

Averaged Word2vec sebagai Ekstraksi Fitur pada Analisis Sentimen Ulasan Film di IMDb menggunakan Artificial Neural Network (ANN)

Averaged Word2vec as Feature Extraction in Sentiment Analysis of Movie Reviews on IMDb with Artificial Neural Network (ANN)

Darwis Alwan¹, Muhammad Ali Ridla*²

¹Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang

²Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Ibrahimy, Situbondo

Email: ¹darwis.alwan@gmail.com, ²el.riedla@gmail.com

*Penulis Koresponden

Received: 05 Desember 2023

Accepted: 19 Januari 2024

Published: 05 Februari 2024



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).
Copyright (c) 2024 JUSTINDO

ABSTRAK

Analisis Sentimen sangat dibutuhkan saat ini, dengan kemajuan dan peningkatan penggunaan internet saat ini orang dapat dengan mudah mengungkapkan pendapatnya melalui internet (online). Salah satu situs paling populer yang dapat digunakan orang untuk mengungkapkan komentarnya adalah IMDb (Internet Movie Database), situs resmi untuk database informasi terkait film dan acara tv. Ada jutaan review orang untuk banyak film di situs IMDb. Pengguna IMDb semakin meningkat dari hari ke hari sehingga komentar terhadap film juga meningkat di situs tersebut. Untuk mengetahui jumlah reaksi positif dan negatif dengan mudah, maka perlu adanya analisis sentimen pada tinjauan data. Pada penelitian analisis sentimen ini, metode ekstraksi ciri yang digunakan adalah rata-rata Word2Vec. Word2Vec adalah teknik penyisipan kata yang dapat menangkap makna semantik kata. Untuk melengkapi proses analisis sentimen penelitian ini menggunakan metode ANN (Artificial Neural Network) sebagai pengklasifikasi. Berdasarkan hasil uji coba dengan dataset review film IMDb, metode yang diusulkan mampu mencapai akurasi sebesar 88,52%. Hasil ini membuktikan bahwa metode yang diusulkan mampu mengklasifikasikan sentimen ulasan film dan juga metode yang diusulkan ini merupakan klasifikasi yang sangat baik dengan luas area di bawah kurva (AUC) sebesar 95%. Selain itu, metode Word2Vec rata-rata memiliki akurasi yang lebih tinggi daripada metode berbasis frekuensi istilah, BOW dan TF-IDF.

Kata kunci: *Sentimen Analisis, NLP, Word2Vec, Jaringan Saraf Tiruan*

ABSTRACT

Sentiment Analysis is really needed at this time, with the advancement and increasing of internet use today people can express their opinion so easily through internet (online). One of the most popular sites that people can express their comments is IMDb (Internet Movie Database), official website for a database of information related to films and tv shows. There are millions of people reviews to many films at the IMDb site. IMDb users are increasing day by day so the comments to the films are also increasing at the site. In order to find out the number of positive and negative reactions easily, it is necessary to have a sentiment analysis on the data reviews. In this sentiment analysis research, the feature extraction method is averaged Word2Vec. Word2Vec is a word embedding technique that can capture the semantic meaning of words. For completing the sentiment analysis process this research uses the ANN (Artificial Neural Network) method as the classifier. Based on the results of trials with the dataset of IMDb film reviews, the proposed method is able to achieve an accuracy of 88,52%. This result prove that the proposed method is able to classify the sentiment of film reviews and also this proposed method is an excellent classification with an area under curve (AUC) of 95%. In addition, the averaged Word2Vec method has higher accuracy than the term frequency based methods, BOW and TF-IDF.

Keywords: *Sentiment Analysis, NLP, Word2Vec, Artificial Neural Network*

1. Pendahuluan

Saat ini *Artificial Intelligence* AI (kecerdasan buatan) terus berkembang dan telah banyak peneliti yang memanfaatkannya guna memecahkan masalah-masalah yang ada di dunia. Salah satu cabang ilmu dari AI tersebut adalah *Natural Language Processing* (NLP) (Sergio & Juan J, 2019). Salah satu penerapan NLP yang cukup terkenal adalah *Sentiment Analysis* (analisis sentimen) (S. Fouzia, Nagaratna, Hegde, & Khaleel, 2018). Analisis sentimen merupakan salah satu metode dari *Text Mining* yang mampu untuk mengidentifikasi dan mengekstrak opini atau sentimen dari perkataan langsung (*speech recognition*) atau dari data teks (Erik, Bjorn, Yunqing, & Catherine, 2013). Analisis sentimen bertujuan menemukan pola sebuah kalimat lalu mengklasifikasikan ke dalam sentimen positif dan negatif (Ananda, Putra, & Sigit, 2019).

Tidak dipungkiri lagi dengan kemajuan dan peningkatan secara drastis penggunaan internet saat ini, publik sudah sangat gampang menyampaikan pendapat atau opininya (Siti, 2019). Salah satunya adalah situs IMDb (*Internet Movie Database*) yang merupakan situs *official* pusat basis data daring informasi yang berhubungan dengan film dan *tv show* (IMDb, 1990-2019). Pengguna situs IMDb semakin hari semakin bertambah, ulasan publik berupa komentar terkait film pun semakin banyak. Agar dapat mengetahui jumlah reaksi positif dan negatif dengan mudah terkait data teks ulasan film yang cukup banyak tersebut maka perlu adanya analisis sentimen pada data review film itu. Sehingga nantinya dapat secara otomatis komputer mengklasifikasikan data ulasan yang negatif dan positif (Ravinder, Aakarsha, Shruti, Shaurya, & Pratyush, 2019). Maka dari itu studi ini mengusulkan untuk melakukan sentimen analisis pada data review film di situs IMDb.

