

**ANALISIS SENTIMEN DESTINASI SUPER PRIORITAS BERBASIS
STREAM PROCESSING DENGAN MENGGUNAKAN BIG DATA
PLATFORM**

SKRIPSI

Diajukan untuk memenuhi sebagian dari
Syarat Memperoleh Gelar Sarjana Komputer
Program Studi Ilmu Komputer



Oleh

Eva Nurkhofifah

1801196

**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PENDIDIKAN INDONESIA
BANDUNG
2024**

Analisis Sentimen Destinasi Super Prioritas Berbasis Stream Processing dengan Menggunakan Big Data Platform

Oleh
Eva Nurkhofifah

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar Sarjana
Komputer pada Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam

© Eva Nurkhofifah 2024
Universitas Pendidikan Indonesia
Agustus 2024

Hak Cipta dilindungi undang-undang.
Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian,
dengan dicetak ulang, difoto kopi, atau cara lainnya tanpa izin dari penulis.

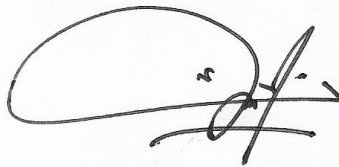
EVA NURKHOFIFAH

1801196

**ANALISIS SENTIMEN DESTINASI SUPER PRIORITAS BERBASIS
STREAM PROCESSING DENGAN MENGGUNAKAN BIG DATA
PLATFORM**

DISETUJUI DAN DISAHKAN OLEH PEMBIMBING:

Pembimbing I,



Prof. Dr. Lala Septem Riza, M.T.

NIP. 197809262008121001

Pembimbing II,



Dr. Fitri Rahmaftria, M.Si.

NIP. 197410182008122001

Mengetahui,
Kepala Program Studi Ilmu Komputer



Dr. Muhammad Nursalman, M.T.

NIP. 197909292006041002

ANALISIS SENTIMEN DESTINASI SUPER PRIORITAS BERBASIS STREAM PROCESSING DENGAN MENGGUNAKAN BIG DATA PLATFORM

Oleh

Eva Nurkhofifah
evanurkhofifah@upi.edu
1801196

ABSTRAK

Media sosial telah menjadi sumber data yang kaya akan opini dan sentimen publik terhadap berbagai topik, termasuk destinasi wisata. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap destinasi super prioritas di Indonesia (antara lain Danau Toba, Borobudur, Mandalika, Labuan Bajo, dan Likupang) melalui platform media sosial Twitter secara real-time menggunakan platform big data berbasis streaming. Penelitian ini, memanfaatkan platform big data yaitu Apache Kafka dan Apache Spark untuk mengumpulkan, memproses, dan menganalisis data dalam skala besar. Proses pengumpulan data dilakukan dengan menggunakan tweet-harvest berdasarkan kata kunci tertentu yang terkait dengan destinasi wisata super prioritas. Data tweet yang diperoleh kemudian diproses secara streaming menggunakan Apache Kafka. Kemudian data tersebut diproses menggunakan Apache Spark untuk melakukan pemrosesan analisis sentimen. Analisis sentimen menggunakan pendekatan berbasis leksikon yaitu Dictionary-Based Sentiment Analysis untuk mengkategorikan tweet menjadi sentimen positif, negatif, atau netral. Hasil analisis ini kemudian divisualisasikan dalam bentuk dashboard interaktif menggunakan Dash Plotly. Visualisasi ini memberikan wawasan real-time yang berharga bagi pemangku kepentingan dalam merespons opini publik dan merumuskan strategi pemasaran serta pengembangan destinasi super prioritas. Penelitian ini menunjukkan bahwa penggunaan Apache Kafka dan Apache Spark Streaming efektif dalam mengelola dan menganalisis data Twitter dalam skala besar secara real-time. Selain itu, visualisasi yang dihasilkan membantu dalam memahami persepsi publik terhadap destinasi wisata super prioritas. Temuan ini dapat digunakan sebagai acuan untuk penelitian lebih lanjut dalam analisis sentimen dan aplikasi big data di sektor pariwisata.

Kata Kunci: Analisis Sentimen, Big Data Platform, Data Streaming, Destinasi Super Prioritas, *Natural Language Processing*

SENTIMENT ANALYSIS OF SUPER PRIORITY DESTINATIONS BASED ON STREAM PROCESSING USING BIG DATA PLATFORM

Arranged by

Eva Nurkhofifah

evanurkhofifah@upi.edu

1801196

ABSTRACT

Social media has become a rich source of data for public opinion and sentiment on various topics, including tourist destinations. This research aims to analyze public sentiment towards super-priority destinations in Indonesia (such as Lake Toba, Borobudur, Mandalika, Labuan Bajo, and Likupang) through the Twitter social media platform in real-time using a streaming-based big data platform. This research utilizes Apache Kafka and Apache Spark big data platforms to collect, process, and analyze data on a large scale. Data collection is conducted using tweet-harvesting based on specific keywords related to these super-priority destinations. The collected tweets are streamed and processed using Apache Kafka and further analyzed using Apache Spark for sentiment analysis. The sentiment analysis employs a lexicon-based approach known as Dictionary-Based Sentiment Analysis to classify tweets into positive, negative, or neutral sentiments. The results of this analysis are then visualized in an interactive dashboard using Dash Plotly. These visualizations provide valuable real-time insights for stakeholders to respond to public opinions and formulate marketing strategies and development plans for super-priority destinations. The research demonstrates the effectiveness of Apache Kafka and Apache Spark Streaming in managing and analyzing Twitter data on a large scale in real-time. Furthermore, the visualizations aid in understanding public perceptions of super-priority tourist destinations. These findings can serve as a reference for further research in sentiment analysis and big data applications in the tourism sector.

