

BAB 4

IMPLEMENTASI DAN PENGUJIAN

Bab ini membahas tentang hasil yang diperoleh dari implementasi metode yang digunakan, yaitu *extreme learning machine* (ELM), untuk melakukan prediksi kualitas air di Danau Toba. Bab ini akan menjabarkan hasil perancangan antarmuka yang digunakan dalam proses prediksi, prosedur operasional dari aplikasi yang dirancang, dan hasil prediksi yang didapat menggunakan ELM.

4.1. Implementasi Sistem

Dalam penelitian ini, tahap *preprocessing* akan diimplementasikan ke dalam sistem menggunakan bahasa pemrograman Python, sedangkan tahap pelatihan hingga visualisasi hasil akhir akan diimplementasikan ke dalam sistem menggunakan bahasa pemrograman MATLAB.

4.1.1. Spesifikasi perangkat keras dan perangkat lunak yang digunakan

Spesifikasi perangkat keras dan perangkat lunak yang digunakan untuk membangun sistem ini adalah sebagai berikut:

1. Prosesor Intel(R) Core(TM) i5-3317U CPU @ 1.70 GHz
2. Kapasitas memori RAM sebesar 4 GB
3. Sistem operasi Windows 10 Pro 64-bit
4. *Hard drive* yang memiliki kapasitas sebesar 500 GB
5. *Software* yang digunakan adalah Python versi 2.7 dan MATLAB versi R2015a (8.5.0.197613)
6. *Library* yang digunakan adalah *library elm_train* dan *elm_predict* yang diterbitkan oleh Zhu *et al.* (2004)

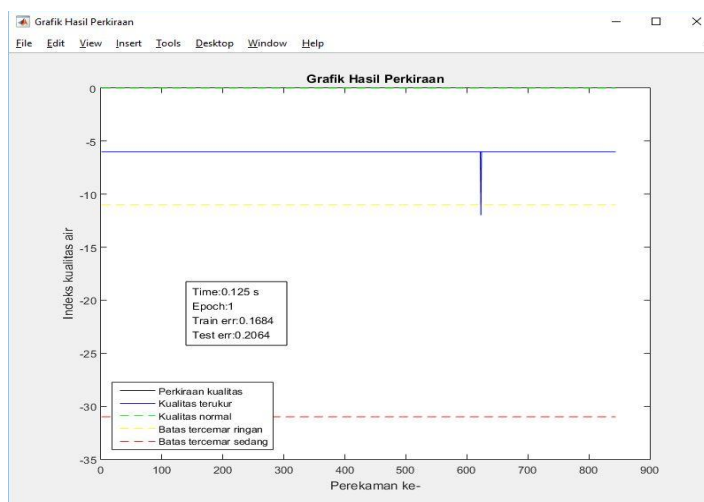
4.1.2. Implementasi perancangan antarmuka

Perancangan antarmuka sistem dibuat berdasarkan rancangan yang telah dilakukan pada bab 3. Antarmuka aplikasi yang telah dirancang pada penelitian ini ditunjukkan oleh Gambar 4.1. Antarmuka yang dirancang terdiri dari halaman utama, di mana dalam halaman utama terdapat menu-menu yang digunakan untuk proses prediksi kualitas air menggunakan *extreme learning machine*.



Gambar 4.1. Hasil rancangan antarmuka aplikasi

Setelah data latih, data uji, jumlah *hidden neuron*, dan fungsi aktivasi untuk proses prediksi telah diberikan kepada aplikasi, proses prediksi akan dimulai dan menghasilkan grafik prediksi seperti yang ditunjukkan oleh gambar 4.2. Pada grafik akan ditampilkan indeks kualitas air yang diperkirakan, indeks kualitas air yang terukur dari hasil pengukuran, batas kualitas air sangat baik, batas kualitas air baik, dan batas kualitas air cukup.



