

Adaptive Neuro Fuzzy Inference System Optimization by Genetic Algorithm pada Time Series

**Agnes Veronika Sinaga¹, Noel Christoper Biutarbutar², Theodora Beata Simamora³,
Junita Amalia⁴**

^{1,2,3,4} Sistem Informasi, Fakultas Informatika dan Teknik Elektro, Institut Teknologi Del
email: junita.amalia@del.ac.id

Abstract: Time series is a collection of observations obtained through repeated measurements over time. Forecasting is one of the applications of time series data which to predict the possibilities that what will be happen in the future based on pre-existing data. One of methods that can be used for time series is Adaptive Neuro Fuzzy Inference System (ANFIS). However, ANFIS has limitations in selecting the hyperparameters. We proposed to use ANFIS and GA (Genetic Algorithm) for optimization algorithm to solve the limitations. The data that will be used in this research is data on electricity consumption. This research aims to determine how the performance of the GA algorithm to optimize ANFIS with the RMSE value as a reference. After conducting several experiments on ANFIS, the researcher got the minimum RMSE result of 0.2323 for the test data using European high electric consumption data. In the ANFIS-GA experiment, the smallest RMSE for the test data is 0,2018; with population = 100, mutation = 0.01 and crossover = 0.5 using Siberian low electricity consumption data.

Keywords: Adaptive Neuro Fuzzy Inference System, Genetic Algorithm, Forecasting

Abstrak: Data deret waktu adalah serangkaian pengamatan yang diambil secara berurutan dari waktu ke waktu. Salah satu kegunaan data *time series* adalah untuk *forecasting*, yaitu memprediksi kemungkinan yang akan terjadi di masa yang akan datang berdasarkan data masa lalu. Salah satu metode yang dapat digunakan untuk *time series forecasting* adalah *Adaptive Neuro Fuzzy Inference System* (ANFIS). Namun, ANFIS memiliki keterbatasan dalam memilih hiperparameternya. Keterbatasan ini diatasi dengan optimasi *Genetic Algorithm* (GA) sehingga penulis mengajukan ANFIS dan dioptimasi dengan GA. Data yang akan digunakan dalam penelitian ini adalah data konsumsi pemakaian listrik. Penelitian ini bertujuan untuk mengetahui bagaimana performa algoritma GA untuk mengoptimasi ANFIS dengan nilai RMSE sebagai acuannya. Setelah dilakukan empat eksperimen pada ANFIS, peneliti mendapatkan hasil RMSE minimum yaitu 0,2323 untuk data *test* dengan menggunakan data *electricity consumption European high*. Untuk eksperimen ANFIS-GA, dilakukan juga empat eksperimen dengan hasil RMSE terkecil dari data *test* 0,2018; dengan populasi = 100, mutasi = 0,01 dan crossover = 0,5 dengan menggunakan data *electricity consumption Siberian low*.

Kata kunci: Adaptive Neuro Fuzzy Inference System, Genetic Algorithm, Forecasting

Pendahuluan

Data deret waktu atau lebih dikenal dengan data *time series* adalah serangkaian pengamatan yang diambil secara berurutan dari waktu ke waktu (Pole, West, & Harrison, Applied Bayesian Forecasting and Time Series Analysis, 2017). Ada banyak contoh penerapan dari data *time series*, salah satunya adalah memprediksi kemungkinan yang terjadi di masa yang akan datang. *Time series forecasting* digunakan untuk

memprediksi data di masa yang akan datang berdasarkan data yang sudah ada. Pengaplikasian *time series forecasting* sangat luas seperti dalam bidang keuangan, prediksi biaya, dan bidang lainnya (Moon & Shin, 2018).

Pada kasus *time series forecasting*, salah satu metode yang dapat digunakan adalah dengan menggunakan *Adaptive Neuro-Fuzzy Inference System* (ANFIS). ANFIS terdiri dari *neural network* dan *fuzzy logic*

(Al-Qaness, Ewess, Fan, & Elaziz, 2020). Selama beberapa dekade terakhir metode ini sudah digunakan dalam beberapa penelitian seperti memperkirakan harga saham, harga penggunaan listrik, dan pemakaian listrik. Namun, ANFIS memiliki keterbatasan yaitu memilih hiperparameter secara manual. Banyak pendekatan yang diterapkan untuk pemilihan hiperparameter tetapi secara umum memakan sumber daya komputasi yang banyak dan memerlukan analisis stabilitas yang kompleks. (Vlasenko, Vlasenko, Vynokurova, Bodyanskiy, & Peleshko, 2019). Penelitian yang berjudul *Investigating the performance of ANFIS model to predict the hourly temperature in Pattani, Thailand* menyatakan ANFIS dapat dioptimalkan kinerjanya dengan tuning parameter dalam fungsi keanggotaannya (Septiarini & Musikasuwana, 2018).

