

## Penentuan Prioritas Penerima Dana Bantuan Operasional Pendidikan Lembaga Pendidikan Anak Usia Dini dengan Metode KNN, TOPSIS dan *K-Means*

Diwahana Mutiara Candrasari<sup>1</sup>, Abdul Syukur<sup>2</sup>, Moch. Arief Soeleman<sup>3</sup>

<sup>123</sup>Pasca Sarjana Teknik Informatika Universitas Dian Nuswantoro

### ABSTRACT

*Education Operational Aid of the Early Childhood Education unit is the financial assistance which is provided to educational institutions, especially for those who engaged in non-formal education, which is used for the process of education in the curriculum of educational institutions in order to give the appropriate and adequate education for students. However, the reality has stated that there is a lot of financial irregularity and inaccuracy of data in the disbursement of the fund of education operational aid at the institute units of Early Childhood Education. On the other hand, there are many complaints that come from the institution itself due to the inaccuracy of data. This research was conducted by applying the KNN, K-Means algorithm and TOPSIS methods. The results of the experiment will be tested with two methods that is using cluster distance performance and K-Means cluster count performance. Meanwhile, to measure the level of ranking of TOPSIS method, the experiment will use the percentage calculation method between the experimental data with the implementation of the data to determine the accuracy of TOPSIS methods.*

*Keywords: Early Childhood Education, Education Operational Aid, KNN, K-Means, TOPSIS.*

### 1. PENDAHULUAN

Bantuan Operasional Pendidikan satuan PAUD adalah pemberian bantuan keuangan yang diberikan kepada lembaga pendidikan khususnya yang bergerak di bidang pendidikan non formal, yang digunakan untuk berlangsungnya proses pendidikan yang ada dalam kurikulum lembaga pendidikan guna memberikan pendidikan yang layak dan memadai bagi peserta didik. Pemberian Dana Bantuan Operasional Pendidikan bagi satuan PAUD, merupakan program yang dicanangkan oleh Pemerintah Pusat melalui Dinas Pendidikan, Pemuda dan Olahraga Pusat dimana program dana bantuan operasional pendidikan diadakan untuk membantu dalam proses berlangsungnya proses pembelajaran lembaga pendidikan non formal yang terkendala biaya operasional.

Pemberian Dana Bantuan Operasional Pendidikan kepada satuan PAUD di Kabupaten Batang dilakukan secara selektif sesuai dengan jenis dana bantuan yang diadakan dan harus memenuhi prinsip adil, jujur, dan tidak diskriminatif dengan tidak membedakan mengenai kondisi lingkungan tempat lembaga pendidikan tersebut menjalankan proses pendidikannya.

Banyak ditemukan penyelewengan dan ketidakakuratan data dalam pencairan dana bantuan operasional pendidikan pada lembaga satuan PAUD. Hal ini dikarenakan dalam proses pengelompokan dan perangkaan data untuk menentukan prioritas penerima dana bantuan belum sesuai dengan ketentuan yang ada, sehingga banyak lembaga satuan PAUD tutup atau tidak beroperasi karena terganjal dari sisi operasional dan kelayakan sarana prasarana yang dimiliki lembaga tersebut.

Guna mengatasi masalah tersebut telah dilakukan penelitian dengan menggunakan metode KNN dan K-Means, namun hasil yang didapatkan kurang akurat dan penggunaan metode TOPSIS dirasa belum mendapatkan hasil yang optimal dalam proses perangkaan antarkelompok. Diperlukan adanya kombinasi dari metode KNN, K-Means, dan TOPSIS dalam proses prioritas penentuan penerima dana bantuan operasional yang tepat sasaran.

Oleh sebab itu penelitian ini bertujuan memberikan masukan bagi Dinas Pendidikan, Pemuda, dan Olahraga dalam penentuan penerimaan dana bantuan operasional pendidikan pada lembaga satuan PAUD. Selain itu penelitian ini akan menguji coba penggunaan kombinasi metode KNN, K-Means, dan TOPSIS agar dapat menunjang proses prioritas penerima dana bantuan operasional pendidikan pada lembaga satuan PAUD agar tepat sasaran, sehingga mendapatkan hasil yang akurat dan tidak terjadi penyelewengan dalam penerimaan dana bantuan operasional pendidikan pada umumnya.

## 2. TINJAUAN PUSTAKA

### 2.1. Penelitian Terkait

Joshua Ignatius, et.al (2012) menggunakan *AHP-FUZZY TOPSIS* untuk mengolah data alokasi proyek pembangunan yang memiliki tujuan agar dapat menemukan sistem pengambilan keputusan yang transparan dalam pengalokasian dana proyek. [7]

Nor Ashikin Mohamed Yusof , et.al (2013) [11] menggunakan metode *K-Means* dikombinasikan dengan metode *AHP* . Hal ini bertujuan untuk mendapatkan hasil ranking dari data clustering yang sudah didapatkan terlebih dahulu. Hasil dari penelitian menghasilkan ranking dari data yang telah tercluster terlebih dahulu menggunakan *K-Means* .

Suhailan Safei, et.al (2015) menggunakan metode *AHP*. Metode *AHP* ini digunakan dalam penelitian untuk mendapatkan hasil ranking dari data lembaga sekolah tinggi yang sudah didapatkan terlebih dahulu. Hasil dari penelitian menghasilkan ranking dari setiap lembaga sekolah tinggi selain itu dengan adanya penelitian ini didapatkan hasil lembaga sekolah tinggi yang berkualitas dan lembaga tinggi yang tidak berkualitas [15].

### 2.2. LANDASAN TEORI

Landasan teori yang digunakan dalam penelitian ini bertujuan sebagai penjelasan dari metode yang digunakan di dalam penelitian ini.