Seperti yang telah dijelaskan sebelumnya, komputer tidak mampu memahami ekspresi manusia pada teks secara langsung layaknya manusia, disini dibutuhkan metode untuk mengubah data teks tersebut menjadi kumpulan angka atau *vector* yang biasa disebut *vectorization* (Sergio & Juan J, 2019). Ada beberapa metode ekstraksi fitur yang telah banyak digunakan, seperti yang sangat terkenal TF-IDF (*Term Frequency–Inverse Document Frequency*) yaitu pembobotan suatu kata pada sebuah dokumen. TF-IDF hanya menghitung bobot kata pada suatu dokumen tanpa peduli pada makna semantik pada kata tersebut (Ravinder, Aakarsha, Shruti, Shaurya, & Pratyush, 2019) (Xiaobo & Qingsong, 2017). Selanjutnya ekstraksi fitur pada teks yang sering digunakan adalah Word2Vec, yaitu algoritma *word embedding* yang mengkonversi data teks menjadi *vector* (Anastasia & Oxana, 2019). Pada Word2Vec setiap kata memiliki *vector* yang dapat diukur kemiripan katanya berdasarkan seberapa sering kata itu terjadi bersamaan. Hal ini mengindikasikan bahwa Word2Vec dapat mengerti makna semantic dari suatu kata, berbeda dengan metode TF-IDF (Nehal, Marwa, Abd, & Aliaa, 2019). *Vector* kata yang dihasilkan Word2Vec nantinya dapat digunakan untuk menyelesaikan beberapa masalah NLP *Natural Language Processing* (Rizka, Retno, & Adi, 2019).

Pada penelitian analisis sentimen review film ini, metode ekstraksi fiturnya akan menggunakan Word2Vec. Telah diketahui Word2Vec hanya *vector* yang merepresentasikan dari suatu kata bukan kalimat. Disini peneliti akan memilih *Average base* Word2vec karena dinilai lebih efektif dan sudah banyak digunakan pada penelitian analisis sentimen (Esther, et al., 2020). Seperti penelitian (Muhammad, Mohammad, Irwan, Radityo, & Andi, 2020) yang melakukan komparasi mengenai *Average base* Word2Vec dan *Bag of Centroid base* Word2Vec untuk analisis sentimen pada dataset komentar mahasiswa terhadap dosen, setelah dilakukan klasifikasi menggunakan SVM hasil akurasi tertinggi diperoleh dengan ekstraksi fitur *Average base* Word2Vec yaitu 84,8%. penelitian (Ananda, Putra, & Sigit, 2019) yang juga menggunakan Word2Vec dengan Naive Bayes sebagai algoritma *classifiernya*. Menggunakan dataset Twitter mengenai komentar publik terhadap pembangunan infrastruktur, akurasi yang didapatkan cukup biasa yaitu 64%. Berikutnya penelitian (Rizka, Retno, & Adi, 2019) menggunakan ekstraksi fitur Word2Vec dengan mencoba melakukan komparasi antara skip-gram model dan CBOV model pada dataset ulasan hotel. Klasifikasinya menggunakan algoritma CNN (*Convolutional Neural Network*) akurasi tertinggi didapat pada ekstraksi fitur Word2Vec Skip-gram model yaitu 89,02%.

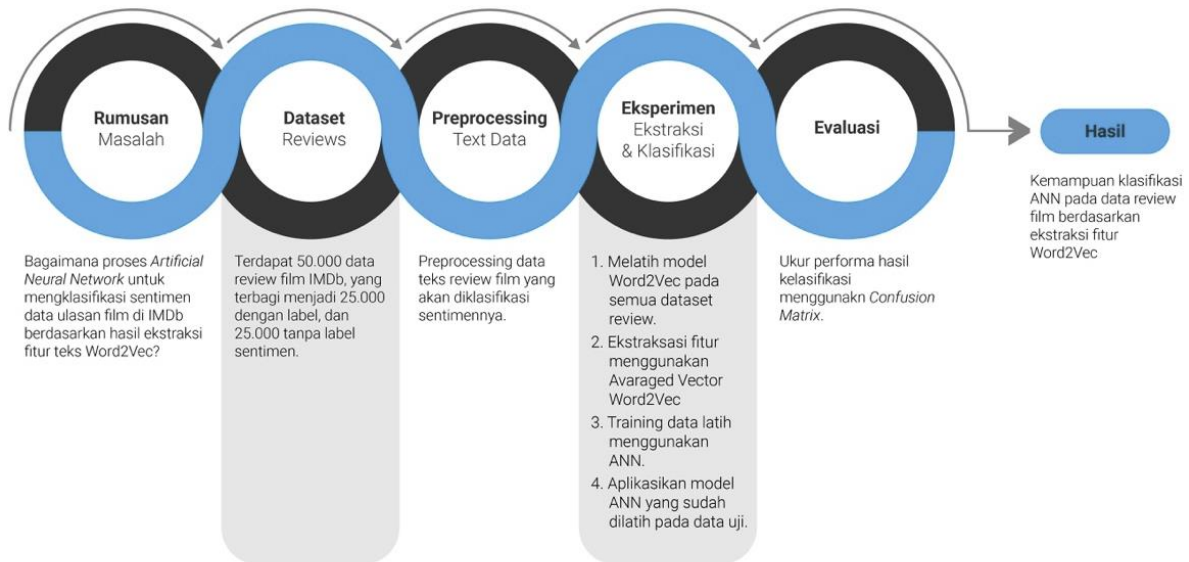
Metode ekstraksi fitur teks Word2Vec telah banyak dan efisien diaplikasikan pada analisis sentimen. Oleh karena itu akan diusulkan analisis dan klasifikasi sentimen pada dataset ulasan atau review film di IMDb menggunakan ekstraksi fitur *avaraged* Word2Vec yang selanjutnya akan diklasifikasi sentimennya menggunakan algoritma ANN. ANN (*Artificial Neural Network*) termasuk model dalam mengolah informasi yang merepresentasikan sistem kerja syaraf atau otak manusia yang saat ini sudah berkembang pesat, juga sebagai jaringan *neuron* yang saling berhubungan (Rahman,

