

# CLUSTERING PENERIMA BANTUAN SANTRI PRA-SEJAHTERA DENGAN MENGGUNAKAN METODE FUZZY C-MEANS BERBASIS WEB STREAMLIT DI PONDOK PESANTREN NURUL JADID

Sudriyanto<sup>1)</sup>, Nur Hatima Inda Arifin<sup>2)</sup>, Nurul Hikmah<sup>3)</sup>

<sup>1,2,3)</sup>Teknik Informatika, Universitas Nurul Jadid

e-mail: [sudriyanto@unuja.ac.id](mailto:sudriyanto@unuja.ac.id)<sup>1)</sup>, [chatiem.indah@gmail.com](mailto:chatiem.indah@gmail.com)<sup>2)</sup>, [nurulhikmah140201@gmail.com](mailto:nurulhikmah140201@gmail.com)<sup>3)</sup>

## ABSTRAK

Bantuan sosial merupakan suatu program yang dilaksanakan dengan cara memberikan (barang/jasa/uang) kepada sekelompok orang tertentu secara percuma dan bersifat sementara dengan tujuan untuk meningkatkan kesejahteraan suatu kelompok masyarakat. Salah satu contoh program bantuan sosial di Pondok Pesantren Nurul Jadid adalah program bantuan santri Pra-Sejahtera yang dikelola oleh Yayasan Bantuan Sosial (YBS). Pada proses penyaluran bantuan di Yayasan Bantuan Sosial (YBS) mengalami kendala dalam menentukan target penerima bantuan. Untuk mengatasi masalah tersebut peneliti memanfaatkan data santri putri tingkat SLTP (Sekolah Lembaga Tingkat Pertama) dan tingkat SLTA (Sekolah Lembaga Tingkat Atas) angkatan tahun 2022 di Pondok Pesantren Nurul Jadid dengan dataset yang berjumlah 843 record yang diperoleh dari admin pedatren (Pengelola Data Pesantren). Algoritma yang digunakan dalam penelitian ini adalah metode Fuzzy C-Means Clustering yang diimplementasikan ke web streamlit. Metode Fuzzy C-Means adalah metode yang memanfaatkan fungsi derajat keanggotaan untuk mengukur seberapa dekat titik data berada dalam kluster yang telah ditentukan. Kelebihan metode Fuzzy C-Means adalah metode yang efisien dan efektif untuk melakukan proses pengelompokan data besar. Hasil dari penelitian ini membentuk 2 cluster, yang mana cluster pertama berjumlah 595 (Keluarga Sejahtera) dan cluster kedua berjumlah 186 (Keluarga Pra-Sejahtera). Pengujian dari hasil clustering tersebut dilakukan dengan menggunakan teknik Silhouette Coefficient yang menghasilkan nilai sebesar 0.5552751407539394 yang menunjukkan bahwa hasil dari proses clustering yang telah dilakukan sudah bagus.

**Kata Kunci:** Clustering, Bantuan Santri, Metode Fuzzy C-Means, Web Streamlit.

## ABSTRACT

Social assistance is a program implemented by providing (goods/services/money) to a certain group of people for free and on a temporary basis with the aim of improving the welfare of a community group. One example of a social assistance program at the Nurul Jadid Islamic Boarding School is the pre-prosperous student assistance program managed by the Social Assistance Foundation (YBS). In the process of distributing aid at the Social Assistance Foundation (YBS), there were problems in determining the target recipients of aid. To overcome this problem, researchers utilized data on female students at the junior high level (first level institutional school) and senior high level (upper level institutional school) class of 2022 at the Nurul Jadid Islamic Boarding School with a dataset of 843 records obtained from the pedatren admin (Islamic boarding school data manager). The algorithm used in this research is the Fuzzy C-Means Clustering method which is implemented on the streamlit web. The Fuzzy C-Means method is a method that utilizes the membership degree function to measure how closely data points fall within a predetermined cluster. The advantage of the Fuzzy C-Means method is that it is an efficient and effective method for carrying out the process of grouping large data. The results of this research formed 2 clusters, where the first cluster numbered 595 (Prosperous Families) and the second cluster numbered 186 (Pre-Prosperous Families). Testing of the clustering results was carried out using the Silhouette Coefficient technique which produced a value of 0.5552751407539394 which shows that the results of the clustering process that has been carried out are good.