Keywords: Big Data Platform, Data Streaming, Sentiment Analysis, Natural Language Processing, Super Priority Destinations

DAFTAR ISI

ABSTRAK.....	i
ABSTRACT.....	ii
KATA PENGANTAR	iii
UCAPAN TERIMA KASIH.....	iv
DAFTAR ISI.....	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN.....	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	5
1.3. Tujuan Penelitian.....	5
1.4. Manfaat Penelitian.....	5
1.5. Batasan Masalah.....	6
1.6. Sistematika Penulisan.....	6
BAB II KAJIAN PUSTAKA.....	8
2.1. Peta Literatur	8
2.2. Pariwisata	8
2.2.1. Pengenalan Pariwisata	9
2.2.2. Unsur dan Dimensi Pariwisata.....	12
2.2.3. Bentuk-Bentuk Pariwisata.....	14
2.2.4. Destinasi Super Prioritas	16
2.3. Twitter	20
2.3.1. Pengenalan Twitter	20
2.3.2. Penggunaan dan Fitur Twitter.....	21
2.3.3. Twitter API	25
2.4. Natural Language Processing.....	30
2.4.1. Pengenalan Natural Language Processing	30
2.4.2. Tahapan Natural Language Processing.....	32
2.4.3. Penerapan Natural Language Processing.....	35
2.5. Analisis Sentimen.....	38
2.5.1. Pengenalan Analisis Sentimen.....	38
2.5.2. Teknik Analisis Sentimen	40

2.5.3.	Tipe-Tipe Analisis Sentimen	44
2.6.	Big Data.....	47
2.6.1.	Pengenalan Big Data.....	47
2.6.2.	Karakteristik Big Data	48
2.6.3.	Tipe Data Big Data	50
2.6.4.	Big Data Platform	51
2.7.	Data Stream	55
2.7.1.	Pengenalan Data Stream	55
2.7.2.	Sumber Data Stream	58
2.7.3.	Penerapan Data Stream	59
2.8.	Apache Kafka.....	61
2.8.1.	Pengenalan Sistem Pengiriman Pesan	61
2.8.2.	Pengenalan Apache Kafka	63
2.8.3.	Cara Kerja Kafka	66
2.9.	Apache Spark	70
2.9.1.	Pengenalan Apache Spark.....	71
2.9.2.	Komponen Apache Spark	72
2.9.3.	Cara Kerja Apache Spark.....	74
2.10.	Penelitian Terkait.....	79
BAB III METODE PENELITIAN		84
3.1.	Desain Penelitian.....	84
3.2.	Metode Penelitian.....	88
3.2.1.	Metode Pengumpulan Data.....	88
3.2.2.	Metode Pengembangan Perangkat Lunak.....	89
3.3.	Alat dan Bahan Penelitian	91
3.3.1.	Alat Penelitian.....	91
3.3.2.	Bahan Penelitian	91
BAB IV HASIL DAN PEMBAHASAN		92
4.1.	Pengumpulan Data	92
4.2.	Perancangan Model Komputasi Analisis Sentimen Destinasi Super Prioritas	93
4.2.1.	Data Collection	95
4.2.2.	Stream Processing dengan Kafka.....	103
4.2.3.	Analisis Sentimen dengan Spark.....	109
4.2.4.	Visualisasi dengan Dash dari Plotly.....	127
4.3.	Pembangunan Aplikasi Analisis Sentimen Destinasi Super Prioritas..	129

4.3.1.	Analisis Aplikasi Analisis Sentimen Destinasi Super Prioritas	129
4.3.2.	Desain Aplikasi Analisis Sentimen Destinasi Super Prioritas	130
4.3.3.	Implementasi Aplikasi Analisis Sentimen Destinasi Super Prioritas 131	
4.3.4.	Pengujian Aplikasi Analisis Sentimen Destinasi Super Prioritas .	162
4.4.	Rancangan Skenario Eksperimen	163
4.5.	Hasil dan Analisis Eksperimen	165
4.5.1.	Hasil Eksperimen	165
4.5.2.	Analisis Eksperimen	175
BAB V KESIMPULAN DAN SARAN		211
5.1.	Kesimpulan.....	211
5.2.	Saran.....	212
DAFTAR PUSTAKA		214
LAMPIRAN.....		221

