# Implementasi Algoritma K-Means Dalam Pengklasteran Untuk Rekomendasi Penerima Beasiswa PPA di UDINUS

by Joins 3350

**Submission date:** 21-Apr-2020 03:16AM (UTC+0700)

Submission ID: 1302959481

File name: 3350-9428-1-BR.docx (289.54K)

Word count: 1784

Character count: 11434

## Implementasi Algoritma K-Means Dalam Pengklasteran Untuk Rekomendasi Penerima Beasiswa PPA di UDINUS

#### Abstrak

Rekomendasi penerima beasiswa Peningkatan Prestasi Akademik (PPA) dikelompokkan menjadi 2 cluster yaitu diterima dan tidak diterima untuk mendapatkan beasiswa. Pengelompokan menjadi 2 cluster ini berguna untuk memudahkan dalam menentukan penerima beasiwa PPA. Algoritma K-Means merupakan algoritma dari teknik clustering yang dapat mengelompokkan data penerima beasiswa. Tujuan dari penelitian ini adalah untuk merekomendasikan penerima beasiswa dengan menggunakan algoritma k-means, hasil rekomendasi ini di lihat dari penempatan anggota ke masing-masing cluster yang dihasilkan. Data yang digunakan dalam penelitian ini adalah data pendaftar beasiswa PPA tahun 2016 di udinus dengan jumlah 441 dataset. Melalui seleksi atribut, k-means ini melakukan perhitungan untuk menempatkan setiap data ke cluster yang sudah ditentukan. Sebanyak 154 mahasiswa direkomendasikan mendapatkan beasiswa PPA sedangkan 287 mahasiswa tidak mendapatkan. Kata Kunci: data mining, beasiswa ppa, clustering, k-means, rekomendasi

#### Abstract

Recommendations for recipients of Peningkatan Prestasi Akademik (PPA) scholarships are grouped into 2 clusters, namely accepted and not accepted for scholarships. Grouping into 2 clusters is useful to facilitate the determination of PPA scholarship recipients. The K-Means algorithm is an algorithm of clustering techniques that can group scholarship recipient data. The purpose of this study is to recommend scholarship recipients using the k-means algorithm, the results of this recommendation are seen from the placement of members to each cluster produced. The data used in this study is the 2016 PPA scholarship registrant data in Udinus with the number of 441 datasets. Through attribute selection, this k-means performs calculations to place each data into the specified cluster. A total of 154 students were recommended to get PPA scholarships while 287 students did not get.

Keywords: data mining, ppa scholarship, clustering, k-means, recommendation

#### 1. PENDAHULUAN

Hak setiap warga Negara adalah berhak mendapatkan kemudahan dalam pelayanan pendidikan yang layak dan bermutu bagi setiap warga negaranya seperti tercantum pada Pasal 31 Undang-Undang Dasar 1945 [1]. Undang-Undang Dasar Negara Indonesia yang berbunyi mencerdaskan kehidupan bangsa telah mewajibkan warga negaranya untuk mengenyam pendidikan setinggi-tingginya. Pemerintah telah berusaha menekan angka putus kuliah bagi mahasiswa yang memiliki prestasi tinggi namun mengalami kesulitan dalam perekonomian [2].

Universitas Dian Nuswantoro adalah perguruan tinggi swasta yang memiliki 20 jenis beasiswa yang ditawarkan kepada semua mahasiswa dari pemerintah maupun non-pemerintah melalui Biro Mahasiswa (Bima). Salah satu beasiswa yang ditawarkan adalah beasiswa Peningkatan Prestasi Akademik (PPA). PPA sendiri adalah jenis beasiswa yang ditawarkan kepada mahasiswa yang kurang mampu tetapi memiliki prestasi akademik maupun non-akademik. Penyelenggara beasiswa PPA adalah Kopertis Wilayah VI dengan persyaratan antara lain fotocopy kartu mahasiswa, ktp dengan umur maksimal 23 tahun, foto ukuran 3x4 dua lembar dengan memakai jas almamater, nilai IPK, minimal semester 1 maksimal semester 6 untuk S1, fotocopy rekening buku tabungan, fotocopy kartu keluarga dan fotocopy krs terbaru dengan legalisir fakultas.

