Institute of Research and Publication Indonesia
MALCOM: Indonesian Journal of Machine Learning and Computer Science
Journal Homepage: https://journal.irpi.or.id/index.php/malcom
Vol. 1 Iss. 1 April 2021, pp: 24-33
P-ISSN: 2797-2313 E-ISSN: 2775-8575

Sentiment Analysis of Online Lectures in Indonesia from Twitter Dataset Using InSet Lexicon

Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon

Desi Musfiroh[1*], Ulfa Khaira[2], Pradita Eko Prasetyo Utomo[3] , Tri Suratno[4]

[1,2,3,4]Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,
Universitas Jambi, Indonesia

E-Mail: [1]musfiroh.desi@gmail.com, [2]ulfa.ilkom@gmail.com,
[3]pradita.eko@unja.ac.id, [4]triel123@gmail.com

Received January 13th 2021; Revised February 18th 2021; Accepted February 24th 2021
Corresponding Author: Desi Musfiroh

Abstract

The implementation of online lectures on various campuses in Indonesia has been emphasized since the outbreak of corona virus. Online lectures are used as a solution to continue teaching and learning activities during pandemic. But the implementation of online lectures raises a variety of opinions in the community, especially among lecturers. It also raises the attitude of the pros and cons from various parties. For this purpose, data mining from Twitter analyzes sentiment on the topic of "online lectures". The data is classified into 3 classes, i.e. positive, negative, and neutral. This research was conducted with a lexicon-based approach technique using InSet Lexicon as an Indonesian opinion dictionary. The determination of the sentiment class for each sentence is obtained from the result of the polarity score calculation. Classification results from 5811 tweet data were found to contain 63.4% negative tweets, 27.6% positive tweets, and 8.9% neutral tweets. Testing of classification results was done by cross-validation method and confusion matrix with a comparison of training data and test data is 8:2 gave accuracy value of 79.2%, precision by 72.9%, recall by 62.8%, and f-measure of 67.4%.

Keyword: InSet Lexicon, .Online Lectures, Sentiment Analysis, Twitter,

Abstrak

Pelaksanaan perkuliahan daring pada berbagai kampus di Indonesia telah dipertegas sejak makin mewabahnya virus corona. Kuliah daring dijadikan solusi untuk terus menjalankan kegiatan belajar-mengajar di masa pandemi. Namun pelaksanaan perkuliahan daring menimbulkan berbagai macam opini dalam masyarakat, khususnya di kalangan pelajar. Hal ini juga memunculkan sikap pro maupun kontra dari berbagai pihak. Untuk itu dilakukan penambangan data dari *twitter* guna menganalisis sentimen terhadap topik "kuliah daring". Data diklasifikasikan ke dalam 3 kelas, yaitu positif, negatif, dan netral. Penelitian ini dilakukan dengan teknik *lexicon-based approach* menggunakan *InSet Lexicon* sebagai kamus kata opini berbahasa Indonesia. Penentuan kelas sentimen untuk setiap kalimat diperoleh dari hasil perhitungan *polarity score.* Hasil klasifikasi dari 5811 data *tweet* ternyata mengandung 63.4% *tweet* negatif, 27.6% *tweet* positif, dan 8.9% *tweet* netral. Pengujian hasil klasifikasi dilakukan dengan metode *cross-validation* serta *confusion matrix* dengan perbandingan data latih dan data uji yaitu 8:2 memberikan nilai *accuracy* 79.2%, *precision* sebesar 72.9%, *recall* sebesar 62.8%, dan *f-measure* sebesar 67.4%.

Keyword: *InSet Lexicon,* Kuliah Daring, *Sentiment Analysis*, *Twitter*

1. PENDAHULUAN

Pandemi COVID-19 yang terjadi belakangan ini telah menimbulkan dampak dan perubahan besar dalam berbagai bidang kehidupan. Salah satu yang menjadi fokus perhatian yaitu pada dunia pendidikan di Indonesia, dampaknya adalah terjadi peralihan sistem pembelajaran menjadi sistem daring atau jarak jauh demi meminimalisir potensi penyebaran virus corona. Melalui Surat Edaran Kementerian Pendidikan dan Kebudayaan RI Tanggal 17 Maret 2020 perihal Pembelajaran secara Daring dan Bekerja dari Rumah dalam Rangka Pencegahan Penyebaran *Coronavirus Disease* (COVID-19), ditetapkan bahwa bagi perguruan tinggi

di bawah naungan Kementerian Pendidikan dan Kebudayaan RI pada daerah yang sudah terdampak COVID-19 harus memberlakukan pembelajaran secara daring [1]. Dengan demikian, aktivitas perkuliahan secara tatap muka ditiadakan untuk sementara waktu hingga kondisi yang memungkinkan. Kegiatan perkuliahan di universitas atau perguruan tinggi dinilai sangat memungkinkan untuk dilakukan secara daring.

Munculnya kebijakan baru tentu bukan hal yang mudah untuk diikuti. Masyarakat khususnya para mahasiswa, dosen, maupun civitas akademika yang terlibat langsung dengan kegiatan perkuliahan perlu beradaptasi dengan kebijakan tersebut. Tak jarang kuliah secara daring dianggap membawa beragam kendala baru dalam perkuliahan, dan tak sedikit pula yang menganggap bahwa kuliah daring sebagai solusi yang paling tepat untuk tetap menjalankan kegiatan perkuliahan di tengah kondisi pandemi yang memprihatinkan.

