

Rekomendasi Grup Pada Website Alumni Teknik Informatika Universitas Muhammadiyah Malang

Hermansyah Adi Saputra^{*1}, Galih Wasis Wicaksono², Yufis Azhar³

^{1,2,3}Universitas Muhammadiyah Malang

e-mail: adisaputra280496@gmail.com^{*1}, galih.w.w@umm.ac.id², yufis@umm.ac.id³

Abstrak

Belakangan ini hampir seluruh universitas yang ada di Indonesia memiliki sistem informasi alumninya sendiri-sendiri. Sistem informasi alumni mampu memberikan informasi tentang kondisi alumninya setelah menyelesaikan masa perkuliahannya. Alumni merupakan aktor yang berperan penting dalam pendidikan. Saat ini jurusan Informatika Fakultas Teknik Universitas Muhammadiyah Malang telah memiliki website alumni. Permasalahannya belum adanya sistem yang memberikan alumni rekomendasi grup pada sistem, sehingga para alumni mampu saling bertukar informasi di dalamnya. Dengan adanya data alumni dan juga di dukung dengan adanya tracer study, dapat di bentuk suatu rekomendasi grup dari data tracer study. K-medoid adalah metode pengelompokan data ke dalam sejumlah cluster tanpa adanya struktur hirarki antara satu dengan yang lainnya. Algoritma k-medoid memiliki nilai coefficient yang lebih tinggi di dibandingkan dengan k-means dalam penelitian ini. Yang mana k-medoid mendapatkan nilai rata-rata Silhouette Score 0.7325888099 dalam pengujian dengan jumlah cluster 5 dan perulangan sebanyak 10 kali. Jika dibandingkan dengan k-means yang hanya memiliki nilai rata-rata Silhouette Score 0.6872873866.

Kata kunci: K-Medoids, Silhouette Score, Clustering, Tracer Study

Abstract

Lately, Almost all universities in Indonesia have their own alumni information systems. The alumni information system is able to provide information about the condition of its alumni after collage graduation. Alumni are actors who play important role in education. Currently, the Department of Informatics, Faculty of Engineering, University of Muhammadiyah Malang has an alumni website. The problem is the absence of system that gives alumni group recommendation on the system, so that alumni are able to exchange information in this website. With the alumni data and also supported by the existence of a tracer study, it can be formed as group recommendation from the data tracer study. Clustering is one of tools in data mining that aims to group object into clusters. K-medoid is a method of grouping data into a number of clusters without hierarchical structure from one another. The k-medoid algorithm has higher coefficient value compared to k-means in this study. This K-medoid gets an average value of Silhouette Score 0.7325888099 in testing with the number of clusters 5 and repetitions 10 times. When compared with k-means which only has an average value of Silhouette Score 0.6872873866.

Keywords: K-Medoids, Silhouette Score, Clustering, Tracer Study

1. Pendahuluan

Belakangan ini hampir seluruh universitas yang ada di Indonesia memiliki sistem informasi alumninya sendiri-sendiri. Sistem informasi alumni mampu memberikan informasi tentang kondisi alumninya setelah menyelesaikan masa perkuliahannya. Alumni merupakan aktor yang berperan penting dalam pendidikan. Karena dari kesuksesan yang di raih oleh alumni mampu memberikan dampak yang positif kepada universitas itu sendiri. Dan juga alumni menjadi sebuah hasil dari sukses atau gagal nya pengajar dalam mendidik setiap mahasiswanya.

Di tahun 2018 saat ini universitas muhammadiyah malang memiliki mahasiswa yang berjumlah 34.123, dan alumni pada tahun 2017 berjumlah 5.225, data tersebut di dapat penulis melalui website resmi universitas muhammadiyah malang (<http://www.umm.ac.id/>). Dengan alumni sebanyak itu seharusnya mampu menjadi jalan pembuka peluang-peluang yang bagus bagi calon alumni. Saat ini jurusan Informatika Fakultas Teknik Universitas Muhammadiyah

Malang telah memiliki website alumni. Permasalahannya belum adanya sistem yang memberikan alumni rekomendasi grup pada sistem, sehingga para alumni mampu saling bertukar informasi didalamnya.

