



Model Prediksi Penyakit Ginjal Menggunakan Algoritma *Neural Network*

Khoirul Mustofa^{1*}, Linda Widyaningrum², Wiji Lestari³, Tri Yusnanto⁴

^{1,2}Fakultas Ilmu Kesehatan, D3-Rekam Medik dan Informasi Kesehatan, Universitas Duta Bangsa Surakarta, Indonesia

³Fakultas Ilmu Komputer, S1-Teknik Informatika, Universitas Duta Bangsa Surakarta, Indonesia

⁴Manajemen Informasi, STMIK Bina Patria Magelang, Indonesia

Email : khoirul_mustofa@udb.ac.id^{1*}, linda_widya@udb.ac.id², wiji_lestari@udb.ac.id³, yusnanto@stmikbinapatria.ac.id⁴

Alamat Kampus: Jl. K.H. Samanhudi No.93, Sondakan, Kec. Laweyan, Kota Surakarta, Jawa Tengah 57147

Korespondensi penulis: khoirul_mustofa@udb.ac.id

Abstract. *Neural networks are the focal point of machine learning research. Their applications span across various fields and solve complex and intricate problems. Neural networks have now been applied in health image processing to detect various diseases such as cancer and diabetes. Another disease that threatens our health is kidney disease. This disease is becoming more common due to the substances and elements we consume. Death is imminent and inevitable within a few days without at least one functioning kidney. Ignoring kidney damage can lead to chronic kidney disease leading to death. Often, Kidney Disease and its symptoms are mild and gradual, often going unnoticed for years and only being realized recently. This study used 400 patients with 10 attributes as our dataset from Bade General Hospital. We used a Neural network model to predict the presence or absence of kidney disease susceptibility in patients. The model yielded an accuracy of 98%. Furthermore, we identified and highlighted the important features to rank the features used in kidney disease prediction. The results revealed that two attributes; Creatinine and Bicarbonate had the highest influence on kidney disease susceptibility prediction.*

Keywords: *Medical data, machine learning, predictive models, neural networks, kidney disease.*

Abstrak. Neural network menjadi titik fokus dalam penelitian machine learning. Aplikasinya merambah ke berbagai bidang dan memecahkan masalah yang rumit dan kompleks. Neural network kini telah diterapkan dalam pemrosesan citra kesehatan untuk mendeteksi berbagai penyakit seperti kanker dan diabetes. Penyakit lain yang mengancam kesehatan kita adalah penyakit ginjal. Penyakit ini semakin umum terjadi karena zat dan elemen yang kita konsumsi. Kematian sudah dekat dan tak terelakkan dalam beberapa hari tanpa setidaknya satu ginjal yang berfungsi. Mengabaikan kerusakan ginjal dapat menyebabkan penyakit ginjal kronis yang berujung pada kematian. Seringkali, Penyakit Ginjal dan gejalanya ringan dan bertahap, sering kali tidak disadari selama bertahun-tahun dan baru disadari akhir-akhir ini. Penelitian ini menggunakan 400 pasien dengan 10 atribut sebagai kumpulan data kami dari Rumah Sakit Umum Bade. Kami menggunakan model Neural network untuk memprediksi ada atau tidaknya kerentanan penyakit ginjal pada pasien. Model tersebut menghasilkan akurasi sebesar 98%. Lebih jauh, kami mengidentifikasi dan menyortir pentingnya Fitur untuk memberikan peringkat fitur yang digunakan dalam prediksi penyakit ginjal. Hasilnya mengungkapkan bahwa dua atribut; Kreatinin dan Bikarbonat memiliki pengaruh tertinggi pada prediksi kerentanan penyakit ginjal.

Kata kunci: Data medis, machine learning, model prediksi, neural network, penyakit ginjal.

1. LATAR BELAKANG

Penyakit ginjal merupakan salah satu penyakit kronis yang prevalensinya terus meningkat secara global, termasuk di Indonesia. Deteksi dini terhadap risiko penyakit ginjal sangat penting untuk mencegah perkembangan penyakit menuju stadium lanjut yang membutuhkan terapi mahal seperti hemodialisis atau transplantasi ginjal. Sayangnya, banyak

kasus penyakit ginjal yang tidak terdeteksi pada tahap awal karena gejalanya sering kali tidak spesifik atau bahkan tidak muncul sama sekali. Dalam upaya mendeteksi risiko secara dini, pemanfaatan data klinis seperti umur, jenis kelamin, tekanan darah, kadar gula darah, kadar kolesterol, serta riwayat genetika, dapat menjadi indikator penting. Variabel-variabel ini berperan signifikan dalam mempengaruhi fungsi ginjal dan merupakan faktor risiko yang sering dikaitkan dengan penyakit ginjal kronis.

