**SKRIPSI**

ANALISIS SENTIMEN TERHADAP PELAKSANAAN PROGRAM MERDEKA BELAJAR KAMPUS MERDEKA MENGGUNAKAN ALGORITMA LONG SHORT TERM MEMORY (LSTM)

*Diajukan Sebagai Syarat Untuk Menyelesaikan*

*Pendidikan Program Strata-1 Pada*

*Program Studi Teknik Informatika*



Oleh:

Pebri Yanti

NIM : 1910031802001

**Program Studi Teknik Informatika**

**Sekolah Tinggi Manajemen Informatika dan Komputer Amik Riau**

**2023**

**LEMBAR PENGESAHAN SKRIPSI**

ANALISIS SENTIMEN TERHADAP PELAKSANAAN PROGRAM MERDEKA BELAJAR KAMPUS MERDEKA MENGGUNAKAN ALGORITMA LONG SHORT TERM MEMORY (LSTM)

Oleh :

Pebri Yanti

NIM : 19.10.031.802.001

|  |  |  |  |
| --- | --- | --- | --- |
|  | | Pekanbaru, ... Maret 2023 | |
| Ketua Program Studi Teknik Informatika | | Pembimbing, | |
|  | |  | |
| Junadhi, M.Kom | | Hadi Asnal, M.Kom | |
| NIDN. | | NIDN. | |
| Mengetahui, | |
| Ketua STMIK Amik Riau | |
|  | |
| Susandri, M.Kom | |
| NIDN. | |

**PENGAKUAN KEASLIAN KARYA**

Saya mengakui bahwa skripsi ini adalah hasil karya sendiri, kecuali kutipan dan beberapa data yang setiap satunya telah saya cantumkan sumbernya.

|  |
| --- |
| Pekanbaru, ... Maret 2023 |
|  |
| Pebri Yanti |
| NIM. 19.10.031.802.001 |

**ABSTRAK**

**KATA PENGANTAR**

Puji syukur kehadirat Tuhan Yang Maha Esa atas segala limpahan Rahmat-Nya sehingga penulis dapat menyelesaikan penulisan skripsi ini dengan judul “Analisis Sentimen Terhadap Pelaksanaan Program Merdeka Belajar Kampus Merdeka Menggunakan Algoritma *Long Short Term Memory* (LSTM)”. Penulis menyadari bahwa dalam penyelesaian skripsi ini telah mendapatkan banyak bantuan, bimbingan, dukungan dan pengarahan yang sangat berharga dari pembimbing, teman-teman dan semua pihak secara langsung maupun tidak langsung oleh karena itu, pada kesempatan ini melalui tulisan ini penulis mengucapkan terimakasih yang sebesar-besarnya kepada:

1. Bapak Dr. Haris Simaremare, S.T,.M.T Ketua Yayasan Komputasi Riau.
2. Bapak Susandri, M.Kom, Ketua STMIK Amik Riau.
3. Bapak Rahmaddeni, M.Kom Wakil Ketua I Bidang Akademik.
4. Ibu Rahmiati, M.Kom, Wakil Ketua II Bidang SDM Keuangan dan Administrasi Umum.
5. Bapak Nurjayadi, M.Kom, Wakil Ketua III Bidang Kemahasiswaan.
6. Bapak Jhunadi, M.Kom Kepala Program Studi Teknik Informatika.
7. Bapak Hadi Asnal, M.Kom selaku Dosen Pembimbing.
8. Seluruh Bapak/Ibu Dosen, serta Karyawan/i di lingkungan STMIK Amik Riau.
9. Serta Orang Tua yang selalu memberikan semangat, nasehat dan do’a
10. Semua pihak yang telah membantu terselesainya penulisan proposal skripsi ini.

Semoga Tuhan Yang Maha Esa dapat memberi balasan yang lebih baik kepada seluruh pihak yang telah membantu penulis. Akhirnya, penulis mengharapkan skripsi ini dapat bermanfaat bagi penulis dan pembaca pada umumnya.

|  |  |
| --- | --- |
|  | Pekanbaru, ... Maret 2023 |
|  |  |
|  | Penulis |

**DAFTAR ISI**

**BAB I PENDAHULUAN**

[1.1 Latar Belakang 1](#_Toc121820360)

[1.2 Rumusan Masalah 3](#_Toc121820361)

[1.3 Batasan Masalah 3](#_Toc121820362)

[1.4 Tujuan Penelitian 3](#_Toc121820363)

[1.5 Manfaat Penelitian 3](#_Toc121820364)

[1.6 Sistematika Penulisan 4](#_Toc121820365)

**BAB II LANDASAN TEORI DAN TINJAUAN PUSTAKA**

[2.1 Landasan Teori 5](#_Toc121820366)

[2.1.1 Analisis Sentimen 5](#_Toc121820367)

[2.1.2 Pemrosesan Teks 6](#_Toc121820368)

[2.1.3 Media Sosial 7](#_Toc121820369)

[2.1.4 Merdeka Belajar Kampus Merdeka 8](#_Toc121820370)

[2.1.5 Phyhon 8](#_Toc121820371)

[2.1.6 Twitter 9](#_Toc121820372)

[2.1.7 Tweepy 9](#_Toc121820373)

[2.1.8 Word2Vec 9](#_Toc121820374)

[2.1.9 Long Short Term Memory (LSTM) 11](#_Toc121820375)

[2.10 K-Fold cross validation 13](#_Toc121820376)

[2.11 Confusion Matrix 14](#_Toc121820377)

[2.2 Tinjauan Pustaka 15](#_Toc121820378)

**BAB III METODOLOGI PENELITIAN**

[3.1 Gambaran Umum Objek Penelitian 21](#_Toc121820379)

[3.2 Alur Penelitian 21](#_Toc121820380)

[3.2.1 Data Collection 22](#_Toc121820381)

[3.2.2 Data Labelling 23](#_Toc121820382)

[3.2.3 Text Preprocessing 23](#_Toc121820383)

[3.2.4 Training Word2vec 24](#_Toc121820384)

[3.2.5 Trainig LSTM 24](#_Toc121820385)

[3.2.6 Model LSTM 25](#_Toc121820386)

[3.2.6 Sentimen Analisis Model 25](#_Toc121820387)

[3.3 Pengujian Model 25](#_Toc121820388)

[**DAFTAR PUSTAKA** 27](#_Toc121820389)

