

# OPTIMASI NAÏVE BAYES MENGGUNAKAN PSO UNTUK TINGKAT KEBERHASILAN CRYOTHERAPY PADA PENYAKIT KUTIL

Rahayu Febryani<sup>1</sup>, Toni Arifin<sup>2</sup>

<sup>1</sup>Universitas Adhirajasa Reswara Sanjaya  
e-mail: febryanirahayu@gmail.com

<sup>2</sup>Universitas Adhirajasa Reswara Sanjaya  
e-mail: toni.arifin@ars.ac.id

## Abstrak

Penyakit kutil menjadi salah satu penyebab masalah kesehatan pada kulit dengan ditandai timbulnya benjolan kecil pada kulit. Penyakit ini disebabkan oleh virus *Human Papilloma Virus* atau biasa disebut dengan virus *HPV*. Ada berbagai cara dalam penyembuhan penyakit kutil, salah satunya dengan melakukan *cryotherapy*. Metode *cryotherapy* ialah teknik pengobatan terapi dengan berendam didalam es atau air yang dingin selama kurang lebih 30 menit, dalam suhu -18 derajat celcius sampai 24 derajat celcius, sehingga sel-sel kanker dalam tubuh pasien membeku serta terbukti dapat memperpanjang umur pasien bahkan ada kemungkinan sembuh. Untuk mengurangi kesalahan deteksi penyakit ini serta menghindari keterlambatan dalam mendiagnosis dengan memanfaatkan teknik *data mining*. Salah satu metode *data mining* yang dapat digunakan yaitu klasifikasi algoritma *Naïve Bayes* serta dalam meningkatkan akurasi akan menggunakan pembobotan optimasi *Particle Swarm Optimization*. Pada penelitian ini, algoritma *Naïve Bayes* berbasis *Particle Swarm Optimization* mengalami peningkatan nilai akurasi sebesar 97.22% dan nilai AUC sebesar 0.991 yang termasuk dalam kategori *Excellent Classification*.

**Kata Kunci:** Penyakit Kutil, *Cryotherapy*, *Data Mining*, *Naïve Bayes*, *Particle Swarm Optimization*

## Abstract

Warts are one of the causes of health problems on the skin characterized by the appearance of small bumps on the skin. This disease is caused by the *Human Papilloma Virus* or commonly known as the *HPV virus*. There are various ways to cure warts, one of them is by doing *cryotherapy*. The *cryotherapy* method is a therapeutic treatment technique by soaking in ice or cold water for approximately 30 minutes, at a temperature of -18 degrees Celsius to 24 degrees Celsius, so that cancer cells in the patient's body freeze and are proven to prolong the patient's life and there is even a possibility of recovery. . To reduce error detection of this disease and avoid delays in diagnosing by utilizing *data mining* techniques. One of the *data mining* methods that can be used is the classification of the *Naïve Bayes* algorithm and in increasing the accuracy will use *Particle Swarm Optimization* optimization weighting. In this study, the *Naïve Bayes* algorithm based on *Particle Swarm Optimization* has increased the accuracy value of 97.22% and the AUC value of 0.991 which is included in the *Excellent Classification* category.

**Keywords:** *Wart Disease*, *Cryotherapy*, *Data Mining*, *Naïve Bayes*, *Particle Swarm Optimization*

## 1. Pendahuluan

Penyakit kutil menjadi salah satu penyebab masalah kesehatan pada kulit dengan ditandai timbulnya benjolan kecil pada kulit. Penyakit ini disebabkan oleh virus *Human Papiloma Virus* atau biasa disebut

virus *HPV* (T. Arifin & Syalwah, 2020), (Supriyatna & Mustika, 2018). Virus ini menginfeksi lapisan atas kulit, seseorang dapat terinfeksi apabila bersentuhan langsung oleh penderita kutil. Akan tetapi, tidak semua orang yang bersentuhan dengan virus *Human Papiloma Virus* (*HPV*) tertular

penyakit kutil. Sistem kekebalan tubuh (imun) setiap orang berbeda sehingga akan menjadi pengaruh dalam penularan penyakit ini, seseorang yang memiliki immunitas rendah rentan tertular oleh penyakit ini (Brawijaya et al., 2019), (Hermawan et al., 2019).

*Human Papiloma Virus* (HPV) merupakan virus yang menginfeksi area kulit organ kelamin. Lebih dari 100 jenis tipe *Human Papiloma Virus* (HPV) yang sudah teridentifikasi. Salah satunya menginfeksi kulit dan menimbulkan kutil pada organ kelamin. Sementara sebagian menginfeksi organ kelamin seperti Kanker mulut rahim, Kanker Vagina, Penis dan Anus. Sehingga kebanyakan dari penderita datang ke dokter ketika sudah terinfeksi *Human Papiloma Virus* (HPV) yang dikarenakan kurangnya pengetahuan masyarakat mengenai virus ini (Tanjung et al., 2016). *Human Papiloma Virus* (HPV) dikategorikan menjadi dua tipe yaitu tipe resiko tinggi dan tipe resiko rendah sehingga dapat dilihat tergantung bagaimana kemampuan virus tersebut menimbulkan infeksi yang berhubungan dengan timbulnya kanker (Marlina et al., 2016).

