

PENERAPAN ALGORITMA LEARNING VECTOR QUANTIZATION DALAM PENGKLASIFIKASIAN TINGKAT PENCEMARAN AIR SUNGAI

Muhammad Ridha Rahimi¹⁾, Hartatik²⁾

^{1), 2)} Teknik Informatika STMIK AMIKOM Yogyakarta
Jl Ring road Utara, Condongcatur, Sleman, Yogyakarta 55281
Email : mridharahimi@gmail.com¹⁾, hartatik@amikom.ac.id²⁾

Abstrak

Pesatnya pembangunan setiap tahunnya pada Kalimantan Selatan memungkinkan terjadinya kerusakan lingkungan. Salah satu kerusakan lingkungan adalah pencemaran air sungai yang disebabkan oleh limbah pabrik dari pabrik-pabrik ataupun dari pembuangan sampah sembarangan yang tidak pada tempatnya. Pencemaran tersebut akan diklasifikasikan kedalam beberapa kelas sesuai parameter-parameter masukannya. Pengklasifikasian ini biasanya dilakukan secara manual dengan metode Indeks Pencemaran.

Dalam mengklasifikasikan suatu permasalahan dengan jumlah data yang besar membutuhkan metode cepat dan akurat. Salah satunya adalah menggunakan metode jaringan syaraf tiruan. Penelitian ini bertujuan untuk mengimplementasikan pengklasifikasian tingkat pencemaran air sungai menggunakan Jaringan Syaraf Tiruan metode Learning Vector Quantization.

Penelitian yang dilakukan menggunakan 22 parameter sebagai neuron input dan 4 neuron output. Data ajar yang digunakan dalam penelitian ini sebanyak 100 dan data uji sebanyak 50. Learning rate yang ditetapkan dalam proses perhitungannya adalah 0,01 dan iterasi maksimal yang ditetapkan berjumlah 20.000. Hasil pengujian menunjukkan bahwa dengan 100 data ajar dan 50 data uji, nilai keakuratan yang didapat sebesar 76%.

Kata kunci:Jaringan syaraf tiruan, learning vector quantization, pencemaran air sungai.

1. Pendahuluan

Kalimantan Selatan terkenal dengan pasar terapungnya yang berada di atas sungai Barito. Tidak hanya berjualan, sungai ini juga dipakai masyarakat sekitar sebagai alat angkut atau transportasi mereka[1]. Selain sungai Barito ada 144 induk sungai lainnya yang tersebar di beberapa kabupaten Kalimantan Selatan[2]. Sungai-sungai di Kalimantan Selatan selain untuk transportasi juga dipakai masyarakat untuk memenuhi kebutuhan sehari-hari.

Air sungai sering dipakai untuk kebutuhan sehari-hari seperti mandi, cuci, kakus selain itu digunakan untuk perikanan, pertanian, perkebunan, dan air minum. Menurut Badan Pusat Statistik (BPS) untuk daerah perkotaan dan perdesaan ada sekitar 17,11% atau 166.845 rumah tangga yang masih menggunakan air sungai sebagai kebutuhan air minum[3]. Nilai ini menempati urutan ke-2 setelah air ledeng karena banyak digunakan di daerah perkotaan. Tapi jika dilihat dari daerah perdesaan air sungai menjadi yang terbanyak digunakan sebagai air minum sekitar 26% dari sumber air minum lainnya. Dikhawatirkan adalah air sungai di Kalimantan Selatan sebagai kebutuhan sehari-hari menjadi sasaran untuk pembuangan limbah para industri nakal, mengingat perkembangan industri dan pembangunan semakin meningkat.

