

Pendekatan Diagnostik Berbasis *Extreme Learning Machine* dengan Kernel Linear untuk Mengklasifikasi Kelainan Paru-Paru

Putu Prima Winangun¹, I Made Oka Widyantara², Rukmi Sari Hartati³

Submission: 09-03-2020, Accepted: 08-06-2020

Abstract— an expert system can be used as a second opinion for comparison or supporting diagnosis from experts. Data mining is used to obtain information applied to this system. Whereas in conducting learning using Artificial Neural Networks which apply the Extreme Learning Machine method so that it can accelerate learning up to thousands of times. In this paper, software development is carried out to test the activation functions used in conducting learning and the variables used as input during learning.

Intisari— Sebuah sistem pakar dapat digunakan sebagai opini kedua untuk membandingkan atau pendukung diagnose dari pakar. Penggalan data digunakan untuk mendapatkan informasi diterapkan pada Sistem ini. Sedangkan dalam melakukan pembelajaran menggunakan Jaringan Saraf Tiruan yang menerapkan metode Extreme Learning Machine sehingga dapat mempercepat pembelajaran hingga ribuan kali lipat. Dalam makalah ini, pengembangan perangkat lunak yang dilakukan untuk menguji fungsi aktivasi yang digunakan dalam melakukan pembelajaran dan variabel yang digunakan sebagai input pada saat pembelajaran.

Kata Kunci— *machine learning, jaringan saraf tiruan, extreme learning machine, kernel linear, artificial intelligence, sistem pakar.*

I. PENDAHULUAN

Kemajuan dalam bidang teknologi terutama pengolahan citra digital memberikan efek yang cukup besar terhadap kemajuan bidang medis. Dengan penemuan tomografi terkomputerisasi (Computerized Tomography) pada tahun 1970an yang hingga kini teknologi tersebut semakin berkembang dan maju pesat. Alat ini mampu digunakan untuk melakukan segmentasi tulang dari otot yang melekat, klasifikasi gigi dan analisis citra mikroskopis. Dengan melakukan analisis terhadap citra yang diperoleh dari CT, para ahli mampu untuk mengidentifikasi bagian - bagian dari hasil citra yang terdapat kelainan dan memberikan diagnosis terhadap citra yang dilihat. Hal ini mendorong ilmuwan untuk membuat sebuah sistem pakar yang dapat memberikan sebuah second opinion (opini kedua) sebagai pembanding atau bahkan pendukung diagnosa dari seorang pakar medis.

Banyak penelitian yang menggunakan teknik penggalan data (data mining) untuk melakukan diagnosa terhadap

penyakit. Penggalan data bisa dilakukan dengan ekstraksi data yang didapat dari citra medis [1] [2]. Fitur dari data tersebut perlu dibangkitkan ataupun diseleksi untuk mengurangi data yang tidak penting masuk ke dalam sistem dan mempengaruhi proses klasifikasi, karena apabila terdapat variabel yang tidak relevan, sistem akan menggunakan informasi data ini untuk data baru yang dapat mengakibatkan sistem mengambil keputusan yang buruk [3].

Salah satu fitur yang di ekstrak adalah fitur tekstur citra karena tekstur merupakan salah satu penanda paling penting dari citra [4]. Terdapat 4 (empat) metode utama yang biasanya digunakan dalam melakukan karakterisasi tekstur pada citra yaitu: metode statistik (metode co-occurrence), metode model basis (Markov random fields), geometri (Voronoi tessellation features, fractal) dan metode pemrosesan sinyal (wavelet transform, curvelets serta Gabor filters) [5]. Dalam penelitian ini, menggunakan metode statistik yang paling banyak dipakai yaitu metode Gray Level Co-occurrence Matrix.

Sementara untuk melakukan klasifikasi dapat menggunakan Support Vector Machine (SVM), Decision Tree (DT), Genetic Algorithm (GA), Jaringan Saraf Tiruan (JST), dan lainnya [6]. Tiap metode klasifikasi memiliki kekurangan dan kelebihan masing-masing. Beberapa dekade terakhir, banyak penelitian yang menunjukkan pendekatan pembelajaran mesin seperti JST dan SVM memiliki nilai akurasi yang lebih tinggi dari pendekatan statistik tradisional [7]. Berdasarkan penelitian yang dilakukan [8], JST memiliki akurasi yang lebih tinggi daripada SVM, akan tetapi memiliki kelemahan dalam waktu pembelajaran dan klasifikasi.