Metode metaheuristik telah digunakan untuk meningkatkan kualitas ANFIS. Sebagai contoh, *hybrid* dari ANFIS dan *Sinus-Cosinus Algorithm* (SCA) untuk perkiraan konsumsi minyak. *Particle Swarm Optimization* (PSO) juga diterapkan dengan ANFIS untuk peramalan hasil *biochar*, harga listrik, dan peramalan keuangan. Algoritma metaheuristik lainnya juga diterapkan untuk meningkatkan kinerja ANFIS, seperti *Multi-Verse Optimizer* (MVO), *Salp Swarm Algorithm* (SSA), *Social-spider Optimization* dan *Genetic Algorithm* (GA) (Al-Qaness, Ewess, Fan, & Elaziz, 2020).

Penelitian dengan judul *Prediction Of Plain Needs Old Telephone Service (POTS) using ANFIS-GA.*, implementasi *Service Government Division*, menggunakan metode ANFIS-GA dapat mengoptimalkan tingkat perkiraan kebutuhan layanan POTS di setiap tahun dengan MAPE - 0,152989097 dibandingkan dengan metode ANFIS (Setiawan & Girsang, 2019). Penelitian lain berjudul Analisis

Kinerja *Adaptive Neuro Fuzzy Inferences System* pada *Feature input* dengan *Tournament Selection* Menggunakan *Algoritma Genetika* menunjukkan terjadinya peningkatan kinerja ANFIS dan mendapatkan akurasi yang lebih tinggi (Siahaan, 2019). Dari penelitian yang berjudul diketahui bahwa ANFIS dan GA menghasilkan *error* yang lebih kecil jika dibandingkan dengan ANFIS tanpa optimasi (Wahyuni, 2017).

Penelitian dengan judul *A Comparative Analysis of Hidden Markov Model, Hybrid Support Vector Machines, and Hybrid Artificial Neural Fuzzy Inference System in Reservoir Inflow Forecasting (Case Study: The King Fahd Dam, Saudi Arabia)* memaparkan RMSE ANFIS-GA lebih kecil dibandingkan ANFIS artinya ANFIS-GA (Alquraish, Abuhasel, Alqahtani, & Khadr, 2021). ANFIS-GA sebagai model yang paling sesuai karena memiliki nilai RMSE terkecil juga diungkapkan dalam penelitian yang berjudul *Application of Genetic Algorithm to predict maximum of air temperature (Case study: Isfahan city)* (Nia, Azad, Farzin, & Karami, 2018). Penambahan GA pada ANFIS memberikan dampak yang baik dengan semakin memperkecil nilai RMSE (Hameed, 2018).

Berdasarkan pemaparan diatas, akan dilakukan penelitian yang berjudul *Analysis of Adaptive Neuro Fuzzy Inference System (ANFIS) Optimization by Genetic Algorithm (GA) in Time Series*. Dengan menganalisis parameter yang terdapat pada metode ANFIS dan juga pembagian data training dan data testing dari metode ANFIS dan ANFIS-GA, diharapkan dapat menentukan model yang optimal untuk *time series forecasting*.

Metode

Pada penelitian ini, penulis menggunakan metode ANFIS dan

mengoptimasi metode tersebut dengan GA untuk mengetahui bagaimana GA mengoptimasi dan seberapa baik hasil optimasinya dengan RMSE sebagai acuan. Data yang digunakan yaitu data Russian Wholesale Electricity Market yang bersumber dari Kaggle. Adapun dataset yang digunakan adalah Russian Wholesale Electricity Market. Dataset yang digunakan merupakan record data dari penggunaan arus listrik mulai September 2006 sampai November 2011.

Dataset Russian Wholesale Electricity Market memiliki empat puluh lima ribu delapan ratus lima belas rows dan lima atribut yaitu *timestep*, *consumption_eur*, *consumption_sib*, *price_eur* dan *price_sib*. Atribut yang digunakan ada tiga yaitu *timestep*, *consumption_eur*, *consumption_sib*. Berikut sampel data yang digunakan pada setiap atribut.

