

## ANALISIS SENTIMEN APLIKASI E-LEARNING SELAMA PANDEMI COVID-19 DENGAN MENGGUNAKAN METODE SUPPORT VECTOR MACHINE DAN CONVOLUTIONAL NEURAL NETWORK

### SENTIMENT ANALYSIS FOR E-LEARNING APPLICATION USING SUPPORT VECTOR MACHINE AND CONVOLUTIONAL NEURAL NETWORK

Nina Ismaya Pangaribuan\*, Alek Sander Simbolon<sup>1</sup>, Nenni Mona Aruan<sup>1</sup>

<sup>1</sup>Institut Teknologi Del, Jl.P.I.Del,Sitoluama, Laguboti, Kabupaten Toba, Sumatera Utara

\*E-mail: ninapangaribuan94870@gmail.com

#### ABSTRAK

*Aplikasi e-learning dibutuhkan masyarakat dalam meningkatkan pendidikan di mana e-learning yang menjadi objek penelitian adalah Ruangguru dan Zenius karena memiliki jumlah pengguna lebih dari 16 juta dengan kepuasan pengguna lebih dari 8.5/10 dan lebih dari 1 juta kali di download di play store. Aplikasi tersebut memberikan ruang bagi pengguna aplikasi untuk mendapatkan tingkat kepuasan dari pengguna aplikasi. Sentiment analysis merupakan natural language preprocessing yang dapat digunakan dalam melakukan ekstraksi opini dari data berupa teks di mana tujuan penelitian ini melakukan evaluasi pada peningkatan hal positif dan memperbaiki hal yang negatif. Data ulasan yang diambil dari Twitter dan play store memiliki promosi dan giveaway yang akan berpengaruh pada pengolahan data dalam penentuan opini dan bukan opini. Penulis menggunakan metode lexicon based dalam memberikan label atau nilai sentiment pada setiap data. Pendekatan yang digunakan algoritma Support Vector Machine (SVM) dan Convolutional Neural Network (CNN) dalam melakukan klasifikasi terhadap data test yang di uji dari model yang telah dibangun. Berdasarkan hasil klasifikasi opini menjadi tiga kelas yaitu kelas positif, negatif, dan netral dari ulasan aplikasi Ruangguru dan Zenius. Dari nilai akurasi dan F-measure diperoleh bahwa klasifikasi yang terbaik adalah menggunakan algoritma CNN dengan akurasi dan F-measure memiliki nilai 86%.*

**Kata kunci:** *Sentiment analysis, lexicon based, Support Vector Machine, Convolutional Neural Network, zenius, Ruangguru, Machine learning*

#### ABSTRACT

*The community needs E-learning applications in improving education. The e-learning that becomes the object of this research are Ruangguru and Zenius because they have more than 16 million users with more than 8.5/10 user satisfaction and more than 1 million downloads from play store. The application provides space for application users to get a level of satisfaction from application users. Sentiment analysis is a natural language preprocess used to extract opinions from data in the form of text. The purpose of this study is to evaluate increasing positive things and fixing negative things. Reviews of data taken from Twitter and play store have promotions and giveaways that will affect data processing in opinions and not opinions. The author uses the lexicon-based method in providing labels or sentiment values for each data. The approach used is the Support Vector Machine (SVM) and Convolutional Neural Network (CNN) algorithms in classifying the test data that tested from the model we have built. Based on opinion classification into three classes, named positive, negative, and*

*neutral classes from the Ruangguru and Zenius application reviews. Of the accuracy and F-measure values, the best classification is using the CNN algorithm with an accuracy and F-measure value of 86%.*

**Keywords:** *Sentiment analysis, lexicon based, Support Vector Machine, Convolutional Neural Network, zenius, Ruangguru, Machine learning*

## PENDAHULUAN

Pendidikan merupakan kebutuhan pokok setiap individu yang memiliki manfaat dalam mengasah cara berpikir untuk menyelesaikan masalah dan bagaimana cara bersikap baik dan benar. Dalam perkembangannya didukung dengan adanya *electronic learning*, atau biasa disingkat dengan *E-learning* yang sebuah konsep atau metode pendidikan yang memanfaatkan teknologi informasi dalam proses belajar secara elektronik. Pada penelitian yang dilakukan ini memilih aplikasi Zenius dan Ruangguru sebagai objek, yang memiliki pengguna lebih dari 16 juta dan lebih dari 1 juta kali di download. Data komentar untuk Ruangguru dan Zenius akan diperoleh dari media *twitter* dan *play store*. Komentar yang telah dikumpulkan dari Januari – Mei 2021 akan menjadi data dalam melakukan sentiment analisis. Dimana data komentar tersebut akan dianalisis untuk mengetahui hal yang dapat dikembangkan dan hal yang diperbaiki dari aplikasi Ruangguru dan Zenius.

