



ASPECT-BASED SENTIMENT ANALYSIS PADA APLIKASI PELACAKAN KASUS COVID-19 (STUDI KASUS: PEDULILINDUNGI)

Suryatin¹, Dhomas Hatta Fudholi², Chandra Kusuma Dewa³, Nur Iman⁴

¹19917016@students.uui.ac.id, ²hattafudholi@uui.ac.id, ³Chandra.kusuma@uui.ac.id,

⁴inang.itc88@gmail.com

^{1,2,3}Universitas Islam Indonesia, ⁴STMIK Bina Bangsa Kendari

Abstrak

Berbagi pengalaman melalui internet dan media sosial dapat menunjukkan sikap dan perasaan dalam bentuk umpan balik. Aplikasi publik yang banyak disoroti pada masa wabah corona virus yaitu aplikasi pedulilindungi yang merupakan aplikasi monitoring perkembangan Corona Virus Disease di Indonesia. Salah satu fenomena timbulnya *Aspect-based sentiment* dalam pada perilaku sentimentil masyarakat terhadap layanan aplikasi pedulilindungi. Penelitian ini bertujuan untuk mengetahui besarnya nilai sentimen pada layanan pedulilindungi dan berfokus pada *aspect based sentiment analysis* (ABSA) pada domain ulasan aplikasi pemerintah. Analisis terdiri dari *user interface*, *user experience*, fungsionalitas dan *work security*. Metode yang digunakan meliputi klasifikasi sentimen dan aspek dengan metode *deep learning* (CNN, GRU dan TCN). Data primer bersumber dari hasil ulasan aplikasi pedulilindungi dengan teknik *scraping* pada situs <https://www.pedulilindungi.id/>. Hasil penelitian menunjukkan bahwa terdapat enam aspek klasifikasi sentimen pada aplikasi pedulilindungi yaitu aplikasi, *user interface*, *user experience*, kode OTP, cek sertifikat vaksin, bukti akses layanan. Hasil penelitian juga menunjukkan bahwa metode CNN memperoleh nilai skor akurasi terbaik pada klasifikasi sentimen sebesar 98% dan klasifikasi aspek sebesar 97%.

Kata kunci: *Aspect Based Sentiment Analysis*, Aplikasi pedulilindungi, *Deep Learning*, Covid-19

Abstract

Sharing experiences via the internet and social media can show attitudes and feelings in the form of feedback. The public application that was widely highlighted during the corona virus outbreak was the careprotect application, which is an application for monitoring the development of Corona Virus Disease in Indonesia. One of the phenomena is the emergence of Aspect-based sentiment in people's sentimental behavior towards the Carelindungi application service. This research aims to determine the magnitude of the sentiment value for protective care services and focuses on aspect based sentiment analysis (ABSA) in the government application review domain. The analysis consists of user interface, user experience, functionality and work security. The methods used include sentiment and aspect classification using deep learning methods (CNN, GRU and TCN). Primary data comes from the results of a review of the Carelindungi application using scraping techniques on the site <https://www.pedulilindungi.id/>. The research results show that there are six aspects of sentiment classification in the Carelindungi application, namely application, user interface, user experience, OTP code, vaccine certificate check, proof of service access. The research results also show that the CNN method obtained the best accuracy score in sentiment classification of 98% and aspect classification of 97%.

Keywords: *Aspect Based Sentiment Analysis*, Careprotection Application, *Deep Learning*, Covid-19

1. Pendahuluan

Aplikasi pedulilindungi merupakan aplikasi yang dirancang Kementerian Komunikasi dan Informatika (kominfo) dan Kementerian BUMN untuk digunakan Kementerian Kesehatan dan Gugus Tugas dengan tujuan untuk memonitor *Coronavirus disease* (Covid-19) di Indonesia. Aplikasi pedulilindungi dirilis pada 28 Maret 2020, di *update* pada 8 Februari 2022 dan sudah di unduh sebanyak lebih dari 50 juta pengguna. Pedulilindungi adalah aplikasi yang dikembangkan untuk menghentikan penelusuran Covid-19 dengan mengandalkan kepedulian dan partisipasi masyarakat untuk saling membagikan data lokasinya saat bepergian agar penelusuran riwayat kontak dengan penderita Covid-19

dapat dilakukan dengan menggunakan fitur *bluetooth* untuk merekam informasi yang dibutuhkan dapat dilakukan pertukaran data terjadi ketika ada *gadget* lain dengan menggunakan fitur *bluetooth* yang terdaftar di pedulilindungi. Selanjutnya pedulilindungi dapat mengidentifikasi apakah seseorang berada di zona aman atau zona merah ketika berada di keramaian.

Kini aplikasi pedulilindungi menjadi penting keberadaannya dan wajib dimiliki oleh masyarakat terutama bagi pelaku perjalanan di Indonesia. Fungsi dari aplikasi pedulilindungi untuk menyimpan dokumen persyaratan perjalanan secara digital atau *password* kesehatan bagi pelaku perjalanan yang telah melakukan dosis vaksin. Pemerintah di Indonesia memberikan layanan berkualitas yang efisien, hemat biaya, dan adil secara sosial dalam menjaga kebutuhan publik di pusat layanan yang diberikan. Sumber digital seperti ulasan yang dibagikan pada pedulilindungi dapat dimanfaatkan sebagai sisi pengembangan pelayanan untuk mengukur kepuasan atas pelayanan yang diberikan. Untuk mengetahui nilai sentimen pada layanan aplikasi pedulilindungi penelitian ini berfokus pada *Aspect Based Sentiment Analysis* (ABSA) pada domain ulasan aplikasi pemerintah.

