

Klasifikasi Data Tidak Lengkap Dengan Pendekatan *Fuzzy Grid Partition*

Murni Marbun¹, Erwin Panggabean², Ricky Martin Ginting³, Robertus Rinaldy Pakpahan⁴

^{1,2,3,4} STMIK Pelita Nusantara Medan
Jl. Iskandar Muda No. 1 Medan, 20154, Indonesia

e-mail: ¹dimpleflorence@yahoo.co.id, ²erwinpanggabean8@gamil.com,
³rickymartinginting503@gmail.com, ⁴naldipakpahan21@gmail.com

Informasi Artikel

Diterima: 00-00-2021

Direvisi: 00-00-2021

Disetujui: 00-00-2021

Abstrak

Klasifikasi data tidak lengkap dapat di proses langsung dengan cara tertentu untuk mendapatkan aturannya atau diperoleh dari pengetahuan para pakar. Ketergantungan terhadap pakar akan sulit memodelkan implikasi logis manusiana, tidak tersedianya *framework* proses pemodelan, dan biaya pakar yang mahal. Kesulitan tersebut dapat diatasi dengan mendapatkan aturan dari data yang bersifat *uncertain* dengan menerapkan metode dari sistem *fuzzy* yang dibangun berdasarkan konsep *fuzzy if-then rules*. Pendekatan metode pada penelitian ini adalah metode *fuzzy grid partition* untuk mengklasifikasikan data tidak lengkap. Data yang digunakan adalah data cuaca yang terdiri data kelembaban udara sebagai konklusi, data temperatur, curah hujan, lamanya peninjoran matahari dan kecepatan angin sebagai antecedent. Tahapan penelitian dimulai dengan menginput data set tidak lengkap, merubah data tidak lengkap menjadi data lengkap, menormalisasi data, membangkitkan aturan dan melakukan proses klasifikasi data. Hasil penelitian menghasilkan 22 aturan untuk mengklasifikasi data dengan tingkat akurasi 66,67%, tingkat error 33,33% dan jumlah data *unclass* adalah 0.

Kata Kunci: klasifikasi data tidak lengkap, *fuzzy if-then rules*, metode *fuzzy grid partition*

Abstract

Classification of incomplete data can be processed directly in a certain way to get the rules or obtained from the knowledge of experts. Dependence on experts will be difficult to model the logical implications of humans, the unavailability of a modeling process framework, and expensive expert fees. This difficulty can be overcome by obtaining rules from uncertain data by applying methods from a fuzzy system that is built on the concept of fuzzy if-then rules. The method approach in this research is the fuzzy grid partition method to classify incomplete data. The data used are weather data consisting of air humidity data as a conclusion, temperature data, rainfall, duration of sunlight and wind speed as antecedents. The research phase begins with inputting incomplete data sets, changing incomplete data into complete data, normalizing data, generating rules and carrying out data classification processes. The research result are 22 rules for classifying data with an accuracy rate of 66.67%, error rate of 33.33% and the number of unclass data is 0.

Keywords: *Classification of incomplete data, Fuzzy if-then rules, Fuzzy grid partition*

1. Pendahuluan

Sebagian besar database tidak terhindar dari masalah ketidaklengkapan informasi yang berkaitan dengan nilai-nilai yang hilang atau salah. Berbagai alasan yang berbeda

mengakibatkan ketidaklengkapan dalam data. Kesalahan prosedur manual entri data, pengukuran yang salah, kesalahan peralatan, dan lainnya merupakan salah satu penyebab ketidaklengkapan data (Sadiq et al., 2012)



Dataset tidak lengkap adalah data yang nilai-nilai atribut dari beberapa objek di dataset tidak diketahui, setidaknya satu objek di dataset mempunyai *missing value* maka diklasifikasikan sebagai dataset tidak lengkap. Dataset tidak lengkap dapat diubah menjadi dataset lengkap dengan pendekatan mengganti nilai atribut yang hilang dengan rata-rata atau median dari atribut untuk semua objek yang memiliki kelas keputusan yang sama, seperti yang dikemukakan Han *et al.* (2012). Data yang telah lengkap tersebut dapat di lakukan klasifikasi untuk menemukan model atau fungsi yang menggambarkan dan membedakan kelas setiap data atau konsep data.

Model data diperoleh berdasarkan analisis dari serangkaian data pelatihan, yaitu kumpulan objek data dimana label kelas dari objek diketahui. Model ini digunakan untuk mengklasifikasi label kelas sebagai objek pada data uji. Klasifikasi merupakan penentuan sebuah *record* data baru ke salah satu beberapa kategori dimana telah ditentukan sebelumnya (Sinambela *et al.*, 2016) Keakuratan klasifikasi pada data uji adalah persentase dari data uji yang telah benar diklasifikasikan oleh model, dimana label kelas masing-masing data uji yang berkaitan dibandingkan dengan hasil klasifikasi kelas dari model (Han J *et al.*, 2014).

Penelitian ini adalah lanjutan dari penelitian sebelumnya yang membahas tentang klasifikasi sistem *fuzzy* berdasarkan aturan *if-then* menggunakan metode *grid partition* (Sitompul *et al.*, 2017).

Pada penelitian sebelumnya data yang digunakan dalam penelitian ini adalah data lengkap yang diambil dari UC *Irvine machine learning repository*, yang merupakan kumpulan dataset yang digunakan oleh *machine learning community* dalam melakukan analisis terhadap suatu algoritma *machine learning*. Namun, melihat dari berbagai media penyedia dataset masih ditemukan ketidaklengkapan informasi yang disebabkan kesalahan prosedur manual entri data, pengukuran yang salah, kesalahan peralatan, dan lainnya.

Model klasifikasi data tidak lengkap dapat direpresentasikan dalam bentuk aturan *fuzzy* (aturan *If-then*). Aturan *fuzzy* akan memetakan ruang input kedalam ruang keluaran sehingga menghasilkan keluaran *fuzzy* (Marbun *et al.*, 2019). Aturan *fuzzy if-then* dapat dibangkitkan melalui pelatihan, antara lain: Liu *et al.* (2013), dengan pendekatan *an axiomatic fuzzy sets decision tree*. Dahal *et al.* (2015) dengan menggunakan algoritma genetika. Elkano *et al.* (2016) dengan menggunakan strategi dekomposisi *biner* sedangkan proses klasifikasi *fuzzy* menggunakan metode One Vs

One (OVO) dan One Vs All (OVA) untuk dekomposisi binernya.

