

**KLASIFIKASI SENTIMEN TERHADAP DATA TEKS
JEJARING SOSIAL DENGAN TOPIK
PEMBELAJARAN DARING MENGGUNAKAN *RIDGE
REGRESSION***



OLEH:

NAUVAL FARIS FADHLULLAH

09011281621037

**JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2021**

**KLASIFIKASI SENTIMEN TERHADAP DATA TEKS
JEJARING SOSIAL DENGAN TOPIK
PEMBELAJARAN DARING MENGGUNAKAN *RIDGE
REGRESSION***

TUGAS AKHIR

**Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer**



Oleh:

NAUVAL FARIS FADHLULLAH

09011281621037

**JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA**

2021

LEMBAR PENGESAHAN

KLASIFIKASI SENTIMEN TERHADAP DATA TEKS JEJARING SOSIAL DENGAN TOPIK PEMBELAJARAN DARING MENGGUNAKAN RIDGE REGRESSION

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat

Memperoleh Gelar Sarjana Komputer

Oleh:

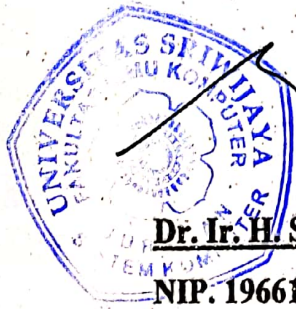
NAUVAL FARIS FADHLULLAH

09011281621037

Indralaya, Agustus 2021

Mengetahui,

Ketua Jurusan Sistem Komputer



Dr. Ir. H. Sukemi, M.T.

NIP. 196612032006041001

Pembimbing Skripsi,

Huda Ubaya, S.T., M.T.

NIP. 198106162012121003

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

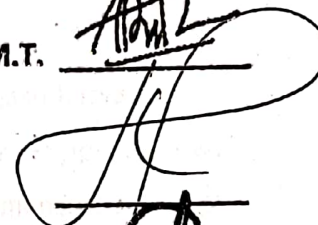
Hari : Senin

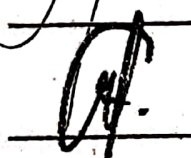
Tanggal : 2 Agustus 2021

Tim Penguji :

1. Ketua : Sarmayanta Sembiring, S.Si., M.T. 

2. Sekretaris : Aditya Putra P Prasetyo, S.Kom., M.T. 

3. Pembimbing : Huda Ubaya, S.T., M.T. 

4. Penguji : Ahmad Zarkasi, M.T. 

Mengetahui,

Ketua Jurusan Sistem Komputer




Dr. Ir. H. Sukemi, M.T.

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Nauval Faris Fadhlullah

NIM : 09011281621037

Judul : Klasifikasi Sentimen Terhadap Data Teks Jejaring Sosial Dengan Topik Pembelajaran Daring Menggunakan Ridge Regression

Hasil Penyecekan *Software iThenticate/Turnitin* : 10 %

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan tugas akhir ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya.

Demikian, pernyataan ini saya buat dalam keadaan sadar dan tanpa paksaan dari siapapun.



Indralaya, Agustus 2021



Nauval Faris Fadhlullah
NIM. 09011281621037

KATA PENGANTAR

Assalamualaikum Warahmatullahi Wabarakatuh,

Puji dan syukur penulis panjatkan kepada Allah Subhanahu Wa Ta'ala, atas segala karunia, berkat, dan rahmat-Nya sehingga penulis dapat menyelesaikan penyusunan Tugas Akhir ini dengan judul "**Klasifikasi sentimen terhadap data teks jejaring sosial dengan topik pembelajaran daring menggunakan *Ridge Regression***".

Dalam Tugas Akhir ini penulis menjelaskan tentang bagaimana cara mengklasifikasi Sentimen dengan menggunakan metode *ridge regression* disertai dengan data-data yang penulis peroleh saat melakukan pengujian. Penulis berharap tulisan ini dapat bermanfaat bagi orang banyak, dan menjadi tambahan bahan bacaan dan referensi untuk para akademisi dan peneliti lain yang juga berada atau sedang menekuni bidang textmining, khususnya pengusaha dan pemerintah.

