

**ANALISIS KLASTERISASI EMOSI DALAM TERJEMAHAN AL-
QUR'AN MENGGUNAKAN *K-MEANS CLUSTERING***

SKRIPSI

Diajukan untuk memenuhi sebagian syarat untuk memperoleh gelar Sarjana
Komputer Program Studi Rekayasa Perangkat Lunak



Oleh:

Muhamad Faisal Fiqri

NIM 2008657

PROGRAM STUDI REKAYASA PERANGKAT LUNAK

KAMPUS UPI DI CIBIRU

UNIVERSITAS PENDIDIKAN INDONESIA

2024

**ANALISIS KLASTERISASI EMOSI DALAM TERJEMAHAN AL-QUR'AN
MENGUNAKAN *K-MEANS CLUSTERING***

Oleh
Muhamad Faisal Fiqri
2008657

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer pada Program Studi Rekayasa Perangkat Lunak

©Muhamad Faisal Fiqri
Universitas Pendidikan Indonesia
Agustus 2024

Hak cipta dilindungi undang-undang
skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian,
dengan dicetak ulang, difotokopi, atau cara lainnya tanpa izin dari penulis

HALAMAN PENGESAHAN

Muhamad Faisal Fiqri

**ANALISIS KLASTERISASI EMOSI DALAM TERJEMAHAN AL-
QUR'AN MENGGUNAKAN *K-MEANS CLUSTERING***

Disetujui dan disahkan oleh pembimbing:

Pembimbing I

Raditya Muhammad, M.T.

NIP: 920190219920507101

Pembimbing II

Mochamad Iqbal Ardimansyah, S.T., M.Kom.

NIP. 920190219910328101

Mengetahui

Ketua Program Studi Rekayasa Perangkat Lunak

Mochamad Iqbal Ardimansyah, S.T., M.Kom.

NIP. 920190219910328101

PERNYATAAN KEASLIAN SKIRPSI DAN BEBAS PLAGIARISME

Dengan ini saya menyatakan bahwa skripsi dengan judul **ANALISIS KLASTERISASI EMOSI DALAM TERJEMAHAN AL-QUR'AN MENGGUNAKAN *K-MEANS CLUSTERING*** ini beserta seluruh isinya adalah benar-benar karya saya sendiri. Saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika ilmu yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung risiko/sanksi apabila di kemudian hari ditemukan adanya pelanggaran etika keilmuan atau ada klaim dari pihak lain terhadap keaslian karya saya ini.

Bandung 11 Agustus 2024
Yang membuat pernyataan,

Muhamad Faisal Fiqri
NIM 2008657

UCAPAN TERIMA KASIH

Segala puji serta rasa syukur kepada Allah SWT yang telah memberikan limpahan berkah, karunia, serta petunjuk-Nya sehingga peneliti dapat melaksanakan dan menyelesaikan skripsi dengan judul “ANALISIS KLASTERISASI EMOSI DALAM TERJEMAHAN AL-QUR’AN MENGGUNAKAN *K-MEANS CLUSTERING*”. Tidak lupa shalawat dan salam semoga senantiasa tercurah limpahkan kepada junjungan Nabi Muhammad SAW, semoga keteladanannya tetap terjaga bagi umat manusia hingga akhir zaman.

Penulis menyadari bahwa selama penyusunan skripsi ini tidak terlepas oleh bantuan serta dukungan dari berbagai pihak. Maka dari itu, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada beberapa pihak yang disebutkan sebagai berikut.

1. Para Direktur dan Wakil Direktur di Kampus Cibiru, Universitas Pendidikan Indonesia, atas segala dukungan dalam pengelolaan akademik selama masa studi.
2. Bapak Raditya Muhammad, S.T., M.T., selaku dosen pembimbing pertama yang telah memberikan arahan serta dorongan motivasi dalam tiap langkah penelitian.
3. Bapak M. Iqbal Ardiansyah, S.T., M.Kom, selaku dosen pembimbing kedua serta kepala program studi Rekayasa Perangkat Lunak yang telah memberikan bimbingan selama pengerjaan skripsi serta selama pelaksanaan perkuliahan.
4. Para dosen beserta staff program studi Rekayasa Perangkat Lunak yang telah memberikan usaha terbaik dalam kegiatan perkuliahan selama menempuh pendidikan.
5. Kakak perempuan penulis sebagai sosok yang berperan sebagai orang tua yang saya hormati, yang telah menanggung berbagai hal selama penulis menjalani perkuliahan, serta selalu memberikan dukungan serta doa selama penelitian ini berlangsung.
6. Seluruh keluarga besar yang telah memberikan dukungan baik berupa moril maupun materil.

