

Algoritma K-Means Dengan Menggunakan Metode Pengukuran Jarak dan Densitas Untuk Peningkatan Strategi Promosi Penerimaan Mahasiswa Baru (PMB)

Riana Safitri

STMIK Widya Utama
rianasafitri07@gmail.com.

Abstract In a company the activity of promoting a product or service is the most needed aspect to support success, so promotion is one way for companies or organizations to develop, maintain, and advance the business that is being managed. In order for the promotion to be carried out in accordance with the expected target, it is necessary to establish market segmentation in which there are many targeted targets, because the targeted target is very dispersed and scattered and varied in the demands of their needs and desires. Data mining is an activity in the database process to find certain previously unknown / hidden patterns that are useful and meaningful and can be processed. One algorithm that can be used for clustering is the k-means algorithm. In order to maximize the k-means algorithm, we use distance and density measurement methods. Grouping data using k-means is done by calculating the closest distance from a data to a centroid point. In this study, the method of calculating the distance on k-means between Euclidean and density will be combined. The test will be carried out using the execution time and the bouldin index davies. From the tests that have been done, the calculation of euclidean distance and density has the most efficient accumulation of time and the value of Davies Index.

Keywords—3-5 Promotion, K-means, Euclidean, Densitas

1. PENDAHULUAN

Kegiatan promosi merupakan kegiatan yang paling penting dalam kegiatan. Agar promosi yang dilakukan sesuai dengan target yang diharapkan maka perlu adanya penetapan segmentasi pasar yang didalamnya terdapat banyak target dituju, karena target yang dituju sangat berpecah dan tersebar serta bervariasi dalam tuntutan kebutuhan dan keinginannya [2].

Salah satu upaya promosi yang dilakukan diantaranya dalam hal Penerimaan Mahasiswa Baru (PMB) merupakan kegiatan yang rutin dilakukan setiap 1 tahun sekali di STMIK Widya Utama. Dari banyak data yang diperoleh data mahasiswa yang diolah untuk dijadikan sebagai dataset strategi promosi STMIK Widya Utama untuk tahun berikutnya.

Data klustering merupakan salah satu hasil dari metode data mining yang bersifat tanpa label atau pembelajaran yang tak terarah (*unsupervised learning*). Ada dua jenis data klustering yang sering

dipergunakan dalam proses pengelompokan data yaitu hirarki data klustering dan non hirarki data klustering.

K-means pada dasarnya adalah metode partisi yang diterapkan untuk menganalisis data dan memperlakukan pengamatan data sebagai objek berdasarkan lokasi dan jarak antara berbagai titik data input. Mempartisi objek menjadi kelompok yang saling eksklusif (K) hingga sedemikian rupa sehingga benda-benda di dalam setiap kluster tetap sedekat mungkin satu sama lain namun sejauh mungkin dari benda-benda di kelompok lain [10]. Algoritma nilai K-means adalah algoritma mudah dan cepat namun belum ada kepastian untuk menentukan nilai K (jumlah *cluster* untuk menghasilkan) ada yang, dan sensitif terhadap nilai awal, untuk nilai awal yang berbeda, sehingga ada kelompok yang berbeda yang dihasilkan [3].

Dataset yang digunakan adalah menggunakan dataset data penerimaan mahasiswa STMIK Widya Utama dari tahun 2014-2016 dengan menggunakan 3 variabel yaitu variabel kecamatan daerah domisili mahasiswa, program studi dan biaya USPI (pengembangan yang dikenakan).

Berdasarkan latar belakang yang sudah disebutkan sebelumnya, dalam penelitian ini akan dirumuskan masalah sebagai berikut:

1. Belum adanya pengelompokan data untuk mendukung keputusan strategi promosi penerimaan mahasiswa baru kedalam 3 status daerah promosi.
2. Belum adanya metode pengelompokan menggunakan algoritma k-means dengan perhitungan jarak paling baik untuk mengelompokkan data mahasiswa baru.