Untuk mengatasi hal tersebut, penelitian ini akan menggunakan salah satu metode baru JST yaitu Extreme Learning Machine dimana metode tersebut memiliki kelebihan dalam kecepatan pembelajaran hingga ribuan kali lipat [9]. Dengan menerapkan single-hidden-layer feedforward neural network, jaringan ini dapat mengatasi kelemahan pada gradient descent learning. Algoritma ini dikembangkan oleh Huang dkk pada tahun 2004. Pelatihan yang cukup cepat mampu dilakukan karena metode ini tidak adanya iterasi dan kemampuan generalisasi yang dihasilkan tetap baik.

Penggunaan fungsi aktivasi dalam melakukan pembelajaran pada metode ELM juga mempengaruhi hasil akurasi dari sistem yang dibuat. Penerapan fungsi aktivasi kernel seperti pada metode SVM mampu meningkatkan hasil akurasi dari sistem jika dibandingkan dengan menggunakan fungsi aktivasi standar seperti fungsi sigmoid, sinus yang biasa dipakai pada ELM standar. Penelitian ini menggunakan fungsi kernel linear dan kernel RBF sebagai fungsi aktivasi. Kernel linear dan RBF yang diterapkan pada SVM mampu memberikan performansi yang cukup tinggi [10] [11].

¹ Mahasiswa, Magister Teknik Elektro, Program Pasca Sarjana Universitas Udayana (e-mail: primawinangun@gmail.com)

^{2, 3} staff pengajar Magister Teknik Elektro, Program Pasca Sarjana Universitas Udayana, JL PB. Sudirman Denpasar Bali (0361-239599); email: ²oka.widyantara@unud.ac.id, ³rukmisari@unud.ac.id

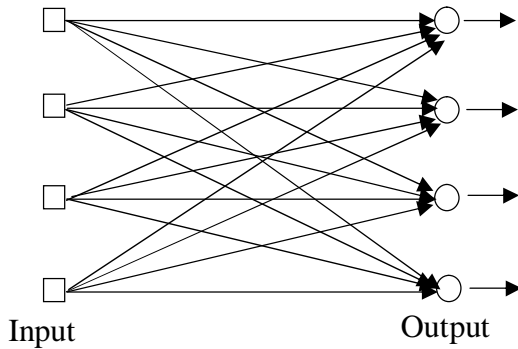


II. LANDASAN TEORI

Adapun teori yang mendukung dan menjadi landasan untuk penelitian ini yaitu:

A. Single Layer Feedforward Network

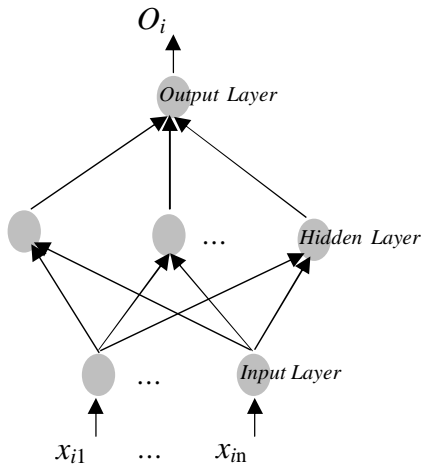
Single Layer Feedforward Networks (SLFN) merupakan arsitektur jaringan yang paling sederhana. Pada jaringan ini hanya terdapat 2 layer yaitu input layer yang kemudian langsung menuju output tanpa melewati lapisan tersembunyi. Akan tetapi input layer tidak dihitung dikarenakan tidak adanya perhitungan yang dilakukan pada layer tersebut [12].



Gambar 1 : Feedforward Network dengan satu lapisan neuron tunggal

B. Extreme Learning Machine

ELM adalah sebuah metode pembelajaran baru dalam JST yang merupakan pengembangan dari JST feedforward yang menggunakan Single Hidden Layer Feedforward Neural Networks (SLFNs) [13] dengan struktur yang kurang lebih sama dengan JST feedforward lainnya terlihat pada Gambar 2. ELM diciptakan untuk mengatasi kelemahan pada JST terutama pada bidang kecepatan pembelajaran. Huang dkk mengemukakan terdapat dua alasan mengapa rendahnya kecepatan pembelajaran JST feedforward lain yaitu penerapan slow gradient based learning algorithm yang diterapkan pada saat training dan jaringan melakukan perhitungan secara iterative pada semua parameter digunakan oleh metode pembelajaran tersebut.



