

DIAGNOSTIK PENYAKIT GINJAL KRONIS MENGGUNAKAN MODEL KLASIFIKASI SUPPORT VECTOR MACHINE

Taryadi¹⁾, Era Yuniyanto²⁾, Kasmari³⁾

STMIK Widya Pratama, Pekalongan ^{1 2)}

Universitas STIKUBANK, Semarang ³⁾

tari_ball@stmik-wp.ac.id ¹⁾, era.yuniyanto@gmail.com ²⁾, fkasmari@edu.unisbank.ac.id ³⁾

Abstrak

Penyakit ginjal atau biasa dikenal dengan gagal ginjal merupakan suatu kondisi menurunnya fungsi ginjal yang dapat mengakibatkan ketidakmampuan ginjal dalam menjalankan tugasnya. Penderita penyakit ginjal berpotensi masuk ke fase kronis. Penyakit ginjal kronik merupakan penurunan fungsi ginjal secara bertahap selama tiga bulan yang mengakibatkan terhentinya fungsi ginjal secara total. Tujuan dari pengembangan ini adalah suatu sistem pendukung keputusan bagi dokter dalam mendiagnosis pasien penyakit ginjal. Sistem menampilkan hasil prediksi apakah pasien penyakit ginjal sudah memasuki fase penyakit ginjal kronis atau belum. Metodologi penelitian ini terdiri dari dua tahap utama: pemodelan klasifikasi dan pengembangan sistem. Pemodelan klasifikasi terdiri dari pengumpulan data, persiapan data, pengelompokan data, klasifikasi, ekstraksi aturan. Pengembangan sistem didasarkan pada aturan yang diekstraksi sebelumnya. Penelitian ini menghasilkan suatu sistem yang dapat mendeteksi suatu kondisi penyakit ginjal kronis berdasarkan beberapa faktor dengan akurasi sebesar 96,34%.

Kata kunci: Penyakit Ginjal; diagnosa; decision support system; support vector machine; klasifikasi;

1. Pendahuluan

Gagal ginjal merupakan suatu keadaan dimana menurunnya fungsi ginjal secara berkala yang dapat mengakibatkan ketidakmampuan ginjal dalam menjalankan tugasnya. Pada kondisi kronis, hal ini akan menyebabkan berkurangnya fungsi ginjal dalam jangka waktu tertentu. Penyakit ginjal kronis dapat berkembang selama bertahun-tahun dan menyebabkan penyakit ginjal stadium akhir. Penyakit ini meningkat pesat dan menjadi ancaman dunia. Pada tahun 2015, 10% populasi di seluruh dunia terkena penyakit ginjal kronis, dan jutaan orang meninggal setiap tahunnya karena pasien tidak memiliki akses terhadap pengobatan yang terjangkau (World Kidney Day, 2015; World Health Organization, 2015). Selanjutnya di negara berkembang seperti Indonesia, pada tahun 2007-2012 tercatat jumlah penderita gagal ginjal baru dan aktif meningkat (Indonesian Renal Registry, 2012).

Penyakit ginjal memerlukan penanganan medis khusus berdasarkan kondisi kronis pasien pada stadium 1 hingga stadium 5. Prosedur yang dilakukan akan berbeda-beda berdasarkan penyebabnya. Perawatan biasanya terdiri dari

tindakan untuk membantu mengendalikan tanda dan gejala, mengurangi komplikasi, dan memperlambat perkembangan penyakit. Dalam mendiagnosisnya, dokter melakukan beberapa pemeriksaan terhadap pasien di laboratorium seperti darah, urin, dan lain-lain. Dari pemeriksaan tersebut, dokter akan menentukan kondisi dan pengobatan pasien.

Untuk meningkatkan penilaian dokter terhadap kondisi pasien, diperlukan sistem pendukung diagnostik. Sistem ini akan membantu dokter dalam menentukan apakah kondisinya kronis atau tidak. Pendekatan sistemnya akan menggunakan teknik pembelajaran mesin pada klasifikasi (Han, J & Kamber, 2006). Terdapat beberapa penelitian sebelumnya dalam bidang ini, seperti Hybrid Modified Cuckoo Search – Neural Network (Chatterjee, Dzitac; Sen, Rohatinovici, Dey, Ashour, Balas, 2017), Multivariate K (Abhinandan, 2015), Naïve Bayes (Dulhare; Ayesha, 2016), dan Decision Tree (Tazin, Sabab; Chowdhury, 2016).