#### 2.2.1 *K-Nearest Neighbor* (KNN)

Termasuk kelompok *instance-based learning*. Algoritma ini juga merupakan salah satu teknik *lazy learning*. KNN dilakukan dengan mencari kelompok *k* objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing. [12][13]

*Algoritma K-Nearest Neighbor* adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. *Nearest Neighbor* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dan kasus lama yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada. [12]

Langkah-langkah untuk menghitung metode *Algoritma K-Nearest Neighbor*:

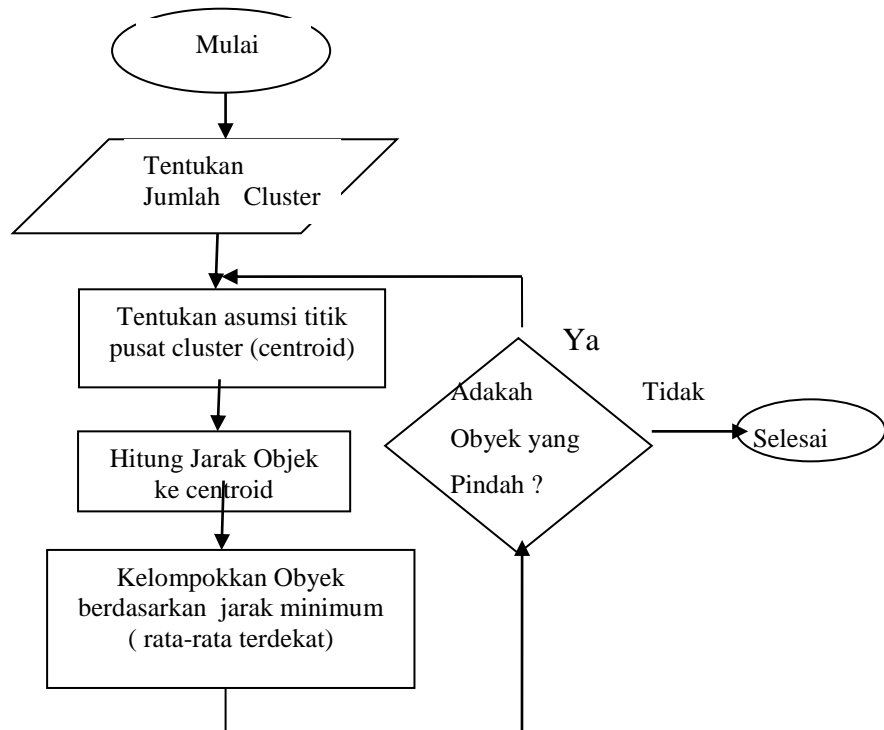
- a. Menentukan Parameter *K* (Jumlah tetangga paling dekat).
- b. Menghitung kuadrat jarak *Euclid* (*queri instance*) masing-masing objek terhadap data sampel yang diberikan.
- c. Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak *Euclid* terkecil.
- d. Mengumpulkan kategori *Y* (*Klasifikasi Nearest Neighbor*)
- e. Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas maka dapat diprediksi nilai *queri instance* yang telah dihitung.

#### 2.2.2 *K-Means*

*K-Means* merupakan algoritma *clustering* yang populer. Algoritma *K-Means* dimulaidengan pemilihan secara acak *K*, *K* di sini merupakan banyaknya *cluster* yang ingin dibentuk. Kemudian tetapkan nilai-nilai *K* secara random, untuk sementara nilai tersebut menjadi pusat dari *cluster* atau biasa disebut dengan centroid, mean atau “means”. Hitung jarak setiap data yang adaterhadap masing-masing centroid

menggunakan rumus *Euclidian* hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. [1][8][10][14]

Teknik dimulai dengan memilih secara acak poin/ benda untuk digunakan sebagai cluster awal pusat. Kemudian memilih jarak terdekat dari objects ke pusat kluster ini menjadi dikelompokkan bersama dalam pusat cluster. ada sebuah kembali banyak teknik untuk menghitung jarak multi-fitur objek, tapi *Euclidean* jarak antara yang terbaik. Klasifikasikan setiap data berdasarkan kedekatannya dengan centroid. Lakukan langkah tersebut hingga nilai centroid tidak berubah (stabil). [14][15][16][17].



Gambar 1. Flowchart *K-Means*

- a. Menentukan k sebagai jumlah cluster yang dibentuk .
- b. Memilih Centroid (titik pusat cluster) awal

$$V = \frac{\sum_{i=1}^n X_i}{n}; i = 1,2,3, \dots, n \quad (2.1)$$

Keterangan:

V = centroid

X<sub>i</sub> = objek ke -i

n = banyaknya obyek

- c. Menghitung jarak setiap objek ke masing-masing *centroid* dari masing-masing cluster. Untuk menghitung jarak antara objek dengan *centroid* menggunakan *Euclidian Distance*.

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1,2,3, \dots, n \quad (2.2)$$

Keterangan:

x<sub>i</sub> : obyek x ke-i

$y_i$  : obyek y ke- i  
 $n$  : banyaknya obyek

- d. Mengelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya.
- e. Menghitung Rasio Untuk Mengetahui Iterasi Selanjutnya
- f. Menentukan posisi centroid baru dengan cara menghitung nilai rata-rata dari data-data yang ada pada centroid yang sama.

$$C_k = \left(\frac{1}{n_k}\right) \sum d_i \tag{2.3}$$

Keterangan: :

Ck : Pusat Cluster Baru  
 nk : Jumlah  
 di : jumlah data dalam cluster

- g. Ulangi langkah 2 jika posisi *centroid tidak sama*, atau nilai rasio iterasi selanjutnya lebih sedikit atau sama dengan nilai rasio iterasi sebelumnya, maka iterasi dinyatakan berhenti.

### 2.2.3 TOPSIS (Technique For Others Reference By Similarity To Ideal Solution)

*Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)* pertama kali ditemukan oleh Yoon (1980) dan Hwang dan Yoon (1981), *TOPSIS* adalah metode yang berguna dalam menangani MCDM atau MADM masalah di dunia nyata. Ini membantu pembuat keputusan (s) mengatur masalah yang harus diselesaikan, dan melaksanakan analisis, perbandingan dan peringkat dari alternatif [5][6][7][9].

*TOPSIS* adalah salah satu metode yang bisa membantu proses pengambilan keputusan yang optimal untuk menyelesaikan masalah keputusan secara praktis. Hal ini disebabkan karena konsepnya sederhana dan mudah dipahami, komputasinya efisien dan memiliki kemampuan untuk mengukur kinerja relatif dari alternatif-alternatif keputusan dalam bentuk matematis sederhana [4].

Prosedur *TOPSIS* didasarkan pada ide intuitif dan sederhana, yaitu bahwa solusi ideal optimal, memiliki manfaat yang maksimal, diperoleh dengan memilih alternatif terbaik yang jauh dari alternatif yang paling cocok, memiliki manfaat minimal solusi ideal harus memiliki pangkat 1 (satu), sedangkan alternatif terburuk harus memiliki peringkat mendekati 0 (nol) (Hu Yonghong) [7][18][19][20].

