

Reduksi Fitur untuk Optimasi Prediksi Penyakit Ginjal Kronis menggunakan Radial Basis Function Neural Network

Candra Bella Vista¹, Wilda Imama Sabilla²

Teknik Informatika, Teknologi Informasi, Politeknik Negeri Malang, Jl. Soekarno Hatta No.9, Jatimulyo,
Kec. Lowokwaru, Kota Malang, Indonesia, 65141
e-mail: ¹bellavista@polinema.ac.id, ²wildaimama@polinema.ac.id

Submitted Date: August 10th, 2021

Revised Date: January 11st, 2022

Reviewed Date: January 05th, 2022

Accepted Date: August 16th, 2022

Abstract

The chronic kidney disease is one of the diseases that causes many deaths in the world. Patient with chronic kidney disease require intensive treatment. However, the symptoms of chronic kidney disease cannot be discovered directly. It appears gradually until the kidney condition is severe. Therefore, it is important to immediately detect as early as possible that the patient suffers from chronic kidney disease. Prediction of chronic kidney disease has been done quite a lot with various data mining and machine learning methods. This study employ radial basis neural network for classification method. To optimize the classification results compared three feature reduction methods, namely Principal Component Analysis, Linear Discriminant Component, and combination of both methods. The experiment results show that the combined feature reduction method of Principal Analysis Component and Linear Discriminant Analysis shows the best performance with accuracy, recall, and precision values of 93.5%, 91.1% and 97.7%, respectively.

Keywords: feature reduction; principal component analysis; linear discriminant analysis; chronic kidney disease

Abstrak

Penyakit ginjal kronis adalah salah satu penyakit tidak menular yang menjadi penyebab kematian terbanyak di dunia. Penderita ginjal kronis memerlukan perawatan yang cukup serius. Kenyataannya gejala dari penyakit ginjal kronis tidak dapat diketahui secara langsung, melainkan secara bertahap hingga kondisi ginjal pasien sudah tidak berfungsi dengan baik dan sulit disembuhkan. Karena itu menjadi penting untuk dapat mendeteksi sedini mungkin kemungkinan pasien menderita penyakit ginjal kronis. Prediksi penyakit ginjal kronis sudah cukup banyak dilakukan dengan berbagai metode *data mining* dan *machine learning*. Penelitian ini membandingkan tiga metode reduksi fitur, yaitu *Principal Component Analysis*, *Linear Discriminant Component*, serta gabungan *Principal Component Analysis* dan *Linear Discriminant Analysis* untuk mengoptimalkan prediksi penyakit ginjal kronis. Pada tahap pelatihan dan pengujian digunakan metode klasifikasi *Radial Basis Function Neural Network*. Hasil uji menunjukkan bahwa metode reduksi fitur gabungan *Principal Analysis Component* dan *Linear Discriminant Analysis* menunjukkan performa terbaik dengan nilai akurasi, *recall*, dan *precision* secara berturut-turut sebesar 93,5%, 91,1% dan 97,7%.

Kata kunci: reduksi fitur; *Principal Component Analysis*; *Linear Discriminant Analysis*; penyakit ginjal kronis

1. Pendahuluan

Berdasarkan survei Organisasi Kesehatan Dunia (WHO) tahun 2017, kematian penduduk di dunia disebabkan oleh penyakit tidak menular sebesar 73% (WHO, 2017). Penyakit Ginjal Kronis

merupakan salah satu dari beberapa penyakit tidak menular yang sering dikatakan sebagai penyebab utama kematian. Survei WHO juga menyatakan bahwa dominasi penderita penyakit ginjal kronis di seluruh dunia adalah sekitar 10% atau dapat

dikatakan bahwa 1 dari 10 orang di seluruh dunia menderita ginjal kronis (WHO, 2017). Penyakit Ginjal Kronis (PGK) ditimbulkan karena adanya gangguan terhadap fungsi ginjal secara progresif, menahun, persisten, dan bersifat *irreversible* (Purwantoro, 2018). Catatan Riset kesehatan Dasar (Risksedas) tahun 2018 menunjukkan kenaikan jumlah penderita penyakit ginjal kronis di Indonesia dari tahun ke tahun. Angka kenaikannya mencapai 0.38% dalam kurun waktu 5 tahun (Kemenkes, 2018).