Tabel 1. Sampel Data

Timestep	Consumpti on_eur	Consumpti on_sib
9 /1/2006 2:00	59685	17161
/1/2006 3:00	59500	17061
/1/2006 4:00	60617	17074
/1/2006 5:00	62262	17074

Data *preprocessing* yang dilakukan adalah data *cleaning* dan data *transformation*. Melakukan data *cleaning* bertujuan untuk mengisi dataset yang bernilai 0 dan Null dengan nilai *mean* dari atribut tersebut. Peneliti menggunakan nilai *mean* untuk melakukan data *cleaning* agar nilai baru yang akan di-input ke dalam dataset tidak melebihi *range* nilai dari dalam data yang sudah ada. Berdasarkan hasil statistika deskriptif yang dilakukan, diketahui bahwa setiap atribut pada dataset tidak berdistribusi secara normal.

Ditemukan adanya data yang memiliki rentang nilai yang sangat tinggi pada dataset. Untuk mengatasi hal ini, maka dilakukan data *transformation* berupa normalisasi pada dataset dengan metode *minmax normalization* (Sangaiah, 2019).

Sebelum memasuki tahapan ANFIS, dataset yang akan digunakan dalam penelitian ini akan dilakukan pembagian data. Pembagian dataset ini dilakukan dengan tujuan untuk mengetahui kinerja algoritma akan lebih baik pada data yang bernilai rendah atau data yang bernilai tinggi. Selain itu, agar model yang dihasilkan lebih tepat, yakni abstraksi dari pola data. Pembagian dataset ini dilakukan dengan membagi dataset menjadi 4 bagian berdasarkan tinggi dan rendahnya nilai dari kedua atribut yang akan digunakan sebagai target yaitu *consumption_eur* dan *consumption_sib*. Pembagian dataset ini dilakukan dengan cara mencari nilai median dari setiap atribut dan menjadikan median sebagai batasan untuk nilai low dan high. Jika nilai lebih kecil dari median, maka akan termasuk ke dalam *electricity consumption* yang tergolong rendah. Dan jika nilai lebih besar dan sama dengan median, maka akan termasuk ke dalam *electricity consumption* yang tergolong tinggi.

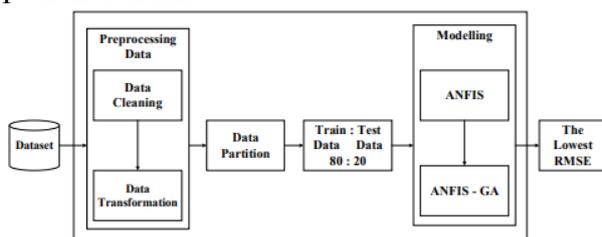
Prosedur untuk membangun ANFIS adalah mengumpulkan data, melakukan pra-pemrosesan data, memilih input, menentukan fungsi keanggotaan, menghasilkan aturan fuzzy, melatih sistem inferensi fuzzy, dan membuat prakiraan. Dalam penelitian ini, nilai pengamatan telah dipisahkan menjadi dua kelompok yaitu 80% data *training* dan 20% data *testing*. Rasio 80%:20% paling umum digunakan (Joseph, 2022). Di samping itu, dibutuhkan juga data *testing* yang berfungsi untuk menguji akurasi ANFIS yang sudah dilatih dengan menggunakan data *training* sebelumnya. Keakuratan model ANFIS

dipengaruhi oleh banyak faktor seperti jumlah fungsi keanggotaan, jenis fungsi keanggotaan, pemilihan input, jumlah iterasi, dan jenis output.

GA digunakan untuk mengoptimasi metode ANFIS dikarenakan GA merupakan salah satu bagian *evolutionary algorithm* yaitu suatu algoritma yang meniru proses evolusi alami dimana konsep utamanya adalah individu yang paling unggul akan bertahan hidup, sedangkan individu yang lemah akan punah. Keunggulan individu diuji dengan *fitness function* dengan menggunakan *fitness value* yang telah diperoleh. *Fitness value* pada GA merupakan nilai kelayakan suatu solusi terhadap suatu permasalahan. *Fitness function* menghasilkan *fitness value* baru yang akan dijadikan referensi untuk proses GA selanjutnya. Parameter utama yang menjadi karakteristik dari GA adalah *population size*, *crossover chance* dan *mutation chance* (Janes, Gordan, & Mladen, 2017). Pemilihan nilai yang tepat dari masing-masing parameter dapat memberikan hasil optimasi yang lebih baik.

Evaluasi yang digunakan adalah *Root Mean Square Error* (RMSE) (Osowski & Siwek, 2019). RMSE memiliki standar eror yang paling dekat dengan standar deviasi, yang mana standar deviasi adalah parameter utama pada kasus analisis statistik, sehingga RMSE adalah teknik yang paling sering digunakan jika proses data mining terkait dengan statistic.

Gambar 1 merupakan rancangan penelitian yang dilakukan pada penelitian ini.



