

0 JIMIK Juni 2024_5

by Jitu STMM MMTC

Submission date: 27-Jun-2024 09:09AM (UTC+0700)

Submission ID: 2409207090

File name: 1719454138_JIMIKJuni20245.docx (608.18K)

Word count: 2906

Character count: 19973

Pemodelan Topik dan Analisis Sentimen pada 'Voices of History: 50 Iconic Speeches' Menggunakan Pendekatan Natural Language Processing

....

.....
^{1,2}Alamat:

*Email korespondensi:

Diterima : || Revisi : ... || Disetujui:

Abstract

This study examines various iconic speeches throughout history using Natural Language Processing (NLP) approaches. The selected talks address a wide range of historical and social issues, from civil rights campaigns to pleas for peace. This analysis is performed to uncover the attitudes, themes, and emotions expressed in each speech, as well as to better understand how orators utilize language to influence and inspire their audience. This study identifies common patterns and themes in these talks using text vectorization methods, topic extraction via Latent Dirichlet Allocation (LDA), and sentiment and emotion analysis. The findings of this analysis suggest that the most influential speeches frequently include strong themes such as social justice, freedom, and unity. Furthermore, good thoughts and powerful emotions such as optimism and courage are frequently featured in the most motivating speeches. This research provides valuable insight for historians, researchers, and oratory enthusiasts into how iconic speeches have shaped public opinion and influenced social change.

Keywords: Natural Language Processing, Sentiment Analysis, Topic Extraction, Text Clustering.

Abstrak

Penelitian ini membahas analisis mendalam terhadap beberapa pidato ikonik dalam sejarah menggunakan teknik Natural Language Processing (NLP). Pidato-pidato yang dipilih mencakup berbagai konteks historis dan sosial, mulai dari perjuangan hak-hak sipil hingga seruan untuk perdamaian. Analisis ini dilakukan untuk mengungkapkan sentimen, tema, dan emosi yang terkandung dalam setiap pidato, serta untuk memahami bagaimana orator menggunakan bahasa untuk mempengaruhi (18) menginspirasi audiens mereka. Dengan memanfaatkan metode vektorisasi teks, ekstraksi topik menggunakan Latent Dirichlet Allocation (LDA), serta analisis sentimen dan emosi, penelitian ini mengidentifikasi pola dan tema yang umum dalam pidato-pidato tersebut. Hasil dari analisis ini menunjukkan bahwa pidato yang paling berpengaruh sering kali mengandung tema-tema yang kuat seperti keadilan sosial, kemerdekaan, dan persatuan. Selain itu, sentimen positif dan emosi kuat seperti harapan dan keberanian sering kali menonjol dalam pidato-pidato tersebut. Penelitian ini memberikan wawasan yang berharga bagi sejarawan, peneliti, dan penggemar oratori tentang bagaimana pidato ikonik telah membentuk opini publik dan mempengaruhi perubahan sosial.

Kata kunci: Pemrosesan Bahasa Alami, Analisis Sentimen, Ekstraksi Topik, Pengelompokan Teks

PENDAHULUAN

Pidato memiliki peran penting dalam sejarah umat manusia, digunakan oleh para pemimpin, aktivis, dan pemikir untuk

menginspirasi, memotivasi, (6) dan mengarahkan masyarakat. Pidato ikonik seperti "I Have a Dream" oleh Martin Luther King Jr. dan "We Shall Fight on the Beaches" oleh Winston

Churchill telah meninggalkan jejak yang mendalam dalam ingatan kolektif masyarakat, berkontribusi pada perubahan sosial dan politik yang signifikan. Meskipun kekuatan pidato-pidato ini sering kali diakui secara luas, memahami elemen-elemen bahasa yang membuat mereka begitu berdampak membutuhkan analisis yang lebih mendalam.