Miftahur, 2019). ANN bekerja dengan menggabungkan neuron buatan untuk memproses informasi (Ridla, Muhammad Ali, 2018).

2. Metode Penelitian

2.1. Kerangka Pemikiran

Kerangka pemikiran dalam penelitian sebelumnya menunjukkan beberapa hasil analisis sentiment menggunakan berbagai macam teknik ekstraksi fitur serta algoritma classifier yang berbeda beda. Berikut ini kerangka pemikiran yang digunakan dalam penelitian ini:

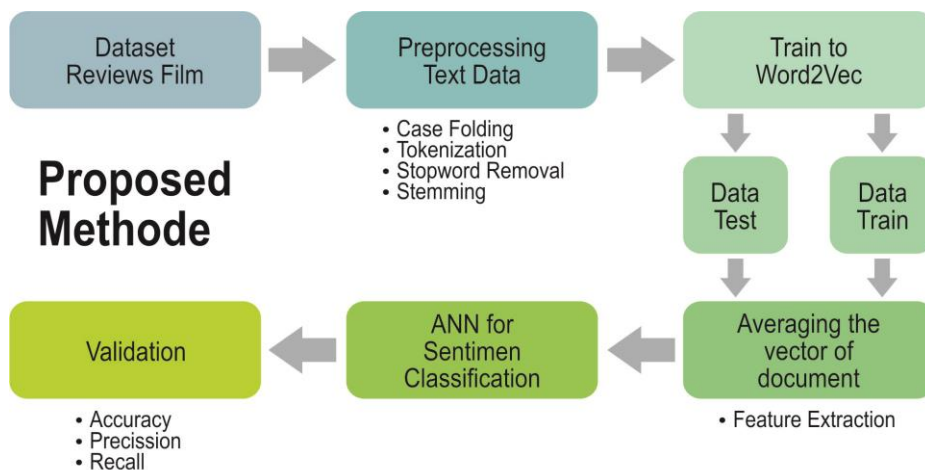


Gambar 1. Kerangka Pemikiran

Tujuan penelitian ini adalah analisis dan klasifikasi sentimen pada data review film di IMDb menggunakan algoritma Artificial Neural Network berdasarkan ekstraksi fitur averaged base Word2Vec, diharapkan mampu menghasilkan nilai akurasi yang tinggi.

2.2. Tahapan Penelitian

Pada penelitian ini untuk mengklasifikasi sentimen yang ada pada data review film IMDb metode penyelesaiannya adalah menggunakan model Artificial Neural Network dengan ekstraksi fitur menggunakan Word2Vec. Tahapan penelitian ini ditunjukkan pada gambar berikut ini:



Gambar 2. Metode yang digunakan

Adapun tahapan-tahapan yang dilakukan dalam penelitian ini, yaitu pengumpulan data, preprocessing data, ekstraksi fitur, eksperimen, dan evaluasi dengan rincian sebagai berikut:

2.2.1 Pengumpulan Data

Pada penelitian ini, data yang digunakan merupakan data skunder, yaitu data yang diperoleh dari data publik yaitu di Kaggle. Data yang diperoleh ialah data review film di IMDb, contoh datanya dapat dilihat seperti tabel 1 dibawah ini:

Tabel 1. Cuplikan Review dari Dataset Review Film di IMDb

ID	Review	Sentiment	Ket
11072	I thought this was a quiet good movie. It was fun to watch it. What I liked best where the 'Outtakes' at the end of the movie. They were GREAT.	1	Positive
...
2166	This is without a doubt the worst movie I have ever seen. It is not funny. It is not interesting and should not have been made.	0	Negative

Data tersebut sudah memiliki label sentiment, yaitu positif dengan label 1 dan negative dengan label 0. Metode penelitian yang digunakan dalam penelitian ini adalah metode kuantitatif, yaitu penelitian dengan melakukan eksperimen terhadap data review film di IMDb yang diperoleh secara global.