Ada berbagai cara dalam penyembuhan penyakit kutil, salah satunya dengan melakukan *cryotherapy*. Metode *cryotherapy* ialah teknik pengobatan terapi dengan berendam didalam es atau air yang dingin selama kurang lebih 30 menit, dalam suhu -18 derajat celcius sampai -24 derajat celcius, sehingga sel-sel kanker dalam tubuh pasien membeku serta terbukti dapat memperpanjang umur pasien bahkan ada kemungkinan sembuh (Brawijaya et al., 2019), (Rahayu et al., 2019). Namun, seperti terapi yang lainnya, *cryotherapy* juga memiliki efek samping yang dapat menyebabkan kerusakan jaringan melalui pembentukan Kristal es ekstra dan intraseluler, distrupsi membran sel serta perubahan sirkulasi bagian kulit. Terapi ini tidak disarankan untuk lesi luas dan memiliki efek lokal yang sering terjadi yaitu nekrosis, nyeri, pembentukan bula, edem dan hipopigmentasi. Terapi ini juga aman untuk wanita hamil, tetapi sangat tidak dianjurkan dilakukan pada anak (Brawijaya et al., 2019).

*Data mining* adalah serangkaian proses dalam menggali nilai tambah pada beberapa informasi yang belum diketahui secara manual dari suatu basis data. Dari sudut pandang keilmuan, *data mining* dapat digunakan untuk mengcapture, menganalisis serta menyimpan data yang bersifat real time

dan sangat besar (Vulandari, 2017). Salah satu proses eksplorasi data, penggunaan *data mining* dapat digunakan dalam pengklasifikasian data (Siswa & Prihandoko, 2018). Klasifikasi data merupakan proses penemuan model atau fungsi untuk membedakan konsep atau kelas data sehingga bertujuan dalam menghasilkan kelas dari suatu objek yang labelnya tidak diketahui (Vulandari, 2017).

Algoritma *Naïve Bayes* merupakan klasifikasi statistik yang dapat digunakan dalam memprediksi probabilitas keanggotaan suatu class. *Naïve Bayes* menjadi salah satu metode klasifikasi yang terkenal dan termasuk kedalam sepuluh algoritma terbaik dalam *Data Mining* (Nurelasari, 2018). *Naïve Bayes* memiliki perhitungan matematik dasar yang kuat serta efisiensi klasifikasinya stabil (Wijayatun & Sulisty, 2016). Namun, metode *Naïve Bayes* memiliki kekurangan dalam memilih atribut yang ada pada data sehingga mempengaruhi hasil akhir nilai akurasi (T. Arifin & Ariesta, 2019). Maka, *Naïve Bayes* perlu dioptimasi dengan memberi bobot pada atribut sehingga *Naïve Bayes* dapat bekerja lebih efektif. Untuk menyelesaikan permasalahan tersebut, optimasi menggunakan *Particle Swarm Optimization* (PSO) dapat digunakan dalam meningkatkan akurasi pada *Naïve Bayes* (Muhamad et al., 2017).

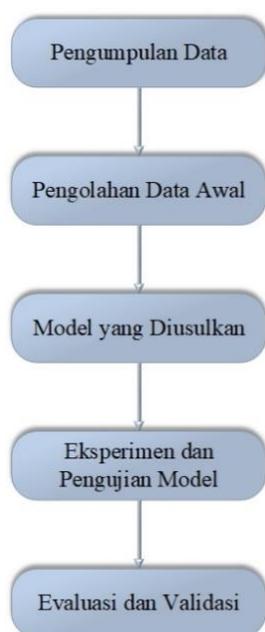
*Particle Swarm Optimization* (PSO) adalah salah satu metode optimasi yang terinspirasi dari gerak serta perilaku hewan dalam mencari mangsa (T. Arifin, 2017). Secara umum karakteristik *Particle Swarm Optimization* (PSO) memiliki konsep sederhana, komputasi efisien, dan implementasinya yang mudah. (T. Arifin & Herliana, 2018). Partikel dalam *Particle Swarm Optimization* (PSO) bergerak dalam ruang pencarian dengan kecepatan dinamis sesuai dengan histori mereka, sehingga cenderung bergerak menuju daerah pencarian yang lebih baik (Muhamad et al., 2017).

Metode ini dipakai berdasarkan beberapa penelitian sebelumnya, dengan judul Analisis Tingkat Keberhasilan *Cryotherapy* Menggunakan *Neural Network* menghasilkan nilai akurasi sebesar 87,78% (Rahayu et al., 2019), Komparasi Algoritma *K-Nearest Neighbor* dan *Naïve Bayes* pada Pengobatan Penyakit Kutil menggunakan *Cryotherapy* dengan nilai akurasi *K-Nearest Neighbor* sebesar 90,00% dan nilai akurasi

*Naïve Bayes* sebesar 86,67% (Brawijaya et al., 2019), Prediksi Keberhasilan *Immunotherapy* Pada Penyakit Kutil Menggunakan Algoritma *Naïve Bayes* menghasilkan nilai akurasi sebesar 81,11% (T. Arifin & Syalwah, 2020), dan Komparasi Algoritma *Naïve Bayes* dan *SVM* Untuk Memprediksi Keberhasilan *Immunotherapy* Pada Penyakit Kutil (Supriyatna & Mustika, 2018).