Pada tahun 2016, tercatat jumlah mahasiswa udinus pendaftar beasiswa PPA sebanyak 441 mahasiswa dari fakultas dan angkatan yang berbeda-beda. Berdasarkan pada kriteria penerimaan beasiswa PPA, terdapat bermacam-macam kriteria yang harus dipenuhi oleh mahasiswa untuk mendapatkan beasiswa tersebut. Permasalahan yang terjadi dalam melakukan seleksi penerimaan beasiswa PPA di udinus adalah sulitnya mengelola data dan atribut dalam jumlah banyak untuk menentukan penerima beasiswa PPA sehingga mengakibatkan beasiswa tersebut tidak tepat sasaran. Maksud dari tidak tepat sasarannya disini adalah mahasiswa yang seharusnya layak mendapatkan beasiswa malah tidak mendapatkannya, justru mahasiswa yang tidak layak malah mendapatkan beasiswa tersebut. Tentunya udinus juga menginginkan agar proses seleksi beasiswa PPA dapat berjalan dengan semestinya.

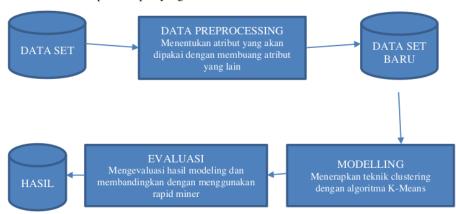
Pada penelitian yang dilakukan oleh Nurul Rohmawati [3] dengan menerapkan algoritma K-Means dalam pengklasteran pelamar beasiswa, peneliti menggunakan Rapidminer Studio 5 untuk melakukan pemodelan. Berdasarkan 36 data mahasiswa yang terdaftar, peneliti melibatkan atribut-atribut yang dia pilih untuk masuk ke tahap pemodelan yaitu NPM, IPK, jumlah SKS yang diambil, pendapatan orang tua dan jumlah tanggungan orang tua. Hasil yang diperoleh ketika mengeksekusi data tersebut dengan algoritma K-Means bahwa cluster mendapatkan hasil sebesar 0.750 atau 75%. Hasil ini membuat bahwa algoritma K-Means memiliki hasil cluster yang baik dalam kasus beasiswa pada tingkat universitas.

Pada penelitian yang dilakukan oleh Fiqih Satria [4] peneliti melakukan perbandingan kinerja Ward dan K-Means dalam proses penentuan cluster data pemohon beasiswa PPA dan BBA. Proses clustering pada beasiswa PPA, algoritma K-Means dan Ward sama-sama memiliki nilai rasio yang sama yaitu sebesar 0,749959584% sedangkan pada beasiswa BBA, metode Ward memiliki rasio yang lebih baik daripada K-Means yaitu sebesar 0,5346668% sedangkan K-Means mendapatkan 0,831525302%. Hasil ini telah menetapkan bahwa Algoritma K-Means dinilai baik dalam menyelesaikan kasus beasiswa PPA sedangkan Ward dinilai baik dalam penyelesaian kasus beasiswa BBA dalam penelitian ini.

Berdasarkan latar belakang diatas, maka dapat dilakukan penelitian tentang peran data mining dalam menyelesaikan permasalahan untuk proses seleksi penerima beasiswa PPA di udinus. Teknik data mining clustering dengan algoritma K-Means akan diterapkan pada penelitian ini untuk memecahkan masalah tersebut serta menghasilkan sebuah rekomendasi penerima beasiswa PPA.

#### 2. METODE PENELITIAN

Untuk melakukan rekomendasi terhadap data pendaftar beasiswa PPA diperlukan sebuah proses untuk mengolah data tersebut menjadi sebuah hasil rekomendasi penerima beasiswa berikut adalah tahapan-tahapan yang dilakukan:



Gambar 1 Metode Penelitian

#### 2.1 Dataset

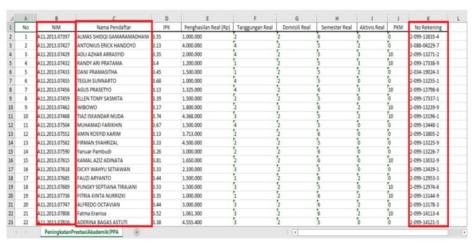
Dataset yang digunakan adalah data dari biro mahasiswa udinus yaitu data pendaftar beasiswa ppa tahun 2016 dengan jumlah 441 dataset.