Clustering atau klusterisasi adalah salah satu alat bantu pada data mining yang bertujuan mengelompokkan objek-objek ke dalam cluster-cluster[1]. Dalam data mining ada dua jenis metode clustering yang digunakan dalam pengelompokan data, yaitu hierarchical clustering dan non-hierarchical clustering[2]. Algoritma Partitioning Around Medoids (Pam) atau yang lebih sering dikenal dengan algoritma k-Medoids adalah sebuah algoritma yang merepresentasikan cluster yang dibentuk menggunakan medoids[3].

Penelitian ini bertujuan untuk membangun sistem rekomendasi grup menggunakan k-Medoids yang nantinya diharapkan mampu memberikan rekomendasi yang berguna bagi para alumni. Dan juga dengan adanya penelitian ini mampu memberikan dampak yang positif bagi para alumni.

2. Metode Penelitian

Metode yang digunakan dalam penelitian ini terdiri dari beberapa tahapan yaitu dimulai dari tahap identifikasi masalah, tahap pengumpulan data, tahap perancangan program, tahap implementasi, dan perancangan laporan.

Yang mana pada penelitian ini perhitungan jarak menggunakan euclidean distance. Metode euclidean distance merupakan metode klasifikasi antar tetangga data terdekat dengan cara menghitung jarak antara dua buah obyek. Adapun rumus dalam perhitungan jarak menggunakan euclidean distance sebagai berikut[4]:

$$d_e = \sqrt{\sum_{k=1}^m (f d_{i,k} - k_j)^2} \quad (1)$$

Silhouette Coefficient adalah metode yang digunakan untuk melihat kualitas dan kekuatan dari sebuah cluster, seberapa baik suatu objek yang ditempatkan dalam suatu cluster. Metode *silhouette coefficient* ini merupakan gabungan dari metode cohesion dan juga separation[5]. *Silhouette Coefficient* dibangun di bawah berguna ketika perkiraan berada pada skala rasio (seperti dalam kasus jarak Euclidean)[6].

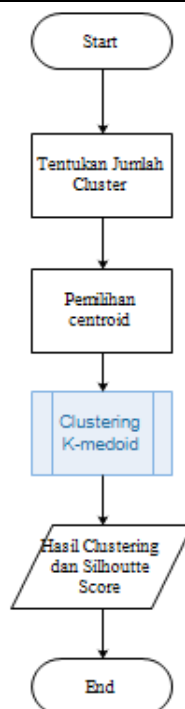
$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2)$$

Nilai dari *silhouette coefficient* dapat dikatakan hasil yang baik jika nilai yang dihasilkan mendekati nilai 1. Semakin nilai *silhouette coefficient* mendekati nilai 1 maka semakin bagus cluster yang dihasilkan[7].

Metode k-Medoids dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw pada tahun 1987. Algoritma k-medoid sering disebut juga algoritma Partitioning Around Medoid (Pam). Metode k-Medoid memiliki kesamaan dengan metode k-means yaitu sama-sama termasuk metode partitioning. Metode partitioning merupakan metode pengelompokan data ke dalam sejumlah cluster tanpa adanya struktur hirarki antara satu dengan yang lainnya. Metode k-medoid memiliki keunggulan dibandingkan dengan metode k-means. k-Medoid memiliki kinerja yang lebih optimal jika jumlah data yang digunakan berjumlah sedikit. Algoritma ini menggunakan objek pada kumpulan objek untuk mewakili sebuah cluster. Objek yang terpilih untuk mewakili sebuah cluster disebut medoid[8].