Tujuan Penelitian adalah mengelompokkan data kedalam 3 status daerah promosi (daerah dengan strategi promosi baik, daerah dengan strategi promosi cukup, dan daerah dengan strategi promosi kurang) menggunakan metode pengukuran jarak dan densitas berdasarkan data mahasiswa tahun 2014, 2015, 2016.

Berdasarkan tujuan penelitian, maka manfaat dari penelitian ini adalah:

1. Manfaat praktis dari hasil penelitian ini adalah dapat digunakan sebagai bahan pertimbangan oleh pengelola dalam memilih maupun menyimpulkan suatu titik pemasangan promosi yang strategis.

2. Manfaat teoritis dari penelitian ini yaitu diharapkan dapat memberikan sumbangan pengembangan teori berkaitan dengan penerapan algoritma K-means dengan menggabungkan metode pengukuran jarak dan densitas untuk memperoleh *centroid* awal terbaik.

2. METODE PENELITIAN

2.1 Pengumpulan Dta

Dalam Penelitian ini akan menggunakan instrumental peralatan yang meliputi peralatan perangkat lunak (*software*) dan peralatan perangkat keras (*hardware*). Masing-masing kebutuhan dalam penelitian ini adalah sebagai berikut :

2.1.1 Perangkat Lunak (*software*)

Dalam penelitian ini membutuhkan peralatan perangkat lunak (*software*) agar implementasi metode dan pengujian dapat dilaksanakan:

- a. Sistem Operasi Untuk mendukung penelitian, dapat menggunakan sistem operasi windows 7, windows 8 ataupun windows 10 dengan versi 64 bit.
- b. RapidMiner Dalam mengimplementasikan algoritma K-means clustering yang akan digunakan software RapidMiner 5.3.
- c. Microsoft Excel 2013 Untuk perhitungan jarak menggunakan density dan dan pengujian model yang diusulkan akan menggunakan software Microsoft Excel 2013.

2.1.2 Jenis Data

Data yang digunakan dalam penelitian ini merupakan data sekunder. Data mahasiswa baru STMIK Widya Utama angkatan 2014, 2015, 2016 yang diperoleh dari STMIK Widya Utama yang berjumlah 27 kecamatan di daerah Banyumas. Dataset variabel yang digunakan memiliki 3 atribut. Atribut-atribut tersebut adalah:

- 1. Nama kecamatan domosili dari mahasiswa baru
- 2. Program studi terdiri dari :
 - a. DIII Komputerisasi Akuntansi
 - b. SI teknik Informatika
- 3. Biaya USPI (uang Gedung)
 - a. Gelombang 1
 - b. Gelombang 2
 - c. Gelombang 3

2.2 Metode yang diusulkan

Metode yang diusulkan adalah sebagai berikut :

Perhitungan densitas pada K-means sebagai berikut :

- 1. Diketahui ($x_{11}, x_{12}, x_{13}, x_{14}, x_{15}$), ($x_{21}, x_{22}, x_{23}, x_{24}, x_{25}$), $x_{1n}, x_{2n}, x_{3n}, x_{4n}, x_{5n}$ dimana adalah x_1 program studi D3 komputerisasi akuntansi, x_2 adalah program studi S1 teknik informatika, x_3 adalah gelombang 1, x_4 adalah gelombang 2 dan x_5 adalah gelombang 3.

- 2. Tentukan nilai k dan inialisasikan, k=3 k merupakan jumlah kluster yang mengelompokkan menjadi daerah dengan promosi baik, daerah dengan promosi cukup dan daerah dengan promosi kurang.
- 3. Menentukan nilai tinggi (daerah dengan promosi baik), nilai sedang (daerah dengan promosi cukup), dan nilai kecil (daerah dengan promosi kurang) nilai tersebut digunakan untuk inialisasi dalam menentukan centroid awal (C1, C2, C3).
- 4. Hitung jarak menggunakan data berdasarkan centroid..

