**ANALISIS SENTIMEN TENTANG WORK FROM HOME PADA MASA TANGGAP DARURAT  
COVID-19 MENGGUNAKAN CONVOLUTIONAL  
NEURAL NETWORK**

**TUGAS AKHIR**

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik

pada Jurusan Teknik Informatika Fakultas Sains dan Teknologi

Universitas Islam Negeri Sunan Gunung Djati Bandung

**Oleh :**

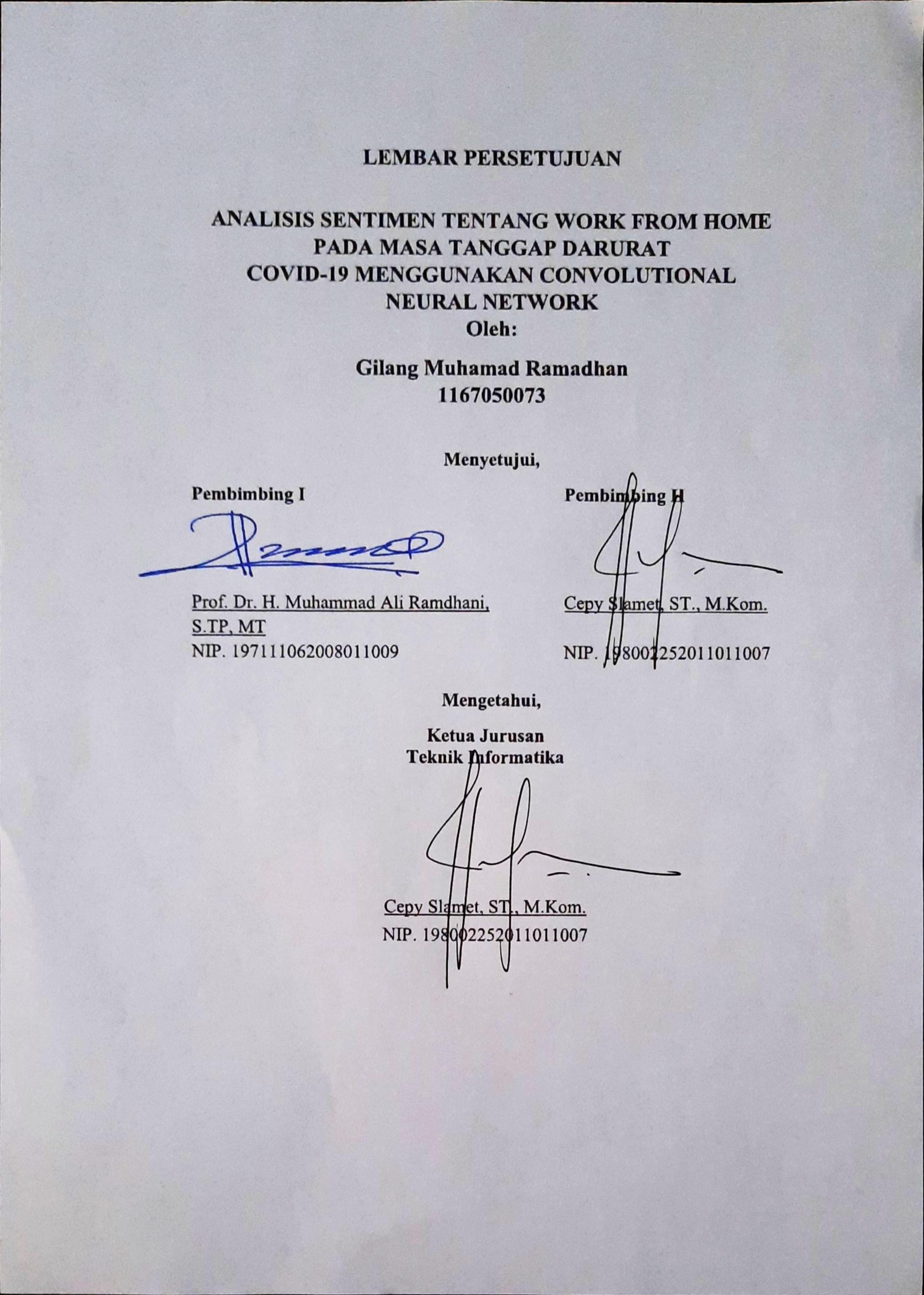
**Gilang Muhamad Ramadhan**

**1167050073**



**BANDUNG**

**2020 M / 1441 H**

**LEMBAR PERSETUJUAN**

**ANALISIS SENTIMEN TENTANG WORK FROM HOME PADA MASA TANGGAP DARURAT  
COVID-19 MENGGUNAKAN CONVOLUTIONAL  
NEURAL NETWORK**

**Oleh:**

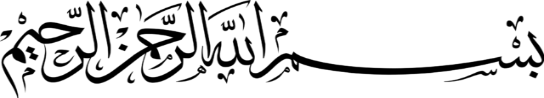
**Gilang Muhamad Ramadhan  
1167050073**

**Menyetujui,**

|  |  |  |  |
| --- | --- | --- | --- |
| **Pembimbing I** |  | **Pembimbing II** | |
|  | | |
| Prof. Dr. H. Muhammad Ali Ramdhani, S.TP, MT |  | Cepy Slamet, ST., M.Kom. | |
| NIP. 197111062008011009 |  | NIP. 198002252011011007 | |

**Mengetahui,**

|  |  |  |
| --- | --- | --- |
|  |  | **Ketua Jurusan**  **Teknik Informatika** |
|  |
|  |  | Cepy Slamet, ST., M.Kom. |
|  |  | NIP. 198002252011011007 |

KATA PENGANTAR

*Assalamu’alaikum Warahmatullahi Wabarakatuh*

Puji dan syukur penulis panjatkan kehadirat Allah SWT yang telah memberikan kekuatan dan kesehatan kepada penulis, sehingga penulis dengan segala kekurangan dan keterbatasannya dapat menyelesaikan skripsi yang berjudul **“Analisis Sentimen Tentang Work From Home Pada Masa Tanggap Darurat Covid-19 Menggunakan Convolutional Neural Network”**. Tak lupa sholawat serta salam semoga tetap terlimpah curahkan kepada Nabi dan Rasul Muhammad SAW, semoga safaatnya sampai kepada kita sebagai pengikut ajarannya, aamiin.

Penyusunan skripsi ini tidak terlepas dengan tujuan penambahan kajian keilmuan secara pribadi juga mudah-mudahan menambah wawasan dan manfaat bagi siapa saja yang membaca penelitian skripsi ini, juga tidak terlepas dari penyelesaian proses perkuliahan sehingga mendapatkan gelar Strata Satu (S1) bidang Teknik Informatika di Fakultas Sains dan Teknologi UIN Sunan Gunung Djati. Penulis menyadari dalam proses pembelajaran dan penyelesaian penyusunan skripsi ini tidak terlepas dari bimbingan, bantuan dan dorongan dari berbagai pihak, untuk itu pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Orang Tua (Bapak Dadang dan Ibu Imas), adik Gibran , dan Galban yang selalu memberikan motivasi dan do’a yang tak henti-hentinya hingga penulis dapat menyelesaikan pendidikan S1 ini.
2. Bapak Cepy Slamet, S.T., M.Kom selaku Ketua Jurusan Teknik Informatika yang telah memberikan dukungan, bimbingan dan kerjasamanya dalam penyelesaian skripsi ini.
3. Bapak Agung Wahana, SE., MT. selaku Pembimbing Akademik yang telah memberikan dukungan, bimbingan dan kerjasamanya selama ini.
4. Bapak Prof. Dr. H. Muhammad Ali Ramdhani, S.TP, MT selaku Pembimbing I yang selama pengerjaan skripsi ini selalu memberikan masukan yang berharga.
5. Bapak Cepy Slamet, S.T., M.Kom selaku Pembimbing II yang selama pengerjaan skripsi ini selalu memberikan masukan yang berharga.

Dan pihak-pihak lain yang tidak dapat disebutkan satu persatunya. Semoga apa yang telah diberikan mendapatkan imbalan yang setimpal dari Allah SWT.

Penulis hanyalah manusia biasa yang tak luput dari salah dan hilaf, maka dari itu untuk perbaikan kekurangan penulisan ini, penulis menerima kritik dan saran yang bersifat membangun dari pembaca untuk pengembangan ke arah yang lebih baik lagi. Akhir kata semoga skripsi ini dapat bermanfaat khususnya bagi penulis dan umumnya bagi pembaca.

*Wassalamualaikum Wr Wb.*

Bandung, 9 Juli 2020

Penulis

# ABSTRAK

Analisis sentimen akan mengelompokkan kontradiksi dari teks yang ada dalam suatu kalimat ataupun dokumen untuk mengetahui pendapat yang dikemukakan tersebut apakah bersifat positif atau negatif, yaitu mengenai *Work From Home* pada masa tanggap darurat *Covid*-19 pada media sosial twitter, dimana istilah *WFH* ini sekarang menjadi isu sekaligus budaya baru bagi kita khusus nya di Indonesia. Berharap budaya *WFH* ini mampu berjalan dengan efektif layaknya bekerja normal yaitu di kantor atau tempat kerja nya itu sendiri. Analisis sentimen merupakan sebuah riset secara terkomputasi terhadap opini sentimen dan emosi yang dikemas secara tekstual. Banyak metode yang dapat diterapkan dalam proses analisis sentimen, salah satunya adalah metode *deep learning.* Tujuan adanya penelitian ini, sistem yang dibangun mampu mengimplementasikan metode Convolutional Neural Network dalam menentukan suatu kalimat itu apakah positif dan negatif terhadap tweet tentang *WFH* dari berbagai orang dan negaranya masing-masing. Dengan begitu kita dapat melihat apakah budaya *WFH* ini efektif atau justru orang-orang kesulitan menjalaninya.

**Kata Kunci** : Implementasi, Work From Home, Sentimen Analisis, Convolutional Neural Network.

# *ABSTRACT*

*Sentiment analysis will classify the contradictions of the text in a sentence or document to find out whether the expressed opinion is positive or negative, namely regarding Work From Home during the Covid-19 emergency response period on social media Twitter, where the term WFH is now an issue as well as a new culture for us specifically in Indonesia. Hoping that this WHF culture will be able to run effectively like normal work in the office or workplace itself Sentiment analysis is a computational research on sentiment and emotion opinions that are textually packaged. Many methods can be applied in the sentiment analysis process, one of them is a deep learning method. The purpose of this research The system built is able to implement the Convolutional Neural Network method in determining whether a sentence is positive and negative about tweets about WFH from various people and their respective countries. That way we can see whether this WFH culture is effective or if people have difficulty living it.*

***Keywords*** *: Implementation, Work From Home, Sentiment Analysis, Convolutional Neural Network.*

DAFTAR ISI

[**LEMBAR PERSETUJUAN** 2](#_Toc47726651)

[KATA PENGANTAR 2](#_Toc47726652)

[ABSTRAK ii](#_Toc47726653)

[*ABSTRACT* iii](#_Toc47726654)

[DAFTAR ISI iv](#_Toc47726655)

[DAFTAR GAMBAR vii](#_Toc47726656)

[DAFTAR TABEL viii](#_Toc47726657)

[BAB I 2](#_Toc47726658)

[PENDAHULUAN 2](#_Toc47726659)

[1.1 Latar Belakang 2](#_Toc47726660)

[1.2 Rumusan Masalah 5](#_Toc47726661)

[1.3 Tujuan Penelitian 5](#_Toc47726662)

[1.4 Batasan Masalah 6](#_Toc47726663)

[1.5 Manfaat Penilitian 6](#_Toc47726664)

[1.6 Kerangka Pemikiran 7](#_Toc47726665)

[1.7 Metodologi Penelitian 9](#_Toc47726666)

[1.7.1 Teknik Pengumpulan Data 9](#_Toc47726667)

[1.7.2 Teknik Pengembangan 10](#_Toc47726668)

[1.8 Sistematika Penulisan 11](#_Toc47726669)

[BAB II 14](#_Toc47726670)

[STUDI PUSTAKA 14](#_Toc47726671)

[2.1 Tinjauan Pustaka 14](#_Toc47726672)

[2.2 Landasan Teori 17](#_Toc47726673)

[2.2.1 Analisis Sentimen 17](#_Toc47726674)

[2.2.2 Twitter 17](#_Toc47726675)

[2.2.3 Python 18](#_Toc47726676)

[2.2.4 Klasifikasi 18](#_Toc47726677)

[2.2.5 Deep Learning 19](#_Toc47726678)

[2.2.6 Keras 19](#_Toc47726679)

[2.2.7 *Tensorflow* 19](#_Toc47726680)

[2.2.8 Algoritma *Convolutional Neural Network* (CNN) 20](#_Toc47726681)

[2.2.9 Model Pengembangan *Waterfall* 21](#_Toc47726682)

[2.2.10 *Preprocessing* 22](#_Toc47726683)

[2.2.11 Metode *Word2Vec* 22](#_Toc47726684)

[2.2.12 Data Flow Diagram (DFD) 23](#_Toc47726685)

[2.2.13 *Confusion Matrix* 25](#_Toc47726686)

[BAB III 28](#_Toc47726687)

[ANALISIS DAN PERANCANGAN 28](#_Toc47726688)

[3.1 System Requirment 28](#_Toc47726689)

[3.1.1 Kebutuhan Fungsional 28](#_Toc47726690)

[3.1.2 Kebutuhan Non Fungsional 28](#_Toc47726691)

[3.2 Software Requirment 29](#_Toc47726692)

[3.2.1 Hardware 29](#_Toc47726696)

[3.2.2 Software 29](#_Toc47726697)

[3.3 Analisis 29](#_Toc47726698)

[3.3.1 Analisis Masalah 29](#_Toc47726699)

[3.3.2 Analisis Data 30](#_Toc47726700)

[3.3.3 Analisis Algoritma 31](#_Toc47726701)

[3.3.4 *Flowchart* Algoritma *Convolutional Neural Network* 31](#_Toc47726704)

[3.3.5 Pemisahan Data 32](#_Toc47726705)

[3.3.6 *Preprocessing* 32](#_Toc47726706)

[3.3.6.1 Casefolding 33](#_Toc47726707)

[3.3.6.2 *Cleaning* 34](#_Toc47726708)

[3.3.6.3 *Stemming* 34](#_Toc47726709)

[3.3.6.4 *Stop Words* 34](#_Toc47726710)

[3.3.7 Pembobotan *Word2Vec* 35](#_Toc47726711)

[3.3.8 Analisis Algoritma *CNN* 35](#_Toc47726712)

[3.4 Desain 37](#_Toc47726713)

[3.4.1 Context Diagram 37](#_Toc47726714)

[3.4.2 Data Flow Diagram Level 1 38](#_Toc47726715)

[3.4.3 Data Flow Diagram Level 2 39](#_Toc47726716)

[3.5 Perancangan User Interface 40](#_Toc47726717)

[BAB IV 41](#_Toc47726718)

[IMPLEMENTASI DAN PENGUJIAN 41](#_Toc47726719)

[4.1 Implementasi Sistem 41](#_Toc47726720)

[4.1.1 *Hardware* 41](#_Toc47726721)

[4.1.2 *Software* 41](#_Toc47726722)

[4.2 Pengujian 42](#_Toc47726723)

[4.2.1 Pengujian Pertama 43](#_Toc47726724)

[4.2.2 Pengujian Kedua 45](#_Toc47726725)

[4.2.3 Pengujian Ketiga 48](#_Toc47726726)

[4.2.4 Pengujian Keempat 50](#_Toc47726727)

[4.2.5 Pengujian Kelima 53](#_Toc47726728)

