Proposal Penelitian :

Deteksi Dini Penyakit Ginjal Kronis menggunakan *Extreme Learning Machine* berbasis *Neural Network*

Moh. Faturrahman

NPM : 1506706276

Fakultas Ilmu Komputer, Universitas Indonesia.

[moh.faturrahman@ui.ac.id](mailto:moh.faturrahman@ui.ac.id)

**Abstrak**

*Prevalensi penyakit Gagal Ginjal Kronis sebesar 0.2 % terjadi di Indonesia dan cenderung terjadi pada masyarakat pedesaan, tidak bersekolah, pekerja swasta, dan petani. Penyakit Gagal Ginjal Kronis meningkatkan resiko terjadinya komplikasi terhadap penyakit lainnya seperti penyakit jantung, hipertensi, anemia, kematian, dan proses menuju pada stadium akhir. Deteksi Dini bertujuan untuk menunda proses penurunan fungsi ginjal, mencegah atau mengatur komplikasi. Penggunaan teknik Pembelajaran Mesin Extreme Learning Machine berpotensi melakukan deteksi dini penyakit Ginjal Kronis secara cepat dan efisien untuk mendukung tujuan diatas. Keunggulan dari teknik ini adalah memiliki proses pembelajaran yang sangat cepat, tidak memiliki banyak parameter yang harus diinisialisasi, tidak terjebak pada local minima, dapat menggunakan non-differentiabale activation function. Tujuan dari penelitian ini mengukur peforma dari Extreme Learning Machine dalam mendeteksi penyakit Ginjal Kronis.*

Kata Kunci : Penyakit Gagal Ginjal, Pembelajaran Mesin, Extreme Learning Machine

1. PENDAHULUAN

Penyakit Ginjal Kronis didefinisikan sebagai adanya tanda-tanda kerusakan pada ginjal ( abnomarlitas dalam darah, urin, atau pada tes gambar) lebih dari 3 bulan atau menurunnya *Glomerular Filtration Rate < 60 mL/min/*1.73 m2[1]⁠. Prevalensi penyakit Gagal Ginjal Kronis sebesar 0.2 % terjadi di Indonesia dan cenderung terjadi pada masyarakat pedesaan, tidak bersekolah, pekerja swasta, dan petani [2]. Penyakit Gagal Ginjal Kronis meningkatkan resiko terjadinya komplikasi terhadap penyakit lainnya seperti penyakit jantung, hipertensi, anemia, kematian, dan proses menuju pada stadium akhir [3]. Deteksi Dini pada penyakit Ginjal dapat dilakukan dengan estimasi nilai Glomerular Filtration Rate dan nilai Albuminnuria, serta menggunakan beberapa tes pendukung seperti Ultra Sound/CT Scan pada saluran kemih, atau tes serologic jika penderita punya riwayat diabetes atau hipertensi. Deteksi Dini bertujuan untuk menunda proses penurunan fungsi ginjal, mencegah atau mengatur komplikasi.

Pemeriksaan laboratorium memiliki kelemahan seperti biaya yang mahal dan proses yang berulang-ulang jika hasil yang didapat perlu dikomfirmasi. Penggunaan teknik Pembelajaran Mesin berpotensi melakukan deteksi dini penyakit Ginjal Kronis untuk mendukung tujuan diatas.

Pemanfaatan teknik Pembelajaran Mesin dalam bidang medis telah banyak dilakukan untuk membantu petugas medis dalam mengambil keputusan, memilih perawatan yang sesuai dengan pasien, membantu merancang pengobatan. Guang Jing et al, menggunakan teknik Pembelajaran Mesin untuk memprediksi mortality rate dari pasien kanker kandung kemih yang telah dilakukan *Radical Cystectomy* (Operasi pengangkatan saluran kandung kemih) dengan akurasi rata-rata 80%, teknik ini menunjukan performa yang lebih baik dibandingkan dengan metode statistik [4]. Penelitian serupa juga dilakukan oleh Vukicevic et al, dengan menggunakan dataset yang berbeda, hasil yang dicapai memiliki performa akuras 92.5% [5].