2.1 Perancangan Program

Perancangan program yang meliputi tahapan berupa analisis, desain, coding, dan pengujian. Yang mana nantinya pada tahap ini penulis menggunakan pemrograman web dengan framework codeigniter. Codeigniter merupakan sebuah framework yang dibuat dengan menggunakan bahasa PHP, yang dapat digunakan untuk pengembangan web secara cepat[9]. Codeigniter itu sendiri adalah framework php yang berjalan pada php 4 dan php 5[10].



Gambar 1. Alur Sistem

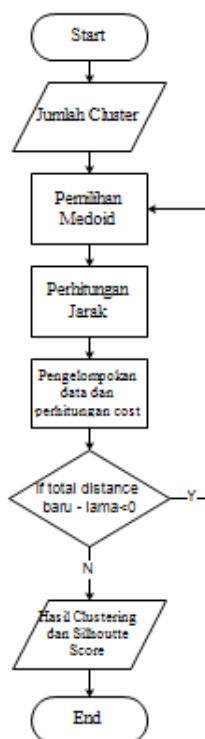
Penjelasan gambar 1 :

- Start – Masuk ke menu Clustering
- Tentukan jumlah cluster yang di inginkan pada data alumni
- Pemilihan centroid pada sistem sesuai dengan jumlah cluster yang sudah di tentukan sebelumnya
- Proses cluster berjalan, alur kerja clustering seperti pada Gambar 1
- Bila proses cluster sudah selesai maka hasil cluster akan di tampilkan beserta score dari silhouette

2.2 Implementasi Algoritma

Pada tahap ini dilakukan implementasi dengan langkah-langkah sebagai berikut [7]:

1. Inisialisasi pusat cluster sebanyak k (jumlah cluster)
2. Alokasikan setiap data (objek) ke cluster terdekat menggunakan persamaan ukuran jarak *Euclidian Distance*.
3. Pilih secara acak objek pada masing-masing cluster sebagai kandidat medoid baru.
4. Hitung jarak setiap objek yang berada pada masing-masing cluster dengan kandidat medoid baru.
5. Hitung total simpangan (S) dengan menghitung nilai total distance baru – total distance lama. Jika $S < 0$, maka tukar objek dengan data cluster untuk membentuk sekumpulan k objek baru sebagai medoid.
6. Ulangi mulai dari langkah 3 sampai 5 hingga tidak terjadi perubahan medoid.



Gambar 2. Flowchart K-Medoid

3. Hasil Penelitian dan Pembahasan

3.1 Data Alumni

Tabel 1 Data Alumni

No	Tahun Lulus	NIM	Nama	Jenis Kelamin	Tanggal Lahir	Asal
1	2009	50001	Alumni 1	L	1987-01-01	Malang
2	2009	50002	Alumni 2	P	1987-01-02	Malang
3	2009	50003	Alumni 3	L	1987-01-03	Malang
4	2009	50004	Alumni 4	L	1987-01-04	Malang
5	2009	50005	Alumni 5	L	1987-01-05	Malang
6	2009	50006	Alumni 6	L	1987-01-06	Malang
7	2009	50007	Alumni 7	L	1987-01-07	Malang
8	2009	50008	Alumni 8	L	1987-01-08	Malang
9	2009	50009	Alumni 9	P	1987-01-09	Malang
10	2009	50010	Alumni 10	P	1987-01-10	Malang

Pada penelitian ini menggunakan data sebanyak 300 data, yang mana data tersebut merupakan alumni angkatan 2005 sampai 2006. Dari 300 data ini akan dilakukan proses clusterisasi berdasarkan kriteria yang sudah di tentukan dan dilakukan pengujian untuk mencari nilai cluster terbaik dengan menggunakan *silhouette coefficient*. Yang mana bila *silhouette score* semakin mendekati nilai 1 maka kualitas *cluster* yang dihasilkan akan lebih baik.

