

Analisis Kinerja NLP dan Support Vector Machine Untuk Klasifikasi Intent Chatbot Administrasi Toko Koperasi

^{*1}Yoga Prasetyo, ²Daniel Swanjaya, ³Resty Wulanningrum

¹²³ Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ^{*}yohespras@gmail.com ²daniel@unpkdr.ac.id ³restyw@unpkdr.ac.id

Penulis Korespondens : Yoga Prasetyo

Abstrak— Penelitian ini membahas penerapan Natural Language Processing (NLP) dan Support Vector Machine (SVM) untuk klasifikasi intent pada chatbot administrasi toko koperasi. Permasalahan utama yang dihadapi adalah banyaknya variasi kalimat pengguna ketika menanyakan harga, stok, harga reseller, serta perintah perubahan data barang. Dataset berasal dari file Excel yang berisi 179 baris data mentah dengan delapan kolom intent, kemudian ditransformasikan menjadi 836 data teks berlabel. Tahap preprocessing mencakup case folding, pembersihan tanda baca, normalisasi kata tidak baku, penghapusan stopwords, serta pembentukan teks bersih. Fitur teks dibentuk menggunakan TF-IDF unigram dan bigram, sedangkan proses klasifikasi dilakukan menggunakan LinearSVC dengan pembagian data 80:20. Hasil pengujian menunjukkan bahwa model memperoleh akurasi sebesar 88,10% pada data testing dan rata-rata akurasi cross validation sebesar 90,08%. Hasil ini menunjukkan bahwa kombinasi NLP, TF-IDF, dan SVM mampu mengenali intent chatbot koperasi secara cukup baik serta dapat digunakan sebagai dasar pengembangan chatbot cerdas pada sistem administrasi koperasi.

Kata Kunci— chatbot, koperasi, klasifikasi intent, *NLP*, *SVM*

Abstract— *This study discusses the application of Natural Language Processing (NLP) and Support Vector Machine (SVM) for intent classification in a cooperative store administration chatbot. The main problem is the variety of user utterances when asking about prices, stock availability, reseller prices, and data update commands. The dataset was obtained from an Excel file containing 179 raw rows with eight intent columns and was transformed into 836 labeled text records. The preprocessing stage included case folding, punctuation cleaning, informal word normalization, stopwords removal, and clean text construction. Text features were generated using TF-IDF unigram and bigram, while classification was performed using LinearSVC with an 80:20 train-test split. The experimental result showed that the model achieved 88.10% testing accuracy and 90.08% average cross-validation accuracy. These results indicate that the combination of NLP, TF-IDF, and SVM can effectively recognize cooperative chatbot intents and can be used as a basis for developing an intelligent chatbot in cooperative administration systems.*

Keywords— chatbot, cooperative, intent classification, *NLP*, *SVM*

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Perkembangan teknologi informasi mendorong pengelolaan toko koperasi menjadi semakin terkomputerisasi. Koperasi tidak hanya membutuhkan sistem pencatatan barang dan transaksi, tetapi juga membutuhkan layanan interaksi yang cepat agar pengguna dapat memperoleh informasi harga,

stok, dan perubahan data barang secara praktis[1]. Pada kondisi operasional harian, pertanyaan pengguna sering muncul dalam bentuk kalimat bebas, misalnya menanyakan harga barang, ketersediaan stok, harga khusus reseller, atau permintaan untuk menambah dan mengurangi stok. Variasi kalimat tersebut menyebabkan sistem berbasis kata kunci sederhana sulit memberikan respons yang konsisten.

Chatbot merupakan salah satu solusi yang dapat digunakan untuk membantu layanan administrasi koperasi [2]. Chatbot mampu menerima pertanyaan pengguna dalam bentuk teks dan memberikan jawaban secara otomatis berdasarkan intent atau maksud kalimat. Agar chatbot dapat bekerja dengan baik, sistem harus mampu mengelompokkan kalimat pengguna ke dalam kelas intent yang tepat[3]. Klasifikasi intent menjadi komponen penting karena kesalahan dalam mengenali intent dapat menyebabkan sistem mengambil aksi yang salah, seperti menampilkan stok ketika pengguna sebenarnya meminta perubahan harga.