Perkembangan pengetahuan dan teknologi di bidang Natural Language Processing (NLP) menyediakan alat yang kuat untuk menganalisis teks secara komprehensif. Dengan NLP, kita dapat mengeksplorasi aspek-aspek seperti sentimen, tema, dan emosi yang terkandung dalam teks, memberikan wawasan yang lebih mendalam tentang bagaimana kata-kata dapat mempengaruhi audiens. Dalam konteks ini, analisis terhadap pidato-pidato ikonik menggunakan teknik NLP dapat mengungkap pola-pola linguistik yang mungkin tidak terlihat melalui analisis tradisional.

Penelitian ini bertujuan untuk mengaplikasikan teknik NLP pada pidato-pidato ikonik dalam sejarah, dengan tujuan memahami bagaimana elemen-elemen linguistik berkontribusi pada kekuatan dan dampak pidato tersebut. Dengan memanfaatkan alat-alat seperti vektorisasi teks, Latent Dirichlet Allocation (LDA) untuk ekstraksi topik, serta analisis sentimen dan emosi, penelitian ini berusaha mengidentifikasi tema-tema umum dan sentimen yang mendasari pidato-pidato yang ¹⁴ling berpengaruh dalam sejarah.

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi signifikan bagi bidang studi retorika, sejarah, dan komunikasi, serta memberikan wawasan praktis bagi para pemimpin dan orator masa kini tentang cara menggunakan bahasa untuk mempengaruhi dan menginspirasi audiens mereka. Selain itu, penelitian ini juga menunjukkan potensi penerapan teknik NLP dalam analisis teks historis dan kontemporer, membuka peluang baru untuk studi yang lebih mendalam tentang kekuatan bahasa dalam konteks sosial dan politik.

KAJIAN PUSTAKA

Penelitian ini berlandaskan pada beberapa teori dan metode dalam bidang Natural Language Processing (NLP), seperti analisis sentimen, dan studi retorika. NLP adalah cabang dari kecerdasan buatan yang berfokus pada interaksi antara

komputer dan bahasa manusia. NLP melibatkan teknik untuk memahami, menafsirkan, dan menghasilkan bahasa alami dengan cara yang bermanfaat (Jurafsky & Martin, 2020).

NLP dapat digunakan untuk mengolah dan menganalisis teks untuk ekstraksi informasi dan penemuan pola-pola linguistik. Beberapa teknik NLP yang sering dimanfaatkan adalah analisis sentimen dan analisis emosi. Analisis sentimen adalah proses mengidentifikasi dan mengkategorikan opini yang diekspresikan dalam suatu teks, terutama untuk menentukan apakah sikap penulis terhadap subjek tertentu bersifat positif, negatif, atau netral. Teknik ini memungkinkan peneliti untuk mengukur emosi dan perasaan yang terkandung dalam pidato, memberikan wawasan tentang bagaimana orator mempengaruhi audiens mereka secara emosional (Liu, B., 2012).

Analisis emosi adalah yang digunakan untuk mengidentifikasi dan mengukur emosi spesifik yang diekspresikan dalam teks. Pendekatan ini membantu dalam memahami bagaimana pidato memanfaatkan emosi untuk menciptakan hubungan dengan audiens dan mempengaruhi mereka. Ekstraksi emosi dalam teks dapat dilakukan menggunakan pendekatan berbasis leksikon atau model pembelajaran mesin (Buechel & Hahn, 2016).

Latent Dirichlet Allocation (LDA) adalah metode generatif dari NLP untuk ekstrak¹¹ topik dari kumpulan dokumen. LDA mengasumsikan bahwa setiap dokumen adalah campuran dari beberapa topik, dan setiap topik adalah distribusi dari kata-kata tertentu (Griffiths & Steyvers, 2004). Dalam penelitian ini, LDA digunakan untuk mengidentifikasi tema-tema utama yang terdapat dalam sejumlah pidato ikonik, sehingga memudahkan pemahaman tentang isu-isu yang ditekankan oleh para orator.