Gambar 2 : Struktur Extreme Learning Machine

Metode ELM memiliki kemampuan yang bagus dalam melakukan klasifikasi hubungan yang kompleks. Hasil klasifikasi hubungan dari ELM sangat tidak linear dan sering menghasilkan model matematika yang cukup besar. Setelah pelatihan dilakukan, ELM mampu melakukan klasifikasi data yang jauh lebih cepat jika dibandingkan dengan melakukan pemecahan model permasalahan secara analitik. Sehingga jika dibandingkan dengan teknik komputasi berbasis algoritma pembelajaran konvensional, ELM dapat memberikan kinerja yang lebih baik pada kecepatan pembelajaran dan dengan sedikit campur tangan manusia [14].

ELM memiliki model matematis yang lebih sederhana dan efektif jika dibandingkan dengan JST feedforward. Model matematis ELM untuk n jumlah sample yang berbeda (X_i, t_i) adalah sebagai berikut:

$$X_i = [X_{i_1}, X_{i_2}, \dots, X_{i_n}]^T \in R^n \tag{1}$$

$$X_t = [X_{t_1}, X_{t_2}, \dots, X_{t_n}]^T \in R^n \tag{2}$$

Sementara untuk standar SLFNs dengan jumlah hidden nodes sebanyak n dan fungsi aktivasi g(x) dapat digambarkan secara matematis dengan rumusan:

$$\sum_{i=1}^n \beta_i g_i(x_j) = \sum_{i=1}^n \beta_i g(w_i \cdot x_j + b_i) = o_j \tag{3}$$

Dimana:

j = 1, 2, ..., N

w_i = [(w_{i1}, w_{i2}, ..., w_{in})]^T

β_i = [(β_{i1}, β_{i2}, ..., β_{im})]^T

b_i merupakan threshold dari hidden nodes ke i

w_i·x_j adalah inner produk dari w_i dan x_j

Berdasarkan penelitian yang dilakukan oleh Guang-Bin Huang [9], ELM mampu memberikan platform pembelajaran terpadu dengan pemetaan fitur yang meluas dan dapat diterapkan untuk melakukan multiclass classification dan regresi secara langsung. Penelitian tersebut juga menyebutkan bahwa dari segi optimasi, kendala yang dimiliki oleh ELM lebih sedikit dibandingkan dengan JST feedforward lainnya. Secara teori, ELM mampu mencapai solusi optimal tanpa memerlukan kompleksitas komputasi yang tinggi, ELM juga cenderung memiliki skalabilitas dan mencapai hasil yang serupa (untuk regresi dan kasus kelas biner) atau mendapatkan hasil yang lebih baik (untuk kasus multiclass) dengan performansi secara general yang jauh lebih cepat (sampai ribuan kali) dibandingkan dengan JST feedforward tradisional lainnya.

Dengan menggunakan ELM sebagai metode untuk pembelajaran dalam menentukan kelainan yang terdapat pada paru-paru maka akurasi dari hasil pembelajaran hasil ekstraksi ciri menggunakan GLCM kedalam sistem yang dibangun.

C. Linear Kernel

Kernel linear merupakan salah satu dari banyak kernel yang sering digunakan dalam penerapan machine learning. Penggunaan kernel linear pada machine learning biasanya untuk pengelompokan atau klasifikasi teks. Hal ini dikarenakan kebanyakan data dari klasifikasi teks dapat dipisahkan secara linear. Akan tetapi tidak menutup kemungkinan penggunaan kernel linear sebagai dasar kernel

untuk melakukan pengelompokkan atau klasifikasi dari data lainnya seperti suara, diagnosa penyakit maupun citra.

Penggunaan kernel linear secara umum dapat dijelaskan seperti membuat sebuah tempat baru yang merupakan hasil pemetaan dari data yang tidak linear. Dalam ruang fitur, sebuah operasi dot product dapat secara langsung digantikan dengan fungsi kernel, sementara kita tidak perlu mengetahui vector eigen konkret dan fungsi pemetaan konkritnya [15].

