

# Analisis Sentimen Ulasan *E-Commerce* Pakaian Berdasarkan Kategori dengan Algoritma *Convolutional Neural Network*

I Made Adi Susilayasa<sup>1</sup>, AAIN Eka Karyawati<sup>2</sup>, Luh Gede Astuti<sup>3</sup>, Luh Arida Ayu Rahning Putri<sup>4</sup>, I Gede Arta Wibawa<sup>5</sup>, I Komang Ari Mogi<sup>6</sup>

<sup>a</sup>Informatics departement, Faculty of Math and Science, University of Udayana  
South Kuta, Badung, Bali, Indonesia

<sup>1</sup>3523adibusilayasa@gmail.com

<sup>2</sup>eka.karyawati@unud.ac.id

<sup>3</sup>lg.astuti@cs.unud.ac.id

<sup>4</sup>rahningputri@unud.ac.id

<sup>5</sup>gede.arta@cs.unud.ac.id

<sup>6</sup>arimogi@cs.unud.ac.id

## Abstract

*Almost everyone looks at reviews before deciding to buy an item in e-commerce. Consumers say that online reviews influence their purchasing decisions. Based on these data, consumers need sentiment reviews to make a decision to choose a product/service. However, the results of the sentiment analysis are still less specific, so the review classification process is carried out based on the review category. Sentiment classification process based on clothing category is carried out using the Convolutional neural network method. The amount of data used is 3384 data with 3 categories. The category classification model shows good performance. When evaluated with testing data (unseen data), the accuracy value is 88%, the precision value is 88%, recall is 88% and the f1-score is 88%. For the sentiment classification model with the bottoms category, the resulting accuracy value is 80%, precision is 81%, recall is 80%, and f1-score is 79%. For the sentiment classification model with the dresses category, the accuracy value is 81%, precision is 81%, recall is 81%, and f1-score is 81%. For sentiment classification with the tops category the resulting accuracy value is 77%, precision is 77%, recall is 77%, and f1-score is 77%.*

**Keywords:** e-commerce, review, sentiment, NLP, CNN

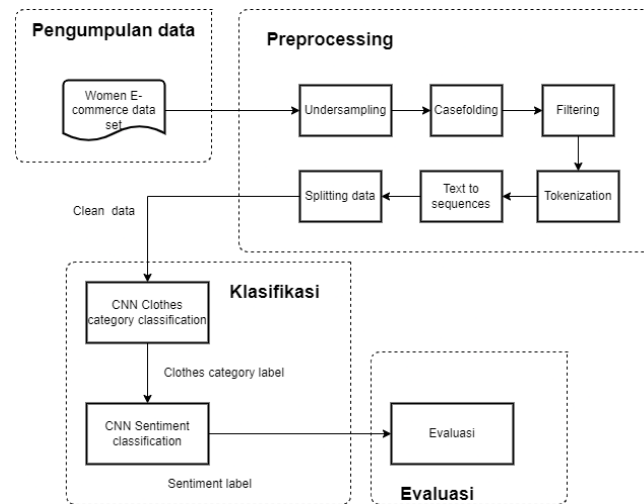
## 1. Pendahuluan

Beberapa pengguna e-commerce dalam menentukan membeli suatu produk umumnya melihat ulasan sebelum memutuskan untuk membeli barang. tepatnya 91% seseorang yang berusia 18-34 tahun memercayai ulasan online seperti halnya rekomendasi pribadi, dan 93% konsumen mengatakan bahwa ulasan online mempengaruhi keputusan pembelian mereka [1]. Berdasarkan data tersebut, konsumen membutuhkan ulasan sentimen guna membuat keputusan untuk memilih suatu produk/jasa. Namun hasil analisis sentiment yang dihasilkan masih kurang spesifik, maka dilakukan proses klasifikasi ulasan tersebut berdasarkan kategori ulasan. Sehingga analisis sentimen yang dihasilkan lebih spesifik. CNN biasa digunakan untuk mendeteksi dan mengenali objek pada sebuah gambar. CNN juga dapat diaplikasikan sebagai Natural Language Processing dengan mengubah representasi dari sebuah kalimat menjadi vektor. Implementasi CNN dalam NLP dapat menghasilkan performa yang baik. Seperti pada penelitian Rahman [2] hasil evaluasi untuk data opini publik pada Instagram mendapatkan hasil rata-rata precision 95.80%, recall 88.12%, f1-score 91.62%.