4	A	В	C	D	E	F	G	H	1	J	K	
1	No	NIM	Nama Pendaftar	IPK	Penghasilan Real (Rp)	Tanggungan Real	Domisili Real	Semester Real	Aktivis Real	PKM	No Rekening	Г
2	1	A11.2013.07397	ALMAS SHIDQI GAMARAMADHANI	3.55	1.000.000	2	2	6	0	0	2-099-13835-4	
3	2	A11.2013.07427	ANTONIUS ERICK HANDOYO	3.13	4.000.000	4	2	5	2	0	3-088-04229-7	
	3	A11.2013.07429	ADLI AZHAR ARRASYID	3.35	2.000.000	4	2	5	3	10	2-099-13271-2	
	4	A11.2013.07432	RANDY ARI PRATAMA	3.4	1.200.000	1	2	5	3	10	2-099-17338-9	
	5	A11.2013.07433	DANI PRAMASITHA	3.45	1.500.000	1	2	5	2	0	2-034-19024-3	
	6	A11.2013.07455	TEGUH SUNNARTO	3.68	1.000.000	4	2	5	0	0	2-099-13255-1	
	7	A11.2013.07456	AGUS PRASETYO	3.13	1.325.000	4	2	6	3	10	2-099-13798-6	
	8	A11.2013.07459	ELLEN TOMY SASMITA	3.39	1.500.000	2	2	5	0	0	2-099-17337-1	
0	9	A11.2013.07462	WIBOWO	3.17	1.800.000	2	1	6	2	10	2-099-13239-9	
ı	10	A11.2013.07468	TIAZ ISKANDAR MUDA	3.74	4.368.000	3	2	5	2	10	2-099-13196-1	
2	11	A11.2013.07504	MUHAMAD FARIKHIN	3.67	1.500.000	4	2	5	0	0	2-099-13448-1	
3	12	A11.2013.07552	AMIN ROSYID KARIM	3.13	3.713.000	2	2	6	0	0	2-099-13805-2	
4	13	A11.2013.07582	FIRMAN SYAHRIZAL	3.33	4.500.000	2	2	5	0	0	2-099-13225-9	
5	14	A11.2013.07590	Yanuar Pambudi	3.26	3.000.000	2	2	6	0	0	2-099-13226-7	
5	15	A11.2013.07615	KAMALAZIZ ADINATA	3.81	1.650.000	2	2	6	0	10	2-099-13032-9	
7	16	A11.2013.07618	DICKY WAHYU SETIAWAN	3.33	2.100.000	3	2	3	0	0	2-099-13439-1	
8	17	A11.2013.07685	FAUZI ARYANTO	3.44	1.500.000	1	2	6	2	0	2-099-12953-3	
9	18	A11.2013.07689	PUNGKY SEPTIANA TIRAJANI	3.53	1.500.000	3	2	5	0	10	2-099-12974-6	
0	19	A11.2013.07736	FITRIA KINTA NURRIZKI	3.35	1.000.000	1	2	6	2	10	2-099-13144-9	
1	20	A11.2013.07747	ALFREDO OCTAVIAN	3.44	3.000.000	3	2	6	2	0	2-099-13178-3	
2	21	A11.2013.07808	Fatma Eranisa	3.52	1.061.300	3	2	6	2	10	2-099-14113-4	
3	22	A11.2013.07816	ADERINA BAGAS ASTUTI	3.38	4.555.400	3	2	5	0	0	2-099-14121-5	

Gambar 2 Data Pendaftar Beasiswa PPA Tahun 2016

#### 2.2 Pre-Processing

Pada tahap preprocesing dilakukan pemilahan atribut yang tidak dipakai dalam penelitian ini, dari 10 atribut yang tidak dipakai berjumlah 3 yaitu nim, nama dan nomor rekening.