7. Teman-teman RPL kelas B sepejuangan yang telah menghabiskan masa-masa perkuliahan dengan penuh kekompakan.

Penulis telah memberikan usaha yang maksimal dalam menyusun penelitian ini, namun dengan segala keterbatasan yang dimiliki, tidaklah ada yang sempurna selain Allah SWT. Penulis berharap dengan dukungan dari pihak-pihak yang telah disebutkan di atas, penelitian ini dapat memberikan kontribusi yang bermanfaat dalam perkembangan ilmu pengetahuan. Terima kasih kepada seluruh pihak yang telah memberikan kontribusi dalam menyelesaikan penelitian ini, semoga segala kebaikan yang telah diberikan akan mendapat balasan yang lebih besar lagi oleh Allah SWT. Aamiin yaa rabbal aalaamiin.

Bandung, 11 Agustus 2024

Muhamad Faisal Fiqri

ANALISIS KLASTERISASI EMOSI DALAM TERJEMAHAN AL-QUR'AN MENGGUNAKAN *K-MEANS CLUSTERING*

MUHAMAD FAISAL FIQRI

NIM 2008657

ABSTRAK

Al-Qur'an sebagai firman Allah yang kaya akan sumber ilmu serta informasi, mencakup aspek spiritual, moral, sosial, dan psikologis, termasuk di dalamnya petunjuk tentang pengenalan emosi yang memiliki dampak signifikan terhadap kecerdasan emosional seseorang. Penelitian ini bertujuan untuk mengidentifikasi dan mengelompokkan ayat-ayat terjemahan Al-Qur'an berbahasa Indonesia yang memuat emosi dasar seperti senang, sedih, takut, marah, jijik, dan terkejut. Proses ini melibatkan preprocessing data, pencarian ayat menggunakan *Vector Space Model*, serta penerapan algoritma *K-Means Clustering*. Hasilnya, ayat-ayat tersebut dapat dikelompokkan ke dalam empat klaster utama. Adapun karakteristik dari klaster yang terbentuk diantaranya, klaster 0 menunjukkan pengelompokan ayat yang memuat kata "senang", klaster 1 dan 2 masing-masing menunjukkan kata "takut", dan klaster 3 menunjukkan kata "sedih". Hasil evaluasi klaster yang diperoleh dengan menggunakan *Silhouette Score* sebesar 0,442 dan *Calinski-Harabasz Index* sebesar 251,653 yang mengindikasikan terbentuknya klaster dengan struktur yang lemah yang berpotensi untuk ditingkatkan lebih lanjut.

Kata Kunci: *K-Means Clustering*, Emosi Dasar, Al-Qur'an Terjemahan, *Text Mining*

CLUSTER ANALYSIS OF EMOTIONS IN QURANIC TRANSLATIONS USING K-MEANS CLUSTERING

MUHAMAD FAISAL FIQRI

NIM 2008657

ABSTRACT

Al-Qur'an as the word of God is a rich source of knowledge and information, covering spiritual, moral, social, and psychological aspects, including instructions on the recognition of emotions that have a significant impact on a person's emotional intelligence. This study aims to identify and categorize verses of the Indonesian translation of the Qur'an that contain basic emotions such as happy, sad, fear, anger, disgust, and surprise. The process involves data preprocessing, verse search using Vector Space Model, and application of K-Means Clustering algorithm. As a result, the verses can be grouped into four main clusters. The characteristics of the clusters formed include, cluster 0 shows the grouping of verses containing the word "happy", clusters 1 and 2 respectively show the word "fear", and cluster 3 shows the word "sad". The cluster evaluation results obtained using the Silhouette Score of 0.442 and the Calinski-Harabasz Index of 251.653 indicate the formation of clusters with a weak structure that has the potential for further improvement.

Keywords: *K-Means Clustering, Basic Emotion, Al-Qur'an Translation, Text Mining*

DAFTAR ISI

HALAMAN PENGESAHAN.....	iii
PERNYATAAN KEASLIAN SKIRPSI DAN BEBAS PLAGIARISME	iv
UCAPAN TERIMA KASIH.....	v
ABSTRAK.....	vii
<i>ABSTRACT</i>	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah Penelitian	4
1.3. Tujuan Penelitian.....	4
1.4. Manfaat Penelitian.....	4
1.5. Batasan Masalah.....	4
1.6. Sistematika Penulisan Skripsi	5
BAB II TINJAUAN PUSTAKA.....	6
2.1. Penelitian Terkait	6
2.2. Emosi Dasar	9
2.3. Terjemahan Al-Qur'an	11
2.4. <i>Text mining</i>	12
2.5. <i>Text Preprocessing</i>	14
2.6. TF-IDF (<i>Term Frequency-Inverse Document Frequency</i>).....	15
2.7. <i>Vector Space Model</i>	16
2.8. <i>Cosine Similarity</i>	17
2.9. <i>Principal Component Analysis (PCA)</i>	17
2.10. <i>K-means Clustering</i>	18
2.11. <i>Elbow Method</i>	18
2.12. <i>Wordcloud</i>	19
2.13. Metrik Evaluasi Kluster	20
BAB III METODE PENELITIAN.....	22
3.1. <i>Design Science Research Method (DSRM)</i>	22
3.2. Instrumen Penelitian.....	24

