

Klasterisasi Tingkat Pendidikan Masyarakat Jawa Timur Menggunakan FCM (*Fuzzy-C-Means*) Clustering

Ngatmari

Prodi. Informatika Fak. Teknik Universitas Islam Malang

Email: ngatmari@unisma.ac.id

ABSTRAK

Dalam menunjang program pemerintah bidang pendidikan, perlu dikaji lebih dalam mengenai partisipasi tingkat pendidikan suatu wilayah. Hasil penelitian ini bisa digunakan sebagai acuan pemerintah untuk meningkatkan partisipasi masyarakat dalam bidang pendidikan agar tidak terjadi ketimpangan antara wilayah satu dengan wilayah yang lain. Data wilayah yang akan dilakukan pemetaan adalah wilayah Jawa Timur dengan menggunakan data sekunder yang diperoleh dari BPS Provinsi Jawa Timur Tahun 2023. Analisis menggunakan algoritma *Fuzzy-c-means clustering* dan untuk mengukur kualitas *clustering* ini menggunakan metode *xie beni indekx* (XBI). Hasil pengujian terhadap seluruh data wilayah yang terdapat 38 wilayah terbagi menjadi 3 klaster dengan *cluster validity* sebesar 0.072178

Kata kunci: *Clustering, Fuzzy-c-means, Xie beni Index, Datamining.*

ABSTRACT

In supporting government programs in the field of education, it is necessary to study more deeply about the participation of the education level of a region. The results of this study can be used as a reference for the government to increase community participation in the education sector so that there is no imbalance between one region and another. The area data to be mapped is the East Java region using secondary data obtained from BPS East Java Province 2023. The analysis uses the Fuzzy-c-means clustering algorithm and to measure the quality of this clustering using the xie beni index (XBI) method. The test results of all regional data contained 38 regions divided into 3 clusters with cluster validity 0.072178.

Keywords: *Clustering, Fuzzy-c-means, Xie beni Index, Data mining.*

I. PENDAHULUAN

Sumber Daya Manusia (SDM) merupakan aset utama dalam membangun suatu bangsa. Ketersediaan sumber daya alam (*natural resource*) yang berlimpah dan adanya sumber daya modal dan teknologi yang semakin canggih, tidak akan mempunyai kontribusi apa pun tanpa di dukung oleh adanya sumber daya manusia (*human resource*) yang berkualitas. Peningkatan kualitas suatu bangsa sesungguhnya bertumpu pada peningkatan kualitas sumber daya manusianya, yang salah satunya adalah meningkatkan dalam bidang pendidikan. Artinya, pendidikan merupakan kontribusi yang sangat berharga dan signifikan dalam meningkatkan kualitas suatu bangsa terutama bangsa Indonesia.

Pendidikan merupakan salah satu indikator penting dalam mengukur tingkat kesejahteraan masyarakat serta kemajuan suatu wilayah. Tingkat pendidikan yang berbeda-beda pada masyarakat dapat memberikan gambaran mengenai kualitas sumber daya manusia di suatu daerah, termasuk di Provinsi Jawa Timur. Oleh karena itu, diperlukan suatu analisis yang mampu mengelompokkan tingkat pendidikan masyarakat secara sistematis agar dapat menjadi dasar dalam pengambilan kebijakan.

Untuk mengoptimalkan kontribusi pendidikan tersebut terhadap peningkatan kualitas suatu bangsa, semua pihak yang terkait (termasuk *stakeholders*) mempunyai kontribusi yang penting, disamping pengelola pendidikan itu sendiri. Menurut Undang-undang Sistem Pendidikan Nasional Nomor 20 Tahun 2003 (UU SISDIKNAS 20 Th 2003)(Nasional 2007) pasal 54, 55,56.

Pada pasal 54, mengatur bentuk dan ruang lingkup peran serta masyarakat sebagai berikut:

1. Peran serta masyarakat dalam pendidikan meliputi peran serta perseorangan, kelompok, keluarga, organisasi profesi, pengusaha, dan organisasi kemasyarakatan dalam penyelenggaraan dan pengendalian mutu pelayanan pendidikan.
2. Masyarakat dapat berperan serta sebagai sumber, pelaksana, dan pengguna hasil pendidikan.

Pasal 55 UU Sisdiknas nomor 20 tahun 2003 mengatur prinsip-prinsip pendidikan berbasis masyarakat. Dalam pasal ini ditetapkan bahwa:

- 1) Masyarakat berhak menyelenggarakan pendidikan berbasis masyarakat pada pendidikan formal dan nonformal sesuai dengan kekhasan agama, lingkungan sosial, dan budaya untuk kepentingan masyarakat.
- 2) Penyelenggara pendidikan berbasis masyarakat mengembangkan dan melaksanakan kurikulum dan evaluasi

pendidikan, serta manajemen dan pendanaannya sesuai dengan standar nasional pendidikan.

- 3) Dana penyelenggaraan pendidikan berbasis masyarakat dapat bersumber dari penyelenggara, masyarakat, Pemerintah, pemerintah daerah dan/atau sumber lain yang tidak bertentangan dengan peraturan perundang-undangan yang berlaku.
- 4) Lembaga pendidikan berbasis masyarakat dapat memperoleh bantuan teknis, subsidi dana, dan sumber daya lain secara adil dan merata dari Pemerintah dan/atau pemerintah daerah.