**Keywords:** Clustering, Student Assistance, Fuzzy C-Means Method, Web Streamlit.

## I. PENDAHULUAN

**B**ANTUAN SOSIAL adalah suatu bantuan (barang/uang/jasa) yang diberikan kepada sekelompok orang tertentu secara percuma dan bersifat sementara dengan tujuan untuk meningkatkan kesejahteraan suatu kelompok masyarakat [1]. Dalam menentukan penerima bantuan sosial perlu adanya filterisasi yang tepat sehingga bantuan tersalurkan kepada orang-orang yang layak untuk menerima bantuan, hal ini dilakukan agar tercapainya suatu tujuan dalam pelaksanaan program bantuan tersebut [2]. Dari berbagai macam program bantuan pemerintah (PKH, BLT, dll) pemerintah menentukan beberapa syarat/kriteria yang berbeda dalam menentukan kelayakan penerima bantuan [3]. Dalam implementasinya banyak dari bantuan yang diberikan kepada pihak yang tidak tepat, hal ini menyebabkan adanya kelompok masyarakat yang seharusnya mendapatkan bantuan tetapi tidak memperoleh bantuan sehingga menyebabkan belum terpenuhi kebutuhan penerima bantuan yang sebenarnya. Demikian pula, ada masyarakat yang tidak layak ternyata menerima bantuan tersebut [2]. Banyak yang menjadi korban dari kesalahan dalam pelaksanaan kebijakan program bantuan sosial dalam berbagai wilayah atau lembaga penyalur program bantuan sosial [4]. Salah satu lembaga penyalur bantuan yang mengalami masalah tersebut adalah Yayasan Bantuan Sosial (YBS) yang merupakan lembaga pendistribusian bantuan di Pondok Pesantren Nurul Jadid yang disebabkan banyak petugas mengalami kesulitan dalam mengevaluasi apakah seorang santri layak menerima bantuan atau tidak yang disebabkan karena kurangnya data santri yang masuk dalam kategori Pra-sejahtera setiap tahunnya, dan pendataan pengajuan bantuan masih diproses secara manual sehingga waktu yang digunakan kurang efisien dan kurang efektif [2].

Upaya dalam mengatasi masalah tersebut peneliti menggunakan metode *Fuzzy C-Means Clustering* yaitu salah satu metode dalam pengolahan data yang memberikan hasil yang lebih fleksibilitas dengan cara menghitung nilai keanggotaan data dari tiap kelompok [5]. Metode *Fuzzy C-Means* memiliki kelebihan yaitu sebagai metode yang efisien dan efektif untuk melakukan proses pengelompokan data karena pada praktiknya metode *Fuzzy C-Means* menggunakan fungsi derajat keanggotaan untuk mengukur seberapa dekat titik data berada dalam kluster yang telah ditentukan sehingga hal ini memungkinkan untuk melakukan pemisah yang lebih tepat dan mengurangi kemungkinan kesalahan dalam pengelompokan data [5]. Penggunaan metode *Fuzzy C-Means* juga memiliki kelemahan yaitu seorang peneliti perlu ketelitian dalam menentukan kluster yang akan dijadikan acuan dalam pengolahan data, jika peneliti membuat kepu-

tusan kluster kurang tepat maka hasilnya mungkin tidak akan akurat [6]. Kemudian peneliti juga menggunakan *platform web streamlit* yang merupakan situs yang memanfaatkan kerangka kerja bahasa pemrograman *python* guna membuat antarmuka pengguna yang bersifat interaktif untuk proyek yang berkaitan dengan pengolahan data dan *machine learning* [7]. Maka, penggunaan metode *Fuzzy C-Means* berbasis *web streamlit* dalam mengelompokkan santri Pra-sejahtera di Yayasan Bantuan Sosial (YBS) bertujuan untuk mengoptimalkan proses pengelompokan data santri pra-sejahtera dan sejahtera agar lebih efisien dan cepat dalam melakukan pencarian dan pengolahan data sehingga meminimalisir ketidaktepatan dalam pemetaan penerima bantuan [2].

## II. STUDI PUSTAKA

### A. Penelitian Terkait

Penelitian pertama dilakukan oleh [2] Penelitian ini dilakukan di Yayasan Bantuan Sosial yang mana yayasan ini berada di Pondok Pesantren Nurul Jadid Paiton Probolinggo Jawa Timur, dengan judul “Penerapan Algoritma K-Means Untuk Clustering Santri Pra-Sejahtera Di Yayasan Bantuan Sosial (YBS) Az-Zainiyah Pondok Pesantren Nurul Jadid” latar belakang penelitian ini dilakukan karena banyaknya petugas yang kesulitan dalam menentukan penerima bantuan yang layak dan tidak layak, masalah selanjutnya juga dikarenakan oleh pendataan pengajuan bantuan masih diproses secara manual (masih menggunakan cara lama) sehingga waktu yang digunakan kurang efisien dan kurang efektif. Dalam penelitian tersebut peneliti menggunakan metode K-Means Clustering berdasarkan data santri Pondok Pesantren Nurul Jadid tingkat SLTP dan Tingkat SLTA angkatan tahun 2021. Hasil clustering dari penelitian ini disimpulkan C1 = 163 (Sejahtera 1), C2 = 215 (Sejahtera 2), C3 = 246 (Sejahtera Plus) dan C4 = 125 (Keluarga Pra-Sejahtera). Hasil dari clustering tersebut dilakukan evaluasi supaya menghasilkan hasil yang lebih akurat dengan menggunakan matriks DBI (Davies Bouldin Index) yang memperoleh akurasi nilai sebesar 0,90 (90%) yang menunjukkan hasil *clustering* menggunakan algoritma k-means sudah sangat baik.