Secara matematis prosedur *TOPSIS* :

- a. Membuat matriks keputusan yang ternormalisasi.

$$DM = \begin{matrix} & C_1 & C_2 & \dots & C_n \\ \begin{matrix} L_1 \\ L_2 \\ \dots \\ L_m \end{matrix} & \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \end{matrix} \tag{2.4}$$

- b. Menghitung matriks keputusan ternormalisasi.

$$NDM = R_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}} \tag{2.5}$$

- c. Menentukan matrik ternormalisasi dengan bobot (W)

$$V = V_{ij} = W_j X R_{ij} \tag{2.6}$$

- d. Mengidentifikasi solusi positif dan solusi negative.

$$PIS = A^+ = \{V_1^+ V_2^+ \dots V_n^+\}, \text{ Where } V_j^+ = \{(max_1 (V_{ij}))\}$$

$$\begin{aligned}
 & \text{if } j \in J, (\min V_{ij} \text{ if } j \in J) \} \\
 NIS = A^- & = \{V_1^-, V_2^-, \dots, V_n^-\}, \text{ Where } V_j^- = \{(\min(V_{ij}) \\
 & \text{if } j \in J, (\min V_{ij} \text{ if } j \in J) \} \quad (2.7)
 \end{aligned}$$

- e. Menghitung jarak nilai setiap alternative dari solusi ideal positif dan solusi ideal negatif.

$$\begin{aligned}
 S^+ & = \sqrt{\sum_{j=1}^n (V^+ - V_{ij})^2} \quad i = 1, \dots, m \\
 S^- & = \sqrt{\sum_{j=1}^n (V^- - V_{ij})^2} \quad i = 1, \dots, m \quad (2.8)
 \end{aligned}$$

- f. Menghitung nilai kemungkinan dari solusi ideal

$$C_i = S_i^- / (S_i^+ + S_i^-), \quad 0 \leq C_i \leq 1 \quad (2.9)$$

- g. Merangking hasil nilai preferensi dari masing-masing altern

#### 2.2.4 Euclidean Distance

Dari beberapa penelitian untuk clustering, pengukuran jarak yang pada umumnya sering digunakan adalah jarak Euclidean. Euclidean Distance dianggap sebagai distance matrix yang mengadopsi prinsip Pythagoras. Hal ini dikarenakan pola perhitungannya yang menggunakan aturan pangkat dan akar kuadrat. Euclidean akan memberikan hasil jarak yang relatif kecil karena menggunakan aturan akar kuadrat [7] [9]

Jarak antara Nilai Random/ Bobot dan data dihitung menggunakan rumus persamaan berikut [7] [9]:

$$D_i = \sum_{i=1}^n (w_{ij} - x_i) \quad (2.10)$$

Keterangan:

- $D_i$  : Jarak Euclidean  
 $w_{ij}$  : Bobot Neuron ke- $i$  (bobot akhir)  
 $x_i$  : input vector ke  $x$

#### 2.2.5 Purity

Purity digunakan untuk menghitung kemurnian dari suatu cluster yang direpresentasikan sebagai anggota cluster yang paling banyak sesuai (cocok) di suatu kelas. Nilai purity yang semakin mendekati 1 menandakan semakin baik cluster yang diperoleh. Untuk menghitung nilai purity setiap cluster dapat menggunakan rumus berikut [18]:

$$Purity(j) = \frac{1}{n_j} \max(n_{ij}) \quad 2.11$$

### 2.3. Penentuan Prioritas Penerimaan Dana Bantuan

Dalam penelitian ini metode yang digunakan adalah KNN digunakan untuk mengelompokkan dan menentukan lembaga yang tidak layak dan layak dalam penerimaan dana bantuan operasional pendidikan, sedangkan penggunaan metode TOPSIS untuk proses perankingan setiap data satuan PAUD, dan K-Means digunakan untuk mengelompokkan satuan PAUD sehingga bisa diketahui lembaga yang mendapatkan dana bantuan operasional pendidikan Negara, daerah tingkat I (Provinsi), dan daerah tingkat II (Kabupaten). Dengan menggunakan metode KNN, TOPSIS, dan K-Means dapat menentukan prioritas penerimaan dana bantuan operasional pada lembaga satuan PAUD dengan lebih akurat dan tepat sasaran, dibandingkan dengan menggunakan metode AHP K-Means.

Penggunaan metode *KNN* dalam penelitian ini untuk menentukan kelayakan setiap lembaga satuan PAUD dalam penentuan dana bantuan operasional pendidikan. Dalam penggunaan metode ini diperoleh dua kelompok yaitu kelompok lembaga yang layak menerima dana bantuan operasional pendidikan dan lembaga yang tidak layak menerima dana bantuan operasional pendidikan. Hasil yang akan digunakan dalam proses selanjutnya yaitu lembaga yang layak mendapatkan dana bantuan operasional pendidikan akan digunakan untuk proses selanjutnya yaitu proses menggunakan metode *TOPSIS*.

a. Cara Kerja dari metode *KNN* secara umum :

- 1) *K Nearest Neighbor* bekerja berdasarkan jarak minimum dari data baru ke data training samples untuk menentukan *K* tetangga terdekat.
- 2) Setelah itu, kita dapatkan nilai mayoritas sebagai hasil prediksi dari data yang baru tersebut. Bagaimana cara kerja Algoritma *K Nearest Neighbor* (*KNN*)?
- 3) Data berisi banyak atribut ( $X_1, X_2, \dots, X_n$ ) digunakan untuk mengklasifikasikan atribut target  $Y$  yang baru.

Berikut ini langkah-langkah dari algoritma *K Nearest Neighbors* (*KNN*):

- 1) Tentukan parameter  $K$  = jumlah banyaknya tetangga terdekat.
- 2) Hitung jarak antara data baru dan semua data yang ada di data training.
- 3) Urutkan jarak tersebut dan tentukan tetangga mana yang terdekat berdasarkan jarak minimum ke- $K$ .
- 4) Tentukan kategori dari tetangga terdekat.
- 5) Gunakan kategori mayoritas yang sederhana dari tetangga yang terdekat tersebut sebagai nilai prediksi dari data yang baru.

Penggunaan metode *TOPSIS* dalam penelitian ini untuk mendapatkan nilai bobot dari nilai evaluasi yang diperoleh setiap lembaga satuan PAUD dan hasil ranking dari setiap lembaga satuan PAUD. Hasil dari metode *TOPSIS* hasil ranking dari nilai tertinggi ke nilai terendah, serta bobot nilai dari matrik ternormalisasi yang terbentuk dari nilai evaluasi yang diperoleh dari setiap lembaga satuan PAUD.