Misalkan $\phi(x_i) \in R^m, m > d$ adalah hasil transformasi dari $x \in R^d$ dimana, fungsi transformasi nonlinear adalah $\phi(x)$.

Parameter x_i termasuk dalam ruang input dimensi, dan

$\phi(x_i)$ termasuk dalam ruang fitur dimensi m . Pada saat ini, masalah optimasi diubah menjadi fungsi seperti:

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \{\phi < (x_i), \phi(x_j) >\} \quad (4)$$

Inner produk dari rumus diatas dilakukan dalam ruang dimensi yang relative tinggi, sehingga terjadi kemungkinan ada masalah pada dimensi. Setelah memperkenalkan fungsi kernel tersebut pada SVM, maka rumus tersebut dapat di optimasi menjadi

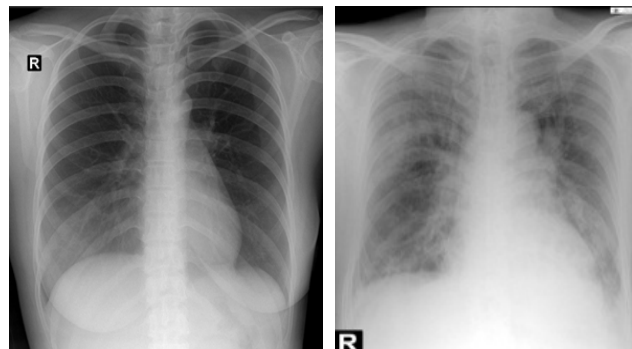
$$Q(a) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (5)$$

Kernel ini adalah fungsi kernel paling sederhana yang dapat dihasilkan dari inner product dan sering kali disetarakan dengan fungsi non-kernel lainnya.

D. Data

Data yang didapatkan dalam penelitian yaitu berupa citra rontgen thorax yang bersifat kuantitatif. Citra ini akan dilakukan standarisasi dengan melakukan kompresi citra ke jpeg dan penyamaan ukuran citra. Data citra merupakan citra grayscale sehingga dalam pemrosesannya tidak diperlukan perubahan warna dari RGB menjadi grayscale.

Citra rontgen thorax normal memiliki bentuk dan gambar yang cukup jelas pada rongga dada, sehingga terlihat mana yang merupakan paru-paru, jantung dan diafragma. Sedangkan pada citra rontgen thorax tidak normal memiliki beberapa kelainan pada bagian dari rongga paru-paru, seperti pada apex terdapat bercak bercak putih, sudut carina berada diluar batas normal, sinus prenicocostalis yang tidak tajam serta terkadang terlihat jelas adanya bercak putih pada rongga paru paru karena terdapatnya cairan pada paru paru. Berikut merupakan contoh citra yang digunakan dalam penelitian

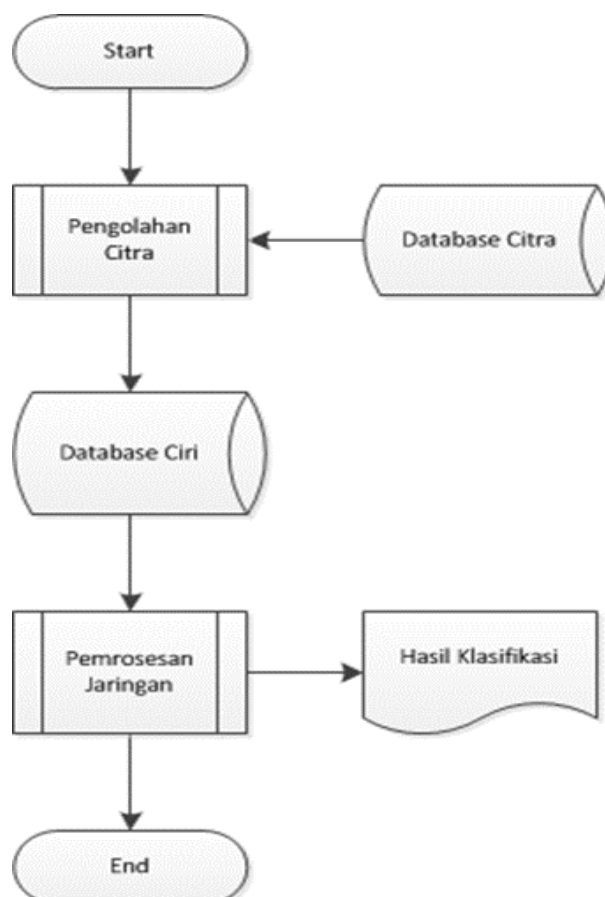


Gambar 3 : Contoh citra thorax normal (kiri) dan tidak normal (kanan)