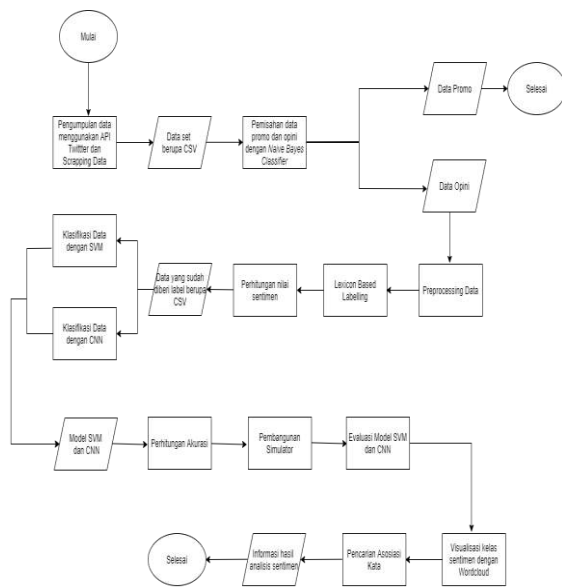
Pada penelitian yang berjudul *Metode Lexicon Based dan Support Vector Machine* untuk Menganalisis Sentimen pada Media Sosial Sebagai Rekomendasi Oleh-Oleh Favorit [1] terbukti bahwa metode *lexicon based* memberikan akurasi dan persisi yang lebih baik, sedangkan metode *support vector machine* memberikan hasil *recall* yang lebih baik. Pada penelitian berjudul *Sentiment Analysis pada Komentar Hotel Dengan Membandingkan Metode Support Vector Machine dan Naïve Bayes Classifier* [2] terbukti bahwa metode *support vector machine* menghasilkan nilai akurasi yang tinggi yaitu senilai 92.2%. Pada penelitian berjudul *Opinion Mining Terhadap Pemberitaan Corona di Instagram menggunakan Convolutional Neural*

*Network*, algoritma CNN berhasil melakukan klasifikasi data komentar menjadi sentimen positif, negatif, dan netral dan menghasilkan persentase presisi mencapai 96%, akurasi mencapai 88% pada data training dan 82% pada data testing, dan recall mencapai 68% [3]. Berdasarkan hasil penelitian terdahulu maka pada penelitian tugas akhir ini dipilih SVM dan CNN dengan pendekatan *lexicon* sebagai metode untuk klasifikasi sentiment karena metode tersebut terbukti efektif untuk pengklasifikasian pada sentimen.

Penelitian ini menggunakan dua algoritma yaitu *Support Vector machine* (SVM) dan *Convolutional Neural Network* (CNN). SVM juga melakukan identifikasi dengan memaksimalkan *margin* dan mendapatkan *hyperplane* yang diinginkan dalam melakukan klasifikasi data. Algoritma CNN telah membawa kemajuan besar dalam CV (*computer vision*) yang dapat diaplikasikan pada mobil kemudi otomatis, robotik, *drone*, *security*, diagnosis penyakit dan lain-lain. Tidak hanya terbatas pada gambar, CNN juga dapat diaplikasikan pada NLP berupa *text analytics* seperti sentimen analisis [4]. Algoritma ini dianggap efektif dalam melakukan klasifikasi karena memiliki lapisan-lapisan yang berfungsi untuk mempelajari fitur data dalam jumlah yang besar [3].

## METODOLOGI

Perencanaan alur analisis sentimen adalah sebuah gambaran umum yang menunjukkan alur dari penelitian yang sedang dilakukan. alur kerja bisa dilihat pada gambar 1 di bawah ini.



Gambar 1 Desain Percobaan

Alur yang dipakai pertama kali dalam penelitian ini adalah pengumpulan data komentar dengan cara crawling dan scrapping pada masing-masing akun Ruangguru dan Zenius. Selanjutnya data yang diperoleh masuk pada tahap pemisahan promo dan opini menggunakan algoritma Naïve Bayes. Setelah data telah terpisah lanjut pada tahap preprocessing yang memiliki beberapa tahap yaitu text cleaning, case folding, tokenization, stemming, normalization dan stop removal. Berikutnya dari hasil preprocessing dilakukan pelabelan menggunakan lexicon based pada data train. Setelah dilakukan pelabelan, data train akan digunakan dalam pembentukan model pada setiap algoritma. Berikutnya data test akan dimasukkan ke dalam model algoritma yang telah dibuat. Kemudian akan dilakukan perhitungan akurasi dan f-measure setiap algoritma. Asosiasi kata dilakukan dalam mempermudah untuk mengambil informasi.