Gambar 1. Rancangan Penelitian

Hasil

Dataset dibagi menjadi 4 bagian yaitu *consumption_eur_high*, *consumption_eur_low*, *consumption_sib_high*, *consumption_sib_low*. Setiap dataset tersebut dilakukan analisis terhadap RMSE menggunakan metode ANFIS dan ANFIS-GA. Pada ANFIS GA, dilakukan eksperimen terhadap *population*, *crossover*, dan *mutation*. Eksperimen pada ANFIS-GA yaitu:

1. Eksperimen 1 dengan nilai *population_size* = 20, *mutation_chance* = 0,02, *crossover_chance* = 0,7.
2. Eksperimen 2 dengan *population_size* = 20, *mutation_chance* = 0,01, *crossover_chance* = 0,5.
3. Eksperimen 3 dengan *population_size* = 100, *mutation_chance* = 0,02, *crossover_chance* = 0,7.
4. Eksperimen 4 dengan nilai *population_size* = 100, *mutation_chance* = 0,01, *crossover_chance* = 0,5.

Perbandingan RMSE ANFIS dan ANFIS-GA sebagai berikut:

Tabel 1. RMSE ANFIS

RMSE ANFIS		
Dataset	Train	Test
cons_eur_low	0,2419	0,2324
cons_eur_high	0,2421	0,2323
cons_sib_low	0,2551	0,2404
cons_sib_high	0,2552	0,2403

Tabel 2. RMSE ANFIS-GA Eksperimen 1

RMSE ANFIS-GA Eks 1		
Dataset	Train	Test
cons_eur_low	0,2419	0,2324
cons_eur_high	0,2419	0,2323
cons_sib_low	0,2551	0,2402
cons_sib_high	0,2551	0,2403

Tabel 3. RMSE ANFIS-GA Eksperimen 2

RMSE ANFIS-GA Eks 2		
Dataset	Train	Test
cons_eur_low	0,2419	0,2323
cons_eur_high	0,2419	0,2323
cons_sib_low	0,2551	0,2402
cons_sib_high	0,2551	0,2403

Tabel 4. RMSE ANFIS-GA Eksperimen 3

RMSE ANFIS-GA Eks 3		
Dataset	Train	Test
cons_eur_low	0,2419	0,2323
cons_eur_high	0,2419	0,2419
cons_sib_low	0,2191	0,2191
cons_sib_high	0,2191	0,2191

Tabel 5. RMSE ANFIS-GA Eksperimen 4

RMSE ANFIS-GA Eks 4		
Dataset	Train	Test
cons_eur_low	0,2419	0,2323
cons_eur_high	0,2419	0,2323
cons_sib_low	0,2212	0,2218
cons_sib_high	0,2212	0,2025

Berdasarkan Tabel 1 sampai Tabel 5 diatas dapat disimpulkan:

1. Data *electricity consumption european low*.
 Nilai RMSE dari ANFIS sebesar 0,2419 untuk data *train* dan 0,2324 untuk data *test*. Nilai RMSE dari ANFIS-GA terkecil yaitu 0,2419 untuk data *train* dan 0,2323 untuk data *test*, dihasilkan dari eksperimen 2, 3 dan 4.
2. Data *electricity consumption European high*.
 Nilai RMSE dari ANFIS sebesar 0,2421 untuk data *train* dan 0,2323 untuk data *test*. Nilai RMSE dari ANFIS-GA terkecil yaitu 0,2419 untuk data *train* dan 0,2323 untuk data *test*, dihasilkan dari eksperimen 1, 2 dan 4
3. Data *electricity consumption Siberian low*.

Nilai RMSE dari ANFIS sebesar 0,2551 untuk data *train* dan 0,2404 untuk data *test*. Nilai RMSE dari ANFIS-GA terkecil yaitu 0,2212 untuk data *train* dan 0,2018 untuk data *test*, dihasilkan dari eksperimen 4.

4. Data *electricity consumption Siberian high*.

Nilai RMSE dari ANFIS sebesar 0,2552 untuk data *train* dan 0,2403 untuk data *test*. Nilai RMSE dari ANFIS-GA terkecil yaitu 0,2191 untuk data *train* dan 0,2191 untuk data *test*, dihasilkan dari eksperimen 3.

Pembahasan

Variabel yang digunakan untuk implementasi ANFIS dan ANFIS-GA adalah *timestep*, *consumption_eur*, *consumption_sib*.

Data *consumption_eur* dan *consumption_sib* dikelompokkan menjadi dua bagian yaitu *high* dan *low*. Kategori *high* adalah data yang nilainya besar sama dengan nilai median dan kategori *low* adalah data yang nilainya kecil dari nilai median. Pengelompokkan ini dilakukan untuk menghindari terjadinya bias karena interval yang yang besar. Selain itu, masing-masing kelompok data akan dibagi menjadi data *train* dan data *test* dengan perbandingan 80:20.

