

PERAMALAN DATA PRODUK DOMESTIK BRUTO DENGAN FUZZY TIME SERIES MARKOV CHAIN

Maria Titah Jatipaningrum
Jurusan Statistika, Fakultas Sains Terapan
Institut Sains & Teknologi AKPRIND Yogyakarta
titahjp@akprind.ac.id

ABSTRACT

In this study, a fuzzy time series-Markov chain approach for analyzing the linguistic or small sample time series data is proposed to further enhance the predictive accuracy, by transferring fuzzy time series data to the fuzzy logic group to derive Markov chain transition matrix, then the matrix obtained is used for forecasting. This paper attempts to improve the accuracy of forecasting by Markov chain and classification states. The method is applied to the time series data set related to Indonesia's GDP. Finally to verify the effectiveness and confirms the potential benefits of the proposed approach with very small MAPE.

Keywords— Fuzzy time series Markov chain, Indonesia's GDP, fuzzy logic group

INTISARI

Penelitian ini membahas pendekatan fuzzy time series Markov chain untuk menganalisis data linguistik atau data time series sampel kecil diusulkan supaya keakuratan prediksi lebih tinggi dengan mentransfer data time series ke grup logika fuzzy, dan menggunakannya untuk mendapatkan matriks transisi Markov chain kemudian digunakan untuk peramalan. Makalah ini berusaha untuk meningkatkan akurasi peramalan dengan Markov chain dan klasifikasi dari ruang keadaan. Metode ini diterapkan pada data time series Produk Domestik Bruto. Pengujian dilakukan untuk melihat akurasi peramalan dengan MAPE (*Mean Average Percentage Error*).

Kata Kunci— Fuzzy time series Markov chain, Produk Domestik Bruto, grup logika fuzzy

PENDAHULUAN

Peramalan adalah proses perkiraan tentang sesuatu yang terjadi pada waktu yang akan datang berdasarkan data empiris. Tujuan utama proses peramalan adalah mengurangi ketidakpastian dan membuat perkiraan lebih baik dari apa yang akan terjadi di masa depan. Sistem peramalan yang handal (*reliable*) dan akurat (*accurate*) dibutuhkan untuk membuat keputusan yang tepat tentang hal-hal yang esensial termasuk variabel-variabel kritical selain melibatkan hal-hal subyektif seperti intuisi dan pengalaman.

Produk Domestik Bruto (PDB) didefinisikan sebagai jumlah total produk berupa barang dan jasa yang dihasilkan oleh unit-unit produksi dalam batas wilayah suatu negara selama satu tahun. Data pergerakan nilai PDB mengikuti suatu runtun waktu didasarkan pada urutan titik data dalam selang waktu tertentu yang dirangkum secara tahunan (*long term time series*).

Pendekatan baru teknik peramalan dengan *soft computing* digunakan pada model-model kompleks dan model non linear yang tidak dapat diselesaikan dengan metode klasik. Logika fuzzy sebagai komponen utama pembangun *soft computing* terbukti memiliki kinerja lebih baik dapat menyelesaikan

masalah termasuk peramalan data runtun waktu.

Peramalan dengan model fuzzy dikenal sebagai fuzzy time series. Fuzzy time series menangkap pola data masa lalu kemudian digunakan untuk memproyeksikan data masa depan. Keunggulannya adalah mendefinisikan relasi fuzzy yang dibentuk dengan menentukan hubungan logika dari data training.

Fuzzy time series dikembangkan oleh Song dan Chissom (Song, 1993) dikenal sebagai fuzzy time series klasik yang menggunakan relasi fuzzy. Relasi fuzzy dibentuk dari hubungan logika data latih yang melibatkan himpunan fuzzy dari partisi himpunan universal. Pembagian partisi himpunan universal berdasarkan distribusi statistik pada setiap partisi. Penggunaan distribusi statistik sebagai pertimbangan partisi ulang dan penggunaan data penyusun himpunan universal masih menjadi masalah terbuka dalam rangka bagaimana menentukan model peramalan yang optimal untuk meningkatkan performansi peramalan.

Permasalahan yang dibahas adalah bagaimana membentuk fuzzy time series Markov chain untuk menganalisis data linguistik atau data time series sampel kecil, diusulkan supaya keakuratan prediksi lebih

tinggi dengan mentransfer data time series ke grup logika fuzzy, dan menggunakannya untuk mendapatkan matriks transisi Markov chain kemudian digunakan untuk peramalan. Selanjutnya dilihat peningkatan kehandalan dan akurasi peramalan dengan MAPE.

Fuzzy time series pertama kali diusulkan oleh Song dan Chissom yang diterapkan dalam konsep logika fuzzy untuk mengembangkan dasar dari fuzzy time series dengan menggunakan metode time invariant dan time variant yang digunakan untuk memodelkan peramalan jumlah pendaftar di suatu Universitas (Song, 1993). Sejak saat itu, banyak metode fuzzy time series yang diusulkan seperti, model Chen (Chen, 1996), model weighted (Yu, 2005), model markov (Sullivan, 1994), model persentase perubahan jumlah pendaftaran universitas (Stevenson, 2009), menggunakan perbedaan dari jumlah pendaftar (Melike, 2005), dan multiple-atribut metode fuzzy time series (Cheng, 2008).