$$D(x,y) = ||x-y||^2 = \sqrt{\sum_{j=1}^n ||x - y||^2}$$

- 5. Cluster-kan entitas sebagai berikut : Untuk i=1 s/d nhitung :

$$C_1 = D(x_I, c_1) = ||x_I - c_1||^2 = \sqrt{\sum_{I=1}^n ||x_{II} - c_{1I}||^2}$$

$$C_2 = D(x_I, c_2) = ||x_I - c_2||^2 = \sqrt{\sum_{I=1}^n ||x_{II} - c_{2I}||^2}$$

$$C_3 = D(x_I, c_3) = ||x_I - c_3||^2 = \sqrt{\sum_{I=1}^n ||x_{II} - c_{3I}||^2}$$

Selanjutnya adalah membandingkan nilai yang dimiliki oleh masing-masing data berdasarkan *centroid*-nya.

$$C_{min-i} = \min (C_1, C_2, C_3)$$

- 6. Hitung Centroid Baru

$$C_1(x_I, x_2) \text{ rerata } C_1$$

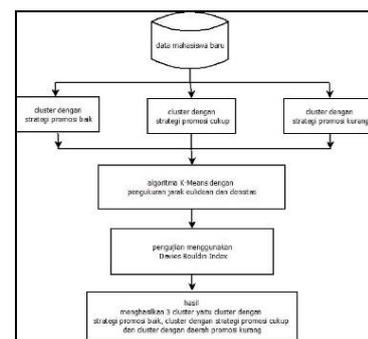
$$C_2(x_I, x_2) \text{ rerata } C_2$$

$$C_3(x_I, x_2) \text{ rerata } C_3$$

$$x = \frac{x_1 + x_2 + \dots + x_n}{n}$$

- 7. Jika $Centroid_{baru} = Centroid_{lama}$ maka proses clustering dengan menggunakan algoritma K-means selesai, jika tidak, ulangi langkah 6.

Dan digambarkan sebagai berikut :



Gambar 3.2. Metode yang diusulkan

Berikut langkah-langkah perhitungan metode algoritma K-means di gabungkan dengan metode perhitungan jarak dan densitas dapat dijelaskan sebagai berikut:

- 1. Input: dataset mahasiswa D = {d1,d2,...,dn} dan k output: k kluster
- 2. Hitung jarak antara setiap pasangan dua objek data yang berada dalam dataset D menggunakan *Euclidean Distance*

$$d(x_i, x_j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{im} - x_{jm})^2} \quad (3)$$

Keterangan :

- d : distance
- i : status cluster ke i
- j : indicator ke-j
- m : banyaknya data

3. Hitung rata-rata jarak.

$$\text{MeanDist} = \frac{1}{c^2} \sum d(x_i, x_j) \quad (4)$$

MeanDist adalah rata-rata jarak dari persamaan

3.

Dimana c merupakan kombinasi pasangan jarak

4. Hitung nilai parameter densitas seluruh objek data yang berada dalam dataset D.

$$\text{Dens}(x_i) = \sum_{j=1}^n u(\text{MeanDist} - d(x_i, x_j)) \quad (5)$$

dimana u(z) merupakan sebuah fungsi bernilai 1 jika z lebih besar dari 0 dan bernilai 0 jika sebaliknya.

5. Hitung nilai parameter rata-rata dari densitas dataset D.

6. Setelah rata-rata dihitung kemudian, maka akan muncul data yang terisolasi dan hapus data yang tidak dipakai dari D dan menghasilkan grup A yang memiliki nilai parameter densitas tertinggi.

$$\text{Dens}(x_i) < \alpha \times \text{MeanDens}(D) \quad (6)$$

dimana α berada pada rentang 0 - 1 dalam menentukan data terisolasi.

7. Pilih objek data dengan nilai parameter densitas tertinggi dari grup A sebagai nilai centroid awal klaster pertama, kemudian gabungkan ke dalam grup B.

8. Data grup A, pilih objek data dengan jarak terjauh dari objek data yang berada dalam B untuk nilai klaster selanjutnya, gabungkan ke dalam B.

