

Penerapan Data Mining Dalam Mengelompokkan Jumlah Usaha Berdasarkan Provinsi Menggunakan K-Means Clustering

Rahel Adelina Hutasoit¹, M.Safii², Iin Parlina²

¹STIKOM Tunas Bangsa Pematangsiantar, Indonesia

²AMIK Tunas Bangsa Pematangsiantar, Indonesia

^{1,2}Jl. Jendral Sudirman Blok A No. 1, 2, 3 Pematangsiantar, Indonesia
rachel.adelina99@gmail.com

Abstract-Industrial revolution 4.0 is an era where technology in the IT field is growing very rapidly, it can also be said a new era for entrepreneurs in Indonesia that must be a golden opportunity to improve business performance and opportunities for millennials to enter the business or business world. In this era people are expected to be able to compete especially in the business world. This study discusses the Application of Data Mining in Grouping the Number of Enterprises by Province Using K-Means Clustering. The source of this research data is collected based on the information documents of the number of businesses in Indonesia produced by the National Statistics Agency. The data used in this study are provincial data consisting of 34 provinces. Data will be processed by clustering in 3 clusters, namely cluster of high number of businesses, cluster of medium number of businesses and cluster of low number of businesses. The results obtained from the assessment process are based on the index of the number of businesses with 4 provinces, the number of high businesses, namely North Sumatra, West Java, Central Java, and East Java, 13 Provinces, the number of medium enterprises and 17 other provinces, including low business numbers. This can be used as input to the community, especially in provinces where the number of businesses is low so they can compete in the business world and input for the government to provide facilities and infrastructure to support entrepreneurs in Indonesia.

Keywords: Clustering, K-Means, Data Mining, Business

Abstrak-Revolusi industri 4.0 merupakan era dimana teknologi dalam bidang IT berkembang sangat pesat, dapat juga dikatakan era baru bagi entrepreneur di Indonesia yang harus dijadikan sebagai peluang emas untuk meningkatkan kinerja usahanya dan peluang bagi millennial terjun ke dunia usaha atau bisnis. Pada era ini masyarakat diharapkan dapat bersaing khususnya dalam dunia usaha. Penelitian ini membahas tentang Penerapan Data Mining Dalam Mengelompokkan Jumlah Usaha Berdasarkan Provinsi Menggunakan K-Means Clustering. Sumber data penelitian ini dikumpulkan berdasarkan dokumen-dokumen keterangan Jumlah Usaha di Indonesia yang dihasilkan oleh Badan Pusat Statistik Nasional. Data yang digunakan dalam penelitian ini adalah data provinsi yang terdiri dari 34 provinsi. Data akan diolah dengan melakukan clustering dalam 3 cluster yaitu cluster jumlah usaha tinggi, cluster jumlah usaha sedang dan cluster jumlah usaha rendah. Hasil yang diperoleh dari proses penilaian berdasarkan indeks jumlah usaha dengan 4 provinsi jumlah usaha tinggi, yaitu Sumatera Utara, Jawa Barat, Jawa Tengah, dan Jawa Timur, 13 Provinsi jumlah usaha sedang dan 17 provinsi lainnya termasuk jumlah usaha rendah. Hal ini dapat dijadikan masukan kepada masyarakat khususnya pada provinsi yang jumlah usahanya rendah agar dapat bersaing di dalam dunia usaha dan masukan untuk pemerintah agar menyediakan sarana dan prasarana yang mendukung para pengusaha di Indonesia.

Kata kunci: Clustering, K-Means, Data Mining, Usaha

1. PENDAHULUAN

Usaha biasanya identik dengan aktivitas bisnis, yang merupakan suatu kegiatan yang dilakukan untuk mendapatkan penghasilan berupa uang atau barang yang diperlukan untuk memenuhi kebutuhan sehari-hari. Indonesia saat ini telah memasuki era revolusi industri 4.0 yang membuat teknologi semakin cepat berkembang. Ini dapat menjadi peluang yang baik bagi usaha kecil, menengah maupun besar untuk meraih keuntungan yang nilainya dapat mencapai miliaran dolar AS. Agar tidak tertinggal dari negara-negara lain masyarakat diharapkan dapat memanfaatkan kemajuan teknologi ini dengan baik, seperti membuka atau membuat usaha yang sudah masuk era digitalisasi. Sebagai contoh seperti bisnis jual beli online yang semakin menjanjikan di era revolusi industri 4.0, dengan promosi yang tidak harus digembar-gemborkan serta tidak perlu biaya operasional yang tinggi dan cara pembayaran yang lebih mudah dilakukan. Penelitian ini membahas tentang membangkitkan masyarakat Indonesia untuk ikut berperan dalam menghadapi era revolusi industri 4.0.

