



ANALISIS SENTIMEN BERBASIS ASPEK PADA WISATA HALAL DENGAN METODE DEEP LEARNING

Risca Naquitasia¹⁾, Dthomas Hatta Fudholi²⁾, Lizda Iswari³⁾

^{1,2,3} Program Studi Informatika, Fakultas Teknologi Industri, Universitas Islam Indonesia

^{1,2,3} Jl. Kaliurang km. 14,5 Sleman, Yogyakarta

Email: ¹ risca.naquitasia@students.uii.ac.id, ² hatta.fudholi@uui.ac.id, ³ lizda.iswari@uui.ac.id

Abstract

Halal tourism is becoming a trend today, along with the improvement of the number of Muslim tourists. Halal tourism is a part of the tourism sector, which offers service by referring to Islam rules. By halal tourism development, start to occur tourism place-review related to the facility that can facilitate Muslim tourists. The facilities involve toilet cleanliness, worship place availability, and halal food availability. However, unfortunately, there are very few who analyze the review regarding this issue. Therefore, through this research, it was done a sentiment analysis based on various aspects of tourism places in Asian countries using the deep learning method. This method was used because it could generate a good performance accuracy. The data used were English reviews taken from the TripAdvisor website. The data then worked and processed so that it could recognize the sentiment and aspect from the review as well. There were three aspects used those were a mosque, halal food, and toilet. After a test was done, the CNN method obtained the highest accuracy result if it was compared with the other methods, both from aspect classification or sentiment classification. With the CNN method, the aspect classification produced an accuracy of 98.299%. Meanwhile, the sentiment classification gained an accuracy of 93.96%. The result of this research is expected to help develop a strategy to promote halal food more.

Keyword: *Aspect based sentiment analysis, CNN, BiLSTM, Deep learning, Halal tourism, Trip advisor.*

Abstrak

Wisata halal kini tengah menjadi tren seiring dengan meningkatnya jumlah populasi wisatawan muslim. Wisata halal merupakan bagian dari sektor pariwisata yang menawarkan pelayanan dengan merujuk kepada aturan-aturan Islam. Dengan berkembangnya wisata halal, mulai muncul ulasan tempat wisata yang berkaitan dengan fasilitas yang dapat memudahkan wisatawan muslim. Fasilitas tersebut meliputi kebersihan toilet, ketersediaan tempat ibadah, serta ketersediaan makanan halal. Namun sayangnya, sangat sedikit yang menganalisis ulasan mengenai persoalan ini. Oleh karena itu, melalui penelitian ini dilakukan analisis sentimen berbasis aspek terhadap tempat wisata di negara Asia menggunakan metode deep learning. Metode ini digunakan karena menghasilkan akurasi performa yang baik. Data yang digunakan adalah ulasan berbahasa inggris yang diambil dari website TripAdvisor. Data tersebut kemudian diolah dan diproses sehingga dapat mengenali sentimen dan juga aspek dari ulasan tersebut. Terdapat tiga aspek yang digunakan yaitu mosque, halal food, dan toilet. Setelah dilakukan pengujian, metode CNN mendapatkan hasil akurasi tertinggi jika dibandingkan dengan metode lainnya baik pada klasifikasi aspek maupun klasifikasi sentimen. Dengan metode CNN, klasifikasi aspek menghasilkan akurasi sebesar 98.299%. Sedangkan klasifikasi sentimen mendapatkan akurasi sebesar 93.96%. Hasil penelitian ini diharapkan dapat membantu mengembangkan strategi untuk lebih memajukan wisata halal.

Kata Kunci: *Analisis sentimen berdasarkan aspek, CNN, BiLSTM, Deep learning, Wisata halal, Trip advisor.*

1. PENDAHULUAN

Sektor pariwisata saat ini tengah hangat diperbincangkan terutama dengan adanya industri halal. Industri halal itu sendiri terdiri dari banyak macam seperti makanan halal, jasa keuangan islami, kosmetik halal, busana muslim, dan juga wisata halal. Setiap sektor tersebut mengadopsi nilai-nilai keislaman dalam setiap produk atau jasa yang ditawarkan. Istilah industri halal ini bahkan tidak tergolong istilah baru karena banyak negara yang mayoritas penduduknya bukan muslim telah menerapkan industri halal seperti Jepang, Australia, USA, Kanada, New Zealand and Afrika Selatan [1]

Industri halal sangat berpengaruh terhadap perkembangan ekonomi Islam. Hal ini terlihat pada pertumbuhan pasar muslim global yang memegang peranan penting terhadap ekonomi global. Pertumbuhan ekonomi Islam sangat pesat terlihat pada wisata halal. Pada tahun 2018, wisatawan global mencapai 140 juta wisatawan, angka ini bertambah 10% apabila dibandingkan pada tahun 2017 yang mencapai 131 juta wisatawan dengan Indonesia mencapai posisi pertama



sebagai tujuan wisatawan muslim [2]. Wisata halal menjadi sektor yang paling cepat berkembang pada pasar pariwisata [3].

