PREDIKSI CURAH HUJAN DENGAN MENGGUNAKAN FUZZY FORECASTING BERBASIS AUTOMATIC CLUSTERING DAN AXIOMATIC FUZZY SET CLASSIFICATION

RAINFALL PREDICTION USING FUZZY FORECASTING BASED ON AUTOMATIC CLUSTERING AND AXIOMATIC FUZZY SET CLASSIFICATION

Imam Aburizal Albana

Prodi SI Ilmu Komputasi, Fakultas Informatika, Universitas Telkom

aburizalalbana@gmail.com

Abstrak

Dalam penelitian ini dilakukan prediksi curah hujan di provinsi Kalimantan Selatan. Pemodelan matematika yang digunakan untuk memprediksi curah hujan diwaktu mendatang yaitu *fuzzy forecasting*. Dalam memprediksi, *fuzzy forecasting* memiliki empat langkah yang akan menghasilkan nilai prediksi. Langkah-langkah dalam *fuzzy forecasting* yaitu *automatic clustering*, Rancang tren fuzzy berlabel *training dataset, axiomatic fuzzy set classification*, dan *forecasting*. Dengan menggunakan langkah-langkah *fuzzy forecasting* tersebut metode ini menghasilkan nilai RMSE sebesar 52.55 dan MAPE sebesar 42.46.

Kata kunci: ***prediksi curah hujan, Automatic Clustering, Axiomatic Fuzzy Set classification.***

Abstract

This study is predicting the rainfall in South Kalimantan province. Fuzzy Forecasting is applied to predict rainfall in the next time. In the way to predicting rainfall, Fuzzy Forecasting has four steps that will give predict value. First step is, Automatic Clustering, second step is Construct Fuzzy Trend Labeled Training Dataset, third step is Axiomatic Fuzzy Set Classification, and the final step is Forecasting. With these steps, this model has 52.55 as RMSE value and 42.46 as MAPE value.

Keyword*: *rainfall predict, Automatic Clustering, Axiomatic Fuzzy Set classification.

1. Pendahuluan

Hujan memiliki peranan penting bagi manusia dan lingkungannya, akan tetapi bisa menyebabkan bahaya pada suatu wilayah jika jumlah curah hujannya terlalu besar. Tjasyono (2004) mendefinisikan curah hujan merupakan jumlah air yang jatuh di permukaan tanah datar selama kurun waktu tertentu yang diukur dengan satuan tinggi millimeter diatas permukaan horizontal, tidak menguap, tidak meresap, dan tidak mengalir. Curah hujan satu millimeter artinya dalam suatu tempat yang seluas satu meter persegi pada tempat yang datar tertampung air setinggi satu millimeter atau tertampung air sebanyak satu liter [3].

Manfaat dari prediksi curah hujan yaitu dapat meramalkan curah hujan dalam kurun waktu (perhari, perbulan, atau pertahun) mendatang, sehingga manusia bisa mempersiapkan diri apa yang harus dilakukan untuk menyambut curah hujan diwaktu mendatang. Manfaat memprediksi curah hujan sangat dirasakan dalam bidang pertanian dalam melakukan kegiatan-kegiatan, seperti perancangan pola tanam, penentuan waktu tanam, pengairan, pemupukan, sampai pendistribusian hasil panen.

Logika Fuzzy merupakan salah satu komponen pembentuk *soft computing*. Logika fuzzy pertama kali diperkenalkan oleh Prof. Lotfi A. Zadeh pada tahun 1994. Teori himpunan fuzzy merupakan dasar dari logika fuzzy. Pada teori himpunan fuzzy, peranan derajat keanggotaan sebagai penentu keberadaan elemen dalam suatu himpunan sangatlah penting. Nilai keanggotaan atau derajat keanggotaan atau *membership function* menjadi ciri utama dari penalaran dengan logika fuzzy tersebut [2].

Algoritma *Automatic Clustering* pertama kali diperkenalkan oleh M Kamel dan B Hadfield (1990) [16]. Kemudian pada tahun 2012 Weina Wang dan Xiaodong Liu menggunakan *Fuzzy Forecasting* berbasis *Automatic Clustering* dan *Axiomatic Fuzzy Set classification* untuk memprediksi nilai Taiwan Capitalization Weighted Stock Index (TAIEX). Pada makalah ini, metode *Fuzzy Forecasting* berbasis *Automatic Clustering* dan *Axiomatic Fuzzy Set* (AFS) *classification* digunakan untuk memprediksi curah hujan [1].

2. Metode

Perancangan sistem yang digunakan dalam memprediksi curah hujan di provinsi Kalimantan Selatan menggunakan metode *Fuzzy Forecasting* berbasis *Automatic Clustering* dan *Axiomatic Fuzzy Set* (AFS) *classification*. Pada tahap pertama dilakukan pengumpulan data sampel yang terdiri dari data *training* dan data uji. Data *training* dimulai dari bulan Januari tahun 2003 sampai dengan bulan Desember tahun 2010, dan data uji dimulai dari bulan Januari tahun 2011 sampai dengan bulan Desember tahun 2013. Pengumpulan data *training* dilakukan untuk membentuk pola prediksi curah hujan, sedangkan pengumpulan data uji dilakukan untuk mengetahui kinerja metode ini dalam memprediksi curah hujan

