

Komparasi Algoritma *Naive bayes* dan SVM Untuk Memprediksi Keberhasilan Imunoterapi Pada Penyakit Kutil

Adi Supriyatna¹, Wida Prima Mustika²

¹AMIK BSI Karawang

Jl. Banten No.1 Karangpawitan, Karawang, 41351, Indonesia.

²STMIK Nusa Mandiri

Jl. Kramat Raya No18 Jakarta Pusat

adi.asp@bsi.ac.id, wida.wpm@nusamandiri.ac.id

Abstract

Warts is a skin health problem that is generally characterized by the appearance of small, rough-textured lumps on the skin surface caused by a virus that is human papilloma virus (HPV). One technique of treatment of wart disease is immunotherapy, this method is a treatment by boosting the immune system to overcome the disease of warts. Naive bayes and Support Vector Machine (SVM) is a method of data mining algorithm used to classify. The aim of this study was to compare the Naive bayes algorithm with Support Vector Machine (SVM) in predicting the success of immunotherapy treatment method in the treatment of wart disease. Tests conducted using the method of Naive bayes and Support Vector Machine (SVM) using the R programming language, then the results are used to do the comparison. The results of this study revealed that the Naive bayes method has superior prediction capability compared to Support Vector Machine (SVM) because Naive bayes can predict all class instances correctly with the accuracy level of 1.

Keywords: *Comparation, Immunotherapy, Naive bayes, Support Vector Machine, Warts.*

Abstrak

Kutil merupakan masalah kesehatan kulit yang pada umumnya ditandai dengan munculnya bejolan kecil bertekstur kasar pada permukaan kulit yang disebabkan oleh virus yaitu human papiloma virus (HPV). Salah satu teknik pengobatan penyakit kutil adalah imunoterapi, metode ini merupakan pengobatan dengan meningkatkan sistem kekebalan tubuh untuk mengatasi penyakit kutil. Naive bayes dan Support Vector Machine (SVM) merupakan metode algoritma data mining yang digunakan untuk melakukan klasifikasi. Penelitian ini dilakukan dengan tujuan untuk melakukan perbandingan antara algoritma Naive bayes dengan Support Vector Machine (SVM) dalam memprediksi keberhasilan metode pengobatan imunoterapi pada pengobatan penyakit kutil. Pengujian yang dilakukan menggunakan metode Naive bayes dan Support Vector Machine (SVM) menggunakan bahasa pemrograman R, kemudian hasilnya digunakan untuk melakukan perbandingan. Hasil dari penelitian ini dinyatakan bahwa metode Naive bayes memiliki kemampuan prediksi lebih unggul dibandingkan dengan Support Vector Machine (SVM) dikarenakan Naive bayes dapat melakukan prediksi terhadap semua instance class secara tepat dengan tingkat akurasi adalah 1.

Kata Kunci: *Imunoterapi, Kutil, Naive bayes, Perbandingan, Support Vector Machine.*

1. PENDAHULUAN

Penyakit kutil merupakan salah satu masalah kesehatan kulit yang biasanya ditandai dengan timbulnya bejolan kecil pada permukaan kulit, penyakit ini yang disebabkan oleh virus yaitu *human papiloma virus (HPV)*. Penularan virus penyebab kutil dapat terjadi dengan mudah, salah satunya hanya bersentuhan langsung dengan seseorang penderita kutil, namun tidak semua orang yang

bersentuhan dengan virus hpv akan menimbulkan penyakit kutil. Imunitas tubuh masing-masing orang sangat berpengaruh dalam penularan penyakit ini, penderita yang memiliki imunitas yang bermasalah akan rentan untuk terserang penyakit kutil. Salah satu teknik pengobatan penyakit kutil adalah imunoterapi, metode ini merupakan pengobatan dengan meningkatkan sistem kekebalan tubuh untuk mengatasi penyakit kutil. Imunoterapi intralesi adalah pengobatan efektif untuk kutil. Metode ini memiliki respons terapeutik yang lebih baik, membutuhkan lebih sedikit sesi, dan mampu mengobati kutil yang sudah lama[1].