Wisata halal adalah layanan pariwisata yang berfokus dalam pemenuhan kebutuhan muslim sesuai dengan ajaran agama islam. Dalam praktiknya mengadaptasi dari nilai-nilai islam. Terdapat beberapa aspek pada wisata halal yaitu ketersediaan makanan halal, fasilitas tempat ibadah yang layak dan tidak jauh dari tempat wisata, toilet dengan air bersih, layanan berbuka puasa saat ramadan, serta larangan pada aktivitas yang tidak sesuai syariah. [4] Ada banyak sekali yang bisa digali dari wisata halal. Namun sayangnya, tidak banyak penelitian yang menganalisis wisata halal berdasarkan aspek.

Saat ini orang menggunakan internet sebagai wadah mereka mengeluarkan pendapat dan isi pikiran mereka. Ada banyak sekali pendapat dari sudut pandang yang berbeda-beda di dalam internet. Contohnya yaitu pada laman TripAdvisor yang menyediakan kolom ulasan pada setiap tempat wisata yang terdaftar pada web tersebut. Ulasan ini dapat dikumpulkan dan diproses menjadi informasi yang bermanfaat. Setelah dataset terkumpul, maka akan dilakukan analisis sentimen. Menurut [5], analisis sentimen atau yang disebut juga dengan *opinion mining* adalah bidang pembelajaran yang melakukan analisis sentimen, opini, emosi, dan evaluasi dari data yang tertulis. Analisis sentimen dilakukan untuk menilai suatu objek apakah cenderung positif atau negatif. Sementara *aspect based sentiment analysis* (ABSA) adalah salah satu bentuk analisis sentimen yang mempertimbangkan sentimen berdasarkan setiap aspek.. Metode untuk menganalisis sentimen sebuah topik bisa dengan berbagai macam. *Deep learning* dipilih untuk menjadi metode penelitian ini.

Pada beberapa tahun ini, *deep learning* menunjukkan pencapaian di berbagai bidang. Jika dibandingkan dengan metode *machine learning* yang masih tradisional, *deep learning* tidak membutuhkan campur tangan manusia. Karena dibutuhkan data yang cukup besar, pendekatan yang dilakukan yaitu mengubah masukan menjadi matriks yang berisikan angka. Matriks ini yang kemudian menjadi bahan untuk melatih algoritma. Metode *deep learning* bekerja dengan mengekstraksi data menggunakan *neural network*. Dari sini model kemudian akan belajar dari *error*. *Neural network* itu sendiri terdiri dari sekumpulan lapisan yang dipetakan oleh fungsi aktivasi. Model *deep learning* yang biasa digunakan untuk menganalisis data berupa teks adalah CNN, RNN, LSTM, dan Gated Recurrent Unit (GRU) [6]. Sebagian besar penelitian saat ini telah menggunakan *deep learning*. Hal ini membuat perkembangan *deep learning* berkembang dengan sangat cepat.

Analisis sentimen berbasis aspek dengan model *deep learning* tentu telah banyak dilakukan oleh peneliti terdahulu, sayangnya hingga saat ini belum ada yang berfokus pada wisata halal khususnya di wilayah negara negara di Asia. Oleh karena itu melalui makalah ini, kontribusi yang akan dilakukan adalah menganalisis sentimen ulasan berbasis aspek mengenai wisata halal dengan mengaplikasikan model *deep learning*.

2. METODE PENELITIAN

2.1 Penelitian Terdahulu

Penelitian mengenai wisata halal telah dilakukan sebelumnya oleh [7] yang menerapkan CNN, RNN, dan LSTM dengan memanfaatkan Word2Vec dan Word2Sec. Setelah melakukan percobaan, didapat hasil tertinggi metode LSTM dan CNN dengan akurasi tertinggi yaitu 93.78%. Selain itu [8] juga mengangkat penelitian yang mengenai wisata halal. Data yang digunakan adalah data dari Twitter. Penelitian ini menggunakan *text mining* dan memanfaatkan *wordcloud* untuk menganalisis data. Jepang menjadi destinasi yang paling sering muncul pada *tweets* mengenai destinasi pariwisata halal yang disusul dengan Malaysia dan Indonesia.

Analisis sentimen berbasis aspek telah menunjukkan kiprahnya pada bidang pembelajaran mesin. Beberapa peneliti terdahulu telah bermunculan mengembangkannya. Salah satunya yaitu [9]. Pada penelitian tersebut dengan menggunakan ulasan wisata berbahasa Indonesia, CNN mendapatkan akurasi tertinggi saat dibandingkan dengan CNN-LSTM dan CNN-GRU. Penelitian lain yang mengangkat mengenai analisis sentimen berbasis aspek adalah [10]. Penelitian ini mengaplikasikan metode CNN-BiLSTM pada ulasan teks bahasa mandarin. Penelitian oleh [11] juga mengangkat tentang hal ini dengan memanfaatkan ulasan berbahasa Arab. Pengaplikasiannya menggunakan metode LSTM.

