

PENERAPAN METODE K-MEANS CLUSTERING UNTUK MENENTUKAN STATUS NEGARA

Alya Aulia Nurdin¹, Shabrina Jamilah², Robby Igfirly Mustaib³

¹Program Studi Sistem Informasi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang

²Program Studi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Gunadarma

³Program Studi Teknik Informatika, Fakultas Teknik, Universitas Mataram

e-mail: 1alyaaulianurdin@gmail.com, 2shabrinajamilah220@gmail.com, 3robbyigfirly007@gmail.com

ABSTRAK

Status dari sebuah negara merupakan hal yang penting untuk diketahui. Hal tersebut dikarenakan status negara dapat menunjang kesuksesan suatu investasi, bisnis, dan juga pemerataan pembangunan oleh pemerintahnya. Status negara yang sering didengar oleh masyarakat, yakni negara maju dan negara berkembang. Dalam studi kasus penentuan status negara ini, penulis mengambil dataset International Monetary Fund (IMF) yang diperoleh dari situs github. Kemudian data mentah dilakukan preprocessing terlebih dahulu berupa impute values dan normalisasi, lalu selanjutnya dilakukan perhitungan klastering dengan metode K-Means. Data akan dibagi menjadi empat klaster dan perhitungan dilakukan dengan dua cara, yakni melalui Microsoft Excel dan Rapid Miner. Berdasarkan perhitungan yang telah dilakukan, 186 negara sudah terbagi menjadi empat klaster yakni, klaster negara agraris, negara industri, negara berkembang, dan negara maju. Penentuan klasterisasi status negara tersebut cukup dikatakan valid karena performance vektornya sebesar 70%.

Kata Kunci: Klastering, K-Means, Status Negara, Penambangan Data, Kecerdasan Bisnis

1. PENDAHULUAN

Suatu negara adalah kelompok sosial yang menempati wilayah tertentu yang terorganisir di bawah institusi politik dan pemerintah yang efektif, memiliki kesatuan politik, berdaulat sehingga berhak menentukan tujuan nasional [1]. Status sebuah negara adalah hal yang penting untuk diketahui guna pemerintah negara tersebut dapat menentukan suatu kebijakan yang tepat demi menunjang kesuksesan investasi, bisnis, maupun pemerataan pembangunan dari suatu negara. Pembangunan ekonomi ini merupakan salah satu topik yang banyak dibahas [2]. Kategorisasi negara dan juga peringkatnya berdasarkan status pembangunan telah menjadi fokus utama dari berbagai lembaga seperti Bank Dunia (World Bank), International Monetary Fund (IMF), dan United Nation Development Programme (UNDP) [3]. Menurut penelitian yang telah dilakukan [3], kategorisasi negara dilakukan supaya lembaga-lembaga tersebut dapat memfasilitasi pembuatan kebijakan yang lebih baik dan penggunaan sumber daya yang efisien bagi suatu negara. Biasanya, suatu negara yang tergolong negara maju atau berkembang dapat dilihat dari banyak faktor seperti ukuran penduduk, tingkat pertumbuhan penduduk, tingkat kejahatan, persentase korupsi, tingkat kelahiran dan kematian, tingkat pengangguran, inflasi, jumlah pengunjung sektor pariwisata, pendapatan per kapita, dan lain-lain [1], [4]. Di banyak negara, sektor pariwisata dianggap sebagai sumber pendapatan devisa utama dan memiliki peluang untuk pembangunan ekonomi [5], [6].

