada struk Indonesia membuat model kesulitan membedakan antara nilai harga dan angka lainnya, seperti nomor telepon atau kode barang. Akibatnya, model sering menghasilkan *false positive* (FP) dan *false negative* (FN) yang lebih tinggi saat memprediksi label *price*. Hal ini berdampak pada penurunan nilai *precision*, *recall*, dan *F1-score* untuk label *price* dibandingkan label lainnya, seperti *address*.

Gambar 5 menunjukkan perbandingan kurva akurasi antara pelatihan dan pengujian. Kurva akurasi pelatihan terus meningkat seiring bertambahnya *epoch*, sementara akurasi pengujian juga meningkat meskipun dengan selisih yang lebih rendah dibandingkan pelatihan. Gambar 6 menampilkan kurva *loss* model untuk pelatihan dan pengujian. Terlihat bahwa *loss* pada data pelatihan terus menurun secara konsisten, begitu juga dengan data pengujian meskipun mengalami penurunan yang sedikit lebih lambat dibandingkan dengan pelatihan.



Gambar 5. Kurva Perbandingan Akurasi Pelatihan dan Pengujian



Gambar 6. Kurva Perbandingan *Loss* Pelatihan dan Pengujian

H. Implementasi Model Bi-LSTM

Model Bi-LSTM digunakan untuk melakukan klasifikasi label pada teks struk belanja. Proses klasifikasi ini bertujuan untuk menandai setiap *token* dalam teks dengan label yang sesuai. Gambar 7 memperlihatkan hasil klasifikasi model pada teks struk belanja, di mana setiap *token* telah diberikan label berdasarkan kelas yang relevan. Setelah proses klasifikasi, dilakukan tahap *post-processing* untuk mengubah hasil klasifikasi tersebut menjadi format JSON, yang dapat dilihat pada Gambar 8.

```
1/1 ----- 0s 68ms/step
pt indomareco prismaatama ['0', '0', '0']
gantung membara indomaret ['0', '0', '0']
pindomaret boulevardpantai indah kapuk ['0', '0', '0', '0']
jakarta utara ['ADDRESS', '0']
rowo133794 092 ['0', 'PRICE']
kalurahan 2 slemang ['ADDRESS', '0', '0']
jl kalurahan km 52 no 25 kel caturtunggal ['ADDRESS', 'ADDRESS', 'ADDRESS', 'ADDRESS', 'ADDRESS', '0', '0']
kac depok kab slemang 55281 ['ADDRESS', 'ADDRESS', 'ADDRESS', '0', '0']
28 04 24 210130141014 14937 ewa_000 ['DATE', 'DATE', 'DATE', 'ITEM_NAME', 'ITEM_NAME', '0']
nescafe ice cf or220 1 11200 11200 ['PRICE', 'ITEM_NAME', 'ITEM_NAME', 'ITEM_NAME', 'ITEM_NAME', 'PRICE', 'PRICE']
nescafe ice cf mc220 1 11200 11200 ['PRICE', 'ITEM_NAME', 'ITEM_NAME', 'ITEM_NAME', 'ITEM_NAME', 'PRICE', 'PRICE']
harga jual 22400 ['PRICE', '0', '0']
total 22400 ['PRICE', 'PRICE']
tunas 22400 ['PRICE', 'PRICE']
dgn dpp2180000 2228 ['0', '0', '0']
layanar konsumen ewa 0811500280 ['0', '0', '0', '0']
call 1500 280 kontakindomaret cd id ['0', '0', '0', '0', '0', '0']
belanja lebih mudah di klikindomaret ['0', '0', '0', '0', '0']
```

Gambar 7. Hasil Klasifikasi Label Menggunakan Model Bi-LSTM

```
{'total': 22400,
'date': '28/04/2024',
'address': 'JL KALURANG KM 52 NO 25 KEL CATURTUNGGAL',
'items': [{'item': 'NESCAFE ICE CF OR220 1', 'price': 11200},
{'item': 'NESCAFE ICE CF MC220 1', 'price': 11200}]}
```

Gambar 8. Konversi Format JSON dari Hasil Klasifikasi Model Bi-LSTM

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa model Bi-LSTM meningkatkan ekstraksi informasi dari struk belanja. Model Bi-LSTM yang dibuat mampu menangkap konteks masa lalu dan masa depan, memberikan keunggulan dibandingkan metode lain untuk pemrosesan dokumen tidak terstruktur. Model Bi-LSTM menunjukkan kinerja yang baik dengan mencapai akurasi 95%, *precision* rata-rata 95%, *recall* rata-rata 95%, dan *F1-score* 95%. Hal ini menjadikannya cocok untuk digunakan dalam aplikasi pelacak keuangan, terutama untuk pengkategorian pengeluaran otomatis dan pengelolaan struk. Namun, kinerja model Bi-LSTM masih dipengaruhi oleh masalah seperti pencahayaan yang buruk, noise pada gambar, dan font yang sulit dibaca. Untuk mengatasi tantangan ini, penelitian di masa depan harus fokus pada peningkatan teknik pemrosesan, menggunakan augmentasi data, dan mengeksplorasi model tambahan.