Pendekatan metode *fuzzy grid-partition* dapat digunakan dalam perancangan sistem *fuzzy* karena lebih mudah digunakan (Hartono, 2016). Beberapa penelitian untuk mengklasifikasi pola dengan pendekatan *fuzzy grid partition* antara lain: Chen *et al.* (2014) membangkitkan beberapa *rules* dengan menggunakan *grid partition* sederhana, kemudian melakukan perubahan bobot *rules* dengan algoritma *particle swarm optimization* dan Takahashi *et al.* (2015) melakukan seleksi *rules* yang dibangkitkan dengan *grid-partition* terdistribusi menggunakan dua fungsi objektif yang saling berotasi. Selain mengklasifikasi pola, metode *grid partition* juga digunakan untuk mengkombinasikan pendekatan tekstur, *grey-level co-occurrence matrices* dengan sistem inferensi *fuzzy* berbasis *grid partition*, bernama ANFIS-GRID untuk klasifikasi radar yang bergema (Sadouki & Haddad, 2016), untuk memilih fitur berdasarkan *Grid Partition Data Field* (FS-GPDF) (Field & Zhao, 2018), dan Mao *et al.* (2020) menerapkan algoritma *Quantum-inspired Quantum-behaved Particle Swarm Optimization* (QiQPSO) untuk membangun klasifikasi awal sistem *fuzzy* dan metode *grid* untuk mempartisi ruang fitur, dan kemudian basis aturan *fuzzy* dioptimalkan lebih lanjut oleh QiQPSO untuk mengurangi jumlah aturan *fuzzy* sehingga meningkatkan interpretabilitas akurasi klasifikasi yang lebih tinggi.

Sitompul *et al.* (2017) menggunakan pendekatan *adaptive distributed grid-partition*. Dengan metode ini, jumlah aturan yang dihasilkan dapat disesuaikan dengan kebutuhan sehingga akurasi sistem serta kompleksitas waktu inferensi dapat dikelola dengan baik. Pendekatan *grid-partition* teradaptasi dapat membangkitkan *fuzzy rules* dengan jumlah yang lebih sedikit bila dibandingkan dengan pendekatan *grid-partition* terdistribusi dengan kriteria berhenti yang sama, yaitu ketika akurasi *rules* yang dibangkitkan terhadap rekam data *learning* telah mencapai 100%.

Tujuan dari penelitian ini adalah mengklasifikasi data tidak lengkap dengan pendekatan metode *fuzzy grid partition*. Klasifikasi data dengan metode *fuzzy grid partition* dimensi n dengan m label adalah $X^p = (X_1^p, X_2^p \dots X_n^p)$, $p = 1, 2 \dots m$, merupakan himpunan pelatihan yang dilambangkan sebagai data pelatihan dan kelas C sebagai $y_1, y_2 \dots y_c$. X_j^p adalah contoh nilai ke-i dari atribut p. Maka pembangkitan aturan *fuzzy if-then* dari data numerik mencakup dua fase yaitu (Borgi, 2018):

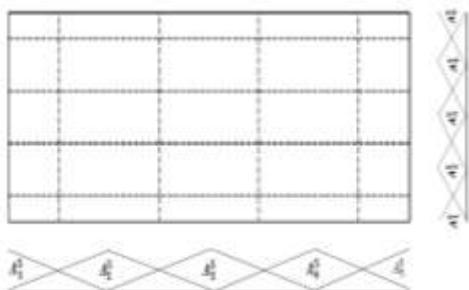
1. Fase Pertama

Pada fase pertama, untuk membangkitkan aturan *fuzzy* menggunakan pendekatan *grid*

partition, ruang pencarian dibagi menjadi beberapa partisi untuk membentuk struktur *grid*. Setiap atribut numerik dipartisi menjadi subset K fuzzy $\{A_1, A_2, \dots, A_K\}$, di mana setiap subset ditentukan oleh fungsi keanggotaan sehingga diperoleh *fuzzy grid*.

2. Fase kedua

Pada fase kedua, aturan *fuzzy if-then* dibangun untuk setiap subruang dari *fuzzy grid*. *Grid* yang terbentuk oleh kombinasi relasi dari setiap partisi akan menghasilkan aturan. Gambar 1 merupakan contoh *grid* fuzzy sederhana di mana semua atribut dipartisi dengan jumlah yang sama yaitu $K = 5$ dari subset fuzzy. Dengan demikian jumlah aturan *fuzzy* yang dihasilkan adalah 5^2 . Secara umum, jika masalah dimensi-n dengan subset K fuzzy untuk partisi fuzzy, maka aturan *fuzzy* yang diperoleh sebanyak K^n .



Sumber: Borgi (2018)

Gambar 1. Partisi Fuzzy Grid

Untuk membangkitkan aturan *fuzzy if-then* yang dibagi menjadi beberapa area ruang pola keputusan maka, aturan *fuzzy if-then* sebagai berikut (Sitompul *et al.* 2017):

$$R_{ij}^K : \text{IF } x_i \text{ is } A_i^K \text{ AND } x_j \text{ is } A_j^K \text{ THEN } C; \text{ CF}_{ij}^K$$

dimana:

R_{ij}^K adalah label aturan

A_i^K dan A_j^K adalah premis

K adalah jumlah partisi

C adalah konklusi

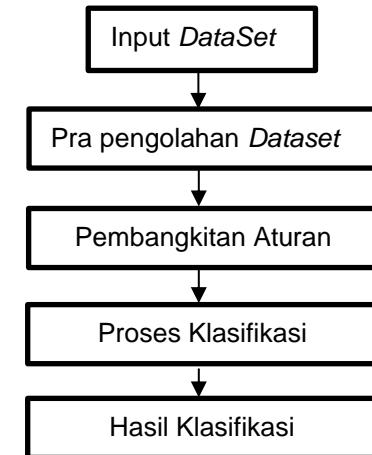
CF adalah bobot aturan

i dan j adalah 1,2, ..., K

2. Metode Penelitian

Metode penelitian adalah kerangka kerja untuk melakukan suatu tindakan, atau suatu kerangka berpikir untuk menyusun gagasan yang terarah dan terkait dengan maksud dan tujuan. Untuk menyelesaikan penelitian ini maka perlu adanya susunan kerangka kerja (*frame work*) yang jelas tahapan-tahapannya seperti

pada gambar 2. Adapun tahapan penelitian sebagai berikut:



Gambar 2. Tahapan Penelitian

Tahapan penelitian sebagai berikut:

1. Input DataSet

Dataset yang di input pada penelitian ini adalah *dataset* yang bersumber dari data sekunder yang dikumpulkan dari Stasiun Meteorologi Maritim Belawan melalui link: https://dataonline.bmkg.go.id/data_iklim.