Penelitian kedua dilakukan oleh [5], penelitian ini bertempat di Dinas Sosial kota Pekanbaru dengan jurnal yang berjudul “Penerapan Fuzzy C-Means Pada Klasterisasi Penerima Bantuan Pangan Non Tunai” masalah pada penelitian tersebut adalah adanya kendala dalam pendistribusian BPNT (Bantuan Pangan Non Tunai) di kota Pekanbaru sehingga pendistribusian bantuan mengalami kesalahan sasaran atau tidak sesuai dengan target tujuan, hal ini juga disebabkan oleh tidak tersediaannya sistem yang dapat mempermudah dalam menentukan penerima bantuan yang layak dan tidak. Dalam menyelesaikan masalah terse-

but peneliti memanfaatkan metode *Fuzzy C-Means Clustering* menggunakan *tools* MATLAB. Dari hasil penelitian ini dilakukan uji coba dengan menggunakan metode *Silhouette Coefficient* dengan hasil rata-rata koefisien sebesar 0.426653079 sehingga dapat disimpulkan bahwa upaya yang dilakukan menggunakan metode *Fuzzy C-Means Clustering* dalam mengolah data BPNT menghasilkan 2 klaster, klaster pertama sebanyak 331 dan klaster kedua sebanyak 351. Diantara kedua klaster, klaster pertama yang diprioritaskan untuk menerima bantuan karena klaster pertama menunjukkan kondisi ekonomi yang lebih rendah daripada klaster kedua.

## B. Landasan Teori

### 1. Santri Pra-Sejahtera

Santri adalah sebutan untuk orang yang mengikuti pendidikan agama islam di lembaga pesantren [8]. Menurut Dhofier pada jurnal [9] santri adalah pelajar yang menjalani semua pembelajaran di pesantren. Santri juga disebut sebagai orang yang sedang belajar ilmu agama di pesantren tertentu dan dalam jangka waktu tertentu [10]

Pra-Sejahtera adalah suatu keadaan yang menunjukkan ketidak mampuan seseorang atau sekelompok orang dalam memenuhi kebutuhan dasarnya [2]. Menurut Sunarti pada jurnal [11] Pra-Sejahtera adalah kondisi tidak dapat terpenuhinya kebutuhan dasar dari seseorang atau sekelompok orang. Pra-Sejahtera adalah kondisi tidak terpenuhinya 6 kriteria kebutuhan dasar (kebutuhan pokok, pangan, sandang, papan, kesehatan, pendidikan agama) [12].

Santri Pra-Sejahtera adalah seorang atau sekelompok santri yang berasal dari keluarga yang kurang secara finansial [13]. Santri Pra-Sejahtera juga disebut sebagai Jadi kesimpulannya santri Pra-Sejahtera adalah seseorang yang menempuh pendidikan di lembaga pesantren yang belum mampu memenuhi kebutuhan kebutuhan dasarnya.

### 2. Data Mining

Data mining adalah salah satu ilmu komputer yang dapat menemukan atau menggali makna tersembunyi dari suatu kumpulan data yang besar. Proses menemukan atau menggali data sangat erat kaitannya dengan beberapa ilmu komputer yang lain yaitu artificial intelligent, *data warehouse*, statistik, *machine learning*, *pattern recognition* dan lain lain [14].

Jenis atau tipe data mining (berdasarkan fungsinya) terdiri dari:

#### a. Deskripsi

Data deskripsi adalah suatu tipe data yang memiliki ciri pada atribut datanya itu boleh nominal boleh numerik, contoh data pembelian barang (barang apa saja yang dibeli secara bersamaan). Metode metode yang

bisa digunakan pada tipe data deskripsi yaitu FP-Growth, A Priori, dan lain lain.

#### b. Estimasi

Estimasi adalah salah satu jenis/tipe data dalam ilmu datamining yang mana pembelajarannya menggunakan metode estimasi (regresi linier) dan ciri datanya itu class dan atributnya itu berbentuk numerik (data yang bisa ditambah kurang kali bagi), contohnya estimasi pengiriman barang.

#### c. Prediksi

Prediksi adalah tipe data mining yang memiliki ciri yaitu salah satu dari atributnya itu terdapat data time series (rentet waktu). Metode yang dapat digunakan pada tipe data ini adalah linear regrestion (LR), Neural Network (NN), Deep Learning (DL), Support Vektor Machine (SVM), dan Generalized Linear Model (GLM). Contohnya tipe data prediksi adalah data harga saham, data inflasi dan lain lain.

#### d. Klasifikasi

Klasifikasi adalah tipe data mining yang pada datanya itu terdapat label, contoh data kelulusan mahasiswa yang mana pada data tersebut terdapat label lulus dan tidak lulus. Metode yang digunakan dalam tipe data klasifikasi yaitu Decision tree, C4.5, Naïve Bayes, KNN, dan lain lain.

#### e. Clustering

Clusterisasi adalah jenis data yang memiliki ciri tidak memiliki label pada susunan datanya, jadi clustering dan klasifikasi itu sama sama pengelompokkan akan tetapi memiliki perbedaan pada susunan datanya, data klasifikasi itu terdapat label sedangkan clustering tidak contoh data jenis pelanggan, clustering tingkat kemiskinan dan lain lain. Untuk metode clustering yang bisa digunakan yaitu K-Means, Self-Organizing Map (SOM) dan Fuzzy C-Means[2] (Khairi et al., 2023).