Gambar 3 Preprocesing

### 2.3 K-Means

Merupakan algoritma yang membagi data pada cluster-cluster yang memiliki kemiripan terhadap satu cluster yang sama dan data yang tidak sama berada pada cluster lainnya. Data yang memiliki karakter yang sama akan dikelompokkan ke dalam cluster yang sama dan data yang memiliki karakter yang berbeda akan dikelompokkan ke dalam cluster yang lain [5]. Menurut Nurul Rohmawati, telah berhasil menjelaskan secara detail bagaimana cara kerja algoritma K-Means sebagai berikut [3]:

- 1. Menentukan jumlah k sebagai cluster yang ingin dibuat.
- Membuat nilai secara acak untuk cluster awal atau centroid awal sebanyak jumlah k yang dibuat.
- Menghitung jarak data terhadap masing-masing centroid dengan menggunakan Euclidean Distance sampai ditemukan jarak yang paling dekat dengan centroid.

Berikut adalah rumus Euclidean Distance:

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (xi - yi)^2}$$

Keterangan:

x = Titik data pertama

y = Titik data kedua

n = jumlah karakteristik (atribut) dalam terminologi data mining

d (x,y) = Euclidean Distance yaitu jarak antara data pada titik x dan titik y

- Melakukan langkah 2 sampai langkah 4 hingga anggota setiap cluster tidak mengalami perubahan.
- 5. Jika sudah tidak ada nilai yang berubah, maka iterasi terakhir digunakan untuk parameter dimana sebagai acuan menentukan prosentase data [3].

#### 2.4 Centroid

Algoritma K-Means adalah model dari *centroid* untuk membuat suatu *cluster*. *Centroid* adalah titik tengah *cluster* yang berupa sebuah nilai yang digunakan untuk menghitung jarak suatu objek data terhadap *centroid*. Penentuan *centroid* awal (initial centroid) adalah mengambil data dari data sumber secara acak sesuai jumlah K yang sudah ditentukan [6].

Sedangkan untuk memperbaharui *centroid* untuk iterasi berikutnya dapat diperoleh dari hasil rata-rata *cluster* yang bersangkutan dengan menggunakan rumus sebagai berikut [3].

$$\mu j(t+1) = \frac{1}{Nsj} \sum\nolimits_{j \in Sj} \! xj \label{eq:muj}$$
 Keterangan:

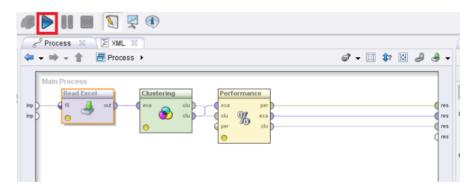
 $\mu j(t+1)$  = centroid baru pada iterasi ke (t+1) Nsj = banyaknya data pada cluster Sj

#### 3. HASIL DAN PEMBAHASAN

Didalam penelitian ini perhitungan algoritma K-Means dilakukan dengan menggunakan tools Rapid Miner Studio.

#### 3.1 Rapid Miner Studio

RapidMiner memudahkan usernya untuk melakukan perhitungan data yang banyak dengan menggunakan operator. Operator itu berfungsi memodifikasidata yang kemudian dihubungkan dengan node-node operator [7]. Lalu tugas user menghubungkan node-node tersebut untuk melihat hasilnya berupa tampilan berupa grafik. RapidMiner merupakan salah satu software untuk melakukan proses ekstraksi data menggunakan metode yang ada pada data mining [8].



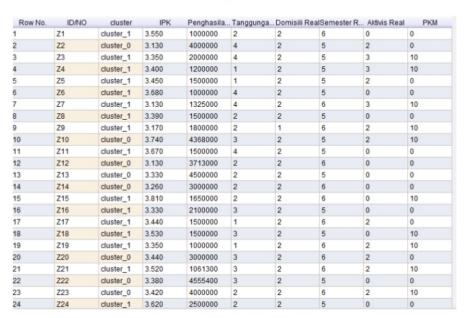
Gambar 4 Konfigurasi Rapid Miner

#### 3.2 Hasil Penelitian

Didapatkan hasil dari 441 dataset, 154 data mengikuti cluster 0 yang artinya diterima sedangkan 287 data lainnya mengikuti cluster 1 yang artinya tidak diterima.