#### A. Pengumpulan Data

Pengambilan data yang dilakukan dengan dua metode yaitu *crawling* menggunakan pemrograman Bahasa *python* pada *twitter*, yang dimana akan mengambil *tweet* dari keyword yang telah ditentukan. Metode yang kedua adalah metode *Scraping* yang menggunakan web

*scrapping* pada web *play store* untuk mengambil komentar dari pengguna aplikasi Zenius dan Ruangguru. Pada tahap pengumpulan data dibutuhkan waktu yang lama mulai dari Januari – Mei 2021, dimana peneliti melakukan pengumpulan data satu kali dalam seminggu untuk mengambil data yang terbaru.

#### B. Pemisahan Promo dan Opini

Pemisahan yang dilakukan dalam penelitian ini yaitu dengan mengumpulkan semua data dari setiap aplikasi. Setelah semua data terkumpul maka akan dilakukan pelabelan manual terhadap 1000 data setiap aplikasi untuk dijadikan sebagai data train. Data train akan digunakan untuk pembentukan model pada algoritma Naïve Bayes. Setelah terbentuknya model pemisahan promo untuk setiap aplikasi, tahap selanjutnya adalah memasukkan data test kedalam model yang telah dibangun. Maka akan didapatkan data opini yang dibutuhkan dan membuang data promo.

#### C. Preprocessing

Data opini yang telah dipisahkan dari promo belum dapat digunakan dalam klasifikasi pada setiap algoritma karena masih memiliki bentuk atau kalimat yang tidak beraturan atau tidak terstruktur dimana masih memiliki *noise*. Oleh karena itu dibutuhkan proses untuk mengekstrak informasi dari data komentar yang dapat digunakan dalam kepentingan analisis pada bidang klasifikasi. Adapun tahapan preprocessing yang dilakukan sebagai berikut;

1. Text cleaning, yaitu proses penghapusan atribut atau tanda baca yang tidak diperlukan pada proses analisis.
2. *Case Folding*, yaitu proses mengubah semua huruf besar menjadi *lowercase* atau huruf kecil.
3. *Tokenization*, yakni proses pemisahan kata perkata dalam satu komentar yang dilakukan berdasarkan karakter spasi yang terdapat dalam komentar.
4. *Stemming*, melakukan perubahan kata menjadi kata dasar.

5. *Normalization*, proses pengubahan kata slang atau gaul menjadi kata baku.
6. *Removing stopword*, merupakan proses penghapusan kata-kata yang dianggap tidak terlalu berpengaruh dalam analisis sentiment.

#### D. Pelabelan dengan Lexicon Based

*Lexicon based* merupakan salah satu cara dalam klasifikasi sentimen berdasarkan kamus kata. Pelabelan dilakukan untuk data train yang dimana data tersebut berfungsi dalam pembentukan model dari setiap algoritma. Pelabelan ini dilakukan dengan memanfaatkan korpus berisi kata positif dan negatif yang telah digunakan pada penelitian sebelumnya yaitu [5]. Setiap kata yang terdapat pada kamus positif akan diberi nilai 1, jika terdapat pada kamus negative akan diberi nilai -1, dan apabila tidak terdapat pada kedua kamus tersebut akan diberi nilai 0. Kemudian setiap ada kata negasi setelah maupun sebelum kata akan dikali dengan -1. Setelah semua kata dari sebuah kalimat telah diberi label, maka semua nilai akan di akumulasi. Apabila jumlahnya bernilai positif maka akan dikategorikan positif, bila negatif akan diklasifikasikan negatif, dan jika bernilai nol akan diklasifikasikan normal.

