

**PENERAPAN ALGORITMA *FUZZY SUBTRACTIVE*
CLUSTERING UNTUK PENGELOMPOKKAN
KABUPATEN/KOTA DI PROVINSI SULAWESI
SELATAN BERDASARKAN DATA LUAS PANEN DAN
PRODUKSI BERAS**

SKRIPSI



NUR ANISA SYAHBANI SALIM

H051181306

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

UNIVERSITAS HASANUDDIN

MAKASSAR

JANUARI 2023

**PENERAPAN ALGORITMA *FUZZY SUBTRACTIVE CLUSTERING* UNTUK
PENGELOMPOKKAN KABUPATEN/KOTA DI PROVINSI SULAWESI
SELATAN BERDASARKAN DATA LUAS PANEN DAN PRODUKSI BERAS**

SKRIPSI

**Diajukan sebagai salah satu syarat memperoleh gelar Sarjana Sains pada
Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu**

Pengetahuan Alam Universitas Hasanuddin

NUR ANISA SYAHBANI SALIM

H051181306

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

UNIVERSITAS HASANUDDIN

MAKASSAR

JANUARI 2023

LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

**PENERAPAN ALGORITMA *FUZZY SUBTRACTIVE CLUSTERING* UNTUK
PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI SULAWESI
SELATAN BERDASARKAN DATA LUAS PANEN DAN PRODUKSI BERAS**

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun.

Makassar, 26 Januari 2023



Nur Anisa Syahbani Salim
NIM H051181306

**PENERAPAN ALGORITMA *FUZZY SUBTRACTIVE CLUSTERING* UNTUK
PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI SULAWESI
SELATAN BERDASARKAN DATA LUAS PANEN DAN PRODUKSI BERAS**

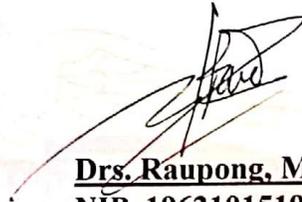
Disetujui Oleh:

Pembimbing Utama,



Anisa, S.Si., M.Si.
NIP. 197302271998022001

Pembimbing Pertama,



Drs. Raupong, M.Si.
NIP. 196210151988101001



Ketua Departemen Statistika



Dr. Nurtiti Sunusi, S.Si., M.Si
NIP. 197201171997032002

Pada 26 Januari 2023

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama : Nur Anisa Syahbani Salim
NIM : H051181306
Program Studi : Statistika
Judul Skripsi : Penerapan Algoritma *Fuzzy Subtractive Clustering*
untuk Pengelompokan Kabupaten/Kota di Provinsi
Sulawesi Selatan Berdasarkan Data Luas Panen dan
Produksi Beras

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

DEWAN PENGUJI

1. Ketua : Anisa, S.Si., M.Si. (.....)
2. Sekretaris : Drs. Raupong, M.Si. (.....)
3. Anggota : Andi Kresna Jaya, S.Si., M.Si. (.....)
4. Anggota : Dr. Nirwan, M.Si. (.....)

Ditetapkan di : Makassar

Tanggal : 26 Januari 2023

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji hanya milik *Allah Subhanallahu Wa Ta'ala* atas segala limpahan rahmat dan hidayah-Nya yang senantiasa diberikan kepada penulis. Shalawat serta salam senantiasa tercurahkan kepada baginda *Rasulullah Muhammad Shallallahu 'Alaihi Wa sallam*, kepada para keluarga, *tabi'in, tabi'ut tabi'in*, serta orang-orang sholeh yang *haq* hingga kadar Allah berlaku atas diri-diri mereka. *Alhamdulillahirobbil'aalamiin*, berkat rahmat dan kemudahan yang diberikan oleh *Allah Subhanallahu Wa Ta'ala*, penulis dapat menyelesaikan skripsi yang berjudul **“Penerapan Algoritma Fuzzy Subtractive Clustering untuk Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Data Luas Panen dan Produksi Beras”** sebagai salah satu syarat memperoleh gelar sarjana pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Dalam penyelesaian skripsi ini, penulis telah melewati perjalanan panjang dengan penuh perjuangan dan pengorbanan yang tidak sedikit. Namun berkat rahmat dan izin-Nya serta dukungan dari berbagai pihak yang turut membantu baik secara moril maupun material sehingga tugas akhir ini dapat terselesaikan di waktu yang terbaik menurut Allah. Oleh karena itu, penulis menyampaikan ucapan terima kasih yang setinggi-tingginya dan penghargaan yang tak terhingga kepada **Ayahanda Mur Salim Saleh, S.Pd., M.Si.** dan **Ibunda Nurhidayah, S.Kep.** yang telah membesarkan dan mendidik penulis dengan penuh kesabaran dan penuh limpahan kasih sayang, cinta dan doa yang senantiasa dilangitkan untuk penulis, adikku tercinta **Nur Afifah Salim** yang selalu menghibur, menyemangati dan mendoakan penulis.

Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan juga penulis ucapkan kepada:

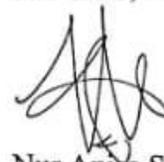
1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.