Pada penyusunan laporan ini, penulis mendapatkan banyak ide dan saran serta bantuan dari berbagai pihak baik secara langsung maupun tidak langsung. Oleh karena itu, pada kesempatan ini penulis mengucapkan rasa syukur kepada Allah Subhanahu Wa Ta'ala dan terimakasih kepada yang terhormat:

1. Ibu penulis Wili Nusa Nara, Ayah penulis Febi Aries Edmardinata, Kedua adik penulis Bintang Muhammad Fadhil dan Muhammad Rayyan Farisly, Sepupu penulis Aditya, Kakek-Nenek, Paman-Bibi, beserta seluruh keluarga yang selalu mendoakan, memberikan dukungan baik moril maupun materil, serta motivasi dan semangat selama hidup penulis.
2. Bapak Jaidan Jauhari, S.Pd. M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Bapak Dr. Ir. H.Sukemi, M.T. selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Alm Bapak Dr. Reza Firsandaya Malik, M.T. selaku Pembimbing Tugas Akhir,
5. Bapak Huda Ubaya, M.T selaku Pembimbing Tugas Akhir dan Pembimbing Akademik Jurusan Sistem Komputer.

6. Bapak Sarmayanta Sembiring, S.Si., M.T., Bapak Aditya Putra P Prasetyo, S.Kom., M.T., dan Ahmad Zarkasi, M.T., selaku Tim Penguji Tugas Akhir Penulis.
7. Seluruh teman-teman SK16 Indralaya khususnya Kelas SK16B.
8. Ilham Solehan, Rizki Ronaldo, Jerry Pratama Saputra, Hisyam Sanusi, Edo Herdianto, dan Siti Jasmine Harnitiah sebagai teman-teman seperjuangan tugas akhir, baik dalam pencarian literatur, pemrograman, serta diskusi mengenai riset dan lainnya.
9. Teman - teman dari Jurusan Sistem Komputer yang tidak bisa disebutkan satu-persatu. Khususnya seluruh teman-teman dari kelas SK16B Indralaya.
10. Para sahabat penulis, Hanif Habibie Supriansyah, Aditya Jovandi, Sarianto Praga Sanjaya, Jerry Pratama Saputra, Hisyam Sanusi, Ichwanul Hakim, Ilham Solehan, Rap Nur Muhammad, Madyus Randika & Muhammad Sultan Alif Utama yang selalu memotivasi penulis untuk cepat lulus.
11. Dan semua pihak yang telah membantu.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari kesempurnaan, oleh karena itu penulis sangat mengharapkan kritik dan saran yang bersifat membangun agar lebih baik lagi dikemudian hari.

Akhir kata dengan segala keterbatasan, penulis berharap semoga laporan ini menghasilkan sesuatu yang bermanfaat bagi kita semua khususnya bagi mahasiswa Fakultas Ilmu Komputer Universitas Sriwijaya secara langsung ataupun tidak langsung sebagai sumbangan pemikiran dalam peningkatan mutu pembelajaran dan penelitian.

Indralaya, Juli 2021


Penulis

KLASIFIKASI SENTIMEN TERHADAP DATA TEKS JEJARING SOSIAL DENGAN TOPIK PEMBELAJARAN DARING MENGUNAKAN RIDGE REGRESSION

Nauval Faris Fadhlullah (09011281621037)

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

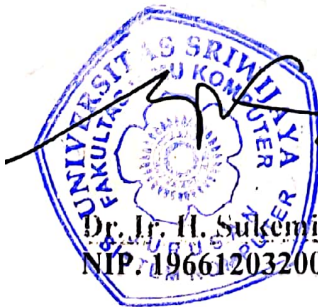
Email : nauvalfaris77@gmail.com

ABSTRAK

Di masa pandemi, sistem dalam jaringan atau online, menjadi hal yang lumrah pada bidang-bidang kehidupan masyarakat. Salah satu bidang yang turut menggunakan pembelajaran daring adalah sektor pendidikan. Pemerintah sebagai pemegang kebijakan tertinggi memutuskan untuk menerapkan pembelajaran daring bagi lembaga-lembaga pendidikan di Indonesia. Hal ini membuat masyarakat memberikan pandangan dan opini yang beragam tentang pembelajaran daring. Pada penelitian ini telah dilakukan percobaan *text mining* berupa sentimen analisis terhadap data teks berupa tweet-tweet masyarakat tentang pembelajaran daring yang berasal dari *platform* Twitter. Data tersebut didapatkan dari Drone Emprit Academic, dengan ketentuan waktu data yang digunakan adalah sejumlah 120128 data dengan rentang waktu 16 Maret 2020 – 4 Juni 2021. Data yang terbagi menjadi tiga sentimen berupa negatif, netral, dan positif tersebut diolah untuk diklasifikasikan sentimennya menggunakan metode *Ridge Regression*.