Analisis sentimen berbasis aspek merupakan analisis sentimen berdasarkan pada aspek tertentu yang memiliki kemampuan nilai sentimen di setiap aspek. Tugas ABSA dalam penelitian [1] menggunakan *post tagger* untuk mengekstrak aspek sebelum proses klasifikasi sentimen. Penelitian ABSA dengan dua sub tugas ekstraksi aspek dan klasifikasi polaritas sentimen menggunakan *Convolutional Neural Networks* (CNN) dengan membandingkan *JASen* terhadap metode *baseline* pada klasifikasi sentimen pada dua kumpulan data *restaruan* dan laptop *JASen* mengungguli semua *baseline* dari kedua kumpulan data memperoleh nilai skor akurasi sebesar 83.83% data *restaurant* 71.01 data laptop untuk klasifikasi aspek, 81.96 data *restaurant* 74.59 data laptop untuk klasifikasi sentimen [2]. Selain itu sentimen analisis berbasis aspek dilakukan dengan menggunakan proses klasifikasi diseluruh aspek [3]. Di beberapa penelitian proses ekstraksi tidak dilakukan karena membutuhkan sumberdaya yang cukup besar [4][5]. Penelitian *analisis sentimen berbasis aspek* (ABSA) dilakukan dengan berbagai metode. ABSA telah diaplikasikan pada domain peningkatan *design product* pada ulasan konsumen tentang produk laptop [6].

Sebuah penelitian menggunakan perbandingan tiga arsitektur *deep learning* yaitu CNN, RNN, dan LSTM akurasi terbaik menunjukkan model LSTM sebesar 93% [7]. ABSA pada domain ulasan data restoran berbahasa Indonesia dengan kombinasi metode BERT-CNN, ELMO-CNN, dan Word2vec-CNN hasil terbaik ditunjukkan pada model ELMO-CNN dengan nilai *micro-average precision* sebesar 0.88%, *micro average recall* sebesar 84%, dan *f1-score* sebesar 0.86% untuk klasifikasi sentimen memberikan hasil terbaik pada BERT-CNN dengan nilai *precision* sebesar 0.89%, *recall* sebesar 0.86 dan *f1-score* sebesar 0.91 [8]. Penelitian serupa ulasan restoran berbahasa Albanian terhadap klasifikasi level aspek dan sentimen pada 2.500 ulasan data restoran menggunakan pendekatan *unsupervised learning* pemodelan algoritma *Non-Negative Matrix factoria* (NMF) dengan skor akurasi 0.81% [9].

Penelitian lain menggunakan empat arsitektur *machine learning* *random forest*, *svm*, *hybrid classifier*, dan *naïve bayes* dengan jumlah *dataset* 1.500 data ulasan hotel berbahasa inggris nilai terbaik rata-rata akurasi 84% ditunjukkan pada model *hybrid classifier* [10]. Beberapa penelitian lain telah berhasil menunjukkan bahwa algoritma *deep learning* pada sentimen analisis dengan menerapkan metode CNN sebagai metode klasifikasi dengan jumlah data 3.051 ulasan akurasi hasil terbaik sebesar 86% pada saat *epoch* ke 100 [11]. Penelitian [12] berhasil menerapkan metode TCN menggunakan *dataset* *sequence modeling task Peen Treebank* pada dua domain data *seq.MNIST* dan *Permuted MNIST* perolehan akurasi pada MNIST sebesar 99.0% dan 97.2% P-MNIST.

Pada penelitian [13] melakukan ABSA pada domain hotel menggunakan dua pendekatan *deep learning* CNN dan CNN-BiLSTM dengan pemilihan parameter terbaik model CNN memberikan hasil terbaik sebesar 95.918% pada epoch ke 15. ABSA dengan kombinasi CNN dan GRU berdasarkan kemampuan ekspansi domainnya. Penelitian ini menggunakan 2 ulasan pada *dataset* yang berbeda yaitu hotel dan otomotif mobil, model dalam penelitian ini mencapai hasil yang baik pada analisis aspek di bidang otomotif namun identifikasi pada polaritas sentimen sedikit lebih buruk daripada di bidang hotel, kedua sub tugas ini mencapai 83.06% dan 83.42% pada *dataset* hotel serta 86.93% dan 77.42% pada *dataset* mobil [14].

Aspect Based Sentiment Analysis (ABSA) banyak dilakukan oleh para peneliti di berbagai domain seperti ulasan hotel, restoran, produk ulasan, film, serta bahasa. Namun ABSA pada domain ulasan aplikasi pemerintah masih terbilang langka. Sudut pandang publik terhadap internet memiliki kontribusi

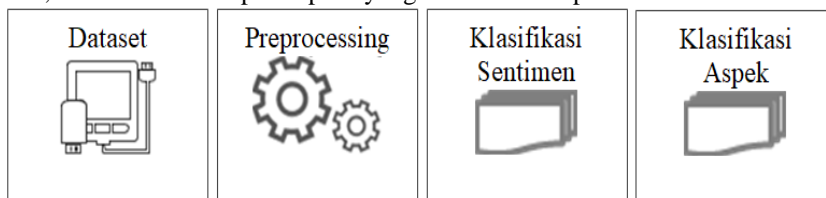
yang besar dalam pertumbuhan kehidupan sehari-hari. Hal ini menunjukkan bahwa pengguna situs web dapat mengemukakan pendapat dan umpan balik. Sentimen analisis adalah salah satu upaya untuk membantu dalam memahami individu sikap terhadap subjek tertentu pada bidang yang berbeda. Selanjutnya, ini dapat membantu dalam mencapai informasi mengenai kualitas manfaat suatu produk dalam mengambil keputusan terkait dengan aspek penting dari layanan [15].