2. **Bapak Dr. Eng. Amiruddin, S.Si., M.Si.**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Nurtiti Sunusi, S.Si., M.Si.**, selaku Ketua Departemen Statistika, segenap Dosen Pengajar dan Staf yang telah membekali ilmu dan kemudahan kepada penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.
4. **Ibu Anisa, S.Si., M.Si.**, dan **Bapak Drs. Raupong, M.Si.**, selaku Pembimbing Utama dan Pembimbing Pertama yang telah meluangkan waktu untuk membimbing penulis dari awal proposal hingga penyelesaian skripsi ini.
5. **Bapak Andi Kresna Jaya, S.Si., M.Si.**, selaku dosen Penguji sekaligus dosen Penasehat Akademik penulis yang telah memberikan saran dan kritikan yang membangun dalam penyempurnaan penyusunan tugas akhir ini.
6. **Bapak Dr. Nirwan, M.Si.**, selaku dosen Penguji yang telah memberikan saran dan kritikan yang membangun dalam penyempurnaan penyusunan tugas akhir ini.
7. **Marsya Anggun Prisila**, sahabat tercinta penulis yang selalu ada, menemani, membantu, dan memberikan semangat kepada penulis sejak menempuh pendidikan di Universitas Hasanuddin hingga penyelesaian skripsi ini serta memberikan nasehat demi kebaikan penulis.
8. **Kak Irwan Budiansyah**, yang selalu ada menemani, membantu, dan memberikan semangat kepada penulis selama penyusunan skripsi ini.
9. **Yustika, Amalia, Naura, Alfiana, Fitra, Juni, Rezki, Annisa**, teman-teman baik penulis yang selalu membantu dan memberikan semangat selama menempuh pendidikan di Universitas Hasanuddin hingga penyelesaian skripsi ini.
10. **INTEGRAL 2018**, terima kasih telah mengajarkan arti kebersamaan dan memberikan kenangan yang indah bagi penulis terkhusus kepada **Kido, Nasmah, Lutfi, Irfan, Fathur, Luthfi, Ell, Ardi, Fuad, Rizky, Jalil**. Tetap “**Satu Kesatuan, Tak Terpisahkan**”.

11. **STATISTIKA 2018**, terima kasih atas kebersamaan, semangat dan bantuan yang diberikan kepada penulis selama menempuh pendidikan di Universitas Hasanuddin. Terkhusus kepada **Acca, Taufiq, Haksar, Aksa, Ishak, Kahar** dan **Fadhil** yang turut membantu penulis dalam proses penyusunan tugas akhir ini.
12. Keluarga besar **HIMASTAT FMIPA UNHAS** dan **HIMATIKA FMIPA UNHAS** yang telah mengajarkan organisasi yang bermanfaat dan memberikan pengalaman hidup yang tak terlupakan bagi penulis
13. **Keluarga Mahasiswa FMIPA UNHAS**, terima kasih atas perbuatan baik yang diberikan, ilmu yang diajarkan serta kenangan yang diberikan kepada penulis selama menempuh pendidikan di Universitas Hasanuddin. Terkhusus kepada **Pengurus BEM FMIPA UNHAS Periode 2021/2022** dan **KM FMIPA 2018**. Salam **“Use Your Mind Be The Best”**.
14. Terima kasih kepada diri sendiri yang tetap bertahan sejauh ini dan berusaha dalam menyelesaikan tugas akhir ini.
15. Kepada seluruh pihak yang tidak dapat penulis sebutkan satu persatu, terima kasih setinggi-tingginya untuk segala dukungan, motivasi, semangat dan partisipasi yang diberikan kepada penulis, semoga bernilai ibadah di sisi Allah *Subhanahu Wa Ta'ala*.

Penulis menyadari bahwa masih banyak kekurangan dalam skripsi ini, untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini memberikan manfaat bagi pembaca maupun penulis secara khusus. *Wassalamu'alaikum Warahmatullahi Wabarakatuh*.

Makassar, 26 Januari 2023



Nur Anisa Syahbani Salim

**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK
KEPENTINGAN AKADEMIK**

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Skripsi ini diajukan oleh :

Nama : Nur Anisa Syahbani Salim
NIM : H051181306
Program Studi : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty-Free Right*)** atas tugas akhir saya yang berjudul:

“Penerapan Algoritma *Fuzzy Subtractive Clustering* untuk Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Data Luas Panen dan Produksi Beras”

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar pada tanggal 26 Januari 2023

Yang menyatakan,



(Nur Anisa Syahbani Salim)

ABSTRAK

Analisis *cluster* mengelompokkan objek sehingga masing-masing objek mempunyai kemiripan dengan yang lain dalam suatu *cluster*. Analisis *cluster* ada dua metode pengelompokkan, yaitu metode hierarki dan metode non-hierarki. Pada proses pengelompokkan hierarki atau non-hierarki, pembentukan kelompok dilakukan sedemikian rupa sehingga setiap objek berada tepat pada satu kelompok. Akan tetapi, pada suatu saat hal itu tidak dapat dilakukan, karena sebenarnya objek tersebut terletak di antara dua atau lebih kelompok yang lain. Penelitian ini bertujuan untuk memperoleh hasil pengelompokkan Kabupaten/Kota di Sulawesi Selatan menggunakan *Fuzzy Subtractive Clustering* berdasarkan data indikator luas panen dan produksi beras. Pada *fuzzy subtractive clustering* jumlah kelompok yang terbentuk dipengaruhi oleh keragaman data. Selain itu, juga dipengaruhi oleh parameter jari-jari, *squash factor*, *accept ratio*, dan *reject ratio*. Digunakan *Cluster Tightness Measure* untuk pemilihan jari-jari terbaik. Diperoleh hasil 4 *cluster* yang merupakan pengelompokkan terbaik. Berdasarkan hasil perhitungan nilai *Cluster Tightness Measure* dari jari-jari 0,1 hingga 0,9 didapatkan nilai *Cluster Tightness Measure* yang paling minimum pada jari-jari 0,3 dengan nilai *Cluster Tightness Measure* sebesar 0,285594.