III. METODOLOGI PENELITIAN

A. Gambaran Umum Sistem

Sistem yang dibangun secara umum terbagi menjadi dua bagian, yaitu proses pengolahan citra yang merupakan sebuah proses untuk mendapatkan data ciri berupa tekstur dari citra rontgen thorax dan proses pengolahan jaringan yang merupakan proses pembelajaran serta simulasi jaringan yang digunakan oleh peneliti sebagai proses klasifikasi citra rontgen thorax.

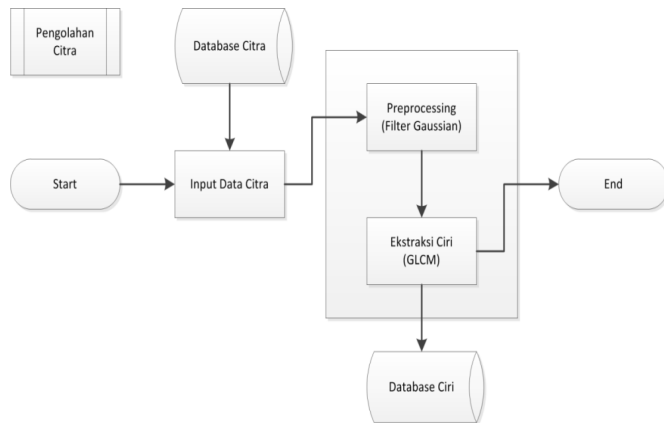


Gambar 4 : Gambaran umum sistem



B. Pengolahan Citra

Berikut ini adalah gambaran umum untuk proses pengolahan citra yang digunakan untuk mencari ciri tekstur dari citra:

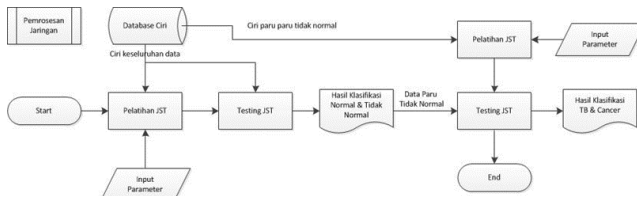


Gambar 5 : Gambaran sistem pengolahan citra

Adapun penjelasan dari gambar diatas yaitu:

1. Data citra dimasukkan kedalam sistem dari database citra yang ada berupa citra digital rontgen thorax. Citra ini berupa citra grayscale dengan ukuran sebesar 512 x 512 piksel.
2. Citra akan diproses menggunakan Filter Gaussian sebagai tahap untuk penghilangan noise pada citra.
3. Hasil keluaran dari proses sebelumnya akan diekstraksi ciri teksturnya menggunakan GLCM dengan mengambil korelasi, energi, kontras dan homogenitasnya. Sudut yang digunakan pada proses GLCM yaitu sudut 0°, 45°, 90° dan 135°.
4. Menyimpan data tersebut kedalam database ciri. Data ciri tersebut berupa matriks hasil konversi metode GLCM berupa matriks korelasi, energi, kontras dan homogenitas dari tiap citra

D. Pelatihan dan Pengujian Jaringan



Gambar 6 : Gambaran sistem pelatihan dan pengujian jaringan

Database ciri akan dibagi menjadi dua bagian, sebagian menjadi data latih untuk pembelajaran jaringan dan sebagian akan digunakan sebagai data uji pada saat melakukan klasifikasi. Data latih dan data uji bisa beririsan satu dengan yang lain.

Data latih akan dimasukkan kedalam jaringan untuk dipelajari oleh jaringan dengan input parameter yang telah ditentukan oleh peneliti. Input yang dimasukkan sebagai parameter adalah jumlah hidden node yang digunakan pada saat pelatihan jaringan, sementara input weight dan hidden bias ditentukan secara acak selama masa pembelajaran sistem. Nilai dari input weight dan hidden bias juga merupakan hasil keluaran dari pelatihan jaringan.

