

Integrasi Kamus Multibahasa pada Feed Forward Neural Network dan IndoBERT dalam Pengembangan Chatbot Mobile

Arba Adhy Pamungkas^{1,*}, Cecep Nurul Alam¹, Aldy Rialdy Atmadja¹ Roby Juliansyah²

¹ Program Studi Informatika, Universitas Islam Sunan Gunung Djati Bandung, Indonesia

² Dinas Komunikasi dan Informatika Provinsi Jawa Barat, Indonesia

* Correspondence: arbaadhipamungkas@gmail.com

Copyright: © 2024 by the authors

Received: 24 Oktober 2024 | Revised: 28 Oktober 2024 | Accepted: 9 November 2024 | Published: 19 Desember 2024

Abstrak

Perkembangan teknologi digital mendorong kebutuhan akan layanan komunikasi efisien dan responsif yang mendukung multibahasa. Penelitian ini bertujuan untuk mengembangkan *chatbot* yang mempermudah komunikasi dan operasional pekerjaan bagi pengguna aplikasi DigiTeam dengan mengintegrasikan kamus multibahasa ke dalam model *Feed Forward Neural Network* (FFNN) dan IndoBERT. Metode penelitian yang digunakan adalah CRISP-DM, yang sistematis dalam eksplorasi, persiapan data, pemodelan, dan implementasi. Aplikasi DigiTeam dikembangkan menggunakan metode *Agile* untuk meningkatkan fitur dan fungsionalitas aplikasi secara bertahap. *Dataset* yang digunakan terdiri dari 456 pola dan 106 *tag* yang berisi seputar pertanyaan umum maupun operasional pekerjaan. Penelitian ini memanfaatkan kamus multibahasa sebanyak 309 kata untuk meningkatkan pemahaman konteks dan akurasi respons *chatbot* terhadap pertanyaan pengguna. Hasil pengujian menunjukkan bahwa integrasi kamus multibahasa ke dalam model FFNN dan IndoBERT menghasilkan akurasi 95,45% dengan presisi dan *recall* yang seimbang, membuktikan kemampuan *chatbot* dalam memahami dan merespons pertanyaan multibahasa secara *real-time*, serta meningkatkan efisiensi operasional dan akses informasi dalam lingkungan kerja.

Kata kunci: aplikasi *mobile*; *chatbot*; *feed forward neural network*; *indobert*; kamus multibahasa

Abstract

The development of digital technology drives the need for efficient and responsive communication services that support multilingual. This study aims to develop a chatbot that facilitates communication and operational tasks for users of the DigiTeam application by integrating a multilingual dictionary into the Feed Forward Neural Network (FFNN) model and IndoBERT. The research method used is CRISP-DM, a systematic approach in data exploration, preparation, modeling, and implementation. The DigiTeam application was developed using the Agile methodology to gradually enhance the features and functionalities of the application. The dataset consists of 456 patterns and 106 tags containing common and operational work-related questions. This study utilizes a multilingual dictionary with 309 words to improve the chatbot's context understanding and response accuracy to user queries. The test results show that integrating the multilingual dictionary into the FFNN and IndoBERT models yields an accuracy of 95.45% with balanced precision and recall, demonstrating the chatbot's ability to understand and respond to multilingual queries in real-time, while also improving operational efficiency and information access in the workplace.

Keywords: mobile apps; chatbot; feed forward neural network; indobert; multilingual dictionaries

PENDAHULUAN

Era digital yang terus berkembang pesat, teknologi telah menjadi tulang punggung bagi individu dan organisasi dalam kehidupan sehari-hari. Dari penggunaan perangkat *mobile*



hingga aplikasi berbasis web, teknologi memudahkan berbagai aktivitas seperti komunikasi dan manajemen waktu. Dalam konteks organisasi dan perusahaan, perkembangan teknologi telah mengubah cara kita berinteraksi, bekerja, dan menjalankan bisnis. Namun, dengan meningkatnya volume data dan informasi yang harus dikelola, pengguna aplikasi sering menghadapi kesulitan dalam mengakses informasi yang relevan secara efisien dan tepat waktu.

Salah satu organisasi yang merasakan dampak perkembangan teknologi adalah Jabar Digital Service (JDS). Organisasi ini berada di bawah Dinas Komunikasi dan Informatika Provinsi Jawa Barat (Setyasih, 2022). Untuk membantu operasional kerja, Jabar Digital Service mengembangkan aplikasi internal bernama DigiTeam, aplikasi ini dirancang untuk digunakan di berbagai perangkat daerah di Jawa Barat. Selain itu, aplikasi ini mempermudah proses absensi melalui *geolokasi*, manajemen *evidence* dan *logbook*, serta pengelolaan cuti. Namun, pengguna sering kali terkendala disebabkan oleh fitur yang kompleks dan terbatasnya panduan, yang mengakibatkan akses informasi terhambat. Hal Ini menunjukkan bahwa, teknologi memiliki potensi besar dan tantangan dalam mendukung efektivitas operasional pekerjaan.

Banyak organisasi telah memanfaatkan teknologi untuk meningkatkan efektivitas operasional dan mengatasi tantangan pengelolaan data. Salah satu solusi tersebut dengan penggunaan *chatbot*. *Chatbot* dirancang untuk mempercepat akses informasi dan menyederhanakan komunikasi di dalam organisasi yang memiliki kemampuan respons mirip interaksi manusia (Putri & Ramadhan, 2024). Sebagai contoh, penerapan *chatbot* untuk mendukung pelaku UMKM sebagai pengganti konsultan bisnis (Guntara, 2022). Penerapan AI, termasuk *chatbot*, dapat meningkatkan analisis data dan pengambilan keputusan berbasis prediksi (Fidiyanti et al., 2023). Selain itu, aplikasi *chatbot* sebagai layanan informasi virtual di situs web *Infinite Learning* berhasil meningkatkan kualitas dan efektivitas pembelajaran digital, serta memberikan metode penyampaian informasi yang lebih intuitif (Syarof & Rasal, 2024).

