

Pemanfaatan Data Mining Dalam Penentuan Penyuluhan Penyakit Stunting Menggunakan Partitioning Around Medoids (PAM)

Allsela Meiriza*¹, Endang Lestari², Pacu Putra³, Nabila Rizky Oktadini⁴, Meitiana Audya⁵

^{1,2}Universitas Sriwijaya; Jl. Palembang- Peabumulih Km.32 Inderalaya , (0711) 7072729

³Fakultas Ilmu Komputer, Universitas Sriwijaya

e-mail: *¹allsela@unsri.ac.id, ²endanglestari@unsri.ac.id, ³pacuputra@ilkom.unsri.ac.id,

⁴nabilarizky@unsri.ac.id, ⁵meitianaaudya9@gmail.com

Abstrak

Stunting merupakan keadaan kurang gizi yang bersifat berbahaya pada masa pertumbuhan dari awal kehidupan manusia. Untuk itu sangat diperlukan pencegahan dini, saat ini Pemerintah khususnya di bidang kesehatan seperti Dinas Kesehatan telah membantu dalam melakukan sosialisasi atau penyuluhan stunting. Namun permasalahan penentuan penyuluhan tersebut terdapat kesulitan dalam penentuannya, khususnya pada Seksi Kesehatan Keluarga Dan Gizi Masyarakat Dinas Kesehatan Kota Palembang, dikarenakan banyaknya Puskesmas di setiap Kecamatan di Kota Palembang, maka untuk membantu hal tersebut tujuan penelitian ini adalah menerapkan metode Partitioning Around Medoids dalam membantu penentuan penyuluhan penyakit stunting dengan tepat. Metode yang digunakan yaitu PAM. Hasil penelitian ini yaitu berupa cluster yang terbagi menjadi dua yaitu kelompok 1 dan 2. Cluster 0 diperoleh 9 Puskesmas, sedangkan cluster 1 terdapat 23 Puskesmas, yang menjadi prioritas dalam penentuan penyuluhan ada pada cluster 0. Hasilnya 0.272 artinya pengujian pengelompokan cukup baik dikarenakan hasil sudah mendekati 0 dari pengujian menggunakan Davies Bouldin Index.

Kata Kunci: Data Mining, PAM, Stunting

Utilization of Data Mining in Determining Stunting Disease Counseling Using Partitioning Around Medoids (PAM)

Abstract

Stunting is a state of malnutrition that is dangerous during the growth period from the beginning of human life. For this reason, early prevention is needed, currently the Government, especially in the health sector such as the Health Office, has assisted in socializing or counseling stunting. However, the problem of determining the distribution there are difficulties in determining it, especially in the Family Health and Community Nutrition Section of the Palembang City Health Office, due to the large number of Puskesmas in each Sub-district in Palembang City, so to help this, the purpose of this study is to apply the Partitioning Around Medoids method in helping to determine stunting disease counseling appropriately. The method used is PAM. The results of this study are in the form of clusters that are divided into two, namely groups 1 and 2. Cluster 0 obtained 9 Puskesmas, while cluster 1 there are 23 Puskesmas, the priority in determining counseling is in cluster 0. The result is 0.272 was obtained, meaning that the cluster evaluation was quite good because the value was close to 0 from the test using the Davies Bouldin Index.

Keywords: Data Mining, PAM, Stunting

1. Pendahuluan

Penyakit *stunting* merupakan suatu pemicu yang dapat menghasilkan manusia yang tidak berkualitas, karena akan dapat mempengaruhi kemajuan bangsa, selain itu dapat menyebabkan resiko kesakitan dan kematian, yang dapat menyerang motorik otak sehingga lambat berkembang dan pertumbuhan mental menjadi terhambat[1]. Di Provinsi Sumatera Selatan persentase *stunting* mencapai 32%[2] dimana menurut WHO angka tersebut telah melewati ambang maksimal yang telah ditentukan yaitu 20%. Oleh karena itu sangatlah perlu menjadi perhatian bagi Pemerintah agar dapat menurunkan angka *stunting* di Indonesia. Saat ini Dinas Kesehatan Kota Palembang telah mendukung program tersebut, namun dikarenakan saat ini terdapat 41 Puskesmas termasuk Posyandu yang tersebar di 18 kecamatan Kota Palembang, hal tersebut membuat pihak Seksi Kesehatan Keluarga dan Gizi Masyarakat kesulitan dalam menentukan penyuluhan *stunting*, untuk itu diperlukan suatu cara yang dapat mendukung penentuan keputusan dengan melakukan pengelompokan puskesmas secara efektif.