Tabel 2.1 Data Curah Hujan provinsi Kalimantan Selatan 2003-2013

Bulan	Tahun										
	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Januari	395.2	626.1	286.9	362.6	240.6	221.7	384.0	324.3	418.9	223.7	355.2
Februari	547.5	375.1	271.8	345.9	239.0	242.0	148.0	320.6	211.8	258.4	414.6
Maret	150.0	303.0	332.5	294.8	482.7	419.4	212.0	285.1	337.1	313.0	308.3
April	197.1	126.9	129.5	219.3	325.6	228.5	279.0	243.0	250.8	319.1	305.5
Mei	50.3	228.0	230.4	72.5	235.3	140.2	237.0	171.0	210.5	149.1	346.5
Juni	115.8	80.0	49.7	188.2	170.9	170.1	22.0	365.7	83.1	58.4	140.7
Juli	47.0	90.1	18.8	24.7	229.3	225.1	73.0	171.7	21.3	193.5	125.7
Agustus	41.6	0.0	49.3	4.6	54.8	157.6	25.0	240.4	26.8	70.3	81.5
September	110.0	32.6	36.1	2.9	30.1	127.5	21.0	338.2	77.3	58.2	33.6
Oktober	171.1	51.7	176.5	16.5	62.4	208.8	189.0	256.5	133.5	157.0	106.0
November	263.4	289.6	203.2	115.6	1641.9	300.2	292.0	317.5	276.4	297.8	439.1
Desember	680.0	415.0	284.4	408.4	255.2	427.2	287.0	354.7	856.4	409.8	349.4

2.1. *Automatic Clustering*

Dalam tahap ini, algoritma *Automatic Clustering* digunakan untuk mengubah data curah hujan kedalam *clustering numerical data*, kemudian *clustering numerical data* diubah kedalam interval-interval yang memiliki berbeda-beda panjang selang pada setiap intervalnya.

Langkah kerja algoritma *Automatic Clustering* antara lain:

Langkah 1: Urutkan data *training* curah hujan dari datum terkecil sampai datum terbesar tanpa ada datum ganda (datum yang nilainya sama). Jika ada datum ganda dalam pengurutan data, maka ambil satu datum dari datum ganda tersebut. Misalkan data yang sudah diurutkan tanpa ada datum ganda digambarkan sebagai $d_1, d_2, d_3, \ldots, d_i, \ldots, d_n$
lalu hitung nilai *"average_dif"* dengan cara berikut:

$$average\_dif = \frac{\sum_{i=1}^{n-1}(di+1 - di)}{n-1} \qquad \textbf{(2.1)}$$

dimana *"average_dif"* adalah nilai rata-rata dari selisih dua datum yang berdekatan dalam data yang sudah diurutkan tanpa ada datum ganda.

Langkah 2: Masukkan datum terkecil yang sudah diurutkan tanpa ada data ganda kedalam *cluster* pertama. Berdasarkan nilai dari *"average_dif"* tentukan apakah datum-datum selanjutnya akan dimasukkan kedalam *current cluster* atau dibuatkan *cluster* baru yang beranggotakan datum-datum tersebut dengan aturan-aturan sebagai berikut:

Aturan 1: Asumsikan *current cluster* adalah cluster pertama dan hanya terdapat satu datum d_1 didalamnya, dan d_2 adalah datum yang nilainya lebih besar dan berdekatan dengan d_1 yang digambarkan sebagai $\{d_1\}, d_2, d_3, \ldots, d_n$.
Jika $d_2 - d_1 \leq average\_dif$, maka masukkan d_2 kedalam *current cluster* dimana d_1 berada, jika tidak dibuatkan *cluster* baru yang beranggotakan d_2 dan jadikan *cluster* baru tersebut menjadi *current cluster*.

Aturan 2: Asumsikan *current cluster* bukan *cluster* pertama, dan hanya terdapat satu datum yaitu datum d_j di dalam *current cluster* $\{d_j\}$. Asumsikan d_k adalah datum yang nilainya lebih besar dan berdekatan dengan datum d_j dan d_i adalah datum terbesar di *antecedent cluster* (*cluster* sebelum *current cluster*), digambarkan sebagai $\{d_1\}, \ldots, \{\ldots, d_i\}, \{d_j\}, d_k, \ldots, d_n$.

Jika $d_k - d_j \leq average\_dif$ dan $d_k - d_j \leq d_j - d_i$, maka masukkan d_k kedalam *cluster* yang dimiliki d_j, jika tidak dibuatkan *cluster* baru yang beranggotakan d_k dan jadikan *cluster* baru tersebut menjadi *current cluster*.

Aturan 3: Asumsikan *current cluster* bukan *cluster* pertama dan asumsikan d_i adalah datum terbesar di *current cluster* dan d_j adalah datum yang nilainya lebih besar dan berdekatan dengan d_i digambarkan sebagai berikut: $\{d_1\}, \ldots, \{\ldots\}, \{\ldots, d_i\}, d_j, \ldots, d_n$.

Jika $d_j - d_i \leq average\_dif$ dan $d_j - d_i \leq cluster\_dif$, maka masukkan d_j kedalam *cluster* yang beranggotakan d_i, jika tidak dibuatkan *cluster* baru yang beranggotakan d_j dan jadikan *cluster* baru tersebut menjadi *current cluster*. Perhitungan *cluster_dif* dapat ditampilkan sebagai berikut:

$$cluster\_dif = \frac{\sum_{i=1}^{n-1}(cli+1 - cli)}{n-1} \qquad \textbf{(2.2)}$$

dimana *cluster_dif* adalah rata-rata dari selisih datum anggota *current cluster* yang berdekatan dan cl_1, $cl_2, \ldots, cl_n$ adalah datum yang menjadi anggota di *current cluster*.

Langkah 3: Perbarui anggota-anggota pada setiap *cluster* yang diperoleh dari Langkah 2 berdasarkan tiga aturan berikut:

Aturan 1: Jika *cluster* memiliki anggota lebih dari dua datum, maka pertahankan datum terkecil dan datum terbesar, kemudian hapus datum lainnya yang berada pada *cluster* tersebut.

Aturan 2: Jika *cluster* memiliki dua anggota datum, maka pertahankan keduanya.

Aturan 3: Jika *cluster* memiliki satu anggota datum d_q, maka masukkan nilai *"d_q – average_dif"* dan *"d_q + average_dif"* kedalam *cluster*, dan hapus datum d_q dari *cluster*. Tetapi juga harus menyesuaikan dengan situasi berikut:

Situasi 1: jika *cluster* pertama, maka hapus *"d_q – average_dif"* dan pertahankan d_q.