## 2. Metode Penelitian

Menurut Dawson (2009:26), metode penelitian paling umum yang biasa digunakan baik secara umum atau gabungan ada empat diantaranya penelitian tindakan, eksperimen, studi kasus serta survei. Penelitian ini akan menggunakan metode penelitian eksperimen dengan pendekatan kuantitatif. Penelitian dilakukan dengan tahapan seperti pada gambar 1.



Gambar 1. Tahapan Penelitian

### Pengumpulan Data

Dalam tahap ini, akan menentukan data yang diteliti. Serta menjelaskan sumber data yang ada dengan cara memperoleh data (T. Arifin & Ariesta, 2019).

*Dataset* yang digunakan dalam penelitian ini ialah *Cryotherapy Dataset* yang termasuk dalam data publik, atau data sekunder yang diperoleh dari penelitian Khozeimeh et al., dalam penelitiannya pada

tahun 2017 dan dipublikasi melalui web *University of California Irvine Machine Learning Repository* (UCI Repository) pada tahun 2018. *Cryotherapy Dataset* terdiri dari 90 data pasien yang memiliki penyakit kutil dan melakukan metode penyembuhan dengan *Cryotherapy*. Data diperoleh dari Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran. Pada *dataset Cryotherapy Dataset* memiliki 7 atribut serta *class* terdiri dari 1 (*Berhasil*) dan 0 (*Gagal*). Di mana 48 *record* pada *class* 1 (*Berhasil*) dan 42 *record* pada *class* 0 (*Gagal*). *Dataset* ini diambil pada halaman situs web UCI *Machine Learning Repository* dengan nama *dataset Cryotherapy Data Set*.

### Pengolahan Data Awal

Pada tahap ini akan dilakukan eksplorasi terhadap *dataset* yang bertujuan untuk memastikan semua atribut serta *class* pada *dataset* valid sehingga dapat digunakan dalam objek penelitian. Sehingga dapat mengetahui hasil klasifikasi terbaik dalam metode penyembuhan penyakit kutil tercapai.

### Data Cleaning

Pada *dataset cryotherapy* penulisan *class* diagnosis menggunakan numerikal 1 (*Berhasil*) dan 0 (*Gagal*). Namun, dalam algoritma *Naïve Bayes* tidak dapat menggunakan *dataset* dengan *class* bertipe numerikal. Sehingga akan dilakukan *data cleaning* dengan mengubah numerikal menjadi menjadi teks. Penggunaan *data cleaning* dipakai untuk mengubah numerikal 1 menjadi *Berhasil* dan numerikal 0 menjadi *Gagal*.

### Data Transformation

Dalam *Cryotherapy Dataset* memiliki 7 atribut dengan jumlah 90 *record*. Akan tetapi, pada proses *data mining* menggunakan algoritma *Naïve Bayes* akan mendapatkan hasil secara optimal jika variabel atau atribut memiliki nilai kategorikal bukan nilai *continus*. Tetapi, pada *dataset cryotherapy* memiliki variabel campuran seperti variabel diskrit atau kategorikal dan variabel *continus* atau numerikal. Sehingga memerlukan proses transformasi data dalam tahap awal pengolahan data. Proses transformasi ini menggunakan *discretization* yang mengubah nilai baku dari atribut numerik seperti *age* menjadi label *interval* (10-20, 21-30). Proses ini bertujuan untuk mengubah semua variabel *continus* atau

baku menjadi variabel kategorikal atau interval.

Di bawah ini merupakan tabel yang menjelaskan nama atribut yang ada pada *Cryotherapy Dataset*.

Tabel 1. Nama-nama atribut

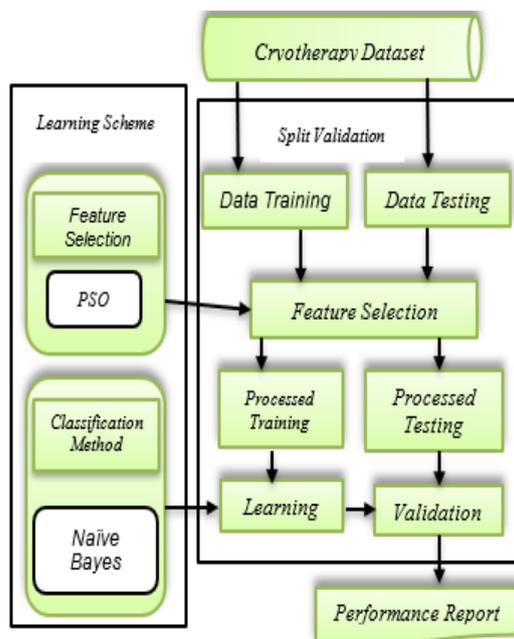
No.	Atribut
1	Sex
2	Age
3	Time
4	Number of Warts
5	Type
6	Area
7	Result of Treatment

Tabel 2. Jumlah data 2 class dataset

Kelas Berhasil	48
Kelas Gagal	42
Total	90

### Model yang Diusulkan

*Dataset Cryotherapy* yaitu data sekunder yang siap diproses pada *data mining*. Model yang digunakan dalam proses ini merupakan model yang belum pernah digunakan oleh para peneliti sebelumnya. Dalam *dataset cryotherapy* ini model yang akan digunakan ialah *Naïve Bayes* serta *Particle Swarm Optimization* sebagai optimasi. Berikut merupakan model yang diusulkan dapat dilihat pada gambar 2.