Kata Kunci: Analisis Cluster, *Fuzzy Subtractive Cluster*, *Cluster Tightness Measure*, Luas Panen dan Produksi Beras

ABSTRACT

Cluster analysis groups objects so that each object is similar to the others in a cluster. There are two cluster analysis methods of grouping, namely the hierarchical method and the non-hierarchical method. In the hierarchical or non-hierarchical grouping process, group formation is done in such a way that each object is in exactly one group. However, at some point this cannot be done, because in fact the object is located between two or more other groups. This study aims to obtain the results of grouping districts/cities in South Sulawesi using Fuzzy Subtractive Clustering based on indicator data of harvested area and rice production. In fuzzy subtractive clustering the number of groups formed is influenced by the diversity of the data. In addition, it is also influenced by the parameters of the radius, squash factor, accept ratio, and reject ratio. Cluster Tightness Measure is used to select the best radius. The results obtained are 4 clusters which are the best grouping. Based on the calculation results of the Cluster Tightness Measure values from radii 0.1 to 0.9, the minimum Cluster Tightness Measure values are obtained at a radius of 0.3 with a CTM value of 0.285594.

Keywords: *Cluster Analysis, Fuzzy Subtractive Cluster, Cluster Tightness Measure, Harvested Area and Rice Production*

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
LEMBAR PERNYATAAN KEOTENTIKAN	iii
HALAMAN PERSETUJUAN PEMBIMBING	iv
HALAMAN PENGESAHAN	v
KATA PENGANTAR	vi
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	ix
ABSTRAK	x
ABSTRACT	xi
DAFTAR ISI	xii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA	5
2.1 Analisis Multivariat	5
2.2 Analisis <i>Cluster</i>	6
2.3 Logika <i>Fuzzy</i>	6
2.4 <i>Fuzzy Subtractive Clustering</i>	7
2.5 <i>Cluster Tightness Measure</i>	13
2.6 Data Luas Panen dan Produksi Pangan	13
BAB III METODOLOGI PENELITIAN	14
3.1 Sumber Data	14
3.2 Deskripsi Variabel.....	14
3.4 Metode Analisis.....	15

BAB IV HASIL DAN PEMBAHASAN.....	18
4.1 Statistik Deskriptif.....	18
4.2 Menetapkan Parameter.....	18
4.3 Data Normalisasi <i>Min-Max</i>	19
4.4 Potensi Awal Data	19
4.5 Mencari Nilai Potensi Tertinggi dan Rasio	21
4.6 Menghitung Potensi Baru.....	21
4.7 Menentukan Pusat <i>Cluster</i>	22
4.8 Denormalisasi Data	23
4.9 Nilai Sigma <i>Cluster</i>	24
4.10 Nilai Derajat Keanggotaan	24
4.11 Menentukan Jari-jari Terbaik Menggunakan CTM.....	26
BAB V KESIMPULAN DAN SARAN	29
5.1 Kesimpulan.....	29
5.2 Saran.....	29
DAFTAR PUSTAKA	30
LAMPIRAN.....	32

DAFTAR TABEL

Tabel 4. 1 Statistik Deskriptif Seluruh Variabel	18
Tabel 4. 2 Data Normalisasi <i>Min-Max</i>	19
Tabel 4. 3 Hasil Perhitungan Nilai Potensi Awal Data untuk $r = 0,3$	21
Tabel 4. 4 Potensi Data Iterasi Kedua untuk $r = 0,3$	22
Tabel 4. 5 Hasil Pusat <i>Cluster</i> untuk $r = 0,3$	23
Tabel 4. 6 Hasil Denormalisasi Data untuk $r = 0,3$	24
Tabel 4. 7 Hasil Nilai Derajat Keanggotaan untuk $r = 0,3$	25
Tabel 4. 8 Anggota <i>Cluster</i> untuk $r = 0,3$	26
Tabel 4. 9 Nilai CTM	28

DAFTAR LAMPIRAN

Lampiran 1. Data Normalisasi <i>Min-Max</i>	33
Lampiran 2. Potensi Awal untuk $r = 0,1$	34
Lampiran 3. Potensi Awal untuk $r = 0,2$	35
Lampiran 4. Potensi Awal untuk $r = 0,3$	36
Lampiran 5. Potensi Awal untuk $r = 0,4$	37
Lampiran 6. Potensi Awal untuk $r = 0,5$	38
Lampiran 7. Potensi Awal untuk $r = 0,6$	39
Lampiran 8. Potensi Awal untuk $r = 0,7$	40
Lampiran 9. Potensi Awal untuk $r = 0,8$	41
Lampiran 10. Potensi Awal untuk $r = 0,9$	42
Lampiran 11. Potensi Data Iterasi Kedua untuk $r = 0,1$	43
Lampiran 12. Potensi Data Iterasi Kedua untuk $r = 0,2$	44
Lampiran 13. Potensi Data Iterasi Kedua untuk $r = 0,3$	45
Lampiran 14. Potensi Data Iterasi Kedua untuk $r = 0,4$	46
Lampiran 15. Potensi Data Iterasi Kedua untuk $r = 0,5$	47
Lampiran 16. Potensi Data Iterasi Kedua untuk $r = 0,6$	48
Lampiran 17. Potensi Data Iterasi Kedua untuk $r = 0,7$	49
Lampiran 18. Potensi Data Iterasi Kedua untuk $r = 0,8$	50
Lampiran 19. Potensi Data Iterasi Kedua untuk $r = 0,9$	51
Lampiran 20. Hasil Pusat <i>Cluster</i>	52
Lampiran 21. Hasil Denormalisasi Data	53
Lampiran 22. Hasil Nilai Derajat Keanggotaan untuk $r = 0,2$	54
Lampiran 23. Hasil Nilai Derajat Keanggotaan untuk $r = 0,3$	55
Lampiran 24. Hasil Nilai Derajat Keanggotaan untuk $r = 0,4$	56
Lampiran 25. Hasil Nilai Derajat Keanggotaan untuk $r = 0,5$	57
Lampiran 26. Hasil Nilai Derajat Keanggotaan untuk $r = 0,6$	58
Lampiran 27. Hasil Nilai Derajat Keanggotaan untuk $r = 0,7$	59
Lampiran 28. Hasil Nilai Derajat Keanggotaan untuk $r = 0,8$	60

