

**IMPLEMENTASI *MINI BATCH K-MEANS* DALAM
MENGELOMPOKKAN DATA DOKUMEN PUTUSAN PERCERAIAN
PENGADILAN AGAMA KABUPATEN INDRAMAYU**

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh
RAUDAH YASMIN GHOZALI
09020221040

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA**

2025

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : RAUDAH YASMIN GHOZALI

NIM : 09020221040

Program Studi : Matematika

Angkatan : 2021

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "IMPLEMENTASI *MINI BATCH K-MEANS* DALAM MENGELOMPOKKAN DATA DOKUMEN PUTUSAN PERCERAIAN PENGADILAN AGAMA KABUPATEN INDRAMAYU". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 10 Juni 2025

Yang menyatakan,



KAUDAH YASMIN GHOZALI

NIM. 09020221040

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

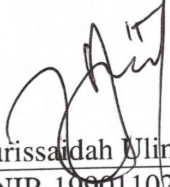
Nama : RAUDAH YASMIN GHOZALI

NIM : 09020221040

Judul skripsi : IMPLEMENTASI *MINI BATCH K-MEANS* DALAM
MENGELOMPOKKAN DATA DOKUMEN PUTUSAN
PERCERAIAN PENGADILAN AGAMA KABUPATEN
INDRAMAYU

telah diperiksa dan disetujui untuk diujikan.

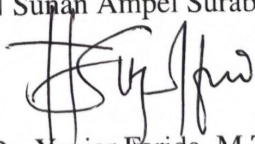
Pembimbing I


Nurissaidah Wlinnuha, M.Kom.
NIP. 199011022014032004

Pembimbing II


Wika Dianita Utami, M.Sc.
NIP. 199206102018012003

Mengetahui,
Ketua Program Studi Matematika
UIN Sunan Ampel Surabaya


Dr. Yuniar Farida, M.T.
NIP. 197905272014032002

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh

Nama : RAUDAH YASMIN GHOZALI
NIM : 09020221040
Judul skripsi : IMPLEMENTASI *MINI BATCH K-MEANS* DALAM
MENGELOMPOKKAN DATA DOKUMEN PUTUSAN
PERCERAIAN PENGADILAN AGAMA KABUPATEN
INDRAMAYU

Telah dipertahankan di depan Tim Penguji
pada tanggal 10 Juni 2025

Mengesahkan,
Tim Penguji

Penguji I



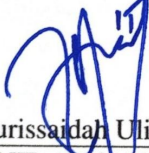
Dr. Dian Candra Rini Novitasari, M.Kom.
NIP. 198511242014032001

Penguji II



Dr. Lutfi Hakim, M.Ag.
NIP. 197312252006041001

Penguji III



Nurissaidah Ulinnuha, M.Kom.
NIP. 199011022014032004

Penguji IV



Wika Dianita Utami, M.Sc.
NIP. 199206102018012003

Mengetahui,
Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya



Dr. A. Saepul Hamdani, M.Pd.
NIP. 196507312000031002



UIN SUNAN AMPEL
SURABAYA

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Raudah Yasmin Ghozali
NIM : 09020221090
Fakultas/Jurusan : Fak. Sains dan Teknologi / Matematika
E-mail address : raudahyasminghozali@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

☒ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :

Implementasi Mini Batch K-Means dalam Mengelompokkan Data Dokumen Putusan

Percerahan Pengadilan Agama Kabupaten Indramayu

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara **fulltext** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 10 Juni 2025

Penulis

(Raudah Yasmin Ghozali)
nama terang dan tanda tangan

ABSTRAK

IMPLEMENTASI *MINI BATCH K-MEANS* DALAM MENGELOMPOKKAN DATA DOKUMEN PUTUSAN PERCERAIAN PENGADILAN AGAMA KABUPATEN INDRAMAYU

