

Klasifikasi Kategori Produk untuk Manajemen Keuangan Remaja menggunakan Algoritma Long Short-Term Memory

Hendra Sutrisno^{1,*}, Nurul Anisa Sri Winarsih

¹ Program Studi Teknik Informatika, Universitas Dian Nuswantoro, Indonesia

* Correspondence: sutrishendra07@gmail.com

Copyright: © 2024 by the authors

Received: 29 Oktober 2024 | Revised: 1 November 2024 | Accepted: 2 Desember 2024 | Published: 19 Desember 2024

Abstrak

Generasi Z sering mengalami kesulitan dalam mengelola keuangan akibat kebiasaan konsumsi impulsif dan kurangnya perencanaan finansial, yang dapat menyebabkan masalah jangka panjang seperti pengeluaran berlebih dan minimnya tabungan. Penelitian ini bertujuan mengembangkan model klasifikasi kategori yang dapat diintegrasikan dalam aplikasi pencatat keuangan untuk membantu generasi muda mengelola uang dengan lebih baik. Aplikasi ini memiliki fitur utama berupa sistem otomatis yang mengklasifikasikan nama produk ke dalam kategori pengeluaran seperti makanan, transportasi, dan belanja menggunakan model *Long Short-Term Memory* (LSTM). Model LSTM dipilih karena kemampuannya memahami urutan kata dan konteks teks, yang penting dalam pengelompokan produk. Dataset yang digunakan terdiri dari 4.499 data produk yang terbagi ke dalam tiga kategori: 1.488 untuk makanan, 1.682 untuk transportasi, dan 1.329 untuk belanja. Model ini dilatih menggunakan pendekatan *supervised learning*, dengan pembagian data untuk pelatihan dan pengujian. Hasilnya, model mencapai akurasi 86% pada data validasi dan data pengujian, dengan metrik tambahan seperti *precision*, *recall*, dan *F1-score* yang menunjukkan performa baik. Penelitian ini berkontribusi dengan menerapkan teknik *preprocessing* inovatif dan *oversampling* untuk menangani ketidakseimbangan data, yang diharapkan dapat meningkatkan akurasi model dalam klasifikasi pengeluaran. Aplikasi yang dihasilkan tidak hanya mempermudah pencatatan keuangan secara otomatis, tetapi juga berpotensi mendukung pengguna dalam membuat keputusan finansial yang lebih bijak.

Kata kunci: aplikasi pencatat keuangan; klasifikasi produk; lstm; manajemen keuangan; pemrosesan teks

Abstract

Generation Z often faces difficulties in managing their finances due to impulsive spending habits and a lack of financial planning, which can lead to long-term issues such as overspending and minimal savings. This research aims to develop a category classification model that can be integrated into a financial tracking application to help young people manage their money more effectively. The main feature of the application is an automated system that classifies product names into expense categories such as food, transportation, and shopping using a *Long Short-Term Memory* (LSTM) model. LSTM was chosen for its ability to understand word sequences and text context, which is essential in product grouping. The dataset used consists of 4,499 product entries divided into three categories: 1,488 for food, 1,682 for transportation, and 1,329 for shopping. The model was trained using a supervised learning approach, with data split for training and testing. The model achieved 86% accuracy on both validation and test data, with additional metrics such as *precision*, *recall*, and *F1-score* indicating good performance. This study contributes by applying innovative preprocessing techniques and oversampling to address data imbalance, which is expected to enhance the model's accuracy in classifying expenses.



Keywords: *financial record keeping application; product classification; lstm; financial management; text processing*

PENDAHULUAN

Keuangan memiliki peran penting dalam kehidupan, terutama bagi remaja atau generasi Z yang mulai menghadapi tanggung jawab finansial lebih besar. Akses mudah ke layanan digital dan platform *e-commerce* mempengaruhi pola transaksi mereka, membuat literasi keuangan menjadi sangat penting bagi remaja (Anjani et al., 2022; Triani & Mulyadi, 2019). Belajar mengelola uang sejak dini dapat membantu remaja mengembangkan kebiasaan finansial yang sehat untuk kesuksesan di masa depan.

Banyak remaja belum sepenuhnya memahami pengelolaan uang dengan baik. Survei OJK pada 2019 menunjukkan indeks literasi keuangan hanya mencapai 38,03 (Rafif et al., 2021). Fenomena *Fear of Missing Out* (FOMO) membuat generasi Z lebih rentan terhadap pengeluaran impulsif dan tren, bahkan 39% dari mereka rela berhutang demi mengikuti tren (Ulum, 2024; Widiantari et al., 2023). Survei juga menunjukkan bahwa generasi muda cenderung melakukan pengeluaran impulsif lebih sering dibandingkan generasi lainnya (Venia et al., 2021).