Dataset E-Commerce pakaian yang digunakan terbagi menjadi dua sentimen yaitu positif dan negatif. Pada penelitian ini melakukan analisis sentimen berdasarkan kategori yang terdapat di dalam dataset seperti tops, bottoms, jackets, dresses, dan intimates. Penelitian ini dilakukan dengan menggunakan dataset berbahasa Inggris yang diperoleh dari Kaggle dengan total keseluruhan data sebanyak 23.485 data dengan perbandingan data positif dan negatif 19314 data dengan label positif dan 4172 data dengan label negatif.

## 2. Metode Penelitian

Pada bagian menggambarkan secara umum bagaimana nantinya pelaksanaan penelitian akan dilakukan. Dimulai dari pengumpulan data yang didapatkan dari website Kaggle. Dataset berbahasa Inggris yang diperoleh dari Kaggle dengan total keseluruhan data sebanyak 23.485 data dengan perbandingan data positif dan negatif 19314 data dengan label positif dan 4172 data dengan label negatif. Data tersebut kemudian melalui *preprocessing*. Pada tahap preprocessing terdapat beberapa tahap seperti *undersampling*, *case folding*, *tokenization*, *filtering*, *text to sequences*, dan *splitting*. Tahap selanjutnya dilanjutkan dengan implementasi arsitektur CNN dan *tuning hyperparameter batch size*, *epoch*, *learning rate* dengan menggunakan metode *k-fold cross validation*. Tahap selanjutnya merupakan evaluasi dengan menggunakan *confusion matrix*.



**Gambar 1.** Alur Penelitian

Secara umum alur penelitian berdasarkan Gambar 1 akan dilakukan dengan berbagai tahapan yaitu pengumpulan data, preprocessing, *word embedding*, klasifikasi, dan evaluasi.

### 2.1. Pengumpulan data

Data yang digunakan pada penelitian ini adalah data sekunder. Data tersebut didapatkan dari website Kaggle. Dataset tersebut berjumlah sebanyak 23486 dengan 10 *data feature* namun hanya 3 *data feature* saja yang digunakan.

**Tabel 1.** Contoh dataset

Review Text	Positive FeedbackCount	Department Name
Absolutely wonderful - silky and sexy and comf...	0	Intimate
Love this dress! it's sooo pretty. i happene...	4	Dresses
Thisshirt is very flattering to all due to th...	6	Tops

**Tabel 2.** Distribusi Data

Kategori pakaian	Label data	Jumlah data
Tops	Positif	564
	Negatif	564
Dresses	Positif	564
	Negatif	564
Bottoms	Positif	564

	Negatif	564
--	---------	-----

Ulasan yang terdapat pada data telah dilabeli dengan label 0 untuk merepresentasikan produk tersebut tidak direkomendasikan dan 1 apabila produk tersebut direkomendasikan. Label pada masing-masing aspek tersebut memiliki sentimen positif dan negatif. Pada Tabel 1 merupakan contoh data yang digunakan pada penelitian ini yang berisikan data judul film, *reviewfilm*, dan nilai sentimen dari *review*. Setelah melalui preprocessing distribusi data menjadi seperti pada Tabel 2

## 2.2. Preprocessing

Sebelum *dataset* diolah menggunakan metode *Convolutional Neural Network*, *dataset* melalui tahapan *pre-processing* yang mana pada tahapan *pre-processing* terdapat beberapa proses didalamnya seperti *tokenization*, *case folding*, *text to sequences*, dan *filtering*. Tahap ini dilakukan untuk memastikan data yang akan digunakan telah bersih, sehingga dapat mencegah *error* atau turunnya performa mesin [3]. Setelah melewati tahapan *pre-processing* data tersebut dikelompokkan berdasarkan pada kategori pada masing-masing data. Kemudian representasi data tersebut diubah menjadi vektor dengan menggunakan algoritma *word embedding*.

