

**SKRIPSI**

**ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP  
KEBIJAKAN MERDEKA BELAJAR PADA MEDIA SOSIAL  
TWITTER**

**Disusun dan diajukan oleh:**

**SUCI AMALIAH MUZAKKIR  
D121171509**



**PROGRAM STUDI SARJANA TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS HASANUDDIN  
GOWA  
2023**

**LEMBAR PENGESAHAN SKRIPSI**

**ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP  
KEBIJAKAN MERDEKA BELAJAR PADA MEDIA SOSIAL  
TWITTER**

Disusun dan diajukan oleh

**SUCI AMALIYAH MUZAKKIR**

**D121171509**

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Studi Program Sarjana Program Studi Teknik Informatika Fakultas Teknik Universitas Hasanuddin pada tanggal 21 Juni 2023 dan dinyatakan telah memenuhi syarat kelulusan.

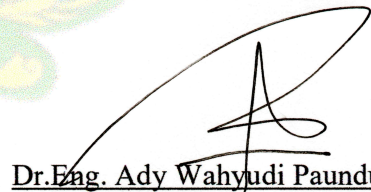
Menyetujui,

Pembimbing Utama,

Pembimbing Pendamping,



Dr. Ir. Zahir Zainuddin, M.Sc.  
Nip. 196404271989101002



Dr. Eng. Ady Wahyudi Paundu, ST., MT.  
Nip. 197503132009121003

Ketua Program Studi,



Prof. Dr. Ir. Indrabayu, ST., MT., M.Bus.Sys., IPM, ASEAN. Eng.  
Nip. 19750716 200212 1 004

## PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini ;

Nama : Suci Amaliyah Muzakkir  
NIM : D121171509  
Program Studi : Teknik Informatika  
Jenjang : S1

Menyatakan dengan ini bahwa karya tulisan saya berjudul

Analisis Sentimen Masyarakat Indonesia Terhadap Kebijakan Merdeka Belajar  
Pada Media Sosial Twitter

Adalah karya tulisan saya sendiri dan bukan merupakan pengambilan alihan tulisan orang lain dan bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri.

Semua informasi yang ditulis dalam skripsi yang berasal dari penulis lain telah diberi penghargaan, yakni dengan mengutip sumber dan tahun penerbitannya. Oleh karena itu semua tulisan dalam skripsi ini sepenuhnya menjadi tanggung jawab penulis. Apabila ada pihak manapun yang merasa ada kesamaan judul dan atau hasil temuan dalam skripsi ini, maka penulis siap untuk diklarifikasi dan mempertanggungjawabkan segala resiko.

Segala data dan informasi yang diperoleh selama proses pembuatan skripsi, yang akan dipublikasi oleh Penulis di masa depan harus mendapat persetujuan dari Dosen Pembimbing.

Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan isi skripsi ini hasil karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Gowa, 23 Juni 2023  
Yang Menyatakan



Suci Amaliyah Muzakkir



## ABSTRAK

**SUCI AMALIYAH MUZAKKIR.** *Analisis Sentimen Masyarakat Indonesia Terhadap Kebijakan Merdeka Belajar Di Media Sosial Twitter* (dibimbing oleh Dr. Ir Zahir Zainuddin, M.Sc dan Dr.Eng. Ady Wahyudi Paundu, S.T., M.T)

Kementrian Pendidikan dan kebudayaan (kemendikbud) mengumumkan kebijakan merdeka belajar yang telah dikeluarkan pada awal tahun 2020. Twitter merupakan salah satu media sosial untuk mengetahui dan memberikan informasi secara luas dan cepat. Sebagian besar informasi yang tersebar dalam media sosial dapat berupa pendapat, opini, masukan, dan juga saran setiap perorangan atau individu terhadap sebuah produk atau objek tertentu yang disebut sentimen. Kebijakan Merdeka Belajar diberlakukan karena banyaknya keluhan dari orangtua pada sistem Pendidikan nasional yang sudah berlaku selama ini, yang menimbulkan pro dan kontra dari berbagai pihak. Dari fenomena tersebut sentiment analisis dapat dilakukan untuk mendapatkan nilai dan value dari objek yang akan dianalisis. Dalam penerapan analisis sentimen membutuhkan algoritma yang dapat melakukan sebuah klasifikasi pendapat atau sentimen masyarakat. Dalam penelitian ini akan dirancang sebuah model analisis sentiment dengan menggunakan metode klasifikasi Convolutional Neural Network (CNN). Keluaran dari model analisis sentiment ini ada 3, yaitu positif, negatif, dan netral. Penelitian ini menerapkan hyperparameter tuning untuk mendapatkan hasil terbaik. Hyperparameter tuning bertujuan untuk memilih satu set hyperparameter yang optimal untuk algoritma pembelajaran. Hasil yang didapatkan pada penelitian ini menunjukkan hasil yang baik yaitu mendapatkan nilai akurasi sebesar 82,54%

Kata Kunci: Sentimen Analisis, Merdeka Belajar, Convolutional Neural Network.



## ABSTRACT

**SUCI AMALIYAH MUZAKKIR.** *Analysis of Indonesian People's Sentiment Against the Freedom to Study Policy on Twitter Social Media* (supervised Dr. Ir Zahir Zainuddin, M.Sc and Dr.Eng. Ady Wahyudi Paundu, S.T., M.T)

The Ministry of Education and Culture (Kemendikbud) announced an independent learning policy that was issued in early 2020. Twitter is one of the social media to find out and provide information widely and quickly. Most of the information spread on social media can be in the form of opinions, opinions, entries, and also suggestions for each individual or individual for a particular product or object, which is called sentiment. The Freedom to Learn Policy was enforced because of the many complaints from parents about the national education system that has been in effect so far, which has raised pros and cons from various parties. From this phenomenon, sentiment analysis can be carried out to obtain the value of the object to be analyzed. In the application of sentiment analysis requires an algorithm that can carry out a classification of public opinion or sentiment. In this research a sentiment analysis model will be designed using the Convolutional Neural Network (CNN) classification method. There are 3 outputs from this sentiment analysis model, namely positive, negative, and neutral. This study applies hyperparameter tuning to get the best results. Hyperparameter tuning aims to select an optimal set of hyperparameters for the learning algorithm. The results obtained in this study showed good results, namely obtaining an accuracy value of 82.54%.