Perkembangan teknologi kecerdasan buatan, khususnya machine learning, memungkinkan pembangunan sistem prediksi berbasis data yang lebih akurat dan efisien. Salah satu algoritma yang umum digunakan dalam klasifikasi adalah algoritma backpropagation, yaitu algoritma pembelajaran terarah pada jaringan saraf tiruan (*neural network*). Algoritma ini mampu mempelajari pola kompleks dari data masukan dan melakukan klasifikasi ke dalam dua kategori, yaitu "rentan penyakit" dan "tidak rentan penyakit". Mesin learning adalah bidang ilmiah yang menangani cara-cara mesin belajar dari pengalaman. Bagi banyak ilmuwan, istilah machine learning identik dengan istilah kecerdasan buatan, mengingat kemungkinan pembelajaran adalah karakteristik utama dari suatu entitas yang disebut cerdas dalam arti kata yang paling luas (Joshi dan Chawan, 2018). Tujuan dari pembelajaran mesin adalah pembangunan sistem komputer yang dapat beradaptasi dan belajar dari pengalaman mereka. Dengan penggunaan pembelajaran mesin untuk mengekstrak informasi yang berguna dapat digunakan dalam memecahkan berbagai masalah, termasuk penyakit ginjal. Purusothaman dan Krishnakumari, (2015) menunjukkan bahwa gagal ginjal termasuk dalam beberapa kelas penyakit seperti penyakit jantung, kebutaan, dll. yang disebabkan oleh diabetes kronis. Dialisis dan transplantasi adalah satu-satunya metode untuk menjaga ginjal berfungsi secara artifisial dan itu juga merupakan proses yang menyakitkan dan mahal. Menurut Luyckx, dan Stanifer, (2018) penyakit ginjal meningkat secara global dari 19 juta pada tahun 1990 menjadi 33 juta pada tahun 2013. Pada tahun 2016 dan tahun 2010, 2,62 juta orang menerima dialisis di seluruh dunia dan kebutuhan akan dialisis diproyeksikan akan berlipat ganda pada tahun 2030.

Dengan menggunakan pendekatan ini, sistem prediksi dapat dikembangkan untuk membantu tenaga medis dalam mengidentifikasi individu yang berisiko tinggi terkena penyakit ginjal. Selain meningkatkan efisiensi dalam skrining awal, sistem ini juga dapat berfungsi sebagai alat bantu dalam pengambilan keputusan klinis berbasis data. Oleh karena itu, penelitian ini bertujuan untuk membangun model prediksi risiko penyakit ginjal menggunakan algoritma backpropagation dengan parameter masukan berupa umur, jenis kelamin, tekanan darah, kadar gula darah, kolesterol, dan riwayat genetika, serta keluaran berupa klasifikasi

individu ke dalam kelompok rentan atau tidak rentan terhadap penyakit ginjal. Metode machine learning mengoordinasikan berbagai analisis statistik dan basis data membantu mengekstrak pola dan hubungan tersembunyi dari data variabel yang besar dan beragam. Keakuratan pengklasifikasi tertentu dipastikan melalui pengujian model atau teknik. Selain itu, atribut seperti Spesifisitas, Sensitivitas, dan Akurasi umum digunakan untuk deteksi penyakit (Padmanaban dan Parthiban 2016). Dalam makalah ini kami menggunakan model *Neural Network* untuk memprediksi penyakit ginjal dan mengevaluasi kinerja model dengan menghitung Specificity, Sensitivity, Recall, ROC Score, Kohens Kappa, F1 Score, dan Precision.