CNN menjadi metode yang banyak dipilih karena performanya yang terbukti sangat baik. Pada penelitian [12] yang membandingkan DNN, CNN, dan RNN menunjukkan bahwa performa CNN memang sangat baik. Di antara ketiga model yang digunakan, CNN menunjukkan hasil terbaik waktu *processing* dan nilai akurasi dibandingkan dengan dua model lainnya. Bahkan hal ini juga ditunjukkan pada penelitian yang dilakukan oleh [13]. Penelitian yang mengangkat tentang *natural language processing* dengan melakukan perbandingan metode antara *naive bayes*, SVM, KNN, RNN, dan CNN membuktikan bahwa metode *multi-CNN* menghasilkan presisi, *recall*, dan skor F1 terbaik

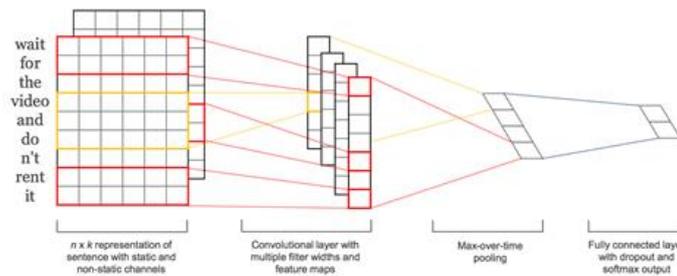
2.2 CNN

Convolutional neural network (CNN) merupakan salah satu bentuk jaringan saraf tiruan yang dapat mendeteksi informasi dengan akurasi yang bagus. CNN sudah banyak diaplikasikan pada klasifikasi gambar dan *natural language processing* lainnya sehingga menjadikan model ini dikenal dengan karakteristiknya sebagai arsitektur untuk pembelajaran [14]. CNN bekerja dengan memasukkan data multi dimensi seperti gambar atau *word embedding* yang kemudian dikirim



ke lapisan konvolusi yang akan dikomposisikan ke dalam beberapa filter untuk dipelajari dengan berbagai fitur. Luarnya biasanya merupakan dimensi yang lebih kecil dan akan dimasukkan ke dalam *full connected layer* [15].

Teks merupakan data terstruktur yang sangat terorganisasi menjadikan dasar pemikiran awal mula analisis teks menggunakan CNN dimulai. Untuk setiap kata diberikan bobot pada *hidden layer* dan kemudian diperiksa untuk kesamaannya dengan proses yang berulang. CNN bekerja dengan logika *sliding window* [16]. Untuk klasifikasi teks, masukan merupakan kumpulan kata yang kemudian diubah menjadi matriks seperti pada gambar 1.



Gambar 1. CNN pada klasifikasi teks

2.3 RNN-BiLSTM

Recurrent neural network (RNN) merupakan jaringan yang memiliki tugas memproses urutan tugas dalam melakukan klasifikasi. Tugas ini dijalankan dengan menyimpan informasi sejarah teks. Tetapi dengan keterbatasan memori dan kapasitas penyimpanan dapat membuat masalah bagi RNN [10]. Untuk mengatasi kekurangan ini, dibentuklah *long short term memory* (LSTM) yang dapat memiliki memori dan penyimpanan yang lebih baik dari RNN. Jaringan LSTM mendapatkan masukan langkah waktu saat ini dan keluaran dari langkah waktu sebelumnya. Kedua masukan itu kemudian menghasilkan keluaran yang akan diberikan kepada langkah waktu berikutnya. *Hidden layer* akan digunakan dari waktu terakhir untuk melakukan klasifikasi. [17]

Pada arsitektur LSTM terdapat sel memori dan tiga gerbang, yaitu *input gate*, *output gate*, *forget gate*. Masukan, *hidden state*, dan tiga gerbang tersebut terhubung dalam perhitungan yang dapat kita lihat seperti dibawah ini:

$$f_t = \sigma (W_f \cdot x_t + U_f \cdot h_{t-1} + b_f) \tag{1}$$

$$i_t = \sigma (W_i \cdot x_t + U_i \cdot h_{t-1} + b_i) \tag{2}$$

$$\bar{c}_t = \tanh(W_c \cdot x_t + U_c \cdot h_{t-1} + b_c) \tag{3}$$

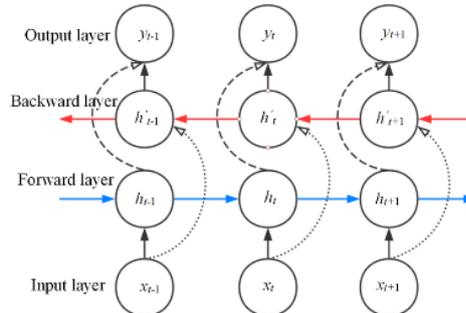
$$c_t = i_t \odot \bar{c}_t + f_t \odot c_{t-1} \tag{4}$$

$$o_t = \sigma (W_o \cdot x_o + U_o \cdot h_{t-1} + b_o) \tag{5}$$

$$h_t = o_t \odot \tanh(c_t) \tag{6}$$

W_f, W_i, W_o, W_c adalah bobot dari LSTM sedangkan b_f, b_i, b_o, b_c melambangkan bias. Fungsi aktivasi *sigmoid* dilambangkan dengan σ dan *tanh* adalah fungsi tangen hiperbola. h_t adalah *hidden state* dari waktu t. [10]