Fuzzy time series markov chain (Tsaur, 2012) merupakan konsep baru dalam penelitiannya untuk menganalisis keakuratan prediksi nilai tukar mata uang Taiwan dengan dolar US. Dalam penelitiannya Tsaur menggabungkan metode fuzzy time series dengan Markov chain, penggabungan tersebut bertujuan untuk memperoleh probabilitas terbesar menggunakan matriks probabilitas transisi. hasil penelitian tersebut menunjukkan bahwa metode fuzzy time series markov chain memberikan akurasi yang cukup baik dibandingkan dengan metode fuzzy time series

METODOLOGI PENELITIAN

Data runtun waktu adalah jenis data yang dikumpulkan menurut urutan waktu dalam suatu rentang waktu tertentu. Tipe data runtun waktu dibagi dalam empat komponen yaitu:

1. Trend (T) adalah pola pergerakan naik atau turun secara bertahap dari waktu ke waktu dalam periode tertentu.
2. Siklus (Cycles/C) adalah pola data yang terjadi dalam beberapa periode tertentu dalam satu siklus.
3. Musiman (Seasonality/S) adalah pola fluktuatif diatas atau dibawah garis yang terjadi setiap periode tertentu.
4. Variasi acak (Random/R) adalah "blip" dalam data yang disebabkan oleh kebetulan dan situasi yang tidak biasa dan tidak mengikuti pola yang jelas.

Menurut (Rosadi,2006) analisis data runtun waktu berarti memecah data lampau

menjadi komponen dan memproyeksikannya ke depan. Analisis data runtun waktu, asumsi yang biasanya digunakan adalah data harus stasioner. Model stasioner yaitu model yang sedemikian hingga semua sifat statistiknya tidak berubah dengan pergeseran waktu. Sedang model non stasioner pergerakan data menjauhi sumbu horizontal, sehingga sifat-sifat statistiknya berbeda bisa dalam rata-rata, variansi, atau kovariansi.

Fuzzy time series adalah suatu proses dinamik dari suatu variabel linguistik yang nilai linguistiknya adalah himpunan fuzzy. Pemodelan fuzzy time series mampu memformulasikan permasalahan yang didasarkan para ahli di bidangnya atau berdasarkan data empiris. Definisi-definisi tentang fuzzy time series yang diusulkan oleh Song dan Chissom (Song, 1993) didefinisikan yaitu:

Definisi 1. Suatu himpunan fuzzy adalah suatu kelas dari objek dengan komponen kesatuan dari derajat keanggotaan (grade of membership). Misalkan U adalah himpunan universal, dimana $U=\{u_1, u_2, \dots, u_n\}$ dimana u_i adalah nilai yang mungkin dari U dan misalkan A merupakan himpunan fuzzy dalam himpunan universal U yang didefinisikan dengan:

$$A = \frac{f_A(u_1)}{u_1} + \frac{f_A(u_2)}{u_2} + \dots + \frac{f_A(u_n)}{u_n} \quad (1)$$

Dimana f_A adalah fungsi keanggotaan (membership function) dari A , $f_A : U \rightarrow [0,1]$, $f_A(u_i)$ mengindikasikan derajat (grade) dari keanggotaan u_i dalam himpunan fuzzy A , $f_A(u_i) \in [0,1]$ dan $1 \leq i \leq n$.

Definisi 2. Misalkan $X(t)$ dimana $(t = \dots, 0, 1, 2, \dots)$ adalah himpunan universal dan menjadi subset dari R , dan misalkan $A_i(t), i=1,2,3, \dots$ adalah himpunan fuzzy yang didefinisikan dalam $X(t)$. $F(t)$ disebut sebagai fuzzy time series pada $X(t)$ ($t = \dots, 0, 1, 2, \dots$) jika $F(t)$ merupakan himpunan semua $A_i(t)$, $i = 1, 2, 3, \dots$

Definisi 3. Andaikan $F(t)$ adalah disebabkan dari $F(t-1)$, dinotasikan dengan $F(t-1) \rightarrow F(t)$, maka ada hubungan fuzzy antara $F(t)$ dan $F(t-1)$ yang dapat direpresentasikan sebagai berikut:

$$F(t) = F(t-1) \circ R(t, t-1) \quad (2)$$

Dimana simbol " \circ " dinotasikan dengan operator komposisi max-min. Relasi $R(t, t-1)$ adalah suatu relasi fuzzy antara $F(t)$ dan $F(t-1)$ dan disebut model order-1 dari $F(t)$.

Definisi 4. Jika $F(t)$ disebabkan oleh lebih kecil dari beberapa himpunan fuzzy $F(t-n), F(t-n+1), \dots, F(t-1)$ maka fuzzy relationshipnya diwakili dengan Persamaan berikut ini:

$$A_{t1}, A_{t2}, \dots, A_{tn} \rightarrow A_j \quad (3)$$

Dimana $F(t-n) = A_{t1}$, $F(t-n+1)=A_{t2}, \dots, F(t-1)=A_{tn}$, hubungan ini disebut order ke-n (n^{th} -order) dari model fuzzy time series.

