

Turnitin Originality Report

Processed on: 25-Jul-2020 12:51 AM WIB

ID: 1361664154

Word Count: 2022

Submitted: 1

Penerapan Metode K-Means Clustering
Pada Data Penerimaan Mahasiswa Baru By
Risa Helilintar

Similarity Index

17%

Similarity by Source

Internet Sources: 6%
Publications: 11%
Student Papers: 9%

3% match (publications)

[Edy Susena, Anista Yulia Ratnawati, Edy Susanto. "ANALISIS DAN PERANCANGAN SISTEM INFORMASI MANAJEMEN PENDATAAN INDUSTRI KECIL DAN MENENGAH \(SIM-IKM\)", Jurnal AKSI \(Akuntansi dan Sistem Informasi\), 2019](#)

1% match (student papers from 08-Jul-2020)

[Submitted to Mandarin High School on 2020-07-08](#)

1% match (Internet from 19-May-2019)

<https://ejournal.akakom.ac.id/index.php/jiko/article/download/9/2>

1% match (publications)

[Riski Sundari, Solikhun Solikhun, Eka Irawan, Edy Satria. "Penerapan Algoritma K-Means Clustering dalam Mengelompokkan Jumlah Posyandu Aktif Berdasarkan Provinsi", Prosiding Seminar Nasional Riset Information Science \(SENARIS\), 2019](#)

1% match (publications)

[Wiga Maulana Baihaqi, Kuat Indartono, Syifaul Banat. "Penerapan Teknik Clustering Sebagai Strategi Pemasaran pada Penjualan Buku Di Tokopedia dan Shopee", Paradigma - Jurnal Komputer dan Informatika, 2019](#)

1% match ()

<http://journal.ppns.ac.id/index.php/CDMA/article/view/348>

1% match (student papers from 08-Feb-2018)

[Submitted to Universiti Malaysia Perlis on 2018-02-08](#)

1% match (Internet from 04-Feb-2020)

<http://ejournal-binainsani.ac.id/index.php/JMBI/article/view/965>

1% match (publications)

[Atina Atina. "Segmentasi Citra Paru Menggunakan Metode k-Means Clustering", Jurnal Pendidikan Fisika dan Keilmuan \(JPFK\), 2017](#)

1% match (student papers from 22-Sep-2018)

[Submitted to UIN Syarif Hidayatullah Jakarta on 2018-09-22](#)

1% match (Internet from 14-Apr-2018)

http://riset.potensi-utama.ac.id/upload/penelitian/penerbitan_jurnal/3615prediksi_curah_hujan_di_kota_medan_menggunakan_metode_backpropagation_neural_network.pdf

1% match (publications)

[Ridho Muktiadi, Abid Yanuar Badharudin. "Metode K-Means untuk Mengelompokkan Alumni Berdasarkan Waktu Mencari Pekerjaan", Sainteks, 2020](#)

< 1% match (Internet from 12-May-2019)

<https://ellytakmp.blogspot.com/2005/03/>

< 1% match (Internet from 22-Jun-2020)

<http://sinta3.ristekdikti.go.id/authors/detail?id=6186755&view=overview>

< 1% match (Internet from 04-Aug-2019)

https://repository.fe.unjani.ac.id/pdf/01. Portofolio_Vita_citra_m_.pdf

< 1% match (publications)

[Yanuar Nurdiansyah, Saiful Bukhori, Rahmad Hidayat. "Sentiment analysis system for movie review in Bahasa Indonesia using naive bayes classifier method", Journal of Physics: Conference Series, 2018](#)

< 1% match (student papers from 11-Oct-2019)

[Submitted to Shasta Union High School District on 2019-10-11](#)