Clustering merupakan salah satu metode data mining yang bersifat tanpa arahan (*unsupervised*) dan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik antara satu data dengan data yang lain [1]. Dalam metode *clustering* konsep utama yang ditekankan adalah pencarian pusat *cluster* secara iteratif, dimana pusat *cluster* ditentukan berdasarkan jarak *minimum* setiap data pada pusat *cluster* [2][3]. Data yang digunakan dalam penelitian ini berdasarkan dokumen-dokumen yang dihasilkan oleh Badan Pusat Statistik Nasional melalui situs <http://www.bps.go.id>. Dalam hal ini peneliti mengangkat topik mengelompokkan jumlah usaha berdasarkan provinsi dimana proses algoritma yang dilakukan adalah *clustering*. Hasil dari *clustering* dapat menjadi masukan bagi masyarakat di provinsi yang masih memiliki minat berusaha yang rendah agar memanfaatkan era revolusi industri 4.0 untuk membuka usaha. Proses *cluster* dibagi kedalam 3 (tiga) *cluster* yakni *cluster* jumlah usaha tinggi, *cluster* jumlah usaha sedang dan *cluster* jumlah usaha rendah.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database* (KDD) secara keseluruhan [4]. Definisi lain diantaranya adalah sebagai proses untuk mendapatkan informasi yang berguna dari Gudang basis data yang besar. Data mining juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan [5].

2.2 Clustering

Clustering adalah salah satu alat bantu pada data minig yang bertujuan mengelompokkan objek-objek ke dalam *cluster-cluster* [6]. *Cluster* adalah sekelompok atau sekumpulan objek-objek data yang similar satu sama lain dalam *cluster* yang sama dan dissimilar terhadap objek-objek yang berbeda *cluster*. Objek akan dikelompokkan ke dalam satu atau lebih *cluster* sehingga objek-objek yang berada dalam satu *cluster* akan mempunyai kesamaan yang tinggi antara satu dengan yang lainnya [7].

2.3 K-Means

K-Means merupakan salah satu metode data *clustering* non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok. Metode ini mempartisi data ke dalam *cluster*/kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam kelompok yang lain [8]–[10]. Adapun tujuan dari data *clustering* ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses *clustering*, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu *cluster* dan memaksimalkan variasi antar *cluster* [11].

2.4 Tahapan Pengumpulan Data

Dalam penerapan algoritma *clustering* dalam mengelompokkan jumlah usaha berdasarkan provinsi, diperlukan data terkait tentang hal tersebut. Sumber data penelitian diperoleh dari data yang dikumpulkan berdasarkan dokumen-dokumen keterangan kontruksi yang dihasilkan oleh Direktorat Jenderal melalui situs <http://www.bps.go.id>. Data yang digunakan dalam penelitian ini adalah data Jumlah Usaha pada setiap provinsi dari tahun 2014-2016 yang terdiri dari 34 provinsi. Variabel yang digunakan jumlah usaha yang sudah diakumulasikan. Data akan diolah dengan melakukan *clustering* jumlah usaha berdasarkan provinsi dalam 3 *cluster* yakni *cluster* jumlah usaha tinggi, *cluster* jumlah usaha sedang dan *cluster* jumlah usaha rendah.

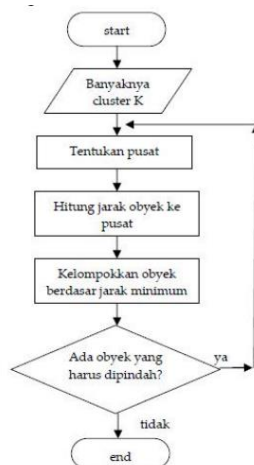
2.5 Tahapan Pengolahan Data

Data yang telah diperoleh akan diolah terlebih dahulu untuk dapat *diclustering*. Dalam tahap sebelumnya, data setiap provinsi akan dijumlah setiap aspeknya sehingga pada tahapan ini sudah diperoleh perhitungan nilai yang akan diproses pada tahap *clustering*.