Terkait dengan undang-undang yang telah disebutkan diatas, maka idealnya pemerintah harus terus berupaya agar tingkat pendidikan yang ada di setiap wilayah tidak terjadi ketimpangan, dimana pada suatu wilayah tingkat partisipasi masyarakat dalam pendidikan itu terlalu tinggi tapi disisi lain terdapat wilayah tertentu yang tingkat partisipasi masyarakat dalam bidang pendidikan itu terlalu rendah. Sehingga ini merupakan bentuk tidak meratanya tingkat pendidikan yang ada di Indonesia.

Pendidikan yang diterapkan di Indonesia sekarang ini adalah pendidikan dimulai dari Pendidikan Anak Usia Dini (PAUD) hingga jenjang yang lebih tinggi yaitu Universitas. Adapun pentingnya PAUD, seperti yang telah diuraikan pada url <http://paud.kemdikbud.go.id/2016/03/30/memahami-pendidikan-anak-usia-dini/>, Pendidikan Anak Usia Dini (PAUD) adalah salah satu upaya pembinaan yangditujukan kepada anak sejak lahir sampai dengan usia enam tahun, yang dilakukan melalui pemberian rangsangan pendidikan untuk membantu pertumbuhan dan perkembangan jasmani dan rohani agar anak memiliki kesiapan dalam memasuki pendidikan lebih lanjut.

Beberapa dasar hukum yang berkaitan dengan Pendidikan Anak Usia Dini (PAUD) adalah sebagai berikut:

1. Pembukaan UUD 1945; Salah satu tujuan kemerdekaan adalah mencerdaskan kehidupan bangsa.
2. Amandemen UUD 1945 pasal 28 C
Setiap anak berhak mengembangkan diri melalui pemenuhan kebutuhan dasarnya, berhak mendapat pendidikan dan memperoleh manfaat dari ilmu pengetahuan dan teknologi, seni dan budaya demi meningkatkan kualitas hidupnya dan demi kesejahteraan umat manusia.
3. UU No. 23/2002 Tentang Perlindungan Anak Pasal 9 ayat (1)
Setiap anak berhak memperoleh pendidikan dan pengajaran dalam rangka pengembangan

pribadinya dan tingkat kecerdasannya sesuai dengan minat dan bakat.

4. UU No 20/2003 pasal 28

- 1) Pendidikan anak usia dini diselenggarakan sebelum jenjang pendidikan dasar.
- 2) Pendidikan anak usia dini dapat diselenggarakan melalui jalur pendidikan formal, non formal, dan/atau informal.
- 3) Pendidikan anak usia dini pada jalur pendidikan formal berbentuk Taman Kanak-Kanak (TK), Raudhatul Athfal (RA), atau bentuk lain yang sederajat.
- 4) Pendidikan anak usia dini pada jalur pendidikan non formal berbentuk kelompok bermain (KB), Taman Penitipan Anak (TPA), atau bentuk lain yang sederajat.
- 5) Pendidikan anak usia dini pada jalur informal berbentuk pendidikan keluarga atau pendidikan yang diselenggarakan oleh lingkungan.

Sedemikian pentingnya pendidikan sehingga pemerintah membuat payung hukum tentang pendidikan, hal ini agar pendidikan di Indonesia bisa merata. Namun demikian, tentunya harus diimbangi dengan sistem kontrol dalam mendampingi antara kebijakan dan realitas yang terjadi di masyarakat apakah pemerataan pendidikan itu sudah benar-benar merata atau belum. Mengingat letak geografis Negara Indonesia yang terdiri dari beberapa pulau, tentunya tidak mudah bagaimana memonitor setiap wilayah dalam pemerataan pendidikan ini tanpa ada kerjasama yang baik antara pemerintah pusat dan pemerintah daerah.

Dan tidak hanya itu, peran serta masyarakat dalam mendukung program pemerintah pun juga harus turut andil minimal mendukung anggota keluarganya yang memiliki usia wajib belajar untuk memacu agar menempuh pendidikan baik mulai tingkat dasar sampai jenjang yang lebih tinggi.

Dalam tingkat pendidikan dasar yang dimaksud diluar Pendidikan Anak Usia Dini, yang lebih ditekankan lagi adalah pendidikan tingkat Sekolah Dasar, tujuan utama pendidikan dasar adalah membantu peserta didik dalam mengembangkan dirinya sehingga menjadi manusia yang intelektualis yang tinggi, mentalitas yang baik, kemandirian, berjiwa sosial, kreatif, dan siap menghadapi perubahan-perubahan (Collier, C.; Houston, J.; Schmatz, R. R.; Walsh 1971).

Sekolah merupakan lembaga formal yang berfungsi sebagai “mitra kerja” keluarga dan masyarakat dalam melaksanakan tugasnya

membentuk warga masyarakat yang diinginkan (ARIKUNTO 19913). Untuk itu, sekolah tidak dapat dipisahkan dari masyarakat. Sekolah tidak dapat terpisahkan dari masyarakat, tetapi sekolah justru diurus oleh masyarakat, hal ini menunjukkan bahwa, sekolah merupakan bagian integral dari masyarakat (Dr. Panuti Sudjiman dan Greta Librata 1989). Meskipun hubungan antara sekolah dan masyarakat demikian erat, masyarakat tidak langsung berperan dalam kegiatan operasional sekolah. Interaksi antara sekolah dan masyarakat merupakan hubungan yang khas. Masyarakat mempengaruhi sekolah secara informal (Snowden and Gorton 1998). Kelompok orang tua murid mengadakan kontak secara individual dengan sekolah. Hubungan antara sekolah dan masyarakat akan berjalan lancar apabila terdapat wadah yang terdiri atas wakil-wakil dari sekolah (kepala sekolah dan guru guru), wakil-wakil orang tua siswa, dan wakil-wakil masyarakat, yang selalu mengadakan kegiatan komunikasi secara rutin. Untuk membahas dan mengevaluasi kegiatan sekolah, partisipasi dari unsur-unsur tersebut sangat mempengaruhi kinerja organisasi wadah tersebut.

