

**MULTI-TARGET KLASIFIKASI TEKS BERDASAR TANGGAPAN  
MASYARAKAT KOTA DENGAN MENGGUNAKAN ALGORITMA CNN  
& RANDOM FOREST**

**ALYA ROHALIA  
D082192002**



**PROGRAM STUDI S2 TEKNIK INFOMATIKA  
DEPARTEMEN TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS HASANUDDIN  
GOWA  
2023**



Optimized using  
trial version  
[www.balesio.com](http://www.balesio.com)

# **PENGAJUAN TESIS**

## **MULTI-TARGET KLASIFIKASI TEKS BERDASAR TANGGAPAN MASYARAKAT KOTA DENGAN MENGGUNAKAN ALGORITMA CNN & RANDOM FOREST**

Tesis

Sebagai Salah Satu Syarat untuk Mencapai Gelar Magister  
Program Studi Teknik Informatika

Disusun dan diajukan oleh:

**ALYA ROHALIA**

**D082192002**

Kepada

**FAKULTAS TEKNIK**

**UNIVERSITAS HASANUDDIN**

**GOWA**

**2023**



# TESIS

## MULTI-TARGET KLASIFIKASI TEKS BERDASAR TANGGAPAN MASYARAKAT KOTA DENGAN MENGUNAKAN ALGORITMA CNN & RANDOM FOREST

**ALYA ROHALIA**  
**D082192002**

Telah dipertahankan di hadapan Panitia Ujian Tesis yang dibentuk dalam rangka penyelesaian studi pada Program Magister Teknik Informatika Fakultas Teknik Universitas Hasanuddin Pada tanggal 4 Desember 2023 dan dinyatakan telah memenuhi syarat kelulusan



Pembimbing Utama



Dr. Ir. Zahir Zainuddin, M.Sc.  
NIP. 19640427 198910 1 002

Pembimbing Pendamping



Dr.Eng. Zulkifli Tahir, S.T., M.Sc.  
NIP. 198404032 01012 1 004

Dekan Fakultas Teknik  
Universitas Hasanuddin



Ummad Isran Ramli, M.T. IPM., ASEAN.Eng.  
NIP. 19730926 200012 1 002

Ketua Program Studi  
S2 Teknik Informatika



Dr. Ir. Zahir Zainuddin, M.Sc.  
NIP. 19640427 198910 1 002



## PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Yang bertanda tangan di bawah ini :

Nama : Alya Rohalia  
Nomor Mahasiswa : D082192002  
Program Studi : Magister Teknik Informatika

Dengan ini menyatakan bahwa, tesis yang berjudul “Multi-Target Klasifikasi Teks Berdasar Tanggapan Masyarakat Kota Dengan Menggunakan Algoritma *CNN-LSTM* dan *Random Forest*” adalah benar karya saya dengan arahan dari komisi pembimbing (Dr. Ir. Zahir Zainuddin, M.Sc dan Dr-Eng. Zulkifli Tahir, S.T., M.Sc.). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apapun kepada perguruan tinggi manapun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini akan dipublikasikan di Jurnal/Prosiding “*International Conference on Artificial Intelligence Robotics, Signal and Image Processing (AiRoSIP)* tahun 2023) sebagai artikel dengan judul “*Classification of Community Responses to Service Offices Using a Combined CNN-LSTM Algorithm and Random Forest*”.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Gowa, 4 Desember 2023

Yang Menyatakan,

  
Alya Rohalia



## KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Allah SWT. karena berkat rahmat dan karunia-Nya sehingga tesis yang berjudul “**Multi-Target Klasifikasi Teks Berdasar Tanggapan Masyarakat Kota Dengan Menggunakan Algoritma CNN dan Random Forest**” ini dapat diselesaikan sebagai salah satu syarat dalam menyelesaikan jenjang Strata-2 pada Departemen Teknik Informatika Fakultas Teknik Universitas Hasanuddin.

Penulis menyadari bahwa dalam penyusunan dan penulisan laporan tesis ini tidak lepas dari bantuan, bimbingan serta dukungan dari berbagai pihak, dari masa perkuliahan sampai dengan masa penyusunan tesis. Oleh karena itu, penulis dengan senang hati menyampaikan terima kasih kepada:

1. Kedua Orang tua penulis, Bapak Hj. Burhan dan Ibu Hj. Wikrayani yang selalu menjadi motivasi terbesar dalam penyelesaian perkuliahan ini yang tidak pernah putus memberikan dukungan, doa, dan semangat serta selalu sabar dalam mendidik penulis sejak kecil;
2. Adik kandung penulis, Muh. Naufal Fauzan yang mendukung dalam segi apapun dan memberikan semangat kepada penulis selama penyusunan tesis;
3. Bapak Dr. Ir. Zahir Zainuddin, M.Sc. selaku pembimbing I dan Bapak Dr-Eng. Zulkifli Tahir, S.T., M.Sc. selaku pembimbing II yang telah memberikan waktu, tenaga, pikiran, dukungan moril maupun materil serta perhatian yang luar biasa untuk mengarahkan penulis dalam penyusunan tesis;
4. Bapak Prof. Dr. Eng. Ir. Syafaruddin, S.T, M.Eng., Bapak Prof. Dr. Ir. Indrabayu, ST., MT., M.Bus.Sys., IPM, ASEAN. Eng., dan Bapak Dr. Amil Ahmad Ilham, S.T., M.IT., selaku dosen penguji yang telah memberikan kritik dan saran yang membangun sehingga laporan tesis ini menjadi lebih baik;
5. Bapak Dr. Ir. Zahir Zainuddin, M.Sc selaku Ketua Program Studi Teknik Informatika Fakultas Teknik Universitas Hasanuddin yang telah memberikan motivasi, bimbingan, dan semangatnya selama masa perkuliahan penulis;



6. Para sahabat, teman-teman, kakak-kakak dan adik-adik di Laboratorium *Computer Based System* Pascasarjana UNHAS yang telah memberikan begitu banyak bantuan, motivasi, semangat dan pengalaman selama proses perkuliahan;
7. Teman-teman Pascasarjana UNHAS Angkatan 1 atas dukungan dan semangat yang diberikan selama ini;
8. Ibu Yuanita serta segenap Staf Departemen Magister Teknik Informatika yang telah banyak membantu penulis selama pengurusan administrasi;
9. Orang-orang terkasih yang tidak sempat dituliskan oleh penulis;

Akhir kata, penulis berharap semoga Allah SWT. Senantiasa berkenan membalas segala kebaikan dari semua pihak yang telah banyak membantu. Semoga Tesis ini dapat memberikan manfaat bagi pengembangan ilmu. Aamiin ya Rabbal Alamin.

Gowa, 4 Desember 2023



**Alya Rohalia**



## ABSTRAK

**Alya Rohalia.** Multi-Target Klasifikasi Teks Berdasar Tanggapan Masyarakat Kota Dengan Menggunakan Algoritma *CNN* dan *Random Forest*. (dibimbing oleh **Zahir Zainuddin**, dan **Zulkifli Tahir**).

Peningkatan pengguna media sosial linier dengan kebebasan masyarakat untuk berpendapat tentang kinerja kota. Dalam meningkatkan kinerja kota, pemerintah membutuhkan tanggapan masyarakat; tanggapan ini merupakan sumber informasi awal dalam menilai kinerja pemerintah terkait tugas dan tanggung jawabnya. Tanggapan masyarakat meliputi komentar di halaman akun media sosial pemerintah dan akun komunitas swasta terkait permasalahan di kota; kami menerima komentar melalui media sosial *Twitter*, *Facebook*, *Instagram*, dan *Youtube*. Riset ini berupa tanggapan masyarakat terhadap 16 dinas, mengklasifikasikannya ke dalam dinas terkait dengan menggunakan lima kategori sentimen. Metode yang digunakan untuk klasifikasi sentimen adalah kombinasi algoritma *CNN* dan *Random Forest*. Penelitian ini menyajikan model *Continuous Bag-of-Word (CBOW)* dan memasukkannya ke dalam tahap ekstraksi fitur *CNN*. Pekerjaan ini menggunakan *Alexnet Convolution Neural Network* dan mengklasifikasikannya menggunakan algoritma *Random Forest*. Keluaran sistem dari kombinasi algoritma *CNN & RF* berupa klasifikasi multi-target dari analisis sentimen memperoleh akurasi sebesar 86%, dan klasifikasi dinas memperoleh akurasi sebesar 81%.

**Kata Kunci:** Klasifikasi Teks, *Convolutional Neural Network (CNN)*, *Random Forest Algorithm*.



## ABSTRACT

**Alya Rohalia.** Multi-Target Classification of Text Based on City Community Responses Using the CNN Algorithm and Random Forest. (Supervised by **Zahir Zainuddin**, and **Zulkifli Tahir**).

The use of social media has increased linearly along with the freedom of the community to express their opinions about the city's performance. In improving the city's performance, the government requires public feedback; this feedback serves as an initial source of information in evaluating the government's performance regarding its duties and responsibilities. Public feedback includes comments on government social media accounts and private community accounts related to city issues; we receive comments through social media platforms such as Twitter, Facebook, Instagram, and Youtube. This research focuses on the public's feedback on 16 departments, classifying them into relevant departments using five sentiment categories. The method used for sentiment classification is a combination of CNN and Random Forest algorithms. This study presents the Continuous Bag-of-Word (CBOW) model and incorporates it into the CNN feature extraction stage. Additionally, this work uses the Alexnet Convolution Neural Network and classifies it using the random forest algorithm. The system output of the CNN & RF algorithm combination is a multi-target classification of sentiment analysis with an accuracy of 86%, while the department classification achieves an accuracy of 80%.

**Keywords:** Klasifikasi Teks, Convolutional Neural Network (CNN), Random Forest Algorithm.



## DAFTAR ISI

	<u>Halaman</u>
PENGAJUAN TESIS .....	i
PERNYATAAN KEASLIAN TESIS .....	iii
KATA PENGANTAR .....	iv
ABSTRAK.....	vi
ABSTRACT.....	vii
DAFTAR GAMBAR .....	x
DAFTAR TABEL.....	xii
BAB I PENDAHULUAN .....	13
I.1 Latar Belakang .....	13
I.2 Rumusan Masalah.....	15
I.3 Tujuan Penelitian .....	15
I.4 Manfaat Penelitian .....	15
I.5 Batasan masalah.....	16
BAB II TINJAUAN PUSTKA.....	17
II.1. Penelitian Terdahulu.....	17
2.1.1 Evaluasi Kinerja Kota .....	17
2.1.2 Analisis Sentiment .....	18
2.1.3 Klasifikasi Teks .....	20
II.2 Kajian Teori.....	21
2.2.1 Dinas Daerah.....	21
2.2.2 Data Set.....	23
2.2.3 Fitur Penilaian .....	27
2.2.3 Preprocessing NLP.....	30
2.2.4 FastText Vectors .....	31
2.2.5 CNN (Convolutional Neural Network) .....	32
2.2.6 Algoritma <i>Random Forest</i> .....	35
II.3 Diagram Alur Penelitian .....	36
II.4. Metode Penyelesaian Masalah.....	38
2.4.1 Metode Klasifikasi .....	38
2.4.2 Metode Analisis .....	38
Metode Pengujian Sistem.....	39
State of the Art Penelitian .....	40
Tingkat Keaslian (Level Orisinalitas) Topik Penelitian yang diusulkan.....	51
Target Hipotesis Penelitian .....	54



BAB III METODOLOGI PENELITIAN .....	57
III.1 Metode .....	57
III.2 Jenis Penelitian.....	57
III.3 Tahapan Penelitian .....	57
III.4 Sumber Data.....	59
III.5 Rancangan Arsitektur .....	59
III.6 Instrumen Penelitian.....	72
III.7 Waktu dan Lokasi .....	72
BAB IV HASIL DAN PEMBAHASAN.....	73
IV.1 Pengambilan Data Set .....	73
4.1.1 Label Dinas dan Sentiment .....	73
IV.2 Model Setting.....	74
IV.2 Data Set Prediction .....	75
IV.3 Validasi Model.....	79
4.3.1 Validasi Model Dinas.....	79
4.3.2 Validasi Model Sentiment.....	81
IV.4 Perbandingan Model .....	85
BAB V PENUTUP .....	87
IV.1 Kesimpulan .....	87
IV.2 Saran .....	87
DAFTAR PUSTAKA .....	88



## DAFTAR GAMBAR

	<u>Halaman</u>
<b>Gambar 1</b> Contoh data set covid-19 indonesian tweets.....	21
<b>Gambar 2</b> Contoh dataset <i>_tweet_sentiment_cellular_service_provider.csv</i> .....	22
<b>Gambar 3</b> Contoh data set <i>sentiment analysis on financial tweets</i> .....	22
<b>Gambar 4</b> contoh komentar yang diambil dari <i>Twitter</i> .....	23
<b>Gambar 5</b> Contoh komentar yang diambil dari <i>Facebook</i> .....	23
<b>Gambar 6</b> Contoh komentar yang diambil dari <i>Instagram</i> .....	24
<b>Gambar 7</b> Contoh komentar yang diambil dari berita <i>Youtube</i> .....	24
<b>Gambar 8</b> Contoh data set yang belum tersedia.....	25
<b>Gambar 9</b> Proses <i>tokenization</i> .....	29
<b>Gambar 10</b> Proses <i>stemming</i> .....	29
<b>Gambar 11</b> <i>Continuous Bag-of-Word (CBOW)</i> .....	30
<b>Gambar 12</b> Arsitektur <i>CNN</i> .....	31
<b>Gambar 13</b> Cara kerja algoritma <i>Random Forest</i> .....	34
<b>Gambar 14</b> Diagram alur klasifikasi multi-target.....	35
<b>Gambar 15</b> Tahapan penelitian.....	56
<b>Gambar 16</b> Data lokasi.....	57
<b>Gambar 17</b> Arsitektur model klasifikasi <i>CNN RF</i> .....	58
<b>Gambar 18</b> Proses dan output <i>tokenization</i> .....	62
<b>Gambar 19</b> Output <i>CNN</i> .....	67
<b>Gambar 20</b> Output fitur.....	68
<b>Gambar 21</b> Data frame fitur.....	68
<b>Gambar 22</b> Accuracy and loss <i>CNN</i> .....	73
<b>Gambar 23</b> Prediksi manusia (dinas dan sentiment).....	74
<b>Gambar 24</b> Prediksi sistem (dinas dan sentiment).....	74
<b>Gambar 25</b> Prediksi manusia & sistem (dinas dan sentiment).....	75
<b>Gambar 26</b> Klasifikasi teks hasil prediksi manusia & sistem (dinas dan sentiment).....	75
 <b>27</b> Confusion matrix dinas.....	77
<b>28</b> Accuracy , precision, recall, etc. Evaluation result matrix dinas.....	77
<b>29</b> Evaluation result matrix sentiment.....	80