DAFTAR TABEL

Tabel 2. 1. Istilah dalam Pariwisata	10
Tabel 2. 2. Terminologi Twitter	22
Tabel 2. 3. Penelitian Terkait	80
Tabel 4. 1. Komponen Fitur Tweet	92
Tabel 4. 2. Hasil Pengumpulan Data Twitter	96
Tabel 4. 3. Data hasil praproses data yang akan menjadi input proses streaming	102
Tabel 4. 4. Pengujian Proses dengan Metode Black Box	162
Tabel 4. 5. Pengujian Proses pada Dashboard dengan Metode Black Box	163
Tabel 4. 6. Hasil Eksperimen	165
Tabel 4. 7. Hasil Penilaian Expert Judgment	175
Tabel 4. 8. hasil korelasi sentimen dengan destinasi dan topik 5A	201
Tabel 4. 9. Pengukuran Waktu Respons dan Latensi.....	202
Tabel 4. 10. Pengukuran waktu respon dan latensi dengan perhitungan sentimen	203
Tabel 4. 11. Perbandingan dengan Penelitian Sebelumnya	207

DAFTAR GAMBAR

Gambar 1. 1. Kontribusi PDB Pariwisata Indonesia (DataIndonesia.Id)	1
Gambar 2. 1. Peta Literatur Penelitian	8
Gambar 2. 2. Bentuk dasar pariwisata (Theobald, 2004)	11
Gambar 2. 3. Tipologi Wisatawan (Eddyono, 2021).....	12
Gambar 2. 4. Destinasi Super Prioritas Danau Toba.....	17
Gambar 2. 5. Destinasi Super Prioritas Borobudur	17
Gambar 2. 6. Destinasi Super Prioritas Mandalika	18
Gambar 2. 7. Destinasi Super Prioritas Labuan Bajo	18
Gambar 2. 8. Destinasi Super Prioritas Likupang	19
Gambar 2. 9. Contoh tweet.....	24
Gambar 2. 10. Contoh style tweet substantif	24
Gambar 2. 11. Contoh style konversasi.....	25
Gambar 2. 12. Workflow OAuth (Kumar et al., 2013)	29
Gambar 2. 13. Komponen Utama NLP (Martinez, 2010)	31
Gambar 2. 14. Teknik Analisis Sentimen (Aftab et al., 2023)	40
Gambar 2. 15. Karakteristik Big Data	49
Gambar 2. 16. Arsitektur HDFS (Aytas, 2021).....	52
Gambar 2. 17. Arsitektur atau eksekusi MapReduce (Aytas, 2021)	53
Gambar 2. 18. Arsitektur YARN (Aytas, 2021).....	53
Gambar 2. 19. Arsitektur Apache Flume (Aytas, 2021).....	55
Gambar 2. 20. Sistem Data Streaming (Psaltis, 2017)	56
Gambar 2. 21. Time-encoded data video (Makota et al., 2021)	59
Gambar 2. 22. Contoh penerapan data stream (Makota et al., 2021)	59
Gambar 2. 23. Komponen Sistem Pengiriman Pesan (Makota et al., 2021)	61
Gambar 2. 24. Model Pengiriman Pesan Publish/Subscribe (M. Kumar & Singh, 2017)	63
Gambar 2. 25. Arsitektur Apache Kafka (Makota et al., 2021)	65
Gambar 2. 26. Partisi dalam Kafka (Shapira et al., 2021).....	67
Gambar 2. 27. Kafka Producer Flow (M. Kumar & Singh, 2017)	68
Gambar 2. 28. Kafka Consumer Flow (M. Kumar & Singh, 2017)	69
Gambar 2. 29. Komponen Apache Spark (Karau et al., 2017).....	72
Gambar 2. 30. Spark RDD flow (Aytas, 2021)	75
Gambar 2. 31. Arsitektur Spark (Aytas, 2021).....	77
Gambar 2. 32. Spark Stage (Aytas, 2021).....	78
Gambar 3. 1. Desain Penelitian	85
Gambar 3. 2. Model Agile (Sharma et al., 2012)	89
Gambar 4. 1. Komponen Fitur tweet pada Twitter.....	92
Gambar 4. 2. Arsitektur Analisis Sentimen Destinasi Super Prioritas Berbasis Stream Processing dengan Big Data Platform.....	93
Gambar 4. 3. Diagram Model Komputasi Analisis Sentimen Berbasis Streaming	94
Gambar 4. 4. Pseudocode Data Collection.....	95
Gambar 4. 5. Inspect laman pencarian Twitter.....	97