Hasil yang akan digunakan dalam proses selanjutnya menggunakan metode *K-Means* yaitu nilai bobot matrik ternormalisasi dari setiap nilai evaluasi dari lembaga satuan PAUD.

b. Cara kerja metode *TOPSIS* :

- 1) *TOPSIS* menggunakan prinsip bahwa alternatif yang terpilih harus mempunyai jarak terdekat dari solusi ideal positif dan terjauh dari solusi ideal negatif dari sudut pandang geometris dengan menggunakan jarak Euclidean untuk menentukan kedekatan relatif dari suatu alternatif dengan solusi optimal.
- 2) Solusi ideal positif didefinisikan sebagai jumlah dari seluruh nilai terbaik yang dapat dicapai untuk setiap atribut, sedangkan solusi negatif-ideal terdiri dari seluruh nilai terburuk yang dicapai untuk setiap atribut.
- 3) *TOPSIS* mempertimbangkan keduanya, jarak terhadap solusi ideal positif dan jarak terhadap solusi ideal negatif dengan mengambil kedekatan relatif terhadap solusi ideal positif. Berdasarkan perbandingan terhadap jarak relatifnya, susunan prioritas alternatif bisa dicapai.

Berikut ini langkah-langkah dari algoritma *TOPSIS*:

- 1) Membuat matriks keputusan yang ternormalisasi
- 2) Membuat matriks keputusan yang ternormalisasi terbobot
- 3) Menentukan matriks solusi ideal positif dan solusi ideal negative
- 4) Menghitung separation measure
- 5) Menentukan jarak antara nilai setiap alternatif dengan matriks solusi ideal positif dan negative
- 6) Menentukan nilai preferensi untuk setiap alternative

Penggunaan metode *K-Means* dalam penelitian ini untuk mengelompokkan nilai lembaga satuan PAUD untuk menentukan penerimaan dana bantuan operasional. Data yang digunakan untuk penentuan

dan pengelompokan menggunakan bobot nilai matrik ternormalisasi dari masing-masing metode *AHP*, dan metode *TOPSIS*. Hasil yang didapatkan yaitu kelompok lembaga satuan PAUD yang mendapatkan dana bantuan operasional pendidikan tingkat negara, provinsi, maupun Kabupaten.

c. Cara kerja metode *K-Means* :

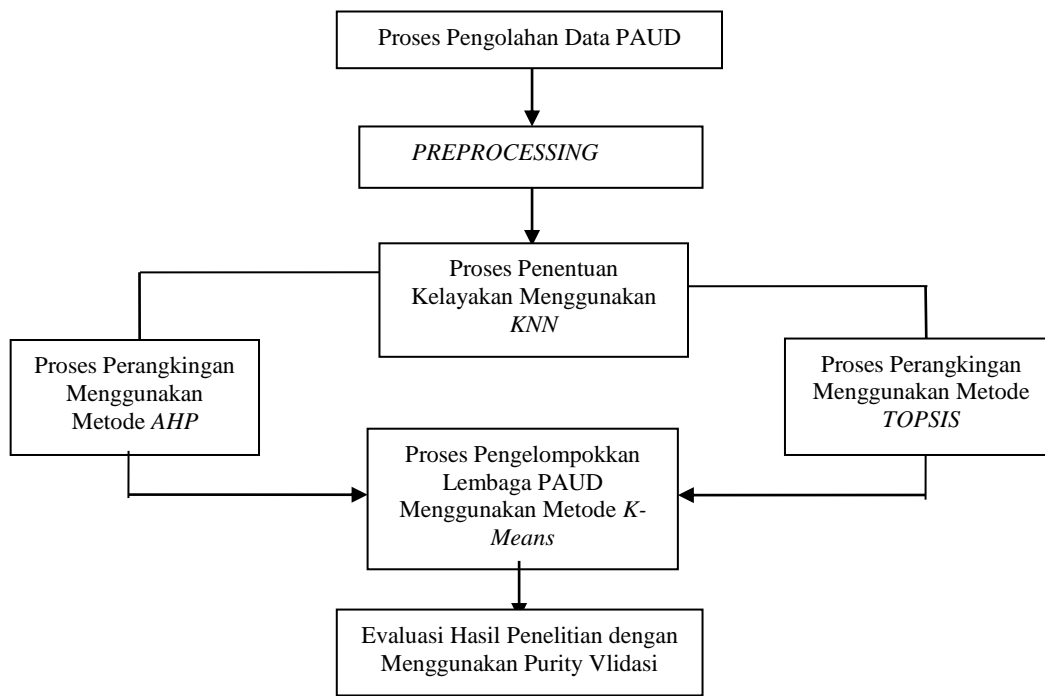
- 1) Mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam cluster yang lain.
- 2) Tujuannya adalah meminimalisasikan fungsi obyektif yang diset dalam proses pengelompokan, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu cluster dan memaksimalkan variasi antar cluster.

Berikut ini langkah-langkah dari algoritma Knearest neighbors (KNN):

- 1) Menentukan jumlah kelompok
- 2) Mengelompokkan data ke dalam kelompok secara acak untuk digunakan sebagai pusat centroid awal.
- 3) Menghitung pusat cluster (centroid/rata-rata) dari data yang ada di masing-masing cluster
- 4) Mengelompokkan masing-masing data ke centroid/rata-rata terdekat
- 5) Kembali ke Langkah 3, apabila masih ada data yang berpindah cluster, atau apabila perubahan nilai centroid ada yang di atas nilai threshold yang ditentukan, atau apabila perubahan nilai pada fungsi obyektif yang digunakan masih di atas nilai threshold yang ditentukan

Untuk meningkatkan nilai akurasi dan mendapatkan hasil yang akurat maka hasil dari *TOPSIS K-Means* akan dievaluasi dengan metode lain dengan menggunakan metode *AHPK-Means*.

### 3. METODE PENELITIAN



Gambar 2. Gambaran Umum Penelitian



### 3.1. Proses Pengolahan Data

Data yang digunakan dalam penelitian ini berupa data primer dan data sekunder:

- a. Data Primer: berupa data tentang mekanisme evaluasi penilaian lembaga satuan PAUD di Kabupaten Batang.
- b. Data Sekunder: berupa data lembaga satuan PAUD Kabupaten Batang yang diperoleh dari Dinas Pendidikan, Pemuda dan Olahraga Kabupaten Batang.

Sampel data yang digunakan dalam penelitian ini adalah data evaluasi penilaian satuan PAUD tahun 2016 yang selama ini diaplikasikan di Dinas Pendidikan, Pemuda dan Olahraga Kabupaten Batang sebanyak 617 data terdapat di 15 kecamatan terdiri dari KB sejumlah 270, TK sejumlah 266, POS PAUD, SPS, dan Tempat Penitipan Anak (TPA) sejumlah 81.