Berbagai penelitian menunjukkan potensi algoritma dalam pengembangan chatbot, seperti pada penelitian yang menggunakan *Feed Forward Neural Network* (FFNN) untuk memproses pertanyaan cepat. Namun penelitian ini memiliki keterbatasan data dalam memahami dan menjawab berbagai jenis pertanyaan (Pratama et al., 2021). FFNN merupakan salah satu dari model *Neural Network* yang banyak digunakan dalam berbagai bidang (Pratiwi et al., 2023). *Chatbot* berbasis *Neural Network* berhasil diterapkan untuk FAQ di situs kuliah dengan akurasi 97,27%, tetapi terbatas pada variasi pertanyaan pelatihan dan belum diintegrasikan ke platform *mobile* atau web (Mustakim et al., 2021). Sementara itu, *chatbot* web untuk informasi akademik di Universitas Bengkulu mencapai akurasi pelatihan 92,47% dan validasi 95,16%, namun hanya menjawab pertanyaan kompleks dengan respons tetap atau *statis* (Faurina et al., 2023). Oleh karena itu, penelitian ini akan mengembangkan *chatbot* dengan respons lebih dinamis yang terhubung dengan *database* aplikasi DigiTeam.

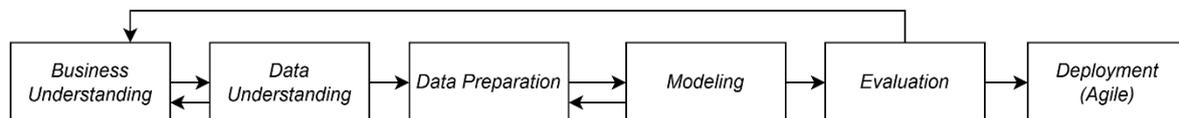
Selanjutnya, penelitian lain yang menggunakan pendekatan *Bidirectional Encoder Representations from Transformers* (BERT) dinilai lebih efektif dalam memahami konteks bahasa secara mendalam (Aljabar & Karomah, 2024). IndoBERT adalah varian BERT yang telah di latih khusus untuk bahasa Indonesia menggunakan pendekatan *transformer* (Simanjuntak et al., 2024), sebagai model bahasa bertopeng yang dilatih menggunakan kerangka kerja *Hugging Face* (Koto et al., 2020). *Chatbot* berbasis BERT telah diterapkan untuk mendukung kesehatan mental pelajar di Indonesia dengan akurasi 98,97%, namun masih terbatas dalam menangani nuansa bahasa dan kesalahan kompleks (Dzaky et al., 2024). IndoBERT juga digunakan untuk meningkatkan kinerja layanan pelanggan *e-wallet*, mencapai akurasi 86,90% dan menunjukkan pentingnya penyesuaian *hyperparameter* dan laju pembelajaran untuk mengoptimalkan kinerja model (Owen & Suhartono, 2024).

Penelitian ini bertujuan untuk mengembangkan *chatbot mobile* yang dapat meningkatkan komunikasi dan operasional pekerjaan di aplikasi DigiTeam dengan mengintegrasikan kamus

multibahasa dalam model FFNN dan IndoBERT, yang mencakup bahasa Indonesia, Sunda, dan Inggris. Integrasi kamus multibahasa ini diharapkan dapat meningkatkan efektivitas *chatbot* dalam memahami dan merespons pertanyaan pengguna dalam berbagai bahasa secara *real-time*, memperbaiki pemahaman konteks serta akurasi respons, dan mempercepat akses informasi, yang pada akhirnya dapat berkontribusi pada peningkatan efisiensi operasional pekerjaan. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan *chatbot*, yang tidak hanya bermanfaat untuk aplikasi DigiTeam, tetapi juga dapat diterapkan di berbagai sektor yang memerlukan solusi teknologi serupa untuk meningkatkan efisiensi operasional pekerjaan.

METODE

Penelitian ini menggunakan pendekatan metodologi *Cross Industry Standard for Data Mining (CRISP-DM)* sebuah pendekatan untuk memenuhi berbagai kebutuhan proyek *data mining* (Pambudi, et al., 2023), yang relevan dalam konteks pengembangan teknologi *chatbot*. Setiap tahapan dalam *CRISP-DM* berperan penting agar kerangka kerja ini berfungsi optimal (Widiansyah et al., 2024). Gambar 1 menunjukkan tahapan *CRISP-DM* yang terdiri dari enam langkah utama yaitu, *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment (Agile)* (Cannavaro, 2023).



Gambar 1. Metode *crisp-dm* dengan pendekatan *agile* pada *deployment*

Pada tahap *business understanding*, dilakukan identifikasi kebutuhan utama JDS untuk meningkatkan operasional pekerjaan melalui aplikasi DigiTeam. Proses ini mencakup analisis terhadap masalah pengguna, seperti lambatnya akses informasi dan kesulitan dalam menggunakan fitur yang kompleks. Tujuan bisnis yang dirumuskan adalah mengembangkan *chatbot* multibahasa yang dapat memberikan respons cepat dan akurat terkait tugas dan informasi pengguna.