### a. Under sampling

Proses ini merupakan proses untuk normalisasi jumlah data menjadi seimbang. Metode ini bekerja dengan mengurangi data mayoritas secara acak, sehingga jumlah data sama dengan data minoritas. Setelah melalui proses ini data yang digunakan sebanyak 3384 data dengan 3 kelas (*tops*, *bottoms*, *dresses*)

### b. Case folding

Pada tahapan *case folding* berfungsi untuk mengubah keseluruhan data teks yang ada dalam *dataset* menjadi huruf kecil sehingga keseluruhan teks memiliki bentuk standar[3]

### c. Filtering

Pada tahap ini dilakukan proses menghilangkan simbol-simbol, *stopwords*, dan angka pada kalimat. Sehingga tidak terdapat simbol, angka, dan *stopwords* dalam data yang digunakan.

### d. Tokenization

Pada tahapan *tokenization* berfungsi untuk merubah kalimat menjadi kata atau bisa juga disebut token. Kemudian pada tahap ini juga dilakukan pembuatan korpus kata berdasarkan kata-kata yang dihasilkan pada tahap ini[3]

### e. Text to sequences

Tahap ini dilakukan untuk merubah representasi kata dalam kalimat menjadi angka. Kata yang dirubah digantikan dengan index kata dalam korpus kata[4]

### f. Splitting dataset

Proses ini dilakukan dengan menggunakan metode *holdout*. Untuk pembagian *dataset* yaitu 90% data digunakan pada proses *training* dan 10% data digunakan untuk *testing*.

## 2.3. Word Embedding

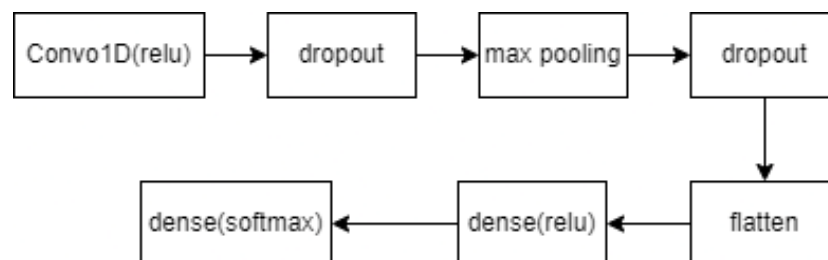
Setiap kata diubah menjadi vektor yang merepresentasikan sebuah titik pada *space* dengan dimensi tertentu. Kata-kata yang memiliki arti serupa memiliki representasi vektor dengan posisi yang tidak berjauhan antara satu dengan lainnya pada vektor *space* tersebut [5]

## 2.4 Convolutional Neural Network (CNN)

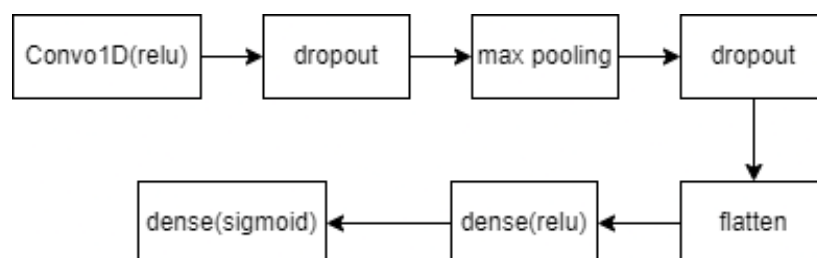
*Convolutional Neural Network* merupakan perkembangan dari *Artificial Neural Network* dengan penambahan *layer* pada *hidden layer*. CNN umumnya yang digunakan untuk proses klasifikasi gambar. CNN dapat diaplikasikan dalam *natural language processing* dengan mengubah representasi dari sebuah kalimat menjadi *matrix*. Dengan konfigurasi *Hyperparameter* dan vektor statis CNN dengan konfigurasi sederhana dapat menghasilkan hasil yang baik pada beberapa percobaan yang berbeda [6]. Arsitektur dari CNN yang akan diimplementasikan oleh penulis diambil berdasarkan referensi pada penelitian [2] Namun dengan sedikit perubahan pada jumlah *units* dan penambahan *layer dropout* setelah *layer konvolusi* dan *layer max pooling*. Setelah melalui proses *word embedding* data dimasukkan ke dalam lapisan *konvolusi*. Pada penelitian ini penulis akan menggunakan *konvolusi 1D*. Pemilihan *konvolusi 1D* dipilih dikarenakan ketika diaplikasikan dengan menggunakan teks sebagai data input, *konvolusi 1D* merupakan pilihan yang biasanya digunakan [7]