Tabel 1. Indikator Penilaian Lembaga Satuan PAUD

NO	INDIKATOR PENILAIAN
1	JUMLAH SISWA
2	JUMLAH GURU
3	SARANA PRASARANA (SARPRAS)
4	IZIN OPERASIONAL
5	GEDUNG
6	JUMLAH KELAS/ RUANGAN
7	TERCATAT DI DAPODIK
8	PROGRAM BELAJAR

### 3.2. Preprocessing

Dalam proses *preprocessing* bertujuan untuk memeriksa data set yang akan digunakan dalam penelitian. Langkah-langkah dalam proses *preprocessing* dalam penelitian ini :

- a. Mempersiapkan dataset yang akan digunakan dalam penelitian.
- b. Memeriksa dataset untuk memastikan dataset siap diolah dan digunakan., serta melihat keakuratan data.
- c. Menghilangkan duplikasi data sehingga data yang akan digunakan akurat dan tidak ada duplikasi data.
- d. Menentukan sampel data set untuk eksperimen pengujian.
- e. Memeriksa setiap indikator penilaian.

### 3.3. Proses Penentuan Kelayakan Menggunakan Metode KNN

Prosedur standar dari algoritma *KNN* :

- a. Menentukan Parameter K (Jumlah tetangga paling dekat).
- b. Menghitung kuadrat jarak *Euclid* (*query instance*) masing-masing objek terhadap data sampel yang diberikan.
- c. Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak *Euclid* terkecil.
- d. Mengumpulkan kategori Y (*Klasifikasi Nearest Neighbor*)
- e. Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas maka dapat diprediksi nilai *query instance* yang telah dihitung.

Setelah mendapatkan hasil dari proses penentuan kelayakan menggunakan metode *KNN*, tahapan selanjutnya adalah proses perbandingan dengan menggunakan metode *TOPSIS*.

### 3.4. Proses Penentuan Rangking Menggunakan Metode TOPSIS

Langkah-langkah dalam proses penentuan rangking dengan menggunakan metode *TOPSIS*:

- a. Membuat matriks keputusan yang ternormalisasi (R).
- b. Membuat matriks keputusan yang ternormalisasi terbobot (Y).



- c. Menentukan matriks solusi ideal positif ( $A^+$ ) dan matriks solusi ideal negatif ( $A^-$ ) menggunakan rumus (7).
- d. Menentukan jarak antara nilai setiap alternatif dengan matriks solusi ideal positif dan matriks solusi ideal negatif. Jarak antara alternatif dengan solusi ideal positif ( $Di^+$ ) dirumuskan sebagai (8).
- e. Menentukan nilai preferensi untuk setiap alternatif.

### 3.5. Proses Pengelompokan Lembaga PAUD Dengan Menggunakan Metode *K-Means*

Prosedur standar dari algoritma klustering *K-Means* [3] [13] :

- a. Menentukan nilai k untuk jumlah cluster yang akan dibentuk.
- b. Memasukkan masing-masing data dengan cara mencari jarak terdekat dari masing-masing data yang paling dekat dengan pusat cluster dengan persamaan (1).
- c. Menghitung masing-masing centroid untuk menentukan centroid baru dengan cara menghitung nilai rata-rata setiap data yang terdapat dalam centroid dengan menggunakan rumus (2).
- d. Jika centroid baru sudah didapatkan, maka ulangi langkah dua, kemudian menuju langkah ke lima.
- e. Mendapatkan hasil kluster.

### 3.6. Evaluasi Hasil Eksperimen

Tahapan akhir dari penelitian ini adalah dengan dilakukannya proses evaluasi, karena proses tersebut memiliki peranan penting dalam penelitian. Evaluasi penelitian dilakukan dengan dua metode untuk proses *AHP K-Means* dan *TOPSIS K-Means* dengan menggunakan purity validasi yang bertujuan untuk menilai kemurnian setiap anggota *cluster* yang terdapat dalam *cluster* tersebut.

## 4. HASIL DAN PEMBAHASAN

### 4.1. Deskripsi Data

Data dalam penelitian ini berjumlah sebanyak 617 data terdapat di 15 kecamatan terdiri dari Kelompok Bermain (KB) sejumlah 270, Taman Kanak-Kanak (TK) sejumlah 266, Pos Pendidikan Anak Usia Dini (POS PAUD), Satuan PAUD Sejenis (SPS), dan Tempat Penitipan Anak (TPA) sejumlah 81 masing-masing diambil 30 record data dari lembaga KB, POS PAUD, SPS dan TPA, serta (TK) yang digunakan sebagai sampel data. *Preprocessing* data digunakan untuk menghilangkan duplikasi data dan menyiapkan data untuk diolah dengan metode tersebut.

Tabel 2. Data Lembaga satuan PAUD

NO	NAMA LEMBAGA	JUMLAH PSRT DIDIK	JUMLAH GURU	SARPRAS	IZIN OP	GEDUNG	JUMLAH KELAS	DAPODIK	PRGM BELAJAR
1	KB TUNAS HATI	1,000	0,875	0,250	0,875	0,750	1,063	0,750	0,750
2	KB GANESHA MULYA	1,125	0,875	0,250	0,875	0,750	0,938	0,750	0,625
3	POS PAUD MAWAR	1,125	1,000	0,625	1,125	0,625	1,188	0,875	0,750
4	POS PAUD ANGGREK 2	1,000	1,143	0,500	1,125	0,625	1,063	0,750	0,625
5	TK ABA SODONG	1,000	1,000	0,250	1,125	0,625	1,188	0,750	0,750
6	TK TAMA BAKTI	1,000	0,875	0,625	1,000	0,500	1,063	0,750	0,625

### 4.2. Proses Data Menggunakan Metode *KNN*

Data satuan PAUD yang akan diolah yaitu data KB, data POS PAUD, SPS, dan data TK. Data ini akan diolah menggunakan metode *KNN*.

Tabel 3. Hasil Pengelompokkan Data KB KNN

NO	NAMA LEMBAGA	C1	C2	C3	C4	C5	C6	C7	C8	KET
1	KB TUNAS HATI	1,000	0,875	0,250	0,875	0,750	1,063	0,750	0,750	BERHAK
2	KB GANESHA MULYA	1,125	0,875	0,250	0,875	0,750	0,938	0,750	0,625	BERHAK
3	POS PAUD ANGGREK 2	1,000	1,143	0,500	1,125	0,625	1,063	0,750	0,625	BERHAK
4	POS PAUD SEJAHTERA 1	1,125	1,000	0,375	1,125	0,625	1,125	0,875	0,500	BERHAK

#### 4.3. Proses Data Menggunakan Metode TOPSIS

Dalam proses ini data yang diperoleh dari hasil pengolahan data menggunakan metode K-Means, selanjutnya hasil dari proses sebelumnya maka akan diolah lagi menggunakan metode TOPSIS untuk menentukan prioritas perangkingan data lembaga satuan PAUD.