3.3.	Alat dan Bahan	24
3.4.	Prosedur Penelitian	25
3.4.1.	Dataset.....	25
3.4.2.	<i>Preprocessing Data</i>	26
3.4.3.	<i>Vector Space Model</i>	26
3.4.4.	Dataset terjemahan Al-Qur'an berdasarekan topik emosi dasar	27
3.4.5.	<i>Clustering</i>	27
3.4.6.	Analisis Hasil	28
BAB IV TEMUAN DAN PEMBAHASAN		29
4.1.	Dataset Terjemahan Al-Qur'an	29
4.2.	<i>Preprocessing Data</i>	29
4.2.1.	<i>Case Folding</i>	29
4.2.2.	<i>Punctual Removal</i>	30
4.2.3.	<i>Tokenization</i>	30
4.2.4.	<i>Stemming</i>	31
4.2.5.	<i>Stopwords Removal</i>	31
4.3.	<i>Vector Space Model</i>	32
4.3.1.	<i>TF-IDF Vectorization</i>	32
4.3.2.	<i>Cosine Similarity</i>	33
4.3.3.	Pencarian Ayat Topik Emosi Dasar	34
4.4.	Dataset Terjemahan Al-Qur'an berdasarkan Topik Emosi Dasar	34
4.5.	<i>Clustering</i>	35
4.5.1.	Penentuan Nilai Kluster Optimal	35
4.5.2.	<i>K-means Clustering</i>	36
4.5.3.	Visualisasi Kluster.....	37
4.6.	Evaluasi Hasil dan Pembahasan	47
4.6.1.	<i>Sillhouette Score</i>	47
4.6.2.	<i>Calinski Harabasz Index</i>	48
BAB V SIMPULAN, IMPLIKASI DAN REKOMENDASI		51
5.1.	Simpulan.....	51
5.2.	Implikasi.....	51
5.3.	Rekomendasi	52
DAFTAR PUSTAKA		53
LAMPIRAN.....		63

DAFTAR GAMBAR

Gambar 2.1 Subjek Area dalam <i>Text Mining</i>	12
Gambar 2.2 Ilustrasi Vector Space Model	16
Gambar 2.3 <i>Elbow Method</i>	19
Gambar 2.4 Contoh <i>Word Cloud</i>	19
Gambar 3.1 Tahapan Design Science Research Method (DSRM)	22
Gambar 3.2 Rancangan Prosedur Penelitian	25
Gambar 3.3 Dataset Terjemahan Al-Qur'an	25
Gambar 3.4 Tahapan <i>Preprocessing Data</i>	26
Gambar 3.5 Tahapan <i>Vector Space Model</i>	27
Gambar 3.6 Tahapan Clustering	27
Gambar 4.1 Data Terjemahan Ayat Al-Qur'an.....	29
Gambar 4.2 Hasil Pembobotan Kata Menggunakan <i>TF-IDF Vectorization</i>	33
Gambar 4.3 Matriks Kesamaan Dokumen	34
Gambar 4.4 Input Kata Kunci Emosi Dasar untuk Pencarian Al-Qur'an	34
Gambar 4.5 Dataset Terjemahan Al-Qur'an Emosi Dasar.....	35
Gambar 4.6 <i>Plot Elbow Method</i>	36
Gambar 4.7 Visualisasi Hasil Klaster	37
Gambar 4.8 <i>Worcloud Cluster 0</i>	38
Gambar 4.9 <i>Worcloud Cluster 1</i>	40
Gambar 4.10 <i>Worcloud Cluster 2</i>	42
Gambar 4.11 <i>Worcloud Cluster 3</i>	44
Gambar 4.12 Plot <i>Sillhouette Score</i> k=3, k=4 dan k=5.....	48
Gambar 4.13 Plot CH Index k=2 hingga k=14	49

DAFTAR TABEL

Tabel 2.1 Kajian Penelitian Terkait	6
Tabel 3.1 Interpretasi <i>Sillhouette Score</i>	24
Tabel 4.1 <i>Input Output Case Folding</i>	29
Tabel 4.2 <i>Input Output Punctual Removal</i>	30
Tabel 4.3 <i>Input Output Tokenization</i>	30
Tabel 4.4 <i>Input Output Stemming</i>	31
Tabel 4.5 <i>Input Output Stopwords Removal</i>	32
Tabel 4.6 Terjemahan Ayat Al-Qur'an dalam Klaster 0	38
Tabel 4.7 Terjemahan Ayat Al-Qur'an dalam Klaster 1	40
Tabel 4.8 Terjemahan Ayat Al-Qur'an dalam Klaster 2	43
Tabel 4.9 Terjemahan Ayat Al-Qur'an dalam Klaster 3	45
Tabel 4.10 Rangkuman Jumlah Klaster	47