Penelitian terkait domain aplikasi pemerintah, penelitian ini menerapkan pendekatan SVM dengan memanfaatkan model *lexicon* untuk mengekstrak aspek sebelum opini dengan mengklasifikasi data sebanyak 7.345 data keseluruhan penelitian tersebut menggunakan lima kelas aspek pada domain aplikasi seluler pemerintah seperti *User Interface (UI)*, *User Experience (UE)*, *Functionality and Performance (F&P)*, *Security (SEC)*, dan *Support and Updates (S&U)* dengan lima kelas polaritas sentimen sangat *positive*, *positive*, netral, *negative*, dan sangat *negative* yang terdiri dari dua elemen evaluasi kinerja aspek eksplisit dengan nilai akurasi sebesar 86.83% dan aspek implisit akurasi sebesar 93.01% [16]. Penelitian ini mencoba untuk menerapkan ABSA dengan mengangkat dua kategori aspek *user interface* dan *user experience* dari penelitian [16] serta menambahkan beberapa kategori aspek diantaranya yaitu aplikasi, kode OTP, kartu vaksin, dan akses layanan.

Dari latar belakang permasalahan yang dijabarkan, penelitian ini mencoba untuk membuat analisis sentimen berbasis aspek (ABSA) menggunakan data pada ulasan aplikasi seluler pemerintah dengan tujuan untuk mendalami teknik *deep learning* dalam meningkatkan kinerja ABSA pada domain ulasan aplikasi pedulilindungi. Di sisi lain, penelitian ini menggunakan pendekatan *Convolutional Neural Network (CNN)*, *Gated Recurrent Unit (GRU)* dan *Temporal Convolutional Network (TCN)*, yang merupakan pendekatan teknik *deep learning*. Model ini di pilih karena terbukti dalam mengatasi permasalahan teks klasifikasi dengan sangat baik [17].

2. Metode

Proses klasifikasi sentimen berbasis aspek memiliki langkah yang beragam dari masing-masing penelitian. Alur penelitian yang dilakukan dalam penelitian ini terdiri dari *dataset*, *preprocessing*, klasifikasi sentimen, dan klasifikasi aspek seperti yang diilustrasikan pada Gambar 1.



Gambar 1. Alur Penelitian

Deskripsi Alur:

a. Dataset

Proses pengumpulan data dengan memanfaatkan beberapa *library* yang tersedia di *python*, lingkup penelitian mencakup aplikasi seluler pemerintah (Pedulilindungi). Pengambilan data dilakukan dengan cara teknik *scraping* data ulasan melalui *Google Play Store*. Sumber data dari <https://www.pedulilindungi.id/>. Ulasan tersebut dilabeli berdasarkan *rating* pengguna terhadap kepuasan layanan (1-5) dipilih dalam eksperimen ini dimana rating (1-3) mengungkapkan sentimen *negative* sedangkan (4-5) mengungkapkan sentimen *positive*. Penentuan kategori aspek penelitian ini mengadopsi dua kategori aspek dari penelitian [15] yaitu *user interface* dan *user experience* serta menambahkan beberapa kategori aspek diantaranya yaitu aplikasi, kode OTP, kartu vaksin, dan akses layanan. Contoh data pelatihan ditunjukkan pada Tabel 1 dan Deskripsi kategori aspek ditunjukkan pada Tabel 2. Pemberian kategori aspek tersebut di verifikasi secara manual bersama seorang ahli di bidang *data mining*.

Tabel 1. Data pelatihan

Konten	Rating	Aspek	Sentimen
Aplikasi sampah tidak terdeteksi lokasi kota hanya terdeteksi 1 wilayah payah bikin aplikasi asal banget developer sorry uninstall aplikasi kasih bintang 1 aplikasi tidak membantu	1	Aplikasi	Negative
Sebagian fitur tidak berfungsi, masa mau login isi tanggal lahir tidak bisa gimana ini	2	User Interface	Negative
Vaksin pertama ada aplikasi muncul sertifikat vaksin2 tidak muncul bingung	2	User experience	Negative
Server nya lambat respon kode otp lelet mohon dikirim cepat otpnya	1	Kode OTP (One time password)	Negative
Membantu proses unduh sertifikat vaksin 1 n 2 cepat	5	Kartu Vaksin	Positive
Call center sulit hubungi sertifikat vaksin muncul lambat layanan	2	Akses layanan	Negative

Tabel 2. Kategori Aspek

Aspek	Deskripsi
Aplikasi	Aspek ini merujuk pada kegunaan dan kemampuan dari app tracking Covid-19
User Interface	Aspek ini menyampaikan fungsionalitas fitur antarmuka pengguna
User Experience	Aspek ini merujuk pada pengalaman pengguna dalam mengaplikasikan perangkat lunak (pedulilindungi)
Kode OTP	Aspek ini terkait dengan kode verifikasi dalam membuat akun pedulilindungi
Akses Kartu Vaksin	Unduh/Cek kartu vaksin apa bila telah melakukan dosis vaksin
Akses Layanan	Aspek ini merujuk pada pemenuhan asas pelayanan publik (Kepentingan umum, keseimbangan hak, keprofesionalan dan partisipasi terhadap tanggap keluhan)

b. Preprocessing

Setelah penentuan kategori aspek selanjutnya *preprocessing*. Tahap *preprocessing* dilakukan dalam data teks guna meminimalisir *noise* pada data teks [18]. Hasil *preprocessing* ditunjukkan pada Tabel 3. Tahapan *preprocessing* penelitian ini terdiri dari:

1. *Casefolding* untuk menyeragamkan huruf menjadi huruf kecil.
2. *Stemming* merupakan tahap menghilangkan imbuhan atau mengubah kata ke dalam kata dasar.
3. *Normalization* tahapan ini berfungsi untuk mengubah kata yang tidak baku menjadi baku.
4. *Remove stop-word* tahapan ini untuk menghilangkan *term* yang tidak memiliki arti dan memperbaiki data yang tidak konsisten.