Menanggapi kebijakan tersebut, banyak masyarakat yang mengutarakan berbagai macam pendapat, opini, maupun pandangan mereka terhadap pelaksanaan perkuliahan daring. Opini tersebut umumnya dikemukakan pada media sosial, salah satunya melalui *twitter*. *Twitter* menjadi situs jejaring sosial yang populer digunakan saat ini. Masyarakat dapat dengan mudah mengungkapkan berbagai macam komentar, pikiran, dan tanggapan mereka berkaitan dengan kondisi yang ada saat ini pada media sosial *twitter* [2]. Berdasarkan laporan yang diterbitkan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) bahwa 1,7% dari keseluruhan jumlah pengguna internet di Indonesia atau sekitar 291.000.417 dari 171.176.716,8 orang merupakan pengguna aktif dari media sosial *twitter* [3].

Twitter menjadi media yang cukup baik dalam memperoleh data karena tingkat akurasi dari kebenaran kalimat opini (*tweet*) yang diunggah ke *twitter* dinilai cukup tinggi jika digunakan untuk mengetahui bagaimana pendapat masyarakat terhadap suatu topik [4]. Melalui *twitter*, pengguna dapat mengunggah konten sesuai keinginan. Konten tersebut berupa opini, sentimen, maupun emoticon, yang bisa menjadi data untuk menganalisis suatu *trend* atau topik tertentu. Upaya menganalisis data tersebut dinamakan *sentiment analysis* atau *opinion mining* [5]. *Sentiment analysis* termasuk cabang ilmu dari *text mining*, *natural language program,* dan *artificial intelegence* yang dilakukan untuk memperoleh informasi yang bermanfaat atau pengetahuan baru dengan cara mengekstrak, memahami, dan mengolah data teks secara otomatis [6]. Melalui proses analisis sentimen akan terlihat bagaimana kecenderungan pendapat atau opini seseorang terhadap suatu topik atau masalah dengan menenetukan klasifikasi sentimen ke dalam dua kelas atau lebih.

Umumnya, ada 2 pendekatan dalam melakukan *sentiment analysis*, yaitu secara *learning-based* (pendekatan menggunakan *machine learning*) dan *lexicon-based* (pendekatan berbasis leksikal). Pendekatan *learning-based* menggunakan dataset yang telah diklasifikasikan secara manual sebelumnya sebagai data latih untuk menghasilkan klasifikasi teks opini secara otomatis. Sedangkan pendekatan *lexicon-based* bergantung pada kamus opini (*lexicon*) untuk penentuan klasifikasi. Kamus opini mengandung sejumlah kata yang digunakan untuk mengidentifikasi jenis opini suatu kalimat [7].

Lexicon-based approach adalah metode ilmiah yang sering digunakan dalam suatu penelitian analisis sentimen. Cara kerja metode ini adalah dengan menggunakan sebuah kamus kata atau *corpus* yang dilengkapi dengan bobot pada setiap katanya sebagai sumber bahasa atau leksikal. Hasil analisis dengan metode ini berupa klasifikasi sentimen positif, negatif, dan netral. Metode ini adalah bagian dari *machine learning* yang bersifat *unsupervised*. Kualitas dari hasil tergantung pada kamus kata atau *corpus* yang digunakan [8].

Lexicon yang digunakan dalam penelitian ini adalah *InSet Lexicon* karena sudah teruji cukup baik untuk analisis sentimen data berbahasa Indonesia. *InSet Lexicon* (Indonesia Sentiment lexicon) terdiri dari 3.609 kata positif dan 6.609 kata negatif berbahasa Indonesia yang telah memiliki bobot nilai atau *polarity score* pada setiap katanya dengan kisaran bobot antara -5 sampai +5. *Polarity score* ini digunakan untuk mengklasifikasikan jenis sentimen. Contoh dari kata negatif dan positif beserta bobotnya yang termuat dalam *InSet lexicon* dapat dilihat pada tabel berikut.

Tabel 1. Contoh Daftar Kata pada *InSet Lexicon*

Kata	Bobot
selisih	-4
sengsara	-5
awas	-3
iri	-4
coretan	-1
wewenang	2
sahabat	4
warisan	3
menarik	4
motivasi	3

InSet lexicon disusun oleh Fajri Koto dan Gemal Y. Rahmaningtyas pada penelitian sebelumnya dengan menggunakan kata-kata yang dikumpulkan dari *twitter* sebagai media sosial yang umum digunakan di

Indonesia. *InSet lexicon* dibangun untuk mengidentifikasi opini tertulis dan mengkategorikannya menjadi opini positif atau negatif yang bisa digunakan untuk menganalisis sentimen publik terhadap topik, acara, atau produk tertentu. Hasil dari tes dan evaluasi penelitian tersebut menunjukkan bahwa *InSet lexicon* mampu memberikan kinerja dan performansi yang cukup memuaskan sebagai kamus sentimen Indonesia dengan tingat akurasi sebesar 65.78% [9].

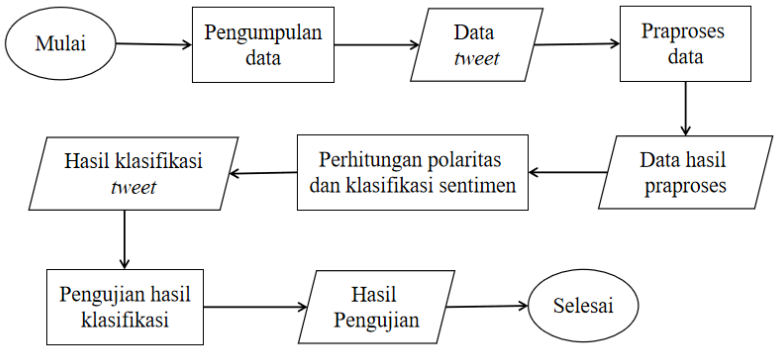
Penelitian ini dilakukan dengan pendekatan berbasis leksikal untuk memperoleh hasil klasifikasi sentimen dari data *tweet* pada topik "kuliah daring" menggunakan *InSet Lexicon* (*Indonesian Sentiment Lexicon*) yang berasal dari penelitian sebelumnya. Penelitian ini difokuskan untuk melakukan analisis sentimen/opini masyarakat khususnya para pengguna *twitter* terhadap topik "kuliah daring" melalui data yang diambil dari *twitter* dengan pengolahan data menggunakan bahasa pemrograman *Phyton*.