**Gambar 30** Diagram perbandingan CNN RF Models.....83



## DAFTAR TABEL

	<u>Halaman</u>
<b>Tabel 1</b> State of the art.....	38
<b>Tabel 2</b> Matriks metode penyelesaian.....	49
<b>Tabel 3</b> Level orisinalitas berdasarkan standar proposal program studi magister teknik informatika universitas hasanuddin.....	52
<b>Tabel 4</b> Kerangka pikir penelitian.....	54
<b>Tabel 5</b> Tahapan Metode.....	55
<b>Tabel 6</b> Dataset Description.....	71
<b>Tabel 7</b> Size of the neural network.....	72
<b>Tabel 8</b> Model Setting.....	72
<b>Tabel 9</b> Prediksi Benar Dan Salah Manusia & Sistem.....	76
<b>Tabel 10</b> Perbandingan model CNN RF.....	83



# BAB I

## PENDAHULUAN

### I.1 Latar Belakang

Dinas daerah merupakan unsur pelaksana pemerintah dalam suatu daerah. Daerah dapat berarti Provinsi, Kabupaten, atau Kota. Dinas Daerah menyelenggarakan fungsi yaitu perumusan kebijakan teknis sesuai dengan lingkup tugasnya, pemberian perizinan dan pelaksanaan pelayanan umum, serta pembinaan pelaksanaan tugas. Pada Dinas Daerah Kabupaten/Kota mempunyai tugas melaksanakan kewenangan desentralisasi yang bertujuan untuk meningkatkan efektivitas dan produktivitas dalam pencapaian tujuan organisasi sehingga dapat melaksanakan peran penting terkait pelayanan publik pada suatu kabupaten/kota.

Seiring berjalannya waktu, peningkatan pengguna media sosial linear dengan kebebasan masyarakat berpendapat terhadap kinerja kota. Setiap warga negara Indonesia yang tinggal di daerah perkotaan penting kiranya untuk menyampaikan setiap keluhan maupun tanggapan ke dinas-dinas terkait. Penyampaian keluhan masyarakat ini sangat penting dalam rangka mendukung evaluasi kinerja kota yang ada di Indonesia. Evaluasi kinerja kota adalah suatu proses untuk mengukur dan mengevaluasi sejauh mana kinerja pemerintah kota dalam memenuhi tujuan dan sasaran yang telah ditetapkan.

Tujuan dari evaluasi kinerja kota adalah untuk menilai efektivitas, efisiensi, dan keterjangkauan kebijakan dan program yang telah diimplementasikan oleh pemerintah kota. Evaluasi kinerja kota dapat dilakukan dengan berbagai metode, termasuk survei publik, analisis data, dan pengamatan lapangan. Evaluasi kinerja kota biasanya melibatkan penilaian terhadap sejumlah aspek, seperti layanan publik, manajemen keuangan, penggunaan sumber daya, dan hasil dari program-program pemerintah.

Hasil dari evaluasi kinerja kota dapat digunakan untuk mengidentifikasi kekuatan dan kelemahan pemerintah kota, serta menentukan prioritas kebijakan dan yang perlu ditingkatkan atau dikembangkan. Evaluasi kinerja kota juga membantu pemerintah kota untuk memperbaiki transparansi dan akuntabilitas menjalankan tugas-tugasnya. Tapi pada umumnya, Analisis terhadap masyarakat terkait kinerja kota relative tidak berkembang begitu juga



dengan kondisi pengelolaan pengaduan pelayanan publik di setiap Dinas penyelenggara belum menunjukkan kondisi pengelolaan yang efektif dan terintegrasi.

Masing-masing Dinas mengelola pengaduan secara parsial, tidak terkoordinir dan terintegrasi dengan baik. Hal ini menyebabkan pengelolaan pengaduan tidak efektif dan tidak efisien, misalnya terjadinya pengaduan secara acak semisal banjir, sampah, jalanan rusak, ataupun terkait dengan pelayanan publik dll yang ada pada laman media social, surat kabar, siaran televisi agar berharap ditangani oleh pemerintah padahal masing-masing dinas juga turut serta menampung aspirasi terkait keluhan masyarakat dan sebisa mungkin memberi solusi sebelum akhirnya disampaikan ke pemerintah pusat terkait dengan keluhan tersebut.

Setiap dinas pemerintahan mempunyai peran dan tugasnya masing-masing dalam meningkatkan layanan masyarakat namun kendala yang di alami yaitu keluhan maupun penilaian kinerja tidak terklasifikasi terhadap dinas-dinas pemerintah kota misalnya keluhan banjir disampaikan ke dinas social, keluhan pelayanan BPJS disampaikan kepada dinas kesehatan dll. Kebanyakan hanya bersifat random dan acak dari berbagai media khususnya media sosial sehingga tiap keluhan menjadi terpecah-pecah, tidak ter'arah, dan tidak jelas tujuan yang tersampaikan sehingga tiap aspirasi kemungkinan akan diabaikan begitu saja oleh dinas tertentu.

Mengangkat dari masalah yang ada maka dari itu penulis mengusulkan suatu sistem klasifikasi multi target. yang dapat menampung keluhan masyarakat, mengelompokkan tiap-tiap keluhan/tanggapan masyarakat ke 16 dinas yang terkait sesuai dengan fungsi dan tugasnya. Penelitian ini mampu memprediksi dan menargetkan sentiment dari 5 kategori yaitu sangat baik, baik, cukup, kurang dan sangat kurang ke dinas yang akan di tujukan. Pada kontribusi penelitian ini berupa klasifikasi multi target menggunakan 2 algoritma yaitu *CNN (Convolutional Neural Network)* dan *Random Forest* yang disatukan..



## I.2 Rumusan Masalah

Berdasarkan pada latar belakang di atas, maka rumusan masalah pada penelitian ini yaitu

1. Bagaimana kinerja model gabungan *CNN* dan *Random Forest* dalam melakukan multi-target klasifikasi teks tanggapan masyarakat kota?
2. Bagaimana perbandingan kinerja model gabungan *CNN* dan *Random Forest* dengan model yang hanya menggunakan satu jenis algoritma?
3. Apa faktor-faktor yang mempengaruhi kinerja model gabungan *CNN* dan *Random Forest* dalam melakukan multi-target klasifikasi teks tanggapan masyarakat kota?

## I.3 Tujuan Penelitian

Berdasarkan rumusan masalah maka tujuan yang akan dicapai dari penelitian ini:

1. Mengembangkan model gabungan *CNN* dan *Random Forest* yang efektif dalam melakukan multi target klasifikasi tanggapan masyarakat kota.
2. Menganalisis kinerja model gabungan *CNN* dan *Random Forest* dalam melakukan multi target klasifikasi tanggapan masyarakat kota dan membandingkannya dengan model yang hanya menggunakan satu jenis algoritma.
3. Mengidentifikasi faktor-faktor yang mempengaruhi kinerja model gabungan *CNN* dan *Random Forest* dalam melakukan multi target klasifikasi tanggapan masyarakat kota untuk meningkatkan akurasi dan efektivitas model.

## I.4 Manfaat Penelitian

Berdasarkan rumusan masalah dan tujuan penelitian maka manfaat penelitian ini yaitu:

1. Mengidentifikasi faktor-faktor yang mempengaruhi kinerja model gabungan *CNN* dan *Random Forest* dalam melakukan multi target klasifikasi tanggapan masyarakat kota untuk meningkatkan akurasi dan efektivitas model.



2. Memberikan manfaat bagi pemerintah dalam mengambil kebijakan dan tindakan yang lebih tepat dan efektif dalam menangani berbagai masalah yang dihadapi masyarakat.
3. Menambah wawasan dan pengetahuan bagi masyarakat dan para peneliti terkait teknologi pengolahan data, khususnya pada bidang analisis sentimen dan klasifikasi multi target.

### **I.5 Batasan masalah**

Mengembangkan model klasifikasi teks yang dapat memprediksi *multiple* target atau *multiple* label dari tanggapan masyarakat kota.



## BAB II

### TINJAUAN PUSTKA

#### II.1. Penelitian Terdahulu

##### 2.1.1 Evaluasi Kinerja Kota

Pada penelitian (Sun and Zhou, 2010), pembangunan berkelanjutan dari kota berbasis sumber daya bergantung pada transformasi ekologi industri yang efektif. Sistem ini menggunakan metode *AHP* membagi system indicator menjadi lapisan target, lapisan kriteria dan lapisan indikator (tiga indikator masing-masing dicatat untuk indikator tingkat A, B, C). Lapisan target mewakili tingkat ekologi industri perkotaan atas nama efektivitas berjalan secara keseluruhan. Lapisan kriteria menyelesaikan sistem indeks ekologi industri menjadi sejumlah indeks evaluasi subsistem yang saling terkait. Lapisan indikator adalah indikator untuk mengevaluasi indikator subsistem di atas, yang digunakan untuk menggambarkan tingkat keadaan variabel tertentu, jumlah keadaannya, untuk mengukur kekuatan. Namun metode ini memiliki beberapa keterbatasan, sehingga pada penelitian selanjutnya peneliti akan melakukan penelitian empiris untuk memverifikasi kualitas ilmiah metode yang dikombinasikan dengan metode evaluasi *komprehensif fuzzy*).

Pada penelitian (Airaksinen *et al.*, 2018), didalam penelitiannya menyajikan kerangka kerja pengukuran kinerja holistik untuk pemantauan dan komparabilitas yang harmonis dan transparan dari aktivitas kota-kota Eropa selama implementasi solusi *Smart City*. Metodologi kerja didasarkan pada kolaborasi dan komunikasi ekstensif dengan kota-kota yang ada di Eropa untuk mengembangkan kerangka Indikator Kinerja Utama (KPI) yaitu masyarakat, planet, kemakmuran, pemerintahan dan perambatan sebagai evaluasi dan komparabilitas inisiatif Kota Cerdas. Baseline ditetapkan dengan analisis dan integrasi hasil yang ada dari inisiatif sebelumnya. Penelitian ini mengembangkan dan memvalidasi kerangka kerja evaluasi kinerja, termasuk definisi KPI, pedoman pengumpulan data, sistem kinerja, dan pengujian di kota-kota yang mempunyai kasus. Selain mengusulkan rekomendasi untuk penerapan sistem kinerja ke dalam proses



pengambilan keputusan kota dan rekomendasi untuk pengembangan peluang bisnis baru.

Pada penelitian (Shen *et al.*, 2018), indikator evaluasi dalam penelitian ini dipilih dengan menerapkan metodologi dari penelitian *hibrida* termasuk tinjauan literatur dan semi- wawancara terstruktur. Adapun indikator holistik untuk memeriksa praktik kota pintar dalam konteks Cina yaitu infrastruktur cerdas, pemerintahan cerdas, ekonomi cerdas, orang cerdas, lingkungan cerdas. Data indikator dikumpulkan dari 44 kota sampel yang ada di China. NS evaluasi dilakukan dengan menerapkan metode *Entropi* dan Teknik *Order Teknik Preference by Similarity to Ideal Solution (TOPSIS)* secara kolektif. Dan hasil penelitian menunjukkan bahwa kinerja keseluruhan praktik kota pintar di China berada pada level yang relatif lebih rendah

Pada penelitian (Limon-Ruiz *et al.*, 2019), berfokus pada bagaimana mendukung pengambilan keputusan untuk berkonsultasi dengan KPI dan memungkinkan masyarakat mengetahui perkembangan kota dalam beberapa waktu melalui platform elektronik. Penelitian ini menggunakan metodologi difusi inovasi melalui praktik pengembangan perangkat lunak, yang menerapkan sebuah temuan ke desain platform yang diusulkan sebagai kontribusi utama, spesifikasi dasar awal untuk desain platform elektronik untuk representasi informasi berharga yang memfasilitasi pengambilan keputusan oleh pemerintah agar dapat meningkatkan keterlibatan warga dalam mendukung kinerja *smart city* dengan baik. Hasilnya, memungkinkan kami untuk mengembangkan dan menguji prototipe fungsional untuk mewakili informasi Kota dalam Antarmuka *Web Visual* berdasarkan KPI yang ada di lingkungan hidup.

### 2.1.2 Analisis Sentiment

Pada penelitian (Prastyo *et al.*, 2020), didalam penelitiannya menganalisis sentiment pada aspek umum menggunakan dataset dua kelas. Hasil penelitian mereka menunjukkan bahwa menggunakan Algoritma *SVM* dengan *Normalized* *nel* dapat memprediksi sentimen di *Twitter* dengan cepat dan akurat. 1 masyarakat sepakat dengan kebijakan pemerintah dalam menangani '9 dari segi ekonomi, namun secara umum tidak puas dengan kinerja



pemerintah. Kinerja algoritma ini mencapai akurasi rata-rata, *presisi*, *recall*, dan *f-measure* tertinggi dengan nilai masing-masing sebesar 82,00%, 82,24%, 82,01%, dan 81,84%.

Pada penelitian (Putri *et al.*, 2020), didalam penelitiannya menganalisis sentiment terhadap pembangunan infrastruktur di Kota Malang dengan menggunakan *Twitter* sebagai media sosial yang dijadikan sumber data. Peneliti mengumpulkan 1980 data *tweet* yang mencantumkan akun pemerintah Kota Malang dan Dinas Pekerjaan Umum Kota Malang pada periode tahun 2018 hingga 2019. Peneliti membagi opini dalam *tweet* tersebut menjadi dua kategori, yaitu positif dan negatif, dengan menggunakan algoritma *Support Vector Machine*. Hasil penelitian menunjukkan bahwa mayoritas *tweet* yang berkaitan dengan pembangunan infrastruktur memiliki sentimen negatif daripada positif. Metode klasifikasi *Support Vector Machine* memiliki akurasi yang berbeda ketika menggunakan 70%, 80%, dan 90% data training, yaitu masing-masing mencapai 73,10%, 72,62%, dan 86,42%.