**DAFTAR TABEL**

Tabel 2.1 Tabel *Confusion matrix* pada dua kelas. 14

Tabel 2.2 Penelitian Terdahulu 15

**DAFTAR GAMBAR**

Gambar 2.1 *One Word Context CBOW* 10

Gambar 2.2 *Mutiple Context Words CBOW* 10

Gambar 2.3 *Skip-Gram* 11

Gambar 2.4 *Layer tanh* pada RNN 11

Gambar 2.5 Perulangan empat layer pada LSTM 12

Gambar 2.6 *Ilustrasi Cell* LSTM 12

Gambar 2.7 *Cell state* pada LSTM 13

Gambar 2.8 *Sigmoid layer* pada LSTM 13

Gambar 3. 2 Alur penelitian yang digunakan. 22

Gambar 3. 2 Alur penelitian yang digunakan 22

**BAB I**

**PENDAHULUAN**

# 1.1 Latar Belakang

Kampus merdeka adalah program terobosan yang dibuat oleh Kementerian Pendidikan dan Kebudayaan (Kemendikbud). Adapun prograng yang diberikan yaitu; studi independen, magang bersertifikat, kampus mengajar, *Indonesian International Student Mobility Awards* (IISMA), petukaran mahasiswa merdeka, membangun desa (KNN Tematik), proyek kemanusiaan, riset dan penelitian, dan wirausaha. Program Kampus merdeka ini bertujuan untuk meningkatkan kompetensi mahasiswa dan memungkinkan mereka untuk beradaptasi dengan kebutuhan zaman dan perubahan sosial budaya dan teknologi. Menurut menteri pendidikan dan kebudayaan Nadiem Makarim, kampus merdeka akan memungkinkan mahasiswa untuk meningkatkan keterampilan yang tidak ada di lingkungan akademik dan dirancang untuk semua mahasiswa sehingga mahasiswa akan dibekali dengan keterampilan yang sama (Kuncoro adhi, 2020).

Media sosial yang biasanya banyak diminati salah satu yaitu twitter. Twitter merupakan salah satu media sosial yang biasa digunakan untuk mengungkapkan ide, pikiran, gagasan dan perasaan tentang kebijakan pemerintah (Rozaq et al., 2022). Pada penelitian ini memebahasa tweet - tweet yang diambil dari twitter terhadap program - program kampus merdeka terdiri dari; Studi independen, magang bersertifikat, kampus mengajar, *Indonesian International Student Mobility Awards* (IISMA), petukaran mahasiswa merdeka, membangun desa (KNN Tematik), proyek kemanusiaan, riset dan penelitian serta wirausaha. Dari delapan program tersebut maka didapatlah program yang paling banyak dikomentari yaitu, program studi independen, magang bersertifikat dan pertukaran mahasiswa merdeka, sehingga diambilah permasalah pada penelitian ini.

Analisis sentimen adalah proses komputasi untuk mengklasifikasikan data teks. Analisis sentimen menjadi semakin populer saat ini karena selain menjadi alasan banyaknya data teks yang tersedia, banyak pemangku kepentingan membutuhkan opini publik tentang topik tertentu (Widayat, 2021). Maka pada penelitian ini penulis perlu melakukan analisis sentimen dengan menggunakan metode *Long short Term Memory* (LSTM) dengan pembobotan kata word2vec dalam mengkasifikasi kata perkata yang didapat pada komentar-komentar yang terdapat di twitter. *Long Short-Term Memory* (LSTM) merupakan sistem penyimpanan data yang dapat memproses, memprediksi, dan mengklasifikasikan informasi yang telah disimpan dalam jangka waktu lama sekali pun. *Long Short-Term Memory* (LSTM) Ini juga merupakan teknik pembelajaran mendalam yang dapat digunakan untuk pemrosesan bahasa alami (NLP) seperti pengenalan ucapan, terjemahan teks dan analisis sentimen (Yahyadi & Latifah, 2022).

LSTM digunkana karena keunggulan dan serta banyaknya peneliti-peneliti yang telah melakukan penelitiannya dengan model LSTM (Uraikan 2 peneliti yang sudah melakukan penelitiannya dengan LSTM).

Berdasarkan permasalahan yang sudah di uraikan diatas serta solusi yang di tawarkan dengan algorimat dari penelitian terdahulu maka penulis berkeinginan untuk meneliti masalah analisis sentimen untuk dituangkan dalam karya tulis berbasis skripsi dengan judul analisis sentimen terhadap merdeka belajar kampus merdekan dengan menggunkan algoritma long short term memory (LSTM).

# 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan di atas, berikut masalah yang akan diselesaikan pada penelitian ini yaitu. Bagaimana analisis sentiment masyarakat pada program kampus merdeka untuk mendapatkan tingkat akurasi yang maksimal menggunkan algoritma *Long Short Term Memory* (LSTM)?

# 1.3 Batasan Masalah

Agar penulisan skripsi ini dapat terarah sesuai dengan keinginan, maksud, dan tujuan yang telah direncanakan, maka penulis membatasi ruang lingkup permasalahan yang akan dibahas antara lain:

1. Data yang dikumpulkan adalah *tweet* dengan kata kunci sebagai berikut: merdeka belajar kampus merdeka, MSIB, #kampusmerdeka, kampus merdeka, MBKM, studi independen, #merdeka belajar dan #magangmerdeka. Data yang didapat pada penelitian ini merupakan data yang bersifat publik dengan banyak data 1.305 tweet
2. Algoritma pembobotan kata yang digunakan yaitu algoritma word2vec.
3. Pada penelitian ini akan dilakukan 4 kali spliting data yang terdiri dari 60:40, 70:30, 80:20, 90:10.
4. Sentimen akan dibagi menjadi tiga yaitu positif, negatif dan netral.
5. Library paython yang digunkan yaitu; pandas, numpy, matplotlib, seaborn, pickle, tensorflow.
6. Evaluasi yang digunkan pada penelitian ini *confusion matrix*.
7. Fitur extration yang digunkana pada penelitian, yaitu fitur extration PCA.

# 1.4 Tujuan Penelitian

Adpun tujuan dari penelitian ini yaitu, untuk menemukan tingkat akurasi dari sentimen di media sosail twitter menggunakan algoritma Long Short Term Memory (LSTM) dengan pembobotan kata Word2Vec.

# 1.5 Manfaat Penelitian

Manfaat yang didapat pada penelitian ini diantaranya sebagai berikut;

1. Untuk mengetahui hasil sentiment masyarakat terhadap pelaksanaan program merdeka belajar kampus medeka yang dilaksanakan.
2. Menunjukan kepuasan dan ke tidak puasan masyarakat terhadap pelaksanaan program merdeka belajar kampus merdeka.
3. Dilakukan sebagai acuan untuk evaluasi analisis sentiment oleh pemerintah.