Gambar 4.2. Grafik hasil perkiraan kualitas air

4.1.3. Implementasi data

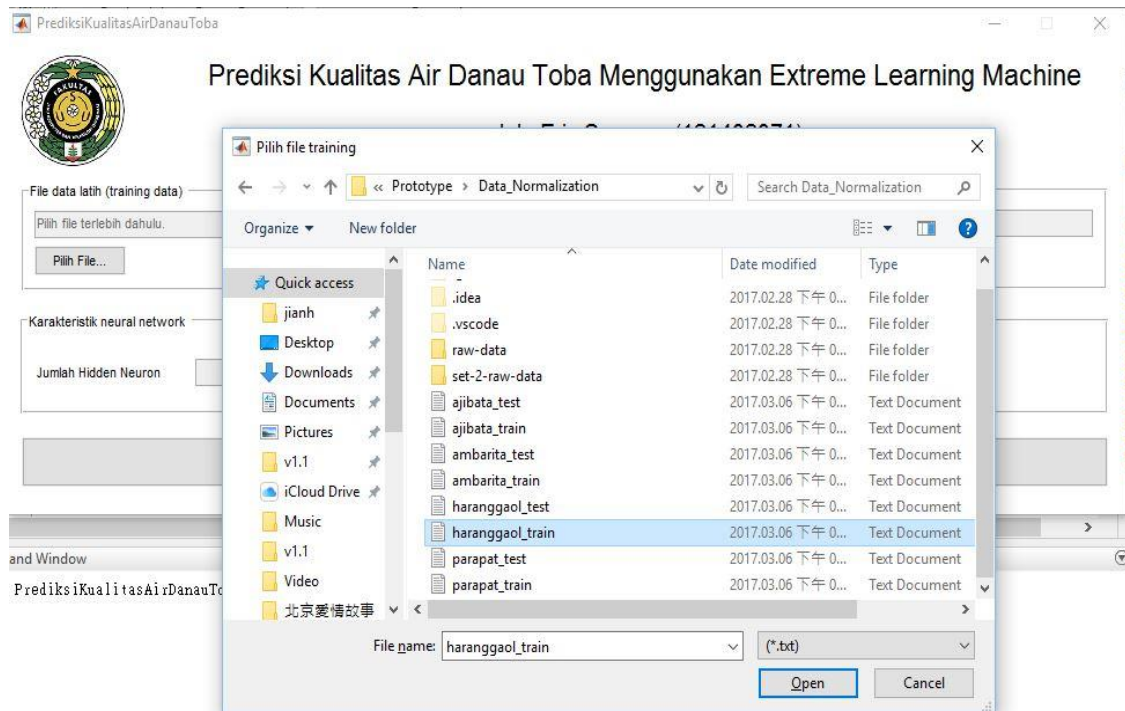
Data yang digunakan dalam penelitian ini bersumber dari penelitian yang dilakukan oleh Rahmat *et al.* (2016), di mana parameter-parameter kualitas air diukur dalam kurun waktu yang sedemikian rupa, dan disimpan dalam format dokumen teks, dengan hasil pengukuran setiap parameter dipisahkan oleh tanda titik koma. Pengukuran dilakukan pada beberapa lokasi, dan dikumpulkan dalam beberapa *file*. *File* tersebut akan diolah terlebih dahulu dalam tahap *preprocessing* sehingga dihasilkan data latih dan data uji yang dapat digunakan oleh *extreme learning machine*. Perincian dari *file* yang digunakan dalam penelitian ini ditunjukkan oleh Tabel 4.1. Pemberian nama “*resume*” pada nama *file* awal menunjukkan bahwa proses pengukuran dilakukan pada waktu yang berbeda.

Tabel 4.1. Perincian data yang digunakan dalam penelitian

| No. | Nama <i>file</i> awal | Lokasi | Jumlah baris data | | | |
|-----|-------------------------|------------|-------------------|---------------------|-------|------|
| | | | Awal | Setelah penyaringan | Latih | Uji |
| 1 | DATA ajibata.txt | Ajibata | 2203 | 2112 | 1268 | 844 |
| 2 | DATA Haranggaol.txt | Haranggaol | 6374 | 3532 | 2120 | 1412 |
| 3 | DATA parapat.txt | Parapat | 2446 | 1452 | 872 | 580 |
| 4 | DATA parapat resume.txt | | | | | |
| 5 | DATA samosir.txt | Ambarita | 6129 | 3113 | 1869 | 1244 |
| 6 | DATA samosir resume.txt | | | | | |

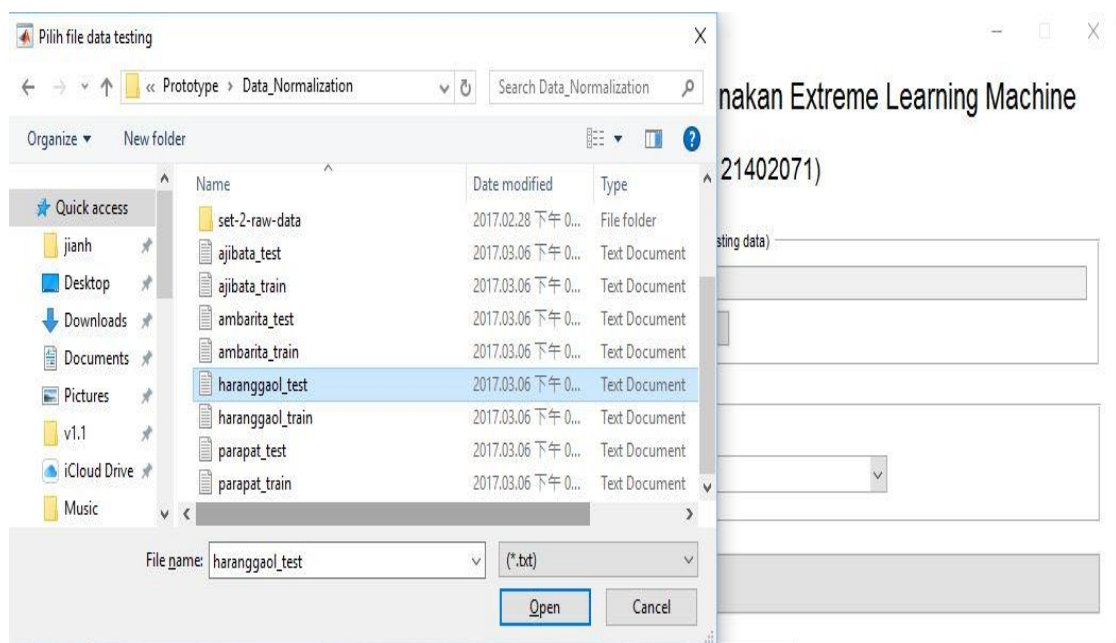
4.2. Prosedur Operasional

Bagian ini akan memaparkan prosedur yang harus dilakukan dalam menggunakan aplikasi yang dikembangkan dalam penelitian ini. Setelah antarmuka tampil di layar, pengguna dapat memilih *file* data latih (*training data*) dengan memilih tombol “Pilih” pada bagian “File data latih (*training data*)”. Tombol “Pilih” akan mengarahkan pengguna pada dialog pemilihan *file* data latih, di mana *file* yang dapat dipilih adalah dokumen teks dengan ekstensi “txt”. Ilustrasi dari tahap pemilihan *file* data latih ditunjukkan oleh Gambar 4.3.



