

Pengembangan Metode Pengelompokan Obyek Penutup Lahan Dan Parameter Sosial-Ekonomi Untuk Daerah Kalimantan Barat

Oleh:

W. Setiawan , A. Murni, B. Kusumoputro, dan D. Hardianto

1. PENDAHULUAN

Penelitian ini bertujuan untuk melakukan perencanaan pemanfaatan lahan berdasarkan analisis potensial lahan dan kondisi sosial-ekonomi. Sumber data yang akan digunakan adalah data sensor Landsat ETM (*Land Satellite Enhanced Thematic Mapper*) yang terdiri dari 7 band multispektral dan 1 band pankromatik serta data sosial-ekonomi yang dikumpulkan dari lapangan. Analisis utama yang akan dilakukan adalah untuk tujuan perencanaan area hutan, permukiman, dan daerah pesisir.

Tiga dasar pertimbangan telah dijadikan dasar dalam pendekatan pengembangan metode pengelompokan. Ketiga dasar pertimbangan tersebut meliputi: (i) penggunaan pendekatan kombinasi metode klasifikasi tanpa pengarahan (*unsupervised*) dan dengan pengarahan (*supervised*); (ii) penggunaan pendekatan jaringan syaraf tiruan untuk metode klasifikasi; dan (iii) penggunaan asumsi distribusi Gaussian untuk data Landsat ETM.

Dasar pertimbangan yang pertama adalah penggunaan kombinasi pengenalan obyek tanpa pengarahan (*unsupervised*) dan pengelompokan obyek dengan pengarahan (*supervised*). Pengguna paket komersial seperti ER Mapper dan ERDAS juga sering menggunakan pendekatan kombinasi ini melalui fasilitas *clustering (unsupervised)* dan fasilitas klasifikasi dengan set sample pembelajaran (*supervised*) seperti metode *maximum-likelihood* atau metode *minimum-distance*. Menurut pengalaman, pendekatan ini memang dapat memperbaiki hasil klasifikasi yang hanya menggunakan pendekatan *supervised*.

Dasar pertimbangan yang kedua adalah penggunaan jaringan syaraf tiruan (*artificial neural network*). Metode klasifikasi (pengelompokan) dengan *neural network* dipilih, karena dianggap dapat lebih baik dalam menyelesaikan masalah-masalah *uncertainties* dibandingkan metode yang konvensional seperti pendekatan statistik. Jaringan *Self-Organizing Mapper* (SOM) merupakan pendekatan *unsupervised*, mempunyai kemampuan yang cepat untuk konvergen tetapi kurang mempunyai

kemampuan pengenalan. Sebaliknya jaringan *Back Propagation* (BP) merupakan pendekatan *supervised* yang lambat konvergen tetapi mempunyai kemampuan pengenalan yang lebih tinggi.

Dasar pertimbangan yang ketiga adalah penggunaan model Gaussian untuk data citra sensor optik. Citra Landsat ETM termasuk sensor optik dan pengklasifikasi Gaussian maximum likelihood telah secara luas digunakan.

Metode pengelompokan data sosial-ekonomi yang diusulkan adalah metode hibrida jaringan SOM (*Self-Organizing Map*) yang bersifat *unsupervised* dan jaringan BP (*Back Propagation*) yang bersifat *supervised*. Jaringan BP ini menggunakan keluaran dari SOM sebagai sampel-sampel yang sudah dikelompokkan.

Metode pengelompokan yang digunakan untuk klasifikasi citra Landsat ETM adalah jaringan PNN (*Probabilistic Neural Network*) dengan model Gaussian yang dioptimisasi dengan algoritma EM (*Expectation-Maximum*). Sebagai perbandingan metode hibrida SOM-BP juga digunakan untuk klasifikasi citra Landsat ETM.

2. PERMASALAHAN DAN TUJUAN PENELITIAN

Tujuan penelitian ini adalah menghasilkan suatu sistem perencanaan tata wilayah dan pengembangan daerah dengan analisis ambang batas yang berdasarkan analisis potensial lahan dan potensial sosial ekonomi.

Fokus penelitian pada periode tahun 2003/2004 adalah pengembangan metode klasifikasi penutup lahan untuk citra Landsat ETM dan metode klasifikasi parameter sosial ekonomi. Fokus penelitian pada periode tahun 2004/2005 adalah menerapkan atau melakukan validasi metode tersebut sesuai dengan kelas obyek penutup lahan yang dibutuhkan untuk analisis. Pada proses validasi dan penyempurnaan sistem akan ditentukan obyek-obyek yang ingin dikenali dan sampel-sampel yang lebih akurat dan diperoleh dari pakar. Selain itu pada studi lebih lanjut diharapkan data untuk parameter sosial ekonomi sudah tersedia.

Pengembangan jaringan hibrida SOM dan BP akan digunakan untuk klasifikasi parameter sosial ekonomi dan untuk klasifikasi citra Landsat ETM. Sedangkan jaringan PNN Gaussian-EM akan digunakan untuk

klasifikasi citra Landsat ETM. Analisis kinerja dan tingkat generalisasi metode akan dilakukan dengan indikator kesalahan klasifikasi berdasarkan *user's accuracy*, *producer's accuracy* dan statistik Kappa.

3. METODE KLASIFIKASI DATA BERDASARKAN JARINGAN HIBRIDA SOM-BP

Untuk mendapatkan keuntungan-keuntungan dari *Self-Organized Mapper* (SOM) dan *Back Propagation* (BP) agar jaringan dapat cepat konvergen dan mempunyai kemampuan pengenalan yang tinggi, maka akan digunakan suatu jaringan yang terdiri dari modul yang *self-organized* dan modul *back propagation* sebagai supervisor. Berikut dibahas arsitektur jaringan hibrida SOM-BP serta cara bekerjanya.

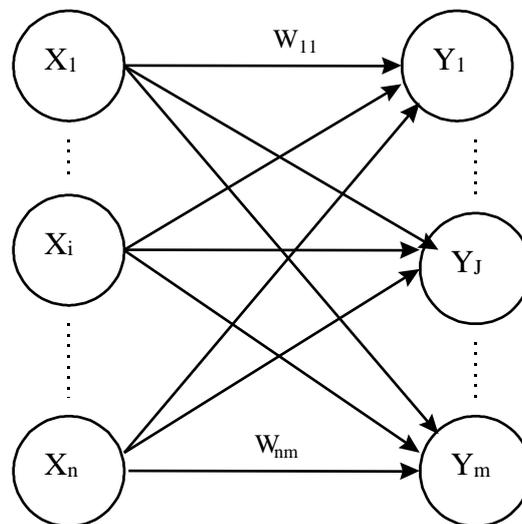
Pelatihan dan pengenalan masih belum dilakukan dengan data sebenarnya, karena data belum tersedia. Rencananya data sosial ekonomi yang akan digunakan meliputi data tenaga kerja, parameter ekonomi dan parameter lingkungan. Data yang sementara digunakan adalah data serupa yang terdiri dari 6 (enam) parameter yang meliputi data sosial, ekonomi dan lingkungan.

3.1. Konsep Jaringan Self-Organizing Mapper (SOM)

Jaringan SOM merupakan jaringan neural buatan yang banyak diterapkan pada persoalan klastering dengan kemampuan melakukan pembelajaran dalam tempo yang singkat. Dalam persoalan klastering, jaringan SOM standar tidak mengenal kemampuan beradaptasi dengan pola masukan. Setiap pola masukan akan dipaksa untuk masuk ke salah satu klaster yang telah ada walaupun sebenarnya pola masukan baru jauh dari klaster yang ada. Berikut dibahas arsitektur, proses pembelajaran dan algoritma jaringan SOM.

3.1.1. Arsitektur Jaringan SOM

Arsitektur jaringan SOM dapat digambarkan dengan skema diagram pada Gambar 1 berikut ini.



Gambar 1. Arsitektur jaringan SOM.

Jaringan SOM merupakan jaringan lapis tunggal yang mem-punyai satu lapis hubungan bobot (W). Neuron dapat dibedakan atas neuron masukan (X) yaitu neuron yang menerima sinyal dari luar dan neuron keluaran (Y) yang merupakan neuron respon dari jaringan. Setiap vektor masukan berkorespondensi dengan satu neuron keluaran dan menyimpan karakteristik atau sifat tertentu yang dimiliki vektor masukan tersebut. Pada suatu saat hanya ada satu neuron keluaran yang aktif, sedangkan neuron keluaran yang lain bersifat pasif. Neuron yang aktif disebut dengan neuron pemenang.