Adapun tahapan dalam data mining sebagai berikut [15]:

#### a. Pembersihan Data

Pembersihan data merupakan proses untuk mengisi nilai nilai yang hilang, menghaluskan noise (data yang yang tidak diinginkan atau data yang salah), menghilangkan outlier, dan memecah ketidak konsistenan data.

#### b. Integrasi Data

Integrasi data merupakan integrasi dari beberapa *database*.

#### c. Seleksi Data

Seleksi data merupakan proses memilih data yang sesuai dengan kebutuhan yang diinginkan dalam suatu penelitian atau analisis.

#### d. Transformasi Data

Tranformasi data adalah proses normalisasi data yang telah dipilih yang selanjutnya akan dilanjut oleh proses *mining*.

e. Proses Mining

Tranformasi data adalah proses normalisasi data yang telah dipilih yang selanjutnya akan dilanjut oleh proses *mining*.

f. Evaluasi Pola

Evaluasi pola adalah proses yang bertujuan untuk menemukan atau menggali suatu pengetahuan atau informasi dari suatu data yang telah melalui proses *mining*.

g. Presentasi Pengetahuan

Presentasi pengetahuan merupakan tahapan terakhir dalam ilmu data mining dengan cara mengimplementasikan pengetahuan yang telah didapat dari suatu metode yang digunakan sehingga mencapai tujuan yang telah diinginkan[14].

3. Metode Fuzzy C-Means

Metode Fuzzy C-Means merupakan salah satu teknik dalam pengelompokan data yang dalam menentukan cluster dari tiap data yang dikelola tergantung dari nilai keanggotaannya [16]. Fuzzy C-Means juga disebut sebagai salah satu metode dalam ilmu data mining yang digunakan untuk pengelompokan data yang tidak hanya memiliki nilai yang sama dalam suatu cluster, tapi juga kelompok kelompok yang memiliki 2 atau lebih kelompok menurut level keanggotaannya [17]. Penentuan titik cluster ini dilakukan dalam beberapa tahap yang mana tahapan tahapan ini dilakukan secara berulang sampai hasil dari pusat cluster telah mencapai pada titik yang telah diinginkan sehingga mendapatkan hasil yang lebih akurat. Berikut merupakan tahapan Fuzzy C-Means :

$$J_m = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|x_j - v_i\|^2 \quad \dots(1)$$

Keterangan :

J = Nilai Fungsi Objektif FCM

U = matriks keanggotaan/derajat keanggotaan

V = matriks pusat cluster

m = parameter fuzzinez

$\|x_j - v_i\|^2$  = jarak euclidean antara titik I dan pusat cluster j

Syarat derajat keanggotaan suatu partisi matriks terhadap semua cluster yaitu :

$$\sum_{i=1}^c u_{ij} = 1, u_{ij} \in [0,1], 0 \leq \sum_{n=1}^n \leq n \quad \dots(2)$$

Tahapan tersebut dilakukan sampai fungsi objektif tercapai, jika belum maka ulangi tahap iterasi dengan melakukan proses perubahan keanggotaan data dengan menggunakan rumus :

$$u_{ij} = \frac{1}{\sum_{i=1}^c \left( \frac{\|x_j - v_i\|}{\|x_j - v_i\|} \right)^2 / m - 1} \quad \dots(3)$$

Proses selanjutnya adalah mengulang proses pembaruan cluster dengan menggunakan rumus :

$$v_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m} \quad \dots(4)$$

Derajat keanggotaan data terhadap cluster dan pusat cluster akan terus mengalami perubahan dan tahap ini akan terus terulang sampai nilai dari fungsi objektif menghasilkan nilai error yang telah diharapkan pada pada saat awal pemrosesan data[17].

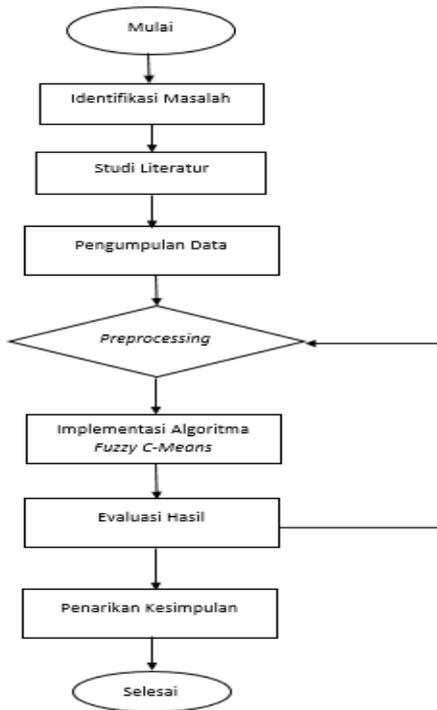
4. Web Streamlit

Streamlit adalah situs yang memanfaatkan kerangka kerja bahasa pemrograman python guna membuat antarmuka pengguna yang bersifat interaktif untuk proyek yang berkaitan dengan pengolahan data dan *machine learning* [7]. Streamlit juga disebut sebagai library python yang dalam implementasinya memanfaatkan open source [18]. Menurut [19] streamlit adalah kerangka python menghasilkan aplikasi berbasis web yang bersifat interaktif. Jadi web streamlit adalah web yang dihasilkan dari framework python yang digunakan untuk mempermudah dalam mengelola proyek dalam bidang ilmu pengolahan data dan *machine learning* [20].