Machine learning dan algoritma data mining digunakan untuk menganalisis dataset besar, menemukan dan mengekstrak pengetahuan dari data tersebut. Algoritma data mining dapat menganalisis data untuk menemukan pola yang tidak diketahui dalam database besar dari beberapa industri seperti perbankan, asuransi, kesehatan, dan pemasaran umumnya menerapkannya untuk mengurangi biaya, meningkatkan kualitas penelitian, dan meningkatkan jumlah penjualan”[2]. Penggunaan algoritma data mining dalam dunia kesehatan biasanya diterapkan dalam melakukan diagnosa berbagai penyakit. Prediksi klinis berkembang sangat pesat dengan mengadopsi ilmu komputer dan teknologi informasi dalam pengolahan data kesehatan[3].

Naive bayes dan *Support Vector Machine* merupakan metode algoritma data mining yang digunakan untuk melakukan klasifikasi. Penelitian ini dilakukan dengan tujuan untuk melakukan perbandingan antara algoritma *Naive bayes* dengan *Support Vector Machine* dalam memprediksi keberhasilan metode pengobatan imunoterapi pada pengobatan penyakit kutil. Berdasarkan penelitian sebelumnya dinyatakan bahwa Tingkat akurasi yang dihasilkan metode SVM tidak selalu lebih unggul dibandingkan dengan metode *Naive bayes*, begitu pula sebaliknya. Pada percobaan melakukan normalisasi data metode *Naive bayes* rata-rata akurasi lebih unggul dari SVM, namun akurasi tertinggi dengan menggunakan token N-Gram adalah metode SVM[4]. Sedangkan pada penelitian yang lain dinyatakan bahwa “performansi akurasi klasifikasi dalam melakukan diagnosis pasien kanker payudara, metode yang memiliki nilai akurasi tertinggi antara Regresi Logistik dengan *Support Vector Machine (SVM)* adalah SVM yaitu sebesar 94,34% sedangkan Regresi Logistik sebesar 84,90%[5]

2. METODOLOGI PENELITIAN

2.1. Data Mining.

Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari basis data yang besar[6]. *Data Mining* merupakan serangkaian proses mendapatkan pengetahuan atau pola dari kumpulan data, data mining memecahkan masalah dengan menganalisis data yang telah ada pada basis data[3][7].

2.2. Machine Learning.

Machine learning adalah bidang yang mempelajari pengembangan algoritma komputer untuk mengubah data menjadi aksi yang cerdas atau secara singkat dapat juga diartikan sebagai proses mengubah data menjadi informasi (*Machine Learning in Action*)[8].

2.3. Klasifikasi.

Menurut Faisal *Klasifikasi* merupakan salah satu tugas utama dalam *machine learning* dan *data mining* yang termasuk ke dalam tipe *supervised learning*. Istilah klasifikasi didapat dari tujuan utama teknik ini yaitu untuk memprediksi sebuah kategori dari masukkan data[8]. Terdapat beberapa metode yang umum digunakan untuk klasifikasi data mining antara lain Algoritma C.45, *Neural Network*, *Naives Bayes*, dan *Support Vector Machine (SVM)*.

2.4. Pemrograman R.

R bukan saja bahasa pemrograman tetapi juga lingkungan/Environment untuk komputasi statistik dan grafik. R merupakan projek GNU yang dikembangkan oleh Bell Laboratories oleh John Chamber dan teman-temannya. Menurut Faisal dalam bukunya yang berjudul Pengenalan Bahasa Pemrograman R, Pemrograman R Menyediakan berbagai macam *tool* statistik dari linier dan memodelan non linier, uji statistik klasik, analisis *time-series*, klasifikasi, clustering dan lain-lain. R juga menyediakan *tool* teknis grafis yang bertujuan untuk menampilkan data yang diolah secara visual dalam bentuk grafik[9]. R adalah bahasa pemrograman open source yang berhubungan dengan komputasi dan pengolahan data untuk Statistika dan yang berhubungan dengan penampilan grafik menggunakan tools yang disediakan oleh paket-paketnya yang sangat berguna di dalam penelitian dan industri. Versi awal dari R dibuat pada tahun 1992 di Universitas Auckland, New Zealand oleh Ross Ihaka dan Robert Gentleman[10].