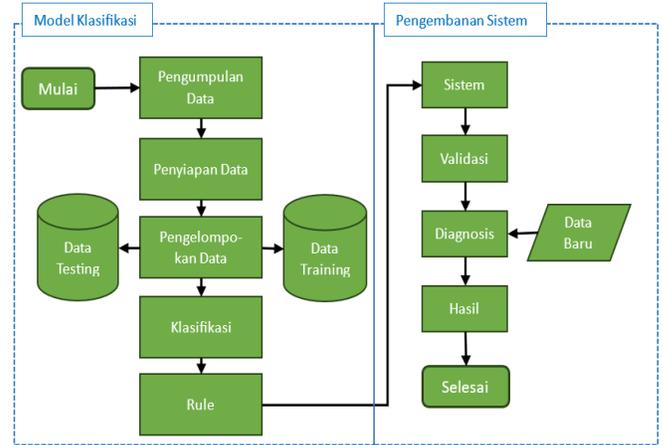
Penelitian ini berkontribusi sebagai alat pendukung keputusan sebagai alat deteksi dini.

Alat ini sangat berguna dalam situasi darurat dimana dokter memerlukan informasi lebih lanjut terkait pasien untuk memutuskan pengobatan selanjutnya. Penelitian ini penting untuk dipersiapkan secara matang karena dalam situasi darurat ada beberapa situasi yang harus dihadapi, yaitu kerusakan infrastruktur rumah sakit, kurangnya informasi (misalnya kehilangan rekam medis) (Tundjungsiari, Yugaswara, 2015). Beberapa penelitian telah meneliti cara merawat pasien penyakit ginjal dalam beberapa situasi darurat (Kopp, Jeffrey, et.al , 2007; Mark, et.al., 2011).

Penelitian ini mengusulkan sistem pendukung keputusan diagnosis penyakit ginjal kronis menggunakan *Support Vector Machine*. *Support Vector Machine* (SVM) adalah algoritma yang digunakan untuk klasifikasi dan clustering (Priyogi, Selviandro, Hasibuan, Ahmad, 2014). Ide dasar dari algoritma *Support Vector Machine* adalah mencari batasan optimal yang digunakan untuk memisahkan dua buah kelas yang disebut hyperplane (Williams, 2008). Para peneliti telah menerapkan SVM untuk melakukan tugas klasifikasi pada berbagai bidang, seperti komputasi afektif (Diana dan Sabiq, 2016). Tujuan dari pengembangan sistem ini adalah sebagai sistem pendukung keputusan dalam mendiagnosis penyakit ginjal dan memprediksi apakah pasien penyakit ginjal sudah memasuki fase penyakit ginjal kronis atau belum.

2. Metode Penelitian

Ada dua fase utama: pemodelan klasifikasi dan pengembangan sistem. Fase pemodelan klasifikasi berfokus pada mendapatkan aturan dan model yang tepat untuk klasifikasi. Selanjutnya aturan-aturan ini akan digunakan pada tahap berikutnya: pengembangan sistem. Metodologi penelitian ini ditunjukkan pada Gambar 1 di bawah ini:



Gambar 1. Metodologi Penelitian

2.1. Pengumpulan Data

Dataset dikumpulkan dari UCI Machine Learning berjudul "Penyakit Ginjal Kronis" dari Dr. P. Soundarapandian. MD, DM. Dataset ini memiliki 25 atribut dan 400 instance.

2.2. Penyiapan Data

Persiapan data merupakan langkah penting untuk menjamin keabsahan data. Peneliti biasanya menerapkan algoritma penyaringan untuk menghilangkan data yang berada di bawah nilai ambang batas, seperti algoritma filter finite impuls respon (FIR) [16]. Pada tahap ini terdapat dua sub tahapan yaitu pemilihan atribut dan pembersihan data. Dalam pemilihan atribut menggunakan metode statistik dan wawancara untuk menentukan atribut apa yang signifikan dalam menentukan kelas. Lima atribut yang digunakan yaitu: tekanan darah, kreatinin serum, volume sel padat, faktor hipertensi, dan faktor anemia. Langkah selanjutnya adalah pembersihan data. Proses ini dengan cara menghilangkan nilai yang hilang dalam kumpulan data. Dari tahap ini, dataset menjadi bersih dan siap digunakan pada tahap berikutnya.