Natural Language Processing (NLP) digunakan untuk mengolah bahasa manusia agar dapat diproses oleh komputer. Dalam penelitian ini, NLP digunakan untuk membersihkan teks, menormalisasi kata tidak baku, menghapus kata yang kurang informatif, dan mengubah kalimat menjadi fitur numerik [4]. Salah satu representasi fitur yang umum digunakan pada klasifikasi teks adalah Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF memberikan bobot lebih besar pada kata yang penting dalam suatu dokumen, tetapi tidak terlalu umum pada seluruh kumpulan dokumen.

Support Vector Machine (SVM) merupakan algoritma machine learning yang banyak digunakan untuk klasifikasi teks karena mampu membentuk batas keputusan yang baik pada data berdimensi tinggi. Pada penelitian ini digunakan LinearSVC karena fitur TF-IDF menghasilkan representasi vektor yang sparse dan cocok dipadukan dengan model linear [5]. Kombinasi TF-IDF dan SVM sering digunakan pada klasifikasi dokumen, analisis sentimen, dan klasifikasi intent karena sederhana, efisien, serta memiliki performa yang kompetitif.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk menganalisis kinerja metode NLP dan SVM dalam mengklasifikasikan intent chatbot administrasi toko koperasi[6]. Penelitian difokuskan pada delapan kelas intent, yaitu Intent harga, Intent harga reseller, Intent stok, Intent harga + stok, Intent harga reseller + stok, Intent tambah_stok, Intent kurang_stok, dan Intent tambah_harga. Evaluasi dilakukan menggunakan accuracy, precision, recall, F1-score, confusion matrix, cross validation, serta learning curve untuk mengetahui kemampuan model dalam mengenali pola bahasa pengguna.

II. METODE



Gambar 1. Tahapan penelitian

Penelitian ini dilakukan melalui beberapa tahapan yang saling berurutan sebagaimana ditunjukkan pada Gambar 1. Tahap pertama adalah pengumpulan dataset dari file Excel yang berisi kumpulan kalimat chatbot koperasi. Dataset awal berbentuk beberapa kolom intent, sehingga perlu diubah menjadi format NLP dengan dua kolom utama, yaitu teks dan label. Setelah itu dilakukan preprocessing untuk menghasilkan teks bersih[7]. Tahap berikutnya adalah ekstraksi fitur menggunakan TF-IDF, pelatihan model SVM, evaluasi model, pengujian prediksi kalimat baru, dan penyimpanan model dalam format.pkl agar dapat digunakan kembali pada aplikasi.