2. KAJIAN TEORITIS

Sistem Prediksi dalam Bidang Kesehatan

Penerapan sistem prediksi dalam sektor kesehatan merupakan hasil integrasi teknologi informasi untuk mendukung proses diagnosis dan estimasi risiko penyakit secara lebih awal dan akurat. Sistem ini bekerja dengan mengolah data rekam medis masa lalu dan menerapkan metode komputasi cerdas untuk mendeteksi pola-pola yang menunjukkan kemungkinan munculnya atau berkembangnya suatu penyakit. Salah satu pendekatan yang cukup populer adalah penggunaan algoritma pembelajaran mesin, khususnya jaringan saraf tiruan (Neural Network), karena kemampuannya dalam mengolah data yang bersifat kompleks dan tidak linear (Jahan et al., 2021).

Artificial Neural Network (ANN)

Artificial Neural Network (ANN) merupakan salah satu teknik dalam machine learning yang meniru mekanisme kerja otak manusia. ANN efektif dalam memetakan hubungan yang rumit antara variabel input dan output, serta mengenali pola dalam data. Struktur ANN terdiri atas tiga bagian utama, yaitu lapisan input, lapisan tersembunyi (hidden layer), dan lapisan output. Dalam konteks prediksi penyakit ginjal, ANN dimanfaatkan untuk menganalisis sejumlah parameter medis seperti tekanan darah, kadar kreatinin, kadar protein dalam urin, dan parameter lainnya guna menentukan tingkat risiko gangguan fungsi ginjal pada pasien (Saxena & Sharma, 2020).

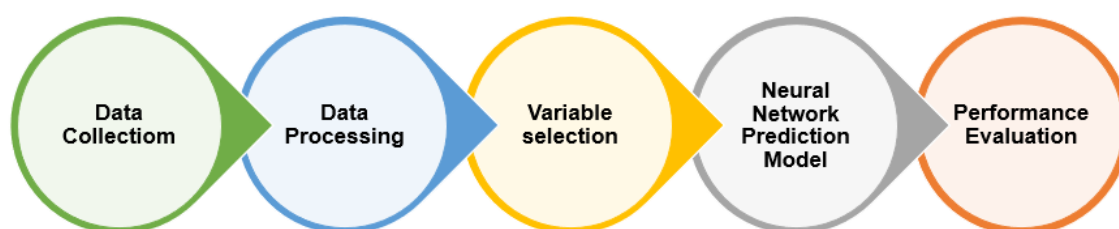
Penyakit Ginjal dan Faktor Risikonya

Penyakit ginjal kronik (PGK) merupakan kondisi medis progresif yang ditandai dengan menurunnya fungsi ginjal dalam jangka waktu yang panjang. Beberapa faktor yang berkontribusi terhadap risiko penyakit ini antara lain adalah hipertensi, diabetes melitus, riwayat keluarga dengan penyakit ginjal, serta gaya hidup yang tidak sehat. Deteksi dini

terhadap faktor-faktor tersebut sangat penting agar penanganan dan pencegahan dapat dilakukan sejak dini. Oleh karena itu, penggunaan model prediksi berbasis kecerdasan buatan sangat membantu tenaga kesehatan dalam mendukung proses pengambilan keputusan klinis yang lebih tepat (Khurana et al., 2019).

3. METODE PENELITIAN

Metode ini menggambarkan desain metode yang akan digunakan untuk melaksanakan eksperimen. Metodologi ini menggabungkan pengumpulan data, praproses data, dan pemilihan variabel target, model prediksi *Neural Network* dan evaluasi. Penelitian ini dilaksanakan dengan pendekatan penelitian dan pengembangan dengan tahapan kinerja seperti yang ditunjukkan pada Gambar 1.



Gambar 1 Tahapan Penelitian

Data Collection

Record penyakit ginjal pasien dipilih sebagai sumber data untuk pekerjaan ini.. Kumpulan data ini berisi catatan 400 pasien dengan 11 atribut/parameter: Usia, Jenis Kelamin, Natrium, Kalium, Klorida, Bikarbonat, Urea, Kreatinin, Asam Urea, Albumin, dan Klasifikasi termasuk variabel target yang diklasifikasikan ke dalam klasifikasi biner penyakit CKD dan non CKD seperti yang ditunjukkan pada Tabel 1.