Pada penelitian (Muslih, Akhmad Rizaldy, Ajib Susanto, Ibnu Utomo Wahyu Mulyono, 2020) didalam penelitiannya menganalisis sentiment terhadap komentar masyarakat di akun Instagram pemerintah dengan menggunakan algoritma *Text Matching Boyer Moore* untuk mengklasifikasikan secara otomatis. Hasil penelitian mereka menghasilkan sebuah aplikasi yang dapat menganalisis komentar masyarakat di akun Instagram instansi pemerintah. Aplikasi ini dapat membantu pengelola akun untuk menganalisis respons masyarakat terhadap postingan, baik itu respons positif, negatif, atau netral. Metode yang diusulkan memiliki performa yang baik dengan mencapai akurasi sebesar 79%.

Pada penelitian (Kastrati *et al.*, 2021), melakukan studi penelitian pada tahun 2021 dengan fokus pada analisis sentimen dari pendapat orang-orang yang diungkapkan di *Facebook* mengenai situasi pandemi saat ini dalam bahasa *Albania* yang sumber daya bahasanya rendah. Mereka membuat kumpulan data skala besar yang terdiri dari 10.742 komentar yang diklasifikasikan secara manual dalam



tersebut. Dalam penelitian ini, mereka melaporkan upaya mereka pada pengembangan penganalisis sentimen yang menggunakan pembelajaran mesin. Hasil temuan eksperimental menunjukkan bahwa penganalisis sentimen

yang mereka usulkan, yang menggabungkan *BiLSTM* dengan mekanisme perhatian, mencapai kinerja tertinggi dengan skor F1 sebesar 72,09%. Selain itu, hasil eksperimen juga menunjukkan bahwa model pembelajaran mendalam (*IDCNN*, *BiLSTM*, *Hybrid*, dan *BERT*) secara umum berkinerja lebih baik daripada model pembelajaran mesin konvensional (*SVM*, *NB*, *DT*, *RF*) dalam pemodelan komentar tekstual.

### 2.1.3 Klasifikasi Teks

Pada penelitian (Gurcan, 2018), melakukan klasifikasi teks dalam bahasa Turki menggunakan metode pembelajaran mesin yang diawasi. Model tersebut diujikan untuk mengklasifikasikan teks berita dalam lima kelas yang telah ditentukan sebelumnya, yaitu ekonomi, politik, olahraga, kesehatan, dan teknologi. Dalam penelitian ini, beberapa metode klasifikasi seperti *Multinomial Naïve Bayes*, *Bernoulli Naïve Bayes*, *Dukungan Mesin Vektor*, *K-Nearest Neighbor*, dan *Pohon Keputusan* digunakan pada teks berita dan dievaluasi dengan berbagai parameter. Hasil penelitian menunjukkan bahwa algoritma *Multinomial Naïve Bayes* menghasilkan tingkat keberhasilan klasifikasi terbaik, mencapai sekitar 90%. Temuan ini menunjukkan bahwa model probabilitas *Naïve Bayes* dapat efektif digunakan sebagai metode pengklasifikasi dalam mengklasifikasikan teks berbahasa Turki jika dibandingkan dengan metode lainnya.

Pada penelitian (Jia *et al.*, 2020), mengajukan sebuah metode yang menggabungkan Representasi teks *Tibet Word2vec* dan *TF-IDF* berdasarkan varians frekuensi kelas. Dalam metode representasi ini, mereka menggunakan Model jaringan *BiLSTM* untuk mengklasifikasikan teks *Tibet*. Hasil eksperimen menunjukkan bahwa dengan menggabungkan representasi teks menggunakan *Word2vec* dan *TF-IDF* berdasarkan varians frekuensi kelas, efek klasifikasi teks dapat meningkat secara efektif. Keakuratan pengklasifikasi teks *Tibet* berdasarkan *BiLSTM* mencapai 89,03%, yang jauh lebih baik daripada menggunakan *RNN LSTM*.



la penelitian (Siji George and Sumathi, 2021), melakukan analisis model baru yang menggabungkan keuntungan dari *Convolutional Neural (CNN)* dan *Random Forest (RF) Classifier*, yang diusulkan untuk

klasifikasi sentimen finegrained. Ide utama dari model yang diusulkan adalah untuk mencapai akurasi maksimum dalam klasifikasi sentimen pada volume data yang besar. Untuk mencapai ini, model *CBOW (Continuous Bag-of-Words)* digunakan untuk mengubah input teks menjadi vektor. Selanjutnya, *Convolutional Neural Network (CNN)* digunakan untuk mengekstrak fitur dari vektor input. Lapisan yang sepenuhnya terhubung dalam Jaringan Saraf Konvolusi digantikan oleh pengklasifikasi Hutan Acak. Kemudian fitur-fitur yang diekstraksi digunakan untuk proses klasifikasi dengan *Random Forest Classifier*. Strategi dropout diterapkan untuk mengatur model *CNNRF* agar menghindari *overfitting*. Analisis sentimen dilakukan pada data review produk dengan menggunakan model *CNN* dan *RF* secara terpisah. Hasil dari model *CNN* dan *RF* dibandingkan dengan hasil dari model *CNNRF* yang diusulkan. Hasil eksperimen menunjukkan bahwa model gabungan *CNNRF* memberikan kinerja yang lebih tinggi daripada model *CNN* dan *RF* yang bekerja secara independen

## II.2 Kajian Teori

### 2.2.1 Dinas Daerah

Dinas Daerah adalah bagian penting dalam pelaksanaan otonomi daerah. Dinas Daerah bertanggung jawab untuk menjalankan tugas-tugas yang terkait dengan Pemerintahan Daerah berdasarkan prinsip otonomi dan tugas-tugas pendukung (Protokol, 2019). Berikut ini adalah tugas dan tanggung jawab masing-masing dari dinas daerah :

1. Dinas Sosial, membantu urusan pemerintahan dalam bidang sosial, advokasi dan pengawasan kesekretariatan berdasarkan asas otonomi dan tugas bantuan
2. Dinas Kearsipan dan Perpustakaan, membantu urusan pemerintahan dalam bidang kearsipan dan perpustakaan.
3. Dinas Pendidikan, Kepemudaan dan Olahraga, membantu urusan pemerintahan dalam bidang pendidikan dan kepemudaan dan olahraga.

Dinas Kependudukan dan Pencatatan Sipil, membantu urusan pemerintahan dalam administrasi kependudukan dan pencatatan sipil serta kesekretariatan



5. Dinas Perhubungan, membantu urusan pemerintahan dalam bidang perhubungan.
6. Dinas Kesehatan, bertanggung jawab untuk melindungi, mempromosikan, dan meningkatkan kesehatan masyarakat.
7. Dinas Komunikasi dan Informatika, membantu urusan pemerintahan dalam bidang komunikasi dan informatika, persandian dan bidang statistik.
8. Dinas Pariwisata, membantu urusan pemerintahan dalam bidang pariwisata.
9. Dinas Perindustrian dan Perdagangan, Koperasi dan Usaha Kecil Menengah, membantu urusan pemerintahan dalam melaksanakan kewenangan otonomi daerah dalam rangka pelaksanaan tugas desentralisasi di bidang Perindustrian, Perdagangan, Koperasi dan UKM.
10. Dinas Kepemudaan dan Olahraga, membantu urusan pemerintahan dalam bidang pendidikan dan kepemudaan dan olahraga.
11. Dinas Lingkungan Hidup dan Kebersihan, membantu urusan pemerintahan dalam pengendalian dampak lingkungan dan konservasi, pengawasan dan penanggulangan kerusakan lingkungan hidup serta kebersihan dan pengelolaan kualitas lingkungan.
12. Dinas Peternakan dan Kesehatan Hewan, membantu urusan pemerintahan dalam bidang pertanian sub urusan peternakan dan kesehatan hewan
13. Dinas Pemberdayaan Masyarakat dan Desa, membantu urusan pemerintahan dalam bidang pemberdayaan masyarakat dan desa.
14. Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu, membantu urusan pemerintahan dalam bidang penanaman modal dan pelayanan terpadu satu pintu.
15. Dinas Tanaman Pangan, Hortikultura dan Peternakan, membantu urusan pemerintahan dalam pertanian sub urusan tanaman pangan, hortikultura dan Peternakan.

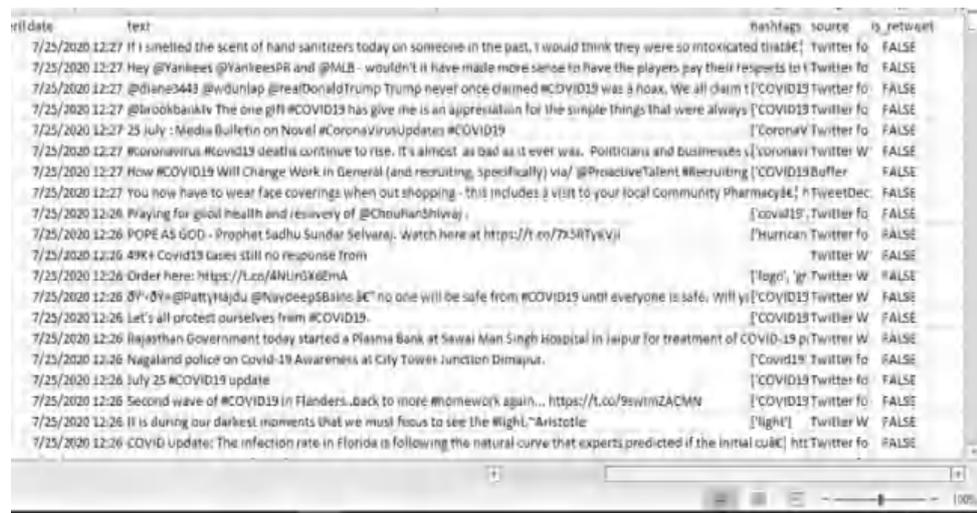


16. Dinas Pekerjaan Umum dan Penataan Ruang, membantu urusan pemerintahan dalam bidang pemerintahan bidang pekerjaan umum dan penataan ruang.

### 2.2.2 Data Set

Dataset adalah kumpulan data yang terorganisir dan terstruktur. Dataset dapat berupa himpunan observasi, informasi, atau fakta yang berkaitan dengan suatu topik atau subjek tertentu.

Pada gambar dibawah 1, 2 dan 3 merupakan data set yang ada di Internet, hanya tersedia untuk satu topik pada suatu permasalahan. Sehingga menjadi tugas peneliti kedepanya untuk mengambil/ menarik serta menyediakan data yang sesuai dan terklasifikasi pada dinas-dinas.



tweetdate	text	hashtags	source	is retweet
7/25/2020 12:27	If I smelled the scent of hand sanitizers today on someone in the past, I would think they were so intoxicated thatâ€¦		Twitter fo	FALSE
7/25/2020 12:27	Hey @Yankees @YankeesPR and @MLB - wouldn't it have made more sense to have the players pay their respects to		Twitter fo	FALSE
7/25/2020 12:27	@diane3443 @wdunlap @realDonaldTrump Trump never once claimed #COVID19 was a hoax. We all claim t	#COVID19	Twitter fo	FALSE
7/25/2020 12:27	@ArrookBairAv The one gift #COVID19 has give me is an appreciation for the simple things that were always	#COVID19	Twitter fo	FALSE
7/25/2020 12:27	25 July : Media Bulletin on Novel #CoronavirusUpdates #COVID19		Twitter fo	FALSE
7/25/2020 12:27	#Coronavirus #COVID19 deaths continue to rise. It's almost as bad as it ever was. Politicians and business	#Coronavirus	Twitter W	FALSE
7/25/2020 12:27	How #COVID19 Will Change Work in General (and recruiting, specifically) via/ @ProactiveTalent #Recruiting	#COVID19	Buffer	FALSE
7/25/2020 12:27	You now have to wear face coverings when out shopping - this includes a visit to your local Community Pharmacy		Twitter Dec	FALSE
7/25/2020 12:26	Praying for global health and resiliency of @ChouhanShivay	#Covid19	Twitter fo	FALSE
7/25/2020 12:26	POPE AS GOD - Prophet Sadhu Sundar Selvaraj. Watch here at <a href="https://t.co/7z5RtYkVji">https://t.co/7z5RtYkVji</a>		Twitter fo	FALSE
7/25/2020 12:26	49K+ Covid19 cases still no response from		Twitter W	FALSE
7/25/2020 12:26	Order here: <a href="https://t.co/4NUn0X6EmA">https://t.co/4NUn0X6EmA</a>		Twitter W	FALSE
7/25/2020 12:26	â€œno one will be safe from #COVID19 until everyone is safe. Will yi	#COVID19	Twitter W	FALSE
7/25/2020 12:26	Let's all protect ourselves from #COVID19.	#COVID19	Twitter W	FALSE
7/25/2020 12:26	Uttarakhand Government today started a Plasma Bank at Sewai Man Singh Hospital in Jaipur for treatment of COVID-19	#COVID-19	Twitter W	FALSE
7/25/2020 12:26	Nagaland police on Covid-19 Awareness at City Tower Junction Dimapur.	#Covid19	Twitter fo	FALSE
7/25/2020 12:26	July 25 #COVID19 update	#COVID19	Twitter fo	FALSE
7/25/2020 12:26	Second wave of #COVID19 in Flanders...back to more homework again... <a href="https://t.co/9swimZACMN">https://t.co/9swimZACMN</a>	#COVID19	Twitter fo	FALSE
7/25/2020 12:26	It is during our darkest moments that we must focus to see the light.~Aristotle	#light	Twitter W	FALSE
7/25/2020 12:26	COVID Update: The infection rate in Florida is following the natural curve that experts predicted if the initial cu		Twitter fo	FALSE