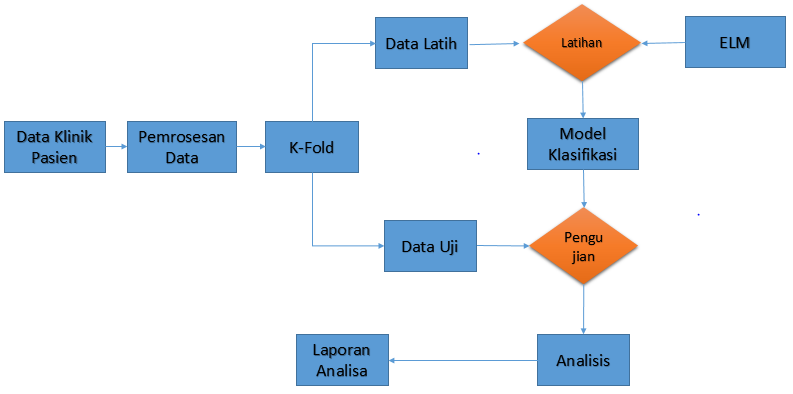
Khusus pada penyakit Ginjal Kronis, penggunaan Teknik Machine Learning Naive Bayes dalam mendeteksi penyakit Gagal Ginjal Kronis sudah dilakukan oleh Chetty et al, dengan menggunakan algoritma seleksi fitur pada pra-pemrosesan data, proses deteksi mampu menghasilkan akurasi 99%[6]. Dengan menggunakan dataset yang sama, Chen Zewei et al, menggunakan *Fuzzy Classfier* untuk mendeteksi penyakit Ginjal Akut, nilai rata-rata dari akurasi yang didapat mencapat 98% [7]. Berdasarkan beberapa hasil penelitain pemanfaatan teknik Pembelajaran Mesin pada bidang medis, menunjukan bahwa teknik Pembelajaran Mesin merupakan salah satu solusi effektif yang dapat mengatasi masalah-masalah pada bidang Medis.

Salah satu Teknik Machine Learning yang dapat digunakan pada bidang medis adalah Extreme Learning Machine [8], [9] [10], yang merupakan variansi dari Single Layer Feedforward Network (*SLFN*). Keunggulan dari teknik ini adalah memiliki proses pembelajaran yang sangat cepat, tidak memiliki banyak parameter yang harus diinisialisasi, tidak terjebak pada *local minima*, dapat menggunakan *non-differentiabale activation function* [11].

Oleh karena itu tujuan dari penelitian ini mengukur peforma dari *Extreme Learning Machine* dalam mendeteksi penyakit Ginjal Kronis. Adapaun penjelasan mengenai *Extreme Learning Machine*  yang akan dijelaskan pada bab selanjutnya.

Pada rancangan proposal penelitian ini ditulis melalui alur penulisan sebagai berikut, BAB II membahas skema penelitian yang akan dilakukan, mencakup tahapan-tahapan penelitian. BAB III, membahas teknik Pembelajaran Mesin yang digunakan, data klinik pasien yang digunakan. BAB IV membahas kesimpulan.

1. METODE



Gambar 1 : Alur Penelitian

Alur dari penelitian ini dapat dilihat pada gambar 1 diatas.