Maka agar dapat membantu dalam penyelesaian penentuan tersebut diperlukan pemanfaatan data mining dalam bentuk pengklusteran. *Data mining* adalah proses mengekstraksi kumpulan data yang diturunkan secara algoritmik dari statistik atau manajemen basis data[3]. Selain itu, juga merupakan domain dari beberapa ilmu yang menggabungkan pembelajaran mesin, pengenalan pola, statistik, basis data, dan metode visualisasi untuk mengatasi masalah pengambilan informasi dari *database* besar [4] atau pola yang ditemukan dalam ilmu yang dapat dikenali datanya [5]. Kemudian pada *data mining* terdapat *clustering*. Klusterisasi adalah alat yang digunakan untuk mengelompokkan objek ke dalam kluster[6]. *Clustering* sering dilakukan langkah awal dalam melakukan proses *data mining* saat melakukan metode analisis [7]. *Clustering* adalah proses pengelompokan objek yang memiliki kemiripan rendah dengan objek lain dalam kluster yang berbeda[8]. Demikian pula, dalam penelitian [9], suatu kelompok dianggap layak jika memiliki kesamaan yang sama di antara anggota *cluster*, keberagaman yang tinggi antar kelompok, dan menerima saran penyuluhan *stunting* yang ditargetkan. Metode yang digunakan untuk pengelompokan tersebut yaitu *Partitioning Around Medoids* (PAM). Penggunaan metode tersebut dikarenakan dapat digunakan dataset yang tidak banyak dan metode ini juga lebih baik dari metode *K-Means*[10]. Begitu juga menurut penelitian[11] metode PAM lebih dikenal dengan *K-Medoids* lebih baik dibandingkan metode *K-Means* karena mampu mengelompokkan sebaran data anak penyandang disabilitas di Riau berdasarkan nilai validitas. Selain itu, metode ini juga dapat meminimalkan jumlah kesamaan antara setiap objek dan titik referensi yang sesuai [12]. Oleh sebab itu penggunaan PAM bertujuan untuk mengklasifikasikan dan memilah data untuk membuat keputusan penyuluhan *stunting* di Kota Palembang secara tepat.

2. Metode Penelitian

Adapun langkah-langkah yang dilakukan guna mencapai keberhasilan riset ini antara lain:

1. Analisa Data
Pada tahapan ini mengumpulkan data dari pihak Seksi KKG berupa data Puskesmas dan data lainnya, kemudian pencarian studi Pustaka.
2. *Preprocessing* data
Langkah selanjutnya setelah data diperoleh yang berkaitan dengan *stunting*, diolah ke dalam tahapan *data mining*.
3. Penerapan metode PAM
Setelah selesai dari tahap sebelumnya dilakukan perhitungan menggunakan metode PAM.
4. Pengujian *cluster* menggunakan DBI
Setelah diperoleh hasil perhitungan, kemudian dilanjutkan dengan evaluasi menggunakan DBI dengan bantuan *tools rapidminer*.
5. Hasil
Diperoleh hasil pengklusteran Puskesmas yang akan digunakan dalam penentuan penyuluhan penyakit *stunting*.

Atau dapat dilihat dalam alur penelitian di bawah ini (Gambar 1):



Gambar 1 Alur Penelitian

3. Hasil

Data Analysis

Berikut adalah data yang diperoleh dari Dinas Kesehatan. Diperoleh 41 Puskesmas yang ada di Kota Palembang beserta variabel lainnya. Berikut data yang akan digunakan:

Tabel 1 Data Penelitian

Puskesmas	Stunting	Bumil KEK	Tenaga Gizi	Posyandu
Makrayu	20	95	2	10
Gandus	0	37	1	14
1 Ulu	54	58	1	14
4 Ulu	80	86	2	20
7 Ulu	15	3	1	30
Pembina	24	83	2	10
Opi	12	28	2	10
Keramasan	0	64	1	23
Kertapati	25	120	2	10
Karya Jaya	24	22	1	0
Nagaswidak	7	57	1	21
Taman Bacaan	58	70	1	21
Plaju	18	67	1	39
Tegal Binangun	1	11	0	0
Kampus	19	14	2	17
Pakjo	8	24	1	12
Padang Selasa	0	87	1	5
Sungai Baung	0	18	1	19
23 Ilir	15	20	1	14
Merdeka	33	15	1	11
Ariodilah	4	16	1	11
Dempo	49	7	2	4
Talang Ratu	17	7	1	17
Basuki Rahmat	9	20	1	24
Sekip	33	35	1	12
5 Ilir	6	48	1	1
Boom Baru	2	22	1	11
Sabokingking	43	94	1	37
11 Ilir	8	44	1	5
Kenten	25	45	1	25
Bukit Sangkal	27	27	2	13
Kalidoni	17	48	2	15
Sei Selincah	53	60	1	22
Multi Wahana	7	50	2	27
Sako	43	62	1	21
Sematang Borang	59	43	1	28
Sosial	9	48	2	14
Sukarame	8	52	1	5

Talang Betutu	0	28	0	21
Punti Kayu	0	2	1	2
Alang-alang Lebar	0	38	3	19

Preprocessing Data

Selanjutnya dilakukan tahapan dalam *data mining*. Berikut langkah-langkahnya:

1. *Cleaning data*, yaitu dilakukan pembersihan data dari tabel 1. Dengan tidak memasukkan data Puskesmas yang tidak memiliki data *stunting*.
2. Transformasi data, tahap ini sudah dilakukan karena data sudah dalam bentuk angka.
3. Normalisasi data.

Tabel 2 Normalisasi Data

Stunting	Bumil KEK	Tenaga Gizi	Posyandu
0.231	0.786	1.000	0.237
0.667	0.470	0.000	0.342
1.000	0.709	1.000	0.500
0.167	0.000	0.000	0.763
0.282	0.684	1.000	0.526
0.128	0.214	1.000	0.237
0.295	1.000	1.000	0.237
0.064	0.462	0.000	0.526
0.718	0.573	0.000	0.526
0.205	0.547	0.000	1.000
0.218	0.094	1.000	0.421
0.077	0.179	0.000	0.289
0.167	0.145	0.000	0.342
0.397	0.103	0.000	0.263
0.026	0.111	0.000	0.263
0.603	0.034	1.000	0.079
0.192	0.034	0.000	0.421
0.090	0.145	0.000	0.605
0.397	0.274	0.000	0.289
0.051	0.385	0.000	0.000
0.000	0.162	0.000	0.263
0.526	0.778	0.000	0.947
0.077	0.350	0.000	0.105
0.295	0.359	0.000	0.632
0.321	0.205	1.000	0.316
0.192	0.385	1.000	0.368
0.654	0.487	0.000	0.553
0.064	0.402	1.000	0.684
0.526	0.504	0.000	0.526
0.731	0.342	0.000	0.711
0.090	0.385	1.000	0.342
0.077	0.419	0.000	0.105

Perhitungan Metode PAM

Kemudian lakukan perhitungan dengan langkah-langkah dibawah ini:

1. Menentukan *medoid* awal.
2. Tentukan *medoid* awal sebanyak k dari n data secara acak, dalam penentuan k pada penelitian ini langsung di tentukan yaitu sebanyak 2 *cluster*, maka k=2, *medoid* awal yang dipilih yaitu Puskesmas Sosial dan Sukarame.
3. Menghitung jarak data ke *medoid*, dengan menggunakan rumus *Euclidean distance*.

4. Beri tanda jarak terdekat objek ke *medoid* dan jumlah keseluruhan adalah 15.27933654 berdasarkan tabel 3.
5. Tentukan anggota *cluster* terhadap medoid sementara.
6. Lakukan pengulangan medoid untuk mendapatkan perbedaan kedekatan. Memilih objek *medoid* sementara baru, yaitu objek bernomor 0 dan 1, mengurangi jumlah pendekatan kedua dari jumlah *medoid* pertama dan mengurangi hasilnya dengan jumlah pendekatan kedua. Hasil selisihnya adalah 1.659892. Artinya nilai lebih besar dari 0 dan hasil *cluster* berhenti pada tahap 1.