#### E. Klasifikasi Support Vector Machine (SVM)

*Support vector machine* merupakan salah satu algoritma machine learning yang memberikan hasil lebih baik dibanding algoritma machine learning lain dalam melakukan pengklasifikasian, regresi, dan prediksi. Dalam melakukan klasifikasi svm bekerja dengan mencari hyperplane yang terbaik, hal ini dibantu dengan pengaturan parameter  $c$  yang berfungsi untuk mengontrol pencarian *hyperplane* [2]. Setelah dilakukan pelabelan menggunakan lexicon based pada data train, data tersebut akan dimasukkan kedalam pembentukan model algoritma svm yang memiliki split data 80:20 berjumlah 1500 data. Kemudian akan dimasukkan data test berjumlah 500 data dimana akan dilakukan pengklasifikasian dengan memberikan

hasil *hyperplane* untuk memisahkan space dari dua buah class, yang secara sederhana metode ini merupakan pengelompokan secara terarah (*linear classifier*) dan kemudian dikembangkan menjadi mampu untuk bekerja secara *problem non linear* [2].

#### F. Klasifikasi Convolutional Neural Network (CNN)

*Convolutional Neural Network* merupakan salah satu algoritma *deep learning* yang biasa digunakan untuk mengklasifikasikan gambar, dapat digunakan untuk mengidentifikasi wajah, tanda-tanda lalu lintas, tumor, dan berbagai aspek data visual. CNN terdiri dari tiga lapisan utama yang saling terhubung untuk memproses data yang dimasukkan. Lapisan ekstraksi fitur terdiri dari lapisan *convolution* dan lapisan *pooling*. Kedua lapisan ini digunakan untuk memproses dan mempelajari fitur pada data yang dimasukkan. Lapisan selanjutnya merupakan lapisan *fully connected* yang digunakan untuk melakukan klasifikasi pada data masukan [6]. Pembentukan model algoritma CNN yang memiliki split data 80:20 berjumlah 1500 data. Kemudian akan dimasukkan data test berjumlah 500 yang akan diklasifikasikan kedalam tiga kategori yaitu positif, negatif, dan netral.

#### G. Wordcloud

Metode *text mining* merupakan metode yang menampilkan kata yang populer dengan menggunakan kata kunci data teks dan internet. Hasil dari metode ini adalah *wordcloud*. *Wordcloud* merupakan visualis dari data teks yang memiliki kegunaan untuk memetakan metadata kata kunci (tag) dalam menggambarkan suatu bentuk teks secara bebas. Secara sederhana *wordcloud* dapat membantu dalam menjelaskan pertanyaan pada penelitian dengan singkat. Besarnya ukuran kata yang ditampilkan pada *wordcloud* sebanding dengan frekuensi kemunculan kata pada sentimen [2].

## H. Asosiasi Kata

Korelasi dapat dikatakan sebagai relasi antara dua atau lebih variabel yang memiliki sifat kuantitatif (data yang dapat diukur) Sedangkan asosiasi diartikan sebagai relasi antara dua variabel atau lebih yang yang memiliki sifat kualitatif (data yang tidak dapat diukur). Jadi pada penelitian ini menggunakan metode pendekatan asosiasi di mana nantinya berguna untuk menemukan hubungan antar kata pada masing-masing sentimen komentar yang bersifat positif dan bersifat negatif yang dapat digunakan untuk memudahkan peneliti menemukan informasi dalam setiap sentimen [7].

## HASIL DAN PEMBAHASAN

### A. Pengumpulan Data

#### 1. Data Crawling

Pengumpulan data dalam penelitian dilakukan dengan metode *crawling* dengan menggunakan Bahasa *python*. Proses pengambilan data dibutuhkan akses dari *twitter* dengan mendaftarkan akun sebagai *developer* dan akan diberikan API agar dapat melakukan akses ke dalam *twitter* untuk mengambil *tweet* tertentu. Hasil dari *crawling* data berjumlah 37.438 data untuk Ruangguru dan 17.350 untuk aplikasi Zenius.

#### 2. Data Scraping

Pengumpulan data menggunakan metode *scraping* dilakukan dengan menggunakan *Web scraping*. Proses ini dilakukan setiap minggu agar mendapat komentar yang terbaru. Setiap aplikasi dilakukan pengumpulan data hingga 1000 data. Pada *playstore* memiliki komentar yang lebih objektif kepada aplikasi yang menjadi objek apabila dibandingkan dengan *twitter*.

### B. Pemisahan Promo dan Opini

Komentar data dipisah menjadi data train dengan split data 80:20 dan data test. Data train berjumlah 1000 data yang telah diberi label 1 sebagai promo dan 0 untuk opini. Kemudian akan dimasukkan kedalam pembentukan model pada algoritma Naïve Bayes. Setelah model

telah dibangun maka akan dimasukkan data test dari Ruangguru dan Zenius. Hasil pemisahan berjumlah 14.338 data opini untuk ruangguru dan 17.350 data untuk opini Zenius. Kemudian melakukan penghapusan data yang memiliki duplikat. Sehingga hasil akhir data untuk Ruangguru 4.923 data dan 11.831 data untuk Zenius. Hasil akurasi pada Ruangguru 77% dan hasil *f-measure* 76% sementara pada Zenius akurasi 78% dan *f-measure* 74%.