Penderita penyakit ginjal kronis memerlukan perawatan yang serius. Gejala yang ditimbulkan dari penyakit ginjal kronis tidak muncul secara langsung, melainkan secara bertahap hingga terjadi penurunan fungsi ginjal yang parah dan sulit diobati (Rostanti et.al., 2016). Untuk itu sangat penting untuk mengetahui sedini mungkin kemungkinan pasien menderita gagal ginjal.

Seiring dengan berkembangnya teknologi, teknik *data mining* dan *machine learning* digunakan untuk menemukan pola dan hubungan antar data, sehingga dapat digunakan untuk melakukan prediksi. Prediksi dibangun dengan membandingkan model dan data masukkan yang diberikan. Berbagai penelitian sebelumnya melibatkan teknik *data mining* dan *machine learning* untuk menyelesaikan permasalahan di bidang kesehatan, salah satunya dalam hal memprediksi berbagai penyakit gagal ginjal kronis. Penelitian sebelumnya memprediksi penyakit ginjal kronis menggunakan algoritma *K-Nearest Neighbor* (Wisnuadji & Waspada, 2020), *Naïve Bayes* dan *Support Vector Machine* (Vijayanti et. al., 2015), serta membandingkan beberapa metode, seperti *Naïve Bayes*, *Decision Rule*, *Support Vector Machine*, dan *Multilayer Perceptron* (Jena & Kamila, 2015), serta membandingkan *K-Nearest Neighbor*, *Random Forest*, dan *Neural Network* (Salekin & Stankovic, 2016). Penelitian-penelitian tersebut berusaha mendapatkan akurasi terbaik dalam hal prediksi penyakit ginjal kronis, sehingga dapat membantu petugas medis dalam menentukan diagnosis dan penanganan tepat sedini mungkin kepada pasien dengan penyakit ginjal kronis.

Permasalahan tersebut yang kemudian menjadi latar belakang dalam penelitian ini. Optimasi teknik pemodelan untuk menghasilkan prediksi penyakit gagal ginjal kronis yang akurat menjadi fokus dalam penelitian ini. Beberapa penelitian sebelumnya juga melakukan optimasi terhadap algoritma tertentu untuk menghasilkan prediksi penyakit ginjal kronis yang lebih handal.

Penelitian sebelumnya memprediksi penyakit ginjal kronis dengan melibatkan optimasi terhadap algoritma *Naïve Bayes* menggunakan algoritma *Particle Swarm Optimization* (Arifin & Ariesta, 2019), dan juga penelitian yang menggunakan metode *symmetrical uncertainty* untuk pengurangan fitur (Kurniawan, 2020). Pada penelitian ini digunakan teknik reduksi dimensi fitur berbasis *Principal Component Analysis* (PCA) dan *Linear Discriminant Analysis* (LDA) yang akan mengoptimalkan prediksi penyakit ginjal kronis menggunakan *Radial Basis Function Neural Network*.

2. Metode Penelitian

2.1. Dataset

Penelitian memanfaatkan dataset *opensource* yang berasal dari UCI *Machine Learning Repository*. Dataset tersebut terdiri atas 400 baris data yang terdiri dari 24 atribut diagnosis pasien penyakit gagal ginjal kronis dan 1 atribut kelas yang merupakan hasil klasifikasi prediksi pasien, yaitu CKD dan NOTCKD dengan sebaran data masing-masing 63% CKD dan 37% NOTCKD. Tabel 1 menunjukkan deskripsi dataset yang digunakan.

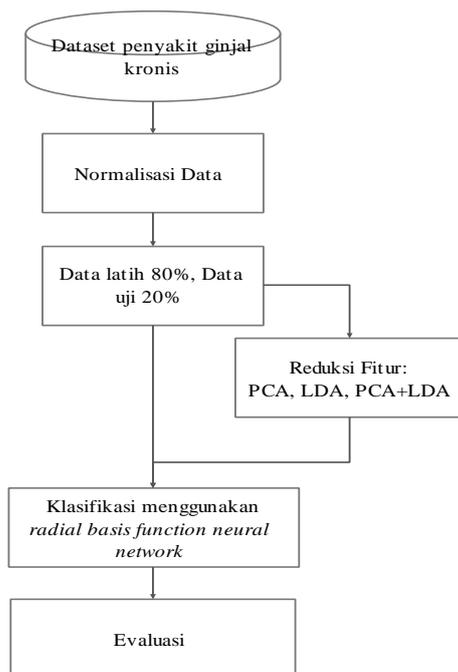
Tabel 1. Deskripsi Dataset
(Sumber: UCI *Machine Learning*)