Gambar 5 Penempatan Data Ke Cluster



Gambar 6 Rincian Cluster

#### 3.3 Hasil Pengujian F-Measure

Pengujian dengan menggunakan F-Measure ini untuk mencari nilai akurasi dengan membandingkan data aktual dan data prediksi.

Data Aktual	Data Prediksi
TP = 113	FP = 154
TN = 328	FN = 287

Gambar 7. Hasil FP-Growth

Keterangan:

TP = True Positif (Hasil dari BIMA Diterima)

TN = True Negatif (Hasil dari BIMA Ditolak)

FP = False Positif (Hasil Rekomendasi Diterima)

FN = False Negatif (Hasil Rekomendasi Ditolak)

$$Precision = \frac{TP}{TP+FP} = \frac{113}{113+154} = 0,42 = 42\%$$

Recall 
$$=\frac{TP}{TP+FN} = \frac{113}{113+287} = 0.28 = 28\%$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{113 + 328}{113 + 328 + 154 + 287} = 0,5 = 50\%$$

#### 4. KESIMPULAN

Berdasarkan penelitian yang dilakukan oleh peneliti, maka didapatkan kesimpulan bahwa metode *clustering* dengan algoritma *K-Means* dapat menghasilkan rekomendasi penerima beasiswa Peningkatan Prestasi Akademik (PPA) dengan melibatkan 7 atribut dan dengan jumlah dataset yang banyak. Penentuan centroid (titik pusat) sangat berpengaruh pada hasil cluster seperti pada hasil pengujian dengan data uji dan data keseluruhan menghasilkan hasil cluster yang berbeda.

Hasil rekomendasi dengan data uji 5 dataset, 4 data direkomendasikan mendapat beasiswa sedangkan 1 data tidak mendapatkan. Hal tersebut juga dibuktikan dengan hasil perhitungan secara manual dan perhitungan rapidminer memiliki hasil yang sama yaitu memiliki jumlah anggota yang sama pada tiap cluster. Dalam pengujian dengan seluruh data berjumlah 441

dataset, sebanyak 154 mahasiswa direkomendasikan mendapatkan beasiswa PPA sedangkan 287 lainnya tidak mendapatkan.

#### 5. SARAN

Untuk meningkatkan kinerja dan penyempurnaan dari penelitian yang telah dibuat, maka peneliti memberikan saran sebagai berikut :

- Dapat dibuat sebuah sistem aplikasi untuk mengimplementasikan penerimaan beasiswa PPA dengan menggunakan algoritma K-Means.
- Metode penelitian ini hanya berfokus pada teknik *clustering* dengan algoritma K-Means.
   Untuk penelitian selanjutnya dapat digunakan teknik dan algoritma lain.
- Nilai akurasi hanya sebesar 50%, untuk penelitian selanjutnya diharapkan dapat meningkatkan nilai akurasi menjadi lebih baik.

#### DAFTAR PUSTAKA

- [1] J. A. D. Guterres, "Kelayakan Algoritma C45 Sebagai Pendukung Keputusan Dalam Pengajuan Penerima Beasiswa," in *Seminar Nasional Teknologi Informasi*, Kupang, 2015.
- [2] P. Utomo, "Analisis Kontribusi Pemberian Beasiswa Terhadap Peningkatan Prestasi Akademik Mahasiswa Fakultas Teknik Universitas Negeri Yogyakarta," *JPTK*, vol. XX, no. 1, pp. 68-87, Mei 2011.
- [3] N. R. W, "Implementasi Algoritma K-Means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa," *Jurnal Ilmiah Teknologi Informasi Terapan*, vol. I, no. 2, pp. 62-68, April 2015.
- [4] F. Satria, "Perbandingan Kinerja Metode Ward Dan K-Means Dalam Menentukan Cluster Data Mahasiswa Pemohon Beasiswa (Studi Kasus: STMIK Pringsewu)," *Jurnal TIM Darmajaya*, vol. II, no. 1, pp. 12-26, Mei 2016.
- [5] Jaroji, "K-Means Untuk Menentukan Calon Penerima Beasiswa Bidik Misi Di Polbeng," Jurnal Inovtek Polbeng Seri informatika, vol. I, no. 1, pp. 87-94, Juni 2016.
- [6] M. Riadi, September 2017. [Online]. Available: https://www.kajianpustaka.com/2017/09/data-mining.html.
- [7] A. E. Wicaksono, "Implementasi Data Mining Dalam Pengelompokan Data Peserta Didik Di Sekolah Untuk Memprediksi Calon Penerima Beasiswa Dengan Menggunakan