Berdasarkan hasil *training* pada ANFIS diketahui bahwa semakin besar *epoch* yang digunakan maka semakin kecil nilai RMSE. Pada ANFIS-GA nilai RMSE antar eksperimen parameter tidak jauh berbeda hal ini disebabkan oleh rentang nilai parameter yang digunakan tidak terlalu besar. Misalnya pada *mutation_chance*, nilai yang digunakan adalah 0,01 dan 0,02 sementara nilai *crossover_chance* juga tidak signifikan perbedaannya yaitu 0,5 dan 0,7. Untuk penelitian berikutnya dapat mencoba membuat rentang nilai yang lebih besar sehingga pengaruh nilai paramater pada ANFIS-GA dapat

dianalisis lebih lanjut. Namun dari hasil eksperimen sudah didapatkan kesimpulan bahwa nilai RMSE ANFIS-GA lebih rendah jika dibandingkan dengan nilai RMSE ANFIS.

Simpulan

Dari hasil eksperimen yang telah dilakukan dapat dilihat bahwa RMSE ANFIS-GA lebih rendah jika dibandingkan dengan nilai RMSE ANFIS. Artinya GA dapat mengoptimasi ANFIS dengan baik. GA digunakan untuk mengoptimasi parameter *membership function* pada ANFIS, yaitu *population*, *crossover* dan *mutation*. Berdasarkan eksperimen parameter pada ANFIS-GA, didapatkan nilai RMSE terkecil pada data test adalah 2,018 dengan nilai parameter *population_size*= 100, *mutation_chance* = 0,01, *crossover_chance* = 0,5.

Referensi

- Al-Qaness, M. A., Ewess, A. A., Fan, H., & Elaziz, M. A. (2020). Optimized Forecasting Method for Weekly Influenza Confirmed Cases. *Int J Environ Res Public Health*.
- Alquraish, M. M., Abuhasel, K. A., Alqahtani, A. S., & Khadr, M. (2021). A Comparative Analysis of Hidden Markov Model, Hybrid. *Water*, 1-18.
- Hameed, I. A. (2018). A GA-Based Adaptive Neuro-Fuzzy Controller for Greenhouse Climate Control System. *Alexandria Engineering Journal*, 773-779.
- Janes, G., Gordan, M., & Mladen, M. (2017). Applying Improved Genetic Algorithm for Solving Job Shop Scheduling Problems. *Tehnički vjesnik*, 1243-1247.
- Joseph, V. R. (2022). *Optimal Ratio for Data Splitting*. Atlanta: H. Milton Stewart School of Industrial and Systems Engineering
- Moon, T., & Shin, D. H. (2018). Forecasting Model of Construction Cost Index Based on VECM with Search Query. *KSCE Journal of Civil Engineering*.
- Nia, M. M., Azad, A., Farzin, S., & Karami, H. (2018). Application of Genetic Algorithm to Optimize the
- Osowski, S., & Siwek, K. (2019). Local dynamic integration of ensemble in prediction of time series. *Bulletin of The Polish Academy of Sciences*, 517-525.
- Pole, A., West, M., & Harrison, J. (2017). *Applied Bayesian Forecasting and Time Series Analysis*. New York: Chapman and Hall/CRC.
- Sangaiah, A. K. (2019). *Deep Learning and Parallel Computing Environment for Bioengineering Systems*. Academic Press.
- Septiarini, T. W., & Musikasawan, S. (2018). Investigating the performance of ANFIS model to predict the . *Journal of Physics: Conference Series*, 1-10.
- Setiawan, A., & Girsang, A. S. (2019). Prediction Of Plain Needs Old Telephone Service. *IJSTR*, 2649-2652.
- Siahaan, W. F. (2019). *Analisis Kinerja Adaptive Neuro Fuzzy Inferences System Pada Feature Input dengan Tournament Selection Menggunakan Algoritma Genetika*. Medan: Universitas Sumatera Utara.
- Vlasenko, A., Vlasenko, N., Vynokurova, O., Bodyanskiy, Y.,

& Peleshko, D. (2019). A Novel Ensemble Neuro-Fuzzy Model for Financial Time Series Forecasting. *MDPI*, 1-12.

Hybrid Adaptive Neuro Fuzzy Inference System (ANFIS) dan Algoritma Genetika. Malang: Universitas Brawijaya.

Wahyuni, I. (2017). *Prediksi Curah Hujan menggunakan Metode*