Bersumber data dari BPS tahun 2004-2013 pembanguan di Kalimantan Selatan mengalami peningkatan setiap tahunnya. Pada tahun 2013 mengalami peningkatan sebesar 0,66% dari tahun sebelumnya[4]. Peningkatan juga terjadi pada sektor industri dalam hal produksi sebesar 0,75% di triwulan I 2015 dibandingkan dengan triwulan I 2014 untuk industri manufaktur besar dan sedang, sebesar 14,72% untuk industri manufaktur mikro dan kecil[5]. Semakin meningkatnya perkembangan industri dan pembangunan maka semakin bertambah pula kemungkinan resiko bahaya pencemaran pada perairan khususnya air sungai di Kalimantan Selatan yang disebabkan oleh hasil buangnya[2]. Maka dikeluarkannya Peraturan Gubernur Kalimantan Selatan No. 05 Tahun 2007. Peraturan ini dibuat bertujuan untuk memelihara kualitas air sungai agar tetap bermanfaat untuk kebutuhan sehari-hari dan mencegah kemungkinan pencemaran air sungai.

Hasil riset dari Greenpeace Indonesia pada salah satu laporannya yang berjudul “ Terungkap:Tambang Batubara Meracuni Air Kalimantan Selatan ” menerangkan 29 sampel yang diambil dari limbah cair, air kolam dan anak sungai sebagian besar telah melampaui batas baku mutu yang di tetapkan pemerintah untuk buangan limbah batubara, karena tingginya kadar besi, mangan dan keasaman[6]. 22 sampel memiliki pH di bawah 6, 17 sampel melewati baku mutu pembanding logam berat mangan dan 7 sampel melewati baku mutu

pembanding logam besi. Sebagian sampel yang diambil merupakan daerah yang berdekatan bahkan terhubung dengan anak sungai mangkaok, martapura, riamkiwa, asam-asam dan batulicin. Yang membahayakan dalam riset ini adalah penampungan air limbah yg berdekatan dengan lingkungan warga sekitar dan beberapa sungai, bahkan terdapat beberapa lubang kecil pada dinding tambang bocoran yang mengalir ke sungai milik masyarakat. Sungai kecil tersebut digunakan masyarakat untuk mengalir perkebunan, mandi dan memasak air. Kondisi tersebut yang membuat masyarakat perlu untuk melakukan pengecekan kualitas air sungai sebelum digunakan sebagai pemenuhan kebutuhan sehari-hari terutama untuk konsumsi.

Penanggulangan pencemaran dan atau kerusakan lingkungan di Indonesia telah diatur pada Undang-Undang Republik Indonesia Nomor 32 Tahun 2009 tentang perlindungan dan pengelolaan lingkungan hidup. Salah satu yang dapat dilakukan mahasiswa teknik informatika dalam pengendalian pencemaran adalah melalui Pasal 53 ayat (2) huruf (d) yang berbunyi "cara lain yang sesuai dengan perkembangan ilmu pengetahuan dan teknologi" [7]. Sesuai dengan poin tersebut, maka dalam penelitian ini mencoba untuk membangun sebuah aplikasi yang dapat mengenali kualitas air sungai berdasarkan parameter kandungan fisika dan kimia pada air sungai yang akan diuji menggunakan jaringan syaraf tiruan algoritma Learning Vector Quantization (LVQ).

2. Pembahasan

Dalam pengklasifikasian tingkat pencemaran air sungai biasanya melalui proses pengambilan sampel air dan dibawa ke laboratorium untuk mengetahui zat apa saja yang terkandung dalam air tersebut. Kadar zat terkandung pada sampel diolah dengan metode indeks pencemaran (*pollution index*) oleh seorang penganalisa air. Data hasil metode tersebutlah yang akan mengklasifikasikan ke dalam tingkat pencemaran air yang diuji. Pada metode tersebut penghitungan data sampel air menggunakan rumus yang cukup rumit, belum lagi ada beberapa parameter air yang penghitungannya berbeda dengan parameter lainnya. Dengan menggunakan Learning Vector Quantization (LVQ) dimaksudkan agar pengujian sampel bisa langsung mendapatkan hasilnya karena proses pembelajarannya berdasarkan pengalaman (*learning by experience*).