Sebuah kernel konvolusi (filter) berukuran tertentu digeser ke seluruh vektor tersebut. Pada setiap pergeseran menghasilkan *dot product*. Dengan menggeserkan filter tersebut ke seluruh vektor maka dihasilkan sebuah feature map. Setelah melalui lapisan konvolusi fitur aktivasi yang digunakan adalah ReLu. Output dari fungsi ReLu kemudian di masukkan ke dalam pooling layer. Untuk arsitektur yang digunakan pada penelitian ini dapat dilihat pada Gambar 2 dan Gambar 3. Arsitektur dari CNN yang akan diimplementasikan oleh penulis diambil berdasarkan referensi pada penelitian [2] dengan sedikit perubahan pada jumlah units dan penambahan layer dropout setelah layer konvolusi dan layer max pooling.

Setelah melalui proses word embedding data dimasukkan ke dalam lapisan konvolusi. Pada lapisan konvolusi terdapat proses perkalian *matrix* yang menghasilkan *feature map*. Setelah melalui lapisan konvolusi fitur aktivasi yang digunakan adalah ReLu. Output dari fungsi ReLu kemudian di masukkan ke dalam pooling layer. Bagian berikutnya merupakan lapisan max pooling layer, lapisan ini digunakan untuk mereduksi input secara spasial dengan menggunakan operasi down-sampling. Penulis menggunakan metode max pooling layer. *Dropout* ditambahkan setelah lapisan konvolusi dan lapisan *max pooling*. Penambahan parameter *dropout* bertujuan untuk mengurangi persentase terjadinya *overfitting* dan *underfitting*, selain itu juga mempercepat proses *learning*. Rentang nilai yang digunakan pada parameter *dropout* berkisar dari 0 sampai 1. Pada penelitian ini parameter *dropout* ditentukan secara eksplisit dengan rentang angka yang sesuai dengan data yang penulis miliki, untuk memberikan akurasi yang maksimal dari model tersebut dan dilakukan secara repetitif hingga ditemukan *Hyperparameter* terbaik. Lapisan *flattening* ditambahkan untuk memastikan dimensi dari vektor yang dihasilkan sebelumnya sesuai dengan dimensi vektor yang dibutuhkan pada layer selanjutnya. Pada lapisan selanjutnya terdapat lapisan *fully-connected* dengan jumlah units sebanyak 500 dan berdasarkan referensi yang diambil pada penelitian [2]. Pada output layer terdapat lapisan *fully-connected* dengan jumlah *neuron* tiga untuk mesin klasifikasi kategori dan fungsi aktivasi *softmax*. Sedangkan untuk mesin klasifikasi sentimen, pada output layer jumlah *neuron* sebanyak dua dan fungsi aktivasi *sigmoid*.



**Gambar 2.** Arsitektur mesin klasifikasi kategori



**Gambar 3.** Arsitektur mesin klasifikasi sentimen

### 3. Hasil dan Pembahasan

#### 3.1. Analisa Parameter

*Hyperparameter* yang akan diuji *batch size*, *learning rate*, dan *epoch* dengan menggunakan *k-fold cross validation* jumlah *fold* yang digunakan sebanyak lima. *K-fold cross validation* bekerja dengan membagi sampel data secara acak dan mengelompokkan data menjadi data train dan data validasi sebanyak nilai *k* pada *k-fold* [8]. Jumlah *fold* yang digunakan sebanyak lima. Proses *tuning*

*hyperparameter* dilakukan secara berurut dimulai dari *Tuning batch size* kemudian *learning rate* dan terakhir *epoch*.