Keywords: Sentiment Analysis, Merdeka Belajar, Convolutional Neural Network.

## DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI.....	i
PERNYATAAN KEASLIAN.....	ii
ABSTRAK .....	iii
ABSTRACT .....	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR .....	vii
DAFTAR TABEL.....	viii
DAFTAR LAMPIRAN .....	ix
DAFTAR SINGKATAN DAN ARTI SIMBOL .....	x
KATA PENGANTAR .....	xi
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	2
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Ruang Lingkup .....	3
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kebijakan Merdeka Belajar .....	5
2.2 Analisis Sentimen.....	6
2.3 Media sosial.....	6
2.4 <i>Data Mining</i> .....	7
2.5 <i>Natural Language Processing (NLP)</i> .....	7
2.6 Metode Klasifikasi.....	7
2.7 <i>Text mining</i> .....	8
2.8 CNN ( <i>Convolutional Neural Network</i> ) .....	8
2.9 <i>Web Scrapping</i> .....	13
2.6 <i>Hyperparameter Tuning</i> .....	13
2.7 <i>Confusion Matrix</i> .....	14
BAB III METODE PENELITIAN.....	18
3.1 Tahapan Penelitian .....	18
3.2 Waktu dan Lokasi Penelitian.....	19
3.3 Instrumen Penelitian .....	19
3.4 Teknik Pengambilan Data .....	20
3.5 Rancangan Sistem .....	21
3.5.1 <i>Preprocessing Data</i> .....	22
3.5.2 <i>Split Data</i> .....	23
3.5.3 <i>Training Data</i> .....	23
3.6 Skenario Uji Performansi .....	29
BAB IV HASIL DAN PEMBAHASAN .....	30
4.1 Pengumpulan Data.....	30
4.2 <i>Preprocessing Data</i> .....	30
4.3 Skenario Pengujian.....	34
4.3.1 <i>Hyperparameter yang menentukan struktur jaringan</i> .....	34

4.3.2	Hyperparameter yang menentukan jaringan yang dilatih (untuk proses training)	34
4.4	Pengujian Hyperparameter Struktur Jaringan	37
4.4.1	Hyperparameter yang menentukan struktur jaringan	37
4.4	Pengujian Hyperparameter Training Data	37
4.5.1	Pengujian Partisi Data	37
4.5.2	Pengujian Nilai <i>Learning Rate</i>	41
4.5.3	Pengujian Nilai <i>Epoch</i>	43
4.5.4	Pengujian Nilai Batch Size	45
4.5.5	Model CNN Terbaik	46
4.5.6	Hasil Pengujian Model Terbaik	47
BAB V KESIMPULAN DAN SARAN		49
5.1	Kesimpulan	49
5.2	Saran	49
DAFTAR PUSTAKA		50
LAMPIRAN		52
	Lampiran 1	52
	Lampiran 2	53
	Lampiran 3	54
	Lampiran 4	60
	Lampiran 5	65



## DAFTAR GAMBAR

Gambar 1 Visualisasi convolution pada matriks.....	9
Gambar 2 Contoh Arsitektur Convolutional Neural Network ID.....	10
Gambar 3 Contoh Input Layer .....	11
Gambar 4 Tahapan penelitian .....	18
Gambar 5 Gambaran umum sistem yang dibangun .....	21
Gambar 6 Contoh dari proses word index.....	23
Gambar 7 Word index ke input layer.....	24
Gambar 8 Proses convolution .....	25
Gambar 9 <i>Output layer</i> 1 untuk <i>activation function</i> .....	26
Gambar 10 <i>Output layer</i> 2 untuk <i>activation function</i> .....	26
Gambar 11 Proses activation function dan max-pooling.....	27
Gambar 12 Hasil <i>scrapping</i> data menggunakan <i>snsrape</i> .....	30
Gambar 13 Grafik pengujian partisi data.....	41
Gambar 14 Grafik pengujian <i>learning rate</i> .....	42
Gambar 15 Grafik pengujian epoch.....	44
Gambar 16 Grafik pengujian <i>batch size</i> .....	46
Gambar 17 Model CNN yang digunakan.....	46
Gambar 18 Grafik Pengujian model .....	47

## DAFTAR TABEL

Tabel 1 Nilai Confusion Matrix .....	15
Tabel 2 Data confusion matrix .....	29
Tabel 3 Hasil perhitungan confusion matrix .....	29
Tabel 4 Proses Cleaning .....	31
Tabel 5 Proses Tokenizing .....	31
Tabel 6 Proses Normalisasi Teks .....	32
Tabel 7 Proses Stop Words .....	32
Tabel 8 Proses Stemming .....	33
Tabel 9 Rincian dataset setelah preprocessing .....	33
Tabel 10 Pengujian partisi data .....	34
Tabel 11 Pengujian <i>learning rate</i> .....	35
Tabel 12 Pengujian <i>epoch</i> .....	35
Tabel 13 Pengujian <i>batch size</i> .....	36
Tabel 14 Hyperparameter Struktur Jaringan .....	37
Tabel 15 Analisis partisi data 50:50 .....	38
Tabel 16 Rangkuman pengujian partisi data .....	40
Tabel 17 Rangkuman pengujian <i>learning rate</i> .....	42
Tabel 18 Rangkuman pengujian <i>epoch</i> .....	44
Tabel 19 Rangkuman pengujian <i>batch size</i> .....	45
Tabel 20 Confusion matrix CNN .....	47
Tabel 21 Hasil Classification Report .....	48

## DAFTAR LAMPIRAN

Lampiran 1 Hasil Preprocessing .....	52
Lampiran 2 Form google proses pelabelan data .....	53
Lampiran 3 Kamus Singkatan.....	54
Lampiran 4 Stopword.....	60
Lampiran 5 Hasil skenario pengujian .....	65



