



Muhammadiyah University of Malang
Faculty of Engineering, Department of Informatics

Authors

Nama : Vita Amalia Herlinda
Nim : 202010370311336
Email : vitaamaliaherlinda@gmail.com
Fakultas : Teknik
Jurusan : Informatika

ANALISIS SENTIMEN MASYARAKAT TERHADAP GENERASI Z DALAM DUNIA KERJA PADA MEDIA SOSIAL TWITTER MENGGUNAKAN METODE NAÏVE BAYES

SKRIPSI



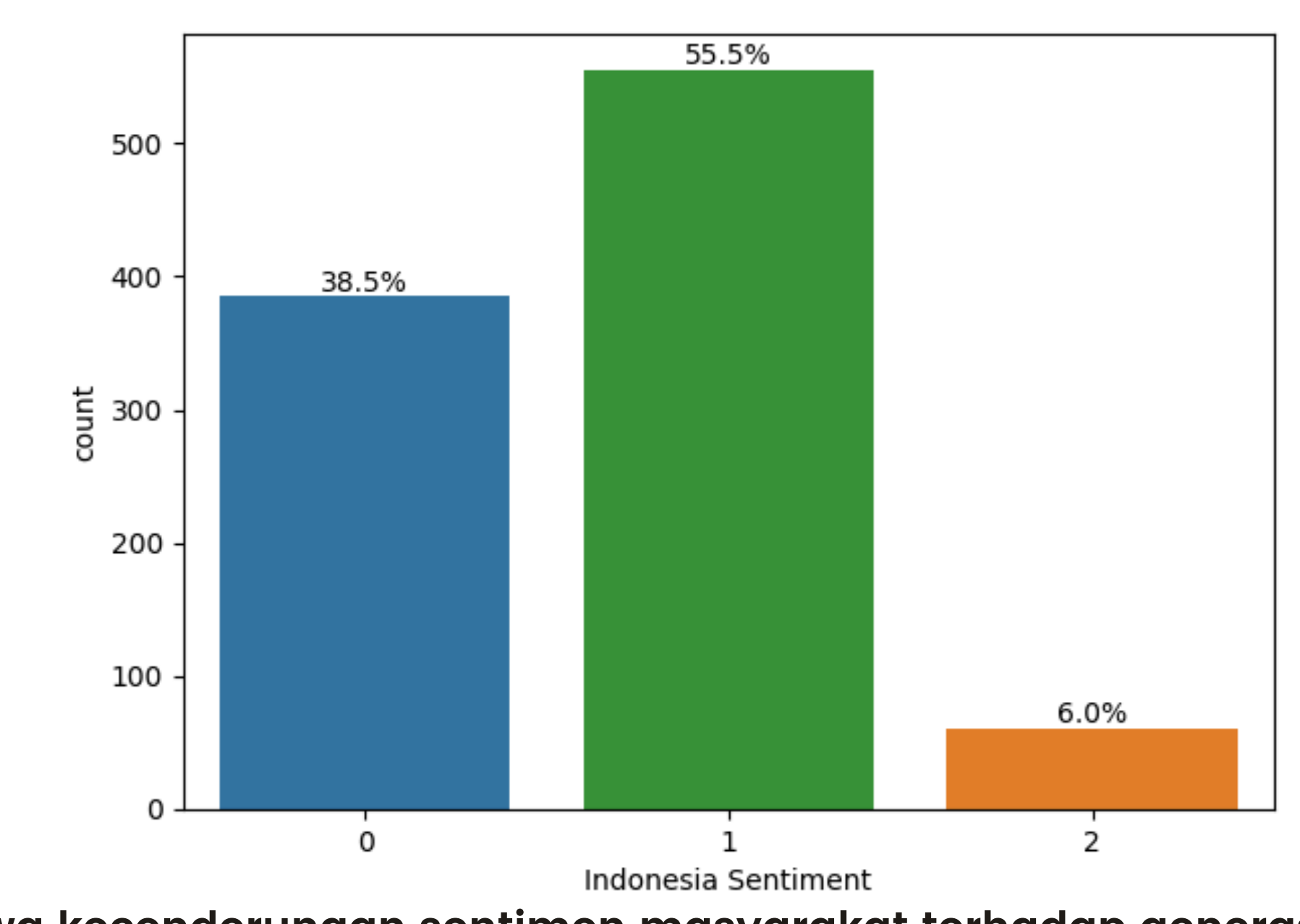
INTRODUCTION

Generasi Z, yang lahir antara tahun 1996 sampai 2010, kini mulai memasuki dunia kerja dan dikenal sebagai "Digital Natives" karena kedekatan mereka dengan teknologi dan media sosial seperti Twitter. Mereka aktif berbagi pemikiran dan pengalaman melalui Twitter, yang telah menjadi platform utama untuk diskusi terbuka dan berbagi perspektif. Penelitian ini menggunakan metode Naïve Bayes untuk menganalisis sentimen tweet terkait Generasi Z dalam dunia kerja, dengan memanfaatkan pelabelan lexicon Inset.