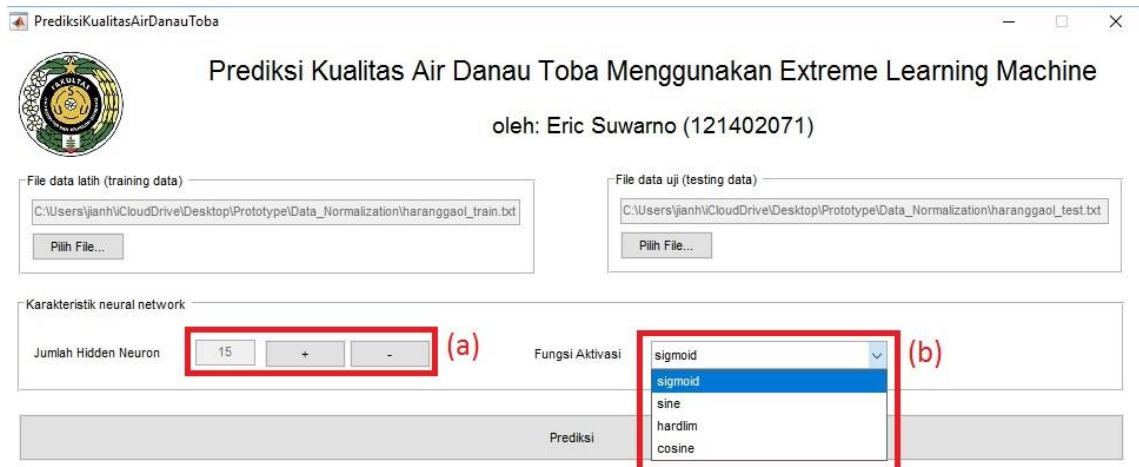
Gambar 4.3. Ilustrasi pemilihan *file* data latih (*training data*)

File data uji (*testing data*) juga dapat diberikan kepada aplikasi dengan menekan tombol “Pilih” pada bagian “File data uji (*testing data*)”. Dialog pemilihan *file* data uji akan muncul setelah tombol “Pilih” ditekan, seperti yang ditunjukkan pada Gambar 4.4. Data uji yang dapat diterima dalam aplikasi ini berupa dokumen teks dengan ekstensi “txt”.



Gambar 4.4. Ilustrasi pemilihan *file* data uji (*testing data*)

Jumlah neuron pada *hidden layer* dapat diatur melalui tombol plus dan minus pada bagian karakteristik *neural network*. Sedangkan, fungsi aktivasi yang akan digunakan oleh *neural network* dapat dipilih melalui *combo box* pada bagian karakteristik *neural network*. Ilustrasi pengaturan fungsi aktivasi dan jumlah neuron pada *hidden layer* ditunjukkan oleh Gambar 4.5.



Gambar 4.5. Ilustrasi pengaturan karakteristik *neural network*, yaitu jumlah neuron pada *hidden layer* (a) dan fungsi aktivasi (b)

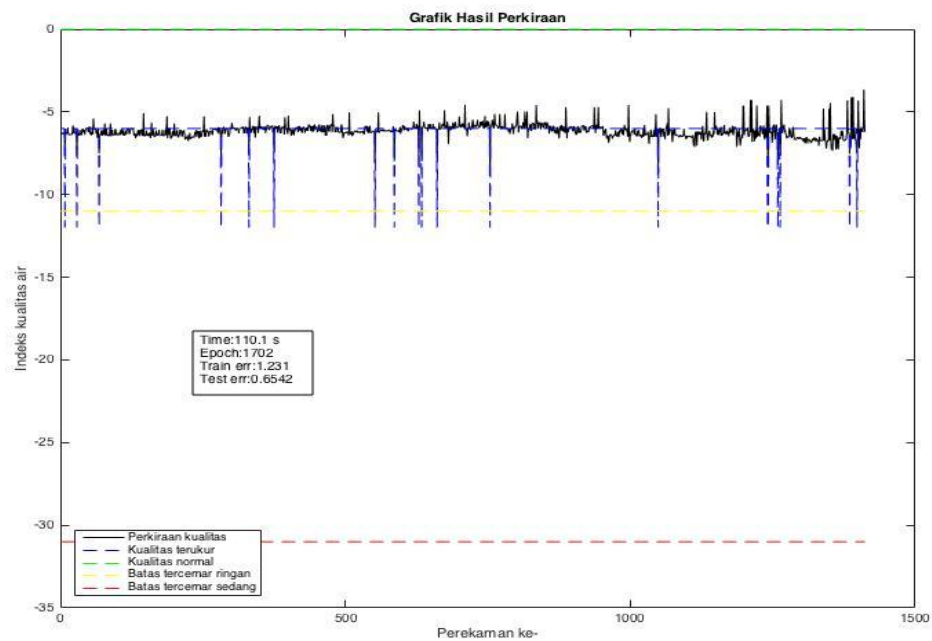
4.3. Hasil Pengujian

Bagian ini akan memaparkan hasil yang didapatkan dari implementasi *Extreme Learning Machine* (ELM) dalam melakukan proses prediksi kualitas air Danau Toba. Proses pengujian akan dilakukan terhadap data uji setelah proses pelatihan telah selesai dilakukan melalui *extreme learning machine* menggunakan data latih yang disediakan. Akurasi dari hasil pengujian yang didapat akan direpresentasikan melalui nilai *root mean square error* (RMSE).

