

## Klasterisasi Tukang Gigi di Jakarta dengan Algoritma K-Means berdasarkan Daerah Asal

Iskandar Zulkarnain<sup>1</sup>, Fauziyah<sup>2</sup>

<sup>1,2</sup>Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Bung Karno, Jakarta

Email: <sup>1</sup>iskandarzulkarnain@ubk.ac.id, <sup>2</sup>fauziyah@ubk.ac.id

### Abstrak

*Tukang gigi yang terdaftar pada organisasi Serikat Tukang Gigi Indonesia (STGI) DPW DKI Jakarta sebanyak 625 anggota. STGI adalah organisasi yang telah Berbadan Hukum KEPMENKUMHAM NOMOR AHU-000083.AH.01.08.Tahun 2020. Dalam penyebarannya tukang gigi yang berada di Jakarta terdiri dari berbagai daerah asal, maka peneliti melakukan penelitian klasterisasi untuk mengetahui penyebaran daerah asal tukang gigi di wilayah Jakarta. Cakupan wilayah data penelitian yaitu: Jakarta Utara, Jakarta Barat, Jakarta Timur, Jakarta Selatan dan Jakarta Pusat. Hasil dari klasterisasi Algoritma K-Means menggunakan aplikasi RapidMiner sebagai berikut: Cluster 0 (C0) : 1 Anggota, Cluster 1 (C1) : 16 Anggota, Cluster 2 (C2) : 1 Anggota. Hasil yang termasuk pada C0 yakni Kabupaten/Kota asal daerah tukang gigi paling banyak yaitu: Pamekasan. Hasil yang termasuk pada C1 merupakan Kabupaten/Kota daerah asal tukang gigi paling beragam terdiri dari 16 Kabupaten/Kota termasuk daerah asal tukang gigi terbanyak ke-3 yaitu: Jakarta. Sedangkan yang termasuk C2 merupakan daerah asal tukang gigi terbanyak ke-2 yaitu: Jember.*

**Kata kunci:** Algoritma K-Means, Clustering, Data Mining, RapidMiner, Tukang Gigi.

### Abstract

*Tukang gigi registered with the Indonesian tukang gigi Union (STGI) DPW DKI Jakarta organization are 625 members. STGI is an organization that has been incorporated by the Ministry of Law and Human Rights number AHU-000083. AH.01.08.Year 2020. In the distribution of tukang gigi in Jakarta consisting of various regions of origin, the researcher conducted clustering research to find out the distribution of the origin of Tukang Gigi in the Jakarta area. The scope of the research data area is: North Jakarta, West Jakarta, East Jakarta, South Jakarta and Central Jakarta. The results of the clustering of the K-Means Algorithm using the RapidMiner application are as follows: Cluster 0 (C0) : 1 Member, Cluster 1 (C1) : 16 Members, Cluster 2 (C2) : 1 Member. The results included in C0 are the Regency/City of origin of the most tukang gigi, namely: Pamekasan. The results included in C1 are the most diverse Regencies/Cities of the origin of tukang gigi, consisting of 16 Regencies/Cities, including the 3rd most Tukang Gigi' origin areas, namely: Jakarta. While those included in C2 are the 2nd most tukang gigi from origin, namely: Jember.*

**Keywords:** K-Means Algorithm, Clustering, Data Mining, RapidMiner, Tukang Gigi



## PENDAHULUAN

Manusia sering mengabaikan pentingnya kesehatan gigi dan mulut, karena kesihatannya gigi dan mulut tidak menjadi perhatian yang utama dalam kesehatan. Seorang profesional dokter gigi yang memiliki keahlian untuk mencegah dan mengobati masalah kesehatan gigi dan mulut. Ditengah maraknya praktik dokter gigi banyak masyarakat kelas menengah kebawah kesulitan dalam hal biaya pencegahan dan pengobatan ke dokter gigi, yang mengarah pada kemampuan sosial ekonomi masyarakat.