Pada tahap *data understanding*, data untuk melatih model *chatbot* dikumpulkan dari pertanyaan-pertanyaan yang diajukan oleh pengguna DigiTeam, yang meliputi topik-topik seperti manajemen proyek, status tugas, dan informasi terkait pengguna. Selain itu, keberagaman bahasa yang digunakan, seperti campuran Bahasa Indonesia, Inggris, dan Sunda, menjadi faktor penting yang perlu diperhatikan agar *chatbot* dapat memberikan respons yang sesuai dan efektif.

Tahap *data preparation* mencakup proses *case folding* untuk mengubah teks menjadi huruf kecil, penghapusan *stop words*, dan *stemming* untuk mereduksi kata-kata. Selanjutnya, *tokenisasi* dilakukan untuk menghitung frekuensi kata dalam representasi *Bag of Words (BoW)*. Pada tahapan ini dilakukan pula pemecahan menjadi sub-kata yang lebih kecil. Selain itu, kamus multibahasa digunakan untuk menerjemahkan kata serta memperkaya pemahaman model terhadap konteks yang lebih luas.

Pada tahap *Modeling*, data yang telah diproses digunakan untuk melatih model *chatbot* dengan pembagian 80% untuk pelatihan dan 20% untuk pengujian. Model dikembangkan dengan mengintegrasikan kamus multibahasa ke dalam FFNN dan IndoBERT, yang dirancang untuk memahami dan memproses Bahasa Indonesia, Sunda, dan Inggris. *Optimizer Adam* digunakan untuk memperbarui bobot model secara efisien, dengan *cross entropy loss* sebagai fungsi kerugian yang mengukur perbedaan antara prediksi dan label. Parameter pelatihan yang digunakan meliputi *num_epochs* 500, *batch_size* 64, *learning_rate* 0.0001, dan *hidden_size*

128. Model ini dirancang untuk memahami konteks pertanyaan pengguna secara *real-time* dan memberikan respons yang akurat dengan memanfaatkan integrasi kamus multibahasa.

Tahap *Evaluation* dilakukan untuk mengukur performa model menggunakan *dataset* pengujian dengan metrik akurasi, presisi, *recall*, dan F1-score, serta pengujian tambahan dengan kalimat *input* yang menggunakan dan tanpa menggunakan kamus multibahasa. Evaluasi ini bertujuan untuk menilai pengaruh integrasi kamus multibahasa terhadap pemahaman konteks dan akurasi respons model, serta memastikan bahwa model dapat memberikan respons yang lebih relevan dan akurat dalam konteks multibahasa.

Pada tahap *deployment*, penerapan model *chatbot* ke dalam aplikasi DigiTeam dilakukan dengan menggunakan pendekatan *Agile*, yang mengutamakan fleksibilitas dan kolaborasi dalam pengembangan aplikasi. Pendekatan ini memungkinkan pengembangan secara *iteratif* melalui sprint mingguan, di mana setiap sprint terdiri dari perencanaan, pengembangan, integrasi, dan pengujian fitur *chatbot*. Pengujian *chatbot* dilakukan dengan menghasilkan data interaksi yang disesuaikan dengan kebutuhan pengguna dan umpan balik yang diperoleh dari pengguna akan digunakan untuk memperbarui dan meningkatkan kinerja *chatbot* sehingga dapat memenuhi ekspektasi dan mendukung efisiensi operasional pekerjaan.

HASIL DAN PEMBAHASAN

Hasil

Hasil dari penelitian ini menunjukkan penetapan tujuan bisnis untuk mengembangkan *chatbot* multibahasa yang dirancang untuk memudahkan akses informasi dalam aplikasi DigiTeam. *Chatbot* ini diharapkan dapat mengatasi kesulitan yang dihadapi pengguna dalam memperoleh informasi dengan menyediakan jawaban yang cepat dan akurat menggunakan kamus multibahasa. Dengan demikian, implementasi *chatbot* ini diharapkan dapat meningkatkan operasional pekerjaan serta pengalaman pengguna secara keseluruhan.