2.5. Support Vector Machine (SVM)

Support Vector Machine merupakan metode yang termasuk kedalam *supervised learning* yang menganalisa data dan mengenali pola untuk klasifikasi[11]. *Support Vector Machine (SVM)* mengharuskan sebuah teks harus diubah menjadi vector sebelum digunakan untuk klasifikasi. *Key idea* dari SVM adalah untuk menemukan permukaan keputusan (*Hyperlane*) yang maksimal dari setiap titik data. Untuk melakukan training mesin yang didukung oleh vector atau biasa disebut *Support Vector Machine (SVM)* memerlukan solusi *Quadratic Programming (QP)* yang sangat besar. *Quadratic Programming* adalah masalah matematika untuk menemukan vector "x" yang meminimalkan fungsi kuadrat, dengan melakukan pembagian kelas menggunakan hyperplane maka masing-masing kelas positif, netral dan negatif dapat dibagi berdasarkan area masing-masing sehingga ketika terdapat data baru dapat ditentukan kelasnya berdasarkan area positif, netral maupun negatif[4].

2.6. Naive Bayes.

Naive bayes merupakan salah satu metode statistik untuk klasifikasi yang memungkinkan untuk menangkap ketidak pastian tentang suatu model dengan cara berprinsip pada mendefinisikan hasil probabilitas. Metode ini digunakan untuk menyelesaikan masalah diagnosa dan prediksi[12]. Teorema Bayes memiliki bentuk umum sebagai berikut[13]:

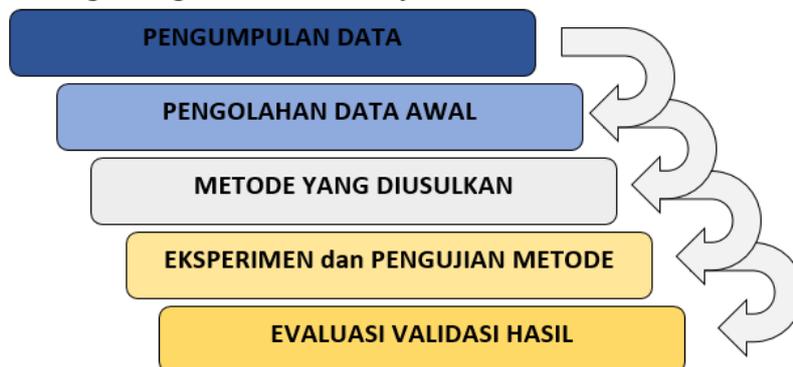
$$P(H | X) = \frac{P(X | H) P(H)}{P(X)} \quad (1)$$

Dimana :

- X = Data dengan class yang belum diketahui.
- H = Hipotesis Data X merupakan suatu class spesifik
- $P(H|X)$ = probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)
- $P(H)$ = Probabilitas hipotesis H (prior prob.)
- $P(X|H)$ = probabilitas X berdasarkan kondisi tersebut
- $P(X)$ = probabilitas dari X

2.7. Desain Penelitian

Berikut ini kerangka penelitian pada prediksi keberhasilan Imunoterapi pada penyakit kutil dengan Algoritma *Naive bayes* dan SVM:



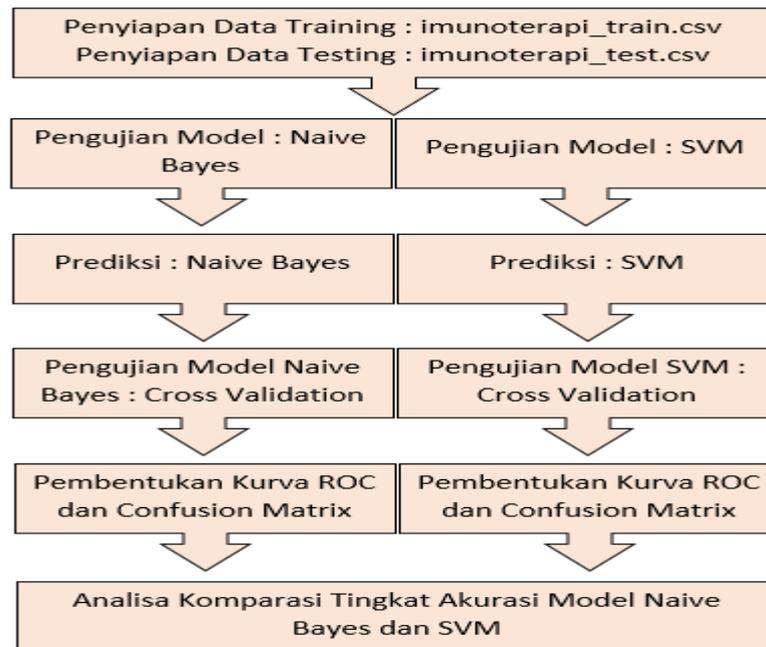
Gambar 1. Desain Penelitian

Pada tahapan awal dilakukan kegiatan pengumpulan data dengan cara melakukan pencarian dataset yang dapat digunakan untuk klasifikasi. Dataset yang digunakan dalam penelitian ini bersumber dari situs *UCI Machine Learning Repository* yaitu *Immunotherapy Dataset*. Dataset tersebut berisi data-data keberhasilan pengobahan imunoterapi dalam pengobatan penyakit kutil. Data tersebut terdiri dari 90 *record* dan tujuh variabel yaitu jenis kelamin, usia pasien, lama mengidap, banyak, jenis, besar dan kekerasan. Selanjutnya tahapan pengolahan data dengan mengidentifikasi atau menganalisa data yang terdapat didalam dataset imunoterapi. Kegiatan yang dikerjakan adalah melakukan *preprocessing data* dengan bahasa pemrograman R agar data yang digunakan untuk menguji metode adalah data yang sudah dalam keadaan baik.

3. HASIL DAN PEMBAHASAN

3.1. Analisa Penerapan *Naïve Bayes* dan SVM

Pada tahapan ini dilakukan percobaan atau pengujian metode algoritma *Naive bayes* dan *Support Vector Machine (SVM)* dengan menggunakan bahasa pemrograman R pada aplikasi Rstudio. Tahapan pengujian metode dapat dilihat pada gambar 2.



Gambar 2. Tahapan Pengujian Metode

a. Pengumpulan Data.

Data yang digunakan dalam penelitian ini merupakan data pengobatan penyakit kutil dengan imunoterapi, data terbagi menjadi 2 jenis yaitu data training merupakan data yang digunakan sebagai dasar atau acuan kemudian data tersebut digunakan untuk melakukan klasifikasi dan prediksi, sedangkan data testing merupakan data yang digunakan untuk mengukur sejauh mana classifier berhasil melakukan klasifikasi dengan benar. Data training dan testing pada penelitian ini ditunjukkan pada tabel 1 dan tabel 2.

Tabel 1. Data training imunoterapi.