Kata Kunci : *text mining*, sentimen analisis, pembelajaran daring, *ridge regression*

Ketua Jurusan Sistem Komputer



Dr. Ir. H. Sukemi, M.T.
NIP. 196612032006041001

Indralaya, Juli 2021
Pembimbing Tugas Akhir

Huda Ubaya, S.T., M.T.
NIP. 198106162012121003

**CLASSIFICATION OF SENTIMENTS ON SOCIAL NETWORK TEXT DATA
WITH ONLINE LEARNING TOPIC USING RIDGE REGRESSION**

Nauval Faris Fadhlullah (09011281621037)

*Dept. of Computer Engineering, Faculty of Computer Science, Sriwijaya
University*

Email : nauvalfaris77@gmail.com

ABSTRACT

During the pandemic, online learning have become commonplace in all areas of people's lives. One of the fields that also uses online learning is the education sector. The government as the highest policy holder decided to implement online learning for educational institutions in Indonesia. This makes people provide diverse views and opinions about online learning. In this study, text mining experiments have been carried out in the form of sentiment analysis of text data in the form of public tweets about online learning originating from the Twitter platform. The data was obtained from Drone Emprit Academic, with the provision that the data time used was 120128 data with a time span of March 16, 2020 - June 4, 2021. The data which is divided into three sentiments in the form of negative, neutral, and positive is processed to classify the sentiment using the Ridge method. Regression.

Keywords: *text mining, sentiment analysis, online learning, ridge regression*

Head of Department Computer Engineering



Dr. Ir. H. Sukemi, M.T.
NIP. 196612032006041001

Indralaya, July 2021
Supervisor

Huda Ubaya, S.T., M.T.
NIP. 198106162012121003

DAFTAR ISI

	Halaman
Halaman Judul	i
Halaman Pengesahan	ii
Halaman Persetujuan	iii
Halaman Pernyataan	iv
Kata Pengantar	v
Abstrak	vii
<i>Abstract</i>	viii
Daftar Isi	ix
Daftar Gambar	xii
Daftar Tabel	xiv
Daftar Lampiran	xv
Bab I. Pendahuluan	1
1.1 Latar Belakang	1
1.2 Tujuan	2
1.3 Manfaat.....	2
1.4 Perumusan Masalah	2
1.5 Batasan Masalah	2
1.6 Metodologi Penelitian	3
1.7 Sistematika Penulisan	3

Bab II. Tinjauan Pustaka	5
2.1 Analisis Sentimen	5
2.2 Jejaring Sosial	5
2.3 Pembelajaran Daring	6
2.4 API Twitter <i>Dataset</i>	6
2.5 <i>Natural Language Processing (NLP)</i>	7
2.6 <i>Machine Learning</i>	8
2.7 <i>Text Mining</i>	8
2.8 <i>Ridge Regression</i>	8
2.9 <i>K-Fold Cross Validation</i>	9
2.10 <i>Confusion Matrix</i>	9
2.10.1 Presisi	10
2.10.2 <i>Recall</i>	10
2.10.3 <i>F1score</i>	10
 Bab III. Metodologi Penelitian	 12
3.1 Pendahuluan	12
3.2 Kerangka Kerja	12
3.3 Pengambilan Data.....	14
3.4 <i>Preprocessing</i>	16
3.4.1 <i>Cleaning</i>	17
3.4.2 <i>Stemming</i>	18
3.4.3 <i>Word tokenizing</i>	20

3.4.4 Pemisahan fitur dan label	21
3.5 Ekstrasi fitur	21
3.6 <i>Cross Validation</i>	22
3.7 <i>Confusion Matrix</i>	22
Bab IV. Hasil dan Pembahasan	24
4.1 Pendahuluan	24
4.2 <i>Text Dataset</i>	24
4.3 Proses <i>Preprocessing</i>	24
4.4 Ekstrasi Fitur Data Teks	27
4.5 Klasifikasi dengan <i>Ridge Regression</i>	27
4.5.1 Eksperimen dengan 20% <i>Data Test</i>	29
4.5.2 Eksperimen dengan 30% <i>Data Test</i>	32
4.5.3 Eksperimen dengan 40% <i>Data Test</i>	35
4.5.4 Eksperimen dengan 50% <i>Data Test</i>	38
4.5.5 Eksperimen dengan 60% <i>Data Test</i>	41
4.5.6 Eksperimen dengan 70% <i>Data Test</i>	44
4.5.7 Eksperimen dengan 80% <i>Data Test</i>	47
4.6 Perbandingan hasil pengujian <i>Data Test</i>	50
4.7 Analisis	51
Bab V. Kesimpulan	53
5.1 Kesimpulan	53