Berdasarkan Definisi-definisi diatas, langkah-langkah untuk membentuk suatu fuzzy time series Markov chain (Tsaur, 2012) sebagai berikut:

1. Mengumpulkan data historikal
2. Mendefinisikan himpunan semesta U dari data, dengan D_1 dan D_2 adalah bilangan positif yang sesuai.
3. Menentukan Menentukan jumlah interval fuzzy, dalam penelitian ini untuk menghitung jumlah interval fuzzy yang terbentuk digunakan metode average based length [Xihao,2008]
4. Mendefinisikan himpunan fuzzy pada universe of discourse U, himpunan fuzzy A_i menyatakan variabel linguistik dari harga saham dengan $1 \leq i \leq n$
5. Fuzzyfikasi data historis. Jika sebuah data time series termasuk ke dalam interval u_i , maka data tersebut di fuzzyfikasi ke dalam A_i .
6. Menentukan fuzzy logical relationship dan Fuzzy Logical Relationships Group (FLRG). Jika himpunan fuzzy sekarang adalah A_i , dan grup relasi logika fuzzy A_i adalah tidak diketahui, misal $A_i \rightarrow \neq$, maka \neq akan merujuk kepada himpunan fuzzy A_i
7. Menghitung hasil peramalan
Untuk data time series, dengan menggunakan FLRG, dapat diperoleh probabilitas dari suatu state menuju ke suatu state berikutnya. Sehingga digunakan matriks transisi probabilitas markov dalam menghitung nilai peramalan, dimensi matriks transisi adalah $n \times n$. Jika state A_i melakukan transisi menuju ke state A_j dan melewati state A_k , $i, j = 1, 2, \dots, n$, maka kita dapat memperoleh FLRG.
8. Menghitung nilai penyesuaian (Dt) pada nilai peramalan. Berikut prinsip-prinsip dalam menghitung nilai penyesuaian.
9. Menghitung nilai peramalan yang telah disesuaikan.

Fuzzy Time Series Markov chain

Peramalan fuzzy time series dilakukan oleh [3] selanjutnya disebut sebagai pendekatan baru fuzzy time series Markov chain. Metode ini mendefinisikan partisi ulang himpunan universal berdasarkan jumlah data yang tersebar setiap partisi sebagai modifikasi metode fuzzy time series klasik.

Perhitungan dengan mengambil sampel data PDB tahun 1978-1997 pada Tabel 1 dengan langkah-langkah metode FTSMC, diuraikan sebagai berikut:

1. Menentukan himpunan universal
Himpunan universal $U = [U_{\min} \quad U_{\max}]$, berdasarkan Tabel 1 nilai terkecil dan terbesar berturut-turut $D_{\min}=219674$ dan $D_{\max} = 627696$. Dimisalkan nilai $D_1 = 9674$, $D_2 = 304$ didapat $U = [21000 \quad 628000]$.
2. Bentuk data input-output sesuai order
Berdasarkan order dibentuk data input dan output dengan susunan $x_{(t-1)}$ sebagai data *current* dan $x_{(t)}$ sebagai data *next*.
3. Partisi himpunan universal
Bagi U dengan panjang interval yang sama sesuai jumlah fm yaitu U_1, U_2, U_3, U_4 , dan U_5 dimana setiap interval memiliki jangkauan sebesar $(628000 - 21000) / 5 = 121400$. Partisi yang dihasilkan adalah: $U_1 = [21000 \quad 142400]$, $U_2 = [142400 \quad 263800]$, $U_3 = [263800 \quad 385200]$, $U_4 = [385200 \quad 506600]$, $U_5 = [506600 \quad 628000]$. Jumlah data setiap partisi adalah $U_1 = 11, U_2 = 4, U_3 = 2, U_4 = 1$, dan $U_5 = 2$.
4. Partisi ulang himpunan universal berdasarkan jumlah data

Berdasarkan jumlah data langkah 3. Partisi dengan jumlah data terbesar dibagi menjadi 4 sub partisi, jumlah data terbesar ke-2 dibagi menjadi 3 sub partisi, jumlah data terbesar ke-3 dibagi menjadi 2 sub partisi. Hasil pembagian partisi ditunjukkan pada Tabel 2.