Berdasarkan hasil *survey* tersebut diperlukan solusi yang dapat membantu generasi Z dalam mencatat dan mengelola pengeluaran mereka secara efektif. Dengan perkembangan teknologi, aplikasi keuangan berbasis digital menawarkan peluang besar untuk membantu individu mengelola keuangan mereka dengan lebih baik. Salah satu pendekatan yang dapat diambil adalah dengan mengembangkan aplikasi pencatat keuangan yang dilengkapi dengan fitur otomatisasi, seperti klasifikasi pengeluaran berdasarkan kategori produk. Untuk meningkatkan efisiensi dalam klasifikasi pengeluaran, kami mengusulkan penggunaan model *Long Short-Term Memory* (LSTM) dalam aplikasi pencatat keuangan. Solusi ini tidak hanya akan memudahkan proses pencatatan, tetapi juga memberikan wawasan yang lebih baik tentang kebiasaan pengeluaran pengguna.

LSTM yang merupakan bagian dari *Recurrent Neural Network* (RNN), dikenal sangat efektif dalam memproses data sekuensial seperti teks jika dibandingkan dengan metode tradisional (Daiman et al., 2024; Mabrouk & Redondo, 2020; Sujjada, 2024), sehingga cocok untuk menangani masalah klasifikasi nama produk dari data transaksi pengguna. Keunggulan LSTM dibandingkan dengan pendekatan *Neural Network* tradisional adalah kemampuannya untuk membuat keputusan tentang informasi mana yang perlu dipertahankan atau dilupakan (Shah et al., 2022). Dengan kemampuan untuk memahami konteks dalam urutan kata, LSTM dapat secara otomatis mengkategorikan produk yang dibeli ke dalam berbagai kategori, seperti makanan, belanja, atau transportasi, berdasarkan nama produk tersebut. Selanjutnya pada penelitian mereka, dataset produk awal akan menjalani proses *preprocessing* untuk memastikan kualitas dan konsistensi data sebelum menerapkan teknik *Machine Learning* (Maia et al., 2024).

Penelitian sebelumnya membandingkan kinerja LSTM dengan SVM dan *Multinomial Naïve Bayes* dalam klasifikasi hoax, namun hasil akurasi LSTM yang rendah menunjukkan bahwa tanpa *preprocessing* yang tepat, kinerjanya kurang optimal (Cahyo & Aesyi, 2023). Penelitian lain menggunakan LSTM untuk klasifikasi topik di Twitter dengan akurasi yang lebih tinggi, tetapi masih menghadapi masalah distribusi data yang tidak seimbang. Kedua penelitian ini menunjukkan pentingnya perbaikan dalam *preprocessing* dan penanganan ketidakseimbangan data (Hani & Ratnasari, 2023).

Berdasarkan penelitian sebelumnya, penelitian ini akan mengembangkan metode dengan fokus pada Teknik normalisasi data dan *oversampling* untuk menangani ketidakseimbangan data. Beberapa contohnya seperti *Remove Vowels for Food Data*, ini berfungsi untuk Menghapus huruf vokal dari nama produk yang berkategori makanan untuk menciptakan

variasi dalam penamaan yang sering ditemukan dalam struk belanja. Selanjutnya ada *One Hot Encoding* untuk Mengubah label data yang bertipe kategorikal menjadi vector biner (Yu et al., 2022). Selanjutnya *case folding* yang berguna untuk menghindari perbedaan antara huruf besar dan huruf kecil (Nofiyani & Wulandari, 2022). Lalu untuk data produk juga akan dilakukan *tokenizing* yang merupakan proses pemotongan teks dalam kalimat menjadi token-token (Hani & Ratnasari, 2023). Selanjutnya untuk mengatasi ketidakseimbangan data akan dilakukan *oversampling* menggunakan *Synthetic Minority Over-sampling Technique* (SMOTE) yang bertujuan untuk menyeimbangkan distribusi kelas dengan menambah jumlah data pada kelas minoritas (Pramayasa et al., 2023). Oversampling merupakan salah satu teknik untuk menangani data yang *imbalance* (Gumelar et al., 2021). Langkah ini penting untuk meningkatkan performa model LSTM dalam mengklasifikasikan pengeluaran ke dalam kategori yang tepat, terutama dalam konteks data yang memiliki distribusi yang tidak merata. Teknik ini juga akan memastikan bahwa setiap kategori pengeluaran mendapatkan representasi yang lebih baik sehingga model dapat bekerja dengan lebih akurat dan konsisten.