III. METODE PENELITIAN

A. Tahapan Penelitian

Proses penelitian ini menggunakan paradigma kuantitatif dengan metode campuran (mixed method). Karena dalam implementasinya penelitian ini menggunakan metode kuantitatif yaitu metode *Fuzzy C-Means* akan tetapi dalam penelitian ini juga menggunakan metode kualitatif untuk membangun *web streamlit* yang akan digunakan untuk memvisualisasikan hasil *clustering* dari perhitungan metode *Fuzzy C-Means* sehingga mempermudah dalam menentukan layak tidaknya seseorang untuk menerima bantuan. Tahapan penelitian seperti gambar 1:



Gambar 1. Bagan Alur Penelitian

Tahapan penelitian dideskripsikan dengan langkah langkah sebagai berikut :

#### A. Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan di atas maka dapat diidentifikasi sebuah masalah dalam penelitian ini yaitu bagaimana meng-clustering santri pra-sejahtera di Pondok Pesantren Nurul Jadid dan bagaimana menentukan santri yang layak dan tidak layak untuk menerima bantuan di Yayasan Bantuan Sosial Pondok Pesantren Nurul Jadid dengan menggunakan metode *Fuzzy C-Means* berbasis *web streamlit*.

#### B. Studi Literatur

Hal pertama yang dilakukan pada saat penelitian ini dimulai adalah mencari masalah yang terjadi di Yayasan Bantuan Sosial kemudian berusaha mencari solusi dari permasalahan tersebut dengan cara memanfaatkan data santri putri tingkat SLTP (Sekolah Lembaga Tingkat Pertama) dan tingkat SLTA (Sekolah Lembaga Tingkat Akhir) angkatan tahun 2022 kemudian peneliti merumuskan tujuan, dan menentukan batasan masalah dalam penelitian ini.

#### C. Pengumpulan Data

Pengumpulan data diperoleh dari pihak admin *pedatren* (Pengelolaan Data Pesantren) Pondok Pesantren Nurul Jadid, kemudian data yang diperoleh dikelola serta dikelompokkan berdasarkan variabel dan kategori yang telah ditentukan sebelumnya.

#### D. Preprocessing

Preprocessing merupan bentuk awal pemrosesan data yang dilakukan dengan cara pembersihan data (menghilangkan data data yang tidak dibutuhkan) yang nantinya akan dilanjutkan dengan proses pengelolaan data dengan menggunakan metode *Fuzzy C-Means clustering*.

#### E. Implementasi Algoritma Fuzzy-C-Means

Data yang telah melalui proses preprocessing akan berlanjut pada tahap implementasi algoritmatma *Fuzzy C-Means* dimana pada tahap ini data akan dikelola dengan cara menggolongkan data yang memiliki nilai yang sama sesuai dengan level keanggotaannya selanjutnya akan diproses menggunakan rumus yang telah ditentukan sehingga menghasilkan informasi yang berguna untuk mengambil suatu keputusan.

#### F. Evaluasi Hasil

Evalusi hasil adalah tahapan dimana data yang telah melalui proses perhitungan akan diimplementasikan ke web streamlit yang kemudian dievaluasi apakah hasil yang telah didapat sudah sesuai dengan tujuan awal atau belum.

#### G. Penarikan Kesimpulan

Pada tahapan ini data yang telah diperoleh dari hasil beberapa tahapan akan dijadikan sebagai acuan untuk menarik suatu kesimpulan atau sebagai tolak ukur untuk mengambil suatu kebijakan sehingga tujuan dari penelitian bisa tercapai.

### IV. HASIL DAN PEMBAHASAN

Hasil dari pengumpulan data yang diperoleh merupakan data santri SLTP (Sekolah Lembaga Tingkat Pertama) dan SLTA (Sekolah Lembaga Tingkat Atas) angkatan tahun 2022 di Pondok Pesantren Nurul Jadid. Data didapat dari pihak admin PEDATREN (Pengelola Data Pesantren) Pondok Pesantren Nurul Jadid dengan dataset sebanyak 843 *record*. Data yang diperoleh berupa nama, jumlah saudara, angkatan, status, lembaga, jurusan, pekerjaan orang tua dan penghasilan orang tua seperti pada tabel 1 :

Tabel 1. Dataset Sampel Santri SLTP & SLTA angkatan 2022

No	Nama	Jumlah Saudara	Angkatan	Status	Lembaga	Jurusan	Pekerjaan Orang Tua	Penghasilan Orang Tua
1	Nabilah Indah Kirana Putri	3	2022	santiripela jar santiripela jar santiripela jar	SMP-NJ	IPA	Petani	Rp.1.000.000
2	Muthia Ziba Layida	3	2022	santiripela jar santiripela jar santiripela jar	SMP-NJ	Reguler	Wiraswasta	Rp.500.000
3	Anindya Maharani	2	2022	santiripela jar santiripela jar	SMP-NJ	LIPS	Pedagang kecil	Rp.1.500.000
4	Novari	2	2022	santiripela jar santiripela jar	SMP-NJ	Reg-	Peda-	Rp.2.000.000



