



SKRIPSI

**KLASTERISASI SISWA BERPRESTASI
BERDASARKAN HASIL NILAI PEMBELAJARAN
MENGUNAKAN METODE *FUZZY*
SUBTRACTIVE CLUSTERING BERBASIS *WEB***

MUHAMMAD BAGAS SATRIO WIBOWO
NPM 21083010071

DOSEN PEMBIMBING

Kartika Maulida Hindrayani, S.Kom, M.Kom.
Trimono, S.Si., M.Si.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**



SKRIPSI

KLASTERISASI SISWA BERPRESTASI BERDASARKAN HASIL NILAI PEMBELAJARAN MENGUNAKAN METODE *FUZZY* *SUBTRACTIVE CLUSTERING* BERBASIS WEB

MUHAMMAD BAGAS SATRIO WIBOWO
NPM 21083010071

DOSEN PEMBIMBING

Kartika Maulida Hindrayani, S.Kom, M.Kom.
Trimono, S.Si., M.Si.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**



SKRIPSI

KLASTERISASI SISWA BERPRESTASI BERDASARKAN HASIL NILAI PEMBELAJARAN MENGUNAKAN METODE *FUZZY* *SUBTRACTIVE CLUSTERING* BERBASIS *WEB*

MUHAMMAD BAGAS SATRIO WIBOWO
NPM 21083010071

DOSEN PEMBIMBING

Kartika Maulida Hindrayani, S.Kom, M.Kom.
Trimono, S.Si., M.Si.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**

LEMBAR PENGESAHAN

KLASTERISASI SISWA BERPRESTASI BERDASARKAN HASIL NILAI PEMBELAJARAN MENGGUNAKAN METODE *FUZZY SUBTRACTIVE* *CLUSTERING* BERBASIS *WEB*

Oleh:
MUHAMMAD BAGAS SATRIO WIBOWO
NPM. 21083010071

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Sidang Skripsi Program Studi
Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur
Pada tanggal 9 Juli 2025;

Kartika Maulida Hindrayani, S.Kom, M.Kom.
NIP. 19920909 202203 2 009



(Pembimbing I)

Trimono, S.Si., M.Si.
NIP. 19950908 202203 1 003



(Pembimbing II)

Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T.,
IPU., Asean, Eng.
NIP. 19801205 200501 1 002



(Ketua Penguji)

Aviolla Terza Damaliana, S.Si., M.Stat.
NIP. 19940802 202203 2 015



(Penguji I)

Mengetahui,
Dekan Fakultas Ilmu Komputer



 Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

KLASTERISASI SISWA BERPRESTASI BERDASARKAN HASIL NILAI PEMBELAJARAN MENGGUNAKAN METODE *FUZZY SUBTRACTIVE* *CLUSTERING* BERBASIS *WEB*

Oleh:
MUHAMMAD BAGAS SATRIO WIBOWO
NPM. 21083010071

Telah disetujui untuk mengikuti Ujian Skripsi

Menyetujui,

**Koordinator Program Studi Sains Data
Fakultas Ilmu Komputer**



Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.
NIP. 19801205 200501 1 002

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Bagas Satrio Wibowo
NPM : 21083010071
Program : Sarjana (S1)
Program Studi : Sains Data
Fakultas : Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiarasi. Apabila di kemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 9 Juli 2025

Yang Membuat Pernyataan



Muhammad Bagas Satrio Wibowo
NPM. 21083010071

ABSTRAK

Nama Mahasiswa / NPM : Muhammad Bagas Satrio Wibowo / 21083010071
Judul Skripsi : Klasterisasi Siswa Berprestasi Berdasarkan Hasil
Nilai Pembelajaran Menggunakan Metode *Fuzzy Subtractive Clustering* Berbasis *Web*
Dosen Pembimbing : 1. Kartika Maulida Hindrayani, S.Kom, M.Kom.
2. Trimono, S.Si., M.Si.