Pada jaringan SOM standar, jumlah neuron di lapis keluaran ditentukan lebih dulu, dan tergantung pada jumlah kluster yang diinginkan, seperti algoritme *K-Mean clustering*. Pada penelitian ini jumlah kluster dilakukan secara adaptif dengan menentukan suatu parameter nilai ambang jarak antar kluster. Nilai ambang ini berpengaruh pada jumlah kluster yang akan diperoleh.

3.1.2. Pembelajaran Jaringan SOM

Pembelajaran pada jaringan SOM merupakan pembelajaran kompetisi. Dalam pembelajaran ini, neuron-neuron keluaran berkompetisi diantara

mereka untuk menjadi neuron pemenang. Berbeda dengan pembelajaran supervisi dimana semua neuron keluaran aktif bersama-sama, pada pembelajaran kompetisi hanya satu neuron yang aktif pada satu saat.

Setiap neuron secara khusus mempelajari suatu set pola tertentu dan menjadi pendeteksi karakteristik pola tersebut. Misalkan neuron yang menjadi pemenang adalah neuron ke- j , nilai aktivasi neuron (y_j) lebih besar dibanding neuron lainnya untuk suatu pola masukan x . Jika sebuah neuron tidak merespon terhadap suatu pola tertentu (neuron yang kalah), pembelajaran tidak dilakukan pada neuron tersebut.

Misalkan w_{ij} menyatakan bobot dari neuron masukan i ke neuron keluaran j , maka koreksi bobot (Δw_{ij}) yang diterapkan pada bobot w_{ij} adalah:

$$\Delta w_{ij} = \begin{cases} \alpha (y_j - w_{ij}) & \text{jika neuron ke } j \text{ menang} \\ 0 & \text{jika neuron ke } j \text{ kalah} \end{cases}$$

Konvergensi jaringan dicapai dengan cara menurunkan nilai α yang merupakan kecepatan pembelajaran jaringan untuk setiap epochnya. Karena hanya ada satu neuron yang aktif pada satu saat, maka harus ditentukan vektor masukan x yang sesuai dengan vektor bobot w_j dengan cara mencari jarak yang paling minimum antara vektor bobot dan vektor masukan. Cara untuk menentukan jarak yang paling minimum bermacam-macam, diantaranya *inner product* dan *jarak euclidean*. Nilai inner product ($w_j^T x$) untuk $j = 1, 2, \dots, N$ yang paling maksimum menyatakan similaritas atau kemiripan yang paling tinggi antara vektor masukan dengan pusat kluster yang dinyatakan dengan vektor bobot. Hasil inner product menentukan pada neuron mana proses komputasi nilai aktivasi akan dilakukan. Proses mencari nilai inner product yang paling *maksimum* setara dengan proses mencari nilai euclidean yang paling *minimum*.

Setiap vektor masukan dilokalisasi pada suatu wilayah tertentu. Misalkan vektor masukan adalah:

$$x_i = [x_1, x_2, \dots, x_p]^T \quad p = \text{jumlah neuron masukan}$$

Vektor bobot dari neuron masukan ke neuron keluaran dinyatakan dengan

$$w_{ij} = [w_{i1}, w_{i2}, \dots, w_{in}]^T, \quad n = \text{jumlah neuron keluaran}$$

Pada jaringan SOM standar biasanya digunakan jarak euclidean yang paling minimum (minimum-euclidean distance norm). Misalkan $i(x)$ menyatakan neuron yang paling sesuai untuk vektor masukan x , secara matematis dapat dinyatakan :

$$i(x) = \arg j \min \|x(n) - w_j\|, \quad j = 1, 2, \dots, N$$

dimana $\|\cdot\|$ menyatakan euclidean norm dari argumen vektor.

3.1.3. Algoritma Pengklasifikasi SOM

Secara ringkas algoritma jaringan SOM dapat dirinci sebagai berikut:

Langkah 0: Inisialisasi bobot-bobot w_{ij} (biasanya random antara 0 s/d 0.5)

Langkah 1: Jika kondisi henti gagal, lakukan langkah 2-8.

Langkah 2: Untuk setiap vektor masukan x , lakukan langkah 3 sampai 5

Langkah 3: Untuk setiap j , hitung:

$$D(j) = \sum_i (w_{ij} - x_i)^2$$

Langkah 4: Temukan indeks j sehingga $D(j)$ minimum

Langkah 5: Untuk setiap neuron j dan untuk semua neuron i :

$$w_{ij}(\text{baru}) = w_{ij}(\text{lama}) + \alpha (x_i - w_{ij}(\text{lama}))$$

Langkah 6: Perbaiki kecepatan pembelajaran (mulai dengan 0.5 dan turunkan 0.01)

Langkah 7: Kurangi radius ketetanggaan klaster bila diperlukan

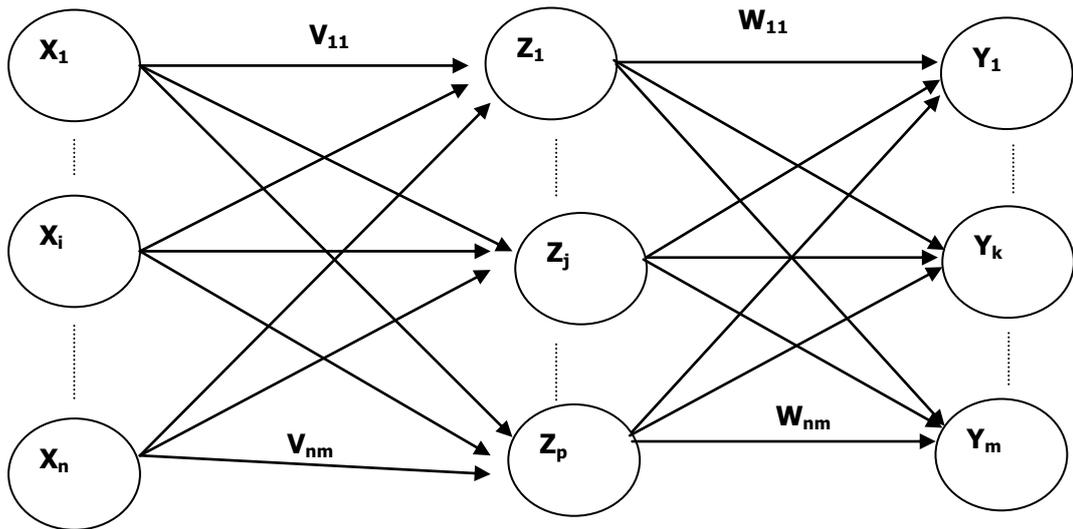
Langkah 8: Tes kondisi henti

3.2. Konsep Jaringan Back Propagation (BP)

Propagasi balik adalah salah satu jenis pembelajaran untuk arsitektur jaringan perseptron lapis jamak. Arsitektur jaringan perseptron lapis jamak, mempunyai lebih dari satu lapis tersembunyi, namun yang akan dibicarakan disini hanya dengan satu lapis tersembunyi. Berikut dibahas arsitektur, proses pembelajaran dan algoritma jaringan BP.

3.2.1. Arsitektur Jaringan BP

Arsitektur jaringan BP dapat digambarkan dengan skema seperti dapat dilihat di Gambar 2.



Gambar 2. Arsitektur Jaringan BP.

3.2.2. Pembelajaran Jaringan BP

Proses pembelajaran BP terdiri dari 3 fase:

- 1) fase maju,
- 2) fase balik, dan
- 3) fase memperbaiki bobot.

Setiap neuron masukan (X_i) menerima sinyal masukan dan menyebarkannya ke setiap neuron pada lapis tersembunyi ($Z_1 \dots Z_p$). Setiap neuron pada lapis tersembunyi menghitung aktivasinya dan mengirim sinyalnya (z_j) ke setiap neuron keluaran. Neuron keluaran (Y_k) menghitung aktivasinya (y_k) membentuk respon jaringan terhadap suatu pola masukan yang diberikan.