1. Tahap pertama dengan mengumpulkan Data Klinik Pasien, yang dapat diunduh pada *UCI Machine Learning Repository*
2. Tahap kedua, melakukan pra-pemrosesan data, yaitu mengisi nilai missing values dengan menggunakan algoritma *KNN Imputation*, melakukan normalisasi [0,1] pada variable yang memiliki nilai *continous*
3. Tahap ketiga, melakukan pembagian data latih dan data uji menggunakan metode *K-Fold Crossvalidation* dengan nilai K= 10. Data Latih selanjutnya digunakan untuk melatih algoritma ELM dalam mengenali dan memprediksi data pasien pada penyakit Ginjal Akut. Data Uji digunakan untuk melakukan validasi dan mengukur performa dalam memprediksi penyakit Ginjal Akut pada pasien. Mekanisme pengujian yang digunakan akan dijelaskan pada bab selanjutnya.
4. Proses Latihan, dimana Algoritma ELM akan belajar mengenali pola data untuk memprediksi apakah suatu pasien mengidap penyakit Gagal Ginjal Akut atau tidak, proses pelatihan ini juga mencari konfigurasi parameter terbaik dalam memprediksi penyakit Gagal Ginjal Akut. Parameter yang akan diubah konfigurasinya adalah, fungsi aktivasi dan jumlah *Hidden Neuron* pada ELM. Hasil keluaran dari proses ini adalah sebuah model klasifikasi ELM yang memprediksi penyakit Gagal Ginjal Akut pada pasien.
5. Proses Pengujian, melakukan uji coba terhadap model ELM yang didapat pada proses latihan dengan memberikan data uji untuk diprediksi, data uji merupakan data yang tidak diketahui sebelumnya pada proses latihan. Algoritma prediksi Support Vector Machine (SVM), Backpropagation Neural Network (BPN), akan digunakan sebagai pembanding terhadap model ELM.
6. Analisa, Tahap analisa yaitu membandingkan hasil eksperimen pengujian model ELM,SVM dan BPN, terhadap nilai Accuracy, Sensitivity, Specifisity, melakukan analisa mendalam temuan-temuan yang terjadi saat proses Eksperimen. Menyajikan hasil Eksperimen dalam bentuk grafik agar mudah dianalisa secara visual.
7. Laporan Analisa, adalah proses penulisan laporan akhir penelitian ditulis dalam format publikasi *paper*, yang terdiri dari Abstrak, Pendahuluan, Metode, Eksperimen, Hasil Eksperiment, Kesimpulan dan Saran.
8. EKSPERIMEN
   1. **Algoritma**

**Extreme Learning Machine**

Extreme Learning Machine (ELM) merupakan salah satu teknik Machine Learning untuk pengenalan pola, proses pembelajaran dalam ELM menggunakan solusi *Moore Penrose Inverse* dan *Smallest Norm Least-Square.* Dengan memodifikasi proses pembelajaran pada *SLFN* (*Single Layer Feed Forward Neural Network*) dengan L *hidden neurons,* dimana L < N jumlah data pada proses latih. Untuk N data, dimana dan SLFN dapat dinotasikan sebagai berikut

|  |  |
| --- | --- |
|  | (1) |

adalah bobot yang menghubungkan *Input Layer*  dan *Hidden Layer.* adalah bobot yang menghubungkan *Hidden Layer* dengan *Output Layer*. Sedangkan adalah *inner product* dari dengan asumsi, output layer adalah linear.

SLFN dengan L Hidden Neurons menggunakan fungsi aktivasi dapat mendekati setiap data dengan meminimalkan *error*  dikarenakan terdapat yang dapat dinotasikan

|  |  |
| --- | --- |
|  | (2) |

Pendekatan ELM diatas dapat ditulis sebagai berikut

|  |  |
| --- | --- |
|  | (3) |

Dimana

|  |  |
| --- | --- |
| **=** | (4) |

|  |  |
| --- | --- |
|  | (5) |

H merupakan *Hidden Layer Output Matrix* dari *Neural Network,* kolom ke-i pada matrix H merupakan hasil output vektor dari input .

Secara singkat algoritma pembelajaran untuk SLFN Extreme Learning Machine dapat dilakukan dalam tiga tahap :

Diberikan Data latih N = {(**xi,ti**)|xi **R**n,ti **R**m , *i =* 1,..., *N*}, fungsi aktivasi *g(x)*, dan jumlah *hidden neuron* L.