Pemilihan parameter pada layer model juga tidak kalah penting untuk menghasilkan model yang maksimal. Proses ini ditentukan dengan *hyperparameter tuning* menggunakan keras tuner, ini merupakan proses mengoptimalkan parameter yang digunakan dalam pembelajaran mesin untuk meningkatkan kinerja model (Purnama, 2024). Fungsi *loss* yang digunakan pada model ini yaitu *categorical crossentropy*. Fungsi *loss* ini dipilih karena sesuai untuk masalah klasifikasi multi-kelas (Nisa & Candra, 2023). Untuk optimasi, digunakan algoritma Adam yang efisien dalam memperbarui bobot, karena dapat menyesuaikan *learning rate* selama proses pelatihan (Fajrina et al., 2024).

Meskipun LSTM menawarkan banyak keunggulan, ada tantangan yang perlu diperhatikan, seperti potensi *overfitting* jika model tidak dioptimalkan dengan baik atau masalah ketidakseimbangan data pada kategori pengeluaran yang berbeda menjadi tantangan dalam penerapan klasifikasi ini. Namun, dengan penanganan yang tepat, seperti regularisasi dan teknik normalisasi data, model ini dapat memberikan hasil yang akurat. Penelitian ini bertujuan untuk mengembangkan model klasifikasi pengeluaran otomatis yang berbasis LSTM, yang diharapkan dapat membantu pengguna, khususnya generasi Z, mengelola pengeluaran secara lebih terstruktur dan efisien. Implikasi praktis dari penelitian ini adalah kemampuannya untuk diterapkan dalam aplikasi keuangan modern yang secara otomatis mengelompokkan pengeluaran pengguna berdasarkan kategori. Hal ini akan mempermudah pengguna dalam memantau pola pengeluaran mereka, memberikan wawasan mengenai kebiasaan belanja, dan membantu mereka membuat keputusan keuangan yang lebih bijak.

METODE

Penelitian ini bertujuan mengembangkan model klasifikasi pengeluaran berbasis LSTM untuk membantu generasi Z mengelola keuangan. Proses penelitian terdiri atas empat tahap utama yaitu *Data Crawling*, *Preprocessing*, *Training*, dan *Evaluasi*. Pada tahap pertama, *Data Crawling* dilakukan dengan dua metode. Metode pertama menggunakan Large Language Model (LLM) seperti ChatGPT untuk menghasilkan data produk Indonesia dalam tiga kategori: makanan, transportasi, dan belanja. Data yang dihasilkan dipilih secara manual untuk memastikan relevansi. Metode kedua dilakukan dengan input manual untuk data yang tidak tersedia. Total dataset yang terkumpul berjumlah 4.499 data, terdiri dari 1.319 data makanan, 1.682 data transportasi, dan 1.329 data belanja. Kategori makanan mencakup nama makanan dan merek produk, kategori transportasi memuat nama agen travel, jenis transportasi, lokasi, dan wilayah, sedangkan kategori belanja meliputi barang non-makanan.

Pada tahap *preprocessing*, dataset dipersiapkan melalui beberapa langkah. Langkah pertama adalah *Drop Duplicate Data* untuk mengurangi risiko *overfitting* dan bias. Kemudian, dilakukan penghapusan huruf vokal pada kategori makanan, yang menghasilkan data tambahan

sehingga total dataset meningkat menjadi 5.246. Selanjutnya, proses *Case Folding* dilakukan untuk menyamakan format huruf menjadi huruf kecil, diikuti oleh proses *Remove Symbol & White Space* untuk mengurangi *noise* dalam data. Data label kategori diubah ke dalam format numerik menggunakan *One Hot Encoding*, di mana makanan diwakili oleh [1., 0., 0.], belanja [0., 1., 0.], dan transportasi [0., 0., 1.]. Nama produk kemudian diubah menjadi urutan angka melalui *Tokenizing*, misalnya *Nasi Goreng* menjadi [99, 107, 0, ...]. Untuk mengatasi ketidakseimbangan data, dilakukan *oversampling* menggunakan *Synthetic Minority Over-sampling Technique*(SMOTE), sehingga jumlah dataset bertambah menjadi 7.914. Dataset ini kemudian dibagi menjadi dua bagian, yaitu 80% untuk pelatihan dan 20% untuk pengujian.

