

*FINAL REPORT*

TUGAS PROYEK SISTEM CERDAS

# Judul Sistem yang Diusulkan

**Topik: Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Menggunakan *K-Means***

## Dosen: Hepatika Zidny Ilmadina, S. Pd., M. Kom

Kelompok :

Niza Fadila (18040030), Muhammad Iqbal Habibi (18040086), A.Furqon Arifin (18040224)

## ABSTRAK

**BREBES** – Berdasarkan data kemiskinan makro Badan Pusat Statistik (BPS) per Maret 2018, tingkat kemiskinan Kabupaten Brebes tahun 2018 sebesar 17,17 persen. Ini berarti mengalami penurunan sebesar 1,97 persen dari tahun sebelumnya yang mencapai 19,14 persen. Prestasi ini luar biasa bagi Kabupaten Brebes dalam upaya penanggulangan kemiskinan.

## Permasalahan

Latar belakang

Masalah kemiskinan merupakan hal yang sangat kompleks. Di wilayah Kecamatan Brebes, seorang warga disebut sebagai keluarga miskin berdasarkan beberapa aspek seperti aspek pangan, sandang, papan, penghasilan, kesehatan, pendidikan, kekayaan, air bersih, listrik maupun jumlah jiwa. Sumber data dari BPS menunjukkan masih tingginya angka kemiskinan di provinsi Jawa Tengah dengan jumlah penduduk miskin 532.590 dengan prosentase kemiskinan 14,55%. Berdasarkan angka tersebut terlihat masih tingginya angka kemiskinan yang ada di wilayah Jawa Tengah. BPS (Badan Pusat Statistik) sebagai badan yang bertugas menghimpun data statistik kemiskinan di Kabupaten Brebes merasa kesulitan dalam pendistribusian berbagai macam bantuan. Pada kondisi saat ini, penentuan pemberian bantuan di Kecamatan Brebes dilakukan berdasarkan status kemiskinan dari total Iqbal menggunakan metode deskriptif kuantitatif dan kualitatif untuk memetakan potensi penduduk miskin perkecamatan di Kabupaten Brebes. Berdasarkan pada kondisi tersebut, maka dalam penelitian ini dibuat model clustering untuk mendapatkan klaster-klaster kemiskinan dengan menganalisa atribut yang berpengaruh maupun tidak. Upaya tersebut dilakukan melalui pembuatan suatu alat bantu berupa aplikasi dengan menggunakan metode *K-Means* untuk mengetahui pola penduduk miskin. skor data yang diperoleh dari hasil pendataan penduduk miskin. Semakin tinggi skornya maka dianggap semakin miskin. Hasil yang diperoleh dengan cara seperti ini mengakibatkan bantuan tidak tersalurkan dengan tepat. Sebagai contoh, terdapat bantuan makanan, seorang penduduk yang dianggap sangat miskin belum tentu dari sisi pangan kesulitan, bisa jadi penduduk tersebut hanya kekurangan dalam aspek sandang, papan, listrik dan yang lainya tetapi dari sisi pangan sudah terpenuhi dengan baik sehingga tidak tepat apabila bantuan makanan diberikan pada penduduk tersebut. Upaya-upaya untuk membantu program pengentasan kemiskinan di daerah Brebes pernah dilakukan melalui penelitian oleh Iqbal, Udin, dan Arifin. Tetapi penelitiannya belum mampu menunjukkan visualisasi penyebaran keluarga miskin. Udin melakukan klasifikasi keluarga miskin menggunakan *Analytical Hierarchy Process* (AHP) tetapi belum memetakan hasil klasifikasinya.

## Dataset dan Feature

Souce dataset serta isi dari dataset.

<https://brebeskab.bps.go.id/pressrelease.html?katsubjek=23&Brs%5Btgl_rilis_ind%5D=&Brs%5Btahun%5D=&yt0=Cari>