Pendidikan merupakan pondasi penting bagi generasi penerus bangsa. Di era digital saat ini, pemanfaatan Teknologi Informasi dan Komunikasi (TIK) dalam bidang pendidikan semakin meningkat, namun tantangan terkait pengelolaan dan analisis data pendidikan juga semakin kompleks. Pertumbuhan jumlah siswa setiap tahunnya mengakibatkan volume data menjadi besar. Pengelolaan data secara manual menjadi tidak efisien, sehingga menghambat identifikasi siswa berprestasi oleh pihak sekolah. Kondisi tersebut menghambat pengembangan potensi siswa secara optimal. Hal ini menunjukkan pada SD XYZ perlunya sistem *Clustering* data untuk memahami siswa berprestasi secara mendalam, dikarenakan saat ini masih dilakukan secara manual tanpa sistem terintegrasi. Penelitian ini mengusulkan pendekatan *Clustering* dengan menggunakan metode *Fuzzy Subtractive Clustering* (FSC) berbasis *web*. Metode *Fuzzy Subtractive Clustering* (FSC) dipilih karena kemampuannya dalam mengidentifikasi kelompok data berdasarkan kerapatan titik data, serta memberikan derajat keanggotaan pada setiap klaster yang terbentuk. Berbeda dengan metode *K-means* yang harus menentukan jumlah klaster ditentukan di awal, FSC lebih fleksibel dalam menemukan pola alami dalam data. Tujuan dari penelitian ini adalah untuk mengelompokkan siswa berprestasi berdasarkan hasil nilai pembelajaran serta mencari parameter terbaik yang digunakan dalam metode FSC guna memperoleh hasil evaluasi *Clustering* yang optimal. Selain itu, penelitian ini juga mengembangkan sistem berbasis *web* untuk mempermudah akses dan penggunaan model *Clustering*. Dari berbagai konfigurasi parameter yang diuji, diperoleh kombinasi parameter terbaik dengan rasio penerimaan = 0,5, rasio penolakan = 0,04, *radius* = 0,3, faktor squash = 1,25 dengan evaluasi terhadap hasil *Clustering* menunjukkan nilai *Silhouette Score* sebesar 0,57 yang termasuk dalam kategori struktur klaster yang baik dengan membentuk lima kelompok siswa.

Kata kunci: *Clustering, Siswa, Fuzzy Subtractive Clustering*

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Muhammad Bagas Satrio Wibowo / 21083010071
Thesis Title : *Clustering Students Based on Learning Score Results Using Fuzzy Subtractive Clustering Method*
Advisor : 1. Kartika Maulida Hindrayani, S.Kom, M.Kom.
2. Trimono, S.Si., M.Si.

Education is an important foundation for the next generation of the nation. In today's digital era, the utilization of Information and Communication Technology (ICT) in education is increasing, but the challenges related to education data management and analysis are also increasingly complex. The growth in the number of students each year results in a large volume of data. Manual data management becomes inefficient, hindering the identification of outstanding students by the school. This condition hinders the optimal development of student potential. This shows SD XYZ the need for a data Clustering system to understand outstanding students in depth, because currently it is still done manually without an integrated system. This research proposes a Clustering approach using the web-based Fuzzy Subtractive Clustering (FSC) method. The Fuzzy Subtractive Clustering (FSC) method was chosen because of its ability to identify data groups based on the density of data points, as well as providing membership degrees to each Cluster formed. Unlike the K-means method that must determine the number of Clusters at the beginning, FSC is more flexible in finding natural patterns in the data. The purpose of this research is to group outstanding students and find the best parameters used in the FSC method to obtain optimal Clustering evaluation results. In addition, this research also developed a web-based system to facilitate access and use of the Clustering model. From various parameter configurations tested, the best parameter combination was obtained with acceptance ratio = 0.5, rejection ratio = 0.04, radius = 0.3, squash factor = 1.25 with an evaluation of the Clustering results showing a Silhouette Score value of 0.57 which is included in the category of good Cluster structure. The results formed five groups of student achievement levels.

Keywords: *Clustering, Student, Fuzzy Subtractive Clustering*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Dengan penuh rasa syukur kepada Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Klasterisasi Siswa Berprestasi Berdasarkan Hasil Nilai Pembelajaran Menggunakan Metode *Fuzzy Subtractive Clustering* Berbasis Web”**. Skripsi ini merupakan salah satu syarat untuk menyelesaikan studi di Universitas Pembangunan Nasional “Veteran” Jawa Timur. Dalam proses penyusunan skripsi ini, penulis mendapatkan banyak dukungan, bimbingan, dan bantuan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis mengucapkan terima kasih kepada :

1. Prof. Dr. Ir. Akhmad Fauzi, M.MT., IPU, selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur
3. Bapak Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU selaku Koordinator Program Studi Sains Data Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Ibu Kartika Maulida Hindrayani, S.Kom, M.Kom., selaku Dosen Pembimbing pertama yang selalu memberikan bimbingan, arahan, dan motivasi selama penulisan skripsi ini.
5. Bapak Trimono S.Si., M.Si., selaku Dosen Pembimbing kedua yang selalu memberikan bimbingan, arahan, dan motivasi selama penulisan skripsi ini.
6. Bapak dan Ibu dosen Program Studi Sains Data UPN “Veteran” Jawa Timur yang telah memberikan ilmu pengetahuan serta kontribusi berharga dalam penyelesaian skripsi ini.
7. Kedua orang tua serta keluarga yang selalu memberikan dukungan moril, materiil, dan doa yang tiada henti.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, baik dari segi materi maupun penyajiannya. Oleh karena itu, penulis sangat mengharapkan saran dan kritik yang membangun demi perbaikan di masa mendatang. Akhir kata, penulis berharap skripsi ini dapat memberikan manfaat bagi

semua pihak yang membacanya, khususnya dalam pengembangan ilmu pengetahuan di bidang teknologi dan pendidikan.