9. Ulangi langkah 8 hingga semua objek data k menjadi data di dalam grup B.

10. Berdasarkan nilai k *centroid* awal klaster, gunakan k-means untuk melakukan pengelompokan terhadap objek data.

3. HASIL DAN PEMBAHASAN

3.1. Analisa Perhitungan

Analisa perhitungan dalam penelitian ini dengan mengambil data sample dari data asli data penerimaan mahasiswa baru tahun 2013 terdiri dari data nama desa sebanyak 27 desa dengan 2 variabel yaitu program studi yang terdiri dari DIII dan biaya USPI terdiri dari gelombang 1, gelombang 2, gelombang 3 untuk dikelompokkan menggunakan algoritma K-Means, kedalam 3 *cluster* dan selanjutnya diinisialisasikan sebagai C1 s/d C3. C1 s/d C3 ini diperoleh dengan menggunakan algoritma k-means dengan metode penghitungan jarak dan densitas.

Proses clautering dengan menggunakan algoritma K-Mean dengan menggunakan Metode Pengukuran Jarak dan Densitas adalah sebagai berikut :

1. Diketahui ($x_{11}, x_{12}, x_{13}, x_{14}, x_{15}$), ($x_{21}, x_{22}, x_{23}, x_{24}, x_{25}$), $x_{1n}, x_{2n}, x_{3n}, x_{4n}, x_{5n}$

dimana adalah x_1 program studi D3 komputerisasi akuntansi, x_2 adalah program studi S1 teknik informatika, x_3 adalah gelombang 1, x_4 adalah gelombang 2 dan x_5 adalah gelombang 3.

2. Langkah selanjutnya menghitung densitas dengan persamaan sebagai berikut :

$$N_{\text{eps}(p)} = \{q \in D \mid \text{dist}(p, q) < \text{eps}\}$$

Keterangan dari rumus diatas adalah $N_{\text{eps}(p)}$ merupakan Eps *neighborhood* objek-objek yang masuk dalam cluster yang ditentukan, q merupakan *core point* (titik pusat) sedangkan D adalah jarak objek dengan *core point* , langkah selanjutnya adalah menentukan densitas dengan mentukan objek-objek yang akan dijadikan pusat centroid. Dalam penelitian ini ditentukan banyaknya cluster adalah 3 : perencanaan kampus dengan mempertimbangkan aspek kapasitas penerimaan mahasiswa baru yaitu 350 mahasiswa setiap tahun dengan komposisi sebagai berikut :

- a. Cluster dengan strategi promosi baik
- b. Cluster dengan strategi promosi cukup
- c. Cluster dengan strategi promosi kurang.

3. Langkah ketiga perhitungan jarak :

Perhitungan jarak dihitung dengan menggunakan *euclidean distance* menggunakan persamaan :

$$D(x, c) = \|x - c\|^2 = \sqrt{\sum_{i=1}^n \|x_{ii} - c_{ii}\|^2}$$

pada perhitungan jarak pertama dihitung dengan menggunakan data penerimaan mahasiswa baru tahun 2014. Dari setiap instance data diukur jarak terhadap setiap instance centroid baru menggunakan Euclidean distance menggunakan persamaan sebagai berikut :

$$d(x_i, x_j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{im} - x_{jm})^2} \quad (3)$$

Keterangan :

- d : distance
- i : status cluster ke i
- j : indicator ke-j
- m : banyaknya data

4. Setelah didapatkan perhitungan jarak, langkah selanjutnya menghitung rata-rata dari data yang memilih pada centroid yang sama. Sehingga di dapat rerata seperti berikut :

Hitung nilai parameter densitas seluruh objek data yang berada dalam dataset D.

$$\text{Dens}(x_i) = \sum_{j=1}^n u(\text{MeanDist} - d(x_i, x_j)) \quad (5)$$

dimana u(z) merupakan sebuah fungsi bernilai 1 jika z lebih besar dari 0 dan bernilai 0 atau sebaliknya.