|  |  |  |  |
| --- | --- | --- | --- |
| **KECAMATAN** | **JUMLAH** | **JUMLAH PENDUDUK MISKIN** | **JUMLAH FAKIR MISKIN** |
| Banjarharjo | 115537.00 | 64491.00 | 17784.00 |
| Bantarkawung | 91704.00 | 55503.00 | 14167.00 |
| Brebes | 155474.00 | 80444.00 | 19432.00 |
| Bulakamba | 156218.00 | 97576.00 | 23962.00 |
| Bumiayu | 100483.00 | 46251.00 | 12398.00 |
| Jatibarang | 79905.00 | 38024.00 | 10052.00 |
| Kersana | 58867.00 | 32085.00 | 8833.00 |
| Ketanggungan | 130622.00 | 90372.00 | 23878.00 |
| Larangan | 136607.00 | 84907.00 | 22886.00 |
| Losari | 122549.00 | 83208.00 | 21121.00 |
| Paguyangan | 91843.00 | 62951.00 | 15688.00 |
| Salem | 55512.00 | 31640.00 | 8958.00 |
| Sirampog | 60733.00 | 32806.00 | 8229.00 |
| Songgom | 73452.00 | 44909.00 | 12165.00 |
| Tanjung | 91272.00 | 65436.00 | 16079.00 |
| Tonjong | 68405.00 | 39561.00 | 10530.00 |
| Wanasari | 133423.00 | 98411.00 | 23370.00 |

## Metode

Tahapan Pengerjaan Sistem yang Diusukan

**Contoh :**

1. Metode Pengumpulan Data

Dalam kegiatan pengumpulan data untuk penelitian ini digunakan metode pengumpulan studi pustaka yang mana pada metode ini kegiatan dilakukan adalah mempelajari, mencari dan mengumpulkan data yang berhubungan dengan penelitian ini. Data yang digunakan dalam pengelompokkan penduduk miskin ini diperoleh dari data BPS Kabupaten Brebes. dimana penelitian ini dilakukan. Data yang diperoleh kemudian akan di olah menggunakan metode kmeans dengan mengambil nilai – nilai dari setiap atribut pada data untuk mengelompokkan data penduduk miskin.

1. Metode Analisis Data

Analisis data merupakan proses mencari dan menyusun secara sistematis data yang telah diperoleh dari wawancara, dokumentasi, dokumen pribadi, observasi, catatan lapangan, gambar foto dan sebagainya, dengan cara mengorganisasikan data tersebut ke dalam kategori, menjabarkan ke dalam unit- unit, melakukan sintesa, menyusun ke dalam pola, memilih mana yang penting dan mana yang akan dipelajari dan kemudian membuat kesimpulan agar dapat dipahami diri sendiri dan orang lain. Dalam penulisan penelitian ini menggunakan analisis data yang bersifat kualitatif, penelitian kualitatif adalah analisis yang dilakukan dengan mengelompokkan data untuk mencari suatu pola dari hal yang dipelajari dan membandingkan konsep – konsep yang ada dalam sumber.

1. Studi Literatur

Dalam penelitian studi literatur adalah kegiatan ilmiah yang dilakukan untuk menemukan jawaban satu permasalahan, dan yang tujuan akhirnya adalah memberikan kontribusi teoritis atau praktis pada pengembangan bidang ilmu yang bersangkutan. Studi literatur yang digunakan disini meliputi pengolahan data penduduk miskin kabupaten brebes dan pemrograman.

1. Instrumen Penelitian

Dalam penelitian selain menggunakan data sebagai bahan penelitian juga diperlukan komponen pendukung seperti software dan hardware sebagai bahan pendukung berlangsungnya penelitian tersebut, komponen pendukung tersebut adalah :

1. Kebutuhan Hardware Kebutuhan perangkat keras (hardware) yang digunakan dalam penelitian ini adalah laptop dengan spesifikasi Processor Intel core I3,Sistem Operasi Windows 7 ,RAM 2 GB dan HDD 500 GB.

2. Kebutuhan Software Kebutuhan perangkat lunak (software) merupakan salah satu faktor yang sangat penting dalam penyusunan penelitian ini, perangkat lunak yang dipakai adalah :

a. Sistem Operasi Dalam penelitian ini sistem operasi yang dipakai adalah sistem operasi windows 7.

b. Microsoft Word Microsoft Word dalam penelitian disini digunakan untuk menyusun laporan penelitian, Microsoft Word yang dipakai adalah versi 2012.

3. Sampel Data-set Data set merupakan kumpulan data tabel dan juga di dalamnya terdapat relasi antar data tabel (data relation ) atau lebih mudahnya di dalam satu dataset bisa terdapat banyak data tabel yang berelasi.