Sebagai bagian dari sistem yang lebih besar, sekolah sebagai suatu organisasi tidak bisa melepaskan diri dari pengaruh sistem-sistem lain yaitu yang secara terus menerus berubah. Hal ini berarti organisasi berada dalam lingkungan yang selalu berubah (Stoner 1982). Keadaan ini menuntut organisasi sekolah untuk menyesuaikan diri dengan perubahan lingkungan itu. Upaya penyesuaian diri tersebut dapat dilakukan melalui pengembangan organisasi (*organizational development*). Pengembangan organisasi sekolah dapat dilakukan dengan cara mengadopsi ide-ide, model, cara, atau metode baru guna meningkatkan keefektifan organisasi (Owens 1991).

Seiring dengan perkembangan teknologi, pemanfaatan teknik *data mining* menjadi salah satu solusi dalam mengolah data dalam jumlah besar untuk informasi yang bernilai. Salah satu teknik dalam data mining yang sering digunakan adalah *clustering*, yaitu proses pengelompokan data berdasarkan kemiripan karakteristik tertentu tanpa adanya label sebelumnya. Metode ini sangat relevan digunakan untuk mengidentifikasi pola tingkat pendidikan masyarakat yang bersifat heterogen.

Dalam penelitian ini, metode *Fuzzy C-Means* (FCM) digunakan karena memiliki keunggulan dibandingkan metode *clustering* konvensional seperti *K-Means*. FCM memungkinkan setiap data memiliki derajat keanggotaan pada lebih dari satu *cluster*, sehingga lebih sesuai untuk data sosial yang cenderung tidak memiliki batasan yang tegas. Dengan pendekatan ini, hasil pengelompokan diharapkan dapat

memberikan representasi yang lebih fleksibel dan realistis terhadap kondisi tingkat pendidikan masyarakat.

Selain itu, untuk mengevaluasi kualitas hasil clustering, digunakan *Xie-Beni Index* (XBI) yang mampu mengukur tingkat kepadatan (*compactness*) dan pemisahan (*separation*) antar *cluster*. Penggunaan indeks ini bertujuan untuk menentukan jumlah *cluster* yang optimal sehingga hasil pengelompokan menjadi lebih valid dan dapat dipertanggungjawabkan.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk melakukan klasterisasi tingkat pendidikan masyarakat di Jawa Timur menggunakan metode *Fuzzy C-Means* serta mengevaluasi hasilnya menggunakan *Xie-Beni Index* (XBI). Hasil dari penelitian ini diharapkan dapat menjadi referensi bagi pemerintah atau pihak terkait dalam merumuskan kebijakan yang tepat dalam bidang pendidikan.

II. TINJAUAN PUSTAKA

A. Clustering

Clustering atau klasterisasi adalah metode pengelompokan data. Menurut (Tan n.d.) *clustering* adalah sebuah proses untuk mengelompokkan data ke dalam beberapa cluster atau kelompok sehingga data dalam satu cluster memiliki tingkat kemiripan yang maksimum dan data antar *cluster* memiliki kemiripan yang minimum. Suatu cluster merupakan sekelompok entitas yang memiliki kesamaan dan memiliki perbedaan dengan entitas dari kelompok lain (Everitt, Landau, and Leese 2001).

Algoritma *Clustering* bekerja dengan mengelompokkan obyek-obyek data (pola, entitas, kejadian, unit, hasil observasi) ke dalam sejumlah *cluster* tertentu (Xu and Wunsch 2009). Dengan kata lain algoritma *Clustering* melakukan pemisahan / pemecahan / segmentasi data ke dalam sejumlah kelompok (*cluster*) menurut karakteristik tertentu.

Metode clustering pada dasarnya mengoptimalkan pusat cluster (*Centroid*) (Kusumadewi and Purnomo 2010). Beberapa metode clustering yang sering digunakan antara lain, yaitu: 1) Berbasis Metode Statistik seperti *Hierarchical clustering method* dan *Non Hierarchical clustering method*, 2) Berbasis *Fuzzy*: *Fuzzy C-Means*, 3) Berbasis *Neural Network*: *Kohonen SOM*, *LVQ*.

Metode lain untuk optimasi *centroid* atau lebar *cluster* seperti *Genetik Algoritma* (GA) (Satriyanto 2010). *Fuzzy C-Means Clustering* (FCM) atau dikenal juga sebagai *Fuzzy ISODATA*, merupakan

salah satu metode *clustering* yang merupakan bagian dari metode *Hard K-Means*. FCM menggunakan model pengelompokan *fuzzy* sehingga data dapat menjadi anggota dari semua kelas atau cluster terbentuk dengan derajat atau tingkat keanggotaan yang berbeda antara 0 hingga 1. Tingkat keberadaan data dalam suatu kelas atau cluster ditentukan oleh derajat keanggotaannya. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981 (Kusumadewi and Purnomo 2010).