5.1 Saran 53

DAFTAR PUSTAKA

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Diagram alur penelitian	13
Gambar 3.2 Bentuk <i>Dataset</i> yang diambil dari Drone Emprit Academy ...	14
Gambar 3.3 <i>Wordcloud</i> sentimen negatif	15
Gambar 3.4 <i>Wordcloud</i> sentimen netral	16
Gambar 3.5 <i>Wordcloud</i> sentimen positif	16
Gambar 3.6 Diagram tahap <i>preprocessing</i>	17
Gambar 3.7 Diagram proses <i>Cleaning</i>	18
Gambar 3.8 Sentimen yang diproses <i>Cleaning</i>	18
Gambar 3.9 Diagram proses <i>Stemming</i>	19
Gambar 3.10 Sentimen yang diproses <i>Stemming</i>	20
Gambar 3.11 Diagram proses <i>Tokenizing</i>	20
Gambar 3.12 Sentimen yang diproses <i>Tokenizing</i>	21
Gambar 3.13 Total data tiap masing-masing tweet	21
Gambar 3.14 Kata-kata yang menjadi skalar.....	22
Gambar 4.1 Diagram prosedur <i>preprocessing</i>	25
Gambar 4.2 Hasil tahap <i>Cleaning</i> terhadap bentuk data teks	26
Gambar 4.3 Hasil tahap <i>Stemming</i> terhadap bentuk data teks	26
Gambar 4.4 Hasil tahap <i>Tokenizing</i> terhadap bentuk data teks	26
Gambar 4.5 Data sampel untuk ekstrasi fitur dari 350 hingga 365	27

Gambar 4.6 Hasil percobaan dengan 20% <i>Data Test</i>	29
Gambar 4.7 Hasil percobaan dengan 30% <i>Data Test</i>	32
Gambar 4.8 Hasil percobaan dengan 40% <i>Data Test</i>	35
Gambar 4.9 Hasil percobaan dengan 50% <i>Data Test</i>	38
Gambar 4.10 Hasil percobaan dengan 60% <i>Data Test</i>	41
Gambar 4.11 Hasil percobaan dengan 70% <i>Data Test</i>	44
Gambar 4.12 Hasil percobaan dengan 80% <i>Data Test</i>	47
Gambar 4.13 Nilai Presisi untuk setiap <i>Data Test</i>	50
Gambar 4.14 Nilai <i>Recall</i> untuk setiap <i>Data Test</i>	51
Gambar 4.15 Nilai <i>F1score</i> untuk setiap <i>Data Test</i>	51

DAFTAR TABEL

	Halaman
3.1 Data kata singkat	14
3.2 <i>Confusion Matrix</i> 3x3	22
4.1 Pembagian Data <i>Training</i> dan Data <i>Test</i>	28
4.2 Hasil perbandingan uji antara jumlah Data <i>Test</i> yang berberda	50

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Infeksi virus corona (Covid-19) menular dengan sangat cepat hingga organisasi kesehatan dunia menetakannya sebagai pandemi pada 11 November 2020 [1]. Merespon hal tersebut, lembaga pendidikan mengubah tata cara pelaksanaan kegiatan menjadi pembelajaran daring [2].

Pelaksanaan pembelajaran daring menghasilkan hasil yang beragam pada tiap lembaga pendidikan. Ada lembaga pendidikan yang berhasil mendapatkan hasil yang baik dengan pembelajaran daring, dan ada pula lembaga pendidikan yang tidak mendapat hasil baik, lalu mengubah kembali cara pembelajaran menjadi tatap muka dengan melaksanakan protokol kesehatan, dan tetap mendapatkan hasil yang baik. Hal ini membuat sebagian masyarakat menilai pelaksanaan pendidikan tidak perlu dilakukan secara daring.

