

KLASIFIKASI PENYAKIT SIROSIS MENGGUNAKAN METODE PCA-BACKPROPAGATION

¹Khofifah Auliyatuz Zahroh, ²Saiful Bahri

¹Departemen Matematika UIN Sunan Ampel, ²Departemen Biologi UIN Sunan Ampel

Jl. Ahmad Yani No. 117, Surabaya, Jawa Timur

¹khofifahauliya12@gmail.com, ²saifulsi@uinsby.ac.id

Abstrak

Penyakit sirosis menjadi penyakit endemik yang berdampak pada sejumlah besar manusia di seluruh dunia. Menurut WHO (World Health Organization) pada tahun 2016 terdapat 1.3 juta kematian yang disebabkan oleh penyakit sirosis di seluruh dunia. Penyakit ini lebih sering terjadi pada pasien lanjut usia dengan usia sekitar 40-60 tahun. Penyakit ini disebabkan oleh bermacam-macam unsur yang membahayakan hati misalnya obesitas, infeksi hepatitis yang tidak terdiagnosis, dan penyalahgunaan alkohol. Tujuan penelitian ini adalah mengklasifikasikan pasien yang terdiagnosis penyakit sirosis. Oleh karena itu, pada penelitian ini dalam melakukan klasifikasi menggunakan metode Backpropagation dengan reduksi fitur menggunakan PCA (Principal Component Analysis). Jumlah data yang digunakan sebanyak 418 data dengan 18 fitur. Kemudian dari data tersebut dilakukan tahap preprocessing serta reduksi fitur dengan PCA dan diperoleh 5 fitur yang berpengaruh. Selanjutnya melakukan analisis perbandingan Backpropagation dengan ekstraksi PCA dan tanpa ekstraksi PCA. Penelitian ini menggunakan parameter uji coba jumlah node dan learning rate. Hasil terbaik yang diperoleh berdasarkan nilai akurasi klasifikasi Backpropagation dengan menggunakan PCA sebesar 76.191% dengan jumlah node 50 dan learning rate 0.1. Sedangkan akurasi klasifikasi Backpropagation tanpa menggunakan PCA sebesar 67.857% dengan jumlah node 50 dan learning rate 0.3.

Kata Kunci: Backpropagation, Klasifikasi, PCA (Principal Component Analysis), Sirosis

Abstract

Cirrhosis is an endemic disease that affects a large number of people worldwide. According to WHO (World Health Organization) in 2016 there were 1.3 million deaths caused by cirrhosis worldwide. This disease is more common in elderly patients aged around 40-60 years. This disease is caused by various factors that damage the liver such as obesity, undiagnosed hepatitis infection, and alcohol abuse. The aim of this study was to classify patients diagnosed with cirrhosis. Therefore, in this study, in carrying out the classification using the Backpropagation method with feature reduction using PCA (Principal Component Analysis). The amount of data used is 418 data with 18 features. Then from the data the preprocessing and feature reduction stages were carried out with PCA and 5 influential features were obtained. Then perform a comparative analysis of Backpropagation with PCA extraction and without PCA extraction. This study uses the trial parameters of the number of nodes and learning rate. The best results were obtained based on the backpropagation classification accuracy value using PCA of 76.191% with 50 nodes and a learning rate of 0.1. While the accuracy of the Backpropagation classification without using PCA is 67.857% with 50 nodes and a learning rate of 0.3.

Keywords: Backpropagation, Cirrhosis, Classification, PCA (Principal Component Analysis)

PENDAHULUAN

Penyakit sirosis merupakan penyakit hati kronis yang ditandai dengan fibrosis dan

regenerasi nodul di hati yang berakibat dapat berkembang menjadi hipertensi portal dan gagal hati [1]. Penyakit ini telah menjadi salah satu penyebab kematian di dunia [2]. Setiap

tahun, sekitar 2 juta orang meninggal dunia karena penyakit hati. Menurut proyek Global Burden of Disease (GBD) yang diterbitkan dalam BMC Medicine, satu juta orang meninggal pada tahun 2010 karena sirosis dan jutaan menderita kanker hati [3]. Prevalensi penyakit sirosis di Amerika Serikat telah meningkat secara konsisten selama tiga dekade terakhir dari 12% pada sekitar awal tahun 1990 dan menjadi 15% pada tahun 2008. Sirosis bertanggung jawab atas 44.358 kematian di Amerika Serikat pada tahun 2019, dengan tingkat kematian kasar 13,5 per 100.000 [4].