# 1.6 Sistematika Penulisan

Penelitian ini dibagi menjadi beberapa bab untuk memudahkan mencari informasi yang dibutuhkan dan menyajikan penyelesaian proposal skripsi yang sistematis. Pembagian babnya adalah sebgai berikut.

|  |  |
| --- | --- |
| **BAB I** | **PENDAHULUAN** |
|  | Bab ini akan membahas dan menjelaskan Latar Belakang,  Perumusan Masalah, Batasan Masalah, Tujuan Penelitian, Manfaat Penelitian, Metode Penelitian dan Sistematika Penulisan. |
| **BAB II** | **LANDASAN TEORI DAN TINJAUAN PUSTAKA** |
|  | Dalam bab ini menjelaskan tentang variabel penelitian, metode penggumpulan data, waktu atau tempat penelitian, serta alur penelitian. |
| **BAB III** | **METODOLOGI PENELITIAN** |
|  | Dalam bab ini menjelaskan tentang variabel penelitian, metode penggumpulan data, waktu atau tempat penelitian, serta alur penelitian. |
| **BAB IV** | **PEMBAHASAN DAN HASIL** |
|  | Berisi tentang implentasi dari hasil dan pembahasan dari analisi sentimen dengan metode yang sudah ada pada penelitian ini. |
| **BAB V** | **PENEUTUP** |
|  | Bab ini merupakan penutup yang meliputi tentang kesimpulan dari  pembahasan yang telah dibahas serta saran yang berguna untuk  melengkapi dan menempurnakan karya tulis ini yang berupa skripsi. |

**BAB II**

**LANDASAN TEORI DAN TINJAUAN PUSTAKA**

# 2.1 Landasan Teori

Landasan teori ini membahas teori-teori dasar yang berkaitan terhadap penelitian yang sedang dilakukan. Terdiri dari pembahasan analisis sentimen, pemrosesan teks, media sosial, merdeka belajar kampus merdeka, python, twitter, tweepy, word2vec, LSTM dan *K-Fold cross validation*.

## **2.1.1 Analisis Sentimen**

Penelitian ini merupakan pengembangan pemrosesan bahasa alami atau disebut juga dengan Natural Laguage Processing (NLP) merupakan algoritma dan sistem yang memungkinkan komputer untuk memahami dan mengeksekusi tugas yang melibatkan bahasa manusia. NLP juga dikenal sebagai linguistik komputasi yang biasanya digunakan sebagai pemrosesan bahasa. NLP dapat menganalisis bahasa manusia dalam bentuk tulisan dan lisan untuk memperoleh informasi. NLP bertujuan untuk memecahkan masalah pemahaman bahasa komputer terhadap seseorang dengan beberapa aturan tata bahasa dan semantik dengan mengubah bahasa alami manusia menjadi *representasi* data yang dapat dimengerti dan dapat diproses oleh komputer (Wicaksana, 2016b).

Analisis sentimen juga dikenal sebagai penambangan opini, ekspresi, dan perasaan audiens tentang entitas. Pendapat tersebut kemudian ditinjau ulang, baik positif maupun negatif. Analisis sentimen yaitu kemampuan untuk menganalisis data tekstual seperti tweet, komentar di media sosial, blog dan situs web. Analisis sentimen dalam bentuk teks ulasan bertujuan untuk menetukan arah ekspresif dari ulasan pengguna atau pelanggan. Analisis Sentimen menjadi populer karena meningkatnya permintaan pelanggan melalui informasi tersembunyi dari ulasan atau pengguna produk dalam bentuk data yang didistribusikan di media jaringan dan situs internet (Wicaksana, 2016b).

**2.1.2 Pemrosesan Teks**

Data teks adalah data tidak beraturan karena ada *loop* kata dan munculnya banyak kata yang tidak berkontribusi pada analisis data. Pra-pemrosesan teks atau disebut juga dengan *text pre-processing* sangat diperlukan untuk membersihkan data teks terlebih dahulu sebelum proses analisis sentimen. Berikut tahapan *pre-processing* data sebagai berikut:

1. *Remove* *Duplicate*

Sebelum mengelolah data adapun yang perlu dilakukan yaitu penghapusan data yang sama atau disebut juga *remove* *duplicate*. *Remove* *duplicate* sangat diperlukan guna untuk menghindari duplikasi data, sehingga data yang memiliki nilai yang sama akan di hapus untuk menghindari *redudansi* (Wibawanto, 2018).

1. *Filtering*

*Filtering* merupakan proses penghapusan kata-kata yang tidak penting. Pemrosesan kata yang akan dihapus seperti nama pengguna, url dan *query*. Kata-kata ini akan dihapus untuk memproleh hasil yang maksimum (Wibawanto, 2018).

1. *Case* *Folding*

Tidak semua dokumen teks konsisten dalam penggunaan huruf *kapital*. Oleh karena itu, peran *case* *folding* diperlukan untuk mengubah keseluruhan mengonversi teks dalam dokumen ke format standar atau huruf kecil. *Case folding* digunakan untuk mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf "a" sampai "z" yang diterima, menampilkan semua huruf kecuali huruf dihilangkan dan diperlakukan sebagai pemisah titik (.), Koma (,), titik koma (;), atau titik dua (:). Proses ini tidak diperlukan beberapa perpustakaan karena dapat menggunakan fungsi *lower* sebagai bahasa pemrograman python (Wibawanto, 2018).

1. *Tokenizing*

*Tokenization* adalah proses memecah dokumen teks menjadi token. Tokenisasi memecah dokumen menjadi kata, frase, simbol atau elemen lain yang bermakna. Dapat melakukan optimasi token. Dengan menghapus karakter ilegal dari file. Seperti tanda baca, simbol, angka, html dan menyebutkan (mention). Contoh menghilangkan karakter ilegal, contohnya sebagai berikut %, &, >, (, {, ], 1-9, @uluwatu, <http://tripadvisor.com> (Wicaksana, 2016a).

1. *Stopword* *Removal*

*Stopwords* digunakan untuk meminimalkan jumlah kata dalam dokumen ini mempengaruhi kinerja pemrosesan bahasa alami. Tahapan ini sangat dibutuhkan dalam menghapus kata-kata yang tidak penting dari tokenisasi. Kata-kata yang sudah di lakuan *stopword* *removal* akan disimpan dalam daftar stopword. Karakteristik utama dari memilih stopword adalah sering adanya kemunculan kata seperti konjungsi “dan”, “atau”, “tetapi”, “akan”. Tidak ada aturan yang jelas untuk menentukan stopword mana yang akan digunakan, tentukan *stopword* dapat disesuaikan berdasarkan penelitian yang dilakukan (Wibawanto, 2018).

1. *Stemming*

Pada tahap *stemming*, perannya adalah mengganti kata awal dan dengan menghilangkan imbuhan pada kata, termasuk awalan, dasar atau akar. Kata perantara (*infix*), kata akhir (*suffix*), awalan yang dihilangkan dan akhiran (*konfiks*) pada kata turunan. Pustaka (*library*) Python yang digunakan dalam proses bahasa Indonesia (Wibawanto, 2018).