Tabel 1. Sample dataset

<i>Belanja</i>	<i>Transportasi</i>	<i>Makanan</i>
Sabun Muka	Traveloka	Yogurt
Tissue Wajah	Tiket.com	Burger
Sabun Mandi	Jakarta	Caramel Machiato
Pasta Gigi	Semarang	Nasi Goreng
Lipstik	Bandung	Chatime

Tabel 2. Arsitektur model

<i>Layer (type)</i>	<i>Output Shape</i>	<i>Param</i>
Embedding	(None, 10, 150)	600.000
Spatial_dropout1d	(None, 10, 150)	0
lstm	(None, 10, 128)	142.848
lstm_1	(None, 512)	20.608
dropout	(None, 512)	0
dense	(None, 3)	90

Tahap *training* melibatkan penggunaan model LSTM dengan arsitektur yang dirancang khusus. Model ini terdiri dari beberapa lapisan, termasuk *Embedding Layer* untuk mengubah token menjadi vektor, *SpatialDropout1D* dan *Dropout Layer* untuk mencegah *overfitting*, serta dua lapisan LSTM dengan masing-masing 128 dan 512 unit untuk menangkap urutan temporal data. Lapisan terakhir adalah *Dense Layer* dengan aktivasi *softmax* yang menghasilkan tiga kelas output. Model dilatih selama 50 *epoch* menggunakan *optimizer Adam* dan fungsi *loss categorical_crossentropy*.

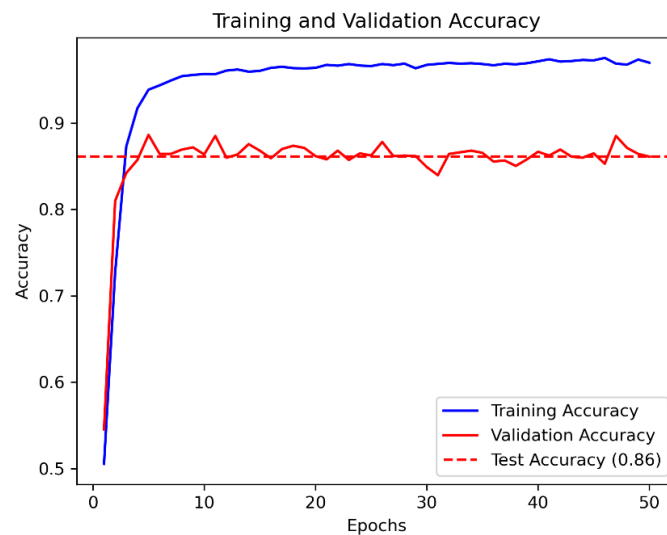
Tahap terakhir adalah evaluasi, dilakukan pengujian dengan *confusion matrix* untuk mengukur *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Dari matriks ini dihitung *precision*, *recall*, dan *F1-score*. *Precision* mengukur akurasi prediksi positif, *Recall* mengukur kemampuan model menemukan data positif relevan, dan *F1-score* memberikan gambaran seimbang performa model. Evaluasi menunjukkan bahwa model dapat mengklasifikasikan produk ke dalam tiga kategori dengan akurasi yang tinggi, menjadikannya layak untuk diintegrasikan dalam aplikasi pencatatan keuangan.

HASIL DAN PEMBAHASAN

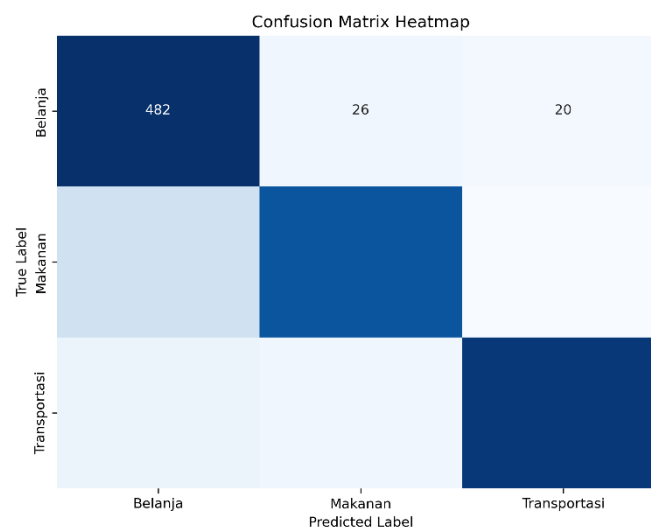
Hasil

Pada gambar 1 merupakan hasil akurasi *training*, validasi, dan *testing* yang dibuat menggunakan *matplotlib*. Gambar tersebut menunjukkan akurasi training dan validasi model LSTM selama 50 *epoch*. Garis biru merepresentasikan akurasi pada data *training*, yang dengan cepat meningkat hingga melebihi 93% dalam sekitar 10 *epoch*, dan kemudian stabil di 96%. Hal ini menunjukkan bahwa model mampu belajar dengan sangat baik dari data latih. Di sisi