Hasil dari pelatihan jaringan akan diuji dengan data uji dengan melakukan simulasi terhadap jaringan yang telah dilatih sebelumnya. Data uji akan dimasukkan ke dalam jaringan yang terlatih dengan parameter input weight, hidden node dan hidden bias yang dikeluarkan pada saat pelatihan jaringan dilakukan. Hasil keluaran dari simulasi ini berupa klasifikasi dari tiap data citra uji yang ada yaitu apakah citra uji tersebut merupakan citra thorax normal atau citra thorax tidak normal.

IV. HASIL PENELITIAN

Data yang digunakan terdiri dari 122 citra dimana 66 citra digunakan sebagai citra latih dan 56 citra sisanya sebagai citra uji. Citra yang diujikan akan dilakukan preprocessing menggunakan filter gaussian dan kemudian dibangkitkan cirinya menggunakan GLCM. Ciri tersebut kemudian dinormalisasi dengan skala -1 sampai dengan 1 untuk mempermudah pelatihan dan pengujian.

Dua jenis fungsi pembangkit digunakan dalam eksperimen yang dilakukan, yaitu fungsi pembangkit sigmoid dan linear kernel, fungsi tersebut dipilih karena dalam sistem yang dibuat memiliki output pada rentang 0 dan 1.

Penggunaan fungsi pembangkit sigmoid dilakukan sebanyak 10 kali dengan melakukan perubahan pada jumlah hidden neuron dengan kelipatan 10 pada saat melakukan pelatihan dan pengujian untuk mendapatkan hasil yang maksimal.

TABEL I
PELATIHAN FUNGSI SIGMOID

Hidden Neuron	Training Time (detik)	Training accuracy (%)
10	0	66,67
20	0	81,82
30	0	87,88
40	0,015625	95,45
50	0	98,48
60	0	100,00
70	0	100,00
80	0	100,00
90	0	100,00
100	0	100,00

Dengan menggunakan fungsi pembangkit sigmoid, waktu yang diperlukan untuk melakukan training menjadi sangat cepat dan tidak mencapai 1 detik untuk semua jumlah hidden neuron yang digunakan. Nilai akurasi untuk training tertinggi didapat pada jumlah hidden neuron mulai dari 60 hingga 100 persen

TABEL III
PENGUJIAN FUNGSI SIGMOID

Hidden Neuron	Training Time (detik)	Training accuracy (%)
10	0	57,14
20	0	66,07
30	0	71,43
40	0	58,93
50	0	66,07
60	0	50,00
70	0	55,36
80	0	53,57

90	0	55,36
100	0	58,93

Pada hasil pengujian diatas, waktu yang diperlukan oleh sistem untuk melakukan pengujian pada data inputan sangat cepat sehingga sistem hanya mencatat nilai 0 sebagai output waktu yang diperlukan. Akurasi pada pengujian menggunakan fungsi aktivasi sigmoid menunjukkan nilai akurasi sebesar 66.07 persen sebagai akurasi tertinggi yang didapat.

Pelatihan dan pengujian selanjutnya dengan menggunakan fungsi pembangkit kernel linear dilakukan sebanyak 10 kali dengan regularization coefficient yang berbeda dan jumlah neuron yang sama untuk mendapatkan hasil yang paling maksimal. Perubahan yang dilakukan pada variabel regularization coefficient akan mempengaruhi bias nilai weight yang didapat pada hidden neuron.

TABEL IIIII
PELATIHAN KERNEL LINEAR

Regularization coefficient	Hidden neuron	Waktu pelatihan (detik)	Akurasi pelatihan (%)
1	60	2.12×10^{-4}	62.12
10	60	3.41×10^{-4}	62.12
100	60	1.28×10^{-4}	63.63
1×10^3	60	1.32×10^{-4}	68.18
1×10^4	60	1.77×10^{-4}	74.24
1×10^5	60	1.75×10^{-4}	75.76
1×10^6	60	1.75×10^{-4}	84.85
1×10^7	60	1.74×10^{-4}	87.88
1×10^8	60	1.88×10^{-4}	95.45
1×10^9	60	1.22×10^{-4}	96.97
1×10^{10}	60	1.22×10^{-4}	96.97

Berdasarkan data yang didapat pada Tabel 1 Hasil Pengujian Citra Latih, akurasi terbaik didapat dengan menggunakan regularization koefisien yang besar yaitu 1×10^9 dan 1×10^{10} hingga mencapai 96.97% akurasi. Kecepatan waktu training tercepat juga terdapat pada koefisien yang sama. Dengan besaran kecepatan yang tidak sampai pada nilai 1 detik.