Perceraian terjadi karena pasangan suami-istri sudah tidak mampu mencapai tujuan utama dari pernikahan. Berdasarkan data Badan Pusat Statistika Jawa Barat, Kabupaten Indramayu menjadi kabupaten dengan jumlah perceraian terbanyak di Jawa Barat selama 2021-2023. Penelitian ini dilakukan untuk mengelompokkan dalil-dalil Penggugat yang tercantum dalam dokumen putusan perceraian Pengadilan Agama Kabupaten Indramayu sehingga mendapatkan informasi terkait faktor-faktor penyebab perceraian. Dalil-dalil Penggugat yang berbentuk tekstual akan dilakukan *pre-processing text* dan pembobotan kata menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Hasil pembobotan kata akan diklusterisasi menggunakan *Mini Batch K-Means*. *Mini Batch K-Means* digunakan untuk mempercepat waktu komputasi dengan menggunakan subset data (*mini batch*) tiap iterasi. Selain itu, metode inisialisasi *centroid* awal adalah pemilihan secara *random* dan menggunakan *K-Means++*. Parameter yang digunakan adalah jumlah kluster (k) dan ukuran *batch* (b). Evaluasi dari *Mini Batch K-Means* diukur menggunakan *Silhouette coefficient*, jumlah iterasi, dan kecepatan waktu komputasi. Penelitian ini menunjukkan hasil bahwa *Mini Batch K-Means* dengan inisialisasi *random* menjadi model terbaik dari segi kualitas kluster dan kecepatan komputasi dibanding dengan model dari algoritma dasar yaitu *K-Means*. Pada uji coba $k=2$ dan $b=100$, model tersebut menghasilkan nilai evaluasi *Silhouette coefficient* sebesar 0,5293 dengan jumlah iterasi sebanyak 4 iterasi dan *running time* sebesar 0,0653 detik. Berdasarkan hasil visualisasi tiap kluster, terdapat 2 kelompok topik, yakni pertengkar dan perselisihan menjadi faktor dominan terjadinya perceraian pada kluster 1 dan pekerjaan, anak, dan keuangan menjadi faktor dominan perceraian pada kluster 2.

Kata kunci: Indramayu, Klusterisasi, *Mini Batch K-Means*, Putusan Perceraian

ABSTRACT

IMPLEMENTATION OF *MINI BATCH K-MEANS* IN CLUSTERING DIVORCE VERDICT DOCUMENT DATA OF INDRAMAYU DISTRICT RELIGIOUS COURT

Divorce occurs because married couples are no longer able to achieve the main goals of marriage. Based on data from the West Java Central Bureau of Statistics, Indramayu Regency has the highest number of divorces in West Java during 2021-2023. This research was conducted to group the Plaintiff's arguments contained in the divorce decree document of the Indramayu District Religious Court in order to obtain information related to the factors causing divorce. The arguments of the Plaintiff in textual form will be pre-processed text and word weighting using Term Frequency-Inverse Document Frequency (TF-IDF). The word weighting results will be clustered using Mini Batch K-Means. Mini Batch K-Means is used to speed up computation time by using a subset of data (mini batch) per iteration. In addition, the initial centroid initialization is randomly and using K-Means++. The parameters used are the number of clusters (k) and batch size (b). The evaluation of Mini Batch K-Means is measured using Silhouette coefficient, number of iterations, and speed of computation time. This research shows that Mini Batch K-Means with random initialization is the best model in terms of cluster quality and computational speed compared to the model from the basic algorithm, namely K-Means. In the k=2 and b=100 trial, the model produced a Silhouette coefficient evaluation value of 0.5293 with a total of 4 iterations and a running time of 0.0653 seconds. Based on the visualization results of each cluster, there are 2 groups of topics, namely quarrels and disputes are the dominant factors of divorce in cluster 1 and work, children, and finances are the dominant factors of divorce in cluster 2.