DAFTAR PUSTAKA

- Abas, N. A. (2021). A Qur'anic Concept of Sadness: Implication for CBT-Q Content Supplement (Part 1). *Ulumuna*, 25(1), Article 1. <https://doi.org/10.20414/ujis.v25i1.438>
- Abdeljaber, H. A. (2021). Automatic Arabic Short Answers Scoring Using Longest Common Subsequence and Arabic WordNet. *IEEE Access*, 9, 76433–76445. IEEE Access. <https://doi.org/10.1109/ACCESS.2021.3082408>
- Abdi, H., & Williams, L. J. (2010). Principal component analysis. *WIREs Computational Statistics*, 2(4), 433–459. <https://doi.org/10.1002/wics.101>
- al-Domi, M. M. (2015). The Cry in the Holy Quran and the Effect on the Human Behavior. *Journal of Education and Practice*, 6(20), 52–61.
- al-Ghazzālī, 1058-1111 (with Princeton Theological Seminary Library). (1910). *The alchemy of happiness by al Ghazzali*. London, J. Murray. <http://archive.org/details/alchemyofhappine00algh>
- Aldeeky, M. (2016). *The word of fear in the Holy Qur'an, a. 6(2)*.
- Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). *A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques* (arXiv:1707.02919). arXiv. <https://doi.org/10.48550/arXiv.1707.02919>
- Anggraeni, I. L., & Rozali, Y. A. (2023). Quarter Life Crisis Ditinjau Dari Kecerdasan Emosional Pada Dewasa Awal. *Psychommunity Seminar Nasional Psikologi Esa Unggul*, 0, Article 0. <https://prosiding.esaunggul.ac.id/index.php/SEMNASPSIKOLOGI/article/view/306>
- Anwar, N. (2018). Ayat-Ayat Tentang Masyarakat: Kajian Konsep Dan Implikasinya Dalam Pengembangan Pendidikan Islam: *Halaqa: Islamic Education Journal*, 2(2), Article 2. <https://doi.org/10.21070/halaqa.v2i2.1778>
- Aritama, R. (2021, Januari 21). Perkembangan Penerjemahan Al-Quran di Indonesia dari Masa ke Masa. *Tafsir Al Quran | Referensi Tafsir di Indonesia*. <https://tafsiralquran.id/perkembangan-penerjemahan-al-quran-di-indonesia-dari-masa-ke-masa/>

- Ariyanti, D., & Iswardani, K. (2020). Teks Mining untuk Klasifikasi Keluhan Masyarakat Menggunakan Algoritma Naive Bayes. *IKRA-ITH Informatika : Jurnal Komputer Dan Informatika*, 4(3), Article 3.
- Ariyanti, M. A., Wibawa, A. P., & Pujiyanto, U. (2019). Metode term frequency— Invers document frequency pada mekanisme pencarian judul skripsi. *TEKNO: Jurnal Teknologi Elektro dan Kejuruan*, 28(2), Article 2. <https://doi.org/10.17977/um034v28i2p177-190>
- Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *Journal of the Learning Sciences*, 13(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1
- Bayati, H., Davoudi, H., & Fatemizadeh, E. (2008). A heuristic method for finding the optimal number of clusters with application in medical data. *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4684–4687. <https://doi.org/10.1109/IEMBS.2008.4650258>
- Cahyani, L., & Arif, M. (2022). Text Mining untuk Pengelompokan Skripsi di Prodi Pendidikan Informatika Universitas Trunojoyo Madura. *Jurnal Ilmiah Edutic : Pendidikan Dan Informatika*, 8(2), Article 2. <https://doi.org/10.21107/edutic.v8i2.13020>
- Cao, J., Fang, J., Meng, Z., & Liang, S. (2023). *Knowledge Graph Embedding: A Survey from the Perspective of Representation Spaces* (arXiv:2211.03536). arXiv. <https://doi.org/10.48550/arXiv.2211.03536>
- Celebi, M. E., Kingravi, H. A., & Vela, P. A. (2013). A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems with Applications*, 40(1), 200–210. <https://doi.org/10.1016/j.eswa.2012.07.021>
- Chu, C.-Y., Park, K., & Kremer, G. E. (2019). Applying Text-mining Techniques to Global Supply Chain Region Selection: Considering Regional Differences. *Procedia Manufacturing*, 39, 1691–1698. <https://doi.org/10.1016/j.promfg.2020.01.271>
- Damiani, C., Maspero, D., Di Filippo, M., Colombo, R., Pescini, D., Graudenzi, A., Westerhoff, H. V., Alberghina, L., Vanoni, M., & Mauri, G. (2019). Integration of single-cell RNA-seq data into population models to