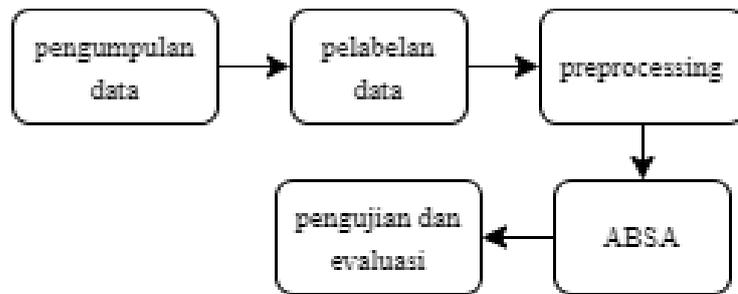
Bidirectional LSTM (BiLSTM) adalah salah satu bentuk terbaru dari LSTM. BiLSTM terjadi secara dua arah dengan melatih dua *hidden layer* pada urutan masukan. Arus yang pertama adalah urutan masukan seperti biasanya, lalu arus kedua adalah urutan masukan secara terbalik. Hal ini akan menambah konteks pada jaringan sehingga pembelajaran akan menjadi semakin cepat dan baik [17]. Berikut ini adalah gambaran dari arsitektur BiLSTM yang tergambar pada gambar 2.



Gambar 2. CNN pada klasifikasi teks

2.4 Tahapan Penelitian

Dalam melakukan penelitian ini, terdapat beberapa tahapan yang perlu dilakukan agar memperoleh hasil yang baik. Tahapan tersebut terdiri dari lima tahap, yaitu tahap pengumpulan data, tahap pelabelan data, tahap *preprocessing*, tahap ABSA, serta tahap pengujian dan evaluasi. Proses penelitian ini dapat diilustrasikan seperti pada gambar 3.



Gambar 3. Tahapan Proses Penelitian

a. Pengumpulan Data

Dataset yang berupa ulasan-ulasan berbahasa inggris pada *website* tripadvisor.com dikumpulkan dan disimpan ke dalam bentuk *file* csv. Pengumpulan data memanfaatkan fitur pencarian pada kolom ulasan dengan menggunakan beberapa kata kunci yang berkaitan, seperti *dirty toilet*, *clean toilet muslims*, *mosque*, *worship*, *halal food*, dan masih banyak lagi. Hal ini dikarenakan pada penelitian ini hanya menggunakan ulasan yang menyinggung tentang toilet, makanan halal, dan tempat beribadah. Pengambilan data dilakukan dengan teknik *scraping*. Beberapa data diambil secara manual dengan tujuan mengefisienkan waktu

b. Pelabelan Data

Setelah dataset telah terkumpul, tahapan selanjutnya adalah melakukan pelabelan data. Pada setiap ulasan akan diberikan dua label yaitu label sentimen dan label aspek. Sentimen ini terbagi menjadi dua, positif dan negatif. Sementara untuk aspek terbagi menjadi tiga yaitu, *mosque*, *halal food*, dan toilet. Untuk setiap data yang memiliki lebih dari satu aspek akan digunakan teknik *multi-labeling*.

c. Preprocessing

Preprocessing adalah tahapan pembersihan data sebelum data akan mulai dilatih. Tujuannya agar komputer dapat lebih mudah mengenal data sehingga hasil yang didapatkan semakin baik. *Preprocessing* memiliki banyak bentuk yang disesuaikan dengan kebutuhan dari dataset yang digunakan. Berikut ini merupakan bentuk *preprocessing* yang dilakukan yaitu

1. Menghapus Tag HTML

Pada saat dataset dikumpulkan melalui teknik *scraping*, terdapat tag HTML pada setiap datanya. Contoh tag HTML yang ada pada data yaitu `
` yang melambangkan terdapat jeda baris antar kalimat.

2. *Casefolding*

Casefolding merupakan tahap penyamakan seluruh huruf pada data. Seluruh huruf akan diubah ke dalam huruf kecil (*lowercase*). Pada proses ini, yang diperbolehkan hanyalah ‘a’ hingga ‘z’.

3. Menghapus Angka

Pada tahap ini, seluruh digit yang ada akan dihapus. Hal ini dikarenakan angka pada penelitian ini tidak diperlukan dan tidak memberikan pengaruh pada hasil.