1. *Padding*

Pembelajaran yang dilakukan oleh jaringan saraf membutuhkan masukkan data dengan panjang yang sama. *Padding* membawa *input* ke panjang yang sama dengan menambahkan kata. Dataset penelitian ini berisi teks dengan panjang yang berbeda. Oleh karena itu, perlu dilakukan *padding* agar vektor memiliki panjang yang sama sebelum diproses di jaringan saraf (Wicaksana, 2016b).

## **2.1.3 Media Sosial**

Media sosial adalah tempat dimana pengguna dapat terlibat dalam interaksi sosial secara tatap muka. Melalui media sosial, mudah untuk menciptakan interaksi seperti komunikasi dengan pengguna lain. Media sosial memudahkan untuk bertemu orang baru yang belum pernah anda temui sebelumnya (Wibawanto, 2018).

Dengan kemajuan itu, media sosial bukan lagi tempat untuk berkomunikasi dengan pengguna lain, tetapi tempat untuk menyuarakan pendapat anda tentang apa yang sedang dibahas dan untuk mengekspresikan diri. Pemahaman tentang media sosial ini menjadikannya sebagai platform yang menghubungkan pengguna dalam hal interaksi sosial dan ekspresi diri, membuka peluang baru bagi orang untuk menyuarakan pendapat mereka dan mendengarkan orang lain. Oleh karena itu, orang-orang dengan media sosial dapat memiliki pendapat dan interaksi yang tidak terbatas dengan orang-orang.

## **2.1.4 Merdeka Belajar Kampus Merdeka**

Merdeka Belajar Kampus Merdeka (MBKM) adalah kebijakan kualifikasi sesuai dengan peraturan menteri pendidikan dan kebudayaan (Permendikbud) nomor 3 tahun 2020 tentang standar nasional pendidikan tinggi. Kebijakan tersebut bertujuan untuk mendorong mahasiswa memasuki dunia kerja dengan berbagai ilmu yang bermanfaat. Selain itu, kebijakan tersebut bertujuan untuk meningkatkan kemampuan lulusan, termasuk *soft* dan *hard skill*, agar lebih adaptif dengan kebutuhan zaman, karena mahasiswa dapat mengembangkan potensinya sesuai kebutuhannya melalui berbagai program berbasis pembelajaran eksperimental serta minat dan bakat masing-masing (Wicaksana, 2016a).

Program MBKM memberikan kebebasan kepada perguruan tinggi dari birokrasi yang kompleks, sehingga memberikan kebebasan seluas-luasnya kepada mahasiswa untuk memilih bidang yang disukai sesuai dengan minatnya. Kebijakan MBKM dapat mengatasi masalah umum bagi lulusan yang sering menjauhkan diri dari dunia usaha. Program MBKM membutuhkan sosialisasi besar-besaran tentang mata kuliah apa saja yang dapat diambil mahasiswa serta keuntungan dan kerugian mengikuti program tersebut. Program MBKM dapat menjawab tantangan perubahan kebutuhan dan tuntutan dunia usaha dan sektor industri untuk menghasilkan bentuk *linkage and match* yang baru (Wibawanto, 2018).

## **2.1.5 Python**

Python adalah bahasa pemrograman yang dimaksudkan untuk memungkinkan pengembangan aplikasi yang cepat. Python merupakan bahasa pemrograman yang mudah dipahami karena memiliki format yang terdefinisi dengan baik untuk penggunanya (Wibawanto, 2018).

Python juga dapat digunakan untuk melakukan analisis statistik, dan dapat juga digunakan untuk membangun model pembelajaran mesin, dapat dikaitkan dengan ilmu, data, dan banyak lagi.

## **2.1.6 Twitter**

Pada tahap ini, dilakukan penghapusan karakter-karakter selain yang dapat menghubungkan pengguna di seluruh dunia. Twitter adalah forum bagi pengguna untuk menyampaikan pendapat dan mengungkapkan pendapat mereka melalui teks. Twitter telah membatasi jumlah tweet hingga 140 kata, twitter juga menyediakan layanan komunikasi yang memungkinkan layanan komunikasi yang memungkinkan pengguna untuk mengirim opini yang dibagikan secara luas. Karena twitter digunakan oleh banyak pengguna, sehingga twitter juga menggunakan sistem berbasis *real-time* dalam layanan komunikasinya (Wibawanto, 2018).

## **2.1.7 Tweepy**

Tweepy adalah sebuah library yang digunakan untuk dapat bisa mengakses API dari twitter, proses ini biasa dilakukan dengan menggunakan bahasa pemrograman python, dengan menggunkan tweepy penarikan data twitter dapat dilakukan dengan cepat, untuk bisa menggunakan. Tweepy diperlukan untuk mendapatkan akses token terlebih dahulu, akses token bisa didapatkan jika melakukan pendaftaran pada twitter development. Akses token yang akan diberikan biasanya berupa *consumer key, consumer secret, access token, access token secret* (Wibawanto, 2018).

## **2.1.8 Word2Vec**

Word2vec adalah algoritma penyematan kata yang memetakan setiap kata teks ke vector. Model penyematan kata ini telah banyak digunakan dalam *studi* NLP sejak awal. Word2vec dapat mempresentasikan kata-kata dalam vektor dan membawa makna semantik dari kata-kata tersebut.

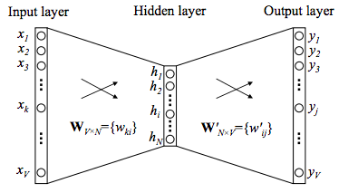
Model penyisipan kata ini adalah aplikasi pembelajaran tanpa pengawasan yang menggunakan jaringan saraf yang terdiri dari lapisan tersembunyi dan terhubungan penuh. Dimensi matrix bobot untuk setiap *layer* adalah jumlah kata dalam korpus dikalikan jumlah bobot lapisan tersembunyi. Matriks bobot lapisan tersembunyi dari model yang dilatih digunakan untuk mengubah kata vektor. Matriks bobot ini seperti tabel pencarian di mana setiap baris mewakili setiap kata dan kolom mewakili vektor untuk kata itu. Word2vec mengandalkan informasi lokal bahasa (Wicaksana, 2016b).

Semantik yang dipelajari dari kata tertentu dipengaruhi oleh kata-kata di sekitarnya. Model ini menunjukan kemampuan untuk menguji pola linguistik sebagai hubungan linier antara vektor kata.

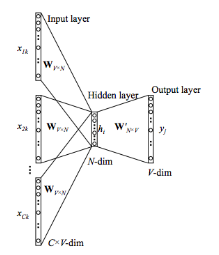
Ada dua algoritma word2vec: *Contiuous Bag-of-Word* (CBOW) dan *Skip-gram*.