Selain konsep di atas, salah satu teori lain yang dimanfaatkan dalam penelitian ini adalah teori retorika yang mempelajari seni berbicara dan menulis dengan tujuan untuk

meyakinkan atau mempengaruhi audiens. Aristoteles mengidentifikasi tiga cara utama untuk meyakinkan audiens: ethos (karakter), pathos (emosi), dan logos (logika) (Aristotle, 2007). Dalam hal ini, teori retorika digunakan untuk memahami bagaimana struktur dan gaya bahasa dalam pidato berkontribusi terhadap efektivitasnya

Beberapa penelitian sebelumnya terkait pemanfaatan NLP sudah dilakukan, diantaranya adalah penggunaan teknik Natural Language Processing (NLP) untuk menganalisis pidato politik dari berbagai tokoh dunia. Studi ini meneliti gaya bahasa, tema utama, dan pola retorika yang digunakan dalam pidato untuk memahami bagaimana politisi mempengaruhi dan berkomunikasi dengan audiens mereka (Sim, dkk, 2013).

Penelitian selanjutnya adalah analisis sentimen terhadap pidato presiden untuk mengukur suasana hati dan perasaan yang disampaikan dalam pidato tersebut. Penelitian ini mengeksplorasi bagaimana sentimen dalam pidato berhubungan dengan konteks politik dan sosial pada saat pidato tersebut disampaikan (Muhammad, dkk., 2016).

Selanjutnya adalah penggunaan teknik analisis emosi untuk mengeksplorasi perasaan yang diekspresikan dalam karya sastra klasik. Studi ini membantu memahami bagaimana penulis sastra menggunakan emosi untuk menciptakan pengalaman membaca yang mendalam dan mengesankan (Kim, & Klinger, 2018).

Penelitian lainnya adalah eksplorasi penggunaan Latent Dirichlet Allocation (LDA) untuk ekstraksi topik dari koleksi dokumen. Studi ini menunjukkan bagaimana LDA dapat digunakan untuk mengidentifikasi tema utama dalam berbagai jenis teks, termasuk artikel ilmiah, berita, dan pidato public (Blei & Jordan, 2003).

Penelitian-penelitian tersebut telah menunjukkan bagaimana berbagai teknik NLP, analisis sentimen, dan analisis topik

telah digunakan dalam berbagai konteks untuk memahami dan mengeksplorasi teks-teks penting.

16 METODE

Metode yang digunakan dalam penelitian ini terdiri dari beberapa langkah penting yang mencakup persiapan data, pra-pemrosesan teks, analisis sentimen, analisis emosi, pembuatan model topik, dan clustering. Berikut adalah uraian dari setiap tahapannya:

1. Persiapan Data

Data yang digunakan adalah kumpulan teks dari pidato-pidato ikonik yang disimpan dalam file CSV. Data diperoleh dari platform open data: Kaggle. Proses awal meliputi pembacaan data dan penghapusan baris yang memiliki nilai kosong.

2. Pra-pemrosesan Teks

Pra-pemrosesan teks melibatkan tokenisasi, penghapusan stopwords, dan penghapusan tanda baca. Ini dilakukan untuk setiap teks dalam dataset.

```
211 stopwords =
set(stopwords.words('english'))
).union(additional_stopwords)
def preprocess_text(text):
    tokens =
word_tokenize(text.lower())
    tokens = [token for token
in tokens if token not in
all_stopwords]
    tokens = [token for token
in tokens if token not in
string.punctuation]
    return tokens
texts = [preprocess_text(text)
for text in df['Speech']]
```

3. Analisis Sentimen

Analisis sentimen dilakukan menggunakan TextBlob untuk mendapatkan polaritas dan subjektivitas dari setiap teks pidato.

```

8
def analyze_sentiment(text):
    analysis = TextBlob(text)
    return analysis.sentiment.
        polarity, analysis.
            sentiment. subjectivity
df['Sentiment_Polarity'],
df['Sentiment_Subjectivity'] =
zip(*df['Speech'].apply(lambda
x: analyze_sentiment('
'.join(preprocess_text(x))))))