**Tabel 3.** *Hyperparameter* untuk mesin klasifikasi kategori

No	Parameter	Ukuran
1	<i>Jumlah filter</i>	100
2	<i>Kemel size</i>	2
3	<i>Batch size</i>	16, 32, 64, 128
4	<i>Learning rate</i>	0.01, 0.001, 0.0001
5	<i>Epoch</i>	3, 5, 10, 15

**Tabel 4.** *Hyperparameter* untuk mesin klasifikasi sentimen

No	Parameter	Ukuran
1	<i>Jumlah filter</i>	100
2	<i>Kemel size</i>	2
3	<i>Batch size</i>	8, 16, 32, 64
4	<i>Learning rate</i>	0.01, 0.001, 0.0001
5	<i>Epoch</i>	3, 5, 10, 15

Untuk nilai parameter yang akan diuji dapat dilihat pada Tabel 3 dan Tabel 4. Terdapat perbedaan pada nilai yang diuji pada parameter *batch size*. Hal tersebut karena pada mesin klasifikasi sentimen berdasarkan kategori jumlah data yang tersedia lebih sedikit daripada data untuk mesin klasifikasi kategori.

a. *Tuning parameter batch size*

Proses ini dilakukan dengan memanfaatkan metode *5-fold cross validation* untuk mendapatkan performa terbaik dari mesin klasifikasi kategori dan mesin klasifikasi sentimen pada tiap kategori. Hasil tuning parameter *batch size* dapat dilihat pada Tabel 5 untuk klasifikasi kategori dan Tabel 6 untuk klasifikasi sentimen pada tiap kategori. Nilai yang diambil merupakan nilai rata-rata akurasi validasi tiap *5-fold* pada masing-masing nilai yang diuji. Berdasarkan Tabel 5 nilai parameter terbaik *batch size* untuk mesin klasifikasi kategori sebesar 64. Untuk mesin klasifikasi sentimen berdasarkan Tabel 6, Tabel 7, dan Tabel 8 sebesar 16.

**Tabel 5.** *Tuning batch size* klasifikasi kategori

<i>Learning Rate</i>	<i>Epoch</i>	<i>Batch Size</i>	<i>5-Fold cross validation</i>					<i>Validation Accuracy</i>
			<i>Fold-1</i>	<i>Fold-2</i>	<i>Fold-3</i>	<i>Fold-4</i>	<i>Fold-5</i>	
0.001	5	16	87,50%	91,00%	84,70%	86,20%	86,70%	87,30%
0.001	5	32	86,70%	91,50%	86,20%	86,20%	87,00%	87,50%
0.001	5	64	86,50%	90,40%	88,70%	87,80%	85,90%	87,90%
0.001	5	128	85,00%	90,60%	85,40%	87,00%	87,20%	87,10%

**Tabel 6.** *Tuning batch size* klasifikasi sentimen *bottoms*

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy Bottoms
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.001	5	8	82,30%	78,80%	81,30%	77,30%	78,30%	79,60%
0.001	5	16	82,30%	79,30%	85,20%	82,30%	88,10%	83,40%
0.001	5	32	81,80%	85,20%	84,20%	82,30%	75,90%	81,90%
0.001	5	64	82,80%	84,70%	80,30%	77,30%	82,30%	81,50%

**Tabel 7.** *Tuning batch size* klasifikasi sentimen *dresses*

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy Dresses
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.001	5	8	76,80%	65,50%	75,90%	86,70%	82,80%	77,30%
0.001	5	16	81,80%	80,30%	80,80%	84,20%	83,70%	82,20%
0.001	5	32	79,30%	77,30%	76,80%	85,70%	82,80%	80,40%
0.001	5	64	78,30%	83,30%	62,60%	80,80%	76,80%	76,40%

**Tabel 8.** *Tuning batch size* klasifikasi sentimen *tops*

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy Tops
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.001	5	8	82,80%	81,30%	80,30%	82,30%	82,80%	81,90%
0.001	5	16	81,80%	83,30%	80,80%	83,30%	87,70%	83,30%
0.001	5	32	82,80%	82,80%	81,30%	83,30%	84,20%	82,90%
0.001	5	64	84,70%	77,80%	76,80%	72,40%	81,80%	78,70%

b. *Tuning parameter learning rate*

Hasil tuning parameter *learning rate* untuk mesin klasifikasi kategori dapat dilihat pada Tabel 9 dengan nilai terbaik 0.001 sedangkan untuk nilai terbaik pada mesin klasifikasi sentimen berdasarkan Tabel 10, Tabel 11, dan Tabel 12 sebesar 0.001