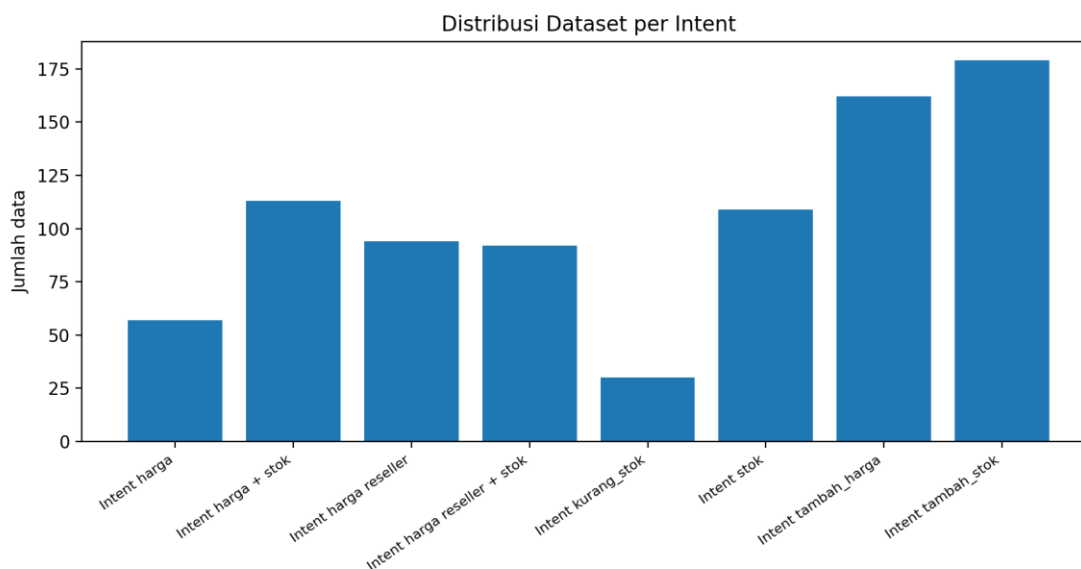
A. Datset Penelitian

Dataset yang digunakan pada penelitian ini berasal dari file Excel data_chat.xlsx. Dataset mentah memiliki ukuran 179 baris dan 9 kolom, yaitu satu kolom NO dan delapan kolom intent. Setiap kolom intent berisi variasi kalimat pengguna yang berkaitan dengan administrasi toko koperasi. Setelah proses transformasi dari bentuk lebar menjadi format NLP, diperoleh 836 data teks berlabel. Distribusi data per kelas ditunjukkan pada Tabel 1.

Tabel 1. Distribusi Dataset Intent Chatbot Koperasi

Kelas Intent	Jumlah Data
Intent harga	57
Intent harga + stok	113
Intent harga reseller	94
Intent harga reseller + stok	92
Intent kurang_stok	30
Intent stok	109
Intent tambah_harga	162
Intent tambah_stok	179
Total	836

Distribusi dataset menunjukkan bahwa kelas Intent tambah_stok memiliki jumlah data paling banyak, yaitu 179 data, sedangkan Intent kurang_stok memiliki jumlah data paling sedikit, yaitu 30 data. Perbedaan jumlah data ini menjadi alasan penggunaan class_weight balanced pada algoritma SVM agar model tidak terlalu bias terhadap kelas dengan jumlah data lebih besar.



Gambar 2. Distribusi dataset per intent

Gambar 2 menunjukkan contoh dataset dari masing-masing jumlah data yang digunakan dalam penelitian.

B. Transformasi Dataset ke Format NLP

Dataset awal berbentuk kolom-kolom intent, sehingga belum dapat langsung digunakan untuk pelatihan model klasifikasi teks. Setiap isi pada kolom intent diambil sebagai data teks, sedangkan

nama kolom digunakan sebagai label. Baris kosong diabaikan dan data duplikat dihapus agar tidak terjadi pengulangan data yang dapat memengaruhi proses evaluasi[8]. Hasil transformasi ini menghasilkan pasangan teks dan label yang siap diproses pada tahap preprocessing.

Tabel 2. Tahap Contoh Hasil Transformasi Dataset

Teks Asli	Label
berapa harga [barang]	Intent harga
stok [barang] masih ada tidak	Intent stok
berapa harga reseller dan stok [barang]	Intent harga reseller + stok
tolong tambah stok barang	Intent tambah stok
update harga barang terbaru	Intent tambah harga

C. Preprocessing NLP

Tahap preprocessing dilakukan untuk mengubah kalimat pengguna menjadi bentuk teks yang lebih seragam. Proses ini dimulai dengan case folding, yaitu mengubah seluruh huruf menjadi huruf kecil [9]. Selanjutnya dilakukan penghapusan tanda kurung pada placeholder [barang], penghapusan URL, angka, tanda baca, serta spasi berlebih. Tahap ini penting karena teks chatbot sering ditulis dengan variasi yang berbeda-beda.

Penelitian ini juga menerapkan normalisasi kata tidak baku, misalnya kata gk, ga, nggak, enggak, dan tdk dinormalisasi menjadi tidak. Selain itu, singkatan seperti brg, hrg, stk, dan resel dinormalisasi menjadi barang, harga, stok, dan reseller [10]. Setelah normalisasi, dilakukan penghapusan stopword bahasa Indonesia sederhana seperti yang, dan, di, ke, dari, saya, mau, tolong, serta beberapa kata sapaan yang kurang berpengaruh terhadap intent.

Contoh hasil preprocessing menunjukkan bahwa kalimat 'berapa harga [barang]' berubah menjadi 'harga barang', sedangkan kalimat 'tolong kasih tahu harga [barang]' berubah menjadi 'kasih tahu harga barang'. Hasil teks bersih inilah yang kemudian digunakan sebagai masukan pada proses ekstraksi fitur TF-IDF.