Gender	Usia	Lama	Jumlah	Jenis	Besar (mm)	Kekerasan	Hasil
LAKI-LAKI	22	2.25	14	Lain-lain	51	50	Berhasil
LAKI-LAKI	15	3	2	Lain-lain	900	70	Berhasil
LAKI-LAKI	16	10.5	2	verruca vulgaris	100	25	Berhasil
LAKI-LAKI	27	4.5	9	Lain-lain	80	30	Berhasil
LAKI-LAKI	20	8	6	verruca vulgaris	45	8	Berhasil
LAKI-LAKI	15	5	3	Lain-lain	84	7	Berhasil
LAKI-LAKI	35	9.75	2	Plantar	8	6	Berhasil
WANITA	28	7.5	4	verruca vulgaris	9	2	Berhasil
WANITA	19	6	2	verruca vulgaris	225	8	Berhasil
WANITA	32	12	6	Lain-lain	35	5	Gagal

Gender	Usia	Lama	Jumlah	Jenis	Besar (mm)	Kekerasan	Hasil
WANITA	33	6.25	2	verruca vulgaris	30	3	Berhasil
WANITA	15	1.75	1	Plantar	49	7	Gagal
...
WANITA	23	6.75	6	verruca vulgaris	19	2	berhasil

Tabel 2. Data testing imunoterapi

Gender	Usia	Lama	Jumlah	Jenis	Besar (mm)	Kekerasan	Hasil
LAKI-LAKI	32	3	2	verruca vulgaris	50	5	Berhasil
WANITA	26	1	2	Plantar	101	10	Berhasil
WANITA	35	5	10	Lain-lain	80	6	Berhasil
LAKI-LAKI	40	10.5	5	verruca vulgaris	65	10	Gagal
WANITA	35	10	13	Plantar	30	45	Berhasil

b. Preprocessing Data.

Proses ini merupakan tahapan awal sebelum melakukan pengujian metode, dimana dataset yang akan digunakan diperiksa terlebih dahulu apakah terdapat *missing value*.

```
preprocessing.R x
1 #Menentukan Folder Lokasi Data Berada
2 setwd("G:/Rapid Miner/R")
3
4 #Membaca Data Yang Akan Di gunakan
5 data_training=read.csv("warts-train.csv")
6
7 #Menampilkan Data
8 view(data_training)
9
10 #Memeriksa apakah ada data yang missing
11 datanull=subset(data_training,is.na(data_training))
12
13 #Menampilkan data yang missing
14 view(datanull)
```

Gambar 3. Preprocessing data menggunakan Rstudio.

Gambar 3. menunjukkan script bahasa pemrograman R pada RStudio untuk melakukan *preprocessing data*. Hasil dari kompilasi script tersebut, dapat dilihat apakah pada dataset yang digunakan terdapat *missing value* yang dapat dilihat pada gambar 4.



Gambar 4. Hasil preprocessing data

Gambar 4. Menunjukkan bahwa tidak ada data yang mengandung *missing values*, berarti dataset training yang digunakan sebagai data training sudah dalam keadaan kondisi baik.

```
> #Menampilkan kesimpulan data
> summary(data_training)
```

Gender	Usia	Lama	Jumlah
LAKI-LAKI:41	Min. :15.00	Min. : 1.000	Min. : 1.000
WANITA :49	1st Qu.:20.25	1st Qu.: 5.000	1st Qu.: 2.000
	Median :28.50	Median : 7.750	Median : 6.000
	Mean :31.04	Mean : 7.231	Mean : 6.144
	3rd Qu.:41.75	3rd Qu.: 9.938	3rd Qu.: 8.750
	Max. :56.00	Max. :12.000	Max. :19.000
Lain-lain	Jenis	Besar..mm.	Kekerasan
Plantar :22	Min. : 6.00	Min. : 2.00	Hasil.terapi
verruca vulgaris:47	1st Qu.: 35.50	1st Qu.: 5.00	Berhasil:71
	Median : 53.00	Median : 7.00	Gagal :19
	Mean : 95.70	Mean :14.33	
	3rd Qu.: 80.75	3rd Qu.: 9.00	
	Max. :900.00	Max. :70.00	

Gambar 5. Summary Dataset Immunotherapy.

Sedangkan Gambar 5. Menampilkan kesimpulan dari dataset imunoterapi dimana terdapat informasi detail dari dataset training imunoterapi yang digunakan untuk melakukan proses pengujian metode.

c. Pengujian Metode.