## DAFTAR SINGKATAN DAN ARTI SIMBOL

Lambang/Singkatan	Arti dan Keterangan
CNN	<i>Convolutional Neural Network</i>
Sosmed	Sosial Media
Kemendikbud	Kementrian Pendidikan dan kebudayaan
mbkm	Program Merdeka Belajar – Kampus Merdeka
USBN	Ujian Sekolah Berstandar Nasional
RPP	Rencana Pelaksanaan Pembelajaran
PPDB	Penerimaan Peserta Didik Baru
PTN	Perguruan Tinggi Negeri
PTS	Perguruan Tinggi Swasta
UN	Ujian Nasional
NLP	<i>Natural Language Processing</i>
TNet	<i>True Netral</i>
TPos	<i>True Positive</i>
TNeg	<i>True Negative</i>
FPosNet	<i>False Positive Neutral</i>
FNegNet	<i>False Negative Neutral</i>
FNetPos	<i>False Neutral Positive</i>
FNegPos	<i>False Negative Positive</i>
FNetNeg	<i>False Neutral Negative</i>
FPosNeg	<i>False Positive Negative</i>
API	Application Programming Interface
Csv	<i>comma separated value</i>
RT	<i>Repost Tweet</i>
NLTK	<i>Natural Language Tool Kit</i>
KBBI	Kamus Besar Bahasa Indonesia

## KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT atas Rahmat dan ridho-Nya, sehingga dapat menyelesaikan tugas akhir skripsi yang berjudul **“ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP KEBIJAKAN MERDEKA BELAJAR DI MEDIA SOSIAL TWITTER”** sebagai salah satu syarat dalam menyelesaikan jenjang Strata-1 di Departemen Teknik Informatika, Fakultas Teknik, Universitas Hasanuddin.

Penulis menyadari banyak kesulitan dan kendala yang dihadapi saat penyusunan tugas akhir ini. Dalam prosesnya, penulis memperoleh banyak bantuan, dukungan, dan bimbingan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih kepada:

1. Allah SWT melalui berkat dan rahmat-Nya sehingga penulis dapat menyelesaikan tugas akhir ini.
2. Kedua orang tua penulis, Bapak Muzakkir Hafid, Mama Hasiah Mangka penulis yang selalu menyertai penulis dalam doanya serta mendukung, membantu, memberi semangat dalam perjalanan penulis menyelesaikan tugas akhir ini.
3. Bapak Dr. Ir Zahir Zainuddin, M.Sc selaku pembimbing I dan Bapak Dr.Eng. Ady Wahyudi Paundu, S.T., M.T selaku pembimbing II, yang senantiasa menyediakan waktu, tenaga, pikiran, dan perhatian yang luar biasa dalam mengarahkan penulis untuk menyelesaikan tugas akhir.
4. Segenap Dosen dan Staff Departemen Teknik Informatika Fakultas Teknik Universitas Hasanuddin yang telah banyak membantu penulis selama masa perkuliahan.
5. Kakak dan adik penulis, Hasim, Nia, Mutia, Kiya, Sari, dan Adly yang selalu memberikan dukungan dan semangat untuk menyelesaikan tugas akhir ini.
6. Muh Nur yang selalu ada dan menyemangati penulis untuk menyelesaikan tugas akhir ini. Six girls, Jum, Ilmi, Priska, Rini, dan Rieka. Sahabat yang selalu menemani dikala suka dan duka. Selalu membantu penulis mengerjakan tugas akhir ini dan selalu menemani dari awal masuk perkuliahan hingga saat ini.
7. Teman-teman Reconn17er selaku rekan yang telah memberi bantuan, dukungan dan penyusunan tugas akhir ini.
8. Serta berbagai pihak atas segala dukungan dan bantuannya yang tidak dapat penulis tuliskan satu persatu.

Penulis berharap semoga Allah SWT berkenan membalas segala kebaikan yang telah diterima oleh penulis dari berbagai pihak yang telah membantu mempermudah penulis dalam mengerjakan tugas akhir ini. Penulis menyadari bahwa masih banyak kekurangan dalam penulisan tugas akhir ini.

Oleh karena itu, penulis menerima segala bentuk masukan, kritik, dan saran untuk kesempurnaan tugas akhir ini. Penulis berharap semoga tugas akhir ini dapat bermanfaat bagi semua pihak yang membacanya.

Gowa, 15 Juni 2023

Penulis,  
Suci Amaliyah Muzakkir



# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Media sosial sebagai sarana informasi dan komunikasi mulai berkembang pesat karena internet mudah diakses dan digunakan untuk mengungkapkan pikiran, minat, dan pendapat orang tentang berbagai topik. Salah satu media sosial yang sering digunakan masyarakat Indonesia adalah Twitter. Indonesia merupakan negara kelima terbesar dalam hal pengguna Twitter, setelah Inggris dan negara besar lainnya (Kominfo, 2021). Media Sosial Publik Twitter sering digunakan untuk mengungkapkan pemikiran tentang sesuatu yang unik yang perlu dibicarakan oleh pengguna lain. Di situs media sosial Twitter, masyarakat Indonesia mengungkapkan pendapat yang beragam tentang kebijakan merdeka belajar.

Kementrian Pendidikan dan kebudayaan (kemendikbud) merupakan salah satu instansi pemerintah yang menggunakan jejaring sosial media Twitter untuk menyebarluaskan informasi dan menerbitkan pedoman tersebut. seperti halnya dalam mengumumkan kebijakan yang telah dikeluarkan. Di awal tahun 2020, kementerian Pendidikan dan Kebudayaan mengeluarkan Kebijakan merdeka belajar, kebijakan ini merupakan salah satu program inisiatif dari Menteri Pendidikan Nadiem Makarim yang bertujuan untuk memerdekakan guru dan siswa. Menurut Nadiem, "Merdeka Belajar" itu perlu di zaman sekarang ini, anak-anak tidak lagi harus mengikuti kurikulum yang tersedia, namun bisa menggunakan metode pembelajaran yang paling sesuai selama ini (Kamil & Irfan, 2020). Kebijakan Merdeka Belajar ini diberlakukan karena banyaknya keluhan dari orangtua pada sistem Pendidikan nasional yang sudah berlaku selama ini. Adapun empat pokok dari kebijakan merdeka belajar ini yang pertama yaitu, perubahan bentuk dan mutu Ujian Sekolah Berstandar Nasional (USBN). kedua, perubahan dalam pembuatan dan perancangan Rencana Pelaksanaan Pembelajaran (RPP). Ketiga, penghapusan Ujian Nasional. Dan yang terakhir, penyempurnaan pada sistem peraturan Penerimaan Peserta Didik Baru (PPDB) Zonasi. Adapun