1. Inisialisasi nilai bobot **wi** dan bias bi, i = 1,...,N.
2. Menghitung keluaran pada Hidden Layer dari matrix **H**
3. Menghitung nilai dari bobot

|  |
| --- |
|  |

Dimana **H,** dan **T,** dapat dilihat pada persamaan 4 dan 5.

* 1. **KNN Imputasi**

KNN-Imputasi atau K-Nearest Neighbour Imputasi merupakan metode untuk mengisi nilai yang tidak tersedia pada suatu variable (*missing values*). Metode ini digunakan jika suatu dataset memiliki proposi *missing* values yang cukup besar. Algoritma ini bekerja dengan memanfaatkan nilai terdekat dari kolom yang terdapat *missing values­*, jika kolom tetangga terdekat memiliki *missing values* maka nilai kolom tentangga selanjutnya yang akan digunakan. Kolom tetangga terdekat adalah kolom dengan jarak *Euclidean*.[12]

* 1. **Dataset**

Dataset klinis penyakit Ginjal Kronis yang digunakan pada penelitian ini diambil dari *UCI Machine Learning Repository* yang tersedia secara online. Data yang dikumpulkan sebanyak 400 data, dengan jumlah atribut 25 termasuk atribut kelas. Terdapat 2 kelas pada dataset, “ckd” dan “notckd” (pasien dengan positif penyakit Ginjal Akut, dan pasien dengan negatif penyakit Ginjal Akut) dengan total jumlah “237” ckd dan 149 “notckd”, setiap atribut terdapat *missing values*.

Tabel 1 : Keterangan Dataset Ginjal Akut

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Input** | **Nama Atribut** | **Tipe data** | **Nilai/Satuan** | **Jumlah Missing Values** |
| 1 | Age | Numerical | Normalisai [0,1] | 9 |
| 2 | Blood Pressure | Numerical | Normalisai [0,1] | 9 |
| 3 | Specific Gravity | Nominal | (1.005,1.010,1.015,1.020,1.025) | 45 |
| 4 | Albumin | Nominal | 0,1,2,3,4,5 | 44 |
| 5 | Sugar | Nominal | 0,1,2,3,4,5 | 46 |
| 6 | Red Blood Cell | Nominal | Normal, Abnormal | 143 |
| 7 | Pus Cell | Nominal | Normal, Abnormal | 62 |
| 8 | Pus Cell Clumps | Nominal | Present, Not Present | 4 |
| 9 | Bacteria | Nominal | Present, Not Present | 4 |
| 10 | Blood Glucose Random | Numerical | mgs/dl | 42 |
| 11 | Blood Urea | Numerical | mgs/dl | 18 |
| 12 | Serum Creatinine | Numerical | mgs/dl | 16 |
| 13 | Sodium | Numerical | mEq/L | 81 |
| 14 | Potassium | Numerical | mEq/L | 82 |
| 15 | Hemoglobin | Numerical | gms | 48 |
| 16 | Packed Cell Volume | Numerical | Normalisai [0,1] | 66 |
| 17 | White Blood Cell Count | Numerical | cells/cmm | 99 |
| 18 | Red Blood Cell Count | Numerical | million/cmm | 123 |
| 19 | Hypertension | Nominal | Yes, No | 2 |
| 20 | Diabetes Melitus | Nominal | Yes, No | 2 |
| 21 | Coronary Artery Disease | Nominal | Yes, No | 2 |
| 22 | Appetite | Nominal | Good, Poor | 1 |
| 23 | Pedal Edema | Nominal | Yes, No | 1 |
| 24 | Anemia | Nominal | Yes, No | 1 |
| 25 | Class (Output) | Nominal | ckd,notckd | 0 |

* 1. ***K-Fold Crossvalidation***

*K-Fold Crossvalidation* merupakan salah satu metode dalam membagi dataset menjadi data latih dan data uji, sebanyak k data. Misalkan dengan memilih nilai k = 10, maka dataset akan dibagi menjadi 10 subset data uji dan data latih yang tidak beririsan. Dengan Proporsi 75 % untuk data latih, 25 % untuk data uji.