B. Fuzzy-C-Means

Fuzzy C-Means (FCM) adalah salah satu teknik peng-cluster-an data yang mana keberadaan tiap-tiap titik data dalam suatu cluster ditentukan oleh derajat keanggotaannya. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981 (Kusumadewi et al. 2006).

Fuzzy Cluster Means (FCM) merupakan algoritma yang digunakan untuk melakukan clustering data sesuai berdasarkan keberadaan tiap-tiap titik data sesuai dengan derajat keanggotaannya (Ahmadi and Hartati 2015).

Menurut (Kusumadewi and Purnomo 2010) *Fuzzy clustering* adalah suatu teknik penentuan cluster optimal dalam suatu vektor yang didasarkan pada bentuk normal euclidian untuk jarak vektor. Menurut (Kusumadewi et al. 2006) *Fuzzy C-means* adalah suatu metode pengelompokkan data dimana tiap-tiap data berada dalam suatu kelompok ditentukan nilai keanggotaan. Menurut karim (2011) *Fuzzy C-Means* adalah suatu metode pengelompokkan yang memungkinkan satu bagian dari data untuk memiliki dua atau lebih kelompok.

Konsep dasar pada *Fuzzy C-Means* adalah menentukan pusat cluster, yang akan menandai lokasi rata-rata untuk tiap cluster. Tiap-tiap titik data memiliki derajat keanggotaan untuk tiap cluster yang terbentuk. Pada kondisi awal pusat cluster masih belum akurat, maka dari itu dilakukan perbaikan pusat cluster dan derajat keanggotaan tiap-tiap titik data secara berulang hingga berada pada titik yang tepat. Perulangan ini didasarkan pada minimasi fungsi objektif yang menggambarkan jarak dari titik data yang diberikan ke pusat cluster yang terbobot oleh derajat keanggotaan titik data tersebut. Dari perulangan tersebut dapat dilihat bahwa semakin lama pusat cluster akan bergerak menuju lokasi yang tepat (Kusumadewi et al. 2006).

C. Xie beni

Salah satu dari beberapa teknik dalam validasi clustering diantaranya adalah menggunakan metode *Xie Beni Index* (XBI). Sesuai dengan namanya Indeks

XB ditemukan oleh Xie dan Beni yang pertama kali dikemukakan pada tahun 1991. Ukuran kevalidan cluster merupakan proses evaluasi hasil *clustering* untuk menentukan *cluster* mana yang terbaik. Ada dua kriteria dalam mengukur kevalidan suatu *cluster*, yaitu: 1. *Compactness*, yaitu ukuran kedekatan antaranggota pada tiap *cluster*. 2. *Separation*, yaitu ukuran keterpisahan antar *cluster* satu dengan *cluster* yang lainnya. Secara umum, nilai yang terbaik XBI adalah nilai indeks yang semakin kecil, semakin kecil nilai XBI maka hasil pengelompokan semakin baik (Zhang et al. 2008).

III. IDENTIFIKASI MASALAH

Berdasarkan tinjauan pustaka tersebut diatas dapat diidentifikasi kesenjangan pengetahuan sebagai berikut:

1. Terdapat ketimpangan tingkat pendidikan antar kabupaten/kota di Jawa Timur yang belum terpetakan secara optimal.
2. Metode analisis konvensional belum mampu mengelompokkan wilayah berdasarkan kemiripan karakteristik pendidikan secara komprehensif.
3. Metode *clustering* klasik seperti *K-Means* memiliki keterbatasan dalam menangani data yang bersifat tidak tegas (*overlapping*).
4. Diperlukan metode *clustering* yang mampu merepresentasikan derajat keanggotaan data secara fleksibel, yaitu *Fuzzy C-Means*.
5. Belum adanya evaluasi yang optimal terhadap jumlah klaster yang terbentuk, sehingga diperlukan metode validasi seperti *Xie-Beni Index*.

Berdasarkan identifikasi tersebut, permasalahan utama dalam penelitian ini adalah bagaimana mengelompokkan tingkat pendidikan masyarakat di Provinsi Jawa Timur secara optimal menggunakan metode *Fuzzy C-Means* serta menentukan jumlah klaster terbaik menggunakan *indeks Xie-Beni*.

IV. METODE PENELITIAN

A. Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining untuk mengekstraksi pola dari data tingkat pendidikan masyarakat di Jawa Timur. Teknik yang digunakan adalah *clustering*, yaitu proses pengelompokan data berdasarkan tingkat kemiripan karakteristik tanpa menggunakan label kelas tertentu. Metode *clustering* yang diterapkan dalam penelitian ini adalah *Fuzzy C-Means* (FCM), yang memungkinkan setiap data

memiliki derajat keanggotaan pada lebih dari satu cluster.

B. Sumber dan Jenis Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang telah dikumpulkan dan disimpan dalam format Microsoft Excel (.xlsx). Data tersebut merupakan hasil penggalan data sebelumnya yang berisi informasi terkait tingkat pendidikan masyarakat di wilayah Jawa Timur.