```

4. Analisis Emosi

Analisis emosi dilakukan dengan menghitung frekuensi kata-kata yang terkait dengan emosi tertentu dalam teks pidato. Empat kategori emosi yang dianalisis adalah 'joy', 'sadness', 'anger', dan 'fear'.

```

def analyze_emotion(text):
    emotion_scores = {'joy':
0, 'sadness': 0, 'anger': 0,
'fear': 0}
    total_words = 0
    tokens =
word_tokenize(text.lower())
    total_words = len(tokens)
    for word in tokens:
        if word in emotions:
            emotion_scores[emo
tions[word]] += 1
        if total_words == 0:
            return {'joy': 0,
'sadness': 0, 'anger': 0,
'fear': 0}
        return {key:
value/total_words for key,
value in
emotion_scores.items()}

```

5. Pembuatan Model Topik

Model topik dibuat menggunakan Latent Dirichlet Allocation (LDA) untuk menemukan topik-topik yang mendasari kumpulan teks.

```

dictionary =
corpora.Dictionary(texts)

```

```

3
corpus =
[dictionary.doc2bow(text) for
text in texts]
lda_model =
models.LdaModel(corpus,
num_topics=5,
id2word=dictionary, passes=15)
for idx, topic in
lda_model.print_topics(-1):
    print(f"Topic: {idx}
\nWords: {topic}")
word_counts =
df['Speech'].apply(lambda x:
len(word_tokenize(x)))

```

7. Clustering

Clustering dilakukan menggunakan KMeans berdasarkan distribusi topik yang diperoleh dari model LDA.

```

# Vektorisasi teks menggunakan
7puntuVectorizer
vectorizer = CountVectorizer
(max_df=0.95,min_df=2,stop_wor
ds='english')
dtm = vectorizer.fit_transform
(df['Speech'])

9 LDA untuk ekstraksi topik
num_topics = 5
lda =LatentDirichletAllocation
(n_components=num_topics,
random_state=42)
lda.fit(dtm)

# Mengambil distribusi topik
untuk setiap dokumen
topic_distribution =
lda.transform(dtm)

# Normalisasi distribusi topik
topic_distribution =
normalize(topic_distribution,
norm='l1')

# Clustering17dokumen
menggunakan K-means
num_clusters = 5

```

```

km = KMeans(n_clusters=
num_clusters, random_state=42,
n_init=10)
clusters = km.fit_predict
(topic_distribution)
df['Cluster'] = clusters

# Reduksi dimensi menggunakan
PCA untuk visualisasi 2D
pca = PCA(n_components=2)
reduced_topic_distribution =
pca.fit_transform(topic_distri
bution)

```

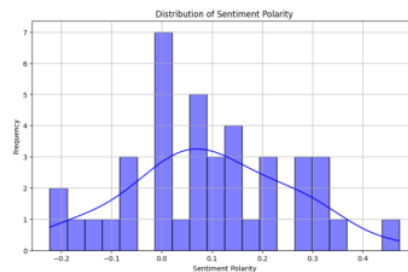
HASIL DAN PEMBAHASAN

Penelitian ini memberikan gambaran tentang isi dan tema pidato-pidato terkenal dalam "Voices of History: 50 Iconic Speeches". Dari distribusi sentimen dan emosi, kita bisa melihat bagaimana pidato-pidato tersebut mencerminkan perasaan dan opini dari pembicara. Model topik LDA dan clustering K-Means membantu mengidentifikasi tema utama dan kelompok dalam pidato-pidato ini, memberikan gambaran umum tentang isu-isu yang paling sering dibahas. Visualisasi seperti awan kata dan histogram frekuensi kata membantu memperkuat temuan ini dengan cara yang mudah dipahami. Secara keseluruhan, analisis ini menunjukkan bahwa pidato-pidato ini sangat berfokus pada isu-isu sosial, politik, dan kemanusiaan, dengan berbagai emosi dan topik yang terlibat. Gambaran detail tentang hasil analisa dapat dijelaskan sebagai berikut:

1. Analisis Sentimen

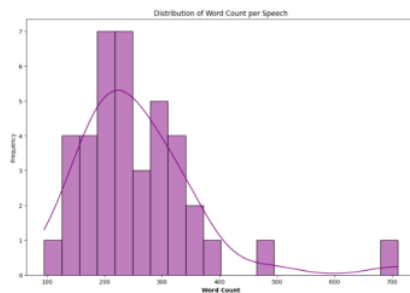
- **Distribusi Sentimen Polaritas:** Kami mengukur polaritas (positif atau negatif) dari setiap pidato menggunakan TextBlob. Nilai polaritas sentimen menunjukkan nuansa emosional dari masing-masing pidato. Polaritas positif (di atas 0) menunjukkan sentimen positif, sedangkan polaritas negatif (di bawah 0) menandakan sentimen negatif. Nilai polaritas yang mendekati 0 meng-

indikasikan sentimen netral. Contoh: Untuk "Saya punya mimpi dari MLK", polaritas sentimennya adalah 0,09776, yang menunjukkan sentimen yang sedikit positif. Hasil menunjukkan bahwa sebagian besar pidato memiliki polaritas yang cukup netral, dengan beberapa pidato menunjukkan polaritas yang sangat positif atau sangat negatif. Hal ini mengindikasikan bahwa pidato-pidato tersebut cenderung memiliki nada yang seimbang. Pidato dengan polaritas sangat positif menekankan harapan dan optimisme, sementara yang sangat negatif menyoroti tantangan dan kritik keras.



- **Distribusi Sentimen Subjektivitas:** Subjektivitas mengukur sejauh mana teks bersifat subjektif atau objektif. Sebagian besar pidato memiliki tingkat subjektivitas yang cukup tinggi, menunjukkan bahwa pidato ini cenderung mengandung opini dan perasaan pribadi daripada hanya fakta. Nilai subjektivitas sentimen berkisar dari 0 hingga 1, di mana 0 sangat objektif (faktual) dan 1 sangat subjektif (berpendapat atau emosional). Contoh: untuk pidato 'Tilbury Speech' oleh Ratu Elizabeth I, subjektivitas sentimennya adalah 0.677366, menunjukkan konten yang lebih subjektif atau emosional. Pidato dengan subjektivitas tinggi sering digunakan untuk menginspirasi atau mempengaruhi audiens dengan menyampaikan perasaan dan keyakinan pembicara.





6. Model Topik dengan LDA

Ekstraksi Topik: Kami menggunakan Latent Dirichlet Allocation (LDA) untuk mengidentifikasi lima topik utama dalam pidato-pidato ini. Topik-topik ini mencakup kata-kata seperti "freedom", "nation", "people", dan "world", mengindikasikan bahwa isu-isu ini sangat dominan dalam pidato-pidato tersebut. Topik-topik seperti kebebasan, perjuangan, harapan, dan persatuan muncul, mencerminkan keprihatinan dan aspirasi yang umum yang diangkat oleh para pembicara.

No.	Model Topik
0	Words: 0.012*"women" + 0.008*"world" + 0.006*"right" + 0.004*"peace" + 0.004*"country" + 0.004*"america" + 0.004*"hour" + 0.004*"men" + 0.004*"free" + 0.003*"last"
1	Words: 0.014*"party" + 0.013*"war" + 0.011*"hope" + 0.006*"russia" + 0.005*"risk" + 0.004*"struggle" + 0.004*"great" + 0.004*"army" + 0.004*"fighting"
2	Words: 0.007*"freedom" + 0.006*"great" + 0.006*"woman" + 0.006*"power" + 0.005*"women" + 0.004*"never" + 0.004*"world" + 0.004*"congress"
3	Words: 0.007*"world" + 0.006*"god" + 0.005*"women" + 0.005*"love" + 0.005*"oligarchy" + 0.005*"fear" + 0.004*"future"