lain, garis merah menggambarkan akurasi validasi, yang juga mengalami peningkatan signifikan di awal proses pelatihan, mencapai stabilitas pada kisaran 86% setelah beberapa *epoch*. Kestabilan ini menandakan bahwa model mampu menggeneralisasi dengan baik pada data baru. Selain itu, garis putus-putus merah yang mewakili akurasi pengujian menunjukkan bahwa model mencapai akurasi sebesar 86% pada data uji, memperkuat kesimpulan bahwa model memiliki performa yang sangat baik dalam melakukan klasifikasi pada data yang tidak pernah dilihat sebelumnya.



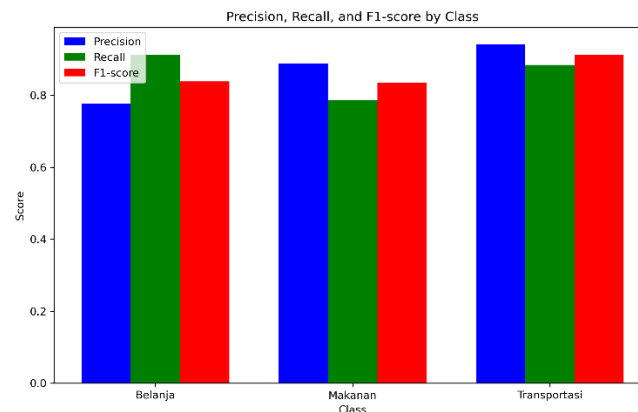
Gambar 1. Hasil akurasi training, validasi, dan *testing*



Gambar 2. *Confusion matrix heatmap*

Selanjutnya pada gambar 2 dapat dilihat performa model LSTM dalam mengklasifikasikan produk ke dalam berbagai kategori. Secara keseluruhan, model LSTM menunjukkan performa yang baik dalam mengklasifikasikan data, terutama untuk kategori transportasi. Namun, model masih mengalami sedikit kesulitan dalam mendeteksi seluruh data yang masuk ke dalam kategori belanja dan makanan. Meskipun telah dilakukan *oversampling*

menggunakan SMOTE untuk mengatasi imbalance data, parameter *random state* pada SMOTE juga berpengaruh pada data yang dihasilkan juga pada akurasi model.



Gambar 3. Evaluasi *precision, recall, f1-score*

Setelah meninjau evaluasi performa model menggunakan metrik *precision, recall,* dan *F1-score,* hasil yang disajikan pada gambar 3 menunjukkan bagaimana model LSTM ini berfungsi dalam mengklasifikasikan pengeluaran ke dalam kategori yang berbeda. Untuk kategori belanja, *precision* tercatat sebesar 78%, yang menunjukkan bahwa dari semua prediksi belanja, 78% adalah benar, meskipun terdapat beberapa kesalahan prediksi. *Recall* untuk kategori ini adalah 91%, yang berarti sebagian besar data belanja berhasil diklasifikasikan dengan benar. *F1-score* untuk kategori belanja mencapai 84%, mencerminkan keseimbangan yang baik antara *precision* dan *recall*. Pada kategori makanan, *precision* sebesar 89% menunjukkan tingkat akurasi yang tinggi dalam prediksi, namun *recall* yang mencapai 79% menunjukkan bahwa masih ada beberapa data makanan yang tidak terdeteksi. *F1-score* untuk kategori makanan adalah 83%, menunjukkan performa yang solid meskipun ada ruang untuk perbaikan. Untuk kategori transportasi, model menunjukkan performa terbaik dengan *precision* sebesar 94% dan *recall* 88%, yang menunjukkan akurasi yang tinggi dalam mendeteksi kategori ini. *F1-score* untuk transportasi adalah 91%, menandakan bahwa model secara keseluruhan sangat efektif dalam mengklasifikasikan pengeluaran dalam kategori ini. Dengan akurasi keseluruhan model mencapai 86% dan rata-rata makro serta rata-rata tertimbang di angka 87%, model LSTM ini menunjukkan kemampuan yang baik dalam mengelola dan mengklasifikasikan data pengeluaran pengguna.