Periode pengumpulan data dimulai dari data cuaca bulan Januari tahun 2021 s/d bulan April 2021. Data set yang tersedia sebanyak 120 rekam data, dengan 4 atribut yaitu Temperatur, Curah Hujan, Lamanya Penyinaran Matahari, Kecepatan Angin dan Kelembaban Udara.

2. Pra Pengolahan Data

a. *Dataset* tidak lengkap dapat diubah menjadi *dataset* lengkap dengan pendekatan mengganti nilai atribut yang hilang dengan rata-rata atau median dari atribut untuk semua objek yang memiliki kelas keputusan yang sama, seperti yang dikemukakan Han *et al.* (2012).

b. Kategorisasi Data

Sebelum menormalisasi data, dataset yang berupa nilai numerik diubah ke dalam bentuk nilai kategorikal. Berikut kategorisasi untuk setiap atribut

1. Kelembaban Udara (KU)

$$KU = \begin{cases} \text{Kering,} & KU \leq 74 \\ \text{Lembab,} & 74 < KU < 80 \\ \text{Basah,} & KU \geq 80 \end{cases}$$

2. Temperatur (T)

$$T = \begin{cases} \text{Dingin,} & T \leq 26,5 \\ \text{Normal,} & 26,5 < T < 29 \\ \text{Panas,} & T \geq 29 \end{cases}$$

3. Curah Hujan (CH)

$$CH = \begin{cases} \text{Ringan,} & CH \leq 20 \\ \text{Normal,} & 20 < CH < 50 \\ \text{Lebat,} & CH \geq 50 \end{cases}$$

4. Lamanya Penyinaran Matahari (LPM)

$$LPM = \begin{cases} Rendah, & LPM \leq 2.16 \\ Normal, & 2.16 < LPM < 8.28 \\ Tinggi, & LPM \geq 8.28 \end{cases}$$

5. Kecepatan Angin (KA)

$$KA = \begin{cases} Lambat, & KA \leq 8.5 \\ Normal, & 8.5 < KA < 19.5 \\ Kencang, & KA \geq 19.5 \end{cases}$$

c. Normalisasi Data

Normalisasi nilai atribut *dataset* dilakukan dengan cara mengubah interval nilai atribut *dataset* yang sebenarnya ke dalam interval tertentu, dalam kasus ini diubah ke dalam interval [0,1].

$$X_{baru} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2.1)$$

3. Pembangkitan aturan dengan pendekatan metode *fuzzy grid partition*

- a. Menentukan titik pusat dan lebar fungsi keanggotaan untuk mendapatkan aturan yang dibangkitkan.

$$a_i^K = \frac{i-1}{K-1}, i = 1, 2, 3, \dots, K \quad (2.2)$$

$$b^K = \frac{1}{(K-1)} \quad (2.3)$$

dimana, K adalah jumlah partisi. a_i^i adalah titik pusat untuk masing-masing fungsi segitiga dan b^K adalah jarak dari titik pusat ke kaki fungsi segitiga

- b. Penghitungan μ pada K pada masing masing atribut.

$$\mu_i^K(x) = \max\{1 - |x - a_i^K|/b^K, 0\} \quad (2.4)$$

dimana:

$$i = 1, 2, \dots, K$$

- c. Menghitung α -predikat pada masing masing kelas terhadap semua aturan yang berpotensi untuk dibangkitkan

$$\beta_{ct} = \sum_{p \in Ct} \mu_i^K(x_{p1}) \cdot (\mu_j^K(x_{p2})) \quad (2.5)$$

- d. menentukan nilai α -predikat yang terbesar pada masing-masing calon aturan yang sama. Nilai α -predikat yang terbesar dinamakan β .

$$\beta_{cx} = \max\{\beta_{c1}, \beta_{c2}, \dots, \beta_{cM}\} \quad (2.6)$$

- e. Penentuan β dengan membandingkan semua α -predikat pada calon aturan yang sama disetiap kelas.

- f. Penentuan konklusi dari aturan berdasarkan nilai β yang terbesar

- g. yang selanjutnya dinamakan β_{max} . Bila terdapat dua atau lebih β_{max} yang memiliki nilai yang sama, atau nilai β_{max} bernilai 0, maka konklusi tidak dapat ditentukan, *rules* tidak dapat dibangkitkan

- h. Penghitungan bobot aturan (CF) untuk mendapatkan aturan.

$$CF_{ij}^K = \frac{|\beta_{cx} - \beta|}{\sum_{t=1}^M \beta_{ct}}, \quad (2.7)$$

dimana:

$$\beta = \sum_{ct \neq cx} \beta_{ct} / (M - 1) \quad (2.8)$$

- i. Menerapkan teknik representase adaptif grid partition untuk pembentukan struktur *grid*-nya berdasarkan bobot CF. Pembangkitan aturan dilakukan dengan 2 fase. Fase pertama akan dibentuk *grid* sehalus mungkin dengan ketentuan semua kemungkinan aturan dapat dibangkitkan, atau iterasi akan berhenti ketika ada aturan yang memiliki bobot bernilai satu. Sedangkan fase kedua akan dibangkitkan aturan baru dengan bobot yang lebih tinggi untuk meningkatkan akurasi dari aturan tersebut.

4. Proses Klasifikasi

Proses klasifikasi dilakukan untuk mengetahui keakuratan aturan yang telah dibangkitkan dengan cara:

- a. Hitung nilai α masing-masing kelas, dimana α merupakan hasil perkalian antara α -predikat dengan bobot aturan.

$$\alpha_{ct} = \max \{(\mu_i^K(x_{p1}) \cdot (\mu_j^K(x_{p2})))$$

$$CF_{ij}^K | C_{ij}^K = Ct; R_{ij}^K \in S_R\} \quad (2.9)$$

- b. Tentukan nilai α yang terbesar pada masing-masing kelas.

$$\alpha_{cx} = \max \{\alpha_{c1}, \alpha_{c2}, \dots, \alpha_{cM}\} \quad (2.10)$$

- c. kemudian tentukan nilai α_{max} dari masing-masing kelas.

- d. Hitung tingkat keakuratan aturan, *error* aturan dan jumlah *unclass* dari aturan yang dibangkitkan.

akurasi (%)

$$= \frac{\sum \text{klasifikasi yang benar}}{\sum \text{klasifikasi keseluruhan}} \times 100\%$$

error (%)

$$= \frac{\sum \text{klasifikasi yang salah}}{\sum \text{klasifikasi keseluruhan}} \times 100\%$$

Unclass

$$= \sum \text{klasifikasi keseluruhan} - (\sum \text{klasifikasi benar} + \sum \text{klasifikasi salah})$$

5. Hasil Klasifikasi

Hasil klasifikasi diperoleh pada saat proses *learning* akan berhenti setelah batas minimal keakuratan yang diinginkan terhadap data *learning* telah dipenuhi.