4. Menghapus Tanda Baca
Di tahap ini segala bentuk tanda baca seperti *#\$/&() akan dihilangkan.
 5. Menghapus Karakter Tidak Terpakai
Karakter yang tidak terpakai pada penelitian ini merupakan karakter yang memiliki huruf kurang dari tiga. Pada tahap ini, karakter tersebut akan dihapus.
 6. Menghapus Non-ASCII
Tahap ini akan menghapus seluruh karakter yang bukan merupakan karakter ascii. Karakter seperti emoji atau karakter dalam bahasa lain yang ikut terambil pada saat pengumpulan data akan terhapus.
 7. *Stopword Removal*
Stopwords adalah sekumpulan kata yang sering sekali muncul dalam sebuah kalimat. Pengolahan *stopwords* tidak diperlukan karena tidak banyak informasi yang dapat diambil. Kata-kata pada daftar *stopwords* biasanya tidak banyak memiliki makna yang berarti. Daftar ini dapat diubah sesuai dengan kebutuhan dari sebuah penelitian. Penelitian ini menggunakan *library* NLTK.
- d. ASBA
Data yang sudah dibersihkan akan dipisah menjadi data *training* dan data *testing*. Setelah itu data *training* akan diproses melalui tahapan ABSA. Model yang sudah dilatih akan diuji kepada data *testing* sehingga nantinya model tersebut dapat memprediksi sentimen dan juga aspek dari ulasan wisata. Model akan membaca data kemudian melakukan ekstraksi dan juga menentukan sentimen melalui polaritasnya. Penelitian ini menggunakan metode *deep learning* dengan memilih model arsitektur CNN dan CNN-BiLSTM. Kedua model ini kemudian akan dibandingkan untuk mendapatkan hasil terbaik.
- Di dalam model CNN terdapat banyak lapisan. Data akan masuk ke *embedding layer* yang akan menginisialisasi kata yang ditetapkan dengan bobot acak dan mempelajarinya untuk ditambahkan ke dalam dataset latih. Kemudian akan menuju ke *convolutional layer*. Pada lapisan ini akan dilakukan ekstraksi fitur. Terjadi perhitungan kepada masukan menggunakan bantuan *pooling layer*. Lapisan *pooling* ini akan membantu mengurangi representasi dari kalimat masukan, parameter, dan komputasi yang ada pada jaringan. Lapisan yang terakhir merupakan *dense layer* atau yang disebut juga sebagai *fully connected layer*. Di lapisan ini akan dilakukan klasifikasi kepada fitur yang telah terekstraksi pada lapisan sebelumnya. Untuk arsitektur model CNN-BiLSTM ditambahkan lapisan *Bidirectional LSTM* sebelum *fully connected layer*.
- Pemilihan model merupakan hal penting karena dapat mempengaruhi hasil luaran. Penelitian ini menggunakan konvolusi 1D karena jenis konvolusi ini cocok untuk pengolahan data berupa teks. Selain itu juga digunakan *global max pooling* 1D untuk memberikan hasil yang baik. Aktivasi yang dipilih adalah RELU. Aktivasi ini akan memberikan nol pada nilai negatif dan akan bertambah untuk nilai positif. Di akhir jaringan akan ditambahkan *softmax* untuk klasifikasi aspek. Sedangkan untuk klasifikasi sentimen akan diberi RELU. Model ini memanfaatkan 64 *batch size* dan 64 *embedding dimentions*. Terakhir model menggunakan *early stopping* yang memantau *val loss*.
- Dalam menemukan hasil terbaik, dibutuhkan pemilihan parameter yang tepat. Parameter tersebut meliputi nilai filter, jumlah filter, ukuran filter, *hidden dimentions*, dropout, dan nilai unit lstm. Berikut ini merupakan tabel parameter yang digunakan

Tabel 1. Tabel parameter

	Klasifikasi aspek		Klasifikasi sentimen	
	Model 1	Model 2	Model 3	Model 4
Model	Model 1	Model 2	Model 3	Model 4
Metode	CNN	CNN-BiLSTM	CNN	CNN-BiLSTM
Kernel	10	10	9	9
Filter	200	200	200	200
Hidden dims	128	128	128	128
Dropout	0.5	0.5	0.25	0.25
Unit lstm	64	64	64	64

- e. Pengujian dan Evaluasi
Setelah model dilatih, tahap selanjutnya adalah pengujian dan evaluasi. Pengujian ini akan mencoba model menggunakan sebuah kalimat yang nantinya akan menghasilkan keluaran berupa aspek dan sentimen dari kalimat tersebut. Sementara untuk evaluasi menggunakan *confusion matrix*. Data testing akan diujikan kepada model kemudian akan menampilkan tabel *confusion matrix* untuk melihat seberapa tepat model yang telah dibangun. *Confusion matrix* terdiri dari *true positive*, *true negative*, *false positive*, dan *false negative*.



3. HASIL DAN PEMBAHASAN

3.1 Dataset

Setelah dilakukan pengumpulan data, didapatkan 1947 ulasan. Ulasan ini diambil dari sejumlah tempat wisata yang berada di wilayah Asia secara acak. Dataset tersebut kemudian diberikan label sesuai dengan aspek (*halal food, mosque, toilet*) dan sentimennya (*positive, negative*). Distribusi dataset ini ditunjukkan pada tabel 2.