3.2 Data Kriteria

Tabel 2 Data Kriteria Alumni

No	Nama	Data
1	Alumni 1	1,3,3,2,3,1,3,3,3,1,1,3,2,2,1,2
2	Alumni 2	2,5,6,4,4,2,1,2,3,4,2,1,4,3,3,4
3	Alumni 3	1,2,1,2,2,1,1,3,2,1,1,1,2,2,2,1
4	Alumni 4	2,2,4,1,1,1,1,1,1,1,1,4,3,2,4,3
5	Alumni 5	1,3,4,1,1,1,1,3,1,1,1,2,2,3,2,2
6	Alumni 6	1,1,2,2,3,1,2,2,2,2,1,1,1,1,1
7	Alumni 7	1,5,4,2,2,1,1,3,4,3,2,2,3,3,1,3

8	Alumni 8	2,4,5,2,2,1,2,2,3,1,1,4,2,1,2,2
9	Alumni 9	1,1,1,2,1,1,1,1,1,2,1,2,3,2,2,2
10	Alumni 10	2,2,2,2,2,2,2,2,2,1,1,1,2,2,2,4

Pada tabel diatas merupakan hasil dari jawaban kriteria yang sudah di beri pembobotan sebelumnya. Sehingga output yang dihasilkan berupa angka dari pembobotan. Yang mana dari data ini nantinya akan di lakukan proses *clustering*. Data yang digunakan diatas merupakan data *dummy* yang mana data tersebut merupakan data acak. Yang mana data acak ini hanya untuk penelitian berfungsi atau tidaknya rekomendasi grup nantinya sebelum di *publish*.

3.3 Data Cluster

Tabel 3 Data Hasil CLuster

No	Nama	Cluster
1	Alumni 1	2
2	Alumni 2	1
3	Alumni 3	2
4	Alumni 4	2
5	Alumni 5	2
6	Alumni 6	2
7	Alumni 7	2
8	Alumni 8	1
9	Alumni 9	2
10	Alumni 10	2

Pada tabel 3 merupakan hasil cluster yang di dapat oleh setiap alumni dan akan di simpan ke dalam database yang mana cluster tersebut berfungsi untuk memberikan rekomendasi grup pada setiap alumni sesuai dengan clusternya.

3.5 Pembuatan Grup

The screenshot shows a web interface for creating a group. At the top, there is a text input field labeled 'Group Baru'. Below it, there is a section titled 'Pilih Anggota:' which includes a dropdown menu currently set to '10' and a search box labeled 'Search...'. Underneath, there is a table with columns 'No', 'Pertanyaan', and 'Aksi'. The table lists 8 members with their names and checkboxes in the 'Aksi' column. Members 5 and 7 have their checkboxes checked.

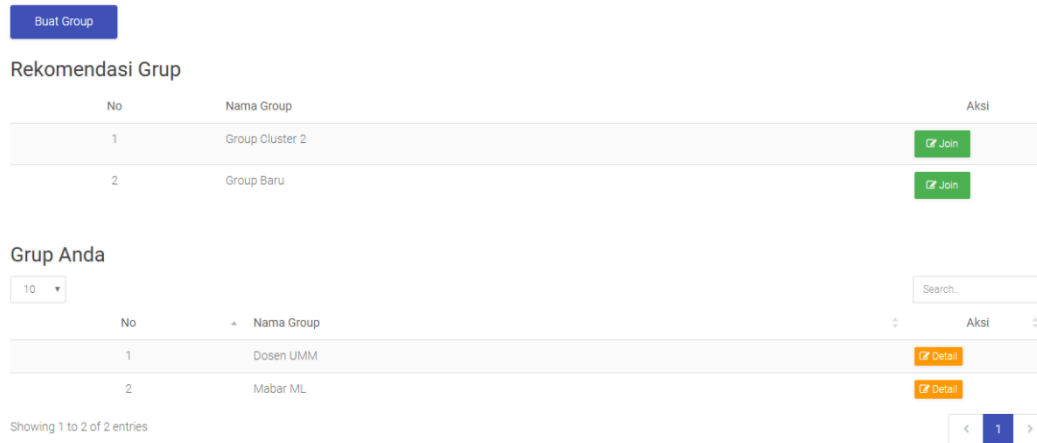
No	Pertanyaan	Aksi
1	Wahyu Andhyka Kusuma	<input type="checkbox"/>
2	Andarta Fardhanul Khoir	<input type="checkbox"/>
3	Galih Wasis Wicaksono	<input type="checkbox"/>
4	Muhammad Tulodo Andarbeni	<input type="checkbox"/>
5	Luqman Hakim	<input checked="" type="checkbox"/>
6	Hardianto Wibowo	<input type="checkbox"/>
7	Evi Dwi Wahyuni	<input checked="" type="checkbox"/>
8	Pepy Maya Fitriyani	<input type="checkbox"/>