Tabel 3 Hasil Hitung Jarak, Kedekatan dan *Cluster*

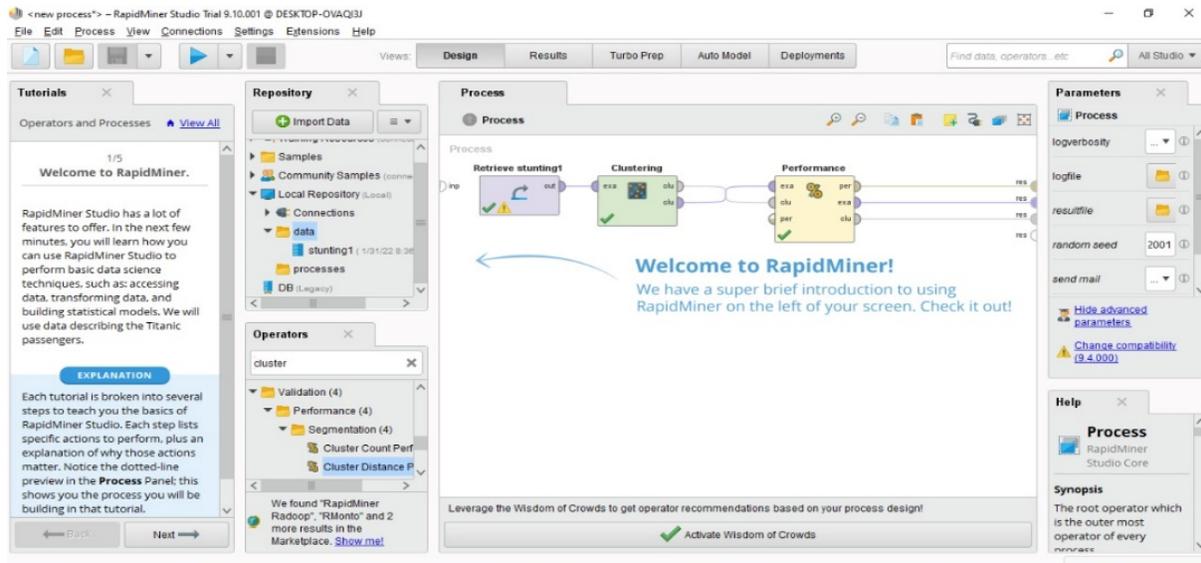
Jarak 1	Jarak 2	Kedekatan	Cluster
0.438565	1.08446	0.438564713	1
1.157646	0.63759	0.637590412	0
0.979277	1.446491	0.979277099	0
1.153747	0.785032	0.785032292	1
0.400504	1.13559	0.400504278	1
0.204402	1.030544	0.204401968	1
0.657158	1.184316	0.65715762	1
1.020053	0.42341	0.423409929	1
1.209931	0.78222	0.782219809	0
1.21347	0.912922	0.9129223	1
0.327287	1.106841	0.327286737	1
1.022258	0.302003	0.302002948	1
1.031111	0.372764	0.372763574	1
1.08649	0.477144	0.47714413	0
1.041704	0.349621	0.349621442	1
0.674564	1.1937	0.67456773	0
1.067498	0.510848	0.510847918	1
1.061379	0.57006	0570060482	1
1.053466	0.397208	0.397207537	1
1.057599	0.113607	0.11360729	1
1.031343	0.310796	0.310795958	0
1.308024	1.019486	1.019485979	1
1.028313	0.068376	0.068376068	1
1.061381	0.572791	0.572790948	1
0.293535	1.071061	0.293534676	1
0.105886	1.041026	0.105886335	1
1.171773	0.733249	0.733248953	0
0.34349	1.155697	0.343490435	1
1.11277	0.621239	0.621239287	0
1.244377	0.894302	0.894301651	0
0	1.028313	0	1
1.028313	0	0	1