Atribut yang digunakan dalam penelitian ini meliputi beberapa indikator tingkat pendidikan, seperti:

- a. Persentase tingkat pendidikan dasar
- b. Persentase tingkat pendidikan menengah
- c. Persentase tingkat pendidikan tinggi
- d. Variabel pendukung lainnya yang relevan

Data yang digunakan bersifat numerik sehingga sesuai untuk proses *clustering* menggunakan metode FCM.

C. Tahapan Penelitian

Metodologi penelitian ini mengadopsi tahapan dalam *Knowledge Discovery in Databases* (KDD) yang terdiri dari beberapa langkah sebagai berikut:

1. Data Selection

Tahap ini dilakukan dengan memilih data yang relevan dari file Excel hasil penggalan data sebelumnya. Data yang dipilih adalah data yang berkaitan langsung dengan indikator tingkat pendidikan masyarakat.

2. Data Preprocessing

Pada tahap ini dilakukan pembersihan data (data cleaning), meliputi:

- Penanganan data yang hilang (*missing value*)
- Penghapusan data duplikat
- Penyesuaian format data

Selain itu, dilakukan proses normalisasi data untuk menyamakan skala antar atribut agar tidak terjadi dominasi variabel tertentu dalam proses *clustering*.

3. Data Transformation

Data yang telah dibersihkan kemudian ditransformasikan ke dalam bentuk yang sesuai untuk proses data mining, khususnya *clustering*. Proses ini meliputi:

- Normalisasi menggunakan metode *Min-Max* atau *Z-score*
- Penyusunan data dalam bentuk matriks numerik

D. Proses Clustering dengan Fuzzy C-Means

Tahapan utama dalam penelitian ini adalah proses *clustering* menggunakan algoritma *Fuzzy C-Means*

(FCM). Langkah-langkahnya meliputi:

1. Menentukan jumlah *cluster* (c)
2. Menentukan nilai parameter *fuzziness* (m)
3. Menginisialisasi matriks keanggotaan awal
4. Menghitung pusat *cluster* (*centroid*)
5. Memperbarui derajat keanggotaan setiap data
6. Melakukan iterasi hingga kondisi konvergen tercapai

Metode FCM dipilih karena mampu menangani data yang bersifat tidak tegas (*fuzzy*), sehingga lebih sesuai untuk data sosial seperti tingkat pendidikan masyarakat.

E. Evaluasi *Cluster* menggunakan *Xie-Beni Index*

Untuk mengukur kualitas hasil *clustering*, digunakan *Xie-Beni Index*. Indeks ini digunakan untuk mengevaluasi:

- Tingkat kepadatan *cluster* (*compactness*)
- Tingkat pemisahan antar *cluster* (*separation*)

Nilai *Xie-Beni* yang lebih kecil menunjukkan hasil *clustering* yang lebih optimal. Evaluasi ini juga digunakan untuk menentukan jumlah *cluster* terbaik.

F. Interpretasi Hasil

Tahap akhir adalah interpretasi hasil *clustering*, yaitu:

- Menganalisis karakteristik masing-masing *cluster*
- Mengelompokkan wilayah berdasarkan tingkat pendidikan (rendah, sedang, tinggi)
- Menarik kesimpulan sebagai dasar pendukung keputusan bagi pihak terkait

V. HASIL DAN PEMBAHASAN

A. Hasil Penelitian

Berdasarkan data yang diperoleh, penelitian ini mengungkap bagaimana mengelompokkan tingkat pendidikan masyarakat di Provinsi Jawa Timur secara optimal menggunakan metode *Fuzzy C-Means* serta menentukan jumlah kluster terbaik menggunakan *indeks Xie-Beni*.

1. Algoritma *Fuzzy C-Means*

Pada kondisi awal, pusat kluster ini masih belum akurat. Tiap-tiap data memiliki derajat keanggotaan untuk tiap-tiap kluster. Dengan cara memperbaiki pusat kluster dan nilai keanggotaan tiap-tiap data secara berulang. Maka akan terlihat bahwa pusat kluster akan bergerak menuju lokasi yang tepat. Perulangan ini di dasarkan pada minimisasi fungsi objektif. Fungsi objektif yang digunakan adalah:

$$J_w(U, V, X) = \sum_{k=1}^n \sum_{i=1}^c (\mu_{ik})^w (d_{ik})^2 \quad (1.1)$$

Dengan $w \in [2, \infty)$

$$(d_{ik}) = d(x_k - v_i) = \left[\sum_{j=1}^m (x_{ij} - v_{ij})^2 \right]^{\frac{1}{2}} \quad (1.2)$$

x adalah data yang akan dikluster:

$$x = \begin{bmatrix} x_{11} & x_{1m} \\ \vdots & \vdots \\ x_{n1} & x_{nm} \end{bmatrix} \quad (1.3)$$

Dan v adalah matriks pusat kluster:

$$v = \begin{bmatrix} v_{11} & v_{1m} \\ \vdots & \vdots \\ v_{n1} & v_{nm} \end{bmatrix} \quad (1.4)$$

Nilai J_w terkecil adalah yang terbaik, sehingga:

$$J_w^*(U^*, V^*; X) = \min J(U, V, X) \quad (1.5)$$

Jika $d_{ik} > 0, \forall i, k; w > 1$ dan setidaknya memiliki m elemen, maka $U, V \in M \times Rmp$ dapat meminimisasi J hanya jika:

$$\mu_{ik} = \frac{\left(\left(\sum_{j=1}^m (x_{ij} - v_{kj})^2 \right)^{\frac{-1}{w-1}} \right)}{\sum_{k=1}^c \left(\left(\sum_{j=1}^m (x_{ij} - v_{kj})^2 \right)^{\frac{-1}{w-1}} \right)}, \forall i, \forall k \quad (1.6)$$