1. *Contiuous Bag-of-Word* (CBOW)

Model ini menggunakan konteks untuk memprediksi kata target atau topik. CBOW memiliki waktu pelatihan lebih cepat dan akurasi lebih baik untuk kata-kata umum.



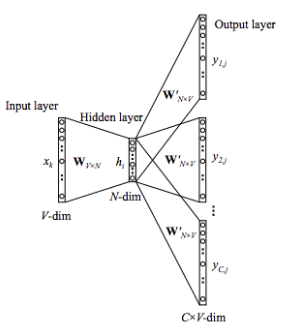
**Gambar 2.1** *One Word Context CBOW* (Fasttext & Klasifikasi, 2020).



**Gambar 2.2***Mutiple Context Words CBOW (Fasttext & Klasifikasi, 2020)*

1. *Skip-gram*

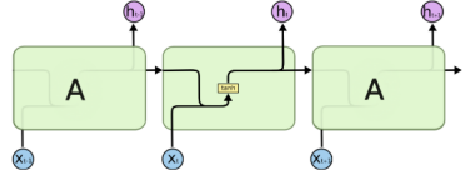
Model ini menguunakan kata-kata untuk memprediksi konteks sasaran. *Skip-gram* bekerja dengan baik pada data pelatihan kecil dan dapat mewakili kata yang dianggap langka.



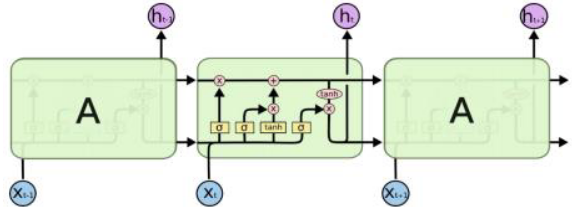
**Gambar 2.3** *Skip-Gram (Fasttext & Klasifikasi, 2020)*

## **2.1.9 Long Short Term Memory (LSTM)**

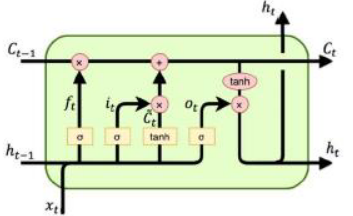
*Long Short Term Memory* (LSTM) diperkenalkan oleh Hotchreiter dan Schmidhuber (1997) untuk memecahkan masalah difusi gradien *Recurrent Neural Network* (RNN). LSTM adalah variasi dari RNN yang dirancang untuk menghindari masalah penyimpanan informasi jangka panjang dari RNN (Wicaksana, 2016b). Jaringan RNN akan melakukan perulangan hanya menggunakan satu lapisan sederhana, yaitu lapisan tanh, seperti yang ditunjukkan pada Gambar 2.4, sedangkan LSTM pada model *loop* memiliki empat lapisan, seperti yang ditunjukkan pada Gambar 2.5.



**Gambar 2.4** *Layer tanh* pada RNN (Wicaksana, 2016b).



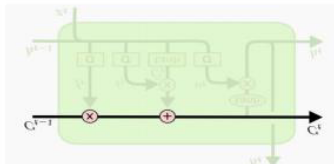
**Gambar 2.5** Perulangan empat layer pada LSTM (Wicaksana, 2016b).



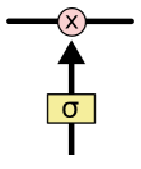
**Gambar 2.6** Ilustrasi *Cell* LSTM(Wicaksana, 2016b).

Sebuah sel dari LSTM ditunjukkan pada gambar 2.6 initerdiri dari lapisan neuron yang ditunjuk oleh kotak, dengan operasi elemen-bijaksana diwakili oleh lingkaran. Panah hitam mewakili aliran informasi di dalam dan di antara sel dan output darisel (output h). Sel LSTM memiliki 2 output, yang pertama adalah informasi stattus tersembunyi yang sebenarnya (ht), diteruskan ke sel kedua adalah keadaan sel (Ct). Status sel adalah kunci utama LSTM (Wicaksana, 2016b).

Status sel adalah garis horizontal yang menghubungkan sema lapisan keluaran pada LSTM seperti yang ditunjukkan pada gambar 2.7. LSTM memiliki kemampuan untuk menambar dan menghapus informasi dari status sel yang disebuat gerbang. Gerbang melayani fungsi mengelola informasi yang masuk, opsional menggunakan lapisan sigmoid yang ditunjukkan pada gambar2.1.8. keluaran dari lapisan sigmoid menunjukan informasi yang dilewatkan atau dibuang. Berhenti di nomor 0, dan di lanjutkan di nomor 1 (Wicaksana, 2016b).



**Gambar 2.7** Cell state pada LSTM (Wicaksana, 2016b)



**Gambar 2.8** Sigmoid *layer* pada LSTM (Wicaksana, 2016b)

Ada tiga jenis gerbang LSTM yaitu gerbang lupa forget (it), gerbang input (ft), dan gerbang output (Ot). Gerbang Forget memiliki fungsi untuk meyelesaikan data yang terhapus dari sel. Fungsi gerbang input adalah yangdihasilkan untuk menentukan nilai input yang akan diperbarui di ruang memori. Gerbang output berfungsi untuk memutuskan apakah output yang dihasilkan atau tidak sesuai dengan input dan memori sel (Wicaksana, 2016b).

## **2.11 *Confusion Matrix***

*Confusion matrix* adalah tabel yang bisa digunakan untuk mengukur performa model klasifikasi *machine learning*. Tabel ini menunjukkan seca lebih rinci berapa banyak informasi yang diklasifikasi dengan benar atau tidak. Memiliki beberapa faktor yang menetukan kinerja klasifikasi (Wibawanto, 2018).

**Tabel 2.1** Tabel confusion matrix pada dua kelas.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | | Prediksi | |
| Negatif (-) | Positif (+) |
| Aktual | Negatif (-) | TN | FN |
| Positif (+) | FP | TP |

Keterangan:

* 1. *True Negative* 𝑇𝑁 jika hasil prediksi negatif dan data aktualnya negatif.
  2. *True Positive* 𝑇𝑃 jika hasil prediksi positif dan data aktualnya positif.
  3. *False Negative* 𝐹𝑁 jika hasil prediksi negatif dan data aktualnya positif
  4. *False Positive* 𝐹𝑃 jika hasil prediksi positif dan data aktualnya negatif

Terdapat beberapa persamaan yang telah ditetapkan pada matriks dua kelas yang memiliki persamaan seperti pada Persamaan 2.1 sampai 2.7.