**Gambar 1** Contoh data set *covid-19 indonesian tweets* (ditujukan pada dinas kesehatan)  
sumber : kaggle.com



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Id	Sentimen	Text	Tweet											
2	1	positive	<USER_MENTION>	#BOIKOT_<PROVIDER_NAME>	Gunakan	Produk	Bangsa	Sendiri	<PROVIDER_NAME>						
3	2	positive	Saktinya	balik	lagi,	alhamdulillah	v	<PROVIDER_NAME>							
4	3	negative	Selamat	pagi	<PROVIDER_NAME>	bisa	bantu	kenapa	di	dalam	kamar	sinyal	4G	hilang	yang
5	4	negative	Dear	<PROVIDER_NAME>	akhir2	ini	jaringan	data	lemot	banget	padahal	H+	!!!!		
6	5	negative	Selamat	malam	PENDUSTA	<PROVIDER_NAME>									
7	6	negative	Untuk	penembakan	paket	dari	<PRODUCT_NAME>	masih	gangguan	ya	<PROVIDER_NAME>				
8	7	positive	<PROVIDER_NAME>	aku	pakai	<PROVIDER_NAME>	,	paket	<PRODUCT_NAME>	nya	udah	off	tapi	gantinya	ada
9	8	negative	RT	<USER_MENTION>	<PROVIDER_NAME>	tak	ada	lagi	dalam	kamus	perhapean	saya.	#Boikot	<PROVIDER_NAME>	<URL>
10	9	negative	keluhan	gak	ditanggapi.	bikin	emosi	aja.	pulsaku	dicuri	terus.	soak	deh.	masa	setiap
11	10	negative	#Billboard	iklan	<PROVIDER_NAME>	kok	ada	pesan	tersembunyi?	#Boikot	<PROVIDER_NAME>	#Ramadhan	<URL>		
12	11	negative	<PROVIDER_NAME>	mampus!!!!	bulan	puasa,	subuh2	mau	sahur	cek	pulsa	tersedot	terus.	dan	sekarang
13	12	negative	<USER_MENTION>	Pernah	coba	pakai	<PROVIDER_NAME>	tapi	parah	koneksinya					
14	13	negative	Setuju,	<USER_MENTION>	silahkan	boikot	saja	pelanggan	<PROVIDER_NAME>	pendukung	#HRS	<URL>			
15	14	negative	Dibela	belain	daftarin	kartu	4G	tunggu	sejam	siang2	sambil	puasa,	sampai	rumah	sinyal
16	15	negative	<PROVIDER_NAME>	datanya	kok	sering	hilang	<PROVIDER_NAME>							
17	16	negative	#internetlelet	<PROVIDER_NAME>	ga	sesuai	iklan								
18	17	positive	<PROVIDER_NAME>	<PROVIDER_NAME>	kangen	nih	dapat	kiriman	pos	marchandes	nya	<PROVIDER_NAME>	langsung	dari	pak
19	18	negative	yang	katanya	150MBPS	kecepatanya	cuma	3kbps,	kuota	full.	Ga	bisa	buka	google?	Udah
20	19	negative	kuota	#internet	<PROVIDER_NAME>	masih	banyak	tapi	koneksi	internetnya	LAMBAT	BANGET!	#<PRODUCT_NAME>	<PROVIDER_NAME>	
21	20	negative	Ampun...	Sinyal	4G	<PROVIDER_NAME>	lemot	seperti	siput.	Acara	live	streaming	jadi	terganggu.	Mundur

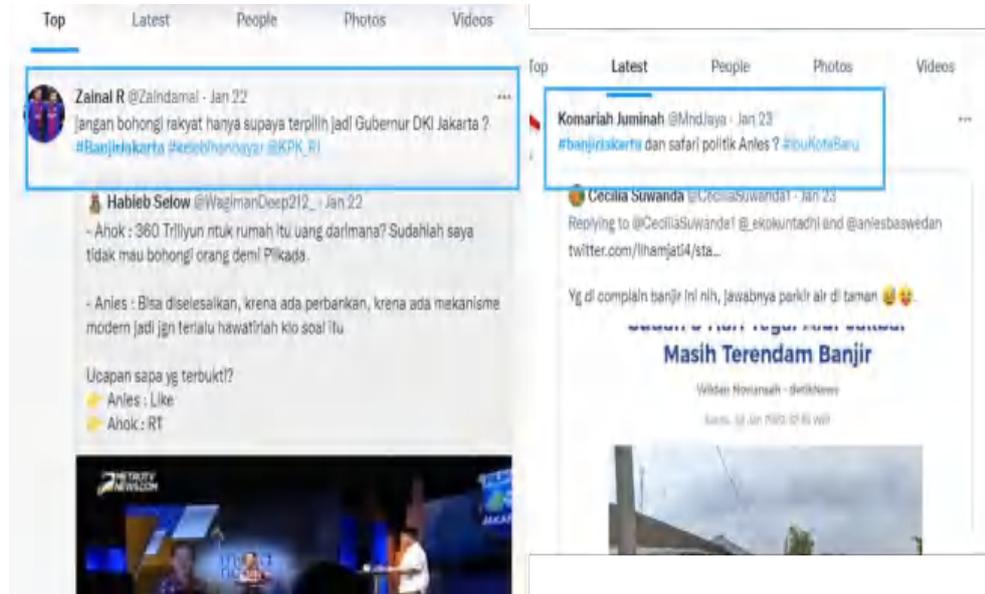
Gambar 2 Contoh dataset *tweet\_sentiment\_cellular\_service\_provider.csv* (ditujukan pada diskominfo) sumber : github

	A	B	C
1	cleaned_tweets		sentiment
2	video offic mind busi david solomon tell gs intern learn wa		0
3	price lumber lb f sinc hit ytd high maci turnaround still happen		0
4	say american dream dead		-1
5	barri silbert extrem optimist bitcoin predict new crypto entrant go zero		1
6	satellit avoid attack space junk circl earth paid		-1
7	david butler favorit fang stock realmoneysod alphabet facebook		0
8	miss convo one favorit thinker		0
9	u intellig document nelson mandela made public		0
10	senat want emerg alert go netflix spotifi etc		1
11	hedg fund manag marc larsi say bitcoin k possibl		0
12	u propos expedit appeal fight amp time warner purchas		-1
13	roger feder uniqlo deal make one athlet earn endors		0
14	bond trader ahead jerom powel come inflat expect via		0
15	alcoa cut adjust ebitda forecast cite tariff share slide		0
16	custom urg boycott mgm resort casino file lawsuit mass shoot victim		-1
17	gap tighten race trillion dollar valuat amazon hit billion via		0
18	presid trump endors brian kemp casey cagl georgia governor race		0
19	white hous struggl contain fallout trump putin summit		0
20	laker guard hart name summer leagu mvp u sport headlin		0
21	two northern white rhino still exist make world endang subspeci mammal		0

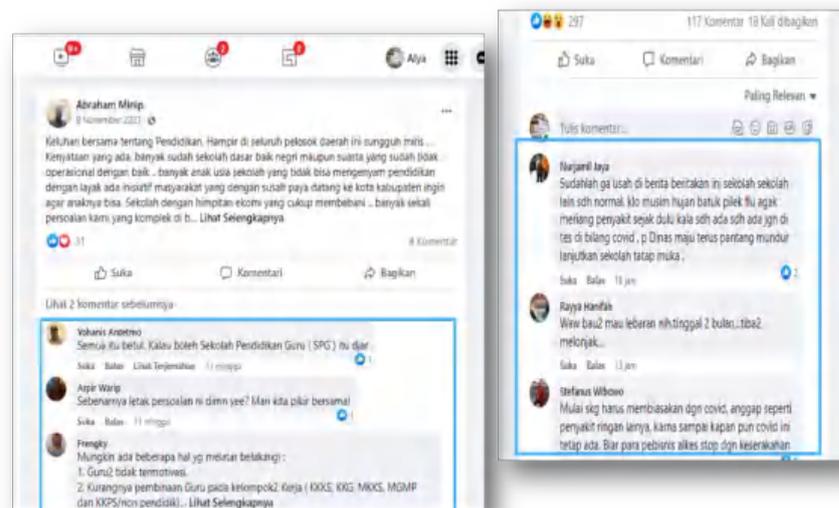
Gambar 3 Contoh data set *sentiment analysis on financial tweets* (ditujukan pada dinas keuangan) sumber : kaggle.com



Adapun beberapa contoh tanggapan masyarakat pada media sosial yang akan ditarik yaitu :



Gambar 4 contoh komentar yang diambil dari *Twitter*



Gambar 5 Contoh komentar yang diambil dari *Facebook*





**Gambar 6** Contoh komentar yang diambil dari *Instagram*



**Gambar 7** Contoh komentar yang diambil dari berita *Youtube*

Berikut ini adalah beberapa contoh kumpulan komentar berupa tanggapan masyarakat yang akan digabungkan dari berbagai laman media sosial dan belum tersedia di internet yaitu :



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	id	Tugas	Sentiment	Tanggapan										
	1	Diskes	Kurang	Pngakit di data terus... yg bisa makan sm yg miskin coba di data tiap hari dan di kasih solusi itu baru MAN										
	2	Diskes	Sangat Kurang	Krtan masih nyalahin orang miskin lg yg nyebar virusnya										
	3	Diskes	Kurang	Trus az skalian perbanyak jalur pelayanan luar negri nya biar wisatawan numpuk. Ga usah peduli in pandemi, si										
	4	Diskes	Baik	Kalau yg tidak mau di vaksin bila terpepar virus Corona, umikron di abaikan aja										
	5	Diskes	Cukup	Inti masalah dari semua ini kemungkinan terbesar nya adalah anggaran										
				Dst...										
	101	Diknas	Kurang	Sebenarnya saya msh kurang setuju 100% tatap muka aturan berahap 50% dulu setiap hari bergantian										
	102	Diknas	Sangat Baik	Terima kasih atas infonya										
	103	Diknas	Cukup	Sebenarnya masih agak deg2an, kenapa gak tetep dibagi, biar gak terlalu banyak kekhawatiran pesti sda minggu										
	104	Diknas	Cukup	Semoga tidak akan muncul kasus baru pada Anak2 sekolah kita nanti, jka ade ujuz2 akan PTM di rumah lagi ken										
	105	Diknas	Sangat Kurang	Peraturan aneh ini hapus aja, mending daring										
				Dst...										
	1001	Diskominfo	Sangat kurang	Amplun... Sinyal 4G <PROVIDER_NAME> lemot seperti siput. Acara live streaming jadi terganggu. Minder terus <PRO										
	1002	Diskominfo	Sangat Kurang	Jaringan <PROVIDER_NAME> di Nabire bagus? Iu hanya berita HOAX #HOAX <PROVIDER_NAME>										
	1003	Diskominfo	Sangat Baik	Selalu pakai <PROVIDER_NAME> belum pernah pakai <PROVIDER_NAME> dan merasalan <PROVIDER_NAME> lemot										
	1004	Diskominfo	Sangat Baik	Terbaik Buat <PROVIDER_NAME> <PROVIDER_NAME>										
	1005	Diskominfo	Kurang	Tolong dong kensinya di benerin di wilayah keramat jati. Kenapa lemot!!!! <PROVIDER_NAME> <PRODUCT_NAME>										
				Dst...										
	1101	Dinsos	Sangat baik	Alhamdulillah... berkat dinsos banyak warga pirwang yg terbantu... terima kasih dinsos selalu ada untuk kami										
	1102	Dinsos	Sangat Baik	Pelaku Sosial yg handal... mantap										
	1103	Dinsos	Sangat Baik	Alhamdulillah pelayanannya sangat baik, pegawai ramah, semoga cmt2 dlu Aamiinnn										
	1104	Dinsos	Baik	Pejuang kemanusiaan										

**Gambar 8** Contoh data set yang belum tersedia

### 2.2.3 Fitur Penilaian

Dalam Kamus Besar Bahasa Indonesia (KBBI), kalimat merupakan satuan bahasa yang relatif berdiri sendiri, mempunyai pola intonasi final dan secara aktual ataupun potensial terdiri atas klausa. Kalimat secara umum digunakan secara lisan maupun tulisan.

Sementara itu, Dilansir dari Jurnal Universitas Pendidikan Indonesia dengan judul 'Pembelajaran Kalimat Bahasa Indonesia Dengan Pola Spiral Pada Program Pendidikan Guru Sekolah Dasar' mengatakan bahwa kalimat adalah satuan bahasa yang lengkap yang mengandung maksud. Kalimat sebagai bentuk bahasa adalah bentuk yang lengkap, bukan bagian dari suatu bentuk bahasa yang lebih besar. Adapun penilaian dari 5 sentiment dapat dikatakan sangat baik, baik, cukup, kurang dan sangat kurang yaitu :

#### 1. Sangat kurang

Ciri-ciri/fitur kalimat yang dikatakan sangat kurang apabila kalimatnya mengandung kata-kata negatif sangat kurang misalnya berupa umpatan, binatang, sumpahan dan kata-kata kotor.

Contoh kalimat :

- Kami sangat kurang setuju, banyak cakap tapi tidak ada bukti (sangat kurang)



- Muak denger bansos.. Yg dpt klg perangkat.. (umpatan)
- Saya sumpahahi semua anggota yg polisi dan dishub yg biasa derek mobil....semua anaknya terlahir idiot (Sumpahian)
- Dishub graup anjing mata duitan,klau di kampung aku udah di bakar hidup hidup, mereka kerja ya korupsi semua rampok yg di pelihara (binatang).

## 2. Kurang

Ciri-ciri/fitur kalimat yang dikatakan kurang apabila kata-katanya mengandung kalimat negatif kurang, tidak terlalu rasis misalnya kalimat yang mengandung kontra/ penolakan, penyangkalan, dan pernyataan yang menyakitkan.