Dan

$$V_{kj} = \frac{\left(\sum_{i=1}^{100} (\mu_{ik})^w \right)}{\sum_{i=1}^{100} (\mu_{ik})^w}, \forall k, \forall j \quad (1.8)$$

2. Xie Beni Index

Rumus kevalidan suatu *cluster* atau indeks *Xie-Beni* XB seperti ditunjukkan pada persamaan (1.9) berikut :

$$XB_c = \frac{\sum_{k=1}^n \sum_{i=1}^c \mu_{ik}^w \|v_i - x_j\|^2}{n * \min_{i,j} \|v_i - v_j\|^2} \quad (1.9)$$

Dimana c =banyaknya *cluster*, n =banyak obyek yang dikelompokkan, μ_{ik} =derajat keanggotaan *fuzzy*, w = pangkat pembobot (*Fuzzifier*), $n * \min_{i,j} \|v_i - v_j\|^2$ = jarak minimum antara pusat *cluster* dengan v_i dan v_j

B. Pembahasan

Temuan penelitian ini akan menjawab apa yang sudah tertuang pada identifikasi masalah yang ada.

1. Langkah-langkah klustering menggunakan FCM:

- 1) Menentukan data yang akan di kluster X . Berupa matriks berukuran $n \times m$ (n adalah jumlah sampel, dan m adalah atribut setiap data). X adalah data sampel ke- I ($i = 1, 2, \dots, n$) atribut ke- j ($j = 1, 2, \dots, m$).
- 2) Menentukan:
 - a. Jumlah kluster : $c \Rightarrow 3$
 - b. Pangkat : $w \Rightarrow 2$
 - c. Maksimum Iterasi : $Maxiter \Rightarrow 50$
 - d. Error terkecil yang diharapkan : ε
 - e. Fungsi Objektif : $P = 0$
 - f. Iterasi awal : $t = 1$
- 3) Membangkitkan bilangan acak μ_{ik} , $i = 1, 2, 3, \dots, 38$; $k = 1, 2, 3$ sebagai elemen-elemen matriks partisi awal U .
Menghitung tiap kolom :

$$Q_i = \sum_{k=1}^c \mu_{ik}$$

Dengan $j = 1, 2, 3$

Menghitung :

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i}$$

- 4) Menghitung pusat ke- k ; V_{kj} dengan $k=1, 2, \dots, c$; $j=1, 2, \dots, m$ menggunakan persamaan (1.8)

- 5) Menghitung fungsi obyektif pada iterasi ke t .

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\sum_{j=1}^m (x_{ij} - v_{kj})^2 (\mu_{ik})^w \right) \quad (1.10)$$

dengan $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$, $k = 1, 2, \dots, c$

- 6) Menghitung perubahan matrik partisi dengan persamaan (1.6)
- 7) Memeriksa kondisi berhenti:
 - a. Jika $P_t - P_{t-1} < \varepsilon$ atau $t > Maxiter$ maka berhenti.
 - b. Jika tidak $t = t + 1$, mengulang langkah ke-4.

Data yang akan dilakukan *clustering* adalah data pendidikan pada provinsi Jawa Timur pada tahun 2023, dengan fitur yang digunakan adalah *Tdk/blm_prnh_sekolah*, *masih_sekolah*, *tdk_sekolah_lagi*. Dataset tersebut ditunjukkan pada Tabel 1:

Tabel 1. Data Pendidikan(Badan Pusat Statistik Provinsi Jawa Timur 2023)

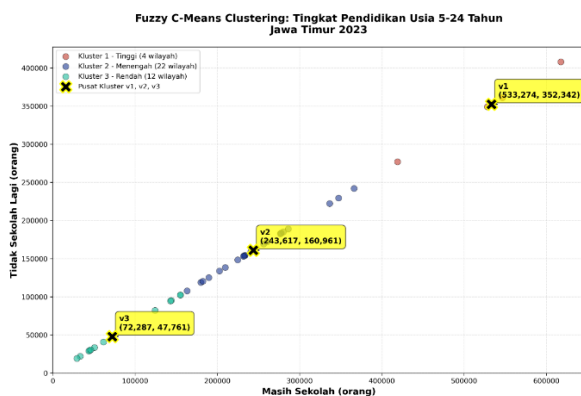
Id_wil	Nama Wilayah	Tdk/blm		
		prnh sekolah	masih sekolah	tdk sekolah lagi
3501	Kab. Pacitan	15.536	124.293	82.122
3502	Kab. Ponorogo	25.326	202.608	133.866
3503	Kab. Trenggalek	19.383	155.064	102.453
3504	Kab. Tulungagung	29.007	232.062	153.326
3505	Kab. Blitar	32.536	260.294	171.98
3506	Kab. Kediri	43.42	347.362	229.507
3507	Kab. Malang	68.331	546.652	361.181
3508	Kab. Lumajang	29.209	233.672	154.39
3509	Kab. Jember	66.037	528.299	349.055
3510	Kab. Banyuwangi	45.778	366.226	241.971
3511	Kab. Bondowoso	20.389	163.114	107.772
3512	Kab. Situbondo	17.993	143.949	95.109
3513	Kab. Probolinggo	30.05	240.401	158.836
3514	Kab. Pasuruan	42.054	336.436	222.288
3515	Kab. Sidoarjo	52.392	419.139	276.931
3516	Kab. Mojokerto	30.31	242.484	160.212
3517	Kab. Jombang	34.581	276.648	182.785
3518	Kab. Nganjuk	29.164	233.318	154.156
3519	Kab. Madiun	19.373	154.99	102.404
3520	Kab. Magetan	17.889	143.115	94.558
3521	Kab. Ngawi	23.696	189.571	125.252
3522	Kab. Bojonegoro	35.023	280.19	185.125
3523	Kab. Tuban	32.029	256.233	169.297
3524	Kab. Lamongan	35.778	286.231	189.117
3525	Kab. Gresik	34.685	277.482	183.336
3526	Kab. Bangkalan	28.097	224.777	148.513
3527	Kab. Sampang	26.196	209.569	138.465
3528	Kab. Pamekasan	22.498	179.988	118.92
3529	Kab. Sumenep	28.982	231.86	153.193
3571	Kota Kediri	7.681	61.454	40.603
3572	Kota Blitar	4.166	33.331	22.022
3573	Kota Malang	22.785	182.28	120.435