Selama pembelajaran, setiap neuron keluaran membandingkan aktivasinya (y_k) dengan target (t_k) besar kesalahan aktivasi jaringan terhadap target. Berdasarkan kesalahan ini, faktor δ_k ($k=1 \dots m$) dihitung. δ_k digunakan untuk mendistribusikan kesalahan pada neuron keluaran Y_k ke semua neuron pada lapis sebelumnya (lapis tersembunyi). δ_k juga akan digunakan untuk memperbaiki bobot antara lapis keluaran dan lapis tersembunyi. Dengan cara yang sama, faktor δ_j ($j=1 \dots p$) dihitung untuk setiap unit pada lapis tersembunyi Z_j .

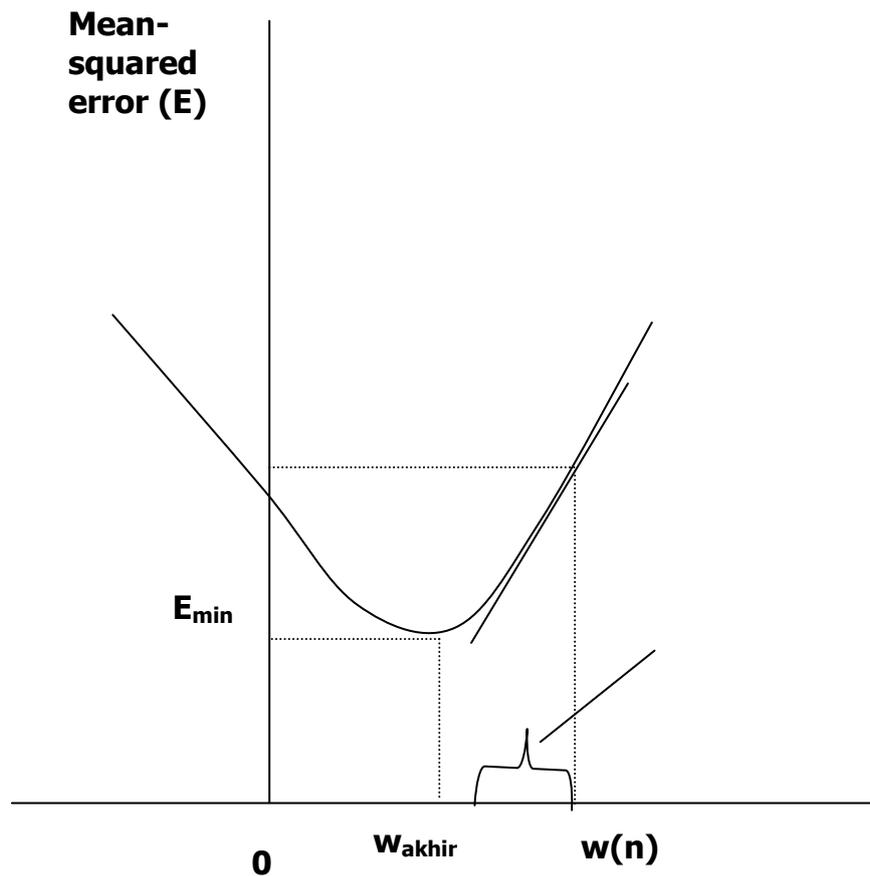
Tidak perlu untuk mempropagasikan balik kesalahan sampai pada lapis masukan, tapi δ_j digunakan untuk memperbaiki bobot antara lapis tersembunyi dan lapis masukan. Setelah semua faktor δ ditentukan, bobot

untuk semua lapis diperbaiki secara serentak. Perbaiki bobot w_{jk} (dari neuron di lapis tersembunyi ke neuron di lapis keluaran) berdasarkan faktor δ_k dan aktivasi (z_j) lapis tersembunyi. Perbaiki bobot v_{ij} (dari neuron masukan ke neuron lapis tersembunyi) berdasarkan faktor δ_j dan aktivasi x_i dari neuron masukan.

3.2.3. Metode Gradient Descent

Jaringan propagasi balik mempunyai paradigma pembelajaran dengan pengarahannya dan algoritma koreksi kesalahan. Setiap pola masukan mempunyai pasangan dengan suatu target yang telah ditentukan. Ide pembelajaran pada jaringan propagasi balik membuat keluaran jaringan menjadi semirip mungkin dengan targetnya, atau kesalahan (perbedaan) antara keluaran jaringan dan target menjadi sekecil mungkin. Metode yang digunakan untuk memperbaiki bobot disebut metode gradient descent.

Dengan metode ini, kesalahan merupakan fungsi dari bobot. Bobot yang diubah secara iteratif akan menyebabkan kesalahan makin kecil. Diperlukan untuk mencari bobot dimana nilai kesalahan paling minimum. Jika bobot lebih dari satu maka didapatkan rata-rata kesalahan (mean-squared error). Untuk satu pola masukan p , didapatkan kesalahan sebesar E_p yang merupakan fungsi bobot, maka kesalahan untuk semua pola adalah $E = E_p(w_1, w_2, \dots, w_n)$, biasanya didefinisikan sebagai kuadrat selisih antara keluaran (y_p) dan target (t_p).



Gambar 3 : Grafik Kesalahan

Maka :

$$E = \frac{1}{2} \sum_{p=1}^n (x_p - y_p)^2$$

Dalam mencari titik terendah dari kurva kesalahan, koreksi bobot adalah gradient kurva dengan arah yang berlawanan dengan vektor bobot.

Didefinisikan:

$$\Delta w = -\eta \frac{\delta E}{\delta w}$$

η : konstanta positif.

3.2.4. Konvergensi

Algoritma propagasi balik adalah pendekatan turunan pertama dari tehnik *steepest descent*, dimana ia tergantung pada gradien permukaan kesalahan pada ruang bobot pada satu waktu tertentu. Algoritma bersifat stochastic, sehingga ia cenderung berjalan zigzag untuk mencari arah yang benar dalam menentukan permukaan kesalahan yang paling minimum.

Pembelajaran propagasi balik merupakan aplikasi metode statistik yang disebut *stochastic approximation* yang sifatnya adalah konvergensi yang lambat dan mengakibatkan komputasi menjadi mahal. Ada 2 hal yang menyebabkan kelambatan ini:

1. Pemilihan permukaan kesalahan agak datar terhadap dimensi bobot.
2. Arah vektor gradien negatif yang jauh dari permukaan kesalahan minimum.

3.2.5. Local Minima

Selain memiliki sifat kelambatan dalam proses pembelajaran, unjuk kerja propagasi balik juga dipengaruhi oleh keberadaan lokal minima dan global minima. Lokal minima yaitu keadaan dimana bobot bernilai minimum akan tetapi perubahan kecil pada bobot akan menambah fungsi kesalahan. Sedangkan global minima yaitu sekumpulan bobot yang bernilai lebih kecil dari lokal minima namun jaringan seperti mengalami sangkutan. Sampai saat ini belum ada penelitian yang dapat mengatasi lokal minima maupun global minima.

3.2.6. Algoritma Jaringan BP

Secara ringkas algoritma BP dapat dirinci sebagai berikut:

Langkah 0: Inisialisasi bobot

Langkah 1: Jika kondisi henti tidak terpenuhi, lakukan langkah 2 - 9

Langkah 2: Untuk setiap pola input, lakukan step 3 – 8

Feedforward:

Pembelajaran pada layer 1

Langkah 3: Setiap neuron masukan (X_i $i= 1\dots n$) menerima sinyal masukan x_i dan menyebarkannya ke semua neuron pada lapis tersembunyi.