Tabel 1. Atribut yang digunakan

No	Atribut	Ukuran	Tipe
1	Kreatin Serum	mgs/dl	Numerik
2	Volume padat sel	%	Numerik
3	Tekanan Darah	mm/hg	Numerik
4	Faktor Hipertensi	angka	Numerik
5	Faktor Anemia	angka	Numerik

2.3. Pengelompokan Data

Pada tahap ini, kumpulan dibagi menjadi dua kelompok: pelatihan (*data training*) dan pengujian (*data test*). Proporsinya adalah 70% untuk pelatihan dan 30% untuk pengujian.

2.4. Klasifikasi

Untuk memanggil metode SVM di Python menggunakan *library* sklearn. Paket harus diinstal terlebih dahulu di jendela konsol. Sebelum memanggil metode SVM, data pada variabel dataset sebelumnya harus diubah menjadi data frame yang dapat dibaca oleh metode SVM. Seperti sebelumnya, data yang telah diubah menjadi data frame harus dimasukkan ke dalam variabel untuk memudahkan pengolahan data lanjutan. Berikut adalah metode pemanggilan bingkai data di python.

```
#inisialisasi variabel dataset
dataset_name = 'chronic_kidney_disease.csv'

#import dan load dataset ke dataframe
chronic_kidney_disease_dataframe = pd.read_csv(dataset_name)
```

Jika data sudah menjadi data frame, maka dapat dibaca oleh SVM. Meneruskan data olahan di atas, selanjutnya dapat memanggil metode SVM di python:

```
#inisialisasi Model Support Vector
support_vector_machine_model = SVC(random_state = 0)

#definisi parameter svc
parameters_grid = {'kernel': ['poly', 'rbf', 'linear', 'sigmoid'],
                  'C': [0.1, 1, 10, 100, 1000],
                  'gamma': ['scale', 'auto'],
                  'shrinking': [True, False]}
```

2.5. Aturan Ekstraksi

Untuk mengekstrak aturan [17] dari dengan menggunakan metode di bawah ini.

```
svm_grid_search = GridSearchCV(support_vector_machine_model,
                              parameters_grid, scoring = 'accuracy')

#fit data ke grid
svm_grid_search.fit(train_features, train_target)
#svm prediksi dalam test features
svm_prediction = svm_grid_search.predict(test_features)
```

Dari prosedur pemanggilan tersebut, akan mendapatkan aturan berikut dan akan diterapkan dalam pengembangan sistem:

```
IF (Tekanan Darah <= 50%) OR (Tekanan Darah > 50%) AND (Kreatin Serum > 1.2) OR ((Kreatin Serum <= 1.2) AND ((Volume Padat Sel <= 30%) OR (Volume Padat Sel > 30%)) AND (Hipertensi = 1)) THEN +Penyakit Ginjal Kronis
```

3. Hasil dan Pembahasan

3.1 Pengembangan Sistem

Sistem dikembangkan menggunakan bahasa Python dengan beberapa paket *library* yang telah disediakan untuk pengolahan data, yaitu paket sklearn, pandas, matplotlib, dan paket seaborn. Paket tersebut memiliki fungsi berbeda dan akan dipanggil di berbagai tempat.

Fitur utama dari sistem ini adalah klasifikasi, prediksi dari data baru dan fitur download untuk menyimpan hasil dalam format CSV.

- Fitur Baca File

Fitur ini dimana pengguna dapat mengunggah file CSV dengan format tertentu.