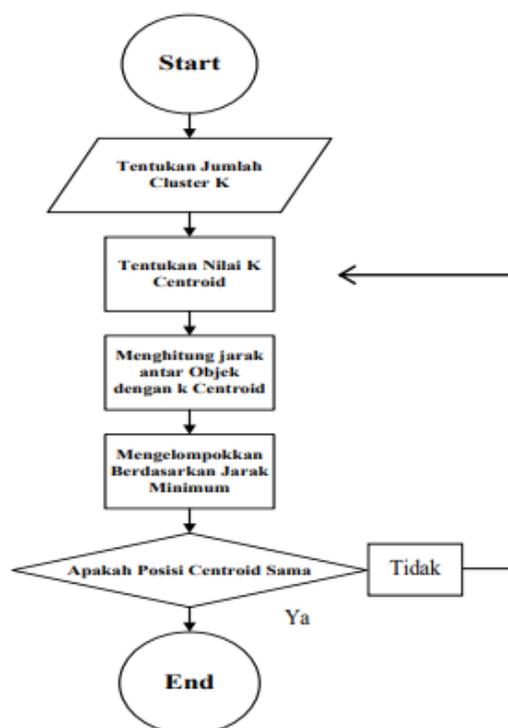
Beberapa penelitian yang sudah dilakukan sebelumnya, telah membahas mengenai klasterisasi sebuah negara. Peneliti di Indonesia [7] berfokus pada kategorisasi yang berdasarkan laju inflasi dengan menggunakan algoritma K-Means. K-Means adalah suatu metode penganalisaan data atau metode *data mining* yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Metode K-Means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok [8], di mana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Dengan kata lain, metode ini berusaha untuk meminimalkan variasi antar data yang ada di dalam suatu klaster dan memaksimalkan variasi dengan data yang ada di klaster lainnya. Penelitian lainnya menggunakan K-Means untuk mengklasterisasi wisata asing di Indonesia [9]. Data pada penelitian tersebut diolah menggunakan Microsoft Excel untuk ditentukan nilai centroid dalam 3 klaster yaitu klaster tinggi (C1), klaster sedang (C2) dan klaster rendah (C3) dan hasil penelitian menyatakan bahwa hampir 90% provinsi di Indonesia masih memiliki potensi wisata yang rendah. Sementara itu, peneliti lain [10] melakukan klasterisasi negara berdasarkan emisi gas di negara-negara benua Eropa menggunakan K-Means. Negara-negara di Eropa tersebut dikategorikan dalam empat klaster. Selain berdasarkan laju inflasi dan emisi gas, penelitian lain juga menggunakan K-Means untuk mengklaster negara berdasarkan data impor beras [11]. Hasilnya adalah penilaian berdasarkan indeks impor beras dengan dua negara klaster impor tinggi yaitu Vietnam dan Thailand, empat kelompok menengah negara impor sedang yaitu China, India, Pakistan dan Lainnya serta empat negara klaster impor rendah yaitu USA, Taiwan, Singapura, dan Myanmar. Dalam [12], *Gross Regional Domestic Product* (GRDP) digunakan sebagai indikator untuk mengklaster Kota Surabaya. GRDP adalah indikator yang digunakan untuk mengukur kinerja ekonomi suatu daerah dalam suatu

periode [13][14]. K-Means juga digunakan untuk mengklaster di beberapa penelitian di berbagai bidang, misalnya untuk *clustering cybercrime* [15], [16], kelompok pelanggan [8], bidang kesehatan [17]–[21], industri manufaktur [22], dan *techno-economic* [23], [24].

Selanjutnya, dalam studi kasus pada penelitian kali ini, penulis bertujuan menentukan kategori atau kluster status dari sebuah negara dengan menggunakan metode K-Means clustering. Status negara yang akan diklasterisasi, yakni negara agraris, negara industri, negara berkembang, dan negara maju. Data yang digunakan merupakan dataset negara dari IMF (International Monetary Fund) yang diperoleh dari situs github. Dataset tersebut berisi 186 negara dengan berbagai informasi, di antaranya jumlah pendapatan dari masing-masing negara, neraca pembayaran, tabungan nasional bruto, serta total investasi masing-masing negara tersebut.