Guna mengetahui tingkat keefektifitasan pembelajaran daring bagi masyarakat Indonesia, penulis akan mencoba mengklasifikasikan sentimen masyarakat dengan menggunakan Natural Language Processing pada data Tweet media sosial Twitter yang penulis dapatkan dari Drone Emprit Academic. Analisis sentimen didefinisikan sebagai ekstraksi informasi dari sepotong teks. Analisis sentimen berarti menceritakan keadaan pikiran pembicara, penulis esai, atau subjek lain tentang fakta atau ekstremitas umum yang relevan atau jawaban antusias terhadap sebuah laporan, komunikasi. Penggunaan estimasi bias dalam analisis data dan pembuatan model dibahas [3]. Tinjauan teori Ridge Regression dan hubungannya dengan regresi terbalik umum disajikan bersama dengan hasil percobaan simulasi dan tiga contoh penggunaan Ridge Regression dalam praktik [4]. Dari penelitian tersebut menunjukkan bahwa metode Ridge Regression menghasilkan akurasi yang cukup baik. Berdasar penelitian tersebut, penulis memutuskan untuk menggunakan

Ridge Regression sebagai metode klasifikasi sentiment analisis dan menggunakan bahasa pemrograman Python. Dengan begitu klasifikasi sentimen yang berupa data teks dapat dilaksanakan.

Keuntungan penggunaan regresi ridge dibandingkan metode lain yaitu regresi ridge mengurangi dampak multikolinearitas dengan menentukan penduga yang bias tetapi mempunyai varians yang lebih kecil dari varians penduga regresi linear berganda [5].

1.2 Tujuan

Penelitian yang dilakukan oleh penulis bertujuan untuk:

1. Mengklasifikasi sentimen dari data berbentuk teks jejaring sosial dengan menggunakan *Ridge Regression*.
2. Mengklasifikasi sentimen dari data berbentuk teks jejaring sosial menjadi 3 jenis sentimen, yaitu negatif, netral, dan positif.

1.3 Manfaat

Penelitian yang dilakukan oleh penulis bertujuan untuk mendapatkan manfaat berupa:

1. Mendapatkan pemahaman tentang pandangan tentang Pembelajaran Daring.
2. Mendapatkan hasil akhir berupa data sentimen, yang dapat digunakan sebagai bahan pertimbangan oleh pemerintah, berkenaan dengan kebijakan untuk masyarakat terkait Pembelajaran Daring.

1.4 Perumusan Masalah

Permasalahan yang dihadapi pada penelitian ini ialah minimnya ketersediaan informasi klasifikasi sentimen terhadap pandangan masyarakat Indonesia tentang Pembelajaran Daring.

1.5 Batasan Masalah

1. Jejaring sosial yang dianalisis sentimennya adalah data teks yang didapatkan dari Drone Emprit Academic, data teks merupakan tweet-

tweet masyarakat tentang Pembelajaran Daring pada jejaring sosial Twitter dari kurun waktu 16 maret 2020 - 4 juni 2021.

2. Pengklasifikasian pada penelitian ini menggunakan data teks berbahasa Indonesia.

1.6 Metodologi Penelitian

Penelitian ini memiliki beberapa tahapan metodologi, yaitu:

1. Perumusan masalah

Merupakan tahap penentuan permasalahan yang dihadapi berkenaan dengan proses pelaksanaan *preprocessing* data teks dan ekstrasi fitur, serta penyelesaian permasalahan tersebut.

2. Studi pustaka dan literatur

Merupakan tahap pembelajaran dan pengolahan sumber bacaan yang memberikan informasi terkait masalah yang diteliti. Dan juga tahap pencarian literatur dengan menggunakan kata kunci yang spesifik sesuai dengan penelitian.

3. Perancangan

Merupakan tahap perancangan sistem berdasarkan masalah yang diteliti. Tahap ini membahas pembuatan program untuk mengolah data teks, seperti *preprocessing*, segmentasi fitur, dan klasifikasi.

4. Pengujian

Merupakan tahap pengujian rancangan.

5. Analisis

Merupakan tahap pengambilan data hasil olahan algoritma dan program yang dibuat, kemudian menganalisisnya.

6. Kesimpulan dan Saran

Merupakan tahap penarikan kesimpulan dari hasil analisa, saran untuk penelitian selanjutnya bagi penulis atau orang lain yang melanjutkan dan mengembangkan penelitian serupa.