Tabel 1. Atribut Dataset

Atribut	Keterangan
Sex	gender/jenis kelamin
Age	umur
Sod	Sodium/Natrium
Pot	Potasium
Chl	Chlorida
Bica	Bicarbonat
Urea	Ureum
Cre	Kreatinin
UA	Asam Urea
Alb	Albumin
Class	Rentan (1), Tidak rentan (0)

Data Processing

Data processing merupakan tugas terpenting dalam teknik *data mining*, yang melibatkan pembersihan, ekstraksi, dan transformasi data ke dalam format yang sesuai untuk eksekusi mesin. Data mentah berisi informasi yang hilang, format yang buruk, informasi yang tidak valid, dan ini menyebabkan bencana dalam prediksi dengan pembelajaran mesin. Kumpulan data yang digunakan memiliki beberapa sel yang hilang yang diganti menggunakan imputasi sederhana dengan nilai rata-rata atribut dan atribut Jenis Kelamin diubah menjadi nilai numerik sebagai '1' dan '0' untuk Pria dan Wanita masing-masing agar memungkinkan mesin untuk memproses karena mesin tidak akan memahami nilai string.

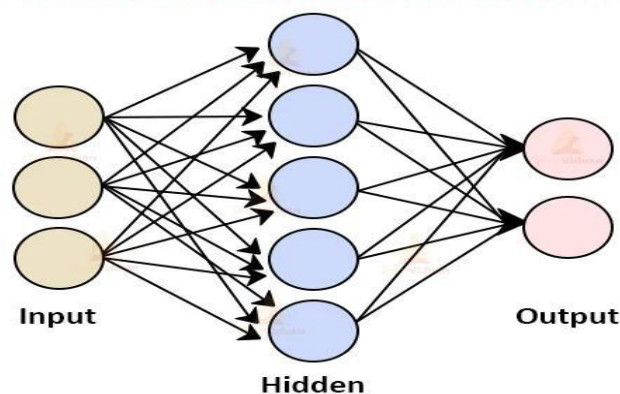
Variable Selection

Dataset penyakit ginjal memiliki banyak variabel berguna yang penting dan diperlukan untuk mengidentifikasi penyakit pada pasien. Kami menggunakan variabel berdasarkan tes dan metode yang digunakan oleh Rumah Sakit dalam menentukan terjadinya penyakit ginjal, yaitu Sepuluh (10) tes darah dan itulah yang kami gunakan sebagai variabel input. Nama juga termasuk dalam tes, tetapi kami memutuskan untuk menghapusnya karena tidak berdampak pada tes kami dan privasi pasien harus dijaga.

Neural Network Prediction Model

Neural network yang mensimulasikan struktur dan fungsi jaringan syaraf biologis yang terdiri dari input, bobot, dan fungsi aktivasi. Struktur neural network memiliki input, lapisan tersembunyi, dan output. Model direpresentasikan sebagai: Model ini menggunakan model yang memiliki sepuluh (10) input, dua (1) lapisan tersembunyi, dan satu (1) output dalam model yang diadopsi. Lapisan input adalah kumpulan neuron yang mengambil data input aktual untuk diproses. Jumlah neuron dapat diputuskan sesuai dengan data input. Untuk melatih model, tiga lapisan digunakan. Untuk input dan lapisan tersembunyi, sigmoid digunakan sebagai fungsi aktivasi. Lapisan output hanya memiliki satu hasil output, baik KD atau NKD dengan Sigmoid sebagai fungsi aktivasi. Gradien Stokastik digunakan sebagai pengoptimal model. Eksperimen dibangun dengan Bahasa Pemrograman Python. Gambar 2 menunjukkan struktur model neural network yang diadopsi untuk penelitian ini.