Pada Gambar 2 ditampilkan struktur *dataset intents* beserta kamus multibahasa, yang menunjukkan bagaimana pola pertanyaan dan terjemahan saling terkait untuk meningkatkan pemahaman *chatbot*. *Dataset intents* ini mencakup 456 pola dan 106 *tag*, yang merepresentasikan berbagai pertanyaan umum dari pengguna aplikasi DigiTeam yang diperoleh melalui survei dan pengumpulan data dari forum diskusi grup di lingkungan JDS.

```

1 {
2   "intents": [
3     {
4       "tag": "get_earlyAttendance",
5       "patterns": [
6         "lihat daftar yang tepat waktu", "siapa aja yang tepat waktu", "Siapa yang datang tepat waktu",
7         "Tampilkan daftar yang tepat waktu", "Lihat yang tepat waktu hari ini", "Siapa yang check in tepat waktu",
8         "Daftar yang datang tepat waktu", "Yang tepat waktu hari ini siapa saja?", "Siapa yang datang tepat waktu?"
9       ],
10      "responses": ["get_earlyAttendance"]
11    },
12    {
13      "tag": "get_users_details",
14      "patterns": [
15        "profile saya", "Nama saya", "Detail pengguna", "Info profile saya", "Informasi akun saya",
16        "Tampilkan profile saya", "Detail akun saya", "Siapa nama saya", "Tampilkan informasi profile saya",
17        "Data pengguna saya"
18      ],
19      "responses": ["get_users_details"]
20    },
21    {
22      "tag": "get_logbook",
23      "patterns": [
24        "Bagaimana caranya melihat daftar logbook?", "Bagaimana cara lihat daftar logbook?", "data logbook?",
25        "lihat daftar logbook?", "perlihatkan daftar logbook?", "lihat data logbook?", "lihat logbook",
26        "Logbook hari ini", "Logbook saya", "Mau lihat logbook", "Data logbook saya", "Cek logbook",
27        "Info logbook saya", "Ada logbook?"
28      ],
29      "responses": ["get_logbook"]
30    }
31  ]
32 }
33
34 {
35   "Pagi": {
36     "Inggris": "morning",
37     "Sunda": ["enjing", "isuk"]
38   },
39   "Siang": {
40     "Inggris": ["afternoon"],
41     "Sunda": ["siang", "beurang"]
42   },
43   "Sore": {
44     "Inggris": "sore",
45     "Sunda": ["sonten"]
46   },
47   "Malam": {
48     "Inggris": "night",
49     "Sunda": ["peuting", "weung"]
50   },
51   "Salam": {
52     "Inggris": "greeting",
53     "Sunda": "punten"
54   },
55   "Apa": {
56     "Inggris": "what",
57     "Sunda": "naon"
58   },
59   "Kabar": {
60     "Inggris": "how are you",
61     "Sunda": ["kumaha damang", "damang"]
62   }
63 }

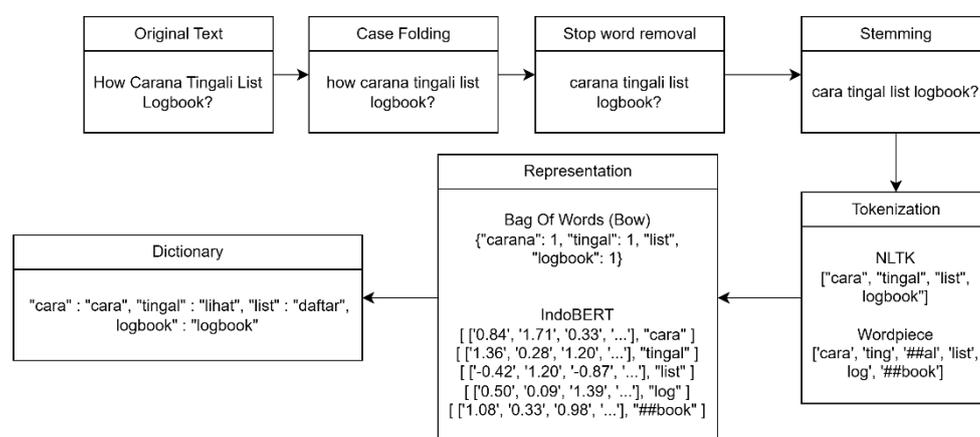
```

Gambar 2. *Dataset intents* dan kamus multibahasa

Selain itu, untuk mendukung penggunaan *chatbot* dalam berbagai bahasa, digunakan kamus multibahasa yang terdiri dari 309 kata, mencakup istilah dalam bahasa Indonesia, Sunda, dan Inggris. Penerjemahan dilakukan dengan memecah pola pertanyaan dari bagian *patterns* dalam data *intents* menjadi kata-kata, yang kemudian diterjemahkan menggunakan

skrip Python dan pustaka *deep_translator*. Hasil terjemahan disimpan dalam *file* JSON, sehingga pengguna dapat memasukkan *input* dalam ketiga bahasa untuk komunikasi yang efektif dan respons *chatbot* yang lebih akurat.

Pada tahapan *Data Preparation* ini dilakukan proses pembersihan untuk pengembangan *chatbot*, seperti terlihat pada gambar 3. Proses ini dimulai dengan *case folding* untuk mengubah teks menjadi huruf kecil, diikuti oleh *stop word removal* untuk menghapus kata-kata yang tidak signifikan, dan *stemming* untuk mengubah kata ke bentuk dasarnya. Pada tahap *tokenisasi*, NLTK digunakan untuk memecah kalimat menjadi kata-kata individual yang digunakan dalam representasi *BoW*, yang menghitung frekuensi kata untuk memahami kemunculan kata dalam teks. Selanjutnya, *AutoTokenizer* dari model IndoBERT digunakan untuk memecah kata menjadi sub-kata dengan algoritma *WordPiece*. Kamus digunakan untuk menerjemahkan pola pertanyaan, menghasilkan data yang bersih dan relevan, serta meningkatkan kemampuan *chatbot* dalam memahami dan merespons pertanyaan pengguna.



Gambar 3. Alur *data preparation*

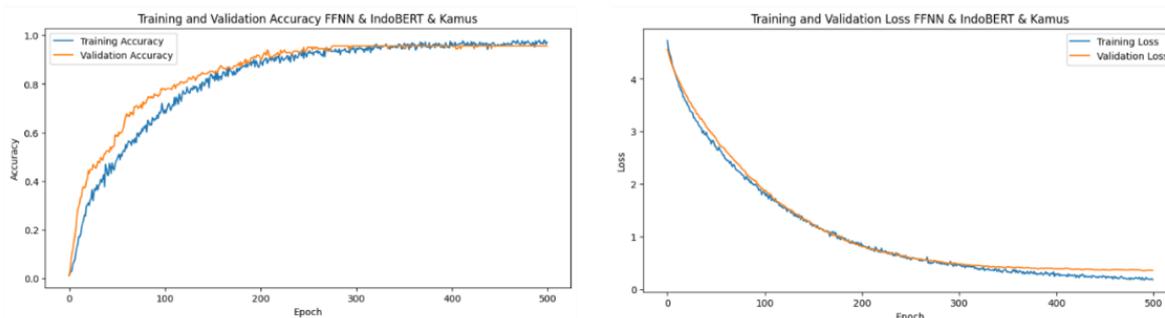
Tabel 1 menunjukkan hasil pengujian model pada *chatbot* DigiTeam. Model FFNN yang menggunakan representasi IndoBERT tanpa integrasi kamus menunjukkan *Train Loss* sebesar 0,5460 dan *Validation Accuracy* sebesar 65,22%. Model ini mampu memahami hubungan kontekstual antar kata dalam bahasa Indonesia, namun masih terbatas dalam menangani variasi bahasa dan istilah lintas bahasa. Setelah integrasi kamus multibahasa, kinerja model mengalami peningkatan signifikan, dengan *Train Loss* yang lebih rendah, yaitu 0,1897, dan *Validation Accuracy* yang mencapai 95,63%. Integrasi kamus multibahasa memperkaya model dengan kemampuan untuk menangani keragaman bahasa dan istilah lintas bahasa, sehingga meningkatkan kemampuan *chatbot* dalam merespons pertanyaan pengguna dengan lebih akurat dalam lingkungan multibahasa.