1.7 Sistematika Penulisan

Penelitian ini memiliki beberapa tahapan Sistematika Penulisan, yaitu:

BAB I PENDAHULUAN

Merupakan bab yang menjawab tentang apa penelitian yang dilakukan dan mengapa penelitian dilakukan. Bab ini berisikan gambaran umum permasalahan dan topik yang akan disajikan.

BAB II TINJAUAN PUSTAKA

Merupakan bab yang berisikan kerangka pemikiran dan dasar teori yang menjadi kerangka umum penelitian. Bab ini berisikan penelitian terdahulu serta literatur-literatur ilmiah.

BAB III METODOLOGI PENELITIAN

Merupakan bab yang menggambarkan tahapan dalam proses penelitian. Bab ini menyajikan sistematika prosedur yang dipakai dalam penelitian.

BAB IV HASIL DAN PEMBAHASAN

Merupakan bab yang berisikan interpretasi atas hasil dari penelitian, yang menjadi acuan untuk dijadikan sebuah teori.

BAB V KESIMPULAN

Merupakan bab yang menjadi akhir dari pembahasan pada penelitian ini. Bab ini berisikan beberapa poin berbentuk ide pemikiran yang dilandasi dari hasil penelitian.

DAFTAR PUSTAKA

- [1] K. Pengantar, "Oleh : Nama : Andriany Putri Kelas : X Mia 3."
- [2] W. A. F. Dewi, "Dampak COVID-19 terhadap Implementasi Pembelajaran Daring di Sekolah Dasar," *Edukatif J. Ilmu Pendidik.*, vol. 2, no. 1, pp. 55–61, 2020, doi: 10.31004/edukatif.v2i1.89.
- [3] U. Rofiqoh, R. S. Perdana, and M. A. Fauzi, "Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1725–1732, 2017, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/628>.
- [4] E. Cambria, S. Poria, A. Gelbukh, I. P. Nacional, and M. Thelwall, "AFFECTIVE COMPUTING AND SENTIMENT ANALYSIS Sentiment Analysis Is a Big Suitcase," *Ieee Intell. Syst.*, 2017.
- [5] X. Han, J. Wang, M. Zhang, and X. Wang, "Using social media to mine and analyze public opinion related to COVID-19 in China," *Int. J. Environ. Res. Public Health*, vol. 17, no. 8, 2020, doi: 10.3390/ijerph17082788.
- [6] L. N. Atiqoh, "Respon Orang Tua Terhadap Pembelajaran Daring Pada Masa Pandemi Covid-19," *Thufuli J. Ilm. Pendidik. Islam Anak Usia Dini*, vol. 2, no. 1, p. 45, 2020, doi: 10.33474/thufuli.v2i1.6925.
- [7] Z. Jianqiang and G. Xiaolin, "Comparison research on text pre-processing methods on twitter sentiment analysis," *IEEE Access*, vol. 5, pp. 2870–2879, 2017, doi: 10.1109/ACCESS.2017.2672677.
- [8] D. Y. Fahmi, Hartoyo, and N. Zulbainarni, "Mining Social Media (Twitter) Data for Corporate Image Analysis: A Case Study in the Indonesian Mining Industry," *J. Phys. Conf. Ser.*, 2021, doi: 10.1088/1742-

6596/1811/1/012107.

- [9] T. Wolf *et al.*, “Transformers : State-of-the-Art Natural Language Processing,” pp. 38–45, 2020.
- [10] N. Burkart and M. F. Huber, “A survey on the explainability of supervised machine learning,” *J. Artif. Intell. Res.*, vol. 70, pp. 245–317, 2021, doi: 10.1613/JAIR.1.12228.
- [11] S. A. Salloum, M. Al-Emran, A. A. Monem, and K. Shaalan, “A survey of text mining in social media: Facebook and Twitter perspectives,” *Adv. Sci. Technol. Eng. Syst.*, vol. 2, no. 1, pp. 127–133, 2017, doi: 10.25046/aj020115.
- [12] W. Wang and B.-Y. Jing, “Convergence of Gaussian process regression: Optimality, robustness, and relationship with kernel ridge regression,” 2021, [Online]. Available: <http://arxiv.org/abs/2104.09778>.
- [13] J. Xu, Y. Zhang, and D. Miao, “Three-way confusion matrix for classification: A measure driven view,” *Inf. Sci. (Ny)*, vol. 507, no. xxxx, pp. 772–794, 2020, doi: 10.1016/j.ins.2019.06.064.