Surabaya, 9 Juli 2025

Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
LEMBAR PENGESAHAN	ii
LEMBAR PERSETUJUAN	iii
SURAT PERNYATAAN BEBAS PLAGIASI.....	iv
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xv
DAFTAR LAMPIRAN	xvii
DAFTAR NOTASI.....	xix
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	5
1.3. Batasan Masalah.....	6
1.4. Tujuan Penelitian	6
1.5. Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA.....	9
2.1. Penelitian Terdahulu	9
2.2. Dasar Teori.....	15
2.2.1. Siswa Berprestasi	15
2.2.2. Penilaian Hasil Belajar	15
2.2.3. Rapor	16
2.2.4. <i>Clustering</i>	17
2.2.5. <i>Fuzzy Subtractive Clustering (FSC)</i>	17
2.2.6. Fungsi <i>Gauss</i>	23
2.2.7. <i>Silhouette Coefficient</i>	24
2.2.8. <i>Website</i>	25

2.2.9. Python.....	26
2.2.9. Streamlit Python	26
BAB III DESAIN DAN IMPLEMENTASI SISTEM.....	29
3.1. Metode dan Langkah Penelitian	29
3.1.1. Variabel Penelitian dan Sumber Data.....	29
3.1.2. Langkah Analisis	31
3.2. Desain Sistem	36
3.2.1. <i>Sequence Diagram</i>	36
3.2.2. Rancangan <i>User Interface</i>	38
BAB IV HASIL DAN PEMBAHASAN	43
4.1. Pengumpulan Data Siswa	43
4.2. Pra Pemrosesan Data	44
4.3. Eksplorasi Data Analisis.....	47
4.4. Transformasi Data	54
4.5. Pemodelan <i>Fuzzy Subtractive Clustering</i>	55
4.6. Eksplorasi Parameter <i>Fuzzy Subtractive Clustering</i>	71
4.7. Model Evaluasi	76
4.8. Implementasi <i>Web Streamlit Clustering FSC</i>	80
4.8.1. Halaman Input Data.....	80
4.8.2. Halaman Visualisasi	81
4.8.3. Halaman <i>Clustering</i>	83
BAB V PENUTUP	87
5.1 Kesimpulan.....	87
5.2 Saran	88
DAFTAR PUSTAKA	91
LAMPIRAN	99

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....	32
Gambar 3. 2 <i>Sequence Diagram</i>	37
Gambar 3. 3 UI Halaman Input Data	39
Gambar 3. 4 UI Halaman Visualisasi.....	39
Gambar 3. 5 UI Halaman <i>Clustering</i> Opsi <i>Cluster</i>	40
Gambar 3. 6 UI Halaman <i>Clustering</i> Opsi Visualisasi	41
Gambar 4. 1 Hasil Duplikasi Data	44
Gambar 4. 2 Pengecekan Nilai Kosong	45
Gambar 4. 3 Visualisasi Nilai Akhir	50
Gambar 4. 4 Visualisasi Jumlah Siswa Perkelas.....	52
Gambar 4. 5 Visualisasi Korelasi Antar Variabel.....	53
Gambar 4. 6 Ilustrasi Model FSC	55
Gambar 4. 7 Pengecekan Kelayakan Pusat <i>Cluster</i>	64
Gambar 4. 8 3D Scatter Plot Hasil Klasterisasi	79
Gambar 4. 9 Hasil Halaman Input Data	81
Gambar 4. 10 Halaman Visualisasi Data	82
Gambar 4. 11 Halaman <i>Cluster</i> Menu Proses <i>Clustering</i>	84
Gambar 4. 12 Menu Visualisasi <i>Cluster</i>	85