Tabel 2. Distribusi data berdasarkan label

Aspek	Positive	Negative	Total
<i>Halal food</i>	691	109	800
<i>Mosque</i>	255	2	257
<i>Toilet</i>	583	307	890
	1529	418	1947

Dataset melalui proses pembersihan sebelum mulai dilakukan pelatihan kepada model. Dengan terjadinya ketidakseimbangan data, dilakukan *random sampling*. Nantinya data yang memiliki jumlah lebih sedikit akan ditumpuk dengan data lainnya sehingga antar kelas menjadi lebih seimbang jumlahnya. Setelah dilakukan *random sampling*, kini setiap kelas aspek terdiri dari 890 data dan setiap kelas sentimen terdiri dari 1529 data. Kedua dataset tersebut baik kelas aspek maupun sentimen akan dibagi menjadi data latih dan data uji secara acak. Pembagian ini menggunakan data uji sebesar 33% dan sisanya digunakan sebagai data latih

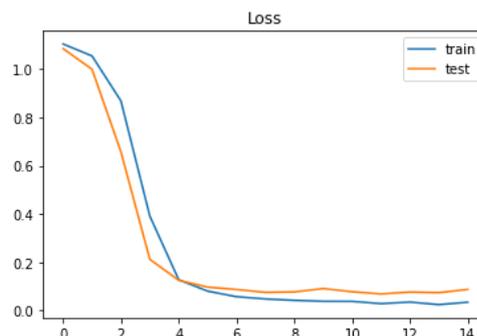
3.2 Klasifikasi Aspek

Untuk melakukan klasifikasi aspek, digunakan beberapa arsitektur dan *hyperparameter*. Selain itu digunakan juga *optimizer* dan fungsi *loss*. Pada klasifikasi aspek, *optimizer* yang digunakan adalah adam. *Optimizer* ini dipilih karena sudah sering digunakan pada penelitian sebelumnya dan menunjukkan hasil yang cukup memuaskan. Fungsi *loss* yang digunakan adalah *categorical crossentropy*. Setelah dilakukan beberapa eksperimen dan dicoba dengan data tes, didapatkan hasil seperti yang tertera pada tabel 3. Tabel tersebut menunjukkan model 1 (CNN) memiliki akurasi paling tinggi saat dijalankan dengan 15 *epoch* yaitu 98.299% sementara model 2 (CNN-BiLSTM) memberikan hasil terbaiknya sebesar 95.918% saat dijalankan dengan 12 *epoch*. Hasil dari model CNN memberikan tingkat akurasi yang lebih baik jika dibandingkan dengan model CNN-BiLSTM.

Tabel 3. Hasil klasifikasi aspek

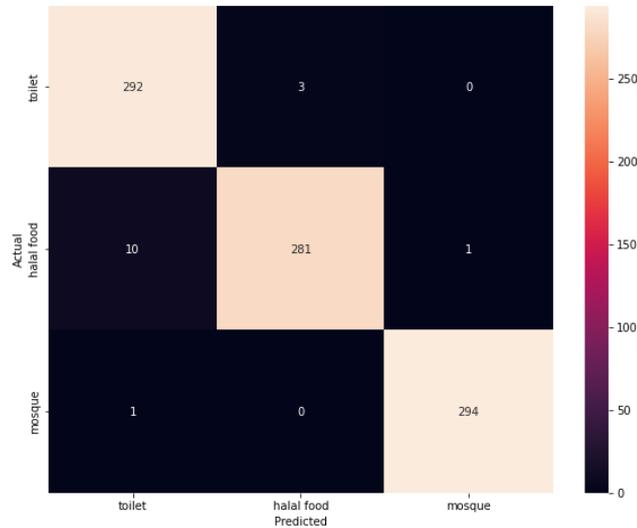
Model	Epoch	Akurasi	F1	Recall	Precision
Model 1	10	0.97732	0.97725	0.97720	0.97732
	13	0.98073	0.98071	0.98071	0.98072
	15	0.98299	0.98296	0.98292	0.98325
Model 2	12	0.96825	0.96832	0.96826	0.96871
	15	0.95692	0.95729	0.95772	0.95716
	17	0.95918	0.95921	0.95949	0.95996

Setelah dilakukan pengujian dengan data tes, kita dapat memantau *loss* yang divisualisasikan pada gambar 4. Model yang sudah dilatih kemudian dievaluasi melalui *confusion matrix* yang dapat terlihat seperti pada gambar 5. Kedua penggambaran ini mengambil dari hasil model terbaik yaitu model 1 dengan *epoch* 15.