Situasi 2: jika *cluster* terakhir, maka hapus *"d_q + average_dif"* dan pertahankan d_q.

Situasi 3: jika *"d_q – average_dif"* lebih kecil dari nilai datum terkecil di *antecedent cluster*, maka Aturan 3 tidak berlaku, sehingga anggota cluster tersebut tetap d_q.

Langkah 4: Asumsikan bahwa hasil *clustering* yang diperoleh dari Langkah 3 adalah sebagai berikut: $\{d_1, d_2\}, \{d_3, d_4\}, \ldots, \{d_r\}, \{d_s, d_t\}, \ldots, \{d_1, d_2\}$. Ubah hasil cluster yang berdekatan berdasarakan *sub* langkah berikut:

Sub Langkah 4.1 ubah *cluster* pertama $\{d_1, d_2\}$ menjadi interval $[d_1, d_2)$.

Sub Langkah 4.2 jika *current interval* $[d_i, d_j)$ dan *current cluster* $\{d_k, d_l\}$, maka:

1) Jika $d_j \geq d_k$, maka bentuk sebuah interval $[d_j, d_l)$. sekarang interval $[d_j, d_l)$ menjadi *current interval* dan *next cluster* $\{d_p, d_q\}$ menjadi *current cluster*.
2) Jika $d_j < d_k$, maka ubah *current cluster* $\{d_k, d_l\}$ menjadi interval $[d_k, d_l)$ dan buat interval baru $[d_j, d_k)$ diantara interval $[d_i, d_j)$ dan $[d_k, d_l)$. Sekarang $[d_k, d_l)$ menjadi *current interval* dan *next cluster* $\{d_p, d_q\}$ menjadi *current cluster*. Jika *current interval* $[d_i, d_j)$ dan *current cluster* adalah $\{d_k\}$, maka ubah *current interval* $[d_i, d_j)$ menjadi $[d_i, d_k)$. Sekarang $[d_i, d_k)$ adalah *current interval* dan *next cluster* menjadi *current cluster*.

Sub Langkah 4. 3 Ulangi *sub* langkah 4.1 dan 4.2 sampai semua *cluster* menjadi interval (u_n) dimana n adalah banyaknya interval yang diperoleh di tahap ini.

Kemudian lakukan klasifikasi data menggunakan algoritma *Automatic Clustering* sehingga menghasilkan interval-interval. Berikut adalah langkah-langkah penerapan *Automatic Clustering*:

Langkah 1: Pertama lakukan pengurutan data *training*, kemudian hapus datum ganda

0.0, 2.9, 4.6, 16.5, 18.8, 21.0, 22.0, 24.7, 25.0, 30.1, 32.6, 36.1, 41.6, 47.0,

49.3, 49.7, 50.3, 51.7, …, 415.0, 419.4, 427.2, 482.7, 547.5, 626.1, 680.0.

Setelah dilakukan pengurutan data, lalu hitung nilai *"average_dif"* yang merujuk pada persamaan **(2.1)**

$$average\_dif = \frac{(2.9-0.0)+(4.6-2.9)+(16.5-4.6)+ \ldots +(680.0-626.1)}{95}$$

$$= 7.16$$

Langkah 2: Berikut adalah hasil *cluster-cluster* yang diperoleh beserta anggotanya:

$cl_1 = \{0.0, 2.9, 4.6\}$

$cl_2 = \{16.5, 18.8, 21, 22\}$

$cl_3 = \{24.7, 25.0\}$

$\vdots$

$cl_{47} = \{680.0\}$

Langkah 3: Berikut adalah hasil *cluster-cluster* yang telah diperbarui beserta anggotanya

cl_1 = {0.0, 4.6}

cl_2 = {16.5, 22}

cl_3 = {24.7, 25.0}

⋮

cl_{47} = {672.84, 680.0}

Langkah 4: Berikut adalah interval-interval yang diperoleh,

u_1 = [0, 4.60)

u_2 = [4.60, 16.50)

u_3 = [16.50, 22.00)

⋮

u_{74} = [672.84, 680.00]

2.2. Rancang Data Sampel Berlabel *Fuzzy Trend*

Tahap ini akan membentuk sebuah fuzzifikasi data *training* beserta label *fuzzy trend*-nya. *Fuzzy classification* dapat diartikan sebagai struktur data dan menggunakan aturan "*if-then"* untuk memperoleh kategori suatu data beserta label kelas.

Fuzzy classification dapat diartikan sebagai struktur data dan menggunakan aturan "*if then"* untuk memperoleh kategori suatu data beserta label kelas.

2.2.1. Proses Fuzzifikasi

Misalkan himpunan fuzzy E_i direpresentasikan sebagai

$Ei = ei1/u1 + ei2/u2 + ... + ein/un$, dimana $e_{ij} \in [0,1]$, $1 \leq i, j \leq n$. Nilai dari e_{ij} menunjukan nilai derajat keanggotaan dari u_j dalam pada himpunan fuzzy E_i.

Data *training* difuzzifikasi berdasarkan representasi himpunan fuzzy. Data *training* difuzzifikasikan kedalam E_i jika derajat keanggotaan terbesar berada di dalam E_j, dalam arti jika $e_i = \max\{e_{i1}, e_{i2}, \ldots, e_{in}\}$, $\{1 \leq j \leq n\}$, kemudian data *training* diklasifikasi kedalam kelas E_i.

Berdasarkan interval-interval yang telah diperoleh dari data curah hujan provinsi Kalimantan Selatan, himpunan fuzzy didefinisikan sebagai berikut:

$E_1 = 1/u_1 + 0.5/u_2 + 0/u_1 + 0/u_3 + 0/u_4 + \ldots + 0/u_{n-1} + 0/u_{74}$,

$E_2 = 0.5/u_1 + 1/u_2 + 0.5/u_1 + 0/u_3 + 0/u_4 + \ldots + 0/u_{n-1} + 0/u_{74}$,

$E_3 = 0/u_1 + 0.5/u_2 + 1/u_1 + 0.5/u_3 + 0/u_4 + \ldots + 0/u_{n-1} + 0/u_{74}$,

⋮

$E_{74} = 1/u_1 + 0.5/u_2 + 0/u_1 + 0/u_3 + 0/u_4 + \ldots + 0/u_{n-1} + 0/u_{74}$,

kemudian setiap datum pada data curah hujan menjadi anggota himpunan fuzzy $E_i (i = 1, 2, 3, \ldots, 74)$.