Beberapa penelitian terkait dengan analisis sentimen telah dilakukan sebelumnya. Diantaranya yaitu penelitian oleh Ibnu Fanhar dkk. dengan judul "Analisis Sentimen Berbasis Leksikon InSet Terhadap Partai Politik Peserta Pemilu 2019 Pada Media Sosial *Twitter*" memperoleh hasil pengujian performansi sistem dengan rata-rata *precision 40%, Recall 42%, F1 35% dan Accuracy 61%* [9]. Selain itu terdapat juga penelitian mengenai "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes" oleh Samsir dkk. menunjukkan bahwa pembelajaran daring memiliki 30% sentimen positif, 69% sentimen negatif, dan 1% netral pada periode tersebut [10]. Dari kedua penelitian tersebut, penelitian ini memiliki perbedaan pada topik yang diangkat dan juga metode yang digunakan. Penelitian ini penting dilakukan agar dapat mengetahui bagaimana opini masyarakat terhadap topik perkuliahan daring. Terlebih lagi dimasa pandemi ini perkuliahan daring menjadi topik yang sedang hangat diperbincangkan oleh berbagai kalangan. Diharapkan adanya penelitian ini dapat menjadi sumber informasi ataupun acuan untuk mengevaluasi pelaksanaan perkuliahan daring di Indonesia.

2. BAHAN DAN METODE

2.1 Prosedur Penelitian

Prosedur yang dilaksanakan dalam penelitian ini terdiri dari beberapa tahapan proses. Dimulai dari tahap pengumpulan data *tweet*, praproses data, perhitungan polaritas dan klasifikasi sentimen, hingga pengujian hasil klasifikasi. Diagram alur untuk prosedur penelitian ditampilkan pada gambar 1.



Gambar 1. Prosedur Penelitian

Adapun tiap tahapan yang ada pada prosedur penelitian akan diproses menggunakan bahasa pemrograman *Phyton* melalui IDE *Jupyter* dengan memanfaatkan sejumlah *library* diantaranya : *pandas, numpy, nltk, sastrawi, textblob, spacy, matplotlib, wordcloud, scikit-learn, dan keras.*

2.2 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini berasal dari *Twitter*. *Twitter* dipilih sebagai sumber data pada penelitian ini dikarenakan *twitter* menjadi media sosial yang cukup digemari masyarakat. *Twitter* lebih unggul sebagai penyalur info/berita tercepat yang berbeda dari media sosial lainnya, karena ketika ada topik baru yang muncul (*trending*) dan menarik perhatian maka akan ada banyak sekali opini masyarakat yang muncul di *twitter*. Opini ini bisa menjadi sumber data yang akurat untuk melakukan analisis sentimen. Pengambilan data *twitter* memanfaatkan *Twitter* API. Kata kunci yang digunakan adalah frasa yang menjadi topik dasar dalam penelitian ini yaitu "kuliah daring", dengan jumlah data awal yang diambil sebanyak 6000 *tweet*. Pembatasan jumlah data awal dilakukan agar langkah selanjutnya dapat berjalan dengan baik karena jika terlalu banyak data justru akan memperlambat praproses data. 6000 *tweet* menjadi jumlah yang cukup ideal untuk mendapat beragam opini. Proses pengumpulan data *tweet* memanfaatkan *twitter scraping tool* bernama *Twint (https://github.com/twintproject/twint)*. Data yang diperoleh mengandung beberapa atribut lengkapnya,

namun untuk melakukan proses analisis sentimen maka cukup data pada atribut *tweet* yang akan diolah lebih lanjut.

2.3 Praproses Data

Data *tweet* yang telah dikumpulkan perlu dilakukan praproses data untuk menghasilkan data yang bersih dan terstruktur sehingga mampu memberikan hasil klasifikasi sentimen yang lebih akurat. Tahapan praproses data yang dilakukan pada penelitian ini meliputi proses *cleaning, tokenizing, filtering*, dan *stemming*.

1. Proses *Cleaning*. adalah tahap pertama praproses teks yang dilakukan untuk membersihkan atau melenyapkan suatu noise pada data. Proses cleaning yang terdiri dari beberapa langkah yaitu :
   a. *Remove punctuation,* atau penghapusan tanda baca. Pada langkah ini hanya huruf alfabet yang diterima sedangkan karakter selain huruf akan dihilangkan.
   b. *Case folding,* adalah proses mengubah keseluruhan teks menjadi huruf kecil / bersifat *lowercase.*
   c. *Drop duplicates*, bertujuan untuk menghilangkan data *tweet* yang berduplikasi atau menghapus *spams tweet*.
   d. *Spelling correction,* yaitu perbaikan ejaan kata.
2. Proses *Tokenizing*. adalah tahap pemotongan string kalimat berdasarkan tiap kata yang menyusunnya. Proses ini sekumpulan karakter akan dipecah menjadi satuan kata.
3. Proses *Filtering* adalah tahap pengambilan kata-kata penting dari hasil tokenisasi. Pada tahapan ini *stopword* akan dihilangkan untuk mengurangi jumlah kata yang disimpan. *Stopword* adalah daftar kata umum yang dianggap tidak memiliki makna. Contohnya antara lain “itu”, “dan”, “yang”, ”atau”.
4. Proses *Stemming* adalah proses penghapusan imbuhan kata untuk mengubah setiap kata ke dalam bentuk dasar.