< 1% match (Internet from 10-Nov-2019) https://www.scribd.com/document/366101148/Jurnal-Siti	
< 1% match (publications) Ranti Andini Margolang, Sundari Retno Andani, Muhammad Ridwan Lubis. "Implementasi Data Mining dalam Mengelompokkan Rumah Tangga Kumuh di Perkotaan Berdasarkan Provinsi Menggunakan Algoritma K-Means", Prosiding Seminar Nasional Riset Information Science (SENARIS), 2019	
< 1% match (publications) Feny Rita Fiantika, I Ketut Budayasa, Agung Lukito. "Membangun Definisi Genetis Kubus Melalui Distorsi dan Manipulasi Objek Spasial", Jurnal Elemen, 2017	
< 1% match (student papers from 08-Oct-2018) Submitted to Universitas Kristen Satya Wacana on 2018-10-08	
< 1% match () https://p3m.sinus.ac.id/jurnal/index.php/TIKomSiN/article/view/435	
< 1% match (Internet from 11-Aug-2019) https://core.ac.uk/download/pdf/148611710.pdf	
< 1% match (Internet from 13-Dec-2019) http://journal2.um.ac.id/index.php/JSPsi/article/download/2860/2138	
< 1% match (Internet from 18-Aug-2018) https://anzdoc.com/skripsi-resti-nurianingsih-f.html	
< 1% match (student papers from 23-Jun-2017) Submitted to Universitas Sebelas Maret on 2017-06-23	
< 1% match (student papers from 02-May-2020) Submitted to Universitas Putera Indonesia YPTK Padang on 2020-05-02	
< 1% match (student papers from 23-Nov-2019) Submitted to Universitas Putera Batam on 2019-11-23	
< 1% match (publications) Suhardi Rustam, Haditsah Annur. "AKADEMIK DATA MINING (ADM) K-MEANS DAN K-MEANS K-NN UNTUK MENGELOMPOKAN KELAS MATA KULIAH KOSENTRASI MAHASISWA SEMESTER AKHIR", ILKOM Jurnal Ilmiah, 2019	
< 1% match (student papers from 19-May-2020) Submitted to Udayana University on 2020-05-19	
PENERAPAN METODE K-MEANS CLUSTERING PADA DATA PENERIMAAN MAHASISWA BARU Risa Heililintar¹, Intan Nur Farida², Rony Heri Irawan³, Universitas Nusantara PGRI Kediri risaheililintar@unpkediri.ac.id, in.nfarida@gmail.com, rony@unpkediri.ac.id ABSTRAK Pembentukan cluster merupakan salah satu teknik yang digunakan dalam megekstrak pola kecenderungan suatu data. Teknik ini digunakan dalam proses knowledge discovery in database (KDD). Data mining biasanya identic dengan proses penggalian data-data yang cukup besar dan dikelompokkan menjadi data yang tersusun rapi. Dalam hal ini penulis mengelompokkan data mahasiswa baru tahun ajaran 2016/2017 dengan clustering. Pengelompokkan yang penulis terapkan menggunakan algoritma K-Means Clustering. Algoritma K-Means clustering mampu mengelompokkan data pada kelompok yang sama dan data yang berbeda pada kelompok yang berbeda. Sehingga akan terlihat kelompok data mahasiswa baru tahunajaran 2016/2017 pada universitas nusantara PGRI yang tidak struktur menjadi terstruktur. Tujuan dari penelitian ini adalah menerapkan algoritma k-Means Clustering pada data penerimaan mahasiswa baru tahunajaran 2016/2017. Hasil K-Means clustering yang diperoleh ada dua kelompok, pusat cluster dengan pusat cluster 1= 1 ; 2.83 ; 1.4 dan cluster 2 = 2.828427124 ; 2.236067977 ; 1.414213562, cluster pertama jika asal sekolah adalah SMA maka rata-rata yang diambil adalah prodi sistem informasi dan kedua jika asal sekolah SMK adalah rata-rata yang diambil adalah Teknik informatika. Kata Kunci : Data Mining, K-Means Clustering, pengelompokan data. PENDAHULUAN Lembaga Pendidikan tertinggi adalah perguruan tinggi yang merupakan penyelenggara Pendidikan yang merupakan penyelenggara Pendidikan akademik bagi mahasiswa. Mahasiswa sering disebut kelompok masyarakat yang memiliki ciri intelektualitas yang lebih luas dibandingkan dengan kelompok seusia mereka yang bukan mahasiswa ataupun kelompok usia lain yang dibawah mereka. Dengan intelektualitasnya mahasiswa akan mampu menghadapi dan mencari permasalahan secara sistematis yang nantinya diterapkan dalam dunia kerja. Perguruan tinggi dituntut untuk menyelenggarakan pendidikan yang berkualitas bagi mahasiswa sehingga menghasilkan sumber daya manusia yang berilmu, cakap, kreatif dan bersaing. Dalam system Pendidikan mahasiswa adalah asset penting bagi institusi Pendidikan dan untuk itu perlu diperhatikan tingkat penyeleksian mahasiswa yang akan masuk ke program studi. Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat	