4. Variabel Penelitian Pada sebuah penelitian data mining terdapat data yang akan diolah dengan metode yang telah ditentukan sebelumnya, pada penelitian ini data yang digunakan adalah data penduduk miskin kabupaten brebes yang akan diolah menggunakan metode k-means untuk mengelompokkan data penduduk miskin tersebut kedalam kelompok penduduk yang berkecukupan atau penduduk miskin berdasarkan beberapa variabel inputan. Variabel inputan yang digunakan dalam pengelompokkan penduduk miskin tersebut adalah nomor ID, Kecamatan, jumlah semua warga, jumlah penduduk miskin, dan jumlah penduduk fakir miskin. Kemudian variabel tersebut akan diolah menggunakan metode k-means yang kemudian menghasilkan output kelompok penduduk berdasarkan hitungan dari metode k-means.

5. Metode yang Diusulkan

Means merupakan suatu algoritma yang digunakan dalam pengelompokkan secara pertisi yang memisahkan data ke dalam kelompok yang berbeda – berda. Algoritma ini mampu meminimalkan jarak antara data ke clusternya. Pada dasarnya penggunaan algoritma ini dalam proses clustering tergantung pada data yang didapatkan dan konklusi yang ingin dicapai di akhir proses.

## Hasil Penelitian dan Pembahasan

## Implementasi Algoritma K-Means

## Dapat dilihat pada gambar di bawah merupakan diagram alur dari metode k-means yang digunakan dalam pengelompokkan penduduk miskin kabupaten brebes, pada umumnya kinerja metode k-means secara berurutan adalah sebagai berikut :

## 1.Menentukan banyaknya cluster (k)

## 2.Menentukan centroid

## 3.Apakah nilai centroidnya berubah? a.Jika ya, hitung jarak data dari centroid b.Jika tidak, selesai. 4.Mengelompokkan data berdasarkan jarak terdekat

Menentukan Banyaknya cluster (k)

Mulai

Menentukan centroid

Selesai

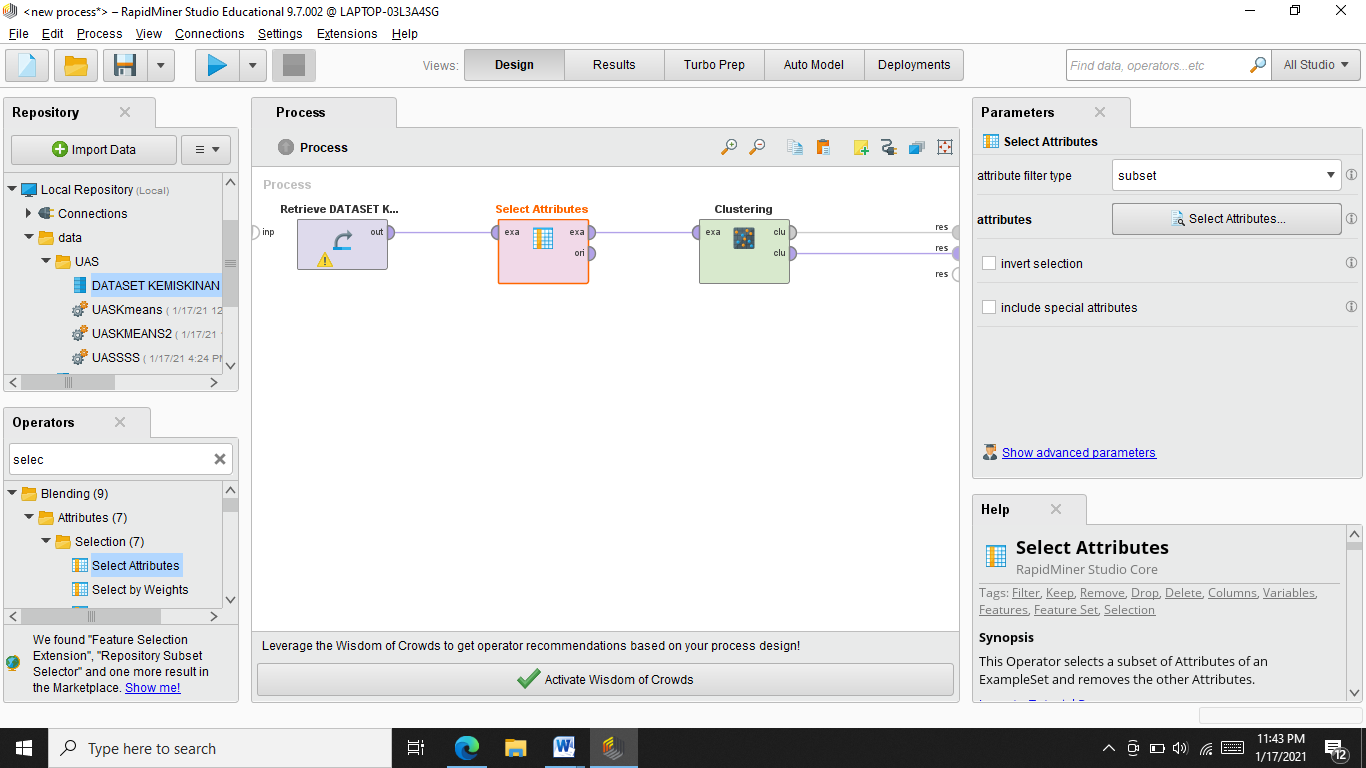
Mengelompokan data berdasarkan jarak terdekat