Tukang gigi sudah hadir sejak zaman penjajahan Belanda dan pada masa tersebut mereka disebut dukun gigi dan dianggap sebagai pahlawan oleh masyarakat menengah kebawah karena pada zaman tersebut masih langkanya profesi dokter gigi dan hanya siap untuk merawat orang Eropa dengan biaya yang mahal (Saropah dkk., 2023).

Tukang gigi merupakan jasa kesehatan tradisional yang didapatkan secara turun temurun, keberadaan tukang gigi di Indonesia semakin pesat dan mudah ditemukan hampir diseluruh wilayah Indonesia salah satunya di wilayah Jakarta (Prawitanto, 2023).

Menurut PERMENKES 39 Tahun 2014 tentang Pembinaan Pengawasan dan Perizinan Pekerjaan Tukang Gigi pada Pasal 1 menyatakan tukang gigi merupakan setiap orang yang memiliki kemampuan dalam membuat dan memasang gigi tiruan lepasan serta memiliki izin tertulis yang diakui oleh Pemerintah Daerah dalam mengerjakan pekerjaannya (Kementerian Kesehatan Republik Indonesia, 2014).

Jumlah tukang gigi di daerah Jakarta yang saat ini terdaftar kurang lebih berjumlah 625 anggota. Angka tersebut didapatkan dari jumlah anggota yang bersertifikat dan terdaftar (memiliki kartu anggota aktif) dari Serikat Tukang Gigi Indonesia (STGI) DPW DKI Jakarta yang telah Berbadan Hukum KEPMENKUMHAM NOMOR AHU-000083.AH.01.08.Tahun 2020. Klasterisasi dilakukan untuk mengetahui penyebaran daerah asal tukang gigi di wilayah Jakarta. Cakupan wilayah penelitian yaitu: Jakarta Utara, Jakarta Barat, Jakarta Timur, Jakarta Selatan dan Jakarta Pusat.

Belum adanya penelitian tentang penyebaran tukang gigi di Jakarta menjadikan alasan Peneliti melakukan penelitian klasterisasi tukang gigi berdasarkan asal dan penyebarannya di Jakarta. Peneliti menggunakan Algoritma K-

Means yang telah tersedia pada aplikasi RapidMiner yang dapat membantu pengelompokan non-hierarki yang bersifat *unsupervised* untuk menentukan kategori klaster kelompok penyebaran asal daerah tukang gigi di wilayah Jakarta (Pratama dkk., 2022).

## TINJAUAN PUSTAKA

### 2.1 Penelitian Terdahulu

Penelitian terkait klasterisasi dengan Algoritma K-Means pernah dilakukan oleh Novita, dkk. Fokus masalah utama yang dibahas dalam penelitian ini adalah perlunya pengelompokan provinsi di Indonesia yang efektif berdasarkan metrik pertanian seperti luas panen, produksi, dan produktivitas komoditas pangan, khususnya dalam konteks meningkatnya permintaan pangan akibat pertumbuhan penduduk. Penelitian ini menggunakan algoritma pengelompokan K-Means, menggunakan metode siku untuk menentukan jumlah cluster yang optimal. Data dari Kementerian Pertanian Indonesia yang mencakup tahun 2017 hingga 2019 dianalisis untuk mengkategorikan provinsi berdasarkan kinerja pertaniannya. Hasil penelitian mengungkapkan tiga cluster yang berbeda: Cluster 0 (potensial sedang) dengan 29 provinsi, Cluster 1 (potensial tinggi) dengan 9 provinsi, dan Cluster 2 (potensial rendah) dengan 64 provinsi. Pengelompokan optimal ditentukan pada  $k=3$ , dengan pengurangan signifikan dalam Sum of Square Error (SSE) yang menunjukkan efektivitas pendekatan pengelompokan (Novita dkk., 2022).