2.6 Tahapan *Clustering*

Clustering merupakan klasifikasi tanpa pengawasan dan merupakan proses partisi sekumpulan objek data dari satu set menjadi beberapa kelas. Hal ini dapat dilakukan dengan menerapkan berbagai persamaan dan langkah-langkah mengenai jarak algoritma, yaitu dengan *Euclidean Distance*. Analisis *cluster* ialah metode yang dipakai untuk membagi rangkaian data menjadi beberapa grup berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya. Dalam menentukan *cluster* berdasarkan data yang telah tersedia, dibutuhkan sebuah *flowchart* untuk memudahkan dalam menentukan alur perhitungan sebagai alur

untuk menemukan hasil dari penerapan *cluster* terhadap data yang akan diproses. Berikut adalah *flowchart* dalam menentukan *cluster* dengan K-Means [12].



Gambar 1. Flowchart K-Means

3. HASIL DAN PEMBAHASAN

Pada tahap ini dilakukan analisis data jumlah usaha. Data yang diperoleh diolah dengan menggunakan perhitungan bobot dari tiap indeks. Pada tahap sebelumnya, telah ditentukan akan di*cluster* kedalam 3 *cluster* yakni *cluster* jumlah usaha tinggi, *cluster* jumlah usaha sedang dan *cluster* jumlah usaha rendah.

Dalam melakukan *clustering*, data yang diperoleh akan dihitung terlebih dahulu berdasarkan hasil rata-rata jumlah usaha berdasarkan provinsi. Data tersebut kemudian diakumulasikan dan diambil nilai rata-ratanya, yaitu :

Tabel 1. Nilai Rata-Rata Jumlah Usaha Berdasarkan Provinsi

Provinsi	Rata-rata	Provinsi	Rata-rata
Aceh	844	Nusa Tenggara Barat	563
Sumatra Utara	1214	Nusa Tenggara Timur	746
Sumatra Barat	757	Kalimantan Barat	518
Riau	432	Kalimantan Tengah	617
Jambi	506	Kalimantan Selatan	581
Sumatra Selatan	662	Kalimantan Timur	356
Bengkulu	430	Kalimantan Utara	224
Lampung	619	Sulawesi Utara	621
Kep.Bangka Belitung	286	Sulawesi Tengah	403
Kep.Riau	288	Sulawesi Selatan	701
DKI Jakarta	295	Sulawesi Tenggara	533
Jawa Barat	988	Gorontalo	259
Jawa Tengah	1169	Sulawesi Barat	218
Di Yogyakarta	166	Maluku	288
Jawa Timur	1439	Maluku Utara	326
Banten	376	Papua Barat	203
Bali	420	Papua	286

Sumber : Badan Pusat Statistik

Setelah diakumulasikan maka akan didapatkan nilai dari seluruh jumlah usaha berdasarkan provinsi. Kemudian data tersebut akan masuk ke tahapan

clustering dengan menerapkan algoritma K-Means untuk meng*cluster* data menjadi tiga *cluster*.

3.1 Centroid Data

Dalam penerapan algoritma K-Means dihasilkan nilai titik tengah atau *centroid* dari data yang didapatkan dengan ketentuan bahwa *clusterisasi* yang diinginkan adalah 3, Penentuan *cluster* dibagi atas 3 bagian yakni *cluster* jumlah usaha tinggi (C1), *cluster* jumlah usaha sedang (C2) dan *cluster* jumlah usaha rendah (C3). Maka nilai titik tengah atau *centroid* juga terdapat 3 titik. Penentuan titik *cluster* ini dilakukan dengan mengambil nilai terbesar (maximum) untuk *cluster* jumlah usaha tinggi (C1), nilai rata-rata (average) untuk *cluster* jumlah usaha sedang (C2) dan nilai terkecil untuk *cluster* jumlah usaha rendah (C3). Nilai titik tersebut dapat diketahui pada Tabel 2 berikut :

Tabel 2. Centroid Data Awal (Iterasi 1)