***𝑨𝒄𝒄𝒖𝒓𝒂𝒄𝒚* = (2.1)**

***𝑻𝒓𝒖𝒆 𝒑𝒐𝒔𝒊𝒕𝒊ve* = (2.2)**

***𝑭𝒂𝒍𝒔𝒆 𝒑𝒐𝒔𝒊𝒕ive* = (2.3)**

***𝑻𝒓𝒖𝒆 𝒏𝒆𝒈𝒂𝒕ive* = (2.4)**

***𝑭𝒂𝒍𝒔𝒆 𝒏𝒆𝒈𝒂𝒕ive* = (2.5)**

***𝑷𝒓𝒆𝒄𝒊𝒔𝒊𝒐𝒏* = (2.6)**

***𝑬𝒓𝒓𝒐𝒓 𝒓𝒂𝒕𝒆* = (2.7)**

Keterangan:

* 1. *Accuracy* 𝐴𝐶,proporsi kasus yang diidentifikasi benar dengan jumlah semua kasus.
  2. *Recall* atau *True positive rate* 𝑇𝑃, proporsi kasus *positif* yang diidentifikasidengan benar.
  3. *False positive rate* 𝐹𝑃, proporsi kasus *negatif* yang salah diklasifikasikan sebagai *positif.*
  4. *True negative rate* 𝑇𝑁, proporsi kasus *negatif* yang diklasifikasikan dengan benar.
  5. *False negative rate* 𝐹𝑁, proporsi kasus *positif* yang salah diklasifikasikan sebagai *negatif.*
  6. *Precision,* proporsi kasus dengan hasil *positif* yang benar*.*
  7. *Error Rate,* proporsi kasus yang diidentifikasi salah dengan sejumlah semua kasus.

**Tabel 2. 1** Penelitian Terdahulu

Pada analisis sentimen ini, terdapat sembilan (9) jurnal yang digunakan sebagai acuan utama dalam menjalankan penelitian ini yaitu:

1. Jurnal ini menganalisis hasil klasifikasi opini pengguna twitter terhadap program MBKM di Indonesia melalui analisis sentimen menggunakan metode Naive Bayes berbasis Particle Swarm Optimization. Hasil penelitian berupa analisis sentimen yang didapatkan adalah sentimen positif sebesar 61.92% dapat disimpulkan bahwa program MBKM ini dapat diterima dengan baik oleh masyarakat pengguna Twitter terutama mahasiswa. Walaupun ada beberapa sentimen negatif yang muncul sekitar 38.08% (Undamayanti et al., 2022).
2. Jurnal membandingkan kinerja dari word embedding seperti Word2Vec, GloVe dan FastText dan diklasifikasikan dengan algoritma Convolutional Neural Network. Sebagai pemahaman dalam melakukan perbandingan untuk menentukan performa terbaik. Performa terbaik menunjukkan FastText unggul dibanding dua metode word embedding lainnya dengan nilai F-Measure sebesar 0.979 untuk dataset 20 Newsgroup dan 0.715 untuk Reuters (Fasttext & Klasifikasi, 2020).
3. Jurnal Analisis sentimen Masyarakat terhadap Program Kampus Merdeka Menggunakan K-Nearest Neighbor dan Convolutional Neural Network (Studi Kasus Twitter)sebagai Pemahaman untuk menenntukan respon positif dan respon negatif dengan menggunakan dua metode (Kuncoro adhi, 2020).
4. Jurnal ini menjelaskan klasifikasi sentimen dengan metode word2vec untuk mengekstraksi fitur dengan mengkonversi data menjadi nilai vektor. Word2Vec memiliki keunggulan dapat melihat hubungan semantik antar kata. Metode klasifikasi pada penelitian ini menggunakan support vector machine (SVM). Dari hasil pengujian, penerapan metode word2vec dan SVM menghasilkan tingkat presisi sebesar 64,4%, recall sebesar 58%, dan f-score sebesar 61,1% (Kurniawan, 2020).
5. Jurnal Analisis Sentimen Tweet COVID-19 menggunakan Word Embedding dan Metode Long Short-Term Memory (LSTM) sebagai pemahaman untuk menganalisis respon positif ataukah negatif dengan metode Long- Short Term Memory (LSTM) dengan penambahan Word Embedding terhadap LSTM. Metode LSTM ini menghasilkan performa yang lebih baik dibandingkan dengan 2 metode lainnya yaitu Naïve Bayes dan Recurrent Neural Network (RNN) dengan selisih sebesar + 8%, dengan perincian 74% accuracy, 72% precision, 74% recall, dan 69% f-measure untuk metode Naïve Bayes dan 71% accuracy, 71% precision, 72% recall, dan 72% f-measure untuk metode RNN (Rahman et al., 2021).
6. Jurnal Analisis Sentimen Terhadap Implementasi Program Merdeka Belajar Kampus Merdeka Menggunakan *Naïve Bayes , K-Nearest Neighboars* Dan *Decision Tree* sebagai Pemahaman algoritma *Naïve Bayes, K-Nearest Neighboars* Dan *Decision Tree.* Dalam pengelompokkan ke kelas positive, negative dan netral menggunakan metode *Naïve Bayes*, *K-Nearest Neighboars* dan *Decision Tree* jumlah data sebagai data *training* dengan hasil akurasi sebesar 99,22% untuk *Naive Bayes, K-Nearest Neighbors* sebesar 96,90% dan *Decision tree* sebesarnya 37,21% (Rozaq et al., 2022).
7. Jurnal ini membahas tentang merdeka belajar kampus merdeka dengan menggunaka metode *Long Short Term Memory* (LSTM) dan Convolutional Neural Network. Untuk menentukan nilai akurasi mana yang terbaik dengan menggunakan metode Long Short Term Memory (LSTM) dan Convolutional Neural Network (Wicaksana, 2016a&b).
8. Jurnal analisis sentimen twitter terhadap kebijakan PPKM di tengah pandemi COVID-19 menggunkan model LSTM. menentukan polaritas sentiment tweet dengan Indonesia sentiment *lexicon*. Dan yang terakhir dilakukan analisis sentiment menggunakan model jaringan syaraf tiruan yaitu *Long Short Term Memory* ( LSTM) kemudian melakukan tuning hyperparameter untuk mendapatkan hyperparameter terbaik dan meningkatkan akurasi model(Yahyadi & Latifah, 2022).
9. Jurnal Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning sebagai pemahman dalam mereview dan digunakan, untuk mengetahui pengaruhnya terhadap akurasi yang dihasilkan. Akurasi terbaik diperoleh pada ukuran dimensi word vector 100 sebesar 88.17% dan akurasi terendah sebesar 85.86% pada ukuran dimensi word vector 500 (Widayat, 2021).