Rp.3.500.000	12
Rp.4.000.000	13
Rp.4.500.000	14
Rp.5.000.000	15
Rp.500.000	16

nilai keanggotaan data. Nilai centroid tiap iterasi dapat dilihat pada tabel 4 :

Tabel 4. Nilai Centroid Tiap Iterasi

Iterasi	Cluster 1	Cluster 2
1	[1.42305654 8.9178311 ]	[1.47431407 9.42203659]
2	[1.36022563 8.24762654]	[ 1.54991503 10.48432545]
3	[1.34147752 7.51028779]	[ 1.59030874 12.10576094]
4	[1.36790325 7.42592313]	[ 1.59766313 13.0316321 ]
5	[1.37935597 7.54639056]	[ 1.60287301 13.53524223]
6	[1.38426549 7.63862495]	[ 1.602966 13.82216599]
7	[1.38701219 7.68938848]	[ 1.60079907 13.98587973]
8	[1.38865362 7.71670775]	[ 1.59871042 14.07598243]
9	[1.38959569 7.73136815]	[ 1.59729833 14.12413784]
10	[1.3901132 7.73914619]	[ 1.59646733 14.14946963]
11	[1.39038977 7.74323058]	[ 1.59600876 14.16268896]
12	[1.39053534 7.74536123]	[ 1.59576357 14.16955944]
13	[1.39061134 7.74646855]	[ 1.59563455 14.17312281]
14	[1.39065085 7.74704286]	[ 1.5955672 14.17496896]
15	[1.39067135 7.74734041]	[ 1.59553219 14.1759249 ]
16	[1.39068197 7.74749448]	[ 1.59551403 14.17641975]
17	[1.39068746 7.74757424]	[ 1.59550462 14.17667587]
18	[1.39069031 7.74761552]	[ 1.59549975 14.17680842]
19	[1.39069178 7.74763688]	[ 1.59549723 14.17687702]
20	[1.39069255 7.74764794]	[ 1.59549593 14.17691252]
21	[1.39069294 7.74765366]	[ 1.59549525 14.17693089]
22	[1.39069315 7.74765662]	[ 1.5954949 14.1769404]
23	[1.39069325 7.74765816]	[ 1.59549472 14.17694532]
24	[1.39069331 7.74765895]	[ 1.59549463 14.17694786]
25	[1.39069333 7.74765936]	[ 1.59549458 14.17694918]
26	[1.39069335 7.74765957]	[ 1.59549455 14.17694986]

Dari perubahan nilai di atas maka di dapatkan data seperti gambar 2 :

	Jumlah Saudara	Penghasilan Orang Tua
0	2	7
1	2	16
2	1	8
3	1	9
4	3	7
..	...	...
776	0	6
777	0	9
778	2	9
779	2	7
780	1	8

[781 rows x 2 columns]

Gambar 2. Data Transformasi

Tahap selanjutnya adalah proses clustering dengan menggunakan metode *Fuzzy C-Means*, langkah pertama dalam metode *Fuzzy C-Means* yaitu menentukan jumlah cluster, parameter fuzziness, maksimal iterasi. Penentuan nilai nilai tersebut dapat dilihat pada gambar 3:

```
K = 2 # Jumlah cluster
m = 2 # Parameter fuzziness
max_iteration = 100 # Maksimum iterasi
```

Gambar 3. Nilai Nilai Fuzzy C-Means

Proses selanjutnya yaitu menentukan nilai random dari setiap cluster. Nilai random yang dihasilkan dalam perhitungan ini dapat dilihat pada gambar 4 :

```
array([[0.58394872, 0.41605128],
       [0.47424506, 0.52575494],
       [0.20467522, 0.79532478],
       ...,
       [0.54880633, 0.45119367],
       [0.88859741, 0.11140259],
       [0.94229254, 0.05770746]])
```

Gambar 4. Data Nilai Random

Data nilai random ini akan dilanjutkan pada tahap perhitungan yang mana perhitungan ini berhenti pada iterasi ke 26 yang menghasilkan *centroid* pada tiap

Hasil *clustering* data pada perhitungan metode *Fuzzy C-Means* dapat dilihat pada gambar 5:

	Jumlah Saudara	Penghasilan Orang Tua	Cluster
0	2	7	0
1	2	16	1
2	1	8	0
3	1	9	0
4	3	7	0
..	...	...	...
776	0	6	0
777	0	9	0
778	2	9	0
779	2	7	0
780	1	8	0

[781 rows x 3 columns]

Gambar 5. Data Hasil Cluster

Pada gambar 5 menunjukkan bahwa *Clustering*

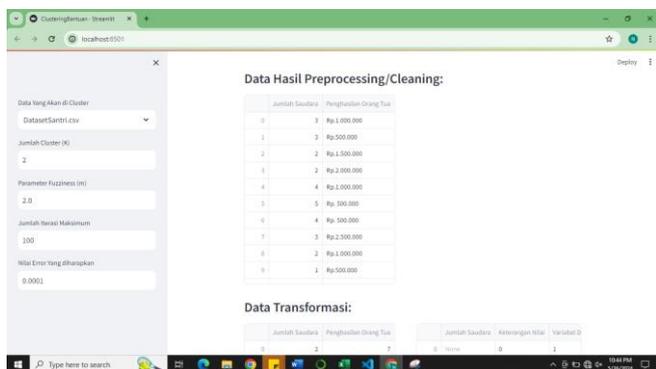
metode *Fuzzy C-Means* menghasilkan 2 cluster. *cluster* pertama berjumlah 595 Dan *cluster* kedua berjumlah 186. Dari kedua *cluster* tersebut *cluster* kedua yang berhak untuk menerima bantuan karena pada data tersebut *cluster* kedua menunjukkan tingkat ekonomi yang lebih rendah daripada cluster pertama.