- characterize cancer metabolism. *PLOS Computational Biology*, 15(2), 1–25. <https://doi.org/10.1371/journal.pcbi.1006733>
- DePaolo, C., & Wilkinson, K. (2014). Get Your Head into the Clouds: Using Word Clouds for Analyzing Qualitative Assessment Data. *TechTrends*, 58, 38–44. <https://doi.org/10.1007/s11528-014-0750-9>
- Dey, L., & Verma, I. (2013). Text-Driven Multi-structured Data Analytics for Enterprise Intelligence. *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, 3, 213–220. <https://doi.org/10.1109/WI-IAT.2013.186>
- D.Napoleon, & S.Pavalakodi. (2011). A New Method for Dimensionality Reduction using K-Means Clustering Algorithm for High Dimensional Data Set. *International Journal of Computer Applications*, 13(7), 41–46.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Gahman, N., & Elangovan, V. (2023). *A Comparison of Document Similarity Algorithms* (arXiv:2304.01330). arXiv. <https://doi.org/10.48550/arXiv.2304.01330>
- HaCohen-Kerner, Y., Miller, D., & Yigal, Y. (2020). The influence of preprocessing on text classification using a bag-of-words representation. *PLOS ONE*, 15(5), 1–22. <https://doi.org/10.1371/journal.pone.0232525>
- Hariharan, S., & Srinivasan, R. (2008). A Comparison of Similarity Measures for Text Documents. *Journal of Information & Knowledge Management*, 07(01), 1–8. <https://doi.org/10.1142/S0219649208001889>
- Hartinah, H., Paundu, A. W., & Ilham, A. A. (2023). Deteksi Malware Ransomware Berdasarkan Panggilan API dengan Metode Ekstraksi Fitur N-gram dan TF-IDF. *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, 9(1), 50–58. <https://doi.org/10.26418/jp.v9i1.58721>
- Hasanah, S. N. (2024). Perbandingan Metode Clustering pada Kelompok Dokumen Teks Untuk Menentukan Nilai Ambang Batas Similaritas. *Nusantara Journal of Multidisciplinary Science*, 1(8), Article 8.

- Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text Mining in Big Data Analytics. *Big Data and Cognitive Computing*, 4(1), Article 1. <https://doi.org/10.3390/bdcc4010001>
- How to Improve Your Emotional Intelligence*. (2019, Agustus 26). Professional & Executive Development | Harvard DCE. <https://professional.dce.harvard.edu/blog/how-to-improve-your-emotional-intelligence/>
- Husaini, A. H., Mayasari, R., & Susilawati, S. (2022). Pengelompokan Ulasan Aplikasi PeduliLindungi dengan Algoritma K-Medoids. *INTECOMS: Journal of Information Technology and Computer Science*, 5(2), 131–139. <https://doi.org/10.31539/intecom.s.v5i2.4410>
- Idaman, I., & Hidayat, S. (2011). Al-Qur'an Dan Kecerdasan Spiritual: Upaya Menyingkap Rahasia Allah Dalam Al-Qur'an. *Khatulistiwa*, 1(1). <https://doi.org/10.24260/khatulistiwa.v1i1.181>
- Imperial, J. M., & Ong, E. (2021). *Under the Microscope: Interpreting Readability Assessment Models for Filipino* (arXiv:2110.00157). arXiv. <https://doi.org/10.48550/arXiv.2110.00157>
- Indriyanto, I., & Sumitra, I. D. (2019). Measuring the Level of Plagiarism of Thesis using Vector Space Model and Cosine Similarity Methods. *IOP Conference Series: Materials Science and Engineering*, 662(2), 022111. <https://doi.org/10.1088/1757-899X/662/2/022111>
- Izutsu, T. (2002). *Ethico-Religious Concepts in the Qur'an*. McGill-Queen's Press - MQUP.
- Januzaj, Y., Beqiri, E., & Luma, A. (2023). Determining the Optimal Number of Clusters using Silhouette Score as a Data Mining Technique. *International Journal of Online and Biomedical Engineering (iJOE)*, 19(04), 174–182. <https://doi.org/10.3991/ijoe.v19i04.37059>
- Jideani, V. (2019). Integrating Text Mining and Network Analysis for Topic Detection from Published Articles on Banana Sensory Characteristics. Dalam *Banana Nutrition—Function and Processing Kinetics*. IntechOpen. <https://doi.org/10.5772/intechopen.84857>