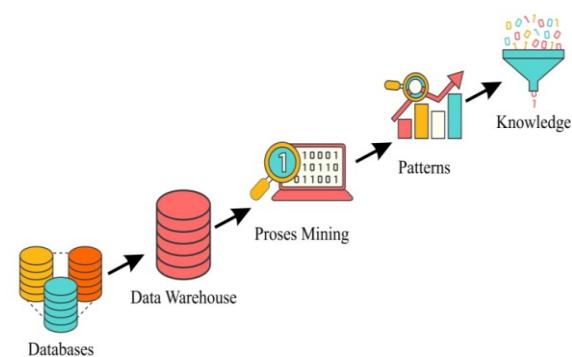
Sarbaini dkk melakukan penelitian dengan pembahasannya adalah pengelompokan kabupaten di Provinsi Riau berdasarkan tingkat pengangguran, menyoroti perlunya kategorisasi yang efektif untuk menginformasikan kebijakan dan intervensi pemerintah untuk meningkatkan kesejahteraan dan penciptaan lapangan kerja. Penelitian ini menggunakan metode pengelompokan Fuzzy K-Means, memanfaatkan data sekunder dari Badan Pusat Statistik (BPS) Provinsi Riau. Indikator utama yang dianalisis meliputi Tingkat Pengangguran Terbuka (TPT), Tingkat Partisipasi Angkatan Kerja (TPAK), Produk Domestik Regional Bruto (PDRB), dan Indeks Pembangunan Manusia (IPM) dari 2017 hingga 2021. Hasil penelitian menunjukkan bahwa kabupaten berhasil dikelompokkan menjadi dua kelompok: Cluster 0 (pengangguran tinggi) dan Cluster 1 (pengangguran rendah),

dengan enam kabupaten di setiap cluster. Klasifikasi ini memberikan wawasan bagi pemerintah untuk fokus pada daerah dengan pengangguran tinggi untuk intervensi yang ditargetkan (Sarbaini dkk., 2022).

Penelitian lain yang dilakukan oleh Dewi dkk, berfokus pada kebutuhan untuk mengelompokkan siswa secara efektif berdasarkan kinerja akademik mereka, seperti yang ditunjukkan oleh nilai kartu laporan mereka dari Semester I hingga Semester IV. Pengelompokan ini bertujuan untuk memfasilitasi pembentukan kelompok belajar yang efektif untuk meningkatkan hasil pembelajaran siswa. Penelitian ini menggunakan algoritma pengelompokan K-Means, yang merupakan metode partisi data non-hierarkis. Studi ini menggunakan data yang dikumpulkan dari kartu laporan dan wawancara dengan pejabat sekolah untuk mengkategorikan siswa menjadi tiga kelompok yang berbeda: Sangat Berprestasi, Berprestasi Tinggi, dan Berprestasi Rendah. Hasil penelitian mengungkapkan tiga kelompok siswa berdasarkan kinerja mereka: Cluster 1 (Very High Achievers), Cluster 2 (High Achievers), dan Cluster 3 (Low Achievers). Klasifikasi ini memberikan kerangka kerja bagi guru untuk membuat keputusan berdasarkan informasi mengenai pembentukan kelompok belajar, memastikan bahwa siswa dikelompokkan dengan cara yang mempromosikan pembelajaran yang efektif (Dewi dkk., 2021).

## 2.2 Data Mining

Merupakan kegiatan dalam menganalisis pengetahuan pada basis data yang biasa disebut *Knowledge Discovery in Database* (KDD). Secara sistematis langkah untuk melakukan data mining secara umum terdiri dari 3 (tiga) tahapan, yaitu: Eksplorasi, Membuat Model dan Validasi dan Penerapan. Sedangkan tahapan data mining sesuai dengan *Knowledge Discovery in Database* (KDD) (Muslim dkk., 2019), dapat dilihat pada gambar 1.