Gambar 2. Model yang Diusulkan

Pada kerangka model yang diusulkan Gambar 2. *dataset cryotherapy* dibagi menjadi dua bagian yaitu *data training* dan *data testing*. *Dataset* dibagi menggunakan *split validation* menjadi 60% untuk *data training* dan 40% untuk *data testing*. Kemudian akan diuji dengan metode *Naïve Bayes* dan *Particle Swarm Optimization*.

### Eksperimen dan Hasil Pengujian

Pada tahap ini, akan dilakukan eksperimen dan pengujian model menggunakan pemodelan data yang diproses dengan metode yang sudah diusulkan dalam sebuah *software*. Penelitian ini, akan melakukan eksperimen untuk *dataset cryotherapy*. Kemudian *dataset* akan dibagi menjadi dua bagian yaitu *data training* sebesar 60% dan *data testing* sebesar 40%. Lalu akan diuji menggunakan algoritma *Naïve Bayes*. Kemudian diuji kembali menggunakan algoritma *Naïve Bayes* dan *Particle Swarm Optimization* dalam meningkatkan nilai akurasi. Model ini akan menghasilkan nilai akurasi dengan tampilan *Confusion Matrix* serta diukur menggunakan *Area Under Curve* (AUC).

### Naïve Bayes

*Naïve Bayes Classifier* termasuk kedalam algoritma *Bayesian Learning* yang cepat dan sederhana. Hal ini berdasarkan peluang dari teorema *Bayes* yang menghasilkan *classifier*. Metode *Naïve Bayes Classifier* termasuk pendekatan statistik dalam melakukan inferensi induksi dalam klasifikasi (T. Arifin & Ariesta, 2019).

*Naïve Bayes* memiliki teknik sederhana yang cocok dalam dimensi dari input tinggi. Untuk membangun *Naïve Bayes* perlu menghitung probabilitas setiap atribut yang diberi *class* (Sulaehani, 2016).

Persamaan Teorema *Bayes* (Sulaehani, 2016) :

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Keterangan :

X : Data *class* yang belum diketahui  
 H : Hipotesis data X dalam suatu *class* spesifik

$P(H|X)$  : Probabilitas hipotesis berdasarkan suatu kondisi (*posteriori probability*)

$P(H)$  : Probabilitas hipotesis H (*prior probability*)

$P(X|H)$  : Probabilitas X berdasarkan kondisi hipotesis H  
 $P(X)$  : Probabilitas X

Untuk membangun model *Naïve Bayes*, perlu menghitung probabilitas pada setiap atribut yang diberi atribut *class*. Maka dari itu, dalam menggunakan atribut dalam jumlah besar, nilai diskrit perlu waktu komputasi lebih lama (Sulaehani, 2016).

### Particle Swarm Optimization (PSO)

*Particle Swarm Optimization* (PSO) adalah salah satu metode optimasi yang terinspirasi dari gerak dan perilaku hewan dalam mencari mangsa (T. Arifin, 2017). Secara umum karakteristik *Particle Swarm Optimization* (PSO) memiliki konsep sederhana, komputasi efisien, dan implementasinya mudah. (T. Arifin & Herliana, 2018). Partikel pada *Particle Swarm Optimization* (PSO) bergerak dalam ruang pencarian dengan kecepatan dinamis sesuai dengan histori mereka, sehingga cenderung bergerak menuju daerah pencarian yang lebih baik (Muhamad et al., 2017).

*Particle Swarm Optimization* (PSO) merupakan teknik komputasi yang mampu menghasilkan solusi secara optimal untuk pencarian melalui interaksi dalam kawanan partikel. Pada setiap partikel mampu mengungkapkan informasi berupa posisi terbaik dalam partikel lain dan menyesuaikan posisi serta kecepatan berdasarkan masing-masing informasi yang diterima pada posisi paling baik (Nurelasari, 2018).

Rumus perhitungan untuk menghitung *Particle Swarm Optimization* (PSO) sebagai berikut (Nurelasari, 2018) :

$$V_{i,m} = W \cdot V_{i,m} + C_1 * R * (pbest_{i,m} - X_{i,m}) + C_2 * R * (gbest_m - X_{i,m})$$

$$X_{i,d} = X_{i,m} + V_{i,m}$$

Keterangan :

$i$  : jumlah partikel kelompok  
 $d$  : dimensi  
 $V_{i,m}$  : kecepatan partikel ke- $i$  pada iterasi ke- $i$   
 $W$  : faktor bobot inersia  
 $C_1, C_2$  : konstanta akselerasi (*learning rate*)  
 $R$  : bilangan random (0-1)  
 $X_{i,d}$  : posisi saat partikel ke- $i$  pada iterasi ke- $i$   
 $pbest_i$  : posisi terbaik sebelum artikel ke- $i$

$gbest_i$  : partikel terbaik antara semua partikel dalam satu kelompok atau populasi

### Evaluasi dan Validasi

Evaluasi digunakan untuk melakukan pengamatan dan analisa hasil kerja algoritma. Sedangkan validasi digunakan untuk melakukan pengukuran pada hasil prediksi (Rahayu et al., 2019). Akurasi dapat diartikan sebagai tingkat kedekatan nilai prediksi dan nilai aktual. Presisi bertujuan untuk menunjuk tingkat ketepatan atau ketelitian dalam klasifikasi. Sedangkan *recall* fungsinya untuk mengukur proporsi positif aktual yang benar untuk diidentifikasi (O. Arifin & Sasongko, 2018).