Penyakit ini ditandai dengan berbagai tingkat peradangan dan nekrosis hati yang berlangsung lebih dari enam bulan [5]. Penyakit hati kronis ini dikenali hanya ketika fibrosis telah mencapai keadaan ireversibel. Fibrosis hati adalah jaringan ikat yang terbentuk dan terjadi sebagai respons terhadap kerusakan hati dan didahului oleh kerusakan hati kronis yang dapat diakibatkan oleh infeksi virus, alkoholisme, perlemakan hati, dan penyebab lainnya. Langkah terbaik dalam menangani penyakit ini adalah dengan melakukan diagnosis dini sehingga dapat dilakukan penanganan lebih cepat [6]. Diagnosis penyakit sirosis dilakukan berdasarkan pemeriksaan laboratorium, pemeriksaan fisik, patologi anatomi, anamnesis, dan radiologis [5].

Pada bidang teknologi terdapat algoritma yang dapat digunakan untuk melakukan diagnosis seperti *Machine Learning*. Algoritma *Machine Learning* telah

memberikan dampak yang signifikan pada bidang biomedis untuk mendiagnosis penyakit sirosis [3]. Salah satu metode pada algoritma *Machine Learning* yang umumnya digunakan untuk melakukan diagnosis atau klasifikasi yaitu metode *Backpropagation*. Metode *Backpropagation* merupakan metode sistematis untuk pelatihan JST (Jaringan Syaraf Tiruan) multilayer. Proses pelatihan *Backpropagation* terdiri dari tiga lapisan yaitu lapisan input, lapisan tersembunyi dan lapisan output, dimana *Backpropagation* merupakan pengembangan dari jaringan lapisan tunggal dengan dua lapisan yaitu lapisan input dan lapisan output [7]. Terdapat penelitian sebelumnya yang melakukan diagnosis penyakit hati menggunakan metode *Naïve Bayes* dengan pengujian pada 40 data dan mendapatkan tingkat akurasi sebesar 87,5% [8]. Penelitian lain yang mengklasifikasi penyakit liver dengan menggunakan metode C4.5 *Adaboost* memperoleh beberapa hasil akurasi dengan metode C4.5 *Adaboost* sebesar 77.12%, metode C4.5 sebesar 65.18%, dan metode SVM (*Support Vector Machine*) sebesar 75.36% [9].

Metode *Backpropagation* mampu menghasilkan tingkat akurasi yang lebih tinggi jika dibandingkan dengan metode lainnya seperti pada penelitian sebelumnya yang menggunakan perbandingan metode *Backpropagation* dan LVQ (*Learning Vector Quantification*) untuk mengklasifikasi kisaran harga smartphone di pasaran yang kemudian menghasilkan akurasi terbaik pada metode

Backpropagation sebesar 88.75% pada data training dan 87.50% pada data testing dibandingkan LVQ dengan akurasi 81.25% pada data training dan 72.50% pada data testing [10]. Pada penelitian lainnya juga membandingkan metode *Backpropagation* dengan beberapa metode klasifikasi lainnya seperti KNN (*K-Nearest Neighbor*), LDA (*Linear Discriminant Analysis*), *Naïve Bayes*, *Decision Tree*, dan SVM untuk melakukan klasifikasi 30 jenis buah. Pada uji coba kombinasi fitur warna dan tekstur, metode *Backpropagation* memperoleh akurasi terbaik sebesar 100% [11]. Selain itu, terdapat penelitian yang membandingkan metode *Backpropagation* dengan Regresi Linier untuk menganalisis peramalan beban listrik jangka pendek didapatkan nilai MSE terkecil pada metode *Backpropagation* sebesar 0.025% sedangkan MSE pada Regresi Linier sebesar 3.118% [12].

Berdasarkan latar belakang masalah dan beberapa penelitian sebelumnya yang melakukan perbandingan metode, dapat disimpulkan bahwa metode *Backpropagation* memiliki tingkat akurasi yang tinggi dibandingkan dengan metode klasifikasi lainnya. Oleh karena itu, penelitian ini menggunakan metode *Backpropagation* untuk mengklasifikasi penyakit sirosis. Sehingga diharapkan agar dapat mengklasifikasi penyakit sirosis dengan akurat dan memperoleh hasil akurasi yang terbaik.