2. METODE PENELITIAN

Dalam penelitian yang bertujuan untuk menentukan status negara ini, digunakan metode K-Means clustering. K-Means merupakan salah satu algoritma clustering. Tujuan algoritma ini yaitu untuk membagi data menjadi beberapa kelompok. Algoritma ini akan mengelompokkan data atau objek ke dalam k buah kelompok tersebut. Pada setiap kluster terdapat titik pusat (centroid) yang merepresentasikan kluster tersebut. Gambar 1 menunjukkan *flowchart* metode K-Means.



Gambar 1. *Flowchart* metode K-Means Clustering [7]

Adapun kelebihan dari algoritma ini, di antaranya mudah dilakukan saat pengimplementasian dan dijalankan, waktu yang dibutuhkan untuk melakukan pembelajaran relatif lebih cepat, sangat fleksibel, adaptasi yang mudah untuk dilakukan, sangat umum penggunaannya, dan menggunakan prinsip yang sederhana dapat dijelaskan dalam non-statistik. Algoritma K-Means pada dasarnya melakukan dua proses, yakni proses pendeteksian lokasi pusat tiap kluster dan proses pencarian anggota dari tiap-tiap kluster. Cara kerja dari algoritma K-Means adalah sebagai berikut.

- 1) Tentukan k sebagai jumlah cluster yang ingin dibentuk.
- 2) Bangkitkan k centroid (titik pusat cluster) awal secara random.
- 3) Hitung jarak setiap data ke masing-masing centroid. Menggunakan perhitungan Euclidean Distance sebagai berikut.

$$D(p, c)_n = \sqrt{\sum_{i=0}^n (p_i - c_i)^2} \quad (1)$$

Keterangan: p = data c = centroid

- 4) Setiap data memilih centroid yang terdekat.
- 5) Tentukan posisi centroid yang baru dengan cara menghitung nilai rata-rata dari data-data yang terletak pada centroid yang sama.

6) Kembali ke langkah-3 jika posisi centroid baru dengan centroid yang lama tidak sama.

Namun, sebelum hal itu dilakukan, dataset yang diperoleh harus dilakukan *preprocessing* terlebih dahulu supaya data yang digunakan dan dihitung adalah data yang memang bagus. Dalam hal ini, dilakukan tahap *preprocessing* data sebagai berikut. Pertama adalah mencari *missing value* (data yang kosong atau hilang), kemudian dilakukan *impute value* (mengganti data yang kosong/hilang dengan rata-rata setiap kolomnya), serta yang terakhir adalah melakukan normalisasi data (agar data yang dihitung tidak terdapat perbedaan yang besar atau tumpang tindih). Adapun karena adanya perkembangan teknologi yang pesat [25] di mana sudah terdapat beberapa *tools* yang dapat dimanfaatkan, maka perhitungan yang dilakukan dalam penelitian ini dilakukan dengan dua *tools*, yakni menggunakan Microsoft Excel dan Rapid Miner 5.3.

3. HASIL DAN PEMBAHASAN

Dalam penelitian yang dilakukan, berikut ini pada Tabel 1 adalah dataset yang dihitung menggunakan *tool* Microsoft Excel. Dalam dataset dari IMF tersebut terdapat beberapa atribut indikator yang digunakan sebagai dasar untuk mengklaster status negara, yaitu *current account balance*, *general government revenue*, *gross national savings*, dan *total investment* dari sebuah negara.

Tabel 1. Dataset IMF

Current account balance	General government revenue	Gross national savings	Total investment	Country
3.877	21.977	30.398	26.521	Afghanistan
-11.372	25.835	14.509	25.886	Albania
7.489	36.458	48.947	41.428	Algeria
9.024	43.479	21.692	12.668	Angola
-13.109	22.43	16.194	29.303	Antigua and Barbuda
0.658	37.199	22.595	24.451	Argentina
-14.653	20.97	16.66	31.313	Armenia
-2.87	31.846	23.925	26.794	Australia
3.009	48.105	24.611	21.602	Austria
28.423	45.652	46.955	18.532	Azerbaijan
...
7.068	19.569	29.635	22.567	Zambia
-28.803	29.585	-4.474	24.33	Zimbabwe

Selanjutnya, setelah dilakukan *preprocessing* berupa *impute values* dan normalisasi (skala 0-1), diperoleh data yang lebih lengkap sebagai berikut pada Tabel 2.