3574	Kota Probolinggo	6.327	50.621	33.446
3575	Kota Pasuruan	5.598	44.788	29.592
3576	Kota Mojokerto	3.645	29.164	19.269
3577	Kota Madiun	5.468	43.747	28.904
3578	Kota Surabaya	77.208	617.668	408.102
3579	Kota Batu	5.754	46.038	30.418

Perhitungan nilai pusat cluster V berdasarkan persamaan (1.3) yang terbentuk tiap cluster adalah seperti ditunjukkan pada matrix berikut :

$$V = \begin{matrix} 378.552,25 & 250.114,83 \\ 219.909,46 & 145.297,00 \\ 91.820,62 & 60.667,08 \end{matrix}$$

Diagram (*ploting*) Bidang Pendidikan seperti ditunjukkan pada Gambar 1.



Gambar 1. *Ploting* dan pusat masing-masing cluster

Perhitungan perubahan matrik partisi U menghasilkan penentuan cluster seperti yang ditunjukkan pada Tabel 2 berikut ini :

Tabel 2. Penentuan Cluster

Data wilayah	μ_{11}	μ_{12}	μ_{13}	Terbesar	cluster
3501	0.0134	0.1575	0.8291	0.8291	3
3502	0.0138	0.8973	0.0889	0.8973	2
3503	0.0249	0.4547	0.5204	0.5204	3
3504	0.0015	0.9933	0.0052	0.9933	2
3505	0.0037	0.9885	0.0078	0.9885	2
3506	0.2142	0.6879	0.0979	0.6879	2
3507	0.9973	0.0019	0.0008	0.9973	1
3508	0.0011	0.9951	0.0038	0.9951	2
3509	0.9996	0.0003	0.0001	0.9996	1
3510	0.3145	0.5839	0.1016	0.5839	2
3511	0.0258	0.5456	0.4286	0.5456	2
3512	0.0218	0.3333	0.6448	0.6448	3
3513	0.0001	0.9995	0.0004	0.9995	2
3514	0.1652	0.743	0.0917	0.743	2
3515	0.6531	0.2762	0.0707	0.6531	1
3516	0	0.9999	0	0.9999	2
3517	0.0159	0.9591	0.0251	0.9591	2
3518	0.0012	0.9948	0.0041	0.9948	2
3519	0.0249	0.4539	0.5212	0.5212	3
3520	0.0215	0.3247	0.6538	0.6538	3

3521	0.02	0.8084	0.1717	0.8084	2
3522	0.0199	0.9507	0.0294	0.9507	2
3523	0.0021	0.9933	0.0047	0.9933	2
3524	0.0278	0.9351	0.0371	0.9351	2
3525	0.0168	0.9572	0.0261	0.9572	2
3526	0.0037	0.9814	0.015	0.9814	2
3527	0.0103	0.9323	0.0573	0.9323	2
3528	0.0235	0.7239	0.2527	0.7239	2
3529	0.0015	0.9931	0.0054	0.9931	2
3571	0.0005	0.0035	0.996	0.996	3
3572	0.0058	0.033	0.9612	0.9612	3
3573	0.0228	0.7454	0.2318	0.7454	2
3574	0.002	0.0124	0.9856	0.9856	3
3575	0.0031	0.0187	0.9782	0.9782	3
3576	0.007	0.0386	0.9544	0.9544	3
3577	0.0033	0.0199	0.9768	0.9768	3
3578	0.9304	0.0474	0.0223	0.9304	1
3579	0.0028	0.0173	0.9799	0.9799	3

2. Xie Beni Index

Untuk menghitung tingkat akurasi klaster (*validity cluster*) pada algoritma *fuzzy-c-means*, menggunakan rumus *Xie Beni Index* (XBI) pada persamaan (1.9) dengan memanfaatkan *library python*, didapatkan nilai XBI = 0.072178 dengan jumlah cluster 3.

VI. KESIMPULAN

Dari perhitungan dataset wilayah provinsi Jawa Timur, untuk mengelompokkan wilayah berdasarkan tingkat partisipasi pendidikan menggunakan algoritma *fuzzy-c-means* terbagi menjadi 3 cluster.