Tabel 3. Hasil *Preprocessing*

Teks Sebelum	Teks Setelah
Aplikasi sampah tidak terdeteksi lokasi kota hanya terdeteksi 1 wilayah payah bikin aplikasi asal banget developer sorry uninstall aplikasi kasih bintang 1 aplikasi tidak membantu	aplikasi sampah tidak deteksi lokasi kota hanya deteksi 1 wilayah payah bikin aplikasi asal banget developer sorry uninstall aplikasi kasih bintang 1 aplikasi tidak bantu
Sebagian fitur tidak berfungsi mau login isi tanggal lahir tidak bisa, gimana ini	sebagian fitur tidak fungsi mau login isi tanggal lahir tidak bisa gimana ini
Vaksin pertama ada aplikasi muncul sertifikat vaksin2 tidak muncul bingung	vaksin pertama ada aplikasi muncul sertifikat vaksin2 tidak muncul bingung
Servernya lambat respon kode otp lelet mohon dikirim cepat otpnya	server lambat respon kode otp lelet mohon kirim otp
Membantu proses unduh sertifikat vaksin 1 n 2 cepat	bantu proses unduh sertifikat vaksin 1 n 2 cepat
Call center sulit dihubungi sertifikat vaksin muncul lambat layanan	Call center sulit hubungi sertifikat vaksin muncul lambat layan

Setelah *preprocessing* selanjutnya adalah klasifikasi sentimen dan klasifikasi aspek. Pada klasifikasi sentimen penelitian ini memberikan dua label sentimen *positive* dan *negative*, sedangkan untuk aspek penelitian ini memetakan enam kategori aspek “aplikasi, *user interface*, *user experience*, kode otp, kartu vaksin, dan akses layanan”. Dalam pengerjaan tugas klasifikasi penulis membagi dataset menjadi dua subset, set pelatihan dan pengujian menggunakan *split* 80/20 di mana rasio 80% untuk melatih model dan 20% untuk menguji model pemisahan ini dipilih karena umum digunakan.

Untuk evaluasi penulis menggunakan *confusion matrix* yang terdiri dari *true positive*, *true negative*, dan *false positive false negative*. Untuk menentukan nilai frekuensi sebuah kata menggunakan TF-IDF solusi pada penelitian ini merujuk pada karya [18]. Serta menggunakan *tuning hyperparameter* yang ditunjukkan pada Tabel 4 untuk memperoleh hasil terbaik [13].

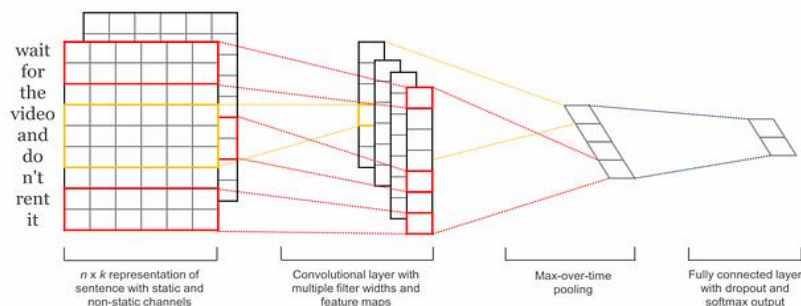
Untuk fungsi *loss* peneliti menggunakan *categorical crossentropy* dan aktivasi *ReLU* untuk memetakan 0 pada nilai *negative* dan akan bertambah apabila *positive*. Di akhir *network* penulis menambahkan *softmax* untuk klasifikasi aspek. Untuk mencapai *threshold* peneliti menerapkan *custom callbacks* dan *fungsi early stopping* untuk memantau *validation loss*. Penelitian ini menggunakan perbandingan tiga metode *deep learning* (CNN, GRU, dan TCN).

Tabel 4. Tuning Hyperparameter

Model	Klasifikasi Sentimen			Klasifikasi Aspek		
	CNN	GRU	TCN	CNN	GRU	TCN
Filter	128	128	-	128	128	-
Kernel	5	-	64	5	-	64
Hidden dims	100	100	100	100	100	100
Input length	193	193	193	193	193	193
Drop out	-	0.1	0.1	-	0.1	0.1
Batch size	16	16	16	64	16	16
epoch	3	3	10	3	3	10

2.1. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan salah satu jenis algoritma *deep learning* yang mampu mengatasi permasalahan seperti *object detection*, *image classification*, maupun *video classification*. CNN sendiri sangat umum digunakan pada data *image*. Meskipun CNN telah banyak diaplikasikan pada *image classification*, CNN juga mampu mengatasi pada permasalahan data teks [7]. CNN terdiri dari beberapa *layer*, *layer* pertama menyimpan kata dalam sebuah *low-dimensional vector*, *layer* kedua menjalankan *convolutions* menggunakan *multiple filter size*, selanjutnya *max-pool* hasil dari *layer convolutional* ke dalam *long feature vector*, menambahkan *fully connected dropout regularization*, dan mengklasifikasi hasil menggunakan *ReLU layer* [21]. Gambar 2 ilustrasi CNN untuk klasifikasi teks.