5. Mendefinisikan himpunan fuzzy
Mendefinisikan himpunan fuzzy berdasarkan pembagian ulang partisi. Himpunan fuzzy dinotasikan dalam variabel linguistik berdasarkan keadaan himpunan semesta U_1, U_2, \dots, U_m . Himpunan fuzzy didefinisikan sebagai berikut:

$$\begin{aligned} A_1 &= \frac{a_{11}}{U_1} + \frac{a_{12}}{U_2} + \dots + \frac{a_{1m}}{U_m} \\ A_2 &= \frac{a_{21}}{U_1} + \frac{a_{22}}{U_2} + \dots + \frac{a_{2m}}{U_m} \\ &\dots \dots \dots \\ A_k &= \frac{a_{k1}}{U_1} + \frac{a_{k2}}{U_2} + \dots + \frac{a_{km}}{U_m} \end{aligned} \quad (4)$$

Dimana $a_{ij} \in [0 \quad 1]$, $1 \leq i \leq k$ dan $1 \leq j \leq m$. Nilai dari a_{ij} menunjukkan derajat keanggotaan U_j dalam himpunan fuzzy A_i . Nilai keanggotaan himpunan fuzzy A_i mengacu pada [3] yang terdiri dari 0, 0,5 atau 1. Penentuan derajat keanggotaan untuk setiap A_i dari $i=1,2,\dots,k$ yaitu jika keanggotaan maksimum dari suatu data dibawah A_i maka nilai fuzzifikasinya didefinisikan sebagai A_i .

Tabel 1. Hasil fuzzifikasi dan defuzzifikasi

Tahun	Data PDB	Hasil Fuzzifikasi	Hasil Defuzzifikasi
1978	21,967.4	A1	
1979	32,025.4	A1	28587.5
1980	45,445.7	A1	74112.5
1981	54,027.0	A2	58937.5
1982	59,362.6	A2	66525
1983	71,214.7	A2	58937.5
1984	77,622.8	A2	89287.5
1985	89,885.1	A3	96875
1986	102,682.6	A3	127225
1987	124,816.9	A4	119637.5
1988	142,104.8	A4	152516.6
1989	167,184.7	A5	203100
1990	195,597.2	A6	243566.6
1991	227,450.2	A7	233450
1992	259,884.5	A7	354850
1993	329,775.8	A9	339675
1994	382,219.7	A9	415550
1995	454,514.1	A10	536950
1996	532,568.0	A11	597650
1997	627,695.5	A12	28587.5

Tabel 2. Hasil Partisi

Interval baru	Umin	Umax	Jumlah data
U ₁₁	21000	51350	3
U ₁₂	51350	81700	4
U ₁₃	81700	112050	2
U ₁₄	112050	142400	2
U ₂₁	142400.0	182866.7	1
U ₂₂	182866.7	223333.3	1
U ₂₃	223333.3	263800.0	2
U ₃₁	263800.0	324500.0	0
U ₃₂	324500.0	385200.0	2
U ₄	385200.0	506600.0	1
U ₅₁	506600.0	567300.0	1
U ₅₂	567300.0	628000.0	1

Berdasarkan hasil partisi baru pada Tabel 2 dapat dibentuk himpunan fuzzy dengan perluasan variabel linguistik sesuai definisi *linguistic hedges* sebagai berikut: A₁=sangat sangat rendah, A₂=sangat sangat rendah, A₃=sangat rendah, A₄= rendah, A₅=sangat sangat menengah bawah, A₆=sangat menengah bawah, A₇=menengah bawah, A₈=sangat menengah, A₉=menengah, A₁₀=menengah atas, A₁₁=tinggi A₁₂=sangat tinggi. Berdasarkan acuan Definisi 1 dan Persamaan 7, nilai derajat keanggotaan antara 0 sampai 1 dapat dibentuk himpunan fuzzy sebagai berikut:

$$\begin{aligned}
 A_1 &= \frac{1}{U_{11}} + \frac{0.5}{U_{12}} + \frac{0}{U_{13}} + \frac{0}{U_{14}} + \frac{0}{U_{21}} + \frac{0}{U_{22}} + \frac{0}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_2 &= \frac{0.5}{U_{11}} + \frac{1}{U_{12}} + \frac{0.5}{U_{13}} + \frac{0}{U_{14}} + \frac{0}{U_{21}} + \frac{0}{U_{22}} + \frac{0}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_3 &= \frac{0}{U_{11}} + \frac{0.5}{U_{12}} + \frac{1}{U_{13}} + \frac{0.5}{U_{14}} + \frac{0}{U_{21}} + \frac{0}{U_{22}} + \frac{0}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_4 &= \frac{0}{U_{11}} + \frac{0}{U_{12}} + \frac{0.5}{U_{13}} + \frac{1}{U_{14}} + \frac{0.5}{U_{21}} + \frac{0}{U_{22}} + \frac{0}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_5 &= \frac{0}{U_{11}} + \frac{0}{U_{12}} + \frac{0}{U_{13}} + \frac{0.5}{U_{14}} + \frac{1}{U_{21}} + \frac{0.5}{U_{22}} + \frac{0}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_6 &= \frac{0}{U_{11}} + \frac{0}{U_{12}} + \frac{0}{U_{13}} + \frac{0}{U_{14}} + \frac{0.5}{U_{21}} + \frac{1}{U_{22}} + \frac{0.5}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_7 &= \frac{0}{U_{11}} + \frac{0}{U_{12}} + \frac{0}{U_{13}} + \frac{0}{U_{14}} + \frac{0}{U_{21}} + \frac{0.5}{U_{22}} + \frac{1}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_8 &= \frac{0}{U_{11}} + \frac{0}{U_{12}} + \frac{0}{U_{13}} + \frac{0}{U_{14}} + \frac{0}{U_{21}} + \frac{0}{U_{22}} + \frac{0.5}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_9 &= \frac{0}{U_{11}} + \frac{0}{U_{12}} + \frac{0}{U_{13}} + \frac{0}{U_{14}} + \frac{0}{U_{21}} + \frac{0}{U_{22}} + \frac{0}{U_{23}} \\
 &\quad + \frac{0.5}{U_{31}} + \frac{1}{U_{32}} + \frac{0.5}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_{10} &= \frac{0}{U_{11}} + \frac{0}{U_{12}} + \frac{0}{U_{13}} + \frac{0}{U_{14}} + \frac{0}{U_{21}} + \frac{0}{U_{22}} + \frac{0}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_{11} &= \frac{0}{U_{11}} + \frac{0}{U_{12}} + \frac{0}{U_{13}} + \frac{0}{U_{14}} + \frac{0}{U_{21}} + \frac{0.5}{U_{22}} + \frac{0.5}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0}{U_{51}} + \frac{0}{U_{52}} \\
 A_{12} &= \frac{0}{U_{11}} + \frac{0}{U_{12}} + \frac{0}{U_{13}} + \frac{0}{U_{14}} + \frac{0}{U_{21}} + \frac{0}{U_{22}} + \frac{0}{U_{23}} \\
 &\quad + \frac{0}{U_{31}} + \frac{0}{U_{32}} + \frac{0}{U_4} + \frac{0.5}{U_{51}} + \frac{1}{U_{52}}
 \end{aligned}$$