Tabel 1. Perbandingan hasil pelatihan model pada *epoch* ke-500

<i>Model</i>	<i>Train Loss</i>	<i>Train Acc (%)</i>	<i>Val Loss</i>	<i>Val Acc (%)</i>
FFNN dengan IndoBERT	0,5460	93,96	1,4613	65,22
FFNN, IndoBERT dan Kamus	0,1897	97,12	0,3645	95,63

Hasil pelatihan model yang terintegrasi dengan kamus multibahasa ditunjukkan dalam gambar 4. Model ini mencatat *Train Loss* terendah sebesar 0,1897 pada *epoch* ke-500, dengan *Train Accuracy* mencapai 97,12%, yang menunjukkan kemampuan model dalam meminimalkan kesalahan selama pelatihan dan memprediksi *output* dengan akurat. *Validation Loss* juga menunjukkan hasil yang baik dengan nilai 0,3645, sementara *Validation Accuracy*

mencapai tingkat tertinggi sebesar 95,63%, mengindikasikan kemampuan model dalam generalisasi terhadap data yang belum dilihat sebelumnya. Hasil ini menunjukkan bahwa integrasi kamus multibahasa meningkatkan pemahaman konteks kalimat dan memperbaiki penanganan keragaman bahasa, yang pada akhirnya meningkatkan akurasi dan kemampuan *chatbot* dalam merespons pertanyaan pengguna di lingkungan multibahasa.



Gambar 4. Hasil *accuracy* dan *loss* model ffnn dengan indobert dan kamus

Tabel 2 menyajikan hasil evaluasi model tanpa integrasi kamus dibandingkan dengan model yang telah terintegrasi dengan kamus multibahasa. Model tanpa integrasi kamus mencatat akurasi sebesar 65,22%, presisi 62,76%, *recall* 65,22%, dan F1-score 60,87%. Hasil ini menunjukkan kemampuan dasar model dalam memahami hubungan kontekstual antar kata, tetapi masih kurang optimal dalam menangani keragaman bahasa dan kosakata. Sebaliknya, integrasi kamus multibahasa secara signifikan meningkatkan performa model. Dengan integrasi ini, akurasi meningkat hingga 95,63%, presisi mencapai 98,59%, *recall* sebesar 95,63%, dan F1-score menjadi 95,99%.

Tabel 2. Perbandingan *Confusion Matrix*

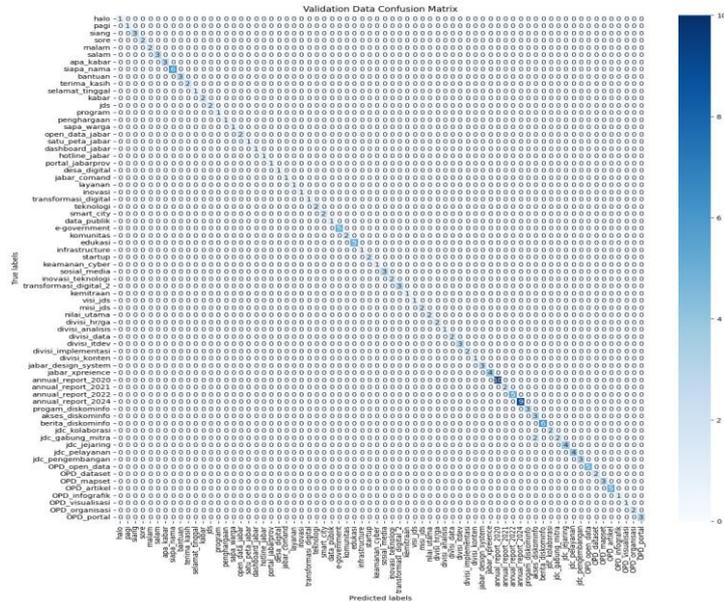
<i>Model</i>	<i>Akurasi (%)</i>	<i>Presisi (%)</i>	<i>Recall (%)</i>	<i>F1-Score (%)</i>
FFNN dengan IndoBERT	65,22	62,76	65,22	60,87
FFNN, IndoBERT dan Kamus	95,63	98,59	95,63	95,99

Peningkatan ini menunjukkan bahwa integrasi kamus multibahasa memperkaya pemahaman model terhadap variasi bahasa dan istilah, sehingga menghasilkan klasifikasi yang lebih akurat dan respons yang relevan. Model dengan presisi yang lebih tinggi menghasilkan lebih banyak prediksi positif yang benar, sementara peningkatan *recall* mencerminkan kemampuan model dalam mengenali dan menangkap lebih banyak contoh sesuai dengan kelas target.