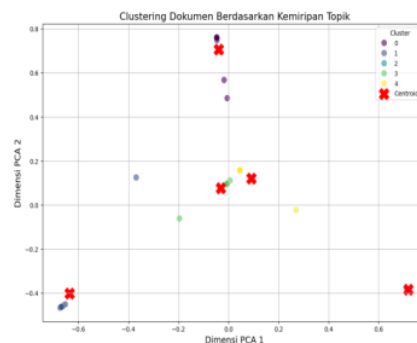
4

Words: 0.009*"black" + 0.007*"freedom" + 0.007*"rise" + 0.006*"country" + 0.006*"right" + 0.006*"men" + 0.006*"struggle" + 0.005*"white" + 0.005*"hope" + 0.005*"world"

Hasil pemodelan topik menunjukkan 5 topik teratas yang diidentifikasi dari pidato-pidato tersebut. Setiap topik diwakili oleh daftar kata kunci (istilah) yang paling khas dari topik tersebut, beserta bobotnya (probabilitas).

7. Clustering Teks Berdasarkan Topik

Dalam penelitian ini, digunakan K-Means clustering untuk mengelompokkan pidato berdasarkan distribusi topik yang dihasilkan oleh LDA. Hasil clustering menunjukkan adanya lima kelompok utama dalam pidato-pidato tersebut. Setiap cluster memiliki karakteristik topik yang berbeda, yang dapat dilihat dari centroid atau pusat massa dari setiap cluster. Misalnya, satu cluster didominasi oleh pidato yang berfokus pada kebebasan dan hak asasi manusia, sementara cluster lain lebih banyak tentang isu-isu nasional dan patriotisme.



SIMPULAN

Penelitian ini menghasilkan pemahaman yang mendalam tentang isi dan tema dari kumpulan pidato terkenal. Berikut adalah beberapa kesimpulan utama dari analisis yang telah dilakukan:

1. Distribusi Sentimen dan Emosi: Pidato-pidato ini menunjukkan variasi dalam polaritas sentimen dan emosi yang digunakan. Sebagian besar pidato memiliki polaritas netral atau sedikit positif, dengan emosi joy dan anger yang dominan. Ini mengindikasikan bahwa pidato-pidato ini mencoba untuk menyampaikan pesan dengan menggabungkan elemen-elemen positif dan kritik.
2. Topik Utama Pidato: Berdasarkan model LDA, kita dapat mengidentifikasi lima topik utama dalam pidato-pidato ini, seperti kebebasan, identitas nasional, hak asasi manusia, dan masalah global. Hal ini menunjukkan bahwa pidato-pidato ini sering mengangkat isu-isu yang penting secara sosial dan politik.
3. Awan Kata dan Frekuensi Kata: Kata-kata seperti "freedom", "nation", "people", dan "world" adalah kata-kata yang paling sering muncul dalam pidato-pidato ini. Ini mencerminkan fokus utama pidato-pidato tersebut pada isu-isu yang mencakup kebebasan, identitas nasional, dan kepedulian terhadap masyarakat global.
4. Clustering Dokumen: Melalui teknik clustering K-Means, pidato-pidato ini dapat dikelompokkan berdasarkan kesamaan topiknya. Ini membantu dalam memahami bagaimana tema-tema yang berbeda saling terkait dan bagaimana pidato-pidato ini bisa dikelompokkan berdasarkan konteks dan fokusnya.

SARAN

Berikut beberapa saran untuk penelitian selanjutnya berdasarkan temuan dari penelitian ini:

1. **Pengembangan Model Sentimen:** Mengembangkan model sentimen yang lebih kompleks untuk menangkap nuansa dan perubahan dalam sentimen

sepanjang pidato. Hal ini dapat memberikan pemahaman yang lebih dalam tentang bagaimana perasaan pembicara berkembang dari awal hingga akhir pidato.