Menghitung jarak dari centroid

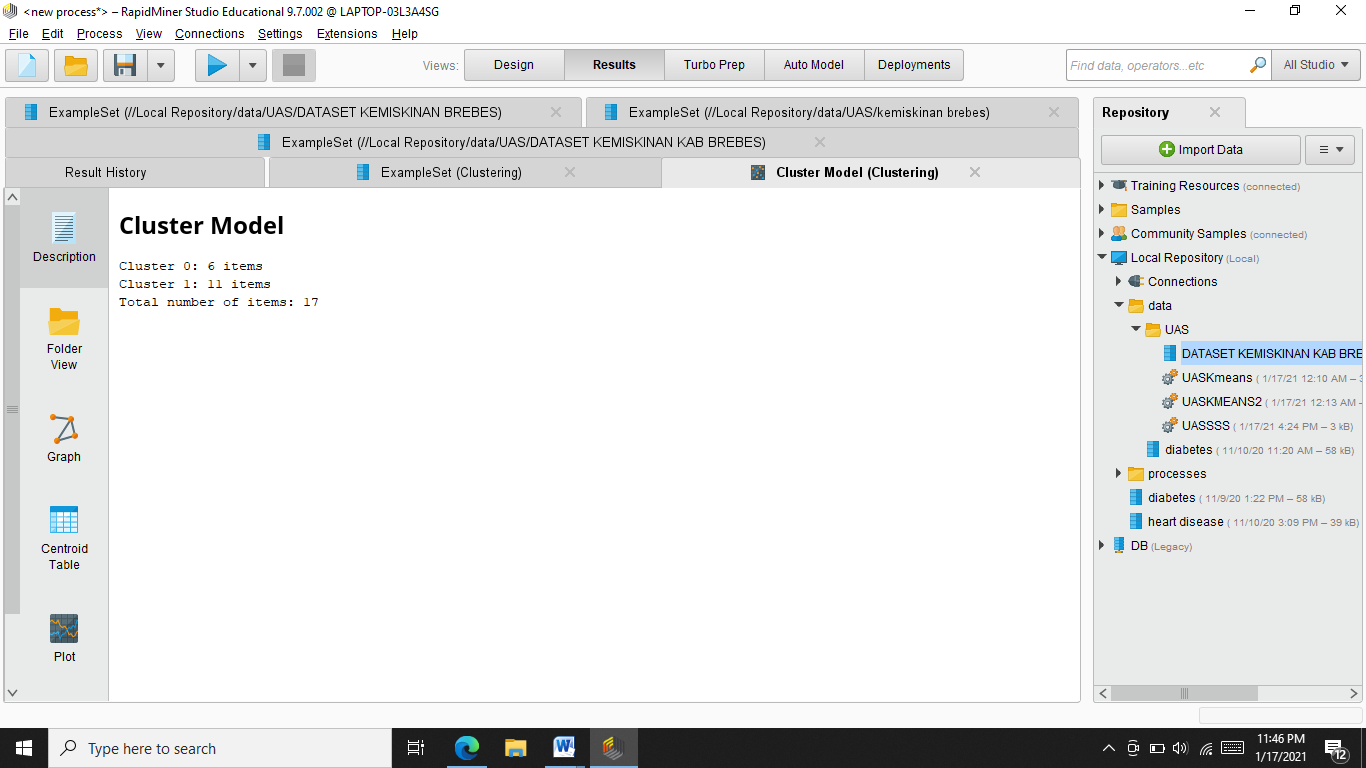
Centroid berubah?

