



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

FORM PENGAJUAN JUDUL

Nama : Fatma Ananta Sari

NIM : 211402036

Judul diajukan oleh* : ☐ Dosen

☒ Mahasiswa



Bidang Ilmu (tuliskan dua bidang) :

1. Data Science and Intelligent System
2. Computer Graphics and Vision

Uji Kelayakan Judul** : ☐ Diterima ☐ Ditolak

Hasil Uji Kelayakan Judul :

Calon Dosen Pembimbing I: Sarah Purnamawati, S.T., M.Sc.

(Jika judul dari dosen maka dosen tersebut berhak menjadi pembimbing I)

Calon Dosen Pembimbing II: Dr. Marischa Elveny, S.TI., M.Kom.

Paraf Calon Dosen Pembimbing I

Medan,

* Centang salah satu atau keduanya

** Pilih salah satu

Fanindia Purnamasari, S.TI., M.IT.

NIP. 198908172019032023



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

RINGKASAN JUDUL YANG DIAJUKAN

*Semua kolom di bawah ini diisi oleh mahasiswa yang sudah mendapat judul

Judul / Topik Skripsi	IDENTIFIKASI PERNYATAAN STEREOTIP TERHADAP GENERASI Z PADA MEDIA SOSIAL X MENGGUNAKAN ALGORITMA GATED RECURRENT UNIT
Latar Belakang dan Penelitian Terdahulu	<p>Latar Belakang</p> <p>Stereotip telah menjadi salah satu fenomena sosial yang melekat dalam kehidupan masyarakat. Menurut Kamus Besar Bahasa Indonesia (KBBI), stereotip adalah konsepsi mengenai sifat suatu golongan berdasarkan prasangka yang subjektif dan tidak tepat. Stereotip merupakan bentuk generalisasi terhadap sekelompok orang dimana karakteristik yang identik diberikan kepada seluruh anggota kelompok tanpa mempertimbangkan adanya perbedaan, kompleksitas dan keunikan tiap individu.</p> <p>Akibat kemajuan teknologi informasi, khususnya media sosial telah mengubah cara stereotip terbentuk dan tersebar di masyarakat modern. Platform media sosial X (sebelumnya dikenal sebagai Twitter) bukan hanya berfungsi sebagai ruang interaksi digital, tetapi juga menjadi wadah utama dalam membentuk dan memperkuat stereotip. Meserole (2018) menunjukkan bahwa ada dua faktor utama yang dapat mempengaruhi penyebaran informasi menyesatkan di media sosial yaitu teknologi dan perilaku manusia. Platform seperti Twitter, Facebook dan YouTube menerapkan algoritma berbasis <i>deep learning</i> bekerja sama dengan kebiasaan pengguna dalam menggunakan media sosial yang cenderung bereaksi pada konten yang sesuai dengan prasangka dan keyakinan mereka, sehingga konten yang bersifat provokatif akan dengan cepat mendapat perhatian. Situasi ini menciptakan lingkungan digital yang kondusif bagi penyebaran stereotip dan memperkuat persepsi yang tidak berdasar secara masif dan cepat.</p> <p>Dalam konteks ini, Generasi Z (kelompok individu yang lahir antara tahun 1995 hingga 2010) menjadi salah satu kelompok yang cukup sering menerima berbagai stereotip di media sosial. Stereotip yang melekat pada Generasi Z, seperti “pemalas”, “bergantung pada teknologi”, atau “hanya menggunakan emoji dalam komunikasi”, tidak selalu</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

mencerminkan kenyataan yang sesungguhnya (Setiadji., et al 2023). Meskipun stereotip dapat bersifat positif, namun dalam beberapa kasus, stereotip positif juga dapat berdampak negatif bagi orang lain. Cheryan & Bodenhausen (2000) menunjukkan bahwa stereotip positif (misalnya, orang Asia umumnya pandai matematika) menciptakan tekanan yang tidak diinginkan pada orang Asia yang akhirnya berdampak negatif karena mengakibatkan penurunan konsentrasi dan penurunan signifikan dalam kinerja pada saat mengerjakan tes matematika (Sayama & Sayama, 2011). Hal ini menunjukkan bahwa stereotip tidak hanya berdampak pada aspek psikologis tetapi juga membatasi peluang dan potensi Generasi Z dalam berbagai aspek kehidupan.