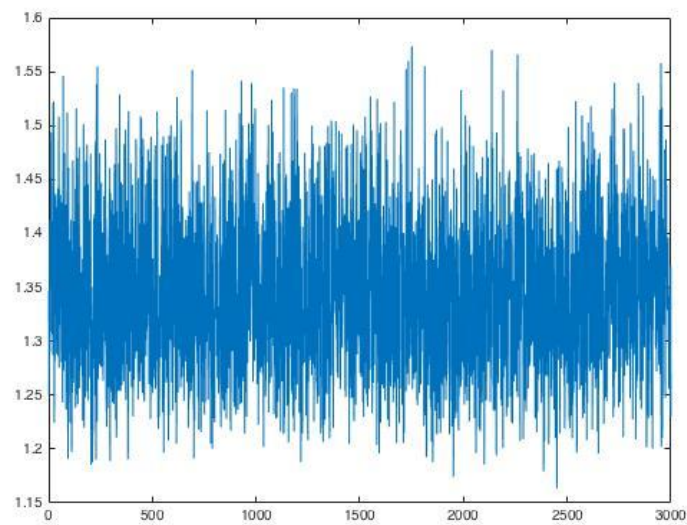
Pengujian pertama dilakukan terhadap data dari proses pengukuran yang dilakukan di Haranggaol. Pengujian dilakukan dengan menggunakan fungsi *sine* sebagai fungsi aktivasi, dan memiliki 15 neuron pada *hidden layer*. Pengujian dilakukan sebanyak 3.000 kali, dengan hasil pengujian yang ditunjukkan oleh Gambar 4.6. Tingkat akurasi prediksi terbaik yang dicapai dalam pengujian ini adalah 0,6542, yang didapat pada pengulangan ke-1.792.

Tingkat error dalam tahap *training* dan *testing* yang didapat dari pengujian pertama ditunjukkan oleh Gambar 4.7 dan Gambar 4.8. *Input weight* dan *bias* pada

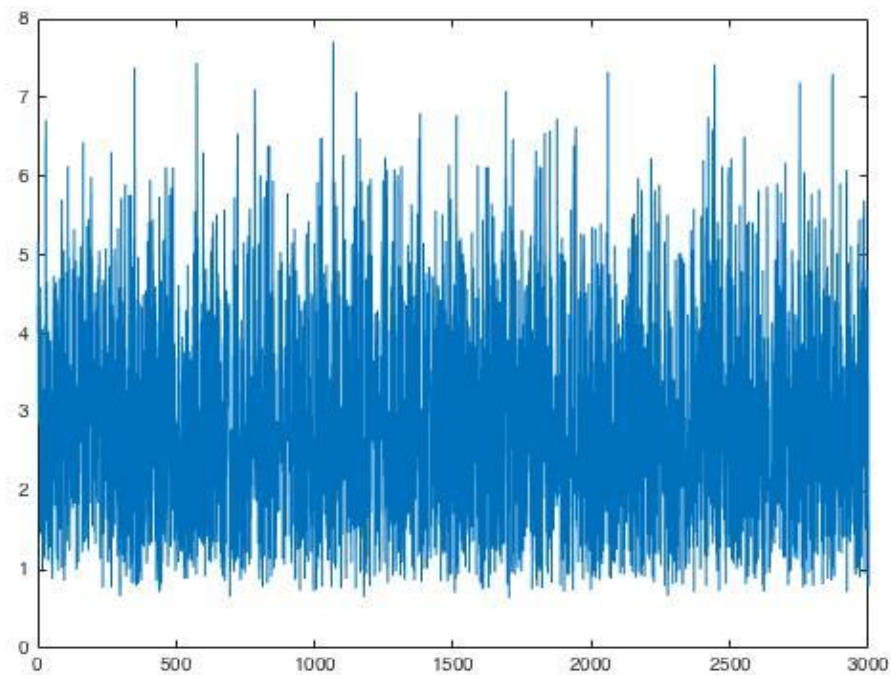
extreme learning machine ditentukan secara acak. Hal ini menyebabkan penghitungan *hidden output matrix* dan *output weight* memberikan hasil yang berbeda untuk setiap pengulangan. Karena perbedaan tersebut, tingkat *error* yang didapat saat pelatihan (*training error*) dapat bervariasi untuk setiap pengulangan. Perbedaan nilai *training error* yang didapat dalam setiap percobaan juga mempengaruhi nilai *error* yang didapat saat pengujian (*testing error*).



Gambar 4.6. Hasil pengujian pertama



Gambar 4.7. Grafik nilai *error* latih pada pengujian pertama



Gambar 4.8. Tingkat *error* uji pada pengujian pertama

Pengujian juga dilakukan dengan menggunakan data hasil pengukuran yang dilakukan di setiap lokasi, dengan menggunakan parameter *artificial neural network* yang berbeda-beda. Parameter-parameter *artificial neural network* yang akan digunakan dalam pengujian ini ditunjukkan pada Tabel 4.2. Setiap pengujian akan dilakukan sebanyak 3.000 pengulangan. Hasil yang didapat dari setiap pengujian terhadap setiap *dataset* akan dibandingkan berdasarkan tingkat *error* latih dan uji yang diperoleh.