Tabel 4. Hasil Rangkaing Data satuan PAUD

NO	V	NAMA LEMBAGA
1	0,7717211	KB ASSAID
2	0,7693477	KB AQILA
3	0,733014368	POS PAUD MELATI
4	0,713357071	SPS PERMATA HATI 03
5	0,79383048	TK AL-AMIN
6	0,721390531	TK SAKTI LESTARI 01

#### 4.4. Proses Data Menggunakan Metode K-Means

Metode *K-Means* memiliki prosedur standar dari algoritma klastering *K-Means* [3] [13] :

- Menentukan nilai k untuk jumlah cluster yang akan dibentuk.
- Memasukkan masing-masing data dengan cara mencari jarak terdekat dari masing-masing data yang paling dekat dengan pusat cluster dengan persamaan (2.1).
- Menghitung masing-masing centroid untuk menentukan centroid baru dengan cara menghitung nilai rata-rata setiap data yang terdapat dalam centroid dengan menggunakan rumus (2.2).
- Jika centroid baru sudah didapatkan, maka ulangi langkah dua, kemudian menuju langkah ke lima.
- Mendapatkan hasil klaster.

Tabel 5. Hasil Klaster I Data satuan PAUD

NO	ID	NAMA LEMBAGA	CLUSTER
1	1	KB TUNAS HATI	C1
2	2	KB GANESHA MULYA	C1
3	8	POS PAUD SEJAHTERA 2	C1
4	11	POS PAUD MUTIARA KASIH	C1
5	18	TK HARAPAN BARU	C1
6	23	TK PAGILARAN II	C1

Tabel 6. Hasil Klaster 2 Data satuan PAUD

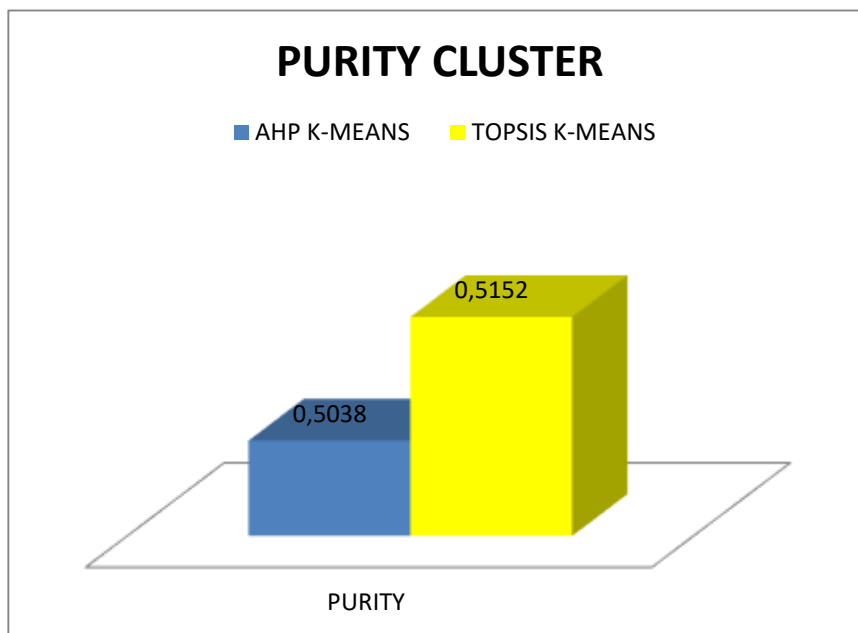
NO	ID	NAMA LEMBAGA	CLUSTER
1	3	KB PELITA ANANDA	C2
2	4	KB TUNAS BANGSA	C2
3	2	POS PAUD ANGGREK 2	C2
4	7	POS PAUD SEJAHTERA 1	C2
5	1	TK ABA SODONG	C2
6	2	TK PERTIWI BANDAR	C2

Tabel 7. Hasil Kluster 3 Data Satuan PAUD

NO	ID	NAMA LEMBAGA	CLUSTER
1	13	KB AL KAROMAH	C3
2	14	KB NURUL IKHSAN	C3
3	9	POS PAUD SEJAHTERA 3	C3
4	13	POS PAUD CERIA	C3
5	2	TK TAMA BAKTI	C3
6	3	TK PERINTIS	C3

#### 4.5. Pengujian Purity Validasi

Penggunaan purity validasi dalam penelitian ini bertujuan untuk memeriksa kemurnian dari setiap cluster yang terdapat dalam penelitian ini. Apabila memperoleh nilai mendekati 1 pada setiap cluster dalam penelitian ini maka cluster tersebut dinyatakan lebih baik dan maksimal.



Gambar 3. Purity Validasi Data KB

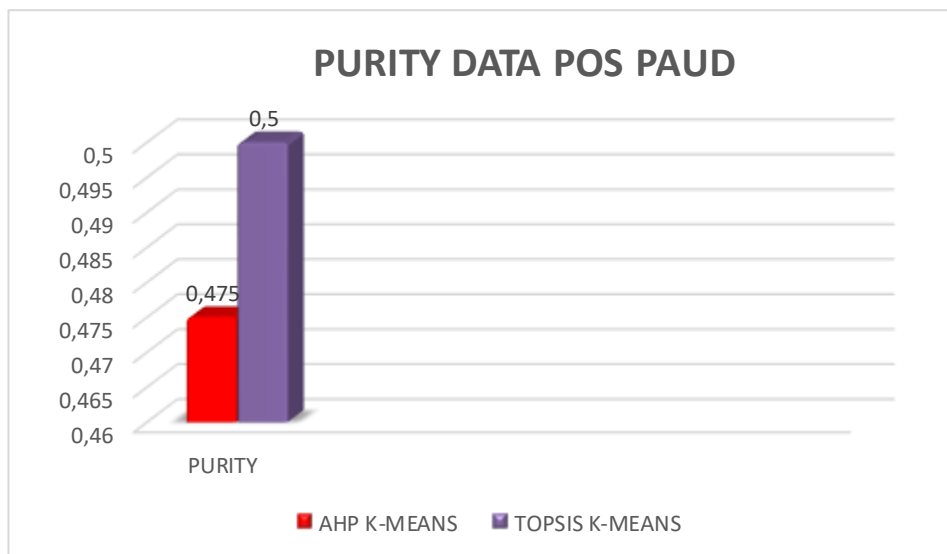
Dari gambar 4.2 diperoleh hasil perhitungan purity data KB dengan menggunakan metode *AHP K-Means* yang mendapatkan nilai purity sebesar 0.5038 sedangkan dalam proses pengolahan data KB menggunakan metode *TOPSIS K-MEANS* mendapatkan nilai purity sebesar 0,5152.