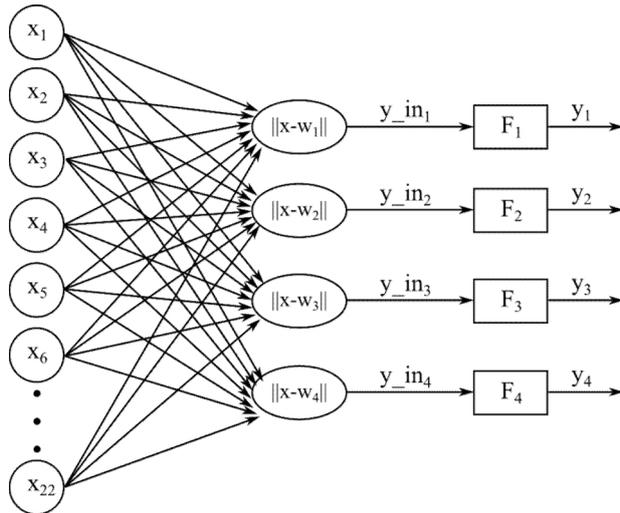
Learning Vector Quantization (LVQ) merupakan suatu metode untuk melakukan pelatihan terhadap lapisan-lapisan kompetitif yang terawasi. Lapisan kompetitif akan belajar secara otomatis untuk melakukan klasifikasi terhadap vektor input yang diberikan. Apabila beberapa vektor input memiliki jarak yang sangat berdekatan, maka vektor-vektor input tersebut akan dikelompokkan dalam kelas yang sama [8].

Dalam pengklasifikasian tingkat pencemaran air ada beberapa parameter yang menjadi pertimbangan dalam analisis pencemaran air. Parameter tersebut meliputi parameter fisika dan kimia. Parameter fisika adalah temperatur atau suhu mewakili neuron 1 (x_1), Zat Padat Tersuspensi (TSS) mewakili neuron 2 (x_2) dan Zat Padat Terlarut (TDS) mewakili neuron 3 (x_3). Parameter kimia adalah Besi (Fe) mewakili neuron 4 (x_4), Kadmium (Cd) mewakili neuron 5 (x_5), Mangan (Mn) mewakili neuron 6 (x_6), pH mewakili neuron 7 (x_7), BOD mewakili neuron 8 (x_8), COD mewakili neuron 9 (x_9), DO mewakili neuron 10 (x_{10}), Amoniak ($\text{NH}_3\text{-N}$) mewakili neuron 11 (x_{11}), Sulfat (SO_4) mewakili neuron 12 (x_{12}), Timbal (Pb) mewakili neuron 13 (x_{13}), Nitrat ($\text{NO}_3\text{-N}$) mewakili neuron 14 (x_{14}), Nitrit ($\text{NO}_2\text{-N}$) mewakili neuron 15 (x_{15}), Minyak dan Lemak mewakili neuron 16 (x_{16}), Klorida (Cl) mewakili neuron 17 (x_{17}), Aluminium (Al) mewakili neuron 18 (x_{18}), Kesadahan (CaCO_3) mewakili neuron 19 (x_{19}), Sulfida (H_2S) mewakili neuron 20 (x_{20}), Arsen (As) mewakili neuron 21 (x_{21}) dan Sianida (CN) mewakili neuron 22 (x_{22}). Maka jumlah dari semua neuron input adalah sebanyak 22 unit neuron. Namun nilai input tidak langsung di proses oleh algoritma LVQ akan tetapi nilai itu akan dinormalisasi terlebih dahulu agar kesenjangan antar data tidak terpaut jauh. Semua nilai parameter akan dinormalisasi ke dalam rentang 0 - 1. Normalisasi dirumuskan sebagai berikut. [9].

$$\text{Nilai Baru} = \left(\frac{\text{nilai asal} - \text{nilai minimal}}{\text{nilai maksimal} - \text{nilai minimal}} \right) \dots (1)$$