V. KESIMPULAN

Penggunaan metode ELM dapat mempercepat pembelajaran dan pengujian pada sistem. Hal tersebut terbukti dari hasil percobaan diatas. Penggunaan fungsi pembangkit mempengaruhi tingkat akurasi sistem, dengan menggunakan fungsi aktivasi kernel dibandingkan dengan menggunakan fungsi aktivasi sigmoid. Hal ini bisa dipengaruhi oleh bentuk data yang diproses oleh sistem. Untuk mengatasi kelemahan pada tingkat akurasi bisa ditingkatkan dengan melakukan proses preprocessing pada data yang akan dimasukkan pada sistem dan memilih fungsi aktivasi yang lebih baik dan lebih sesuai dengan data inputan untuk melakukan pelatihan dan pengujian.

REFERENSI

[1] U. Gupta, V. Kumar and S. Sharma, "Improving medical image Putu Prima Winangun: Pendekatan Diagnostik Berbasis Extreme ...

classification model using data mining technique," in *Communication and computing systems*, London, 2017.

- [2] N. S and P. M. Goel Dr, "Comparison of Classification Techniques on Data Mining," *International Journal of Emerging Technology and Innovative Engineering*, vol. 5, no. 5, 2019.
- [3] S. A. Lashari, R. Ibrahim, N. Senan and N. S. A. M. Taujuddin, "Application of Data Mining Techniques for Medical Data," in *MATEC Web of Conferences 150*, 2018.
- [4] A. Vidyarthi and N. Mittal, "Texture based feature extraction method for classification of brain tumor MRI," *Journal of Intelligent and Fuzzy Systems*, vol. 32, no. 4, pp. 1-12, 2017.
- [5] L. Armi and S. Fekri-Ershad, "Texture image analysis and texture classification methods - A Review," *International Online Journal of Image Processing and Pattern Recognition*, vol. 2, no. 1, pp. 1-29, 2019.
- [6] T. Sharma, A. Sharma and P. Mansotra, "Performance Analysis of Data Mining Classification Techniques on Public Health Care Data," *International Journal of Innovative Research in Computer and Communication Engineering*.
- [7] S. Makridakis, E. Spiliotis and V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward," *PLOS ONE Statistical and ML forecasting methods*, 2018.
- [8] A. Bakhshipour and A. Jafari, "Evaluation of support vector machine and artificial neural networks in weed detection using shape features," *Computers and Electronics in Agriculture*, vol. 145, pp. 153-160, 2018.
- [9] G.-B. Huang, H. Zhou, X. Ding and R. Zhang, "Robust Classification of Brain Tumor in MRI Images using Salient Structure Descriptor and RBF Kernel- SVM," *TAGA Journal*, vol. 14, pp. 718-737, 2018.
- [10] V. Sharma, D. Baruah, D. Chutia and P. L. N. Raju, "An Assessment of Support Vector Machine Kernel Parameters using Remotely Sensed Satellite Data," in *IEEE International Conference On Recent Trends In Electronics Information Communication Technology*, India, 2016.
- [11] P. K. Intan, "Comparison of Kernel Function on Support Vector Machine in Classification of Childbirth," *Jurnal Matematika MANTIK*, vol. 5, no. 2, pp. 90-99, 2019.
- [12] P. B. S. Hendrayana, R. L. Rahardian and M. Sudarma, "Application of Neural Network Overview In Data Mining," *International Journal of Engineering and Emerging Technology*, vol. 2, pp. 94-96, 2017.
- [13] J. Cao, K. Zhang, M. Luo, C. Yin and X. Lai, "Extreme learning machine and adaptive sparse representation for image classification," *Neural Networks*, 2016.
- [14] N. Nietoa, F. Ibarrolaa , V. Petersonb , H. L. Rufinera and R. Spiesb, "Extreme Learning Machine design for dealing with unrepresentative features," *Preprint submitted to Journal of LATEX Templates*, 2019.
- [15] Arash A. Amini and Zahra S. Razaee, "Concentration of kernel matrices with application to kernel spectral clustering," *arxiv.org*, 2019.



{ Halaman ini sengaja dikosongkan }