Kebijakan Kampus Merdeka yang merupakan kelanjutan dari konsep Merdeka Belajar. Terdapat empat penyesuaian kebijakan di lingkup pendidikan tinggi. Kebijakan pertama yaitu, otonomi bagi Perguruan Tinggi Negeri (PTN) dan Swasta (PTS). Kedua, adalah program reakreditasi yang bersifat otomatis untuk seluruh peringkat dan bersifat sukarela bagi perguruan tinggi dan prodi yang sudah siap naik peringkat. Ketiga, terkait kebebasan bagi PTN Badan Layanan Umum (BLU) dan Satuan Kerja (Satker) untuk menjadi PTN Badan Hukum (PTN BH). Keempat, memberikan hak kepada mahasiswa untuk mengambil mata kuliah di luar prodi dan melakukan perubahan definisi Satuan Kredit Semester (sks). Setelah kebijakan ini diberlakukan selama kurang lebih satu tahun, tentunya menimbulkan opini yang sangat beragam mulai dari opini positif yang mendukung kebijakan merdeka belajar ini maupun opini negatif yang menentang kebijakan ini.

Berdasarkan pokok bahasan yang menjadi dasar dalam penelitian ini yaitu mengetahui sentimen masyarakat terhadap penerapan kebijakan Merdeka belajar. Analisis sentimen merupakan bidang penelitian berkelanjutan yang berada di antara berbagai bidang seperti Data Mining, Natural Language Processing (NLP) yang umumnya digunakan untuk menganalisis data tekstual dalam bentuk opini yang mengandung polaritas sehingga nantinya menghasilkan sebuah informasi yang memiliki nilai positif, negatif, atau netral. Dalam penelitian ini dibutuhkan algoritma agar dapat bekerja dengan baik dalam mengklasifikasi kalimat. Convolutional Neural Network (CNN) merupakan salah satu algoritma yang dapat diterapkan untuk melakukan analisis sentimen, yang saat ini memiliki performa cukup bagus untuk melakukan klasifikasi kalimat. CNN pada umumnya diterapkan untuk klasifikasi gambar, namun dengan mengadopsi metode yang sama dan penyesuaian singkat CNN dapat diterapkan untuk analisis sentimen dan mendapatkan nilai akurasi jauh lebih tinggi dibandingkan dengan algoritma lain.

## **1.2. Rumusan Masalah**

Bagaimana cara mengklasifikasikan setiap opini terhadap kebijakan merdeka belajar dari twitter menjadi beberapa kategori sentimen yaitu positif, netral, dan negatif menggunakan algoritma Convolutional Neural Network?

### **1.3. Tujuan Penelitian**

Mengklasifikasikan opini terhadap kebijakan merdeka belajar twitter menjadi beberapa kategori sentimen yaitu positif, netral, dan negatif menggunakan Convolutional Neural Network.

### **1.4 Manfaat Penelitian**

Dengan dilakukannya penelitian ini, diharapkan manfaat yang didapatkan antara lain:

1. Membantu pemerintah mengevaluasi kebijakan Merdeka Belajar.
2. Memberikan informasi polaritas sentimen pengguna media sosial twitter terhadap kebijakan Merdeka Belajar
3. Menjadi referensi bagi peneliti mendatang apabila akan melakukan penelitian dengan tema terkait

### **1.5. Ruang Lingkup**

1. Kumpulan data yang bersumber dari media sosial Twitter yang menggunakan Bahasa Indonesia.
2. Penelitian ini menggunakan data dengan topik kebijakan Merdeka Belajar yang berasal dari media sosial twitter.
3. Dataset yang digunakan telah divalidasi oleh ahli Bahasa



## 1.6 Sistematika Penulisan

Pada bagian ini memberikan gambaran singkat mengenai isi tulisan secara keseluruhan pada tugas akhir ini, maka akan diuraikan beberapa tahapan dari penulisan secara sistematis, yaitu:

### **BAB 1 PENDAHULUAN**

Pada bab ini diuraikan mengenai latar belakang masalah dari judul penelitian Analisis Sentimen Menggunakan Algoritma Convolutional Neural Network, disertai dengan rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup dan sistematika penulisan.

### **BAB II TINJAUAN PUSTAKA**

Dalam bab ini berisi teori-teori yang berhubungan dengan proses penelitian yang dilakukan untuk menganalisis masalah yang akan diteliti serta hal-hal lain yang berhubungan dengan variabel-variabel data yang akan digunakan, *Preprocessing* dan *Convolutional Neural Network*

### **BAB III METODOLOGI PENELITIAN**

Pada bab ini berisi tentang apa saja yang akan dilakukan pada saat penelitian, yang meliputi tahapan penelitian, waktu dan lokasi, instrumen penelitian, perancangan sistem dan analisis kerja sistem

### **BAB IV HASIL DAN PEMBAHASAN**

Dalam bab ini berisi tentang hasil pengolahan data serta pembahasan yang disertai tabel hasil penelitian.

### **BAB V PENUTUP**

Dalam bab ini berisi tentang kesimpulan yang didapatkan berdasarkan hasil penelitian yang dilakukan, serta memberikan saran-saran untuk penelitian/pengembangan selanjutnya.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Kebijakan Merdeka Belajar**

Kebijakan publik adalah serangkaian keputusan yang saling terkait (termasuk keputusan-keputusan yang tidak bertindak) yang dibuat oleh Lembaga pemerintah. Kebijakan pada umumnya digunakan untuk menyeleksi dan menunjukkan pilihan terpenting untuk mempererat kehidupan, baik dalam kehidupan organisasi pemerintahan maupun privat. Kebijakan tidak boleh memiliki konotasi atau nuansa yang mengandung kata politis, yang diduga mengandung bias kepentingan.