sangat tergantung pada tujuan dan proses KDD secara keseluruhan. [Beberapa penulis terdahulu telah menerapkan teknik K-Means Clustering sebagai penelitian](#) dalam hal pengelompokan data. Dari [penelitian yang dilakukan oleh](#) (Wuryanto, [et.al.](#), 2017) [dengan judul](#) Implementasi K- Means Clustering untuk pemetaan desa dan kelurahan di kabupaten bangkalan berdasarkan contraceptive prevalence rate dan tingkat Pendidikan dapat disimpulkan yangtelah dijabarkan, strategi advokasi yang dapat dilakukan oleh DKBPP dan PA yaitu membedakan intensitas advokasi pada tiap cluster. Intensitas advokasi pada cluster 3 dapat dilakukan lebih banyak karena tingkat kesuksesan cluster-nya rendah. Pada cluster kedua dengan tingkat kesuksesan cluster sedang, intensitas advokasi dapat dilakukan lebih sedikit daripada cluster 3 namun lebih banyak daripada cluster. Intensitas advokasi pada cluster 1 dapat dilakukan paling sedikit karena tingkat kesuksesan clusternya paling tinggi, dan hasilnya cukup efisien dan efektif. Pada penelitian [yang dilakukan oleh](#) (Syam, 2017) [dengan judul](#) Implementasi [metode](#) klastering [k-means](#) untuk mengelompokkan hasil evaluasi mahasiswa hasil penelitian ini berupa pengelompokkan mahasiswa berdasarkan cluster mahasiswa berprestasi, cluster mahasiswa berpotensi berprestasi, cluster mahasiswa berpotensi bermasalah. Selanjutnya setelah dilakukan pengujian dengan aplikasi RapidMiner dan hasilnya sama dengan perhitungan Analisa Algoritma K-Means yang dilakukan, dan hasilnya cukup efisien da efektif. Dari [penelitian yang dilakukan oleh](#) (Primajaya, [et.al.](#), 2017) [dengan judul](#) [Implementasi ALgoritma K-means clustering untuk](#) pemetaan produktivitas panen padi di kabupaten karawang. Berdasarkan uji coba yang dilakukan dan analisis yang dijelaskan data produktivitas panen padi dapat dipetakan menggunakan Teknik pengelompokan data mining menjadi 3 kelompok. Hasil perbandingan aplikasi data mining Rapidminer menunjukkan hasil pengelompokan dengan klistar sama. KAJIAN TEORI KDD Istilah data mining dan knowledge discovery in database (KDD) sering kali digunakan secara bergantian untuk menjelaskan pDarroses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapn dalam keseluruhan proses KDD adalah data mining. Proses KDD secara garis besar dapat dijelaskan sebagai berikut : 1. Data selesction Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil selesksi yang akan digunakan untuk proses data mining disimpan dalam suatu berkas, terpisah dari basis data operasional. 2. Pre- processing / cleaning Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses pembersihan pada data yang menjadi focus KDD. Proses pembersihan mencakup antara lain membuang duplikasi data, memeriksa data yang inkosisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi) 3. Transformation Coding adalah transformation pada data yang telah dipilih sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data. 4. Data mining Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan Teknik atau metode tertentu. 