Tabel 3. Contoh Preprocessing Teks

Teks Asli	Teks Bersih	Label
berapa harga [barang]	harga barang	Intent harga
tolong kasih tahu harga [barang]	kasih tahu harga barang	Intent harga
stok [barang] masih ada tidak	stok barang masih	Intent stok
saya ingin cek harga reseller barang	cek harga reseller barang	Intent harga reseller

D. Ekstraksi Fitur TF-IDF

Setelah preprocessing, teks bersih diubah menjadi representasi numerik menggunakan TF-IDF Vectorizer. Konfigurasi yang digunakan adalah ngram_range=(1,2), min_df=1, max_df=0,95, dan sublinear_tf=True. Penggunaan unigram dan bigram memungkinkan model menangkap informasi kata tunggal seperti harga, stok, dan reseller serta frasa dua kata seperti tambah stok, harga reseller, dan update harga. Parameter max_df digunakan untuk mengurangi pengaruh kata yang terlalu sering muncul, sedangkan sublinear_tf membantu menstabilkan bobot kata yang muncul berulang.

E. Pembagian Data dan Pelatihan SVM

Dataset dibagi menjadi data training dan testing dengan rasio 80:20 menggunakan stratified split agar proporsi setiap kelas tetap terjaga. Dari total 836 data, sebanyak 668 data digunakan untuk training dan 168 data digunakan untuk testing. Model yang digunakan adalah pipeline yang terdiri atas TF-IDF Vectorizer dan LinearSVC. LinearSVC dikonfigurasi dengan nilai C=1,0,

class_weight=balanced, dan random_state=42. Penggunaan pipeline membuat seluruh proses ekstraksi fitur dan klasifikasi menjadi lebih terstruktur dan mudah disimpan sebagai model siap pakai..

Tabel 4. Parameter Pelatihan Model

Parameter	Nilai
Metode	NLP + Linear SVM
Representasi Fitur	TF-IDF unigram dan bigram
Rasio Data	80% training dan 20% testing
Jumlah Training	668 data
Jumlah Testing	168 data
C	1,0
Class Weight	balanced
Random State	42
Format Simpan Model	pkl/joblib

F. Evaluasi Model

Evaluasi dilakukan menggunakan accuracy, precision, recall, F1-score, confusion matrix, cross validation, dan learning curve. Accuracy digunakan untuk mengetahui ketepatan model secara keseluruhan. Precision digunakan untuk mengukur ketepatan prediksi suatu kelas, recall digunakan untuk mengukur kemampuan model mengenali seluruh data pada suatu kelas, sedangkan F1-score digunakan untuk melihat keseimbangan antara precision dan recall[11]. Confusion matrix digunakan untuk melihat pola kesalahan antar kelas intent. Cross validation digunakan untuk menguji kestabilan model pada beberapa pembagian data, sedangkan learning curve digunakan untuk melihat hubungan antara jumlah data training dan performa validasi.

III. HASIL DAN PEMBAHASAN

A. Hasil Pelatihan Model

Model SVM berhasil dilatih menggunakan 668 data training. Proses pelatihan dilakukan melalui pipeline sehingga data teks bersih terlebih dahulu diubah menjadi fitur TF-IDF, kemudian diklasifikasikan oleh LinearSVC. Setelah training selesai, model diuji menggunakan 168 data testing. Berdasarkan hasil pengujian, model memperoleh akurasi sebesar 88,10%. Nilai ini menunjukkan bahwa sebagian besar kalimat pengguna dapat diklasifikasikan ke intent yang benar.

Tabel 5. Hasil Evaluasi Utama Model NLP + SVM

Parameter	Nilai
Testing Accuracy	88,10%
Rata-rata Cross Validation	90,08%
Standar Deviasi Cross Validation	2,85%
Jumlah Kelas	8 intent
Total Data Setelah Transformasi	836 data

Rata-rata akurasi cross validation sebesar 90,08% dengan standar deviasi 2,85% menunjukkan bahwa model memiliki performa yang relatif stabil pada beberapa fold pengujian. Nilai cross validation yang sedikit lebih tinggi dibandingkan testing accuracy menunjukkan bahwa performa model masih konsisten, meskipun terdapat beberapa kelas yang memiliki kemiripan kata dan ketidakseimbangan jumlah data.