Misalkan pada bulan Januari tahun 2003 memiliki nilai curah hujan 395.2, dengan merujuk pada *Tabel 4.1*, maka nilai curah hujan 395.2 terdapat pada selang interval E_{62}, kemudian ubah nilai curah hujan 395.2 menjadi E_{62}. Setelah itu fuzzyfikasi data *training* sehingga menghasilkan,

395.2 → E_{62},

547.5 → E_{70}

150.0 → E_{24}

⋮

354.7 → E_{56}

2.2.2 Bangun *Second Order Fuzzy Relationship*

Dalam langkah ini dibuat data curah hujan bahwa curah hujan sekarang (f_t) dipengaruhi oleh curah hujan dua bulan sebelum bulan sekarang (f_{t-2}), dan satu bulan sebelum sekarang (f_{t-1}) yang digambarkan E_{i2}, E_{i1} → E_j, dimana E_{i2} adalah nilai fuzzyfikasi dua bulan sebelum sekarang, E_{i1} dalah nilai fuzzyfikasi satu bulan sebelum sekarang, dan E_j adalah nilai fuzzyfikasi bulan sekarang. sehingga menghasilkan *second order fuzzy relationship* sebagai berikut:

$E_{i2}, E_{i1} \rightarrow E_j$

$E_{62}, E_{70} \rightarrow E_{24}$

$E_{70}, E_{24} \rightarrow E_{30}$

$E_{24}, E_{30} \rightarrow E_9$

⋮

$E_{50}, E_{56} \rightarrow E_{64}$

2.2.3. *Capture Fuzzy Trends of Historical Samples*

Langkah 1: Mencari nilai $\blacktriangledown z^+$ dan nilai $\blacktriangledown z^-$

Hitung selisih antara dua nilai curah hujan yang berdekatan dalam data *training*, misal $\blacktriangledown z_t = z_t - z_{t-1}$, kemudian diperoleh rata-rata selisih positif dan selisih negatif yaitu threshold $\blacktriangledown z^+$ dan threshold $\blacktriangledown z^-$. Berikut adalah perhitungan untuk mencari nilai threshold $\blacktriangledown z^+$ dan threshold $\blacktriangledown z^-$:

$$\blacktriangledown z^+ = \frac{\sum_{t=1}^{n} (z+)}{\varepsilon + n} \qquad (2.4)$$

$$\blacktriangledown z^- = \frac{\sum_{t=1}^{n} (z-)}{\varepsilon + n} \qquad (2.5)$$

dengan nilai ε = 0.00001

Langkah 2: setelah diperoleh nilai dari threshold $\blacktriangledown z^+$ dan threshold $\blacktriangledown z^-$ ada aturan-aturan untuk memberi label kelas pada tiap datum.

Situasi 1: Jika $\blacktriangledown mt_t > \blacktriangledown Z^+$ maka tren fuzzy dari *current state* ke *state* berikutnya adalah "*upward*".

Situasi 2: Jika $\blacktriangledown Z^- \leq \blacktriangledown mt_t \leq \blacktriangledown Z^+$ maka tren fuzzy dari *current state* ke *state* berikutnya adalah "*unchanged*".

Situasi 3: Jika $\blacktriangledown mt_t < \blacktriangledown Z^-$ maka tren fuzzy dari *current state* ke *state* berikutnya adalah "*downward*".

dimana nilai $\blacktriangledown mt_t$ adalah selisih *median* E_{i1} – *median* E_j.

Kemudian, hasil dari data sampel berlabel *fuzzy trend* dapat direpresentasikan sebagai berikut:

$$[i_2, i_1, (i_2\text{-}i_1), y] \qquad (2.6)$$

dimana i_1 dan i_2 adalah *subscript* dari E_{i1} dan E_{i2}, dan *y* adalah fuzzy tren "*upward*", "*unchanged*", dan "*downward*". **[i_2, i_1, (i_2-i_1)]** menyatakan fitur vektor dari karakter pemodelan untuk data *training* curah hujan dan ***y*** menyatakan sebuah label kelas. Berdasarkan hasil trend yang diperoleh dari **Langkah 2** ini, kemudian bangun vektor **[i_2, i_1, (i_2-i_1), *y*]** sebagai berikut:

62, 70, 8, *Downward*

70, 24, -46, *Unchanged*

24, 30, 6, *Downward*

⋮

43, 50, 7, *Unchanged*.

2.3. *Axiomatic Fuzzy Set* (AFS) *classification*

Axiomatic Fuzzy Sets (AFS) classification merupakan algoritma yang mencari nilai derajat keanggotaan pada tiap-tiap kelas. Dalam metode ini juga diperoleh kelas-kelas dan nilai derajat keanggotaan di tiap kelasnya.

2. 3. 1 AFS *Algebras*

Data sampel berlabel *fuzzy trend* yang telah diperoleh, memiliki tiga fitur, antara lain:

Fitur pertama (f_1) adalah nilai-nilai dari i_2, dimana nilai i_2 adalah interval-interval ke-*n*.

Fitur kedua (f_2) adalah nilai-nilai dari i_1, dimana nilai i_1 adalah interval-interval ke-*n*.

Fitur ketiga (f_3) adalah nilai selisih dari i_1 - i_2.

Langkah 1: Ubah nilai fitur pertama, fitur kedua, dan fitur ketiga, kedalam kategori "*large*", "*medium*", dan "*small*".