Lampiran 29. Hasil Nilai Derajat Keanggotaan untuk $r = 0,9$ 61
Lampiran 30. Hasil Anggota *Cluster* untuk Setiap Jari-jari..... 62

BAB I PENDAHULUAN

1.1 Latar Belakang

Analisis *cluster* adalah suatu metode statistik yang mengidentifikasi kelompok sampel berdasarkan karakteristik yang sama dengan tingkat homogenitas yang tinggi antar objek menjadi *cluster* yang berbeda dengan heterogenitas objek yang tinggi antar *cluster* (Prasetyo dkk, 2016). Tujuan analisis *cluster* adalah mengelompokkan objek atas dasar karakteristik yang dimiliki. Analisis *cluster* mengelompokkan objek sehingga masing-masing objek mempunyai kemiripan dengan yang lain dalam suatu *cluster* (Wijaya dan Budiman, 2016).

Jang (dalam Azizah dkk, 2018) menjelaskan bahwa analisis *cluster* ada dua metode pengelompokkan, yaitu metode hierarki dan metode non-hierarki. Pada proses pengelompokkan hierarki atau non-hierarki, pembentukan kelompok dilakukan sedemikian rupa sehingga setiap objek berada tepat pada satu kelompok. Akan tetapi, pada suatu saat hal itu tidak dapat dilakukan, karena sebenarnya objek tersebut terletak di antara dua atau lebih kelompok yang lain. Sehingga perlu dilakukan pengelompokkan dengan menggunakan *fuzzy clustering* dimana dalam melakukan pengelompokkan mempertimbangkan tingkat keanggotaan himpunan *fuzzy* sebagai dasar pembobotan.

Ada beberapa metode yang telah dikembangkan dalam analisis *fuzzy clustering*, salah satunya adalah metode *fuzzy subtractive clustering*. (Kusumadewi dan Purnomo, 2013). *Fuzzy subtractive clustering* merupakan algoritma *clustering* tidak terawasi, dalam arti jumlah *cluster* yang akan dibentuk belum diketahui. *Fuzzy subtractive clustering* didasarkan atas ukuran potensi titik-titik data dalam suatu variabel. Konsep dasar dari *fuzzy subtractive clustering* adalah menentukan daerah-daerah dalam suatu variabel yang memiliki potensi tinggi terhadap titik-titik disekitarnya (Kusuma dkk, 2017). Pada *fuzzy subtractive clustering* jumlah kelompok yang terbentuk dipengaruhi oleh keragaman data. Selain itu, juga dipengaruhi oleh parameter jari-jari, *squash*

factor, *accept ratio*, dan *reject ratio*. Nilai jari-jari akan menentukan seberapa besar pengaruh pusat *cluster* terhadap data di sekitarnya, sehingga jari-jari mempengaruhi banyak kelompok yang terbentuk (Anizar, 2017).

Penelitian *fuzzy subtractive clustering* telah dilakukan oleh beberapa peneliti sebelumnya yaitu, Nur Azizah (2018) dengan judul “Penerapan Metode *Fuzzy Subtractive Clustering* (Studi Kasus: Pengelompokan Kecamatan di Provinsi Kalimantan Timur Berdasarkan Luas Daerah dan Jumlah Penduduk Tahun 2015)” hasil analisis menunjukkan terdapat 8 sampai 24 kecamatan yang memiliki kepadatan penduduk tinggi. Dari validitas *cluster* diperoleh hasil terbaik untuk pengkalsteran kecamatan di Kalimantan Timur berdasarkan luas kecamatan dan jumlah penduduk pada tahun 2015 adalah 2 *cluster*, yaitu luas kecamatan kecil tetapi banyak penduduk dan luas kecamatan besar dengan sedikit penduduk. Dalam penelitian lain oleh Vianti Mala Anggraeni Kusuma (2017) dengan judul penelitian “Implementasi Metode *Fuzzy Subtractive Clustering* Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan” dengan hasil analisis dalam penentuan kebakaran hutan/lahan diperoleh hasil potensi tinggi dengan nilai rata-rata *brightness* 335,727°K, FRP 57,248 dan rata-rata *confidence* 83,47% dan potensi sedang dengan nilai rata-rata *brightness* 318,934°K, FRP 23,330 dan rata-rata *confidence* 58,08%. Penelitian oleh Onny Kartika Hitasari (2015) dengan judul “Pengelompokan Kabupaten/Kota di Provinsi Jawa Tengah Berdasarkan Angka Partisipasi Pendidikan Jenjang SMA/MA/Paket C dengan *Fuzzy Subtractive Clustering*” hasil penelitian menunjukkan angka partisipasi kelompok *cluster* terbaik di provinsi Jawa Tengah yang terdiri dari 4 *cluster* dengan nilai varian *cluster* 0,00749 dan radius antara 0,35 sampai dengan 0,50. *Fuzzy subtractive clustering* juga biasanya diimplementasikan pada berbagai bidang salah satunya adalah bidang pertanian.