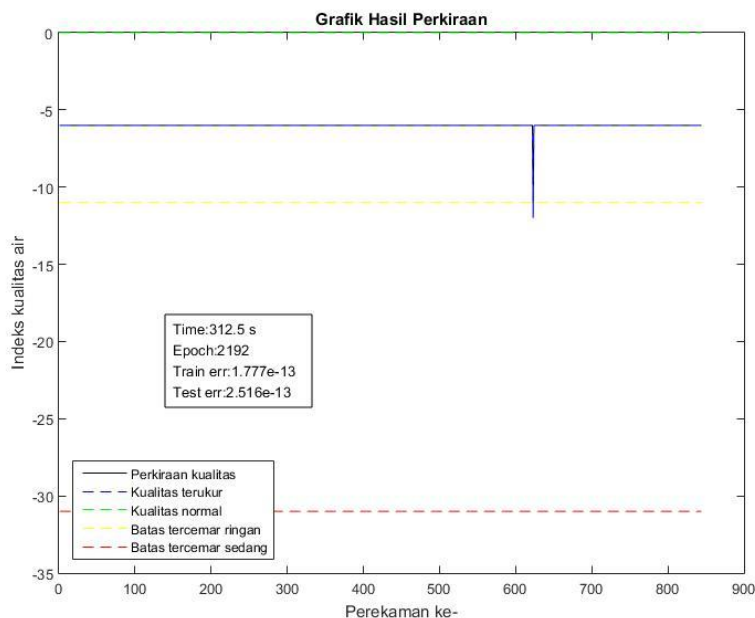
Tabel 4.2. Parameter *artificial neural network* pada pengujian setiap *dataset*

| Pengujian <i>dataset</i> ke- | Fungsi aktivasi | Jumlah neuron pada <i>hidden layer</i> |
|---------------------------------|--------------------|---|
| 1 | <i>Sigmoid</i> | 12 |
| 2 | <i>Sigmoid</i> | 48 |
| 3 | <i>Sine</i> | 12 |
| 4 | <i>Sine</i> | 48 |
| 5 | <i>Hardlim</i> | 25 |
| 6 | <i>Hardlim</i> | 50 |
| 7 | <i>Hardlim</i> | 75 |
| 8 | <i>Cosine</i> | 12 |
| 9 | <i>Cosine</i> | 48 |

Hasil pengujian terhadap data yang diperoleh dari hasil pengukuran di Ajibata ditunjukkan pada Tabel 4.3. Hasil dari tahap pengujian terhadap data ini menunjukkan bahwa dari sembilan pengujian yang dilakukan, pengujian dataset dengan parameter kelima, yaitu menggunakan fungsi *hardlim* sebagai fungsi aktivasi, dan menggunakan 25 neuron pada *hidden layer*, memberikan hasil yang paling baik, dengan nilai *error* uji sebesar $2,516 \times 10^{-13}$ pada perulangan ke-2.192. Hasil pengujian dataset Ajibata dengan parameter kelima sesuai Tabel 4.2 ditunjukkan oleh Gambar 4.8.

Tabel 4.3. Hasil pengujian pada *dataset* Ajibata

| Pengujian <i>dataset</i> ke- | Hasil uji terbaik | | |
|---------------------------------|------------------------------|----------------------------|-------------------|
| | <i>Error</i> latih (RMSE) | <i>Error</i> uji (RMSE) | Perulangan ke- |
| 1 | 0,1612 | 0,08525 | 404 |
| 2 | 0,6776 | 2,32 | 1818 |
| 3 | 0,1623 | 0,07849 | 2045 |
| 4 | 0,6353 | 2,006 | 1711 |
| 5 | $1,777 \times 10^{-13}$ | $2,516 \times 10^{-13}$ | 2192 |
| 6 | $2,625 \times 10^{-13}$ | 0,0243 | 1249 |
| 7 | $1,311 \times 10^{-13}$ | 0,02688 | 1444 |
| 8 | 0,1486 | 0,07035 | 2353 |
| 9 | 0,5736 | 2,192 | 1898 |



Gambar 4.9. Hasil prediksi *dataset* Ajibata pada pengujian kelima

Hasil pengujian pada Tabel 4.3 menunjukkan bahwa jumlah neuron pada *hidden layer* mempengaruhi tingkat akurasi yang diperoleh dalam proses prediksi. Hal ini dibuktikan melalui perbandingan antara hasil uji terbaik pada pengujian pertama dan kedua terhadap *dataset* Ajibata.

Dalam pengujian terhadap *dataset* Ajibata, pengujian pertama memiliki 12 neuron pada *hidden layer*, sedangkan pengujian kedua memiliki 48 neuron pada *hidden layer*. Pengujian pertama dan kedua menggunakan fungsi aktivasi *sigmoid* sebagai fungsi aktivasi. Hasil uji terbaik yang didapatkan pada pengujian pertama memiliki *error* latih 0,1612 dan *error* uji 0,08525, yang didapat pada perulangan ke-404. Sedangkan pada pengujian kedua, hasil uji terbaik yang diperoleh memiliki *error* latih 0,6776 dan *error* uji 2,32, yang didapat pada perulangan ke-1.818.

Hasil pengujian kelima, keenam, dan ketujuh pada Tabel 4.3 didapat dengan menggunakan fungsi *hardlim* sebagai fungsi aktivasi setiap neuron, dengan jumlah neuron pada *hidden layer* yang berbeda-beda. Jumlah neuron pada *hidden layer* dalam pengujian kelima, keenam, dan ketujuh, masing-masing adalah 25, 50, dan 75 buah. Hasil prediksi terbaik dari pengujian kelima, keenam, dan ketujuh menunjukkan bahwa implementasi fungsi *hardlim* sebagai fungsi aktivasi dalam *extreme learning machine* memberikan hasil prediksi yang lebih akurat pada *dataset* ini.