**Tabel 9.** *Tuning learning rate* klasifikasi kategori

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.01	5	64	83,40%	90,30%	83,70%	83,90%	82,40%	84,80%

0.001	5	64	86,70%	91,10%	87,00%	85,90%	87,30%	87,60%
0.0001	5	64	68,90%	59,90%	73,60%	71,10%	67,50%	68,20%

**Tabel 10.** Tuning learning rate klasifikasi sentimen bottoms

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy bottoms
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.01	5	64	64,00%	79,30%	77,30%	79,30%	78,30%	75,70%
0.001	5	64	82,80%	86,70%	82,30%	84,20%	84,20%	84,00%
0.0001	5	64	61,60%	65,00%	63,50%	62,00%	59,60%	62,40%

**Tabel 11.** Tuning learning rate klasifikasi sentimen dresses

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy dresses
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.01	5	64	77,80%	75,90%	72,40%	75,40%	73,90%	75,10%
0.001	5	64	77,80%	78,80%	76,80%	82,30%	80,80%	79,30%
0.0001	5	64	64,00%	48,80%	57,60%	57,10%	72,90%	60,00%

**Tabel 12.** Tuning learning rate klasifikasi sentimen tops

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy tops
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.01	5	64	77,80%	65,50%	80,80%	73,40%	78,80%	75,30%
0.001	5	64	82,80%	81,30%	80,80%	79,30%	85,20%	81,90%
0.0001	5	64	61,00%	61,00%	57,60%	56,16%	59,60%	59,10%

c. Tuning parameter epoch

Proses *tuning parameter epoch* dilakukan sama seperti proses *tuning parameter batch size* dan *learning rate*. Berdasarkan Tabel 13 nilai *epoch* terbaik dihasilkan ketika *epoch* sebesar 5. Berdasarkan Tabel 14 dan Tabel 16 untuk nilai *epoch* terbaik pada mesin klasifikasi sentiman pada kategori *bottoms* dan *tops* sebesar 10 sedangkan untuk kategori *dresses* berdasarkan Tabel 15 nilai *epoch* terbaik sebesar 5.

**Tabel 13.** Tuning epoch klasifikasi kategori

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.001	3	64	87,80%	89,30%	85,90%	86,20%	87,50%	87,40%
0.001	5	64	87,80%	91,10%	87,40%	86,40%	88,00%	88,10%
0.001	10	64	86,90%	91,60%	88,90%	85,60%	87,00%	88,00%

0.001	15	64	88,20%	90,50%	87,40%	86,70%	87,20%	87,90%
-------	----	----	--------	--------	--------	--------	--------	--------

**Tabel 14.** Tuning epoch klasifikasi sentimen bottoms

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.001	3	16	82,80%	79,30%	78,80%	74,90%	81,80%	79,50%
0.001	5	16	82,80%	79,80%	81,30%	74,40%	85,70%	80,80%
0.001	10	16	82,80%	83,30%	80,80%	82,30%	82,80%	82,40%
0.001	15	16	83,30%	81,30%	81,30%	75,90%	82,30%	80,80%

**Tabel 15.** Tuning epoch klasifikasi sentimen dresses

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.001	3	16	83,70%	78,30%	77,30%	74,40%	83,20%	79,40%
0.001	5	16	83,30%	81,30%	78,80%	79,30%	84,70%	81,50%
0.001	10	16	82,80%	80,30%	76,80%	79,30%	85,70%	80,90%
0.001	15	16	80,80%	78,30%	80,80%	79,80%	81,80%	80,30%

**Tabel 16.** Tuning epoch klasifikasi sentimen tops

Learning Rate	Epoch	Batch Size	5-Fold cross validation					Validation Accuracy
			Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	
0.001	3	16	73,40%	80,80%	77,80%	82,30%	84,20%	79,70%
0.001	5	16	82,30%	81,30%	75,90%	87,20%	74,90%	80,00%
0.001	10	16	81,80%	80,30%	77,80%	81,80%	81,80%	80,70%
0.001	15	16	82,30%	78,30%	78,30%	75,40%	81,80%	79,80%