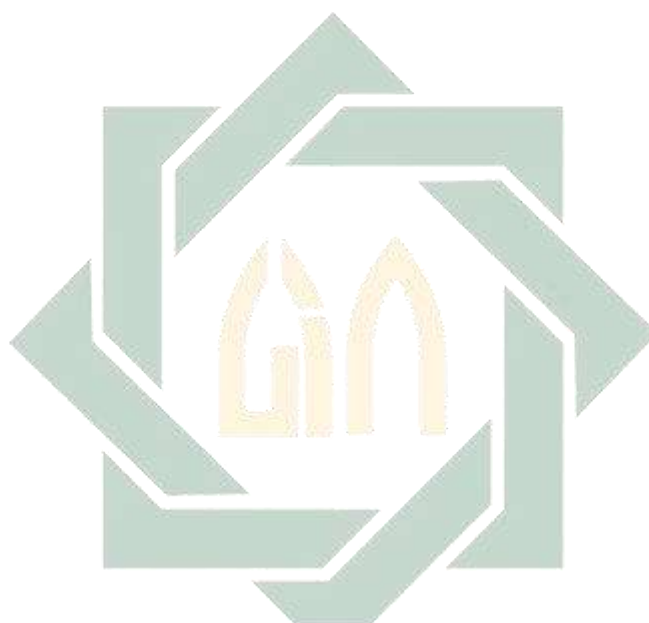
Keywords: Indramayu, Clustering, Mini Batch K-Means, Divorce Verdict

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN PEMBIMBING	ii
PENGESAHAN TIM PENGUJI SKRIPSI	iii
HALAMAN PERNYATAAN KEASLIAN	iv
MOTTO	v
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xv
ABSTRAK	xvi
ABSTRACT	xvii
I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	8
1.3. Tujuan Penelitian	8
1.4. Manfaat Penelitian	9
1.5. Batasan Masalah	10
1.6. Sistematika Penulisan	10
II TINJAUAN PUSTAKA	12
2.1. Perceraian	12
2.2. Populasi dan Sampel	15
2.3. <i>Simple Random Sampling</i>	16
2.4. <i>Text Mining</i>	17
2.5. <i>Pre-processing Text</i>	18
2.5.1. <i>Correcting Typo</i>	18
2.5.2. <i>Case Folding</i>	18

2.5.3. <i>Remove Punctuation</i>	18
2.5.4. <i>Lemmatization</i>	19
2.5.5. <i>Stopword Removal</i>	19
2.5.6. <i>Tokenizing</i>	19
2.6. <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	19
2.7. <i>K-Means++</i>	21
2.8. <i>Mini Batch K-Means</i>	22
2.9. <i>Silhouette Coefficient</i>	24
2.10. <i>Word Cloud</i>	26
2.11. <i>Integrasi Keislaman</i>	27
III METODE PENELITIAN	31
3.1. <i>Jenis Penelitian</i>	31
3.2. <i>Sumber Data</i>	31
3.3. <i>Tahapan Penelitian</i>	31
IV HASIL DAN PEMBAHASAN	37
4.1. <i>Sampling Data</i>	37
4.2. <i>Pre-processing Text</i>	37
4.2.1. <i>Correcting Typo</i>	38
4.2.2. <i>Case Folding</i>	39
4.2.3. <i>Remove Punctuation</i>	40
4.2.4. <i>Lemmatization</i>	41
4.2.5. <i>Stopword Removal</i>	42
4.2.6. <i>Tokenizing</i>	43
4.3. <i>Pembobotan Kata</i>	43
4.4. <i>Klasterisasi dengan Mini Batch K-Means</i>	47
4.4.1. <i>Proses Inisialisasi Centroid dengan Algoritma K-Means++</i>	47
4.4.2. <i>Proses Algoritma Mini Batch K-Means</i>	54
4.5. <i>Evaluasi Hasil Klaster</i>	72
4.6. <i>Visualisasi Hasil Klaster dan Interpretasi</i>	79
4.7. <i>Diskusi Hasil dan Pembahasan</i>	85