Langkah 4: Setiap neuron pada lapis tersembunyi (Z_j , $j=1\dots p$) menjumlahkan bobot-bobot sinyal masukan,

$$z_in_j = \sum_{i=1}^n x_i v_{ij}$$

dan menerapkannya pada fungsi aktivasi untuk menghitung sinyal keluarannya,

$$z_j = f(z_in_j)$$

dan mengirim sinyal ini ke semua neuron pada lapis keluaran

Langkah 5: Setiap neuron keluaran (Y_k , $k=1\dots m$) menjumlahkan bobot sinyal yang masuk,

$$y_in_k = \sum_{j=1}^p z_j w_{jk}$$

dan mengaplikasikan fungsi aktivasinya untuk menghitung sinyal yang akan dikeluarkannya

$$y_k = f(y_in_k)$$

Backpropagasi error

Langkah 6: Setiap neuron keluaran (Y_k , $k=1\dots m$) menerima sebuah pola target yang berhubungan dengan pola masukan pelatihan dan menghitung kesalahan informasi dengan mengalikan dengan turunan fungsi aktivasinya

$$\delta_k = (t_k - y_k) f'(y_in_k)$$

menghitung koreksi bobot (yang akan digunakan untuk memperbaiki w_{jk} nanti),

$$\Delta w_{jk} = \alpha \delta_k z_j$$

dan mengirim δ_k ke neuron lapis tersembunyi

Langkah 7: Setiap neuron tersembunyi (Z_j , $j=1\dots p$) menjumlahkan bobot setiap neuron yang telah dikali dengan kesalahan informasinya,

$$\delta_in_j = \sum_{k=1}^m \delta_k w_{jk}$$

Langkah 7 (lanjutan):

mengalikan dengan turunan fungsi aktivasinya untuk menghitung kesalahan informasinya,

$$\delta_j = \delta_{in_j} f'(z_{in_j})$$

menghitung koreksi bobotnya (yang akan digunakan untuk memperbaiki v_j nanti)

$$\Delta v_{ij} = \alpha \delta_j x_i$$

Perbaiki bobot untuk proses selanjutnya

Langkah 8: Setiap neuron keluaran ($Y_k, k=1\dots m$) memperbaiki bobotnya ($j=0\dots p$)

$$w_{jk}(\text{new}) = w_{jk}(\text{old}) + \Delta w_{jk}$$

Setiap neuron tersembunyi ($Z_j, j=1\dots p$) memperbaiki bobotnya

$$v_{ij}(\text{new}) = v_{ij}(\text{old}) + \Delta v_{ij}$$

Langkah 9: Lakukan pengujian kondisi henti.

3.3. Konsep Jaringan Hibrida SOM dan BP

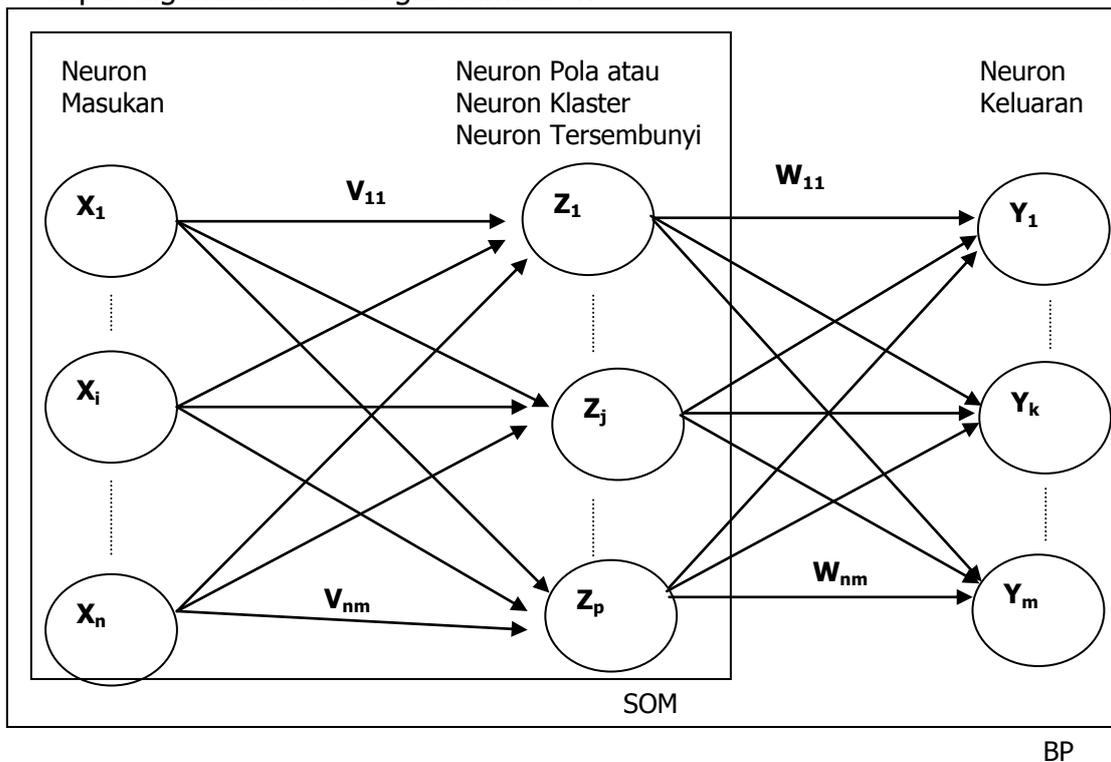
Jaringan hibrida SOM dan BP pertama kali diperkenalkan oleh Miyanaga et al., 1998). Masalah yang dihadapi pada penelitian tersebut adalah terjadinya pembengkakan dalam hal jumlah neuron klaster, sehingga menyulitkan proses back propagation. Pada jaringan SOM standar, jumlah neuron di lapis keluaran ditentukan lebih dulu, dan tergantung pada jumlah klaster yang diinginkan, seperti algoritme *K-Mean clustering*. Pada penelitian ini jumlah klaster tidak dibatasi. Untuk menghindari terjadinya pembengkakan jumlah neuron klaster, digunakan pendekatan jumlah neuron hidden yang lebih besar dari suatu konstanta k dikali jumlah neuron masukan. Pada penelitian ini mula-mula k diberi harga 2 dan ditambah sampai diperoleh hasil ketelitian yang cukup baik. Hasil eksperimen menunjukkan bahwa $k = 5$ telah memberikan hasil yang baik.

Jaringan hibrida ini terdiri dari 2 (dua) lapis, yaitu:

- Lapis pertama merupakan *self-organized layer* (bagian SOM) untuk data masukan

- Lapis kedua merupakan *supervised layer* (bagian BP) yang menentukan kelas suatu pola

Pada konsep jaringan hibrida SOM-BP, penggunaan jaringan SOM dimaksud untuk melakukan proses pengelompokan tahap awal. Hasil pengelompokan ini seolah-olah digunakan sebagai sample untuk proses klasifikasi dengan pendekatan *supervised* dengan jaringan BP. Hal tersebut dilakukan dengan menggunakan lapis keluaran dari SOM sebagai lapis tersembunyi dari jaringan BP. Arsitektur jaringan hibrida SOM-BP dapat digambarkan sebagai Gambar 4.



Gambar 4. Jaringan Hibrida SOM-BP.

3.3.1. Lapis Pertama (*Self-Organized Network / SOM*)

Penambahan dan penghapusan suatu neuron pola (klaster) berkaitan dengan percepatan konvergensi. Misalkan vektor data masukan dapat dinyatakan sebagai:

$$x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$$

dengan t suatu waktu instan, n jumlah sensor (atau feature) yang digunakan, dan T artinya transpose.

Setiap neuron kluster j pada SOM menyimpan 3 (tiga) buah informasi, yaitu:

- Jumlah keanggotaan q_j
- Pusat kluster $y_j(t)$
- Pusat matriks varian kesalahan (error) $S_j(t)$

dengan $e_j = x_j(t) - y_j(t)$ adalah perbedaan antara nilai data masukan dan nilai pusat kluster dan p adalah jumlah kluster, maka

$$S_j(t) = \text{diag}[\sigma_{j,1}^2(t), \sigma_{j,2}^2(t), \dots, \sigma_{j,p}^2(t)]$$

Tingkat keserupaan (similarity value) μ_j antara data masukan dan pusat kluster dapat dinyatakan dengan jarak Mahalanobis M_j sebagai berikut:

$$\mu_j(t) = \exp\left(-\frac{M_j}{2n}\right) \quad \text{dan} \quad M_j = \alpha e_j(t)^T S_j^{-1} e_j(t)$$

dimana $0 < \alpha < 1$ merupakan rectification factor.

3.3.2. Lapis Kedua (*supervised perceptron / BP*)

Supervised layer bertugas untuk melakukan proses pengenalan. Lapis ini mempunyai struktur yang sama dengan struktur output dari multi-layered perceptron neural network. Input dari lapis ini adalah nilai similarity μ_j yang diperoleh dari *self-organized layer*. Output activation dihitung dengan menggunakan sigmoid activation function dan besarnya kesalahan (error) ditentukan berdasarkan prosedur updating bobot koneksi. Artinya algoritme supervised layer sama dengan algoritme *back propagation* untuk layer keluaran. Setelah bobot-bobot di-update pada *supervised layer*, vektor kluster pada *self-organized module* juga diupdate dengan propagated error dari aktivasi neuron output dan vektor target (lihat prinsip algoritme BP).