METODE PENELITIAN

Fillmissing

Fillmissing bertujuan untuk mengisi data kosong yang dapat mempengaruhi tingkat akurat dan kualitas data saat diolah [13]. Pada penelitian ini melakukan proses *fillmissing* menggunakan nilai median untuk data numerik dan nilai modus untuk data kategori. Rumus perhitungan median dan modus terdapat pada Persamaan (1) dan (2).

$$Median = Tb + \left(\frac{\frac{1}{2}n - f_{kk}}{f_i} \right) \ell \quad (1)$$

$$Modus = L + \frac{d_1}{d_1 + d_2} \times c \quad (2)$$

Dimana, Tb adalah batas bawah kelas median, n adalah banyaknya data, f_{kk} adalah frekuensi kumulatif kurang dari kelas median, f_i adalah frekuensi kelas median, ℓ adalah panjang kelas median, L adalah tepi bawah kelas modus, d_1 adalah selisih antara frekuensi kelas modus dengan frekuensi kelas sebelumnya, d_2 adalah selisih antara frekuensi kelas modus dengan frekuensi kelas sesudahnya, dan c adalah interval kelas.

Normalisasi Data

Normalisasi data merupakan salah satu tahap preprocessing dimana proses penskalaan nilai atribut menjadi rentang yang lebih kecil [14]. Keberhasilan proses klasifikasi bergantung pada kualitas data untuk mendapatkan model prediksi umum dari masalah klasifikasi [15].

Min-max normalization menjadi salah satu metode yang dapat digunakan untuk normalisasi. Normalisasi *min-max* merupakan

metode normalisasi dengan memproses data asli untuk transformasi linier sehingga dapat menghasilkan keseimbangan nilai perbandingan antar data saat sebelum dan sesudah proses [16]. Rumus normalisasi *min-max* terdapat pada Persamaan (3).

$$x = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (3)$$

Dimana, x merupakan data yang di normalisasi, x_i merupakan data ke- i , x_{max} merupakan nilai maksimum dari data, x_{min} merupakan nilai minimum dari data.

Principal Component Analysis (PCA)

Metode PCA digunakan untuk melakukan reduksi jumlah variabel pada data, sehingga dari seluruh variabel yang terdapat pada data hanya akan diketahui beberapa variabel yang berpengaruh dan dari beberapa variabel yang dihasilkan dengan metode PCA dapat mewakili seluruh variabel yang ada pada data. Terdapat beberapa solusi yang dapat digunakan untuk menentukan komponen utama menggunakan PCA diantaranya yaitu dilihat dari variansi yang memiliki total 80%, dilihat dari nilai eigen yang nilainya lebih dari satu, dan dapat dilihat pada scree plot [17].

Algoritma Backpropagation

Backpropagation merupakan metode JST (jaringan syaraf tiruan) yang diawasi. Metode backpropagation menghitung gradien dari loss atau error fungsi terhadap setiap bobot menggunakan aturan rantai [18]. Tahapan mengolah data menggunakan metode

backpropagation diantaranya yaitu tahap inisialisasi, tahap aktivasi, mengubah bobot dari input layer ke hidden layer, mengubah bobot dari hidden layer ke output layer, dan tahap iterasi [19]. Berikut ini tahapan pelatihan backpropagation [20].

1. Menginisialisasi bobot dan bias.
2. Selama kondisi berhenti tidak terpenuhi, lakukan langkah 3 – 10.
3. Untuk setiap proses training, lakukan langkah 4 – 9.