Tabel 2. Hasil dataset yang telah dilakukan *preprocessing*

Current account balance (W)	General government revenue (X)	Gross national savings (Y)	Total investment (Z)	Country
0,456009059	0,220858626	0,580988356	0,405530273	Afghanistan
0,268315199	0,277291011	0,356195974	0,39086142	Albania
0,500467727	0,43267754	0,843413551	0,749890272	Algeria
0,51936143	0,535376289	0,457818712	0,085518261	Angola
0,246935158	0,363241425	0,380034803	0,469796022	Antigua and Barbuda
0,246935158	0,443516419	0,470594061	0,357712121	Argentina
0,227930678	0,363241425	0,470594061	0,516228141	Armenia
0,246935158	0,365216119	0,489410466	0,411836725	Australia
0,445325193	0,603042493	0,49911577	0,291898635	Austria
0,758135985	0,567161559	0,815231385	0,220979926	Azerbaijan
...

0,495285806	0,185635925	0,570193682	0,314190672	Zambia
0,05376397	0,332143641	0,087630689	0,314190672	Zimbabwe

Kemudian, pada Tabel 3 menunjukkan data centroid awal yang ditentukan secara random sebagai berikut. Lalu hasil perhitungan iterasi pertama dapat dilihat pada Tabel 4.

Tabel 3. Data centroid awal

Centroid (C)	W	X	Y	Z
C1	0,9	0,8	0,7	0,8
C2	0,7	0,8	0,5	0,7
C3	0,5	0,6	0,6	0,4
C4	0,3	0,5	0,4	0,4

Tabel 4. Hasil perhitungan iterasi pertama

Data ke-	C1	C2	C3	C4	Jarak Terpendek	Cluster
1	0,838035	0,698719	0,382198	0,367486	0,367486193	C4
2	0,978696	0,75883	0,4662	0,229359	0,229358947	C4
3	0,563587	0,543287	0,457898	0,603124	0,457897788	C3
...
184	1,146261	0,92669	0,598974	0,432762	0,432762019	C4
185	0,891121	0,757052	0,42423	0,416282	0,416281693	C4
186	1,243387	0,97745	0,735363	0,440166	0,440165613	C4

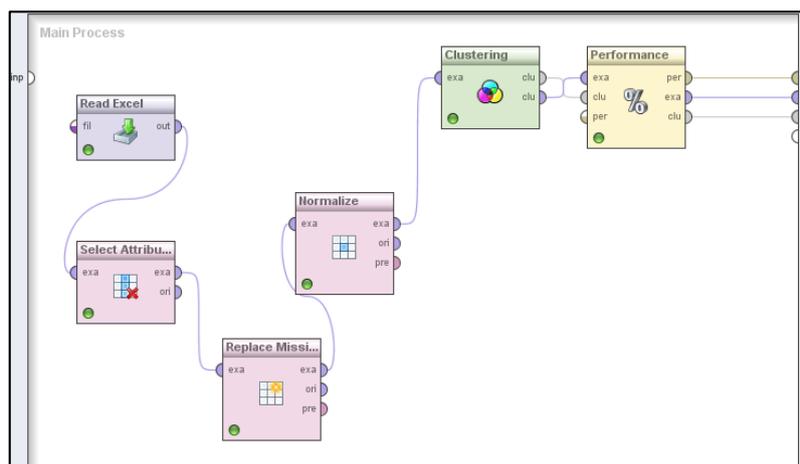
Dari iterasi pertama, didapatkan hasil yakni pada cluster 1 terdapat satu item negara yaitu Libya, pada cluster 2 terdapat satu item negara yaitu Timor Leste, pada cluster 3 terdapat 30 item negara salah satunya adalah Algeria, dan pada cluster 4 terdapat 154 item negara salah satunya yaitu Afghanistan. Selanjutnya dari setiap cluster tersebut dihitung lagi centroid nya, dan menghasilkan centroid baru sebagai berikut pada Tabel 5.