Dalam mengukur evaluasi, akurasi, presisi dan *recall* dapat menggunakan *confusion matrix* yang bertujuan untuk mempermudah menganalisis performa algoritma. *Confusion matrix* menghasilkan informasi dalam bentuk angka, sehingga dapat melihat rasio keberhasilan suatu data. Evaluasi pada model berdasarkan pada pengujian dalam memperkirakan objek yang benar dan salah. *Confusion matrix* termasuk dalam salah satu alat ukur berbentuk matrik 2x2 yang digunakan dalam mendapatkan jumlah ketepatan klasifikasi *dataset* pada *class* aktif dan tidak aktif dalam algoritma yang dipakai (Sulaehani, 2016),(O. Arifin & Sasongko, 2018).

Tabel 3. *Confusion Matrix*

Classification	Predicted Class		
	Class = Yes	Class = No	
Observed Class	Class = Yes	True Positif (TP)	False Negatif (FN)
	Class = No	False Positif (FP)	True Negatif (TN)

Sumber : (Sulaehani, 2016)

Keterangan :

*True Positif* (TP) : Sampel yang bernilai *true* diprediksi secara *true*  
*False Positif* (FP) : Sampel yang bernilai *false* diprediksi secara *true*  
*False Negatif* (FN) : Sampel yang bernilai *true* diprediksi secara *false*  
*True Negatif* (TN) : Sampel yang bernilai *true* diprediksi secara *false*

Rumus perhitungan untuk menghitung nilai *accuracy*, *precision* dan *recall* yaitu sebagai berikut (O. Arifin & Sasongko, 2018) :

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

$$precision = \frac{TP}{TP + FP} \times 100\%$$

$$recall = \frac{TP}{TP + FN} \times 100\%$$

*Receiver Operating Characteristic* (ROC) ialah grafik yang digunakan untuk menilai suatu model. Sedangkan *Area Under Curve* (AUC) merupakan luas area bawah kurva. Luas AUC selalu ada diantara nilai 0 hingga 1. AUC dapat dihitung menurut gabungan luas trapesium titik-titik (*sensitivity* dan *specificity*) (Arifin & Sasongko, 2018) (Sulaehani, 2016). Sedangkan pada klasifikasi, ROC merupakan teknik visualisasi, mengatur serta memilih klasifikasi berdasarkan kinerjanya (Gorunescu, 2011).

Kategori klasifikasi berdasarkan nilai AUC yaitu sebagai berikut (O. Arifin & Sasongko, 2018), (Gorunescu, 2011) :

1. 0.90 – 1.00 = *Excellent classification*
2. 0.80 – 0.90 = *Good classification*
3. 0.70 – 0.80 = *Fair classification*
4. 0.60 – 0.70 = *Poor classification*
5. 0.50 – 0.60 = *Failure classification*

### 3. Hasil dan Pembahasan

Penerapan dalam algoritma *Naïve Bayes* menggunakan *Particle Swarm Optimization* pada *Cryotherapy Dataset* bertujuan untuk mengetahui serta mendapatkan hasil akurasi yang lebih baik pada klasifikasi penyakit kutil dari penelitian sebelumnya. Eksperimen ini menggunakan metode validasi *Split Validation*.

*Dataset* yang telah diolah akan dilakukan perhitungan menggunakan *RapidMiner* untuk mengetahui hasil akurasi metode yang digunakan pada penelitian ini.

#### 3.1. Hasil Eksperimen *Naïve Bayes* dengan *Split Validation*

*Dataset* dibagi menjadi 60% *data training* sebesar 54 data dan 40 % *data testing* sebesar 36 data. Kemudian dihitung menggunakan *Naïve Bayes* dan menghasilkan *confusion matrix* seperti Tabel 4.

Tabel 4. Hasil Evaluasi Model *Split Validation* dengan Perbandingan 60 : 40

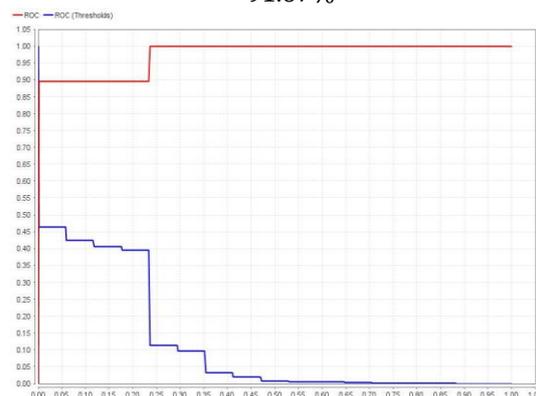
	<i>True 0</i>	<i>True 1</i>	<i>Class Precision</i>
<i>Pred. 0</i>	17	3	85.00%
<i>Pred. 1</i>	0	16	100.00%
<i>Class Recall</i>	100.00%	84.21%	

Dari *confusion matrix* pada tabel 4. dapat diukur tingkat akurasi dari klasifikasi sebagai berikut :

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

$$= \frac{17 + 16}{17 + 16 + 3 + 0} \times 100\%$$

$$= 91.67\%$$



Gambar 3. Grafik ROC pada eksperimen Metode *Naïve Bayes*

Selain menghasilkan *confusion matrix*, pengujian tersebut menghasilkan kurva ROC (*Receiver Operating Characteristic*) seperti pada gambar 3 dengan menghasilkan nilai AUC sebesar 0.975 serta termasuk dalam *Excellent Classification*.