5. Proses perhitungan nilai rata-rata dari data-data yang memilih pada centroid yang sama dirumuskan sebagai berikut :

$$\text{MeanDist} = \frac{1}{c_i} \sum d(x_i, x_j)$$

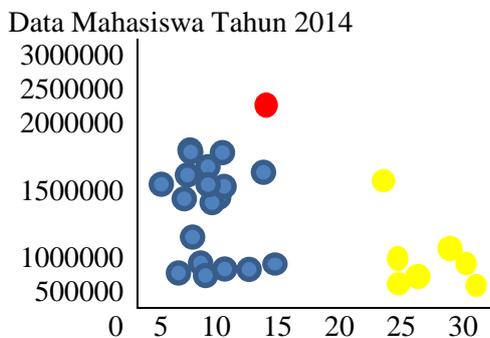
MeanDist adalah rata-rata jarak dari persamaan 3. Dimana c merupakan kombinasi pasangan jarak centroid baru cluster 1

Perhitungan centroid baru adalah sebagai berikut :

- Cluster 1, 1 untuk centroid daerah dengan strategi promosi baik
- Cluster 1,2 daerah dengan strategi promosi baik
- Cluster 1,3 daerah dengan strategi promosi baik
- Cluster 1,4 daerah dengan strategi promosi baik
- Cluster 1,5 daerah dengan strategi promosi baik

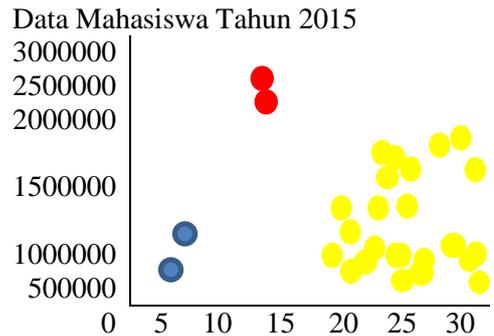
Selanjutnya melakukan perhitungan 5 untuk menentukan cluster2 dan cluster 3 untuk mendapatkan cendroid awal. Merupakan centroid baru untuk melakukan perhitungan pengulangan langkah ke 3. Dan setelah perhitungan dilakukan iterasi kedua tidak ditemukan adanya perpindahan data antara cluster, sehingga proses iterasi pada clustering menggunakan algoritma K-Means dengan pengukuran jarak dan densitas dihentikan. Berdasarkan perhitungan rerata pada centroid yang sama menunjukan tabel dibawah ini menunjukan lokasi cendroid yang sudah tetap (konvergen)

a. Grafik Hasil Metode K-Means dan Densitas Data Mahasiswa Tahun 2014



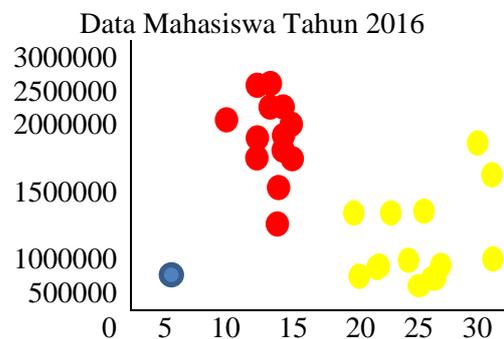
Gambar 3.1 Grafik Clustering pada Iterasi 2

b. Grafik Hasil Metode K-Means dan Densitas Data Mahasiswa Tahun 2015



Gambar 3.2 Grafik Clustering pada Iterasi 2

c. Grafik Hasil Metode K-Means dan Densitas Data Mahasiswa Tahun 2016



Gambar 3.3 Grafik Clustering pada Iterasi 2

Dari gambar 3.2 dilihat hasil clustering bahwa titik berwarna biru merupakan cluster daerah strategi promosi baik yang terdapat 1 kecamatan, titik berwarna merah merupakan cluster daerah strategi promosi cukup terdapat 13 kecamatan, titik dengan warna kuning adalah cluster daerah strategi kurang terdapat 13 kecamatan. Rincian kecamatan sebagai berikut :

- Cluster daerah strategi promosi baik : Ajibarang
- Cluster daerah strategi promosi cukup : Wangon, Tambak, Somagede, Purwokerto Timur, Purwokerto Barat, Purwojati, Pekuncen, Patikraja, Kedungbanteng, Kebumen Jatilawang, Gumelar, Banyumas.
- Cluster daerah strategi promosi kurang : Sumpiuh, Sumbang, Sokaraja, Rawalo, Purwokerto Utara, Purwokerto Selatan, Lumbir, Kemranjen, Kembaran, Karanglewas, Kalibagor, Cilongok, Baturaden.