Algoritma K-Means (Studi Kasus SMAN 16 Bekasi)," *Jurnal Teknologi Rekayasa*, vol. XXI, no. 3, pp. 206-216, Desember 2016.

- [8] B. Rahmat, "Implementasi K-Means Clustering Pada RapidMiner Untuk Analisis Daerah Rawan Kecelakaan," in *Seminar Nasional Riset Kuantitatif Terapan 2017*, Kendari, 2017.
- [9] A. Chandra, "Peningkatan Performa Algoritma Apriori Untuk Aturan Asosiasi Data Mining," in *Seminar Nasional Teknologi Informasi dan Multimedia*, Yogyakarta, 2017.
- [10] M. Iqbal and Muatin, "Analisa Keranjang Belanja Konsumen Pada Data Penjualan Bulan Ramadhan Menggunakan Algoritma Apriori (Studi Kasus: Citro Coffepark Clothes Pekanbaru)," in Seminar Nasional Teknologi Informasi, Komunikasi dan Industri(STNKI), Riau, 2017.

# Implementasi Algoritma K-Means Dalam Pengklasteran Untuk Rekomendasi Penerima Beasiswa PPA di UDINUS

ORIGIN	IALITY REPORT			
3 SIMIL	% ARITY INDEX	29% INTERNET SOURCES	18% PUBLICATIONS	% STUDENT PAPERS
PRIMA	RY SOURCES			
1	docobool Internet Source			5%
2	unisbank Internet Source			3%
3	www.neli			2%
4	id.scribd. Internet Source			2%
5	jurnal.uns			2%
6	garuda.ri Internet Source	stekdikti.go.id		2%
7	Angraini. Provinsi	Sukamto, Ibnu I "Penentuan Dae Riau Menggunak a K-Means", JUI	erah Rawan Ti kan Clustering	tik Api di

8	ejournal.bsi.ac.id Internet Source	1%
9	repository.uksw.edu Internet Source	1%
10	Jaroji Jaroji, Danuri Danuri, Fajri Profesio Putra. "K-MEANS UNTUK MENENTUKAN CALON PENERIMA BEASISWA BIDIK MISI DI POLBENG", INOVTEK Polbeng - Seri Informatika, 2016 Publication	1%
11	jurnal.umt.ac.id Internet Source	1%
12	journal.universitaspahlawan.ac.id Internet Source	1%
13	docplayer.info Internet Source	1%
14	journal.uny.ac.id Internet Source	1%
15	id.dinus.ac.id Internet Source	1%
16	eprints.dinus.ac.id Internet Source	1%
17	worldwidescience.org Internet Source	1%

adoc.tips Internet Source	1%
repository.upiyptk.ac.id Internet Source	1%
20 www.reportworld.co.kr Internet Source	1%
asepkurniawan58.blogspot.com Internet Source	<1%
jurnal.staialhidayahbogor.ac.id Internet Source	<1%
www.kuliahan.com Internet Source	<1%
etheses.uin-malang.ac.id Internet Source	<1%
repository.unimus.ac.id Internet Source	<1%
riko-ilmu.blogspot.com Internet Source	<1%
www.scribd.com Internet Source	<1%
id.123dok.com Internet Source	<1%

Exclude quotes Off Exclude matches Off

Exclude bibliography Off