Untuk menentukan nilai f_1 dan f_2 yaitu berdasarkan banyaknya interval-interval u_n (n = 1, 2, 3, …, n), dalam arti terdapat interval sebanyak n. Kemudian banyaknya interval tersebut dibagi menjadi tiga bagian, bagian pertama adalah mulai dari *1* sampai dengan $\frac{n}{3}$. Bagian kedua adalah mulai dari $\frac{n}{3}$ + *1* sampai dengan $\frac{2n}{3}$. Bagian ketiga adalah mulai dari $\frac{2n}{3}$ + *1* sampai *n*. Bagian pertama dikategorikan *"small",* bagian kedua dikategorikan *"medium"*, dan bagian ketiga dikategorikan *"large".*

Kemudian untuk menentukan nilai f_3 yaitu berdasarkan banyaknya rentang selisih $i_2 - i_1$ (*p, p+1, p+2, …, q*), dimana *p* adalah nilai selisih terkecil dan *q* adalah nilai selisih terbesar. Kemudian banyaknya rentang selisih tersebut dibagi menjadi tiga bagian. Setelah dilakukan pembagian, bagian pertama adalah mulai dari *p* sampai dengan $\frac{(q-p)}{3}$, bagian kedua adalah mulai dari $\frac{(q-p)}{3}$ + 1 sampai dengan $\frac{2(q-p)}{3}$, dan bagian ketiga adalah dimulai dari $\frac{2(q-p)}{3}$ + 1 sampai dengan *p*.

Untuk merepresentasikan langkah ini, dapat digambarkan dalam aturan-aturan sebagai berikut:

Situasi 1: berlaku untuk f_1 dan f_2

$i_2 \leq \frac{n}{3}$, maka f_1 = *"small"*, $i_1 \leq \frac{n}{3}$, maka f_2 = *"small"*.

$\frac{n}{3} < i_2 \leq \frac{2n}{3}$, maka f_1 = *"medium"*, $\frac{n}{3} < i_2 \leq \frac{2n}{3}$, maka f_1 = *"medium"*,

$i_2 \leq \frac{2n}{3}+1$, maka f_1 = *"large"*, $i_1 \leq \frac{2n}{3}+1$, maka f_2 = *"large"*.

Situasi 2: berlaku untuk f_3

$i_1 - i_2 \leq \frac{(q-p)}{3}$, maka f_3 = *"small"*,

$\frac{(q-p)}{3} < i_1 - i_2 \leq \frac{2(q-p)}{3}$, maka f_3 = *"medium,*

$i_1 - i_2 \leq \frac{2(q-p)}{3}+1$, maka f_3 = *"large"*.

Setelah diperoleh nilai-nilai dari *f1*, *f2*, dan f_3 kemudian diikuti dengan nilai *trend* ***y*** (*Downward, Unchanged, Upward*).

Langkah 2: bangun pola x = [x_1, x_2, x_3], dimana x_i adalah fitur ke-*i*. Ubah nilai-nilai *f1*, *f2*, dan f_3 kedalam bentuk m_{jk} dimana *j* adalah fitur ke-*i* (*i* = 1, 2, 3) dan *k* adalah *term* (*small, medium, large*). Jika *term* = *"small"*, maka *k* = 3. Jika *term* = *"medium",* maka *k* = 2. Jika *term* = *"large",* maka *k* = 1.

Setelah terbangun pola *x*, kemudian pola *x* tersebut, kemudian diikuti dengan nilai *trend* ***y***.

Langkah 3: bangun *fuzzy rules* dari pola *x* yang telah didapat pada Langkah 2. Cara membangun *fuzzy rules* yaitu dengan cara menbentuk $A_1 = x_1.x_3$ dan $A_2 = x_2.x_3$ diikuti dengan *trend* ***y*** untuk setiap pola *x*. Kemudian pada setiap pola *x* akan membentuk *fuzzy rules*

$$\sum_{u=1}^{2}(\prod_{m\in \mathrm{Ai}} m) = \boldsymbol{y} \qquad \textbf{(2.7)}$$

Langkah 4: bentuk konsep fuzzy **EM. EM** adalah himpunan yang beranggotakan tiga *fuzzy rules* (*Downward, Unchanged, Upward*) dengan cara mengklasifikasikan *fuzzy rules* yang dimiliki oleh setiap ***y*.** Setelelah diklasifikasikan, tentukan *fuzzy rules* di setiap ***y*** dengan cara membuang pola *x* ganda yang jumlahnya lebih sedikit dibanding pola *x* ganda pada ***y*** yang lain, jika jumlah *fuzzy rules*-nya sama, maka pilih salah satu ***y*** yang boleh memiliki fuzzy rules tersebut. Hal ini dilakukan agar setiap ***y*** memiliki *fuzzy rules* yang berbeda. Kemudian hapus *fuzzy rules* ganda pada setiap ***y*.** Buang *x* ganda yang terdapat pada setiap ***y*** lalu terbentuklah konsep fuzzy **EM.**

Setelah membangun *fuzzy rules*, lakukan pembentukan konsep fuzzy **EM** dengan cara merujuk *AFS Algebras* **Langkah 4,** sehingga diperoleh hasil berikut:

(*Downward*) $= m_{11}m_{32} + m_{21}m_{32} + m_{12}m_{31} + m_{21}m_{31} + m_{13}m_{31} + m_{13}m_{32} + m_{22}m_{32}$;

(*Unchanged*) $= m_{11}m_{32} + m_{22}m_{32} + m_{11}m_{33} + m_{21}m_{33} + m_{12}m_{32} + m_{21}m_{32} + m_{23}m_{32} + m_{12}m_{33} + m_{23}m_{33} + m_{13}m_{32}$;

(*Upward*) $= m_{11}m_{33} + m_{23}m_{33} + m_{12}m_{31} + m_{21}m_{31} + m_{12}m_{33} + m_{22}m_{33} + m_{13}m_{31}$