Penelitian terkait identifikasi stereotip terhadap suatu kelompok sudah pernah dilakukan sebelumnya. Penelitian oleh (Cryan et al., 2020) melakukan deteksi stereotip terhadap gender dengan membandingkan metode Lexicon dengan Supervised Learning. Selanjutnya penelitian (Fraser et al., 2022) yang melakukan ekstraksi stereotip terhadap berbagai kelompok usia, termasuk generasi baby boomers, generasi X, milenial, dan generasi Z, serta kategori "tua" dan "muda" secara umum yang bertujuan untuk mengidentifikasi dan memahami stereotip terkait usia yang tersebar di masyarakat, khususnya pada platform media sosial X (Twitter). Lalu penelitian oleh (Bosco et al., 2023) dengan judul “Detecting racial stereotypes: An Italian social media corpus where psychology meets NLP” yang mengembangkan deteksi stereotip ras pada postingan Facebook dan Twitter berbahasa Italia.

Berdasarkan latar belakang diatas, penulis mengajukan sebuah penelitian untuk mengidentifikasi stereotip terhadap Gen Z menggunakan natural language processing yang berjudul **IDENTIFIKASI PERNYATAAN STEREOTIP TERHADAP GENERASI Z PADA MEDIA SOSIAL X MENGGUNAKAN ALGORITMA GATED RECURRENT UNIT.**

Penelitian Terdahulu



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	No.	Penulis	Judul	Tahun
	1.	Cristina Bosco, Viviana Patti, Simona Frenda, Alessandra Teresa Cignarella, Marinella Paciello, dan Francesca D'Errico	Detecting racial stereotypes: An Italian social media corpus where psychology meets NLP	2023
	2.	Kathleen C. Fraser, Svetlana Kiritchenko, dan Isar Nejadgholi	Extracting Age-Related Stereotypes from Social Media Texts	2022
	3.	Fabiola Jeldes- Delgado, Tiago Alves Ferreira, David Diaz, dan Rodrigo Ortiz	Exploring Gender Stereotypes in Financial Reporting: An Aspect-Level Sentiment Analysis Using Big Data and Deep Learning	2024
	4.	Jenna Cryan, Shiliang Tang, Xinyi Zhang, Miriam Metzger, Haitao Zheng, dan Ben Y. Zhao	Detecting Gender Stereotypes: Lexicon vs. Supervised Learning Methods	2020
	5.	Prashant Kapil dan Asif Ekba	A deep neural network based multi-task learning approach to hate speech detection	2020



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

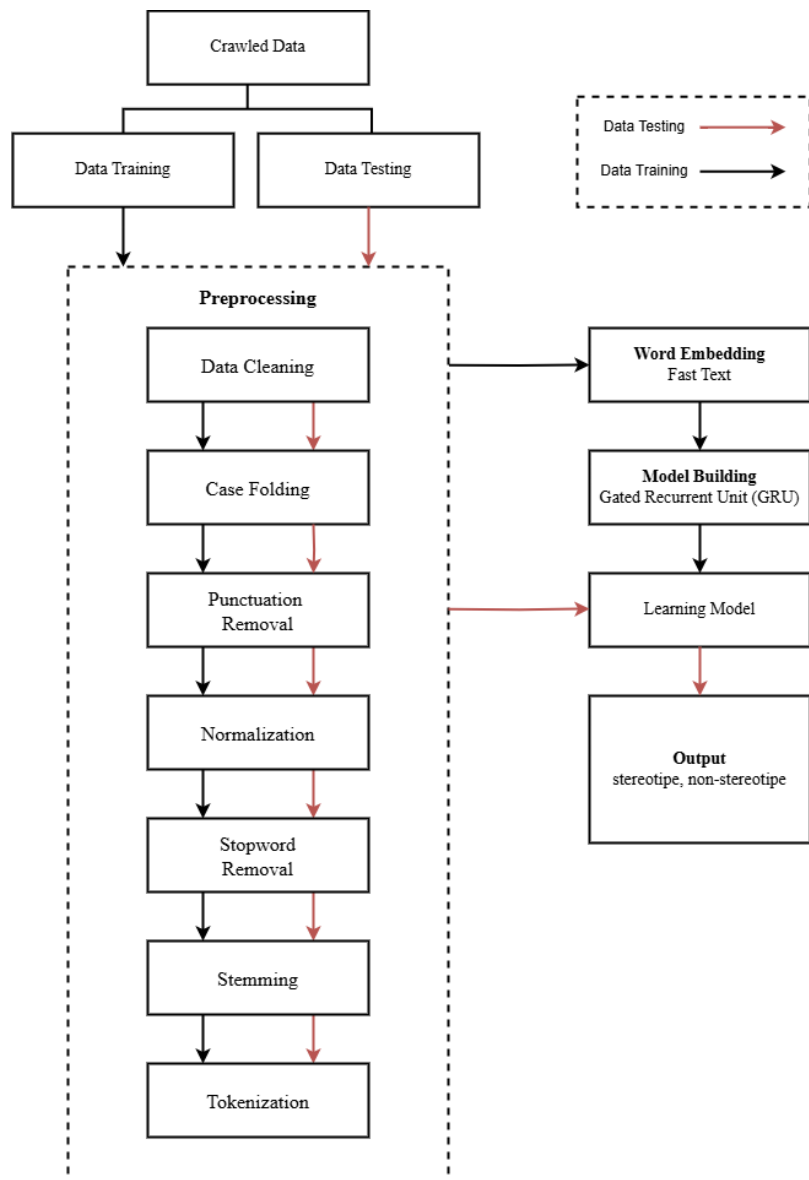
Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Rumusan

