



Fine-Tuning Model Indobert (*Indonesian Bidirectional Encoder Representations from Transformers*) untuk Analisis Sentimen Berbasis Aspek pada Aplikasi M-Paspor

Muhammad Widiansyah^{1*}, Fathia Frazna Az-Zahra², Agung Pambudi³.

¹Teknik Informatika, Universitas Muhammadiyah Sukabumi,
Jl. R. Syamsudin, S.H. No. 50 43113, Kota Sukabumi Jawa Barat, Indonesia
Widiansyah1@ummi.ac.id

²Teknik Informatika, , Universitas Muhammadiyah Sukabumi,
Jl. R. Syamsudin, S.H. No. 50 43113, Kota Sukabumi Jawa Barat, Indonesia
fathiafrazna@ummi.ac.id

³Teknik Informatika, Universitas Muhammadiyah Sukabumi,
Jl. R. Syamsudin, S.H. No. 50 43113, Kota Sukabumi Jawa Barat, Indonesia
agungpambd@ummi.ac.id

*Corresponding Author

Abstrak:

M-paspor adalah aplikasi untuk membantu masyarakat dalam proses permohonan paspor. Di Google Play Store aplikasi ini telah diunduh lebih dari 1 juta pengguna dengan ulasan yang diperoleh sebanyak 29 ribu. Data ulasan ini dapat dimanfaatkan untuk mengetahui kelebihan dan kekurangan aplikasi berdasarkan pengalaman nyata. Dengan Teknik analisis sentiemen berbasis aspek, ulasan pengguna dapat dimanfaatkan untuk mengevaluasi aplikasi dan pengembangan aplikasi. Sehingga kualitas layanan aplikasi dapat meningkat. Penelitian ini bertujuan untuk menganalisis sentimen pada ulasan dengan mengelompokannya berdasarkan aspek-aspek yang diulas oleh pengguna. Teknik pemodelan aspek menggunakan LDA (*Latent Dirichlet Allocation*) dan teknik pemodelan sentimen menggunakan IndoBERT(*Indonesian Bidirectional Encoder Representations From Transformers*). Pemodelan aspek menghasilkan *coherence score* tertinggi pada topik ke3 dengan nilai 0.435, sehingga aspek dibagi menjadi 3 kategori yaitu *Usability, Reliability, Efficiency*. Pemodelan sentimen menggunakan IndoBERT menghasilkan nilai *Accuracy* tertinggi pada 94%. Dan hasil evaluasi terhadap kinerja model menggunakan ROC mendapatkan nilai 94%.

Kata Kunci:

M-Paspor, analisis sentimen berbasis aspek , IndoBERT, LDA, *fine-tuning*

Abstract:

M-passport is an application to assist people in the passport application process. On the Google Play Store, this application has been downloaded by more than 1 million users with 29 thousand reviews. This review data can be used to find out the advantages and disadvantages of the application based on real experience. With the aspect-based sentiment analysis technique, user reviews can be leveraged to evaluate the application and application development. So that the quality of application services can improve. This study aims to analyze the sentiment on reviews by grouping them based on the aspects reviewed by users. The aspect modeling technique uses LDA (*Latent Dirichlet Allocation*) and the sentiment modeling technique uses IndoBERT (*Indonesian Bidirectional Encoder Representations From Transformers*). Aspect modeling produced the highest *coherence score* in the 3rd topic with a value of 0.435, so that the aspects were divided into 3 categories, namely *Usability, Reliability, Efficiency*. Sentiment modeling using IndoBERT resulted in the highest accuracy value at 94%. And the results of the evaluation of the model's performance using ROC got a score of 94%.

Keywords:

M-Paspor, Aspect Based Sentiment Analysis, IndoBERT, LDA fine-tuning

1. Pendahuluan

Pada tanggal 27 januari 2022 di Jakarta, Menteri hukum dan HAM RI resmi meluncurkan aplikasi M-Paspor sebagai media untuk membantu masyarakat dalam mengajukan permohonan pembuatan paspor. Masyarakat cukup mengunggah *scan* berkas data yang dibutuhkan ke aplikasi, sehingga tidak perlu datang keimigrasi untuk sekedar mengirimkan berkas, aplikasi ini telah dilengkapi fitur validasi sehingga dapat memastikan dokumen yang diunggah sudah benar. Tidak hanya itu pengguna juga dapat memantau status permohonan paspor melalui aplikasi.