Centroid	
Max(C1)	1202
Average(C2)	562
Min(C3)	258

3.2 Cluster Data

Dengan menggunakan *centroid* tersebut maka dapat di*cluster* data yang telah didapat menjadi 3 *cluster*. Proses *cluster* dengan mengambil jarak terdekat dari setiap data yang diolah. Dari data jumlah usaha berdasarkan provinsi didapatkan pengelompokan pada iterasi 1 untuk 3 *cluster* tersebut. Proses pencarian jarak terpendek, pengelompokan data pada iterasi 1 dan *clustering* atas dapat digambarkan pada table dan gambar berikut :

Tabel 3. Perhitungan Jarak Pusat Cluster Iterasi 1

Provinsi	Rata-Rata	Iterasi1			Jarak terpendek
		C1	C2	C3	
Aceh	844	594	305	678	305
Sumatra Utara	1214	225	675	1048	225
Sumatra Barat	757	682	218	591	218
Riau	432	1007	108	265	108
Jambi	506	932	33	340	33
Sumatra Selatan	662	777	122	495	122
Bengkulu	430	1009	109	264	109
Lampung	619	820	79	452	79
Kep.Bangka Belitung	286	1153	253	120	120
Kep.Riau	288	1151	252	121	121
DKI Jakarta	295	1144	245	128	128
Jawa Barat	988	451	448	821	448
Jawa Tengah	1169	270	630	1003	270
Di Yogyakarta	166	1272	373	0	0
Jawa Timur	1439	0	899	1272	0
Banten	376	1063	163	210	163
Bali	420	1019	120	253	120
Nusa Tenggara Barat	563	875	24	397	24

Provinsi	Rata-Rata	Iterasi1			Jarak terdekat
		C1	C2	C3	
Nusa Tenggara Timur	746	692	207	580	207
Kalimantan Barat	518	920	21	352	21
Kalimantan Tengah	617	822	78	451	78
Kalimantan Selatan	581	858	41	414	41
Kalimantan Timur	356	1083	184	189	184
Kalimantan Utara	224	1215	315	58	58
Sulawesi Utara	621	817	82	455	82
Sulawesi Tengah	403	1036	137	236	137
Sulawesi Selatan	701	737	162	535	162
Sulawesi Tenggara	533	905	6	367	6
Gorontalo	259	1179	280	93	93
Sulawesi Barat	218	1221	321	52	52
Maluku	288	1150	251	122	122
Maluku Utara	326	1113	213	160	160
Papua Barat	203	1236	337	36	36
Papua	286	1152	253	120	120

Tabel 4. Hasil Pengelompokan Iterasi 1

Provinsi	C1	C2	C3	Provinsi	C1	C2	C3
Aceh		1		Nusa Tenggara Barat		1	
Sumatra Utara	1			Nusa Tenggara Timur		1	
Sumatra Barat		1		Kalimantan Barat		1	
Riau		1		Kalimantan Tengah		1	
Jambi		1		Kalimantan Selatan		1	
Sumatra Selatan		1		Kalimantan Timur		1	
Bengkulu		1		Kalimantan Utara			1
Lampung		1		Sulawesi Utara		1	
Kep.Bangka Belitung			1	Sulawesi Tengah		1	
Kep.Riau			1	Sulawesi Selatan		1	
DKI Jakarta			1	Sulawesi Tenggara		1	
Jawa Barat	1			Gorontalo			1
Jawa Tengah	1			Sulawesi Barat			1
Di Yogyakarta			1	Maluku			1
Jawa Timur	1			Maluku Utara			1
Banten		1		Papua Barat			1
Bali		1		Papua			1

Proses K-Means akan terus beriterasi sampai pengelompokan data sama dengan pengelompokan data iterasi sebelumnya. Dengan kata lain, proses akan terus melakukan iterasi sampai data pada iterasi terakhir sama dengan iterasi sebelumnya. Setelah mendapatkan nilai titik tengah atau *centroid*, proses sama dilakukan dengan mencari jarak terdekat. Proses pencarian jarak terdekat, pengelompokan data pada iterasi terakhir dan *clustering* data dapat digambarkan pada tabel berikut:

Tabel 5. Centroid Data Iterasi 4

Centroid	
Max(C1)	1202
Average(C2)	636

Min(C3)	309
---------	-----

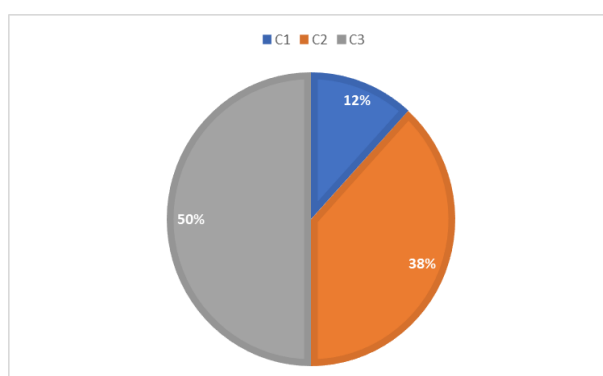
Tabel 6. Perhitungan Jarak Pusat *Cluster* Iterasi 4

Provinsi	Rata-Rata	Iterasi1			Jarak terpendek
		C1	C2	C3	
Aceh	844	358	208	535	208
Sumatra Utara	1214	12	578	905	12
Sumatra Barat	757	445	121	448	121
Riau	432	771	204	123	130
Jambi	506	696	130	197	130
Sumatra Selatan	662	541	26	353	26
Bengkulu	430	772	206	121	121
Lampung	619	584	17	310	17
Kep.Bangka Belitung	286	916	350	23	23
Kep.Riau	288	915	348	21	21
DKI Jakarta	295	908	341	14	14
Jawa Barat	988	215	352	679	215
Jawa Tengah	1169	33	533	860	33
Di Yogyakarta	166	1036	470	143	143
Jawa Timur	1439	236	803	1130	236
Banten	376	826	260	67	67
Bali	420	783	216	111	111
Nusa Tenggara Barat	563	639	73	254	73
Nusa Tenggara Timur	746	456	110	437	110
Kalimantan Barat	518	684	118	209	118
Kalimantan Tengah	617	585	19	308	19
Kalimantan Selatan	581	622	55	272	55
Kalimantan Timur	356	847	280	47	47
Kalimantan Utara	224	978	412	85	85
Sulawesi Utara	621	581	15	312	15
Sulawesi Tengah	403	800	233	94	94
Sulawesi Selatan	701	501	65	392	65
Sulawesi Tenggara	533	669	103	224	103
Gorontalo	259	943	377	50	50
Sulawesi Barat	218	984	418	91	91
Maluku	288	914	348	21	21
Maluku Utara	326	876	310	17	17
Papua Barat	203	1000	433	106	106
Papua	286	916	350	23	23

Tabel 7. Hasil Pengelompokan Iterasi 4

Provinsi	C1	C2	C3	Provinsi	C1	C2	C3
Aceh		1		Nusa Tenggara Barat		1	
Sumatra Utara	1			Nusa Tenggara Timur		1	
Sumatra Barat		1		Kalimantan Barat		1	
Riau			1	Kalimantan Tengah		1	
Jambi		1		Kalimantan Selatan		1	
Sumatra Selatan		1		Kalimantan Timur			1
Bengkulu			1	Kalimantan Utara			1
Lampung		1		Sulawesi Utara		1	
Kep.Bangka Belitung			1	Sulawesi Tengah			1
Kep.Riau			1	Sulawesi Selatan		1	

Provinsi	C1	C2	C3	Provinsi	C1	C2	C3
DKI Jakarta			1	Sulawesi Tenggara		1	
Jawa Barat	1			Gorontalo			1
Jawa Tengah	1			Sulawesi Barat			1
DI Yogyakarta			1	Maluku			1
Jawa Timur	1			Maluku Utara			1
Banten			1	Papua Barat			1
Bali			1	Papua			1



Gambar 2. Clustering Data Iterasi 4

3.3 Analisa Data

Pada iterasi 4, pengelompokan data yang dilakukan terhadap 3 *cluster* dengan iterasi 3 dihasilkan sama. Dari 34 data jumlah usaha berdasarkan provinsi dapat diketahui bahwa provinsi *cluster* tinggi (C1) sebanyak 4 provinsi, untuk *cluster* sedang (C2) sebanyak 13 provinsi dan untuk *cluster* rendah (C3) sebanyak 17 provinsi.