Dijaminnya kebebasan masyarakat untuk menyampaikan aspirasi dan berpartisipasi dalam penyusunan seperti kebijakan publik di daerah, agar kebijakan publik memenuhi rasa keadilan dan tidak menimbulkan kontroversi di masyarakat. Oleh karena itu, perumusan kebijakan publik dimulai dari dan oleh rakyat, serta untuk rakyat, terutama di sebuah negara demokrasi (Anggara, 2014).

Kebijakan merdeka belajar yang diumumkan oleh Kemendikbud memiliki empat pokok kebijakan. Pokok kebijakan yang pertama menyangkut Ujian Sekolah Berbasis Nasional (USBN) yang akan berubah. Kedepannya ujian akan diselenggarakan oleh sekolah sehingga guru dan sekolah lebih merdeka dalam penilaian hasil belajar siswa. Pokok kedua bahwa pelaksanaan Ujian Nasional (UN) akan diubah menjadi Asesmen Kompetensi Minimum dan Survei Karakter. Pokok ketiga adalah menyederhanakan penyusunan Rencana Pelaksanaan Pembelajaran (RPP) dan pokok yang keempat adalah kebijakan zonasi tetap digunakan dalam penerimaan peserta didik baru (PPDB) karena untuk mengakomodasi ketimpangan akses dan kualitas di berbagai daerah.

## 2.2 Analisis Sentimen

Analisis sentimen adalah proses menganalisis teks dari sumber informasi dengan tujuan mengumpulkan opini publik. Data yang dikumpulkan adalah pendapat tentang produk, layanan, politik, dan lain lain. Sentimen analisis yang biasa disebut sebagai penambangan opini, adalah bidang studi yang menganalisis opini, sentimen, evaluasi, penilaian sikap, dan emosi individu terhadap produk, layanan, peristiwa, dan atributnya (B. Liu, 2012).

*Sentiment Analysis* juga disebut sebagai *opinion mining*. *Sentiment analysis* dapat memberikan gambaran apakah opini publik terhadap suatu aspek bisnis. Kelemahan dan kekuatan terhadap suatu produk untuk dapat melakukan perbaikan produk, atau untuk melakukan analisis terhadap pembuatan *new marketing campaign*.

Dengan kata lain, *Sentiment Analysis* merupakan garis penelitian yang memanfaatkan opini dan sikap orang-orang dalam kaitannya dengan berbagai topik. Ini adalah perpanjangan data mining di domain NLP (*Natural Language Processing*). Konsep ini melibatkan kemampuan untuk dapat mempelajari setiap kata atau frasa apapun dalam teks dan memberi label sebagai positif, negatif atau netral. Analisis sentiment meliputi beberapa langkah seperti mendefinisikan dataset, preprocessing, featured selection, pelabelan, klasifikasi dan evaluasi (Suryani dkk, 2019).

## 2.3 Media sosial

Media sosial adalah sebuah website atau aplikasi berbasis teknologi mobile yang telah berkembang menjadi mekanisme yang efektif yang dapat berfungsi sebagai strategi pemasaran perusahaan; khususnya dalam interaksi perusahaan dengan pelanggan, seperti memfasilitasi pencarian informasi, interaktivitas, mempromosikan dan meningkatkan perilaku pembelian pelanggan (Rustian, 2012). Media sosial adalah teknologi berbasis website atau aplikasi yang dapat meningkatkan komunikasi dua arah antara perusahaan dan pelanggan. Selain itu media sosial digunakan sebagai sarana promosi digital suatu produk perusahaan,

dikarenakan media sosial bisa tersaji konten yang dapat dipublikasikan secara tekstual, visual dan verbal (Rustian, 2012).

## **2.4 Data Mining**

*Data mining*, atau penambangan data adalah disiplin ilmu yang menggabungkan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk memecahkan masalah pengambilan informasi dari berbagai sumber (Mardi, 2017). Data mining memiliki berbagai metode yang dapat digunakan untuk menggali informasi, salah satunya adalah klasifikasi. Salah satunya adalah klasifikasi. Klasifikasi adalah tentang menemukan model atau fungsi yang dapat membedakan antara konsep atau kelas informasi untuk memperkirakan kelas dari objek yang tidak diketahui (Mustafa dkk., 2017). Misalnya pengklasifikasian ulasan positif, netral, dan negatif.

## **2.5 Natural Language Processing (NLP)**

*Natural language processing* adalah proses pemrosesan teks yang tidak dapat dipahami bahasa alami manusia oleh komputer (Munasatya & Novianto, 2020), oleh karena itu, model Komputasi untuk berpindah dari bahasa yang biasa digunakan manusia ke interaksi manusia yang lebih alami dengan komputer (Dharmesta dkk., 2017). Dalam pembuatan model komputasi tersebut dibutuhkan beberapa tahap yaitu *case folding*, *emoticon removal*, *remove punctuation*, *stopwords removal*, *frequent word removal*, *rare word removal*, *word stemming*, *lemmatization*, dan data labeling.

## **2.6 Metode Klasifikasi**

Klasifikasi adalah fungsi yang digunakan untuk memetakan atau mengklasifikasikan objek ke dalam salah satu kelas dengan label yang telah di definisikan sebelumnya berdasarkan karakteristik objek tersebut. Terdapat proses membangun classifier (model) dan pengenalan dalam melakukan klasifikasi. Model akan menyimpan memori atau hasil training untuk kemudian menentukan kategori dari suatu objek, model tersebut dapat berupa

rule, pohon keputusan, neural network, ataupun formula matematika (Abdullah dkk., 2017).