Pertanian merupakan kontributor terbesar dalam pembentukan nilai tambah Sulawesi Selatan. Pada tahun 2019, pertanian menyumbangkan 21,28% atau sebesar Rp107,42 triliun terhadap nilai tambah Sulawesi Selatan. Disusul perdagangan besar dan eceran yang memberikan kontribusi sebesar 14,8% atau

senilai Rp74,7 triliun. Sementara itu tempat ketiga di tempati oleh konstruksi yang memberikan kontribusi terbesar 14,18% atau senilai Rp71,56 triliun. Tingginya kontribusi nilai tambah sektor pertanian bukan saja menjadikan sektor ini sebagai motor penggerak perekonomian Sulawesi Selatan namun juga menobatkan Sulawesi Selatan sebagai salah satu wilayah lumbung pangan di Indonesia.

Saat ini pemerintah sedang gencar melaksanakan berbagai program untuk mendukung pencapaian tujuan kedua *Sustainable Development Goals* (SDG's) yaitu mengakhiri kelaparan melalui upaya mencapai ketahanan pangan, meningkatkan gizi dan mendukung pertanian berkelanjutan. Upaya tersebut tentunya membutuhkan dukungan data pangan yang akurat. Salah satu data yang sangat penting untuk pengambilan kebijakan di bidang pangan adalah luas panen dan produksi beras. Sekitar 38,08% kebutuhan kalori penduduk Indonesia masih dipenuhi dari beras. Hal ini menjadikan beras sebagai salah satu komoditas pangan yang strategis secara nasional. Polemik kebijakan terkait masalah beras seringkali terjadi karena simpang siurnya data luas panen, produksi, konsumsi hingga distribusi beras (BPS Sulsel, 2020)

Dari uraian di atas menandakan *fuzzy subtractive clustering* dapat diterapkan pada beberapa jenis kasus, salah satunya luas lahan panen dan produksi beras. Berdasarkan penelitian sebelumnya, maka skripsi ini akan dikaji mengenai *fuzzy subtractive clustering* dalam mengelompokkan kabupaten di Sulawesi Selatan berdasarkan luas panen dan produksi beras pada tahun 2020.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, masalah yang akan dikaji dalam penelitian ini adalah bagaimana pengelompokkan Kabupaten/Kota di Sulawesi Selatan menggunakan *Fuzzy Subtractive Clustering* berdasarkan data indikator luas panen dan produksi beras?

1.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini yaitu:

1. Data yang digunakan dalam penelitian ini terbatas pada data tahun 2020 dengan indikator Luas Panen Padi, Produksi Padi, dan Produksi Beras dari 24 kabupaten/kota di Sulawesi Selatan.
2. Pemilihan jari-jari terbaik menggunakan *Cluster Tighness Cluster*.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah untuk memperoleh hasil pengelompokan Kabupaten/Kota di Sulawesi Selatan menggunakan *Fuzzy Subtractive Clustering* berdasarkan data indikator luas panen dan produksi beras.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat akademisi yaitu menambah pengetahuan mengenai *fuzzy subtractive clustering*. Selain itu hasil analisis juga dapat digunakan sebagai informasi mengenai luas panen dan produksi beras di provinsi Sulawesi Selatan pada tahun 2020.

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Multivariat

Analisis multivariat berhubungan dengan metode-metode statistik yang secara bersama-sama melakukan analisis terhadap lebih dari dua variabel pada setiap objek. Dapat dikatakan analisis multivariat merupakan perluasan dari analisis univariat atau bivariat (Santoso, 2017). Analisis multivariat berasal dari akar kata *multi* dan *variate*. Dari dua akar kata ini maka analisis statistika tertentu yang menganalisis banyak variabel secara simultan. Variabel di dalam analisis multivariat dibagi menjadi variabel dependen dan variabel independen. Sebagai sebuah alat analisis statistika yang bersifat general, analisis multivariat terdiri dari beberapa jenis metode. Metode dalam multivariat dapat dikelompokkan menjadi metode dependen dan independent. Metode dependen yaitu ketika variabel dependen dipengaruhi oleh variabel independent. Sedangkan metode independent merupakan metode dimana semua variabel saling berpengaruh. Dengan kata lain semua adalah variabel independent (Ramadhani, dkk., 2018).

Teknik analisis multivariat secara dasar diklasifikasi menjadi dua, yaitu analisis dependensi dan analisis interdependensi. Analisis dependensi berfungsi untuk menerangkan atau memprediksi variabel tergantung dengan menggantungkan dua atau lebih variabel bebas. Adapun yang termasuk dalam klasifikasi analisis dependensi ialah analisis regresi linear berganda, analisis diskriminan, analisis varian multivariat (MANOVA), dan analisis korelasi kanonikal. Sedangkan analisis interdependensi berfungsi untuk memberikan makna terhadap seperangkat variabel atau membuat kelompok-kelompok secara bersama-sama. Adapun yang termasuk dalam klasifikasi analisis interdependensi ialah analisis faktor, analisis *cluster*, dan *multidimensional scaling* (Wijaya dan Budiman, 2016).

2.2 Analisis Cluster

Analisis *cluster* adalah salah satu analisis multivariat yang digunakan untuk mengelompokkan objek-objek menjadi beberapa *cluster* berdasarkan kemiripan variabel-variabel yang diamati, sehingga diperoleh kemiripan objek dalam *cluster* yang berbeda (Goreti, dkk., 2016). Prosedur analisis *cluster* digunakan untuk mengidentifikasi kelompok kasus yang secara relatif sama yang didasarkan pada karakteristik yang sudah dipilih dengan menggunakan algoritma yang dapat mengatur kasus dalam jumlah besar. Hasil *cluster* suatu objek harus memiliki internal (*within cluster*) homogenitas yang tinggi dan memiliki eksternal (*between cluster*) heterogenitas yang tinggi. Jika pengelompokan berhasil, maka objek dalam suatu *cluster* akan saling dekat satu sama lain jika diplot secara geometri dan *cluster* yang berbeda akan saling menjauh satu sama lain (Wijaya dan Budiman, 2016).