DATASET

pengumpulan dataset dilakukan dengan metode crawling menggunakan Tweet Harvest dan autentikasi API Cookie di Twitter dengan keyword utama "Kerja Gen z" berbahasa Indonesia pada rentang tahun 2020-2024.

HASIL DISTRIBUSI SENTIMEN



Hasil analisis menunjukkan bahwa kecenderungan sentimen masyarakat terhadap generasi Z dalam dunia kerja adalah negatif/1 sebesar 55.5%, diikuti dengan sentimen positif/ 0 sebesar 38.5%, dan sentimen netral/ 0 sebesar 6.0%. Hasil ini mengindikasikan bahwa sebagian besar teks yang diambil dari sumber data menunjukkan sentimen negatif.

WORDCLOUD SENTIMEN

POSITIF



Berdasarkan visualisasi wordcloud sentimen positif di atas dapat dilihat 3 kata yang disoroti: "mental", "Resign", dan "Toxic". Kehadiran kata-kata ini mengindikasikan adanya permasalahan atau tantangan yang dihadapi oleh Generasi Z dalam lingkungan kerja. Kata "mental" mungkin mencerminkan tekanan atau stres yang dialami dalam bekerja, dimana banyak kalangan menganggap bahwa Generasi Z memiliki ketahanan mental yang rendah atau kurang kuat dalam menghadapi tekanan di tempat kerja, sering dianggap sebagai "mental tempe" atau "strawberry". "Resign" menunjukkan kecenderungan untuk mengundurkan diri dari pekerjaan, seringkali karena lingkungan kerja tidak sesuai dengan harapan. Sedangkan "Toxic" merujuk pada lingkungan kerja yang tidak sehat, dimana banyak Generasi Z mengeluhkan perilaku toksik dari rekan kerja atau atasan di tempat kerja atau sebaliknya.

NEGATIF



Berdasarkan visualisasi WordCloud sentimen negatif di atas, terdapat 3 kata kunci yang menonjol: "anak", "kerja", dan "healing". Kata "anak" dalam konteks Generasi Z yang merasa akan selalu diberi dukungan finansial oleh orang tua mereka, mencerminkan pengakuan atas realitas bahwa mereka memiliki jaringan pengamanan keuangan dari orang tua, sehingga mereka merasa lebih aman. Kata "kerja" mengacu pada masuknya Generasi Z ke dalam dunia kerja, menunjukkan adaptasi mereka terhadap lingkungan kerja yang baru. Sedangkan kata "healing" menunjukkan bahwa Generasi Z sangat menyukai healing dan menjadi kebutuhan wajib

NETRAL



Berdasarkan visualisasi WordCloud sentimen netral di atas, terdapat 3 kata kunci yang menonjol: "anak", "kerja", dan "healing". Kata "anak" dalam konteks Generasi Z yang merasa akan selalu diberi dukungan finansial oleh orang tua mereka, mencerminkan pengakuan atas realitas bahwa mereka memiliki jaringan pengamanan keuangan dari orang tua, sehingga mereka merasa lebih aman. Kata "kerja" mengacu pada masuknya Generasi Z ke dalam dunia kerja, menunjukkan adaptasi mereka terhadap lingkungan kerja yang baru. Sedangkan kata "healing" menunjukkan bahwa Generasi Z sangat menyukai healing dan menjadi kebutuhan wajib

UJI COBA

1. PENGUJIAN PENGARUH FITUR STOPWORD

hasil ini menunjukkan penggunaan fitur stopwords dapat menghilangkan beberapa kata atau term bersentimen dari hasil fitur lexicon-based karena terdapat pada stopwords

Pengujian	Dengan	Tanpa
	Fitur Stopword	Fitur Stopword
Accuray	0.76	0.96
Precision	0.75	0.96
Recall	0.76	0.96
F1-Score	0.75	0.96

2. PENGUJIAN PELABELAN MANUAL

Hasil menunjukkan bahwa pelabelan manual lebih stabil dan cenderung memberikan akurasi yang lebih tinggi dibandingkan pelabelan lexicon saat menggunakan stopwords. Akan tetapi, pelabelan lexicon mampu mencapai akurasi yang sangat tinggi ketika stopwords dihilangkan, menandakan bahwa model berbasis lexicon sangat sensitif terhadap penggunaan stopwords dan dapat memanfaatkan variasi term yang lebih luas untuk meningkatkan akurasinya.