2.4 Perhitungan Polaritas dan Klasifikasi Sentimen

Pada tahap ini, setiap *tweet* yang ada akan dianalisis satu per satu. Penentuan klasifikasi sentimen untuk tiap data *tweet* dilakukan dengan metode *lexicon-based approach* menggunakan *InSet Lexicon. InSet Lexicon* mengandung sejumlah kata berbahasa Indonesia yang bersifat positif dan negatif disertai bobot dari tiap kata tersebut. Bobot kata berkisar antara -5 sampai 5, nilai minus (-) menunjukkan bahwa kata memiliki sentimen negatif sedangkan nilai plus menunjukkan bahwa kata memiliki sentimen positif.

Masing-masing kata yang terdapat pada kalimat *tweet* akan dicocokkan dengan kata pada *lexicon* untuk selanjutnya dilakukan perhitungan *polarity score* pada setiap kalimat. Proses perhitungan *polarity score* dilakukan dengan cara menjumlahkan keseluruhan bobot dari kata yang terdeteksi oleh sistem dan kemudian data tweet akan diklasifikasikan ke dalam jenis sentimen melalui algoritma yang diterapkan. Secara umum dinyatakan dengan algoritma sebagai berikut :

If sentiment score > 0	*then* Sentimen Positif	(1)
If sentiment score < 0	*then* Sentimen Netral	(2)
If sentiment score $= 0$	*then* Netral.	(3)

Klasifikasi kalimat *tweet* ke dalam sentimen positif, negatif, dan netral ditentukan berdasarkan bobot *polarity score (sentiment score)* yang diperoleh. Kalimat *tweet* tergolong sebagai kelas positif jika bobot *polarity score*-nya lebih besar dari 0, dan tergolong kelas negatif apabila *polarity score*-nya lebih kecil dari 0. Sedangkan *tweet* dengan *polarity score* sama dengan 0 akan tergolong sebagai kelas netral.

2.5. Pengujian Hasil Klasifikasi

Pada bidang klasifikasi, ukuran akurasi dari suatu model klasifikasi merupakan hal yang penting untuk diperhatikan. Nilai akurasi dapat menggambarkan bagus tidaknya suatu model klasifikasi. Dalam penelitian ini dilakukan pengujian akurasi dengan teknik *cross-validation*, dimana dataset akan dibagi menjadi 2 bagian yaitu *training set* (data latih) dan *testing set* (data uji). *Training set* digunakan untuk melatih model, sedangkan *testing set* digunakan untuk mengevaluasi performa dari model. Teknik *cross-validatio*n dengan sejumlah perulangan (*epoch*) dilakukan untuk menghindari terjadinya *overfitting* dan *overlapping* pada data uji. Data uji kemudian diproses dalam pembuatan *confusion matrix*.

Confusion Matrix adalah sebuah matriks yang memuat data klasifikasi yang dilakukan oleh sistem klasifikasi baik secara aktual maupun prediktif. Dengan mengevaluasi data pada matriks akan diketahui bagaimana performa suatu model. *Confusion matrix* untuk 2 kelas ditampilkan pada tabel berikut.

Tabel 2. Confusion Matrix 2 Kelas

		Prediksi	
		Positif	Negatif
Aktual	Positif	TP (True Positive)	FN (False Negative)
	Negatif	FP (False Positive)	TN (True Negative)

Dari pengolahan nilai-nilai yang ada pada kolom matriks (*True Negative* (TN), *False Positive* (FP), *False Negative* (FN), dan *True Positive* (TP)) maka dapat diketahui nilai *accuracy, precision*, *recall,* dan *F-measure*.

a. *Accuracy* menunjukkan kedekatan hasil klasifikasi dengan nilai sesungguhnya. Akurasi diperoleh dari perbandingan antara data yang berhasil diklasifikasikan secara benar dengan keseluruhan data.

$$accuracy = \frac{TP + TN}{TP + FP + FP + FN} \times 100\% \quad (1)$$

b. *Precision* adalah tingkat ketepatan yang menunjukkan seberapa dekat perbedaan nilai tiap kali dilakukan pengulangan. Dari nilai *precision* kita dapat mengetahui kedekatan hasil antara informasi yang diminta dengan jawaban yang sistem berikan.

$$precision = \frac{TP}{TP + FP} \times 100\% \quad (2)$$

c. *Recall* atau biasa disebut sensitifitas merupakan nilai persentase suatu model memprediksi data ke bukan kelas aktualnya.

$$recall = \frac{TP}{TP + FN} \times 100\% \quad (3)$$

d. F-*measure* adalah perhitungan yang mendapatkan informasi dengan memadukan nilai *recall* dan *precission*.

$$f - measure = \frac{2 \times recall \times precision}{recall + precision} \times 100\% \quad (4)$$

Prediksi yang benar maupun salah dari model klasifikasi yang dibuat akan terlihat pada *Confusion matrix.* Terdapat 3 kelas dalam model klasifikasi sehingga *confusion matrix* yang dihasilkan akan memiliki ordo 3x3, ditunjukkan pada tabel 3. Tabel matriks terdiri dari data aktual dan data prediksi. Dari *confussion matrix* tersebut diperoleh nilai rata-rata *accuracy, precision, recall, dan f-measure.*

Tabel 3. Confusion Matrix 3 Kelas

		Prediksi		
		Kelas A	Kelas B	Kelas C
Aktual	Kelas A	AA	AB	AC
	Kelas B	BA	BB	BC
	Kelas C	CA	CB	CC

2. HASIL DAN ANALISIS

2.1 Praproses Data

Data *tweets* yang diperoleh dari *twitter* masih berupa data mentah sehingga perlu dilakukan tahap *praproses* data untuk memperoleh data yang bersih dan terstruktur agar dapat digunakan untuk analisis sentimen.