[4.2.6 Pengujian Keenam 55](#_Toc47726729)

[4.2.7 Pembahasan Hasil Pengujian 58](#_Toc47726730)

[BAB V 60](#_Toc47726731)

[PENUTUP 60](#_Toc47726732)

[5.1 Kesimpulan 60](#_Toc47726733)

[5.2 Saran 61](#_Toc47726734)

[DAFTAR PUSTAKA 62](#_Toc47726735)

DAFTAR GAMBAR

[Gambar 1. 1 Kerangka Pemikiran 8](#_Toc45205015)

[Gambar 2. 1 Model Waterfall 21](#_Toc45205113)

[Gambar 2. 2 Proses Preprocessing Data 22](#_Toc45205114)

[Gambar 2. 3 Confusion Matrix 25](#_Toc45205115)

[Gambar 3. 1 Data Preparation Process 30](#_Toc46254571)

[Gambar 3. 2 Flowchart Algoritma CNN 31](#_Toc46254572)

[Gambar 3. 3 Data Structure 32](#_Toc46254573)

[Gambar 3. 4 Arsitektur Model Algoritma Convolutional Neural Network 36](#_Toc46254574)

[Gambar 3. 5 Context Diagram 37](#_Toc46254575)

[Gambar 3. 6 Data Flow Diagram Level 1 38](#_Toc46254576)

[Gambar 3. 7 Data Flow Diagram Level 2 39](#_Toc46254577)

[Gambar 3. 8 User Interface Jupyter Notebook 40](#_Toc46254578)

[Gambar 4. 1 Persentase Data Negatif dan Positif 42](#_Toc45205971)

[Gambar 4. 2 Confusion Matrix Pengujian Pertama 43](#_Toc45205972)

[Gambar 4. 3 Grafik F1 Score Pengujian Pertama 44](#_Toc45205973)

[Gambar 4. 4 Confusion Matrix pengujian kedua 45](#_Toc45205974)

[Gambar 4. 5 Grafik F1 Score pengujian kedua 47](#_Toc45205975)

[Gambar 4. 6 Confusion Matrix pengujian ketiga 48](#_Toc45205976)

[Gambar 4. 7 Grafik F1 Score pengujian ketiga 49](#_Toc45205977)

[Gambar 4. 8 Confusion Matrix pengujian keempat 50](#_Toc45205978)

[Gambar 4. 9 Grafik F1 Score pengujian keempat 52](#_Toc45205979)

[Gambar 4. 10 Confusion Matrix pengujian kelima 53](#_Toc45205980)

[Gambar 4. 11 Grafik F1 Score pengujian kelima 54](#_Toc45205981)

[Gambar 4. 12 Confusion Matrix pengujian keenam 55](#_Toc45205982)

[Gambar 4. 13 Grafik F1 Score pengujian keenam 57](#_Toc45205983)

[Gambar 4. 14 Grafik hasil pengujian 58](#_Toc45205984)

DAFTAR TABEL

[Tabel 2. 1 State of The Art 16](#_Toc45205204)

[Tabel 2. 2 DFD Symbol 24](#_Toc45205205)

[Tabel 3. 1 Data Set Labeling 33](#_Toc45205236)

[Tabel 3. 2 Case Folding 33](#_Toc45205237)

[Tabel 3. 3 Cleaning Data 34](#_Toc45205238)

[Tabel 3. 4 Stemming 34](#_Toc45205239)

[Tabel 3. 5 Stop Words 35](#_Toc45205240)

[Tabel 4. 1 Hasil pengujian 59](#_Toc46254582)

BAB I

PENDAHULUAN

## Latar Belakang

Akhir-akhir ini media pemberitaan di seluruh dunia dipenuhi mengenai wabah atau *pandemic Covid-19.* Wabah ini pertama terdeteksi dan menyebar di Kota Wuhan, Tiongkok pada bulan desember 2019. Pasalnya *pandemic* ini sudah menyebar hamper ke seluruh dunia termasuk sekarang menyerang ke negara kita yaitu Indonesia [1].

Penggunaan media informasi yang awalnya hanya berupa media cetak (koran, majalah dan sejenisnya) sekarang beralih ke media elektronik karena di anggap lebih mudah untuk diakses dan pembaharuan suatu beritanya cepat. Saat ini media informasi yang banyak digunakan diantaranya *Twitter, Facebook, Instgram, dan Youtube*. Menurut data yang diperoleh dari *Webershandwick*, perusahaan pemberi layanan jasa komunikasi dan *public relations*, di Indonesia sekitar 65 juta pengguna aktif Facebook dan sejumlah 33 juta pengguna yang aktif per harinya. Lalu utuk pengguna Twitter berdasarkan data dari PT Bakrie Telecom, yaitu sebanyak 19,5 juta pengguna. Bahkan jumlah pengguna Twitter ini secara global terus meningkat. Pada tahun 2019, pengguna aktif harian di platform ini bertambah 17 persen menjadi ke angka 145 juta pengguna dan faktanya Indonesia salah satu negara yang pertumbuhan pengguna aktif *Twitter* paling besar. Oleh angka yang terus bertambah setiap tahunnya yang dimana dapat dikatakan sebagian rakyat Indonesia sudah melek dengan teknologi [2].

Istilah *Work From Home* muncul ketika merebaknya wabah corona, sering disingkat WFH dan kian akrab di telinga kita. Cara ini, selain dapat mengurangi atau memutus rantai penularan virus corona, dampak lain yang dirasakan adalah berkontribusi terhadap mengurangi tingkat kemacetan sehingga polusi menurun dan kesehatan lingkungan meningkat menjadi lebih baik. Pilihan *WFH* ini adalah logis, karena dianggap sebagai salah satu ikhtiar dari musibah yang sedang terjadi sat ini. Karena pada saat wabah seperti ini, keramaian atau kerumunan sangat dilarang keras, menjadi tempat penyebaran virus yang ada. Menjaga jarak dan kontak fisik atau saat ini kita dengan dengan istilah *physical distancing* merupakan tindakan yang sangat dianjurkan untuk mencegah tertularnya virus covid-19 ini. Di Eropa sana, budaya *work from home* ini sering diistilahkan dengan *telework,* ada juga yang menyebutnya *telecommuting* yang pada intinya bekerja dengan cara jarak jauh (tidak berada di kantor). Kemajuan teknologi informasi dan komunikasi telah memungkinkan manusia mengerjakan banyak hal dari jarak jauh. Namun berbeda dengan keadaan di Indonesia yang mana *WFH* belum terlalu terbiasa ataupun tabu. Meskipun dalam keadaan yang belum terbiasa, kita tetap harus berikhtiar dalam mencari nafkah karena pada dasarnya Allah akan memberikan jalan bagi hamba-Nya yang berusaha seperti dalam hadits Rasulullah;   
"إِذَا سَبَّبَ اللهُ لِأَحَدِكُمْ رِزْقًا مِنْ وَجْهٍ فَلَا يَدَعْهُ حَتَّى يَتَغَيَّرَ لَهُ أَوْ يَتَنَكَّرَ لَهُ " رواه ابن ماجه   
yang artinya : “Jika Allah memberikan jalan bagi seseorang di antara kamu untuk memperoleh rezeki dari suatu arah, maka janganlah dia meninggalkannya sampai dia berubah atau hilang darinya” (HR. Ibnu Majah). Dari hadits tersebut kita dapat menyimpulkan bahwa pada saat ini salah satu arah datangnya rezeki yang diberikan Allah adalah bekerja, baik itu di rumah maupun di kantor, selama pekerjaan itu masih halal. Budaya ini pun memeliki kekurangan dan kelebihannya [3].

Dengan segala kekurangan dan kelebihan yang ditimbulkan dari *Work From Home* dan itu pun tergantung dari jenis pekerjaan yang dilakukan, orang-orang mencoba untuk mengungkapkan baik itu opini maupun pengalaman *WFH* mereka masing-masing ke sosial media termasuk Twitter. Ada yang bersifat positif maupun negatif. Oleh karena itu, manfaat dari analisis sentimen yaitu agar dapat mengelompokkan dan mengetahui mana sentimen positif dan mana yang sentiment negatif.

*Deep Learning* merupakan cabang dari keilmuan *Machine Learning.* Salah satu metode  *Deep Learning* adalah *Recurrent Neural Network* (RNN), *Convolutional Neural Network* (CNN), dan *Long Short Term Memory* (LSTM). Dengan menggunakannya *Deep Learning* dapat menutupi kekurangan metode sebelumnya. Metode ini dipilih karena dapat dilatih dan konsusten terhadap data yang *besar(big data).* CNN(*Convolutional Neural Network*) ialah salah satu metode *machine learning* berasal dari pengembangan *MLP*  (Multi Layer Protocol) di desain untuk pengolah data satu maupun dua dimensi. *Convolutional Neural Network* termasuk kedalam jenis *Deep Neural Network* karena dalamnya tingkat jaringan serta banyak diimplementasikan dalam data citra gambar. *Convolutional Neural Network* memiliki dua metode, yaitu klasifikasi menggunakan *feedforward* serta dapat mempelajari menggunakan *backpropagtion.* Cara kerja *Convolutional Neural Network* ini mempunyai persamaaan dengan *Multi Layer Protocol*, namun pada *Convolutional Neural Network* setiap neuron direpresentasikan dalam bentuk dua dimensi, tidak seperti *Multi Layer Protocol* yang pada setiap neuron hanya berukuran satu dimensi. Disini menggunakan *Convolutional Neural Network* karena lebih cocok dan dari hasil yang telah dicoba hasilnya lebih tinggi tingkat akurasinya dibandingkan algoritma lain [4].

Berdasarkan latar belakang diatas, serta adanya teknologi analisis sentimen diharapkan dapat membantu masyarakat untuk mengukur seberapa besar dampak wabah ini khususnya pada sektor pekerjaan serta menjadikan bahan pertimbangan kepada Pemerintah maupun Intansi untuk mencari alternatif lain selain kebijakan *Work From Home* ini. Penulis berharap dapat mengimplementasikan algoritma CNN pada penelitian ini yang berjudul **“Analisis Sentimen Tentang *Work From Home* Pada Masa Tanggap Darurat Covid-19 Menggunakan Convolutional Neural Network”**.

## Rumusan Masalah

Berdasarkan latar belakang tersebut, dapat dirumuskan permasalahan sebagai berikut :

1. Bagaimana mengelompokkan opini publik mengenai *Work From Home* menjadi sentimen positif atau sentimen negatif ?
2. Bagaimana akurasi algoritma *Convolutional Neural Network* dalam pengklasifikasian opini teresebut ?

## Tujuan Penelitian

Adapun tujuan dari pembangunan sistem klasifikasi opini mengenai kebijakan pubik ini yaitu :

1. Untuk mengetahui respon masyarakat terhadap *Work From Home* yang dikelompokkan menjadi sentimen positif dan sentimen negatif;
2. Untuk mengetahui tingkat akurasi algoritma *Convolutional Neural Network* dalam mengklasifikasian sentimen positif dan sentimen negatif.

## Batasan Masalah

Agar penelitian ini lebih terarah dan sesuai dengan tujuan yang diinginkan maka berikut adalah beberapa batasan masalah dari pembangunan analisis sentimen tentang *work from home* pada masa tanggap darurat covid-19 menggunakan algortima *Convolutional Neural Network*.

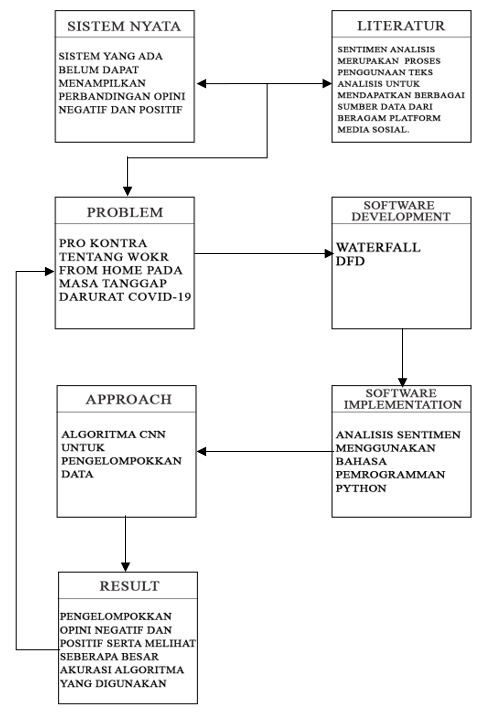
1. Data yang digunakan diambil dari media sosial twitter dengan menggunakan beberapa tagar (#wfh, #covid19, #physicaldistancing, #workfromhome) dan dari kata kunci ( workfromhome, office);
2. Data yang diambil dimulai dari bulan April 2020 hingga Mei 2020;
3. Input dari sistem ini berupa data tweet yang terkait dengan *Work From Home*;
4. *Tweet* yang digunakan hanya *tweet* yang berbahasa Inggris;
5. Pengembangan sistem ini menggunakan bahasa pemrograman *python;*
6. Sistem ini menggunakan algoritma *Convolutional Neural Network.*

## Manfaat Penilitian

Manfaat yang diberikan pada penelitian ini, yaitu:

1. Mengetahui dampak dari wabah *covid-19* salah satunya dari sektor pekerjaan;
2. Mengukur tingkat akurasi algoritma CNN.

## Kerangka Pemikiran

Gambar dibawah menjelaskan tentang kerangka pemikiran berdasarkan dari sistem yang dirancang saat ini, dimana *user* menggunakan media sosial sebagai wadah dalam mengungkapkan argument atau pendapat mereka dalam menanggapi suatu permasalahan sosial yang ada. Ketika wabah *covid-19* ini merebak di mulai pada akhir tahun 2019 hingga saat ini, mengakibatkan segala sektor terhambat salah satunya yaitu pekerjaan. Banyak sekali pengguna media sosial *twitter* yang mengungkapkan opininya mengenai pekerjaan mereka yang terganggu bahkan kehilangannya. Solusi yang dibutuhkan dalam permasalahan ini adalah analisis dari *tweet* pengguna *twitter* untuk mengetahui seberapa besar perbedaan opini dari data yang dihasilkan. Maka dari itu diguakan pendekatan menggunakan *Convolutional Neural Network* yang dapat mengklasifikasikan *tweet* tersebut dengan metode pengembangan *waterfall* yang mana akhir produk ini adalah pola-pola dari analisis sentiment masyarakat terkait *work from home* pada masa tanggap darurat *covid-19.*

Gambar 1. 1 Kerangka Pemikiran

Gambar diatas menjelaskan tentang kerangka pemikiran berdasarkan dari sistem yang dirancang saat ini, dimana *user* menggunakan media sosial sebagai wadah dalam mengungkapkan argument atau pendapat mereka dalam menanggapi suatu permasalahan sosial yang ada. Ketika wabah *covid-19* ini merebak di mulai pada akhir tahun 2019 hingga saat ini, mengakibatkan segala sektor terhambat salah satunya yaitu pekerjaan. Banyak sekali pengguna media sosial *twitter* yang mengungkapkan opininya mengenai pekerjaan mereka yang terganggu bahkan kehilangannya. Solusi yang dibutuhkan dalam permasalahan ini adalah analisis dari *tweet* pengguna *twitter* untuk mengetahui seberapa besar perbedaan opini dari data yang dihasilkan. Maka dari itu diguakan pendekatan menggunakan *Convolutional Neural Network* yang dapat mengklasifikasikan *tweet* tersebut dengan metode pengembangan *waterfall* yang mana akhir produk ini adalah pola-pola dari analisis sentiment masyarakat terkait *work from home* pada masa tanggap darurat *covid-19.*

## Metodologi Penelitian

### 1.7.1 Teknik Pengumpulan Data

Tahap ini dilakukan bertujuan untuk mengumpulkan data yang dibutuhkan dalam penelitian ini dengan menggunakan dua cara yaitu pertama menggunakan *twitterscraper* dan kedua melalui website *Get Tags Explorer*. Terdapat banyak cara untuk melalukan *crawling data,* namun dua cara ini paling efektif. Untuk cara pertama tidak membutuhkan *API Twitter,* sedangkan pada cara kedua membutuhkan *API Twitter* yang aksesnya diperoleh melalui website *Twitter Developer*.