- Joshi, P., & Patel, S. V. (2016). A Novel Approach to Filter Topic Related Important Terms Within Document. Dalam S. C. Satapathy, A. Joshi, N. Modi, & N. Pathak (Ed.), *Proceedings of International Conference on ICT for Sustainable Development* (hlm. 697–704). Springer. https://doi.org/10.1007/978-981-10-0135-2_67
- Kafatan, S., Riyanto, D. E., & Saputra, R. (2014). Sistem Informasi Pengelolaan Arsip Statis Pada Badan Arsip Dan Perpustakaan Provinsi Jawa Tengah Menggunakan Vector Space Model. *Jurnal Masyarakat Informatika*, 5(9), 45–52. <https://doi.org/10.14710/jmasif.5.9.8438>
- Kalra, V., & Agrawal, R. (2018). *Importance of Text Data Preprocessing & Implementation in RapidMiner* (hlm. 75). <https://doi.org/10.15439/2017KM46>
- Kaufman, L., & Rousseeuw, P. (1990). Finding Groups in Data: An Introduction To Cluster Analysis. Dalam *Wiley, New York. ISBN 0-471-87876-6*. <https://doi.org/10.2307/2532178>
- Kirişci, M. (2023). New cosine similarity and distance measures for Fermatean fuzzy sets and TOPSIS approach. *Knowledge and Information Systems*, 65(2), 855–868. <https://doi.org/10.1007/s10115-022-01776-4>
- Kurniawan, S., Gata, W., Puspitawati, D. A., Parthama, I. K. S., Setiawan, H., & Hartini, S. (2020). Text Mining Pre-Processing Using Gata Framework and RapidMiner for Indonesian Sentiment Analysis. *IOP Conference Series: Materials Science and Engineering*, 835(1), 012057. <https://doi.org/10.1088/1757-899X/835/1/012057>
- Latif, U. (2015). Konsep Amarah Menurut Al-Qur'an. *Jurnal Al-Bayan: Media Kajian Dan Pengembangan Ilmu Dakwah*, 21(2), Article 2. <https://doi.org/10.22373/albayan.v21i32.422>
- LeDoux, J. E. (2000). *Emotion Circuits In The Brain*.
- Liang, H., Sun, X., Sun, Y., & Gao, Y. (2017). Text feature extraction based on deep learning: A review. *EURASIP Journal on Wireless Communications and Networking*, 2017(1), 211. <https://doi.org/10.1186/s13638-017-0993-1>

- Liu, H. M., & Lu, J. G. (2015). Brief Survey of K-Means Clustering Algorithms. *Applied Mechanics and Materials*, 740, 624–628. <https://doi.org/10.4028/www.scientific.net/AMM.740.624>
- Mardan, M. (2010). *Al-Qur'an Sebuah Pengantar*. Pustaka Mapan Jakarta. <https://repositori.uin-alauddin.ac.id/1661/>
- Martin, R. E., & Ochsner, K. N. (2016). The Neuroscience of Emotion Regulation Development: Implications for Education. *Current opinion in behavioral sciences*, 10, 142–148. <https://doi.org/10.1016/j.cobeha.2016.06.006>
- Matsumoto, D., Frank, M. G., & Hwang, H. C. (2015). The Role of Intergroup Emotions in Political Violence. *Current Directions in Psychological Science*, 24(5), 369–373. <https://doi.org/10.1177/0963721415595023>
- Messakh, G. C., Hayati, M. N., & Sifriyani, S. (2023). Comparison K-Means and Fuzzy C-Means In Regencies/Cities Grouping Based on Educational Indicators. *Jurnal Varian*, 7(1), Article 1. <https://doi.org/10.30812/varian.v7i1.2879>
- Miner, G., IV, J. E., Hill, T., Nisbet, R., Delen, D., & Fast, A. (2012). *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications* (1st edition). Academic Press.
- Mustofa, A. (2018). Pemikiran Harun Yahya Dalam Nilai Nilai Moral Al-Qur'an (Studi Analisis Nilai Nilai Pendidikan Akhlaq). *Jurnal Pendidikan Islam*, 4(1), Article 1. <https://doi.org/10.37286/ojs.v4i1.29>
- Nasr, S. H. (with Internet Archive). (2002). *The heart of Islam: Enduring values for humanity*. [San Francisco, Calif.]: HarperSanFrancisco. <http://archive.org/details/heartofislamendu0000nasr>
- Nisar, F., & Saeed, A. (2020). “Happiness” as an Idea Presented by Media and the Criteria given by Islam (An Overview under the light of Quran and Hadith). *Al-Wifaq*, 3(1), Article 1.
- Nofiyani, N., & Wulandari, W. (2022). Implementasi Electronic Data Processing Untuk meningkatkan Efektifitas dan Efisiensi Pada Text Mining. *Jurnal Media Informatika Budidarma*, 6(3), Article 3. <https://doi.org/10.30865/mib.v6i3.4332>