Tabel 8. Purity Data KB dengan Metode *AHP K-Means*

KB					
CLUSTER	JUMLAH	t1 (min)	t2 (max)	t3 (average)	Purity
C1	25	1	15	9	0.6000
C2	30	1	15	14	0.5000
TOTAL	78	1	37	40	0.4744
PURITY KB	133				0.5038

Tabel 9. Purity Data KB dengan Menggunakan Metode K-Means TOPSIS

KB					
CLUSTER	JUMLAH	t1 (min)	t2 (max)	t3 (average)	Purity
C1	43	1	22	20	0.5116
C2	58	1	30	27	0.5172
C3	31	1	16	14	0.5161
TOTAL	132				
PURITY KB					0.5152



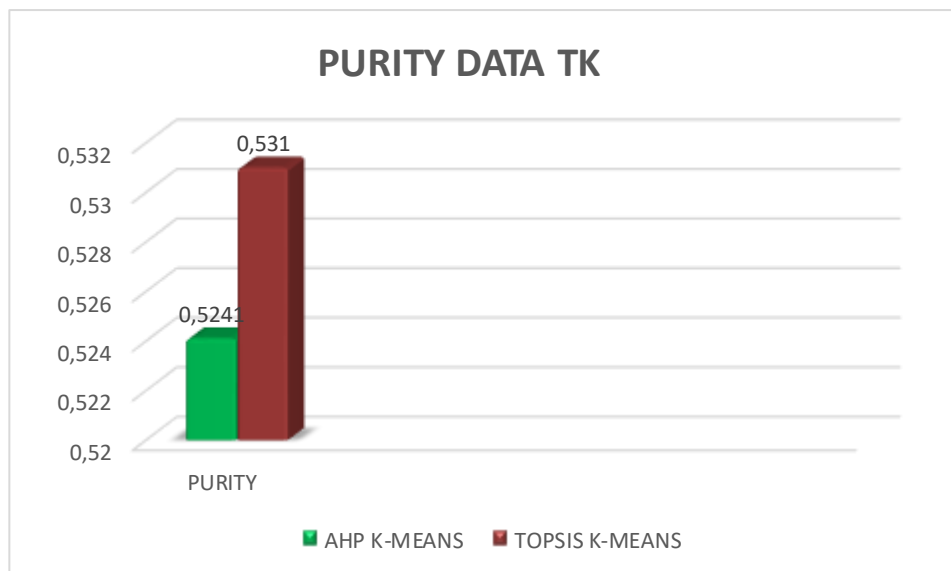
Gambar 4. Purity Validasi Data POS PAUD

Tabel 10. Purity Data POS PAUD dengan Menggunakan Metode AHP K-Means

POS PAUD					
CLUSTER	JUMLAH	t1 (min)	t2 (max)	t3 (average)	Purity
C1	9	1	4	4	0.4444
C2	14	1	7	6	0.5000
C3	17	1	8	8	0.4706
TOTAL	40				
PURITY POS PAUD					0.4750

Tabel 11. *Purity* Data POS PAUD dengan Menggunakan Metode *TOPSIS K-Means*

POS PAUD					
CLUSTER	JUMLAH	t1 (min)	t2 (max)	t3 (average)	Purity
C1	15	1	8	6	0.5333
C2	12	1	7	4	0.5833
C3	13	1	5	7	0.3846
TOTAL	40				
PURITY POS PAUD					0.5000



Gambar 5. Purity Validasi Data TK

Tabel 12. *Purity* Data TK dengan menggunakan metode *AHP K-Means*

TK					
CLUSTER	JUMLAH	t1 (min)	t2 (max)	t3 (average)	Purity
C1	31	1	17	13	0.5484
C2	87	1	45	41	0.5172
C3	27	1	14	12	0.5185
TOTAL	145				
PURITY TK					0.5241

Tabel 13. *Purity* Data TK dengan Menggunakan Metode *TOPSIS K-Means*

TK					
CLUSTER	JUMLAH	t1 (min)	t2 (max)	t3 (average)	Purity
C1	37	1	16	20	0.5405
C2	45	1	22	22	0.4889
C3	63	1	35	27	0.5556
TOTAL	145				
PURITY TK					0.5310

#### 4.6. Analisis Hasil Pengujian

Dari pengujian dalam prioritas penentuan dana bantuan operasional satuan PAUD dengan menggunakan algoritma *K-Means* dan *TOPSIS* yang telah dilakukan dihasilkan:

- a. Dalam pengujian data satuan PAUD menggunakan metode *KNN* yang digunakan untuk mengelompokkan data dan menentukan data lembaga yang berhak mendapatkan dan tidak berhak mendapatkan dana bantuan operasional pendidikan dapat berjalan dengan efisien .
- b. Penggunaan metode *TOPSIS* dalam setiap data lembaga satuan PAUD bertujuan untuk mencari nilai bobot matrik ternormalisasi dan mencari nilai preferensi yang akan digunakan sebagai pedoman dalam penentuan perankingan setiap data satuan PAUD. Dalam hal ini nilai preferensi tertinggi dari data KB yaitu sebesar 0,7717211 yang diperoleh oleh KB Assaid, sedangkan untuk nilai preferensi tertinggi data POS PAUD, SPS diperoleh lembaga POS PAUD Melati sebesar 0,733014368, dan untuk nilai preferensi tertinggi data TK yaitu sebesar 0,79383048 diperoleh lembaga TK Al Amin, didapatkannya nilai preferensi ini sebagai pedoman dalam penentuan prioritas dana bantuan operasional pendidikan.
- c. Penggunaan pengujian *purity* dalam metode *AHP K-Means* dan *TOPSIS K-Means* menunjukkan bahwa kemurnian suatu cluster dalam data satuan PAUD yang memiliki nilai *purity* paling optimal sebesar 0,5152 diperoleh pada data KB pada saat menggunakan metode *TOPSIS K-Means*, 0,5000 pada data POS PAUD, SPS pada metode *TOPSIS K-Means*, dan 0,5310 pada data TK dengan menggunakan metode *TOPSIS K-Means*.

Dari pengujian yang sudah dilakukan dalam penelitian ini dengan menggunakan data satuan PAUD menggunakan metode *TOPSIS K-Means* diketahui bahwa pengelompokan data satuan PAUD mendapatkan nilai *purity* paling baik dan optimal daripada penelitian sebelumnya yang menggunakan metode *AHP K-Means*.