Gambar 2. CNN Untuk Klasifikasi Teks

2.2. Gate Recurrent Unit (GRU)

GRU *Neural* merupakan sebuah pengembangan dari model LSTM yang dapat mengatasi permasalahan data panjang dengan arsitektur yang lebih sederhana. Berbeda dengan LSTM yang memiliki tiga jenis parameter, GRU mempunyai sedikit parameter (*update gate* dan *reset gate*) sehingga pelatihan dan konvergensinya lebih cepat dari pada LSTM. Di mana *update gate* (r_t) merepresentasikan berapa banyak informasi yang tersimpan, sedangkan pada *reset gate* (z_t) berfungsi untuk menentukan banyaknya informasi yang tidak diperlukan. Kandidat aktivasi $h(t)$ dihitung seperti *recurrent* unit menyimpan informasi sebelum (r_t) digunakan, selanjutnya *hidden state* $\hat{h}(t)$ memori akhir yang akan disimpan dan σ adalah fungsi *tanh* dengan hitungan input tersembunyi saat ini dari langkah waktu sebelumnya untuk mengontrol berapa banyak porsi memori baru dan lama yang harus digabung dalam memori akhir [22]. Representasi matematis pada model dasar GRU ditunjukkan pada persamaan 1, 2, 3, dan persamaan 4.

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r) \quad (1)$$

$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_z) \quad (2)$$

$$\hat{h}(t) = \tanh(W_h[r_t * h_{t-1}, x_t] + b_h) \quad (3)$$

$$h_t = (1 - z_t) * h_t + z_t * \hat{h}(t) \quad (4)$$

2.3. Temporal Neural Network (TCN)

TCN adalah salah satu jenis algoritma *deep learning* yang mampu mengatasi permasalahan data *sequential* menggunakan 1D *fully-convolutional network* pada teks klasifikasi, berbeda dengan CNN, TCN memiliki karakteristik seperti konvolusi dalam arsitektur bersifat kausal yaitu tidak ada informasi kebocoran dari masa depan ke masa lalu, dapat mengambil urutan panjang berapapun, dan memetakan urutan *output* dengan panjang yang sama [12]. Gambar 3 arsitektur TCN yang terdiri dari:

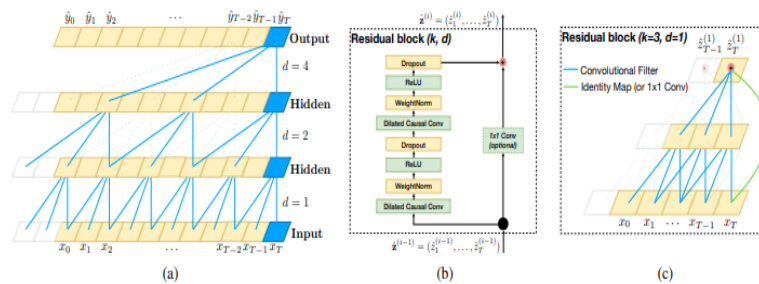
1. *Sequence modelling* sebagai tugas pemodelan urutan *input* yang di inginkan x_0, \dots, x_t dan prediksi *output* yang sesuai y_0, \dots, y_t pada setiap waktu. Kendala utama adalah untuk prediksi keluaran y_t selama beberapa waktu (t) menggunakan input yang diamati sebelumnya: x_0, \dots, x_t secara formal urutan pemodelan jaringan setiap fungsi $f : x^{t+1} \rightarrow y^{t+1}$ itu menghasilkan pemetaan:

$$\hat{y}_0, \dots, \hat{y}_t = f(x_0, \dots, x_t) \quad (5)$$

Jika memenuhi batas kausal bahwa y_t hanya bergantung pada x_0, \dots, x_t dan bukan pada *input* masa depan x_{t+1}, \dots, x_T . Tujuannya untuk menemukan jaringan f yang meminimalkan beberapa kerugian yang diharapkan antara *output* aktual dan prediksi, $L(y_0, \dots, y_t, f(x_0, \dots, x_t))$, di mana urutan dan keluaran berdasarkan distribusi.

2. *Causal convolutiontions* untuk mencapai poin pertama, TCN menggunakan arsitektur jaringan 1D *fully-convolutional network* (FCN) [23] di mana tiap *hidden layer* mempunyai panjang yang sama dengan *input layer* dan panjang *padding* 0 (*kernel size* - 1) ditambah untuk menjaga *hidden* berikutnya yang memiliki panjang sama dengan sebelumnya. Untuk mencapai poin kedua, TCN menggunakan *causal convolutions* di mana *output* pada waktu (t) di konvolusi hanya dengan elemen waktu (t) dan sebelumnya pada *layer* tersembunyi.
Sederhananya: TCN = 1D FCN + konvolusi kausal.
3. *Dilated convolutions* untuk melihat kembali sejarah dengan ukuran *linear* dalam jaringan. Secara lebih formal, untuk 1-D panjang input $X \in R^n$ dan *filter* $f : \{0, \dots, k - 1\} \rightarrow R$, konvolusi yang melebar operasi F pada elemen S yang didefinisikan sebagai:

$$f(s) = (x * f)(s) = \sum_{t=0}^{k-1} f(i) \cdot X_{s-d.i} \quad (6)$$



Gambar 3. Arsitektur TCN

Gambar 3. Elemen arsitektur dalam TCN. (a) *dilated causal convolution* dengan *dilated* faktor $d = 1; 2; 4$ dan ukuran *filter* $k= 3$. Bidang reseptif mampu mencakup semua nilai dari urutan *input*. (b) blok residu TCN. Konvolusi 1X1 ditambahkan ketika *input* dan *output* residual memiliki dimensi yang berbeda. (c) contoh koneksi residual di TCN. Garis biru adalah *filter* dalam fungsi residual, dan garis hijau merupakan pemetaan identitas.