Untuk Cluster 1 terdiri dari 4 wilayah: Kab. Malang, Kab. Jember, Kab. Sidoarjo, dan Kota Surabaya. Cluster ini memiliki karakteristik jumlah penduduk usia 5-24 tahun tertinggi. Rata-rata komposisi pendidikan: tidak/belum pernah sekolah 65,992 orang (7.00%), masih sekolah 527,940 orang (56.00%), dan tidak sekolah lagi 348,817 orang (37.00%).

Cluster 2 mencakup 22 wilayah: Kab. Ponorogo, Kab. Tulungagung, Kab. Blitar, Kab. Kediri, Kab. Lumajang, Kab. Banyuwangi, Kab. Bondowoso, Kab. Probolinggo, Kab. Pasuruan, Kab. Mojokerto, Kab. Jombang, Kab. Nganjuk, Kab. Ngawi, Kab. Bojonegoro, Kab. Tuban, Kab. Lamongan, Kab. Gresik, Kab. Bangkalan, Kab. Sampang, Kab. Pamekasan, Kab. Sumenep, dan Kota Malang. Cluster ini merupakan kelompok terbesar. Rata-rata komposisi pendidikan: tidak/belum pernah sekolah 30,982 orang (7.00%), masih sekolah 247,855 orang (56.00%), dan tidak sekolah lagi 163,761 orang (37.00%).

Cluster 3 terdiri dari 12 wilayah: Kab. Pacitan, Kab. Trenggalek, Kab. Situbondo, Kab. Madiun, Kab.

Magetan, Kota Kediri, Kota Blitar, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, dan Kota Batu. Kluster ini didominasi kota-kota dan kabupaten kecil. Rata-rata komposisi pendidikan: tidak/belum pernah sekolah 10,734 orang (7.00%), masih sekolah 85,880 orang (56.00%), dan tidak sekolah lagi 56,742 orang (37.00%).

Akurasi klaster *fuzzy-c-means* menggunakan perhitungan *Xie Beni Index* (XBI) sebesar 0.072178 dengan 3 klaster. Angka tersebut merupakan akurasi klaster terbaik karena nilai yang dihasilkan paling kecil.

DAFTAR PUSTAKA

- Ahmadi, Aziz, and Sri Yuni Hartati. 2015. "PENERAPAN FUZZY C-MEANS DALAM SISTEM PENDUKUNG KEPUTUSAN UNTUK PENENTUAN PENERIMA BANTUAN LANGSUNG MASYARAKAT (BLM) PNPM-MPd(Studi Kasus : PNPM-MPd Kec. Ngadirojo Kab. Pacitan)." <https://api.semanticscholar.org/CorpusID:109589755>.
- ARIKUNTO, Suharsimi. 19913. *Manajemen Pengajaran Secara Manusiawi Suharsimi Arikunto*. Cet. 2. Jakarta: Rineka Cipta.
- Badan Pusat Statistik Provinsi Jawa Timur. 2023. *Jawa Timur Dalam Angka 2023*. Surabaya.
- Collier, C.; Houston, J.; Schmatz, R. R.; Walsh, W. J. 1971. *Teaching in the Modern Elementary School*. New York: Macmillan.
- Dr. Panuti Sudjiman dan Greta Librata, MA. 1989. *Sosiologi Pendidikan: Perspektif Pendahuluan Yang Analitis*. Jakarta: Bhratara.
- Everitt, B S, S Landau, and M Leese. 2001. *Cluster Analysis*. Wiley. <https://books.google.co.id/books?id=htZzDGI CnQYC>.
- Kusumadewi, Sri, Sri Hartati, Agus Harjoko, Retantyo Wardoyo, and others. 2006. "Fuzzy Multi-Attribute Decision Making (Fuzzy Madm)." *Yogyakarta: Graha Ilmu* 74.
- Kusumadewi, Sri, and Hari Purnomo. 2010. "Aplikasi Logika Fuzzy Untuk Pendukung Keputusan." *Yogyakarta: Graha Ilmu* 2.
- Nasional, Departemen Pendidikan. 2007. "Undang-Undang SISDIKNAS (Sistem Pendidikan Nasional) UU RI No. 20 Tahun 2003 Dan Undang-Undang Guru Dan Dosen UU RI Nomor 14 Tahun 2005." *Jakarta: Depdiknas*.
- Owens, R G. 1991. *Organizational Behavior in Education*. Prentice Hall. <https://books.google.co.id/books?id=dFokAQ AAMAAJ>.
- Satriyanto, Edi. 2010. "Clustering." *Diakses pada 9*.
- Snowden, P E, and R A Gorton. 1998. *School Leadership and Administration: Important Concepts, Case Studies, and Simulations*. McGraw-Hill. <https://books.google.co.id/books?id=KZidAAA AAMAAJ>.
- Stoner, James A.F. 1982. *Management*. New Delhi: Prentice Hall.
- Tan, P.N.M.S. *Introduction to Data Mining*. Pearson. <https://books.google.co.id/books?id=64GVEjpT WIAC>.
- Xu, R, and D Wunsch. 2009. *Clustering*. Wiley. <https://books.google.co.id/books?id=XC4nAQA AIAAJ>.
- Zhang, Yunjie, Weina Wang, Xiaona Zhang, and Yi Li. 2008. "A Cluster Validity Index for Fuzzy Clustering." *Information Sciences* 178: 1205–18. doi:10.1016/j.ins.2007.10.004.