2.3 Logika Fuzzy

Logika *fuzzy* pertama kali diperkenalkan oleh Lotfi A. Zadeh pada tahun 1965. Logika *fuzzy* merupakan suatu logika yang memiliki nilai kekaburan atau kesamaran antara benar atau salah. Dalam teori logika *fuzzy* suatu nilai bias bernilai benar atau salah secara bersama. Namun berapa besar keberadaan dan kesalahannya tergantung pada derajat keanggotaan yang dimilikinya. Logika *fuzzy* memiliki derajat keanggotaan dalam rentang 0 hingga 1 ((Nasution, 2012)

Menurut Cox (Kusumadewi dan Purnomo, 2013), ada beberapa alasan mengapa orang menggunakan logika *fuzzy*, antara lain:

1. Konsep logika *fuzzy* mudah dimengerti. Karena logika *fuzzy* menggunakan dasar teori himpunan, maka konsep matematis yang mendasari penalaran *fuzzy* tersebut cukup mudah untuk dimengerti.
2. Logika *fuzzy* sangat fleksibel, artinya mampu beradaptasi dengan perubahan-perubahan, dan ketidakpastian yang menyertai permasalahan.
3. Logika *fuzzy* memiliki toleransi terhadap data yang tidak tepat. Jika diberikan sekelompok data yang cukup homogen, dan kemudian ada

beberapa data yang “eksklusif”, maka logika *fuzzy* memiliki kemampuan untuk menangani data eksklusif tersebut.

4. Logika *fuzzy* mampu memodelkan fungsi-fungsi nonlinear yang sangat kompleks.
5. Logika *fuzzy* dapat membangun dan mengaplikasikan pengalaman-pengalaman para pakar secara langsung tanpa harus melalui proses pelatihan. Dalam hal ini, sering dikenal dengan nama *Fuzzy Expert Systems* menjadi bagian terpenting.
6. Logika *fuzzy* dapat bekerjasama dengan teknik-teknik kendali secara konvensional. Hal ini umumnya terjadi pada aplikasi di bidang teknik mesin maupun teknik elektro.
7. Logika *fuzzy* didasarkan pada bahasa alami. Logika *fuzzy* menggunakan Bahasa sehari-hari sehingga mudah dimengerti.

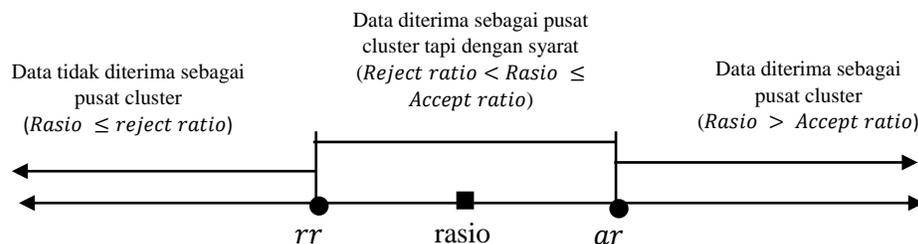
2.4 Fuzzy Subtractive Clustering

Fuzzy subtractive clustering didasarkan atas ukuran potensi titik-titik data dalam suatu ruang variabel. Konsep dasar dari *subtractive clustering* adalah menentukan daerah-daerah dalam suatu variabel yang memiliki potensi tinggi terhadap titik-titik di sekitarnya. Titik dengan jumlah tetangga terbanyak akan dipilih sebagai pusat *cluster*. Titik yang sudah terpilih sebagai pusat *cluster* ini kemudian akan dikurangi potensinya. Kemudian algoritma akan memilih titik lain yang memiliki tetangga terbanyak untuk dijadikan pusat *cluster* yang lain. Hal ini akan dilakukan berulang-ulang hingga semua titik diuji (Kusumadewi dan Purnomo, 2013).

Pada implementasinya, bisa digunakan 2 faktor pembanding, yaitu *Accept ratio* dan *Reject ratio*. Baik *accept ratio* maupun *reject ratio* keduanya merupakan suatu bilangan antara 0 sampai 1. *Accept ratio* merupakan batas bawah dimana suatu titik data yang menjadi calon pusat cluster diperbolehkan untuk menjadi pusat *cluster*. Sedangkan *reject ratio* merupakan batas atas dimana suatu titik data menjadi calon pusat *cluster* tidak diperbolehkan untuk menjadi

pusat cluster. Pada suatu iterasi, apabila telah ditemukan suatu titik data dengan potensi tertinggi (misalkan X_k dengan potensi D_k), kemudian dilanjutkan dengan mencari rasio potensi titik data tersebut dengan potensi tertinggi suatu titik data pada awal iterasi (misalkan X_h dengan potensi D_h). Hasil bagi antara D_k dengan D_h ini kemudian disebut dengan Rasio ($Rasio = D_k/D_h$). Ada 3 kondisi yang bisa terjadi dalam suatu iterasi (Kusumadewi dan Purnomo, 2013):

1. Apabila $Rasio > Accept\ ratio$, maka titik data tersebut diterima sebagai pusat cluster baru.
2. Apabila $Reject\ ratio < Rasio \leq Accept\ ratio$, maka titik data tersebut akan diterima sebagai pusat *cluster* baru apabila data tersebut terletak pada jarak yang cukup jauh dengan pusat *cluster* yang lainnya. Apabila hasil penjumlahan antara rasio dan jarak terpanjang titik data tersebut dengan pusat *cluster* lainnya kurang dari 1, maka selain titik data tersebut tidak akan diterima sebagai pusat cluster dan tidak akan dipertimbangkan lagi untuk menjadi pusat cluster baru.
3. Apabila $Rasio \leq reject\ ratio$, maka sudah tidak ada lagi titik data yang akan dipertimbangkan untuk menjadi calon pusat *cluster* atau dengan kata lain iterasi dihentikan.