1. Proses *Cleaning*

   Proses *cleaning* dilakukan dengan tujuan untuk membersihkan *tweet* dari karakter atau elemen yang tidak diperlukan sehingga *noise* pada proses klasifikasi akan berkurang. Adapun elemen yang dihilangkan dari kalimat *tweet* diantaranya:

   a. *Hashtag twitter* (#)
   b. *Retweet*
   c. @username (*Mentions* dengan *username twitter*)
   d. Link URL
   e. Simbol, angka, dan tanda baca lainnya

Langkah yang menjadi bagian dalam proses cleaning data diantaranya adalah *remove punctuation*, *case folding*, *spelling correction*, dan *drop duplicates*. Data awal yang diperoleh dari proses *crawling* sebelumnya berjumlah 6000 data, setelah dilakukan *drop duplicates* maka tersisa 5811 data *tweets* yang akan digunakan pada tahap selanjutnya. Proses *remove punctuation*, *case folding*, dan *spelling correction* ditampilkan pada tabel 4.

Tabel 4. Hasil Proses *Cleaning* Data

Proses	Sebelum	Sesudah
Remove Punctuation	@collegemenfess Yg paling menyebalkan adalahh pada saat lgi Kuliah daring, ditanya dosen... ketika mau menjawab,, SINYALnya hilangg /tidak mendukungg. :((*#sial*	Yg paling menyebalkan adalahh pada saat lgi Kuliah daring ditanya dosen ketika mau menjawab SINYALnya hilangg tidak mendukungg
Case Folding	Yg paling menyebalkan adalahh pada saat lgi Kuliah daring ditanya dosen ketika mau menjawab SINYALnya hilangg tidak mendukungg	yg paling menyebalkan adalahh pada saat lgi kuliah daring ditanya dosen ketika mau menjawab sinyalnya hilangg tidak mendukungg
Spelling Correction	yg paling menyebalkan adalahh pada saat lgi kuliah daring ditanya dosen ketika mau menjawab sinyalnya hilangg tidak mendukungg	yang paling menyebalkan adalah pada saat lagi kuliah daring ditanya dosen ketika mau menjawab sinyalnya hilang tidak mendukung

2. Proses *Tokenizing*
   Tokenizing dilakukan untuk memecah string dalam kalimat *tweet* menjadi satuan kata yang menyusunnya. Pada dasarnya *tokenizing* ini adalah proses pemenggalan kalimat menjadi kata. Proses *tokenizing* ditampilkan pada tabel dibawah ini.

Tabel 5. Hasil Proses Tokenizing

Sebelum	Sesudah
yang paling menyebalkan adalah pada saat lagi kuliah daring ditanya dosen ketika mau menjawab sinyalnya hilang tidak mendukung	['yang', 'paling', 'menyebalkan', 'adalah', 'pada', 'saat', 'lagi', 'kuliah', 'daring', 'ditanya', 'dosen', 'ketika', 'mau', 'menjawab', 'sinyalnya', 'hilang', 'tidak', 'mendukung']

3. Proses *Filtering*
   Pada proses *filtering* dilakukan pengambilan kata-kata penting dari hasil proses sebelumnya. *Stopwords* atau kata yang kurang memiliki makna akan dihilangkan karena tidak diperlukan untuk analisis sentimen. Proses *filtering* ditampilkan pada tabel berikut.

Tabel 6. Hasil Proses Filtering

Sebelum	Sesudah
['yang', 'paling', 'menyebalkan', 'adalah', 'pada', 'saat', 'lagi', 'kuliah', 'daring', 'ditanya', 'dosen', 'ketika', 'mau', 'menjawab', 'sinyalnya', 'hilang', 'tidak', 'mendukung']	['menyebalkan', 'kuliah', 'daring', 'ditanya', 'dosen', 'menjawab', 'sinyalnya', 'hilang', 'tidak', 'mendukung']

4. Proses *Stemming*
   Pada proses ini dilakukan untuk mengubah kata menjadi kata dasar (stem) dengan cara menghilangkan imbuhan kata berupa awalan maupun akhiran. Untuk melakukan proses stemming digunakan salah satu *library Phyton* yaitu *Sastrawi*. Proses ini memakan waktu yang cukup lama terlebih lagi apabila dataset yang diolah sangat banyak. Proses *stemming* ditunjukkan pada tabel 7.

Tabel 7. Hasil Proses Stemming

Sebelum	Sesudah
['menyebalkan', 'kuliah', 'daring', 'ditanya', 'dosen', 'menjawab', 'sinyalnya', 'hilang', 'tidak', 'mendukung']	['sebal', 'kuliah', 'daring', 'tanya', 'dosen', 'jawab', 'sinyal', 'hilang', 'tidak', 'dukung']

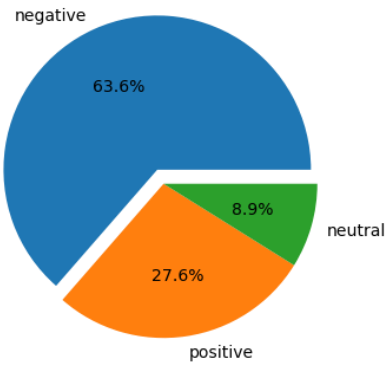
2.2 Hasil Klasifikasi

Proses analisis sentimen dilakukan dengan pendekatan *lexicon-based* menggunakan *InSet Lexicon* yang menghasilkan model klasifikasi dalam 3 kelas, yaitu positif, negatif, dan netral. Data hasil *preprocessing* yang berjumlah 5811 *tweet* kemudian diklasifikasikan secara otomatis menggunakan algoritma yang menerapkan *InSet lexicon* sebagai kamus opini. Apabila kata dalam kalimat *tweet* merupakan kata yang ada pada *opinion words* di leksikon, maka bobot dari kata tersebut diakumulasikan untuk nilai *polarity score* pada *tweet*. Penentuan kalimat *tweet* ke dalam kelas yang sesuai didasarkan pada perhitungan *polarity score*. Jumlah *polarity score* yang bernilai positif akan menjadikan data *tweet* tergolong ke dalam sentimen positif, begitu pula sebaliknya, jika *polarity score* bernilai negatif atau *minus* maka *tweet* termasuk ke dalam sentimen negatif. Kelas netral untuk *tweet* dengan *polarity score* sama dengan 0. Contoh hasil perhitungan *polarity score* dari sejumlah *tweet* ditunjukkan pada tabel 8.