Pengumpulan data dilakukan dengan cara *scraping* dengan memanfaatkan fitur yang ada di masing-masing cara. Perbedaan pada cara pertama dan kedua yaitu jika cara pertama tidak ada batas periode untuk pengambilan data sedangkan cara kedua di batas hanya dapat mengambil data satu minggu sebelumnya.

Lalu untuk mendapatkan akses *API Twitter* kita dapat daftar pada website *developer.twitter.com* lalu ikuti Langkah dan syarat yang tertera, jika kita tepat mengisinya maka hanya butuh waktu kurang lebih satu jam untuk dikonfirmasi dari pihak *twitter*nya. Jika sudah maka akan diberi *Consumer API Keys* dan *Access Token,* untuk *Access Token* hanya dapat dilihat satu kali saja, disarankan untuk menyalinnya. Hasil dari *scraping data* kita disimpan dalam ekstensi file (.csv).

Data yang digunakan pada penelitian adalah data *tweet* pada bulan April 2020 hingga Mei 2020 yang memiliki keterkaitan terhadap *work from home* pada masa tanggap darurat *covid-19.* Berikut ini hal-hal yang berkaitan dengan *inputan* data :

1. Data *tweet* berbahasa Inggris diperoleh dengan cara *scrapping* dengan menggunaan API twitter;
2. Data kata baku Bahasa Inggris berdasarkan model *Word2Vec* yang disediakan oleh *Google* khusus untuk bahasa Inggris.

### 1.7.2 Teknik Pengembangan

Pengembangan sistem ini menggunakan metode *waterfall* atau air terjun*. Waterfall* menggunakan pendekatan alur hidup perangkat lunak yang *sequential* mulai dari tahap Analisa hingga ke proses desain pengkodean dan pengujian.Berikut merupakan tahapan dari metode *waterfall* :

1. Analisa Kebutuhan

Komponen apa saja yang diperlukan system dapat dijabarkan sebagai berikut :

1. Input sistem;
2. Output sistem;
3. Proses yang berjalan dalam sistem.
4. Desain sistem

Tahapan dimana dilakukan penuangan pikiran dan perancangan sistem terhadap solusi dari permasalahan yang ada dengan menggunakan perangkat pemodelan sistem seperti diagram alir data (data flow diagram), diagram hubungan entitas (entity relationship diagram) serta struktur dan bahasan data.

1. Penulisan Kode Program

Penulisan kode program atau coding merupakan penerjemahan design dalam bahasa yang bisa dikenali oleh komputer.

1. Pengujian Program

Tahapan akhir dimana sistem yang baru diuji kemampuan dan keefektifannya sehingga didapatkan kekuranganp dan kelemahan sistem yang kemudian dilakukan pengkajian ulang dan perbaikan terhadap aplikasi menjadi lebih baik dan sempurna.

## Sistematika Penulisan

Sistematika penulisan laporan tugas akhir ini disusun dalam beberapa bab yang masing-masing bab menguraikan beberapa pokok pembahasan. Adapun sistematika penulisan laporan ini yaitu sebagai berikut:

**BAB I PENDAHULUAN**

Bab ini berisikan tentang latar belakang permasalahan yang diambil penulis, perumusan masalah yang dihadapi, batasan masalah, tujuan, state of the art, kerangka pemikiran, metodologi penelitian serta bagaimana sistematika penulisan.

**BAB II LANDASAN TEORI**

Bab ini menjelaskan tentang konsep/teori apa saja yang berkaitan dengan topik yang diangkat, yang telah dibuat berdasarkan hasil penelitian dan hal-hal yang berguna dalam proses penulisan tugas akhir ini.

**BAB III ANALISIS DAN PERANCANGAN**

Pada bagian ini merupakan bagian yang menjadi alur bagaimana penelitian dilakukan, mulai dari cara mengumpulkan data, mengolah atau menganalisis data dan menyimpulkan atau menetapkan simpulan dari sebuah hipotesis semua tercantum di dalam BAB 3.

**BAB IV IMPLEMENTASI DAN PENGUJIAN**

Bab ini menjelaskan tentang pengujian sistem secara umum maupun spesifik. Pengujian sistem secara umum akan membahas mengenai lingkungan uji coba untuk menggunakan sistem ini. Selanjutnya secara lebih spesifik dijelaskan dalam pengujian sistem meliputi skenario pengujian beserta langkah - langkah dalam uji coba sistem untuk mengetahui aplikasi tersebut telah dapat menyelesaikan permasalahan yang dihadapi sesuai dengan yang diharapkan.

**BAB V PENUTUP**

Berisi tentang pernyataan singkat berupa kesimpulan dari pembahasan perangkat lunak yang dibuat secara keseluruhan dan saran untuk mengembangan perangkat lunak yang lebih baik.

BAB II

STUDI PUSTAKA

## Tinjauan Pustaka

Sebelumnya, terdapat beberapa penelitian yang searah dengan penelitian ini baik yang mengangkat tema serupa maupun tidak. Tabel 1 merupakan tabel perbandingan penelitian yang sedang dilakukan dengan penelitian sebelumnya. Berikut ini adalah penelitian yang telah dilakukan sebelumnya dan ada keterkaitan dengan penelitian yang sedang dibahas, antara lain:

1. Maudina Agustin (2017) dalam penilitiannya yang berjudul “PERBANDINGAN KEPUASAN PENGGUNA APLIKASI CHATTING BERDASARKAN ANALISA SENTIMEN DENGAN METODE NAÏVE BAYES”. Sistem yang dibuat untuk membandingkan tingkat kepuasan konsumen terhadap aplikasi chatting, dengan menggunakan algoritma *Naïve Bayes,* hasilnya 400 opini positif dan 500 opini negative serta tingkat akurasi sebesar 68% [5];
2. Wafi Muhamad (2017) dalam penelitiannya yang berjudul “PENINGKATAN HASIL ANALISA SENTIMEN MENGGUNAKAN POS TAGGER UNTUK MELIHAT TANGGAPAN MASYARAKAT TERHADAP FULL DAY SCHOOL”. Sistem yang dibuat untuk melihat respon masyarakat terutama pelajar terhadap kebijakan *Full Day School.* Hasil dari analisis menemukan beberapa kendala terutama pada opini yang tidak seimbang dikarenakan keterbatasan dalam mengambil data, dan rata-rata akurasi sebesar 40% [6];
3. Ahmad Fathan (2014) dalam penelitiannya yang berjudul “ANALISIS SENTIMEN DAN KLASIFIKASI KATEGORI TERHADAP TOKOH PUBLIK PADA TWITTER”. Sistem ini bertujuan untuk klasifikasi *tweet*  berasarkan opini dan kategori menggunakan algoritma *Naïve Bayes Classifier.* Hasil akurasi yang didapatkan sebesar 79,8% [7];
4. Ira Zulfa, Edi Winarko (2017) dalam penelitiannya yang berjudul “Sentimen Analisis Tweet Berbahasa Indonesia dengan Deep Belief Network”, system ini bertujuan untuk melakukan pengklasifikasian terhadap sentimen positif, negatif, dan netral terhadap data uji dan untuk mengetahui akurasi model klasifikasi dengan menggunakan metode Deep Belief Network ketika diaplikasikan pada klasifikasi tweet untuk menandai kelas sentimen data training tweet berbahasa Indonesia [8];
5. Abdan Syakuro (2017) dalam penelitiannya yang berjudul “Analisis Sentimen Masyarakat Terhadap E-Commerce pada Media Sosial menggunakan Metode Naive Bayes Classifier (NBC) dengan Seleksi Fitur Information Gain”, sistem ini bertujuan untuk membuktikan bahwa metode *Naïve Bayes Classifier* dengan seleksi Fitur Information Gain dapat digunakan dalam pengklasifikasian Analisis Sentimen *E-Commerce* [9]*;*
6. Aditya Santoso (2018) dalam penelitiannya yang berjudul “IMPLEMENTASI DEEP LEARNING BERBASIS KERAS UNTUK PENGENALAN WAJAH”, sistem ini bertujuan untuk membentuk suatu model untuk pengenalan citra wajah pada computer [10];
7. Muhamad Gilang (2020) dalam penelitiannya yang berjudul “ANALISIS SENTIMEN TENTANG WORK FROM HOME PADA MASA TANGGAP DARURAT COVID-19”. Sistem ini bertujuan untuk mengetahui respon masyarakat terhadap kebijakan *Work From Home* dan mengetahui seberapa besar tingkat akurasi algoritma yang digunakan pada penilitian ini. Akurasi paling tinggi yang didapat sebesar 87%.

Tabel 2. 1 State of The Art

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Peneliti | Objek Penelitian | Algoritma yang digunakan | Tujuan |
| 1 | Maudina Agustin | Perbandingan Kepuasan Pengguna Aplikasi Chatting. | Naïve Bayes | Untuk mengetahui opini atau sentimen pengguna terhadap masing-masing aplikasi chatting dengan menerapkan metode naive bayes untuk mengklasifikasi komentar-komentar pengguna yang digunakan untuk mencari nilai probabilitas tertinggi. |
| 2 | Wafi Muhamad | Melihat Tanggapan Masyarakat Terhadap Full Day School. | Pos Tagger | Untuk meningkatkan hasil analisa sentimen opini publik tentang full day school. Peningkatan dilakukan dengan menambah proses POS Tagging sebelum masuk ke proses klasifikasi dengan Sentiwordnet. |
| 3 | Ahmad Fathan | Kategori Terhadap Tokoh Publik. | Naïve Bayes | Membangun model untuk melakukan klasifikasi tweet berdasarkan sentimen dan kategori dengan Naive Bayes Classifier. Hasil akurasi pengujian klasifikasi dengan fitur term frequency diperoleh sebesar 79,91% sedangkan fitur TF-IDF didapatkan akurasi sebesar 79,68%. |
| 4 | Ira Zulfa, Edi Winarko | Sentimen Analisis Tweet Berbahasa Indonesia dengan Deep Belief Network | Deef Belief Network | Untuk melakukan pengklasifikasian terhadap sentimen positif, negatif, dan netral terhadap data uji dan untuk mengetahui akurasi model klasifikasi dengan menggunakan metode Deep Belief Network ketika diaplikasikan pada klasifikasi tweet untuk menandai kelas sentiment data training tweet berbahasa Indonesia. |
| 5 | Abdan Syakuro | Analisis Sentimen Masyarakat Terhadap E-Commerce pada Media Sosial menggunakan Metode Naive Bayes Classifier (NBC) dengan Seleksi Fitur Information Gain | Naïve Bayes Classifier | Membuktikan bahwa metode *Naïve Bayes Classifier* dengan seleksi Fitur Information Gain dapat digunakan dalam pengklasifikasian Analisis Sentimen *E-Commerce.* |
| 6 | Aditya Santoso | Implementasi Deep Learning berbasis Keras untuk Pengenalan Wajah | Deep Learning | Membentuk suatu model untuk pengenalan citra wajah pada komputer. |
| 7 | Muhamad Gilang | Analisis tentang Work From Home Pada Masa Tanggap Darurat Covid-19 | Convolutional Neural Network | Mengetahui respon masyarakat terhadap kebijakan Work From Home, apakah kebijakan ini sudah relevan dan dapat berjalan di berbagai sektor pekerjaan*.* |

Oleh karena itu disini peneliti menggunakan algoritma yang berbeda dan ingin mengetahui seberapa besar tingkat akurasi yang didapatkannya nanti.

## Landasan Teori

### Analisis Sentimen

Analisis sentimen (juga dikenal sebagai penambangan opini atau emosi AI) mengacu pada penggunaan pemrosesan bahasa alami, analisis teks, linguistik komputasi, dan biometrik untuk secara sistematis mengidentifkasi, mengekstraksi, mengukur, dan mempelajari status afektif dan informasik subjektif. Tugas dasar dalam analisis sentiment adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, atau kalimat, apakah bersifat positif atau negatif [11].

Tujuan dari Analisa sentiment adalah untuk menentukan perilaku atau opini dari seorang penulis dengan memperhatikan suatu topik tertentu. Perilaku bisa mengindikasikan alasan, opini atau penilaian, kondisi kecenderungan (bagaimana si penulis ingin mempengaruhi pembaca).

### Twitter

Twitter merupakan sebuah alat jaringan sosial yang dapat digunakan secara gratis yang dapat memungkinkan orang untuk berbagi ataupun bertukar informasi secara *real-time.* Para penggunanya dapat bertukar informasi dalam bentuk teks, video, gambar maupun halaman website. Namun untuk teks dibatas hanya dapat 140 karakter. Twitter dirilis pada November 2010, awalnya hanya sebagai media bertukar informasi namun saat ini twitter menjadi berkembang tidak hanya untuk info semata. Terkadang berita paling hangat yang sedang muncul awalnya terdapat pada *twitter*. Twitter memungkinkan pertukaran ide bebas secara nasional dan global, antara orang yang tertarik pada bidang keahlian serupa, serta debat kritis. *Tweet* yang muncul sangat beragam mulai dari kategori politik, olahraga, movie/film, otomotif, teknologi, kesehatan, lifestyle dan berbagai topik menarik lainnya [12].

### Python

Dibuat oleh Guide Van Rossum dan pertama kali dirilis pada 1991, filosopi desain *python* menekankan pembacaan kode dengan penggunaannya yang penting dari ruang signifikan. Bahasa kontruksi dan pendekatan berorientasi objek bertujuan untuk membantu programmer menulis jelas, kode logis untuk proyek kecil maupun besar. *Python* itu dinamis, program ini mendukung beberapa paradigma pemrograman, termasuk terstruktur, berorientasi objek, dan pemrograman fungsional [13].

### Klasifikasi

Klasifikasi merupakan salah satu metode yang sering digunakan dalam *data mining* selain *clustering* dan asosiasi*,* terutama terhadap penerapan prediksi. Klasifikasi adalah sebuah metode pengelompokan data dengan mempelajari data latih dengan dengan menggunakan *algoritma* pengklasifikasian[14]. Dalam klasifikasi, terdapat beberapa target kategori variabel, contohnya penggololongan gaji atau pendapatan, dipisahkan dalam tiga kategori, yaitu pendapatan rendah, pendapatan sedang dan pendapatan tinggi [15].

### Deep Learning

Pembelajaran mendalam memungkinkan model komputasi yang terdiri dari beberapa lapisan pengolahan untuk mempelajari representasi data dengan beberapa tingkat abstraksi. Metode ini telah secara dramatis meningkatkan *State-of-The-Art* dalam pengenalan ucapan, pengakuan objek visual, objek Deteksi dan banyak domain lainnya seperti penemuan obat dan Genomics. Pembelajaran mendalam menemukan struktur rumit dalam kumpulan data besar dengan menggunakan algoritma backpropagasi untuk menunjukkan bagaimana mesin harus mengubah parameter internal yang digunakan untuk menghitung representasi di setiap lapisan dari representasi di lapisan sebelumnya. *Deep Convolutional Nets* telah membawa terobosan dalam pengolahan gambar, video, pidato dan audio, sedangkan jaring berulang memiliki cahaya bersinar pada data berurutan seperti teks dan ucapan [16].