6. Menentukan nilai keanggotaan setiap data runtun waktu.

Perhitungan menggunakan himpunan fuzzy pada langkah 5, dapat ditentukan nilai keanggotaan setiap data berdasarkan fungsi keanggotaan segitiga. Misalnya data x = 21967,4 maka perhitungan nilai keanggotaan

pada setiap himpunan fuzzy adalah: Nilai x akan diinputkan pada setiap interval U_i , dari $i=1,2,\dots, m$ dengan batas bawah a dan batas atas c . Jika nilai x jatuh pada interval antara batas bawah dan nilai tengah b maka nilai keanggotaan = $\frac{x-a}{b-a}$

Nilai x ada pada interval $U_{11} = [21000 \ 51350]$ maka nilai keanggotaan x adalah $=(21967,4 - 21000)/(36175- 21000) = 0,063$. Selanjutnya dengan prosedur yang sama dihitung nilai keanggotaan U_{12} , U_{13} , sampai U_{52} . Perhitungan yang sama dilakukan untuk himpunan fuzzy A_2 sampai A_{12} . Nilai keanggotaan terbesar adalah pada A_1 .

7. Fuzzifikasi berdasarkan data runtun waktu

Nilai fuzzifikasi diambil berdasarkan nilai keanggotaan terbesar pada himpunan fuzzy. Hasil nilai keanggotaan x pada langkah 6 jatuh pada A_1 , sehingga dikatakan nilai x termasuk dalam variabel "sangat sangat rendah". Hasil fuzzifikasi keseluruhan ditunjukkan pada Tabel 1.

8. Menentukan hubungan logis fuzzy

Berdasarkan nilai fuzzifikasi, dapat dibentuk aturan fuzzy berbasis komposisi berdasarkan data input-output langkah 2.

Hasil *Fuzzy Logical Relationship* (FLR) berdasarkan nilai fuzzifikasi pada Tabel 1 sebagai berikut:

$A_1 \rightarrow A_1$, $A_1 \rightarrow A_1$, $A_1 \rightarrow A_2$, $A_2 \rightarrow A_2$, $A_2 \rightarrow A_2$, $A_2 \rightarrow A_2$, $A_2 \rightarrow A_3$, $A_3 \rightarrow A_3$, $A_3 \rightarrow A_4$, $A_4 \rightarrow A_4$, $A_4 \rightarrow A_5$, $A_5 \rightarrow A_6$, $A_6 \rightarrow A_7$, $A_7 \rightarrow A_7$, $A_7 \rightarrow A_9$, $A_9 \rightarrow A_9$, $A_9 \rightarrow A_{10}$, $A_{10} \rightarrow A_{11}$, $A_{11} \rightarrow A_{12}$.

9. Bagi partisi baru masing-masing menjadi 4 sub partisi

Partisi baru pada Tabel 2 dibagi masing-masing menjadi 4 sub bagian dengan panjang interval yang sama. Dari ke-4 titik interval yang dibentuk, titik interval 0,25 digunakan sebagai titik peramalan ke bawah dan titik interval 0,75 digunakan sebagai titik peramalan ke atas.

10. Bentuk aturan Markov chain

Aturan Markov chain digunakan untuk menentukan nilai peramalan dengan data inputan adalah FLR *current* A_i dan FLR *next* A_j didefinisikan sebagai $A_i \rightarrow A_j$.