Contoh kalimat :

- Kami masih belum kebagian, bagaimana sih bantuannya masih kurang merata ? (kurang)
- Dana dikorupsi oleh tikus-tikus berdasi, lebih mementingkan isi perut sendiri dari pada kami masyarakat kecil (Pernyataan yang menyakitkan)
- Di hukum di lporin tindak pnganiayaan, gx di hukum ngelawan... kids jaman now (penyangkalan)
- kami kurang setuju tenaga jaringan 3G di hapus (penolakan)

## 3. Cukup

Ciri-ciri/fitur kalimat yang dikatakan cukup apabila kata-katanya mengandung kalimat netral atau cukup, misalnya kalimat yang mengandung informasi, saran, pertanyaan maupun pernyataan yang netral.

Contoh kalimat :

- Kami sudah cukup bersabar dan saat ini kami masih sabar (cukup)
- Sebaiknya pelajaran di sekolah itu nggak melulu mentingin materi tapi sikap dan moral para siswa juga harus di didik (saran)
- Apakah daerah kami sudah ada dana alokasi khusus untuk penerimaan bantuan ? (Pertanyaan)
- Informasi Penting untuk Penerima Bantuan Pemerintah mulai 1 Maret 2022 (Informasi)



#### 4. Baik

Ciri-ciri/fitur kalimat yang dikatakan Baik apabila kata-katanya mengandung kalimat positif baik, misalnya apabila kalimatnya mengandung persetujuan terhadap sesuatu/ pembelaan, mengandung saran perbaikan, pernyataan terima kasih dan pujian terhadap kinerja pemerintah.

- Kami kedatangan Pak Jokowi minggu lalu berkunjung ke pasar, beliau begitu baik, humble dan ramah (baik)
- Perpustakaan keliling memang bagus untuk membuat pintar masyarakat, terutama masyarakat yang terbatas tentang informasi (Pujian)
- Terima kasih telah berbagi (Terima kasih)
- Saya setuju pertahankan wilayah kalian suku Punan jgn sampai di rusak orang dari luar (Pembelaan)
- Bismillah, subhanallah panorama sempurna dan bangunannya bagus, lebih bagus lagi jika ada lambang rumah adat digapura jadi ada ciri khas Sumatra Barat ini hanya saran saja dan juga pelayanan yang ramah jadi benar benar sumbar banget (saran)

#### 5. Sangat baik

Ciri-ciri/fitur kalimat yang dikatakan Sangat Baik apabila kata-katanya mengandung kalimat positif sama dengan sentiment baik, yang membedakan jika kalimatnya mengandung ucapan terima kasih sekaligus pujian ataupun doa baik. Dan jika kalimatnya juga mengandung kata sangat setuju terhadap sesuatu.

- Mantap, alhamdulillah kami masyarakat daerah sangat bangga memiliki pemimpin yang amanah, semoga Bapak diberi kesehatan selalu (Pujian dan doa)
- Terima kasih saya ucapkan kepada Perpustakaan Dan Kearsipan yang telah memberikan Pembinaan dan Pendampingan kepada kami SDN 004 Bontang Barat, sehingga kami mendapatkan Akreditasi Perpustakaan Sekolah dengan predikat A. (Terima kasih)
- Sangat setuju tindakan yang diambil oleh Dinas Tetkait dan maunya perlu ditinjau kembali yang diterapkan apa ukuran bak truk atau



khusus roda 6 keatas, karena sekarang banyak truk yang ukuran besar memasang roda 4 asalkan bisa bongkar muat di jalan yang disebut di atas yang mungkin menurutnya disamakan dengan pick up (sangat setuju)

### 2.2.3 Preprocessing NLP

Data yang diperoleh dari laman web umumnya berupa kalimat yang tidak terstruktur seperti terdapat karakter non alfanumerik, link url, tata bahasa tidak teratur, dll. Adanya *noise* dan ketidakteraturan data memengaruhi kinerja mesin. *Preprocessing NLP* adalah langkah penghapusan noise yang tidak diinginkan dari dataset twitter sehingga data yang dihasilkan menjadi lebih terstruktur. Adapun tahap-tahap dari *preprocessing* yaitu :

#### 1. *Character Removal*

Dataset yang telah diberi label selanjutnya dibersihkan dari karakter yang tidak relevan seperti link url, hastag, username, tanda baca dan karakter non alfanumerik (kecuali spasi) seperti ! @ # \$ & % dll. Karakter tersebut sebagian besar dapat memengaruhi akurasi proses analisis, karena bukan termasuk teks yang dapat ditemukan di kamus. Penghapusan karakter menggunakan *library regular expression* milik *python*

#### 2. *Case Folding*

*Case Folding* Proses menyamaratakan seluruh kata, dengan mengkonversi huruf kapital di dataset menjadi huruf kecil

#### 3. *Tokenzation*

*Tokenzation* Suatu proses pemotongan atau penguraian teks menjadi per kata. *Stopwords Removal* Penerapan *stop word* bertujuan untuk menghapus kata-kata yang dianggap tidak mempunyai arti penting, yang kemungkinan berpengaruh terhadap kecepatan dan kinerja proses analisis.



Teks Asli	ayo pilkada jangan takut corona
Setelah Tokenisasi	'ayo', 'pilkada', 'jangan', 'takut', 'corona'

**Gambar 9** Proses *tokenization*

#### 4. *Stemming*

*Stemming* merupakan langkah menghilangkan imbuhan berupa awalan, akhiran maupun campuran (awalan-akhiran) menjadi bentuk kata dasar sesuai kaidah bahasa Indonesia

Sebelum di <i>Stemming</i>	Sesudah di <i>Stemming</i>
Mengatur	
Diatur	Atur
Teratur	
Bayangkan	
Membayangkan	Bayang
Bayangan	
Berkerumun	
Kerumunan	Kerumun

**Gambar 10** Proses *stemming*

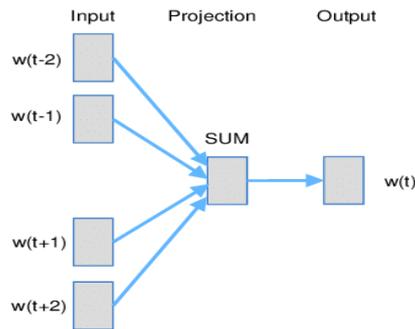
#### 2.2.4 FastText Vectors

Algoritma pembelajaran mesin dan arsitektur pembelajaran mendalam tidak dapat memproses input teks. Teks biasa harus dikonversi ke bentuk numerik sebelum dapat diklasifikasikan. Kita dapat menggunakan metode penyisipan kata untuk tujuan ini. Dengan kata lain, *Word2vec* adalah memahami dan memvektor kata menjadi dokumen; dengan kata lain, kata-kata dengan makna yang sama dalam konteks tertentu menunjukkan kedekatan.

*Word2vec* adalah model yang diusulkan oleh mikolov. Metode penyisipan kata berbasis prediksi ini terdiri dari beberapa jenis yaitu (*CBOW*) dan *Skip-gram*, namun yang paling banyak digunakan adalah menggunakan model *Continuous Bag-of-Word*. Secara umum kedua algoritma pembelajaran ini menghadirkan



input, lapisan proyeksi, dan lapisan output, perbedaannya hanya terletak pada lokasi derivasi output (Jang, Kim and Kim, 2019). Berikut adalah gambar *CBOW* yang kami terapkan yaitu:



**Gambar 11** *Continuous Bag-of-Word (CBOW)*

Gambar 2 menunjukkan lapisan input, lapisan proyeksi, dan lapisan output, tetapi proses deviasi output berbeda. Lapisan input menerima  $W_n = \{W(t-2), W(t-1), \dots, W(t+1), W(t+2)\}$  sebagai argumen. di mana  $W_n$  sesuai dengan lapisan kata. Menunjukkan *proyeksi array* multidimensi dari vektor dan menyimpan jumlah dari beberapa vektor. Lapisan keluaran kemudian mengeluarkan hasil vektor dari lapisan proyeksi (Jang, Kim and Kim, 2019). Secara khusus, *CBOW* mirip dengan model bahasa jaringan saraf.

Model ini menyajikan kata-kata dan konteksnya secara eksplisit dalam sebuah matriks, model yang dibangun di atas konsep prediksi ini mengandalkan jendela yang relatif sempit yang meluncur melalui korpus, dengan mengubah bobot jaringan, model pembelajaran memprediksi kata saat ini yang diberikan kata konteks (Mandera, Keuleers and Brysbaert, 2017). Model ini tidak hanya dapat meningkatkan akurasi vektor, tetapi juga secara signifikan meningkatkan kecepatan pelatihan (Xu *et al.*, 2019).

### 2.2.5 CNN (Convolutional Neural Network)

*CNN* adalah sebuah algoritma dari *Deep Learning* yang digunakan untuk pemrosesan gambar. Selain pemrosesan gambar juga termasuk dalam pemrosesan teks. *CNN* menyatukan lapisan dan kecanggihannya karena memberikan arsitektur standar untuk memetakan kalimat dengan panjang variabel menjadi kalimat vektor yang tersebar dengan ukuran tetap (Tul *et al.*, 2017). Ada beberapa

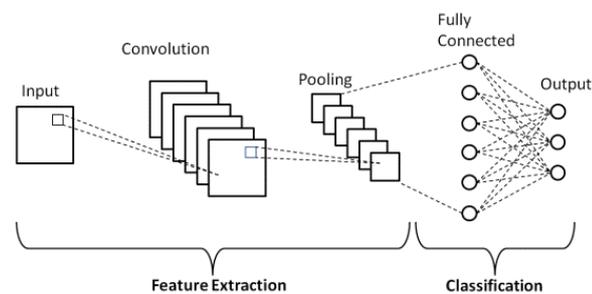
utama dalam arsitektur *CNN* yaitu :

Alat konvolusi yang memisahkan dan mengidentifikasi berbagai fitur untuk dianalisis dalam proses yang disebut Ekstraksi Fitur.



2. Jaringan ekstraksi fitur ini terdiri dari banyak pasang lapisan *convolutional* atau *pooling*.
3. *Fully connected layer* memanfaatkan keluaran dari proses konvolusi dan memprediksi kelas citra berdasarkan fitur yang diekstraksi pada tahap sebelumnya.
4. Model ekstraksi fitur *CNN* ini bertujuan untuk mengurangi jumlah fitur yang ada dalam sebuah dataset, menciptakan fitur baru yang merangkum fitur yang ada yang terkandung dalam satu set fitur asli.

Ada banyak lapisan *CNN* seperti yang ditunjukkan pada diagram arsitektur *CNN* (MK Gurucharan, 2022):



**Gambar 12** Arsitektur *CNN*

### 1. *Convolutional Layer*

Lapisan ini merupakan lapisan pertama yang digunakan untuk mengambil berbagai fitur dari citra masukan. Dalam lapisan ini, dilakukan operasi konvolusi matematis antara gambar input dan filter dengan ukuran tertentu, misalnya  $M \times M$ . Dengan memindahkan filter di seluruh gambar input, dilakukan perkalian titik antara filter dan bagian gambar input yang sesuai dengan ukuran filter ( $M \times M$ ).

Hasilnya disebut sebagai peta fitur, yang memberikan informasi tentang gambar seperti sudut dan tepi. Selanjutnya, peta fitur ini diteruskan ke lapisan lain untuk mempelajari fitur-fitur lain dari gambar masukan. Lapisan konvolusi dalam *CNN* mengirimkan keluarannya ke lapisan berikutnya setelah menerapkan operasi

i pada input. Lapisan konvolusional dalam *CNN* sangat penting karena hubungan spasial antara piksel tetap utuh.



## 2. Pooling Layer

Umumnya, setelah *Convolutional Layer*, diikuti oleh *Pooling Layer*. Fokus utama dari lapisan ini adalah untuk mengurangi ukuran peta fitur yang dihasilkan oleh lapisan konvolusi, sehingga mengurangi beban komputasi. Lapisan ini mencapai tujuannya dengan mengurangi koneksi antara lapisan dan beroperasi secara independen pada setiap peta fitur. Terdapat beberapa jenis operasi *Pooling* tergantung pada metode yang digunakan. Pada dasarnya, operasi ini merangkum fitur yang dihasilkan oleh lapisan konvolusi..

Dalam *Max Pooling*, elemen terbesar dipilih dari peta fitur. Pada *Average Pooling*, rata-rata elemen dihitung dalam bagian gambar/teks dengan ukuran sebelumnya yang telah ditentukan. Jumlah total elemen dihitung dalam *Sum Pooling*. *Layer Pooling* umumnya berfungsi sebagai penghubung antara *Convolutional Layer* dan *FC Layer*.

Model *CNN* ini menggeneralisasi fitur yang diekstrak oleh lapisan konvolusi, dan membantu jaringan dalam mengenali fitur secara otonom. Dengan bantuan ini, komputasi dalam jaringan juga dapat dikurangi.

## 4. Fully Connected Layer

Lapisan *Fully Connected (FC)* terdiri dari bobot, bias, dan neuron yang digunakan untuk menghubungkan neuron antara dua lapisan yang berbeda. Lapisan ini umumnya ditempatkan sebelum lapisan output dan menjadi bagian dari beberapa lapisan terakhir dalam Arsitektur *CNN*.

Dalam konteks ini, gambar/teks input dari lapisan sebelumnya diubah menjadi bentuk vektor dan disampaikan ke lapisan *FC*. Vektor yang dihasilkan kemudian melewati beberapa lapisan *FC* lagi, di mana operasi fungsi matematika biasanya dilakukan.

Pada tahap ini, proses klasifikasi dimulai. Alasan untuk menghubungkan an adalah karena dua lapisan yang terhubung secara penuh cenderung lebih baik daripada satu lapisan yang terhubung. Lapisan-lapisan ini dalam membantu mengurangi kebutuhan pengawasan manusia.



#### 4. Dropout

Biasanya, saat semua fitur terhubung ke lapisan *FC*, hal ini dapat menyebabkan overfitting pada dataset pelatihan. Overfitting terjadi ketika model bekerja sangat baik pada data pelatihan, tetapi berdampak negatif pada kinerja model saat digunakan pada data baru.

Untuk mengatasi masalah ini, digunakan lapisan *dropout* di mana beberapa *neuron* dijatuhkan dari jaringan saraf selama proses pelatihan, sehingga ukuran model menjadi lebih kecil. Dengan *dropout* 0,3, misalnya, 30% dari node dikeluarkan secara acak dari jaringan saraf.