Dalam penelitian ini LVQ diminta untuk dapat mengenali 4 tingkat pencemaran air sungai yaitu, memenuhi baku mutu (y_1), tercemar ringan (y_2), tercemar sedang (y_3) dan tercemar berat (y_4), sehingga LVQ mempunyai 4 kelas output. Setiap unit input mempunyai bobot (w) pada kelasnya masing-masing, jadi bobot pada penelitian ini sebanyak 22 unit untuk setiap kelas LVQ. Karena penelitian akan dilakukan untuk 4 kelas pencemaran air (Kelas I, II, III dan IV) maka banyak bobot adalah sebanyak $22 \times 4 \times 4$. Data yang akan dijadikan data pembelajaran pada aplikasi ini berjumlah 100 unit dimaksudkan agar aplikasi dapat mengenali pola dengan baik.

Arsitektur LVQ ini dapat digambarkan seperti gambar berikut.



Gambar 1. Arsitektur Aplikasi LVQ Kualitas Air

Teknisnya dalam pembelajaran atau pelatihan data ada beberapa tahap-tahap yang dilakukan untuk mendapatkan hasil akhir, tahap-tahap tersebut sebagai berikut.[8]

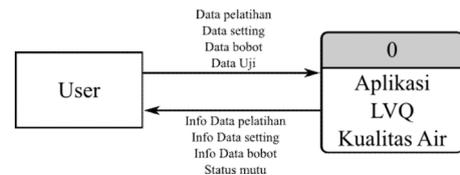
- a. Tetapkan :
 - 1) bobot awal awal variabel input ke-j menuju ke kelas (cluster) ke-i: W_{ij} , dengan $i = 1,2,\dots,K$; dan $j = 1,2,\dots,m$.
 - 2) Parameter learning rate (α), pengurangan learning rate ($dec\alpha$), minimal learning rate ($min\alpha$) dan maksimal iterasi ($MaxIter$).
- b. Masukkan :
 - 1) Data input: x_{ij} , dengan $i = 1,2,\dots,n$; dan $j = 1,2,\dots,m$.
 - 2) Target berupa kelas: T_k , dengan $k = 1,2,\dots,n$.
- c. Tetapkan kondisi awal : $epoch = 0$.
- d. Kerjakan jika : ($\alpha \geq min\alpha$) dan ($epoch < MaxIter$)
 - 1) $Epoch = epoch + 1$.
 - 2) Kerjakan untuk $i = 1$ sampai n .
 - a) Tentukan J sedemikian hingga $\|x_i - w_j\|$ minimum, dengan $j = 1,2,\dots,k$ (2)
 - b) Perbaiki W_j dengan ketentuan.
 - Jika $T = C_j$ maka $w_j = w_j + \alpha (x_i - w_j)$... (3)
 - Jika $T \neq C_j$ maka $w_j = w_j - \alpha (x_i - w_j)$... (4)
 - 3) Kurangi nilai α . Pengurangan α dilakukan dengan cara $\alpha = \alpha * dec\alpha$.

Tahap diatas akan diulang 4 kali untuk kelas pencemaran air lainnya. Hasil keluar dari aplikasi berupa informasi tingkat pencemaran air pada ke empat kelas pencemaran air.