Gambar 4. Grafik *loss* model 1 dengan *epoch* 15



Gambar 5. *Confusion matrix* model 1 dengan *epoch* 15

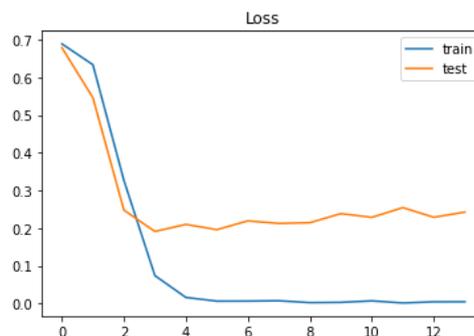
3.3 Klasifikasi Sentimen

Pelaksanaan klasifikasi sentimen memiliki proses yang sama seperti klasifikasi aspek. Sama halnya dengan klasifikasi aspek, model ini juga membutuhkan *optimizer* dan fungsi *loss*. Adam masih digunakan sebagai *optimizer* pada klasifikasi ini, sementara fungsi *loss* menggunakan *binary crossentropy*. Model dilatih dengan mencoba beberapa parameter sehingga pada saat dicoba dengan hasil tes menghasilkan akurasi, *F1 score*, *recall*, *precision* (tabel 4). Pada hasil kita dapat melihat model 3 (CNN) memiliki hasil tertinggi pada saat *epoch* ditetapkan menjadi 14 dengan akurasi sebesar 93.69%. Model 4 (CNN-BiLSTM) mendapatkan hasil terbaik pada saat *epoch* 11 dengan akurasi sebesar 93,465%. Apabila kita membandingkan kedua model ini, kita dapat mengetahui bahwa model CNN memiliki hasil akurasi yang lebih baik daripada model CNN-BiLSTM.

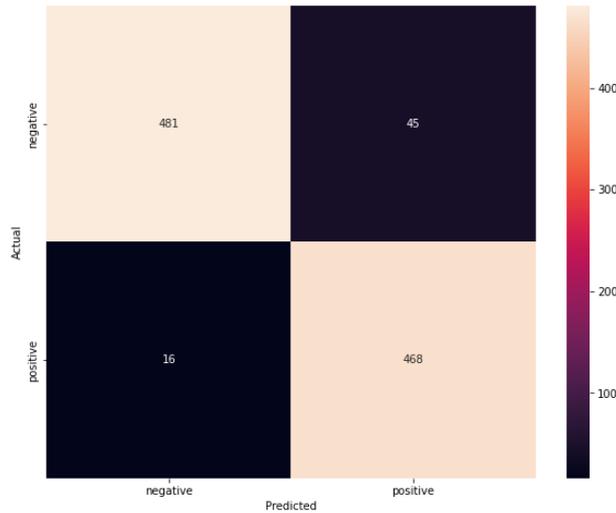
Tabel 4. Hasil klasifikasi sentimen

Model	Epoch	Akurasi	F1	Recall	Precision
Model 3	9	0.92871	0.92869	0.93006	0.92919
	11	0.93069	0.93069	0.93112	0.93099
	14	0.93960	0.93959	0.94069	0.94004
Model 4	7	0.92970	0.92969	0.92972	0.92982
	10	0.92871	0.92870	0.92868	0.92876
	11	0.93465	0.93464	0.93462	0.93467

Loss dari pelatihan model ini kemudian digambarkan seperti pada gambar 6. Selanjutnya model memasuki proses evaluasi yang bisa terlihat di gambar 7.



Gambar 6. Grafik *loss* model 3 dengan *epoch* 14



Gambar 7. *Confusion matrix* model 3 dengan *epoch* 14

3.4 Pengujian

Pengujian dilakukan dengan mencoba menebak aspek dan sentimen dari kalimat yang diberikan. Pada hal ini, kami mengambil model terbaik yang sudah diuji coba sebelumnya. Berikut ini merupakan hasil dari pengujian.

```

only sell non halal food
Sentiment : positive
Aspect: mosque

nice toilet
Sentiment : positive
Aspect: toilet
    
```

Gambar 8. Pengujian model terhadap kalimat

Pada gambar 8 terlihat model melakukan kesalahan saat memprediksi sentimen dan aspek dari kalimat “only sell non halal food”. Aspek dari kalimat tersebut terdeteksi sebagai *mosque* dengan sentimen positif.

Kesalahan ini dapat terjadi karena banyak hal, salah satunya yaitu data yang sangat sedikit seperti yang kita lihat pada tabel 6. Model CNN merupakan model yang dapat menghasilkan performa yang bagus apabila data yang diujikan tergolong banyak. Semakin banyak data, maka akan semakin bagus pula performanya. Tetapi sayangnya dengan waktu yang singkat, peneliti hanya dapat mengumpulkan 1947 data. Faktor lainnya adalah adanya kesalahan pada tahap *preprocessing* yang telah dilakukan. Selain itu penelitian ini tidak menggunakan *POS Tagging* dan *negation handling* sehingga model belum dapat mengekstrak sentimen dan aspek secara sempurna. Masih banyak lagi faktor-faktor yang memungkinkan lainnya.