Tahap *feed-forward*

4. Menghitung nilai *input* (Z_j). Kemudian nilai Z_j dipetakan oleh fungsi aktivasi sigmoid.

$$z_{net_j} = v_{0_j} + \sum_{i=1}^n x_i v_{ij} \quad (4)$$

$$z_j = f(z_{net_j}) = \frac{1}{1 + e^{-z_{net_j}}} \quad (5)$$

5. Menghitung nilai *output* (y_k). Kemudian nilai y_k dipetakan oleh fungsi aktivasi sigmoid.

$$y_{net_k} = W_{0_j} + \sum_{i=1}^n z_i w_{ij} \quad (6)$$

$$y_k = f(y_{net_k}) = \frac{1}{1 + e^{-y_{net_k}}} \quad (7)$$

Tahap *backward*

6. Menghitung nilai δ pada unit keluaran untuk menentukan tingkat kesalahan

$$\delta_k = (t_k - y_k) f'(y_{net_k}) \quad (8)$$

7. Menentukan nilai bobot untuk memperbarui nilai bobot dan bias.

$$\Delta w_{jk} = \alpha \delta_k z_j \quad (9)$$

$$\Delta w_{0k} = \alpha \delta_k \quad (10)$$

8. Menghitung nilai δ_j untuk menghitung tingkat *error*.

$$\delta_{net_j} = \sum_{k=1}^m \delta_k w_{jk} \quad (11)$$

$$\delta_j = \delta_{net_j} f'(z_{net_j}) \quad (12)$$

9. Menghitung koreksi bobot dan bias.

$$\Delta v_{ij} = \alpha \delta_j x_i \quad (13)$$

$$\Delta v_{0j} = \alpha \delta_j \quad (14)$$

Tahap *update* bobot dan bias

10. Menghitung bobot baru dan bias.

$$w_{jk}(\text{baru}) = w_{jk}(\text{lama}) + \Delta w_{jk} \quad (15)$$

11. Mencari nilai bobot dan bias baru.

$$v_{ij}(\text{baru}) = v_{ij}(\text{lama}) + \Delta v_{ij} \quad (16)$$

Confusion Matrix

Confusion matrix merupakan algoritma yang digunakan untuk mengevaluasi hasil kinerja klasifikasi dengan menghitung nilai akurasi. *Confusion matrix* berisi informasi tentang data aktual dan prediksi pada sistem klasifikasi. Terdapat tiga hasil yang merupakan hasil utama dari *confusion matrix* diantaranya yaitu akurasi, sensitifitas, dan spesifitas [21]. Untuk mendapatkan nilai tersebut, perlu dilakukan perhitungan untuk mendapatkan parameter True Negative (TN), False Positive

(FP), False Negative (FN), True Positive (TP) pada *confusion matrix*. Tabel *confusion matrix* dapat dilihat pada Tabel 1.

Perhitungan nilai akurasi, sensitifitas, dan spesifitas dapat menggunakan rumus pada Persamaan (17), (18), dan (19).

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (17)$$

$$\text{Sensitifitas} = \frac{TP}{TP+FN} \quad (18)$$

$$\text{Spesifitas} = \frac{TN}{TN+FP} \quad (19)$$

Data Penelitian

Penelitian ini termasuk dalam penelitian kuantitatif. Dalam penelitian ini menggunakan data sekunder penyakit sirosis yang diperoleh dari situs *Kaggle* dengan format data yang berupa csv. Data yang akan diproses sebanyak 418 data yang terdiri dari 19 atribut yaitu *N_Days*, *Drug*, *Ascites*, *Cholesterol*, *Albumin*, *Sex*, *Age*, *Hepatomegaly*, *Spiders*, *Edema*, *Bilirubin*, *Copper*, *Alk_Phos*, *SGOT*, *Triglycerides*, *Platelets*, *Prothrombin*, *Stage*, dan *Status* [22]. Sampel data penyakit sirosis ditunjukkan pada Tabel 2.

Tabel 1. *Confusion matrix*

Aktual	Klasifikasi	
	+	-
+	True Positif (TP)	False Negatif (FN)
-	False Positif (FP)	True Negatif (TN)

Tabel 2. Sampel Data Penyakit Sirosis

<i>N_Days</i>	<i>Drug</i>	<i>Ascites</i>	<i>Cholesterol</i>	<i>Albumin</i>	<i>Sex</i>	...	<i>Status</i>
400	D-penicillamine	Y	261	2.600	F	...	1
4500	D-penicillamine	N	302	4.140	F	...	0
3445	Placebo	N	252	3.830	M	...	1
...
51	Placebo	Y	200	2.740	F	...	1

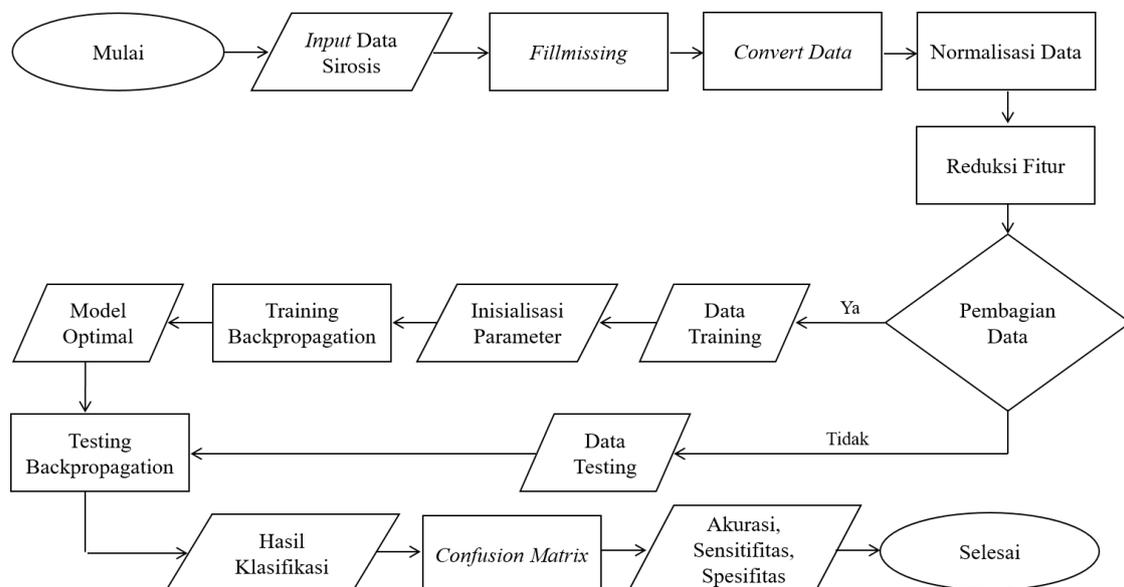
Keterangan : Y adalah *yes*, N adalah *no*, F adalah *female*, M adalah *male*.