Tabel 1 Akurasi dan *F-measure* Pemisahan Opini

	Ruangguru	Zenius
Akurasi	0.77	0.78
F-measure	0.76	0.74

### C. Pelabelan Data dengan Lexicon Based

#### 1. Pelabelan Data Ruangguru

Data dari *preprocessing* yang berjumlah 5500 berkurang menjadi 5346 dikarenakan adanya sentimen yang hanya berisi kata yang terdapat pada *stopword*. Hasil pelabelan data dengan menggunakan *lexicon based* pada Ruangguru dari total data sebanyak 5346 data terdapat 1952 data yang positif, 811 data yang negatif, dan 2583 data yang netral.

#### 2. Pelabelan Data Zenius

Data dari *preprocessing* yang berjumlah 5500 berkurang menjadi 5310 dikarenakan adanya sentimen yang hanya berisi kata yang terdapat pada *stopword*. Hasil pelabelan data dengan menggunakan *lexicon based* pada Zenius dari total 5310 data terdapat 2692 data yang positif, 559 data yang negatif, dan 2059 data yang netral.

### D. Klasifikasi Algoritma Data

#### 1. Klasifikasi Algoritma SVM

##### a. Klasifikasi Sentimen Ruangguru

Pada pemodelan klasifikasi data Ruangguru dengan menggunakan algoritma SVM, setelah dilakukan pembagian dataset menjadi data train dan data test dengan perbandingan 80:20.

Akurasi yang dihasilkan untuk klasifikasi data ruangguru dengan menggunakan algoritma SVM adalah sebesar 83% dengan *f-measure* sebesar 83%. Hasil

akurasi sebesar 83% termasuk ke dalam kategori *good classifier*.

#### b. Klasifikasi Sentimen Zenius

Pada pemodelan klasifikasi data Zenius dengan menggunakan algoritma SVM, setelah dilakukan pembagian dataset menjadi data train dan data test dengan perbandingan 80:20.

Akurasi yang dihasilkan untuk klasifikasi data Ruangguru dengan menggunakan algoritma SVM adalah sebesar 76% dengan F-Measure sebesar 76%. Hasil akurasi sebesar 76% termasuk ke dalam kategori *fair classifier*.

## 2. Klasifikasi Algoritma CNN

### a. Klasifikasi Sentimen Ruangguru

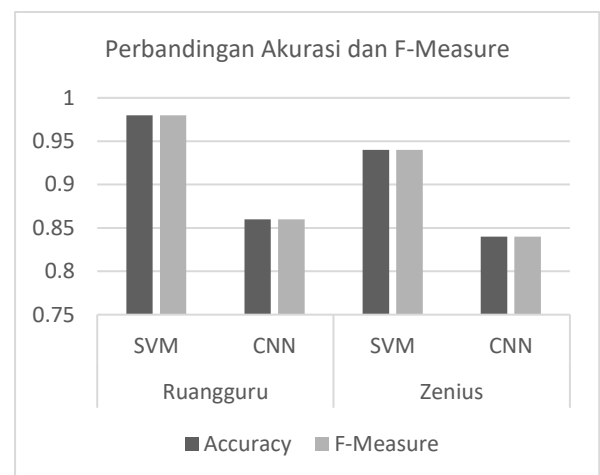
Pada pemodelan klasifikasi data Ruangguru dengan menggunakan algoritma CNN, setelah dilakukan pembagian dataset menjadi data train dan data test dengan perbandingan 80:20. Untuk performansi hasil klasifikasi dengan menggunakan metode CNN pada Ruangguru diinterpretasikan ke dalam tabel berikut. Akurasi yang dihasilkan untuk klasifikasi data Ruangguru dengan menggunakan algoritma CNN adalah sebesar 86% dengan F-Measure sebesar 86%. Hasil akurasi sebesar 86% termasuk ke dalam kategori *good classifier*.