M-Paspor telah tersedia di Google Play Store, Aplikasi di Google Play Store akan memiliki predikat terbaik apabila memiliki riwayat unduhan dan jumlah *rating* tertinggi[1]. aplikasi M-Paspor telah diunduh oleh lebih dari 1juta pengguna dan mendapatkan *rating* 2,4 hal ini menunjukkan aplikasi m-paspor memiliki potensi tetapi masih memiliki tugas untuk meningkatkan kualitas aplikasi dan kepuasan pengguna terhadap aplikasi, sehingga *rating* aplikasi dapat meningkat. Aplikasi M-paspor telah mendapatkan ulasan sebanyak 29 ribu, kolom ulasan sering kali menjadi sarana berbagi informasi pengguna terhadap aplikasi baik mengenai kelebihan maupun kekurangan aplikasi. data ulasan ini dapat dimanfaatkan sebagai acuan untuk mengevaluasi aplikasi berdasarkan pengalaman nyata pengguna, namun dengan jumlah data yang banyak dan bervariasi pengambilan keputusan berdasarkan pengalaman pengguna akan mengalami kesulitan dan menyita banyak waktu. Sehingga diperlukan pemanfaatan berupa ilmu komputer untuk meniru dan mempelajari bagaimana manusia memmenyelesaikan masalah dengan baik[2], [3].

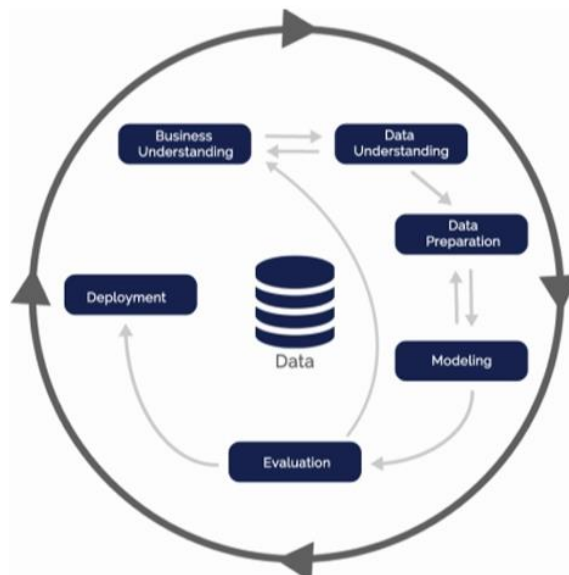
Mengolah data ulasan termasuk kedalam pemrosesan NLP(*Natural Language Processing*) yang memiliki kaitan dengan computer dalam mempelajari , memproses dan menghasilkan Bahasa berdasarkan proses pembelajaran Bahasa alami[4]. Aplikasi M-paspor memiliki ulasan yang banyak dan bervariasi, sehingga untuk memproses ulasan tersebut dapat dilakukan dengan menganalisis sentimen atau menemukan polaritas sebuah teks kemudian mengklasifikasikannya berdasarkan sentimen yang terkandung pada ulasan[5], [6]. Klasifikasi Teks merupakan proses pemberian kelompok dokumen

terhadap kelas yang berbeda [7]. penelitian ini menggunakan pendekatan analisis sentiment berbasis aspek. analisis sentimen berbasis aspek atau ABSA (*Aspect Based Sentiment Analysis*) merupakan teknik menganalisis kalimat yang mengandung emosi atau sentimen dan menghubungkannya dengan aspek-aspek atau topik yang bersangkutan[8]. penelitian sebelumnya telah melakukan pendekatan yang serupa dengan judul Analisis Sentimen Ulasan Aplikasi DLU Ferry Pada Google Play Store Menggunakan Bidirectional Encoder Representations from Transformers[9]. Pada penelitian ini telah menghasilkan nilai akurasi sebesar 86% hal ini tentunya sudah cukup baik, namun, pengelompokan aspek belum diterapkan pada penelitian ini.

Penelitian ini berfokus pada analisis sentimen berbasis aspek pada ulasan aplikasi M-Paspor di Google Play Store. Model untuk pengelompokan topik pada penelitian ini menggunakan *Latent Dirichlet Allocation* atau disebut LDA. LDA digunakan untuk merangkum dan mengelompokkan data menjadi beberapa kelas[10]. Setelah mengelompokkan aspek, klasifikasi sentimen dilakukan menggunakan model IndoBERT. IndoBERT digunakan karena model ini telah melalui pelatihan dengan lebih dari 220 juta data kata Bahasa Indonesia yang bersumber pada Wikipedia (70 juta), artikel-artikel Indonesia(50 juta) dan Corpus Indonesi (90juta)[11]. Untuk menghasilkan nilai kinerja model terbaik akan dilakukan pengujian sekenario dengan menambahkan dan mengatur parameter-parameter untuk pemrosesan data. pemrosesan penelitian ini dilakukan pada *platform* yang menyediakan pustaka python yaitu google collaboratory[12]. Python merupakan bahasa pemrograman yang sederhana, fleksibel dan dinamis[13].