Tabel 5. Data centroid baru

Centroid (C)	W	X	Y	Z
C1	0,6521	0,86471	1	0,72171
C2	1	1	0,43244	0,3309
C3	0,35481	0,3649	0,43773	0,32858
C4	0,35459	0,36133	0,43716	0,32957

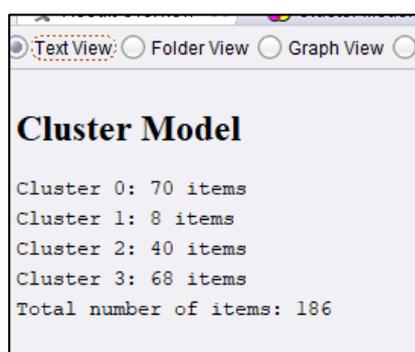
Kemudian, dilakukan perhitungan iterasi ke-2 dan menghasilkan data pada cluster 1 terdapat lima item negara yaitu Algeria, Bhutan, China, Kuwait, dan Libya, pada cluster 2 terdapat dua item negara yaitu Brunei Darussalam dan Timor Leste, pada cluster 3 terdapat 73 item negara salah satunya adalah Angola, dan pada cluster 4 terdapat 106 item negara salah satunya adalah Afghanistan. Selanjutnya, dilakukan perhitungan centroidnya lagi dan menghasilkan centroid baru yang berbeda dengan centroid sebelumnya.

Dengan adanya perbedaan centroid tersebut, maka perlu dilakukan iterasi hingga iterasi ke-n sampai nilai centroid tetap sama dan tidak berubah. Namun, karena data yang digunakan sangat banyak, penulis memutuskan untuk melakukan iterasi di Microsoft Excel hanya sampai iterasi ke-2. Sementara itu, untuk perhitungan selanjutnya dilakukan dengan *tool* Rapid Miner 5.3 pada Gambar 2.



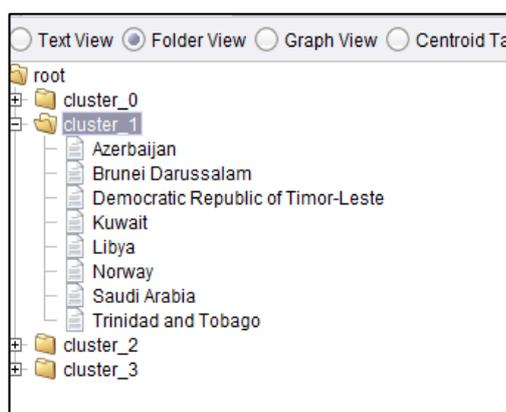
Gambar 2. Perhitungan dengan tool Rapidminer 5.3

Pertama, dataset yang ada di Microsoft Excel diinputkan ke Rapid Miner, kemudian memilih atribut dan id yang akan digunakan. Dalam hal ini, penulis menggunakan nama negara sebagai id, kemudian menggunakan empat item data (jumlah pendapatan dari masing-masing negara, neraca pembayaran, tabungan nasional bruto, serta total investasi masing-masing negara tersebut) sebagai atribut. Setelah itu, dilakukan *preprocessing* berupa *impute values* dan normalisasi. Kemudian, pada menu clustering, bagi menjadi empat kluster dan diperoleh hasil sebagai berikut pada Gambar 3.