4.8. Integrasi Keislaman	87
V PENUTUP	92
5.1. Kesimpulan	92
5.2. Saran	93
DAFTAR PUSTAKA	94



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR TABEL

2.1	Kriteria <i>Silhouette Coefficient</i>	26
4.1	Contoh Teks Dokumen Putusan Perceraian	38
4.2	Contoh Tahapan <i>Correcting Typo</i>	39
4.3	Contoh Tahapan <i>Case Folding</i>	40
4.4	Contoh Tahapan <i>Remove Punctuation</i>	41
4.5	Contoh Tahapan <i>Lemmatization</i>	42
4.6	Contoh Tahapan <i>Stopword Removal</i>	42
4.7	Contoh Tahapan <i>Tokenizing</i>	43
4.8	<i>Indexing</i> Fitur Kata	43
4.9	Hasil Pembobotan Kata <i>Term Frequency</i> (TF)	44
4.10	Hasil Pembobotan Kata TF-IDF	46
4.11	<i>Centroid</i> Pertama pada <i>K-Means++</i>	47
4.12	Hasil Perhitungan Jarak Kuadrat	51
4.13	Hasil Perhitungan Jarak Kuadrat	53
4.14	Inisialisasi <i>Centroid</i> Awal	54
4.15	Jarak <i>Euclidean</i> pada Iterasi Ke-1	55
4.16	Hasil <i>Updating Centroid</i> pada Iterasi Pertama	57
4.17	Jarak <i>Euclidean</i> pada Iterasi Pertama	58
4.18	Hasil <i>Updating Centroid</i> pada Iterasi Kedua	60
4.19	Jarak <i>Euclidean</i> pada Iterasi Pertama	61
4.20	Hasil <i>Updating Centroid</i> pada Iterasi Pertama	63
4.21	Jarak <i>Euclidean</i> pada Iterasi Kedua	64
4.22	Hasil <i>Updating Centroid</i> pada Iterasi Kedua	66
4.23	Jarak <i>Euclidean</i> Minimum pada Iterasi Pertama	67
4.24	Hasil <i>Updating Centroid</i> pada Iterasi Pertama	69
4.25	Jarak <i>Euclidean</i> Minimum pada Iterasi Kedua	70

4.26 Hasil <i>Updating Centroid</i> pada Iterasi Kedua	72
4.27 Hasil Jarak <i>Euclidean</i> Antardata	73
4.28 Perbandingan Hasil Evaluasi <i>Mini Batch K-Means</i>	76
4.29 Perbandingan Hasil Evaluasi <i>K-Means</i>	77
4.30 Perbandingan <i>Mini Batch K-Means</i> dan <i>K-Means</i> dengan Inisialisasi secara Acak dan dengan <i>K-Means++</i> ($k=2$)	78



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR GAMBAR

3.1	Diagram Alir Penelitian	32
3.2	Diagram Alir <i>Mini Batch K-Means</i>	34
3.3	Diagram Alir Inisialisasi <i>K-Means++</i>	35
4.1	Visualisasi <i>Mini Batch K-Means</i> Klaster 1	79
4.2	Diagram <i>Fishbone</i> Klaster 1	81
4.3	Visualisasi <i>Mini Batch K-Means</i> Klaster 2	82
4.4	Diagram <i>Fishbone</i> Klaster 2	84



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR PUSTAKA

- Alanezi, M. A. and Hewahi, N. M. (2020). Tweets Sentiment Analysis during COVID-19 Pandemic. In *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, pages 1–6. IEEE.
- Aldalalah, O. M. A. A. (2022). Employment the Word Cloud in Brainstorming Via the Web and Its Effectiveness in Developing the Design Thinking Skill. *International journal of instruction*, 15(1):1045–1064.
- Amri, Dewi, D. R. C., and Kurniawati, L. S. (2022). Disfungsi Nilai Maqhasid Syariah terhadap Perceraian karena Alasan Ekonomi di Masa Pandemi Covid-19. *Tasamuh: Jurnal Studi Islam*, 14(2):295–315.
- Arthur, D. and Vassilvitskii, S. (2006). K-Means++: The Advantages of Careful Seeding. Technical report, Stanford.
- Arunadevi, J. and Sowjanya, A. M. (2022). Recognizing Fake Headlines Using Clustering Algorithms. *Mathematical Statistician and Engineering Applications*, 71(2):111–121.
- Aslamiah, N., Ramadhianisha, S., Azahra, S. J., et al. (2023). Kekerasan dalam Rumah Tangga sebagai Alasan Pemicu Perceraian di Pengadilan Agama Bandung Kelas 1A. *Al-Mashlahah Jurnal Hukum Islam dan Pranata Sosial*, 11(02).
- Auliya, I., Fitri, F., Amalita, N., and Mukhti, T. O. (2024). Comparison of K-Means and Fuzzy C-Means Algorithms for Clustering Based on Happiness Index