2. Metode

CRISP-DM digunakan agar penelitian ini dapat terstruktur. CRISP-DM (*Cross Industry Standard Process for Data Mining*) merupakan kerangka kerja data *mining* yang memiliki enam fase yaitu *business understanding, data understanding, data preparation, modelling, deployment, dan evaluasi*[14]. Berikut gambaran alur CRISP-DM:



Gambar 1: Metode CRISP-DM[15]

Setiap tahapan pada metode CRISP-DM memiliki peran yang sangat penting, sehingga kerangka kerja ini dapat beroperasi dengan baik. berikut penerapannya dalam metode CRISP-DM:

1. *Business Understanding*

Tahap ini berperan untuk menetapkan tujuan penelitian dalam menganalisis sentimen berbasis aspek pada ulasan M-paspor. Pada tahapan ini diperlukan akan pentingnya dalam memahami dan memanfaatkan data yang akan diolah, Selain itu, tahap ini diperlukan untuk menghasilkan tujuan penelitian, sehingga evaluasi dan peningkatan kinerja aplikasi berdasarkan pengalaman pengguna dapat di *realisasikan*.

2. *Data Understanding*

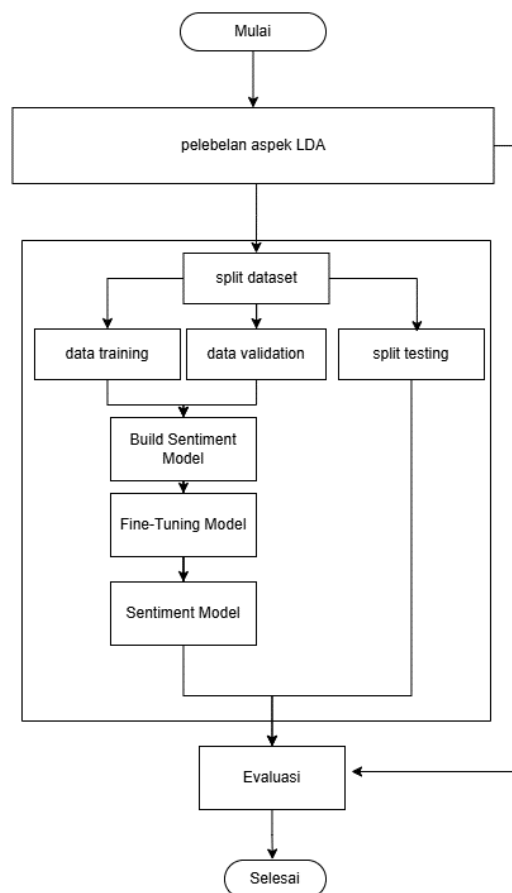
Pemahaman data merupakan fase untuk memahami data yang dimiliki, seperti data disimpan dalam format CSV dimana field yang digunakan hanya *rating* dan ulasan,

3. *Data Preparation*

Pada tahapan ini dilakukan pembersihan data berupa pemilihan dan pengolahan data yang bertujuan untuk pemodelan dapat memberikan hasil yang maksimal. Tahapan ini merubah data mentah dan tidak terstruktur menjadi terstruktur. Didalam tahap data *preparation* terdapat proses *cleaning* data dari noise, angka, tanda baca, link, menghapus karakter berulang, melakukan normalisasi kata baku dan *stop words removal*.

4. *Modeling*

Setelah data dibersihkan selanjutnya dilakukan tahapan *modeling*, tahapan ini terdiri di 2 tahapan pemodelan yaitu aspek *modeling* kemudian *sentiment modeling*. Berikut gamabaranya:



Gambar 2. Alur Tahapan Model

Pada gambar 3. Dilakukan pemodelan terhadap data yang telah dibersihkan, dengan melakukan pelebelan aspek menggunakan LDA data akan dikelompokkan berdasarkan interpretasi data yang

sesuai dari pelatihan model aspek, tidak hanya itu pada tahap ini model akan menentukan jumlah topik atau aspek terbaik berdasarkan data yang digunakan. Setelah data melalui aspek labeling kemudian proses selanjutnya adalah tahapan analisis sentimen, data yang digunakan akan dibagi menjadi 3 sub data, yaitu data Data testing, data validation dan data testing dengan rasio yang ditentukan, selanjutnya melakukan *build sentiment model* pada tahap ini dilakukan Pembangunan model menggunakan IndoBERT dan data akan disesuaikan berdasarkan format IndoBERT sehingga dapat diolah.

5. Evaluasi

Pada tahapan Evaluasi akan dilakukan perhitungan kinerja model menggunakan *Confusion Matrix* dimana akan menghasilkan nilai *accuracy*, *precision*, *Recall* dan F-1 Score, berikut rumus perhitungannya:

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

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

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

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

Untuk memberikan hasil evaluasi kinerja yang maksimal, pada pengujian sentiment dirancang skenario berdasarkan pengaturan parameter-parameter berikut:

Tabel 1: Skenario Uji

Skenario	Model	Rasio pembagian data	Drop out	optimizer	Learning rate
1	IndeoBERT	Data latih 80 % Data uji 10%	0.2 – 0,3	Adam	0.0005
		Data validasi 10% Data latih 90 %		RMSprop	0.00005
2		Data uji 5% Data validasi 5%			

Setelah mendapatkan model terbaik dilakukan visualisasi menggunakan *Wordcloud*, *Wordcloud* akan menunjukkan kata-kata yang sering muncul berdasarkan sentiment pada setiap aspek.