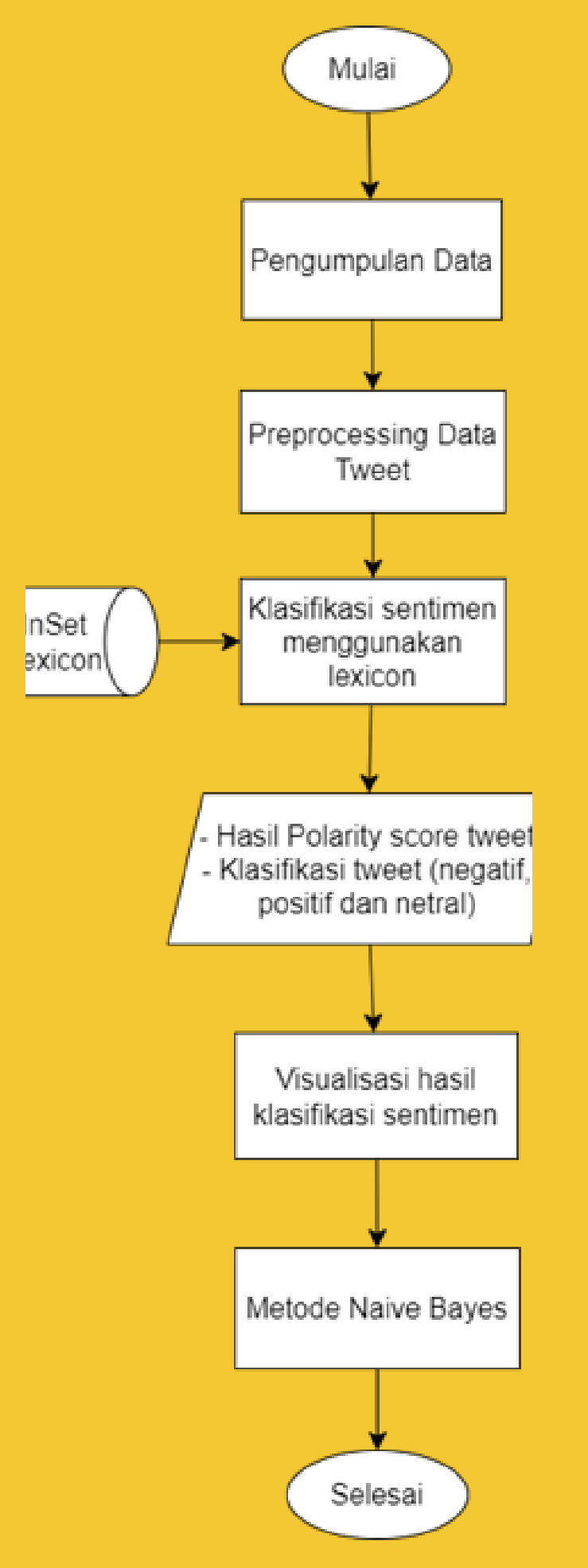
	Label Lexicon		Label Manual	
	Stopword	Tanpa stopwords	Stopword	Tanpa stopwords
Accuray	76%	96%	87%	90%
Precision	75%	96%	88%	91%
Recall	76%	96%	86%	90%
F1-Score	75%	96%	86%	89%

KESIMPULAN & SARAN

Pada eksperimen yang dilakukan oleh peneliti, ditemukan bahwa pelabelan lexicon, yang bergantung pada kamus inset lexicon bahasa Indonesia, memiliki akurasi yang rendah dibanding pelabelan manual. Hal ini disebabkan karena lexicon cenderung kurang sensitif terhadap nuansa dan konteks dalam teks yang kompleks, sehingga tidak mampu menangkap ekspresi sentimen dengan akurat seperti yang dapat dilakukan oleh pelabelan manual. Selain itu, fitur stopwords memainkan peran penting dalam peningkatan kinerja model. Penghapusan stopwords memungkinkan model untuk mempertimbangkan setiap kata dalam teks secara lebih komprehensif, termasuk kata-kata yang mungkin diabaikan dalam analisis sentimen. Repositor Vol. , No. , Bulan Tahun: halaman 10 Beberapa kata stopwords mungkin mengandung informasi yang bermanfaat dalam konteks tertentu, yang dapat membantu model dalam memahami dan mengklasifikasikan sentimen dengan lebih baik. Oleh karena itu, ketika fitur stopwords dihapus, model memiliki akses lebih baik ke informasi ini, yang mengarah pada peningkatan akurasi dalam klasifikasi sentimen

- Untuk meningkatkan analisis, peneliti mengusulkan untuk melakukan penelitian lanjutan dengan saran sebagai berikut .
1. Menambah jumlah dataset teks yang diambil dari media sosial Twitter terkait dengan pemikiran dan pengalaman kerja Generasi Z.
 2. Mengeksplorasi ekstraksi fitur selain dari metode TF-IDF.
 3. Menguji pendekatan lexicon-based yang berbeda selain menggunakan kamus kata.
 4. Melakukan eksperimen dengan metode klasifikasi lainnya selain Naive Bayes.

METODOLOGI



Pada penelitian ini menggunakan metode Naive Bayes yaitu ialah salah satu metode klasik terkenal yang telah banyak digunakan untuk kategorisasi teks yang memiliki struktur sederhana dan tingkat efektifitasnya yang tinggi sehingga digunakan dalam data mining