3.2 Pengujian Metode yang diusulkan

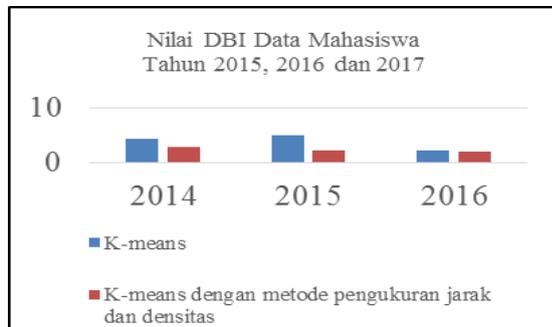
3.2.1. Model Pengujian

Model pengujian yang digunakan dalam penelitian ini adalah model clustering yang divalidasi menggunakan Davies Bouldin Index (DBI) dan waktu eksekusi Clustering dilakukan dengan menggunakan algoritma k-means dimana nilai k adalah 3. Penentuan nilai k

didasarkan pada kebutuhan dari pihak Sekolah Tinggi dalam pengelompokan daerah kedalam 3 status promosi.

3.2.2. Pengujian Davies Bouldin Index

Diperoleh bahwa :



Gambar 3.4. Index Davies

Pada gambar 3.4 dapat dilihat bahwa DBI dengan Kmeans dengan metode pengukuran jarak dan densitas lebih optimal dibandingkan dengan Kmeans biasa.

Tabel 3.1 Data Index Davies

Data tahun mahasiswa	K-means	K-means dengan metode pengukuran jarak dan densitas
2014	4,27	2,7
2015	4,9	2,1
2016	2,07	1,95

Dari tabel 3.1 dapat diketahui bahwa nilai yang paling optimal dari algoritma k-means dan algoritma K-Means dengan pengukuran jarak dan densitas adalah algoritma K-means dengan pengukuran jarak dan densitas dengan nilai Index Davies pada data mahasiswa tahun 2014 sebesar 2,7, pada data mahasiswa tahun 2015 sebesar 2,1 dan pada data mahasiswa tahun 2016 sebesar 1,95.

4. KESIMPULAN

Dari pembahasan dan evaluasi pada bab-bab sebelumnya, pengelompokan data penerimaan mahasiswa tahun 2014, tahun 2015 dan tahun 2016 kedalam 3 kelompok menggunakan algoritma K-Means dengan metode perhitungan jarak dan densitas diperoleh kesimpulan yaitu:

1. Telah dikelompokkannya data penerimaan mahasiswa tahun 2014 kedalam 3 status daerah dengan diperoleh jumlah daerah strategi promosi untuk setiap cluster yaitu cluster strategi promosi baik sebanyak 19 kecamatan, daerah dengan strategi promosi cukup 1 kecamatan, dan daerah dengan strategi promosi kurang sebanyak 7 kecamatan.

2. Telah dikelompokkannya data penerimaan mahasiswa tahun 2015 kedalam 3 status daerah dengan diperoleh jumlah daerah strategi promosi untuk setiap cluster yaitu cluster strategi promosi baik sebanyak 2 kecamatan, daerah dengan strategi promosi cukup 2 kecamatan, dan daerah dengan strategi promosi kurang sebanyak 23 kecamatan
3. Telah dikelompokkannya data penerimaan mahasiswa tahun 2016 kedalam 3 status daerah dengan diperoleh jumlah daerah strategi promosi untuk setiap cluster yaitu cluster strategi promosi baik sebanyak 1 kecamatan, daerah dengan strategi promosi cukup 13 kecamatan, dan daerah dengan strategi promosi kurang sebanyak 13 kecamatan
4. Pengelompokan data penerimaan mahasiswa tahun 2014, tahun 2015 dan tahun 2016 kedalam 3 status daerah promosi menggunakan algoritma k-algoritma k-means dengan metode perhitungan jarak dan densitas memiliki nilai Index Davies paling optimal dibanding algoritma k-means biasa.