Tabel 8. Hasil Perhitungan Polarity Score dan Klasifikasi

Tweet	*Tweet Preprocessed*	*Polarity Score*	*polarity*
SEMESTER DEPAN KAMPUSKU MASIH KULIAH DARING! ALHAMDULILLAH.	['semester', 'kampus', 'kuliah', 'daring', 'alhamdulillah']	4	positive
Kuliah daring di rumah benar benar melatihku menjadi ibu rumah tangga ya	['kuliah', 'daring', 'rumah', 'latih', 'rumah', 'tangga', 'ya']	10	positive
Di tengah kesibukan kuliah daring, banyak mahasiswa yang menyisihkan waktunya untuk menjadi sukarelawan mengajar siswa. #Muda #adadikompas https://t.co/VFxmFOvLnN	['sibuk', 'kuliah', 'daring', 'mahasiswa', 'sisih', 'sukarelawan', 'ajar', 'siswa']	0	neutral
@adiechh_ki Kuliah daring, suara ayam lebih nyaring.	['kuliah', 'daring', 'suara', 'ayam', 'nyaring']	0	neutral
Sa'at kuliah masih tatap muka, semua terasa baik-baik saja, namun ketika kuliah daring, semua terasa berantakan, dalam benak muncul sebuah pertanyaan, apakah orang tua akan mema'afkan anaknya ini?	['kuliah', 'tatap', 'muka', 'baik', 'kuliah', 'daring', 'berantak', 'benak', 'muncul', 'orang', 'tua', 'maaf', 'ananda']	-12	negative
Ekspektasi kuliah daring : -Belajar sambil rebahan -Uang saku tetep jalan -Belajar sambil ngedrakor Realita : -Tugas, tugas dan tugas -Otak tegang, stres, mumet -Kurus kering -Makin kere Tinggal Gila nya :)	['ekspektasi', 'kuliah', 'daring', 'ajar', 'rebah', 'uang', 'saku', 'jalan', 'ajar', 'ngedrakor', 'realitas', 'tugas', 'tugas', 'tugas', 'otak', 'tegang', 'stres', 'mumet', 'kurus', 'kering', 'kere', 'tinggal', 'gila']	-29	negative

Keseluruhan data *tweet* telah diklasifikasikan ke dalam kelasnya berdasarkan perhitungan *polarity score* dari tiap kata menggunakan *InSet lexicon*. Selanjutnya untuk mempermudah melihat hasil dari proses klasifikasi, maka penyajian data ditampilkan dalam bentuk diagram lingkaran dengan keterangan persentase tiap kelas yang diperoleh.

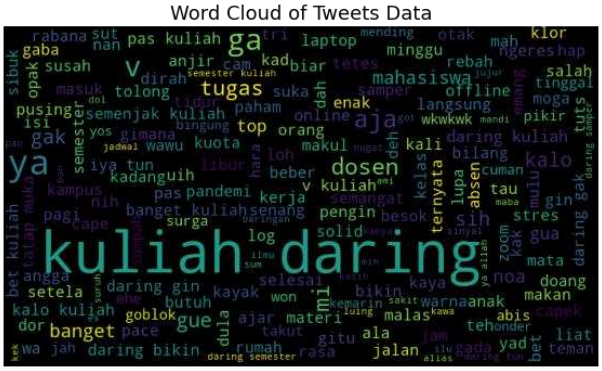


Gambar 2. Diagram Persentase Hasil Klasifikasi Sentimen

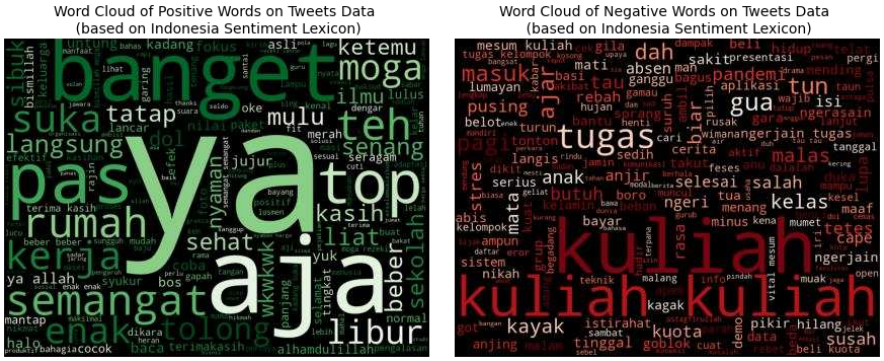
Berdasarkan diagram yang disajikan pada gambar x, dapat terlihat bahwa sentimen negatif merupakan kelas dengan persentase terbesar yaitu 63,6%, artinya sebanyak 3696 *tweet* menunjukkan opini, keluhan, atau pun pandangan negatif terhadap pelaksanaan perkuliahan daring di Indonesia. Sedangkan jumlah kelas sentimen positif memiliki persentase 27,6% atau sebanyak 1604 *tweet* memberikan opini positif baik berupa rasa senang, setuju, pendapat yang menerima, merasakan manfaat, atau pun dukungan terhadap pelaksanaan

perkuliahan daring di Indonesia. Sisanya adalah 8.9% atau sebanyak 517 *tweet* termasuk ke dalam kelas yang netral.