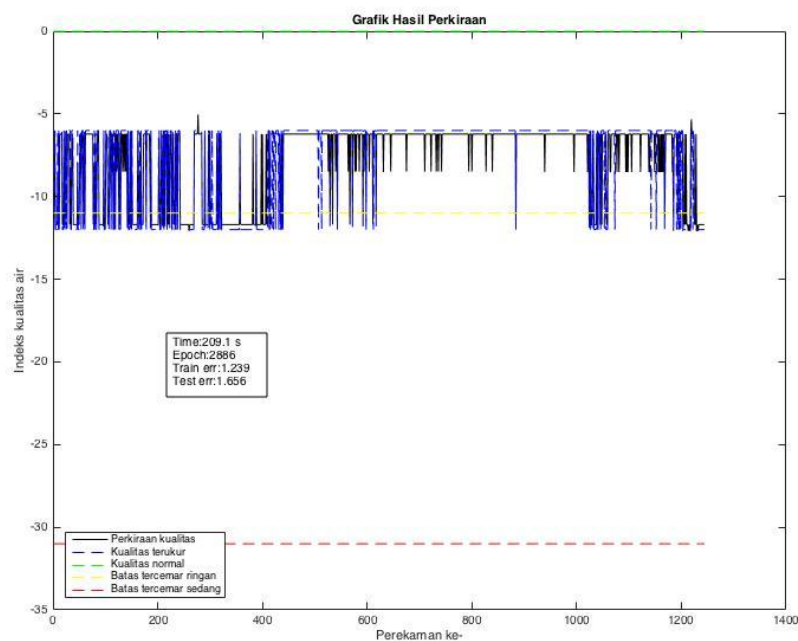
Pada Tabel 4.4 ditunjukkan hasil pengujian terhadap data yang diperoleh dari hasil pengukuran di Ambarita, Samosir. Berdasarkan hasil pengujian pada Tabel 4.4, diketahui bahwa hasil pengujian ketujuh, di mana *artificial neural network* yang digunakan memiliki 75 neuron pada *hidden layer*, serta menggunakan fungsi *hardlim* sebagai fungsi aktivasi pada setiap neuron, memiliki hasil uji terbaik, dengan *error* uji sebesar 1,6559 pada perulangan ke-2.886. Hasil ini juga menunjukkan bahwa dengan menggunakan fungsi *hardlim* sebagai fungsi aktivasi, bertambahnya jumlah neuron pada *hidden layer* dapat meningkatkan akurasi prediksi. Grafik hasil prediksi yang dilakukan pada pengujian ketujuh terhadap *dataset* Ambarita ditunjukkan oleh Gambar 4.10.

Berdasarkan Tabel 4.4, diketahui pula bahwa dari empat fungsi aktivasi yang digunakan dalam proses pengujian ini, fungsi *hardlim* memberikan hasil yang lebih baik dibandingkan dengan fungsi aktivasi lainnya. Hal ini dibuktikan melalui hasil uji terbaik yang diperoleh pada pengujian kelima, keenam, dan ketujuh terhadap *dataset* Ambarita.

Pada pengujian menggunakan fungsi aktivasi *hardlim*, *error* uji terbaik yang diperoleh lebih rendah daripada hasil uji terbaik yang diperoleh melalui fungsi aktivasi lainnya.

Tabel 4.4. Hasil pengujian pada *dataset* Ambarita

| Pengujian <i>dataset</i> ke- | Hasil uji terbaik | | |
|---------------------------------|------------------------|------------------|----------------|
| | Error latih (RMSE) | Error uji (RMSE) | Perulangan ke- |
| 1 | 1,285 | 3,486 | 2726 |
| 2 | 1,134 | 3,668 | 232 |
| 3 | 1,325 | 3,668 | 450 |
| 4 | 1,135 | 3,433 | 264 |
| 5 | 1,4796 | 2,1304 | 753 |
| 6 | $9,025 \times 10^{-1}$ | 1,6704 | 1837 |
| 7 | 1,2392 | 1,6559 | 2886 |
| 8 | 1,3955 | 3,5507 | 1493 |
| 9 | 1,1442 | 3,3985 | 1194 |



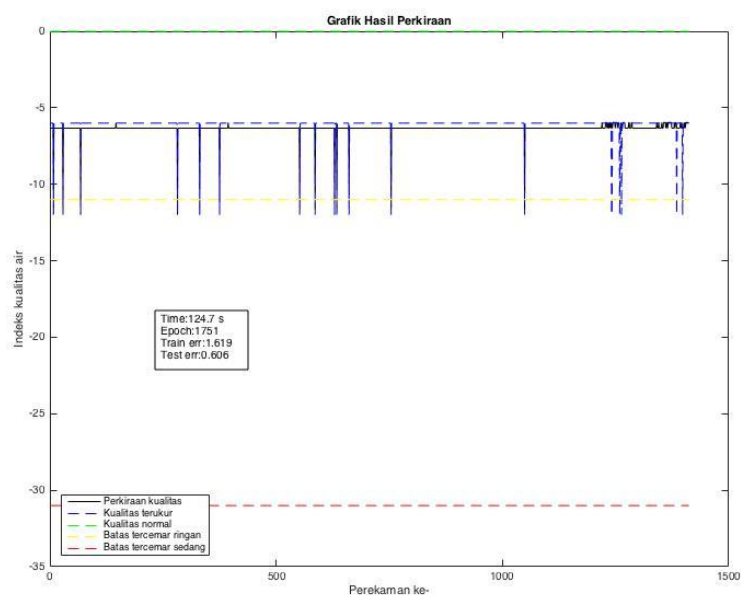
Gambar 4.10. Grafik hasil prediksi *dataset* Ambarita pada pengujian ketujuh

Hasil pengujian terhadap *dataset* hasil pengukuran yang dilakukan di Haranggaol ditunjukkan pada Tabel 4.5. Berdasarkan hasil pengujian yang dilakukan, diketahui bahwa hasil terbaik yang didapat dalam pengujian dataset Haranggaol didapat pada percobaan keenam. Pada percobaan keenam, terdapat 50 neuron pada *hidden layer*, di mana setiap neuron menggunakan fungsi *hardlim* sebagai fungsi aktivasi. Gambar

4.11 menunjukkan hasil prediksi terhadap *dataset* Haranggaol yang diperoleh pada pengujian keenam.