| Atribut | Nilai/Satuan |
|-------------------------|------------------------------|
| Umur | Tahun |
| Tekanan Darah | mm/Hg |
| Specific Gravity | (1.005, 1.010, 1.020, 1.025) |
| Albumin | (0,1,2,3,4,5) |
| Sugar | (0,1,2,3,4,5) |
| Red Blood Cells | (normal, abnormal) |
| Pus Cell | (normal, abnormal) |
| Pus Cell Clumps | (present, notpresent) |
| Bacteria | (present, notpresent) |
| Blood Glucose Random | mgs/dL |
| Blood Urea | mgs/dL |
| Serum Creatinine | mEq/dL |
| Sodium | mEq/L |
| Potassium | mEq/L |
| Hemoglobin | Gms |
| White Blood Cell Count | cells/cmm |
| Red Blood Cell Count | million/cmm |
| Hypertension | yes, no |
| Diabetes Mellitus | yes, no |
| Coronary Artery Disease | yes, no |
| Appetite | good, poor |
| Pedal Edema | yes, no |
| Anemia | yes,no |
| Class | ckd, notckd |

Pada dataset ditemukan 70% dari dataset memiliki atribut kosong. Untuk mengisi nilai atribut kosong digunakan metode *mean imputation*. Metode tersebut mengubah nilai atribut kosong (*missing value*) dengan cara membagi jumlah keseluruhan nilai atribut diketahui dengan total atribut atau fitur yang digunakan (Maheswari et. al, 2019).

2.2. Arsitektur Sistem

Penelitian ini menggunakan dataset penyakit ginjal kronis yang terdiri dari 400 baris data, dimana terdapat 11 atribut numerik dan 14 atribut nominal. Untuk mendapatkan interpretasi yang seragam, maka sebelum dilakukan proses klasifikasi, terlebih dahulu dilakukan normalisasi data. Proses selanjutnya adalah membagi data menjadi data latih dan data uji. Pembagian data latih dan data uji ini adalah 4:1 atau 80% data latih dan 20 data uji. Untuk mengetahui efektivitas model yang diusulkan dibangun *baseline system*. *Baseline system* dibangun dengan memodelkan data secara langsung menggunakan *radial basis function neural network*. Kemudian hasil klasifikasi dari *baseline system* akan dibandingkan dengan hasil klasifikasi dari data yang telah melewati proses seleksi fitur yang melibatkan metode *Principal Component Analysis* (PCA), *Linear Discriminant Analysis* (LDA) dan gabungan keduanya. Gambar 1 menunjukkan gambaran usulan sistem yang dibangun pada penelitian ini.



Gambar 1. Arsitektur Sistem

2.2.1. Normalisasi Data

Atribut-atribut yang terdapat pada dataset penyakit ginjal kronis terdiri dari dua tipe data, yaitu numerik dan nominal. Agar diperoleh interpretasi yang seragam, diperlukan suatu proses normalisasi data. Normalisasi data dilakukan menggunakan teknik *minmax scaler*, dimana data numerik dan nominal terkecil akan diubah menjadi 0, data terbesar diubah menjadi 1, sedangkan data diantaranya diubah pada jangkauan nilai antara 0 dan 1 (Raju et al., 2020). Untuk data yang berupa pernyataan *yes* atau *no* secara berturut-turut akan diubah juga menjadi nilai 1 dan 0.

2.2.2. Reduksi Fitur

Salah satu strategi yang dapat digunakan untuk mengoptimasi algoritma klasifikasi adalah dengan mereduksi fitur. Dua tipe reduksi fitur yang dapat digunakan adalah seleksi dan ekstraksi fitur. Seleksi fitur merupakan suatu teknik untuk memilih fitur yang dianggap penting dalam proses klasifikasi, sedangkan proses ekstraksi dilakukan dengan membandingkan data yang akan di klasifikasikan dengan fitur baru yang memiliki jumlah yang lebih sedikit. Penelitian ini melibatkan dua metode ekstraksi fitur, yaitu *Principal Component Analysis* (PCA) dan *Linear Discriminant Analysis* (LDA).