6. Deployment

Pada tahapan *deployment* ini dialukan untuk tujuan pengujian data baru, sehingga data ulasan dapat langsung diproses tanpa perlu memperbaiki atau memperbarui model.

3. Hasil dan Pembahasan

Hasil dan pembahasan untuk analisis sentiemn berbasis aspek pada aplikasi M-Paspor menggunakan pendekatan model LDA dan IndoBERT. data didapatkan menggunakan *package* Google play store, kemudian difilter dan menghilangkan content ulasan dengan *rating* 3 karena tidak memberikan wawasan yang signifikan sehingga dihasilkan data sebanyak 888. Agar menghasilkan

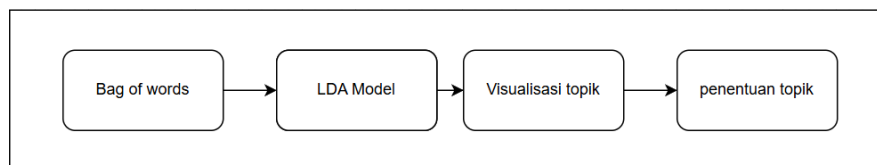
model yang baik maka data harus dipersiapkan sebelum dapat diolah, berikut hasil dari persiapan data sebelum diolah:

Tabel 2: Preprocessing Data

Preprocessing	Content
Data Mentah	Untuk kuota harus sering2 liat sosmed sama aplikasinya, jadi sangat menguras waktu. Kan gak mungkin kita setiap saat liat aplikasi sama sosmed. Mohon ditingkatkan kembali
data cleaning	Untuk kuota harus sering liat sosmed sama aplikasinya jadi sangat menguras waktu Kan gak mungkin kita setiap saat liat aplikasi sama sosmed Mohon ditingkatkan kembali
Case folding	untuk kuota harus sering liat sosmed sama aplikasinya jadi sangat menguras waktu kan gak mungkin kita setiap saat liat aplikasi sama sosmed mohon ditingkatkan kembali
Tokenizing	Untuk, kuota, harus, sering, liat, sosmed, sama, aplikasinya, jadi, sangat, menguras, waktu, kan, gak, mungkin, kita, setiap, saat, liat, aplikasi, sama, sosmed, mohon, ditingkatkan, kembali
Stopword	kuota, sering, liat, sosmed, sama, aplikasinya, jadi, sangat, menguras, waktu, kan, gak, mungkin, liat, aplikasi, sama, sosmed, mohon, ditingkatkan
Stemming	kuota, sering, liat, sosmed, sama, aplikasi, jadi, sangat, uras, waktu, kan, gak, mungkin, liat, aplikasi, sama, sosmed, mohon, tingkat

3.1 Pemodelan Aspek

pemodelan aspek atau topik dilakukan menggunakan algoritma LDA (*Latent Dirichlet Allocation*) sebagai metode pengelompokan ulasan aplikasi M-paspor, berikut alur pemodelan Aspek :



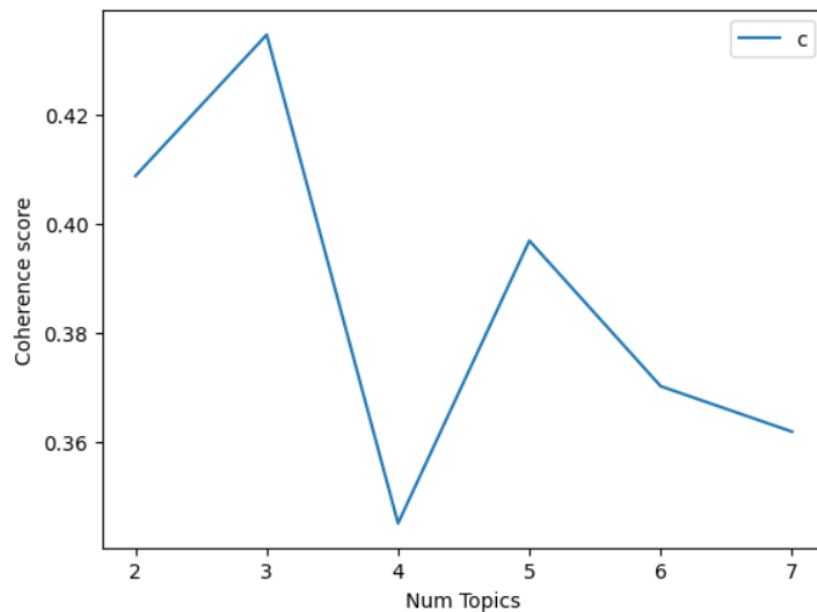
Gambar 3: Alur Pemodelan Aspek

Pada alur diatas dilakukan *Bag of words* atau pembobotan kata menggunakan library `table188`, kemudian dibuat model LDA untuk mengekstrak topik yang terdapat pada ulasan Aplikasi M-Paspor, pada tahapan ini akan menentukan jumlah topik dengan mengevaluasi nilai koherensi, model LDA akan mengidentifikasi pola dari topik, pada penelitian ini penentuan topik dibatasi diawali start dua topik sampai limit tujuh topik dan step satu untuk setiap penambahan topik, berikut hasil dari nilai koherensi untuk setiap topik:

Tabel 3: Nilai Koherensi

Num Topics	Coherence Value
2	0.409
3	0.435
4	0.345
5	0.397
6	0.37
7	0.362

Pada tabel diatas ditunjukkan bahwa nilai koherensi tertinggi terdapat pada topik tiga dengan nilai koherensi 0,435, sehingga jumlah topik yang ditentukan adalah tiga berikut tampilan visualisasi nilai koherensi pada Gambar :

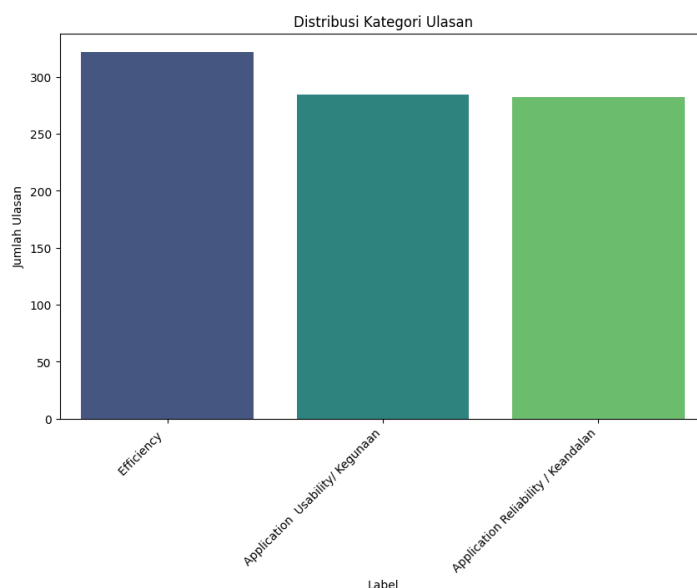


Gambar 4: Grafik Coherence Value

Penentuan nama dan label pada aspek diinterpretasikan berdasarkan ISO 25010:2011 standar internasional dalam pengujian *software* sehingga yaitu *Usability*, *Reliability* dan *Efficiency*. Penentuan label aspek ini pernah dilakukan pada penelitian sebelumnya Penggunaan Latent Dirichlet Allocation (LDA) dan Support Vector Machine (SVM) Untuk Menganalisis Sentimen Berdasarkan Aspek Dalam Ulasan Aplikasi EdLink[10], berikut adalah hasil dari pemodelan topik dan visualisasi sebaran data :

Tabel 4: Interpretasi Topik

Topic	Keywords	Topic Label
0	0.050*"aplikasi" + 0.013*"gak" + 0.012*"daftar" + 0.010*"imigrasi" + 0.010*"data" + 0.009*"masuk" + 0.009*"membantu" + 0.009*"susah" + 0.008*"tanggal" + 0.007*"kode" + 0.007*"aplikasinya" + 0.007*"onlen" + 0.007*"kantor" + 0.006*"email" + 0.006*"pakai"	Application Usability/ Kegunaan
1	0.022*"aplikasi" + 0.014*"gak" + 0.010*"tanggal" + 0.009*"time" + 0.008*"out" + 0.008*"bayar" + 0.008*"email" + 0.008*"paspor" + 0.007*"pembayaran" + 0.007*"bagus" + 0.007*"imigrasi" + 0.007*"request" + 0.006*"kedatangan" + 0.006*"data" + 0.006*"daftar"	Application Reliability / Keandalan Aplikasi
2	0.054*"aplikasi" + 0.012*"data" + 0.011*"tanggal" + 0.011*"gak" + 0.008*"onlen" + 0.008*"paspor" + 0.008*"email" + 0.008*"iya" + 0.007*"aplikasinya" + 0.007*"pakai" + 0.007*"daftar" + 0.007*"susah" + 0.007*"imigrasi" + 0.007*"pembayaran" + 0.006*"banget"	Efficiency



Gambar 5: Distribusi Kategori Ulasan

Untuk mengetahui nilai kinerja dari aspek digunakan model IndoBERT, berikut hasilnya:

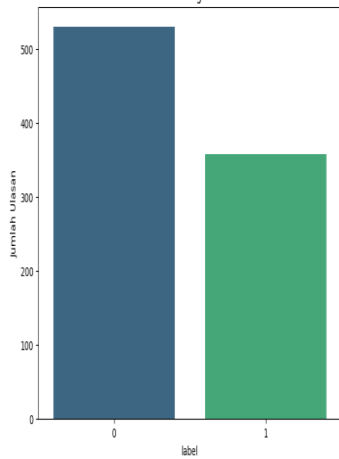
Tabel 5: Akurasi Aspek

Pelebelan	Aspek	Accuracy (%)	Precision(%)	Recall(%)	F1-Score(%)
IndoBERT	Usability	90	80	67	73
	Reliability	83	67	89	76
	Efficiency	85	92	73	81