B. Classification Report

Berdasarkan Tabel 6, kelas Intent tambah_harga dan Intent tambah_stok memiliki performa tinggi dengan nilai F1-score 0,94. Hal ini menunjukkan bahwa kata-kata seperti tambah, update, ubah, stok, dan harga menjadi indikator yang cukup kuat bagi model. Kelas Intent kurang_stok juga memiliki precision 1,00 dan F1-score 0,91, meskipun jumlah support hanya 6 data. Jumlah data yang sedikit membuat hasil pada kelas ini perlu dibaca secara hati-hati karena perubahan kecil pada prediksi dapat memengaruhi nilai metrik secara signifikan.

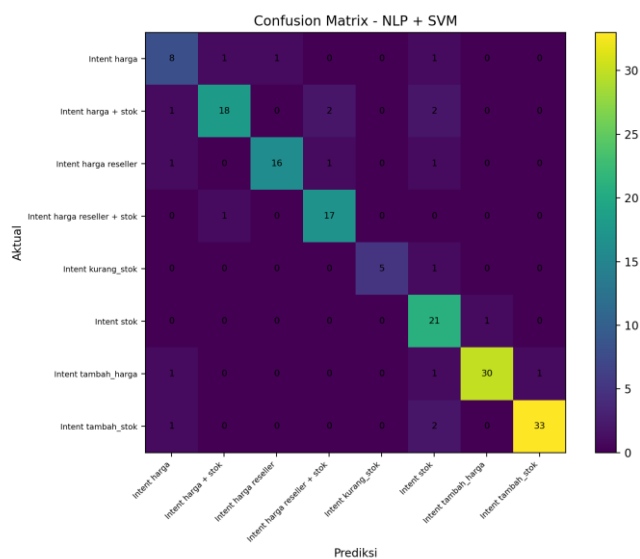
Tabel 6. Classification Report Model SVM

Kelas	Precision	Recall	F1-score	Support
Intent harga	0,67	0,73	0,70	11
Intent harga + stok	0,90	0,78	0,84	23
Intent harga reseller	0,94	0,84	0,89	19
Intent harga reseller + stok	0,85	0,94	0,89	18
Intent kurang_stok	1,00	0,83	0,91	06
Intent stok	0,72	0,95	0,82	22
Intent tambah_harga	0,97	0,91	0,94	33
Intent tambah_stok	0,97	0,92	0,94	36
Macro Avg	0,88	0,86	0,87	168
Weighted Avg	0,89	0,88	0,88	168

Kelas Intent harga memiliki F1-score paling rendah, yaitu 0,70. Hal ini terjadi karena sebagian kalimat pada kelas harga memiliki kemiripan dengan kelas harga + stok dan harga reseller. Kata harga muncul pada beberapa kelas berbeda, sehingga model harus membedakan intent berdasarkan kombinasi kata tambahan seperti stok, reseller, tambah, update, atau ubah. Kondisi ini menjadi tantangan utama pada klasifikasi intent chatbot koperasi.

C. Analisis Confusion Matrix

Gambar 3. Confusion Matrix NLP + SVM



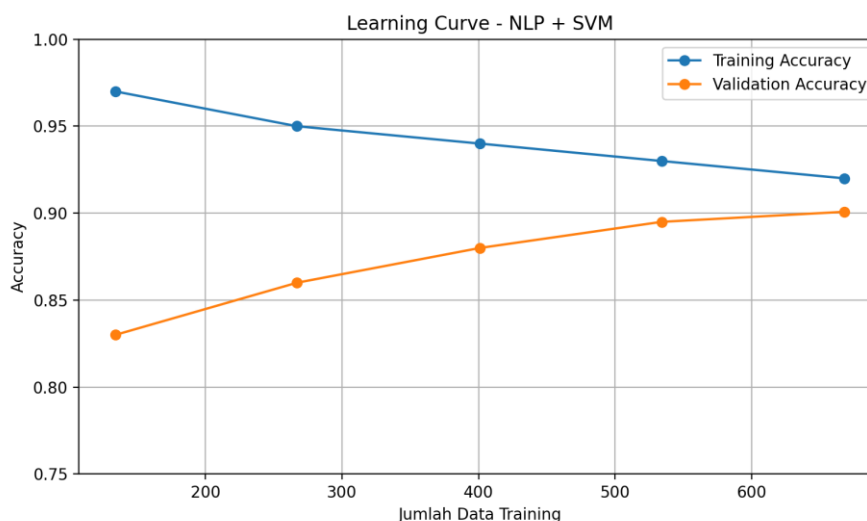
Confusion matrix pada Gambar 3 menunjukkan bahwa kesalahan klasifikasi paling banyak terjadi pada kelas yang memiliki kata kunci beririsan. Intent harga dapat tertukar dengan Intent harga