**Gambar 1.** Proses Data Mining sesuai *Knowledge Discovery in Database* (KDD)

## 2.3 Clustering/Klasterisasi

Clustering/Klasterisasi adalah teknik pengolahan banyak data pada *database* untuk dikelompokkan berdasarkan pada kriteria yang telah ditentukan. Clustering termasuk kedalam metode *unsupervised learning* yakni dapat memberikan label pada kelas data yang belum diketahui. Secara umum *clustering* adalah mengelompokkan objek pada *cluster* yang mempunyai kesamaan, dimana semakin tinggi kesamaan dari objek maka semakin akurat hasil dari *cluster* tersebut (Apriyani dkk., 2023).

## 2.4 Algoritma K-Means

K-Means merupakan salah satu metode *clustering* non-hierarki berbasis jarak yang dibagi kedalam sejumlah *cluster*. Pusat klaster dipilih secara acak dari kumpulan populasi data (Wahyudi dkk., 2020).

Langkah-langkah dalam melakukan klasterisasi dengan Algoritma K-Means (Harahap & Zulvia, 2021), sebagai berikut:

1. Menentukan jumlah *cluster* k.
2. Lanjutkan dengan inisialisasikan *cluster* k ke pusat *cluster*.

Dalam penginisialisasikan dilakukan secara acak (*random*), pusat *cluster* diberi nilai awal dengan angka acak.

$$V = \frac{\sum_{i=1}^n x_i}{n}$$

Keterangan:

*V* : *Centroid* pada *cluster*

*x<sub>i</sub>* : Objek ke-*i* ; dimana *i* = 1,2,3, ..., n

*n* : Banyaknya objek/jumlah objek yang menjadi anggota *cluster*

3. Menentukan jarak Euclidean yang ditentukan menggunakan rumus berikut.

$$D_{(x,y)} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

Keterangan:

x : Objek x ke-i

y : Daya y ke-i

n : banyaknya objek

4. Mengalokasikan masing-masing dari objek ke *centroid* baru.

## METODE

Dalam melaksanakan penelitian, peneliti menggunakan langkah-langkah pada gambar 2 yang dapat membantu klasterisasi asal daerah tukang gigi di Jakarta.

a. Identifikasi Masalah

Merupakan kegiatan awal yaitu menemukan permasalahan penelitian dari objek penelitian. Tujuan dari tahap ini adalah agar peneliti dapat memahami ruang lingkup masalah dan hal-hal yang dilakukan sebagai pemecahan masalah (Faria dkk., 2023).

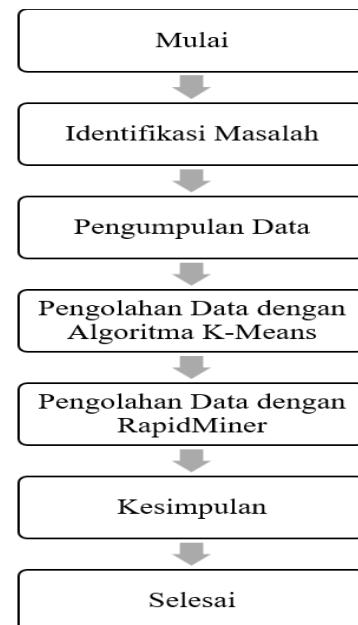
b. Pengumpulan Data

Pada pengumpulan data, peneliti melakukan 3 (tiga) tahapan, yaitu:

a) Wawancara: Peneliti melakukan wawancara dengan Ketua Serikat Tukang Gigi Indonesia (STGI) DPW DKI JAKARTA yaitu Bapak H. Muhammad Zubaidi.

b) Observasi: Kegiatan mengamati objek penelitian (Firmansyah dkk., 2021). Peneliti berkunjung ke kantor Serikat Tukang Gigi Indonesia (STGI) DPW DKI Jakarta yang beralamat di Jl. Ulujami Raya RT.09/RW.01, Kel. Pesanggrahan, Kec. Pesanggrahan, Jakarta Selatan.

c) Studi Literatur: Merupakan salah satu teknik pengumpulan data yang itu dengan menganalisis data yang telah dikumpulkan serta menganalisis artikel terkait dengan topik penelitian (Fitriyani dkk., 2021).