### b. Klasifikasi Sentimen Zenius

Pada pemodelan klasifikasi data Zenius dengan menggunakan algoritma CNN, setelah dilakukan pembagian dataset menjadi data train dan data test dengan perbandingan 80:20. Hasil klasifikasi dapat diinterpretasikan ke dalam tabel berikut. Untuk performansi hasil klasifikasi dengan menggunakan metode CNN pada Zenius diinterpretasikan ke dalam tabel berikut. Akurasi yang dihasilkan untuk klasifikasi data Zenius dengan menggunakan algoritma CNN adalah sebesar 84% dengan F-Measure sebesar 84%. Hasil akurasi sebesar 84% termasuk ke dalam kategori *good classifier*.

### E. Perbandingan Metode SVM dan CNN

Parameter dalam perbandingan algoritma *Support Vector Machine* dan *Convolutional Neural Network*. Nilai akurasi memiliki fungsi dalam mengukur tingkat kebenaran dari kasus yang sedang diidentifikasi. Sementara *f-measure* yang merupakan hasil dari kombinasi dari nilai *precision* dan nilai *recall* yang digunakan untuk mengukur kinerja dari sistem klasifikasi. Hasil perbandingan akurasi dan *f-measure* antara kedua algoritma pada masing-masing divisualisasikan pada diagram berikut.

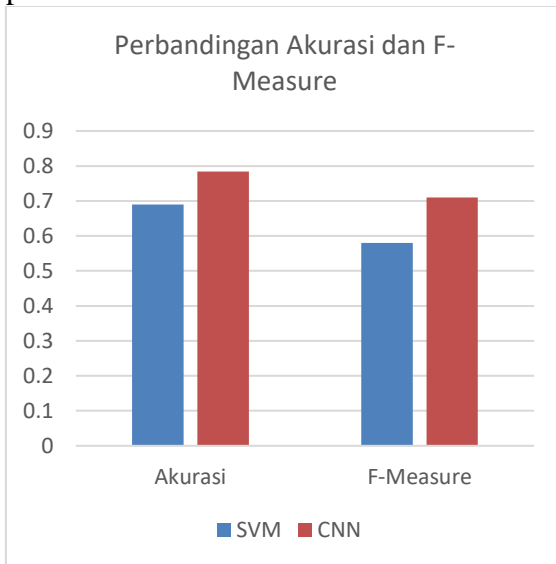


Gambar 2 Diagram Perbandingan Akurasi dan F-Measure

Dari gambar di atas dapat disimpulkan bahwa algoritma CNN memiliki performansi yang lebih baik dalam mengklasifikasikan sentimen pada data dari ke dua aplikasi baik Ruangguru maupun Zenius. CNN merupakan algoritma yang dianggap efektif dalam melakukan klasifikasi karena memiliki lapisan-lapisan yang berfungsi untuk mempelajari fitur data. Dalam melakukan pemodelan pada data yang telah dilatih pada kedua algoritma diketahui bahwa algoritma SVM termasuk ke dalam *good classifier* dan CNN termasuk ke dalam *good classifier* dan dapat diketahui juga bahwa kedua algoritma sudah baik dalam melakukan pemodelan.

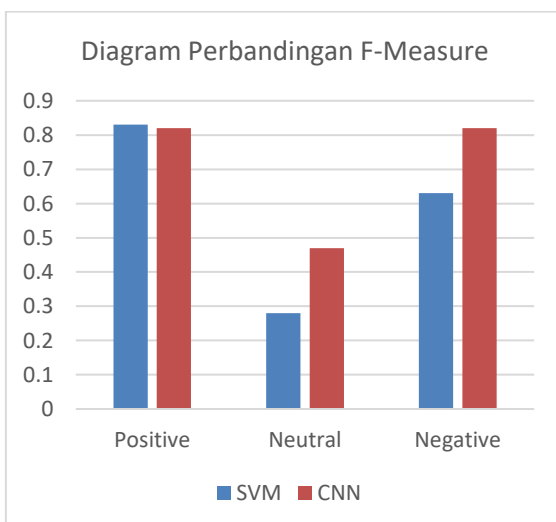
### F. Evaluasi Model Algoritma

Untuk mengevaluasi model masing-masing algoritma digunakan 500 data baru. Data baru dibersihkan terlebih dahulu dengan melakukan *preprocessing* sebelum masuk ke pelabelan dengan *lexicon* dan klasifikasi dengan algoritma. Data baru tersebut juga dipastikan merupakan data opini yang sudah dibersihkan dari data yang berupa promo.



Dari hasil di atas dapat diketahui pada evaluasi algoritma SVM dan CNN untuk data baru diketahui bahwa nilai akurasi dan *f-measure* CNN lebih tinggi dibandingkan dengan SVM dan hal ini sesuai dengan hasil perbandingan klasifikasi pada pemodelan.