- Components Across Provinces in Indonesia. *UNP Journal of Statistics and Data Science*, 2(1):114–121.
- Badan Pusat Statistik Indonesia (2022). Indeks Kebahagiaan Menurut Provinsi, 2014-2021.
- Badan Pusat Statistik Indonesia (2024a). Jumlah Perceraian Menurut Provinsi dan Faktor, 2023.
- Badan Pusat Statistik Indonesia (2024b). Nikah dan Cerai Menurut Provinsi, 2023.
- Badan Pusat Statistik Jawa Barat (2024). Jumlah Nikah dan Cerai, 2022-2023.
- Basri, R. (2020). *Fikih Munakahat 2*, volume 17. IAIN Parepare Nusantara Press.
- Chaerudin, A., Murdiansyah, D. T., and Imrona, M. (2021). Implementation of K-Means++ Algorithm for Store Customers Segmentation Using Neo4J. *Indonesia Journal on Computing (Indo-JC)*, 6(1):53–60.
- Damayanti, P. and Yuliawan, D. (2023). Determinasi Happiness Index Melalui Pendekatan Sosial Ekonomi. *Jurnal Manajemen, Ekonomi, Hukum, Kewirausahaan, Kesehatan, Pendidikan dan Informatika (MANEKIN)*, 1(03: Maret):90–101.
- Darujati, C. and Gumelar, A. B. (2012). Pemanfaatan Teknik Supervised untuk Klasifikasi Teks Bahasa Indonesia. *Jurnal Bandung Text Mining*, 16(1):5–1.
- Dewi, C., Indriawan, F. A., and Christanto, H. J. (2023). Spam Classification Problems Using Support Vector Machine and Grid Search. *International Journal of Applied Science and Engineering*, 20(4):1–10.

Eser, G. and Sahin, C. (2024). Sentiment Analysis and Rating Prediction for App Reviews Using Transformer-based Models.

Fahmi, N. (2023). Pengaplikasian Mediasi dalam Perceraian di Era Pandemi COVID-19 pada Pengadilan Agama Pasangkayu (Tinjauan Yuridis-Normatif). *Bilancia: Jurnal Studi Ilmu Syariah dan Hukum*, 17(1):45–74.

Faishol, I. and Azzahrah, F. (2022). Perceraian disebabkan Perselingkuhan (Studi Kasus Satu Keluarga di Kelurahan Oesapa Kupang). *Mutawasith: Jurnal Hukum Islam*, 5(1):54–67.

Fajrussalam, H., Rahman, E. A., Hafizha, F. Z., Ulhaq, S., et al. (2023). Hakikat dan Eksistensi Manusia sebagai Mahluk yang Bermoral. *Innovative: Journal Of Social Science Research*, 3(2):1706–1721.

Farida, Y., Khariri, A. F., Yuliati, D., and Khaulasari, H. (2022). Clustering Couples of Childbearing Age to Get Family Planning Counseling using K-Means Method. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 22(1):189–200.

Fauzy, A. (2019). *Metode Sampling*. Universitas Terbuka, Tangerang Selatan, kedua edition.

Feizollah, A., Anuar, N. B., Salleh, R., and Amalina, F. (2014). Comparative Study of K-Means and Mini Batch K-Means Clustering Algorithms in Android Malware Detection Using Network Traffic Analysis. In *2014 International Symposium on Biometrics and Security Technologies (ISBAST)*, pages 193–197.