**BAB III**

**METODOLOGI PENELITIAN**

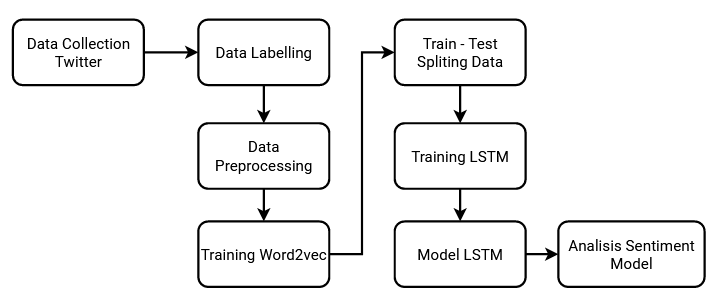
# 3.1 Gambaran Umum Objek Penelitian

Objek penelitian ini adalah pendapat masyarakat dengan menggunakan bahasa indonesia tentang pendapat masyarakat terhadapa pelaksanaan program merdeka belajar kampus merdeka yang akan di bagi dua kategori yaitu positif dan negatif. Tujuan penelitian ini adalah mengetahui hasil dari sentimen masyarakat dalam pelaksanaan program merdeka belajar kampus merdeka, dimana hasil dapat dicapai dengan melakukan analisis sentimen.

Proses analisis sentimen dilakukan dengan mengumpulkan informasi dari situs media sosial yang menyebarkan topik berita yang sedang berkembang dengan cukup cepat dalam masyarakat. Dengan banyaknya pengguna twitter di indonesia, twitter dapat merespon opini publik mengenai topik hangat, yang salah satunya program merdeka belajar kampus merdeka . oleh karena itu, data untuk analisis sentimen dikumpulkan melalui media sosial twitter. Pengambilan data tweet merupakan pandapat masyarakat indonesia tentang merdeka belajar kampus merdeka.

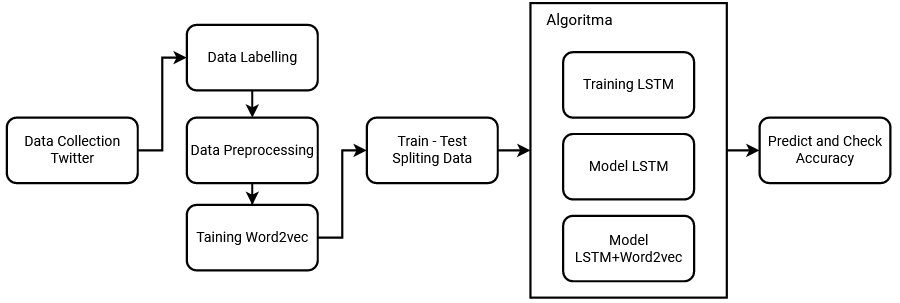
# 3.2 Alur Penelitian

Alur penelitian yang akan di lakukan terdiri dari data *collection* twitter, data *labelling*, *processing* data, *training* word2vec, *splitting* data, *training* LSTM, serta tahapan akhir adalah *predict and check accuracy.* Gambar dibawah ini adalah alur penelitian secara detail.



**Gambar 3. 1** Alur penelitian yang digunakan.

Untuk alur pemodelan di atas yang dibutuhkan 2 model algoritma yang berbeda dengan begitu proses *training* dan testing dilakukan untuk kedua model tersebut. Beriku merupakan dua alur model algoritma yang akan dilakukan.



**Gambar 3.3** Model Algoritma Word2Vec dan LSTM.

## **3.2.1 Data Collection**

Pada proses data collection, data yang akan digunakan akan dikumpulkan menjadi satu, berikut tahap yang akan dilakukan pada proses data colection adalah.

1. Crawling data adalah proses pengambilan data melalui library tweepy, baik berupa data nama pengguna atau data postingan pengguna. Pengambilan data dilakukan menggunakan google collaboratory tools dalam bahasa pemrograman python.
2. Data yang di ambil adalah data non labeling sebanyak 1.476 *tweet*.
3. Pengambilan data di mulai data tanggal 26 juli 2022 sampai 19 november 2022.
4. Pencarian data twitter menggunakan kata kunci dan tagar sehubungan dengan penelitian ini yaitu program kampus merdeka. Di bawah ini adalah contoh kata kunci dan tagar yang digunakan.
5. Merdeka belajar kampus merdeka
6. #kampus merdeka
7. MSIB
8. Kampus merdeka
9. MBKM
10. Studi independen
11. #merdekabelajar
12. #magangmerdeka

Kata kunci di atas sebuah pencarian yang digunakan untuk menarik data pada twitter karena merupakan program yang ada di indonesia.

## **3.2.2 Data Labelling**

Informasi yang sebelumnya dikumpulkan dalam penandaan atau di sebut juga dengan data labeling adalah pengkasifikasian yang berdasarkan pendapat masyarakat. Pendapat itu sendiri memiliki dua kategori yaitu pendapat negatif dan positif, tujuannya adalah untuk mengetahui seberapa baik kinerja saat pengujian model algoritma.

## **3.2.3 Text Preprocessing**

Tahapan text preprocessing merupakan lanjutan dari tahap penambangan data. Hasil penambangan data tentunya belum rapi dan tidak terstruktur. Maka dari itu dibutuhkan tahap text preprocessing dimana tahap ini akan mengolah informasi yang kotor tadi, akan melalui serangkaian proses seperti berikut untuk menghasilkan data yang dapat digunakan untuk penelitian. Berikut adalah penjabaran dari tahap text preprocessing:

1. *Remove* *Duplicate*

Sebelum data di proses, untuk menghindari data yang berulang maka perlu dilakukan *remove* *duplikcate*, data yang memiliki nilai yang sama akan di hapus agar tidak terja redudansi.

1. *Case* *Folding*

Proses *case* *folding* adalah merubah teks menjadi huruf kecil untuk menyederhanakan penggunaan, agar dapat diolah.

1. *Filtering*

*Filtering* adalah proses menyaring kata-kata yang dianggap tidak terlalu penting pemrosesan data menghapus seperti nama pengguna, URL, *query*. Kata-kata tersebut akan di hapus sehingga pemrosesan mendapat hasil yang baik (Wibawanto, 2018).

1. *Remove* *Stopword*

Pada proses *stopword* akan menghapus beberapa kata untuk mengurangi noise sehingga data bisa dijadikan untuk data *train* dan data *test*.

1. *Tokenizing*

Pada proses *tokenizing*, kalimat pada tweets akan di potong menjadi sebuah kumpulan kata-kata agar lebih mudah udah dilakukan *stemming* (Wicaksana, 2016b).

1. *Stemming*

Pada proses *stemming*, kata-kata yang sudah dipotong pada proses *tokenizing*, selanjutnya kata akan dipecah menjadi kata dasar tanpa imbuhan. Proses *stemming* pada penelitian kali ini menggunakan *library* sastrawi (Wicaksana, 2016b).