Perbandingan ketepatan prediksi untuk masing-masing kelas dapat dilihat melalui diagram perbandingan akurasi berikut.



Dari gambar di atas dapat diketahui bahwa model yang telah dibuat pada masing-masing algoritma SVM dan CNN memiliki ketepatan prediksi paling baik pada kelas positif, diikuti dengan kelas negatif, dan kelas netral memiliki ketepatan prediksi yang paling buruk. Hal yang mempengaruhi prediksi kelas netral merupakan yang paling buruk dikarenakan oleh proses *learning* yang dibuat oleh peneliti dalam proses pembuatan model, tidak ada karakteristik khusus untuk kelas netral pada saat proses belajar. Kelas positif dan kelas negatif memiliki karakteristik khusus di mana kelas positif memiliki korpus positif di dalamnya dan kelas negatif memiliki korpus negatif di dalamnya.

### G. Visualisasi Wordcloud dan Asosiasi Kata

Wordcloud merupakan sistem yang mampu memvisualisasikan kata-kata dengan menyoroti frekuensi kata yang digunakan dalam suatu teks tertulis dan sebagai alat bantu untuk melakukan analisis guna memperoleh informasi tentang hal yang perlu dikembangkan dan diperbaiki dari Ruangguru dan Zenius. Data yang diaplikasikan pada wordcloud adalah data hasil klasifikasi dari algoritma CNN dikarenakan memiliki akurasi dan *f-measure* yang lebih baik daripada SVM.



Gambar 3 Visualisasi Wordcloud Positif Ruangguru





Tabel 5 Hasil dari Tabel Asosiasi Kata Negatif Ruangguru

Update (132)		Bug (50)	
Kata korelasi	Nilai korelasi	Kata korelasi	Nilai korelasi
Versi	0.22	Gambar	0.28
Offline	0.18	Resah	0.19
Error	0.16	Frame	0.17
Rusak	0.15	Gelap	0.14
Bug	0.14	Update	0.14
Lag	0.13	Gangguan	0.10

Dari hasil tabel asosiasi kata negatif di atas terdapat hal-hal yang harus diperbaiki oleh pihak Ruangguru seperti berikut:

- Aplikasi Ruangguru memiliki *bug* seperti layer menjadi gelap, menyatakan tidak ada jaringan, *error* setelah melakukan *update* aplikasi.
- Harga yang ditawarkan Ruangguru kurang dapat dijangkau oleh penguuna.
- Setelah melakukan *download* vidio, pengguna tidak dapat menonton vidio tersebut secara offline dikarenakan adanya *corrupt file*.

Tabel 6 Asosiasi Kata Positif Zenius

Ajar (411)		Aplikasi (359)		Bagus (251)	
Kata korelasi	Nilai korelasi	Kata korelasi	Kata korelasi	Nilai korelasi	Kata korelasi
Asyik	0.18	Bagus	0.36	Aplikasi	0.36
Materi	0.14	Materi	0.12	Tentor	0.16
Paham	0.13	Bantu	0.16	Vidio	0.12
Metode	0.14	Admin	0.15	Ajar	
Konsep	0.11	Lengkap	0.11		

Tabel 7 Hasil dari Tabel Asosiasi Kata Positif Zenius

Materi (144)		Vidio (127)	
Kata korelasi	Nilai korelasi	Kata korelasi	Nilai korelasi
Paham	0.31	Gratis	0.25
Lengkap	0.2	Paham	0.16
Konsep	0.18	Kreatif	0.16
Rangkum	0.18	Materi	0.12
		Langgan	0.14

Dari hasil tabel asosiasi kata di atas diperoleh informasi dari ulasan positif yang dapat dipertahankan atau

ditingkatkan oleh aplikasi Zenius seperti berikut:

- Zenius memiliki pengajaran yang asyik, mudah dipahami, dan memiliki tentor yang bagus.
- Vidio aplikasi Zenius kreatif, gratis dan mudah dipahami oleh pengguna.
- Zenius Memiliki materi yang lengkap, rangkum, dan mudah dipahami oleh pengguna.