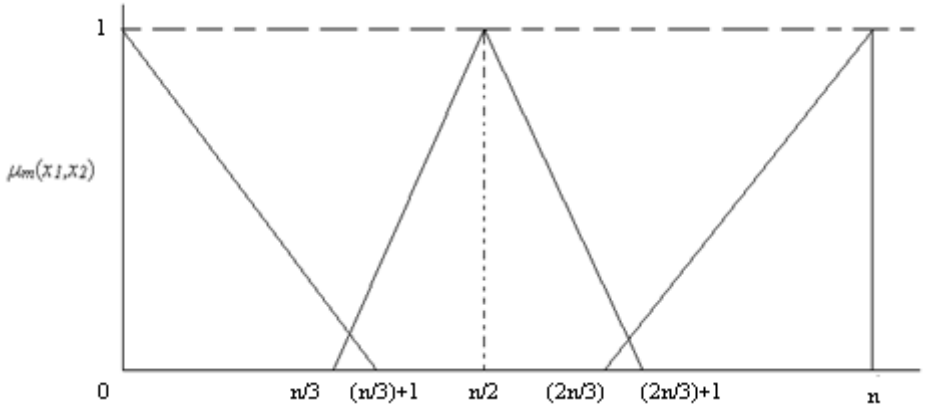
Setelah diperoleh hasil di atas, maka konsep fuzzy **EM** berisi himpunan {$m_{11}m_{32} + m_{21}m_{32} + m_{12}m_{31} + m_{21}m_{31} + m_{13}m_{31} + m_{13}m_{32} + m_{22}m_{32}$, $m_{11}m_{32} + m_{22}m_{32} + m_{11}m_{33} + m_{21}m_{33} + m_{12}m_{32} + m_{21}m_{32} + m_{23}m_{32} + m_{12}m_{33} + m_{23}m_{33} + m_{13}m_{32}$, $m_{11}m_{33} + m_{23}m_{33} + m_{12}m_{31} + m_{21}m_{31} + m_{12}m_{33} + m_{22}m_{33} + m_{13}m_{31}$}

2. 3. 2 *A Classifier Design based on* AFS *Fuzzy Logic*

Klassifikasi dalam tahap ini adalah sebuah *c-class problem* dalam m-dimensi ruang data sampel, dan vector x_p = (x_{p1}, x_{p2}, …, x_{pn}), p = 1, 2, …, n adalah ditetapkan sebagai data sampel dari *c-class*, dimana *c-class* adalah banyaknya kelas X_1, X_2, …, X_c; c = 1, 2, 3 (*Downward, Unchanged, Upward*), dan X adalah gabungan dari semua X_c.

Langkah-langkah detail dari tahap *A Classifier Design based on AFS Fuzzy Logic* adalah sebagai berikut:

Langkah 1: Hitung derajat keanggotaan setiap μ_m, $m \in M$ ($M = \{m_{11}, m_{12}, m_{13}, m_{21}, m_{22}, m_{23}, m_{31}, m_{32}, m_{33}\}$, $x_1 = \{m_{11}, m_{12}, m_{13}\}$, $x_2 = \{m_{21}, m_{22}, m_{23}\}$, $x_3 = \{m_{31}, m_{32}, m_{33}\}$) yang membutuhkan perhitungan nilai fungsi keanggotaan. Untuk menghitung $\mu_m(x_1,x_2)$ adalah menggunakan fungsi keanggotaan yang direpresentasikan sebagai berikut:



Gambar 2.1 Representasi $\mu_m(x_1,x_2)$

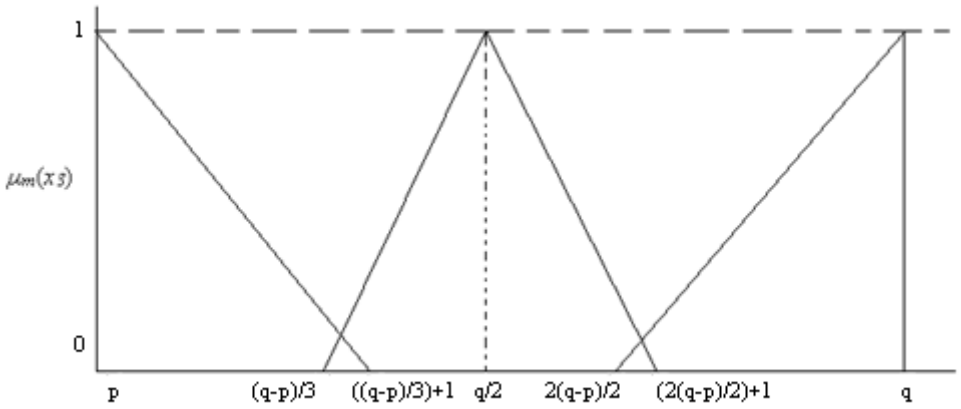
Jika $m < \frac{n}{3} + 1$, maka $\mu_m(x_1,x_2) = \frac{(\frac{n}{3}+1) - m}{(\frac{n}{3}+1) - 0}$;

Jika $\frac{n}{3} < m \leq \frac{n}{2}$, maka $\mu_m(x_1,x_2) = \frac{(m - \frac{n}{3})}{(\frac{n}{2} - \frac{n}{3})}$

Jika $\frac{n}{2} < m \leq \frac{2n}{3} + 1$, maka $\mu_m(x_1,x_2) = \frac{\left(\frac{2n}{3}+1\right) - m}{\left(\frac{2n}{3}+1\right) - \frac{n}{2})}$;

Jika $\frac{2n}{3} < m$, maka $\mu_m(x_1,x_2) = \frac{(m - \frac{2n}{3})}{(n - \frac{2n}{3})}$

dimana m $\in x_1$ dan x_2. Kemudian untuk menghitung $\mu_m(x_3)$ adalah menggunakan fungsi keanggotaan yang direpresentasikan sebagai berikut:



Gambar 2.2 Representasi $\mu_m(x_3)$

Jika $m < \frac{q-p}{3} + 1$, maka $\mu_m(x_3) = \frac{(\frac{q-p}{3}+1) - m}{(\frac{q-p}{3}+1) - 0)}$;