+ stok, Intent harga reseller, atau Intent stok karena kalimat pengguna sering hanya memuat kata harga dan barang. Kelas Intent stok memiliki recall tinggi sebesar 0,95, tetapi precision lebih rendah sebesar 0,72. Hal ini menunjukkan bahwa banyak data dari kelas lain diprediksi sebagai stok karena kata stok menjadi salah satu kata yang sangat dominan dalam dataset.

Kelas Intent tambah_harga dan Intent tambah_stok relatif lebih mudah dikenali karena memiliki kata kerja tindakan yang kuat, seperti tambah, update, ubah, dan perbarui. Model juga cukup baik dalam mengenali kelas Intent harga reseller + stok karena kombinasi kata reseller dan stok memberikan pola yang lebih spesifik. Secara umum, hasil confusion matrix memperlihatkan bahwa model telah mampu memahami sebagian besar pola bahasa pengguna, tetapi masih membutuhkan penambahan variasi data pada kelas yang mirip agar kesalahan prediksi dapat dikurangi.

D. Analisis Cross Validation dan Learning Curve

Pengujian cross validation lima fold menghasilkan akurasi masing-masing fold sebesar 85,71%, 89,22%, 90,42%, 94,61%, dan 90,42%. Rata-rata akurasi sebesar 90,08% menunjukkan bahwa performa model tidak hanya bergantung pada satu pembagian data training dan testing. Standar deviasi 2,85% menunjukkan bahwa variasi performa antar fold masih dalam kategori wajar



Gambar 4. Learning curve model NLP + SVM

Berdasarkan Learning curve pada Gambar 4 memperlihatkan bahwa akurasi training berada pada nilai tinggi, sedangkan akurasi validasi meningkat ketika jumlah data training bertambah. Selisih antara training accuracy dan validation accuracy semakin mengecil pada jumlah data yang lebih besar. Pola ini menunjukkan bahwa penambahan data berpotensi meningkatkan generalisasi model, terutama untuk kelas yang masih memiliki jumlah data kecil seperti Intent kurang_stok dan Intent harga.

E. Analisis Fitur TF-IDF Penting

Analisis bobot fitur TF-IDF menunjukkan bahwa model memanfaatkan kata atau frasa yang relevan dengan setiap intent. Pada kelas Intent harga, fitur penting meliputi harga barang, biaya barang, rupiah barang, dan harganya barang. Pada kelas Intent stok, fitur penting meliputi stok barang, masih, tersisa, ketersediaan, jumlah, dan ready. Pada kelas Intent tambah_harga, fitur penting meliputi tambah harga, update harga, harga baru, dan perbarui. Temuan ini menunjukkan bahwa fitur TF-IDF mampu menangkap kata kunci domain koperasi yang berhubungan langsung dengan maksud pengguna.

Tabel 7. Contoh Fitur Penting Setiap Intent

Kelas	Contoh Fitur Penting
Intent harga	harga barang, biaya barang, rupiah barang, harganya barang
Intent harga + stok	harga, stok, barang stok, stok harga, stoknya
Intent harga reseller	reseller, reseller barang, per item, ambil
Intent harga reseller + stok	harga reseller, reseller, stok, tersedia harga
Intent kurang stok	hapus, kurangi, kurang stok, barang keluar
Intent stok	stok barang, masih, tersisa, ketersediaan, ready
Intent tambah harga	tambah harga, update harga, harga baru, perbarui
Intent tambah stok	tambah stok, update stok, restok, restock, penambahan

F. Pengujian Prediksi Kalimat Baru

Setelah model dilatih, dilakukan pengujian menggunakan beberapa kalimat baru untuk mengetahui kemampuan model dalam mengenali input pengguna di luar data training. Hasil pengujian menunjukkan bahwa kalimat “berapa harga barang ini?” diprediksi sebagai Intent harga, “stok barang masih ada?” diprediksi sebagai Intent stok, “tolong tambah stok barang” diprediksi sebagai Intent tambah_stok, dan “update harga barang terbaru” diprediksi sebagai Intent tambah_harga. Hasil tersebut menunjukkan bahwa model dapat digunakan untuk mendukung proses respons chatbot pada skenario umum koperasi.