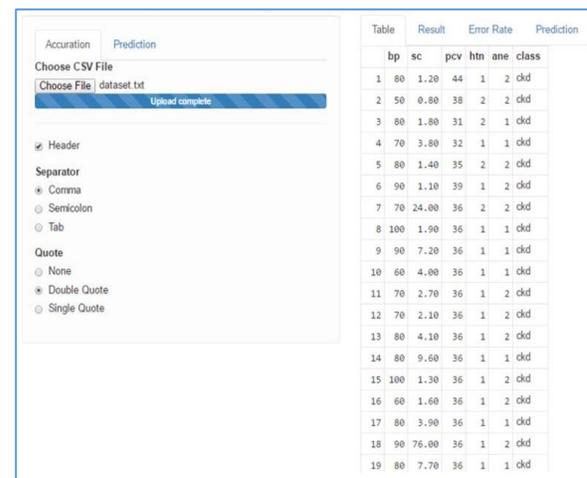
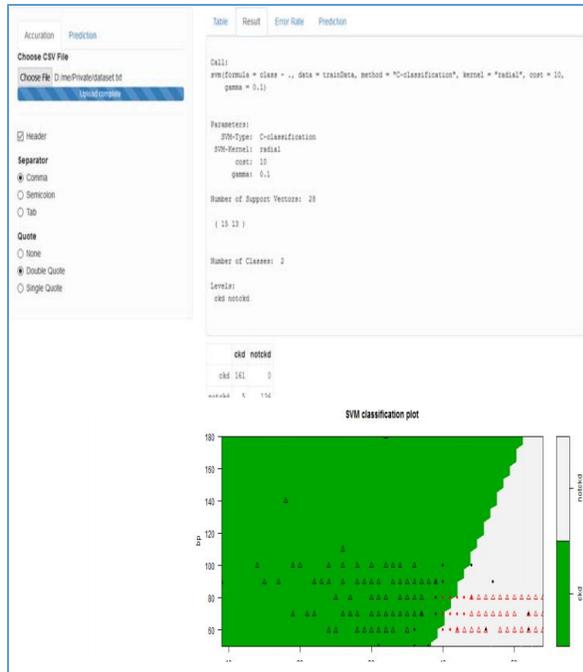


Table	Result	Error Rate	Prediction			
bp	sc	pcv	htn	ane	class	
1	80	1.20	44	1	2	ckd
2	50	0.80	38	2	2	ckd
3	80	1.80	31	2	1	ckd
4	70	3.80	32	1	1	ckd
5	80	1.40	35	2	2	ckd
6	90	1.10	39	1	2	ckd
7	70	24.00	36	2	2	ckd
8	100	1.90	36	1	1	ckd
9	90	7.20	36	1	1	ckd
10	60	4.00	36	1	1	ckd
11	70	2.70	36	1	2	ckd
12	70	2.10	36	1	2	ckd
13	80	4.10	36	1	2	ckd
14	80	9.60	36	1	1	ckd
15	100	1.30	36	1	2	ckd
16	60	1.60	36	1	2	ckd
17	80	3.90	36	1	1	ckd
18	90	76.00	36	1	2	ckd
19	80	7.70	36	1	1	ckd

Gambar 2. Fitur Baca File/Data

- Fitur Klasifikasi

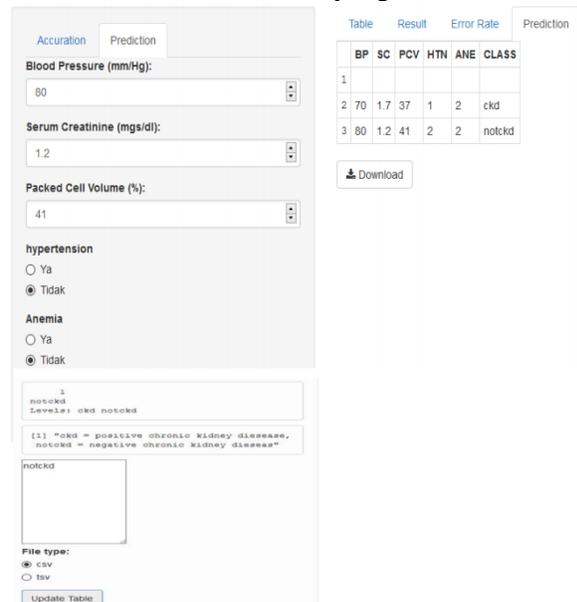
Fitur ini akan menampilkan hasil klasifikasi dari data yang diunggah.