Gambar 2. 1 Rasio, Accept Ratio, dan Reject Ratio

Adapun algoritma dari *fuzzy subtractive clustering* adalah sebagai berikut (Kusumadewi dan Purnomo, 2013):

1. Menginput data yang akan dicluster berupa matriks X_{ij} berukuran $n \times m$, dengan $i = 1,2, \dots, n$; dan $j = 1,2, \dots, m$ dimana n adalah banyaknya data dan m adalah banyaknya variabel.

2. Menentukan nilai:

a. Jari-jari.

Jari-jari (r) merupakan konstanta positif berupa vektor yang akan menentukan seberapa besar pengaruh pusat *cluster* pada tiap variabel. Jari-jari bernilai antara 0 sampai 1.

b. *Squash factor*

Squash factor (q) merupakan suatu konstanta untuk menetapkan besar jarak titik-titik data sekitar pusat *cluster* yang akan diukur penurunan potensi datanya. Pada beberapa penelitian sebelumnya, seperti penelitian oleh Nur Azizah (dkk) menggunakan nilai *squash factor* sebesar 1,25 (Azizah dkk, 2018), penelitian oleh Hitasari (dkk) juga menggunakan nilai *squash factor* sebesar 1,25 (Hitasari dkk, 2015), dan penelitian oleh Andriany juga menggunakan nilai *squash factor* sebesar 1,25 (Andriany dkk, 2021).

c. *Accept ratio*

Accept ratio (ar) merupakan batas bawah dimana suatu titik data yang menjadi calon pusat *cluster* diperbolehkan untuk menjadi pusat *cluster*. Pada beberapa penelitian sebelumnya, seperti penelitian oleh Nur Azizah (dkk) menggunakan nilai *accept ratio* sebesar 0,5 (Azizah dkk, 2018), penelitian oleh Hitasari (dkk) juga menggunakan nilai *accept ratio* sebesar 0,5 (Hitasari dkk, 2015), dan penelitian oleh Andriany juga menggunakan nilai *accept ratio* sebesar 0,5 (Andriany dkk, 2021).

d. *Reject ratio*

Reject ratio (rr) merupakan batas atas dimana suatu titik data yang menjadi calon pusat *cluster* tidak diperbolehkan untuk menjadi pusat *cluster*. Pada beberapa penelitian sebelumnya, seperti penelitian oleh Nur Azizah (dkk) menggunakan nilai *reject ratio* sebesar 0,15 (Azizah dkk, 2018), penelitian oleh Hitasari (dkk) juga menggunakan nilai *reject ratio* sebesar 0,15 (Hitasari dkk, 2015), dan penelitian oleh Andriany juga menggunakan nilai *reject ratio* sebesar 0,15 (Andriany dkk, 2021).

- e. Nilai minimum dan maximum data.
3. Melakukan normalisasi data dengan *min-max normalization*:

$$X_{ijnorm} = \frac{X_{ij} - X_{minj}}{X_{maxj} - X_{minj}}, i = 1, 2, \dots, n; j = 1, 2, \dots, m \quad (2.1)$$

Keterangan:

X_{ijnorm} = data ternormalisasi.

X_{ij} = data ke- i variabel ke- j

X_{minj} = data minimum variabel ke- j

X_{maxj} = data maximum variabel ke- j

4. Menentukan potensi awal tiap-tiap titik data berdasarkan rumus berikut,

Jika $m = 1$, maka

$$D_i = \sum_{i=1}^n \exp \left(-4 \left(\frac{X_{jnorm}^i - X_{ijnorm}}{r} \right)^2 \right) \quad (2.2)$$

Jika $m > 1$, maka

$$D_i = \sum_{i=1}^n \exp \left(-4 \left(\sum_{j=1}^m \left(\frac{X_{jnorm}^i - X_{ijnorm}}{r} \right)^2 \right) \right) \quad (2.3)$$

Keterangan:

D_i = nilai potensi awal dari data ke- i

X_{jnorm}^i = nilai ternormalisasi variabel ke- j pada data tetap ke- i

r = jari-jari

5. Mencari titik dengan nilai potensi tertinggi,

$$M = \max [D_i | i = 1, 2, \dots, n]; \text{ untuk iterasi ke-1.} \quad (2.4)$$

$$Z = \max [D_i^t | i = 1, 2, \dots, n]; \text{ untuk iterasi ke-2, 3, dan seterusnya.} \quad (2.5)$$

6. Menghitung nilai rasio,

$$R = \frac{Z}{M} \quad (2.6)$$

Dimana untuk iterasi pertama nilai $Z = M$ sehingga nilai rasio untuk iterasi pertama adalah 1.