1. Naive bayes.

```
Prediksi_Warts_NBR
1 library(c1071)
2 library(caret)
3 library(ROCR)
4
5 setwd("G:/RAPID MINER/R")
6
7 datatrain=read.csv("warts-train.csv")
8 datates=read.csv("warts-test.csv")
9
10 model=naiveBayes(Hasil~, datatrain)
11
12 predict_result = predict(model, datates[,-8])
13
14 comparation_result = cbind(prediction=as.character(predict_result), actual=as.character(datates[,8]))
15 print(comparation_result)
16
17 performance.value=confusionMatrix(predict_result, datates[,8])
18 print(performance.value)
19
20 roc.prediction=prediction(as.numeric(as.factor(predict_result)),as.numeric(as.factor(datates[,8])))
21
22 roc.tpr.fpr=performance(roc.prediction,"tpr","fpr")
23 roc.auc=performance(roc.prediction,"auc")
24 plot(roc.tpr.fpr,col="blue", lty=1)
25 abline(a=0, b=1)
26 print(paste("Luas AUC :", roc.auc@y.values))
27
```

Gambar 6. Pengujian Metode Naive bayes Menggunakan RStudio.

Gambar 6. Menunjukkan proses pengujian metode Naive bayes dengan menggunakan bahasa pemrograman R pada RStudio. Proses yang dilakukan pada script diatas dimulai dari memanggil library, menentukan folder penyimpanan data, membaca data, menentukan model Naive bayes, melakukan prediksi dengan, mengukur performance model sampai dengan menampilkan kurva ROC.

```
> comparation_result = cbind(prediction=as.character(predict_result), actual=as.character(datates[,8]))
> print(comparation_result)
  prediction actual
[1,] "Berhasil" "Berhasil"
[2,] "Berhasil" "Berhasil"
[3,] "Berhasil" "Berhasil"
[4,] "gagal"    "gagal"
[5,] "Berhasil" "Berhasil"
```

Gambar 7. Hasil Prediksi Naive bayes.

Gambar 7. menunjukkan prediksi yang dihasilkan dengan menggunakan metode Naive bayes, dari data tersebut dapat ditentukan bahwa prediksi dan actual menghasilkan nilai yang sama, artinya metode Naive bayes dapat memprediksi class untuk semua instance secara tepat.

```
> performance.value=confusionMatrix(predict_result, datates[,8])
> print(performance.value)
Confusion Matrix and Statistics

          Reference
Prediction Berhasil  gagal
Berhasil      4         0
Gagal         0         1

      Accuracy : 1
      95% CI   : (0.4782, 1)
    No Information Rate : 0.8
    P-value [Acc > NIR] : 0.3277

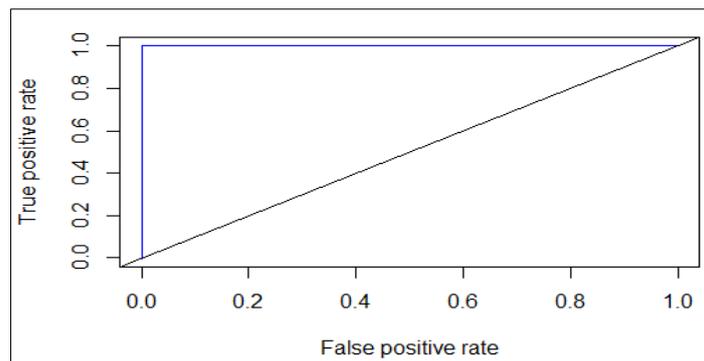
      Kappa : 1
  Mcnemar's Test P-value : NA

      Sensitivity : 1.0
      Specificity : 1.0
    Pos Pred Value : 1.0
    Neg Pred Value : 1.0
    Prevalence : 0.8
    Detection Rate : 0.8
    Detection Prevalence : 0.8
    Balanced Accuracy : 1.0

'Positive' Class : Berhasil
```

Gambar 8. Pengukuran Performance Metode *Naive bayes*.