### Keras

Keras adalah *API deep learning* yang ditulis dengan *python,* berjalan diatas *platform machine learning TensorFlow*. Keras mengikuti praktik terbaik untuk mengurangi muatan kognitif, ia menawarkan *API* yang konsisten dan sederhana, meminimalkan jumlah Tindakan penggunaan umum serta memberikan pesan kesalahan yang jelas dan dapat ditindak lanjuti [17].

### *Tensorflow*

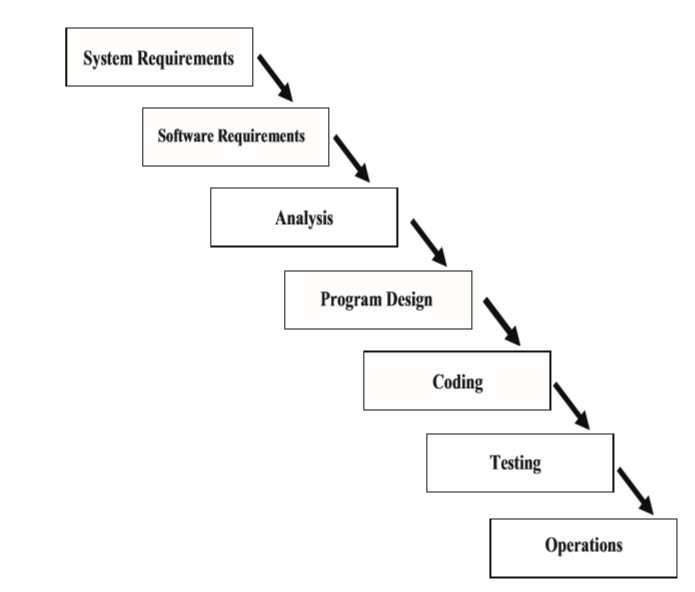
*Tensorflow* yaitu perpustakaan bagi perangkat lunak, pengembangannya dilakukan oleh tim *Google Brain* dalam divisi penelitian *Machine Learning Google,* bertujuan untuk melakukan penelitian jaringan syaraf dalam dan *Machine Learning*.Fitur – fitur didalamnya meliputi:

* + 1. Mengoptimalkan, mendefinisikan, dan menghitung secara efisien ekspresi matematis yang melibatkan *Array Multidimension;*
    2. Pemrograman pendukung jaringan syaraf dalam dan teknik *Machine Learning.*
    3. Penggunaan *Graphics Processing Unit* (GPU) yang transparan, dapat mengoptimalisasi memori yang sama dan mengotomatisasi manajemen data yang digunakan. *Tensorflow* bisa menulis kode yang sama dan menjalankannya baik di *Central Processing Unit* (CPU) atau *Graphics Processing Unit* (GPU). Lebih khususnya lagi, *Tensorflow* akan mengetahui bagian perhitungan yang harus dipindahkan ke *Graphics Processing Unit* (GPU).
    4. *Scalablity* komputasi yang tinggi di seluruh mesin serta kumpulan data yang besar [18].

### Algoritma *Convolutional Neural Network* (CNN)

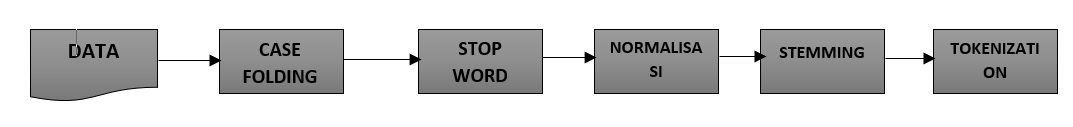
Convolutioal Neural Network adalah salah satu dari algoritma *neural network* yang pada umumnya digunakan dalam proses pengolahan data *image* atau gambar, dan ada pula digunakan pada teks. Convolutional neural network (CNN), kelas jaringan saraf tiruan yang telah menjadi dominan tugas visi computer. CNN dirancang untuk secara otomatis dan beradaptasi belajar hirarki spasial fitur melalui backpropagation dengan menggunakan beberapa blok bangunan, seperti lapisan lilitan, penggabungan lapisan, dan lapisan sepenuhnya terhubung. Algoritma CNN ini mempunyai beberapa layer yang digunakan untuk melakukan tahap filtering pada setiap prosesnya. Proses disini dapat disebut sebagai proses training dimana tahapannya adalah *Convolutional Layer, Pooling Layer,* dan *Fully Connected Layer* [19]*.*

### Model Pengembangan *Waterfall*

Model *Waterfall* adalah model klasik yang digunakan dalam siklus hidup pengembangan sistem untuk menciptakan sistem dengan pendekatan linier dan berurutan. Hal ini disebut sebagai air terjun karena model berkembang secara sistematis dari satu fase ke yang lain dalam mode ke bawah. Model ini dibagi menjadi fase yang berbeda dan output dari satu fase digunakan sebagai input dari fase berikutnya. Setiap fase harus diselesaikan sebelum fase berikutnya dimulai dan tidak ada tumpang tindih dari fase [20].

Gambar 2. 1 Model Waterfall

### *Preprocessing*

*Preprocessing* adalah suatu proses yang dimana fungsinya untuk membersihkan data dari *noise,* menyemakan bentuk kata serta mengurangi daya muat kata. Fase yang dilakukan dari data *preprocessing* dapat dilihat pada gambar berikut :

Gambar 2. 2 Proses Preprocessing Data

Pada fase *preprocessing* terdiri dari *Case folding, Stopword, Normalisasi, Stemming,* dan *Tokenization.*

* + 1. *Case folding* adalah proses dimana merubah huruf kapital menjadi hurup kecil;
    2. *Stopword* adalah proses menghilangkan kata yang sering muncul atau kata yang maknanya sedikit seperti ‘at’ dan yang lainnya;
    3. *Normalisasi* adalah proses pendefinisian singkatan seperti ‘ur’ menjadi ‘your’;
    4. *Stemming* adalah proses menghilangkan kata imbuhan [21].

### Metode *Word2Vec*

*Word2Vec* adalah algoritma *word embedding,* yaitu memetakan kata menjadi sebuah vector. Nantinya vector ini digunakan untuk beberapa macam tugas *NLP (Natural Language Processing).* Dalam mempresentasikan sebuah kata, *Word2Vec* mengimplementasi atau menerapkan neural network utuk menghitung kontekstual dan kesamaan semantik ini dapat merepresentasikan relasi sebuah kata dengan kata lainnya, contohnya kata ‘Sepakbola’ dekat kepada kata ‘Stadion’;

### Data Flow Diagram (DFD)

DFD adalah diagram aliran data digunakan untuk merepresentasikan aliran data dalam sistem informasi bisnis secara grafis. DFD menjelaskan proses yang terlibat dalam sistem untuk mentransfer data dari input ke penyimpanan file dan laporan generasi.

Diagram aliran data dapat dibagi menjadi logis dan fisik. Diagram aliran data Logis menjelaskan aliran data melalui sistem untuk melakukan fungsi tertentu dari suatu bisnis. Diagram aliran data fisik menjelaskan penerapan aliran data logis [22]. Diagram aliran data dipopulerkan pada akhir 1970-an, yang timbul dari buku Structured Design, oleh para perintis komputasi Ed Yourdon dan Larry Constantine. Mereka mendasarkan pada "*data flow graph*" model komputasi oleh David Martin dan Gerald Estrin. Konsep desain terstruktur lepas landas di bidang Rekayasa perangkat lunak, dan metode DFD melepas dengan itu. Hal ini menjadi lebih populer di kalangan bisnis, seperti yang diterapkan pada analisis bisnis, daripada di kalangan akademik [23]. Dengan cara menggambarkan setiap proses yang berjalan pada sistem yang akan di rancang dan dibangun maka akan meringankan atau mempermudah dalam pengembangan sistem itu sendiri. Beritkut ini merupakan notasi-notasi DFD :

Tabel 2. 2 DFD Symbol

|  |  |
| --- | --- |
| ***Symbol*** | ***Element Name*** |
|  | Proses |
|  | Data Flow |
|  | Data Store |
|  | External Entity |

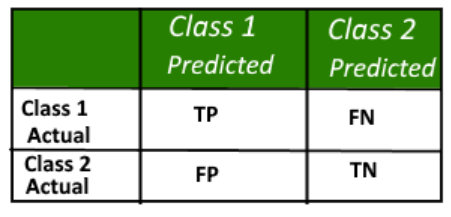
Tahapan-tahapan perancangan dengan menggunakan DFD:

1. DFD Level 0 disebut *Context Diagram* yaitu menggambarkan entitas tunggal yang berinteraksi dengan orang maupun sistem lain. DFD menggambarkan interaksi antara sistem yang akan dikembangkan dengan entitas luar;
2. DFD Level 1 menggambarkan modul yang ada didalam sistem yang akan dikembangkan. DFD Level 1 merupakan hasil *breakdown* DFD Level 0 yang sudah dibuat;
3. DFD Level 2 merupakan hasil *breakdown* DFD dari proses di level 1.

### *Confusion Matrix*

Dalam bidang pembelajaran mesin dan khususnya masalah klasifikasi statistik, *confusion matrix*, juga dikenal sebagai matriks kesalahan. *Confusion matrix* adalah tabel yang sering digunakan untuk menggambarkan kinerja model klasifikasi (atau "*classifier*") pada satu set data pengujian yang nilai sebenarnya diketahui. Hal ini memungkinkan visualisasi kinerja algoritma. Hal ini memungkinkan identifikasi mudah kebingungan antara kelas misalnya satu kelas umumnya salah label sebagai yang lain. Sebagian besar tindakan kinerja dihitung dari *confusion matrix*.

Ketika mengklasifikasi sebuah data perlu untuk mengetahui bagaimana caranya agar data dapat di klasifikasikan scara tepat dan benar, maka perlu sebuah pengukuran didalamnya. *Confusion Matrix* adalah ringkasan hasil prediksi pada masalah klasifikasi. Jumlah prediksi yang benar dan salah diringkas dengan nilai menghitung dan dipecah oleh setiap kelas. Ini adalah kunci untuk *Confusion Matrix*. *Confusion Matrix* menunjukkan cara-cara di mana model klasifikasi Anda bingung ketika membuat prediksi. Ini memberi kita wawasan tidak hanya ke dalam kesalahan yang dibuat oleh classifier tetapi yang lebih penting jenis kesalahan yang sedang dibuat [24].



Gambar 2. 3 Confusion Matrix

Keterangan ,

* Class 1 : Positif;
* Class 2 : Negatif.

Definisi syarat:

* *Positive* (P) : pengamatan positif (contoh: ini adalah benda mati);
* *Negative* (N) : pengamatan tidak positif (contoh: ini bukan bendamati);
* *True positive* (TP) : pengamatan positif, dan diprediksi akan positif;
* *False negative* (FN) : pengamatan positif, tetapi diprediksi negative;
* *True negative* (TN) : pengamatan negatif, dan diprediksi akan negative;
* *False positive* (FP) : pengamatan negatif, tetapi diprediksi positif.

**Tingkat klasifikasi/akurasi:**

Tingkat klasifikasi atau akurasi yang diberikan oleh relasi:

Akurasi adalah sebuah persentasi pada ketepatan dalam proses klasifikasi.

**Recall:***Recall* adalah tingkat keberhasilan pada sistem dalam mendapatkan kembali suatu informasi. Persamaannya seperti dibawah ini **(2.2)**.

*Recall* dapat didefinisikan sebagai rasio dari jumlah total yang benar diklasifikasikan contoh positif membagi jumlah total contoh positif. Tinggi ingat menunjukkan kelas dikenali dengan benar (sejumlah kecil FN).

**Precision:**

**Untuk mendapatkan nilai presisi kita membagi jumlah total dengan benar diklasifikasikan contoh positif dengan jumlah total prediksi contoh positif. Presisi tinggi menunjukkan contoh yang berlabel positif memang positif (sejumlah kecil FP).**

***High Recall*, presisi rendah: ini berarti bahwa sebagian besar contoh positif dikenali dengan benar (FN rendah) tetapi ada banyak positif palsu.**

***Law Recall* , presisi tinggi: ini menunjukkan bahwa kita melewatkan banyak contoh positif (tinggi FN) tetapi mereka kita memprediksi sebagai positif memang positif (FP rendah)**

***F-Measure*:**

**Karena kita memiliki dua langkah (presisi dan recall) membantu untuk memiliki pengukuran yang mewakili keduanya. Kita menghitung F-ukuran yang menggunakan *mean Harmonic* di tempat mean aritmatika seperti itu menghukum nilai ekstrim lebih. *F-Measure* akan selalu lebih dekat ke nilai yang lebih kecil dari Precision atau Recall. Dapat dilihat dari persamaan berikut (2.3).**

BAB III

ANALISIS DAN PERANCANGAN

Bab ini menjelaskan tentang tahapan implementasi dari metode yang digunakan dalam pengembangan analisis sentimen tentang *Work From Home* pada masa tanggap darurat *covid-19*.

## **3.1** System Requirment

### **3.1.1** Kebutuhan Fungsional

* + 1. Sistem dapat membaca dataset;
    2. Sistem dapat melakukan preproses data;
    3. System dapat menampilkan kata hasil preproses;
    4. Sistem dapat membuat file baru hasil preproses;
    5. Sistem dapat melakukan pembobotan dengan model *Word2Vec*;
    6. Sistem dapat melakukan proses pembelajaran (*training);*
    7. Sistem dapat melakukan proses pengujian (*testing);*
    8. Sistem dapat menghitung *precission*, *recall, f1-score* dan *support;*
    9. Sistem dapat melakukan proses *scoring* dan *evaluation.*

### 3.1.2 Kebutuhan Non Fungsional

1. System dikembangkan menggunakan Bahasa pemrograman Python;
2. Data disimpan pada file dengan format .csv dan .xlsx;
3. Data yang digunakan harus memiliki sumber pengetahuan dengan dilabelkan *positive*, dan *negative*.

## 3.2 Software Requirment



### Hardware

Kebutuhan minimal perangkat keras (*hardware*) untuk menjelaskan aplikasi ini adalah sebagai berikut:

1. *Processors* :Intel Dual-Core
2. *Harddisk* :10 GB
3. RAM : 4 GB

### Software

Adapun kebutuhan minimal perangkat lunak (*software*) yang diperlukan diantaranya:

1. Sistem Operasi : Windows 7 or *higher*, macOS dan Linux;
2. Versi Python : 3.X;
3. *Tools*  : Jupyter Notebook, conda, conda-env;
4. Perangkat : Pycharm, Atom;
5. *Python Packages* : NumPy, WordCloud, Matplotlib.pyplot, Stopwords, Stemming, Keras, Pandas, Nltk, Gensim, Tokenizer, Image Color Generator.

## Analisis

### 3.3.1 Analisis Masalah

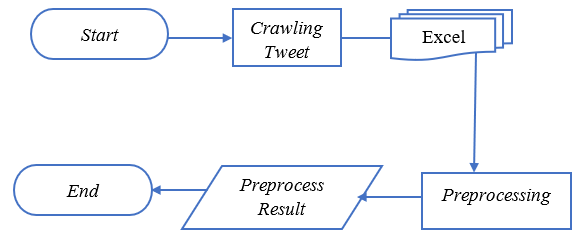
Berdasarkan penelitian yang dilakukan, para pengguna atau *user* media sosial media, mengungkapkan argument atau opini mereka melalui media sosial. Hal ini mengakibatkan sebuah isu yang sedang hangat timbul perdebatan antara pro dan kontra.

Oleh karena itu untuk mengatasi dan mengetahui permasalahan apa saja yang terjadi terutama isu-isu seputar dampak wabah *covid-19* di seluruh dunia. Maka diajukanlah sistem yang dapat menganalisis berbagai statemen masyarakat. Selain itu site mini dapat menghitung akurasi dari sistem tersebut.