Pengujian hasil *cluster* dilakukan dengan menggunakan teknik *Silhouette Conficient* yang menghasilkan nilai sebesar 0.5552751407539394 yang menunjukkan bahwa hasil dari proses clustering yang telah dilakukan sudah bagus.

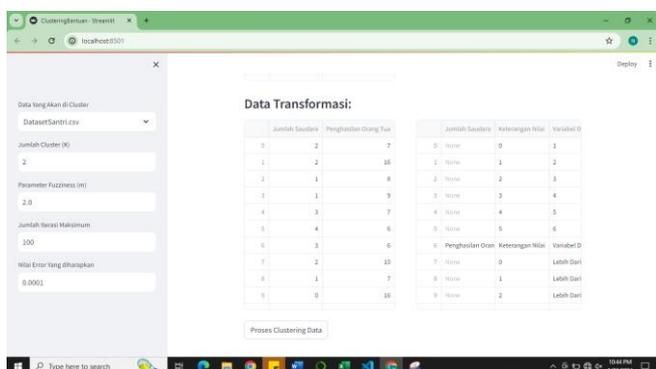
Dari perhitungan diatas peneliti akan melanjutkan pada tahapan implementasikan ke *framework streamlit* dengan menggunakan *Visual Studi Code*. Hasil implementasi dapat dilihat pada gambar 6,7,8,9 :



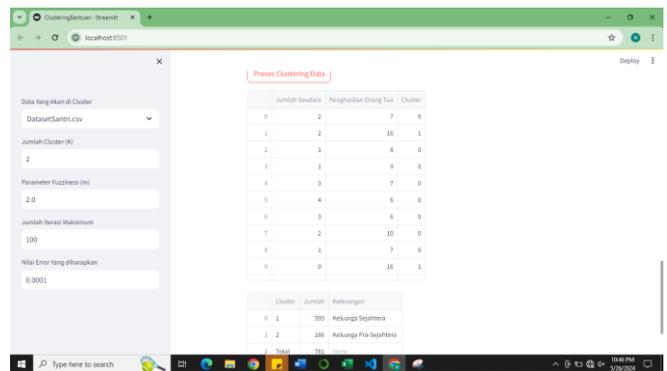
Gambar 6. Hasil Implementasi Streamlit



Gambar 7. Lanjutan1 Hasil Implementasi Streamlit



Gambar 8. Lanjutan2 Hasil Implementasi Streamlit



Gambar 9. Lanjutan3 Hasil Implementasi Streamlit

Pada gambar 9 keika meng-klik tombol Proses Clustering Santri maka akan tampil hasil clustering seperti gambar 10 tersebut.

### V. KESIMPULAN

Berdasarkan dari penelitian yang telah dilakukan dan hasil dari clustering metode *Fuzzy C-Means* pada data santri di Pondok Pesantren Nurul Jadid Tingkat SLTP (Sekolah Lembaga Tingkat Pertama) & SLTA (Sekolah Lembaga Tingkat Atas) angkatan tahun 2022 maka dapat disimpulkan dari data sejumlah 781 record menghasilkan 2 *cluster*. *Cluster* pertama berjumlah 595 (Keluarga Sejahtera) dan *cluster* kedua berjumlah 186 (Keluarga Pra-Sejahtera). Hasil dari *clustering* tersebut dilakukan pengujian dengan menggunakan teknik *Silhouette Conficient* yang menghasilkan nilai sebesar 0.5552751407539394 yang menunjukkan bahwa hasil dari proses *clustering* yang telah dilakukan sudah bagus. Hasil dari penelitian ini masih belum bisa dikatakan sempurna sehingga perlu untuk diperbaiki dan dikembangkan lagi. Saran peneliti untuk kedepannya *Clustering* pada data bantuan santri ini dapat dikembangkan lagi dengan menggunakan metode yang berbeda yang mungkin dapat menghasilkan tingkat akurasi yang lebih baik daripada penelitian ini dan implementasinya bukan hanya berbasis *web* tapi bisa dikembangkan lagi ke aplikasi berbasis *android* sehingga proses *Clustering* bisa lebih efisien dan lebih efektif lagi.