Jika $\frac{q-p}{3} < m \leq \frac{n}{2}$, maka $\mu_m(x_1,x_2) = \frac{(m - \frac{q-p}{3})}{(\frac{q}{2} - \frac{q-p}{3})}$

Jika $\frac{q}{2} < m \leq \frac{2(q-p)}{3} + 1$, maka $\mu_m(x_1,x_2) = \frac{\left(\frac{2(q-p)}{3}+1\right) - m}{\left(\frac{2(q-p)}{3}+1\right) - \frac{q}{2})}$;

Jika $\frac{2(q-p)}{3} < m$, maka $\mu_m(x_1,x_2) = \frac{(m - \frac{2(q-p)}{3})}{(q-p) - \frac{2q-p}{3})}$

Langkah 2: untuk setiap data sampel $x_0 \in X$, tentukan derajat keanggotaan $\mu_\theta(x_0)$, dimana θ adalah derajat keanggotaan terbesar yang berada di dalam **EM.** Kemudian hitung nilai Bx_0 dengan perhitungan sebagai berikut:

$$Bx_0 = \{m \in M \mid \mu_m(x_0) \geq \mu_\theta(x_0) - 0.5\} \quad \textbf{(2.8)}$$

Bx_0 adalah himpunan dari semua kemungkinan *fuzzy terms* di dalam M yang dikarakterisasikan x_0.

Langkah 3: asumsikan x_0 dimiliki oleh kelas X_c, kemudian tentukan sebuah himpunan yang memiliki anggota sebagai berikut:

$$\Lambda x_0 \{\gamma = \prod_{m \in H} m \mid H \in BX_0 ; \mu_\gamma(x_0) \geq 0.5 ; \forall \gamma \in X - X_i \ \mu_\gamma(y) < 0.5\} \quad \textbf{(2.9)}$$

Sub Langkah 3.1: tentukan nilai dari ζx_i dengan cara berikut:

Jika $\Lambda x_0 = \varnothing$, maka $\zeta x_i = 0$;

Jika $\Lambda x_0 \neq \varnothing$, maka $\zeta x_i = arg\ max\ \{\mu_\gamma(x_0)\}$, $\gamma \in \Lambda x_0$.

Selanjutnya tentukan anggota himpinan pada CX_i dengan cara sebagai berikut:

$$CX_i = \{\zeta_x \mid \zeta_x \neq 0, x \in X_i\} \quad \textbf{(2.10)}$$

Langkah 4: Bangun karakterisasi dari kelas X_i sebagai berikut:

$$\zeta_{Xi} = \bigvee_{\zeta \in CXi} \zeta = \sum_i (\prod_{m \in Ai} m) \in \textbf{EM,}\ i = 1, 2, 3. \quad \textbf{(2.11)}$$

untuk setiap data uji (x_{test}) memiliki:

$$A^{\succeq}_k (x_{test}) = \{x \in X \mid x_{test} \succeq_m x,\ m \in A_k\} \subseteq X,\ k \in I \quad \textbf{(2.12)}$$

kemudian tentukan derajat keanggotaan x_{test} yang ada didalam ζ_{Xi} denga cara berikut:

$$\mu_{\zeta_{X_i}}(x_{test}) = \sup_{i \in I}\left(\prod_{\gamma \in A_i} m_\gamma\left(A_i^{\succeq}(x_{test})\right)\right) = \sup_{i \in I}\left(\prod_{\gamma \in A_i} \frac{\sum_{x \in A_i^{\succeq}(x_{test}) \cup x_{test}} \rho_\gamma(x)}{\sum_{x \in X \cup x_{test}} \rho_\gamma(x)}\right) \quad \textbf{(2.13)}$$

Setelah itu, label kelas dari x_{test} adalah nilai terbesar dari $\mu\zeta_{Xi}$ (x_{test}). Berdasarkan persamaan **(2.13)**, diperoleh $\mu\zeta_{Xdownward}$ (x_{test1}) = 0.430, $\mu\zeta_{Xunchanged}$ (x_{test1}) = 0.199, $\mu\zeta_{Xupward}$ (x_{test1}) = 0.436. Kemudian nilai terbesar dijadikan label kelas dari data uji bulan Januari tahun 2011 atau bisa disebut x_{test1}, dengan demikian label kelas dari x_{test1} adalah "*upward*" dengan nilai $\mu\zeta_{Xupward}$ (x_{test1}) sebesar 0.436

2.4. *Forecasting*

Pada tahap ini, label dengan nilai derajat keanggotaan tertinggi akan digunakan untuk memperoleh nilai prediksi curah hujan. Untuk memperoleh nilai predeksi curah hujan bulan mendatang, hitung nilai $F_{forecast}$ dengan cara berikut ini:

jika $\mu\zeta_{Xi}(x_{test}) = \mu\zeta_{Xdownward}$, maka $F_{forecast} = mt_{test} + \mu\zeta_{Xdownward} * \blacktriangledown z^{+}$;