Tahapan Penelitian

Tahap awal pada penelitian ini yaitu melakukan *input* data penyakit sirosis. Kemudian melalui tahap *preprocessing* yang meliputi *fillmissing*, konversi data kategori ke data numerik, dan normalisasi data. Karena pada data tersebut terdapat beberapa data yang kosong, maka dilakukan proses *fillmissing* menggunakan nilai median untuk data numerik dan nilai modus untuk data kategori. Tahap *preprocessing* yang selanjutnya yaitu melakukan konversi data yang berupa kategori menjadi data numerik dengan nilai antara 0 atau 1 agar dapat dilakukan proses klasifikasi. Tahap *preprocessing* yang terakhir yaitu menormalisasikan seluruh data numerik menggunakan normalisasi *min-max*. Setelah itu, melakukan reduksi fitur menggunakan PCA kemudian dilanjutkan melakukan

pembagian data menjadi dua dengan persentase 80% untuk data pelatihan dan 20% untuk data pengujian.

Setelah itu, dilakukan proses klasifikasi menggunakan metode *Backpropagation*. Dalam penelitian ini, *input layer* yang digunakan sebanyak 18 atribut dan satu *hidden layer* dengan uji coba jumlah node 10, 20, 30, 40, 50 dan *learning rate* 0.1, 0.2, 0.3. Fungsi aktivasi yang digunakan pada *hidden layer* menggunakan fungsi aktivasi sigmoid. Pada bagian *output layer* akan memberikan hasil klasifikasi berupa 2 label kelas yaitu 0 (*censored*) dan 1 (death) yang terdapat pada atribut Status. Tahap terakhir yaitu melakukan evaluasi hasil menggunakan *Confusion Matrix* untuk memperoleh nilai akurasi, sensitifitas, dan spesifitas. Urutan langkah penelitian tersebut ditunjukkan pada Gambar 1.



Gambar 1. *Flowchart* Penelitian

HASIL DAN PEMBAHASAN

Penelitian ini dimulai dengan mengumpulkan data penyakit sirosis, dilanjutkan dengan *preprocessing* data. Tahap *preprocessing* terdiri dari *fillmissing* yang bertujuan untuk mengisi data yang kosong, *convert data* bertujuan untuk mengubah data kategori menjadi data numerik dengan inisialisasi nilai antara 1 atau 0. Beberapa atribut yang dilakukan konversi yaitu *Drug*, *Ascites*, *Sex*, *Hepatomegaly*, *Spiders*, *Edema*, dan *Status*. Setelah dilakukan konversi data, kemudian normalisasi data

menggunakan *min-max normalization*. Hasil *preprocessing* terdapat pada Tabel 3.

Langkah selanjutnya yaitu melakukan reduksi jumlah variabel pada data menggunakan metode PCA (*Principal Component Analysis*). Hasil dari *loading factor* PCA ditunjukkan pada Tabel 4.

Pada Tabel 4 dapat dilihat nilai *loading factor* yang memiliki nilai lebih dari 0.5 dapat disimpulkan bahwa *component* 1 dengan fitur yang memenuhi yaitu *N_Days*, *Age*, *Edema*, dan *Stage*. *Component* 2 dengan fitur yang memenuhi yaitu *Bilirubin*, *Cholesterol*, dan *Alk_Phos*.

Tabel 3. Sampel Data Hasil *preprocessing*

<i>N_Days</i>	<i>Drug</i>	<i>Ascites</i>	<i>Cholesterol</i>	<i>Albumin</i>	<i>Sex</i>	...	<i>Status</i>
0.076	0	1	0.085	0.239	1	...	1
0.938	0	0	0.109	0.813	1	...	0
0.204	1	0	0.034	0.567	0	...	1
...
0.002	1	1	0.048	0.291	1	...	1