Hamdi, A., Asyhar, A., Farida, Y., Ulinnuha, N., Novitasari, D., and Arifin, A. (2019). Sentiment Analysis of Regional Head Candidate's Electability from the

- National Mass Media Perspective Using the Text Mining Algorithm. *Advances in Science, Technology and Engineering Systems*, 4:134–139.
- Hamid, H. (2018). Perceraian dan Penanganannya. *Jurnal Ilmiah Wahana Pendidikan*, 4(3):24–29.
- Handayani, L. (2022). Prosedur Perceraian Ditinjau berdasarkan Hukum Fiqh dan Hukum Positif. *Journal of Legal and Cultural Analytics*, 1(1):1–18.
- Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., and Yeganegi, M. R. (2020). Text Mining in Big Data Analytics. *Big Data and Cognitive Computing*, 4(1):1.
- Hilman, M. F., Passarella, R., Kurniawan, D., and Sutarno, S. (2024). Automatic Text Summarization on Aviation Traffic Accident Report Synopsis Using Term Frequency-Inverse Document Frequency (TF-IDF) Algorithm. *Available at SSRN 4758981*.
- Igescu, S., Sanca, V., Zapridou, E., and Ailamaki, A. (2023). Improving K-Means Clustering Using Speculation. In *Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW'23)*.
- Januari, N. (2023). Menggali Akar Masalah: Analisis Kasus Perceraian di Indonesia. *AKADEMIK: Jurnal Mahasiswa Humanis*, 3(3):120–130.
- Jarana, B. (2024). *An Unsupervised Approach to COVID-19 Fake Tweet Detection*. PhD thesis.
- Kamyab, M., Liu, G., and Adjeisah, M. (2021). Attention-based CNN and Bi-LSTM Model Based on TF-IDF and Glove Word Embedding for Sentiment Analysis. *Applied Sciences*, 11(23):11255.

- Karo, I. M. K., Dewi, S., Mardiana, M., Ramadhani, F., and Harliana, P. (2023). K-Means and K-Medoids Algorithm Comparison for Clustering Forest Fire Location in Indonesia. *Jurnal Ecotipe (Electronic, Control, Telecommunication, Information, and Power Engineering)*, 10(1):86–94.
- Khan, R., Qian, Y., and Naeem, S. (2019). Extractive Based Text Summarization Using K-Means and TF-IDF. *International Journal of Information Engineering and Electronic Business*, 12(3):33.
- Khurin'In, D., Muta'alimin, M., Maulana, A., and Musyafa'ah, N. L. (2022). Perceraian akibat Perselisihan dan Pertengkaran Perspektif Hukum Islam. *Ma'mal: Jurnal Laboratorium Syariah Dan Hukum*, 3(1):18–37.
- Kumar, L. and Bhatia, P. K. (2013). Text Mining: Concepts, Process and Applications. *Journal of Global Research in Computer Science*, 4(3):36–39.
- Kusmiran, A. (2023). Clustering and Risk Analysis of The Earthquake in Sulawesi Using Mini Batch K-Means, K-Medoids, and Maximum Likelihood Method. *Elkawanie: Journal of Islamic Science and Technology*, 9(1):1–23.
- Miner, G. (2012). *Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications*. Academic Press.
- Momahhed, S. S., Emamgholipour Sefiddashti, S., Minaei, B., and Shahali, Z. (2023). K-Means Clustering of Outpatient Prescription Claims for Health Insureds in Iran. *BMC public health*, 23(1):788.
- Mukti, F. S., Junikhah, A., Putra, P. M. A., Soetedjo, A., and Krismanto, A. U. (2022). A Clustering Optimization for Energy Consumption Problems in