### 3.3.2 Analisis Data

Data yang digunakan adalah *tweet* yang terkait dengan istilah *work from home* atau kerja dari rumah di Seluruh dunia dengan kata kunci *WFH* dan *work from home.* Pengambilan sampel data tweet dimulai dari bulan Maret 2020 ketika wabah *covid19* ini masuk dan menyebar di Indonesia sampai dengan bulan Mei 2020.

Sebelum data digunakan dilakukan pengolahan terlebih dahulu dengan proses *preprocessing* sehingga menghasilkan dataset yang bersih dan terbebas dari *noise* sehingga siap digunakan untuk kemudian dilakukan pembobotan kata*.*

*Dataset* yang telah melalui proses embedding selanjutnya akan dimasukan kedalam system untuk dilakukan proses *training* dan *testing* tepisah sehingga menghasilkan akurasi dari proses klasifikasi yang dilakukan oleh algoritma *CNN.* Adapun gambaran dari proses ini adalah sebagai berikut:

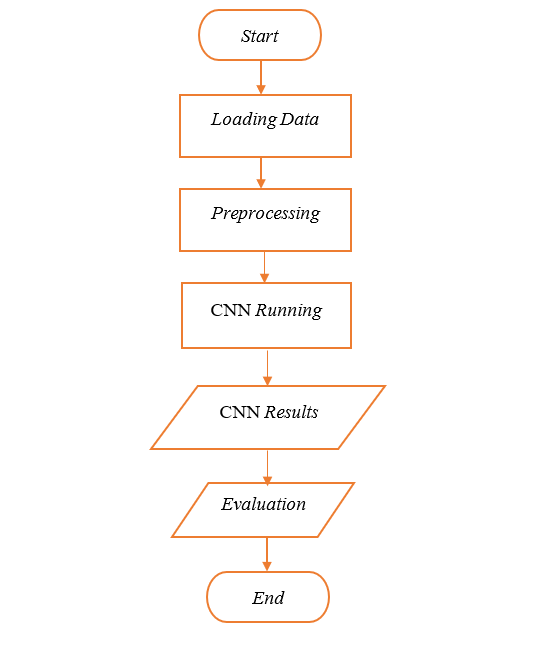
Gambar 3. 1 Data Preparation Process

### 3.3.3 Analisis Algoritma

Algoritma yang digunakan dalam penelitian ini adalah Convolutional Neural Network dengan menggunakan metode pendekatan Supervised. Data yang digunakan adalah data yang sudah dipersiapkan terlebih dahulu dan tidak menggunakan data eksternal. Data yang ada akan di acak untuk dijadikan data test dan data training .



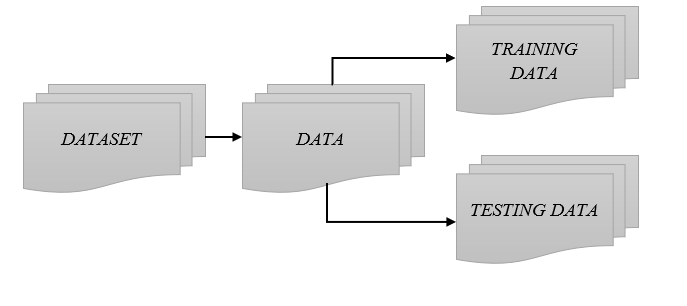
### *Flowchart* Algoritma *Convolutional Neural Network*



Gambar 3. 2 Flowchart Algoritma CNN

Gambar diatas menjelaskan alur proses algoritma *CNN* berjalan, proses *Loading* ini yaitu memasukan *crawling data* yang sudah dilabel, lalu masuk ketahap *preprocessing* dimana data yang sudah ada akan di *cleaning* dari berbagai simbol dan karakter diluar konteks opini. Selanjutnya akan di analisis oleh algoritma *CNN* yang hasilnya berupa grafik dari opini. Hasil akurasi akan di evaluasi kembali agar mendapatkan angka yang terbaik.

### Pemisahan Data

Setelah data diberikan label berupa negatif dan positif, kemudian dibagi menjadi 2 bagian yaitu data *train* dan data *testing* pada masing masing dataset, seperti yang kita dapat lihat pada gambar 3.2*.*

Gambar 3. 3 Data Structure

Data latih adalah data yang akan digunakan sebagai data *training* terhadap sistem yang nantinya akan menghasilkan model. Sedangkan data uji adalah data yang akan digunakan sebagai data *testing* pada sistem, lalu hasilnya akan diketahui tingkat keakurasian dari proses klasifikasi dengan algoritma *Convolutional Neural Network* (CNN)*.*

### *Preprocessing*

Data yang sudah dibagi atau dipisahkan menjadi dua bagian yaitu data *testing* dan data *train* kemudian lanjut ke pada tahap *preprocessing. Preprocessing* yaitu tahap pemrosesan data yang tidak terstruktur menjadi data yang *terstruktur* sehinga nanntinya data dapat diolah dengan lebih efisien. Dalam tahap *preprocessing* ini terdapat tahapan-tahapan untuk menghasilkan data yang sempurna, yaitu:

Tabel 3. 1 Data Set Labeling

|  |  |  |
| --- | --- | --- |
| No | *Content* | *Label* |
| 1 | Working from home is actually more stressful than work. Bosses and clients are on you ANYTIME THEY WANT! | 0 |
| 2 | work From home enjoy !!!! | 1 |

#### **3.3.6.1** Casefolding

Data yang di dapat dari proses *Scraping* langsung dari media sosial *twitter* beragam bentuk penulisan opininya, ada yang menggunakan huruf kapital adapula yang huruf kecil. Oleh karena itu untuk menghasilkan data yang lebih baik dilakukannya proses *Case Folding* ini sendiri merupakan sebuah proses dimana mengubah kalimat yang didalamnya mengandung hurup besar atau *Uppercase* menjadi huruf kecil atau *Lowercase.* Seperti yang dapat kita lihat pada Table di bawah ini:

Tabel 3. 2 Case Folding

|  |  |  |
| --- | --- | --- |
| No | **Before *Case Folding*** | **After *Case Folding*** |
| 1 | Working from home is actually more stressful than work. Bosses and clients are on you ANYTIME THEY WANT! | working from home is actually more stressful than work. bosses and clients are on you anytime they want! |
| 2 | work From home enjoy !!!! | work form home enjoy !!!! |

#### **3.3.6.2** *Cleaning*

*Cleaning* yaitu suatu proses pembersihan data *delimeter* seperti ($, #, @, titik, *url,* koma, angka dan tanda baca lainnya), symbol tersebut akan terhapus dan akan tersisa kata kata saja. Proses penghapusan data yang tidak konsisten dan kotor berfungsi agar data menjadi bersih dan terbebas dari noise dalam proses klasifikasi.

Tabel 3. 3 Cleaning Data

|  |  |  |
| --- | --- | --- |
| No | *Before Cleaning Data* | *After Cleaning Data* |
| 1 | @johnedavid No too bad, thanks. Yes, WFH - prefer it to be honest. Music is better! Will request to carry on even if we return to 'normal', as my sister is still having chemo and we are in contact occasionally, sorting out the aftermath from our mum. Hope all is swell up your end? | no too bad thanks yes wfh prefer it to be honest music is better will request to carry on even if we return to normal as my sister is still having chemo and we are in contact occasionally sorting out the aftermath from our mum hope all is swell up your end |
| 2 | Is working remotely bad for a persons career? I THINK YES | is working remotely bad for a persons career i think yes |

#### **3.3.6.3** *Stemming*

*Stemming* adalah merubah kata ke bentuk dasar atau awalnya. Contohnya *driving* menjadi drive.

Tabel 3. 4 Stemming

|  |  |  |
| --- | --- | --- |
| No | *Before Stemming* | *After Stemming* |
| 1 | I hate working from home’s | i hate work from home |
| 2 | I’am Stressed because work from home !! | I’am stress because work from home |

#### **3.3.6.4** *Stop Words*

*Stop Words* adalah fungsi untuk menghilangkan kata yang sedikit maknanya atau dengan kata lain kata imbuhan, seperti *to, in and etc.*

Tabel 3. 5 Stop Words

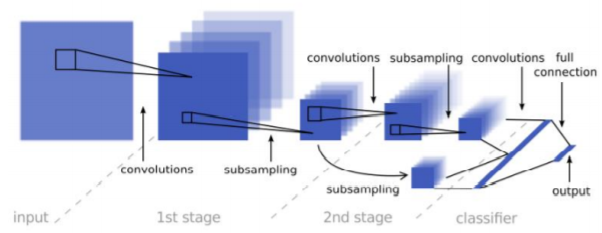
|  |  |  |
| --- | --- | --- |
| No | *Before Stop Words* | *After Stop Words* |
| 1 | A Profession post #Covid\_19 would be in big demand is "ENT"  Since most of the WFH people would have earing problems after attending concalls ; Eye problem after starring at the devices ( Mobile, Laptop ; TV) | Profess post covid would big demand ent sinc wfh people would ear problem attend concal eye problem star devic mobile laptop tv |
| 2 | I can not working at home, help me !! | i can not work home help me |

### Pembobotan *Word2Vec*

Salah satu *word vector* representasi yang dibuat oleh *google* yang bernama *Word2Vec.* Dengan *Word2Vec* ini kita dapat mengukur vektor sebagai perbandingan. Jika kita mengukur jarak antara vektor dari kata “Bola” dan “Olahraga” atau “Bandung” dan “Jawa Barat”, maka kita akan menemukan bahwa jaraknya akan muncul dengan angkat yang berdekatan. Hal ini disebabkan oleh diantara kata diatas merupakan sesuatu hal yang berdekatan, maka nilai vector nya pun berdekatan.

### Analisis Algoritma *CNN*

*Convolutional Neural Network (CNN)* ini dapat dianalogikan sebagai kelanjutan untuk model jaringan saraf tradisional, seperti *Multilayer Perceptron (MLP).* Sebuah model arsitektur *Convolutional Neural Network (CNN)* terdiri dari lapisan khusus untuk fitur ekstrak dari ruang *input raw* dan model jaringan saraf yang sepenuhnya terhubung dengan logistic pengklasifikasi regresi. Fitur yang biasanya disebut sebagai peta, fitur diperoleh dari lapisan khusus tersebut dan kemudian menjadi input untuk sepenuhnya terhubung jaringan sarap yang sebenarnya model *Multilayer Perceptron (MLP).* Gambar dibawah menunjukkan contoh dari arsitektur model *Convolutional Neural Network* [25]*.*

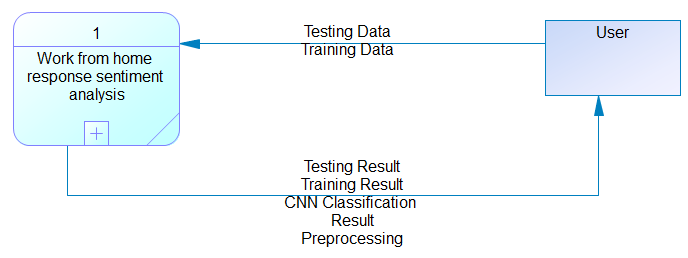
Gambar 3. 4 Arsitektur Model Algoritma Convolutional Neural Network

Dari gambar ini, dapat dilihat bahwa sebelum diproses dalam dua tahap, kedua tahapan ini terdiri dari konvolusi dan subsampling operasi. Hal ini juga dapat dilihat bahwa baik konvolusi dan operasi subsampling akan mengurangi dimensi input. Sebuah operasi lilitan akan mengubah satu input dua dimensi matriks menjadi beberapa matriks dua dimensi yang lebih kecil. Pada proses akhir, fungsi dilakukan untuk meratakan atau mengubah setiap peta fitur dari dua dimensi matriks ke matriks satu dimensi sehingga peta fitur siap diklasifikasikan dengan *Multilayer Perceptron* atau jaringan saraf yang terhubung sepenuhnya. Tujuan utama dari konvolusi dan operasi *subsampling* adalah untuk mengekstrak fitur dari data input mentah. Konvolusi beroperasi merubah menjadi multiplikasi kernel kecil dan area tertenu dari input dua dimensi matriks bekerja [25].

## 3.4 Desain

*Context Diagram* merupakan pemodelan dari sistem – sistem yang akan dibuat. Pada pemodelan yang ada dibawah hanya terdapat satu entitas yaitu entitas user, yang memiliki peran dalam menjalankan klasifikasi dalam proses analisis sentimen respon masyarakat terhadap *work from home* pada media sosial Twitter*.*

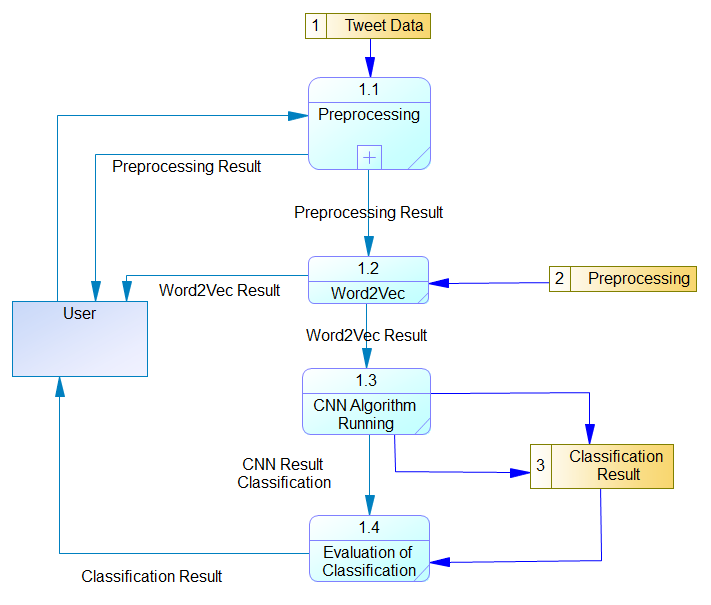
### 3.4.1 Context Diagram



Gambar 3. 5 Context Diagram

Dapat kita lihat pada gambar diatas yang merupakan pemodelan sistem secara umum yang kini sedang dirancang. Pada pemodelan diatas hanya terdapat satu entitas yaitu *user* yang akan mengatur dan menjalankan sistem klasifikasi ini.

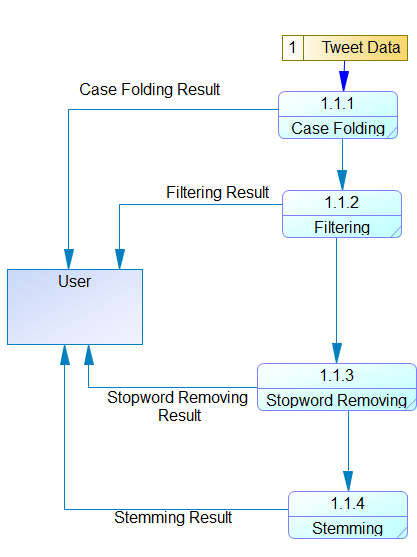
### 3.4.2 Data Flow Diagram Level 1



Gambar 3. 6 Data Flow Diagram Level 1

*Data Flow Diagram* level 1 ini memiliki empat proses utama yaitu *preprocessing, Word2Vec,* algoritma *CNN* dan yang terakhir merupakan suatu proses *evaluating* dari algoritma *CNN* yang akan digunakan sebagai hasil dari analisis sentimen lalu di tampilkan dalam berbentuk grafik.