Pembahasan

Secara keseluruhan, model LSTM yang digunakan dalam penelitian ini berhasil mencapai akurasi tinggi baik pada data *training, validasi,* maupun pengujian. Meskipun terdapat ruang untuk peningkatan dalam kategori belanja terutama pada *precision,* serta *recall* pada kategori makanan, performa model pada kategori transportasi menunjukkan keseimbangan *precision* dan *recall* yang sangat baik. Hal ini terlihat dari hasil metrik evaluasi yang menunjukkan bahwa model mampu mengenali dan membedakan transaksi terkait transportasi dengan cukup akurat. Namun, untuk kategori belanja dan makanan, performa model masih kurang memuaskan. Pada kategori belanja, *precision* dan *recall* yang lebih rendah mengindikasikan bahwa model sering salah mengklasifikasikan produk dari kategori ini, terutama dengan memprediksi kategori belanja sebagai makanan. Kesalahan ini juga terlihat jelas pada *heatmap confusion matrix,* di mana banyak data dari kategori belanja diprediksi sebagai makanan. Kondisi ini menunjukkan bahwa meskipun model mampu memahami

konteks data transportasi dengan baik, masih diperlukan peningkatan pada *preprocessing* atau pelatihan model untuk lebih mengoptimalkan akurasi dalam kategori belanja dan makanan.

Penerapan *oversampling* menggunakan SMOTE dalam penelitian ini, meskipun telah membantu dalam menangani ketidakseimbangan data, juga memiliki beberapa keterbatasan yang perlu diperhatikan. SMOTE efektif pada data numerik, namun penerapannya pada data teks kurang optimal. Data sintetik yang dihasilkan dari interpolasi token yang ada cenderung hanya menghasilkan variasi terbatas dan tidak memperkaya model dengan informasi baru. Hal ini menyebabkan model tetap kesulitan dalam membedakan kategori yang mirip, seperti belanja dan makanan. Oleh karena itu, meskipun akurasi keseluruhan cukup baik, hasil ini menggarisbawahi pentingnya penggunaan teknik lain yang lebih sesuai untuk data teks.

Model LSTM dalam penelitian ini memiliki *precision*, *recall*, dan *F1-score* yang lebih stabil dibanding 2 model pada penelitian sebelumnya. Namun, kesulitan dalam membedakan kategori belanja dan makanan tetap menjadi tantangan, terutama karena adanya kemiripan kata atau frasa antara produk-produk dalam kedua kategori tersebut. Panjang input teks juga memengaruhi performa model, dimana *input* yang lebih panjang meningkatkan risiko kesalahan dalam klasifikasi. Hal ini mengindikasikan perlunya pendekatan yang lebih cermat dalam menangani data teks, terutama dalam memastikan bahwa model mampu membedakan karakteristik unik dari setiap kategori dengan lebih baik.

Penelitian ini juga memberikan pandangan penting mengenai teknik penanganan ketidakseimbangan data teks. Meskipun SMOTE digunakan sebagai solusi, hasil ini menegaskan bahwa teknik ini tidak ideal untuk teks. Penelitian ini membuka peluang untuk mengadopsi metode lain, seperti augmentasi teks berbasis sinonim atau *back-translation*, yang dapat menghasilkan variasi data yang lebih relevan dan meningkatkan kemampuan model dalam mengklasifikasikan data yang tidak seimbang. Dengan menggunakan teknik-teknik ini, penelitian di masa depan dapat menghasilkan klasifikasi yang lebih baik dan akurat, tidak hanya dalam domain keuangan tetapi juga pada domain lain yang melibatkan data teks.

Hasil klasifikasi model ini memberikan pandangan tentang persentase belanja pada setiap kategori pengeluaran, sehingga pengguna dapat mengevaluasi apakah alokasi dana mereka sudah sesuai dengan prioritas finansial yang diinginkan. Model ini membantu pengguna memahami distribusi pengeluaran berdasarkan kebutuhan utama, seperti makanan, transportasi, dan belanja. Dengan demikian, pengguna dapat lebih bijak dalam merencanakan pengeluarannya, sekaligus meningkatkan literasi finansial mereka.