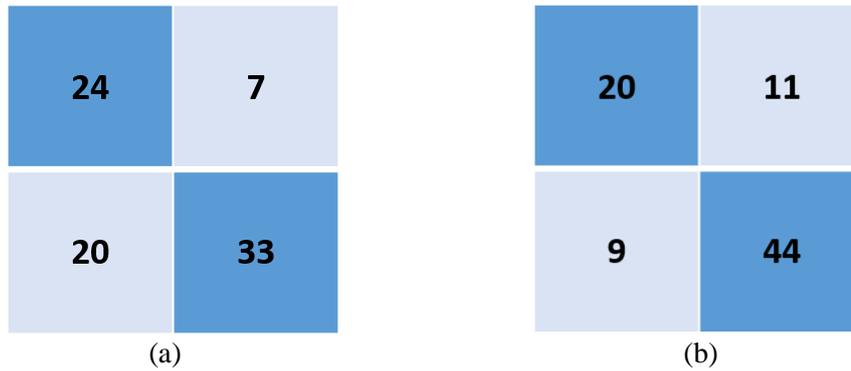
Tabel 4. Hasil *Loading factor*

	<i>Component</i>				
	1	2	3	4	5
<i>N-Days</i>	-.689	-.177	-.218	.104	.142
<i>Drug</i>	-.008	-.046	.032	.146	-.853
<i>Ascites</i>	-.172	.041	.015	.656	.019
<i>Cholesterol</i>	.381	.609	.313	.155	-.088
<i>Albumin</i>	.145	.011	.612	-.006	-.028
<i>Sex</i>	.122	-.143	.457	-.482	.275
<i>Age</i>	.642	.041	.025	-.100	.068
<i>Hepatomegaly</i>	.136	.152	.696	.065	-.022
<i>Spiders</i>	.230	.097	.262	.587	-.091
<i>Edema</i>	-.527	-.143	-.382	.075	.118
<i>Bilirubin</i>	.156	.679	-.297	.068	-.061
<i>Copper</i>	-.236	.423	.109	.184	.266
<i>Alk_Phos</i>	-.034	.674	.221	-.055	.067
<i>SGOT</i>	-.479	.421	-.115	-.155	-.272
<i>Triglycerides</i>	.411	.435	-.182	.312	.060
<i>Platelets</i>	.274	.461	.207	.228	.411
<i>Prothrombin</i>	.141	-.020	.512	.314	.184
<i>Stage</i>	.596	-.046	.247	.177	.136

Component 3 dengan fitur yang memenuhi yaitu *Hepatomegaly*, *Albumin*, dan *Prothrombin*. *Component 4* fitur yang memenuhi yaitu *Ascites*, dan *Spiders*. Sedangkan *Component 5* fitur yang memenuhi hanya *Drug*. Selanjutnya, dilakukan pembagian data hasil PCA dengan pembagian jumlah data training sebanyak 334 data dan data testing sebanyak 84 data. Setelah itu, melakukan klasifikasi menggunakan metode *Backpropagation*. Hasil uji coba terdapat pada Tabel 5.

Tabel 5. Hasil Uji Coba Penelitian

Metode	Jumlah node	Learning rate	Akurasi (%)	Sensitifitas (%)	Spesifitas (%)	Error (MSE)
<i>Backpropagation</i>	10	0.1	64.286	51.282	64.516	0.102
		0.2	61.905	49.057	83.871	0.106
		0.3	63.095	50	67.742	0.115
	20	0.1	60.714	48.077	80.645	0.108
		0.2	64.286	51.02	80.645	0.106
		0.3	63.095	50	70.968	0.088
	30	0.1	61.905	48.971	77.419	0.102
		0.2	63.095	50	77.419	0.091
		0.3	61.905	49.019	80.645	0.086
	40	0.1	60.714	48	77.419	0.108
		0.2	65.476	52.273	74.194	0.099
		0.3	64.286	50.98	83.871	0.091
	50	0.1	60.714	47.826	70.968	0.107
		0.2	63.095	50	83.871	0.088
		0.3	67.857	54.546	77.419	0.093
PCA- <i>Backpropagation</i>	10	0.1	70.238	63.636	45.161	0.136
		0.2	63.095	50	77.419	0.14
		0.3	72.619	61.111	70.968	0.133
	20	0.1	76.191	67.742	67.742	0.132
		0.2	70.238	57.143	77.419	0.138
		0.3	67.857	55.556	64.516	0.143
	30	0.1	65.476	52.083	80.645	0.137
		0.2	72.619	61.111	70.968	0.143
		0.3	73.801	62.857	70.968	0.108
	40	0.1	63.095	50	67.742	0.114
		0.2	72.619	62.5	64.516	0.125
		0.3	72.619	61.111	70.968	0.136
	50	0.1	76.191	68.966	64.516	0.115
		0.2	72.619	61.111	70.968	0.134
			0.3	71.429	58.974	74.194