Kombinasi *hyperparameter* terbaik untuk mesin klasifikasi kategori didapatkan dengan nilai *epoch* 5, *learning rate* 0.001, dan *batch size* 64. Untuk mesin klasifikasi sentimen kategori *bottoms* kombinasi *hyperparameter* dengan performa terbaik didapatkan dengan nilai *batch size* 16, *learning rate* 0.001, dan *epoch* 10. Untuk mesin klasifikasi sentimen kategori *dresses* kombinasi *hyperparameter* dengan performa terbaik didapatkan dengan nilai *batch size* 16, *learning rate* 0.001, dan *epoch* 5. Untuk mesin klasifikasi sentimen kategori *tops* kombinasi *hyperparameter* dengan performa terbaik didapatkan dengan nilai *batch size* 16, *learning rate* 0.001, dan *epoch* 10.

### 3.2. Evaluasi

Evaluasi dari mesin dengan parameter terbaik dilakukan dengan menggunakan metode *confusion matrix*, karena mesin *deep learning* yang dilatih merupakan mesin klasifikasi [9]. Terdapat beberapa *metrics* yang dijadikan patokan seperti nilai *recall*, *precision*, *f1-score*. Pengukuran ketiganya nilai



tersebut dilakukan untuk mendapatkan informasi secara detail bagaimana performa dari mesin yang dilatih.

a. Mesin klasifikasi kategori

Hasil evaluasi model terbaik untuk klasifikasi kategori pakaian dengan tuning *Hyperparameterbatch size* 64, *learning rate* 0.001, dan *epoch* 5. Berdasarkan pada Tabel 17 rata-rata nilai recall, precision, dan f-score yang dihasilkan yaitu 88%.

**Tabel 17.** Evaluasi mesin klasifikasi kategori

	precision	recall	f1-score	Jumlah data
Bottoms	0.87	0.92	0.90	113
Dresses	0.96	0.82	0.89	113
Tops	0.82	0.89	0.86	113
accuracy			0.88	339
average	0.88	0.88	0.88	339

b. Mesin klasifikasi sentimen

Hasil evaluasi model klasifikasi sentimen kategori bottoms dengan tuning *Hyperparameterbatch size* 16, *learning rate* 0.001, dan *epoch* 10 dapat dilihat pada Tabel 18 mesin diuji dengan menggunakan data testing menunjukkan performa yang kurang dibandingkan mesin klasifikasi sentimen pada kategori lainnya. Rata-rata nilai recall yang dihasilkan 81%, precision 80%, dan f1-score 79%.

**Tabel 18.** Evaluasi mesin klasifikasi sentimen *bottoms*

	precision	recall	f1-score	Jumlah data
negative	0.74	0.91	0.82	56
positive	0.89	0.68	0.77	57
accuracy			0.80	113
average	0.81	0.80	0.79	113

Hasil evaluasi mesin klasifikasi *dresses* dengan tuning *Hyperparameterbatch size* 16, *learning rate* 0.001, dan *epoch* 5 pada Tabel 19 menunjukkan performa yang lebih baik dibandingkan dengan hasil evaluasi pada model klasifikasi sentimen kategori bottoms. Nilai rata-rata *precision*, *recall*, dan *f1-score* yang dihasilkan sama yaitu dengan tingkat persentase sebesar 81%.

**Tabel 19.** Evaluasi mesin klasifikasi sentimen *tops*

	precision	recall	f1-score	Jumlah data
negative	0.81	0.79	0.80	56
positive	0.80	0.82	0.81	57

accuracy			0.81	113
average	0.81	0.81	0.81	113

Berdasarkan pada Tabel 20 hasil evaluasi model klasifikasi *tops* dengan tuning hyperparameter batch size 16, learning rate 0.001, dan epoch 10 pada Tabel 20 menunjukkan performa yang baik. Nilai rata-rata precision, recall, f1-score, dan accuracy yang dihasilkan sama yaitu dengan tingkat persentase sebesar 77%.