#### 3.2. Hasil Eksperimen Menggunakan *Naïve Bayes* dan *Particle Swarm Optimization*

Pada pembobotan atribut akan menggunakan algoritma *Particle Swarm Optimization*. Sehingga untuk mengetahui akurasi pada *Naïve Bayes* berbasis *Particle Swarm Optimization* akan dilakukan perhitungan menggunakan *RapidMiner* dan menghasilkan *weight* seperti Tabel 5.

Tabel 5. Hasil Seleksi Atribut

Attribute	Weight
<b>Sex</b>	0.241
<b>Age</b>	0.401
<b>Time</b>	0.189
<b>Number of Warts</b>	0.029
<b>Type</b>	0.142
<b>Area</b>	0.077

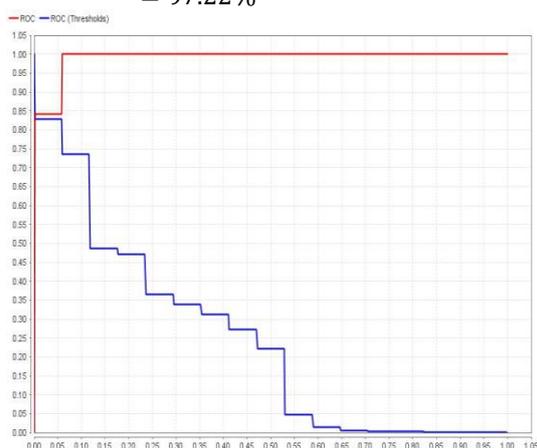
Hasil seleksi atribut memperoleh 1 atribut terendah yang memiliki *weight* 0.029 yaitu *Number Of Warts* yang berarti atribut tersebut tidak memiliki pengaruh terhadap nilai akurasi. Sehingga jika atribut ini dihilangkan akan meningkatkan nilai akurasi.

Tabel 6. Hasil Evaluasi Model *Split Validation* dengan Perbandingan 60 : 40

	True 0	True 1	Class Precision
Pred. 0	17	3	100.00%
Pred. 1	0	16	95.00%
Class Recall	94.12%	100.00%	

Dari *confusion matrix* pada tabel 6. dapat diukur tingkat akurasi dari klasifikasi sebagai berikut :

$$\begin{aligned}
 accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \\
 &= \frac{16 + 19}{16 + 19 + 0 + 1} \times 100\% \\
 &= 97.22\%
 \end{aligned}$$

Gambar 4. Grafik ROC pada eksperimen Metode *Naive Bayes*

Selain menghasilkan *confusion matrix*, pengujian tersebut menghasilkan

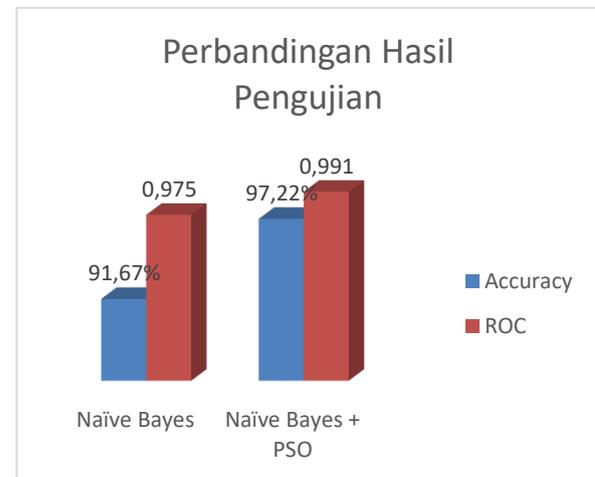
kurva ROC (*Receiver Operating Characteristic*) seperti pada gambar 4 dengan nilai AUC sebesar 0.991 dan termasuk *Excellent Classification*.

### 3.3. Pembahasan

Berdasarkan hasil penelitian di atas, pada tahap ini akan dilakukan evaluasi sehingga dapat mengetahui peningkatan nilai akurasi dari klasifikasi metode *Naive Bayes* sebelum dilakukan seleksi fitur dan *Naive Bayes* setelah dilakukan seleksi fitur *Particle Swarm Optimization* dengan cara menghitung *data testing* dalam bentuk nilai akurasi serta grafik ROC. Berikut merupakan perbandingan nilai akurasi *Naive Bayes* sebelum dilakukan seleksi fitur dan *Naive Bayes* setelah seleksi fitur menggunakan *Particle Swarm Optimization*.

Tabel 7. Perbandingan Hasil Pengujian

No.	Metode	Accuracy	ROC
1	<i>Naive Bayes</i>	91.67%	0.975
2	<i>Naive Bayes + Particle Swarm Optimization</i>	97.22%	0.991



Gambar 5. Grafik Perbandingan Hasil Pengujian

Pada Gambar 5. dapat dilihat bahwa hasil klasifikasi metode *Naive Bayes* mengalami peningkatan akurasi sebesar 5.55% setelah dilakukan seleksi fitur menggunakan *Particle Swarm Optimization*. Sehingga terbukti bahwa seleksi fitur menggunakan *Particle Swarm Optimization* dapat meningkatkan akurasi dalam metode *Naive Bayes* pada penyakit kutil.