5. Interpretation / Evaluation Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan tahap ni merupakan bagian dari proses KDD yang disebut interpretation. Gambar 1. Aliran informasi [dalam data mining Algoritma K-MEANS](#) Pengertian [dari K-Means Clustering adalah](#), K dimaksudkan [sebagai jumlah cluster](#) yang diinginkan, Means dalam hal ini berarti nilai suatu rata-rata dari suatu grub data yang dalam hal ini didefinisikan sebagai cluster, sehingga [K-Means clustering adalah](#) suatu [metode](#) penganalisaan [data](#) atau [metode data mining yang melakukan](#) proses pemodelan [tanpa](#) supervise dan merupakan salah satu metode yang melakukan pengelompokkan data dengan system partisi. Metode K-Means berusaha mengelompokkan data yang ada kedalam satu kelompok, dimana data dalam satu kelompok [mempunyai karakteristik yang berbeda dengan data](#) yang ada didalam [kelompok yang lain](#). Dasair algoritma k-measn adalah sebagai berikut : 1. Tentukan nilai k sebagai jumlah klaster yang ingin dibentuk. 2. Inisialisasi k sebagai centroid yang dapat dibangkitkan secara random. 3. Hitung [jarak setiap data ke masing-masing centroid](#) menggunakan persamaan Euclidean Distance yaitu sebagai berikut : (1) 4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya. 5. Tentukan posisi centroid baru (k) 6. Kembali ke langkah 3 jika posisi centroid baru dengan centroid lama tidak sama. METODE PENELITIAN Metode yang dilakukan pada [penelitian ini](#) menggunakan [SDLC \(Software Development Life Cycle\) dengan model waterfall](#), dimulai [dengan analisis](#) kebutuhan pada implementasi algoritma k-means clustering untuk pengklasifikasian Program studi yang dipiiah oleh maba. Metode yang dilakukan dalam analisis kebutuhan meliputi studi literature, pengumpulan data, serta memodelkan algoritma k-means clustering ke Pseudocode. Gambar 2. Tahapan metode [penelitian HASIL DAN PEMBAHASAN Penelitian ini bertujuan untuk](#) mengelompokkan data [mahasiswa baru tahun ajaran 2016/2017 dengan menggunakan Algoritma K-Means Clustering](#). Untuk mencapai tujuan tersebut, penulis akan melakukan pengujian dengan [menggunakan data mahasiswa baru tahun ajaran 2016/2017](#) pada [Universitas Nusantara PGRI](#) . Pemodelan k-means dimodelkan dalam algoritma flowchart Gambar 3. Flochart algoritma data mining Adapun pengelompokan data tersebut digunakan langkah-langkah [sebagai berikut : Sumber Data](#) [Sumber data dalam penelitian ini](#) langsung diambil dari Universitas Nusantara PGRI Kediri dimana data tersebut merupakan data sekunder yang terdiri atas data mahasiswa baru angkatan 2016/2017 dengan banyak data 140 data dengan field yaitu data daftar, NO. Formulir, No Ujian, Nama Mahasiswa baru, Prodi, Nilai UAN, Asal Sekolah, berikut adalah contoh data pendaftaran Mahasiswa baru Tabel 1. Data Pendaftaran Mahasiswa Baru Transformasi Data [Pada tahap ini dilakukan proses](#) perubahan [data](#), tujuannya adalah agar data dapat diolah [dengan menggunakan](#) algoritma [K-Means Clustering](#). Variabel [yang](#) dipilih pada data pendaftaran mahasiswa baru yaitu data Nilai UAN, Asal Sekolah dan program studi yang dipilih. Untuk variable Nilau UAN dikelompokkan menjadi 3 kelompok yang pertama untuk nilai UAN dengan rata-rata <3 ditarnsformasikan dengan nilai 1, >3 dan data <= 4 ditransformasikan dengan nilai 2 and >=4 ditransformasikan dengan nilai 3. Untuk data program studi dikelompokkan menjadi 2 kelompok, jika program studi system informasi ditransformasikan dengan nilai 1, prodi Teknik informatika ditransformasikan dengan nilai 2 . Hasil transformasi [dapat dilihat pada table 2 berikut ini : Tabel 2. Data](#) Hasil Transformasi Pengolahan data Setelah proses transformasi langkah selanjutnya adalah proses pengolahan data [menggunakan metode](#) algoritma [K-Means Clustering](#). Tahapan proses algoritma