Teknik reduksi fitur yang umum digunakan salah satunya adalah *Principal Component Analysis*. Teknik *Principal Component Analysis* akan mentransformasikan data ke dalam sistem koordinat baru, sehingga hasil reduksi tidak akan menghilangkan informasi penting yang menjadi karakteristik suatu data (Sehgal et. al., 2014). Langkah-langkah reduksi data menggunakan metode *Principal Component Analysis* dimulai dari menghitung rata-rata dari nilai vektor yang diinputkan. Kemudian dihitung matriks *covariance* untuk setiap vektor. Dari hasil tersebut dihitung *eigen* vektor dan *eigen value*. Vektor yang dipilih adalah vektor fitur yang memiliki nilai *eigen* tertinggi.

Teknik reduksi fitur lainnya yang tergolong umum digunakan dan memiliki akurasi yang baik adalah *Linear Discriminant Analysis*. Metode ini menghitung kombinasi linear dari fitur yang akan dimodelkan dengan tujuan untuk mengklasifikasikan *sample* sesuai dengan kelasnya (Khalid et. al., 2016). Tujuan menggunakan *Linear Discriminant Analysis* adalah untuk mereduksi dimensi data dengan cara menghitung jarak antar

kelas dan jarak data dalam suatu kelas. Langkah-langkah seleksi fitur dengan *Linear Discriminant Analysis* dimulai dari input dataset, selanjutnya menghitung vektor rata-rata (*mean*) dengan mencari rata-rata untuk setiap atribut di dalam dataset untuk setiap kelas. Dari nilai rata-rata vektor kemudian dihitung nilai *covariance* kelas dan nilai *covariance* antar kelas. Selanjutnya dihitung *eigen* vektor dan *invers* dari matriks *covariance*. Langkah selanjutnya adalah mengurutkan nilai hasil perhitungan secara *descending*. Reduksi fitur dilakukan pada vektor yang memiliki nilai *eigen* terendah.

Untuk mereduksi fitur pada dataset penyakit ginjal kronis, pada penelitian ini digunakan *hybrid* metode *Principal Component Analysis* dan *Linear Discriminant Analysis*. Hasil matriks *eigen* pada proses reduksi *Principal Component Analysis* digunakan sebagai input untuk reduksi menggunakan *Linear Discriminant Analysis*, sebelum pada akhirnya diklasifikasikan menggunakan metode *Radial Basis Function Neural Network*.

2.2.3. Radial Basis Function Neural Network

Penerapan metode *radial basis function neural network* telah banyak dilakukan untuk memprediksi beberapa kasus dalam bidang kesehatan. Contohnya ada pada penelitian sebelumnya yang memanfaatkan metode *radial basis function neural network* untuk mendiagnosis penyakit diabetes (Hosseini & Bardsiri, 2019) dan mengevaluasi laju filtrasi glomerulus untuk menentukan diagnosis penyakit ginjal (Xu et. al., 2020). Melihat keberhasilan penelitian-penelitian tersebut, maka pada penelitian ini prediksi penyakit ginjal kronis digunakan metode *radial basis function neural network*. Secara umum metode ini adalah turunan dari metode *neural network*, perbedaannya adalah pada pemrosesan sinyal dan fungsi aktivasinya (Qasem & Mariyam, 2011). Sinyal yang menuju lapisan tersembunyi diproses secara non linear, sedangkan sinyal yang keluar menuju lapisan keluaran akan diproses secara linear. Pada metode ini fungsi aktivasi yang digunakan berbasis radial. Perbedaan lainnya adalah jaringan bersifat searah atau tidak bolak-balik.

Proses pelatihan *radial basis function neural network* mengalami dua kali proses, yaitu *unsupervised* dan *supervised*. Proses *unsupervised* melibatkan algoritma *K-means* menghasilkan pusat *cluster* dan sebaran. Pusat *cluster* dihitung

berdasarkan jarak *euclidian* seperti pada persamaan 1.

$$\sqrt{\sum_{j=1}^N (\mu_j - x_j)^2} \quad (1)$$

Di mana x_j adalah data masukan ke- j dan μ_j adalah rata-rata dari kelas ke- j . Pada nilai sebaran dihitung secara *random* berdasarkan nilai data. Pusat *cluster* dan sebaran digunakan untuk mengaktifkan fungsi aktivasi pada lapisan tersembunyi. Pada penelitian ini digunakan fungsi *Gaussian* untuk aktivasi pada lapisan tersembunyi seperti pada persamaan 2.