Tabel 4.5. Hasil pengujian pada *dataset* Haranggaol

| Pengujian <i>dataset</i> ke- | Hasil uji terbaik | | |
|---------------------------------|------------------------------|----------------------------|-------------------|
| | <i>Error</i> latih (RMSE) | <i>Error</i> uji (RMSE) | Perulangan ke- |
| 1 | 1,479 | $6,679 \times 10^{-1}$ | 1124 |
| 2 | 1,027 | 1,14 | 1852 |
| 3 | 1,613 | $6,546 \times 10^{-1}$ | 1693 |
| 4 | 1,052 | 1,101 | 1989 |
| 5 | 1,65 | $6,448 \times 10^{-1}$ | 251 |
| 6 | 1,619 | $6,06 \times 10^{-1}$ | 1751 |
| 7 | 1,683 | $6,181 \times 10^{-1}$ | 2000 |
| 8 | 1,487 | $6,578 \times 10^{-1}$ | 2566 |
| 9 | 1,072 | $9,528 \times 10^{-1}$ | 70 |



Gambar 4.11. Hasil prediksi *dataset* Haranggaol pada pengujian keenam

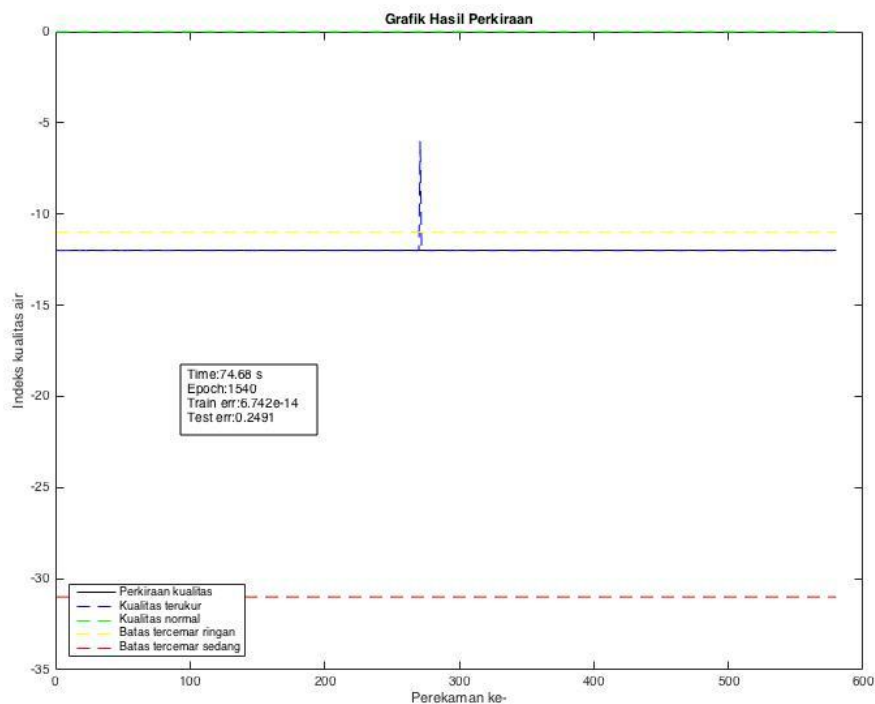
Tabel 4.6 menunjukkan hasil pengujian yang dilakukan terhadap *dataset* hasil pengukuran yang dilakukan di Parapat. Berdasarkan hasil pengujian yang ditunjukkan pada tabel, dapat diketahui bahwa *error* latih yang dihasilkan dalam proses *training* relatif berkurang seiring dengan bertambahnya jumlah neuron yang terdapat pada *hidden layer*. Namun, hal tersebut tidak selalu berakibat pada penurunan nilai *error*

yang dihasilkan pada tahap uji. Hal itu dibuktikan pada hasil uji terbaik yang diperoleh pada pengujian pertama dan kedua. Pada pengujian pertama, yang menggunakan *artificial neural network* dengan 12 neuron pada *hidden layer*, hasil uji terbaik yang diperoleh memiliki *error* latih sebesar 0,5279 dan *error* uji sebesar 0,7823, yang didapat pada perulangan ke-148. Sedangkan pada pengujian kedua, yang menggunakan *artificial neural network* dengan 48 neuron pada *hidden layer*, hasil uji terbaik yang diperoleh memiliki *error* latih sebesar 0,1478 dan *error* uji sebesar 0,8266, yang didapat pada perulangan ke-1.743.

Pada pengujian terhadap *dataset* Parapat, hasil uji terbaik didapat pada pengujian keenam. Sebanyak 50 neuron terdapat pada *hidden layer* dari *artificial neural network* yang digunakan dalam pengujian ini. Fungsi aktivasi yang diterapkan oleh setiap neuron pada pengujian ini adalah fungsi *hardlim*. *Error* uji yang didapat pada pengujian keenam adalah 0,2491, yang didapat pada perulangan ke-1.540. Grafik dari hasil prediksi yang didapat pada pengujian keenam terhadap *dataset* Parapat ditunjukkan oleh Gambar 4.12.