Gambar 3. Hasil cluster model

Karena penulisan hasil di Rapid Miner berbentuk array, berdasarkan Gambar 3 di atas, maka dapat diinterpretasikan bahwa, pada cluster pertama, terdapat 70 item negara salah satunya ada Algeria, Bhutan, dan China, pada cluster kedua terdapat 8 item negara yakni Azerbaijan, Brunei Darussalam, Timor Leste, Kuwait, Libya, Norway, Saudi Arabia, dan Trinidad and Tobago, pada cluster ketiga terdapat 40 item negara salah satunya adalah Central African Republic, Cyprus, dan Czech Republic, dan pada cluster keempat terdapat 68 item negara salah satunya adalah Burkina Faso, Cambodia, dan Cameroon. Total negara yang diklusterisasi ada 186. Hasil klusterisasi dapat dilihat juga dalam folder yang sudah dikelompokkan sebagai berikut pada Gambar 4.

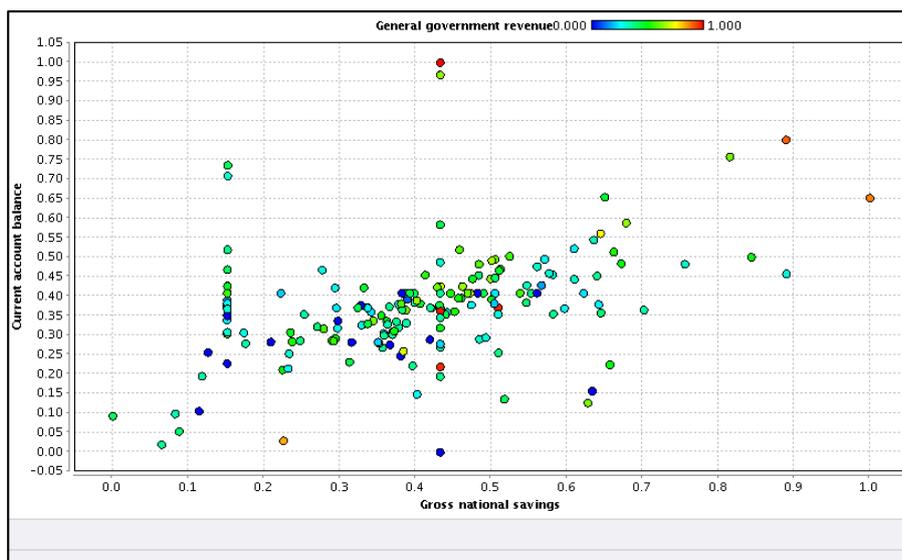


Gambar 4. Folder hasil cluster

Centroid akhir yang ada pada tiap kluster juga dapat diketahui, sebagai berikut pada Gambar 5 dan hasil juga dapat dilihat dalam bentuk grafik sebagai berikut pada Gambar 6.

Attribute	cluster_0	cluster_1	cluster_2	cluster_3
Current accc	0.314	0.748	0.395	0.381
General gov	0.223	0.735	0.339	0.528
Gross nator	0.283	0.693	0.574	0.378
Total investn	0.317	0.373	0.625	0.383

Gambar 5. Centroid akhir



Gambar 6. Grafik persebaran atribut dalam cluster

Adapun, persentase validnya hasil klustering di atas adalah +- 70%. Hal tersebut dapat dilihat pada hasil *performance vector* (Davies Bouldin) sebanyak 0,307.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, maka dapat disimpulkan bahwa pembagian klusterisasi status dari 186 negara di dunia adalah pada cluster pertama, terdapat 70 item negara di antaranya ada Algeria, Bhutan, dan China, pada cluster kedua terdapat 8 item negara yakni Azerbaijan, Brunei Darussalam, Timor Leste, Kuwait, Libya, Norway, Saudi Arabia, dan Trinidad and Tobago, pada cluster ketiga terdapat 40 item negara di antaranya adalah Central African Republic, Cyprus, dan Czech Republic, dan pada cluster keempat terdapat 68 item negara di antaranya adalah Burkina Faso, Cambodia, dan Cameroon. Namun, pemberian nama status (negara agraris, negara industri, negara berkembang, dan negara maju) belum dapat dilakukan secara eksplisit karena belum ada metode untuk menentukannya secara valid dalam arti dibutuhkan indikator-indikator lain yang lebih detail untuk menentukannya. Maka dari itu, penentuan klusterisasi status negara ini hanya dilakukan secara umum, yakni sudah terbagi ke dalam empat kluster dengan *performance vector* sebesar 70%.