Gambar 3. Pembuatan Grup

Pada Gambar diatas mahasiswa melakukan pembuatan grup yang memiliki anggota 3 orang yaitu: Fahmi samia selaku pembuat dan juga anggota, Luqman Hakim serta anggota prima Evi Dwi Wahyuni. Yang mana nantinya grup yang sudah di buat ini akan di rekomendasikan kepada mahasiswa yang memiliki cluster yang sama dengan cluster pembuat grup

3.6 Hasil Rekomendasi

Pada Gambar 6 dapat dilihat bahwa grup yang sudah di buat akan di rekomendasikan kepada alumni yang memiliki cluster yang sama. Maka grup dari cluster lain tidak akan tampil pada rekomendasi grup. Dari rekomendasi itupun alumni dapat bergabung dengan grup yang di rekomendasikan.



Gambar 4. Rekomendasi

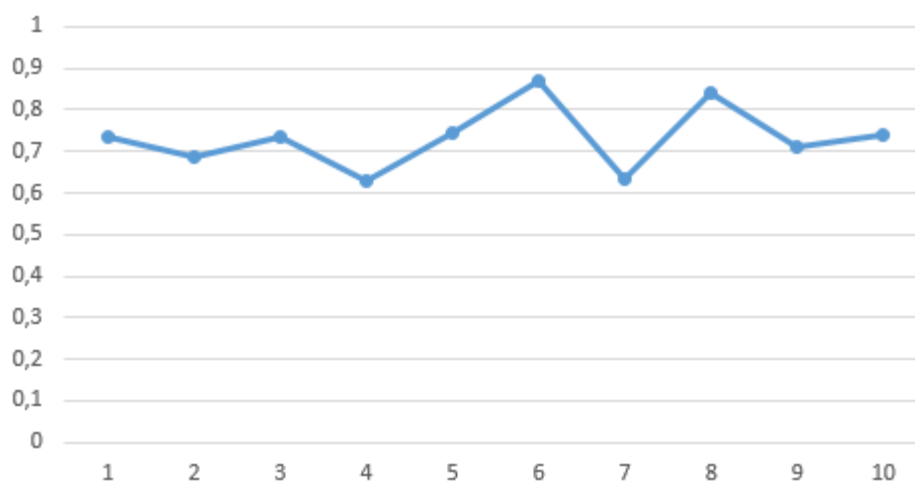
3.7 Pengujian

Pada pengujian ini dilakukan perhitungan nilai *Silhouette Score* dengan menggunakan pergantian medoid sebanyak 10 kali dengan jumlah cluster 5, yang mana data acak yang digunakan pada K-medoid dan K-mean adalah sama. Selanjutnya akan di ambil rata-rata hasil dari *Silhouette Score* dan akan dibandingkan hasil rata-ratanya.