Visualisasi dari kata-kata yang muncul pada keseluruhan data *tweet* kemudian ditampilkan dalam bentuk *Word Cloud*. *Word cloud* mempermudah penggambaran kata dengan memberikan visualisasi data teks yang menarik dan informatif. Frekuensi kata dalam keseluruhan data mempengaruhi besar atau kecilnya kata tersebut pada *word cloud*. Semakin tinggi frekuensi kata maka ukuran kata tersebut pada *word cloud* akan semakin besar, begitu pula sebaliknya. Kata dengan frekuensi yang sedikit akan ditampilkan dengan ukuran yang lebih kecil. Visualisasi *word cloud* yang diperoleh disajikan pada gambar 3 dan 4.



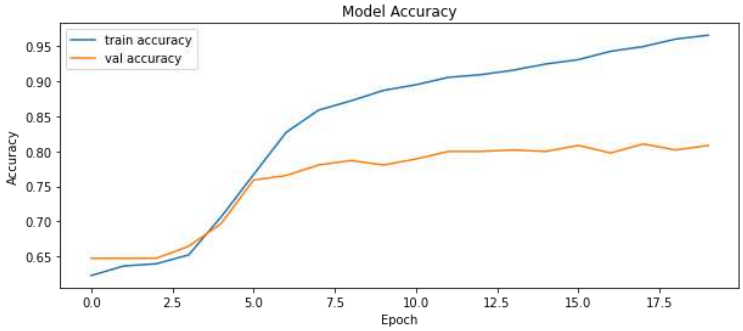
Gambar 3. *Word Cloud* Data *Tweet*



Gambar 4. *Word Cloud* dari Kata dengan Sentimen Positif dan Negatif

2.3 Pengujian dan Hasil Klasifikasi

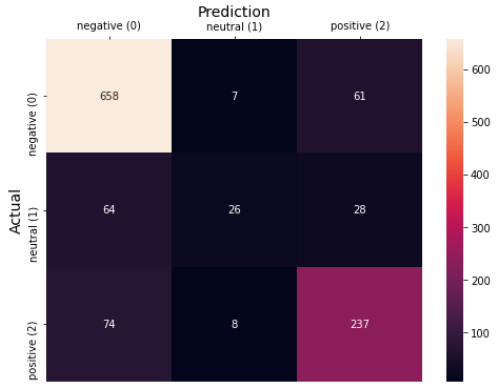
Data hasil klasifikasi dibagi menjadi 2 bagian, yaitu data latih dan data uji. Perbandingan data latih dan data uji yang digunakan pada penelitian ini adalah 8:2. 80% dari keseluruhan data atau sejumlah 4648 *tweet* menjadi data latih, sedangkan 20% sisanya atau 1163 *tweet* menjadi data uji. Hasil pengujian model klasifikasi dengan menggunakan teknik *cross-validation* ditunjukkan pada gambar 5.



Gambar 5. Hasil Pengujian Akurasi Model dengan Cross Validation

Gambar 5 merepresentasikan perbandingan antara nilai akurasi pada proses *training* maupun proses *testing* dalam 20 perulangan (*epoch*). Dari grafik terlihat bahwa nilai akurasi antara penggunaan *epoch* yang tinggi akan berpengaruh dalam menghasilkan nilai akurasi yang tinggi jika jumlah data yang digunakan dalam

proses *training* banyak. Namun terlihat juga bahwa terjadi *overfitting* di mana nilai akurasi pada proses *training* cukup tinggi, akan tetapi nilai akurasi sangat rendah pada saat proses pengujian model. Hasil pengujian 1163 data uji dengan *confusion matrix* ditampilkan pada gambar 6.



Gambar 6. Hasil Confusion Matrix dari Model Klasifikasi

Dengan melihat *confusion matrix* tersebut, maka dapat dihitung nilai *accuracy, precision, recall,* dan *f-measure* dari model klasifikasi.

1. *Accuracy* menunjukkan seberapa akurat model mengklasifikasikan data dengan benar. Maka, untuk mengetahui nilai *accuracy* di hitunglah rasio data yang diprediksi benar dengan keseluruhan data.

$$accuracy = \frac{658 + 26 + 237}{1163} \times 100\% = 79.2\%$$

Berdasarkan perhitungan, diketahui bahwa hasil klasifikasi sentimen menggunakan *InSet Lexicon* pada penelitian ini memberikan *accuracy* secara keseluruhan sebesar 79,2%.

2. *Precision* menggambarkan ketepatan hasil prediksi yang diberikan model. Nilai *precision* diperoleh dari perbandingan data yang diklasifikasi benar dengan jumlah data yang diprediksi pada kelas itu. Perhitungan nilai *precision* pada model klasifikasi ini adalah sebagai berikut.

$$P(positif) = \frac{658}{658 + 64 + 74} = 0.826$$
$$P(netral) = \frac{26}{7 + 26 + 8} = 0.634$$
$$P(negatif) = \frac{237}{61 + 28 + 237} = 0.727$$
$$precision = \frac{P(positif) + P(netral) + P(negatif)}{total\_kelas} \times 100\% = \frac{0.826 + 0.634 + 0.727}{3} \times 100\% = 72.9\%$$

Berdasarkan perhitungan, diperoleh nilai rata-rata *precision* dari keseluruhan nilai *precision* tiap kelas sentimen adalah sebesar 72,9%.

3. *Recall* menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi. *Recall* didapat dengan menghitung perbandingan data yang mampu terklasifikasi dengan benar dari keseluruhan data yang seharusnya masuk ke kelas itu. Perhitungan *recall* adalah sebagai berikut.

$$R(positif) = \frac{658}{658 + 7 + 61} = 0.906$$
$$R(netral) = \frac{26}{64 + 26 + 28} = 0.237$$
$$R(negatif) = \frac{237}{74 + 8 + 237} = 0.742$$
$$recall = \frac{R(positif) + R(netral) + R(negatif)}{total\_kelas} \times 100\% = \frac{0.906 + 0.237 + 0.742}{3} \times 100\% = 62.8\%$$

Berdasarkan perhitungan, diperoleh nilai rata-rata *recall* dari keseluruhan nilai *recall* tiap kelas sentimen adalah sebesar 62,8%.