Gambar 3. Fitur Klasifikasi

• **Fitur Diagnosis**

Fitur ini akan mendiagnosis kondisi pasien (kronis atau tidak) dari data yang diisi.



Gambar 4. Fitur Diagnosis

Parameter yang dioptimasi pada metode SVM adalah :

1. Kernel – Nilai kernel menggunakan poly, rbf, linear dan sigmoid. Kernel Values used included poly, rbf, linear dan sigmoid.
2. C – Nilai C yang digunakan 0.1, 1, 10, 100, 1000
3. Gamma – Nilai Gamma yang digunakan yaitu skala dan auto.
4. Shrinking – Nilai Shrinking yang digunakan yaitu True dan False.

Akurasi menggunakan evaluasi matrik matrik karena *class* data target sangat seimbang (rasio dari notckd dan ckd adalah 5:3) dan akurasi merukan masalah umum dalam model klasifikasi.

3.2 Hasil dan Analisis

Untuk menghitung tingkat kesalahan sistem, menggunakan model *Precision*, *Accuracy*, *Recall* dan *F1-Score* menggunakan *confusion matrix*. Dari pemanggilan paket tersebut, menunjukkan tingkat kesalahan sistem ini adalah 4 %.

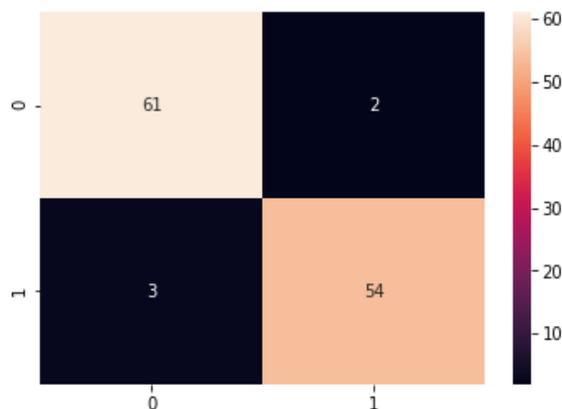
Precision: 0.9642857142857143
 Accuracy: 0.9583333333333334
 Recall: 0.9473684210526315
 F1-score: 0.9557522123893805

Classification Report:

	precision	recall	f1-score	support	
	0	0.95	0.97	0.96	63
	1	0.96	0.95	0.96	57
accuracy				0.96	120
macro avg	0.96	0.96	0.96	0.96	120
weighted avg	0.96	0.96	0.96	0.96	120

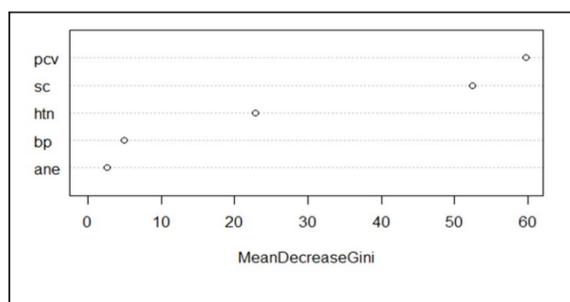
Gambar 5. Hasil pengujian akurasi sistem

Dari informasi diatas dapat disimpulkan bahwa keakuratan sistem adalah 96,42%. Selanjutnya visualisasi tingkat kesalahan dapat dilihat pada Gambar 6 di bawah ini.



Gambar 6. Visualisasi Error Rate

Selain tingkat akurasi dan *error* perlu dilihat pentingnya setiap atribut dari *Mean Decrease Gini*. Semakin besar nilainya maka semakin besar peranan atribut tersebut. Begitu pula sebaliknya, semakin kecil nilainya maka semakin kecil peranan atribut-atribut tersebut dalam proses klasifikasi. Dari Gambar 6 di bawah ini menunjukkan bahwa *Packed Cell Volume* merupakan atribut yang paling penting.