2.2.2 Preprocessing Data

Preprocessing adalah teknik yang akan mentransformasi dataset yang tidak terstruktur ke bentuk yang terstruktur. Tujuan transformasi tersebut agar dataset dapat diproses untuk tugas-tugas natural language processing berikutnya. Berikut urutan preprocessing pada teks dalam NLP (Novelty & Adiwijaya, 2020): (1) *Case Folding*: Tujuan case folding adalah merubah semua huruf capital pada sebuah dataset teks menjadi huruf kecil. (2) *Tokenization*: adalah tahapan pada preprocessing untuk memisahkan setiap kata pada kalimat atau dokumen menjadi potongan-potongan kata yang biasanya disebut dengan token. (3) *Stopword Removal*: adalah proses untuk menghilangkan kata-kata yang tidak memiliki makna atau tak dianggap penting serta menghilangkan tanda baca atau huruf tertentu yang dianggap tidak memiliki arti. (4) *Stemming*: adalah proses untuk menghilangkan infleksi (imbuhan) dari kata ke bentuk dasarnya (*root*).

2.2.3 Ekstraksi Fitur Menggunakan Averaged Word2Vec

Tahap berikutnya adalah *feature extraction* menggunakan Word2Vec dengan model *Skip-Gram* dan CBOW. Hasil akhir dari pelatihan Word2Vec ini adalah sebuah rangkaian vector dari setiap kata. Karena Word2Vec hanya menghasilkan vector yang hanya merepresentasikan sebuah kata maka untuk mendapatkan fitur dari satu ulasan yang terdiri dari banyak kata, digunakanlah teknik *averaging Word2Vec* guna mencari rata-rata beberapa kata dalam satu kalimat atau paragraph, digunakan sebagai masukan atau fitur pada algoritma ANN.

Model Word2Vec pertama kali dikenalkan oleh Mikolov di Google sebagai model ekstraksi fitur teks menjadi vector yang juga dapat mengenali arti semantik bahasa alami manusia pada kumpulan kata (Xiaobo & Qingsong, 2017). Selain itu, Word2Vec merupakan algoritma untuk merepresentasikan kata menjadi sebuah vector yang dapat memiliki kemampuan terbaik dalam tugas-tugas NLP dengan mengelompokkan kata yang mirip satu sama lain. Kata yang mirip akan memiliki vector yang sama (Rizka, Retno, & Adi, 2019).

Arsitektur model Word2Vec diproses menggunakan *Neural Network* dengan input berupa one hot encoding dari teks kata dan outputnya adalah ruang vector. Vector yang dihasilkan dari proses Neural Network ini adalah vector dengan dimensi kecil namun tetap dapat menangkap arti semantik dari setiap kata. Ada dua tipe arsitektur model Word2Vec yaitu model Skip-Gram dan Continuous Bag of Words (CBOW). Model Skip-Gram adalah model yang efisien untuk mempelajari vector kata yang besar pada teks yang tidak terstruktur. Model Skip-Gram bertujuan untuk memprediksi probabilitas konteks kata berdasarkan target kata, sedangkan model CBOW bertujuan untuk memprediksi konteks kata berdasarkan target kata (Ananda, Putra, & Sigit, 2019) (Rizka, Retno, & Adi, 2019).

Metode evaluasi pada model Word2Vec yaitu menggunakan teknik *Softmax* pada output layer. Nilai pada output layer adalah hasil distribusi probabilitas dari semua kata konteks berdasarkan kata

target menggunakan metode *Softmax*. Berikut persamaan dari teknik *Softmax* pada output layer ini (Rizka, Retno, & Adi, 2019):

$$P(w_{konteks}|w_{target}) = \frac{\exp(W_{output} \cdot h)}{\sum_{i=1}^V \exp(W_{output(i)} \cdot h)} \in \mathbb{R}^V \tag{1}$$

Pada persamaan (1) W_{output} merupakan vector baris ke i dari vector berdimensi $1 \times N$ dari matriks output. $W_{output(i)}$ juga merupakan baris vector berdimensi $1 \times N$ dari matriks output sesuai dengan kata konteksnya. V adalah jumlah seluruh vocab di corpus, dan h adalah jumlah neuron pada hidden layer. Hasil dari persamaan ini adalah nilai skalar probabilitas 1×1 dari kisaran $[0,1]$.

2.2.4 Eksperimen dan Pengujian Model *Artificial Neural Network*

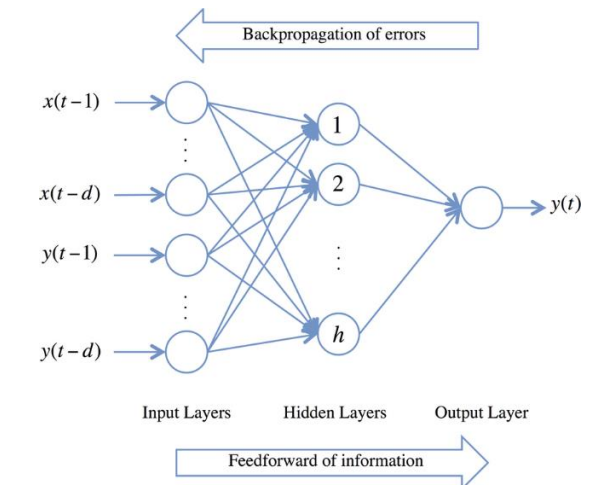
Pada penelitian ini dilakukan eksperimen menggunakan alat bantu tool library di python untuk mengimplementasikan model neural network dalam mengklasifikasi sentimen data review film di IMDb. Adapun langkah – langkah pengujian data pada model *neural network* yaitu dengan menentukan parameter neural network untuk mencari hasil klasifikasi terbaik.