Misalkan vektor target pada layer output adalah:

$$g(t) = [g_1(t), g_2(t), \dots, g_m(t)]$$

dimana m adalah jumlah kelas yang ditentukan.

Masukan ke supervised layer adalah nilai similarity yang diperoleh dari output SOM

$$\mu(t) = [\mu_1(t), \mu_2(t), \dots, \mu_m(t)]$$

Neuron activation pada layer output dihitung dengan

$$z(t) = [z_1(t), z_2(t), \dots, z_m(t)]$$

$$z_j(t) = \text{sigmoid}[\sum_{j=1}^m w_{jk}(t)\mu_j(t) + \theta_k(t)]$$

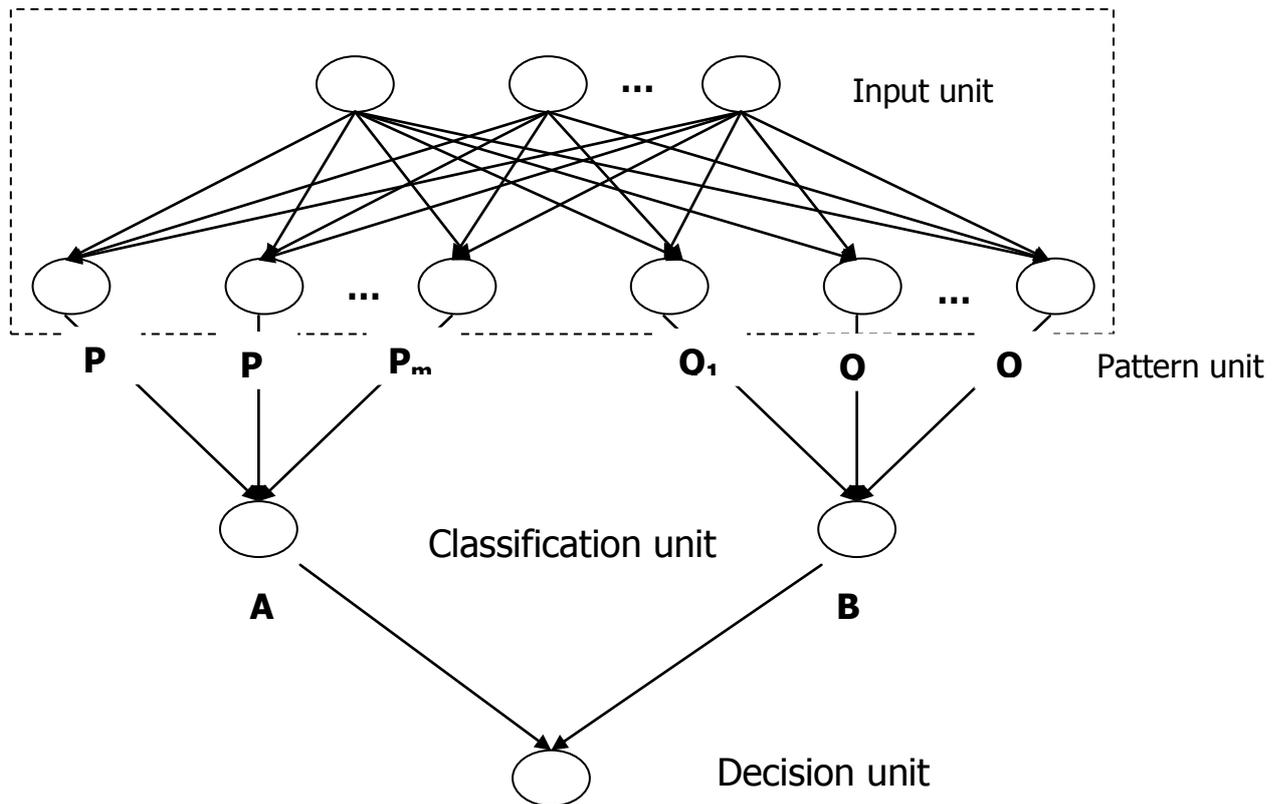
$$\text{sigmoid}(f) = \frac{1}{1 + e^{-f}}$$

dimana w_{jk} adalah bobot koneksi antara neuron kluster dan neuron output dan θ_k menyatakan suatu nilai ambang
Update bobot pada neuron output dan selanjutnya mengikuti algoritme backpropagation:

4. METODE KLASIFIKASI DATA BERDASARKAN JARINGAN PNN GAUSSIAN DENGAN OPTIMASI ALGORITME EM

4.1. Arsitektur Jaringan PNN Gaussian dengan EM

Skema jaringan *Probabilistic Neural Network* (PNN) Gaussian dapat dilihat di Gambar 5. Proses optimasi dengan algoritme *Expectation-Maximum* (EM) merupakan sub-sistem dari sistem PNN Gaussian secara keseluruhan. PNN Gaussian yang diterapkan adalah yang dideskripsi oleh Donald *et al.*, 1991. Algoritme EM yang diterapkan adalah sesuai dengan penjelasan Dempster *et al.*, 1997; Bruzzone *et al.*, 1999; dan Jackson *et al.*, 2001.



Gambar 5. Skema Jaringan PNN.

Algoritme PNN standar (Donald *et al.*, 1991) dapat dijelaskan sebagai berikut:

- Jaringan terdiri dari 3 (tiga lapis), yaitu: satu lapis unit masukan, satu lapis unit pola, dan satu lapis unit keluaran
- Terdapat korespondensi satu-satu antara unit pola dan vector data masukan.
- Setiap unit keluaran menyatakan satu kelas kategori onyek
- Unit pola terhubung ke unit keluaran jika dan hanya jika data yang bersangkutan termasuk kategori obyek yang sama
- Inisialisasi bobot
 - Vektor bobot dari unit pola ke- i di-inisialisasi dengan vektor input ke- i yang dinormalisasi menjadi bernilai antara 0 dan 1.
 - Vektor bobot dari unit keluaran di-inisialisasi dengan vector bernilai 1.
- Perhitungan aktvasi
 - Aktivasi dari unit masukan ditentukan oleh data masukan yang diberikan ke jaringan dan telah dinormalisasi
 - Aktivasi y_j dari unit pola j diberikan oleh:

$$O_j = \exp\left[\frac{\sum_i W_{ji} X_i - 1}{\sigma^2}\right]$$

Smoothing parameter diberi nilai 0.02. Dimana W_{ji} adalah bobot hubungan antara unit masukan ke- i dengan unit pola ke- j dan X_i adalah aktivasi dari input ke- i

- o Aktivasi O_j dari unit keluaran j diberikan oleh:

$$O_j = \frac{1}{m} P(w_j) \sum_i W_{ji} O_i$$

dimana W_{ji} adalah bobot hubungan antara unit pola ke- i dengan unit keluaran ke- j , dan m adalah jumlah data pelatihan per kelas yang diestimasi oleh w_k yang berhubungan dengan unit keluaran k , dan $P(w_k)$ adalah probabilitas awal $w_k = 1/K$ dimana K adalah jumlah kelas obyek.

- o Keputusan jaringan adalah kelas yang mempunyai unit keluaran dengan aktivasi maksimum

Pembelajaran Jaringan PNN adalah sebagai berikut:

- Nilai bobot tidak pernah diubah
- Ketika ada data baru masuk, sebuah unit pola baru ditambahkan dan bobot yang berhubungan di-inisialisasi dengan cara diatas
- Pelatihan jaringan sifatnya inisialisasi semua variabel, dan keputusan jaringan adalah kelas yang mempunyai unit keluaran dengan aktivasi maksimum

4.2. Optimisasi Dengan Algoritma Expectation-Maximum (EM)

Algoritme Expectation-Maximum (EM) merupakan metode iterative untuk mendapatkan estimator maximum likelihood dari data pengamatan yang tidak lengkap. Permasalahan yang ingin diselesaikan oleh EM adalah mengestimasi parameter $\theta = \bar{\theta}$ sehingga memaksimalkan fungsi logaritma likelihood $\log L(\theta | t) = f(t | \theta)$. Dalam hal ini algoritme EM akan membangkitkan barisan estimasi $\{\theta^{(m)}\}$ dari suatu nilai awal $\{\theta^{(0)}\}$.