*Dropout* membantu meningkatkan kinerja model pembelajaran mesin dengan mencegah *overfitting* melalui penyederhanaan jaringan. Hal ini dilakukan dengan menjatuhkan beberapa *neuron* dari jaringan saraf selama pelatihan.

#### 5. Activation Function

Fungsi aktivasi merupakan salah satu parameter penting dalam model *CNN*. Fungsi ini digunakan untuk mempelajari dan mendekati hubungan yang kompleks dan kontinu antara variabel dalam jaringan. Secara sederhana, fungsi aktivasi ini menentukan informasi mana yang harus diteruskan ke depan dalam model dan mana yang tidak.

Fungsi aktivasi memberikan non-linearitas ke jaringan. Terdapat beberapa fungsi aktivasi umum yang digunakan, seperti *ReLU*, *Softmax*, *tanH*, dan *Sigmoid*. Setiap fungsi memiliki kegunaan tertentu. Pada model *CNN* untuk klasifikasi biner, fungsi *sigmoid* dan *softmax* lebih disukai, sedangkan *softmax* umumnya digunakan untuk klasifikasi multi-kelas. Dalam hal yang sederhana, fungsi aktivasi dalam model *CNN* menentukan apakah suatu *neuron* harus diaktifkan atau tidak. Fungsi ini memutuskan apakah masukan tersebut penting atau tidak untuk diprediksi menggunakan operasi matematika.



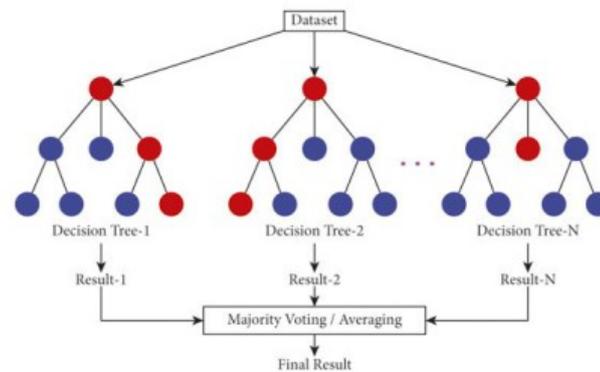
#### Algoritma *Random Forest*

*Random Forest* adalah algoritma pembelajaran diawasi yang digunakan untuk masalah regresi dan klasifikasi. *Random Forest* berupa kumpulan pohon dan pohon berbeda satu sama lain. Hal tersebut membangun beberapa pohon

keputusan dan akhirnya menggabungkannya bersama-sama untuk mendapatkan nilai absolut dan stabil, yang terutama digunakan pada saat melatih dan menghasilkan output kelas (Karthika, Murugeswari and Manoranjithem, 2019).

### Cara Kerja Algoritma *Random Forest*

Random Forest bekerja dalam dua fase. Fase pertama yaitu menggabungkan sejumlah  $N$  decision tree untuk membuat *Random Forest*. Kemudian fase kedua adalah membuat prediksi untuk setiap *tree* yang dibuat pada fase pertama.



**Gambar 13** Cara kerja algoritma *Random Forest*

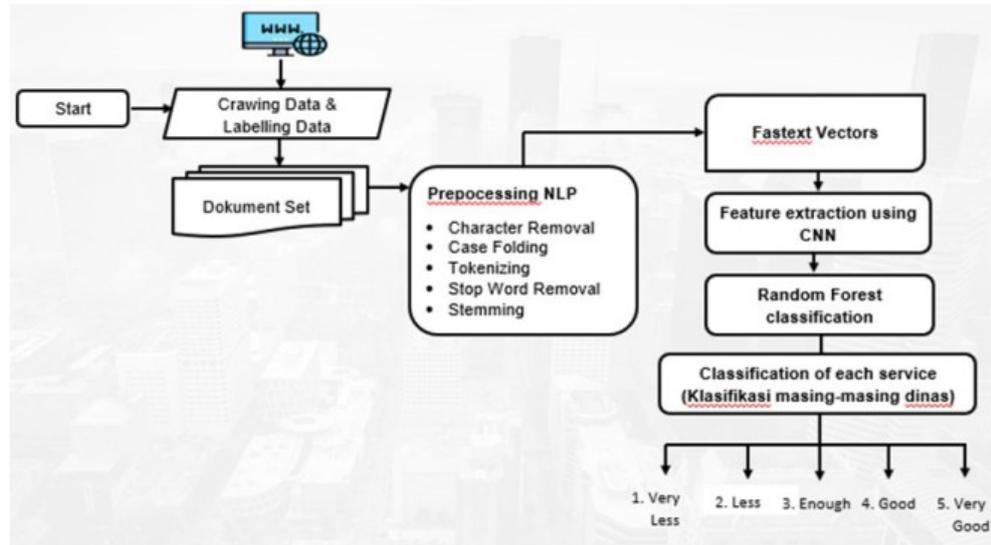
Cara kerja algoritma *Random Forest* dapat dijabarkan dalam langkah-langkah berikut:

1. Algoritma memilih sampel acak dari dataset yang disediakan.
2. Membuat *decision tree* untuk setiap sampel yang dipilih. Kemudian akan didapatkan hasil prediksi dari setiap *decision tree* yang telah dibuat.
3. Dilakukan proses voting untuk setiap hasil prediksi. Untuk masalah klasifikasi menggunakan modus (nilai yg paling sering muncul), sedangkan untuk masalah regresi akan menggunakan mean (nilai rata-rata).
4. Algoritma akan memilih hasil prediksi yang paling banyak dipilih (vote terbanyak) sebagai prediksi akhir.



### ram Alur Penelitian

apun alur penelitian dari analisis sentimen menggunakan *CNN* (*Convolutional Neural Networks*) dan *RF* (*Random Forest*) adalah sebagai berikut:



**Gambar 14** Diagram alur klasifikasi multi-target

*Crawling* data diambil dari berbagai sumber web baik itu dari komentar media sosial, *Facebook*, *Twitter*, *Instagram*, dan berita di *Youtube* dengan menggunakan bahasa pemrograman *Python* juga aplikasi pendukung untuk menarik data yaitu *web scrapper*. Kemudian data diberi label, pemberian label data dilakukan secara manual dengan bantuan pakar. Data diberi 5 label sentiment dan 16 klasifikasi dinas yang akan disatukan kedalam satu dokumen dalam bentuk format csv yang akan di proses.

Setelah itu kita menggunakan data tadi untuk melakukan *preprocessing NLP* yang terdiri dari *character removal*, *case folding*, *tokenizing*, *stop word removal*, dan *stemming* sebelum akhirnya di klasifikasi, kita melewati beberapa tahap setelah *preprocessing* yaitu *fasttext vector* yaitu salah satu library pada *Phyton* untuk mengubah text menjadi bentuk vektor. Kita hanya mengambil bagian hingga fitur CNN yang akan di ekstraksi dan dimasukkan dalam klasifikasi *Random Forest*

Pada tahap klasifikasi, setelah melewati tahap *Preprocessing* dari *NLP*, *fasttext vector*, dan ekstraksi fitur *CNN*, selanjutnya data diklasifikasikan dengan pendekatan *Deep Learning* yaitu menggunakan algoritma *Random Forest* sehingga mengeluarkan output klasifikasi 16 ke masing-masing dinas dan output 5 kategori sentiment yaitu sangat baik, baik, cukup, kurang dan sangat kurang.



## II.4. Metode Penyelesaian Masalah

### 2.4.1 Metode Klasifikasi

Klasifikasi terkait dengan dinas-dinas sesuai dengan tiap keluhan dan sentimennya yaitu menggunakan klasifikasi *NLP (Natural language processing)* Menurut Textmetrics, *NLP* digunakan untuk mengukur sentimen dan menentukan bagian mana dari bahasa manusia yang penting. (*NLP*) adalah cabang dari kecerdasan buatan yang berhubungan dengan interaksi antara komputer dan manusia menggunakan bahasa alami [12].

### 2.4.2 Metode Analisis

Klasifikasi analisis sentimen dengan pendekatan Modifikasi Algoritma *Deep Learning* bagian *NLP* menggunakan algoritma *Convolutional Neural Networks (CNN)*. *CNN* merupakan bagian dari *Artificial Neural Network (ANN)* yang mampu mendeteksi informasi dengan akurasi tinggi (Kastrati *et al.*, 2021), begitu juga dengan algoritma *RF (Random Forest Classifier)*.

Bagian Algoritma yang penulis akan modifikasi yaitu pada bagian *Preprocessing NLP (Natural Language Processing)* pada tahapan cleansing data untuk menghilangkan tanda baca, karakter, symbol, emoticon, hastag, link url dll yang dapat mempengaruhi tingkat akurasi bukan hanya itu penulis juga melakukan beberapa penambahan parameter/kategori yang biasanya hanya 3 kategori yaitu positif, netral negative. Penulis menambah menjadi 5 kategori yaitu sangat baik, baik, cukup, kurang, sangat kurang.

Pada document data set analisis sentiment *Deep Learning CNN* Penulis mengubah bobot angka dari bilangan biner pada *CNN* menjadi bobot angka *decimal* yaitu berkisar dari bobot 1-5, dimana 5 = Sangat Baik, 4 = Baik, 3 = Cukup, 2 = Kurang, 1 = Sangat Kurang. Dimana bobot 1, 2 dapat kita namakan sentiment negative, bobot 3 yaitu netral dan bobot 4 dan 5 yaitu positif sebagai bentuk penyederhanaan dari teks untuk mengecek prediksi dari teks *maining*.



Selain itu penulis juga akan mencoba mengutak-ngatik/mengganti angka bobot dari model yang kita terapkan pada *layer (type) CNN* dan juga melakukan berbagai *epoch/* jangka waktu secara bertahap untuk meningkatkan tingkat akurasi yang tinggi selanjutnya penulis juga coba

membandingkan dengan algoritma *Random Forest* biasa tanpa modifikasi lalu mencari cara penggabungan antara keduanya dalam rangka mendapatkan tingkat akurasi yang lebih baik.

### 2.2.3 Metode Pengujian Sistem

Menurut (Vanaja, 2018), kita dapat menilai keefektifan setiap algoritma klasifikasi dengan memeriksa skor presisi, daya ingat, akurasi, dan F1-nya. Metode pengujian ini juga dapat digunakan untuk mengevaluasi sistem klasifikasi dengan beberapa kelas selain klasifikasi biner. Pendekatan ini disebut *confusion matrix* yang akan diuji dengan cara berikut:

1. Akurasi yaitu mengevaluasi kedekatan nilai prediksi dengan nilai sebenarnya. Klasifikasi data yang benar memungkinkan untuk menentukan keakuratan hasil prediksi.
2. Presisi adalah pendekatan evaluasi yang mengukur proporsi informasi terkait yang diperoleh oleh suatu sistem dibandingkan dengan keseluruhan informasi yang diambil, terlepas dari relevansinya.
3. Recall adalah teknik evaluasi yang mengukur jumlah informasi terkait yang diambil oleh sistem dibandingkan dengan jumlah total informasi relevan yang tersedia dalam pengumpulan informasi, terlepas dari apakah sistem mengambilnya.
4. F1-Score adalah harmonic mean dari precision dan recall :

Berikut ini adalah rumus persamaan dari confusion matrix:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = 100\% \quad (7)$$

$$Precision = \frac{TP}{FP+TP} = 100\% \quad (8)$$

$$Recall = \frac{TP}{FN+TP} = 100\% \quad (9)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision+Recall} = 100\% \quad (10)$$



P yaitu jumlah data positif yang terklasifikasi dengan benar oleh *system*  
 N yaitu jumlah data negatif yang terklasifikasi dengan benar oleh *system*  
 N yaitu jumlah data negatif namun terklasifikasi salah oleh *system*

4. FP FN yaitu jumlah data Positif namun terklasifikasi salah oleh *system*

### 2.2.4 State of the Art Penelitian

Penelitian-penelitian terkait yang telah dilakukan sebelumnya dapat dilihat pada Tabel 1 dibawah ini.