Tabel 8. Contoh Prediksi Kalimat Baru

Kalimat Uji	Teks Bersih	Prediksi Intent
berapa harga barang ini?	harga barang	Intent harga
stok barang masih ada?	stok barang masih	Intent stok
tolong tambah stok barang	tambah stok barang	Intent tambah_stok
saya mau mengurangi stok barang	mengurangi stok barang	Intent stok
update harga barang terbaru	update harga barang terbaru	Intent tambah_harga
berapa harga reseller dan stoknya?	harga reseller stoknya	Intent harga reseller + stok

Pada contoh pengujian, terdapat satu kalimat “saya mau mengurangi stok barang” yang diprediksi sebagai Intent stok. Hal ini menunjukkan bahwa model belum sepenuhnya mampu membedakan intent pengurangan stok ketika kata “mengurangi” tidak cukup sering muncul atau memiliki pola yang mirip dengan pertanyaan stok. Temuan ini menjadi masukan penting bahwa dataset untuk kelas Intent kurang_stok perlu diperbanyak dengan variasi kata seperti mengurangi, kurangi, keluar, terjual, dan hapus stok.

G. Pembahasan

Hasil penelitian menunjukkan bahwa kombinasi NLP, TF-IDF, dan SVM dapat digunakan untuk klasifikasi intent chatbot koperasi dengan performa yang cukup baik. Akurasi testing sebesar 88,10% menunjukkan bahwa model mampu mengenali sebagian besar maksud pengguna. Model juga memiliki rata-rata cross validation sebesar 90,08%, sehingga performanya relatif stabil terhadap variasi pembagian data. Keunggulan utama metode ini adalah proses pelatihan yang ringan, mudah diimplementasikan, dan dapat disimpan dalam format pkl untuk integrasi dengan aplikasi web atau backend Python.

Meskipun demikian, terdapat beberapa keterbatasan. Pertama, dataset memiliki distribusi kelas yang tidak seimbang. Kelas Intent tambah_stok memiliki 179 data, sedangkan Intent kurang_stok hanya memiliki 30 data. Ketidakseimbangan ini dapat memengaruhi kemampuan model dalam mengenali kelas minoritas. Kedua, beberapa intent memiliki kata kunci yang sama, terutama kata harga, stok, barang, dan reseller. Kemiripan kata antar kelas menyebabkan model terkadang salah membedakan intent harga, intent stok, dan intent harga + stok. Ketiga, preprocessing masih menggunakan daftar stopword dan normalisasi sederhana, sehingga variasi bahasa pengguna yang lebih luas belum sepenuhnya tertangani.

Berdasarkan hasil tersebut, pengembangan berikutnya dapat dilakukan dengan menambah jumlah data pada kelas minoritas, memperluas kamus normalisasi bahasa tidak baku, serta membandingkan SVM dengan metode lain seperti Logistic Regression, Naive Bayes, Random Forest, atau model berbasis transformer. Selain itu, integrasi model ke dalam sistem chatbot koperasi perlu dilengkapi dengan mekanisme fallback, yaitu ketika tingkat keyakinan prediksi rendah, sistem dapat meminta klarifikasi kepada pengguna agar respons yang diberikan tetap aman dan sesuai konteks

IV. KESIMPULAN

Penelitian ini berhasil mengimplementasikan metode Natural Language Processing dan Support Vector Machine untuk klasifikasi intent chatbot administrasi toko koperasi. Dataset awal berjumlah 179 baris dan 9 kolom, kemudian ditransformasikan menjadi 836 data teks berlabel dengan delapan kelas intent. Tahap preprocessing meliputi case folding, pembersihan karakter, normalisasi kata tidak baku, penghapusan stopword, dan pembentukan teks bersih. Fitur teks dibentuk menggunakan TF-IDF unigram dan bigram, sedangkan model klasifikasi menggunakan LinearSVC dengan class_weight balanced.