Pengujian Kelompok

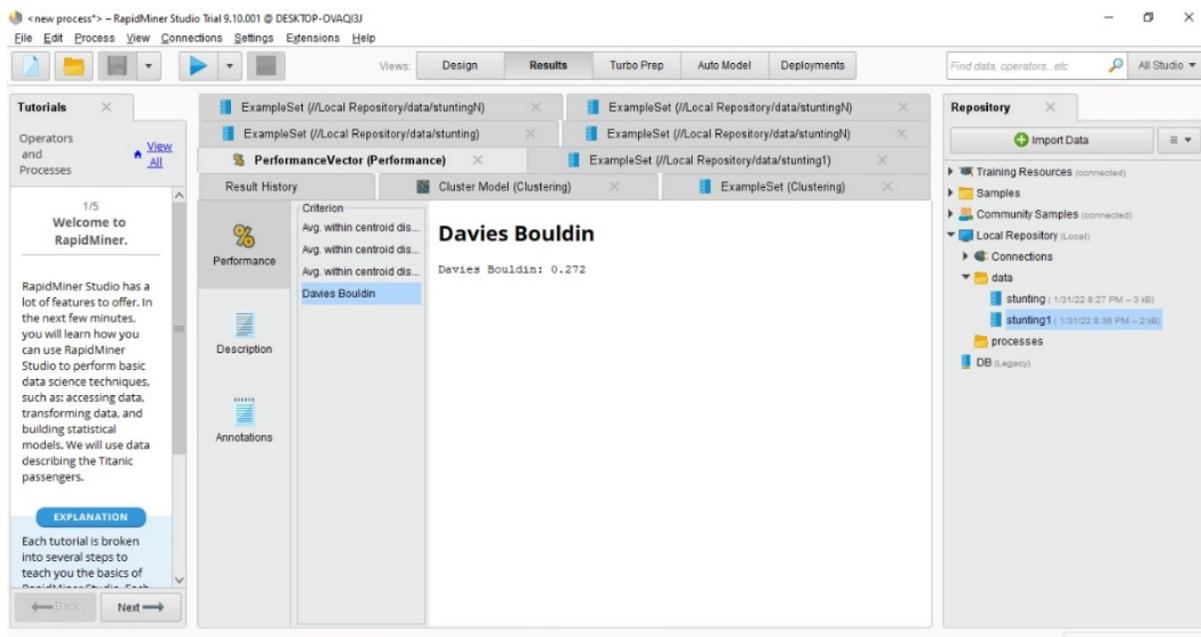
Davies Bouldin Index berfungsi dalam menentukan banyaknya *cluster* yang terbaik agar mengoptimalkan *distance* antara kelompok yang pertama dengan yang lainnya dan pada waktu yang bersamaan mencoba untuk meminimalkan interval antara objek dalam suatu *cluster*. Dalam penelitian ini dilakukan uji secara internal dengan menggunakan informasi apa adanya dari data set. DBI minimum didasarkan pada akuisisi, dan nilai DBI yang diperoleh (non-negatif ≥ 0) menunjukkan *cluster* yang lebih baik dari kelompok K-Medoids yang digunakan. Pengujian ini menggunakan *tools rapidminer* sebagai berikut:

1. Pilih file data set yang telah disiapkan pada *repository*, kemudian drag ke *process*.
2. Drag K-Medoids ke *process*.

3. Pilih *cluster distance performance*.
4. Pilih *normalize* dan *maximize* dalam parameters.
5. Pilih *execution*, maka akan tampil *result* dari data yang kita olah tadi, lalu pilih di tab *performances*.



Gambar 2 Proses pemasukan data, Clustering dan Pengujian



Gambar 3 Hasil Pengujian Davies Bouldin

4. Pembahasan/Kesimpulan

Berdasarkan hasil kedekatan pada tabel 3, maka diperoleh Puskesmas yang termasuk dalam kelompok 0 dan 1 .

Tabel 4. Puskesmas *Cluster* 0 dan 1

<i>Cluster</i> 0	<i>Cluster</i> 1
1 Ulu	Makrayu
4 Ulu	7 Ulu
Taman Bacaan	Nagaswidak
<i>Cluster</i> 0	<i>Cluster</i> 1
Merdeka	Pembina
Dempo	Opi
Sabokingking	Plaju
Sei Selincah	Pakjo
Sako	23 Ilir
Sematang Borang	Ariodilah
	Kertapati
	Kampus
	Talang Ratu
	Bukit Sangkal
	Kalidoni
	Multi Wahana
	Sosial
	Basuki Rahmat
	Sekip
	5 Ilir
	Boom Baru
	11 Ilir
	Kenten
	Sukarame

Berdasarkan tabel 4 di atas terdapat 9 Puskesmas pada *cluster* 0 dan 23 Puskesmas pada *cluster* 1, maka Puskesmas yang diprioritaskan untuk di lakukan penyuluhan adalah Puskesmas yang termasuk dalam *cluster* 0.