Model yang terintegrasi dengan kamus multibahasa menunjukkan kinerja unggul dengan akurasi 95,63%, presisi 98,59%, *recall* 95,63%, dan *F1-score* 95,99%. Hasil ini mencerminkan kemampuan model yang telah ditingkatkan dalam memahami konteks dan variasi bahasa melalui penggunaan kamus multibahasa. Integrasi kamus tidak hanya memperkaya pemahaman terhadap kosakata lintas bahasa tetapi juga mengurangi kesalahan klasifikasi, seperti yang ditunjukkan dalam gambar 5 melalui *confusion matrix*. Peningkatan presisi, *recall*, dan *F1-score* ini mendukung kemampuan *chatbot* untuk memberikan respons yang lebih akurat dan relevan, sehingga meningkatkan kepuasan dan efektivitas dalam penggunaan nyata melalui pemahaman yang lebih mendalam terhadap kompleksitas bahasa dan hubungan antar kata.

Model dilatih dengan *threshold* 0,75 untuk menyeimbangkan presisi dan *recall* dalam pengenalan bahasa, guna mengurangi *false positives* yang dapat membingungkan pengguna. Integrasi IndoBERT dan kamus multibahasa diharapkan meningkatkan akurasi dan relevansi

respons, terutama pada kalimat dengan variasi bahasa. Tabel 3 menunjukkan bahwa penggunaan kamus multibahasa secara signifikan meningkatkan kualitas respons model.



Gambar 5. Confusion matrix ffnn dan indobert dengan integrasi kamus multibahasa

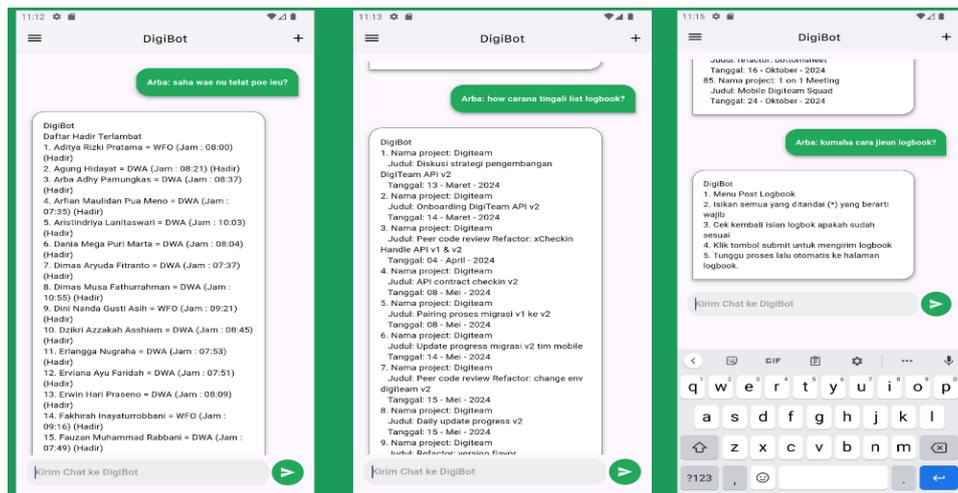
Tabel 3. Hasil pengujian

No.	Kalimat Input	Terjemahan	Probabilitas	Respons
1	How Cara Tingali List Logbook?	None	0,2535	None
		Gimana Cara Lihat Daftar Logbook?	0,9867	Daftar Logbook
		None	0,0839	None
2	Tembongken Information Profile Abdi	Perlihatkan Informasi Profile Saya	0,8253	Data Profile
		None	0,2055	None
3	Saha Wae Nu Telat Poe Ieu?	Siapa Saja Yang Telat Hari Ini?	0,9767	Daftar absen telat
		None	0,3657	None
4	Daptar Logbook Empty	Daftar Logbook Kosong	0,8981	Logbook Kosong

Setelah tahap pengembangan dan pengujian, integrasi chatbot dilakukan dengan menghubungkan model dengan database aplikasi DigiTeam, memungkinkan akses data proyek secara real-time dan memastikan kinerja optimal di perangkat mobile. Penerapan pendekatan Agile memberikan fleksibilitas dan responsivitas tinggi terhadap perubahan kebutuhan pengguna, dengan setiap sprint mingguan yang mengintegrasikan fitur baru dan peningkatan kinerja chatbot secara bertahap. Hasil integrasi yang ditampilkan pada gambar 6 menunjukkan bahwa chatbot dapat memberikan respons yang lebih relevan dan akurat berkat akses langsung ke database DigiTeam, sementara proses pengujian dan evaluasi berkelanjutan memastikan perbaikan yang konsisten, mendukung perkembangan model untuk memenuhi ekspektasi pengguna, serta meningkatkan efisiensi operasional pekerjaan.

Penelitian ini berhasil mengembangkan sebuah sistem chatbot multibahasa untuk aplikasi DigiTeam, yang memberikan kontribusi signifikan dalam meningkatkan akses informasi dan efisiensi operasional di lingkungan kerja. Melalui integrasi kamus multibahasa

dengan model FFNN dan IndoBERT, *chatbot* yang dikembangkan berhasil mencapai tingkat akurasi sebesar 95,63%, serta menunjukkan kemampuan yang superior dalam memahami konteks dan menghasilkan respons yang relevan. Evaluasi yang dilakukan dalam lingkungan aplikasi DigiTeam membuktikan bahwa sistem *chatbot* ini efektif dalam situasi nyata, memungkinkan pengguna untuk mengakses informasi dengan cepat dan akurat, yang berdampak positif terhadap produktivitas pengguna. Temuan ini menunjukkan bahwa *chatbot* ini dapat diterapkan di lingkungan kerja multibahasa, sehingga meningkatkan akses informasi dan efisiensi operasional secara keseluruhan.