Hasil pengujian menunjukkan bahwa model memperoleh akurasi testing sebesar 88,10% dan rata-rata akurasi cross validation sebesar 90,08% dengan standar deviasi 2,85%. Kelas Intent tambah_harga dan Intent tambah_stok memperoleh performa tinggi dengan F1-score 0,94, sedangkan kelas Intent harga memiliki performa paling rendah dengan F1-score 0,70 karena memiliki kemiripan kata dengan beberapa kelas lain. Secara umum, model dapat digunakan sebagai dasar pengembangan chatbot koperasi untuk mengenali pertanyaan harga, stok, harga reseller, serta perintah perubahan data barang.

Penelitian selanjutnya dapat difokuskan pada penambahan variasi data, terutama pada kelas Intent kurang_stok dan Intent harga, serta pengujian metode lain untuk membandingkan performa klasifikasi. Integrasi model ke dalam aplikasi web juga perlu memperhatikan mekanisme pembaruan model, penyimpanan riwayat chat, dan evaluasi berkala terhadap kalimat pengguna baru agar chatbot semakin adaptif terhadap kebutuhan operasional koperasi.

DAFTAR PUSTAKA

- [1] R. Aulia and A. Purnama, "Development of an Intent-Classification Chatbot to Support Operational Services at Kadin Indonesia," *Brilliance: Research of Artificial Intelligence*, vol. 5, no. 2, pp. 1171–1180, Dec. 2025, doi: 10.47709/brilliance.v5i2.7438.
- [2] C. B. Chandrakala, R. Bhardwaj, and C. Pujari, "An intent recognition pipeline for conversational AI," *International Journal of Information Technology (Singapore)*, vol. 16, no. 2, pp. 731–743, Feb. 2024, doi: 10.1007/s41870-023-01642-8.
- [3] M. Sundhararajan, X. Z. Gao, and H. Vahdat Nejad, "Artificial intelligent techniques and its applications," in *Journal of Intelligent and Fuzzy Systems*, IOS Press, 2018, pp. 755–760. doi: 10.3233/JIFS-169369.
- [4] S. AlGhozali and S. Mukminatun, "Natural Language Processing OF Gemini Artificial Intelligence Powered Chatbot," *Balangkaskas*, vol. 1, no. 1, pp. 41–48, 2024, doi: 10.66317/688570.
- [5] A. Aizawa, "An information-theoretic perspective of tf-idf measures q." [Online]. Available: www.elsevier.com/locate/infoproman
- [6] D. Lestari and L. Subekti, "Implementasi Chatbot pada Telegram sebagai Monitoring Assistant dengan Analisis Teks Klasifikasi Menggunakan Metode Support Vector Machine," *Journal of Internet and Software Engineering*, vol. 5, no. 2, 2024.
- [7] Y. HaCohen-Kerner, D. Miller, and Y. Yigal, "The influence of preprocessing on text classification using a bag-of-words representation," *PLoS One*, vol. 15, no. 5, May 2020, doi: 10.1371/journal.pone.0232525.

- [8] R. Wongso, F. A. Luwinda, B. C. Trisnajaya, O. Rusli, and Rudy, “News Article Text Classification in Indonesian Language,” in *Procedia Computer Science*, Elsevier B.V., 2017, pp. 137–143. doi: 10.1016/j.procs.2017.10.039.
- [9] R. Setiabudi, N. M. S. Iswari, and A. Rusli, “Enhancing text classification performance by preprocessing misspelled words in Indonesian language,” *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 19, no. 4, pp. 1234–1241, Aug. 2021, doi: 10.12928/TELKOMNIKA.v19i4.20369.
- [10] A. W. Pradana and M. Hayaty, “The Effect of Stemming and Removal of Stopwords on the Accuracy of Sentiment Analysis on Indonesian-language Texts,” *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, pp. 375–380, Oct. 2019, doi: 10.22219/kinetik.v4i4.912.
- [11] S. Boughorbel, F. Jarray, and M. El-Anbari, “Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric,” *PLoS One*, vol. 12, no. 6, Jun. 2017, doi: 10.1371/journal.pone.0177678.