Architecture of Artificial Neural Network



Gambar 2. Model Prediksi neural network

Performance Evaluation

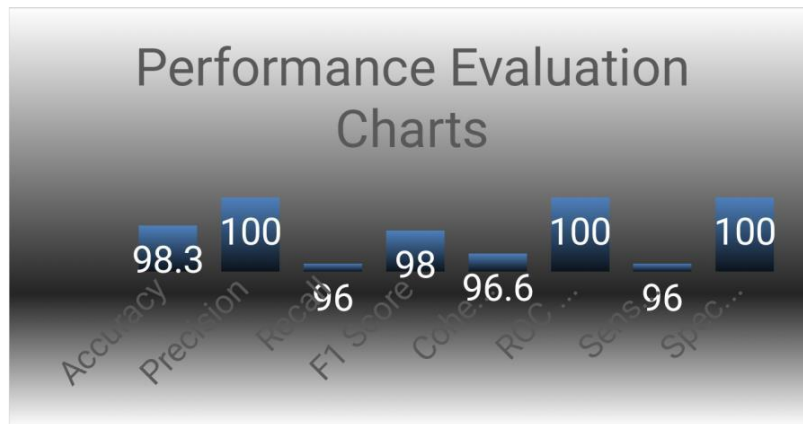
Dalam model ini, kinerja diukur berdasarkan Akurasi, spesifisitas, sensitivitas, statistik kappa, presisi, skor F1, Skor ROC, dan recall yang dijelaskan sebagai berikut. Matriks Confusion: matriks confusion menunjukkan kesesuaian statistik model dan kompatibilitasnya dengan dataset; matriks confusion juga dapat didefinisikan sebagai tata letak tabel yang secara khusus digunakan untuk visualisasi kinerja algoritma.

4. HASIL DAN PEMBAHASAN

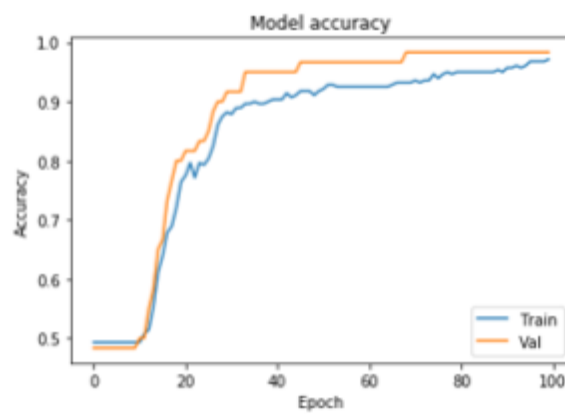
Proses identifikasi penyakit ginjal ini menggunakan algoritma neural network dengan arsitektur lapisan input 10 node, lapisan tersembunyi 4 node dan lapisan output 1 node. Lapisan input merupakan variabel-variabel dataset dan lapisan output merupakan label kelas yaitu 1 (rentan penyakit) dan 0 (tidak rentan penyakit). Bagian ini memberikan hasil algoritma Neural Network yang dihasilkan berdasarkan dataset.

Tabel 2. Matrik Confusion

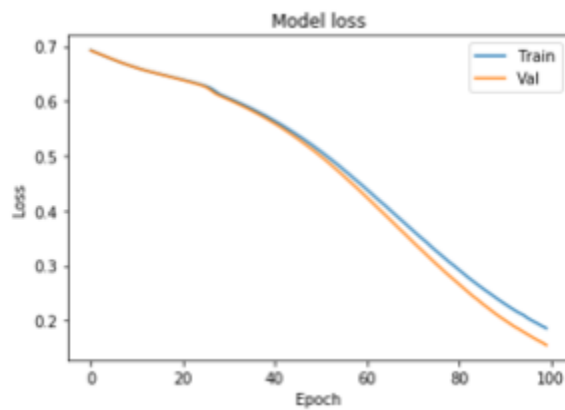
Classification		Prediction	
		Absence	Presence
Observation	Absence	29	0
	Presence	1	30