Tabel 4 Hasil *Silhouette Score* K-medoid

Data Uji	Jumlah Cluster	Data Acak Ke	<i>Silhouette Score</i>
300	5	1	0.7356700685562
		2	0.684313856669
		3	0.73661192980268
		4	0.63019703191096
		5	0.74599310590409
		6	0.87027599298055
		7	0.63188912271229
		8	0.84094524717836
		9	0.70884692180482
		10	0.74113745357973

Silhouette Score K-medoid

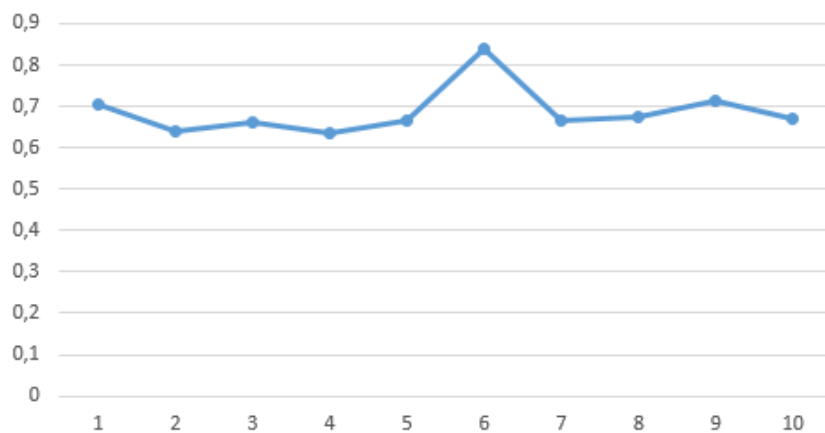


Gambar 5 Grafik *Silhouette Score* K-medoid

Tabel 5 Hasil Silhouette Score K-mean

Data Uji	Jumlah Cluster	Data Acak Ke	Silhouette Score
300	5	1	0.70570298419394
		2	0.64052231507391
		3	0.66329438750793
		4	0.63537322538457
		5	0.66454522704668
		6	0.83758947820878
		7	0.66779227924359
		8	0.67373334719126
		9	0.71379020394966
		10	0.67053041807394

Silhouette Score K-mean

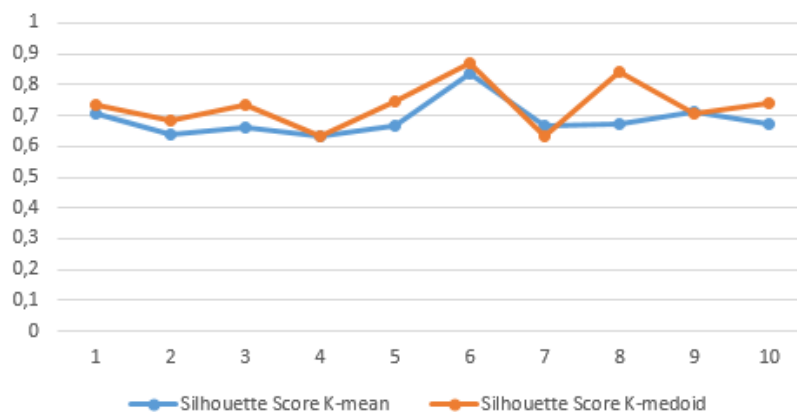


Gambar 6 Grafik Silhouette Score K-mean

Tabel 6 Perbandingan Hasil

Perbandingan	Rata-rata
K-medoids	0.7325888099
K-means	0.6872873866

Grafik Silhouette Score



Gambar 7 Grafik Perbandingan Silhouette Score

Dari hasil pengujian di atas dapat kita lihat bahwa hasil rata-rata *Silhouette Score* dari k-medoid lebih tinggi di bandingkan dengan k-means. Hal itu dikarenakan k-medoid lebih unggul dalam pengolahan data dalam skala kecil dan juga k-medoid lebih handal di bandingkan dengan k-means ketika ada data noise dan pencilan karena k-medoids tidak terlalu dipengaruhi oleh data

pencilan atau data ekstrem lainnya dibandingkan dengan k-means. Dengan melakukan 10 kali percobaan menggunakan 5 cluster didapatkan rata-rata k-medoid yang lebih tinggi yaitu 0.7325888099 di bandingkan dengan k-mean yang hanya 0.6872873866

4. Kesimpulan

Setelah penelitian dilakukan didapatkan beberapa kesimpulan diantaranya sebagai berikut :