Gambar 2. *Confusion Matrix* (a) *Backpropagation*, (b) *PCA- Backpropagation*

Hasil uji coba terbaik dilihat berdasarkan nilai *error* terkecil pada Tabel 4 menggunakan metode *Backpropagation* dengan jumlah node 50 dan *learning rate* 0.3 memperoleh hasil akurasi, sensitifitas, dan spesifitas yang terbaik berturut-turut sebesar 67.857%, 54.546%, dan 77.419%. Sedangkan jika menggunakan metode *PCA-Backpropagation* dengan jumlah node 50 dan *learning rate* 0.1 memperoleh hasil akurasi, sensitifitas, dan spesifitas berturut-turut sebesar 76.191%, 68.966%, dan 64.516%. Setelah didapatkan hasil uji coba, kemudian melakukan evaluasi hasil pada data klasifikasi menggunakan *Confusion Matrix*. Hasil dari *Confusion Matrix* ditunjukkan pada Gambar 2.

Berdasarkan pada Gambar 2(a), dapat dilihat hasil *Confusion Matrix* yang hanya menggunakan *Backpropagation* mendapatkan parameter TP, FP, FN, dan TN berturut-turut yaitu 24, 7, 20, dan 33. Sedangkan untuk Gambar 2(b), hasil *Confusion Matrix* yang menggunakan *PCA* dengan *Backpropagation* mendapatkan parameter TP, FP, FN, dan TN berturut-turut yaitu 20, 11, 9, dan 44.

KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian untuk klasifikasi penyakit sirosis menggunakan metode *PCA-Backpropagation* dengan melakukan uji coba jumlah node 10, 20, 30, 40, 50 dan *learning rate* 0.1, 0.2, 0.3 maka diperoleh peningkatan akurasi menjadi sebesar 76.191% jika dibandingkan dengan proses klasifikasi tanpa dilakukan *PCA* dengan nilai akurasi sebesar 67.857%. Oleh karena itu, dapat disimpulkan bahwa penggunaan *PCA* untuk mereduksi fitur memberi dampak pada peningkatan akurasi klasifikasi penyakit sirosis.

Pada penelitian ini melakukan uji coba hanya menggunakan 2 parameter, sehingga diharapkan untuk penelitian selanjutnya dapat dikembangkan dengan menambah jumlah parameter uji coba. Selain itu, untuk penelitian lebih lanjut juga dapat dilakukan dengan menambahkan metode klasifikasi lainnya sehingga dapat digunakan sebagai perbandingan metode agar dapat diketahui metode yang memberikan hasil akurasi terbaik.

DAFTAR PUSTAKA

- [1] C. Min Lee, A. Burgio, G. Young Geong, dan S. Hyung Kang, "An updated review on the epidemiology, pathophysiology, etiology, and diagnosis of liver cirrhosis," *J Gastroenterol*, vol. 56, no. March, hal. 593–619, 2021, [Daring]. Tersedia pada: <https://www.researchgate.net/publication/n/331752269>
- [2] A. M. Moon, A. G. Singal, dan E. B. Tapper, "Contemporary Epidemiology of Chronic Liver Disease and Cirrhosis," *Clin. Gastroenterol. Hepatol.*, vol. 18, no. 12, hal. 2650–2666, 2020.
- [3] A. K. M. S. Rahman, F. M. Javed Mehedi Shamrat, Z. Tasnim, J. Roy, dan S. A. Hossain, "A comparative study on liver disease prediction using supervised machine learning algorithms," *Int. J. Sci. Technol. Res.*, vol. 8, no. 11, hal. 419–422, 2019.
- [4] J. J. Alukal, H. A. Naqvi, dan P. J. Thuluvath, "Vaccination in Chronic Liver Disease: An Update," *J. Clin. Exp. Hepatol.*, vol. 12, no. 3, hal. 937–947, 2022, doi: 10.1016/j.jceh.2021.12.003.
- [5] P. P. A. Dewi, "Penyakit Hati Kronis Pada Anak," *Univ. Udayana*, hal. 9, 2019.
- [6] A. D. Kamila, P. Kedokteran, F. Kedokteran, dan U. S. Maret, "Analisis Edukasi Mahasiswa Mengenai Sirosis Sebagai Strategi Pencegahan Sirosis".
- [7] F. Zola, "Jaringan Syaraf Tiruan Menggunakan Algoritma Backpropagation Untuk Memprediksi Prestasi Siswa," *J. Teknol. Dan Open Source*, vol. 1, no. 1, hal. 58–72, 2018, doi: 10.36378/jtos.v1i1.12.
- [8] N. D. Prayoga, N. Hidayat, dan R. K. Dewi, "Sistem Diagnosis Penyakit Hati Menggunakan Metode Naïve Bayes," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 8, hal. 2666–2671, 2018.
- [9] Wiwid Wahyudi, "Implementasi Data Mining Untuk Klasifikasi Penyakit Liver Dengan C4.5 Adaboost," *J. Ilm. Tek. Inform. dan Komun.*, vol. 1, no. 3, hal. 71–76, 2021, doi: 10.55606/juitik.v1i3.120.
- [10] M. Y. Anshori, D. Rahmalia, dan T. Herlambang, "Comparison Backpropagation (BP) and Learning Vector Quantification (LVQ) on classifying price range of smartphone in market," *J. Phys. Conf. Ser.*, vol. 1836, no. 1, 2021, doi: 10.1088/1742-6596/1836/1/012040.
- [11] S. Ghazal, W. S. Qureshi, U. S. Khan, J. Iqbal, N. Rashid, dan M. I. Tiwana, "Analysis of visual features and classifiers for Fruit classification problem," *Comput. Electron. Agric.*, vol. 187, no. February, hal. 106267, 2021, doi: 10.1016/j.compag.2021.106267.
- [12] F. R. Rizqulloh, "Analisa Perbandingan Peramalan Beban Listrik Jangka Pendek antara Metode Backpropagation Neural