Dalam setiap iterasi EM terdiri dari 2 (dua) tahap, yaitu (Dempster *et al.*, 1997):

1. Tahap ekspektasi: merupakan perumusan nilai ekspektasi fungsi likelihood untuk pengamatanyang cukup / lengkap sebagai:
 $E(\log L(\theta(t) | x, \theta^{(m)}) = Q(\theta, \theta^{(m)})$. Dalam hal ini X adalah vector pengamatan yang tidak lengkap yang bersesuaian dengan t .
2. Tahap maksimalisasi: nilai $\theta = \theta^{(m-1)}$ dirumuskan secara iterative, sehingga memaksimumkan $Q(\theta, \theta^{(m)})$. Perumusan nilai θ tergantung pada karakterisasi distribusi campuran kategori obyek yang teramati.

Algoritme EM dapat dinyatakan sebagai berikut:

1. Menentukan estimasi awal parameter $\theta^{(0)}$;
2. $m = 0$;
3. Evaluasi $Q(\theta, \theta^{(0)})$ (tahap ekspektasi)
4. Tentukan parameter $\theta^{(m+1)}$,
 sehingga $Q(\theta^{(m+1)}, \theta^{(m)}) \geq Q(\theta^{(m)}, \theta^{(m-1)})$
 untuk setiap vector θ (tahap maksimalisasi)
5. Jika konvergensi tercapai, stop, dan $\theta^{(m+1)}$ sebagai hasil estimasi. Jika belum lakukan langkah 6.
6. $m = m + 1$ dan kembali ke langkah 4.

Pada penelitian ini dilakukan optimisasi dengan EM untuk mendapatkan optimum probabilitas posterior. Mekanisme EM yang digunakan dapat dijelaskan sebagai berikut.

Distribusi pengamatan x dinyatakan sebagai (Jackson *et al.*, 2001):

$$f(x | \Phi) = \sum_{i=1}^L \alpha_i f_i(x | \phi_i)$$

dengan

- | | |
|---|------------------------------|
| $\alpha_1, \dots, \alpha_L$, | adalah probabilitas prior |
| $\Phi = (\alpha_1, \dots, \alpha_L, \phi_1, \dots, \phi_L)$ | merupakan parameter campuran |
| $x = (x_1, \dots, x_n)$ | adalah sample pelatihan |
| (ϕ_1, \dots, ϕ_L) | kategori kelas obyek |

Tahap ekspektasi dinyatakan dengan persamaan:

$$\tau_{ij} = \tau_i(x_j | \phi_{ij}^c) = \frac{\alpha_i f_i(x_j | \phi_{ij}^c)}{\sum_{i=1}^L \alpha_i f_i(x_j | \phi_{ij}^c)}$$

dimana τ_{ij} adalah probabilitas posterior

Tahap maksimalisasi dinyatakan dengan persamaan

$$\phi_i^+ \in \arg \max_{\phi \in \Omega} \left(\sum_{k=1}^L \tau_{ik}^c \ln f_i(x_k | \phi_i) \right)$$

$$\alpha_i^+ = \frac{\sum_{j=1}^L \tau_{ij}^c}{n}$$

Algoritma pelatihan adalah sebagai berikut:

1. Menentukan factor bobot

$$\omega_{ij}^c = \tau_i(x_j | \phi_{ij}^c) = \frac{f_i(x_j | \phi_i^c)}{\sum_{k=1}^L f_i(x_{ij} | \phi_k^c)}$$

2. Memaksimalkan fungsi log likelihood

$$\phi_i^+ \in \arg \max_{\phi \in \Omega} \left(\sum_{k=1}^n \omega_{ik}^c \ln f_i(x_k | \phi_i) \right)$$

3. Menentukan klasifikasi

$$x \in i \leftrightarrow i = \arg \max_{1 \leq i < L} \ln(f_i(x | \phi_{ij}^+))$$

$$x \in i \leftrightarrow i = \arg \min_{1 \leq i < L} d_i(x)$$

dimana $d_i(x)$ adalah

$$d_i(x) = (x - \mu_i^+) (\sum_i^+)^{-1} (x - \mu_i^+)^T + \ln |\sum_i^+|$$

$$\mu_i^+ = \frac{\sum_{j=1}^n \omega_{ij}^c x_{ij}}{m_i + \sum_{j=1}^n \omega_{ij}^c}$$

$$\sum_i^+ = \frac{\sum_{j=1}^n \omega_{ij}^c (x_j - \mu_i^+)^T}{m_i + \sum_{j=1}^n \omega_{ij}^c}$$

Proses dilakukan secara iterative sampai nilai perbedaan probabilitas posterior pada suatu iterasi dan iterasi sebelumnya sudah minimal.

5. PENERAPAN ALGORITME

Algoritme jaringan hibrida SOM-BP dan jaringan PNN Gaussian dengan optimasi EM diterapkan dengan memanfaatkan alat bantu paket MATLAB Versi 6. Hal ini dilakukan karena MATLAB Versi 6 telah banyak memuat fasilitas pustaka dasar jaringan syaraf tiruan seperti SOM dan BP. Pengembangan software dilakukan di Komputer PC Pentium IV.

6. HASIL EKSPERIMEN

Berikut akan disampaikan hasil eksperimen yang telah dilakukan, terdiri dari:

1. Pengklasifikasi jaringan hibrida SOM-BP dengan data sementara parameter sosial, ekonomi dan lingkungan
2. Pengklasifikasi jaringan SOM-BP dengan data Landsat ETM daerah Kabupaten Bengkayang
3. Pengklasifikasi jaringan PNN Gaussian dengan optimasi EM dengan data Landsat ETM daerah Kabupaten Bengkayang dan Sambas.

6.1. Pengelompokan Parameter Sosial Ekonomi dengan Jaringan Hibrida SOM-BP

Data sementara parameter konstrain sosial dan ekonomi serta lingkungan dapat dilihat di Tabel 1.

Tabel 1. Data Sementara Parameter Konstrain.

Parameter	Tolok Ukur	Komponen
Ekonomi	Pertumbuhan Ekonomi	1. Produk Domestik Regional Brutto 2. Daya tarik investasi
Tenaga Kerja	Ketersediaan Tenaga Kerja	1. Jumlah 2. Mobilitas
Sumber Daya Alam	Potensi Sumber Daya Alam	1. Lahan untuk pengembangan 2. Air, enersi, sda lainnya
Sarana dan Prasarana	Pasar dan Jaringan	1. Aksesibilitas pasar 2. Jaringan jalan lokal
Minat Investasi	Swasta / BUMN / Koperasi / Masyarakat / Pemerintah	Investasi yang ada
Lingkungan	Dampak Lingkungan	1. Degradasi Lingkungan 2. Gangguan permukiman

Karakteristik dari jaringan yang digunakan adalah sebagai berikut:

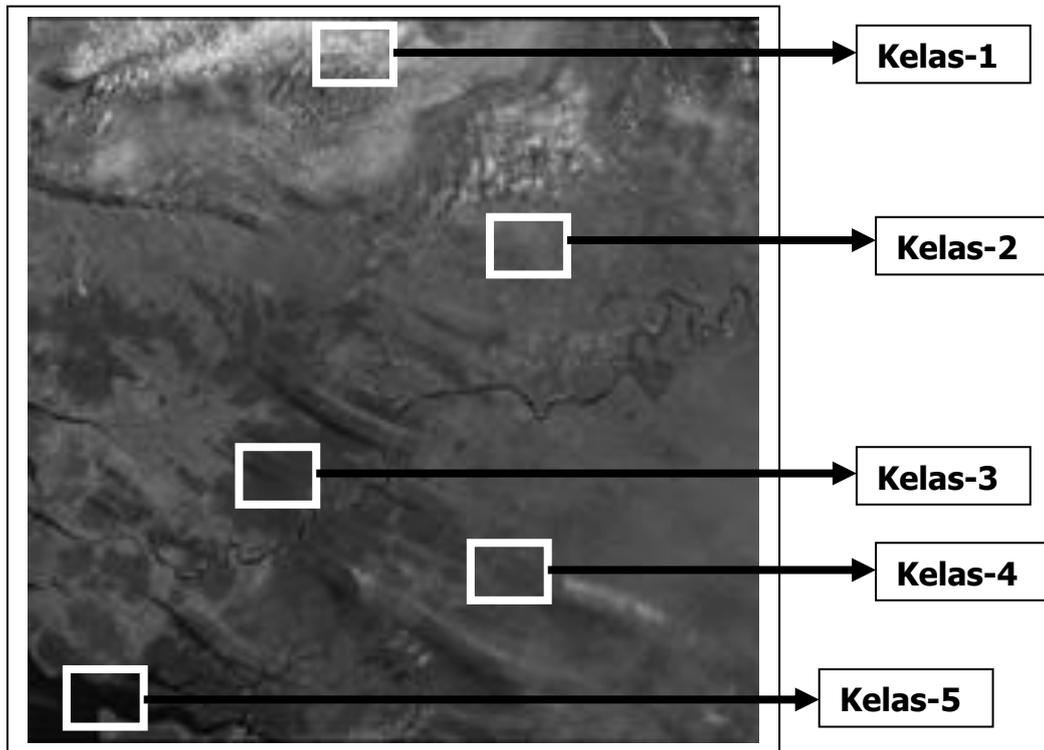
- Neuron masukan : 6 buah
- Neuron pola : 147 buah
- Neuron keluaran : 5 buah

Tabel 2. Data Ketelitian Hasil Klasifikasi Data Sosial Ekonomi dengan SOM-BP.

Kelas	Kelas-1	Kelas-2	Kelas-3	Kelas-4	Kelas-5	Piksel	Producer's Accuracy
Kelas-1	119					119	100
Kelas-2		214	16			230	93.04
Kelas-3		46	212	3		261	81.23
Kelas-4				55	10	65	84.62
Kelas-5					102	102	100
Piksel	119	260	228	58	112	Overall Accuracy (OA) = 90.347%	PA _{avr} = 91.778% UA _{avr} = 92.232%
User's Accuracy	100	82.31	92.98	94.8	91.07		

6.2. Klasifikasi Citra Landsat ETM dengan Jaringan Hibrida SOM-BP

Gambar 6 menunjukkan salah satu band (yaitu band 5) dari data Landsat ETM daerah kabupaten Bengkayang, Kalimantan, yang digunakan untuk eksperimen.



Gambar 6. Citra Landsat ETM Band 5 daerah kabupaten Bengkayang.

Jumlah sample pelatihan dipilih secara manual dan visual berdasarkan petunjuk informasi pendukung dari pakar atau dari peta yang ada. Ukuran jendela sample adalah 10 x 10 piksel per kelas obyek penutup lahan. Untuk melakukan proses pelatihan dan pengenalan telah diambil 25% (125 piksel) dari sample untuk pelatihan dan 75% (375 piksel) untuk proses pengenalan.

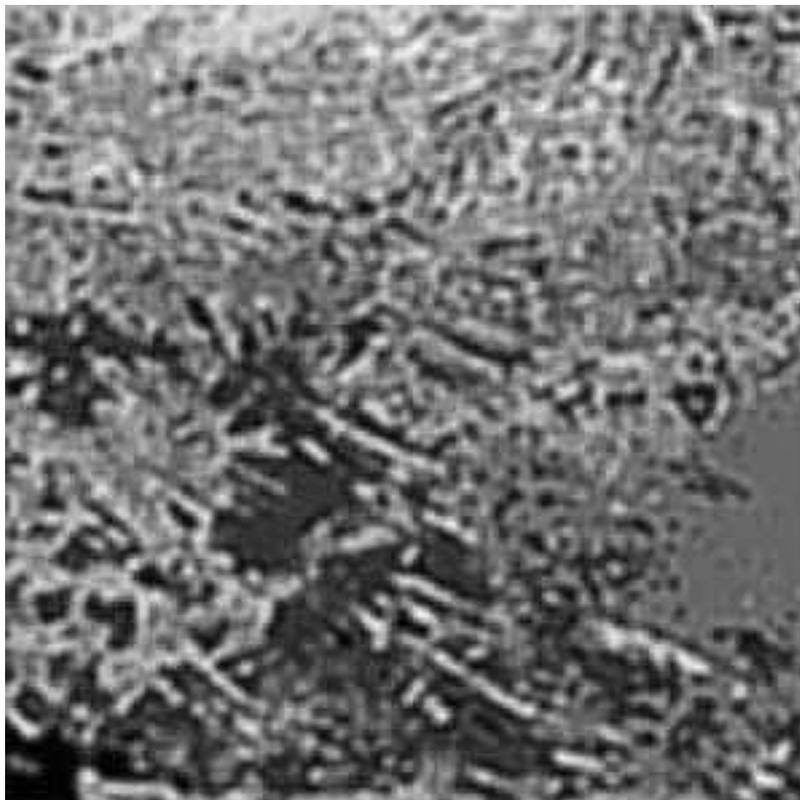
Karakteristik dari jaringan SOM-BP yang digunakan adalah sebagai berikut:

- Neuron masukan : 12 buah
- Neuron pola : 89 buah
- **Neuron keluaran : 5 buah**

Tabel 3. Data Ketelitian Hasil Klasifikasi Data Landsat ETM daerah kabupaten Bengkayang dengan jaringan hibrida SOM-BP.

Kelas	Kelas-1	Kelas-2	Kelas-3	Kelas-4	Kelas-5	Piksel	Producer's Accuracy
Kelas-1	73	1	-	-	-	74	98.65
Kelas-2	2	71	-	1	1	75	94.67
Kelas-3	-	3	72	5		80	90.00
Kelas-4	-	-	-	69	2	71	97.18
Kelas-5	-	-	3	-	72	75	96.00
Piksel	75	75	75	75	75	Overall Accuracy (OA)	PA _{avr} = 95,30% UA _{avr} = 95.20%
User's Accuracy	97.33	94.67	96.00	92.00	96.00		

Hasil dari proses klasifikasi citra adalah suatu gambar tematik yang terdiri dari pewilayahan obyek-obyek penutup lahan. Gambar tematik yang diperoleh dari eksperimen ini dapat dilihat pada Gambar 7 (ada kesalahan pengiriman gambar dan akan diperbaiki).



Gambar 7. Citra tematik Landsat ETM daerah kabupaten Bengkayang.

6.3. Klasifikasi Citra Landsat ETM dengan Jaringan PNN Gaussian Dengan Optimasi Algoritme EM

Jumlah sample pelatihan dipilih secara manual dan visual berdasarkan petunjuk informasi pendukung dari pakar atau dari peta yang ada. Sampel yang digunakan untuk proses pelatihan dan proses pengenalan pada eksperimen ini sama dengan sample yang digunakan pada eksperimen pada bagian 6.2 dan dapat dilihat pada Gambar 6.

Jaringan PNN dengan model Gaussian tanpa optimasi atau dengan optimasi algoritme EM mempunyai karakteristik sebagai berikut:

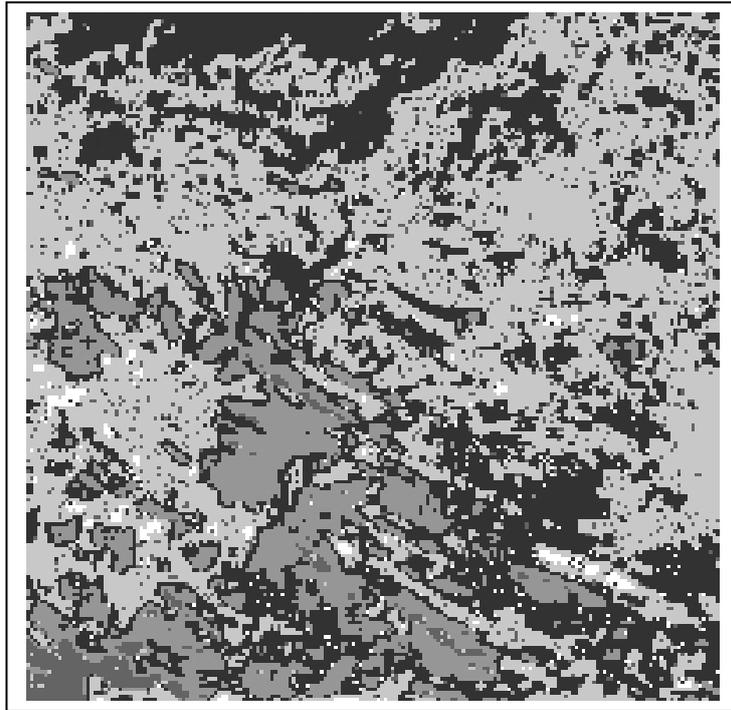
- Neuron masukan : 12 buah
- Neuron pola : 125 buah
- Neuron keluaran : 5 buah

Tabel 4. Data Ketelitian Hasil Klasifikasi Data Landsat ETM daerah kabupaten Bengkayang dengan jaringan PNN Gaussian.