Gambar 4. 6. Pseudocode crawling data Twitter.....	98
Gambar 4. 7. Pseudocode preprocessing data	98
Gambar 4. 8. Menangani missing value	99
Gambar 4. 9. Mengubah format data tanggal	100
Gambar 4. 10. Sorting data sesuai tanggal	101
Gambar 4. 11. Hapus dan reindex kolom	102
Gambar 4. 12. Perintah untuk menjalankan zookeeper	104
Gambar 4. 13. Perintah untuk menjalankan server kafka.....	104
Gambar 4. 14. Perintah untuk membuat topik kafka.....	104
Gambar 4. 15. Pseudocode Modul Konfigurasi	105
Gambar 4. 16. Pseudocode pada Modul Producer dan Consumer Kafka.....	105
Gambar 4. 17. Pseudocode fungsi producer.....	106
Gambar 4. 18. Pseudocode fungsi consumer.....	107
Gambar 4. 19. Pseudocode modul producer.....	107
Gambar 4. 20. Pseudocode tahap chek producer pada modul producer.....	108
Gambar 4. 21. Perintah menjalankan konsumen kafka.....	108
Gambar 4. 22. Hasil data yang dikirimkan producer kafka dari file csv	109
Gambar 4. 23. Pseudocode modul `consumer`	109
Gambar 4. 24. Pseudocode membuat sesi Spark.....	110
Gambar 4. 25. Pseudocode membuat skemai Spark.....	111
Gambar 4. 26. Pseudocode membaca data dari Kafka	111
Gambar 4. 27. Pseudocode untuk menjalankan modul Preprocessing	112
Gambar 4. 28. Pseudocode modul preprocessing.....	112
Gambar 4. 29. Pseudocode fungsi `clean_tweet`	113
Gambar 4. 30. Regex menghapus url dengan awalan http	113
Gambar 4. 31. Regex menghapus url dengan awalan bit.ly	114
Gambar 4. 32. Regex untuk menghapus username pada teks tweet.....	115
Gambar 4. 33. Regex untuk menghapus hashtag pada teks tweet.....	116
Gambar 4. 34. Regex menghapus tanda baca.....	117
Gambar 4. 35. Regex menghapus karakter angka pada teks tweet.....	118
Gambar 4. 36. Regex untuk menghapus newline dan spasi yang berlebihan.....	119
Gambar 4. 37. Mengonversi semua huruf menjadi huruf kecil	120
Gambar 4. 38. Pseudocode fungsi `remove_stepwords`	120
Gambar 4. 39. Contoh hasil menghapus stopwords	121
Gambar 4. 40. Pseudocode untuk menjalankan modul `sentiment_analysis`.....	122
Gambar 4. 41. Pseudocode modul sentiment analysis.....	122
Gambar 4. 42. Pseudocode modul sentiment dictionaries.....	123
Gambar 4. 43. Hasil analisis sentimen dictionary-based.....	123
Gambar 4. 44. Pseudocode untuk menjalankan modul `categories`.....	124
Gambar 4. 45. Pseudocode modul pengkategorian destinasi dan topik	125
Gambar 4. 46. Pseudocode fungsi `categorize_destination`	126
Gambar 4. 47. Pseudocode fungsi `categorize_topics`	126
Gambar 4. 48. Pseudocode modul `topic_dictionaries`	127
Gambar 4. 49. Contoh data hasil proses analisis dengan Spark	127

Gambar 4. 50. Contoh pie chart.....	128
Gambar 4. 51. Contoh bar chart	128
Gambar 4. 52. Contoh line chart	129
Gambar 4. 53. Desain Skema Data.....	131
Gambar 4. 54. Modul `main.py` untuk menjalankan aplikasi	156
Gambar 4. 55. Modul `home.py` menampilkan halaman utama dashboard.....	157
Gambar 4. 56. Modul `dataset.py` untuk menampilkan halaman data.....	158
Gambar 4. 57. Halaman utama dashboard.....	161
Gambar 4. 58. Contoh halaman dashboard	162
Gambar 4. 59. Persentase dan jumlah sentimen pengguna Twitter terkait DSP	166
Gambar 4. 60. Frekuensi diskusi Destinasi Super Prioritas dan Topik 5A	167
Gambar 4. 61. Distribusi sentimen pengguna Twitter tiap Destinasi Super Prioritas	168
Gambar 4. 62. Kategori Sentimen tiap Destinasi Super Prioritas dan Topik 5A	169
Gambar 4. 63. Heatmap Sentimen tiap Destinasi Super Prioritas dan Topik 5A	169
Gambar 4. 64. Frekuensi kata tiap Destinasi Super Prioritas	171
Gambar 4. 65. Frekuensi kata tiap ketegori sentimen	172
Gambar 4. 66. Tren Sentimen Destinasi Super Prioritas	173
Gambar 4. 67. Real Time Tren Sentimen Destinasi Super Prioritas	174
Gambar 4. 68. Frekuensi diskusi DSP tiap kuartal.....	177
Gambar 4. 69. Frekuensi diskusi aspek topik berdasarkan 5A.....	178
Gambar 4. 70. Distribusi sentimen tiap destinasi periode akhir tahun	180
Gambar 4. 71. Distribusi sentimen topik atraksi tiap destinasi	181
Gambar 4. 72. Frekuensi kata tiap kategori sentimen destinasi Borobudur	183
Gambar 4. 73. Frekuensi kata tiap kategori sentimen destinasi Danau Toba.....	184
Gambar 4. 74. Frekuensi kata tiap kategori sentimen destinasi Labuan Bajo.....	185
Gambar 4. 75. Frekuensi kata tiap kategori sentimen destinasi Likupang	187
Gambar 4. 76. Frekuensi kata tiap kategori sentimen destinasi Mandalika	188
Gambar 4. 77. Tren sentimen destinasi Borobudur tiap Kuartal	190
Gambar 4. 78. Tren sentimen destinasi Danau Toba tiap Kuartal.....	192
Gambar 4. 79. Tren sentimen destinasi Labuan Bajo tiap Kuartal.....	194
Gambar 4. 80. Tren sentimen destinasi Likupang tiap Kuartal	196
Gambar 4. 81. Tren sentimen destinasi Mandalika tiap Kuartal	198
Gambar 4. 82. Real Time Tren Sentimen.....	199
Gambar 4. 83. Pendekatan time-based trigger dengan interвал waktu 1 detik	204
Gambar 4. 84. Pendekatan time-based trigger dengan interвал waktu 10 detik	205
Gambar 4. 85. Pendekatan time-based trigger dengan interвал waktu 60 detik	206