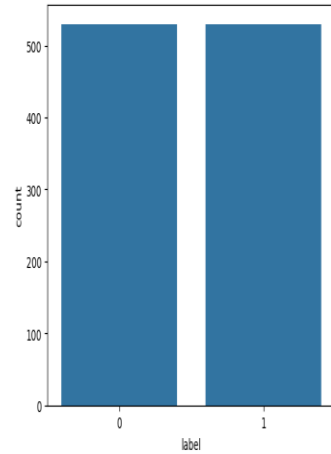
Berdasarkan hasil evaluasi nilai kinerja model dihasilkan nilai *accuracy* pada setiap aspek *Usability* 90%, *Reliability* 83% dan *Efficiency* 85%.

3.2 Pemodelan Analisis Sentimen

pada pelebelan analisis table190ent menggunakan IndoBERT, dalam kinerjanya rasio sebaran data sangat mempengaruhi hasil kinerja sehingga sebaran data harus seimbang, pada gambar dibawah menunjukkan bahwa sebaran data belum seimbang kemudian dilakukan penyeimbangan sebaran data menggunakan *library imbalance*, berikut visualisasi sebaran data:



Gambar 6: Data awal Sebelum *Balancing*



Gambar 7: Data setelah *Balancing*

Pada klasifikasi dilakukan uji coba sekenario dengan menentukan parameter-parameter untuk meningkatkan kinerja hasil, berikut hasilnya:

Tabel 7: Kinerja Hasil

Rasio data	optimizer	Model sekenario	latih		validation		uji				
			Acc (%)	Loss (%)	Acc (%)	Loss (%)	Acc (%)	Loss (%)	Precision (%)	Recall (%)	F-1 (%)
80%	Adam	1	87	28	91	19	94	20	97	94	95
10%	RMSprop		87	31	85	30	86	42	89	84	86
90%	Adam	2	91	23	90	34	87	33	88	74	80
5%	RMSprop		79	45	84	36	76	45	70	75	78

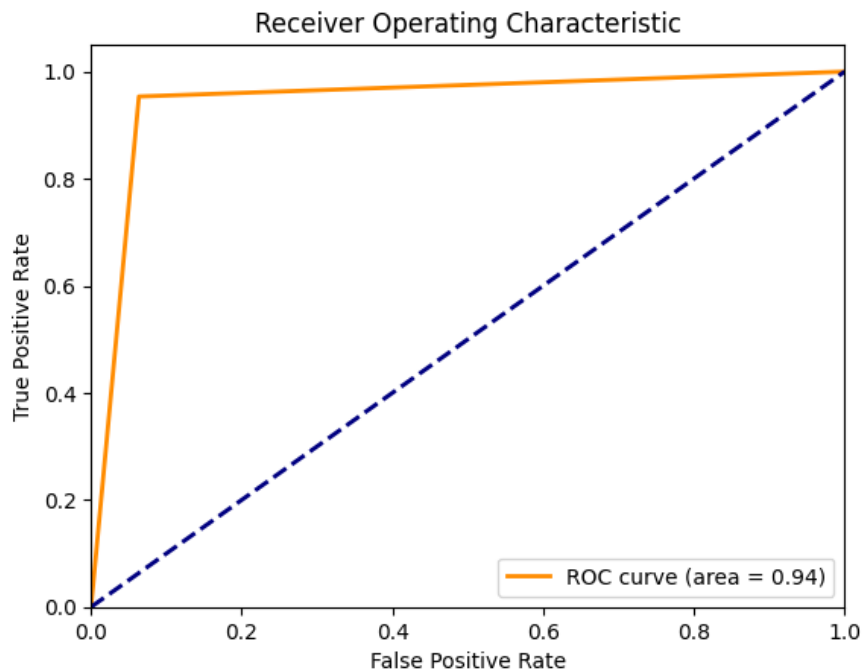
Berdasarkan uji coba pada model sentimen menggunakan IndoBERT terhadap ulasan aplikasi M-paspor mendapatkan hasil akurasi tertinggi pada sekenari pertama dengan rasio 80:10:10, optimizer Adam, dropout 0,3 dan *learning-rate* 5e-5 menghasilkan *accuracy* 94%. Dilakukan evaluasi kinerja model menggunakan confusion matrix berikut gambaranya:

Table 8: Confusion Matrix

59	4
2	41

$$\begin{aligned}
 \text{Accuracy} &= \frac{(41 + 59)}{(59 + 4 + 2 + 41)} &= & 0.94 \\
 \text{precision} &= \frac{59}{59 + 2} &= & 0.97 \\
 \text{recall} &= \frac{59}{59 + 4} &= & 0.94 \\
 \text{F1 Score} &= 2 \times \frac{0.94 \times 0.97}{0.94 + 0.97} &= & 0.95
 \end{aligned}$$