Inputan algoritma pembentukan aturan Markov chain terdiri dari:

- X_n = data peramalan ke- n sebanyak m data
- U_{ab} = partisi baru himpunan universal
Indeks a adalah jumlah partisi baru yang ditentukan berdasarkan nilai fuzzifikasi A_j dengan nilai keanggotaan terbesar. Indeks b adalah titik-titik dari masing-masing sub partisi terdiri dari 4 titik partisi yaitu titik awal: nilai terkecil interval, titik ke-1: $\frac{1}{4}$

panjang interval, titik tengah: $\frac{1}{2}$ panjang interval, titik ke-3: $\frac{3}{4}$ panjang interval, titik ke-4 adalah nilai terbesar partisi.

- Fuzzy logical relationship* (FLR) $A_i \rightarrow A_j$, di mana A_i adalah hasil fuzzifikasi data tahun $(n-1)$ dan A_j hasil fuzzifikasi dari tahun ke- n . Nilai i, j adalah jumlah himpunan fuzzy baru.

Defuzzifikasi berdasarkan FLR dan aturan Markov chain

Sebagai contoh dihitung nilai defuzzifikasi dengan aturan 2 atau 3 untuk nilai $n=3$. Digunakan FLR $A_1 \rightarrow A_2$ dengan data $(n-1)$ adalah data tahun 1980 dan data n adalah tahun 1981. Data $n = 54027,0$, data $(n-1) = 45445,7$, data $(n-2) = 32025,4$, data $(n-3) = 21967,4$. Berdasarkan algoritma nilai variabel selisih didapat dari $(X_{(n-1)} - X_{(n-2)}) - (X_{(n-2)} - X_{(n-3)})$. Hasil perhitungan selisih adalah 3362.3 adalah positif dan $j > i$ maka digunakan aturan 2 untuk menghitung nilai peramalan. Berdasarkan kondisi pada aturan 2 dihitung nilai $|\text{selisih}| \times 2 + X_{(n-1)}$ menghasilkan 52170.3. Nilai $X_{(n-1)} - |\text{selisih}| \times 2$ menghasilkan 38721.1, nilai $|\text{selisih}| / 2 + X_{(n-1)}$ menghasilkan 47126.9, dan nilai $X_{(n-1)} - |\text{selisih}| / 2$ menghasilkan 43764.55. Dari ke-4 nilai yang masuk dalam interval subpartisi yang bersesuaian dengan nilai fuzzifikasi A_j yaitu partisi $U_{12} = [51350 \ 81700]$ adalah 52170.3. Sesuai algoritma aturan 2, perhitungan masuk pada kondisi pertama dan nilai peramalan jatuh pada titik interval U_{a3} yaitu pada subpartisi $a= 2$ pada titik ke-3 atau U_{23} yaitu = 74112.5. Perhitungan yang sama dilakukan pada setiap FLR yang dibentuk, dapat dihitung nilai defuzzifikasi berdasarkan aturan Markov chain dan didapat hasil defuzzifikasi seperti pada Tabel 1.

Tabel 3. Data PDB, aturan fuzzy, FLRG, defuzzifikasi

Data PDB	Aturan fuzzy	FLRG	Hasil Defuzzifikasi
21967,4	A_1	A_1, A_2	42662,9
32025,4	A_1	A_1, A_2	56883,8
45445,7	A_1	A_1, A_2	56883,8
54027,0	A_2	A_1, A_2, A_3	58797,7
59362,6	A_2	A_1, A_2, A_3	58797,7
71214,7	A_2	A_1, A_2, A_3	58797,7
77622,8	A_2	A_1, A_2, A_3	58797,7
89885,1	A_3	A_2, A_3, A_4	91875,5
102682,6	A_3	A_2, A_3, A_4	91875,5
124816,9	A_4	A_3, A_4, A_5	124256,2
142104,8	A_4	A_3, A_4, A_5	124256,2
167184,7	A_5	A_4, A_5, A_6	159480,9
195597,2	A_6	A_5, A_6, A_7	198987,0

227450,2	A ₇	A ₆ ,A ₇ ,A ₉	241916,8
259884,5	A ₇	A ₆ ,A ₇ ,A ₉	241916,8
329775,8	A ₉	A ₇ ,A ₉ ,A ₁₀	354658,2
382219,7	A ₉	A ₇ ,A ₉ ,A ₁₀	354658,2
454514,1	A ₁₀	A ₉ ,A ₁₀ ,A ₁₁	436406,2
532568,0	A ₁₁	A ₁₀ ,A ₁₁ ,A ₁₂	523517,9
627695,5	A ₁₂	A ₁₁ ,A ₁₂	575947,2

PEMBAHASAN

Sistem peramalan diimplementasikan dengan program *Matrix Laboratory* (MatLab) 7.6.0 (R2008a) berbasis GUI. Pengujian dilakukan dengan mengambil 6 sampel data sebagai data latih yang terdiri dari:

1. Data latih 10 dari 1967-1976 s/d 2000-2009 sebanyak 34
2. Data latih 20 dari 1967-1986 s/d 1990-2009 sebanyak 24
3. Data latih 30 dari 1967-1996 s/d 1980-2009 sebanyak 14
4. Data latih 35 dari 1967-2001 s/d 1975-2009 sebanyak 9
5. Data latih 40 dari 1967-2006 s/d 1970-2009 sebanyak 4
6. Data latih 43 yaitu keseluruhan data dari 1967-2009.