$$\varphi_j(x) = \exp\left(-\frac{(x-c_j)^2}{2\sigma_j^2}\right) \quad (2)$$

Keluaran dari *radial basis function neural network* dihitung dengan mengalikan *neuron* ke- j pada lapisan tersembunyi dan bobotnya seperti pada persamaan 3.

$$y = \sum_{j=1}^p w_j \varphi_j(x) \quad (3)$$

Untuk setiap iterasi bobot baru didapatkan dengan menggunakan persamaan *least means square* seperti pada persamaan 4.

$$w(k+1) = w(k) + \alpha (d(k) - y(k)) x(k) \quad (4)$$

Di mana α adalah *learning rate* dan $w(k+1)$ adalah bobot baru. *Radial basis function neural network* dievaluasi menggunakan parameter akurasi berdasarkan *Mean Square Error* (MSE) seperti pada persamaan 5.

$$f_j = \frac{1}{N} \sum_{j=1}^N (t_j - o_j)^2 \quad (5)$$

Di mana o_j adalah unit output dan t_j adalah target.

2.3. Evaluasi

Dalam penelitian ini evaluasi performa sistem digunakan parameter akurasi, *precision*, dan *recall* (Christopher et. al., 2008). *Confusion matrix* disusun dengan empat representasi hasil klasifikasi, yaitu *True Positif* (TF), *True Negative* (TN), *False Positif* (FP), dan *False Negative* (FN). Hasil *confusion matrix* akan digunakan sebagai dasar menghitung akurasi, *precision*, dan *recall*.

Akurasi menunjukkan performa sistem dalam memprediksi penyakit ginjal kronis, yang didefinisikan sebagai perbandingan data yang diklasifikasikan benar dengan keseluruhan data.

Persamaan 6 menunjukkan bagaimana akurasi didapatkan.

$$akurasi = \frac{TP+TN}{TP+FP+FN+TN} \quad (6)$$

Precision didapatkan dari membagi jumlah data positif yang diprediksi secara benar dengan seluruh data yang diprediksi positif. *Recall* adalah ukuran yang menyatakan perbandingan data positif yang dikelompokkan ke dalam kelas yang benar dengan keseluruhan data positif.

3. Hasil dan Pembahasan

Tahapan awal dalam penelitian ini dimulai dari tahap persiapan data. Dataset dikumpulkan dari UCI *Machine Learning Repository*, terdiri dari 400 data. Setelah melewati serangkaian proses normalisasi, maka data dibagi menjadi data latih dan data uji dengan rasio 4:1. Total data latih sebanyak 320, sedangkan data uji sebanyak 80 data. Selanjutnya pengujian untuk memprediksi penyakit ginjal kronis digunakan metode *Radial Basis Function Neural Network*.

Skenario uji pertama adalah membuat *baseline system*. Dataset akan langsung dimodelkan menggunakan *radial basis function neural network* tanpa melalui proses reduksi fitur. Dari pengujian didapatkan hasil *confussion matrix* seperti yang ditunjukkan pada Tabel 2.

Tabel 2. *Confussion Matrix Baseline System*

| | True: CKD | True: NOTCKD |
|--------------|-----------|--------------|
| Pred: CKD | 211 | 8 |
| Pred: NOTCKD | 32 | 149 |

Untuk mengetahui kinerja metode *Principal Component Analysis* (PCA) dan *Linear Discriminant Analysis* (LDA) terhadap tingkat akurasi prediksi penyakit ginjal kronis, maka dilakukan pengujian dengan melibatkan metode reduksi fitur secara terpisah. Didapatkan hasil seperti pada Tabel 3 dan Tabel 4.

Evaluasi selanjutnya dilakukan dengan membandingkan *baseline system* terhadap usulan sistem yang menerapkan reduksi fitur menggunakan gabungan metode *Principal Component Analysis* dan *Linear Discriminant Analysis* (PCA+LDA). Tabel 5 menunjukkan *confussion matrix* hasil prediksi penyakit ginjal kronis dengan teknik reduksi fitur *Principal Component Analysis* dan *Linear Discriminant Analysis* (PCA+LDA).