- Nur Taufiq, H. (2014). Pembelajaran Al-Qur'an Melalui Pendekatan Metode Tafsir Tematik Bagi Guru Tpq Di Kota Malang. *Progresiva : Jurnal Pemikiran dan Pendidikan Islam*, 5(1). <https://doi.org/10.22219/progresiva.v5i1.2058>
- Nurdiana, O., Jumadi, J., & Nursantika, D. (2016). Perbandingan Metode Cosine Similarity Dengan Metode Jaccard Similarity Pada Aplikasi Pencarian Terjemah Al-Qur'an dalam Bahasa Indonesia. *Jurnal Online Informatika*, 1(1), Article 1. <https://doi.org/10.15575/join.v1i1.12>
- Nurrohm, A., & Sidik, I. N. (2020). Hikmah Dalam Al-Qur'an: Studi Tematik Terhadap Tafsir Al-Mizān. *Profetika: Jurnal Studi Islam*, 20(2), Article 2. <https://doi.org/10.23917/profetika.v20i2.9954>
- Oppi Anda Resta, Aditya, A., & Febry Eka Purwiantono. (2021). Plagiarism Detection in Students' Theses Using The Cosine Similarity Method. *Sinkron: jurnal dan penelitian teknik informatika*, 5(2), 305–313. <https://doi.org/10.33395/sinkron.v5i2.10909>
- Otten, N. V. (2023, September 7). *Vector Space Model Made Simple With Examples & Tutorial In Python*. Spot Intelligence. <https://spotintelligence.com/2023/09/07/vector-space-model/>
- Oyelade, O. J., Oladipupo, O. O., & Obagbuwa, I. C. (2010). *Application of k Means Clustering algorithm for prediction of Students Academic Performance* (arXiv:1002.2425). arXiv. <https://doi.org/10.48550/arXiv.1002.2425>
- Palacio-Niño, J.-O., & Berzal, F. (2019). *Evaluation Metrics for Unsupervised Learning Algorithms* (arXiv:1905.05667). arXiv. <https://doi.org/10.48550/arXiv.1905.05667>
- Phalkey, M. (2023, Mei 12). *What is Clustering in Machine Learning? (With Examples)*. AlmaBetter. <https://www.almabetter.com/bytes/articles/what-is-clustering-in-machine-learning>
- Pham, D. T., Dimov, S. S., & Nguyen, C. D. (2005). Selection of K in K-means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 219(1), 103–119. <https://doi.org/10.1243/095440605X8298>

- Pradana, M. G. (2020). Penggunaan Fitur Wordcloud Dan Document Term Matrix Dalam Text Mining. *Jurnal Ilmiah Informatika*, 8(01), Article 01. <https://doi.org/10.33884/jif.v8i01.1838>
- Putra, D. I. A., & Hidayaturrahman, M. (2020). The roles of technology in al-Quran exegesis in Indonesia. *Technology in Society*, 63, 101418. <https://doi.org/10.1016/j.techsoc.2020.101418>
- Qibtiyah, M. (2023). *Emosi Dalam Perspektif Al Qur'an* [Skripsi, Fakultas Ushuluddin Dan Pemikiran Islam]. <https://repository.ptiq.ac.id/id/eprint/1393/>
- Ramadhani, S., Azzahra, D., & Z, T. (2022). Comparison of K-Means and K-Medoids Algorithms in Text Mining based on Davies Bouldin Index Testing for Classification of Student's Thesis. *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, 13(1), 24–33. <https://doi.org/10.31849/digitalzone.v13i1.9292>
- Reddy, C. K., & Vinzamuri, B. (2018). A Survey of Partitional and Hierarchical Clustering Algorithms. Dalam C. C. Aggarwal & C. K. Reddy (Ed.), *Data Clustering* (1 ed., hlm. 87–110). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315373515-4>
- Resyanto, F., Sibaroni, Y., & Romadhony, A. (2019). Choosing The Most Optimum Text Preprocessing Method for Sentiment Analysis: Case:iPhone Tweets. *2019 Fourth International Conference on Informatics and Computing (ICIC)*, 1–5. <https://doi.org/10.1109/ICIC47613.2019.8985943>
- Rofiqi, M. A., Fauzan, A. C., Agustin, A. P., & Saputra, A. A. (2019). Implementasi Term-Frequency Inverse Document Frequency (TF-IDF) Untuk Mencari Relevansi Dokumen Berdasarkan Query. *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, 1(2), Article 2. <https://doi.org/10.28926/ilkomnika.v1i2.18>
- Rosid, M. A., Fitriani, A. S., Astutik, I. R. I., Mulloh, N. I., & Gozali, H. A. (2020). Improving Text Preprocessing For Student Complaint Document Classification Using Sastrawi. *IOP Conference Series: Materials Science and Engineering*, 874(1), 012017. <https://doi.org/10.1088/1757-899X/874/1/012017>