Neural network atau yang dikenal dengan jaringan syaraf buatan merupakan salah satu algoritma yang terinspirasi dengan cara kerjanya otak manusia dan algoritma Neural Network ini dibangun dengan jaringan yang lebih sederhana dibandingkan dengan jaringan syaraf manusia sebenarnya (Ridla, 2018). Pada Neural Network terdapat komponen terkecil yang dikenal dengan neuron yang mana cara kerja neuron tersebut adalah mentransformasikan informasi yang diperoleh menuju neuron-neuron berikutnya. Artificial Neural Network (ANN), dikenal juga dengan istilah multilayer perceptron, memiliki persamaan umum sebagai berikut (ST. Aminah, Purwanto, & Catur, 2018):

$$y_k(x, w) = \sigma \left(\sum_{j=1}^M w_{kj}^{(2)} h \left(\sum_{i=1}^D w_{ji}^{(1)} x_i + w_{j0}^{(1)} \right) + w_{k0}^{(2)} \right) \tag{2}$$

Pada persamaan (2) W_{kj} , W_{ji} merupakan parameter bobot, dan W_{k0} , W_{j0} adalah parameter bias sedangkan $H(\cdot)$, $\sigma(\cdot)$ merupakan fungsi aktivasi.

Neural network dapat melakukan pembelajaran pada proses training menggunakan teknik backpropagation. Backpropagation merupakan suatu model yang menggunakan algoritma pembelajaran yang dapat diawasi serta struktur terhadap artificial neural network dan pada biasanya digunakan oleh perceptron dengan banyak layer lapisan untuk merubah beberapa bobot yang ada pada lapisan tersembunyi. Backpropagation merupakan jenis training dengan menggunakan mekanisme bobot yang dapat di kontrol untuk tercapainya nilai kesalahan yang kecil antara output hasil prediksi dengan output yang nyata (Zeeshan, Abdul, Chuangbai, Muhammad, & Tariq, 2020).



Gambar 3. Arsitektur Artificial Neural Network Backpropagation

2.2.5 Evaluasi

Evaluasi kinerja klasifikasi dari algoritma Artificial Neural Network yaitu menggunakan beberapa parameter yaitu Accuracy, precision dan recall dengan menggunakan Confusion Matrix untuk

mengetahui seberapa banyak data yang berhasil di klasifikasi dan yang gagal diklasifikasi menggunakan metode ANN.

3. Hasil dan Pembahasan

3.1. Ekstraksi Fitur Averaged Word2Vec

Proses ekstraksi fitur teks review menggunakan Word2Vec akan dilakukan menggunakan beberapa parameter Word2Vec yang berbeda seperti arsitektur Word2Vec CBOW/ Skip Gram, jumlah window, dan jumlah dimensi vektornya pada proses pembentukan *vector* dari setiap kata yang ada di *corpus* review film di IMDb. *Training* Word2Vec disini akan menggunakan 152.000 data review dari 3 dataset. Berikut cuplikan *vector* dari setiap kata hasil dari pelatihan Word2Vec:

Tabel 2. Cuplikan Vector Setiap Kata dari Seluruh Vocabulary dengan Jumlah Dimensi 100

No	Vocab	V1	V2	V...	V99	V100
1	plot	-0.0113	-0.0132	-0.3060	0.0369	-0.0360
2	movie	0.0454	0.0389	-0.1032	0.0327	-0.1104
3	couple	0.1778	-0.2619	0.0925	0.0750	-0.0716
...
28100	spilsbury	-0.0302	-0.0585	-0.0996	0.0744	-0.0619
28101	beller	-0.1131	0.0211	0.0249	0.0393	-0.1928

Selanjutnya melakukan ekstraksi fitur dari setiap ulasan film, untuk menggunakan *vector* setiap kata hasil dari pelatihan Word2Vec digunakanlah teknik *averaging* Word2Vec, yaitu menjumlahkan semua *vector* kata pada dokumen lalu digunakan nilai rata-ratanya (*averaging*). Berikut cuplikan implementasi dari teknik *averaging* Word2Vec dari data ulasan film yang memiliki label sentiment negative dan positive.

Tabel 3. Tampilan Dataset dengan Fitur dan Label Terstruktur

ID Doc	V1	V2	V...	V99	V100	Sentiment
3907	0.0289	-0.0041	-0.0263	-0.0118	0.0057	0
5355	-0.0005	-0.0593	0.0619	0.0354	-0.0280	0
...
10376	-0.0240	-0.0153	0.0190	0.0055	-0.0350	1
11072	-0.0353	-0.0102	-0.0923	0.0068	-0.0386	1

3.2. Klasifikasi Menggunakan ANN (Sentiment Analysis)

Eksperimen klasifikasi untuk sentiment analisis review data film menggunakan neural network akan dilakukan dengan beberapa tahapan dalam menggunakan parameter-parameter yang berbeda pada *neural network*. Eksperimen ini akan melakukan pengujian pada parameter-parameter tersebut sehingga nanti akan didapatkan parameter terbaik dalam mengklasifikasi sentiment pada dataset review film dengan ekstraksi fitur Word2Vec. Pertama akan dilakukan pengujian *neural network* terhadap dataset dengan metode ekstraksi fitur word2vec yang berbeda (*model*, *window size*, dan *vector size*). Berikut hasil pengujian neural network dengan model dan parameter Word2Vec yang berbeda pada dataset review film IMDb 25.000 data latih dan 25.000 data uji, hasil pengujian ditampilkan pada tabel 4.28 di bawah ini:

Tabel 4. Hasil Pengujian NN Berdasarkan Model & Parameter Word2vec

Model	Arsitektur Word2Vec			Neural Network		
	Window Size	Vector Size	Hidden Layer	Learning Rate	Training cycles	Accuracy
CBOW	20	500	22	0,01	100	81,35
CBOW	10	500	22	0,01	100	78,54
CBOW	5	500	22	0,01	100	78,62
...
CBOW	20	100	10	0,01	100	85,95
CBOW	10	100	10	0,01	100	85,62
CBOW	5	100	10	0,01	100	84,66
Skip Gram	20	500	22	0,01	100	82,34
Skip Gram	10	500	22	0,01	100	81,86
Skip Gram	5	500	22	0,01	100	81,85
...
Skip Gram	20	100	10	0,01	100	87,12
Skip Gram	10	100	10	0,01	100	87,01
Skip Gram	5	100	10	0,01	100	86,10

Dari hasil pengujian *neural network* pada tabel di atas, *Word2Vec model Skip Gram* dengan *window size: 20* dan *vector size: 100* menghasilkan tingkat akurasi tertinggi. Arsitektur model *Word2Vec* tersebut yang akan digunakan pada tahap eksperimen selanjutnya yaitu menggunakan pengujian parameter *neural network* yang berbeda. Berikut ini adalah penentuan parameter *neural network* untuk mendapatkan setting parameter terbaik:

Tabel 5. Hasil Pengujian Jumlah Neuron Hidden Layer
Arsitektur Word2Vec: Skip Gram 20, 100
Neural Network:

Hidden Layer	Learning Rate	Training cycles	Accuracy
10	0,01	100	87,12
20	0,01	100	87,03
30	0,01	100	86,16
40	0,01	100	86,16
50	0,01	100	86,06
60	0,01	100	85,50
70	0,01	100	85,39
80	0,01	100	84,80
90	0,01	100	84,00
100	0,01	100	84,84

Tabel 6. Hasil Pengujian Jumlah Trainig Cycle
Arsitektur Word2Vec: Skip Gram 20, 100
Neural Network:

Hidden Layer	Learning Rate	Training cycles	Accuracy
10	0,01	10	87,85
10	0,01	20	86,37
10	0,01	30	87,71
10	0,01	40	87,23
10	0,01	50	87,40
10	0,01	60	87,07
10	0,01	70	87,67
10	0,01	80	87,40
10	0,01	90	87,09
10	0,01	100	87,12
10	0,01	110	87,03
...
10	0,01	180	87,22
10	0,01	190	87,14
10	0,01	200	86,77

Tabel 7. Hasil Pengujian Nilai *Learning Rate*
Arsitektur Word2Vec: Skip Gram 20, 100
Neural Network:

Hidden Layer	Learning Rate	Training cycles	Accuracy
10	0,1	10	69,57
10	0,09	10	74,15
10	0,08	10	81,96
10	0,07	10	84,35
10	0,06	10	85,14
10	0,05	10	86,21
10	0,04	10	85,72
10	0,03	10	87,15
10	0,02	10	87,50
10	0,01	10	87,85
10	0,009	10	86,43
10	0,008	10	86,97
10	0,007	10	88,33
10	0,006	10	88,06
10	0,005	10	88,33
10	0,004	10	88,04
10	0,003	10	88,28
10	0,002	10	87,87
10	0,001	10	88,04

...
10	0,00003	10	81,01
10	0,00002	10	80,62
10	0,00001	10	76,23

Tabel 8. Hasil Pengujian Nilai Momentum

Arsitektur Word2Vec: Skip Gram 20, 100
Neural Network:

Hidden Layer	Learning Rate	Training cycles	Momentum	Accuracy
10	0,005	10	0,1	88,52
10	0,005	10	0,2	88,00
10	0,005	10	0,3	88,10
10	0,005	10	0,4	87,60
10	0,005	10	0,5	87,43
10	0,005	10	0,6	87,78
10	0,005	10	0,7	87,71
10	0,005	10	0,8	86,95
10	0,005	10	0,9	86,52

3.3. Evaluasi

Setelah melakukan pengujian parameter-parameter terbaik pada *neural network* yang digunakan untuk mengklasifikasi sentiment pada data review film maka diperoleh arsitektur terbaik dari setiap parameter, yaitu: jumlah *neuron* pada *hidden layer* 10, *learning rate* 0,005, *momentum* 0,1 dan *training cycle* 10. Hasil akurasi dari nilai parameter tersebut yaitu 88,52%. Arsitektur model neural network dari hasil eksperimen ini memiliki pola inputan 100 fitur berdasarkan hasil eksperimen neural network terhadap jenis model Word2Vec sebelumnya yaitu Word2Vec model skip gram dan memiliki dimensi vector 100 dengan window size 20. 100 fitur tersebut menjadi inputan neural network yaitu: x1, x2, x3,...,x100 dan jumlah neuron pada hidden layer 10 dan satu output layer untuk klasifikasi biner pada label sentiment.