Gambar 6. Implementasi dalam aplikasi digiteam

Pembahasan

Penelitian ini menggunakan *dataset* yang terdiri dari 456 pola dan 106 *tag*. Model FFNN dengan representasi IndoBERT menunjukkan kinerja yang baik dengan akurasi pelatihan 93,96% dan *Validation Accuracy* sebesar 65,22%. Meskipun model ini mampu memahami konteks antar kata dalam bahasa Indonesia, ia masih terbatas dalam menangani keragaman bahasa dan istilah lintas bahasa. Hal ini tercermin dalam akurasi yang lebih rendah pada data validasi dibandingkan dengan data pelatihan. Model ini mengandalkan kemampuan IndoBERT untuk mengenali konteks kata dalam kalimat, namun tantangan utama terletak pada variasi bahasa dan istilah yang digunakan dalam lingkungan multibahasa. Kemampuan model *transformer* dalam menangani variasi kata dan frasa dalam berbagai konteks bahasa juga diungkapkan dalam penelitian yang dilakukan oleh Pramana & Romadhony (2024).

Peningkatan yang signifikan terlihat pada model FFNN dengan IndoBERT yang terintegrasi kamus multibahasa. Dengan integrasi kamus multibahasa, *Train Loss* turun menjadi 0,1897 dan *Validation Accuracy* meningkat menjadi 95,63%. Kamus multibahasa memperkaya pemahaman model terhadap variasi kata dan istilah lintas bahasa. Model ini menunjukkan kemampuan lebih baik dalam merespons pertanyaan pengguna yang menggunakan bahasa campuran. Hal ini sejalan dengan penelitian model multilingual BERT yang dapat memproses dan memahami teks dalam beberapa bahasa dengan baik (Astuti et al., 2024).

Dibandingkan dengan penelitian sebelumnya, temuan ini menunjukkan perbedaan signifikan, terutama dalam fleksibilitas dan variasi bahasa yang digunakan. Penelitian ini mengintegrasikan kamus multibahasa ke dalam model FFNN dan IndoBERT, yang memungkinkan *chatbot* untuk memproses variasi bahasa Indonesia, Sunda, dan Inggris, serta menangani interaksi yang lebih dinamis dan kompleks. Sementara itu, penelitian Matalarens & Setyowatie (2023) lebih terbatas pada penggunaan bahasa Indonesia saja dan berfokus pada domain yang sempit, memberikan jawaban yang cenderung statis. Penelitian Dzaky et al.

(2024) menghadapi masalah dengan *dataset* yang kurang dinamis dan beragam, yang mengurangi kemampuannya dalam menangani keragaman bahasa dan konteks yang lebih luas.

Performa model diukur melalui presisi dan *recall*, yang memastikan relevansi dan kemampuan *chatbot* dalam memahami berbagai variasi pertanyaan, sehingga respons yang dihasilkan cepat dan akurat. Hasil penelitian menunjukkan bahwa integrasi kamus multibahasa secara signifikan meningkatkan akurasi dan relevansi *chatbot*, memungkinkan pengguna DigiTeam mengakses informasi dengan lebih cepat. Hal ini mendukung pengambilan keputusan yang lebih efektif dan meningkatkan efisiensi komunikasi serta operasional pekerjaan. Untuk pengembangan selanjutnya, disarankan agar model ini diperluas dengan menambahkan bahasa-bahasa lain yang umum digunakan, serta melakukan pengelolaan kamus multibahasa secara berkala agar tetap relevan dengan dinamika bahasa sehari-hari.

SIMPULAN

Penelitian ini berhasil mengembangkan *chatbot* multibahasa yang efektif untuk aplikasi JDS menggunakan metodologi CRISP-DM dan mencapai akurasi 95,45% melalui integrasi kamus multibahasa terhadap model FFNN dan IndoBERT. Hasil penelitian menunjukkan bahwa model ini memiliki keseimbangan presisi dan *recall* yang optimal, memungkinkan *chatbot* memberikan respons yang cepat dan akurat terhadap berbagai pertanyaan pengguna DigiTeam dalam konteks multibahasa. Peningkatan ini tidak hanya memperkuat efisiensi komunikasi dan operasional pekerjaan di DigiTeam, tetapi juga memperkaya pengalaman pengguna dalam mengakses informasi terkait berbagai proses pekerjaan. Integrasi tersebut menciptakan dasar yang kuat untuk pengembangan lanjutan, seperti penambahan bahasa baru dan pembaruan kamus secara berkala, guna meningkatkan fleksibilitas dan relevansi interaksi di lingkungan kerja yang multibahasa.