## 2.7 Text mining

Merupakan suatu proses analisis teks yang dilakukan secara otomatis oleh komputer untuk menggali informasi yang berkualitas dari sekumpulan teks yang digabungkan menjadi sebuah dokumen (Denny, 2015). Text Mining bekerja dengan cara mengubah sebuah kata maupun kalimat dalam data yang tidak terstruktur menjadi sebuah data dengan topik yang lebih terstruktur dan dapat lebih mudah dianalisis (Yulia & Agus, 2016).

*Text mining* terbagi menjadi tiga kegiatan utama, yaitu: (i) pencarian informasi dengan mengumpulkan teks-teks yang terkait (ii) mengekstraksi informasi dari beberapa tipe spesifik yang dicari, dan (iii) data mining yang menemukan beberapa potongan informasi yang diekstraksi dari beberapa teks yang berbeda (Rozi dkk, 2012).

## 2.8 CNN (*Convolutional Neural Network*)

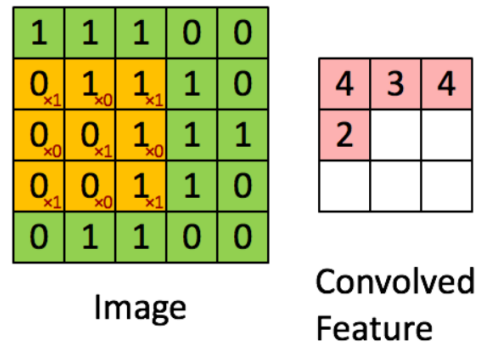
CNN (*Convolutional Neural Network*) merupakan salah satu klasifikasi utama untuk melakukan pengenalan gambar, klasifikasi gambar, mendeteksi objek, pengenalan wajah dll. CNN banyak digunakan pada inputan gambar namun bisa juga digunakan dalam kalimat atau teks (Denny, 2015).

- *Convolutional Neural Network* untuk klasifikasi teks

### 1. *Convolution Layer*

Convolution adalah cara untuk menggabungkan beberapa matriks menjadi matriks yang baru dengan batasan-batasan (*filter, padding, dll*) yang sudah ditentukan, seperti fungsi *sliding window* yang biasa digunakan pada sebuah matriks.

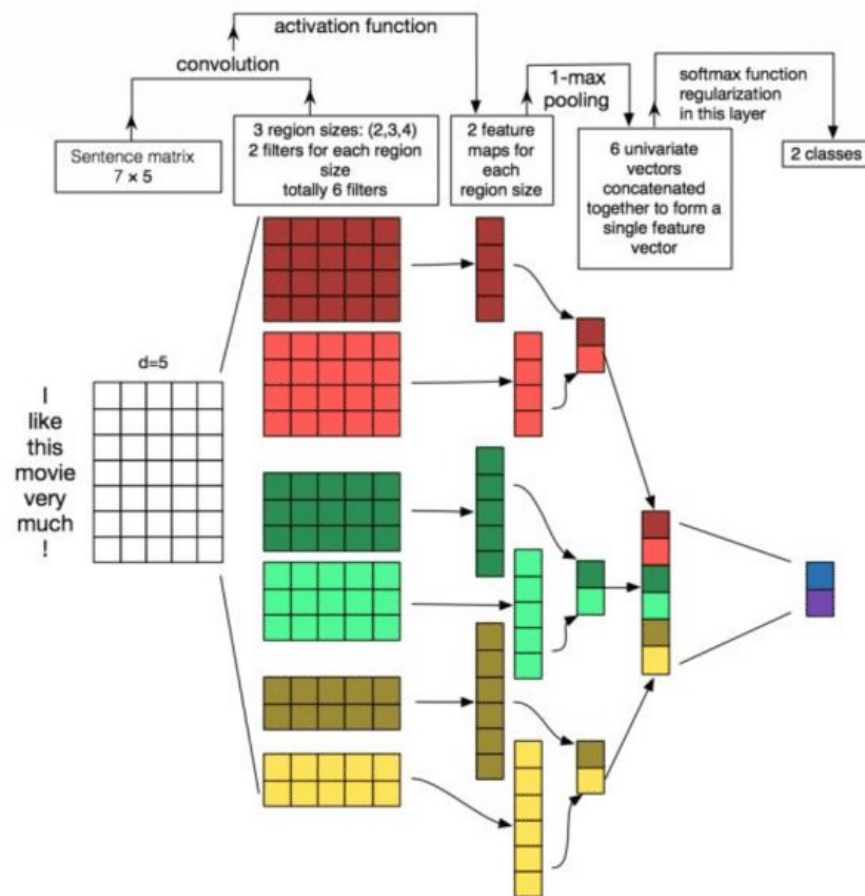
Berikut visualisasi dari proses konvolusi:



Gambar 1 Visualisasi convolution pada matriks

Pada dasarnya CNN hanya mempunyai beberapa lapisan dari konvolusi yang memiliki *nonlinear activation functions* seperti ReLU atau tanh. Secara tradisional, *fully connected layer* adalah *feedforward neural network* yang berhubungan satu sama lain dari setiap masukan neuron ke keluaran neuron pada lapisan berikutnya.

Setiap lapisan berlaku filter yang berbeda beda, biasanya ratusan hingga ribuan dan digabungkan hasilnya. Pada saat *training* dilakukan, CNN secara otomatis mempelajari nilai dari sebuah filter tergantung pada pekerjaan yang dilakukan.



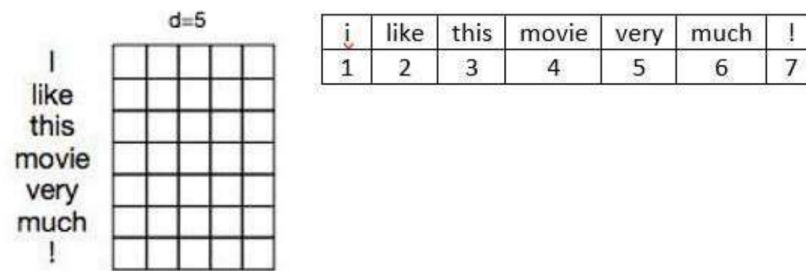
Gambar 2 Contoh Arsitektur *Convolutional Neural Network 1D*

Pada gambar 2 terdapat jaringan yang terdiri dari input layer, convolutional, Activation Function & Feature Map layer, max pooling, dan softmax function. Sebelum memasukkan teks ke dalam input layer CNN, harus ditentukan jumlah dahulu berapa jumlah matriks yang dibutuhkan oleh kalimat teks tersebut yang nantinya akan diproses pada Convolutional Layer.