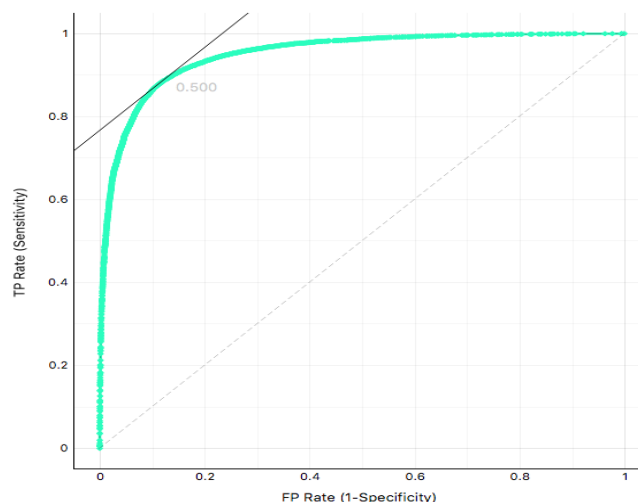
Selanjutnya akan ditampilkan hasil evaluasi dari pengujian klasifikasi dengan arsitektur terbaik tersebut menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, AUC (*Area Under Curve*) seperti pada Tabel 9 dan 10 serta gambar ROC pada Gambar 4.7 di bawah ini:

Tabel 9. Confusion Matrix

	1 (Positif)	0 (Negatif)
1 (Positif)	11.014	1.486
0 (Negatif)	1.384	11.116

Tabel 10. Accuracy, Precision, dan Recall

	Precision	Recall	Accuracy
0	89	88	88
1	88	89	89
Accuracy			88,52%
AUC (Area Under Curve)			95%

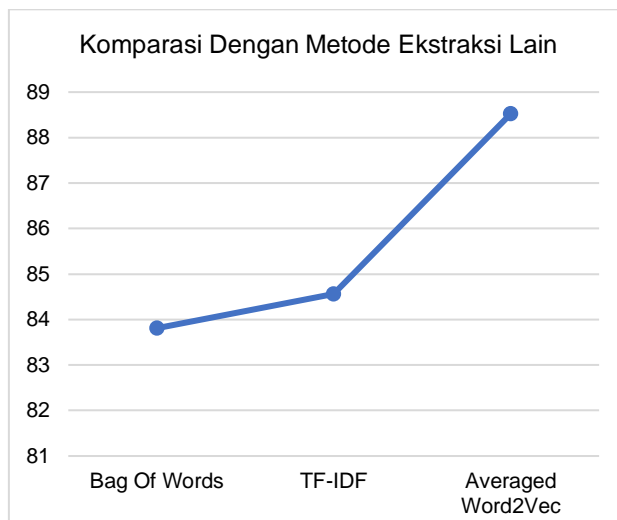


Gambar 4. ROC (Receiver Operating Characteristic)

Hasil penelitian menunjukkan bahwa metode yang diusulkan berhasil memperoleh akurasi 88,52%, hasil ini membuktikan bahwa metode yang diusulkan mampu mengklasifikasi sentiment ulasan film dengan *area under curve* (AUC) 95%. Nilai tersebut masuk dalam kategori *excellent classification* menurut (ST. Aminah, Purwanto, & Catur, 2018).

Tabel 11. State of The Art Terhadap Metode Ekstraksi Fitur Lain

No	Metode Ekstraksi Fitur	Dataset	Classifier	Akurasi
1	Bag Of Words	IMDb	ANN	83,81%
2	TF-IDF	IMDb	ANN	84,56%
3	Averaged Word2Vec	IMDb	ANN	88,52%



Gambar 5. Grafik Komparasi Terhadap Metode Ekstraksi Fitur Lain

Selain itu, metode ekstraksi fitur *averaged Word2Vec* memiliki akurasi yang lebih tinggi daripada metode berbasis *term frequency* yaitu BOW dan TF-IDF dengan perbandingan yang cukup jauh.

4. Kesimpulan

Pada penelitian analisis sentiment ulasan film IMDb ini dapat disimpulkan bahwa, penentuan parameter terbaik pada model Word2Vec dan neural network dalam melakukan analisis sentiment tersebut yaitu Word2Vec model *skip gram* dengan *window size* 20 dan dimensi *vector* 100, dan algoritma *neural network* dengan parameter jumlah neuron di *hidden layer* 10, *training cycles* 10, *learning rate* 0,005, dan *momentum* 0,1. Menggunakan parameter terbaik dari hasil eksperimen tersebut, metode yang diusulkan menghasilkan nilai akurasi 88,52%. Penelitian ini membuktikan bahwa metode yang diusulkan mampu mengklasifikasi sentiment pada ulasan film dengan baik dan masuk dalam kategori *excellent classification* dengan *area under curve* (AUC) 95%.

Saran untuk penelitian terkait untuk dapat juga menggunakan fitur *bigram* dan *trigram (phrase)* pada proses pelatihan Word2Vec guna mendapatkan hubungan semantic makna kata yang lebih baik dan natural terhadap bahasa manusia. Selain itu juga dapat menambahkan metode optimasi berupa *dimensional reduction* untuk meningkatkan kecepatan komputasi pada fitur *vector* teks yang pada umumnya sangat besar.

