

Prediksi kualitas air Danau Toba menggunakan *Extreme Learning Machine*

Eric Suwarno, Romi Fadillah Rahmat, Maya Silvi Lydia

Fakultas Ilmu Komputer dan Teknologi Informasi

Universitas Sumatera Utara

Medan, Indonesia

E-mail: 121402071.es@gmail.com | romi.fadillah@usu.ac.id | maya2@usu.ac.id

Abstrak—Haro *et al.* [1] menyatakan bahwa telah terjadi pencemaran sumber daya air di permukaan Danau Toba, khususnya di wilayah Kecamatan Haranggaol Horison, Kabupaten Simalungun, Sumatera Utara. Pengukuran kualitas air dilakukan dengan mengambil sampel air dari beberapa lokasi di permukaan Danau Toba untuk menjalani pengujian di laboratorium. Hasil pengujian tersebut menjadi acuan dalam penilaian kualitas air Danau Toba. Dengan perkembangan teknologi komputer, pengukuran kualitas air dapat dilakukan melalui berbagai alat dan sensor, sehingga proses pengukuran menjadi lebih efisien dari segi biaya dan waktu. Dalam penelitian ini, prediksi kualitas air Danau Toba akan dilakukan menggunakan *extreme learning machine*. Data yang digunakan berasal dari hasil penelitian yang dilakukan oleh Rahmat *et al.* [2], yang menggunakan berbagai sensor pada Arduino dalam mengukur parameter kualitas air. Hasil pengujian menunjukkan bahwa secara umum, proses prediksi kualitas air menggunakan *extreme learning machine* dapat memberikan hasil dengan durasi antara () hingga () detik. *Training error* dan *testing error* terbaik yang dihasilkan dari pengujian diperoleh dengan menggunakan fungsi *hard-limit* sebagai fungsi aktivasi.

Kata kunci—prediksi kualitas air, *artificial neural network*, *extreme learning machine*

I. PENDAHULUAN

Penelitian yang dilakukan oleh Haro *et al.* [1] menunjukkan bahwa berdasarkan hasil pengukuran yang dilakukan di Kecamatan Haranggaol Horison, Kabupaten Simalungun, terjadi pencemaran air di perairan Danau Toba. Tingkat pencemaran air yang terukur di Danau Toba bervariasi dari ringan hingga sedang. Limbah yang dihasilkan dari aktivitas rumah tangga dan aktivitas industri, yang diikuti dengan perkembangan ekosistem sekitar, menjadi sumber utama dari pencemaran air di Danau Toba.

Pengukuran kualitas air yang dilakukan di Danau Toba dilakukan dengan melakukan pengambilan sampel air di beberapa lokasi di Danau Toba. Setiap sampel air yang dikumpulkan akan diuji melalui pemeriksaan laboratorium, sehingga dapat diketahui tingkat kualitas air yang terukur di setiap lokasi. Dengan metode pengukuran ini, terjadi kemungkinan peningkatan biaya dan waktu yang diperlukan untuk proses pengukuran. Karena itu, diperlukan sebuah metode

yang dapat melakukan prediksi kualitas air, dengan mempertimbangkan efisiensi waktu dan biaya yang dilakukan pada saat pengukuran.

Seiring dengan perkembangan teknologi komputer, berbagai perangkat seperti mikrokontroler dan sensor dapat digunakan untuk melakukan pengukuran kualitas air. Rahmat *et al.* [2] melakukan pengukuran kualitas air Danau Toba menggunakan Arduino yang dihubungkan dengan berbagai sensor-sensor, di mana hasil pengukuran akan dimutakhirkan setiap interval tertentu. Untuk mengolah data pengukuran kualitas air tersebut, sebuah metode diperlukan untuk melakukan proses prediksi kualitas air dengan mengutamakan akurasi prediksi dan waktu komputasi.

Kasabov [3] menyatakan bahwa *artificial neural network* adalah salah satu metode yang dapat digunakan untuk melakukan proses prediksi secara berkelanjutan. Namun, masalah yang dihadapi dalam implementasi *artificial neural network* adalah waktu yang digunakan dalam pemrosesan data, terutama pada proses *training*. Werbos [4] dan Rumelhart *et al.* [5] mengembangkan algoritma *backpropagation* untuk meningkatkan kecepatan komputasi pada *artificial neural network*. Namun, algoritma ini memiliki kelemahan dalam pemrosesan data dengan jumlah besar.