## 2. Input Layer

Langkah pertama sebelum memasukkan teks ke dalam *input layer* CNN adalah menentukan jumlah matriks yang dibutuhkan dalam kalimat teks tersebut yang selanjutnya akan diproses oleh *Convolution Layer*.

Terdapat contoh kalimat “*I like this movie very much!*” yang berupa 7 kata dimana tanda seru juga dianggap sebagai kata. Tapi tanda seru dapat dihapus pada tahap *preprocessing*. Pada ilustrasi diatas ditetapkan 5 dimensi sebagai kata vector. Sehingga hasil akhir akan didapatkan yaitu berupa input layer berbentuk 7 x 5.



Gambar 3 Contoh Input Layer

Setelah mendapatkan matriks input layer CNN maka dilakukan proses konvolusi pada convolution layer. Proses ini dilakukan untuk mendapatkan nilai matriks yang baru dengan hyperparameter filter, padding, dan stride. Stride merupakan jumlah langkah pergeseran filter. Semakin kecil nilai stride maka informasi yang didapat dari sebuah input akan semakin detail. Padding merupakan jumlah piksel yang akan ditambahkan pada sisi dari matriks yang telah di inputkan. Padding bertujuan untuk mengurangi informasi yang terbuang sehingga ukuran matriks input dan matriks output tetap sama (Ihromi F, 2019).

### 3. Activation Function & Feature Map

Pada tahap ini setiap hasil output dari Convolution Layer yang sudah mengalami proses konvolusi akan di proses kembali pada activation function.



Sebelum itu, pemetaan ukuran layer output membutuhkan perhitungan rumus dibawah ini:

$$Output = \frac{N + 2p - f}{s} + 1 \quad (1)$$

Keterangan:

$N$  = ukuran input layer

$p$  = padding

$f$  = filter

$s$  = stride

Salah satu tipe Activation Function yang umum digunakan pada CNN adalah ReLU yang biasanya diterapkan secara element-wise pada output dari beberapa fungsi lain, seperti matriks-vektor (Denny, 2015).

#### 4. *Max-Pooling*

Setelah dilakukan proses *Activation Function* dan *Feature map* maka dilakukan pengambilan nilai terbesar pada layer - layer yang sudah difilter dengan persamaan dibawah ini:

$$c = \max\{c\} \quad (2)$$

#### 5. *Softmax*

Proses *softmax* dilakukan perhitungan pada nilai probabilitas dari hasil *Max pooling* yang matriksnya digabung dengan persamaan berikut:

$$S(y_i) = \frac{e^{x_i}}{e^{y_i}} \quad (3)$$

Setelah itu dilakukan perbandingan hasil softmax dengan masing-masing layer klasifikasi akhir.

## 2.9 Web Scrapping

Web scaping adalah metode yang digunakan untuk mengambil dokumen semi-terstruktur dari Internet biasanya dalam bentuk halaman web dalam format HTML atau XHTML. Dokumen dianalisis dan kemudian beberapa data dikumpulkan untuk digunakan untuk keperluan lain (Mitra dkk, 2017). Web scraping melibatkan *locating data* (lokasi data) dan kemudian mengekstraksinya. Proses web scraping tidak melakukan copy-paste, tetapi secara instan mengambil data dengan presisi dan akurasi.

## 2.10 Hyperparameter Tuning

Algoritma *machine learning* menggunakan beberapa parameter yang biasa digunakan untuk mengontrol nilai *gradient descent* dalam melakukan proses *training data*. Parameter ini dapat diatur sesuai dengan kebutuhan data dan terhadap hasil klasifikasi saat melakukan *training data*. Model CNN terdiri dari berbagai parameter yang perlu untuk di konfigurasi (Aszemi & Dominic, 2019). nilai-nilai parameter pada CNN dapat mempengaruhi hasil dari klasifikasi CNN. Hyperparameter pada CNN dibagi menjadi dua jenis, yaitu hyperparameter yang menentukan struktur jaringan dan hyperparameter yang menentukan jaringan yang dilatih. Adapun hyperparameter yang menentukan struktur jaringan sebagai berikut:

- a. *Kernel / Filter Size* merupakan ukuran filter untuk mengekstrak input feature
- b. *Embedding* adalah ukuran dimensi vektor kata dari word embedding yang telah dilakukan sebelumnya yaitu sebesar 300 dimensi
- c. *Number filter* adalah jumlah filter yang dihasilkan oleh setiap convolutional layer sebagai output agar mereduksi dimensi tetapi tetap menangkap makna kata.
- d. *Output Size* merupakan jumlah nilai kelas klasifikasi. Pada penelitian ini ada 3 output.

Nilai parameter *training* kadang berpengaruh terhadap hasil uji data nantinya. Hyperparameter yang menentukan jaringan yang digunakan sebagai berikut:

a. *Batch Size*

Total dari contoh training (*training example*) yang direpresentasikan dalam satu *batch*. Neural network tidak bisa melewati semua dataset ke dalam jaringan saraf dalam satu kali proses, sehingga dataset harus dibagi menjadi beberapa *batch*.

b. *Learning rate*

*Learning rate* adalah salah satu hyperparameter yang berfungsi untuk mengkonfigurasi rentang weight yang akan dianalisa. Semakin kecil nilai *learning rate*, maka semakin rinci pelatihannya.

c. *Epoch*

*Epoch* adalah banyaknya pengulangan pada pelatihan data yang akan dilakukan untuk mendapatkan hasil akurasi yang diharapkan. Semakin banyaknya pengulangan pada pelatihan data tidak menjamin akan mendapatkan akurasi yang lebih tinggi. Ketika akurasi mengalami penurunan maka pelatihan tersebut akan mengalami *overfitting*.