### REFERENCES

- [1] M. L. A. hafiz A. Habibi, Dorlan. Irham, "Dorlan+Habibi+1134," *J. Ilm. Ekon. Dan Manaj. Vol.2*, vol. 2, pp. 241–247, 2024.
- [2] A. Khairi, A. Shohibul Hikam, T. Informatika, and U. Nurul Jadid Pondok Pesantren Nurul Jadid Karanganyar Paiton Probolinggo, "PENERAPAN ALGORITMA K-MEANS UNTUK CLUSTERING SANTRI PRA-SEJAHTERA DI YAYASAN BANTUAN SOSIAL (YBS) AZ-ZAINIYYAH PONDOK PESANTREN NURUL JADID," 2023.
- [3] D. Kinasih Widiyati, M. Wati, and H. Santoso Pakpahan, "Penerapan Algoritma ID3 Decision Tree Pada Penentuan Penerima Program

- Bantuan Pemerintah Daerah di Kabupaten Kutai Kartanegara,” 2018.
- [4] D. Herdiana, “PENGAWASAN KOLABORATIF DALAM PELAKSANAAN KEBIJAKAN BANTUAN SOSIAL TERDAMPAK COVID-19,” *JDP (JURNAL Din. PEMERINTAHAN)*, vol. 3, no. 2, pp. 85–99, Aug. 2020, doi: 10.36341/jdp.v3i2.1323.
- [5] Solahuddin, “PENERAPAN FUZZY C-MEANS PADA KLASTERISASI PENERIMA BANTUAN PANGAN NON TUNAI TUGAS AKHIR,” Pekanbaru, 2023.
- [6] Y. S. Firdaus *et al.*, “Klaster Daerah Kesejahteraan pada Masa Pandemi Covid-19 di Jawa Timur dengan Metode Fuzzy C-Means Clustering,” 2022.
- [7] T. Abu, *Buku Refrensi Implementasi Algoritma Machine Learning Berbasis Web dengan Framework Streamlit*. Pustaka Nurja, 2023. [Online]. Available: <https://pustakanurja.unuja.ac.id>
- [8] Admin Website Pesantren Al-Irsyad Tenganan, “Apa Itu Santri?,” [pesantrenalirsyad](https://pesantrenalirsyad.org/2023/03/16/Apa-Itu-Santri). [Online]. Available: <https://pesantrenalirsyad7.org/2023/03/16/Apa-Itu-Santri>
- [9] I. A. Gufron, “Santri dan Nasionalisme,” *Islam. Insights J.*, vol. 1, no. 1, pp. 41–45, 2019, doi: 10.21776/ub.ij.2019.001.01.4.
- [10] N. Subiantoro and P. Pangeran, “Implementasi Pelayanan Prima terhadap Santri pada Pondok Pesantren,” *All Fields Sci. J. Liaison Acad. Society*, vol. 2, no. 4, pp. 143–150, 2022, doi: 10.58939/afosj-las.v2i4.460.
- [11] F. N. Alvianta, A. A. Prabowo, and A. Komarudin, “Pembinaan Kesejahteraan Keluarga (Pkk) Dalam Pemberdayaan Keluarga Prasejahtera,” *JISIP (Jurnal Ilmu Sos. dan Pendidikan)*, vol. 5, no. 3, pp. 137–151, 2021, doi: 10.36312/jisip.v5i3.2095.
- [12] bidin A, “Kondisi Sosio-Ekonomi Keluarga Prasejahtera Dengan BalitaGizi Buruk,” *Вестник Росздравнадзора*, vol. 4, no. 1, pp. 9–15, 2017.
- [13] M. A. Aziz and I. Ariyani, “Pemberdayaan Santri melalui Profetik Filantropreneur di Pesantren Planet Nufo Rembang,” *Islam. Manag. Empower. J.*, vol. 2, no. 2, pp. 115–132, 2020, doi: 10.18326/imej.v2i2.115-132.
- [14] A. Nur Khomarudin, “Teknik Data Mining : Algoritma K-Means Clustering,” 2003. [Online]. Available: <https://agusnkhom.wordpress.com>
- [15] D. Marlina and M. Bakri, “Penerapan Data Mining Untuk Memprediksi Transaksi Nasabah Dengan Algoritma C4.5,” *J. Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 23–28, 2021.
- [16] V. Herlinda and D. Darwis, “Analisis Clustering Untuk Recredesialing Fasilitas Kesehatan Menggunakan Metode Fuzzy C-Means,” *Darwis, Dartono*, vol. 2, no. 2, pp. 94–99, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [17] S. R. Hardiyanto Ivan, Purwanto Yudhi, “Implementasi Segmentasi Citra dengan Menggunakan Metode Generalized Fuzzy C-Means Clustering Algorithm with Improved Fuzzy Partitions,” 2012.
- [18] A. Putranto, N. L. Azizah, and I. R. I. Astutik, “Sistem Prediksi Penyakit Jantung Berbasis Web Menggunakan Metode Svm Dan Framework,” *J. Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 4, no. 2, pp. 442–452, 2023, [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/heart+disease>
- [19] H. Judul, F. T. Industri, and U. I. Indonesia, “IDENTIFIKASI HOTSPOT AREA DAN WAKTU SPASIAL CLUSTERING BERBASIS,” 2023.
- [20] M. Syafiih, “Klasifikasi Kategori Berdasarkan Tingkat Ketergantungan Siswa Terhadap Penggunaan Smartphone Di SMK Negeri 1 Suboh Situbondi,” *JEECOM J. Electr. Eng. Comput.*, vol. 5, no. 2, pp. 329–338, 2023, doi: 10.33650/jeeecom.v5i2.6833.