- Network dengan Metode Regresi Linier,” hal. 69–77, 2019.
- [13] S. Zulaikhah Hariyanti Rukmana, A. Aziz, dan W. Harianto, “Optimasi Algoritma K-Nearest Neighbor (Knn) Dengan Normalisasi Dan Seleksi Fitur Untuk Klasifikasi Penyakit Liver,” *JATI (Jurnal Mhs. Tek. Inform.,* vol. 6, no. 2, hal. 439–445, 2022, doi: 10.36040/jati.v6i2.4722.
- [14] D. A. Nasution, H. H. Khotimah, dan N. Chamidah, “PERBANDINGAN NORMALISASI DATA UNTUK KLASIFIKASI WINE MENGGUNAKAN ALGORITMA K-NN,” *J. Comput. Eng. Syst. Sci.,* vol. 4, 2019.
- [15] D. Singh dan B. Singh, “Investigating the impact of data normalization on classification performance,” *Appl. Soft Comput.,* vol. 97, no. xxxx, hal. 105524, 2020, doi: 10.1016/j.asoc.2019.105524.
- [16] D. A. Nasution, H. H. Khotimah, dan N. Chamidah, “Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN,” *Comput. Eng. Sci. Syst. J.,* vol. 4, no. 1, hal. 78, 2019, doi: 10.24114/cess.v4i1.11458.
- [17] N. M. Sunariadi, S. N. Fadilah, dan D. C. R. Novitasari, “Analisis Resiko Kanker Serviks Menggunakan PCA-ANFIS Berdasarkan Historical Medical Record,” *J. Media Inform. Budidarma,* vol. 6, no. 3, hal. 1349, 2022, doi: 10.30865/mib.v6i3.3901.
- [18] I. P. B. D. Purwanta, N. P. N. P. Dewi, dan C. K. Adi, “Backpropagation Neural Network for Book Classification Using the Image Cover,” *Int. J. Appl. Sci. Smart Technol.,* vol. 2, no. 2, hal. 89–106, 2020, doi: 10.24071/ijasst.v2i2.2653.
- [19] D. Chen, X. Li, dan S. Li, “A Novel Convolutional Neural Network Model Based on Beetle Antennae Search Optimization Algorithm for Computerized Tomography Diagnosis,” *IEEE Trans. Neural Networks Learn. Syst.,* vol. 34, hal. 1418–1429, 2023.
- [20] E. S. Putri, A. Hamid, dan D. C. R. Novitasari, “Prediction of the Tuban Regency Open Unemployment Rate in 2022 with the Backpropagation Method,” *Numer. J. Mat. dan Pendidik. Mat.,* vol. 6, hal. 25–36, 2022, doi: 10.25217/numerical.v6i1.2150.
- [21] D. C. R. Novitasari, A. Z. Foady, M. Thohir, A. Z. Arifin, K. Niam, dan A. H. Asyhar, “Automatic Approach for Cervical Cancer Detection Based on Deep Belief Network (DBN) Using Colposcopy Data,” *2020 Int. Conf. Artif. Intell. Inf. Commun. ICAIIC 2020,* hal. 415–420, 2020, doi: 10.1109/ICAIC48513.2020.9065196.
- [22] Fedesoriano, “Cirrhosis Prediction Dataset,” 2021.