Tabel 8 Asosiasi Kata Negatif Zenius

Ajar (213)		Vidio (149)		Materi (94)	
Kata korelasi	Nilai korelasi	Kata korelasi	Kata korelasi	Nilai korelasi	Kata korelasi
Mood	0.30	Buffer	0.18	Beda	0.26
Masalah	0.18	Urut	0.20	Sumber	0.11
Paksa	0.17	Load	0.18	Malas	0.15
Paham	0.19	Lag	0.17	Nalar	0.17
Langgan	0.13			keluh	0.21

Tabel 9 Hasil dari Tabel Asosiasi Kata Negatif Zenius

Bingung (57)		Download (53)	
Kata korelasi	Nilai korelasi	Kata korelasi	Nilai korelasi
Tampil	0.16	Hapus	0.39
Rangku	0.12	Cache	0.37
m	0.10	Nyesal	0.23
Siswa	0.25	Aplikasi	0.17
Campur		Daftar	0.16

Dari hasil tabel asosiasi kata negatif di atas terdapat hal-hal yang harus diperbaiki oleh pihak Zenius seperti berikut:

- Aplikasi Zenius memiliki pelayanan yang membuat pengguna resah dengan penawaran untuk berlangganan.
- Pada materi memiliki perbedaan sumber dari aplikasi android dan website.
- Vidio yang disediakan oleh aplikasi Zenius memiliki buffer lag, dan vidio tidan berurut.

## KESIMPULAN

Sentimen analisis pada komentar *e-learning* dapat diterapkan dengan melalui pemisahan komentar promo dengan komentar opini, *preprocessing*, pelabelan

komentar dengan *lexicon based*, klasifikasi komentar dengan metode *support vector machine* dan *convolutional neural network*, visualisasi *wordcloud* dan asosiasi kata untuk pencarian informasi sentimen. Hasil akurasi tertinggi *Convolutional Neural Network* adalah 86% dan *F-Measure* 86%. Hasil klasifikasi sentiment dan hasil asosiasi kata diperoleh informasi yang dapat dipertahankan dari ulasan positif dan hal yang diperbaiki dari ulasan negatif pada aplikasi Ruangguru dan Zenius. Pada pemodelan yang telah dibangun memiliki ketepatan prediksi paling baik pada kelas negatif dan positif.

### SARAN

Saran untuk penelitian seterusnya dapat menggunakan data set yang lebih besar dengan perbandingan setiap label yang seimbang pada data positif, negatif, dan neutral yang dimana akan meningkatkan ketepatan dalam melakukan klasifikasi pada sentimen analisis. Kemudian pada penelitian yang akan dilanjutkan disarankan menggunakan tingkat pelabelan yang lebih tinggi dalam melakukan pelabelan. Untuk mengurangi adanya *noise* yang lolos ada baiknya menggunakan algoritma pendeteksi *typo* agar dataset lebih baik.

### UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih kepada pada berbagai pihak yang telah membantu peneliti dalam menyelesaikan penelitian ini. Untuk itu peneliti mengucapkan terimakasih kepada Ibu Nenni Mona Aruan, S.Pd., M.Si selaku dosen pembimbing, Bapak Lit Malem Ginting, S.Si, MT dan Bapak Anthon Roberto Tampubolon, S.Kom, M.T sebagai dosen penguji yang telah memberikan kritik dan saran yang membangun dalam penyelesaian penelitian ini.

### DAFTAR PUSTAKA

- [1] H. Himawan, D. G. P. D.D.R and W. Kaswidjanti, "METODE LEXICON BASED DAN SUPPORT VECTOR MACHINE UNTUK MENGANALISIS SENTIMEN PADA MEDIA SOSIAL SEBAGAI REKOMENDASI OLEH-OLEH FAVORIT," *Seminar Nasional Informatika 2018 (semnasIF 2018)*, p. 1, 2018.
- [2] F. Siagian and C. N. Manalu, C. Sentimen Analisis pada Komentar Hotel dengan Membandingkan Metode Support Vector Machine dan Naïve Bayes Classifier, Laguboti, 2020.
- [3] A. R. Maulana and N. Rochmawati, "Opinion Mining Terhadap Pemberitaan Corona di Instagram menggunakan Convolutional Neural Network," *Journal of Informatics and Computer Science*, 2020.
- [4] C. Nicholson, "pathmind," A.I. Wiki, 2020. [Online]. Available: <https://wiki.pathmind.com/convolutional-network>. [Accessed April 2021].
- [5] D. H. Wahid and A. SN, "Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity," *IJCCS*, 2016.
- [6] Y. E. Mahendra, R. Ilyas and F. Kasyidi, "Klasifikasi Kalimat Ilmiah Menggunakan 1D Convolutional Neural Networks," *Prosiding The 11th Industrial Research Workshop and National Seminar*, p. 2, 2020.
- [7] U. Rofiqoh, R. S. Perdana and M. A. Fauzi, "Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexicon Based Features," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, p. 1727, 2017.