DAFTAR PUSTAKA

- [1] L. E. I. Xu, C. Jiang, and J. Wang, "Information Security in Big Data : Privacy and Data Mining," pp. 1149–1176, 2014.
- [2] Lubis, Arlina , "Strategi Pemasaran dalam Persaingan Bisnis", 2014
- [3] M. Verma, M. Srivastava, N. Chack, A. K. Diswar, and N. Gupta, "A Comparative Study of Various Clustering Algorithms in Data Mining Manish Verma , Mauly Srivastava , Neha Chack , Atul Kumar Diswar , Nidhi Gupta," vol. 2, no. 3, pp. 1379–1384, 2012.
- [4] P. Thangaraju, B. Deepa, and T. Karthikeyan, "Comparison of Data mining Techniques for Forecasting Diabetes Mellitus," vol. 3, no. 8, pp. 7674–7677, 2014.
- [6] L. Xu, C. Jiang, J. Wang, J. Yuan and Y. Ren, "Information Security in Big Data: Privacy and Data Mining," IEEE Access: The Journal for Rapid Open Access Publishing, vol. 1, pp. 1149-1176, 9 October 2014
- [7] Purwanto, 2008
- [8] N. C. a. J. Saichon, "Opinion Mining for Thai Restaurant Reviews using KMeans Clustering and MRF Feature Selection," Knowledge and Smart Technology (KST), 2015.
- [9] S. Ghosh, "Comparative Analysis of K-Means and Fuzzy C- Means Algorithms," vol. 4, no. 4, pp. 35–39, 2013.
- [10] C. Zhang and Z. Fang, "An Improved K-means Clustering Algorithm Traditional K-mean Algorithm," vol. 1, pp. 193–199, 2013.

- [11] Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining. [Morgan Kaufmann series in data management systems]*. [https://doi.org/10.1002/1521-3773\(20010316\)40:6<9823::AID-ANIE9823>3.3.CO;2-C](https://doi.org/10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C)
- [12] Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., ... Dan, J. H. (2008). *Top 10 algorithms in data mining*. <https://doi.org/10.1007/s10115-007-0114-2>
- [13] Zhang, C., & Fang, Z. (2013). An Improved K-means Clustering Algorithm Traditional K-mean Algorithm, *1*, 193–199.
- [14] J. Pengembangan, T. Informasi, and D. Ilmu, “Implementasi Metode Improved K-Means Untuk Mengelompokkan Dokumen Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer,” no. February, 2018.
- [15] Informatika, J. I., Julianto, V., Permadi, J., Informatika, J. T., Negeri, P., Laut, T., ... Laut, T. (2017). Baru Politeknik Negeri Tanah Laut Menggunakan Metode K-Means Clustering, *2(1)*, 99–105.
- [16] Abdurasyid, M., Indriati, I., & Perdana, R. S. (2018). Implementasi Metode Improved K-Means Untuk Mengelompokkan Dokumen Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *2(10)*, 3939–3947. Retrieved from <http://j-ptiik.ub.ac.id>
- [17] Xiong, C., & Lv, K. (2016). An Improved K-means text clustering algorithm By Optimizing initial cluster centers, *272–275*. <https://doi.org/10.1109/CCBD.2016.29>
- [18] Rahman, M. A., Islam, M. Z., & Bossomaier, T. (2015). ModEx and Seed-Detective: Two novel techniques for high quality clustering by using good initial seeds in K-Means. *Journal of King Saud University - Computer and Information Sciences*, *27(2)*, 113–128. <https://doi.org/10.1016/j.jksuci.2014.04.002>