### 3.4.3 Data Flow Diagram Level 2

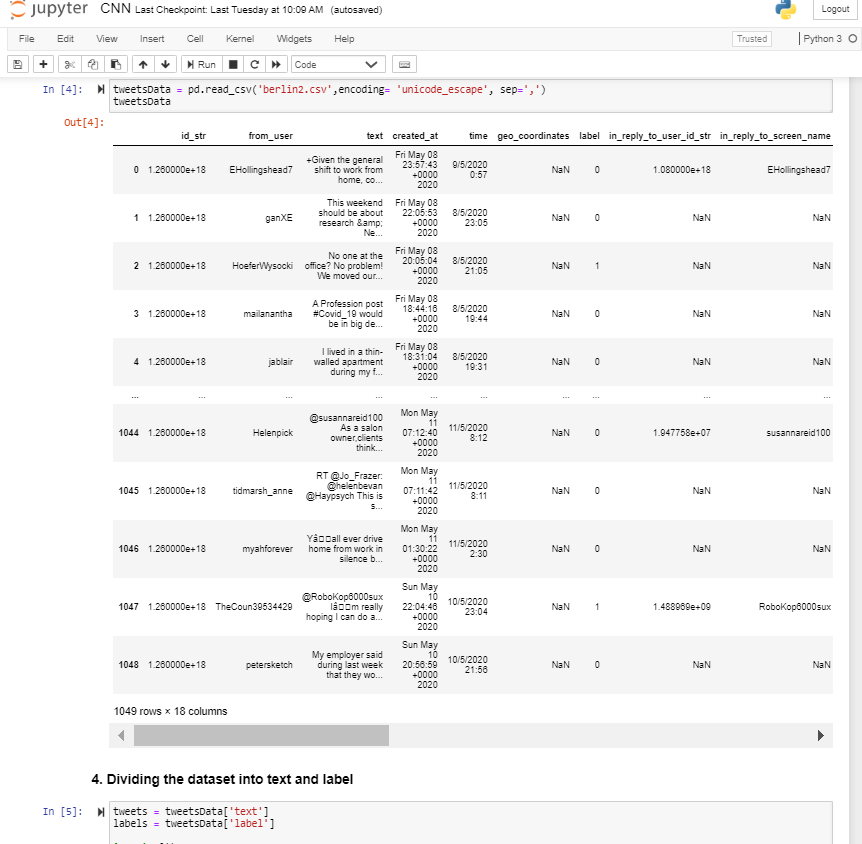


Gambar 3. 7 Data Flow Diagram Level 2

Pada level ini merupakan turunan *filtering* dari proses *prepocessing*. Proses yang ada pada *prepocessing* yang pertama adalah *case folding* yaitu merubah semua huruf *Uppercase* menjadi huruf kecil atau *lowercase*, selanjutnya setelah merubah ke bentuk huruf kecil lalu data yang mengandung simbol atau karakter aneh harus dihapus seperti @,#, angka dan lain sebagainya proses ini merupakan proses *filtering.* Setelah proses *filtering* adalah proses *stopwords* yaitu membuang kata-kata yang tidak memiliki atau sedikit maknanya seperti *at, in, a* dan lain sebagainya. Selanjutnya yang terakhir *stemming* yaitu proses menghilangkan kata imbuhan atau mengembalikan kata ke kata dasarnya seperti *working* menjadi *work.*

## 3.5 Perancangan User Interface

Perancangan *User Interface* atau lebih dikenal dengan perancangan antar muka, antara pengguna dan aplikasi pada penelitian ini adalah dengan aplikasi *Jupyter Notebook* sebagai media untuk menampilkan hasil analisis sentimen dari data yang sudah diolah. Pada perancangan *user interface* kita dapat mengetahui hasil analisis dari data yang sudah di olah sebelumnya, dan akan ditampilkan dalam berbagai macam bentuk (vector, teks, dan grafik).

Berikut merupakan contoh gambar, output hasil yang akan ditampilkan pada *Jupyter Notebook.*

Gambar 3. 8 User Interface Jupyter Notebook

BAB IV

IMPLEMENTASI DAN PENGUJIAN

## 4.1 Implementasi Sistem

Tahapan implementasi merupakan suatu proses yang dimana menerapkan suatu sistem perancangan yang sebelumnya sudah pernah dibangun. Pada penelitian ini perancangan pada bab yang sebelumnya diterapkan kedalam sebuah aplikasi seperti yang dapat kita lihat dibawah ini.

### 4.1.1 *Hardware*

*Hardware* yang digunakan dalam membangun system ini memiliki spesifikasi sebagai berikut :

1. *Processor* : AMD Ryzen 5 2.1GHz 4core
2. RAM : 8.00 GB
3. SSD M2 : 256 GB
4. Tipe Sistem : 64-Bit

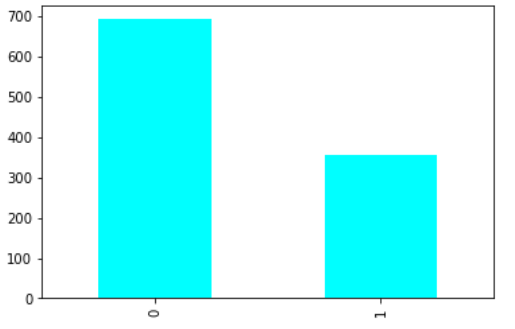
### 4.1.2 *Software*

Adapula perangkat lunak yang diperlukan dalam membangun aplikasi ini adalah sebagai berikut:

1. Sistem Operasi : Windows 10 Home Edition
2. Bahasa Pemprograman : Python versi 3.6
3. Pemodelan Sistem : Power Designer
4. *Tools* : Jupyter Notebook

## 4.2 Pengujian

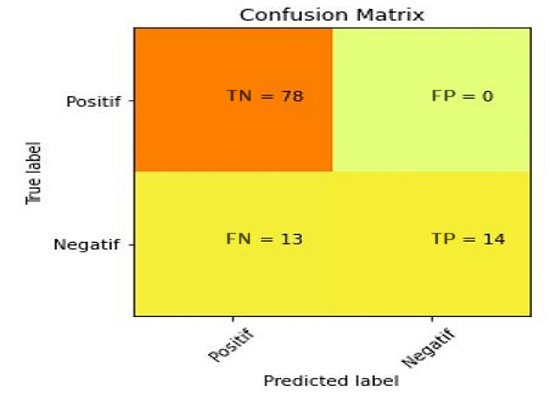
Pengujian dilakukan terhadap algoritma *Convolutional Neural Network (CNN)* sebanyak enam kali pengujian dengan membagi-bagi data kedalam *data training* dan *data testing* serta membaginya kedalam 90% *data training* 10% *data testing*, kemudian80% *data training* 20% *data testing*, kemudian 70% *data training* 30% *data testing*, kemudian 60% *data training* 40% *data*, kemudian 50% *data training* 50% *data testing*, dan yang terakhir 40% *data training* 60% *data testing,* semua pengujian menggunakan pengulangan iterasi sebanyak 5 kali.

Jadi terdapat 1049 dataset yang digunakan untuk penelitian ini terhadap respon masyarakat pada kebijakan *work from home* di media sosial *twitter.* Berikut adalah persentase dari data yang digunakan dalam proses analisis ini.

Gambar 4. 1 Persentase Data Negatif dan Positif

### 4.2.1 Pengujian Pertama

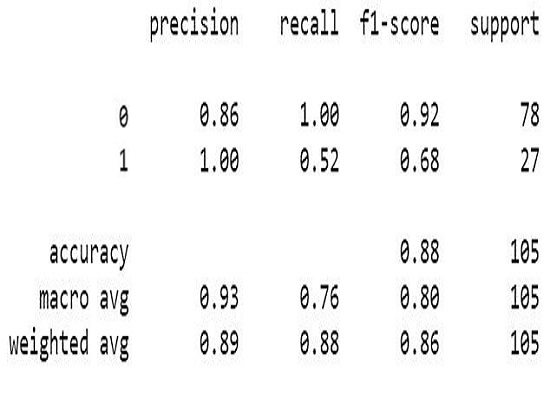
Pengujian pertama dilakukan jumlah *data training* 90%, *data testing* 10%dan *epoch* 5*.*



Gambar 4. 2 Confusion Matrix Pengujian Pertama

Dari hasil pengujian pengujian yang ditunjukan dengan gambar diatas, sehingga dengan menggunakan persamaan (2.1) didapatkan :

Kemudian menghitung *recall* pada tiap – tiap kelas yang di prediksi. Hasil perhitungan dari persamaan (2.2) dan disubstitusikan dalam perhitungan dibawah ini :

 Perhitungan selanjutnya yaitu perhitungan *F1-Score* atau ­*F-Measure* yang dimerupakan salah satu perhitungan evaluasi klasifikasi yang mengkombinasikan *recall* dan *precission.* Nilai *F1-Score* dapat diketahui dengan menggunakan persamaan (2.3) bab sebelumnya.

Sehingga dapat dilihat grafik pada gambar dibawah yang dihasilkan dari perhitungan dari persamaan (2.1), (2.2) dan (2.3) dihasilkan grafik seperti berikut:

Gambar 4. 3 Grafik F1 Score Pengujian Pertama

Dari grafik diatas dihasilkan *precission* dari positif sebesar 1.00 dan negatif sebesar 0.86. Sehingga nilai rata – rata dari presisinya yaitu *avg* = (1.00 + 0.86) / 2 = 0.93. Nilai *precission* didapatkan dengan cara menggunakan persamaan (2.1).  
 Kemudian hasil *recall* dari label negatif sebesar 1.00 dan positif 0.52. Sehingga nilai rata – rata dari *recall* yaitu: *recall* = (1.00 + 0.52) / 2 = 0.76.

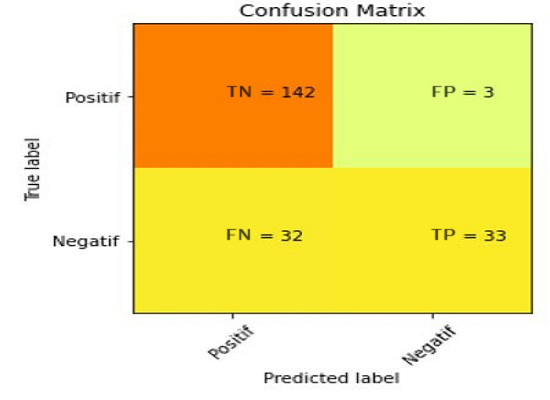
Kemudian hasil dari *F1-Score* dari label negatif sebesar 0.92 dan positif sebesar 0.68, sehingga nilai rata – rata *F1-Score* yaitu (0.92 + 0.68) / 2 = 0.80.

*Support* dari negatif yaitu sebanyak 78 dokiumen dan positif sebanyak 27 dokumen. Sehingga total dokumen yang diklasifikasi pada percobaan yaitu 105 dokumen.

Terakhir yaitu menghitung akurasi dari aplikasi ini. *Accuracy* dapat diketahui dengan cara membagi jumlah klasifikasi benar dengan seluruh dokumen yang diklasifikasi, seperti pada persamaan (2.4). Sehingga :

Dari perhitungan didapat nilai akurasi sebesar 87% yang dihasilkan dari 90% *data training,* 10% *data testing* serta *epoch* 5.

### 4.2.2 Pengujian Kedua

Pengujian kedua dilakukan jumlah *data training* 80%, *data testing* 20%dan *epoch* 5*.*

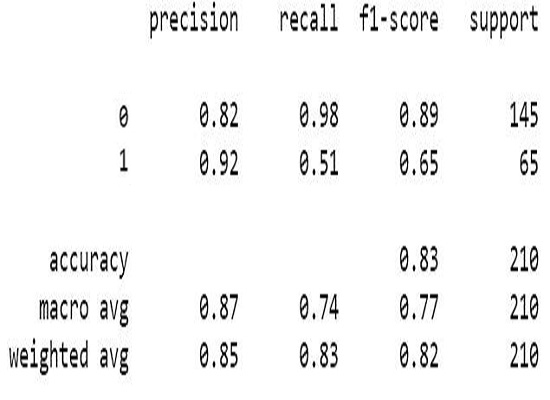
Gambar 4. 4 Confusion Matrix pengujian kedua

Dari hasil pengujian pengujian yang ditunjukan dengan gambar diatas, sehingga dengan menggunakan persamaan (2.1) didapatkan :

Kemudian menghitung *recall* pada tiap – tiap kelas yang di prediksi. Hasil perhitungan dari persamaan (2.2) dan disubstitusikan dalam perhitungan dibawah ini :

Perhitungan selanjutnya yaitu perhitungan *F1-Score* atau ­*F-Measure* yang dimerupakan salah satu perhitungan evaluasi klasifikasi yang mengkombinasikan *recall* dan *precission.* Nilai *F1-Score* dapat diketahui dengan menggunakan persamaan (2.3) bab sebelumnya.

Sehingga dapat dilihat grafik pada gambar dibawah yang dihasilkan dari perhitungan dari persamaan (2.1), (2.2) dan (2.3) dihasilkan grafik seperti berikut:



Gambar 4. 5 Grafik F1 Score pengujian kedua

Dari grafik diatas dihasilkan *precission* dari positif sebesar 0.92 dan negatif sebesar 0.82. Sehingga nilai rata – rata dari presisinya yaitu *avg* = (0.92 + 0.82) / 2 = 0.87. Nilai *precission* didapatkan dengan cara menggunakan persamaan (2.1).  
 Kemudian hasil *recall* dari label negatif sebesar 0.98 dan positif 0.51. Sehingga nilai rata – rata dari *recall* yaitu: *recall* = (0.98 + 0.51) / 2 = 0.74.

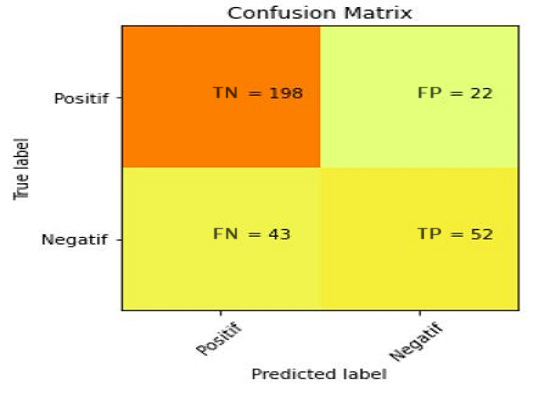
Kemudian hasil dari *F1-Score* dari label negatif sebesar 0.89 dan positif sebesar 0.65, sehingga nilai rata – rata *F1-Score* yaitu (0.89 + 0.65) / 2 = 0.77.

*Support* dari negatif yaitu sebanyak 145 dokiumen dan positif sebanyak 65 dokumen. Sehingga total dokumen yang diklasifikasi pada percobaan yaitu 210 dokumen.

Terakhir yaitu menghitung akurasi dari aplikasi ini. *Accuracy* dapat diketahui dengan cara membagi jumlah klasifikasi benar dengan seluruh dokumen yang diklasifikasi, seperti pada persamaan (2.4). Sehingga :

Dari perhitungan didapat nilai akurasi sebesar 83% yang dihasilkan dari 80% *data training,* 20% *data testing* serta *epoch* 5.

