

PERBANDINGAN METODE DATA MINING SVM DAN NN UNTUK KLASIFIKASI PENYAKIT GINJAL KRONIS

Hilda Amalia

Program Studi Manajemen Informatika

AMIK BSI Jakarta

www.bsi.ac.id

hilda.ham@bsi.ac.id

Abstract—Data mining is a method that has been widely used to make scientific discoveries from a collection of datasets that have been only stored without further management. In the world of health the use of data mining methods has greatly helped the health world in making predictions about health problems encountered. One of the most deadly diseases is chronic kidney disease. Chronic kidney disease can cause many other deadly diseases. The rate of development of chronic kidney disease is also increasing from year to year. In research data of chronic kidney disease will be processed by data mining method that is Support Vector Machine and Neural network. Both are data mining methods that are known to perform well for data with many attributes and parameters. The result of neural network method yields 93.37% accuracy value and SVM with value 95.16%

Keyword: data mining, neural network, chronic kidney disease

Intisari—Data mining merupakan suatu metode yang telah banyak digunakan untuk melakukan penemuan ilmu pengetahuan dari kumpulan dataset yang selama ini hanya disimpan tanpa dikelola lebih lanjut. Dalam dunia kesehatan penggunaan metode data mining telah banyak membantu dunia kesehatan dalam membuat prediksi mengenai masalah kesehatan yang dihadapi. Salah satu penyakit yang sangat mematikan yaitu penyakit ginjal kronik. Penyakit ginjal kronik dapat menyebabkan banyak penyakit mematikan lainnya. Tingkat perkembangan penyakit ginjal kronik ini juga terus meningkat dari tahun ke tahunnya. Dalam penelitian data penyakit ginjal kronis akan diolah dengan metode data mining yaitu Support Vector Machine dan Neural network. Keduanya merupakan metode data mining yang diketahui memiliki kinerja yang baik untuk data dengan atribut dan parameter yang banyak dan beragam. Dari Hasil penelitian diperoleh hasilnya yaitu metode neural network menghasilkan nilai akurasi 93.37% dan SVM dengan nilai 95.16%.

Kata Kunci: penyakit ginjal kronis, neural network, genetika algoritma

PENDAHULUAN

Ginjal merupakan organ tubuh yang berfungsi untuk mempertahankan stabilitas tubuh dengan cara mengatur cairan tubuh, keseimbangan elektrolit, pengeluaran hasil metabolisme (Lucida, Trisnawati, & Suardi, 2011). Karena itu penting bagi manusia untuk memperhatikan kesehatan ginjalnya. salah satu penyakit ginjal yang menakutkan adalah gagal ginjal kronik (Satoso, A.E, & Asbullah, 2016). Saat ginjal tak mampu lagi menjalankan fungsinya maka akan muncul masalah kesehatan yang berkaitan dengan penyakit ginjal kronis (). Gagal ginjal kronik terjadi apabila laju filtrasi glomeruler (LFG) kurang dari 60 ml/menit/ 1,73 m² selama tiga bulan atau lebih (Hidayati, Kushadiwijaya, & Suhardi, 2008)

Masalah berikutnya yang dihadapi penderita adalah mahalnya biaya perawatan terhadap penderita Penyakit ginjal kronik dan dengan hasil yang tidak baik serta belum ditemukannya obat untuk penyakit ginjal kronik (Satoso et al., 2016). Di Indonesia diperkirakan terdapat 70.000 penderita penyakit gagal ginjal kronik akan terus bertambah setiap tahunnya, menurut survey WHO di Indonesia akan terdapat peningkatan jumlah penderita penyakit ginjal kronik sebesar 46% dari tahun 1955-2025 (Melti, Arthur, & Manoppo, 2014). Penyakit ginjal kronik juga dapat menyebabkan penyakit mematikan lainnya seperti diabetes, tekanan darah tinggi, sakit jantung dan lupus (Kumar, 2016)

Untuk itu penting bagi setiap orang untuk memperhatikan kesehatan ginjal. Perlu adanya suatu sistem pendukung keputusan yang mampu melakukan penunjang pengambilan keputusan penyakit ginjal kronik. Data mining merupakan metode yang telah dikenal sebagai metode yang mampu menggali pola atau ilmu pengetahuan dari tumpukan data yang besar (Amalia, 2017).