Gambar 2. Tahapan Penelitian

## HASIL DAN PEMBAHASAN

### 4.1 Sumber Data

Data yang digunakan pada penelitian ini merupakan hasil dari pengumpulan data tukang gigi wilayah Jakarta yang terdaftar dan aktif pada Serikat Tukang Gigi Indonesia (STGI) per Januari 2024, yang dapat dilihat pada tabel 1 berikut ini.

Tabel 1. Sumber Penelitian

Kab/Kota	JU	JT	JB	JP	JS
Pamekasan	92	136	103	37	84
Jember	12	21	50	2	24
Surabaya	1	2	1	0	0
Mekkah	0	2	2	2	0
Sampang	0	2	2	2	3
Lamongan	0	1	0	0	0
Jakarta	5	3	9	5	0
Pasuruan	0	0	0	0	3
Tapanuli	0	0	0	1	0
Sumenep	2	0	4	1	1
Nganjuk	0	0	1	0	0
Banyuwangi	0	0	2	0	0
Tangerang	0	0	2	0	0
Pacitan	0	0	1	0	0

<b>Malang</b>	0	0	1	0	0
<b>Sumedang</b>	0	0	1	0	0
<b>Sukabumi</b>	0	0	0	0	1
<b>Lumajang</b>	0	0	0	0	1

Keterangan:

- a. JU = Jakarta Utara
- b. JT = Jakarta Timur
- c. JB = Jakarta Barat
- d. JP = Jakarta Pusat
- e. JS = Jakarta Selatan

#### 4.2 *Centroid* Awal

Penerapan Algoritma K-Means dalam penelitian diawali dengan menentukan nilai *centroid* awal dari sumber data. Nilai *centroid* awal diambil secara acak (Wijayanto & Fathoni, 2021). Nilai *centroid* awal ditampilkan pada tabel 2 berikut.

**Tabel 2.** Nilai *Centroid* Awal

Kab/Kota	JU	JT	JB	JP	JS
<b>Pamekasan</b>	92	136	103	37	84
<b>Jember</b>	12	21	50	2	24
<b>Jakarta</b>	5	3	9	5	0

#### 4.3 Menentukan Jarak Euclidean

Setelah menentukan nilai *centroid* awal dilanjutkan dengan menghitung jarak antara *centroid* dengan setiap data. Setiap data penelitian dihitung jarak dengan *centroid* sehingga didapatkan hasil jarak euclidean yang ditunjukkan pada tabel 3.

**Tabel 3.** Nilai Jarak Euclidean

C0	C1	C2
92	0.5	12
136	0.625	21
103	1.625	50
37	0.6875	2
84	0.5625	24

#### 4.4 *Clustering* Data

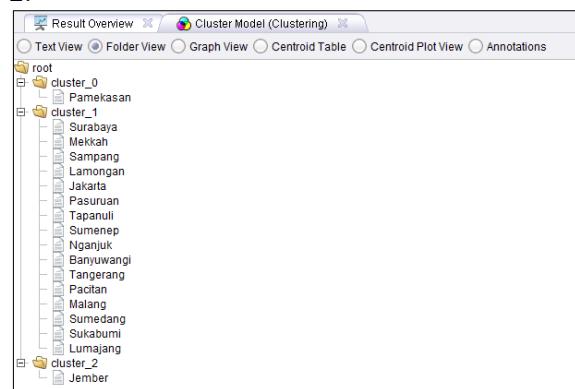
*Clustering* data ialah menentukan letak *cluster* dari setiap data yang ditentukan dengan

memilih nilai yang paling kecil diantara C0, C1, C2 (Sirait dkk., 2019). Hasil dari klasterisasi menggunakan Algoritma K-Means sebagai berikut:

- a. *Cluster* 0 (C0) : 1 Anggota
- b. *Cluster* 1 (C1) : 16 Anggota
- c. *Cluster* 2 (C2) : 1 Anggota

#### 4.5 Hasil Perhitungan RapidMiner

Perhitungan *cluster* K-Mean dari sebaran daerah asal tukang gigi di Jakarta sebanyak 18 data menggunakan aplikasi RapidMiner 5.3.000. Diperoleh hasil yang dapat dilihat pada gambar 3 yang menampilkan kabupaten/kota yang termasuk dalam *Cluster* 0, *Cluster* 1 dan *Cluster* 2.



**Gambar 3.** Hasil *Clustering* K-Means Menggunakan RapidMiner

#### SIMPULAN

Penelitian yang telah dilakukan terkait dengan klasterisasi daerah asal tukang gigi di Jakarta menggunakan Algoritma K-Means dan berhenti pada iterasi ke-5. Hasil yang termasuk pada C0 yakni Kabupaten/Kota asal daerah tukang gigi paling banyak yaitu: Pamekasan. Hasil yang termasuk pada C1 merupakan Kabupaten/Kota daerah asal tukang gigi paling beragam terdiri dari 16 Kabupaten/Kota termasuk daerah asal tukang gigi terbanyak ke-3 yaitu: Jakarta. Sedangkan yang termasuk C2 merupakan daerah asal tukang gigi terbanyak ke-2 yaitu: Jember.

Sebagai saran untuk penelitian selanjutnya dapat menambahkan variabel tambahan seperti pengalaman bertahun-tahun, spesialisasi atau demografi pasien sehingga dapat memberikan faktor-faktor yang

mempengaruhi distribusi dan pengelompokan tukang gigi.

**UCAPAN TERIMA KASIH**

Peneliti mengucapkan terima kasih kepada yang telah berkontribusi dalam membantu penelitian ini. Terima kasih kepada Bapak H. Muhammad Zubaidi yang telah menjembatani dalam pengumpulan data penelitian ini dan juga terima kasih kepada LPPM Universitas Bung Karno serta Jurnal INFORMATIKA dan Rekayasa Perangkat Lunak yang telah mempublish artikel penelitian ini.

**DAFTAR PUSTAKA**

Apriyani, P., Dikananda, A. R., & Ali, I. (2023). Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi. *Hello World Jurnal Ilmu Komputer*, 2(1), 20–33. <https://doi.org/10.56211/helloworld.v2i1.230>

Dewi, S., Defit, S., & Yunus, Y. (2021). Akurasi Pemetaan Kelompok Belajar Siswa Menuju Prestasi Menggunakan Metode K-Means. *Jurnal Sistem Informasi dan Teknologi*, 3(1), 28–33. <https://doi.org/10.37034/jsisfotek.v3i1.40>

Faria, N., Noviasri, R., Perwitasari, W., Pangestu, F. H., Tyasa, T., Majid, A., Ilham, M. A., & Pratama, M. J. (2023). Sosialisasi Pemilahan Sampah dan Pelatihan Pengolahan Sampah Organik dan Anorganik untuk Mewujudkan Desa Wisata Zero Waste Pulau Gili. *Community Engagement & Emergence Journal*, 4(3), 2023.

Firmansyah, Y., Maulana, R., & Maulana, M. S. (2021). Implementasi Metode SDLC Prototype pada Sistem Informasi Indeks Kepuasan Masyarakat (IKM) Berbasis Website Studi Kasus Dinas Kependudukan dan Catatan Sipil. *Jurnal Sistem dan Teknologi Informasi (Justin)*, 9(3), 315. <https://doi.org/10.26418/justin.v9i3.46964>

Fitriyani, Sari, R., & Faiz, S. (2021). SISTEM INFORMASI PENGELOLAAN ZAKAT DAN INFAQ MENGGUNAKAN METODE RAD. *SISTEM INFORMASI PENGELOLAAN ZAKAT DAN INFAQ MENGGUNAKAN METODE RAD*, 2(2), 197–210.