1. Implementasi rapid miner

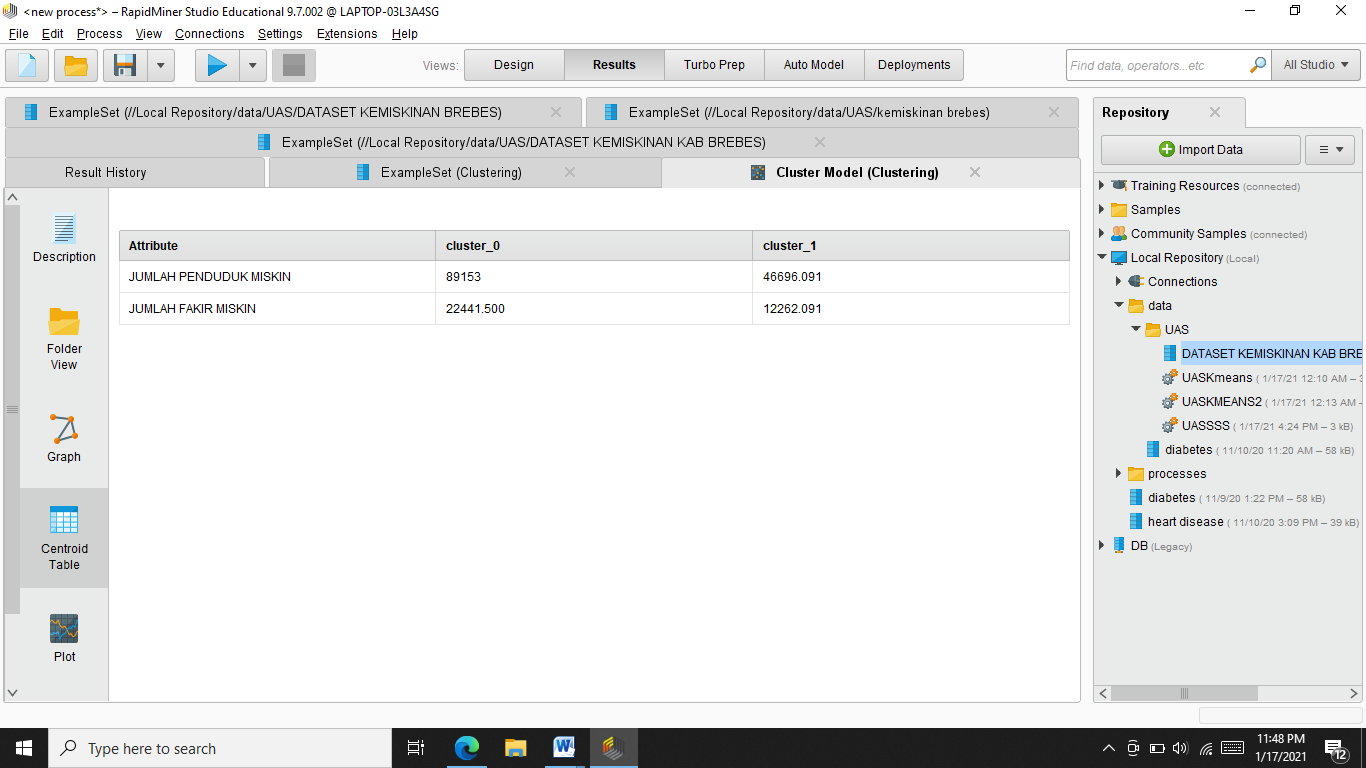
Berikut adalah pengolahan data dengan menggunakan k-means pada RapidMiner :