3. Hasil dan Pembahasan

Pembangkitan aturan fuzzy menggunakan sampel data sebanyak 6 data yang terdapat pada 2 data di kelas kering, 2 data di kelas lembab dan 2 data di kelas basah. Sampel data seperti pada tabel 2 berikut.

Tabel 2. Enam Sampel Dataset

No. Rekam Data	Atribut				Kelas
	T	CH	LPM	KA	
55	0.67	0.00	0.70	0.50	Kering
91	1.00	0.00	0.34	1.00	Kering
37	0.64	0.01	0.41	1.00	Lembab
46	0.73	0.00	0.83	1.00	Lembab
86	0.58	0.03	0.66	0.50	Basah
98	0.60	0.20	0.03	1.00	Basah

3.1 Pembangkitan Aturan

Pembangkitan aturan dengan pendekatan metode *fuzzy grid partition* sebagai berikut:

- a. Perhitungan titik pusat dan lebar fungsi

Dengan mengikuti sertakan semua variabel linguistik pada K= 2, sebagai contoh:

a_1^2 dan b^2 dihitung sebagai berikut:

$$a_1^2 = (1 - 1)/(2 - 1) = 0/1 = 0$$

$$b^2 = 1/(2 - 1) = 1/1 = 1$$

dengan cara yang sama, akan diperoleh a_2^2 dan b^2

$$a_2^2 = (2 - 1)/(2 - 1) = 1/1 = 1$$

$$b^2 = 1/(2 - 1) = 1/1 = 1$$

Maka terdapat 16 aturan yang berpotensi untuk dibangkitkan, yang diperoleh dari Kd, dimana d adalah jumlah atribut dataset. Calon aturan tersebut pada K= 2 adalah sebagai berikut:

1. $R_{1111}^2: IF T \text{ is } A_1^2 \text{ AND } CH \text{ is } A_1^2 \text{ AND } LPM \text{ is } A_1^2 \text{ AND } KA \text{ is } A_1^2 \text{ THEN } KU \text{ is } C$
2. $R_{1112}^2: IF T \text{ is } A_1^2 \text{ AND } CH \text{ is } A_1^2 \text{ AND } LPM \text{ is } A_1^2 \text{ AND } KA \text{ is } A_2^2 \text{ THEN } KU \text{ is } C$
3. $R_{1121}^2: IF T \text{ is } A_1^2 \text{ AND } CH \text{ is } A_2^2 \text{ AND } LPM \text{ is } A_2^2 \text{ AND } KA \text{ is } A_1^2 \text{ THEN } KU \text{ is } C$
4. $R_{1122}^2: IF T \text{ is } A_1^2 \text{ AND } CH \text{ is } A_2^2 \text{ AND } LPM \text{ is } A_2^2 \text{ AND } KA \text{ is } A_2^2 \text{ THEN } KU \text{ is } C$
5. $R_{1211}^2: IF T \text{ is } A_1^2 \text{ AND } CH \text{ is } A_2^2 \text{ AND } LPM \text{ is } A_1^2 \text{ AND } KA \text{ is } A_1^2 \text{ THEN } KU \text{ is } C$
6. $R_{1212}^2: IF T \text{ is } A_1^2 \text{ AND } CH \text{ is } A_2^2 \text{ AND } LPM \text{ is } A_1^2 \text{ AND } KA \text{ is } A_2^2 \text{ THEN } KU \text{ is } C$
7. $R_{1221}^2: IF T \text{ is } A_1^2 \text{ AND } CH \text{ is } A_2^2 \text{ AND } LPM \text{ is } A_2^2 \text{ AND } KA \text{ is } A_1^2 \text{ THEN } KU \text{ is } C$
8. $R_{1222}^2: IF T \text{ is } A_1^2 \text{ AND } CH \text{ is } A_2^2 \text{ AND } LPM \text{ is } A_2^2 \text{ AND } KA \text{ is } A_2^2 \text{ THEN } KU \text{ is } C$
9. $R_{2111}^2: IF T \text{ is } A_2^2 \text{ AND } CH \text{ is } A_1^2 \text{ AND } LPM \text{ is } A_1^2 \text{ AND } KA \text{ is } A_1^2 \text{ THEN } KU \text{ is } C$
10. $R_{2112}^2: IF T \text{ is } A_2^2 \text{ AND } CH \text{ is } A_1^2 \text{ AND } LPM \text{ is } A_1^2 \text{ AND } KA \text{ is } A_2^2 \text{ THEN } KU \text{ is } C$
11. $R_{2121}^2: IF T \text{ is } A_2^2 \text{ AND } CH \text{ is } A_1^2 \text{ AND } LPM \text{ is } A_2^2 \text{ AND } KA \text{ is } A_1^2 \text{ THEN } KU \text{ is } C$
12. $R_{2122}^2: IF T \text{ is } A_2^2 \text{ AND } CH \text{ is } A_1^2 \text{ AND } LPM \text{ is } A_2^2 \text{ AND } KA \text{ is } A_2^2 \text{ THEN } KU \text{ is } C$
13. $R_{2211}^2: IF T \text{ is } A_2^2 \text{ AND } CH \text{ is } A_2^2 \text{ AND } LPM \text{ is } A_1^2 \text{ AND } KA \text{ is } A_1^2 \text{ THEN } KU \text{ is } C$
14. $R_{2212}^2: IF T \text{ is } A_2^2 \text{ AND } CH \text{ is } A_2^2 \text{ AND } LPM \text{ is } A_1^2$

- AND KA is A_2^2 THEN KU is C
15. $R_{2221}^2: IF T \text{ is } A_2^2 \text{ AND } CH \text{ is } A_2^2 \text{ AND } LPM \text{ is } A_2^2 \text{ AND } KA \text{ is } A_1^2 \text{ THEN } KU \text{ is } C$
 16. $R_{2222}^2: IF T \text{ is } A_2^2 \text{ AND } CH \text{ is } A_2^2 \text{ AND } LPM \text{ is } A_2^2 \text{ AND } KA \text{ is } A_2^2 \text{ THEN } KU \text{ is } C$

- b. Derajat Keanggotaan pada K = 2

Perhitungan μ data temperatur ke 55 pada K = 2 sebagai berikut:

$$\mu_1^2(0.67) = \frac{1.00 - 0.67}{1.00 - 0.00} = 0.33$$

Perhitungan μ untuk semua rekam data dapat dilihat pada tabel 3.