**Tabel 20.** Evaluasi mesin klasifikasi sentimen *tops*

	precision	recall	f1-score	Jumlah data
negative	0.77	0.77	0.77	56
positive	0.77	0.77	0.77	57
accuracy			0.77	113
average	0.77	0.77	0.77	113

#### 4. Kesimpulan

Mesin klasifikasi kategori ulasan produk pakaian performa terbaik didapatkan dengan kombinasi hyperparameter batch size sebesar 64, learning rate sebesar 0.001, epoch sebesar 5 dengan akurasi validasi yang dihasilkan sebesar 88,1%. Evaluasi mesin dengan data *testing* menunjukkan *performa* yang baik dengan akurasi, *precision*, *recall* dan *f1-score* yang dihasilkan cukup tinggi ketika dievaluasi dengan *data testing (unseen data)*. Nilai akurasi yang dihasilkan sebesar 88%, nilai *precision* yang dihasilkan sebesar 88%, *recall* 88% dan *f1-score* sebesar 88%.

Performa terbaik pada mesin klasifikasi sentimen dengan kategori *bottoms* didapatkan dengan kombinasi hyperparameter batch size sebesar 16, learning rate sebesar 0.001, dan epoch sebesar 10 dengan akurasi validasi sebesar 82,36%. Pada mesin klasifikasi sentimen kategori *dresses* performa terbaik didapatkan dengan kombinasi hyperparameter batch size sebesar 16, learning rate sebesar 0.001, dan epoch sebesar 5 dengan akurasi validasi sebesar 81,5%. Untuk mesin klasifikasi sentimen kategori *tops* performa terbaik didapatkan dengan kombinasi hyperparameter batch size sebesar 16, learning rate sebesar 0.001, dan epoch sebesar 10 dengan akurasi validasi 80,7%. Hasil evaluasi mesin klasifikasi sentimen pada tiap-tiap kategori menunjukkan performa yang baik. Dimana mesin klasifikasi sentimen dengan kategori *bottoms*, nilai akurasi yang dihasilkan sebesar 80%, *precision* 81%, *recall* 80%, dan *f1-score* 79%. Pada kategori *dresses* nilai akurasi yang dihasilkan sebesar 81%, *precision* 81%, *recall* 81%, dan *f1-score* 81%. Pada kategori *tops* nilai akurasi yang dihasilkan sebesar 77%, *precision* 77%, *recall* 77%, dan *f1-score* 77%.

#### Referensi

- [1] D. Kaemingk, "Online reviews statistics to know in 2021," 2020. <https://www.qualtrics.com/blog/online-review-stats/> (accessed Jun. 12, 2021).
- [2] M. A. Rahman *et al.*, "Aspect Based Sentimen Analysis Opini Publik Pada Instagram dengan Convolutional Neural Network," *J. Intell. Syst. Comput.*, pp. 1–9, 2020.
- [3] K. D. Yonatha Wijaya and A. A. I. N. E. Karyawati, "The Effects of Different Kernels in SVM Sentiment Analysis on Mass Social Distancing," *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 9, no. 2, p. 161, 2020, doi: 10.24843/jlk.2020.v09.i02.p01.
- [4] M. T. Ari Bangsa, S. Priyanta, and Y. Suyanto, "Aspect-Based Sentiment Analysis of Online Marketplace Reviews Using Convolutional Neural Network," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 14, no. 2, p. 123, 2020, doi: 10.22146/ijccs.51646.
- [5] J. W. Gotama Putra, *Pengenalan Konsep Pembelajaran Mesin dan Deep Learning*. 2020.

- [6] K. Yoon, "Convolutional neural networks for sentence classification," *EMNLP 2014 - 2014 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, pp. 1746–1751, 2014, doi: 10.3115/v1/d14-1181.
- [7] Y. Goldberg, *Yoav Goldberg Book - Neural Network Methods in Natural Language Processing*. 2017.
- [8] A. Rohani, M. Taki, and M. Abdollahpour, "A novel soft computing model (Gaussian process regression with *K-fold* cross validation) for daily and monthly solar radiation forecasting (Part: I)," *Renew. Energy*, vol. 115, pp. 411–422, Jan. 2018, doi: 10.1016/j.renene.2017.08.061.
- [9] M. Swamynathan, *Mastering Machine Learning with Python in Six Steps*. 2019.

*This page is intentionally left blank*