2. **Analisis Komparatif:** Melakukan analisis komparatif antara pidato-pidato dari berbagai pembicara atau waktu yang berbeda untuk melihat perubahan dan konsistensi dalam tema dan pesan yang disampaikan.
3. **Analisis Sentimen terhadap Spesifik Topik:** Memfokuskan analisis sentimen pada topik-topik tertentu yang sering muncul dalam pidato-pidato tersebut, seperti kebebasan, hak asasi manusia, atau isu-isu global. Ini dapat memberikan wawasan yang lebih khusus tentang bagaimana sentimen berkembang dalam konteks topik yang spesifik.
4. **Penggunaan Teknik NLP Lanjutan:** Mengintegrasikan teknik-teknik NLP lanjutan seperti pengenalan entitas, pemodelan bahasa yang lebih canggih, atau analisis sintaksis untuk menggali lebih dalam struktur dan hubungan dalam teks pidato.
5. **Pengujian terhadap Data yang Lebih Luas:** Memperluas kumpulan data untuk mencakup pidato-pidato dari berbagai budaya atau konteks politik yang berbeda untuk melihat bagaimana perbedaan ini mempengaruhi tema dan sentimen dalam pidato.

DAFTAR PUSTAKA

- Jurafsky, D., & Martin, J. H. (2020). *Speech and Language Processing* (3rd ed.). Prentice Hall.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.

Griffiths, T. L., & Steyvers, M. (2004). Finding Scientific Topics. *Proceedings of the National Academy of Sciences*, 101(Suppl 1), 5228-5235.

Aristotle. (2007). *On Rhetoric: A Theory of Civic Discourse* (G. A. Kennedy, Trans.). Oxford University Press.

Buechel, S., & Hahn, U. (2016). Emotion Analysis as a Regression Problem – Dimensional Models and Their Implications on Emotion Representation and Metrical Evaluation. *EACL 2016*.

Sim, Y., Acree, B. L., Gross, J. H., & Smith, N. A. (2013). Measuring Ideological Proportions in Political Speeches. *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 91-101.

Muhammad, A., Wiratunga, N., Lothian, R., & Osborn, G. (2016). Sentiment Analysis of Informal Political Discourse. *Proceedings of the International Conference on Knowledge Management and Information Sharing*, 67-76.

Kim, E., & Klinger, R. (2018). A Survey on Sentiment and Emotion Analysis for Computational Literary Studies. *Zeitschrift für Digitale Geisteswissenschaften*,

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.

ORIGINALITY REPORT

12%
SIMILARITY INDEX

9%
INTERNET SOURCES

4%
PUBLICATIONS

9%
STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Universitas Budi Luhur Student Paper	4%
2	Submitted to University of North Texas Student Paper	1%
3	dspace.lib.ntua.gr Internet Source	1%
4	ojs.mmtc.ac.id Internet Source	1%
5	Submitted to Universitas Maritim Raja Ali Haji Student Paper	1%
6	www.unisbank.ac.id Internet Source	1%
7	Submitted to University of Kent at Canterbury Student Paper	1%
8	Submitted to University College London Student Paper	1%
9	Submitted to The Robert Gordon University Student Paper	<1%

10	repository.itny.ac.id Internet Source	<1 %
11	Submitted to Unika Soegijapranata Student Paper	<1 %
12	Submitted to University of Winchester Student Paper	<1 %
13	ejournal.iaida.ac.id Internet Source	<1 %
14	repository.mercubuana.ac.id Internet Source	<1 %
15	www.lap-publishing.com Internet Source	<1 %
16	www.researchgate.net Internet Source	<1 %
17	Usman Qamar, Muhammad Summair Raza. "Applied Text Mining", Springer Science and Business Media LLC, 2024 Publication	<1 %
18	Pangky Putra Aziztiya, Muhammad Habibi, Netania Indi Kusumaningtyas. "Analisis Sentimen Berdasarkan Topik Terkait Wabah Covid-19 di Twitter Menggunakan Latent Dirichlet Allocation (LDA) dan Naive Bayes Classifier (NBC)", Teknomatika: Jurnal Informatika dan Komputer, 2022 Publication	<1 %

Exclude quotes Off

Exclude matches Off

Exclude bibliography On