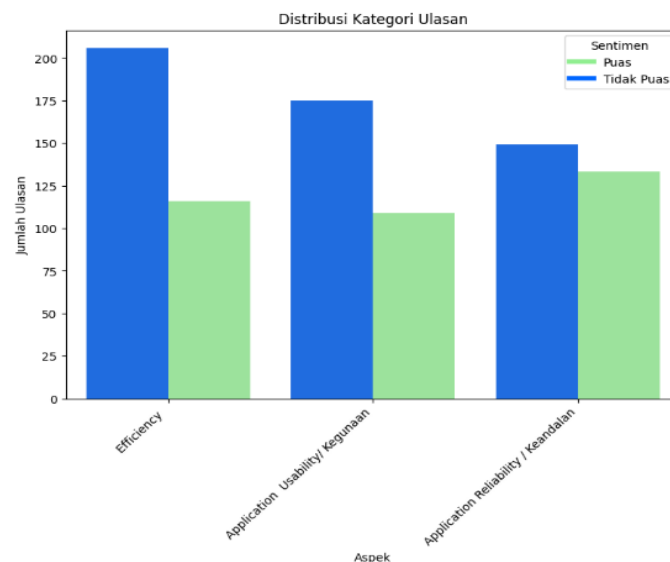
kemudian evaluasi menggunakan ROC *analyst* pada gambar



Gambar 8: Curva ROC

3.3 Interpretasi Data

Berdasarkan pada hasil pemodelan aspek dan sentimen, sehingga dapat diinterpretasikan kedalam bentuk visualisasi hasil, berikut hasil visualisasi hasil model:



Gambar 9: Interpretasi data

Dari pelatihan model akan ditunjukkan visualisasi *Wordcloud* pada setiap aspek dengan sentimen positif dan negatif yang terdapat didalamnya:

- a. Word cloud Aspek usability



Gambar 10: Wordcloud positif aspek usability



Gambar 11 Wordcloud negatif aspek usability

Pada gambar menunjukkan sentimen puas(positif) menunjukkan kata kunci membantu, bagus, terima kasih, aplikasinya menunjukkan bahwa pengguna aplikasi M-paspor merasa terbantu dengan adanya aplikasi ini, mereka berterimakasih dan merasa aplikasi M-paspor ini bagus.gambar menunjukkan sentiment tidak puas(negatif) dengan kata kunci email, daftar, kode otp, aplikasi, Hal ini menunjukkan email pada aspek ini sering kali tidak menerima kode otp atau pengalaman pengguna yang tidak memadai, dalam proses outentikasi.

b. Word cloud Aspek Reliability



Gambar 12: Wordcloud positif aspek Reliability



Gambar 13: Wordcloud negatif aspek Reliability

Pada gambar menunjukkan sentimen puas(positif) menunjukkan kata kunci Aplikasi, paspor, membantu, alhamdulillah, imigrasi. menunjukkan bahwa pengguna merasa aplikasi dapat diandalkan dalam membantu permohonan pembuatan paspor, pengguna bersyukur dengan adanya aplikasi M-Paspor dapat diandalkan dalam proses imigras. gambar menunjukkan sentimen tidak puas(negatif) dengan kata kunci aplikasi, daftar, lambat, data menunjukkan aplikasi sering mengalami masalah teknis, seperti masalah dalam proses daftar, lambat dalam kinerjanya, data yang sering tidak terproses.

c. Word cloud Aspek Efficiency



Gambar 14: Wordcloud positif aspek Efficiency



Gambar 15: Wordcloud negatif aspek Efficiency

Pada gambar menunjukkan sentimen puas(positif) menunjukkan kata kunci aplikasi, paspor, memudahkan, pembuatan, langsung, Menunjukkan aplikasi M-paspor bekerja dengan efisien,

memudahkan proses administrasi, pengurusan dan pembuatan paspor gambar menunjukkan sentimen tidak puas(negatif) dengan kata kunci aplikasi, ribet, tanggal kedatangan menunjukan bahwa sebagian pengguna masih merasa aplikasi kurang efisien dengan berbagai masalah yang dialami, sehingga dianggap rumit, seperti kesulitan pengaturan tanggal.

4. Kesimpulan

Berdasarkan penelitian analisis sentiment berbasis aspek menggunakan model LDA dan IndoBERT, yang telah dilaksanakan berdasarkan skenario didapatkan Kesimpulan sebagai berikut:

1. Rancangan arsitektur model IndoBERT terbaik pada penentuan parameter pembagian data 80% data latih, 10% data validasi dan 10% data uji, penambahan dropout 0.3 atau sebanyak 30%, penggunaan optimizer Adam dan *learning-rate* sebanyak 5e-5 menghasilkan *accuracy* 94%.
2. Berdasarkan pemodelan LDA dihasilkan nilai *Coherence Value* tertinggi 0.453 pada topic ke 3. sehingga ditentukan topik menjadi *Usability* atau Kegunaan, *Reliability* atau Keandalan dan *Efficiency* atau Efisiensi. Dengan hasil acc