Extreme learning machine [6] adalah salah satu algoritma yang digunakan untuk meningkatkan kecepatan *training* pada *artificial neural network*. Dengan menggunakan invers matriks Moore-Penrose, *extreme learning machine*, yang diterapkan pada sebuah *feedforward neural network* dengan satu *hidden layer*, dapat memberikan hasil yang lebih akurat dengan kecepatan *training* yang tinggi.

Pada penelitian ini, data hasil pengukuran yang dilakukan oleh Rahmat *et al.* [2] akan diolah menggunakan *extreme learning machine* untuk memperkirakan indeks kualitas air yang ditunjukkan pada setiap baris data. Metode penghitungan kualitas air yang digunakan dalam penelitian ini mengacu pada Keputusan Menteri Negara Lingkungan Hidup Nomor 115 Tahun 2003 tentang Pedoman Penentuan Status Mutu Air. Hasil akhir dari penelitian ini adalah sebuah grafik yang menunjukkan indeks kualitas air yang dihasilkan dari proses prediksi menggunakan *extreme learning machine*.

II. RUMUSAN MASALAH

Pengukuran kualitas air di Danau Toba selama ini dilakukan dengan pengujian laboratorium terhadap sampel air yang diambil dari beberapa lokasi di sekitar Danau Toba, sehingga dapat meningkatkan biaya dan waktu yang digunakan dalam proses pengukuran. Karena itu, diperlukan sebuah metode untuk melakukan prediksi kualitas air dengan mempertimbangkan efisiensi biaya dan waktu yang digunakan dalam proses penelitian.

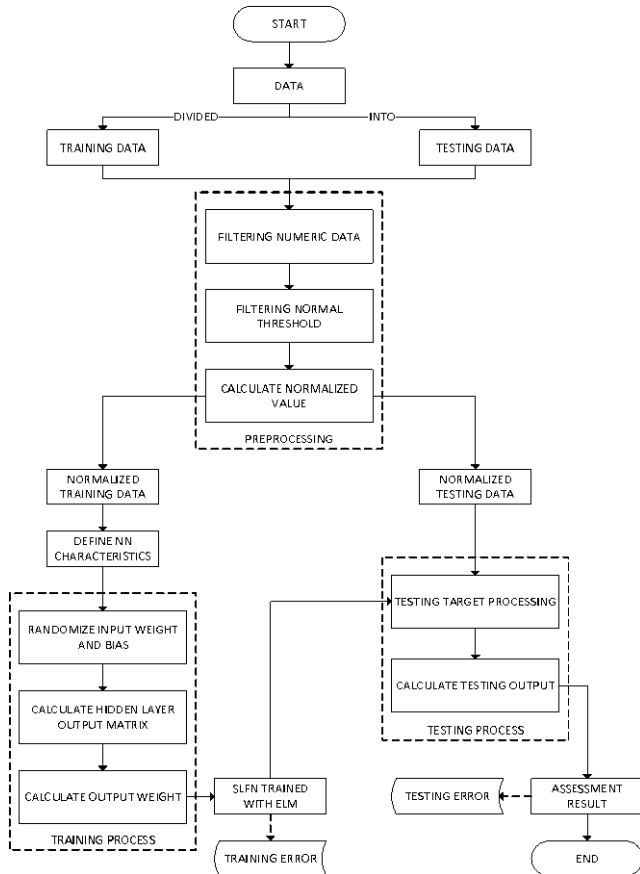
III. METODOLOGI PENELITIAN

Bagian ini akan menjelaskan sumber data yang digunakan dan arsitektur umum dari penelitian yang diterapkan. Bagian A akan menjelaskan sumber data yang digunakan dalam penelitian ini, sedangkan Bagian B akan menjelaskan arsitektur umum dari penelitian ini.

A. Sumber Data

Data yang digunakan pada penelitian ini adalah data hasil pengukuran yang dilakukan oleh Rahmat *et al.* [2] pada beberapa lokasi di sekitar perairan Danau Toba. Parameter yang diukur pada penelitian ini adalah dissolved oxygen (DO), tingkat keasaman (pH), tingkat potensi redoks, suhu air, suhu udara, dan kelembaban udara.