Daftar Pustaka

- Ananda, F. N., Putra, P. A., & Sigit, A. (2019). Analisis Sentimen Pembangunan Infrastruktur di Indonesia dengan Automated Lexicon Word2Vec dan Naive-Bayes. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, III(3), 2673-2679.
- Anastasia, D., & Oxana, V. (2019). Analytical Review of Methods for Identifying Emotions in Text Data. *VEUR-WS*. St. Petersburg.
- Cindy, A. P., Adiwijaya, & Said, A. F. (2020). Analisis Sentimen Review Film Berbahasa Inggris Dengan Pendekatan Bidirectional Encoder Representations from Transformers. *Jurnal Teknik Informatika dan Sistem Informasi*, VI(2), 181-193.

- Erik, C., Bjorn, S., Yunqing, X., & Catherine, H. (2013). New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent System*(13), 1541-1672.
- Esther, I. S., Adriel, F., Joan, S., Yosi, K., Gunawan, Surya, S., & Mauridhi, H. P. (2020). Analisis Pendapat Masyarakat terhadap Berita Kesehatan Indonesia menggunakan Pemodelan Kalimat berbasis LSTM. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, IX(1), 8-17.
- Farhan, H. K., Usman, Q., & Saba, B. (2017). SentiMI: Introducing Point-Wise Mutual Information with SentiWordNet to Improve Sentiment Polarity Detection. *Applied Soft Computing* , XXXIX(11), 140-153.
- IMDb. (1990-2019). *What is IMDb?* (IMDb) Retrieved 05 28, 2020, from <https://help.imdb.com/article/imdb/general-information/what-is-imdb>
- Mansi, L., Romy, K., Rachna, J., & Preeti, N. (2020). Sentiment classification based on public reviews of different states towards the police. *International Conference on Innovative Computing and Communication (ICICC 2020)*. New Delhi.
- Muhammad, R., Mohammad, R. F., Irwan, B., Radityo, A. N., & Andi, F. (2020). Ekstraksi Fitur Menggunakan Model Word2Vec Pada Sentiment Analysis Kolom Komentar Kuisisioner Evaluasi Dosen Oleh Mahasiswa. *Kumpulan jurnaL Ilmu Komputer (KLIK)*, VII(1), 35-47.
- Nehal, M. A., Marwa, M., Abd, E. H., & Aliaa, Y. (2019). Sentiment Analysis For Movies Reviews Dataset Using Deep Learning Models. *International Journal of Data Mining & Knowledge Management Process (IJDMP)* , IX(2), 19-27.
- Novelty, O. F., & Adiwijaya. (2020). Sentiment Analysis on Movie Reviews Using Information Gain and K-Nearest Neighbor. *Journal Of Data Science And Its Applications*, III(1), 1-7.
- Pozzi, F., Fersini, E., Messina, E., & Liu, B. (2017). Challenges Of Sentiment Analysis In Social Networks: An Overview. *Sentiment Analysis in Social Networks.*, 1-11.
- Rahman, Miftahur. (2019). Prediksi Pembayaran Tagihan Listrik Menggunakan Model Artificial Neural Network. *JUSTINDO (Jurnal Sistem dan Teknologi Informasi Indonesia)*, Vol. 4, No.1,, 7-12.
- Ravinder, A., Aakarsha, C., Shruti, K., Shaurya, G., & Pratyush, A. (2019). The Impact of Features Extraction on the Sentiment Analysis. *Procedia Computer Science*, 152, 341-348.
- Ridla, Muhammad Ali. (2018). Particle Swarm Optimization Sebagai Penentu Nilai Bobot Pada Artificial Neural Networkberbasis Backpropagation untuk Prediksi Tingkat Penjualan Minyak Pelumas Pertamina. *Jurnal Ilmiah Informatika Vol. 3 No.1*, 183-192.
- Rizka, P. N., Retno, K., & Adi, W. (2019). Word2Vec for Indonesian Sentiment Analysis towards Hotel Reviews: An Evaluation Study. *4th International Conference on Computer Science and Computational Intelligence 2019* . Semarang.
- S. Fouzia, S., Nagaratna, Hegde, P., & Khaleel, U. R. (2018). Wilcoxon Signed Rank Based Feature Selection for Sentiment Classification. *Second International Conference on Computational Intelligence and Informatics, Advances in Intelligent Systems and Computing*. Singapore.
- Sergio, A. L., & Juan J, C.-G. (2019). Supervised Learning Methods Application to Sentiment Analysis. *IDEAS 2019*. Athens.
- Siti, N. (2019). Analisis Sentimen Aplikasi Transportasi Online KRL Access Menggunakan Metode Naive Bayes. *Jurnal Swabumi*, VII(1), 31-38.
- ST. Aminah, D. G., Purwanto, & Catur, S. (2018). Neural Networkberbasis Algoritma Genetika Untuk Prediksi Kesempatan Kerja. *Jurnal Pendidikan Informatika Universitas Ivet*, 1(1), 55-65.
- Xiaobo, Z., & Qingsong, Y. (2017). Hotel Reviews Sentiment Analysis Based on Word Vector Clustering. *IEEE International Conference on Computational Intelligence and Applications*. Shanghai: IEEE.
- Zeeshan, S., Abdul, A. Z., Chuangbai, X., Muhammad, A., & Tariq, M. (2020). Sentiment analysis on IMDB using lexicon and neural networks. *Springer Nature Switzerland AG 2020*, II(148).