Tahap pengujian mempunyai kesamaan dengan tahap pengujian namun tahap ini lebih relatif singkat karena tidak ada inialisasi bobot awal, parameter seperti learning rate, pengurangan α , minimal α , maksimal iterasi dan perbaikan nilai bobot. Nilai bobot diambil dari hasil pembelajaran yang sebelumnya sudah dilakukan dan tersimpan di database. Untuk lebih jelasnya berikut adalah tahap-tahapnya.[8]

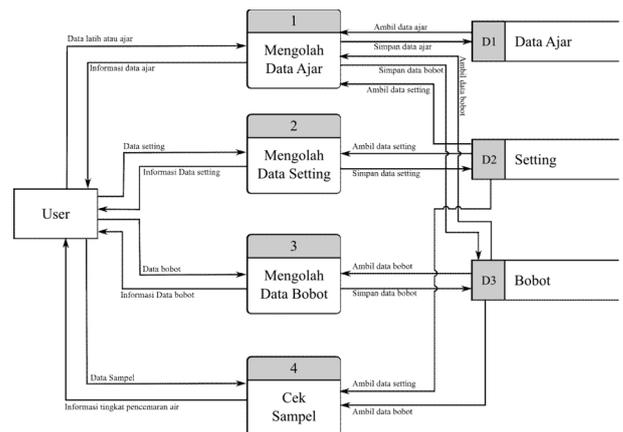
- a. Masukkan data yang akan diuji, misal : x_{ij} , dengan $i = 1,2,\dots,np$, dan $j = 1,2,\dots,m$.
- b. Kerjakan untuk $i = 1$ sampai np .
 - 1) Tentukan J sedemikian hingga $\|x_i - w_j\|$ minimum. Dengan $j = 1,2,\dots,k$.
 - 2) J adalah kelas untuk x_i .

Aplikasi yang akan dibangun menggunakan web base dan bahasa pemrogramannya menggunakan PHP. Untuk penyimpanan data ajar, inialisasi dan data lainnya menggunakan MySQL. Berikut adalah alur atau proses aplikasi secara garis besar yang memperlihatkan masukan, proses, dan keluaran dari sistem yang akan dirancang. Aplikasi akan menghasilkan keluaran berupa tingkat pencemaran air dari data sampel yang di inputkan.



Gambar 1. Diagram Konteks Aplikasi LVQ Kualitas Air

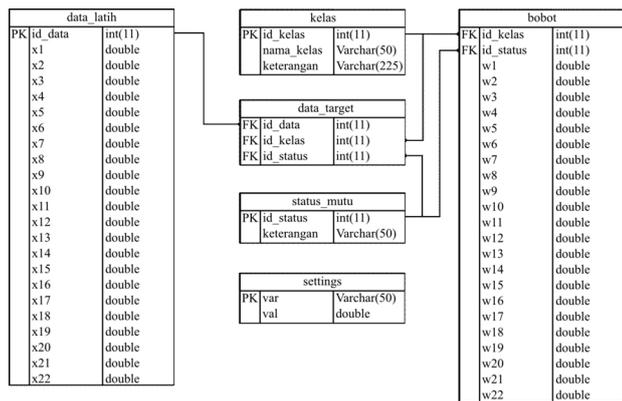
Gambar 1 diatas merupakan diagram konteks dari aplikasi. User adalah entitas luar yang dapat mengakses aplikasi. Diagram ini akan dikembangkan lagi ke DFD level 0 seperti gambar berikut.



Gambar 2. DFD Level 0

Gambar 2 merupakan DFD level 0 dari aplikasi. Diagram ini adalah pengembangan dari diagram konteks.

Berikut adalah relasi antar tabel aplikasi LVQ yang menjadi database penyimpanan data.



Gambar 3. Relasi Antar Tabel

Pada Gambar 3 berisikan tentang tabel-tabel yang digunakan dan relasi diantara tabel-tabel tersebut. Tabel-tabel yang dibuat berelasi satu sama lainnya seperti tabel data_latih dengan data target, tabel kelas dengan bobot dan data_target dan tabel status_mutu dengan data_target dan bobot, hanya tabel setting yang tidak berelasi dengan tabel lainnya.

Setelah aplikasi selesai dirancang dan dibangun maka tahap selanjutnya adalah tahap pengoperasian aplikasi dari pembelajaran hingga pengujian.