Harahap, A. S., & Zulvia, P. (2021). Klasterisasi Desa dengan Menggunakan Algoritma K-Means pada Data Potensi Desa. *JURIKOM (Jurnal Riset Komputer)*, 8(6), 237. <https://doi.org/10.30865/jurikom.v8i6.3724>

Kementerian Kesehatan Republik Indonesia. (2014). *Permenkes Nomor 39 Tahun 2014*.

Muslim, M. A., Prasetyo, B., Mawarni, E. L. H., Herowati, A. J., Mirqotussa'adah, M., Rukmana, S. H., & Nurzahputra, A. (2019). *Data Mining Algoritma C4.5 Disertai Contoh Kasus dan Penerapannya dengan Program Komputer* (E. Listiana & N. Cahyani, Ed.; 1 ed.).

Novita, A., Ernawati, Ii., & Chamidah, N. (2022). Klasterisasi Provinsi Di Indonesia Berdasarkan Produktivitas Komoditas Pangan Menggunakan Algoritma K-Means. *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, 462–471.

Pratama, Y., Pramudya, Y., Albert, E., S, M. I., Ferdinand, R., Juniansyah, V., & Rasywir, E. (2022). Klasterisasi Data Pertanian di Tingkat Provinsi Jambi Tahun 2021 Menggunakan Algoritma K-Means. *Bulletin Of Computer Science Research*, 3(1), 57–63. <https://doi.org/10.47065/bulletincsr.v3i1.205>

Prawitanto, N. (2023). Praktek Tukang Gigi Ditinjau dari Permenkes Nomor 39 Tahun 2014 tentang Pembinaan, Pengawasan dan Perizinan Pekerjaan Tukang Gigi (Studi di Tempat Praktek Tukang Gigi Kec. Banjarsari, Kota Surakarta). *Jurnal Bevinding*, 1(1), 94–101.

Sarbaini, S., Saputri, W., Nazaruddin, N., & Muttakin, F. (2022). Cluster Analysis Menggunakan Algoritma Fuzzy K Means Untuk Tingkat Pengangguran Di Provinsi Riau. *Jurnal Teknologi dan Manajemen Industri Terapan (JTMIT)*, 1(2), 78–84.

Saropah, S., Istijab, I., & Sukron, A. (2023). Tanggung Jawab Pekerjaan Tukang Gigi Ditinjau Dari Peraturan Menteri Kesehatan Republik Indonesia Nomor 39 Tahun 2014 Tentang Pembinaan, Pengawasan Dan Perizinan, Pekerjaan Tukang Gigi. *YURIJAYA-Jurnal Ilmiah Hukum*, 5(1), 53–58. <http://repository.unissula.ac.id/9453/4/BA%20I.pdf>

Sirait, W., Defit, S., & Nurcahyo, G. W. (2019). Algoritma K-Means Untuk Klasterisasi Tugas Akhir Mahasiswa Berdasarkan Keahlian. *Jurnal Sistim Informasi dan Teknologi*, 1(3), 25–30. <https://doi.org/10.35134/jsisfotek.v1i3.6>

Wahyudi, M., Msitha, M., Saragih, R., & Solikhun, S. (2020). *Data Mining: Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering* (J. Simarmata, Ed.; 1 ed.). Yayasan Kita Menulis.

Wijayanto, S., & Fathoni, M. Y. (2021). Pengelompokan Produktivitas Tanaman Padi di Jawa Tengah Menggunakan Metode Clustering K-Means. *Jurnal JUPITER*, 13(2), 212–219.