Dibandingkan dengan dua penelitian sebelumnya yang menghadapi masalah data tidak seimbang tanpa menerapkan *oversampling* dan *preprocessing*, penelitian ini menunjukkan hasil yang lebih baik dan stabil (Cahyo & Aesy, 2023; Hani & Ratnasari, 2023). Penelitian mereka cenderung menghasilkan akurasi yang rendah dan tidak konsisten karena tidak adanya penanganan khusus terhadap data tidak seimbang dan kurang optimalnya pengolahan data teks. Dengan menerapkan *preprocessing* yang cermat dan eksperimen dengan *oversampling*, penelitian ini menawarkan pendekatan yang lebih sistematis, menghasilkan performa model LSTM yang lebih stabil.

Penelitian kami menunjukkan bahwa model LSTM dapat berfungsi dengan baik untuk klasifikasi produk dalam aplikasi pencatatan keuangan. Penggunaan teknik *preprocessing* dan pengelolaan data teks yang tepat berperan penting dalam keberhasilan model. Temuan ini tidak hanya berkontribusi dalam konteks manajemen keuangan, tetapi juga membuka peluang eksplorasi lebih lanjut untuk pengembangan model yang lebih kompleks pada masalah sekuensial lain, seperti klasifikasi sentimen atau pengenalan entitas dalam teks yang lebih besar.

SIMPULAN

Hasil Penelitian ini menunjukkan bahwa model LSTM mampu mencapai akurasi 86% dalam klasifikasi nama produk, dengan kinerja terbaik pada kategori transportasi. Namun, model masih kesulitan membedakan kategori belanja dan makanan, terutama karena kemiripan kata-kata pada nama produk. Kontribusi utama penelitian ini adalah penerapan LSTM untuk otomatisasi pencatatan keuangan, yang meningkatkan akurasi pelacakan pengeluaran pengguna. Implikasi dari penelitian ini adalah potensi penerapan model dalam aplikasi keuangan berbasis teknologi untuk membantu pengguna memahami pola pengeluaran mereka dan membuat keputusan keuangan yang lebih baik. Penerapan model ini akan membantu pengguna mengelola transaksi dengan lebih efisien. Untuk pengembangan selanjutnya, disarankan penggunaan transfer learning, penggunaan teknik augmentasi data, dan eksplorasi model yang lebih kompleks seperti BERT atau Transformer guna meningkatkan akurasi, khususnya pada kategori yang lebih ambigu.

REFERENSI

- Anjani, D., Robiah, S., Khotimah, L., & Adinugraha, H. (2022). Pelatihan Manajemen Keuangan Guna Mengatur Keuangan Pribadi serta Investasi Masa Depan Bagi Remaja. *Journal of Applied Community Engagement*, 2(1), 61–69. <https://doi.org/10.52158/jace.v2i1.320>
- Cahyo, P. W., & Aesy, U. S. (2023). Perbandingan LSTM dengan Support Vector Machine dan Multinomial Naïve Bayes pada Klasifikasi Kategori Hoax. *Jurnal Transformatika*, 20(2), 23–29. <https://doi.org/10.26623/transformatika.v20i2.5880>
- Daiman, C. N., Rahman, A. Y., Nudiyansyah, F., Studi, P., Informatika, T., Teknik, F., Malang, U. W., Teks, K., News, B., & Tinggi, A. (2024). Klasifikasi Teks Berita Breaking News Di Manggarai Menggunakan Long Short Term Memory. *Jurnal Mnemonic*, 7(2), 170–174. <https://doi.org/https://doi.org/10.36040/mnemonic.v7i2.9939>
- Fajrina, A. N., Pradana, Z. H., Purnama, S. I., & Romadhona, S. (2024). Penerapan Arsitektur EfficientNet-B0 Pada Klasifikasi Leukimia Tipe Acute Lymphoblastik Leukimia. *Jurnal Riset Rekayasa Elektro*, 6(1), 59–68. <https://doi.org/10.30595/jrre.v6i1.22090>
- Gumelar, G., Ain, Q., Marsuciati, R., Agustanti Bambang, S., Sunyoto, A., & Syukri Mustafa, M. (2021). Kombinasi Algoritma Sampling dengan Algoritma Klasifikasi untuk Meningkatkan Performa Klasifikasi Dataset Imbalance. *SISFOTEK : Sistem Informasi Dan Teknologi*, 250–255.
- Hani, D. S., & Ratnasari, C. I. (2023). Klasifikasi Masalah Pada Komunitas Marah-Marah di Twitter Menggunakan Long Short-Term Memory. *Jurnal Media Informatika Budidarma*, 7, 1829–1837. <https://doi.org/10.30865/mib.v7i4.6755>
- Mabrouk, A., & Redondo, R. P. D. (2020). Deep Learning-Based Sentiment Classification : A Comparative Survey. *IEEE Access*, 8, 85616–85638. <https://doi.org/10.1109/ACCESS.2020.2992013>
- Maia, W. F., Carmignani, A., Bortoli, G., Maretti, L., Luz, D., Guzman, D. C. F., ... & Neto, F. L. (2024). Multi-level Product Category Prediction through Text Classification. *arXiv preprint arXiv:2403.01638*.
- Nisa, C., & Candra, F. (2023). Klasifikasi Jenis Rempah-Rempah Menggunakan Algoritma Convolutional Neural Network. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(1), 78–84. <https://doi.org/10.57152/malcom.v4i1.1018>
- Nofiyani, N., & Wulandari, W. (2022). Implementasi Electronic Data Processing Untuk meningkatkan Efektifitas dan Efisiensi Pada Text Mining. *Jurnal Media Informatika Budidarma*, 6(3), 1621–1629. <https://doi.org/10.30865/mib.v6i3.4332>
- Pramayasa, K., Maysanjaya, I. M. D., & Indradewi, I. G. A. A. D. (2023). Analisis Sentimen Program Mbkm Pada Media Sosial Twitter Menggunakan KNN Dan SMOTE. *SINTECH*