Tabel 3. Derajat keanggotaan pada K=2

Class (No. rekam data)	Label μ	Derajat Keanggotaan			
		T	CH	LPM	KA
Kering (55)	μ_1^2	0.33	1	0.3	0.5
	μ_2^2	0.67	0	0.7	0.5
Kering (91)	μ_1^2	0	1	0.66	0
	μ_2^2	1	0	0.34	1
Lembab (37)	μ_1^2	0.36	0.99	0.59	0
	μ_2^2	0.64	0.01	0.41	1
Lembab (46)	μ_1^2	0.27	1	0.17	0
	μ_2^2	0.73	0	0.83	1
Basah (86)	μ_1^2	0.42	0.97	0.34	0.5
	μ_2^2	0.58	0.03	0.66	0.5
Basah (98)	μ_1^2	0.4	0.8	0.97	0
	μ_2^2	0.6	0.2	0.03	1

- c. Nilai $\alpha - predikat$

Perhitungan α -predikat dilakukan pada dataset dengan rekam data nomor 55 pada label aturan R_{1112}^2 dengan persamaan:

$$R_{1112}^2 = \mu_1^2(T) \cdot \mu_1^2(CH) \cdot \mu_1^2(LPM) \cdot \mu_2^2(KA)$$

$$R_{1112}^2 = (0.33)(1)(0.3)(0.5)$$

$$R_{1112}^2 = 0.0495$$

dengan cara yang sama, semua α -predikat dihitung untuk kelas kering, lembab dan basah dapat dilihat pada tabel 4, tabel 5 dan tabel 6.

Tabel 4. $\alpha - predikat$ Kelas Kering

R_{ijk}^K	No. Data	T is A_i^K	CH is A_j^K	LPM is A_k^K	KA is A_l^K	α
R_{111}^2	55	0.33	1	0.3	0.5	0.0495
	91	0	1	0.66	0	0
R_{111}^2	55	0.33	1	0.3	0.5	0.0495
	91	0	1	0.66	1	0
.
.
.
R_{222}^2	55	0.67	0	0.7	0.5	0
	91	1	0	0.34	1	0

Tabel 5. α – predikat Kelas Lembab

R_{ijkl}^K	No. Data	T is A_i^K	CH is A_j^K	LPM is A_k^K	KA is A_l^K	α
R_{1111}^2	37	0.36	0.99	0.59	0	0
	46	0.73	0	0.83	1	0
R_{1112}^2	37	0.36	0.99	0.59	1	0.2103
	46	0.27	1	0.17	1	0.0459
.						
.						
R_{2222}^2	37	0.64	0.01	0.41	1	0.0026
	46	0.73	0	0.83	1	0

Tabel 8. Nilai β Kelas Lembab

R_{ijkl}^K	α – predikat Pada Rekam Data		β
	37	46	
R_{1111}^2	0	0	0
R_{1112}^2	0.21028	0.0459	0.21028
.			
.			
R_{2222}^2	0.00262	0	0.00262

Tabel 6. α – predikat Kelas Basah

R_{ijkl}^K	No. Data	T is A_i^K	CH is A_j^K	LPM is A_k^K	KA is A_l^K	α
R_{1111}^2	86	0.42	0.97	0.34	0.5	0.069
	98	0.4	0.8	0.97	0	0
R_{1112}^2	86	0.42	0.97	0.34	0.5	0.069
	98	0.4	0.8	0.97	1	0.310
.						
.						
R_{2222}^2	86	0.58	0.03	0.66	0.5	0.006
	98	0.6	0.2	0.03	1	0.004

d. Penentuan β

Perhitungan nilai β dilakukan dengan membandingkan semua α -predikat pada calon aturan yang sama. Sebagai simulasi, calon aturan R_{1112}^2 pada kelas kering, dilakukan dengan membandingkan nilai α -predikat pada dataset dengan rekam data nomor 55 dan rekam data nomor 91, seperti pada persamaan berikut:

$$\beta = \max \{0.0495; 0\} = 0.0495$$

$$\beta = 0.0495$$

dengan cara yang sama, nilai β pada masing-masing kelas kering, kelas lembab dan kelas basah dapat dilihat pada tabel 7, tabel 8 dan tabel 9.

Tabel 7. Nilai β Kelas Kering

R_{ijkl}^K	α – predikat Pada Rekam Data		β
	55	91	
R_{1111}^2	0.0495	0	0.0495
R_{1112}^2	0.0495	0	0.0495
.			
.			
R_{2222}^2	0	0	0

Tabel 9. Nilai β Kelas Basah

R_{ijkl}^K	α – predikat Pada Rekam Data		β
	86	98	
R_{1111}^2	0.06926	0	0.06926
R_{1112}^2	0.06926	0.3104	0.3104
.			
.			
R_{2222}^2	0.00574	0.0036	0.00574

e. Penentuan Konklusi

Konklusi-konklusi dari aturan ditentukan berdasarkan nilai β yang terbesar diantara β kering, β lembab, dan β basah, yang selanjutnya dinamakan β max. Bila terdapat dua atau lebih β max yang memiliki nilai yang sama, atau nilai β max bernilai 0, maka konklusi tidak dapat ditentukan, aturan tidak dapat dibangkitkan. Nilai β max untuk R_{1111}^2 ditentukan sebagai berikut:

$$\begin{aligned}\beta_{\text{max}} &= \max \{\beta_{\text{kering}}, \beta_{\text{lembab}}, \text{ dan } \beta_{\text{basah}}\} \\ &= \max \{0.0495, 0, 0.06926\} \\ &= 0.06926\end{aligned}$$

Rekapitulasi hasil perhitungan nilai β max dapat dilihat pada tabel 10.