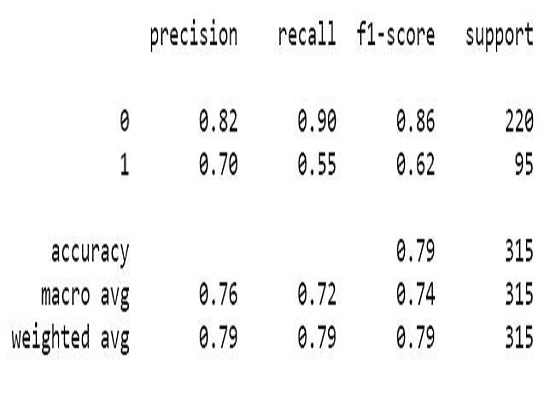
### 4.2.3 Pengujian Ketiga

Pengujian ketiga dilakukan jumlah *data training* 70%, *data testing* 30%dan *epoch* 5*.*

Gambar 4. 6 Confusion Matrix pengujian ketiga

Dari hasil pengujian pengujian yang ditunjukan dengan gambar diatas, sehingga dengan menggunakan persamaan (2.1) didapatkan :

Kemudian menghitung *recall* pada tiap – tiap kelas yang di prediksi. Hasil perhitungan dari persamaan (2.2) dan disubstitusikan dalam perhitungan dibawah ini :

 Perhitungan selanjutnya yaitu perhitungan *F1-Score* atau ­*F-Measure* yang dimerupakan salah satu perhitungan evaluasi klasifikasi yang mengkombinasikan *recall* dan *precission.* Nilai *F1-Score* dapat diketahui dengan menggunakan persamaan (2.3) bab sebelumnya.

Sehingga dapat dilihat grafik pada gambar dibawah yang dihasilkan dari perhitungan dari persamaan (2.1), (2.2) dan (2.3) dihasilkan grafik seperti berikut:

Gambar 4. 7 Grafik F1 Score pengujian ketiga

Dari grafik diatas dihasilkan *precission* dari positif sebesar 0.70 dan negatif sebesar 0.82. Sehingga nilai rata – rata dari presisinya yaitu *avg* = (0.70 + 0.82) / 2 = 0.76. Nilai *precission* didapatkan dengan cara menggunakan persamaan (2.1).  
 Kemudian hasil *recall* dari label negatif sebesar 0.90 dan positif 0.55. Sehingga nilai rata – rata dari *recall* yaitu: *recall* = (0.90 + 0.55) / 2 = 0.72.

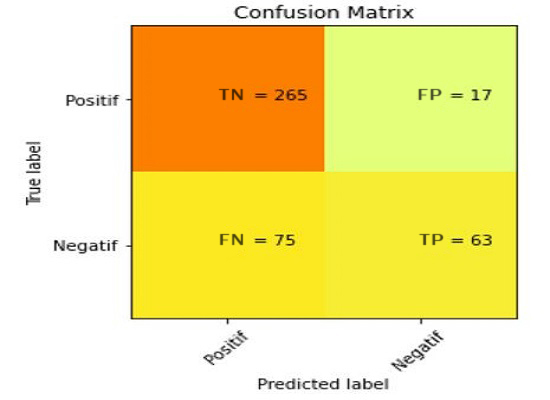
Kemudian hasil dari *F1-Score* dari label negatif sebesar 0.86 dan positif sebesar 0.62, sehingga nilai rata – rata *F1-Score* yaitu (0.86 + 0.62) / 2 = 0.74.

*Support* dari negatif yaitu sebanyak 220 dokiumen dan positif sebanyak 95 dokumen. Sehingga total dokumen yang diklasifikasi pada percobaan yaitu 315 dokumen.

Terakhir yaitu menghitung akurasi dari aplikasi ini. *Accuracy* dapat diketahui dengan cara membagi jumlah klasifikasi benar dengan seluruh dokumen yang diklasifikasi, seperti pada persamaan (2.4). Sehingga :

Dari perhitungan didapat nilai akurasi sebesar 79% yang dihasilkan dari 70% *data training,* 30% *data testing* serta *epoch* 5.

### 4.2.4 Pengujian Keempat

Pengujian keempat dilakukan jumlah *data training* 60%, *data testing* 40%dan *epoch* 5. 

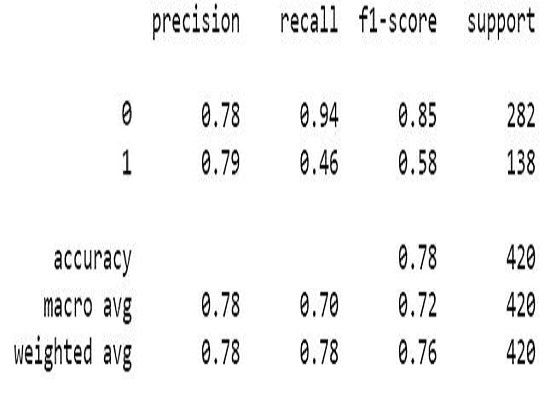
Gambar 4. 8 Confusion Matrix pengujian keempat

Dari hasil pengujian pengujian yang ditunjukan dengan gambar diatas, sehingga dengan menggunakan persamaan (2.1) didapatkan :

Kemudian menghitung *recall* pada tiap – tiap kelas yang di prediksi. Hasil perhitungan dari persamaan (2.2) dan disubstitusikan dalam perhitungan dibawah ini :

Perhitungan selanjutnya yaitu perhitungan *F1-Score* atau ­*F-Measure* yang dimerupakan salah satu perhitungan evaluasi klasifikasi yang mengkombinasikan *recall* dan *precission.* Nilai *F1-Score* dapat diketahui dengan menggunakan persamaan (2.3) bab sebelumnya.

Sehingga dapat dilihat grafik pada gambar dibawah yang dihasilkan dari perhitungan dari persamaan (2.1), (2.2) dan (2.3) dihasilkan grafik seperti berikut:



Gambar 4. 9 Grafik F1 Score pengujian keempat

Dari grafik diatas dihasilkan *precission* dari positif sebesar 0.79 dan negatif sebesar 0.78. Sehingga nilai rata – rata dari presisinya yaitu *avg* = (0.79 + 0.78) / 2 = 0.785. Nilai *precission* didapatkan dengan cara menggunakan persamaan (2.1).  
 Kemudian hasil *recall* dari label negatif sebesar 0.94 dan positif 0.46. Sehingga nilai rata – rata dari *recall* yaitu: *recall* = (0.94 + 0.46) / 2 = 0.70.

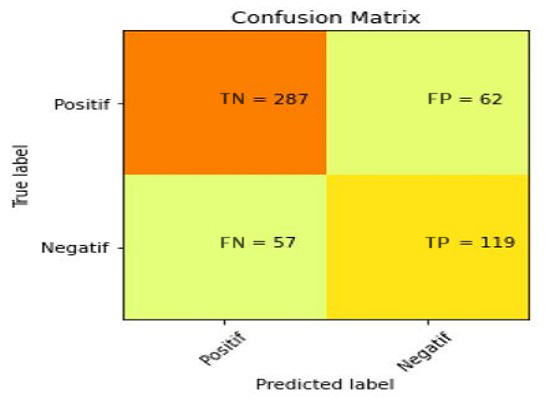
Kemudian hasil dari *F1-Score* dari label negatif sebesar 0.85 dan positif sebesar 0.58, sehingga nilai rata – rata *F1-Score* yaitu (0.85 + 0.58) / 2 = 0.72.

*Support* dari negatif yaitu sebanyak 282 dokiumen dan positif sebanyak 138 dokumen. Sehingga total dokumen yang diklasifikasi pada percobaan yaitu 420 dokumen.

Terakhir yaitu menghitung akurasi dari aplikasi ini. *Accuracy* dapat diketahui dengan cara membagi jumlah klasifikasi benar dengan seluruh dokumen yang diklasifikasi, seperti pada persamaan (2.4). Sehingga :

Dari perhitungan didapat nilai akurasi sebesar 78% yang dihasilkan dari 60% *data training,* 40% *data testing* serta *epoch* 5.

### 4.2.5 Pengujian Kelima

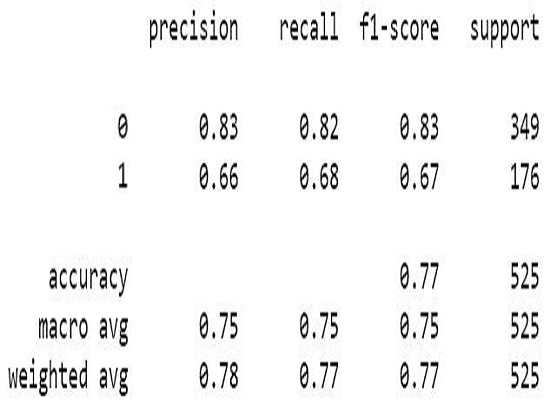
Pengujian kelima dilakukan jumlah *data training* 50%, *data testing* 50%dan *epoch* 5.

Gambar 4. 10 Confusion Matrix pengujian kelima

Dari hasil pengujian pengujian yang ditunjukan dengan gambar diatas, sehingga dengan menggunakan persamaan (2.1) didapatkan :

Kemudian menghitung *recall* pada tiap – tiap kelas yang di prediksi. Hasil perhitungan dari persamaan (2.2) dan disubstitusikan dalam perhitungan dibawah ini :

Perhitungan selanjutnya yaitu perhitungan *F1-Score* atau ­*F-Measure* yang dimerupakan salah satu perhitungan evaluasi klasifikasi yang mengkombinasikan *recall* dan *precission.* Nilai *F1-Score* dapat diketahui dengan menggunakan persamaan (2.3) bab sebelumnya.

  
 Sehingga dapat dilihat grafik pada gambar dibawah yang dihasilkan dari perhitungan dari persamaan (2.1), (2.2) dan (2.3) dihasilkan grafik seperti berikut:

Gambar 4. 11 Grafik F1 Score pengujian kelima

Dari grafik diatas dihasilkan *precission* dari positif sebesar 0.66 dan negatif sebesar 0.83. Sehingga nilai rata – rata dari presisinya yaitu *avg* = (0.66 + 0.83) / 2 = 0.75. Nilai *precission* didapatkan dengan cara menggunakan persamaan (2.1).  
 Kemudian hasil *recall* dari label negatif sebesar 0.82 dan positif 0.68. Sehingga nilai rata – rata dari *recall* yaitu: *recall* = (0.82 + 0.68) / 2 = 0.75.

Kemudian hasil dari *F1-Score* dari label positif sebesar 0.83 dan negatif sebesar 0.67, sehingga nilai rata – rata *F1-Score* yaitu (0.83 + 0.67) / 2 = 0.75.

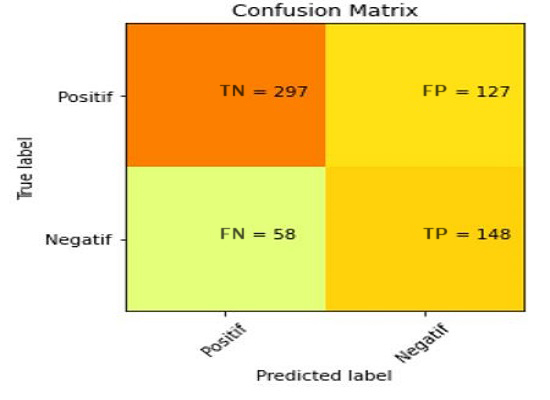
*Support* dari negatif yaitu sebanyak 349 dokiumen dan positif sebanyak 176 dokumen. Sehingga total dokumen yang diklasifikasi pada percobaan yaitu 525 dokumen.

Terakhir yaitu menghitung akurasi dari aplikasi ini. *Accuracy* dapat diketahui dengan cara membagi jumlah klasifikasi benar dengan seluruh dokumen yang diklasifikasi, seperti pada persamaan (2.4). Sehingga :

Dari perhitungan didapat nilai akurasi sebesar 77% yang dihasilkan dari 50% *data training,* 50% *data testing* serta *epoch* 5.

### 4.2.6 Pengujian Keenam

Pengujian keenam dilakukan jumlah *data training* 40%, *data testing* 60%dan *epoch* 5*.*

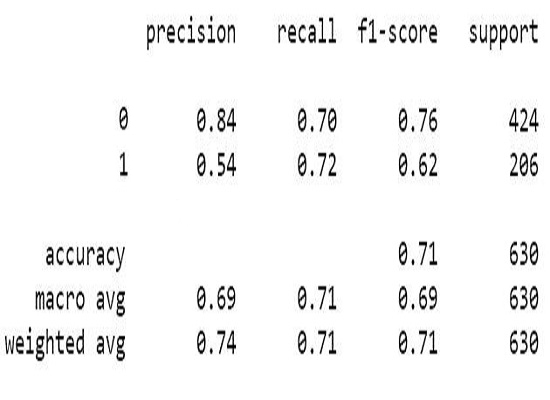


Gambar 4. 12 Confusion Matrix pengujian keenam

Dari hasil pengujian pengujian yang ditunjukan dengan gambar diatas, sehingga dengan menggunakan persamaan (2.1) didapatkan :

Kemudian menghitung *recall* pada tiap – tiap kelas yang di prediksi. Hasil perhitungan dari persamaan (2.2) dan disubstitusikan dalam perhitungan dibawah ini :

Perhitungan selanjutnya yaitu perhitungan *F1-Score* atau ­*F-Measure* yang dimerupakan salah satu perhitungan evaluasi klasifikasi yang mengkombinasikan *recall* dan *precission.* Nilai *F1-Score* dapat diketahui dengan menggunakan persamaan (2.3) bab sebelumnya.

Sehingga dapat dilihat grafik pada gambar dibawah yang dihasilkan dari perhitungan dari persamaan (2.1), (2.2) dan (2.3) dihasilkan grafik seperti berikut:

Gambar 4. 13 Grafik F1 Score pengujian keenam

Dari grafik diatas dihasilkan *precission* dari positif sebesar 0.54 dan negatif sebesar 0.84. Sehingga nilai rata – rata dari presisinya yaitu *avg* = (0.54 + 0.84) / 2 = 0.69. Nilai *precission* didapatkan dengan cara menggunakan persamaan (2.1).  
 Kemudian hasil *recall* dari label negatif sebesar 0.70 dan positif 0.72. Sehingga nilai rata – rata dari *recall* yaitu: *recall* = (0.70 + 0.72) / 2 = 0.71.

Kemudian hasil dari *F1-Score* dari label negatif sebesar 0.76 dan positif sebesar 0.62, sehingga nilai rata – rata *F1-Score* yaitu (0.76 + 0.62) / 2 = 0.69.

*Support* dari negatif yaitu sebanyak 424 dokiumen dan positif sebanyak 206 dokumen. Sehingga total dokumen yang diklasifikasi pada percobaan yaitu 630 dokumen.

Terakhir yaitu menghitung akurasi dari aplikasi ini. *Accuracy* dapat diketahui dengan cara membagi jumlah klasifikasi benar dengan seluruh dokumen yang diklasifikasi, seperti pada persamaan (2.4). Sehingga :

Dari perhitungan didapat nilai akurasi sebesar 70% yang dihasilkan dari 40% *data training,* 60% *data testing* serta *epoch* 5.

### 4.2.7 Pembahasan Hasil Pengujian

Pengujian yang telah dilakukan dengan total sebanyak 6 kali dengan menghasilkan nilai akurasi terbaik, oleh karena itu hasil pengujian tersebut dilakukan dianalisis.

Gambar 4. 14 Grafik hasil pengujian

Grafik diatas menunjukan pengujian pertama berada di angka diatas 80% tepatnya 87%, yang kemudian dilakukan pengujian kedua mengalami penurunan akurasi menjadi 83%, lalu pengujian ketiga mengalami penurunan lagi menjadi 79%. Pengujian keempat pun mengalami penurunan mnejadi 78%, selanjutnya pengujian kelima masih mengalami penurunan menjadi 77%, lalu pengujian yang terakhir menjadi 70%.