- Wireless Sensor Networks Using Modified K-Means++ Algorithm. *International Journal of Intelligent Engineering & Systems*, 15(3).
- Mulyadi, E. S. (2023). Peran Pengadilan Agama Kota Banjar sebagai Penegak Hukum dalam Penanganan dan Penanggulangan Perceraian. *Journal Presumption of Law*, 5(1):60–72.
- Mulyadi, S., Insani, F., Agustian, S., and Afriyanti, L. (2024). Pengelompokan Data Pendistribusian Listrik Menggunakan Algoritma Mini Batch K-Means Clustering: Grouping Electricity Distribution Data Using The Mini Batch K-Means Clustering Algorithm. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(3):1051–1062.
- Musyarrafa, N. I. and Khalik, S. (2020). Batas Usia Pernikahan dalam Islam: Analisis Ulama Mazhab terhadap Batas Usia Nikah. *Shautuna: Jurnal Ilmiah Mahasiswa Perbandingan Mazhab*.
- Ningtias, I. S. (2022). Faktor yang Mempengaruhi Penurunan Angka Pernikahan di Indonesia. *Jurnal Registratie*, 4(2):87–98.
- Noor, S., Tajik, O., and Golzar, J. (2022). Simple Random Sampling. *International Journal of Education & Language Studies*, 1(2):78–82.
- Nugroho, N. and Adhinata, F. D. (2022). Penggunaan Metode K-Means dan K-Means++ sebagai Clustering Data Covid-19 di Pulau Jawa. *Teknika*, 11(3):170–179.
- Nurhalisa, R. (2021). Tinjauan Literatur: Faktor Penyebab dan Upaya Pencegahan Sistematis terhadap Perceraian. *Media Gizi Kesmas*, 10(1):157.

- Olabanjo, O., Wusu, A., Afisi, O., Asokere, M., Padonu, R., Olabanjo, O., Ojo, O., Folorunso, O., Aribisala, B., and Mazzara, M. (2023). From Twitter to Aso-Rock: A Sentiment Analysis Framework for Understanding Nigeria 2023 Presidential Election. *Heliyon*, 9(5).
- Park, J., Park, K. V., Yoo, S., Choi, S. O., and Han, S. W. (2020). Development of The WEEE Grouping System in South Korea Using The Hierarchical and Non-Hierarchical Clustering Algorithms. *Resources, Conservation and Recycling*, 161:104884.
- Prabowo, W. A. and Azizah, F. (2020). Sentiment Analysis for Detecting Cyberbullying Using TF-IDF and SVM. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(6):1142–1148.
- Pradipta, N. Y. and Soetanto, H. (2024). Sentiment Classification of General Election 2024 News Titles on Detik. com Online Media Website Using Multinomial Naive Bayes Method. *Journal of Applied Science, Engineering, Technology, and Education*, 6(1):43–55.
- Pragholapati, A. (2020). Dampak Perceraian di Indonesia: Systematic Literature Review.
- Prastyo, P. H., Sumi, A. S., Dian, A. W., and Permanasari, A. E. (2020). Tweets Responding to The Indonesian Government's Handling of COVID-19: Sentiment Analysis Using SVM with Normalized Poly Kernel. *J. Inf. Syst. Eng. Bus. Intell*, 6(2):112.
- Pratiwi, S. I. (2022). 6 Kota Penyumbang Perceraian Tinggi di Indonesia, Salah Satunya Indramayu.