**Tabel 1 State of the Art**

No	Judul Karya Ilmiah, Nama, Tahun Terbit Dan Penerbit	Objek Dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <=,>
1.	<b>Topik yang diusulkan :</b> Multi-Target Klasifikasi Teks Berdasar Tanggapan Masyarakat Kota Dengan Menggunakan Algoritma <i>CNN &amp; Random Forest</i>	<b>Objek :</b> Keluhan Masyarakat  <b>Masalah :</b> Kondisi pengelolaan pengaduan pelayanan publik di setiap Dinas penyelenggara belum menunjukkan kondisi pengelolaan yang efektif dan terintegrasi.	Klasifikasi Teks menggunakan metode <i>CNN &amp; Random Forest Algorithm</i>	Penulis membuat system analisis klasifikasi teks multi-target berdasar tanggapan masyarakat kota dengan menggabungkan dua algoritma sebagai evaluasi pemerintah untuk meningkatkan kinerja kota.	
2.	<b>Judul :</b> Performance Evaluation of Transformation of Industrial Ecology in the Resource-Based Cities  <b>Nama :</b> Sun Xiumei, Zhou Min  <b>Tahun :</b> 2010  <b>Penerbit :</b> IEEE	<b>Objek :</b> Ekologi Industri  <b>Masalah :</b> Banyak kota yang bergantung pada sumber daya alam mengalami kesulitan dalam menghadapi penipisan sumber daya dan kinerja kota yang menurun.	<i>Analytic Hierarchy Process (AHP)</i>	<i>AHP</i> adalah sebuah teknik yang ilmiah dan rasional untuk mengevaluasi kinerja, dan sangat penting untuk digunakan dalam transformasi ekologi industri di kota-kota yang bergantung pada sumber daya alam. Namun, meskipun metode ini bermanfaat, masih ada beberapa kelemahan yang perlu diperhatikan (Sun and Zhou, 2010).	
3.	<b>Judul :</b> Analisis  asan k di emko igan	<b>Objek :</b> Pemko Pekanbaru  <b>Masalah :</b> Dalam rangka mewujudkan tata	<i>Regression Analysis General Linear Model</i>	Dari hasil analisis data, ditemukan bahwa rata-rata tingkat kepuasan masyarakat terhadap pelayanan publik di	

No	Judul Karya Ilmiah, Nama, Tahun Terbit Dan Penerbit	Objek Dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <,=,>
	<p>Menggunakan <i>Regression Analysis General Linear Model</i></p> <p><b>Nama :</b> Warnia Nengsih</p> <p><b>Tahun :</b> 2013</p> <p><b>Penerbit :</b> ReserachGate</p>	<p>kelola yang baik (<i>good governance</i>) di lingkungan pemerintah, ranah pelayanan publik menjadi sebuah isu dasar yang harus menjadi fokus dalam birokrasi pemerintahan.</p>		<p>lingkungan Pemerintah Kota (Pemko) Pekanbaru adalah 2,68 dari skala 4 dengan nilai rata-rata yang diperoleh untuk setiap variabel yang telah didefinisikan dengan jelas (X1-X15). Selain itu, terungkap bahwa variabel yang paling mempengaruhi kepuasan masyarakat adalah variabel X1, yang terkait dengan prosedur pelayanan (Nengsih, 2017).</p>	
4	<p><b>Judul :</b> Smart City performance measurement framework CITYkeys</p> <p><b>Nama :</b> Miimu Airaksinen, Isabel Pinto Seppä, Aapo Huovila, Peter Bosch</p> <p><b>Tahun :</b> 2017</p> <p><b>Penerbit :</b> IEEE</p>	<p><b>Objek :</b> Kota Eropa</p> <p><b>Masalah :</b> Populasi kota dengan kepadatan tinggi meningkatkan beban energi, transportasi, air, bangunan, dan ruang publik</p>	<i>Key Performance Indicators (KPIs)</i>	<p>Sebagian besar KPI proyek dan kota telah diuji dalam beberapa sumber daya data, dan sebagian besar KPI proyek (72%) dan KPI kota (82%) telah dinyatakan berhasil. Untuk KPI kota kuantitatif, tingkat ketersediaan data rata-rata di kota-kota Eropa diharapkan lebih dari 70% (sekitar 25% sebagai data terbuka) dan mendekati 100% untuk semua KPI kota yang bersifat kualitatif (Airaksinen <i>et al.</i>, 2018)</p>	
	 <p>laint stem</p>	<p><b>Objek :</b> Public Complaint</p> <p><b>Masalah :</b></p>	<i>Google API Integration Web/Android:</i>	<p>Dalam aplikasi ini, warga dimudahkan untuk menyampaikan</p>	

No	Judul Karya Ilmiah, Nama, Tahun Terbit Dan Penerbit	Objek Dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <=,>
	<p>using web services and android</p> <p><b>Nama :</b> Mareeswari V, Gopalakrishnan V</p> <p><b>Tahun :</b> 2017</p> <p><b>Penerbit :</b> IOP Publishing</p>	<p>Saat ini, badan penyelenggara utama di kota MC yang bekerja sama telah mengetahui masalah dan kekurangan yang terjadi di kota melalui pemasangan sensor/kamera CCTV. Selain itu, warga kota juga dapat mengajukan keluhan melalui beberapa media yang tersedia.</p>	<ul style="list-style-type: none"> <li>• <i>Google Places API</i></li> <li>• <i>Firebase Cloud Messaging (FCM)</i></li> </ul>	<p>pengaduan langsung dari smartphone kepada pejabat tinggi. Banyak <i>API</i> yang berfungsi sebagai layanan web yang sangat penting untuk mempermudah pendaftaran keluhan seperti <i>Google Places API</i> untuk mendeteksi lokasi Anda saat ini dan menampilkannya di Peta. Portal Web digunakan untuk memproses berbagai keluhan dengan baik didukung dengan layanan web yang berbeda(Mareeswari and Gopalakrishnan, 2017).</p>	
6.	<p><b>Judul :</b> Smart City Service System Engineering Based on Microservices Architecture Case Study: Government of Tangerang City</p> <p><b>Nama :</b> Taufiq Hidayat, Suhardi, Novianto Budi Kurniawan</p> <p><b>Tahun :</b> 2017</p> <p><b>Penerbit :</b> IEEE</p>	<p><b>Objek :</b> Pemerintah Kota Tangerang</p> <p><b>Masalah :</b> permasalahan perkotaan menjadi pelik, kemacetan lalu lintas, banjir, sampah, dan lain-lain semakin menumpuk. Sebuah kota pintar sangat bergantung pada layanan, dan ini menjadi tujuan utama di mana semua kombinasi dan integrasi layanan menjadi dasar untuk kota yang benar-benar cerdas</p>	<p><b>Service System Engineering Framework :</b> Metodologi Layanan Rekayasa Sistem pada penelitian ini 3 tahapan</p> <ol style="list-style-type: none"> <li>1.Mengidentifikasi</li> <li>2.Merancang</li> <li>3.Mengembangkan</li> </ol>	<p>Aplikasi ini memudahkan warga untuk mengirimkan keluhan melalui smartphone secara langsung ke pejabat yang berwenang. Untuk mempermudah pendaftaran keluhan, terdapat berbagai <i>API</i> yang berfungsi sebagai layanan web penting, seperti <i>Google Places API</i> yang digunakan untuk mendeteksi lokasi saat ini dan menampilkannya di Peta. <i>Portal web</i> juga digunakan untuk memproses berbagai keluhan</p>	



No	Judul Karya Ilmiah, Nama, Tahun Terbit Dan Penerbit	Objek Dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <,=>
				dan didukung dengan berbagai layanan web (Hidayat, Suhardi and Kurniawan, 2018).	
7.	<p><b>Judul :</b> A holistic evaluation of smart city performance in the context of China</p> <p><b>Nama :</b> Liyin Shen, Zhenhua Huang, Siu Wai Wong, Shiju Liao, Yingli Lou</p> <p><b>Tahun :</b> 2018</p> <p><b>Penerbit :</b> Journal of Cleaner Production</p>	<p><b>Objek :</b> Pemerintahan China</p> <p><b>Masalah :</b> Perkembangan urbanisasi yang pesat telah menghadirkan berbagai permasalahan, seperti kekurangan pasokan energi, polusi lingkungan, kemacetan lalu lintas, kesenjangan sosial, kurangnya akses atau ketersediaan layanan publik, serta berkurangnya lahan yang tersedia.</p>	<p><b>Entropi dan Teknik Order Teknik Preference by Similarity to Ideal Solution (TOPSIS)</b></p>	<p>Secara keseluruhan, tingkat kinerja kota pintar di Cina masih relatif rendah. Selain itu, terdapat ketidakseimbangan yang signifikan dalam kinerja lima dimensi kota pintar, yakni infrastruktur cerdas, tata kelola, sumber daya manusia, ekonomi, dan lingkungan (Shen <i>et al.</i>, 2018).</p>	
8.	<p><b>Judul :</b> Towards Smart Riyadh: Riyadh Wiki Information And Complaining System</p> <p><b>Nama :</b> Eman Al Helal, Hala Mokhtar</p> <p><b>Tahun :</b> 2018</p> <p><b>Penerbit :</b></p>	<p><b>Objek :</b> Pengaduan <i>Wiki Riyadh</i></p> <p><b>Masalah :</b> Banyak tantangan yang dihadapi oleh pemerintah, seperti masalah transparansi antara pemerintah dan masyarakat, kepercayaan masyarakat pada pemerintah, proses pengambilan keputusan, dan akuntabilitas. Penyelesaian masalah ini</p>	<p><b>1. Pendekatan crowd sourcing</b> memungkinkan warga bertindak sebagai sumber data untuk mendukung pemerintah dan memperbaiki kota mereka.</p> <p><b>2. Pendekatan co-design</b> menjadi platform open source yang memungkinkan warga bekerja sama untuk membangun sistem dan menambahkan layanan baru.</p>	<p>Dengan sistem ini, masalah pemerintahan seperti transparansi, kepercayaan, pengambilan keputusan, dan akuntabilitas dapat dipecahkan dengan biaya yang lebih murah dan meningkatkan kualitas hidup warga negara (Al Helal and Mokhtar, 2018).</p>	



No	Judul Karya Ilmiah, Nama, Tahun Terbit Dan Penerbit	Objek Dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <,>
	Technology (IJMIT)	membutuhkan biaya yang besar dan memerlukan upaya serta sumber daya yang signifikan.			
9.	<p><b>Judul :</b> Multi-Class Classification of Turkish Texts with Machine Learning Algorithms</p> <p><b>Nama :</b> Fatih Gürçan</p> <p><b>Tahun :</b> 2018</p> <p><b>Penerbit :</b> IEEE</p>	<p><b>Objek :</b> <i>Turkish Texts</i></p> <p><b>Masalah :</b> Masalah klasifikasi teks merupakan sebuah proses dimana dokumen teks yang diawasi diberikan tugas untuk dikategorikan ke dalam satu atau lebih kategori atau kelas yang telah ditentukan sebelumnya</p>	<i>Machine Learning</i>	Kinerja klasifikasi dari algoritma <i>Multinomial Naïve Bayes</i> , <i>Bernoulli Naïve Bayes</i> , Dukungan Mesin Vektor, <i>K-Nearest Neighbor</i> , dan Pohon Keputusan pada teks berita dalam bahasa Turki telah dibandingkan dan dievaluasi berdasarkan hasil yang diperoleh dengan menggunakan parameter yang berbeda. Hasil penelitian menunjukkan bahwa algoritma <i>Multinomial Naïve Bayes</i> menghasilkan tingkat keberhasilan klasifikasi terbaik, yaitu sekitar 90%(Gurcan, 2018).	
10.	<p><b>Judul :</b> User-oriented representation of Smart Cities indicators to support citizens &amp; governments decision-making processes</p>	<p><b>Objek :</b> Amerika Latin</p> <p><b>Masalah :</b> Diperlukan penetapan metrik pada layanan publik untuk memahami cara meningkatkan kinerja dan mengoptimalkannya, serta untuk mengetahui bagaimana teknologi dapat meningkatkan</p>	<ul style="list-style-type: none"> <li>• <i>Metodologi difusi inovasi</i></li> <li>• <i>Key Performance Indicators (KPI)</i></li> </ul>	Dalam proses menyesuaikan platform dengan kasus penggunaan, sebuah metodologi diusulkan yang mudah diikuti oleh kota untuk membangun model indikator Kota Cerdas mereka. (Limon-Ruiz <i>et al.</i> , 2019).	



No	Judul Karya Ilmiah, Nama, Tahun Terbit Dan Penerbit	Objek Dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <,=,>
	Maciel, Raul Beltran  <b>Tahun : 2019</b>  <b>Penerbit : IEEE</b>	efisiensi layanan publik.			
11.	<b>Judul :</b> Public Complaint Sorting Using Image Processing  <b>Nama :</b> Manisha Adhude, Faizan Sayyed, Wassyatullah Sayed, Assistant Prof. Kanchan V. Wankhade  <b>Tahun :</b> 2019  <b>Penerbit :</b> Journal of Analysis and Computation (JAC)	<b>Objek :</b> Public Complaint  <b>Masalah :</b> Warga harus melewati prosedur dan formalitas yang panjang untuk melaporkan keluhan mereka. Namun, tidak ada jaminan bahwa otoritas kota akan menangani atau menyelesaikan keluhan tersebut. Selain itu, masyarakat biasanya sibuk dengan pekerjaan sehari-hari sehingga tidak memiliki inisiatif untuk mendaftarkan masalah.	<b>Image Processing – SVM :</b> Setelah pengaduan terdaftar, seluruh warga dapat memberikan suara terhadapnya dan pengaduan dengan prioritas tertinggi akan diprioritaskan terlebih dahulu dalam penyelesaiannya, sementara pengaduan dengan prioritas rendah akan ditangani dengan waktu yang lebih lama.	Dalam makalah ini, diusulkan sebuah arsitektur konseptual yang serbaguna, fleksibel, dan hemat biaya untuk memantau masalah warga. Aplikasi ini menggunakan pembelajaran mesin dan pemrosesan gambar untuk memungkinkan warga mendaftarkan atau memposting masalah sipil mereka secara online. Warga juga dapat memastikan bahwa masalah mereka akan diselesaikan dalam waktu tertentu yang diberikan oleh mereka atau oleh otoritas (Manisha Adhude, Faizan Sayyed, Wassyatullah Sayed, 2019).	
12.	<b>Judul :</b> Tweets Responding to the Indonesian Government's Handling of COVID-19:  lysis ith oly	<b>Objek :</b> Covid-19  <b>Masalah :</b> Trending di Twitter di Indonesia adalah penanganan COVID-19 (Corona Virus Disease-2019). Ada banyak pendapat pro dan kontra dalam	<b>SVM</b> Salah satu algoritma klasifikasi	Metode analisis SVM pada sentimen mengenai aspek umum dengan dataset dua kelas mencapai performa paling optimal dalam hal akurasi rata-rata, presisi, recall, dan f-measure dengan	





No	Judul Karya Ilmiah, Nama, Tahun Terbit Dan Penerbit	Objek Dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <=,>
	Instagram Instansi Pemerintah  <b>Nama :</b> Muslih, Akhmad Rizaldy, Ajib Susanto, Ibnu Utomo Wahyu Mulyono  <b>Tahun :</b> 2020  <b>Penerbit :</b> Semnas LPPM	masyarakat dan pembuat kebijakan memerlukan umpan balik dari masyarakat sebagai sumber data dan informasi sebagai upaya dalam meningkatkan kinerja dan pelayanan	instansi pemerintah pada media social Instagram	netral terhadap suatu postingan. Metode yang digunakan dalam analisis ini mencapai performa yang baik dengan akurasi mencapai 79 persen (Muslih, Akhmad Rizaldy, Ajib Susanto, Ibnu Utomo Wahyu Mulyono, 2020).	
15.	<b>Judul :</b> Tibetan Text Classification Method Based on BiLSTM Model  <b>Nama :</b> Li Jia, Tao Jiang, Jia Hao Meng, TingTing Zhang  <b>Tahun :</b> 2020  <b>Penerbit :</b> IEEE	<b>Objek :</b> <i>Tibetan</i> Teks  <b>Masalah :</b> Dibutuhkan Klasifikasi teks untuk memecahkan masalah kekacauan informasi	<b>BiLSTM</b>	Dari hasil pengklasifikasian, terbukti bahwa penggunaan representasi teks yang menggabungkan metode <i>Word2vec</i> dan <i>TF-IDF</i> berdasarkan kelas varians frekuensi secara efektif meningkatkan efek klasifikasi teks. Keakuratan pengklasifikasi teks <i>Tibet</i> berdasarkan penggunaan <i>BiLSTM</i> dapat mencapai 89,03%, yang jauh lebih baik dibandingkan dengan penggunaan <i>RNN LSTM</i> (Jia et al., 2020).	
16.	<b>Judul :</b> A Deep Learning Sentimen Analyser for Social Media Comments in Low-Resource	<b>Objek :</b> Situasi pandemi  <b>Masalah :</b> Dalam masa pandemi, menjaga jarak fisik menjadi penting. Di tengah situasi ini, platform media sosial telah	<b>Deep Learning</b>	Kami mengusulkan analisis sentimen yang memiliki kinerja yang baik dan bahkan melampaui klasifikasi dasar pada kumpulan data yang dikumpulkan. Secara khusus, skor	