Adapun kesimpulan yang diperoleh dari penelitian ini bahwa telah diterapkan metode PAM dalam penentuan penyuluhan penyakit stunting memperoleh dua *cluster* Puskesmas yaitu 0 dan 1, dimana cluster 0 berjumlah 9 Puskesmas, sedangkan cluster 1 berjumlah 23 Puskesmas, dan *cluster* yang diprioritaskan untuk dilakukan penyuluhan adalah *cluster* 0 yaitu Puskesmas 1 Ulu, 4 Ulu, Taman Bacaan, Merdeka, Dempo, Sabokingking, Sei Selincah, Sako, Sematang Borang. Kemudian hasil dari pengujian *cluster* diperoleh angka 0.272, yang berarti *cluster* yang dihasilkan cukup baik, karena nilai DBI mendekati 0.

5. Daftar Pustaka

- [1] M. Mitra, "Permasalahan Anak Pendek (Stunting) dan Intervensi untuk Mencegah Terjadinya Stunting (Suatu Kajian Kepustakaan)," *J. Kesehat. Komunitas*, vol. 2, no. 6, pp. 254–261, 2015.
- [2] K. K. Indonesia, "Warta Kesmas Gizi Seimbang, Prestasi Gemilang," 2019.
- [3] S. M. Andri, Yesi Novaria Kunang, "Implementasi Teknik Data Mining Untuk Memprediksi Tingkat Kelulusan," *Semin. Nas. Inform. 2013 (semnasIF 2013) UPN "Veteran"*, vol. 2013, no. June 2016, pp. 56–63, 2013.
- [4] S. Syahidatul Helma *et al.*, "Clustering pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma K-Means," *Puzzle Res. Data Technol. Fak. Sains dan Teknol.*, vol. 1, no. November, p. 4, 2019.

- [5] D. A. Silitonga, A. P. Windarto, and D. Hartama, "Penerapan Metode K-Medoid pada Pengelompokan Rumah Tangga Dalam Perlakuan Memilah Sampah Menurut Provinsi," *Semin. Nas. Sains Teknol. Inf. SENSASI 2019 ISBN*, pp. 313–318, 2019.
- [6] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. Ilmi R.H.Zer, and D. Hartama, "Analisis algoritma K-Medoids clustering dalam pengelompokan penyebaran Covid-19 di Indonesia," *Jti (Jurnal Teknol. Informasi)*, vol. 4, no. 1, pp. 166–173, 2020.
- [7] Y. H. Chrisnanto and G. Abdillah, "Penerapan Algoritma Partitioning Around Medoids (Pam) Clustering Untuk Melihat Gambaran Umum Kemampuan," in *Seminar Nasional Teknologi Informasi dan Komunikasi*, 2015, vol. 2015, no. Sentika, pp. 444–448.
- [8] E. Okta, N. Satyahadewi, and N. N. Debataraaja, "Penerapan Metode K-Medoids Pada Pengelompokan," vol. 08, no. 4, pp. 813–820, 2019.
- [9] H. Februriyanti *et al.*, "Algoritma Partitioning Around Medoids (PAM) Clustering untuk Melihat Gambaran Umum Skripsi Mahasiswa," vol. 21, no. 1, pp. 25–31, 2016.
- [10] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, p. 64, 2018.
- [11] M. Baru and T. Di, "Implementasi Metode K-Medoids Clustering Untuk Mengetahui Pola Pemilihan Program Studi," *J. Terap. Sains Teknol.*, vol. 1, no. 3, pp. 54–69, 2019.
- [12] Ogunde and Ajibade, "A Data Mining System for Predicting University Students' Graduation Grades Using ID3 Decision Tree Algorithm Ogunde A. O 1 . and Ajibade D. A 1 .," *Comput. Sci. Inf. Technol.*, vol. 2, no. 1, pp. 21–46, 2014.