Gambar 4. Form Settings (inisialisasi)

Pada tahap pembelajaran, penelitian ini nilai learning rate yang diberikan pada inisialisasi adalah 0,01. Pengurangan learning rate sebesar 0,001. Minimal learning rate 0 dan maksimal iterasi 20.000.

Gambar 5. Aplikasi memproses 100 data

Setelah proses pembelajaran 100 data selesai, maka dihasilkan bobot akhir untuk ke empat kelas. Bobot inilah yang akan digunakan pada tahap pengujian untuk mengenali tingkat pencemaran air.

Gambar 6. Aplikasi memproses pengujian

Pengujian dilakukan untuk menilai seberapa akuratnya aplikasi yang telah dibuat untuk mengenali tingkat pencemaran air. Pengujian dilakukan dengan data uji 50 unit dengan target yang sudah diketahui sebelumnya.

Tabel 1. Tabel hasil pengujian aplikasi

No	Nilai Kebenaran			
	Kelas I	Kelas II	Kelas III	Kelas IV
1	Benar	Benar	Benar	Salah
2	Salah	Benar	Benar	Benar
3	Benar	Benar	Benar	Benar
4	Benar	Salah	Benar	Benar
5	Benar	Salah	Salah	Benar
6	Salah	Benar	Benar	Benar
7	Benar	Salah	Salah	Benar
8	Benar	Salah	Salah	Benar
9	Salah	Benar	Benar	Salah
10	Benar	Benar	Benar	Benar
11	Benar	Benar	Benar	Salah
12	Benar	Benar	Benar	Salah
13	Benar	Benar	Benar	Benar
14	Benar	Benar	Benar	Benar
15	Salah	Benar	Benar	Benar
16	Benar	Benar	Benar	Salah
17	Benar	Benar	Benar	Salah
18	Benar	Benar	Benar	Benar
19	Benar	Benar	Benar	Salah
20	Salah	Salah	Salah	Benar
21	Benar	Benar	Benar	Salah
22	Benar	Benar	Benar	Salah
23	Benar	Benar	Benar	Salah
24	Benar	Benar	Benar	Benar
25	Salah	Salah	Benar	Benar
26	Salah	Salah	Benar	Benar
27	Benar	Benar	Benar	Benar
28	Benar	Benar	Benar	Benar
29	Benar	Benar	Benar	Benar
30	Salah	Benar	Benar	Benar
31	Benar	Benar	Benar	Benar
32	Benar	Benar	Salah	Benar
33	Benar	Benar	Benar	Benar
34	Benar	Benar	Benar	Benar
35	Benar	Benar	Benar	Benar
36	Benar	Benar	Salah	Salah
37	Benar	Benar	Benar	Benar
38	Salah	Benar	Benar	Salah

39	Benar	Benar	Benar	Benar
40	Benar	Benar	Benar	Benar
41	Benar	Salah	Benar	Salah
42	Salah	Salah	Salah	Benar
43	Benar	Salah	Benar	Benar
44	Benar	Salah	Benar	Salah
45	Salah	Salah	Benar	Benar
46	Benar	Salah	Benar	Benar
47	Benar	Salah	Benar	Benar
48	Salah	Benar	Benar	Salah
49	Benar	Benar	Benar	Benar
50	Benar	Benar	Benar	Benar
Benar	38	36	43	35
Salah	12	14	7	15

Tabel 1 diatas merupakan hasil pengujian pada 50 unit data yang menghasilkan nilai kebenaran aplikasi dalam mengklasifikasikan data masukan.

Setelah diketahui jumlah nilai kebenaran pada setiap kelas maka selanjutnya adalah menilai keakuratan untuk setiap kelasnya. Penghitungan persentase akurasi per kelas dirumuskan sebagai berikut.

$$\text{Persentase Akurasi} = \left(\frac{\text{banyak data benar}}{\text{banyak data uji}} \right) \times 100 \quad \dots(5)$$

Dari hasil pengujian menunjukkan bahwa keakuratan aplikasi untuk pengenalan tingkat pencemaran air pada kelas I baru mencapai 76%, kelas II 72%, kelas III 86%, kelas IV 70%.