Berdasarkan *confussion matrix* pada Tabel 3, 4, dan 5 dihitung nilai akurasi, *recall*, dan *precision*

untuk prediksi penyakit ginjal kronis. Komparasi akurasi, *recall*, dan *precision* dari keempat model ditunjukkan pada Tabel 6.

Tabel 3. *Confussion Matrix PCA*

| | True: CKD | True: NOTCKD |
|--------------|-----------|--------------|
| Pred: CKD | 215 | 5 |
| Pred: NOTCKD | 30 | 150 |

Tabel 4. *Confussion Matrix LDA*

| | True: CKD | True: NOTCKD |
|--------------|-----------|--------------|
| Pred: CKD | 218 | 7 |
| Pred: NOTCKD | 29 | 146 |

Tabel 5. *Confussion Matrix PCA+LDA*

| | True CKD | True NOTCKD |
|--------------|----------|-------------|
| Pred: CKD | 217 | 5 |
| Pred: NOTCKD | 21 | 157 |

Tabel 6. Komparasi Akurasi, *Recall*, dan *Precision*

| | baseline | PCA | LDA | PCA+LDA |
|-----------|----------|------|------|---------|
| Akurasi | 90 | 91 | 91 | 93.5 |
| Recall | 86.8 | 88.2 | 87.7 | 91.1 |
| Precision | 96.3 | 96.8 | 97.2 | 97.7 |

Hasil pengujian pada Tabel 6 menunjukkan bahwa pengujian dengan metode reduksi fitur meningkatkan akurasi rata-rata sebesar 1,9% dibandingkan dengan model *baseline*. Hasil uji dengan parameter *recall* dan *precision* juga menunjukkan rata-rata nilai yang lebih tinggi dibandingkan model *baseline*. Diketahui pada Tabel 6 reduksi fitur meningkatkan nilai *recall* dan *precision* prediksi penyakit ginjal kronis rata-rata sebesar 2.2% dan 0.93%.

Pada Tabel 6 terlihat bahwa dibandingkan dengan model *baseline*, model dengan reduksi fitur *Principal Analysis Component* akurasinya meningkat sebesar 1%. Namun jika dibandingkan dengan metode reduksi fitur *Linear Analysis Discriminant*, akurasi metode reduksi fitur dengan *Linear Analysis Discriminant* sedikit lebih baik. Selisih akurasinya mencapai 0,2%. *Recall* dan *Precision* model prediksi penyakit ginjal kronis dengan reduksi fitur *Principal Component Analysis* meningkat sebesar 2,6% dan 0,5% dibandingkan model *baseline*. Berbanding lurus dengan akurasinya, reduksi fitur dengan metode *Principal*

Analysis Component juga menghasilkan nilai *precision* yang lebih rendah dibandingkan dengan metode *Linear Analysis Discriminant*, yaitu selisih sebesar 0,4%, namun nilai *recall* model dengan reduksi fitur *Principal Analysis Component* lebih baik dibanding dengan model dengan reduksi fitur *Linear Component Analysis*, yaitu meningkat sebesar 0,5%.

Metode reduksi fitur menggunakan gabungan metode *Principal Analysis Component* dan *Linear Discriminant Analysis* menghasilkan akurasi terbaik dalam kasus prediksi penyakit ginjal kronis. Akurasinya mencapai 93,5%. Akurasi metode *hybrid* dalam reduksi fitur 3,5% lebih baik dibanding model *baseline*, 2,5% dan 2,3% lebih tinggi dibanding menggunakan metode reduksi fitur *Principal Analysis Component* dan *Linear Discriminant Analysis* secara terpisah. Peningkatan akurasi yang dihasilkan sebanding dengan *recall* dan *precision*. Penggunaan metode gabungan *Principal Analysis Component* dan *Linear Discriminant Analysis* memberikan perbaikan *recall* dan *precision* rata-rata sebesar 3% dan 0,93% dibanding ketiga model lainnya.