4. KESIMPULAN

Dari permasalahan diatas untuk melakukan penilaian jumlah usaha berdasarkan provinsi dapat menerapkan metode *clustering* K-Means. Data tersebut diolah menggunakan Ms.Excel untuk ditentukan nilai *centroid* dalam 3 *cluster* yaitu berdasarkan *cluster* jumlah usaha tinggi, *cluster* jumlah usaha sedang dan *cluster* jumlah usaha rendah. *Centroid* data untuk *cluster* jumlah usaha tinggi 1202, *centroid* data untuk *cluster* jumlah usaha sedang 636 dan *centroid* data untuk *cluster* jumlah usaha rendah 309. Sehingga diperoleh penilaian berdasarkan indeks jumlah usaha dengan 4 provinsi *cluster* tinggi, 13 provinsi *cluster* sedang dan 17 provinsi *cluster* rendah. Hasil dari penelitian ini dapat dijadikan masukan bagi masyarakat khususnya provinsi yang masuk kedalam jumlah usaha terendah agar dapat bersaing di era revolusi industri 4.0 khususnya dalam dunia usaha.

DAFTAR PUSTAKA

- [1] I. Parlina, A. P. Windarto, A. Wanto, and M. R. Lubis, "Memanfaatkan Algoritma K-Means dalam Menentukan Pegawai yang Layak Mengikuti Asessment Center untuk Clustering Program SDP," *CESS (Journal of Computer Engineering System and Science)*, vol. 3, no. 1, pp. 87-93, 2018.

- [2] A. P. Windarto, P. Studi, S. Informasi, and D. Mining, "Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering," vol. 16, no. 4, pp. 348-357, 2017.
- [3] M. G. Sadewo, A. P. Windarto, and S. R. Andani, "PEMANFAATAN ALGORITMA CLUSTERING DALAM MENGELOMPOKKAN JUMLAH DESA / KELURAHAN YANG MEMILIKI SARANA KESEHATAN," vol. I, pp. 124-131, 2017.
- [4] P. Bidang, K. Sains, Y. Mardi, J. Gajah, M. No, and S. Barat, "Jurnal Edik Informatika Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD) . Jurnal Edik Informatika."
- [5] S. Kasus, U. Dehasen, S. Haryati, A. Sudarsono, and E. Suryana, "IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI MASA STUDI MAHASISWA MENGGUNAKAN ALGORITMA C4 . 5," vol. 11, no. 2, pp. 130-138, 2015.
- [6] S. Defiyanti, M. Jajuli, T. Informatika, F. Ilmu, K. Universitas, and S. Karawang, "IMPLEMENTASI ALGORITMA K-MEANS DALAM," vol. I, no. 2, pp. 62-68, 2015.
- [7] B. M. Metisen and H. L. Sari, "ANALISIS CLUSTERING MENGGUNAKAN METODE K-MEANS DALAM PENGELOMPOKKAN PENJUALAN PRODUK PADA SWALAYAN FADHILA," vol. 11, no. 2, pp. 110-118, 2015.
- [8] S. Sudirman, A. P. Windarto, and A. Wanto, "Data Mining Tools | RapidMiner : K-Means Method on Clustering of Rice Crops by Province as Efforts to Stabilize Food Crops In Indonesia," *IOP Conference Series: Materials Science and Engineering*, vol. 420, no. 12089, pp. 1-8, 2018.
- [9] M. G. Sadewo, A. P. Windarto, and A. Wanto, "Penerapan Algoritma Clustering dalam Mengelompokkan Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi/ Mitigasi Bencana Alam Menurut Provinsi dengan K-Means," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 2, no. 1, pp. 311-319, 2018.
- [10] R. W. Sari, A. Wanto, and A. P. Windarto, "Implementasi Rapidminer dengan Metode K-Means (Study Kasus : Imunisasi Campak pada Balita Berdasarkan Provinsi)," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 2, no. 1, pp. 224-230, 2018.
- [11] M. Anggara, H. Sujiani, and H. Nasution, "Pemilihan Distance Measure Pada K-Means Clustering Untuk Pengelompokkan Member Di Alvaro Fitness," vol. 1, no. 1, pp. 1-6, 2016.
- [12] M. G. Sadewo, A. P. Windarto, and D. Hartama, "Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means," *InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan)*, vol. 2, no. 1, pp. 60-67, 2017.