## **3.2.4 Training Word2vec**

Pada penelitian ini akan dilakukan representasi kata dengan menggunakan word2vec yang dibuat oleh google. Dengan menggunkan word2vec setiap kata dapat di representasikan ke dalam bentuk *vector* dengan ukuran N dimensi tertentu (Widayat, 2021). Maka dari situ data yang akan di gunakan untuk representasi kata yaitu data *training*. Data training ini akan digunakan dalam melakukan representasi kata dengan menggunakan algoritma word2vec.

## **3.2.5 Trainig LSTM**

Pada tahap ini akan dilakukan pembagian komposisi dataset untuk *training* dan *testing* model LSTM yang dibuat akan menggunakan pembagian data sebesar 70% : 30% sampai dengan 90% : 10%. Dengan catatan porsi dataset untuk training lebih besar.

## **3.2.6 Model LSTM**

Pada tahap ini dilakukan pelatihan lebih lanjut terhadap model LSTM yang dibuat dengan menggunakan data masukan yang telah disiapkan sebelumnya. Data input merupakan kumpulan vektor hasil dari tahap preprocessing dan pembuatan vektor kata pada tahap sebelumnya. Alur proses pelatihan model LSTM ini adalah setiap kata yang direpresentasikan dalam bentuk vektor dimasukkan ke dalam model LSTM satu per satu sesuai dengan urutannya. Proses ini berulang hingga urutan terakhir, setelah itu model LSTM menggunakan fungsi aktivasi sigmoid untuk menentukan nilai probabilitas sentimen evaluasi. Setiap kata yang direpresentasikan dalam bentuk vektor (word vector) diinputkan satu per satu ke dalam unit LSTM sesuai urutan pemeriksaan data sequence. Pencocokan parameter dilakukan terhadap beberapa parameter, seperti panjang urutan input, jumlah unit LSTM yang digunakan, dan ukuran dimensi vektor kata untuk setiap kata input. Analisis dilakukan untuk mengetahui dampak penyetelan parameter tersebut terhadap hasil akurasi yang diperoleh dari model LSTM (Widayat, 2021).

## **3.2.6 Sentimen Analisis Model**

Pada tahap ini akan dilakukan analisis sentimen terhadap tweet untuk mengetahuinya tweet mana yang termasuk dalam kategori sentimen positif dan negatif. Pada langkah ini peneliti akan menggunakan dataset traning dengan model algoritma word2vec dan long short term memory, dengan menggunakan pemabagian data 70% : 30% sampai dengan 90% : 10%. hasil dari perbandingan akan digunakan untuk melakukan pengukuran dari dua jenis algoritma yang akan digunakan.

# 3.3 Pengujian Model

Keakuratan model algoritma akan diuji pada tahap ini. Uji akurasi  
berfungsi sebagai tolak ukur keberhasilan penggunaan model algoritma. Dalam  
memperoleh hasil tersebut, uji akurasi menggunakan Confusion Matrix.  
Semakin tinggi akurasi, semakin tinggi tingkat keberhasilan algoritma.  
Tingkat keberhasilan dari sebuah pengujian model algortima terletak pada hasil  
pengujian akurasi algoritma itu sendiri. Pada penelitian sebelumnya hasil  
akurasi yang didapatkan pada algoritma LSTM dengan pembobotan Word2Vec  
sebesar 81% (Rahman et al., 2021). Pada penelitian ini tingkat keberhasilan dapat dikatakan berhasil jika hasil akurasi dari pengujian model melebihi dari nilai 81% untuk LSTM.

**BAB IV**

**HASIL DAN PEMBAHASAN**

**BAB V**

**PENUTUP**

# 5.1 Kesimpulan

Berdasarkan implementasi yang telah dilakukan, kesimpulan yang dapat diambil sebagai berikut :

# 5.2 Saran

# DAFTAR PUSTAKA

Fasttext, D. A. N., & Klasifikasi, P. (2020). *PERBANDINGAN KINERJA WORD EMBEDDING WORD2VEC , GLOVE ,*. *14*(2).

Kuncoro adhi. (2020). Bab I .... ﻢﻠﺳو ﮫﯿﻠﻋ ﷲ ﻰﻠﺻ ﻲ ﱠ ِ ﺒﱠﻨﻟا ﺎ َ ﻨ ﯿ ْ َ َ ر ﺎ َ ﻤ ﻛ َ اﻮﱡﻠ َ ﺻ َ و. *Bab I*, 1–16.

Kurniawan, F. W. (2020). Analisis Sentimen Twitter Bahasa Indonesia dengan Word2Vec. *Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *2*(2), 4704–4713. https://openlibrary.telkomuniversity.ac.id/home/catalog/id/159923/slug/analisis-sentimen-twitter-bahasa-indonesia-dengan-word2vec.html%0A/home/catalog/id/159923/slug/analisis-sentimen-twitter-bahasa-indonesia-dengan-word2vec.html

Rahman, M. Z., Sari, Y. A., & Yudistira, N. (2021). Analisis Sentimen Tweet COVID-19 menggunakan Word Embedding dan Metode Long Short-Term Memory (LSTM). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *5*(11), 5120–5127. http://j-ptiik.ub.ac.id

Rozaq, A., Yunitasari, Y., Sussolaikah, K., Resty, E., Sari, N., & Syahputra, R. I. (2022). Analisis Sentimen Terhadap Implementasi Program Merdeka Belajar Kampus Merdeka Menggunakan Naïve Bayes , K-Nearest Neighboars Dan Decision Tree. *Jurnal Media Informatika Budidarma*, *6*(April), 746–750. https://doi.org/10.30865/mib.v6i2.3554

Undamayanti, E., Iman Hermanto, T., Kaniawulan, I., Studi, P., Informatika, T., Teknologi, S. T., & Purwakarta, W. (2022). Analisis Sentimen Menggunakan Metode Naive Bayes Berbasis Particle Swarm Optimization Terhadap Pelaksanaan Program Merdeka Belajar Kampus Merdeka. *Jurnal Sains Komputer & Informatika (J-SAKTI*, *6*(2), 916–930.

Wibawanto. (2018). Bab Ii Landasan Teori. *Journal of Chemical Information and Modeling*, *53*(9), 8–24.

Wicaksana, A. (2016a). 済無No Title No Title No Title. *Https://Medium.Com/*, 2017–2019. https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf

Wicaksana, A. (2016b). 済無No Title No Title No Title. *Https://Medium.Com/*, *1*, 4–25. https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf

Widayat, W. (2021). Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning. *Jurnal Media Informatika Budidarma*, *5*(3), 1018. https://doi.org/10.30865/mib.v5i3.3111

Yahyadi, A., & Latifah, F. (2022). Analisis Sentimen Twitter Terhadap Kebijakan PPKM di Tengah Pandemi COVID-19 Menggunakan Mode LSTM. *Journal of Information System, Applied, Management, Accounting and Research*, *6*(2), 464–471. https://doi.org/10.52362/jisamar.v6i2.791