Kemudian semua persentase dijumlahkan dan dicari rata-rata ke empat kelas menjadi persentase akhir. Penghitungan persentase akhir dirumuskan sebagai berikut.

$$\text{Persentase Akhir} = \left(\frac{\text{jumlah persentase}}{\text{banyak kelas}} \right) \quad \dots(6)$$

Maka diketahui rata-rata keakuratan aplikasi hanya mencapai 76%.

3. Kesimpulan

Kesimpulan pada penelitian ini adalah algoritma Learning Vector Quantization dapat diterapkan pada pengklasifikasian tingkat pencemaran air sungai. Dengan learning rate 0,01, pengurangan learning rate 0,001 dan maksimal iterasi 20.000 mendapatkan keakuratan aplikasi sebesar 76%. Nilai ini tidak bersifat tetap, nilai ini dapat dipengaruhi oleh perubahan nilai learning rate, deca, mina dan maksimal iterasi pada saat pembelajaran. Nilai ini juga dipengaruhi oleh banyaknya parameter masukan, semakin banyak data pembelajaran diharapkan aplikasi dapat mengenali lebih banyak pola pengklasifikasian air sungai yang dapat meningkatkan nilai akurasi aplikasi.

Daftar Pustaka

- [1] Ministerie van Marine (Afdeling Hydrografie), "Zeemansgids voor den Oost-Indischen Archipel Jilid III", Mouton & Co, 'sGravenhage, 1973.
- [2] Peraturan Gubernur Provinsi Kalimantan Selatan Nomor 05 Tahun 2007 tentang Peruntukan dan Baku Mutu Air Sungai. 29 Januari 2007. Banjarmasin.
- [3] Badan Pusat Statistik (BPS). Persentase Rumah Tangga Menurut Provinsi dan Sumber Air Minum 2000-2013. Berbagai Tahun Penerbitan.
- [4] Badan Pusat Statistik (BPS) Provinsi Kalimantan Selatan. Indeks Pembangunan Manusia. Berbagai Tahun Penerbitan.
- [5] Badan Pusat Statistik. Berita Resmi Statistik BPS Provinsi Kalimantan Selatan No.28/05/63/Th.XIX, 04 Mei 2015.
- [6] Greenpeace, "Terungkap:Tambang Batubara Meracuni Air Kalimantan Selatan", Greenpeace Indonesia, Jakarta, 2014.
- [7] Undang-Undang Republik Indonesia Nomor 32 Tahun 2009 tentang Perlindungan dan Pengelolaan Lingkungan Hidup. 3 Oktober 2009. Lembaran Negara Republik Indonesia Tahun 2009 Nomor 140. Jakarta.
- [8] Kusumadewi, S. & Hartati, S., "Neuro-Fuzzy Integrasi Sistem Fuzzy & Jaringan Syaraf, Edisi 2", Graha Ilmu, Yogyakarta, 2010.
- [9] Atthina, N., Iswari, I., "Klasterisasi Data Kesehatan Penduduk untuk Menentukan Rentang Derajat Kesehatan Daerah dengan Metode K-Means". *Jurnal SNASTI*, 2015, B-52 - B-59

Biodata Penulis

Muhammad Ridha Rahimi, saat ini menjadi Mahasiswa di STMIK AMIKOM Yogyakarta.

Hartatik, S.T., M.Cs., memperoleh gelar Sarjana Teknik (S.T), Jurusan Teknik Informatika Universitas Muhammadiyah Cirebon, lulus tahun 2006. Memperoleh gelar *Master of Computer Science* (M.Cs) Program Pasca Sarjana Ilmu Komputer Universitas Gajah Mada Yogyakarta, lulus tahun 2011. Saat ini menjadi Dosen di STMIK AMIKOM Yogyakarta.