DAFTAR LAMPIRAN

Lampiran 1. Kode Program Modul `consumer.py`	221
Lampiran 2. Daftar Kamus Kata Sentimen	223
Lampiran 3. Daftar Kamus Kata Kategori Topik 5A	229
Lampiran 4. Kode Program Modul `overview.py`	231
Lampiran 5. Kode Program Modul `relationship.py`	233
Lampiran 6. Kode Program Modul `sentiment_analysis.py`	235
Lampiran 7. Kode Program Modul `word_trends.py`	238
Lampiran 8. Kode Program Modul `realtime_trends.py`	239
Lampiran 9. Hasil Penilaian Expert Judgment	242

DAFTAR PUSTAKA

- Abdulaziz, M., Alotaibi, A., Alsolamy, M., & Alabbas, A. (2021). Topic based Sentiment Analysis for COVID-19 Tweets. *International Journal of Advanced Computer Science and Applications*, 12(1), 626–636. <https://doi.org/10.14569/IJACSA.2021.0120172>
- Abdullah, N. A. S., & Rusli, N. I. A. (2021). Multilingual sentiment analysis: A systematic literature review. *Pertanika Journal of Science and Technology*, 29(1), 445–470. <https://doi.org/10.47836/pjst.29.1.25>
- Aftab, F., Bazai, S. U., Marjan, S., Baloch, L., Aslam, S., Amphawan, A., & Neo, T. K. (2023). A Comprehensive Survey on Sentiment Analysis Techniques. *International Journal of Technology*, 14(6), 1288–1298. <https://doi.org/10.14716/ijtech.v14i6.6632>
- Alqaryouti, O., Siyam, N., Abdel Monem, A., & Shaalan, K. (2024). Aspect-Based Sentiment Analysis Using Smart Government Review Data. *Applied Computing and Informatics*, 20(1–2), 142–161. <https://doi.org/10.1016/j.aci.2019.11.003>
- Ankam, Venkat. (2016). Big data analytics: a handy reference guide for data analysts and data scientists to help obtain value from big data analytics using Spark on Hadoop clusters. In *Springer*.
- Anzela, A. I., & Rachmawati, I. (2023). Diplomasi Komersial 5 Destinasi Super Prioritas: Promosi dan Kerja Sama. *Jurnal Pena Wimaya*, 3(2). <https://bakri.uma.ac.id/kualitas-merupakan-menurut-pakar/>
- Anupama, Rakshith, Kumar, R., & Navaneeth. (2020). Real Time Twitter Sentiment Analysis using Natural Language Processing. *International Journal of Engineering Research & Technology (IJERT)*, 9(07), 1107–1112. www.ijert.org
- Appiahene, P., Afrifa, S., Kyei, E. A., & Nimbe, P. (2022). Understanding the Uses, Approaches and Applications of Sentiment Analysis. *Research Square, ML*, 1–20.
- Aytas, Y. (2021). *Designing Big Data Platforms*. Wiley.
- Bagheri, H., & Jahirul, M. (2018). Sentiment analysis of twitter data. *Proceedings - 2018 International Conference on Computational Science and Computational Intelligence, CSCI 2018*, 1301–1302. <https://doi.org/10.1109/CSCI46756.2018.00252>
- Bhadane, C., Dalal, H., & Doshi, H. (2015). Sentiment analysis: Measuring opinions. *Procedia Computer Science*, 45(C), 808–814. <https://doi.org/10.1016/j.procs.2015.03.159>
- Cahyono, Y. (2017). Analisis Sentiment pada Sosial Media Twitter Menggunakan Naïve Bayes Classifier dengan Feature Selection Particle Swarm Optimization dan Term Frequency. In *Jurnal Informatika Universitas Pamulang* (Vol. 2, Issue 1, p. 14). <https://doi.org/10.32493/informatika.v2i1.1500>

- Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2017). *A practical guide to sentiment analysis*.
- Cha, S., & Wachowicz, M. (2015). Developing a Real-Time Data Analytics Framework using Hadoop. *2015 IEEE International Congress on Big Data*, 0–3. <https://doi.org/10.1109/BigDataCongress.2015.102>
- Chandarana, P., & Vijayalakshmi, M. (2014). Big data analytics frameworks. *2014 International Conference on Circuits, Systems, Communication and Information Technology Applications, CSCITA 2014*, 430–434. <https://doi.org/10.1109/CSCITA.2014.6839299>
- Chung, S. S., & Aring, D. (2018). Integrated Real-Time Big Data Stream Sentiment Analysis Service. *Journal of Data Analysis and Information Processing*, 06(02), 46–66. <https://doi.org/10.4236/jdaip.2018.62004>
- Cielen, D., Meysman, A. D. B., & Ali, M. (2016). *Introducing Data Science. 1st edn*.
- Crawford, A., Richter, T., Hohner, J., Greipl, S., & Schulze, H. (2023). *International Events and Social Media Activities on X (formerly Twitter): Towards the Establishment of a Monitoring Dashboard*. 398–409.
- Damji, J. S., Wenig, B., Das, T., & Lee, D. (2020). Learning Spark: Lightning-Fast Data Analytics. In *Learning Spark: Lightning-Fast Data Analytics*.
- Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A. Y. A., Gelbukh, A., & Zhou, Q. (2016). Multilingual Sentiment Analysis: State of the Art and Independent Comparison of Techniques. *Cognitive Computation*, 8(4), 757–771. <https://doi.org/10.1007/s12559-016-9415-7>
- Eddyono, F. (2021). Pengelolaan Destinasi Wisata. In *Pengelolaan Destinasi Pariwisata*. Uwais Inspirasi Indonesia.
- Eisenstein, J. (2018). *Introduction to Natural Language Processing*. <http://www.nber.org/papers/w16019>
- Ellis, B. (2014). *Real-Time Analytics Techniques to Analyze and Visualize Streaming Data*.
- Elragal, A. (2014). ERP and Big Data: The Inept Couple. *Procedia Technology*, 16(February), 242–249. <https://doi.org/10.1016/j.protcy.2014.10.089>
- Elzayady, H., Badran, K. M., & Salama, G. I. (2019a). Sentiment Analysis on Twitter Data using Apache Spark Framework. *Proceedings - 2018 13th International Conference on Computer Engineering and Systems, ICCES 2018, December*, 171–176. <https://doi.org/10.1109/ICCES.2018.8639195>
- Elzayady, H., Badran, K. M., & Salama, G. I. (2019b). Sentiment Analysis on Twitter Data using Apache Spark Framework. *Proceedings - 2018 13th International Conference on Computer Engineering and Systems, ICCES 2018, December*, 171–176. <https://doi.org/10.1109/ICCES.2018.8639195>
- Fanissa, S., Fauzi, M. A., & Adinugroho, S. (2018a). *Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Analisis Sentimen Pariwisata di*

- Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking. August.*
- Fanissa, S., Fauzi, M. A., & Adinugroho, S. (2018b). *Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking. August.*
- Hidayat, T. F. T., Garno, G., & Ridha, A. A. (2021). Analisis Sentimen Opini Pemindahan Ibu Kota Pada Twitter Dengan Metode Support Vector Machine. *Jurnal Ilmu Komputer*, 14(1), 49. <https://doi.org/10.24843/jik.2021.v14.i01.p06>
- Hoang, M., Alija Bihorac, O., & Rouces, J. (2019). Aspect-Based Sentiment Analysis Using BERT. *Proceedings of the 22nd Nordic Conference on Computational Linguistics*, 187–196.
- Isdarmanto. (2017). Dasar-Dasar Kepariwisata dan Pengelolaan ODTW Pariwisata. In *Yogyakarta: Gerbang Media Aksara.*
- Kale, V. (2020). *Parallel Computing Architectures and APIs: IoT Big Data Stream Processing.* <http://www.nber.org/papers/w16019>
- Kamal, R., & Saxena, P. (2019). *Big Data Analytics: Introduction to Hadoop, Spark, and Machine Learning.*
- Karau, H., Konwinski, A., Wendell, P., & Zaharia, M. (2017). Learning Spark: Lightning-Fast Data Analysis. In *Journal of Crystal Growth* (Vol. 466). <https://doi.org/10.1016/j.jcrysgro.2017.03.020>
- Karau, H., & Warren, R. (2017). High Performance Spark: Best Practices for Scaling & Optimizing Apache Spark. In *Nihon Reoroji Gakkaishi (Journal of the Society of Rheology, Japan)*. https://doi.org/10.1678/rheology1973.15.2_62
- Kim, S. (2023). Analysis of Sentiment Analysis Research Trends using Text Mining Analysis. *TechRxiv.*
- Kolchyna, O., Souza, T. T. P., Treleaven, P., & Aste, T. (2015). *Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination. July.* <http://arxiv.org/abs/1507.00955>
- Kumar, M., & Singh, C. (2017). *Building Data Streaming Applications with Apache Kafka.* Packt Publishing.
- Kumar, S., Morstatter, F., & Liu, H. (2013). *Twitter Data Analytics.* Springer, 89.
- Kusal, S., Patil, S., Kotecha, K., Aluvalu, R., & Varadarajan, V. (2021). AI Based Emotion Detection for Textual Big Data: Techniques and Contribution. *Big Data and Cognitive Computing*, 5(3), 5, 43. <https://doi.org/10.3390/bdcc5030043>
- Kwak, H., Lee, C., Hosung, P., & Sue, M. (2011). What is Twitter, a Social Network or a News Media? *Archivos de Zootecnia*, 60(230), 297–300. <https://doi.org/10.4321/S0004-05922011000200015>

- Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Opinion Analysis for Online Reviews, May*, 79–94. https://doi.org/10.1142/9789813100459_0007
- Liu, B. (2015). *Sentiment Analysis Mining Opinions, Sentiments, and Emotions*. Cambridge University Press.
- Lye, S. H., & Teh, P. L. (2021). Customer Intent Prediction using Sentiment Analysis Techniques. *Proceedings of the 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2021, 1*(November), 185–190. <https://doi.org/10.1109/IDAACS53288.2021.9660391>
- Maclean, F., Jones, D., Carin-Levy, G., & Hunter, H. (2013). Understanding twitter. *British Journal of Occupational Therapy, 76*(6), 295–298. <https://doi.org/10.4276/030802213X13706169933021>
- Makota, T., Maguire, B., Cagne, D., & Chakrabarti, R. (2021). *Scalable Data Streaming with Amazon Kinesis Design and Secure Highly*. Packt Publishing Limited.
- Martinez, A. R. (2010). Natural language processing. In *Wiley Interdisciplinary Reviews: Computational Statistics* (Vol. 2, Issue 3). <https://doi.org/10.1002/wics.76>
- Mistry, V. (2002). Critical care training in Spain. *Thorax, 57*(7), 658-a-658. <https://doi.org/10.1136/thorax.57.7.658-a>
- Mohanty, S., Jagadeesh, M., & Srivatsa, H. (2013). *Big Data Imperatives*. Apress.
- Mollett, A., Moran, D., & Dunleavy, P. (2015). Using Twitter in university research , teaching and impact activities. *The London School of Economics and Political Science, July*.
- Negara, E. S., Andryani, R., & Saksono, P. H. (2016). Analisis Data Twitter: Ekstraksi dan Analisis Data Geospasial. *Jurnal INKOM, 10*(1), 27. <https://doi.org/10.14203/j.inkom.433>
- Obembe, D., Kolade, O., Obembe, F., Owoseni, A., & Mafimisebi, O. (2021a). Covid-19 and the tourism industry: An early stage sentiment analysis of the impact of social media and stakeholder communication. *International Journal of Information Management Data Insights, 1*(2). <https://doi.org/10.1016/j.jjime.2021.100040>
- Obembe, D., Kolade, O., Obembe, F., Owoseni, A., & Mafimisebi, O. (2021b). Covid-19 and the tourism industry: An early stage sentiment analysis of the impact of social media and stakeholder communication. *International Journal of Information Management Data Insights, 1*(2), 100040. <https://doi.org/10.1016/j.jjime.2021.100040>
- Otviriyanti, G., Tilottama, R. D., Meilani, S. S., Fani, A. M., & Purwanta, W. (2023). Kajian Strategi Pengurangan Sampah dan Potensi Penerapan Ekonomi Sirkuler Pada Pengelolaan Sampah di Kabupaten Toba Sumatera Utara. *Jurnal Teknologi Lingkungan, 24*(2), 220–227. <https://doi.org/10.55981/jtl.2023.653>

- Özköse, H., Arı, E. S., & Gencer, C. (2015). Yesterday, Today and Tomorrow of Big Data. *Procedia - Social and Behavioral Sciences*, 195(January 2016), 1042–1050. <https://doi.org/10.1016/j.sbspro.2015.06.147>
- Pambudi, A. S., Fikri Masteriarsa, M., Dwifebri, A., Wibowo, C., Amaliyah, I., Adhitya, D., & Ardana, K. (2020). Strategi Pemulihan Ekonomi Sektor Pariwisata Pasca Covid-19. *Majalan Media Perencanaan*, 1(1), 1–21. <https://mediaperencana.perencanapembangunan.or.id/index.php/mmp/article/view/1>
- Pandu, A., Nababan, R., Lumenta, A. S. M., Rindengan, Y. D. Y., Pontoh, F. J., Akay, Y. V., Elektro, T., Sam, U., Manado, R., & Manado, J. K. B. (2020). Analisis Sentimen Twitter Pasca Pengumuman Hasil Pilpres 2019 Menggunakan Metode Lexicon Analysis. *Jurnal Teknik Informatika*, 15(1), 33–44. <https://doi.org/10.35793/jti.15.1.2020.29030>
- Pfeffer, J., Matter, D., Jaidka, K., Varol, O., Mashhadi, A., Lasser, J., Assenmacher, D., Wu, S., Yang, D., Brantner, C., Romero, D. M., Otterbacher, J., Schwemmer, C., Joseph, K., Garcia, D., & Morstatter, F. (2023). Just Another Day on Twitter: A Complete 24 Hours of Twitter Data. *Proceedings of the International AAAI Conference on Web and Social Media*, 17(Icws), 1073–1081. <https://doi.org/10.1609/icws.v17i1.22215>
- Philander, K., & Zhong, Y. Y. (2016). Twitter sentiment analysis: Capturing sentiment from integrated resort tweets. *International Journal of Hospitality Management*, 55, 16–24. <https://doi.org/10.1016/j.ijhm.2016.02.001>
- Psaltis, A. G. (2017). *Streaming Data: Understanding the real-time pipeline*.
- Purwaningrum, H., & Ahmad, H. (2021). Evaluasi Pengelolaan Wisata Jati Larangan Dan Taman Sengon Melalui Indikator 5A Di Dusun Iroyudan. *Kepariwisata : Jurnal Ilmiah*, 15(2), 107–129.
- Rahma, A. A. (2020). Potensi Sumber Daya Alam dalam Mengembangkan Sektor Pariwisata Di Indonesia. *Jurnal Nasional Pariwisata*, 12(1), 1. <https://doi.org/10.22146/jnp.52178>
- Riza, L. S., Pratama, F. D., Piantari, E., & Fahsi, M. (2020). Genomic repeats detection using Boyer-Moore algorithm on Apache Spark Streaming. *Telkomnika (Telecommunication Computing Electronics and Control)*, 18(2), 783–791. <https://doi.org/10.12928/TELKOMNIKA.V18I2.14883>
- Rochmah, W. Y., & Lentera, C. (2023). Analisis eWOM terhadap Destinasi Wisata Super Prioritas Menggunakan Text Mining Indonesian Journal of Multidisciplinary on Social and Technology Analisis eWOM terhadap Destinasi Wisata Super Prioritas Menggunakan Text Mining. *Indonesian Journal of Multidisciplinary on Social and Technology*, 1(3), 270–276. <https://doi.org/10.31004/ijmst.v1i3.227>
- Sagum, R. A. (2021). An Application of Emotion Detection in Sentiment Analysis on Movie Reviews. *Turkish Journal of Computer and Mathematics Education*