Tabel 10. Rekapitulasi β max

R_{ijkl}^K	$\sum \{\mu_i^K(T), \mu_j^K(CH), \mu_k^K(LPM), \mu_l^K(KA)\}$	β_{max}	Konklusi	
	β_{kering}	β_{lembab}	β_{basah}	
R_{1111}^2	0.0495	0	0.06926	Basah
R_{1112}^2	0.0495	0.21028	0.3104	Basah
R_{1211}^2	0.1155	0	0.13444	Basah
R_{1212}^2	0.1155	0.2241	0.13444	Lembab
R_{1211}^2	0	0	0.00214	Basah
R_{1212}^2	0	0.00212	0.0776	Basah
R_{1221}^2	0	0	0.00416	Basah
R_{1222}^2	0	0.00148	0.00416	Basah
R_{2111}^2	0.1005	0	0.09564	Kering
R_{2112}^2	0.66	0.37382	0.4656	0.66
R_{2121}^2	0.2345	0	0.18566	Kering
R_{2122}^2	0.34	0.6059	0.18566	0.6059
R_{2211}^2	0	0	0.00296	Basah
R_{2212}^2	0	0.00378	0.1164	Basah
R_{2221}^2	0	0	0.00574	Basah
R_{2222}^2	0	0.00262	0.00574	Basah

f. Perhitungan Bobot Aturan

Sebagai simulasi, perhitungan bobot *aturan* pada R_{1111}^2 dilakukan sebagai berikut:

$$CF_{1111}^2 = \frac{|0.069258 - (0.0495 + 0)/2|}{0.069258 + 0.0495 + 0}$$

$$CF_{1111}^2 = \frac{0.0445}{0.1188}$$

$$CF_{1111}^2 = 0.3748$$

Rekapitulasi untuk perhitungan bobot aturan CF_{ijkl}^K pada $K = 2$ dapat dilihat pada tabel 11.

Tabel 11. Bobot Aturan

No.	R_{ijkl}^K	Konklusi	CF_{ijkl}^K
1	R_{1111}^2	Basah	0.374779
2	R_{1112}^2	Basah	0.31659
3	R_{1121}^2	Basah	0.306839
4	R_{1122}^2	Lembab	0.209114
5	R_{1211}^2	Basah	1
6	R_{1212}^2	Basah	0.960037
7	R_{1221}^2	Basah	1
8	R_{1222}^2	Basah	0.607029
9	R_{2111}^2	Kering	0.512384
10	R_{2112}^2	Kering	0.315513
11	R_{2121}^2	Kering	0.558123
12	R_{2122}^2	Lembab	0.303185
13	R_{2211}^2	Basah	1
14	R_{2212}^2	Basah	0.952869
15	R_{2221}^2	Basah	1
16	R_{2222}^2	Basah	0.529524

Dengan cara yang sama seperti pada $K=2$, rekapitulasi pembangkitan aturan fuzzy pada $K=3$ dapat dilihat pada tabel 12, dan rekapitulasi aturan fuzzy yang hanya dapat dibangkitkan pada $K=4$ dapat dilihat pada tabel 13.

Tabel 12. Pembangkitan Aturan Fuzzy K = 3

No.	R_{ijkl}^K	Konklusi	CF_{ijkl}^K
1	R_{2113}^3	Basah	0.670513
2	R_{2122}^3	Basah	0.363295
3	R_{2123}^3	Lembab	0.928876
4	R_{2132}^3	Kering	0.255481
5	R_{2133}^3	Lembab	1
6	R_{2213}^3	Basah	0.987185
7	R_{2222}^3	Basah	1
8	R_{2223}^3	Basah	0.428793
9	R_{2232}^3	Basah	1
10	R_{2322}^3	Basah	1
11	R_{2332}^3	Basah	1
12	R_{3113}^3	Kering	0.495454
13	R_{3122}^3	Kering	0.499112

No.	R_{ijkl}^K	Konklusi	CF_{ijkl}^K
14	R_{3123}^3	Kering	0.618166
15	R_{3132}^3	Kering	0.607925
16	R_{3133}^3	Lembab	1
17	R_{3213}^3	Basah	0.98016
18	R_{3222}^3	Basah	1
19	R_{3223}^3	Basah	0.266661
20	R_{3232}^3	Basah	1
21	R_{3322}^3	Basah	1
22	R_{3332}^3	Basah	1

Tabel 13. Pembangkitan Aturan Fuzzy K = 4

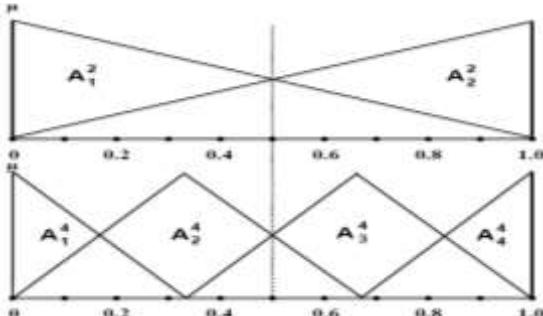
No.	R_{ijkl}^K	Konklusi	CF_{ijkl}^K
1	R_{2112}^4	Basah	1
2	R_{2113}^4	Basah	1
3	R_{2114}^4	Basah	1
4	R_{2122}^4	Basah	1
5	R_{2123}^4	Basah	1
6	R_{2124}^4	Lembab	0.808615
7	R_{2132}^4	Basah	1
8	R_{2133}^4	Basah	1
9	R_{2134}^4	Lembab	1
10	R_{2142}^4	Basah	1
11	R_{2143}^4	Basah	1
12	R_{2212}^4	Basah	1
13	R_{2213}^4	Basah	1
14	R_{2114}^4	Basah	1
15	R_{2222}^4	Basah	1
16	R_{2223}^4	Basah	1
17	R_{2224}^4	Basah	0.817073
18	R_{2232}^4	Basah	1
19	R_{2233}^4	Basah	1
20	R_{2234}^4	Lembab	1
21	R_{2242}^4	Basah	1
22	R_{2243}^4	Basah	1
23	R_{3112}^4	Basah	1
24	R_{3113}^4	Basah	1
25	R_{3114}^4	Basah	1
26	R_{3122}^4	Basah	1
27	R_{3123}^4	Basah	1
28	R_{3124}^4	Lembab	0.938905
29	R_{3132}^4	Kering	0.3314
30	R_{3133}^4	Kering	0.3314
31	R_{3134}^4	Lembab	1
32	R_{3142}^4	Basah	0.786671
33	R_{3143}^4	Basah	0.786671
34	R_{3144}^4	Lembab	1
35	R_{3212}^4	Basah	1
36	R_{3213}^4	Basah	1
37	R_{3214}^4	Basah	1
38	R_{3222}^4	Basah	1
39	R_{3223}^4	Basah	1
40	R_{3224}^4	Basah	0.514614
41	R_{3232}^4	Basah	1
42	R_{3233}^4	Basah	1

No.	R_{ijkl}^K	Konklusi	CF_{ijkl}^K
43	R_{3242}^4	Basah	1
44	R_{3243}^4	Basah	1
45	R_{4124}^4	Kering	0.376341
46	R_{4132}^4	Kering	1
47	R_{4133}^4	Kering	1
48	R_{4134}^4	Lembab	0.819003
49	R_{4142}^4	Kering	1
50	R_{4143}^4	Kering	1
51	R_{4144}^4	Lembab	1
52	R_{4224}^4	Lembab	1
53	R_{4234}^4	Lembab	1