Perbandingan hasil penelitian dengan peneliti sebelumnya juga dilakukan sehingga dapat mengevaluasi hasil keseluruhan penelitian. Penelitian ini merupakan penelitian lanjutan dari penelitian-penelitian sebelumnya dengan menggunakan data yang sama yaitu *Cryotherapy Dataset*.

Penelitian pertama dilakukan oleh (Cüvitoğlu & Işık, 2018) dengan menggunakan metode *Naïve Bayes*, *Random Forest*, *Support Vector Machine* (SVM), *K-Nearest Neighbor* (K-NN) dan *Artificial Neural Network* (ANN). Penelitian kedua dilakukan oleh (Rahayu et al., 2019) menggunakan metode *Neural Network*. Penelitian ketiga dilakukan oleh (Brawijaya et al., 2019) dengan menggunakan *K-Nearest Neighbor* (K-NN) dan *Naïve Bayes*. Berikut ini merupakan perbandingan hasil penelitian ini dengan peneliti-peneliti sebelumnya yang dapat dilihat pada tabel 8.

Tabel 8. Perbandingan Hasil Penelitian

NO.	Judul Penelitian	Metode	Akurasi
1	<i>Evaluation Machine Learning Approaches for Classification of Cryotherapy and Immunotherapy Datasets</i>	<i>Naïve Bayes</i>	86%
		<i>Random Forest</i>	95%
		<i>Support Vector Machine (SVM)</i>	91%
		<i>K-Nearest Neighbor (K-NN)</i>	93%
		<i>Artificial Neural Network (ANN)</i>	89%
2	<i>Analisis Tingkat Keberhasilan Cryotherapy Menggunakan Neural Network</i>	<i>Neural Network</i>	87.78% dan AUC 0.955
3	<i>Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes pada Pengobatan Penyakit Kulit</i>	<i>K-Nearest Neighbor (K-NN)</i>	90.00% dan AUC 0.500
		<i>Naïve Bayes</i>	86.67% dan AUC 0.932

	<i>Menggunakan Cryotherapy</i>		
--	--------------------------------	--	--

Pada tabel 8. merupakan perbandingan dari hasil penelitian yang dilakukan peneliti sebelumnya sehingga menunjukkan bahwa hasil klasifikasi Penyakit Kulit menggunakan metode *Naïve Bayes* dan *Particle Swarm Optimization* menghasilkan akurasi sebesar 97.22% dan AUC 0.991.

#### 4. Kesimpulan dan Saran

##### Kesimpulan

Pada penelitian ini dilakukan eksperimen menggunakan *Naïve Bayes* berbasis *Particle Swarm Optimization* untuk mengklasifikasi *Cryotherapy Dataset* yang diambil dari *UCI Repository*. Untuk mendapatkan hasil yang optimal dari klasifikasi ini dilakukan metode validasi performa klasifikasi metode *Split Validation*. Kesimpulan dari hasil penelitian ini yaitu :

1. Telah diterapkan algoritma klasifikasi *Naïve Bayes* berbasis *Particle Swarm Optimization* pada *Cryotherapy Dataset* menggunakan metode *Split Validation*. Setelah dilakukan optimasi pada *Cryotherapy Dataset* menggunakan *Naïve Bayes* berbasis *Particle Swarm Optimization* menghasilkan peningkatan akurasi dan performa.
2. Hasil akurasi dengan algoritma *Naïve Bayes* pada *Cryotherapy Dataset* menggunakan *Split Validation* menghasilkan nilai akurasi *confusion matrix* sebesar 91.67% dan AUC sebesar 0.975 setelah dilakukan optimasi menggunakan *Particle Swarm Optimization* untuk pembobotan atribut pada *Cryotherapy Dataset* serta *Split Validation* hasil akurasi dari *Naïve Bayes* meningkat menjadi sebesar 97.22% dan AUC sebesar 0.991. sehingga dapat disimpulkan bahwa *Particle Swarm Optimization* dapat meningkatkan akurasi dari algoritma *Naïve Bayes* pada *Cryotherapy Dataset*.

##### Saran

Penelitian ini hanya sampai pada eksperimen algoritma *Naïve Bayes* berbasis *Particle*

*Swarm Optimization* serta analisa terhadap hasil eksperimen tersebut tidak diimplementasikan menjadi sebuah aplikasi.

1. Pada penelitian selanjutnya dapat dilakukan eksperimen menggunakan algoritma klasifikasi serta algoritma optimasi lainnya untuk mengetahui performa dalam mengklasifikasi penyakit kutil pada *Cryotherapy Dataset*.
2. Pada penelitian berikutnya dapat dilakukan eksperimen menggunakan algoritma *Naïve Bayes* berbasis *Particle Swarm Optimization* pada dataset lain untuk mengetahui performa *Naïve Bayes* berbasis *Particle Swarm Optimization* dengan jumlah data yang beragam.