Proses penambangan data ini nantinya akan menghasilkan pengetahuan yang berharga kemudian dapat diimplementasikan kedalam aplikasi sistem penunjang keputusan. Dalam dunia kesehatan data mining telah digunakan untuk

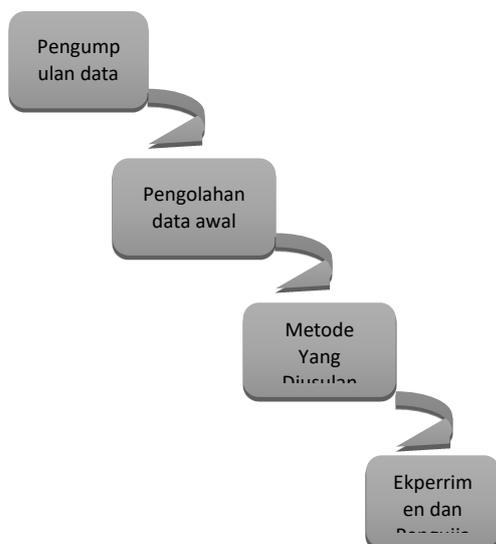
menghasilkan prediksi-prediksi mengenai diagnosa suatu penyakit, Penelitian sebelumnya menggunakan dataset penyakit ginjal dengan metode data mining telah dilakukan oleh beberapa peneliti yaitu menggunakan Prediksi penyakit ginjal kronis dengan menggunakan KNN dan SVM (Sinha & Poonam, 2015) Radom Forest Machine Learning Algorithm (Kumar, 2016), menggunakan Back-Proparation Neural Network, Radial Basis Function dan Random Forest (Ramya & Radha, 2016)

Dari studi literature diketahui bahwa metode neural network memiliki nilai akurasi yang tertinggi untuk pengolahan dataset penyakit ginjal kronik sehingga dalam penelitian ini akan dilakukan peningkatan nilai akurasi metode neural network. Metode data mining lainnya yang mampu menghasilkan nilai akurasi yang baik untuk data dan atribut atau parameter yang banyak dan beragam yaitu Support Vector Machine. VM berada dalam satu kelas dengan Artificial Neural Network (ANN) dalam hal fungsi dan kondisi permasalahan yang bisa diselesaikan. Keduanya masuk dalam kelas supervised learning (Amalia, Lestari, & Puspita, 2017)

Untuk itu dalam penelitian ini akan dilakukan pengolahan dataset penyakit ginjal kronis dengan menggunakan metode neural network kemudian dibandingkan dengan metode data mining lainnya yang memiliki kemampuan sama baik dengan nueral network yaitu Support Vector Machine.

BAHAN DAN METODE

Dalam penelitian ini dilakukan beberapa tahapan penelitian, dibawah ini disajikan tahapan penelitian yang digunakan:



Sumber: Amalia (2018)

Gambar 1 Tahapan Penelitian yang Digunakan

Pengumpulan data

Data yang digunakan dalam penelitian ini merupakan data sekondr dan diperoleh dari situs penyedia data untuk penelitian yaitu uci repository. Data yang digunakan sebanyak 400 record dan terdiri 24 atribut dan satu label. Dataset yang digunakan adalah data sekunder yaitu data yang diperoleh bukan dari pengumpulan data secara langsung oleh peneliti. Dataset diperoleh dari situs penyedia data yaitu uci repository dengan laman [http://archive.ics.uci.edu/ml/datasets/Chronic Kidney Disease](http://archive.ics.uci.edu/ml/datasets/Chronic_Kidney_Disease).