Masalah

Stereotip terhadap Generasi Z yang sering kali muncul di sosial media ternyata memiliki dampak signifikan, terutama dalam membentuk persepsi publik. Banyak dari stereotip tersebut cenderung bias dan tidak akurat, yang pada akhirnya mempengaruhi cara masyarakat berinteraksi dengan Generasi Z, baik dalam konteks pekerjaan, pendidikan, maupun hubungan sosial. Persepsi yang keliru ini dapat berdampak negatif pada kesempatan dan pengalaman Gen Z di berbagai aspek kehidupan. Oleh karena itu, diperlukan pendekatan yang dapat mengotomatisasi proses mengidentifikasi pernyataan stereotip terhadap generasi Z secara efisien dan akurat.

Metodologi





KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Penelitian ini akan melewati beberapa tahapan proses. Tahapan pertama yaitu pengumpulan data dari media sosial X(Twitter) dan disimpan dalam file berekstensi .csv (*comma separated values*). Selanjutnya masuk kedalam tahap *preprocessing*. Pada tahapan ini, dataset akan melewati proses *Cleaning*, *Case Folding*, *Punctual Removal*, *Normalization*, *Stopword Removal*, *Stemming* dan *Tokenization*. Data hasil tokenisasi akan memasuki tahapan *word embedding*, dimana data diubah kedalam bentuk vektor menggunakan *fastText*. Setelah itu, *matrix* yang berisikan vektor akan dijadikan sebagai input terhadap model untuk melalui proses *training* dan dipakai untuk mengidentifikasi pernyataan stereotipe terhadap Gen Z dengan menggunakan data *testing*.

1. Data input

Data yang digunakan dalam penelitian ini merupakan kumpulan komentar pengguna media sosial X (Twitter). Data yang dikumpulkan merupakan komentar dalam bahasa Indonesia yang berkaitan dengan stereotipe terhadap Gen Z. Setelah dikumpulkan, data akan disimpan dalam file berekstensi .csv (*comma separated values*).

2. *Splitting Data*

Dataset akan dibagi kedalam dua dataset, yaitu data *training* dan data *testing*. Data *training* akan digunakan dalam proses membangun model, sedangkan data *testing* digunakan untuk menguji model.

3. *Data Labelling*

Setelah data dibagi, selanjutnya akan dilakukan pelabelan pada data *training*. Data komentar akan diberi label stereotipe dan non-stereotipe.

4. *Preprocessing*

Tahapan *preprocessing* merupakan aspek penting yang perlu diperhatikan dalam melakukan *mining data*. Pada tahap inilah data mentah diproses dan diolah terlebih dahulu. Hal ini berguna untuk mempersiapkan data agar lebih siap dan lebih mudah dipelajari oleh mesin. Berikut tahapan *preprocessing* yang dilakukan pada penelitian ini :

a. *Cleaning Data*



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Cleaning Data merupakan tahapan untuk pembersihan data mentah. Data akan dibersihkan dari *term* yang tidak perlu dari data seperti karakter HTML, emoji, tanda baca, *hashtag* (#), *username* (@), serta URL.

b. *Case Folding*

Case Folding adalah proses menyamaratakan penggunaan huruf kapital pada data. Semua huruf kapital dari 'A' sampai 'Z' yang ada pada data akan diubah menjadi huruf kecil (*lowercase*). Karakter selain huruf juga akan dihilangkan dan dianggap *delimiter*. Tahap ini bertujuan agar menjaga konsistensi dan menghindari perbedaan makna atau interpretasi yang disebabkan oleh penggunaan huruf besar atau kecil.

c. *Punctual Removal*

Punctual Removal adalah tahapan menghapus tanda baca, simbol dan angka yang ada pada kalimat. Penghapusan karakter bertujuan untuk menghilangkan karakter yang tidak relevan dalam proses analisis teks atau pemodelan, sehingga mempermudah proses pelatihan model.