Data latih terbaik setiap metode yang akan diuji dengan jumlah himpunan fuzzy berbeda beda dari 5,7,15,20, dan 25. Pengujian dilakukan untuk melihat kehandalan peramalan dengan MSE (*Mean Square Error*) dan akurasi peramalan dengan MAPE (*Mean Average Percentage Error*) yang dirumuskan (Tsaur, 2012) sebagai berikut:

$$MSE = \frac{\sum_{t=1}^n (X_t - \hat{X}_t)^2}{n} \quad (5)$$

dimana n menyatakan jumlah data, X_t adalah nilai data aktual dan \hat{X}_t adalah nilai ramalannya.

$$MAPE = \frac{\sum_{t=1}^n \left| \frac{e_t}{\hat{X}_t} \right|}{n} \times 100 \quad (6)$$

dimana n adalah jumlah data dan e_t adalah nilai kesalahan ramalan yang diperoleh dari $X_t - \hat{X}_t$, X_t adalah nilai data aktual dan \hat{X}_t adalah nilai ramalannya.

Berdasarkan hasil MSE dan MAPE terbaik pada Tabel 5 persentasi peningkatan MSE tertinggi 99,92% dan persentasi peningkatan MAPE tertinggi 99,63%. Dapat disimpulkan secara rata-rata metode modifikasi (Melike Sah-Konstain) lebih handal dan akurat dari metode yang lain. Peningkatan performansi dari metode Melike Sah-Konstain ke modifikasi Melike Sah-Konstain berdasarkan nilai MSE dan MAPE terlihat ada peningkatan

kehandalan 42,04% dan akurasi 34,99%. Hal ini menunjukkan algoritma modifikasi Melike Sah-Konstain terbukti handal dan akurat untuk peramalan data PDB Indonesia.

Berdasarkan hasil MSE dan MAPE rata-rata pada Tabel 5, persentasi peningkatan MSE tertinggi 99,60% dan persentasi peningkatan MAPE tertinggi 98,94%. Dapat disimpulkan secara rata-rata metode Markov chain lebih handal dan lebih akurat dari metode yang lain. Peningkatan performansi dari metode Melike Sah dan Konstain ke Markov chain berdasarkan nilai nilai rata-rata MSE dan MAPE terlihat ada peningkatan kehandalan 29,79% dan akurasi 17,41%. Hal ini menunjukkan fuzzy time series Markov chain terbukti handal dan akurat untuk peramalan data PDB Indonesia.

Tabel 5. Perbandingan peningkatan MSE dan MAPE

Perbandingan Metode	Peningkatan kehandalan (%)		Peningkatan akurasi (%)	
	Rerata MSE terbaik	Rerata MSE rata-rata	Rerata MAPE terbaik	Rerata MAPE rata-rata
Melike Sah dan Konstain	45.57	67.15	68.13	74.36
J Sullivan dan Woodall	99.85	99.59	99.14	98.94
Markov chain	99.92	99.60	99.63	98.86

Tabel 6 Perbandingan MSE untuk himpunan fuzzy berbeda

Metode Peramalan	Perbedaan fuzzy set				
	5	10	15	20	25
Melike Sah dan Konstain	4.96E-03	8.35E-03	7.85E-03	5.64E-03	6.66E-03
J Sullivan dan Woodall	2.88E-03	5.33E-03	0.400E-03	4.44E-03	4.42E-03
Markov chain	6E-06	2.9E-05	6.57E-04	6.8E-05	4.99E-03

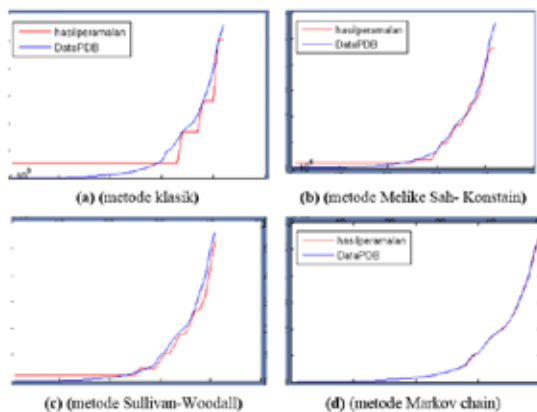
Hasil pengujian berdasarkan jumlah himpunan fuzzy yang berbeda-beda yakni: 5,7,15,20,25 untuk rata-rata MSE terbaik setiap metode. Berdasarkan Tabel 6 metode yang paling handal adalah Markov chain dengan nilai MSE ternormalisasi= 6.10^{-6} dan jumlah himpunan fuzzy 5.