5. SARAN

Saran yang disampaikan oleh peneliti dalam penelitian ini adalah pada penelitian berikutnya dapat menggunakan kombinasi algoritma yang lebih akurat untuk melakukan klusterisasi sehingga nantinya hasil yang didapatkan bisa lebih mendalam dan valid berdasarkan indikator-indikator pada dataset yang digunakan pula. Penelitian berikutnya yang lebih lanjut dapat menggunakan dataset yang sama dengan menggunakan tambahan beberapa indikator lagi untuk dapat mengklusterisasi negara-negara di dunia bukan hanya berdasarkan data ekonomi saja melainkan berdasarkan atribut atau indikator lainnya selain ekonomi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada laboratorium Artificial Intelligence & Data Mining Centre (AIDM) Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang yang telah memberi dukungan bimbingan pengetahuan dan penulisan dalam penelitian ini.

DAFTAR PUSTAKA

- [1] L. Wulandari and B. O. Yogantara, "Algorithm Analysis of K-Means and Fuzzy C-Means for Clustering Countries Based on Economy and Health," vol. 15, no. 2, pp. 109–116, 2022.
- [2] M. Wielechowski, D. Cherevyk, K. Czech, P. Kotyza, Ł. Grzęda, and L. Smutka, "Interdependence between human capital determinants and economic development: K-means regional clustering approach for Czechia and Poland," *Entrep. Bus. Econ. Rev.*, vol. 9, no. 4, pp. 173–194, 2021, doi: 10.15678/EBER.2021.090411.
- [3] S. Basel, K. U. Gopakumar, and R. P. Rao, "Classification of countries based on development indices by using K-means and grey relational analysis," *GeoJournal*, vol. 2, no. 2011, 2021, doi: 10.1007/s10708-021-10479-2.
- [4] U. A. Gani, R. Bambang, and K. Umam, "Discriminant Analysis of Classify Developed Countries and Developing Countries With Fisher Method," *J. Geuthèè Penelit. Multidisiplin*, vol. 01, no. 01, pp. 1–12, 2018, [Online]. Available: <http://www.journal.geutheeinstitute.com>.
- [5] A. A. Nurdin, "Decision support system for geuthing the best tourist attractions using simple additive weighting (SAW) method," *J. Soft Comput. Explor.*, vol. 2, no. 2, pp. 77–85, 2021.
- [6] D. Barbe and L. Pennington-Gray, "Social Media and Crisis Communication in Tourism and Hospitality," *Handb. e-Tourism*, pp. 1–27, 2020, doi: 10.1007/978-3-030-05324-6_130-1.
- [7] Y. Prayoga, H. S. Tambunan, and I. Parlina, "Penerapan Clustering Pada Laju Inflasi Kota Di Indonesia Dengan Algoritma K-Means," *BRAHMANA J. Penerapan Kecerdasan Buatan*, vol. 1, no. 1, pp. 24–30, 2019, doi: 10.30645/brahmmana.v1i1.4.
- [8] N. Hazimah, S. Harahap, A. Amirullah, M. B. Saputro, and I. A. Tamaroh, "Classification of potential customers using C4 . 5 and k-means algorithms to determine customer service priorities to maintain loyalty," pp. 123–130, 2022, doi: 10.52465/josce.v3i2.89.
- [9] R. W. Sari and D. Hartama, "Data Mining : Algoritma K-Means Pada Pengelompokan Wisata Asing ke Indonesia Menurut Provinsi," *Semin. Nas. Sains Teknol. Inf.*, pp. 322–326, 2018.
- [10] A. Kijewska and A. Bluszcz, "Research of varying levels of greenhouse gas emissions in European countries using the k-means method," *Atmos. Pollut. Res.*, vol. 7, no. 5, pp. 935–944, 2016, doi: 10.1016/j.apr.2016.05.010.