* 1. **Evaluasi**

Pada proses evaluasi Algoritma, diukur dengan menghitung nilai Akurasi, Sensitivity, Spesificity dengan menggunakan formula dibawah ini

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

TP adalah True Positif yang merujuk pada data pasien penyakit Ginjal Akut yang diprediksi secara benar. TN adalah True Negatif yang merujuk pada data pasien yang tidak mengidap Ginjal Akut yang diprediksi secara benar. FP adalah False Positive yang merujuk pada pasien yang tidak mengidap Ginjal Akut, tetapi diprediksi mengidap Ginjal Akut. FN adalah False Negative yang merujuk pada data pasien Gagal Ginjal Akut tetapi diprediksi mengidap penyakit Gagal Ginjal Akut.

* 1. **Perangkat Pendukung**

Seluruh Eksperimen pada penelitian ini menggunakan satu Perangkat Laptop dengan Processcor Core i7 5th, RAM 4.GB, Windows Operating System 64 Bit. *Source Code* program ditulis menggunakan bahasa pemrograman R, dengan menggunakan R Studio sebagai IDE. Untuk *Source Code* ELM, SVM,dan BPN menggunakan library package caret yang tersedia *online* pada R Studio.

1. KESIMPULAN

**Daftar Pustaka**

[1] “Chronic kidney disease - ClinicalKey.” [Online]. Available: https://remote-lib.ui.ac.id:2194/#!/content/medical\_topic/21-s2.0-1014826. [Accessed: 25-Mar-2016].

[2] Ministry of Health, “Riset Kesehatan Dasar (Basic Health Research),” p. 303, 2013.

[3] J. A. Vassalotti, R. Centor, B. J. Turner, R. C. Greer, M. Choi, and T. D. Sequist, “A Practical Approach to Detection and Management of Chronic Kidney Disease for the Primary Care Clinician.,” *Am. J. Med.*, vol. 129, no. 2, pp. 153–162.e7, 2015.

[4] G. Wang, K.-M. Lam, Z. Deng, and K.-S. Choi, “Prediction of mortality after radical cystectomy for bladder cancer by machine learning techniques,” *Comput. Biol. Med.*, vol. 63, pp. 124–132, 2015.

[5] A. M. Vukicevic, G. R. Jovicic, M. M. Stojadinovic, R. I. Prelevic, and N. D. Filipovic, “Evolutionary assembled neural networks for making medical decisions with minimal regret: Application for predicting advanced bladder cancer outcome,” *Expert Syst. Appl.*, vol. 41, no. 18, pp. 8092–8100, 2014.

[6] N. Chetty, “Role of Attributes Selection in Classification of Chronic Kidney Disease Patients,” pp. 0–5, 2015.

[7] Z. Chen, Z. Zhang, R. Zhu, Y. Xiang, and P. B. Harrington, “Diagnosis of patients with chronic kidney disease by using two fuzzy classifiers,” *Chemom. Intell. Lab. Syst.*, 2016.

[8] Z. You, L. Li, Z. Ji, M. Li, and S. Guo, “Prediction of protein-protein interactions from amino acid sequences using extreme learning machine combined with auto covariance descriptor,” *Memetic Comput. (MC), 2013 …*, pp. 80–85, 2013.

[9] M. Fathurachman and U. Kalsum, “Heart disease diagnosis using extreme learning based neural networks,” *… Concept, Theory …*, 2014.

[10] S. Ismaeel, A. Miri, and D. Chourishi, “Using the Extreme Learning Machine ( ELM ) Technique for Heart Disease Diagnosis,” no. 1, pp. 1–3, 2015.

[11] G.-B. H. G.-B. Huang, Q.-Y. Z. Q.-Y. Zhu, and C.-K. S. C.-K. Siew, “Extreme learning machine: a new learning scheme of feedforward neural networks,” *2004 IEEE Int. Jt. Conf. Neural Networks (IEEE Cat. No.04CH37541)*, vol. 2, pp. 25–29, 2004.