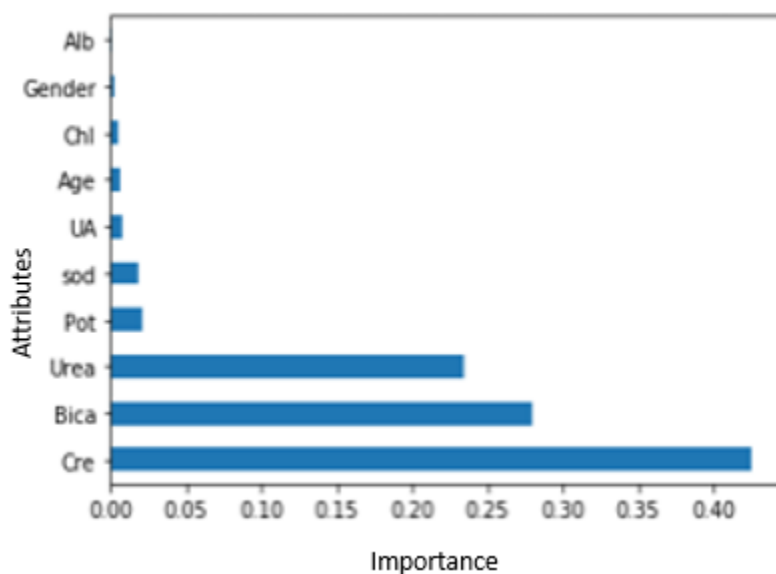
Gambar 3. Evaluasi Performa Model



Gambar 4. Akurasi Model



Gambar 5. Validation loss Model



Gambar 6. Attributes Importance

Penelitian ini membagi dataset kami menjadi 70% (280) dan 30% (60) untuk pengujian dan pelatihan, Tabel 3 menunjukkan matriks kebingungan yang memberikan ringkasan visualisasi model Neural Network pada dataset, model tersebut berhasil mengklasifikasikan 34 sampel dengan benar yang bukan CKD dan 25 sampel dengan benar yang merupakan Penyakit ginjal tetapi 1 sampel yang merupakan penyakit ginjal diklasifikasikan secara salah. Totalnya menjadi 59 sampel yang diklasifikasikan dengan benar yang menunjukkan bahwa model neural network bekerja dengan baik pada dataset kami karena hanya satu sampel dari sampel yang dipilih untuk pengujian yang tidak diklasifikasikan.

Gambar 3 menunjukkan evaluasi kinerja model dan menunjukkan bahwa model tersebut berhasil memprediksi rekam medis pasien ginjal pada rentan dan tidak rentan dengan akurasi prediksi keseluruhan sebesar 98% ketika model diuji dengan data yang tidak terlihat. Berdasarkan akurasi sebesar 98% ini, model menunjukkan bahwa model tersebut sangat cerdas untuk membuat prediksi secara mulus pada rentan dan tidak rentan. Metrik lain yang digunakan untuk mengevaluasi kinerja model Neural Network adalah Presisi: 100%, Ingat: 96%, Skor F1: 98%, Cohens Kappa: 96,6%, Skor ROC: 100%, Sensitivitas: 96%, dan Spesifisitas: 100%.

Gambar 4 menunjukkan tingkat keakuratan model kita saat diulang pada set pelatihan dan set validasi. Semakin banyak pengulangan, keakuratan pada kedua set meningkat dan dengan tingkat kemajuan yang hampir sama dengan set validasi, bahkan melampaui set pelatihan, yang menunjukkan keefektifan model kita dan tidak terjadi overfitting.

Gambar 5 menunjukkan hilangnya akurasi set validasi dan set pelatihan dari model kita saat diulang 100 kali dan berdasarkan hasil pada gambar. Ini menunjukkan tidak ada overfitting

karena ada peningkatan karena saat diulang akurasi kerugian menurun pada set pelatihan dan pengujian dengan tingkat penurunan yang hampir sama.

Gambar 6 menunjukkan peringkat fitur dalam kumpulan data yang memiliki pengaruh lebih besar dalam prediksi penyakit ginjal dari 10 atribut, Kreatinin dan Bikarbonat memiliki peringkat tertinggi diikuti oleh Urea, Kalium, Natrium, Asam Urea, Usia, Klorida, Jenis Kelamin dan Albumin. Kreatin dan Bikarbonat memiliki pengaruh banyak selama prediksi dengan model Neural Network

5. KESIMPULAN DAN SARAN

Tujuan utama penelitian ini adalah menggunakan model Neural network untuk prediksi penyakit ginjal dengan tingkat akurasi yang tinggi. Kami berhasil mengklasifikasikan dataset penyakit ginjal menjadi rentan penyakit ginjal dan tidak rentan penyakit ginjal dengan akurasi keseluruhan 98% ketika model diuji dengan serangkaian data yang tidak digunakan selama proses pelatihan. Model Neural network yang diadopsi terbukti efisien dan cocok untuk prediksi penyakit ginjal. Penelitian ini juga menyoroti pentingnya fitur yang digunakan dalam prediksi penyakit ginjal. Hal ini mengungkapkan bahwa dari 10 atribut, Kreatinin dan Bikarbonat adalah atribut dengan pengaruh tertinggi pada kerentanan penyakit ginjal.