3. Hasil dan Pembahasan

3.1 Dataset

Dataset yang digunakan pada penelitian ini sebanyak 23.183 ulasan. Dari hasil pelabelan diperoleh hasil distribusi dengan sebaran sentimen yang tidak seimbang dengan rentang sebanyak 9.217-13.966, jadi metode yang digunakan peneliti disini menggunakan *undersampling*. Dimana pengambilan sampel baris pada setiap aspek berdasarkan distribusi aspek yang paling kecil. Sehingga jumlah antar kelas data jadi lebih seimbang. Distribusi label aspek ditunjukkan pada Tabel 5.

Tabel 5. Distribusi Label Aspek

Aspek	Positive	Negative	Total
Aplikasi	2534	3180	5714
User Interface	1171	1901	3072
User Experience	1006	2235	3241
Kode OTP	1938	332	2270
Unduh Sertifikat Vaksin	2138	5348	7486
Bukti Akses Layanan Publik	430	970	1400
	9217	13966	23183

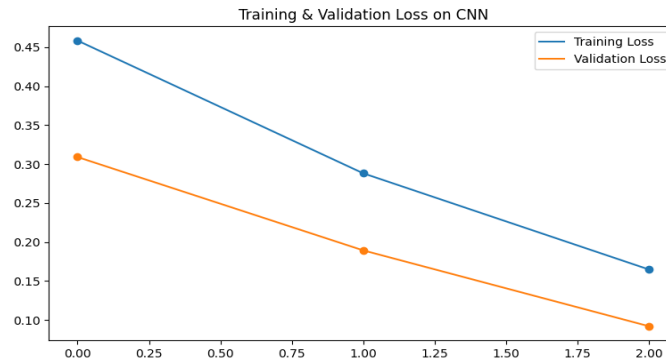
3.2 Klasifikasi Sentimen

Proses klasifikasi sentimen, penulis memanfaatkan beberapa *tuning hyperparameter* serta *optimizer*. Untuk *optimizer* penulis menggunakan ADAM karena dapat mencapai hasil yang baik dan telah banyak diterapkan oleh para peneliti terdahulu [13]. *Hyperparameter* yang digunakan dalam penelitian ini meliputi jumlah *filter*, *kernel*, *hidden dimension*, *input length*, *drop out*, *batch size*, dan *epoch*. Hasil evaluasi dari ketiga model yang dibangun dapat dilihat pada Tabel 6. Dari masing-masing model menunjukkan nilai terbaik tidak berbeda jauh. CNN memperoleh akurasi tertinggi sebesar 0.97% dari kedua model lainnya. Arsitektur CNN memperoleh akurasi tertinggi dalam penelitian ini terdiri dari 128 jumlah *filter*, 5 jumlah *kernel*, 100 jumlah *hidden dims*, 193 *input length*, 16 *batch size*. Model berlatih menggunakan pengakhiran awal *epoch* untuk menetapkan jumlah *train* saat model membaik.

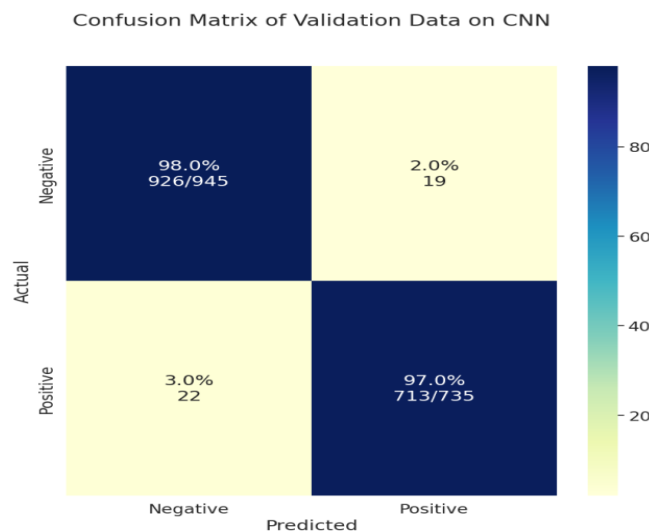
Tabel 6. Hasil Klasifikasi Sentimen

Model	Precision	Recall	F1 Score	Accuracy
CNN	0.98	0.97	0.98	0.98
GRU	0.96	0.96	0.96	0.96
TCN	0.93	0.93	0.93	0.93

Setelah proses pelatihan dilakukan pada data tes, kita dapat melihat validasi *loss* untuk memantau nilai *error* hasil prediksi. Model yang telah dilatih selanjutnya di evaluasi dengan *confusion matrix*. *Validation loss* dan *confusion matrix* pada klasifikasi sentimen dengan model terbaik. Ditunjukkan pada Gambar 5 dan 6 secara berurutan.



Gambar 5. Grafik *Loss* Klasifikasi Sentimen



Gambar 6. *Confusion Matrix* Klasifikasi Sentimen

3.3 Klasifikasi Aspek

Tugas klasifikasi aspek serupa dengan pengerjaan klasifikasi sentimen dengan memanfaatkan *tuning hyperparameter*. Untuk *optimizer* ADAM masih digunakan serta fungsi *loss categorical_crossentropy*. Hasil klasifikasi aspek ditunjukkan pada Tabel 7. Skor akurasi terbaik pada klasifikasi aspek masih diungguli pada model CNN dengan nilai skor akurasi sebesar 0.9702%.

Tabel 7. Hasil Klasifikasi Setiap Aspek

Aspek	Precision	Recall	F1 Score
Aplikasi	0.95	0.97	0.96
User Interface	1	1	1
User Experience	0.97	0.91	0.94
Kode OTP	0.99	1	1
Kartu vaksin	0.94	0.91	0.94
Akses Layanan	0.99	0.97	0.98
Accuracy	0.97		

Model CNN memperoleh akurasi terbaik pada klasifikasi aspek terdiri dari 128 jumlah *filter*, 5 jumlah *kernel*, 100 jumlah *hidden dims*, 193 *input length*, 64 *batch size* dan 3 *epoch*. Hasil prediksi label dengan data sebenarnya dapat dilihat pada tabel 8.