- Saida, Y. E. (2007). *Pengelompokan terjemahan ayat-ayat Al-Qur'an dalam Bahasa Indonesia dengan Algoritma K-Means Clustering* [Sarjana, Universitas Brawijaya]. <https://repository.ub.ac.id/id/eprint/151725/>
- Sakthi Vel, S. (2021). Pre-Processing techniques of Text Mining using Computational Linguistics and Python Libraries. *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, 879–884. <https://doi.org/10.1109/ICAIS50930.2021.9395924>
- Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Commun. ACM*, 18(11), 613–620. <https://doi.org/10.1145/361219.361220>
- Setiawan, F. A., Sadikin, M., & Kaburuan, E. R. (2022). Analisis Permasalahan Perangkat Pencetak Menggunakan Metode Algoritma K-Means dan K-Medoids. *Teknika*, 11(2), Article 2. <https://doi.org/10.34148/teknika.v11i2.471>
- Suparlan, S. (2008). PSIKOLOGI DAN KEPERIBADIAN PERSPEKTIF AL-QURAN. *Humanika: Kajian Ilmiah Mata Kuliah Umum*, 8(1), Article 1. <https://doi.org/10.21831/hum.v8i1.21005>
- Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering*, 336(1), 012017. <https://doi.org/10.1088/1757-899X/336/1/012017>
- Syukron, H., Fayyad, M. F., Fauzan, F. J., Ikhsani, Y., & Gurning, U. R. (2022). Perbandingan K-Means K-Medoids dan Fuzzy C-Means untuk Pengelompokan Data Pelanggan dengan Model LRFM: Comparison K-Means K-Medoids and Fuzzy C-Means for Clustering Customer Data with LRFM Model. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 2(2), Article 2. <https://doi.org/10.57152/malcom.v2i2.442>
- Tabassum, A., & Patil, D. R. R. (2020). *A Survey on Text Pre-Processing & Feature Extraction Techniques in Natural Language Processing*. <https://api.semanticscholar.org/CorpusID:235211496>

- Talebinamvar, M., & Zarrabi, F. (2022). Clustering students' writing behaviors using keystroke logging: A learning analytic approach in EFL writing. *Language Testing in Asia*, 12(1), 6. <https://doi.org/10.1186/s40468-021-00150-5>
- Thaib, H. Z. bin H. (2016). Tadarus Alquran: Urgensi, Tahapan, dan Penerapannya. *Almufida: Jurnal Ilmu-Ilmu Keislaman*, 1(1), Article 1. <https://doi.org/10.46576/almufida.v1i1.103>
- Umargono, E., Suseno, J. E., & Gunawan, S. K. V. (2020). *K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula*. 121–129. <https://doi.org/10.2991/assehr.k.201010.019>
- vom Brocke, J., & Maedche, A. (2019). The DSR grid: Six core dimensions for effectively planning and communicating design science research projects. *Electronic Markets*, 29(3), 379–385. <https://doi.org/10.1007/s12525-019-00358-7>
- Wegmann, M., Zipperling, D., Hillenbrand, J., & Fleischer, J. (2021). *A review of systematic selection of clustering algorithms and their evaluation* (arXiv:2106.12792). arXiv. <https://doi.org/10.48550/arXiv.2106.12792>
- Wishnu Hardi, A. (2019). *Pengelompokan topik dokumen berbasis text mining dengan algoritme k-means: Studi kasus pada dokumen kedutaan besar Australia Jakarta*. Universitas Indonesia Library; Pusat Jasa Perpustakaan dan Informasi. <https://lib.ui.ac.id>
- Yahya, M. D. (2015). *Nilai-Nilai Pendidikan Dalam Alquran*. <https://idr.uin-antasari.ac.id/5246/>
- Ys, I. A. F., & Rohmanudin, D. (2021). Sebab-sebab Historis Lahirnya Metode Maudu'i dalam Tafsir. *Jurnal Iman Dan Spiritualitas*, 1(3), Article 3. <https://doi.org/10.15575/jis.v1i3.12944>
- Zahid, M. (2012). Perbedaan Pendapat Para Ulama Tentang Jumlah Ayat Al-Qur'an Dan Implikasinya Terhadap Penerbitan Mushaf Al-Qur'an Di Indonesia. *Nuansa: Jurnal Penelitian Ilmu Sosial Dan Keagamaan Islam*, 9(1), Article 1. <https://doi.org/10.19105/nuansa.v9i1.20>