4. *F-measure* didapat dengan menggabungkan nilai *precision* dan *recall.* Perhitungannya dengan mencari perbandingan 2 kali *recall* dan *precision* dengan hasil penjumlahan *recall* dan *precision*.

$$f-measure = \frac{2 \times recall \times precision}{recall + precision} \times 100\%$$
$$= \frac{2 \times 0.628 \times 0.729}{0.628 + 0.729} \times 100\%$$
$$= \frac{0.915}{1.357} \times 100\% = 67.4\%$$

Berdasarkan perhitungan, diperoleh nilai rata-rata *F-measure* adalah sebesar 67,4%.

4. KESIMPULAN

Berdasarkan hasil pengujian dan analisis yang telah dilakukan terhadap data *tweet* dengan topik perkuliahan daring untuk menentukan klasifikasi sentimen, maka diperoleh kesimpulan bahwa *InSet Lexicon* dapat digunakan dalam proses analisis sentimen melalui perhitungan *polarity score* pada data tekstual yang menghasilkan klasifikasi sentimen dalam 3 kelas, yaitu positif, negatif, dan netral. Hasil menunjukkan bahwa tingkat sentimen negatif memiliki persentase paling tinggi yaitu sebesar 63.6%. Sedangkan persentase sentimen positif sebesar 27.6% dan netral sebesar 8.9%. Tingkat akurasi yang diperolah adalah 79.2%, *precision* sebesar 72.9%, *recall* sebesar 62.8% dan *f-measure* sebesar 67.4% dengan komposisi data latih 80% dan data uji 20%. Dari hasil tersebut diketahui bahwa sentimen negatif merupakan kelas yang paling dominan. Hal ini menandakan cukup banyak opini masyarakat yang mengutarakan ketidakpuasan, keluhan atau pun pandangan negatif terhadap pelaksanaan perkuliahan daring di Indonesia. Oleh karenanya, perkuliahan daring dirasa belum maksimal diterapkan di Indonesia sehingga perlu adanya upaya untuk meningkatkan kualitas maupun efektivitas perkuliahan daring agar tingkat kekecewaan publik dapat diminimalisir. Upaya tersebut bisa dengan menghadirkan pembelajaran yang interaktif, mengatasi kendala jaringan untuk akses informasi, atau pun mengevaluasi kembali capaian pembelajaran yang harusnya diperoleh, dan banyak lagi upaya lainnya. Penulis menadari bahwa masih terdapat banyak kekurangan pada penelitian ini, oleh karena itu sebagai saran dan masukan bagi penelitian selanjutnya, ada beberapa poin yang menjadi catatan untuk perbaikan yaitu dengan upaya meningkatkan kualitas hasil dari tahapan *preprocessing* data, memperkaya kamus kata, mencari formula yang dapat menangani singkatan dari bahasa tidak formal, menghilangkan *emoticon* atau memanfaatkan *emoticon* sebagai salah satu indikator kecenderungan sentimen, serta menyingkirkan *tweet* yang bukan kalimat opini seperti iklan dan berita.

REFERENSI

[1] SE Mendikbud, “Surat Edaran Kementerian Pendidikan dan Kebudayaan RI Tanggal 17 Maret 2020 Perihal Pembelajaran secara Daring dan Bekerja dari Rumah dalam Rangka Pencegahan Penyebaran *Coronavirus Disease* (COVID-19),” Jakarta, No. 36962/MPK.A/HK/2020.

[2] G.A. Buntoro, Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter, *INTEGER: Journal of Information Technology,* Vol. 2, Ed. 1, 2017.

[3] Polling Indonesia, "Laporan Survei Penetrasi & Profil Perilaku Pengguna Internet Indonesia," Asosiasi Penyelenggara Jasa Internet Indonesia, Jakarta, 2018.

[4] M. R. Ma’arif, “Analisis Konten Interaksi Pengguna Twitter pada Masa 100 Hari Pertama Pemerintahan Baru DKI Jakarta Menggunakan Text Mining,” *Jurnal Pekomnas*, Vol. 3 No. 2, 137-142, 2018.

[5] S. H. P. E. A. D. Imam Fahrur Rozi, "Implementasi Opinion Mining (Analisis Sentimen) untuk Ekstraksi Data Opini Publik pada Perguruan Tingg," *Jurnal EECCIS* , vol. 6, p. 37, 2012.

[6] Akbari, M. I. H. A. D., Astri Novianty S.T., M. & Casi Setianingsih S.T., M., 2012. Analisis Sentimen Menggunakan Metode Learning Vector Quantization. Telkom University.

[7] Y. Azhar, “Metode Lexicon Learning Based untuk Identifikasi Tweet Opini Berbahasa Indonesia, ” *Jurnal Nasional Pendidikan Teknik Informatika*, Vol. 6, No. 3, 2017.

[8] I. F. Nur, A. Herdiani, dan W. Astuti, Analisis Sentimen Berbasis Leksikon InSet Terhadap Partai Politik Peserta Pemilu 2019 pada Media Sosial Twitter, ” *e-Proceeding of Engineering*, Vol. 6, No. 3, 2019.

[9] F. Koto dan G. Y. Rahmaningtyas, “InSet Lexicon: Evaluation of a Word List for Indonesian Sentiment Analysis in Microblogs, ” *International Conference on Asian Language Processing (IALP),* 2017.

[10] Samsir dkk. “Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes, ” *Jurnal Media Informatika Budidarma*, Vol. 5, No. 1, 2021.

[11] B. K. Santra dan C.J. Christy, Genetic Algorithm and Confusion Matrix for Document Clustering, ” *International Journal of Computer Science Issues*, Vol. 9, Issue 1, No 2, 2012.