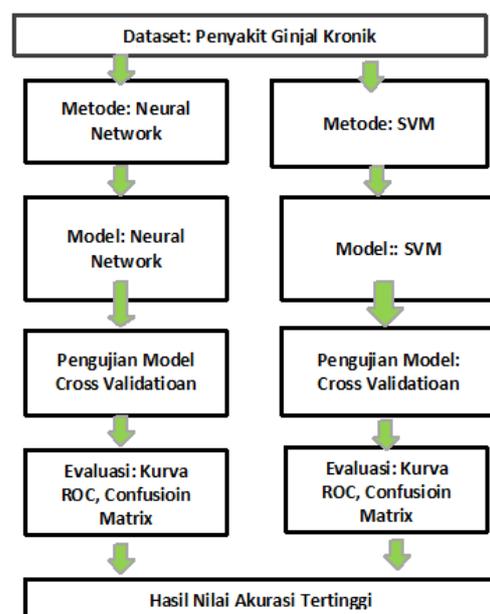
Pengolahan data awal

Dalam tahapan pengolahan data awal, dilakukan tiga jenis yaitu: (1) Validasi data,(2) Intregrasi data dan (3)Pengubahan data, ketiga tahapan diatas dilakukan untuk mempeoleh dtaset yang mewakili dan diharapkan dapat digunakan untuk mendapatkan informasi yang berharga dari dataset yang ada (Amalia & Evicienna, 2017) Dalam penelitian ini pengolahan data awal dilakukan dengan menggunakan tiga teknik pengolahan data diatas. Dari 400 record yang ada diolah menjadi 204 record.

HASIL DAN PEMBAHASAN

Metode yang diusulkan

Pada tahapan ini dibuatkan suatu metode yang digunakan dalam penelitian dataset penyakit ginjal kronik. berikut metode yang digunakan:

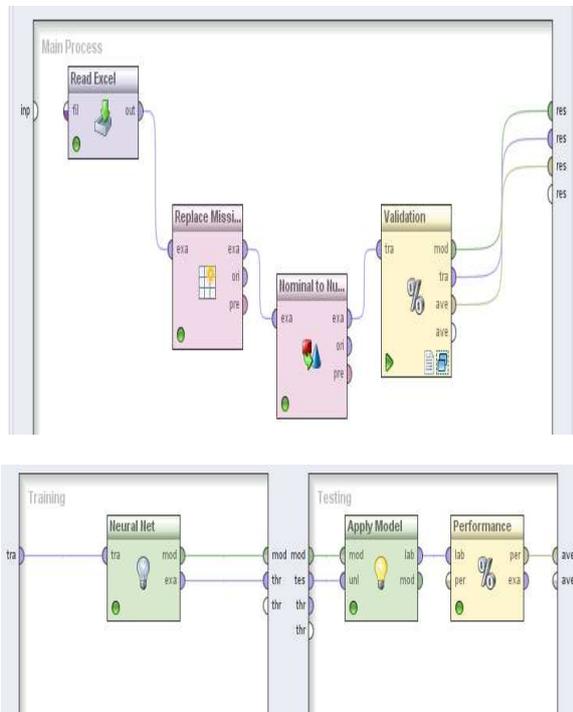


Sumber: Amalia(2018)

Gambar 2 Metode yang Diusulkan

Eksperimen dan Pengujian Hasil

Dalam penelitian ini pengolahan dataset penyakit ginjal kronis dengan menggunakan tools RapidMiner. Setelah menetapkan metode yang diusulkan untuk digunakan dalam penelitian ini maka dilakukan eksperimen dan pengujian hasil dengan menggunakan tool mRapidMiner. Dataset penyakit ginjal kronis diletakan pada modul read Excel, dihubungkan dengan modul missing value untuk dilakukan pengecekan terhadap dataset, dan dihubungkan dengan modul binominal to numerical untuk mengubah isi data yang berubah binominal ke dalam nilai numeric, kemudian dihubungkan dengan modul validation yang berfungsi untuk melakukan pengujian hasil eksperimen yang telah dilakukan. Tampilan pengolahan data seperti gambar 3.



Sumber: Amalia(2018)
Gambar 3 Pengolahan Data dengan Neural Network

Berikut tabel 1 Confusion Matrix yang dihasilkan:

Tabel 1 Confusion Matrix Neural Network
Accuracy: 93,36%

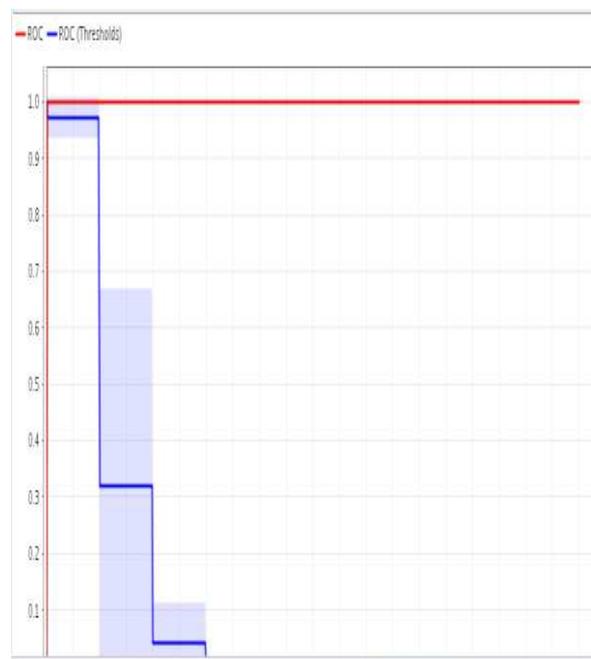
| | True ckd | True notckd |
|------------------|----------|-------------|
| Pred. true ckd | 82 | 0 |
| Pred true notckd | 6 | 138 |
| Class recall | 93,18% | 100.00% |

Sumber: Amalia(2018)

Perfarmace vector yang dihasilkan:

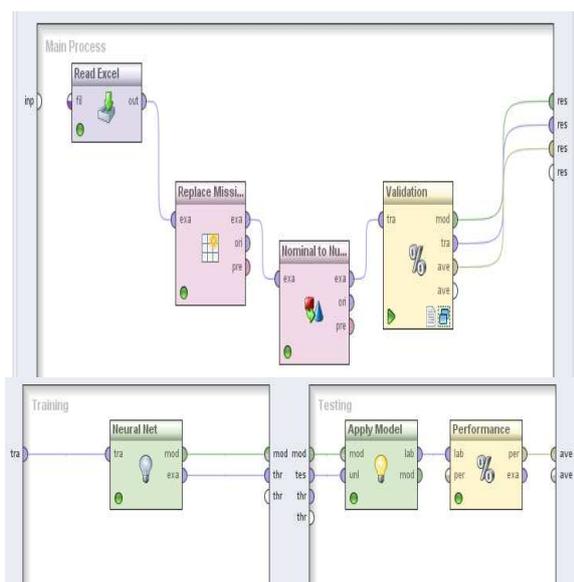
PerformanceVector:
accuracy: 98.73% +/- 1.94% (mikro: 98.74%)
ConfusionMatrix:
True: ckd notckd
ckd: 97 0
notckd: 3 139
precision: 97.95% +/- 3.13% (mikro: 97.89%)
(positive class: notckd)
ConfusionMatrix:
True: ckd notckd
ckd: 97 0
notckd: 3 139
recall: 100.00% +/- 0.00% (mikro: 100.00%)
(positive class: notckd)
ConfusionMatrix:
True: ckd notckd
ckd: 97 0
notckd: 3 139
AUC (optimistic): 1.000 +/- 0.000 (mikro: 1.000)
(positive class: notckd)
AUC: 1.000 +/- 0.000 (mikro: 1.000) (positive class: notckd)
AUC (pessimistic): 1.000 +/- 0.000 (mikro: 1.000)
(positive class: notckd)

Kurva ROC yang dihasilkan oleh metode neural network untuk pengolahan dataset penyakit ginjal kronis, seperti dibawah ini:



Sumber: Amalia(2018)
Gambar 4 Kurva ROC Metode Neural Network

Tahapan selanjutnya adalah melakukan pengolahan data dengan menggunakan metode SVM, berikut tampilannya:



Sumber: Amalia (2018)

Gambar 5 Pengolahan Data dengan Suport Vector Machine (SVM)

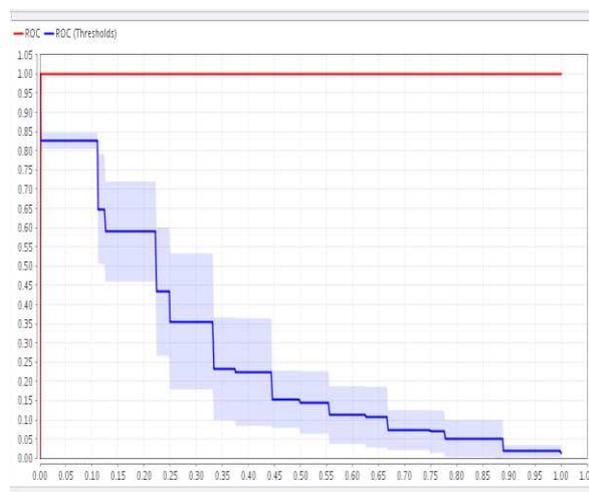
Confusion Matrix yang dihasilkan dari metode Support Vector Machine:

Tabel 2 Confusion Matrix Support Vector Machine Accuracy: 95,16%

| | True ckd | True notckd |
|------------------|----------|-------------|
| Pred. true ckd | 77 | 0 |
| Pred true notckd | 11 | 138 |
| Class recall | 87.50% | 100.00% |

Sumber: Amalia (2018)

Dari hasil pengolahan dataset dengan menggunakan Suport Vector Machine diperoleh kurva ROC, seperti ditampilkan gambar 6



Sumber: Amalia(2018)

Gambar 6 Korva ROC pengolahan SVM

KESIMPULAN

Penelitian ini melakukan pengolahan dataset penyakit ginjal kronis yang diperoleh dari uci repository. Peneliti melakukan pengolahan data untuk memperoleh metode mana yang memiliki nilai akurasi yang lebih tinggi bagi dataset penyakit ginjal kronis. Diketahui dari hasil penelitian yaitu dengan metode Neural Network diperoleh nilai akurasi 93,36% dan metode Support Vector Machine (SVM) diperoleh nilai akurasi 95.16%. Hasil yang diperoleh termasuk dalam jenis klasifikasi sangat baik. Sehingga dapat disimpulkan bahwa SVM dan NN memiliki performa kinerja yang baik untuk pengolahan dataset penyakit ginjal kronis. Dan dari hasil penelitian diketahui untuk dataset ginjal kronis bahwa metode SVM menghasilkan nilai akurasi yang lebih tinggi dari metode Neural Network.

REFERENSI

- Amalia, H., & Evicienna, E. (2017). Komparasi metode Data Mining Untuk Penentuan Proses Persalinan Ibu Melahirkan. *Jurnal Sistem Informasi (Journal of Information Systems)*, 13, 103–109. <https://doi.org/http://dx.doi.org/10.21609/jsi.v13i2.545>
- Amalia, H., Lestari, A. F., & Puspita, A. (2017). PENERAPAN METODE SVM BERBASIS PSO UNTUK PENENTUAN KEBANGKRUTAN PERUSAHAAN. *Jurnal Techno Nusa Mandiri*, 14(2), 131–136.
- Hidayati, T., Kushadiwijaya, H., & Suhardi, S. (2008). HUBUNGAN ANTARA HIPERTENSI, MEROKOK DAN MINUMAN SUPLEMEN ENERGI DAN KEJADIAN PENYAKIT GINJAL KRONIK. *Berita Kedokteran Masyarakat*, 24(2), 90–102.
- Kumar, M. (2016). Prediction of Chronic Kidney Disease Using Random Forest Machine Learning Algorithm. *International Journal of Computer Science and Mobile Computing*, 5(2), 24–33.
- Lucida, H., Trisnawati, R., & Suardi, M. (2011). Analisis Aspek Farmakokinetika Klinik Pasien Gagal Ginjal Pada IRNA Penyakit Dalam RSUP DR. M. DJAMIL PADANG. *Jurnal Sains Dan Teknologi Farmasi*, 16(2), 144–155.

- Melti, T., Arthur, M., & Manoppo, F. (2014). HUBUNGAN ANTARA DERAJAT PENYAKIT GINJAL KRONIK DENGAN NILAI AGREGASI TROMBOSIT. *Jurnal E-Biomedik (eBM)*, 2, 509–513.
- Ramya, S., & Radha, N. (2016). A novel method to segment blood vessels and optic disc in the fundus retinal images. *Diagnosis of Chronic Kidney Disease Using Machine Learning Algorithms*, 4(1), 812–820. <https://doi.org/10.15680/IJIRCCE.2016>.
- Satoso, R. B., A.E, M. Y., & Asbullah. (2016). HUBUNGAN LAMA HEMODIALISIS DENGAN PENURUNAN NAFSU MAKAN PADA PASIEN GAGAL GINJAL KRONIK DI UNIT HEMODIALISA RSUD ULIN BANJARMASIN. *Dinamika Kesehatan*, 7(1), 139–151.
- Sinha, P., & Poonam, S. (2015). Comparative Study of Chronic Kidney Disease Prediction using KNN and SVM, 4(12), 608–612.