4. Kesimpulan

Penelitian ini membandingkan tiga metode reduksi fitur untuk mengoptimalkan prediksi penyakit ginjal kronis. Ketiga metode reduksi fitur yang digunakan adalah *Principal Component Analysis*, *Linear Discriminant Component*, serta gabungan *Principal Component Analysis* dan *Linear Discriminant Analysis*. Ketiga model dilatih dan diuji menggunakan metode klasifikasi *Radial Basis Function Neural Network*. Hasil pengujian menunjukkan bahwa reduksi fitur mempengaruhi performa klasifikasi. Reduksi fitur meningkatkan akurasi, *recall*, dan *precision* rata-rata sebesar 1,9%, 2,2%, dan 0,93% dibanding model *baseline*. Akurasi terbaik diperoleh dengan menggunakan metode reduksi fitur gabungan *Principal Component Analysis* dan *Linear Discriminant Analysis*, yaitu sebesar 93,5%. Hasil *recall* dan *precision* juga menunjukkan performa terbaik, yaitu sebesar 91,1% dan 97,7%.

References

Purwanto D. (2013). Penyakit Ginjal Kronik yang Terjadi pada Pasien dengan Faktor Risiko Hipertensi. *MEDULA* vol.1, no. 1.
Kemenkes. (2018). Situasi Penyakit Ginjal Kronik. Pusat Data dan Informasi Kemenkes RI, 2018.

Rostanti, A., Bawotong, J., dan Onibala, F. (2016). Faktor-Faktor yang Berhubungan dengan Kepatuhan Menjalani Terapi Hemodialisa pada Penyakit Ginjal Kronik di Ruang Dahlia dan Melati RSUP Prof. Dr. R. D Kandau Manado. *Ejournal Keperawatan*, 4 nomor 2.
Wisnuadji dan Waspada. (2020). Implementasi Data Mining untuk Deteksi Penyakit Ginjal Kronis (PGK) Menggunakan K-Nearest Neighbor (KNN) dengan Backward Elimination. *JTIK*, vol. 7, no. 2.
Vijayarani, S., Dhayanand et al. (2015). Data mining classification algorithms for kidney disease prediction. *International Journal on Cybernetics & Informatics (IJCI)*, vol. 4, no. 4, pp. 13–25.
Jena, L. and Kamila, N. K. (2015). Distributed data mining classification algorithms for prediction of chronic-kidney-disease. *International Journal of Emerging Research in Management & Technology*, vol. 4, no. 11, pp. 110–118.
Salekin A. and Stankovic J. (2016) Detection of chronic kidney disease and selecting important predictive attributes. *2016 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE, 2016, pp. 262–270.
Arifin T. dan Ariesta, D. (2019). Prediksi Penyakit Ginjal Kronis menggunakan Algoritma Naïve Bayes Classifier berbasis Particle Swarm Optimization. *Jurnal Tekno Insentif* vol. 13, no. 1, pp. 26-30.
Kurniawan, M. (2020). Prediksi Penyakit Ginjal Kronis dengan Metode Pengurangan Fitur Symmetrical Uncertainty. *Jnanaloka*.
Maheswari, K., Priya, Ramkumar S., and Arun. (2019). Missing data handling by mean imputation method and statistical analysis of classification algorithm. *International Conference on Big Data Innovation for Sustainable Cognitive Computing*.
Raju, Lakshmi, Scholar, Kalidindi, and Padma. (2020). Study the Influence of Normalization/Transformation Process on the Supervised Classification. *3rd International Conference on Smart Systems and Inventive Technology*.
Sehgal, S., Singh, H., Agarwal, M., Bhasker, V., and Shantanu. (2014). Data Analysis using Principal Component Analysis. *2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*.
Khalid, M.I., Alotaiby, T., Aldosari, S.A., Alshebeili, S.A., Al-Hameed, Almohammed, and Alotaibi. (2016). Epileptic MEG Spikes Detection using Common Spatial Patterns and Linear Discriminant Analysis. *IEEE Access* Vol 4.
Hosseini and Bardsiri, A.K. (2019). Improving Diagnosis Accuracy of Diabetic Disease using Radial Basis Function Network and Fuzzy Clustering. *Hamara Journals*.

- Xu, J., Guo, B., and Liu, C. (2020). Evaluation of Glomerular Filtration Rate in Chronic Kidney Disease by Radial Basis Function Neural Network. *Transplantation Proceedings*.
- Qasem, S.N and Mariyam, S.S. (2011). Radial basis function based on time variant multi objective particle swarm optimization for medical diseases diagnosis. *Applied Soft Computing* vol. 11 no 1.
- Christopher, D., Prabhakar, R., and Hinrich, S. (2008). Introduction to information retrieval”, *An Introduction To Information Retrieval*, pp.151-177.