V. SARAN

Penelitian yang dilakukan memberikan beberapa saran, yaitu dengan menggunakan metode *deep learning* lainnya sebagai perbandingan terhadap Bi-LSTM, menggunakan metode pelabelan dengan format IOB dengan tujuan untuk mengetahui letak awal dan akhir kalimat, serta pengembangan aspek analisis selain aspek kinerja sebagai arah penelitian di masa mendatang. Selain itu, untuk penelitian berikutnya dapat mempertimbangkan penggunaan koordinat letak teks sebagai *input* tambahan dalam pelatihan model, sehingga model dapat lebih akurat dalam mengenali posisi teks pada struk atau dokumen lain yang dianalisis.

DAFTAR PUSTAKA

- [1] T. Saout, F. Lardeux, and F. Saubion, "An Overview of Data Extraction From Invoices," *IEEE Access*, vol. 12, pp. 19872–19886, 2024, doi: 10.1109/ACCESS.2024.3360528.
- [2] M. Ylisiurunen, "Extracting Semi-Structured Information from Receipts," Master's thesis, Aalto University, Espoo, 2022.
- [3] O. Yeung, J. John, and N. Muton, "Digitizing Receipts with OCR," *International Journal of Computing and Digital Systems*, vol. 14, no. 1, pp. 413–422, 2023, doi: 10.12785/ijcds/140132.
- [4] C. Zong, R. Xia, and J. Zhang, *Text Data Mining*. Gateway East: Springer Nature Singapore, 2021.
- [5] P.-H. Li, T.-J. Fu, and W.-Y. Ma, "Why Attention? Analyze BiLSTM Deficiency and Its Remedies in the Case of NER," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, pp. 8236–8244, 2020, doi: 10.1609/aaai.v34i05.6338.
- [6] Z. Hameed and B. Garcia-Zapirain, "Sentiment Classification Using a Single-Layered BiLSTM Model," *IEEE Access*, vol. 8, pp. 73992–74001, 2020, doi: 10.1109/ACCESS.2020.2988550.
- [7] M. Salvi, U. R. Acharya, F. Molinari, and K. M. Meiburger, "The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis," *Comput Biol Med*, vol. 128, p. 104129, Jan. 2021, doi: 10.1016/j.combiomed.2020.104129.
- [8] Z. Bin Faheem *et al.*, "Image Watermarking Using Least Significant Bit and Canny Edge Detection," *Sensors*, vol. 23, no. 3, p. 1210, 2023, doi: 10.3390/s23031210.
- [9] E. A. Sekehravani, E. Babulak, and M. Masoodi, "Implementing canny edge detection algorithm for noisy image," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 4, pp. 1404–1410, Aug. 2020, doi: 10.11591/eei.v9i4.1837.
- [10] S. Drobac and K. Lindén, "Optical character recognition with neural networks and post-correction with finite state methods," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 23, no. 4, pp. 279–295, Dec. 2020, doi: 10.1007/s10032-020-00359-9.
- [11] M. Liao, Z. Zou, Z. Wan, C. Yao, and X. Bai, "Real-Time Scene Text Detection With Differentiable Binarization and Adaptive Scale Fusion," *IEEE Trans Pattern Anal Mach Intell*, vol. 45, no. 1, pp. 919–931, Jan. 2023, doi: 10.1109/TPAMI.2022.3155612.
- [12] W. Yu, M. Ibrayim, and A. Hamdulla, "Scene Text Recognition Based on Improved CRNN," *Information*, vol. 14, no. 7, p. 369, Jun. 2023, doi: 10.3390/info14070369.
- [13] Y. Li, "Synergizing Optical Character Recognition: A Comparative Analysis and Integration of Tesseract, Keras, Paddle, and Azure OCR," Master's thesis, University of Sydney, Sydney, 2024.
- [14] W. K. Sari, D. P. Rini, R. F. Malik, and I. S. B. Azhar, "Multilabel Text Classification in News Articles Using Long-Term Memory with Word2Vec," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 2, pp. 276–285, Apr. 2020, doi: 10.29207/resti.v4i2.1655.
- [15] M. M. A. Busst, K. S. M. Anbananthen, S. Kannan, J. Krishnan, and S. Subbiah, "Ensemble BiLSTM: A Novel Approach for Aspect Extraction From Online Text," *IEEE Access*, vol. 12, pp. 3528–3539, 2024, doi: 10.1109/ACCESS.2023.3349203.
- [16] N. Afrianto, D. H. Fudholi, and S. Rani, "Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 41–46, Feb. 2022, doi: 10.29207/resti.v6i1.3676.
- [17] S. Khomsah and A. S. Aribowo, "Text-Preprocessing Model Youtube Comments in Indonesian," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 4, pp. 648–654, Aug. 2020, doi: 10.29207/resti.v4i4.2035.
- [18] B. B. Sihotang, Y. H. Chrisnanto, and M. Melina, "Klasifikasi Penyakit Stroke Menggunakan Metode Naïve Bayes Classification dan Chi-Square Feature Selection," *Voteteknika (Vocational Teknik Elektronika dan Informatika)*, vol. 12, no. 3, p. 292, Sep. 2024, doi: 10.24036/voteteknika.v12i3.129022.
- [19] M. Heydarian, T. E. Doyle, and R. Samavi, "MLCM: Multi-Label Confusion Matrix," *IEEE Access*, vol. 10, pp. 19083–19095, 2022, doi: 10.1109/ACCESS.2022.3151048.