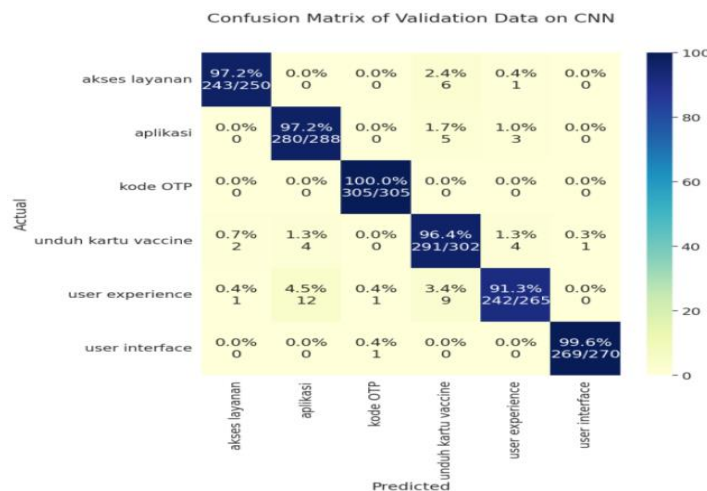
Tabel 8. Hasil Prediksi Label Dengan Data Sebenarnya

Reviews	Sebenarnya Aspek	Prediksi Aspek
Aplikasi sampah tidak deteksi lokasi kota hanya deteksi 1 wilayah payah bikin aplikasi asal banget developer sory uninstall aplikasi kasih bintang 1 aplikasi tidak bantu	Aplikasi	Aplikasi
Sebagian fitur tidak fungsi masa mau login isi tanggal lahir tidak bisa gimana ini	User Interface	User Interface
Vaksin pertama tapi aplikasi muncul sertifikat vaksin2 tidak muncul bingung	User Experience	User Experience
Server nya lambat respon kode otp lelet mohon dikirim cepat otpnya	Kode OTP (One time password)	Kode OTP (One time password)
Membantu proses unduh sertifikat vaksin 1 n 2 cepat	Kartu Vaksin	Kartu Vaksin
Call center sulit hubung sertifikat vaksin muncul lambat layanan	Akses layanan	User experience

Validasi *loss* dan *confusion matrix* pada klasifikasi aspek model terbaik ditunjukkan pada Gambar 7 dan 8 secara berurutan.



Gambar 7. Grafik Loss Klasifikasi Aspek



Gambar 8. Confusion Matrix Klasifikasi Aspek

Pada gambar 7 dan 8 di atas dapat menjelaskan proses pelatihan dilakukan pada data tes, kita dapat melihat validasi *loss* untuk memantau nilai *error* hasil prediksi. Model yang telah dilatih selanjutnya di evaluasi dengan *confusion matrix*. *Validation loss* dan *confusion matrix* pada klasifikasi sentimen dengan model terbaik dan selanjutnya dapat dijelaskan pada hasil pengujian pada pokok bahasan 3.4 pengujian hasil.

3.4 Pengujian Hasil

Berdasarkan pengujian yang dilakukan dalam menebak klasifikasi sentimen model sudah cukup baik saat memprediksi kelas sentimen. Namun, model belum cukup baik saat memprediksi kelas aspek dengan benar dapat dilihat pada evaluasi *confusion matrix* pada klasifikasi aspek hal ini dapat dilihat dari nilai *recall* yang rendah pada aspek *user experience* dimana model tidak dapat menebak kelas aspek yang bukan termasuk aspeknya dengan benar. Model melakukan kesalahan saat memprediksi kalimat “call center sulit hubung sertifikat vaksin muncul lambat layanan” sebagai aspek layanan namun mesin mengidentifikasi sebagai aspek *user experience*.

Kesalahan ini terjadi karena faktor ketidakseimbangan data sehingga model lebih cenderung memprediksi kelas mayoritas dari pada kelas minoritas seperti pada kalimat “vaksin pertama tapi aplikasi muncul sertifikat vaksin2 tidak muncul bingung” terdapat kata “muncul” sebanyak 156 kali pada aspek *user experience* dan 43 kali pada aspek layanan faktor lain adalah kesalahan tahap *preprocessing* yang dilakukan, serta pengambilan data dilakukan tiap dokumen. Sedangkan satu dokumen terdapat sejumlah kalimat dan pada satu kalimat bisa mewakili beberapa aspek.

4. Kesimpulan dan Saran

Pada penelitian ini, dilakukan analisis sentimen berbasis aspek dengan perbandingan tiga model *deep learning* CNN, GRU, dan TCN pada domain ulasan aplikasi pedulilindungi. Sebanyak 23.183 ulasan dengan dua polaritas sentimen “*positive* dan *negative*” dan aspek yang dipetakan “aplikasi *user interface*, *user experience*, kode OTP, kartu vaksin, dan akses layanan”. Hasil akurasi yang didapatkan oleh model CNN untuk klasifikasi sentimen sebesar 97.56 dan untuk klasifikasi aspek sebesar 97,02. model terbaik CNN terdiri dari 128 jumlah *filter*, 3 jumlah *epoch*, 5 jumlah *kernel*, 100 jumlah *hidden dims*, 193 *input length*, 16 *batch size*.