No	Judul Karya Ilmiah, Nama, Tahun Terbit Dan Penerbit	Objek Dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <,=>
	Kurti, Fatbardh Kadriu, Doruntina Murtezaj, Fatbardh Gashi  <b>Tahun : 2021</b>  <b>Penerbit :</b> Elektronics	menjadi salah satu tempat di mana orang-orang dapat menyampaikan pendapat, pikiran, sentimen, dan emosi mereka terkait dengan pandemi.		F1 sebesar 72,09% berhasil dicapai dengan menggabungkan mekanisme perhatian lokal dengan <i>BiLSTM</i> (Kastrati <i>et al.</i> , 2021).	
17	<b>Judul :</b> A Novel Deep Learning Approach of Convolutional Neural Network and Random Forest Classifier for Fine-grained Sentiment Classification  <b>Nama :</b> Siji George C. G., B. Sumathi  <b>Tahun :</b> 2021  <b>Penerbit :</b> International Journal on Electrical Engineering and Informatics	<b>Objek :</b> Online Shop  <b>Masalah :</b> Kebanyakan konsumen meninggalkan komentar mereka secara online, baik di platform media sosial maupun di situs <i>web e-commerce</i> . Komentar tersebut kemudian diposting, dibagikan, di-tweet, dan ditinggalkan pada halaman produk atau toko online.	<b>Convolutional Neural Network (CNN), Random Forest (RF) &amp; CNNRF</b>	Dari hasil analisis, terlihat bahwa model yang diusulkan memiliki kinerja yang lebih baik dibandingkan dengan model <i>CNN</i> dan <i>RF</i> secara individual untuk klasifikasi sentimen yang lebih terperinci. Selain itu, strategi dropout juga berkontribusi dalam meningkatkan kemampuan model <i>CNNRF</i> untuk menghasilkan performa yang lebih baik secara umum. (Siji George and Sumathi, 2021)	

**Note : baris terakhir berisi usulan penelitian**

**< kinerja topik yang diusulkan lebih buruk**

**= kinerja topik yang diusulkan sama, dan dapat digunakan metode penyelesaian masalahnya**

**> kinerja topik yang diusulkan akan lebih baik karena metodenya juga lebih baik**



·dasarkan dari penjelasan pada bagian latar belakang, penjelasan penelitian an uraian pada tabel State Of The Art diatas, diperoleh hasil mengenai

tantangan dan peluang untuk penelitian ini berkaitan dengan analisis sentiment, klasifikasi dan analisis survey, yaitu :

- a. Penelitian (Sun and Zhou, 2010), (Airaksinen *et al.*, 2018) mengalami kesulitan dalam pengumpulan data yang efektif di tingkat kabupaten dan kota karna lokalisasi data diambil secara manual dan satu-satu pada tiap lokasi yang dituju pada unit pemerintahan sehingga data yang diambil memakan waktu yang lama serta tidak seimbang jika ada data yang tidak tersedia. Pada penelitian kali ini penulis mencoba mengambil data lansung dari tanggapan masyarakat kota di sosial media secara cepat dan tepat.
- b. Pada penelitian (Sun and Zhou, 2010), (Shen *et al.*, 2018), (Limon-Ruiz *et al.*, 2019), (Mareeswari and Gopalakrishnan, 2017), (Hidayat, Suhardi and Kurniawan, 2018), (Al Helal and Mokhtar, 2018), (Manisha Adhude, Faizan Sayyed, Wassyatullah Sayed, 2019), (Prastyo *et al.*, 2020), (Muslih, Akhmad Rizaldy, Ajib Susanto, Ibnu Utomo Wahyu Mulyono, 2020) , (Kastrati *et al.*, 2021), (Siji George and Sumathi, 2021), (Listyarini and Anggoro, 2021), (Aidasani, Bhadkamkar and Kashyap, 2017), (Nengsih, 2017) belum tersedianya data set terkait tanggapan masyarakat yang ada di kota. Pada penelitian ini penulis mencoba *crawling*/ menarik data pada media sosial dan menyediakan data set secara lengkap terkait tanggapan masyarakat kota dalam bentuk format file .csv
- c. Penelitian (Prastyo *et al.*, 2020), (Muslih, Akhmad Rizaldy, Ajib Susanto, Ibnu Utomo Wahyu Mulyono, 2020), (Kastrati *et al.*, 2021), (Putri *et al.*, 2020) adanya pengaruh akurasi dari metode sentiment yang berbeda-beda. Pada penelitian ini tingkat akurasi yang dihasilkan dari metode-metode tersebut masih bisa ditingkatkan dengan menggunakan metode yang diusulkan peneliti yaitu modifikasi *CNN* dan algoritma *Random Forest* dan penggabungan dari kedua metode tersebut.
- d. Penelitian (Nengsih, 2017), hasil penelitian tidak lebih akurat karna Unit Pelayanan Teknis yang dijadikan sebagai objek survey dibatasi. Pada penelitian ini tanggapan masyarakat kota pada dinas-dinas terkait tidak batasi pada suatu lokasi dan unit.



- e. Penelitian (Prastyo *et al.*, 2020), (Muslih, Akhmad Rizaldy, Ajib Susanto, Ibnu Utomo Wahyu Mulyono, 2020), (Kastrati *et al.*, 2021), (Putri *et al.*, 2020) parameter analisis sentiment sangat umum dan kurang begitupun klasifikasi (Gurcan, 2018), (Jia *et al.*, 2020), (Siji George and Sumathi, 2021). Pada penelitian ini menambahkan beberapa parameter dari umumnya 3 parameter menjadi 5 parameter dan klasifikasi menjadi 16 kategori dinas.
- f. Penelitian (Kastrati *et al.*, 2021) jumlah data yang digunakan tidak seimbang karna kelas netral memiliki lebih dari setengah komentar. Pada penelitian ini penulis mencoba untuk menyeimbangkan komentar pada proses *crawling* data (penarikan data) pada media sosial dan menyeimbangkan data dengan teknik *oversampling*.
- g. Analisis sentiment (Prastyo *et al.*, 2020), (Muslih, Akhmad Rizaldy, Ajib Susanto, Ibnu Utomo Wahyu Mulyono, 2020), (Kastrati *et al.*, 2021), (Putri *et al.*, 2020) dan klasifikasi teks (Gurcan, 2018), (Jia *et al.*, 2020) menggunakan algoritma tersendiri. Belum adanya gabungan algoritma dari keduanya. Pada penelitian ini penulis akan menggabungkan dua metode dari analisis sentiment dan klasifikasi dinas serta menggabungkan keduanya menjadi klasifikasi multi target.



### 2.2.5 Tingkat Keaslian (Level Orisinalitas) Topik Penelitian yang diusulkan

Pengembangan topik penelitian yang diusukan ini dapat dilihat dari Tabel 2. Tabel ini berisi kesesuaian antara referensi pada tabel State of the Art dengan metode penyelesaian yang digunakan dalam penelitian.

**Tabel 2 Matriks metode penyelesaian**

No. Referensi dari tabel 2	Metode Penyelesaian																
	a	b	c	d	e	f	G	h	I	J	k	l	m	n	o	p	q
Metode yang diusulkan	✓															✓	✓
2		✓															
3			✓														
4				✓													
5					✓												
6						✓											
7							✓										
8								✓									
9									✓								
10										✓							
11											✓						
12												✓	✓				
13												✓	✓				
14														✓			
15															✓		
16																✓	
17																	✓

#### Keterangan:

##### Nomor Referensi:

1. Sistem Evaluasi Kinerja Kota Berdasar Tanggapan Masyarakat Dengan Kombinasi Analisis Sentimen Dan Analisis Survey
2. Performance Evaluation Of Transformation Of Industrial Ecology In The Resource-Based Cities
3. Analisis Data Mining Tingkat Kepuasan Layanan Publik Di Lingkungan Pemko Pekanbaru Dengan Menggunakan Regresion Analysis General Linear Model  
Smart City Performance Measurement Framework Citykeys



5. Complaint Go: An Online Complaint Registration System Using Web Services And Android
6. Smart City Service System Engineering Based On Microservices Architecture Case Study: Government Of Tangerang City
7. A Holistic Evaluation Of Smart City Performance In The Context Of China
8. Towards Smart Riyadh: Riyadh Wiki Information And Complaining System
9. Multi-Class Classification Of Turkish Texts With Machine Learning Algorithms
10. User-Oriented Representation Of Smart Cities Indicators To Support Citizens & Governments Decision-Making Processes
11. Public Complaint Sorting Using Image Processing
12. Tweets Responding To The Indonesian Government's Handling Of COVID-19: Sentimen Analysis Using SVM With Normalized Poly Kernel
13. Analisis Sentimen Opini Masyarakat Terhadap Pembangunan Infrastruktur Kota Malang Melalui Twiter Dengan Menggunakan Metode Support Vector Machine
14. Tibetan Text Classification Method Based On Bilstm Model
15. Implementasi Algoritma Boyer Moore Untuk Analisis Sentimen Komentar Di Akun Instagram Instansi Pemerintah
16. A Deep Learning Sentimen Analyser For Social Media Comments In Low-Resource Languages
17. A Novel Deep Learning Approach Of Convolutional Neural Network And Random Forest Classifier For Fine-Grained Sentiment Classification

Metode Penyelesaian:

- a. Deep Learning & Klasifikasi
  - b. Analytic Hierarchy Process (AHP)
- Regression Analysis General Linear Model
- Key Performance Indicators (KPIs)
- Google API Integration



- f. Service System Engineering Framework
- g. Entropi dan Teknik Order Teknik Preference by Similarity to Ideal Solution (TOPSIS)
- h. Pendekatan crowd sourcing & co-design
- i. Machine Learning
- j. Metodologi difusi inovasi & Key Performance Indicators (KPI)
- k. Image Processing
- l. SVM with Normalized Poly Kernel
- m. SVM (Support Vector Machine)
- n. Boyer Moore
- o. BiLSTM Model
- p. Deep Learning
- q. CNNRF

Berdasarkan pengembangan metode yang dapat dilihat dari tabel 2. topik penelitian ini memenuhi level keaslian tingkat 4 sesuai tingkat keaslian dari pedoman penulisan proposal yang dikeluarkan oleh program studi Magister Informatika UNHAS (tabel 3). Tingkat keaslian ini dapat dipenuhi karena dari judul yang di usulkan “Multi-Target Klasifikasi Teks Berdasar Tanggapan Masyarakat Kota Dengan Menggunakan Algoritma *CNN* & *Random Forest*“ merupakan dikategorikan sebagai metode penyelesaian pengembangan karena menggabungkan beberapa metode lama (poin b) untuk menyelesaikan objek dan permasalahan yang sama



**Tabel 3 Level orisinalitas berdasarkan standar proposal program studi magister teknik informatika universitas hasanuddin**

Level orisinalitas	Objek penelitian	Jenis permasalahan	Metode penyelesaian masalah	Status orisinal proposal/penelitian
1	xxxxx	Tidak jelas, sangat minim untuk level S2	xxxxxx	Orisinalitas tidak ada
2	Terpakai (sudah digunakan sebelumnya)	Serupa dengan objek yang sama	Sudah digunakan untuk objek dan jenis permasalahan sebelumnya	Sangat tidak orisinal (replikasi)
3	Baru	Serupa dengan sebelumnya	Sama untuk masalah yang serupa	Kurang
4	Terpakai (sudah digunakan sebelumnya)	Serupa dengan objek yang sama	Lain yang tersedia (belum digunakan untuk masalah yang sama)	Minimalis
5	Baru	Serupa dengan sebelumnya	Lain yang tersedia	Minimalis
6	Terpakai	Baru	Tersedia	Orisinal
7	Baru	Baru	Tersedia	Orisinal
8	Terpakai	Serupa	Baru	Orisinal+Novelty
9	Baru	Serupa	Baru	Orisinal+Novelty
10	Terpakai	Baru	Baru	Sangat Orisinal+Novelty
11	Baru	Baru	Baru	Sangat Orisinal+Novelty

Kategori metode penyelesaian baru:

- Modifikasi dari metode yang lama
- Penggabungan beberapa metode lama
- Metode yang baru dikembangkan dan atau belum pernah digunakan sebelumnya
- Selalu menghasilkan keluaran yang lebih baik

### 2.2.6 Target Hipotesis Penelitian

Berdasarkan penelitian yang akan dikerjakan maka target penelitian an mampu menghasilkan suatu analisis multi target yaitu sentiment n masyarakat kota beserta klasifikasi dinas. Penelitian ini juga diharapkan memodifikasi algoritma yang ada terkait analisis sentiment dan



mengelompokkan tiap-tiap tanggapan/keluhan ke dinas-dinas yang terkait dengan tugasnya yang akan menjadi bahan evaluasi bagi pemerintahan dan sebagai sumber pengetahuan bagi akademisi dan peneliti-peneliti selanjutnya yang ingin melakukan penelitian serupa





Optimized using  
trial version  
[www.balesio.com](http://www.balesio.com)