K-Means Clustering adalah sebagai berikut ; 1. Tentukan nilai k dari jumlah cluster yang ingin dibentuk. Cluster yang akan dibuat adalah 2 cluster. 2. Tentukan pusat awal dari setiap cluster. Dalam penelitian ini titik pusat awal ditentukan secara random dan didapat titik pusat dari setiap cluster dapat dilihat pada tabel 3. Tabel 3. Titik Pusat awal setelah cluster 3. Dalam penelitian ini digunakan metode hard k-means untuk mengalokasikan setiap data ke dalam suatu cluster yang memiliki jarak paling dekat dengan titik pusat setiap cluster. Untuk mengetahui cluster mana yang paling dekat dengan data, maka perlu dihitung jarak setiap data dengan titik pusat setiap cluster. Sebagai contoh akan dihitung jarak dari data mahasiswa pertama ke pusat cluster pertama : Dari hasil perhitungan diatas didapatkan hasil bahwa jarak data mahasiswa pertama dengan pusat cluster pertama adalah 0. 4. Jarak data mahasiswa pertama ke pusat cluster kedua : 1 Dari hasil perhitungan diatas didapatkan hasil bahwa jarak data mahasiswa pertama dengan pusat cluster adalah 1.414214. Berdasarkan hasil kedua perhitungan diatas dapat disimpulkan bahwa jarak data mahasiswa pertama yang paling dekat adalah dengan cluster , sehingga data mahasiswa pertama dimasukkan kedalam cluster 1, hasil perhitungan selengkapnya untuk 20data mahasiswa pertama dapat dilihat pada tabel dibawah ini : Tabel 4. Hasil perhitungan setiap data ke setiap cluster Iterasi 1 5. Setelah semua data ditempatkan ke dalam cluster yang terdekat, kemudian hitung kembali pusat cluster yang baru berdasarkan rata-rata anggota yang ada pada cluster tersebut. Dalam hal ini jika centroid yang lama maka hentikan iterasi, jika tidak maka lanjutkan iterasi berikutnya. Penghentian iterasi dilakukan karena saat centroid baru yang dibangkitkan dengan centroid yang lama akan menyebabkan konvergensi pada grub atau cluster sehingga tidak perlu menghitung distance space (menghitung jarak) data terhadap centroid-nya lagi. 6. Pengelompokkan hasil cluster pada iterasi pertama belum kovergen, dalam hal ini bangkitkan ulang centroid baru dengan rumus ;(2) Dimana : C : Centroid data m : Anggota data yang termasuk ke dalam centroid tertentu n : Jumlah data yang menjadi anggota centroid tertentu. Tabel 5. Hasil perhitungan setiap data ke setiap cluster Iterasi 2 Pada iterasi ke-2 ini, titik pusat dari setiap cluster sudah tidak lagi berubah dan tidak ada lagi data yang berpindah dari satu cluster ke cluster yang lain. Maka pusat cluster pada iterasi 2 adalah sebagai berikut ; Cluster 1 = 1 ; 2.83 ; 1.4 Cluster 2 = 2.828427124 ; 2.236067977 ; 1.414213562 KESIMPULAN Dari hasil penelitian yang dilakukan, maka didapatkan beberapa kesimpulan sebagai berikut : 1. Pengujian yang dilakukann dalam penelitian ini, iterasi clustering data mahasiswa terjadi sebanyak 2 kali iterasi 2. Berdasarkan dari hasil cluster kesimpulan yang dapat diambil adalah bahwa jika asal sekolah adalah SMA atau sekolah menengah atas maka rata-rata jurusan yang diambil adalah sistem informasi dan jika asal sekolahnya SMK atau sekolah menengah kejuruan rata-rata yang diambil adalah jurusan Teknik informatika. 3. Pada penelitian ini bisa dikembangkan dengan metode-metode yang lainnya. DAFTAR PUSTAKA Cahyati, D., Herawatie, D., and Wuryanto, E. 2017. Implementasi K-Means Clustering pemetaan desa dan kelurahan di kabupaten bangkalan berdasarkan contraceptive prevalence rate dan tingkat Pendidikan. Syam, A., F. 2017. Implementasi Metode klastering K-Means untuk mengelompokkan hasil evaluasi mahasiswa. Primajaya, A., Defiyanti, S., Ridlo, R., M. 2017. Implementasi Algoritma K-Means Untuk pemetaan produktivitas panen padi di kabupaten karawang. Kusrini and Lutfi E., T. 2009. Algoritma Data Mining. Yogyakarta. Seminar Nasional Teknologi Informasi dan Komunikasi Literasi Digital pada Era Revolusi Industri 4.0 SENATIK|2018| Seminar Nasional Teknologi Informasi dan Komunikasi Literasi Digital pada Era Revolusi Industri 4.0 SENATIK|2018| Seminar Nasional Teknologi Informasi dan Komunikasi Literasi Digital pada Era Revolusi Industri 4.0 SENATIK|2018| Seminar Nasional Teknologi Informasi dan Komunikasi Literasi Digital pada Era Revolusi Industri 4.0 SENATIK|2018| Seminar Nasional Teknologi Informasi dan Komunikasi Literasi Digital pada Era Revolusi Industri 4.0 SENATIK|2018| Seminar Nasional Teknologi Informasi dan Komunikasi Literasi Digital pada Era Revolusi Industri 4.0 SENATIK|2018|

=====

PROGRAM STUDI TEKNIK INFORMATIKA - UNIVERSITAS PGRI MADIUN | 14

=====

PROGRAM STUDI TEKNIK INFORMATIKA - UNIVERSITAS PGRI MADIUN | 15

=====

PROGRAM STUDI TEKNIK INFORMATIKA - UNIVERSITAS PGRI MADIUN | 16

=====

PROGRAM STUDI TEKNIK INFORMATIKA - UNIVERSITAS PGRI MADIUN | 17

=====

PROGRAM STUDI TEKNIK INFORMATIKA - UNIVERSITAS PGRI MADIUN | 18

=====

PROGRAM STUDI TEKNIK INFORMATIKA - UNIVERSITAS PGRI MADIUN | 19

=====

PROGRAM STUDI TEKNIK INFORMATIKA - UNIVERSITAS PGRI MADIUN | 20

=====