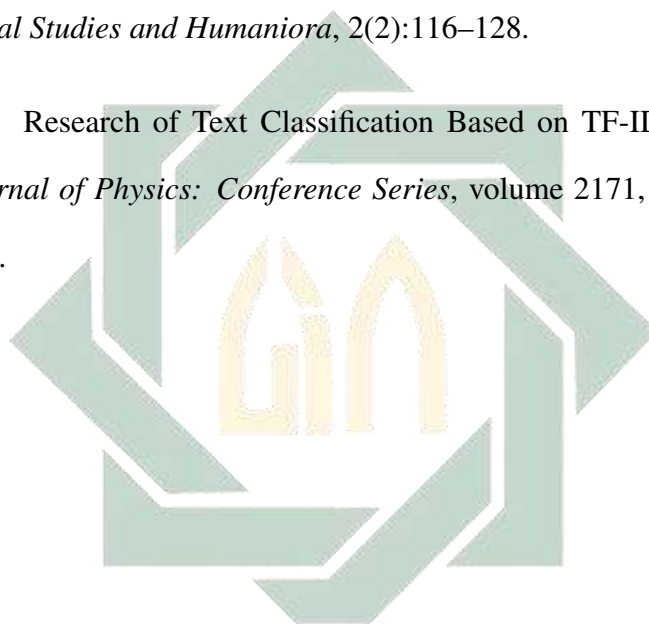
- Punhani, A., Faujdar, N., Mishra, K. K., and Subramanian, M. (2022). Binning-based Silhouette Approach to Find the Optimal Cluster Using K-Means. *IEEE Access*, 10:115025–115032.
- Purnajaya, A. R., Lieputra, V., Tayanto, V., and Salim, J. G. (2022). Implementasi Text Mining untuk Mengetahui Opini Masyarakat tentang Climate Change. *Journal of Information System and Technology (JOINT)*, 3(3):320–328.
- Qaiser, S. and Ali, R. (2018). Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents. *International Journal of Computer Applications*, 181(1):25–29.
- Qorib, M., Oladunni, T., Denis, M., Ososanya, E., and Cota, P. (2023). COVID-19 Vaccine Hesitancy: Text Mining, Sentiment Analysis and Machine Learning on COVID-19 Vaccination Twitter Dataset. *Expert Systems with Applications*, 212:118715.
- Rahim, R., Sao'dah, Tiring, S. S. N. D., Asman, Fitriyah, L. A., Dewi, M. S., R, I. H., Ferawati, Mutia, Pamungkas, M. D., Sutrisno, E., Wulandari, H., Trimurtini, and Wicaksono, A. B. (2021). *Metodologi Penelitian (Teori dan Praktik)*. Perkumpulan Rumah Cemerlang Indonesia, Tasikmalaya.
- Rasmadi, S. (2023). Masalah 'Klasik' Pemicu Tingginya Perceraian di Indramayu.
- Sculley, D. (2010). Web-Scale K-Means Clustering. In *Proceedings of the 19th international conference on World wide web*, pages 1177–1178.
- Shahapure, K. R. and Nicholas, C. (2020). Cluster Quality Analysis Using Silhouette Score. In *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*, pages 747–748. IEEE.

- Sugiyono (2013). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta, Bandung.
- Sukenti, D., Tambak, S., and Charlina (2020). Developing Indonesian Language Learning Assessments : Strengthening The Personal Competence and Islamic Psychosocial of Teachers. *International Journal of Evaluation and Research in Education (IJERE)*, 9(4):1079–1087.
- Sun, B., Mao, H., and Yin, C. (2020). Male and Female Users' Differences in Online Technology Community Based on Text Mining. *Frontiers in Psychology*, 11:806.
- Surianto, D. F. (2023). Clustering Tweets Data on Twitter Social Media Using K-Means Method. *Journal of Security, Computer, Information, Embedded, Network, and Intelligence System*, pages 44–51.
- Susanto, E. B. (2016). Evaluasi Hasil Klaster pada Dataset Iris, Soybean-small, Wine Menggunakan Algoritma Fuzzy C-Means dan K-Means++. *Jurnal Surya Informatika*, 2(1):6–13.
- Syahrial, M. (2023). Perkara Perceraian di Indramayu Masih Tinggi, Ada yang Masih 19 Tahun Sudah Dua Kali Bercerai.
- Syaifuddin, M., Turatmiyah, S., and Yahanan, A. (2022). *Hukum Perceraian*. Sinar Grafika.
- Syawaluddin, M., Hidayatullah, S., and Putra, M. Y. (2024). Motif Perceraian pada Masyarakat Desa Wawonduru dan Dampak terhadap Psikologi Anak. *NALAR: Journal Of Law and Sharia*, 2(3):195–202.

Ulinnuha, N. and Indriyani, J. G. (2023). Ekstraksi Topik Pantun di Twitter Menggunakan K-Means Clustering. *KUBIK: Jurnal Publikasi Ilmiah Matematika*, 8(1):24–34.

Wijaya, M. D. P., Saputri, R. P. M., and Danugroho, A. (2023). Konflik Disintegrasi di Indonesia dan Dampaknya Bagi Nasionalisme Bangsa Indonesia. *Antroposen: Journal of Social Studies and Humaniora*, 2(2):116–128.

Zhou, H. (2022). Research of Text Classification Based on TF-IDF and CNN-LSTM. In *Journal of Physics: Conference Series*, volume 2171, page 012021. IOP Publishing.



UIN SUNAN AMPEL
S U R A B A Y A