REFERENSI

- Aljabar, A., & Karomah, B. M. (2024). Mengungkap Opini Publik: Pendekatan BERT-based-caused untuk Analisis Sentimen pada Komentar Film. *Journal of System and Computer Engineering (JSCE)*, 5(1), 2723–1240. <https://doi.org/10.61628/jsce.v5i1.1060>
- Astuti, W., Wibawa, A. P., Haviluddin, & Herdianti, D. (2024). DIET Classifier Model Analysis for Words Prediction in Academic Chatbot. *ILKOM Jurnal Ilmiah*, 16(1), 59–67. <https://doi.org/10.33096/ilkom.v16i1.1598.59-67>
- Cannavaro, N. (2023). Aplikasi Chatbot untuk Layanan Akademik Menggunakan Platform RASA Open Source dengan Fitur Two Stage Fallback. *Jurnal Ilmu Komputer Dan Informatika*, 3(1), 53–64. <https://doi.org/10.54082/jiki.73>
- Dzaky, A. A., Zeniarja, J., Supriyanto, C., Shidik, G. F., Paramita, C., Subhiyakto, E. R., & Rakasiwi, S. (2024). Optimization Chatbot Services Based on DNN-Bert for Mental Health of University Students. *Journal of Applied Informatics and Computing (JAIC)*, 8(1), 2548–6861. <https://doi.org/10.30871/jaic.v8i1.7403>
- Faurina, R., Revanza, D., & Sopran, A. (2023). Pengembangan Chatbot Menggunakan Deep Feed-Forward Neural Network sebagai Pusat Layanan Informasi Akademik. *Jurnal Eksplora Informatika*, 11(2), 120–129. <https://doi.org/10.30864/eksplora.v11i2.833>
- Fidiyanti, F., Subagja, A. R., Wachyu, R. P., & Madiistriyatno, H. (2023). Analisis Strategi Pengembangan Bisnis Menggunakan Teknologi Artificial Intelligence. *Journal of Comprehensive Science (JCS)*, 2(7), 1994–2001. <https://doi.org/10.59188/jcs.v2i7.425>
- Guntara, R. G. (2022). Aplikasi Chatbot Konsultan Bisnis untuk UMKM Berbasis Dialogflow pada Platform Android. *Indonesian Journal of Digital Business*, 2(1), 9–16. <https://doi.org/10.17509/ijdb.v2i1.49405>
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. In D. Scott,

- N. Bel, & C. Zong (Eds.), *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 757–770). International Committee on Computational Linguistics. <https://doi.org/10.18653/v1/2020.coling-main.66>
- Matalarens, A., & Setyowatie, D. (2023). Perancangan Aplikasi Chatbot FAQ Berbasis Aplikasi Android. *Prosiding Seminar Nasional Pemanfaatan Sains Dan Teknologi Informasi, 1*(1), 281–286.
- Mustakim, F., aFauziah, F., & Hayati, N. (2021). Algoritma Artificial Neural Network pada Text-based Chatbot Frequently Asked Question (FAQ) Web Kuliah Universitas Nasional. *Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi), 5*(4), 438–446. <https://doi.org/10.35870/jtik.v5i4.261>
- Owen, C., & Suhartono, D. (2024). Intent Classification in Artificial Intelligence-Based Customer Service Chatbot for E-Wallet Service Providers. *International Journal of Computing and Digital Systems, 17*(1), 1–10.
- Pambudi, A., Abidin, Z., & Permata. (2023). Penerapan Crisp-Dm Menggunakan Mlr K-Fold Pada Data Saham Pt. Telkom Indonesia (Persero) Tbk (Tlkm)(Studi Kasus: Bursa Efek Indonesia Tahun 2015-2022). *Jurnal Data Mining Dan Sistem Informasi, 4*(1), 1–14. <https://doi.org/10.33365/jdmsi.v4i1.2462>
- Pramana, R. A., & Romadhony, A. (2024). Identifikasi Similar Question dengan IndoBERT (Studi Kasus Dataset QAS Covid-19). *Jurnal Penelitian Informatika, 2*(1), 12–17. <https://doi.org/10.25124/logic.v2i1.7437>
- Pratama, R., Pradana, F. A., Chandra, M., & Bonita, A. (2021). Chatbot interaksi rumah sakit menggunakan FFNN. *Indonesian Journal of Data and Science, 2*(2), 62–68. <https://doi.org/10.56705/ijodas.v3i1.36>
- Pratiwi, A. S., Siswono, G. O., & Saputri, P. D. (2023). Financial Distress Classification Using Feedforward Neural Network Based on Altman and Ohlson Financial Ratios. *Jurnal Matematika, Statistika Dan Komputasi, 20*(1), 184–195. <https://doi.org/10.20956/j.v20i1.27742>
- Putri, T. E., & Ramadhan, G. (2024). Penerapan Chatbot sebagai Alat Pembelajaran untuk Pengembangan Pendidikan Karakter. *Indonesian Journal of Computer Science and Engineering, 1*(01), 32–38.
- Setyasih, E. T. (2022). Transformasi Digital Pemerintah Daerah Di Era Society 5.0: Studi Kasus Di Provinsi Jawa Barat. *PAPATUNG: Jurnal Ilmu Administrasi Publik, Pemerintahan Dan Politik, 5*(3), 59–66. <https://doi.org/10.54783/japp.v5i3.657>
- Simanjuntak, A., Lumbantoruan, R., Sianipar, K., Gultom, R., Simaremare, M., Situmeang, S., & Pangabeang, E. (2024). Research and Analysis of IndoBERT Hyperparameter Tuning in Fake News Detection. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi, 13*(1), 60–67. <https://doi.org/10.22146/jnteti.v13i1.8532>
- Syarof, H. I., & Rasal, I. (2024). Aplikasi Chatbot sebagai Layanan Informasi Virtual pada Website Infinite Learning. *Edumatic: Jurnal Pendidikan Informatika, 8*(1), 56–64. <https://doi.org/10.29408/edumatic.v8i1.25215>
- Widiansyah, M., Az-Zahra, F. F., & Pambudi, A. (2024). Fine-Tuning Model Indobert (Indonesian Bidirectional Encoder Representations from Transformers) untuk Analisis Sentimen Berbasis Aspek pada Aplikasi M-Paspor. *Journal of Informatic Engineering (JOUTICA), 9*(2), 183–195.