## 2.11 *Confusion Matrix*

Pengujian dilakukan menggunakan *Confusion matrix* yang merupakan sebuah metode yang biasa digunakan untuk perhitungan akurasi, *recall*, *precision*, dan *f1-score*. Dimana, *Accuracy* merupakan perbandingan kasus yang diidentifikasi benar dengan jumlah seluruh kasus. *Precision* mengevaluasi kemampuan sistem untuk menemukan diterima dan benar-benar relevan terhadap *query*. *Recall* mengevaluasi kemampuan sistem untuk menemukan semua item yang relevan terhadap *query* dan *f1-score* didapatkan dari hasil *Precision* dan *Recall*. *Confusion matrix* akan membandingkan nilai aktual dengan nilai hasil prediksi model. Dalam penelitian ini menggunakan *multi class confusion matrix 3x3* dikarenakan memiliki tiga *value* sentimen yaitu: netral, positif, dan negatif.

Tabel 1 Nilai Confusion Matrix

		PREDIKSI		
		NETRAL	POSITIF	NEGATIF
AKTUAL	NETRAL	TNet	FNetPos	FNetNeg
	POSITIF	FPosNet	TPos	FPosNeg
	NEGATIF	FNegNet	FNegPos	TNeg

Tabel 1 menjelaskan bahwa dimana biasanya *confusion matrix* terhadap nilai aktual dan prediksi. Nilai TNet (*True Netral*) merupakan nilai prediksi dan aktualnya yang bersifat netral. Nilai TPos (*True Positive*) merupakan nilai prediksi dan aktualnya yang bersifat positif. Nilai TNeg (*True Negative*). Nilai FPosNet (*False Positive Neutral*) merupakan nilai prediksi bersifat positif sedangkan nilai aktualnya yang bersifat netral. Nilai FNegNet (*False Negative Neutral*) merupakan nilai prediksi bersifat negatif sedangkan nilai aktualnya yang bersifat netral. Nilai FNetPos (*False Neutral Positive*) merupakan nilai prediksi bersifat netral sedangkan nilai aktualnya yang bersifat positif. Nilai FNegPos (*False Negative Positive*) merupakan nilai prediksi bersifat negatif sedangkan nilai aktualnya yang bersifat positif. Nilai FNetNeg (*False Neutral Negative*) merupakan nilai prediksi bersifat netral sedangkan nilai aktualnya yang bersifat negatif. Nilai FPosNeg (*False Positive Negative*) merupakan nilai prediksi bersifat positif sedangkan nilai aktualnya yang bersifat negatif.

Selanjutnya untuk pengujian peformansi akan diuji dengan perhitungan *accuracy*, *precision*, *recall*, dan *f1-score*.

1. *Accuracy*

*Accuracy* adalah nilai dari data latih dan data uji yang berpengaruh pada kedekatannya. Persamaan *accuracy* yaitu:

$$Accuracy = \frac{TNet+TPos+TNeg}{TNet+FNetPos+FNetNeg+FPosNet+TPos+FPosNeg+FNegNet+FNegPos+TNeg} \quad (4)$$

$TNet = True\ Neutral$

$TPos = True\ Positive$

$TNeg = True\ Negative$

## 2. Precision

*Precision* adalah nilai ketepatan dari data aktual dengan data prediksi. (W. Koehrsen, 2019) Persamaannya *Precision* yaitu:

$$Precision\ (Netral) = \frac{TNet}{TNet + FNetPos + FNetNeg} \quad (5)$$

*Precision* dapat digunakan sebagai acuan performansi algoritma jika dataset memiliki banyak jumlah data *TNet* (*True Neutral*), sedikit jumlah data *FNetPos* (*False Neutral Positive*). dan *FNetNeg* (*False Neutral Negative*).

Setiap kelas akan dihitung menggunakan rumus 2.7 lalu dihitung rata-ratanya menggunakan rumus:

$$\frac{Precision\ label\ 0 + Precision\ label\ 1 + Precision\ label\ 2}{3} \times 100\% \quad (6)$$

## 3. Recall

*Recall* adalah tingkat keberhasilan sistem dalam mengklasifikasikan. Persamaannya *Recall* yaitu:

$$Recall\ (Netral) = \frac{TNet}{TNet + FPosNet + FNegNet} \quad (7)$$

*Recall* dapat digunakan sebagai acuan performansi algoritma jika dataset memiliki lebih banyak jumlah data *TNet* (*True Netral*), sedikit jumlah data *FPosNet* (*False Positive Neutral*) dan *FNegNet* (*False Negative Neutral*).

Setiap kelas akan dihitung menggunakan rumus 2.8 lalu dihitung rata-ratanya menggunakan rumus:

$$\frac{\text{Recall label 0} + \text{Recall label 1} + \text{Recall label 2}}{3} \times 100\% \quad (8)$$

#### 4. *F1-Score*

*F1-Score* adalah nilai evaluasi gabungan antar *precision* dan *recall*.

Persamaannya *F1-Score* yaitu

$$F1 - Score (\text{Netral}) = 2 \times \frac{\text{RecallNet} \times \text{PrecisionNet}}{\text{RecallNet} + \text{PrecisionNet}} \quad (9)$$

Setiap kelas akan dihitung menggunakan rumus 2.9 lalu dihitung rata-ratanya menggunakan rumus.

$$\frac{F1 - score \text{ label 0} + F1 - score \text{ label 1} + F1 - score \text{ label 2}}{3} \times 100\% \quad (10)$$

Perbedaan antara *macro average* dan *weight average* adalah *macro average* menghitung matriks secara bebas untuk setiap kelas kemudian mengambil rata-ratanya. Sedangkan *weight average*, menghitung rata-rata dengan memperhitungkan bobot pada setiap datanya. Adapun rumus untuk mencari *weight average* f1-score sebagai berikut :

$$\frac{f1 - score (0) \times W0 + f1 - score (1) \times W1 + f1 - score (2) \times W2}{(W0 + W1 + W2)} \times 100\% \quad (11)$$

Adapun  $w_0$ ,  $w_1$ , dan  $w_2$  merupakan bobot dari masing- masing label netral, positif, dan negative.