- (*TURCOMAT*), 12(3), 5468–5474.
<https://doi.org/10.17762/turcomat.v12i3.2204>
- Salukh, J. Y., Pandie, D. B. W., & Fanggidae, A. H. J. (2023). Analisis Evaluasi Kebijakan Pengembangan Pariwisata Super Prioritas Berkelanjutan di Labuan Bajo Nusa Tenggara Timur. *Jurnal Ilmiah Multidisiplin*, 1(7), 191–214.
<https://doi.org/10.5281/zenodo.8205683>
- Samah, K. A. F. A., Jailani, N. S., Hamzah, R., Aminuddin, R., & Riza, L. S. (2024a). Aspect-Based Classification and Visualization of Twitter Sentiment Analysis Towards Online Food Delivery Services in Malaysia. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 1(1), 139–150.
- Samah, K. A. F. A., Jailani, N. S., Hamzah, R., Aminuddin, R., & Riza, L. S. (2024b). Aspect-Based Classification and Visualization of Twitter Sentiment Analysis Towards Online Food Delivery Services in Malaysia. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 1(1), 139–150.
- Samah, K. A. F. A., Misdan, N. F. A., Jono, M. N. H. H., & Riza, L. S. (2022). The Best Malaysian Airline Companies Visualization through Bilingual Twitter Sentiment Analysis: A Machine Learning Classification. *International Journal on Informatics Visualization*, 6(1), 130–137.
<https://doi.org/10.30630/joiv.6.1.879>
- Sharma, S., Sarkar, D., & Gupta, D. (2012). Agile Processes and Methodologies: A Conceptual Study. *International Journal on Computer Science & Engineering*, 4(5), 892–898.
<http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=82397457&site=ehost-live>
- Shapira, G., Palino, T., Sivaram, R., & Petty, K. (2021). Kafka: The Definitive Guide, 2nd Edition. In *O'Reilly Media, Inc.*
- Singh, J. (2023). Natural Language Processing in the Real World. In *Natural Language Processing in the Real World*.
<https://doi.org/10.1201/9781003264774>
- Somantri, O., & Dairoh, D. (2019). Analisis Sentimen Penilaian Tempat Tujuan Wisata Kota Tegal Berbasis Text Mining. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 5(2), 191. <https://doi.org/10.26418/jp.v5i2.32661>
- Sotiris, Aristidis, & Vrahitis. (2018). *Real Time Sentiment Change Detection of Twitter Data Streams*.
- Steele, J., Loukides, M., & Hill, C. (2012). Big Data Now: Current Perspective form O'Reilly Radar. In *Big data now: 2012 edition*.
- Sugiyono. (2011). *Metode Penelitian Kuantitatif, Kualitatif dan R&D*.
- Sukhdeve, Dr. S. R., & Sukhdeve, S. S. (2023). Google Cloud Platform for Data Science. In *Google Cloud Platform for Data Science*.
<https://doi.org/10.1007/978-1-4842-9688-2>

- Sulistiyadi, Y., Eddyono, F., & Hasibuan, B. (2017). *Pariwisata Berkelanjutan: Pengelolaan Destinasi Wisata Berbasis Masyarakat*. AURA (Anugrah Utama Raharja).
- Tumbel, C. Z., Sitepu, H., & Hutagalung, M. (2017). Analisis Big Data Berbasis Stream Processing Menggunakan Apache Spark. *Jurnal Telematika*, 11(1), 6. <http://journal.ithb.ac.id/telematika/article/view/145>
- Wang, Z. (2017). Fine-grained sentiment analysis of social media with emotion sensing. *Future Technologies Conference (FTC)*, 1361–1364. <https://doi.org/10.1109/FTC.2016.7821783>
- Wu, S., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Who says what to whom on twitter. *Proceedings of the 20th International Conference on World Wide Web, WWW 2011, March 2011*, 705–714. <https://doi.org/10.1145/1963405.1963504>
- Yadranjiaghdam, B., Yasrobi, S., & Tabrizi, N. (2017). Developing a Real-Time Data Analytics Framework for Twitter Streaming Data. *Proceedings - 2017 IEEE 6th International Congress on Big Data, BigData Congress 2017*, 329–336. <https://doi.org/10.1109/BigDataCongress.2017.49>