Pustaka

- [1] S. Rahayu, Y. MZ, J. E. Bororing, and R. Hadiyat, "Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP," *Edumatic J. Pendidik. Inform.*, vol. 6, no. 1, pp. 98–106, 2022, doi: 10.29408/edumatic.v6i1.5433.
- [2] A. Cucus, R. Y. Endra, and T. Nalrita, "Chatter Bot Untuk Konsultasi Akademik Di Perguruan Tinggi," *Explor. J. Sist. Inf. dan Telemat.*, vol. 10, no. 1, 2019, doi: 10.36448/jsit.v10i1.1214.
- [3] K. R. Ririh, N. Laili, A. Wicaksono, and S. Tsurayya, "Studi Komparasi dan Analisis Swot Pada Implementasi Kecerdasan Buatan (Artificial Intelligence) di Indonesia," *J. Tek. Ind.*, vol. 15, no. 2, pp. 122–133, 2020, [Online]. Available: <https://ejournal.undip.ac.id/index.php/jgti/article/view/29183>
- [4] F. Rumaisa, Y. Puspitarani, A. Rosita, A. Zakiah, and S. Violina, "Penerapan Natural Language Processing (NLP) di bidang pendidikan," *J. Inov. Masy.*, vol. 1, no. 3, pp. 232–235, 2021, doi: 10.33197/jim.vol1.iss3.2021.799.
- [5] A. S. P. Braja and A. Kodar, "Implementasi Fine-Tuning BERT untuk Analisis Sentimen terhadap Review Aplikasi PUBG Mobile di Google Play Store," *J I M P - J. Inform. Merdeka Pasuruan*, vol. 7, no. 3, p. 120, 2023, doi: 10.51213/jjimp.v7i3.779.
- [6] A. Nurian, "Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naive Bayes," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3s1, pp. 829–835, 2023, doi: 10.23960/jitet.v11i3s1.3348.
- [7] Yudi Widhiyana, Transmissia Semiawan, Ilham Gibran Achmad Mudzakir, and Muhammad Randi Noor, "Penerapan Convolutional Long Short-Term Memory untuk Klasifikasi Teks Berita Bahasa Indonesia," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 4, pp. 354–361, 2021, doi: 10.22146/jnteti.v10i4.2438.
- [8] E. Subowo, F. Adi Artanto, I. Putri, and W. Umaedi, "BLTSM untuk analisis sentimen berbasis aspek pada aplikasi belanja online dengan cicilan," *J. Fasilkom*, vol. 12, no. 2, pp. 132–140, 2022.

- [9] E. P. A. Akhmad, "Analisis Sentimen Ulasan Aplikasi DLU Ferry Pada Google Play Store Menggunakan Bidirectional Encoder Representations from Transformers," *J. Apl. Pelayaran Dan Kepelabuhanan*, vol. 13, no. 2, pp. 104–112, 2023, doi: 10.30649/japk.v13i2.94.
- [10] Y. Kustyaningsih, Y. Permana, P. Studi, S. Informasi, U. Trunojoyo, and J. Timur, "Penggunaan Latent Dirichlet Allocation (LDA) dan Support- Vector Machine (SVM) Untuk Menganalisis Sentimen Berdasarkan Aspek Dalam Ulasan Aplikasi EdLink The Use of Latent Dirichlet Allocation (LDA) and Support-Vector Machine (SVM) to Analyze Sent," vol. 13, no. 1, pp. 127–136, 2024, doi: 10.34148/teknika.v13i1.746.
- [11] K. A. Wijaya, A. Romadhony, and ..., "Implementasi Model IndoBERT pada Dashboard Sentimen Media Sosial (Studi Kasus Universitas XYZ)," *eProceedings ...*, vol. 10, no. 4, pp. 3910–3926, 2023, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/20853%0Ahttps://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/download/20853/20321>
- [12] R. G. Guntara, "Visualisasi Data Laporan Penjualan Toko Online Melalui Pendekatan Data Science Menggunakan Google Colab," *J. Ilm. Multidisiplin*, vol. 2, no. 6, pp. 2091–2100, 2023.
- [13] D. A. Budi, "Perancangan Sistem Login pada Aplikasi Berbasis GUI Menggunakan Qtdesigner Python," *J. SIMADA (Sistem Inf. dan Manaj. Basis Data)*, vol. 4, no. 2, pp. 92–100, 2021, doi: 10.30873/simada.v4i2.2961.
- [14] Y. Christian and K. O. Y. R. Qi, "Penerapan K-Means pada Segmentasi Pasar untuk Riset Pemasaran pada Startup Early Stage dengan Menggunakan CRISP-DM," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, p. 966, 2022, doi: 10.30865/jurikom.v9i4.4486.
- [15] Y. Suhandi, I. Kurniati, and S. Norma, "Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik," *J. Teknol. Inform. dan Komput.*, vol. 6, no. 2, pp. 12–20, 2020, doi: 10.37012/jtik.v6i2.299.