Data hasil pengujian lebih rinci dari algoritma CNN dapat dilihat ditabel berikut ini:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Training Data*** | | ***Testing Data*** | | ***Precision*** | | ***Recall*** | | ***F1-Score*** | | ***Accuracy*** |
| ***%*** | ***Data*** | ***%*** | ***Data*** | ***1*** | ***0*** | ***1*** | ***0*** | ***1*** | ***0*** | ***%*** |
| 90% | 944 | 10% | 105 | 100% | 86% | 52% | 100% | 68% | 92% | 87 |
| 80% | 839 | 20% | 210 | 92% | 82% | 51% | 98% | 65% | 89% | 83 |
| 70% | 734 | 30% | 315 | 70% | 82% | 55% | 90% | 62% | 86% | 79 |
| 60% | 629 | 40% | 420 | 79% | 78% | 46% | 94% | 58% | 85% | 78 |
| 50% | 524 | 50% | 525 | 66% | 83% | 68% | 82% | 67% | 83% | 77 |
| 40% | 419 | 60% | 630 | 54% | 84% | 72% | 70% | 62% | 76% | 70 |
| ***Mean of Accurary*** | | | | **76.8%** | **82,5%** | **57.3%** | **89%** | **64%** | **85.1%** | **79** |

Tabel 4. 1 Hasil pengujian

Pada tabel diatas dapat dilihat bahwa dari enam pengujian yang telah dilakukan dengan beberapa variasi untuk data latih dan data uji mendapatkan hasil yang berbeda. Dengan demikian terlihat jelas bahwa prediksi yang dihasilkan dari algortima *Convolutional Neural Network* bergantung pada jumlah data yang digunakan sebagai *data training* dan *data testing,* namun banyak faktor lain pula yang dapat mempengaruhi hasil akurasi, bahasa, jumlah pengulangan iterasi (*epoch*), dan data yang kita labeli. Dari setiap perhitungan *precision, recall* serta *f1-score* terdapat 2 perhitungan yaitu negatif dan positif. Semakin banyak data latih maka semakin tinggi akurasi, begitu pula mempengaruhi terhadap algoritma untuk mempelajari pola – pola yang diajarkan. Karena pada dasarnya sebuah algoritma akan semakin pintar ketika semakin sering untuk belajar pola ataupun model yang dipelajarinya.

BAB V

PENUTUP

## 5.1 Kesimpulan

Sentimen analisis terhadap data *Tweet* mengenai kebijakan *Work from home* pada masa tanggap darurat *Covid-19* menggunakan algoritma *Convolutional Neural Network* telah berhasil dilakukan dengan cara mengambil (*scraping*) data dari media sosial Twitter, yang kemudian diolah dengan beberapa proses diantara lain; *preprocessing,* pembobotan kata menggunakan metode *Word2Vec* dan selanjutnya di proses dengan dibagi menjadi 2 bagian yaitu *data training* dan *data testing.* Kemudian data dipisahkan dengan tujuan memprediksi terhadap *data testing* yang sebelumnya sudah dipisahkan dan diberikan pelabelan secara manual.

Algoritma *Convolutional Neural Network* ini memberikan akurasi tertinggi sebesar 87% hal ini terjadi karena data yang diambil pada saat pengujian dengan secara acak atau *random* yang mengakibatkan antara percobaan satu dengan lainnya menggunakan data yang berbeda. Dari hasil penelitian ini dapat disimpulkan bahwa masih banyak respon masyarakat yang negatif terhadap kebijakan *Work from home* pada media sosial Twitter yang dimana hal ini membuktikan bahwa masyarakat belum siap terhadap kebijakan baru ini. Terlihat pula dari jumlah banyaknya sentimen negatif sebanyak 66% mengandung kata – kata yang tidak setuju dengan kebijakan *Work from home.* Tingkat akurasi dipengaruhi beberapa faktor:

- Kamus data *Word2Vec* yang digunakan dari mulai 50D, 100D, dan 200D, dimana semakin besar dimensi semakin baik terhadap pembobotan kata;

*- Epoch* atau pengulangan iterasi, yang mempengaruhi terhadap pembelejaran pola pada algoritma;

- *Preprocessing* data dimana semakin bersih data akan mempengaruhi terhadap akurasi;

## 5.2 Saran

1. Lebih banyaknya data yang digunakan maka akan menghasilkan tangkat keakuratan yang lebih baik, terutama pada data *training,* agar sistem dapat belajar dengan lebih baik lagi;

2. Sebaiknya dilakukan pengecekan secara rinci dan secara manual terhadap *data set* yang akan dipergunakan pada proses *training* dan *testing,* agar data yang akan diolah benar benar bersih dari *noise* atau gangguan.

# DAFTAR PUSTAKA

[1] ROHDE RODNEY, “2019 Novel Coronavirus (2019-nCoV) Update: Uncoating the Virus,” *American Society for Microbiology*, 2020. .

[2] Rustiana Yohana, “Persepsi Digital Dependent terhadap Pemanfaatan Media Sosial dan Dampak Sosial Ekonominya,” *J. Ilmu Komun.*, vol. 15, no. 1, pp. 17–32, 2018.

[3] Rahmadi R, “Budaya Work from Home dan Perbaikan Kualitas Lingkungan : Mongabay.co.id,” Apr. 06, 2020. https://www.mongabay.co.id/2020/04/06/budaya-work-from-home-dan-perbaikan-kualitas-lingkungan/ (accessed Apr. 16, 2020).

[4] W. S. Eka Putra, “Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101,” *J. Tek. ITS*, vol. 5, no. 1, 2016, doi: 10.12962/j23373539.v5i1.15696.

[5] W. Febrianti, P. S. Informatika, F. Komunikasi, D. A. N. Informatika, and U. M. Surakarta, “Perbandingan Kepuasan Pengguna Aplikasi Chatting,” 2017.

[6] “Peningkatan hasil analisa sentimen menggunakan pos tagger untuk melihat tanggapan masyarakat terhadap full day school,” 2017.

[7] A. Fathan Hidayatullah and A. Sn, “ISSN: 1979-2328 UPN "Veteran,” *Semin. Nas. Inform.*, vol. 2014, no. semnasIF, pp. 115–122, 2014, [Online]. Available: http://www.situs.com.

[8] I. Zulfa and E. Winarko, “Sentimen Analisis Tweet Berbahasa Indonesia Dengan Deep Belief Network,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 11, no. 2, p. 187, 2017, doi: 10.22146/ijccs.24716.

[9] A. (UNIVERSITAS I. N. M. M. I. M. SYAKURO, “PADA MEDIA SOSIAL MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER ( NBC ) DENGAN SELEKSI FITUR INFORMATION GAIN ( IG ) HALAMAN JUDUL SKRIPSI Oleh : ABDAN SYAKURO,” *Anal. sentimen Masy. terhadap e-commerce pada media Sos. menggunakan Metod. naive bayes Classif. dengan Sel. fitur Inf. gain*, pp. 1–89, 2017.

[10] A. Santoso and G. Ariyanto, “Implementasi Deep Learning Berbasis Keras Untuk Pengenalan Wajah,” *Emit. J. Tek. Elektro*, vol. 18, no. 01, pp. 15–21, 2018, doi: 10.23917/emitor.v18i01.6235.

[11] T. A. Lorosae and B. D. Prakoso, “Analisis Sentimen Berdasarkan Opini Masyarakat Pada,” *Semin. Nas. Teknol. Inf. dan Multimed. 2018 Univ. AMIKOM Yogyakarta, 10 Februari 2018*, pp. 25–30, 2018.

[12] F. Maclean, D. Jones, G. Carin-Levy, and H. Hunter, “Understanding twitter,” *Br. J. Occup. Ther.*, vol. 76, no. 6, pp. 295–298, 2013, doi: 10.4276/030802213X13706169933021.

[13] R. Murri, “Performance of Python runtimes on a non-numeric scientific code,” Apr. 2014, Accessed: Apr. 18, 2020. [Online]. Available: http://arxiv.org/abs/1404.6388.

[14] K. Baktha and B. K. Tripathy, “Investigation of recurrent neural networks in the field of sentiment analysis,” *Proc. 2017 IEEE Int. Conf. Commun. Signal Process. ICCSP 2017*, vol. 2018-Janua, no. April, pp. 2047–2050, 2018, doi: 10.1109/ICCSP.2017.8286763.

[15] Y. Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4.5,” *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017.

[16] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[17] Keras, “About Keras,” 2020. https://keras.io/about/ (accessed Jul. 09, 2020).

[18] S. R. DEWI, “Deep Learning Object Detection Pada Video,” *Deep Learn. Object Detect. Pada Video Menggunakan Tensorflow Dan Convolutional Neural Netw.*, 2018, [Online]. Available: https://dspace.uii.ac.id/bitstream/handle/123456789/7762/14611242\_Syarifah Rosita Dewi\_Statistika.pdf?sequence=1.

[19] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” *Insights Imaging*, vol. 9, no. 4, pp. 611–629, 2018, doi: 10.1007/s13244-018-0639-9.

[20] E. Times, “Definition of Waterfall Model | What is Waterfall Model ? Waterfall Model Meaning - The Economic Times,” 2015. https://economictimes.indiatimes.com/definition/waterfall-model (accessed May 12, 2020).

[21] F. Ugm and F. Ugm, “Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 8, no. 1, pp. 91–100, 2014, doi: 10.22146/ijccs.3499.

[22] VisualParadigm, “What is Data Flow Diagram?,” 2019. https://www.visual-paradigm.com/guide/data-flow-diagram/what-is-data-flow-diagram/ (accessed May 12, 2020).

[23] L. Chart, “What is a Data Flow Diagram | Lucidchart,” *Lucidchart.Com*, 2019. https://www.lucidchart.com/pages/data-flow-diagram (accessed May 12, 2020).

[24] K. S. Nugroho, “Confusion Matrix untuk Evaluasi Model pada Supervised Learning,” *Medium*, 2019. https://medium.com/@ksnugroho/confusion-matrix-untuk-evaluasi-model-pada-unsupervised-machine-learning-bc4b1ae9ae3f (accessed Jul. 06, 2020).

[25] C. K. Dewa, A. L. Fadhilah, and A. Afiahayati, “Convolutional Neural Networks for Handwritten Javanese Character Recognition,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 12, no. 1, p. 83, 2018, doi: 10.22146/ijccs.31144.

[1] ROHDE RODNEY, “2019 Novel Coronavirus (2019-nCoV) Update: Uncoating the Virus,” *American Society for Microbiology*, 2020. .

[2] Rustiana Yohana, “Persepsi Digital Dependent terhadap Pemanfaatan Media Sosial dan Dampak Sosial Ekonominya,” *J. Ilmu Komun.*, vol. 15, no. 1, pp. 17–32, 2018.

[3] Rahmadi R, “Budaya Work from Home dan Perbaikan Kualitas Lingkungan : Mongabay.co.id,” Apr. 06, 2020. https://www.mongabay.co.id/2020/04/06/budaya-work-from-home-dan-perbaikan-kualitas-lingkungan/ (accessed Apr. 16, 2020).

[4] W. S. Eka Putra, “Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101,” *J. Tek. ITS*, vol. 5, no. 1, 2016, doi: 10.12962/j23373539.v5i1.15696.

[5] W. Febrianti, P. S. Informatika, F. Komunikasi, D. A. N. Informatika, and U. M. Surakarta, “Perbandingan Kepuasan Pengguna Aplikasi Chatting,” 2017.

[6] “Peningkatan hasil analisa sentimen menggunakan pos tagger untuk melihat tanggapan masyarakat terhadap full day school,” 2017.

[7] A. Fathan Hidayatullah and A. Sn, “ISSN: 1979-2328 UPN "Veteran,” *Semin. Nas. Inform.*, vol. 2014, no. semnasIF, pp. 115–122, 2014, [Online]. Available: http://www.situs.com.

[8] I. Zulfa and E. Winarko, “Sentimen Analisis Tweet Berbahasa Indonesia Dengan Deep Belief Network,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 11, no. 2, p. 187, 2017, doi: 10.22146/ijccs.24716.

[9] A. (UNIVERSITAS I. N. M. M. I. M. SYAKURO, “PADA MEDIA SOSIAL MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER ( NBC ) DENGAN SELEKSI FITUR INFORMATION GAIN ( IG ) HALAMAN JUDUL SKRIPSI Oleh : ABDAN SYAKURO,” *Anal. sentimen Masy. terhadap e-commerce pada media Sos. menggunakan Metod. naive bayes Classif. dengan Sel. fitur Inf. gain*, pp. 1–89, 2017.

[10] A. Santoso and G. Ariyanto, “Implementasi Deep Learning Berbasis Keras Untuk Pengenalan Wajah,” *Emit. J. Tek. Elektro*, vol. 18, no. 01, pp. 15–21, 2018, doi: 10.23917/emitor.v18i01.6235.

[11] T. A. Lorosae and B. D. Prakoso, “Analisis Sentimen Berdasarkan Opini Masyarakat Pada,” *Semin. Nas. Teknol. Inf. dan Multimed. 2018 Univ. AMIKOM Yogyakarta, 10 Februari 2018*, pp. 25–30, 2018.

[12] F. Maclean, D. Jones, G. Carin-Levy, and H. Hunter, “Understanding twitter,” *Br. J. Occup. Ther.*, vol. 76, no. 6, pp. 295–298, 2013, doi: 10.4276/030802213X13706169933021.

[13] R. Murri, “Performance of Python runtimes on a non-numeric scientific code,” Apr. 2014, Accessed: Apr. 18, 2020. [Online]. Available: http://arxiv.org/abs/1404.6388.

[14] K. Baktha and B. K. Tripathy, “Investigation of recurrent neural networks in the field of sentiment analysis,” *Proc. 2017 IEEE Int. Conf. Commun. Signal Process. ICCSP 2017*, vol. 2018-Janua, no. April, pp. 2047–2050, 2018, doi: 10.1109/ICCSP.2017.8286763.

[15] Y. Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4.5,” *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017.

[16] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[17] Keras, “About Keras,” 2020. https://keras.io/about/ (accessed Jul. 09, 2020).

[18] S. R. DEWI, “Deep Learning Object Detection Pada Video,” *Deep Learn. Object Detect. Pada Video Menggunakan Tensorflow Dan Convolutional Neural Netw.*, 2018, [Online]. Available: https://dspace.uii.ac.id/bitstream/handle/123456789/7762/14611242\_Syarifah Rosita Dewi\_Statistika.pdf?sequence=1.

[19] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” *Insights Imaging*, vol. 9, no. 4, pp. 611–629, 2018, doi: 10.1007/s13244-018-0639-9.

[20] E. Times, “Definition of Waterfall Model | What is Waterfall Model ? Waterfall Model Meaning - The Economic Times,” 2015. https://economictimes.indiatimes.com/definition/waterfall-model (accessed May 12, 2020).

[21] F. Ugm and F. Ugm, “Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 8, no. 1, pp. 91–100, 2014, doi: 10.22146/ijccs.3499.

[22] VisualParadigm, “What is Data Flow Diagram?,” 2019. https://www.visual-paradigm.com/guide/data-flow-diagram/what-is-data-flow-diagram/ (accessed May 12, 2020).

[23] L. Chart, “What is a Data Flow Diagram | Lucidchart,” *Lucidchart.Com*, 2019. https://www.lucidchart.com/pages/data-flow-diagram (accessed May 12, 2020).

[24] K. S. Nugroho, “Confusion Matrix untuk Evaluasi Model pada Supervised Learning,” *Medium*, 2019. https://medium.com/@ksnugroho/confusion-matrix-untuk-evaluasi-model-pada-unsupervised-machine-learning-bc4b1ae9ae3f (accessed Jul. 06, 2020).

[25] C. K. Dewa, A. L. Fadhilah, and A. Afiahayati, “Convolutional Neural Networks for Handwritten Javanese Character Recognition,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 12, no. 1, p. 83, 2018, doi: 10.22146/ijccs.31144.