1. Algoritma K-medoid dapat diaplikasikan kedalam clustering data mahasiswa sehingga di dapatkan rekomendasi grup sesuai data kuesioner.
2. Algoritma k-medoid memiliki nilai coefficient yang lebih tinggi di bandingkan dengan k-means dalam penelitian ini. Yang mana k-medoid mendapatkan nilai rata-rata *Silhouette Score* 0.7325888099 dalam pengujian dengan jumlah cluster 5 dan perulangan sebanyak 10 kali. Jika dibandingkan dengan k-means yang hanya memiliki nilai rata-rata *Silhouette Score* 0.6872873866.

Saran yang diberikan penulis untuk pengembangan penelitian selanjutnya yaitu :

1. Pengembangan selanjutnya dapat mengimplementasi rekomendasi pertemanan dari hasil clustering atau rekomendasi teman dari grup yang diikuti.
2. Pengembangan selanjutnya juga bisa menambah fitur yang sekira dapat di implementasikan kedalam penelitian ini.
3. Penelitian selanjutnya coba menggunakan algoritma lain untuk membandingkan *coefficient* sehingga didapatkan algoritma apa yang lebih cocok diterapkan pada penelitian ini.

Referensi

- [1] S. Defiyanti, M. Jajuli, and N. Rohmawati, "Optimalisasi K-MEDOID dalam Pengklasteran Mahasiswa Pelamar Beasiswa dengan CUBIC CLUSTERING CRITERION," *J. Teknol. dan Sist. Inf.*, vol. 3, no. 1, pp. 211–218, 2017.
- [2] H. Sulastri and A. I. Gufroni, "Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia," *J. Teknol. dan Sist. Inf.*, vol. 03, no. 02, pp. 299–305, 2017.
- [3] Y. H. Chrisnanto and G. Abdillah, "Penerapan Algoritma Partitioning Around Medoids (PAM) Clustering untuk Melihat Gambaran Umum Kemampuan Akademik Mahasiswa," *J. Semin. Nas. Teknol. Inf. dan Komun. 2015*, vol. 2015, no. Sentika, pp. 444–448, 2015.
- [4] R. Wulanningrum, S. Kom, A. Rachmad, and S. T. Mt, "Pengenalan Rumput Laut Menggunakan Euclidean Distance Berbasis," vol. 2012, no. Snati, pp. 15–16, 2012.
- [5] M. Anggara, H. Sujiani, and N. Helfi, "Pemilihan Distance Measure Pada K-Means Clustering Untuk Pengelompokkan Member Di Alvaro Fitness," *J. Sist. dan Teknol. Inf.*, vol. 1, no. 1, pp. 1–6, 2016.
- [6] J. G. Pearce, Z. Shaar, and R. E. Crosbie, "Scattering of energetic ions by solids — a simulation," *Simulation*, vol. 29, no. 4, pp. 97–104, 1977.
- [7] D. F. Pramesti, M. T. Furqon, and C. Dewi, "Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan / Lahan Berdasarkan Persebaran Titik Panas (Hotspot)," *Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 9, pp. 723–732, 2017.
- [8] A. Y. Rofiqi, "Clustering Berita Olahraga Berbahasa Indonesia Menggunakan Metode K-Medoids Bersyarat," vol. 6, no. 1, 2017.
- [9] L. Afuan, J. Soeparno, K. Mipa, and U. Karangwangkal, "dalam Pengembangan Sistem Informasi Pendataan Laporan Kerja Praktek Mahasiswa Program Studi Teknik Informatika Unsoed (Codeigniter Framework Used in Information System Development in Informatics Engineering Program Study of Unsoed)," vol. I, pp. 39–44, 2010.
- [10] D. Rosmala, M. Ichwan, and M. I. Gandalisha, "Komparasi Framework MVC (Codeigniter, Dan Cakephp) Pada Aplikasi Berbasis Web," *J. Inform.*, vol. 2, no. 8, pp. 22–30, 2011.