4. KESIMPULAN

Pada penelitian ini dilakukan analisis sentimen berbasis aspek terhadap wisata halal. Sebanyak 1947 ulasan dipetakan ke dalam tiga aspek (*mosque, halal food, toilet*) dan dua sentimen (positif dan negatif). Model CNN dan CNN-BiLSTM dibandingkan dengan mencoba berbagai parameter. Setelah diujikan, model CNN menunjukkan hasil yang terbaik apabila melihat dari kedua hasil klasifikasi. Untuk klasifikasi aspek didapat akurasi 98.299% sedangkan untuk klasifikasi sentimen didapat akurasi sebesar 93.96%. Perbedaan hasil dari berbagai skenario yang diujikan pun tidak jauh berbeda. Hanya saja pada saat diujikan dengan sebuah kalimat, model terlihat melakukan kesalahan saat mencoba memprediksi kalimat tersebut. Hal ini diasumsikan terjadi karena data yang diujikan sangat sedikit dan adanya kesalahan pada tahap *preprocessing*.

**DAFTAR PUSTAKA**

- [1] R. S. Asa, "an Overview of the Developments of Halal Certification Laws in Malaysia, Singapore, Brunei and Indonesia," *J. Syariah*, vol. 27, no. 1, pp. 173–200, 2019.
- [2] M. CrescentRating, "Global Muslim Travel Index 2019." April, 2019.
- [3] H. El-Gohary, "Halal tourism, is it really Halal?," *Tour. Manag. Perspect.*, vol. 19, pp. 124–130, Jul. 2016, doi: 10.1016/j.tmp.2015.12.013.
- [4] V. Paramarta, R. R. V. K. Dewi, F. Rahmanita, S. Hidayati, and D. Sunarsi, "Halal Tourism in Indonesia: Regional Regulation and Indonesian Ulama Council Perspective," *Int. J. Criminol. Sociol.*, vol. 10, pp. 497–505, 2021.
- [5] B. Liu, "Sentiment Analysis and Opinion Mining," *Synth. Lect. Hum. Lang. Technol.*, vol. 5, no. 1, pp. 1–167, May 2012, doi: 10.2200/S00416ED1V01Y201204HLT016.
- [6] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *WIREs Data Min. Knowl. Discov.*, vol. 8, no. 4, p. e1253, Jul. 2018, doi: 10.1002/widm.1253.
- [7] A. Feizollah, S. Ainin, N. B. Anuar, N. A. B. Abdullah, and M. Hazim, "Halal Products on Twitter: Data Extraction and Sentiment Analysis Using Stack of Deep Learning Algorithms," *IEEE Access*, vol. 7, pp. 83354–83362, 2019, doi: 10.1109/ACCESS.2019.2923275.
- [8] S. Ainin, A. Feizollah, N. B. Anuar, and N. A. Abdullah, "Sentiment analyses of multilingual tweets on halal tourism," *Tour. Manag. Perspect.*, vol. 34, p. 100658, Apr. 2020, doi: 10.1016/j.tmp.2020.100658.
- [9] R. A. N. Nayoan, A. F. Hidayatullah, and D. H. Fudholi, "Convolutional Neural Networks for Indonesian Aspect-Based Sentiment Analysis Tourism Review," in *2021 9th International Conference on Information and Communication Technology (ICoICT)*, 2021, pp. 60–65.
- [10] K. Zhou and F. Long, "Sentiment Analysis of Text Based on CNN and Bi-directional LSTM Model," in *2018 24th International Conference on Automation and Computing (ICAC)*, 2018, pp. 1–5, doi: 10.23919/ICoAC.2018.8749069.
- [11] M. Al-Smadi, B. Talafha, M. Al-Ayyoub, and Y. Jararweh, "Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 8, pp. 2163–2175, 2019.
- [12] N. C. Dang, M. N. Moreno-García, and F. De la Prieta, "Sentiment Analysis Based on Deep Learning: A Comparative Study," *Electronics*, vol. 9, no. 3, 2020, doi: 10.3390/electronics9030483.
- [13] Y.-C. Chang, C.-H. Ku, and C.-H. Chen, "Using deep learning and visual analytics to explore hotel reviews and responses," *Tour. Manag.*, vol. 80, p. 104129, 2020, doi: <https://doi.org/10.1016/j.tourman.2020.104129>.
- [14] M. Rhanoui, M. Mikram, S. Yousfi, and S. Barzali, "A CNN-BiLSTM model for document-level sentiment analysis," *Mach. Learn. Knowl. Extr.*, vol. 1, no. 3, pp. 832–847, 2019.
- [15] P. M. Sosa, "Twitter sentiment analysis using combined LSTM-CNN models," *Eprint Arxiv*, pp. 1–9, 2017.
- [16] B. S. Lakshmi, P. S. Raj, and R. R. Vikram, "Sentiment analysis using deep learning technique CNN with KMeans," *Int. J. pure Appl. Math.*, vol. 114, no. 11, pp. 47–57, 2017.
- [17] S. Minaee, E. Azimi, and A. Abdolrashidi, "Deep-sentiment: Sentiment analysis using ensemble of cnn and bi-lstm models," *arXiv Prepr. arXiv1904.04206*, 2019.