Gambar 7. Mean Decrease Gini

4. Kesimpulan dan Saran

Diperlukan adanya sistem pendukung keputusan diagnostik untuk meningkatkan penilaian dokter dalam menentukan kondisi kronis pasien penyakit ginjal. Penelitian ini mengusulkan sebuah sistem yang akan membantu dokter merawat pasien penyakit ginjal. Sistem ini dikembangkan berdasarkan teknik pembelajaran mesin yaitu *Support Vector Machine* (SVM). Metodologi penelitian ini dibagi menjadi dua tahap utama: pemodelan klasifikasi dan pengembangan sistem. Dari pemodelan klasifikasi, menggunakan aturan dan model dalam klasifikasi penyakit ginjal. Selanjutnya

aturan-aturan tersebut diimplementasikan ke dalam suatu sistem dengan menggunakan bahasa pemrograman Python. Sistem pendukung keputusan diagnosis ini memiliki beberapa fitur utama seperti baca file, klasifikasi, tingkat kesalahan dan diagnosis. Akurasi sistem ini adalah 96,42%. Sistem ini diharapkan dapat mendukung dokter dalam menentukan kondisi kronis pasien penyakit ginjal dengan akurasi yang baik.

DAFTAR PUSTAKA

World Kidney Day, 2015, "Chronic Kidney Disease".

World Health Organization (WHO), 2015, "Country Statistics and Global Health Estimates. 2015".

Indonesian Renal Registry, 2012, "5 th Report Of Indonesian Renal Registry", Program Indonesian Ren Regist, 2012:12-13.

Han, J & Kamber, M 2006, "Data Mining Concepts and Techniques", Second Edition, Printed in the United States of America. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

Chatterjee, S., Dzitac S.; Sen S., Rohatinovici N.C., Dey N., Ashour A.S, Balas V.E., 2017 "Hybrid modified Cuckoo Search-Neural Network in chronic kidney disease classification". 14th International Conference on Engineering of Modern Electric Systems (EMES), pp 164 – 167

Abhinandan, D, 2015. "A Classification of CKD Cases Using MultiVariate K-, International Journal and Research Publications, Volume5, Issue 8, August 2015, ISSN 2250-3153.

Dulhare U.N., Ayesha M., 2016, "Extraction of action rules for chronic kidney disease using Naïve bayes classifier". IEEE International Conference on Computational Intelligence and Computing Research (ICCIC).

Tazin N., Sabab S.A., Chowdhury M.T., 2016,

- “Diagnosis of Chronic Kidney Disease using effective classification and feature selection technique”. International Conference on Medical Engineering, Health Informatics and Technology (MediTec).
- Tundjungsari V., Yugaswara H., 2015, "Supporting collaborative emergency response system wit reputation-based trust peer-to-peer file sharing.", Technology, Informatics, Management, Engineering & Environment (TIME-E), International Conference on. IEEE.
- Kopp, Jeffrey B., et al., 2007, "Kidney patient care in disasters: lessons from the hurricanes and earthquake of 2005." *Clinical Journal of the American Society of Nephrology* 2.4 (2007): 814-824.
- Foster, Mark, et al., 2011, "Personal disaster preparedness of dialysis patients in North Carolina." *Clinical Journal of the American Society of Nephrology* 6.10 (2011): 2478-2484.
- Priyogi B., Selviandro N., Hasibuan Z.A., Ahmad M., 2014, “Image Clustering Using Multi-visual Features”. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2014.
- Williams, C, 2008, "Support Vector Machine", 1(October), 1–8.
- Diana N.E., Sabiq A., 2016. “Cognitive-Affective Emotion Classification: Comparing Features Extraction Algorithm Classified by Multi-class Support Vector Machine”. *International Journal of Computer and Communication Engineering*, Vol. 5, No. 5, pp. 350-357
- Soundarapandian P., Rubini L.J., Eswaran P., 2020, “UCI Machine Learning Repository : Chronic_Kidney_Disease Data Set”,
- Diana N.E., Kalsum U., Sabiq A., Jatmiko W., Mursanto P., 2016, “Comparing windowing methods on finite impulse response (FIR) filter algorithm in electroencephalography (EEG) data processing”. *Journal of Theoretical and Applied Information Technology*, Vol. 88, No. 3, pp.558-567
- Zhu, P., Hu, Q, 2013, "Rule Extraction From Support Vector Machine Based on Consistent Region Covering Reduction”. *Knowledge-Based Systems*, 42, 1–8.