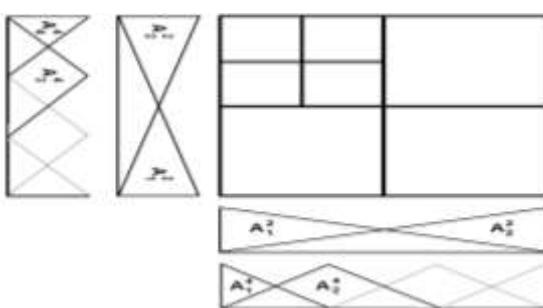
g. Grid Partition Teradaptasi

Aturan yang dapat dibangkitkan dengan rekam data seperti pada table 5.11 untuk fase 1 adalah $SR = \{R_{1111}^2, R_{1112}^2, R_{1121}^2, R_{1122}^2, R_{1211}^2, R_{1212}^2, R_{1221}^2, R_{1222}^2, R_{2111}^2, R_{2112}^2, R_{2121}^2, R_{2122}^2, R_{2211}^2, R_{2212}^2, R_{2221}^2, R_{2222}^2\}$.

Berdasarkan table 11, R_{1122}^2 dipilih karena memiliki bobot aturan yang terendah. Nilai K dan K pada fase 1 diinisialisasi dari R_{1122}^2 , kemudian $K=2K$ sehingga nilai $K=4$. Pada fase kedua ini, aturan yang berpotensi untuk dibangkitkan dari R_{1122}^2 dapat dijelaskan pada gambar 3 dan gambar 4.



Gambar 3. Pelebelan A_i^K pada $K = 2$ dan $K = 2K$ Pada Satu Atribut



Gambar 4. Pelebelan R_{ij}^K pada $K = 2$ dan $K = 2K$ Pada Dua Atribut

Sesuai gambar 3, label A_i^K pada $K=2$ adalah A_2^2 , setelah $K = 2K$, akan terbentuk A_4^3 dan A_4^4 . Dengan menggunakan prinsip yang sama

seperti pada gambar 4, aturan yang berpotensi untuk dibangkitkan dari R_{1122}^2 pada $K=2K$ adalah seperti pada tabel 14.

Tabel 14. Pembangkitan Aturan Fuzzy dari R_{1122}^2 Pada $K=2K$

No.	R_{ijkl}^K	β_{Max}	Konklusi	CF_{ijkl}^K
1	R_{2133}^4	0.106854	Basah	1
2	R_{2134}^4	0.014247	Lembab	1
3	R_{2143}^4	0.106854	Basah	1
4	R_{2233}^4	0.010685	Basah	1
5	R_{2234}^4	0.000445	Lembab	1
6	R_{2243}^4	0.010685	Basah	1

Berdasarkan langkah-langkah tersebut, maka aturan yang dibangkitkan adalah $SR = \{R_{1111}^2, R_{1112}^2, R_{1121}^2, R_{1122}^2, R_{1211}^2, R_{1212}^2, R_{1221}^2, R_{1222}^2, R_{2111}^2, R_{2112}^2, R_{2121}^2, R_{2122}^2, R_{2211}^2, R_{2212}^2, R_{2221}^2, R_{2222}^2, R_{2133}^4, R_{2134}^4, R_{2143}^4, R_{2233}^4, R_{2234}^4, R_{2243}^4\}$. SR merupakan himpunan aturan yang dapat dibangkitkan yaitu sejumlah 22 aturan.

3.2 Proses Klasifikasi

Proses klasifikasi dilakukan untuk mengetahui keakuratan aturan yang telah dibangkitkan. Simulasi hasil perhitungan untuk rekam data no 55 pada kelas kering, lembab dan basah dapat dilihat pada table 15, table 16 dan table 17 berikut:

Tabel 15. Nilai α_{kering} Berdasarkan Aturan Yang Dibangkitkan

R_{ijkl}^K	T is A_i^K	CH is A_j^K	LPM is A_k^K	KA is A_l^K	CF_{ijkl}^K	α
R_{2111}^2	0.67	1	0.3	0.5	0.5124	0.051496
R_{2112}^2	0.67	1	0.3	0.5	0.3155	0.031708
R_{2121}^2	0.67	1	0.7	0.5	0.5581	0.130874
α_{max}						0.130874

Tabel 16. Nilai α_{lembab} Berdasarkan Aturan Yang Dibangkitkan

R_{ijkl}^K	T is A_i^K	CH is A_j^K	LPM is A_k^K	KA is A_l^K	CF_{ijkl}^K	α
R_{1122}^2	0.33	1	0.7	0.5	0.2091	0.024153
R_{2122}^2	0.67	1	0.7	0.5	0.3032	0.071097
R_{2134}^4	0.06	0.97	0.24	1	1	0.014247
R_{2234}^4	0.06	0.03	0.24	1	1	0.000445
α_{max}						0.071097

Tabel 17. Nilai α_{basah} Berdasarkan Aturan Yang Dibangkitkan

R_{ijkl}^K	T is A_i^K	CH is A_j^K	LPM is A_k^K	KA is A_l^K	CF_{ijkl}^K	α
R_{1111}^2	0.33	1	0.3	0.5	0.3748	0.0185
R_{1112}^2	0.33	1	0.3	0.5	0.3166	0.0156

R_{ijkl}^K	T is A_i^K	CH is A_j^K	LPM is A_k^K	KA is A_l^K	CF_{ijkl}^K	α
R_{1121}^2	0.33	1	0.7	0.5	0.3068	0.0354
R_{1211}^2	0.33	0	0.3	0.5		0
R_{1212}^2	0.33	0	0.3	0.5	0.9600	0
R_{1221}^2	0.33	0	0.7	0.5	1	0
R_{1222}^2	0.33	0	0.7	0.5	0.6070	0
R_{2211}^2	0.67	0	0.3	0.5	1	0
R_{2212}^2	0.67	0	0.3	0.5	0.9528	0
R_{2221}^2	0.67	0	0.7	0.5	1	0
R_{2222}^2	0.67	0	0.7	0.5	0.5295	0
R_{2133}^4	0.06	0.97	0	0	1	0
R_{2143}^4	0.06	0.03	0.24	0	1	0
R_{2233}^4	0.06	0.03	0	0	1	0
R_{2243}^4	0.33	1	0.3	0.5	0.3748	0.0185
α_{max}						0.0354

3.3 Hasil Klasifikasi

Berdasarkan proses klasifikasi yg telah dilakukan, diperoleh nilai α yang terbesar (α_{max}) pada setiap rekam data. Pada nomor rekam data 55 nilai α_{max} terletak pada kelas kering. Hasil klasifikasi menunjukkan benar yang berarti aturan yang dibangkitkan dapat mengklasifikasi data rekam nomor 55. Dengan cara yang sama, hasil klasifikasi yang dilakukan pada data dengan rekam data nomor 55, nomor 91, nomor 37, nomor 46, nomor 86, dan nomor 98, dengan menggunakan aturan yang dibangkitkan dapat dilihat pada tabel 18.