## Referensi

- Arifin, O., & Sasongko, T. B. (2018). Analisa perbandingan tingkat performansi metode support vector machine dan naïve bayes classifier. *Seminar Nasional Teknologi Informasi dan Multimedia 2018*, 6(1), 67–72.
- Arifin, T. (2017). Implementasi Algoritma PSO Dan Teknik Bagging Untuk Klasifikasi Sel Pap Smear. *Jurnal Informatika*, 4(2), 155–162. <https://doi.org/10.31311/JI.V4I2.2129>
- Arifin, T., & Ariesta, D. (2019). Prediksi Penyakit Ginjal Kronis Menggunakan Algoritma Naive Bayes Classifier Berbasis Particle Swarm Optimization. *Jurnal Tekno Insentif*, 13(1), 26–30. <https://doi.org/10.36787/jti.v13i1.97>
- Arifin, T., & Herliana, A. (2018). Optimasi Metode Klasifikasi Dengan Menggunakan Particle Swarm Optimization Untuk Identifikasi Penyakit Diabetes Retinopathy. *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika*, 4(2), 77–81. <https://doi.org/10.23917/khif.v4i2.6825>
- Arifin, T., & Syalwah, S. (2020). Prediksi keberhasilan immunotherapy pada penyakit kutil dengan menggunakan algoritma naïve bayes. *JURNAL RESPONSIF*, 2(1), 38–43.
- Brawijaya, H., Samudi, S., & Widodo, S. (2019). Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes Pada Pengobatan Penyakit Kutil Menggunakan Cryotherapy. *JUITA: Jurnal Informatika*, 7(2), 93–99. <https://doi.org/10.30595/juita.v7i2.5609>
- Cüvitoğlu, A., & Işık, Z. (2018). Evaluation Machine-Learning Approaches for Classification of Cryotherapy and Immunotherapy Datasets. *International Journal of Machine Learning and Computing*, 8(4), 331–335. <https://doi.org/10.18178/ijmlc.2018.8.4.707>
- Dawson, C. W. (2009). *Project in Computing and Information Systems A Student's Guide* (2 ed.). ADDISON-WESLEY.
- Gorunescu, F. (2011). *Data Mining Concepts, Models and Techniques* (P. J. Kacprzyk & P. L. C. Jain (ed.); 12 ed.). Springer.
- Hermawan, A., Sukma, A. R., & Halfis, R. (2019). Analisis Algoritma Klasifikasi C4.5 Untuk Memprediksi Keberhasilan Immunotherapy Pada Penyakit Kutil. *Jurnal Teknik Komputer AMIK BSI*, V(2), 155–160. <https://doi.org/10.31294/jtk.v4i2>
- Khozeimeh, F., Alizadehsani, R., Roshanzamir, M., Khosravi, A., Layegh, P., & Nahavandi, S. (2017). An expert system for selecting wart treatment method. *Computers in Biology and Medicine*, 81, 167–175. <https://doi.org/10.1016/j.combiomed.2017.01.001>
- Marlina, M., Aldi, Y., Putra, A. E., Sopianti, D. S., Hari, D. G., Arfiandi, A., Djamaan, A., & Rustini, R. (2016). Identifikasi Type Human Papillomavirus (HPV) pada Penderita Kanker Serviks. *Jurnal Sains Farmasi & Klinis*, 3(1), 54–63. <https://doi.org/10.29208/jsfk.2016.3.1.100>
- Muhamad, H., Prasojo, C. A., Sugianto, N. A., Surtiningsih, L., & Cholissodin, I. (2017). Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 4(3), 180–184. <https://doi.org/10.25126/jtiik.201743251>
- Nurelasari, E. (2018). Komparasi Algoritma Naive Bayes Dengan Support Vector Machine Berbasis Particle Swarm Optimization untuk Prediksi Kesuburan. *Bina Insani ICT Journal*, 5(1), 61–70.
- Rahayu, S., Nugraha, F. S., & Shidiq, M. J. (2019). Analisis Tingkat Keberhasilan Cryoterapy Menggunakan Neural Network. *Jurnal Pilar Nusa Mandiri*, 15(2), 141–148.

- 
- <https://doi.org/10.33480/pilar.v15i2.599>  
Siswa, T. A. Y., & Prihandoko. (2018). Analisis Penerapan Optimasi Perbandingan Kinerja Algoritma C4.5 dan Naïve Bayes Berbasis Particle Swarm Optimization ( PSO ) Untuk Mendeteksi Kanker Payudara. *Bangkit Indonesia*, 2(VII), 1–9.
- Sulaehani, R. (2016). Prediksi Keputusan Klien Telemarketing Untuk Deposito Pada Bank Menggunakan Algoritma Naive Bayes Berbasis Backward Elimination. *ILKOM Jurnal Ilmiah*, 8(3), 182–189.  
<https://doi.org/10.33096/ilkom.v8i3.83>.  
182-189
- Supriyatna, A., & Mustika, W. P. (2018). Komparasi Algoritma Naive bayes dan SVM Untuk Memprediksi Keberhasilan Imunoterapi Pada Penyakit Kutil. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 2(2), 152–161.  
<https://doi.org/10.30645/j-sakti.v2i2.78>
- Tanjung, J., Lumenta, A. S. M., & Sugiarso, B. A. (2016). Sistem Pakar Mendiagnosa Penyakit HPV (Human Papilloma Virus). *E-Journal Teknik Elektro Dan Komputer*, 5(4), 18–23.
- Vulandari, T. R. (2017). *DATA MINING TEORI DAN APLIKASI RAPIDMINER* (Cetakan I). PENERBIT GAVA MEDIA.
- Wijayatun, R., & Sulisty, Y. (2016). Prediksi Rating Film Menggunakan Metode Naive Bayes. *Jurnal Teknik Elektro*, 8(2), 60–63.