- [11] A. P. Windarto, "Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method," *Int. J. Artif. Intell. Res.*, vol. 1, no. 2, p. 26, 2017, doi: 10.29099/ijair.v1i2.17.
- [12] N. A. Febriyati, A. D. GS, and A. Wanto, "GRDP Growth Rate Clustering in Surabaya City uses the K-Means Algorithm," *IJISTECH (International J. Inf. Syst. Technol.)*, vol. 3, no. 36, pp. 276–283, 2020.
- [13] A. B. M. Bintang, "Pengaruh PDRB, Pendidikan, Kesehatan, dan Pengangguran Terhadap Tingkat Kemiskinan di Jawa Tengah (2011-2015)," *MEDIA Ekon. DAN Manaj.*, vol. 33, no. 1, pp. 20–28, 2018.
- [14] R. Giovanni, "Analisis Pengaruh PDRB, Pengangguran dan Pendidikan Terhadap Tingkat Kemiskinan di Pulau Jawa Tahun 2009-2016," *Econ. Dev. Anal. J.*, vol. 7, no. 1, pp. 23–31, 2018.
- [15] S. Rustam, "Analisa Clustering Phising Dengan K-Means Dalam Meningkatkan Keamanan Komputer," *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 175–181, 2018, doi: 10.33096/ilkom.v10i2.309.175-181.
- [16] A. Kigerl, "Cyber crime nation typologies: K-means clustering of countries based on cyber crime rates," *Int. J. Cyber Criminol.*, vol. 10, no. 2, pp. 147–169, 2016, doi: 10.5281/zenodo.163399.
- [17] 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 Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 224–230, 2018, doi: 10.30865/komik.v2i1.930.
- [18] N. Rofiqo, A. P. Windarto, and D. Hartama, "Penerapan Clustering Pada Penduduk Yang Mempunyai Keluhan Kesehatan Dengan Datamining K-Means," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 216–223, 2018, doi: 10.30865/komik.v2i1.929.
- [19] V. Chandu, "Identification of spatial variations in COVID-19 epidemiological data using K-Means clustering algorithm: a global perspective," *medRxiv*, p. 2020.06.03.20121194, 2020, [Online]. Available: <http://medrxiv.org/content/early/2020/06/05/2020.06.03.20121194.abstract>.
- [20] J. Hutagalung, N. L. W. S. R. Ginantra, G. W. Bhawika, W. G. S. Parwita, A. Wanto, and P. D. Panjaitan, "COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm," *J. Phys. Conf. Ser.*, vol. 1783, no. 1, 2021, doi: 10.1088/1742-6596/1783/1/012027.
- [21] M. E. Al Rivian and R. A. Sonaru, "Perbandingan Metode K-Means Dan GA K-Means Untuk Clustering Dataset Heart Disease Patients," *J. Tek. Inform. dan Sist. Inf.*, vol. 9, no. 3, pp. 2585–2597, 2022.
- [22] Z. Karaca, "the Cluster Analysis in the Manufacturing Industry With K-Means Method: an Application for Turkey," *Eurasian J. Econ. Financ.*, vol. 6, no. 3, pp. 1–12, 2018, doi: 10.15604/ejef.2018.06.03.001.
- [23] A. Jha and D. Saha, "Examining categorization of Telecom Circles in India using unsupervised k-means

- clustering on techno-economic indicators,” *Decision*, vol. 46, no. 4, pp. 365–383, 2019, doi: 10.1007/s40622-019-00225-6.
- [24] A. M. and T. Georgiadis, “Economic clustering of countries in the Asia-Pacific region,” *Electronic Libr.*, vol. 34, no. 1, pp. 1–5, 2004.
- [25] A. A. Nurdin and D. Djuniadi, “Securing audio chat with cryptool-based twofish algorithm,” *J. Soft Comput. Explor.*, pp. 37–43, 2022, doi: 10.52465/josce.v3i1.65.