Daftar Pustaka

- [1] Laskari, N. K., & Sanampudi, S. K. (2016). Aspect Based Sentiment Analysis Survey Deep Learning for NLP and IR View project Extraction of Events, Times from natural language text and mapping of the relations between them View project Aspect Based Sentiment Analysis Survey. 18(2), 24–28. <https://doi.org/10.9790/0661-18212428>
- [2] Huang, J., Meng, Y., Guo, F., Ji, H., & Han, J. (2020). Weakly-Supervised Aspect-Based Sentiment Analysis via Joint Aspect-Sentiment Topic Embedding. <https://github.com>.
- [3] Arianto, D., & Budi, I. (2020). Aspect-based Sentiment Analysis on Indonesia’s Tourism Destinations Based on Google Maps User Code-Mixed Reviews (Study Case: Borobudur and Prambanan Temples). <https://maps.google.com/localguides>
- [4] Gu, S., Zhang, L., Hou, Y., & Song, Y. (2018). A Position-aware Bidirectional Attention Network for Aspect-level Sentiment Analysis.
- [5] D. Ekawati and M. L. Khodra, “aspect-based sentiment analysis for Indonesia restaurant reviews,” (2017). International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA), 2017, pp. 1-6, doi:10.1109/ICAICTA.2017.8090963
- [6] Bhargav, J. (2019). Product Design Improvement System using Aspect Based Sentiment Analysis of Consumer Reviews Learning Interactions and Dynamics of Swarms View project Aspect Based Sentiment Analysis of Consumer Reviews View project. <https://www.researchgate.net/publication/343726171>
- [7] Dhabekar, S., & Patil, M. D. (2021). Implementation of Deep Learning Based Sentiment Classification and Product Aspect Analysis. *ITM Web of Conferences*, 40, 03032. <https://doi.org/10.1051/itmconf/20214003032>
- [8] Rizki Amalia, P., & Winarko, E. (2021). Aspect-Based Sentiment Analysis on Indonesian Restaurant Review Using a Combination of Convolutional Neural Network and Contextualized Word Embedding.
- [9] Axhiu, M., & Aliu, A. (2021). Semi-Supervised Aspect-Based Sentiment Analysis On Albanian Restaurant Reviews. In *KNOWLEDGE-International Journal* (Vol. 45, Issue 3).

-
- [10] Utama, I. P. A. M., Prasetyowati, S. S., & Sibaroni, Y. (2021). Multi-Aspect Sentiment Analysis Hotel Review Using RF, SVM, and Naïve Bayes based Hybrid Classifier. *Jurnal Media Informatika Budidarma*, 5(2), 630. <https://doi.org/10.30865/mib.v5i2.2959>
- [11] Parameswari, P. L., & Prihandoko. (2022). Penggunaan Convolutional Neural Network Untuk Analisis Sentimen Opini Lingkungan Hidup Kota Depok Di Twitter. *Jurnal Ilmiah Teknologi Dan Rekayasa*, 27(1), 29–42. <https://doi.org/10.35760/tr.2022.v27i1.4671>
- [12] Bai, S., Kolter, J. Z., & Koltun, V. (2018). An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. <http://arxiv.org/abs/1803.01271>
- [13] Naqitasia, Risca & Fudholi, Dhomas & Iswari, Lizda. (2022). Analisis Sentimen Berbasis Aspek pada Wisata Halal dengan Metode Deep Learning. *Jurnal Teknoinfo*. 16. 156. [10.33365/jti.v16i2.1516](https://doi.org/10.33365/jti.v16i2.1516).
- [14] Zhao, N., Gao, H., Wen, X., & Li, H. (2021). Combination of convolutional neural network and gated recurrent unit for aspect-based sentiment analysis. *IEEE Access*, 9, 15561–15569. <https://doi.org/10.1109/ACCESS.2021.3052937>
- [15] Siyam, B. Y. (2018). Deep Learning for Aspect-Based Sentiment Analysis of Government Mobile Apps Reviews. *ل العميق التعلم استخدام المشاعر تحليل السمات على المستند للت على عبقات الذكية الحكومية التطبيقات*.
- [16] Alqaryouti, O., Siyam, N., Monem, A. A., & Shaalan, K. (2019). Aspect-based sentiment analysis using smart government review data. *Applied Computing and Informatics*. <https://doi.org/10.1016/j.aci.2019.11.003>
- [17] Ramdhani, M. A., Maylawati, D. S. adillah, & Mantoro, T. (2020). Indonesian news classification using convolutional neural network. *Indonesian Journal of Electrical Engineering and Computer Science*, 19(2), 1000–1009. <https://doi.org/10.11591/ijeecs.v19.i2.pp1000-1009>
- [18] Hidayatullah, A. F., & Ma'Arif, M. R. (2017). Pre-processing Tasks in Indonesian Twitter Messages. *Journal of Physics: Conference Series*, 801(1). <https://doi.org/10.1088/1742-6596/801/1/012072>
- [19] Antonio, V. D. (2021) Analisis Kinerja Ekstraksi Fitur Tf-Idf (Term Frequency Inverse Document Frequency) untuk Algoritma Klasifikasi Stochastic Gradient Descent Pada Analisis Sentimen Teks Indonesia. <https://repositori.usu.ac.id/handle/123456789/48179>
- [20] Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. <http://arxiv.org/abs/1408.5882>
- [21] Moch Ari Nasichuddin. (2017). Implementasi CNN untuk Klasifikasi Teks Menggunakan Tensorflow.
- [22] Biswas, S., Chadda, E., & Ahmad, F. (2015). Sentiment Analysis with Gated Recurrent Units. *Advances in*
- [23] *Computer Science and Information Technology (ACSIT)*, 2(11), 59–63. <http://www.krishisanskriti.org/acsit.html> Long, J., Shelhamer, E., & Darrell, T. (2016). Fully Convolutional Networks for Semantic Segmentation.