Tabel 4.6. Hasil pengujian pada *dataset* Parapat

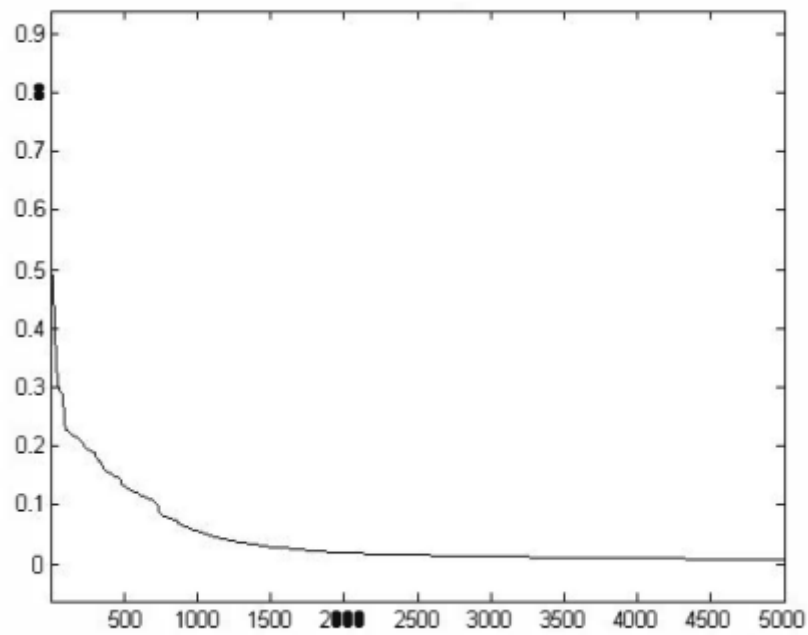
| Pengujian <i>dataset</i> ke- | Hasil uji terbaik | | |
|---------------------------------|------------------------------|----------------------------|-------------------|
| | <i>Error</i> latih (RMSE) | <i>Error</i> uji (RMSE) | Perulangan ke- |
| 1 | 0,5279 | 0,7823 | 148 |
| 2 | 0,1478 | 0,8266 | 1743 |
| 3 | 0,4615 | 0,5504 | 1185 |
| 4 | 0,1484 | 0,8346 | 2936 |
| 5 | 0,7243 | 0,2876 | 1272 |
| 6 | $6,742 \times 10^{-14}$ | 0,2491 | 1540 |
| 7 | 0,4978 | 0,309 | 2921 |
| 8 | 0,5351 | 0,7558 | 177 |
| 9 | 0,1724 | 0,8924 | 2173 |



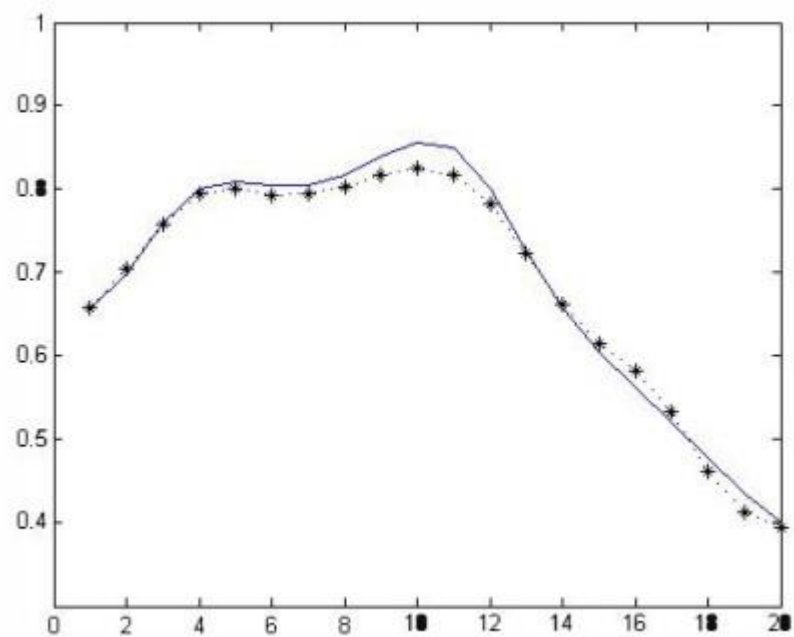
Gambar 4.12. Grafik hasil pengujian keenam pada prediksi *dataset* Parapat

Secara umum, dapat disimpulkan bahwa tingkat pencemaran air di wilayah Danau Toba, bervariasi dari tingkat ringan ke tingkat sedang. Hal ini dapat dilihat dari indeks kualitas air hasil prediksi menurut indeks Scoret, yang berkisar antara -6 hingga -11. Hasil yang didapat sesuai dengan hasil penelitian yang dilakukan Haro *et al.* (2013), yang menyimpulkan bahwa pencemaran air di Danau Toba berada dalam tingkat ringan hingga sedang.

Hal berbeda ditunjukkan oleh penelitian yang dilakukan oleh Ming (2012), yang menggunakan *backpropagation neural network* dan *fuzzy logic* untuk melakukan penilaian kualitas air. Proses penilaian dilakukan melalui masukan berupa hasil pengukuran *dissolved oxygen*, *biological oxygen demand*, indeks permanganat, kadar zat phosphor, amonia, dan nitrat. Penelitian yang dilakukan oleh Ming menunjukkan bahwa nilai *error* yang dihasilkan dari proses pelatihan menunjukkan penurunan seiring dengan bertambahnya jumlah pengujian yang dilakukan, seperti yang ditunjukkan oleh Gambar 4.13. Proses pelatihan yang dilakukan oleh Ming memberikan hasil paling baik pada pengulangan ke-1.610, dengan parameter *learning rate* sebesar 0,7 dan koefisien impuls sebesar 0,5, seperti yang ditunjukkan pada Gambar 4.14.



Gambar 4.13. Grafik hasil pelatihan dari *backpropagation neural network*
(Ming, 2012)



Gambar 4.14. Grafik hasil pengujian dengan *backpropagation neural network*
(Ming, 2012)