Hasil pengujian berdasarkan jumlah himpunan fuzzy yang berbeda-beda untuk mendapatkan rata-rata MAPE terbaik setiap metode. Berdasarkan Tabel 7 metode yang akurasi paling tinggi adalah metode Markov chain dengan nilai MAPE=0,60% pada jumlah himpunan fuzzy 5.

Tabel 7 Perbandingan MAPE untuk himpunan fuzzy berbeda

Metode Peramalan	Perbedaan <i>fuzzy set</i>				
	5	10	15	20	25
Melike Sah&Konstain	231.16	164.63	74.96	56.72	52.86
J Sullivan& Woodall	0.90	0.86	1.59	1.85	2.13
Markov chain	0.60	1.41	2.54	2.49	2.34

Salah satu contoh output grafik perbandingan hasil peramalan dan data runtun waktu PDB Indonesia dengan metode fuzzy time series Markov chain ditunjukkan pada Gambar 1. Secara grafis dapat ditunjukkan metode peramalan yang memiliki akurasi tinggi pada Gambar 5. Perbandingan hasil peramalan setiap metode terlihat pada garis biru adalah data aktual dan garis merah adalah data hasil peramalan. Secara grafis terlihat metode klasik Gambar 5.a tidak mampu meramalkan perubahan data PDB, hal ini terlihat pada garis merah mendatar yang menunjukkan tetapnya data hasil peramalan. Metode Melike Sah-Konstain dan J. Sullivan-Woodall masih menunjukkan garis mendatar yang tetap namun lebih baik dari metode klasik., sedangkan pada metode Markov chain sudah terlihat garis merah mengikuti pola garis biru yang menunjukkan akurasi peramalan tinggi.



Gambar 1. Hasil perbandingan peramalan

KESIMPULAN

Metode fuzzy time series klasik, Melike Sah dan Konstain, J. Sullivan Woodal, dan Markov chain dapat digunakan untuk meramalkan data PDB Indonesia. Metode pendekatan baru terbukti memberikan peningkatan performansi peramalan dengan tingkat kehandalan dan akurasi yang lebih tinggi dibandingkan dengan fuzzy time series klasik. Secara rata-rata metode modifikasi pendekatan baru Markov chain lebih handal

99,92 % jika dihitung dengan nilai MSE dan tingkat akurasi 99.63 % jika dihitung dengan MAPE.

Berdasarkan uji perbedaaan terhadap jumlah himpunan fuzzy, metode Markov chain lebih handal dan akurasi tinggi pada jumlah himpunan fuzzy sebanyak 5.

DAFTAR PUSTAKA

- Berutu, S.S 2013. Peramalan Penjualan dengan Metode Fuzzy Time Series Ruey Chyn Tsaur, *Thesis*, Program Studi Magister Sistem Informasi, Universitas Diponegoro, Semarang.
- Chen, S. 1996. Forecasting Enrollment Based on Fuzzy Time Series. *Fuzzy sets and systems*, 81(3): 311-319.
- Cheng, C. H., Cheng, G. W., dan Wang, J. W. 2008. Multi-attribute Fuzzy Time Series Method Based on Fuzzy Clustering. *Expert systems with applications*, 34(2), pp.1235-1242.
- Li, S. T. dan Cheng, Y. C. 2007. Deterministic Fuzzy Time Series Model For Forecasting Enrollment. *Computers and mathematics with application*, 53(12): 1904-1920.
- Melike, S. dan Degtiarev, K. Y. 2005. Forecasting Enrollment Model Based on First Order Fuzzy Time Series. *Proceedings of world academy of science, engineering and technology*, 1: 132-135.
- Rosadi, D., 2006, *Pengantar Analisa Runtun Waktu*, <http://dedirosadi.staff.uqm.ac.id>, diakses tanggal 18 November 2015.
- Singh, S. R. 2007. A Simple Method of Forecasting Based on Fuzzy Time Series. *Applied mathematic and computation*, 186(1): 330-339.
- Song, Q. dan Chissom, B. S. 1993. Forecasting Enrollment With Fuzzy Time Series- Part I. *Fuzzy sets and systems*, 54(1): 1-9.
- Stevenson, M. dan Porter, J. E. 2009. Fuzzy Time Series Forecasting Using Percentage Change As the Universe of Discourse. *World academy of science, engineering and technology*, 55: 154-157.
- Sullivan, J. and W.H. Woodall, (1994), A comparison of fuzzy forecasting and Markov modeling, *Fuzzy Sets and Systems*. 64 279–293

- Tsaur, R. C. 2012. A Fuzzy Time Series-Markov Chain Model With an Application to Forecast the Exchange Rate Between the Taiwan and US Dollar. *International journal of innovative computing, information and control*, 8(7B): 4931-4942.
- Tsaur, R. C., Yang, J.C.O., dan Wang, H. F. 2005. Fuzzy relation Analysis in Fuzzy Time Series Model. *Computer and mathematics with applications*, 49(4): 539-548.
- Yu, H. 2005. Weighted Fuzzy Time Series Models For Taiex Forecasting. *Physica A*, 349: 609-624.