Halaman ini sengaja dikosongkan

DAFTAR TABEL

Tabel 2. 1 Studi Literatur	9
Tabel 2. 2. <i>Silhouette Coefficient Details</i>	24
Tabel 3. 1 Variabel Data Siswa.....	30
Tabel 4. 1 Hasil Pengumpulan Data Siswa	43
Tabel 4. 2 Hasil Pra Pemrosesan Data	47
Tabel 4. 3 Hasil Statistika Deskriptif	48
Tabel 4. 4 Hasil <i>Label Encoding</i>	54
Tabel 4. 5 Hasil Standarisasi.....	57
Tabel 4. 6 Potensi Awal Tertinggi	59
Tabel 4. 7 Hasil Pengurangan Potensi Setiap Titik Data	61
Tabel 4. 8 Derajat Keanggotaan.....	69
Tabel 4. 9 Hasil Evaluasi Model	71
Tabel 4. 10 Jarak Parameter yang Dieksplorasi	72
Tabel 4. 11 Hasil Eksplorasi Parameter <i>Radius</i>	73
Tabel 4. 12 Hasil Eksplorasi Parameter <i>Squash factor</i>	73
Tabel 4. 13 Eksplorasi Parameter <i>Accept Ratio</i>	74
Tabel 4. 14 Eksplorasi Parameter <i>Reject Ratio</i>	75
Tabel 4. 15 Eksplorasi parameter terbaik untuk setiap jumlah <i>Cluster</i>	76
Tabel 4. 16 Profil Setiap <i>Cluster</i> Numerik.....	77
Tabel 4. 17 Profil Klaster Berdasarkan Spiritual dan Sosial.....	78

Halaman ini sengaja dikosongkan

DAFTAR LAMPIRAN

Lampiran 1. <i>Letter Of Acceptance</i>	99
Lampiran 2. Kode Program.....	100

Halaman ini sengaja dikosongkan

DAFTAR NOTASI

X_{ij}	:	Data sampel ke-i ($i= 1, 2, \dots, n$) dan atribut ke-j ($j=1, 2, \dots, m$).
n	:	Banyaknya jumlah sampel data
m	:	Banyaknya atribut setiap data
r		<i>Radius</i>
q	:	<i>Squash factor</i>
X_{ijnorm}	:	Data sampel ke-i dan atribut ke-j setelah standarisasi
$Dist_{ij}$	:	Jarak antara titik pusat klaster atau calon pusat dengan data ke-i fitur ke-j
T_j	:	nilai pusat klaster atau calon pusat pada dimensi ke- j , dalam skala normalisasi.
D_i	:	Potensi yang diperoleh dari titik data ke-i
M	:	Data dengan potensi tertinggi iterasi 1
Z		Data dengan potensi tertinggi iterasi 2, 3 dan seterusnya
R	:	Nilai rasio
S_{ij}	:	Selisih antara pusat klaster dan komponen titik data pada dimensi ke-j
C_{lj}	:	Pusat klaster ke-l dengan atribut ke-j
D_{cl_i}	:	Nilai pengurangan potensi dari titik data ke-i
D_i^t	:	Potensi data baru sampel ke-i dengan t menunjukkan banyak iterasi
D_i^{t-1}	:	Potensi data baru sampel ke-i pada iterasi sebelumnya
Sd_l	:	Jarak Standarisasi Kuadrat antara Calon Pusat Klaster dan Pusat Klaster ke-l
V_j	:	Calon pusat klaster

Mds	:	Jarak terdekat data calon pusat klaster dengan pusat klaster
Md	:	Jarak Minimum yang Ditemukan pada Iterasi
$C_{lj}denorm$	:	Pusat klaster ke- l pada atribut ke- j setelah distandarisasi
$x_{max\ j}$	:	Data maksimum pada atribut ke- j
$x_{min\ j}$	:	Data minimum pada atribut ke- j
σ_j	:	Nilai penyebaran data pada dimensi/fitur ke- j
μ_{li}	:	Derajat keanggotaan klaster ke- l pada sampel ke- i
σ_j^2	:	Kuadrat nilai penyebaran data pada dimensi/fitur ke- j
$ A $	:	Jumlah objek dalam klaster yang sama dengan objek i
i,j	:	Indeks yang berkaitan dengan objek
$d_{i,j}$	:	Jarak antara objek ke- i dan objek ke- j
$a(i)$	:	Rata-rata pada jarak objek data i dengan seluruh objek data yang berada pada satu <i>Cluster</i> yang sama
$a(i)$	:	Rata-rata jarak antara objek i dan seluruh objek lain dalam <i>Cluster</i> yang sama
$b(i)$	:	Rata-rata jarak antara objek i dan objek-objek dalam klaster terdekat yang berbeda (<i>Cluster</i> yang tidak mengandung objek i)
$S(i)$	:	<i>Silhouette Coefficient</i> untuk objek i , dengan nilai yang bervariasi antara -1 hingga 1.
p	:	Banyaknya pusat klaster yang ada
Mds	:	Jarak terdekat data calon pusat klaster dengan pusat klaster
Md	:	Jarak antara pusat klaster yang baru ditemukan dan pusat- pusat klaster yang sudah ada sebelumnya.

Sd_l : Merupakan jarak dari calon pusat klaster (yang sedang dipertimbangkan) ke pusat klaster yang ada.

Halaman ini sengaja dikosongkan