7. Menentukan pusat *cluster* dari nilai rasio yang telah diperoleh, maka terdapat tiga kondisi yang kemungkinan terjadi:

1) Jika rasio $>$ *accept* rasio, calon pusat *cluster* dapat diterima sebagai pusat *cluster* baru dan ditulis sebagai C_l . Kemudian mengurangi potensi dari titik-titik di dekat pusat *cluster*:

$$D_i^t = D_i^{t-1} - \left(M \cdot \exp \left(-4 \left(\sum_{j=1}^m \left(\frac{C_{lj} - X_{ijnorm}}{r \cdot q} \right)^2 \right) \right) \right) \quad (2.7)$$

Keterangan:

D_i^t = potensi data baru sampel ke- i pada iterasi ke- t

D_i^{t-1} = potensi data baru sampel ke- i pada iterasi sebelumnya

M = potensi data tertinggi pada iterasi pertama

C_{lj} = pusat *cluster* ke- l dengan variabel ke- j

q = *squash factor*

2) Jika *reject* rasio $<$ rasio \leq *accept* rasio, calon pusat *cluster* baru akan diterima sebagai pusat *cluster* jika letaknya cukup jauh dengan pusat *cluster* lain yang telah ada dan dihitung dalam persamaan berikut:

$Md = -1$ untuk $l = 1$ sampai $l = p$; $p =$ banyaknya *cluster*

$$Sd_l = \sum_{j=1}^m \left(\frac{V_j - C_{lj}}{r} \right)^2 \quad (2.8)$$

Keterangan:

V_j = calon pusat *cluster*

C_{lj} = pusat *cluster* ke- l pada variabel ke- j

Jika ($Md < 0$) atau ($Sd < Md$) maka $Md = Sd_l$. Jika ($Sd_l > Md$) maka Md tidak berubah,

$$Mds = \sqrt{Md} \quad (2.9)$$

dengan Mds adalah jarak terdekat data calon pusat *cluster* dengan pusat *cluster*. Jika ($\text{rasio} + Mds$) $<$ 1, calon pusat *cluster* tidak diterima dan tidak akan dipertimbangkan kembali sebagai pusat *cluster* dan potensi

data tersebut diset menjadi 0. Jika $(\text{rasio} + Mds) \geq 1$, calon pusat *cluster* diterima sebagai *cluster* baru.

- 3) Apabila $\text{rasio} \leq \text{reject rasio}$, maka sudah tidak ada lagi titik data yang akan dipertimbangkan untuk menjadi kandidat pusat *cluster*, iterasi dihentikan.

Kemudian kembali mencari nilai rasio dan mengulangi langkah ke-7.

8. Mengembalikan pusat *cluster* dari bentuk ternormalisasi ke bentuk semula (denormalisasi).

$$C_{ljdenorm} = C_{lj} \cdot (X_{maxj} - X_{minj}) + X_{minj} \quad (2.10)$$

Keterangan:

$C_{ljdenorm}$ = pusat *cluster* ke-*l* pada variabel ke-*j* setelah didenormalisasi

C_{lj} = pusat *cluster* ke-*l* pada variabel ke-*j* (bentuk ternormalisasi)

$XMax_j$ = data maksimum pada variabel ke-*j*

$XMin_j$ = data minimum pada variabel ke-*j*

9. Menghitung nilai sigma *cluster*

$$\sigma_j = r \cdot \frac{X_{maxj} - X_{minj}}{\sqrt{8}} \quad (2.11)$$

Hasil dari algoritma *Fuzzy Subtractive Clustering* ini berupa matriks pusat *cluster* (C) dan sigma (σ) akan digunakan untuk menentukan nilai parameter fungsi keanggotaan Gauss. Derajat keanggotaan suatu titik data X_i pada *cluster* ke-*l* adalah (Kusumadewi dan Purnomo, 2013):

$$\mu_{li} = \exp\left(-\sum_{j=1}^m \frac{(X_{ij} - C_{lj})^2}{2\sigma_j^2}\right) \quad (2.12)$$

Keterangan:

μ_{li} = derajat keanggotaan *cluster* ke-*l* pada data ke-*i*

σ_j^2 = kuadrat sigma *cluster*

2.5 Cluster Tightness Measure

Hasil pengelompokan yang optimal adalah pada saat anggota-anggota dari suatu *cluster* memiliki kemiripan yang dekat. Hal tersebut ditentukan menggunakan *Cluster Tightness Measure* atau *CTM*. Metode *CTM* ini merupakan perhitungan yang didasarkan oleh simpangan baku dari beberapa kelompok dengan beberapa variabel. Nilai *CTM* terkecil menunjukkan metode pengelompokan menghasilkan *cluster* yang terbaik (Andriany dkk, 2021).

$$CTM = \frac{1}{p} \sum_{l=1}^p \left(\frac{1}{m} \sum_{j=1}^m \frac{\sigma_j^l}{\sigma_j} \right) \quad (2.13)$$

Keterangan:

- p = banyaknya *cluster* yang terbentuk
- m = banyaknya variabel
- σ_j^l = simpangan baku *cluster* ke- l dengan variabel ke- j
- σ_j = simpangan baku seluruh data pada variabel ke- j .

2.6 Data Luas Panen dan Produksi Pangan

Berdasarkan publikasi BPS tahun 2020, data luas panen selama ini dikumpulkan digunakan metode konvensional menggunakan daftar isian Statistik Pertanian. Berdasarkan metode tersebut, pengumpulan data luas panen didasarkan pada hasil pandangan mata petugas pengumpul data. Meskipun secara praktikal, metode tersebut mudah diterapkan, tetapi penggunaan metode tersebut memiliki kekurangan. Rendahnya akurasi dan waktu pengumpulan data yang cukup lama menjadi beberapa kekurangan dari penggunaan metode tersebut.

Sehingga digunakan metode baru yaitu Kerangka Sampel Area (KSA) yang bertujuan untuk memperbaiki metode pengumpulan data menjadi lebih objektif dan modern dengan melibatkan peranan teknologi di dalamnya, sehingga data pertanian yang dikumpulkan menjadi lebih akurat dan tepat waktu.