B. Arsitektur Umum



Gambar 1. Arsitektur Umum

Gambar 1 menunjukkan arsitektur umum yang diterapkan pada penelitian ini. Data yang diterima akan diolah melalui tahap-tahap yang dijelaskan sebagai berikut:

- *Preprocessing*: Data yang digunakan dalam penelitian ini akan diproses terlebih dahulu melalui tahap *preprocessing*. Hal ini dilakukan agar setiap baris data dapat digunakan oleh *extreme learning machine*. Penyaringan akan dilakukan terhadap tiap baris data menurut nilai setiap parameter. Metode normalisasi yang digunakan pada tahap ini adalah *min-max normalization* [7], di mana untuk setiap parameter, normalisasi dilakukan menggunakan persamaan 1:

$$A' = \left(\frac{A - A_{min}}{A_{max} - A_{min}} \right) * (D - C) + C \quad (1)$$

di mana A' merupakan hasil normalisasi dari data bernilai A yang memiliki rentang $[C, D]$, A_{min} adalah nilai terendah dari A , dan A_{max} adalah nilai tertinggi dari A . Hasil dari tahap *preprocessing* adalah training dataset dan testing dataset yang dapat digunakan oleh *extreme learning machine*.

- *Penentuan karakteristik neural network*: Sebelum proses training dilakukan, dua karakteristik *neural network*, yaitu jumlah *hidden neuron* dan fungsi aktivasi yang diterapkan oleh setiap neuron pada *neural network* akan ditentukan terlebih dahulu. Pada penelitian ini, fungsi aktivasi yang akan digunakan adalah fungsi *sigmoid*, *sine*, *cosine*, dan *hard-limit*.
- *Training*. Pada tahap ini, *extreme learning machine* akan diterapkan dalam proses training pada *neural network*, yang terbagi atas tiga tahap, yaitu:
 - Pengacakan *input weight* dan *bias*. Pada tahap ini, *input weight* dari vektor yang menghubungkan setiap *input neuron* dan *hidden neuron*, dan *bias* pada *hidden neuron*, akan diperoleh secara acak. Nilai dari hasil pengacakan tersebut akan digunakan untuk kalkulasi keluaran dari *hidden layer*.
 - Penghitungan keluaran *hidden layer*. Pada tahap ini, input yang diterima dari *input neuron* dan *bias* akan dihitung sesuai fungsi aktivasi yang digunakan, di mana hasil penghitungan akan disimpan dalam sebuah matriks H .
 - Penghitungan nilai *output*. Pada tahap ini, nilai output yang diberikan akan dihitung sesuai persamaan 2:

$$\beta = H^+ T \quad (2)$$

di mana β adalah output matrix, H^+ adalah invers matriks Moore-Penrose dari matriks H , dan T adalah target keluaran yang diterima dari *input neuron*. Hasil dari proses ini adalah sebuah *neural network* yang telah dilatih, diikuti dengan *training error* yang diperoleh.

- *Testing*. Pada tahap ini, *artificial neural network* yang telah dilatih menggunakan *extreme learning machine*

akan diuji dengan memberikan testing dataset. Target keluaran akan diproses terlebih dahulu sebelum nilai parameter dari setiap baris data dihitung menggunakan *neural network* yang dihasilkan dari proses *training*. Hasil akhir dari penelitian ini adalah grafik yang menunjukkan indeks kualitas air dari setiap baris data.

IV. HASIL PENGUJIAN

V. KESIMPULAN

Pada penelitian ini, proses prediksi kualitas air di Danau Toba dilakukan dengan menggunakan data yang diperoleh dari penelitian yang dilakukan oleh Rahmat *et al.* [2] menggunakan *extreme learning machine*.

DAFTAR PUSTAKA

- [1] D. Haro, Y. Djayus and Z. Harahap, "Kondisi Kualitas Air Danau Toba di Kecamatan Haranggaol Horison Kabupaten Simalungun Sumatera Utara", in *AQUACOASTMARINE*, vol. 1, no. 1, 2013.
- [2] R. F. Rahmat, Athmanathan, M. F. Syahputra and M. S. Lydia, "Real Time Monitoring System for Water Pollution in Lake Toba," in *Proceedings of ICIC 2016*, Lombok, 2016.
- [3] N. Kasabov, *Evolving Connectionist Systems*, 1st ed. Dordrecht: Springer, 2007.
- [4] P. J. Werbos, "Beyond regression: new tools for prediction and analysis in the behavioral science", in *Ph. D. Thesis, Harvard University*, 1974.
- [5] D. Rumelhart, G. Hinton and R. Williams, "Learning representations by back-propagating errors", *Nature*, vol. 323, no. 6088, pp. 533-536, 1986.
- [6] G. Huang, Q. Zhu and C. Siew, "Extreme learning machine: Theory and applications", *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501, 2006.
- [7] S. Patro and K. sahu, "Normalization: A Preprocessing Stage", *IARJSET*, pp. 20-22, 2015.
- [8]