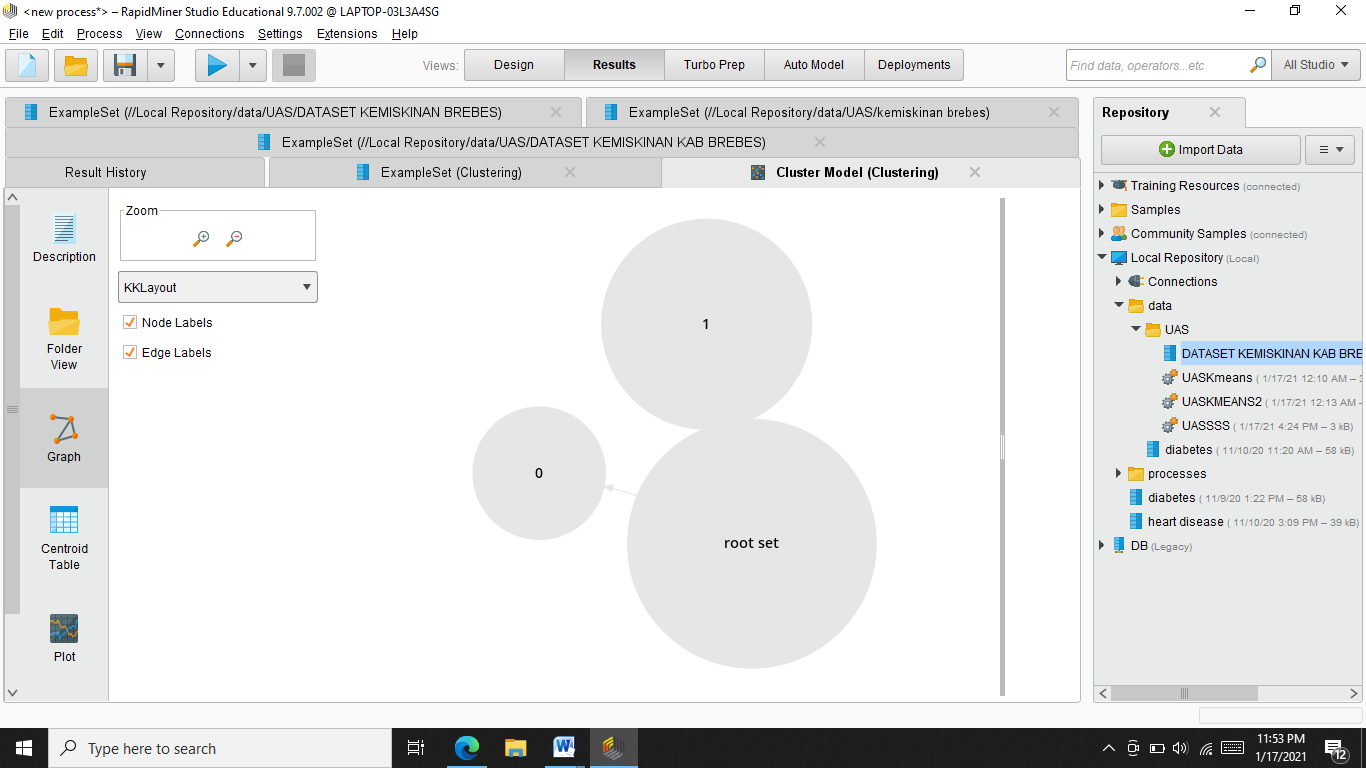
Dengan menggunakan pemodelan k-means clustering seperti gambar diatas, dengan inisialisasi jumlah cluster sebanyak 2 buah, maka didapatkan hasil dengan cluster yang terbentuk adalah 2, sesuai dengan pendefinisian nilai k dengan jumlah cluster\_0 ada 6 item, cluster\_1 ada 11 item dengan total jumlah data adalah 17.



Hasil perhitungan antara jarak cluster dan centroid



Grafik dari hasil cluster



## Kontribusi

Penjelasan kontibusi masing-masing anggota kelompok dalam pengerjaan tugas akhir proyek

1. Furqon Arifin : Membuat Dataset .

Niza Fadila : Membuat Laporan

M. Iqbal Habibi : Mencari referensi

1. **Referensi**

[1] Agustin F.E.M., Fitria A., dan S.A.H., (2015).“Implementasi Algoritma K-Means Untuk Menentukan Kelompok Pengayaan Materi Mata Pelajaran Ujian Nasional (Studi Kasus: Smp Negeri 101 Jakarta)”. Jurnal Teknik Informatika Vol. 8 No. 1 April 2015, (Page 73-78).

[2] Arta I.K.J., Indrawan G., dan Dantes. R., (2016).“Data Mining Rekomendasi Calon Mahasiswa Berprestasi Di Stmik Denpasar Menggunakan Metode Technique For Others Reference By Similarity To Ideal Solution”. Jurnal Sains dan Teknologi, Vol. 5, No. 2, Oktober 2016. ISSN: 2303-3142 (page 747- 760).

[3] Dalhatu K., Sim H.T.A., (2016). “Density base K-Mean’s Cluster Centroid Initialization Algorithm”. International Journal of Computer Applications (0975 – 8887) Volume 137 – No.11, March 2016, (pages 48-51).