DAFTAR REFERENSI

- Ahmed, R. M., & Alshebly, O. Q. (2019). Prediction and Factors Affecting of Chronic Kidney Disease Diagnosis using Artificial Neural Networks Model and Logistic Regression Model. *Iraqi Journal of Statistical Sciences*, 16(28), 140-159.
- Arafat, F., Fatema, K., & Islam, S. (2018). Classification of chronic kidney disease (ckd) using data mining techniques (Doctoral dissertation, Daffodil International University).
- Arbain, A. N., & Balakrishnan, B. Y. P. (2019). A Comparison of Data Mining Algorithms for Liver Disease Prediction on Imbalanced Data. *International Journal of Data Science and Advanced Analytics*, 1(1), 1-11.
- Ayon, S. I., & Islam, M. (2019). Diabetes Prediction: A Deep Learning Approach. *International Journal of Information Engineering & Electronic Business*, 11(2).
- Başar, M. D., & Akan, A. (2018). Chronic Kidney Disease Prediction with Reduced Individual Classifiers. *Journal of Electrical and Electronics Engineering*, 18(2), 249-255.
- Chahal, A., & Gulia, P. (2019). Machine learning and deep learning. *International Journal of Innovative Technology and Exploring Engineering*, 8(2), 2278-3075.

- Ge, Y., Wang, Q., Wang, L., Wu, H., Peng, C., Wang, J. & Yi, Y. (2019). Predicting post-stroke pneumonia using deep neural network approaches. *International Journal of Medical Informatics*, 132, 103986.
- Geri, G., Stengel, B., Jacquelinet, C., Aegerter, P., Massy, Z. A., & Vieillard-Baron, A. (2018). Prediction of chronic kidney disease after acute kidney injury in ICU patients: study protocol for the PREDICT multicenter prospective observational study. *Annals of Intensive Care*, 8(1), 77.
- Jahan, N., Rahman, M. M., & Islam, M. R. (2021). Application of Machine Learning Algorithms in Healthcare Sector for Disease Prediction. *Journal of Biomedical Informatics*, 118, 103799.
- Jain, D., & Singh, V. (2018). Feature selection and classification systems for chronic disease prediction: A review. *Egyptian Informatics Journal*, 19(3), 179-189.
- Joshi, T. N., & Chawan, P. M. (2018). Logistic Regression and SVM Based Diabetes Prediction System. *International Journal For Technological Research In Engineering*, 11(5), 2347-4718, July 2018
- Khurana, M., Sharma, N., & Singh, V. (2019). Prediction of Chronic Kidney Disease Using Machine Learning Algorithms. *Procedia Computer Science*, 167, 674–683.
- Kriplani, H., Patel, B., & Roy, S. (2019). Prediction of Chronic Kidney Diseases Using Deep Artificial Neural Network Technique. In *Computer Aided Intervention and Diagnostics in Clinical and Medical Images* (pp. 179-187). Springer, Cham. Chicago.
- Kumar, S. (2018). Chronic Kidney Disease Prediction Using Machine Learning. *International Journal of Computer Science and Information Security (IJCSIS)*, 16(4).
- Luyckx, V. A., Tonelli, M., & Stanifer, J. W. (2018). The global burden of kidney disease and the sustainable development goals. *Bulletin of the World Health Organization*, 96(6), 414.
- National Kidney Foundation (NKF). (2015). Global Facts: About Kidney Disease. In National Kidney Foundation.
- Saxena, S., & Sharma, S. (2020). Artificial Neural Network in Medical Diagnosis: A Review. *International Journal of Computer Applications*, 176(29), 1–5.
- Scholar, P. G. (2018). Chronic Kidney Disease Prediction Using Machine Learning. *International Journal of Computer Science and Information Security (IJCSIS)*, 16(4).
- Shafi, N., Bukhari, F., Iqbal, W., Almustafa, K. M., Asif, M., & Nawaz, Z. (2020). Cleft prediction before birth using deep neural network. *Health Informatics Journal*, 1(18) 1460458220911789.
- Sharma, S., & Parmar, M. (2020). Heart Diseases Prediction using Deep Learning Neural Network Model. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 9(3).

WHO. (2006). World Health Organization: Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycemia. Geneva, World Health Org. WHO2.

World Kidney Day. (2017). Chronic Kidney Disease - World Kidney Day. ISN – Global Operations Center