- (*Science and Information Technology*) *Journal*, 6(2), 89–98. <https://doi.org/10.31598/sintechjournal.v6i2.1372>
- Purnama, J. J. (2024). Penerapan Algoritma K-Means Untuk Mengelompokkan Kepadatan. *Journal of Applied Computer Science and Technology (Jacost)*, 5(1), 50–55. <https://doi.org/doi.org/10.52158/jacost.v5i1.809>
- Rafif, M. F., Patria, A. S., & Surabaya, U. N. (2021). Perancangan Mobile Game. *ANDHAPURA: Jurnal Desain Komunikasi Visual & Multimedia*, 07(02), 268–281. <https://doi.org/10.33633/andharupa.v7i2.3966>
- Shah, S. A. A., Masood, M. A., & Yasin, A. (2022). Dark web: E-commerce information extraction based on name entity recognition using bidirectional-LSTM. *IEEE Access*, 10, 99633–99645. <https://doi.org/10.1109/ACCESS.2022.3206539>
- Sujjada, A. (2024). Prediksi Harga Bitcoin Menggunakan Algoritma Long ShortTerm Memory. *JURNAL INOVTEK POLBENG*, 9, 450–459. <https://doi.org/10.35314/isi.v9i1.4247>
- Triani, A., & Mulyadi, H. (2019). Peningkatan Pengalaman Keuangan Remaja Untuk Literasi Keuangan Syariah Yang Lebih Baik. *I-Finance: A Research Journal on Islamic Finance*, 5(1), 9–22. <https://doi.org/10.19109/ifinace.v5i1.3714>
- Ulum, M. T., & Yuhertiana, I. (2024). Studi Literatur: Relevansi Perilaku Keuangan Dan Nilai-Nilai Bela Negara Pada Generasi Z. *Journal of Economic, Bussines and Accounting (COSTING)*, 7(4), 7728–7738. <https://doi.org/10.31539/costing.v7i4.10075>
- Venia, M., Marzuki, F., & Yuliniar. (2021). Analisis Faktor yang Mempengaruhi Perilaku Impulse Buying (Studi Kasus pada Generasi Z Pengguna E-commerce). *Korelasi Riset Nasional Ekonomi, Manajemen, Dan Akuntansi*, 2(1), 929–941.
- Widiantari, K. S., Mahadewi, I. A. G. D. F., Suidarma, I. M., & Arlita, I. G. A. D. (2023). Pengaruh Literasi Keuangan, E-Money Dan Gaya Hidup Terhadap Perilaku Keuangan Generasi Z Pada Cashless Society. *Jurnal Ilmiah Manajemen, Ekonomi, & Akuntansi (MEA)*, 7(3), 429–447. <https://doi.org/10.31955/mea.v7i3.2802>
- Yu, L., Zhou, R., Chen, R., & Lai, K. K. (2022). Missing Data Preprocessing in Credit Classification: One-Hot Encoding or Imputation? *Emerging Markets Finance and Trade*, 58(2), 472–482. <https://doi.org/10.1080/1540496X.2020.1825935>