## 5. KESIMPULAN DAN SARAN

### 5.1. Kesimpulan

Dari hasil penelitian mengenai prioritas penentuan dana bantuan operasional pada satuan PAUD dengan menggunakan metode *KNN*, *TOPSIS*, dan *K-Means* dapat disimpulkan hal-hal sebagai berikut:

- a. Penggunaan metode *KNN*, *TOPSIS*, dan *K-Means* dalam penelitian ini mendapatkan hasil yang akurat dan tepat dalam menentukan prioritas penerima dana bantuan operasional pendidikan lembaga satuan PAUD.
- b. Dengan penggunaan kombinasi metode *KNN*, *TOPSIS* dan *K-Means* dapat menunjang proses prioritas penerima dana bantuan operasional pendidikan pada lembaga satuan PAUD tepat sasaran, sehingga tidak terjadi penyelewengan dalam penerimaan dana bantuan operasional pendidikan .

## 5.2. Saran

Setelah mengadakan penelitian dan mengamati masalah yang timbul dalam prioritas penentuan dana bantuan operasional pendidikan pada satuan PAUD, penulis ingin memberikan saran antara lain :

- a. Dalam menentukan pemberian dana bantuan operasional pendidikan pada satuan PAUD, Dinas Pendidikan, Pemuda dan Olahraga harus lebih berhati-hati dan waspada dalam memberikan keputusan dalam penentuan dana operasional pendidikan .
- b. Metode *KNN*, *TOPSIS* dan *K-Means* dapat digunakan untuk menentukan prioritas penentuan dana bantuan operasional pendidikan, tetapi juga dapat digunakan untuk menyelesaikan masalah di bidang lain terutama dalam bidang penentuan dana bantuan.

## PERNYATAAN ORISINALITAS

“Saya menyatakan dan bertanggung jawab dengan sebenarnya bahwa artikel ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya”

[Diwahana Mutiara Candrasari].

## DAFTAR PUSTAKA

- [1] A. K. Jain, Jun. 2010, *Data clustering: 50 years beyond K-means*, Pattern Recognition Letters , vol. 31, no. 8, pp. 651–666.
- [2] D. Sameer Kumar, S. Radhika and K. N. S.Suman. (2013). *MADM Methods for Finding The Right Personnel in Academic Institutions*. International Journal of u- and e- Service, Science and Technology , Vol.6, No.5 (2013), pp.133-144.
- [3] Fitriana Nur A, et.al. (Februari 2015 – April 2015). *Sistem Pendukung Keputusan Untuk Menentukan Prestasi Akademik Siswa dengan Metode TOPSIS*. Citec Journal, ISSN: 2354-5771 , Vol. 2, No. 2
- [4] Gjorgji Madjarov, Dragi Kocev, Dejan Gjorgjevikj, and SasO Dzeroski. *An extensive experimental comparison of methods for multi-label learning*. Pattern Recogn., 45(9):3084{310}.
- [5] Grigorios Tsoumakas, Ioannis Katakis, and Ioannis Vlahavas. *Random k-labelsets for multilabel classification*. IEEE Transactions on Knowledge and Data Engineering, 23(7):1079{1089, 2011 }
- [6] Hwang C.L, Yoon K . (1981). *Multiple Attribute Decision Making Methods and Application* . New York : Springer - Verlag .
- [7] Ignatius Joshua , et.al . (ICIC International @ 2012 ISSN 1349-4198). *Modelling Funding Allocatin Problems Via AHP-FUZZY TOPSIS* International Journal of Innovative Computing, Information and Control , Volume 8, Number 5(A), May 2012.
- [8] Irhamni Firlil, Damayanti Fitri, et.al. *Optimalisasi Pengelompokan Kecamatan Berdasarkan Indikator Pendidikan Menggunakan Metode Clustering dan Davies Bouldin Index* .
- [9] I Wayan Supriana, 2012, *Sistem Pendukung Keputusan Dalam Pemilihan Tempat Kost Dengan Metode Pembobotan ( Studi Kasus : Sleman Yogyakarta)*Program Pascasarjana Ilmu Komputer Fakultas MIPA Universitas Gadjah Mada , Jurnal Ilmu Komputer-Volume 5 - No 2– September 2012
- [10] Muzakir, A. (2014). *Analisa Dan Pemanfaatan Algoritma K-Means Clustering Pada Data Nilai Siswa Sebagai Penentuan Penerima Beasiswa*.Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2014 , ISSN: 1979-911X.
- [11] Nor Ashikin Mohamed Yusof, S. H. (2013). *Analytical Hierarchy Process in Multiple Decisions Making for Higher Education in Malaysia* . Procedia - Social and Behavioral Sciences , 389 – 394 .
- [12] Pang Ning Tan, Michael Steinbach, Vipin Kumar . (2006). *Introduction to Data Mining*. Addison Wesley.



- [13] R.Listia dan A. Harjoko, *Klasifikasi Massa pada Citra Mamogram Berdasarkan Gray Level Coocurrence Matrix (GLCM)*, IJCSS, vol 8, no. 1, pp. 59-68, Januari 2014
- [14] S. Chormunge., S. Jena . (2014). *Evaluation of Clustering Algorithms for High Dimensional Data Based on Distance Functions*. in Proceeding of th3 2014 International Conference on Information and Communication Technology for Competitive Strategies , Article No. 78 .
- [15] Safei, Suhailan, et .all (2015). *Targeted Ranking-Based Clustering Using AHP K-Means* . Int. J. Advance Soft Compu. Appl, Vol. 7, No. 3, November 2015 . ISSN 2074-8523
- [16] Singh A, Yadav, and Rana A. (2013). *K-Means with Three Different Distance Metrics* . International Journal of Computer Application, vol 67, no 10, pp. 13-17.
- [17] T. Rismawan dan S. Kusumadewi, *Aplikasi K-Means Untuk Pengelompokkan Mahasiswa Berdasarkan Nilai Body Mass Index & Ukuran Kerangka*, SNATI, pp. 43-48, Juni 2008
- [18] Tukan, E. A., & Kennedy, J. (2016). *Sistem Pendukung Keputusan Prestasi Siswa*. Seminar Nasional Teknologi Informasi dan Multimedia 2016 .
- [19] Turban, 2005, *Decision Support System and Intelligent System (Terjemahan: Sistem Pendukung Keputusan dan Sistem Cerdas) Jilid 1*, Andi Offset, Yogyakarta
- [20] Zagoto, D. (ISSN : 2301-9425 , 2015). *Sistem Pendukung Keputusan Seleksi Penerima Bantuan Siswa Miskin (BSM) Pada Sekolah Yayasan Pendidikan Berlianta (SD) Marendal Menggunakan Metode TOPSIS*. Pelita Informatika Budi Darma , Volume : IX, Nomor: 3, April 2015.