Kelas	Kelas-1	Kelas-2	Kelas-3	Kelas-4	Kelas-5	Piksel	Producer's Accuracy
Kelas-1	75					75	100
Kelas-2		75				75	100
Kelas-3			75	15		90	83,3
Kelas-4				60		75	100
Kelas-5					75	75	100
Piksel	75	75	75	75	75	Overall Accuracy (OA) = 96.0%	PA _{avr} = 96,6% UA _{avr} = 96,0%
User's Accuracy	100	100	100	80	100		

Tabel 5. Data Ketelitian Hasil Klasifikasi Data Landsat ETM daerah kabupaten Bengkayang dengan PNN Gaussian optimasi algoritme EM.

Kelas	Kelas-1	Kelas-2	Kelas-3	Kelas-4	Kelas-5	Piksel	Producer's Accuracy
Kelas-1	75					75	100
Kelas-2		74				74	100
Kelas-3			75	2	1	78	96.1
Kelas-4				73		73	100
Kelas-5		1			72	73	98.6
Piksel	75	75	75	75	73	Overall Accuracy (OA) = 98.9%	PA _{avr} = 98.9% UA _{avr} = 98.9%
User's Accuracy	100	98.6	100	97.3	98.6		

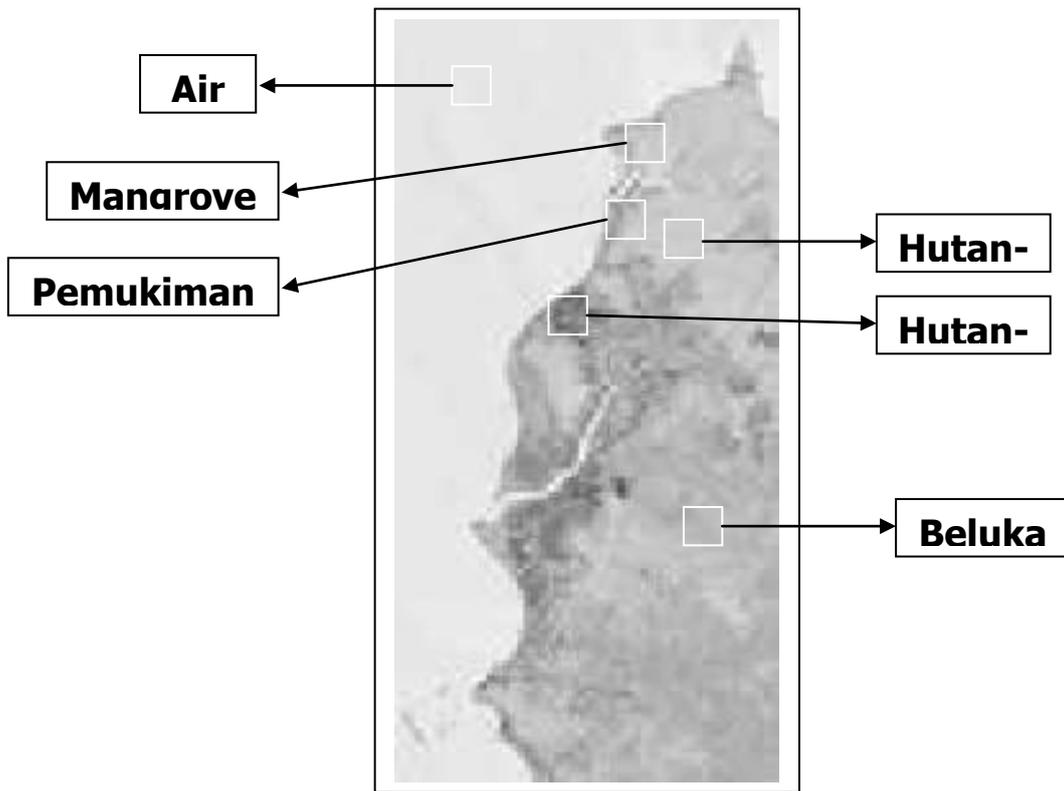


Gambar 8. Citra tematik Landsat ETM daerah kabupaten Bengkayang, hasil klasifikasi PNN Gaussian dengan optimasi algoritme EM.

Eksperimen dengan data lain, yaitu data Landsat ETM daerah kabupaten Sambas juga telah dilakukan. Citra masukan dan daerah sampel yang diambil dapat dilihat di Gambar 9. Perbandingan jumlah sampel pelatihan dan testing adalah 25% dan 75%.

Karakteristik dari jaringan PNN Gaussian dengan optimasi EM yang digunakan adalah sebagai berikut:

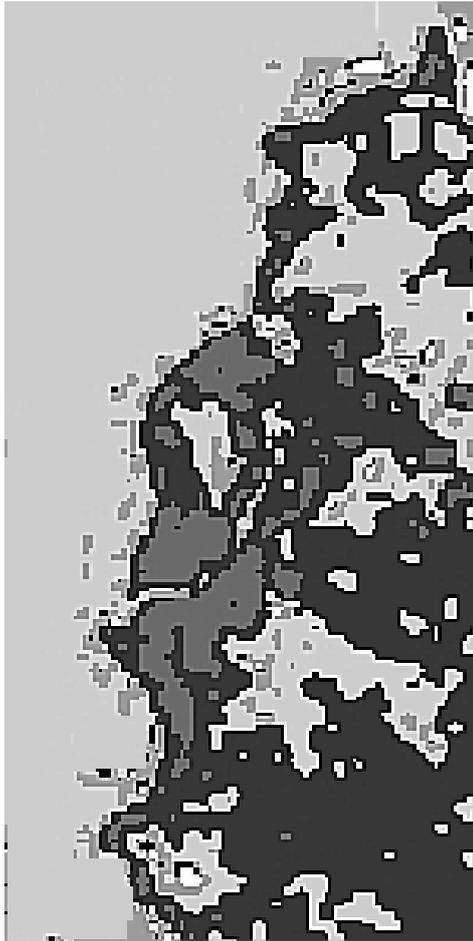
- Neuron masukan : 12 buah
- Neuron pola : 150 buah
- Neuron keluaran : 6 buah



Gambar 9. Citra Landsat ETM daerah kabupaten Sambas.

Tabel 6. Data Ketelitian Hasil Klasifikasi Data Landsat ETM daerah kabupaten Sambas dengan PNN Gaussian optimasi algoritme EM

Object Class	1	2	3	4	5	6	Producer's Accuracy - PA (%)
1	2,523	-	-	-	-	303	89.29
2	-	2,523	606	-	-	-	80.65
3	-	-	1,917	-	-	-	100.00
4	-	-	-	2,523	-	-	100.00
5	-	-	-	-	2,523	-	100.00
6	-	-	-	-	-	2,220	100.00
User's Accuracy - UA (%)	100.00	100.00	76.00	100.00	100.00	88.00	Overall Accuracy - OA (%) = 93.99% <i>PA = 94.99%</i> <i>UA = 94.00%</i>



Gambar 10. Gambar Tematik Landsat ETM daerah kabupaten Sambas.

7. KESIMPULAN

Dari hasil eksperimen dapat diberikan beberapa rekomendasi sebagai berikut:

1. Jaringan hibrida SOM dan BP dapat direkomendasikan sebagai metode pengelompokan parameter sosial, ekonomi dan lingkungan.
2. Jaringan SOM dapat digunakan secara terpisah untuk membantu memberikan nilai kuantitas terhadap kualitas atau kondisi sosial, budaya dan lingkungan.
3. Jaringan PNN Gaussian dengan optimasi EM dapat digunakan sebagai alternatif yang baik untuk klasifikasi citra Landsat ETM