d. *Normalization*

Normalization pada tahap *preprocessing* merupakan proses untuk mengubah kata tidak baku menjadi kata baku yang sesuai dengan KBBI. Pada tahap ini, kata-kata singkatan, salah penulisan kata hingga perulangan huruf juga akan diubah menjadi kata yang baku. Data yang sudah di normalisasikan akan lebih mudah dipahami.

e. *Stopword Removal*

Stopword Removal merupakan tahapan untuk menghapus kata-kata umum yang biasa muncul dan dianggap tidak memiliki makna. Contoh *stopword* dalam bahasa Indonesia adalah "yang", "dan", "di", "dari", dll. Tujuannya adalah menghapus kata-kata yang memiliki informasi rendah dan berfokus pada kata-kata yang lebih bermakna dan informatif dalam sebuah teks.

f. *Stemming*



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	<p><i>Stemming</i> adalah tahapan menghilangkan imbuhan yang ada dalam kata sehingga menjadi kata dasar. Tahap ini berfungsi untuk mengurangi variasi kata yang memiliki makna yang sama.</p> <p>g. <i>Tokenization</i>.</p> <p><i>Tokenization</i> pada <i>preprocessing</i> merupakan tahapan memecah teks menjadi unit-unit yang lebih kecil yang disebut token.</p> <p>5. <i>Word Embedding</i></p> <p>Tahap <i>word embedding</i> adalah proses merepresentasikan token kedalam bentuk vektor. Dalam penelitian ini, proses <i>word embedding</i> akan dilakukan menggunakan <i>pre-trained word vectors fastText</i>.</p> <p>6. <i>Model Building</i></p> <p>Untuk membentuk model, pada penelitian ini akan digunakan algoritma <i>Gated Recurrent Unit</i> (GRU). Berikut beberapa layer yang digunakan dalam pembentukan model.</p> <p>a. <i>Embedding Layer</i></p> <p>b. <i>Gated Recurrent Unit Layer</i></p> <p>c. <i>Dense Layer</i></p> <p>7. <i>Learned Model</i></p> <p>Tahap ini akan dilakukan pengujian terhadap model yang telah dihasilkan pada saat proses <i>training data</i>.</p> <p>8. <i>Output</i></p> <p><i>Output</i> yang dihasilkan adalah berupa identifikasi komentar yang mengandung stereotipe terhadap Gen Z atau tidak.</p>
Referensi	<p>Sayama, M., & Sayama, H. 2011. <i>Positive stereotyping and multicultural awareness: An online experiment</i>. Current Research in Social Psychology, 16(6), 1-13.</p> <p>Meserole, Chris (2018, 9 May). <i>How misinformation spreads on social media—And what to do about it</i>. Brookings. Diakses pada 27 Oktober 2024, dari https://www.brookings.edu/articles/how-misinformation-spreads-on-social-media-and-what-to-do-about-it/</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Setiadji, A. R. A., Kusumaningtyas, S., & Juniarti, J. E. 2023. *Persepsi Milenial terhadap Stereotype Gen Z*. In Prosiding Seminar Nasional Ilmu Ilmu Sosial (SNIIS) (Vol. 2, pp. 103-113).

Bosco, C., Patti, V., Frenda, S., Cignarella, A. T., Paciello, M., & D'Errico, F. 2023. *Detecting racial stereotypes: An Italian social media corpus where psychology meets NLP*. Information Processing & Management, 60(1), 103118.

Fraser, K. C., Kiritchenko, S., & Nejadgholi, I. 2022. *Extracting Age-Related Stereotypes from Social Media Texts*. In Proceedings of the Thirteenth Language Resources and Evaluation Conference (pp. 3183-3194).

Jeldes, F., Ferreira, T., Diaz, D., & Ortiz, R. 2024. Exploring Gender Stereotypes in Financial Reporting: An Aspect-Level Sentiment Analysis Using Big Data and Deep Learning. Heliyon.

Cryan, J., Tang, S., Zhang, X., Metzger, M., Zheng, H., & Zhao, B. Y. 2020. *Detecting gender stereotypes: Lexicon vs. supervised learning methods*. In Proceedings of the 2020 CHI conference on human factors in computing systems (pp. 1-11).

Kapil, P., & Ekbal, A. 2020. *A deep neural network based multi-task learning approach to hate speech detection*. Knowledge-Based Systems, 210, 106458.

Medan, 28 Oktober 2024
Mahasiswa yang mengajukan,

Fatma Ananta Sari
NIM. 211402036