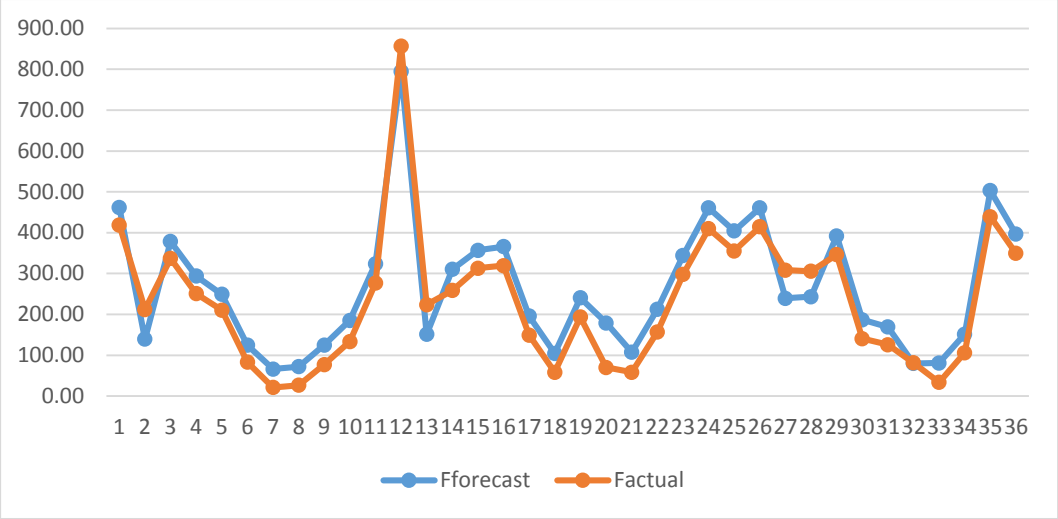
jika $\mu\zeta_{Xi}(x_{test}) = \mu\zeta_{Xunchanged}$, maka $F_{forecast} = mt_{test}$;

jika $\mu\zeta_{Xi}(x_{test}) = \mu\zeta_{Xupward}$, maka $F_{forecast} = mt_{test} + \mu\zeta_{Xupward} * \blacktriangledown z^{-}$

dimana nilai mt_{test} adalah nilai *median* dari interval yang dimiliki datum curah hujan pada data uji yang sudah difuzifikasi. x_{test1} memiliki label kelas "*upward"*, maka hasil prediksi curah hujan bulan Januari tahun 2011adalah $F_{forecast}$ (x_{test1}) = 413.90 + (0.436*96.35) = 455.98.

3. Hasil dan Pembahasan

Pada tahap ini dilakukan perbandingan data *forecast* dan data aktual atau data uji menggunakan skema grafik, dimana sumbu *x* adalah kurun waktu perbulan, dan sumbu *y* adalah nilai curah hujan.



Gambar 3.1 Grafik perbandingan data *forecast* dengan data aktual

Berdasarkan grafik tersebut, dapat disimpulkan bahwa metode ini mampu menangkap pola nilai curah hujan di waktu mendatang. Setelah diperoleh data hasil prediksi, kemudian dilakukan perhitungan nilai *error* menggunakan perhitungan RMSE dan MAPE.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(Fforecast(t)-Factual(t))^2}{n}}$$

$$= 52.32$$

$$MAPE = \frac{100}{n} \cdot \sum_{t=1}^{n} \frac{|Factual-Forecast|}{Factual}$$

$$= 43.53$$

4. Kesimpulan dan Saran

4.1. Kesimpulan

Berdasarkan hasil dari prediksi curah hujan yang telah dilakukan, maka dapat disimpulkan:

1. Metode *Fuzzy Forecasting* berbasis *Automatic Clustering* dan *Axiomatic Fuzzy Sets* (AFS) *classification* berhasil menangkap pola curah hujan di waktu mendatang.
2. Metode ini menghasilkan nilai *error* RMSE sebesar 52.32 dan nilai *error* MAPE sebesar 43.53

4.2. Saran

Untuk memperkecil nilai *error* dalam penelitian selanjutnya, disarankan untuk memperbanyak pengumpulan data sampel terutama pada data *training*.

5. Daftar Pustaka

[1] Weina Wang, Xiaodong Liu. 2014. *Fuzzy forecasting based on automatic clustering and axiomatic fuzzy set classification.*

[2] Sri Kusumadewi, Hari Purnomo. 2010. *Aplikasi logika fuzzy untuk pendukung keputusan.*

[3] Bayong Tjasyono. 2004, *Klimatologi*.

[4] Badan Pusat Statistik Provinsi Kalimantan Selatan. *https://kalsel.bps.go.id/linkTabelStatis/view/id/1193*

[5] Yusuf Priyo Anggodo, Wayan Firdaus Mahmudy. 2016. *Peramalan butuhan hidup minimum menggunakan automatic clustering dan fuzzy logical relationship.*

[6] Poosapati Padmaja. 2016. *A study of notations and illustrations of axiomatic fuzzy set theory.*

[7] Xiaodong Liu, Witold Pedrycz. 2009. *Axiomatic fuzzy set theory and its applications.*

[8] S.M. Chen, *Forecasting enrollments based on fuzzy time series, Fuzzy Sets Syst*. 81 (3) (1996) 311–319.

[9] S.M. Chen, *Forecasting enrollments based on high-order fuzzy time series*, Cybernet. Syst. 33 (1) (2002) 1–16.

[10] T.H.K. Huarng, K.H. Yu, Y.W. Hsu, *A multivariate heuristic model for fuzzy time-series forecasting, IEEE Trans. Syst., Man*, Cybernet.-Part B: Cybernet. 37 (4) (2007) 836–846.

[11] N.Y. Wang, S.M. Chen, *Temperature prediction* dan *TAIFEX forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series*, Exp. Syst. Appl. 36 (2) (2009) 2143–2154.

[12] S.M. Chen, C.D. Chen, TAIEX *forecasting based on fuzzy time series and fuzzy variation groups*, IEEE *Trans*. Fuzzy Syst. 19 (1) (2011) 1–12.

[13] S.M. Chen, N.Y. Wang, *Fuzzy forecasting based on fuzzy-trend logical relationship groups, IEEE Trans. Syst., Man, Cybernet.-Part B: Cybernet*. 40 (5) (2010) 1343–1358.

[14] K. Huarng, T.H.K. Yu, *Ratio-based lengths of intervals to improve fuzzy time series forecasting, IEEE Trans. Syst., Man, Cybernet.-Part B: Cybernet*. 36 (2) (2006) 328–340.

[15] J. Sullivan, W.H. Woodall, *A comparison of fuzzy forecasting and Markov modeling, Fuzzy Sets Syst*. 64 (3) (1994) 279–293.

[16] Robert Kurniawan, *Metode Automatic-Fuzzy Logic relationship untuk Peramalan Data Univariate*.