Tabel 18. Hasil Klasifikasi

No. Rekam Data	α_{max}	Kelas	Hasil Klasifikasi
55	0.1309	Kering	Benar
91	0.2082	Kering	Benar
37	0.1179	Kering	Salah
46	0.1837	Lembab	Benar
86	0.1069	Basah	Benar
98	0.1469	Kering	Salah

Berdasarkan tabel 18, tingkat keakuratan rules yang dibangkitkan adalah sebagai berikut:

$$\text{akurasi (\%)} = \frac{\sum \text{klasifikasi yang benar}}{\sum \text{klasifikasi keseluruhan}} \times 100$$

$$\text{akurasi (\%)} = \frac{4}{6} \times 100\%$$

$$\text{akurasi (\%)} = 66,67\%$$

sedangkan tingkat error (kesalahan) dari rules yang telah dibangkitkan sebagai berikut:

$$\text{error (\%)} = \frac{\sum \text{klasifikasi yang salah}}{\sum \text{klasifikasi keseluruhan}} \times 100\%$$

$$\text{error (\%)} = \frac{2}{6} \times 100\%$$

$$\text{error (\%)} = 33,33\%$$

Jumlah unclass (data yang tidak dapat diklasifikasi) sebagai berikut:

$$\text{unclass} = \Sigma \text{klasifikasi keseluruhan} - (\Sigma \text{klasifikasi benar} + \Sigma \text{klasifikasi salah})$$

$$\text{unclass} = 6 - (4 + 2)$$

$$\text{unclass} = 0$$

4. Kesimpulan

Metode *fuzzy grid partition* dapat mengklasifikasikan data tidak lengkap. Data tidak lengkap terlebih dahulu di proses dengan pendekatan mengganti nilai atribut yang hilang dengan rata-rata atau median dari atribut untuk semua objek yang memiliki kelas keputusan yang sama. Aturan *fuzzy* yang dibangkitkan dengan metode *grid partition* teradaptasi menghasilkan 22 himpunan aturan. Dari pengujian enam rekam data, akurasi aturan yang dibangkitkan menghasilkan 66,67 % hasil klasifikasi dinyatakan benar, 33,33% hasil klasifikasi dinyatakan error dan nilai *unclass* adalah 0 yang berarti semua data dapat diklasifikasi.

Referensi

- Agarwal, S. (2014). Data mining: Data mining concepts and techniques. In *Proceedings - 2013 International Conference on Machine Intelligence Research and Advancement, ICMIRA 2013*. <https://doi.org/10.1109/ICMIRA.2013.45>
- Borgi, A. (2018). Attributes regrouping in Fuzzy Rule Based Classification Systems: an intra-classes approach. *2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, 1–7.
- Chen, T., Shen, Q., Su, P., & Shang, C. (2016). Fuzzy rule weight modification with particle swarm optimisation. *Soft Computing*, 20(8), 2923–2937. <https://doi.org/10.1007/s00500-015-1922-z>
- Dahal, K., Almejalli, K., Hossain, M. A., & Chen, W. (2015). GA-based learning for rule identification in fuzzy neural networks. *Applied Soft Computing Journal*, 35, 605–617. <https://doi.org/10.1016/j.asoc.2015.06.046>
- Elcano, M., Galar, M., Sanz, J., & Bustince, H. (2016). Fuzzy Rule-Based Classification Systems for multi-class problems using binary decomposition strategies: On the influence of n-dimensional overlap functions in the Fuzzy Reasoning Method. *Information Sciences*, 332, 94–114. <https://doi.org/10.1016/j.ins.2015.11.006>
- Field, D., & Zhao, L. (2018). Feature Selection Method based on Grid Partition + IEEE. 52–

- 57.
- Hartono. (2016). Optimization of Tsukamoto Fuzzy Inference System using Fuzzy Grid Partition. *IJCSN International Journal of Computer Science and Network*, 5(5), 2277–5420. Retrieved from www.IJCSN.org
- Liu, X., Feng, X., & Pedrycz, W. (2013). Extraction of fuzzy rules from fuzzy decision trees: An axiomatic fuzzy sets (AFS) approach. *Data and Knowledge Engineering*, 84, 1–25. <https://doi.org/10.1016/j.datak.2012.12.001>
- Mao, L., Chen, Q., & Sun, J. (2020). *Construction and Optimization of Fuzzy Rule-Based Classifier with a Swarm Intelligent Algorithm*. 2020. <https://doi.org/10.1155/2020/9319364>
- Marbun, M., Ramdhani, W., Priyanto, D., & Zarlis, M. (2019). *Philosophy of Fuzzy Logic as Fundamental of Decision Making Based On Rule Philosophy of Fuzzy Logic as Fundamental of Decision Making Based On Rule*. <https://doi.org/10.1088/1742-6596/1230/1/012021>
- Sadiq, A. T., Duaimi, M. G., & Shaker, S. A. (2012). Data missing solution using rough set theory and swarm intelligence. *Proceedings - 2012 International Conference on Advanced Computer Science Applications and Technologies, ACSAT 2012*, 3, 173–180. <https://doi.org/10.1109/ACSAT.2012.29>
- Sadouki, L., & Haddad, B. (2016). *Adaptive Neuro-Fuzzy Inference System for Echoes Classification in Radar Images*. 4(Visigrapp), 159–166. <https://doi.org/10.5220/0005717401590166>
- Sinambela, M. S., Rosely, E., & Mayadewi, R. P. (2016). *THE DECISION SUPPORT SYSTEM FOR SUBJECT SPECIALIZATION STUDENTS IN*. 2(3), 858–866.
- Sitompul, opim salim; Nababan, Erna Budhiarti; Alim, Z. (2017). *Adaptive Dsitibuted Grid-Partition in Generating Fuzzy Rules*. 119–124.
- Takahashi, Y., Nojima, Y., & Ishibuchi, H. (2015). Rotation effects of objective functions in parallel distributed multiobjective fuzzy genetics-based machine learning. *2015 10th Asian Control Conference: Emerging Control Techniques for a Sustainable World, ASCC 2015*, (C), 1–6. <https://doi.org/10.1109/ASCC.2015.7244890>