Gambar 8. menunjukkan confusion matrix dan tingkat akurasi metode *Naive bayes* dengan bahasa pemrograman R pada RStudio, dari output tersebut dapat menyatakan bahwa berdasarkan confusion matrix prediksi yang dilakukan metode *Naive bayes* sudah sesuai dengan class actual, tingkat akurasinya adalah 1 dan class positif adalah berhasil.



Gambar 9. Kurva ROC Imunoterapi *Naive bayes*.

2. Support Vector Machine (SVM).

```
Prediksi_Warts_SVM.R
1 library(kernlab)
2 library(caret)
3 library(ROCR)
4
5
6 setwd("G:/RAPID MINER/R")
7
8 datatrain=read.csv("warts-train.csv")
9 datates=read.csv("warts-test.csv")
10
11 model=ksvm(Hasil--, datatrain)
12
13 predict_result = predict(model, datates[,-8])
14
15 comparation_result = cbind(prediction=as.character(predict_result), actual=as.character(datates[,8]))
16 print(comparation_result)
17
18 performance.value=confusionMatrix(predict_result, datates[,8])
19
20 print(performance.value)
21
22 roc.prediction=prediction(as.numeric(as.factor(predict_result)),as.numeric(as.factor(datates[,8])))
23
24 roc.tpr.fpr=performance(roc.prediction,"tpr","fpr")
25 roc.auc=performance(roc.prediction,"auc")
26 plot(roc.tpr.fpr,col="blue", lty=1)
27 abline(a=0, b=1)
28 print(paste("Luas AUC :", roc.auc$y.values))
```

Gambar 10. Pengujian Metode SVM Menggunakan RStudio.

Sama halnya seperti pengujian metode *Naive bayes*, Gambar 10. Menunjukkan pengujian metode SVM menggunakan bahasa pemrograman R. Proses pengujiannya sama yang membedakan adalah penggunaan metode pada baris 11.

```
> comparison_result = cbind(prediction=as.character(predict_result), actual=as.character(datates[,8]))
> print(comparison_result)
  prediction actual
[1,] "Berhasil" "Berhasil"
[2,] "Berhasil" "Berhasil"
[3,] "Berhasil" "Berhasil"
[4,] "gagal"    "Gagal"
[5,] "gagal"    "Berhasil"
```

Gambar 11. Hasil prediksi SVM.

Gambar 11. menunjukkan prediksi yang dihasilkan dengan menggunakan metode SVM, dari data tersebut dapat ditentukan bahwa prediksi dan actual menghasilkan nilai yang berberda, artinya metode SVM tidak dapat memprediksi class untuk semua instance secara tepat dikarenakan terdapat 1 instance yang diprediksi salah dan 1 instance lainnya yang diprediksi secara benar.

```
> performance.value=confusionMatrix(predict_result, datates[,8])
> print(performance.value)
Confusion Matrix and Statistics

      Reference
Prediction Berhasil Gagal
Berhasil    3      0
Gagal       1      1

      Accuracy : 0.8
      95% CI   : (0.2836, 0.9949)
      No Information Rate : 0.8
      P-Value [Acc > NIR] : 0.7373

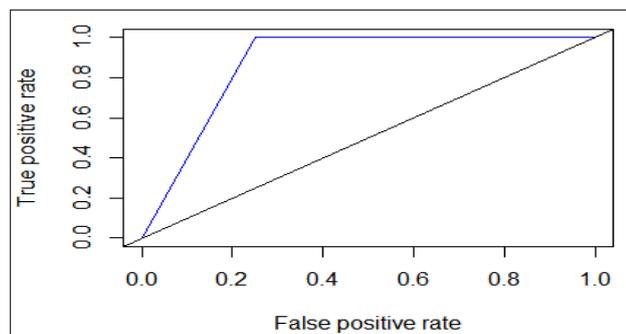
      Kappa : 0.5455
      Mcnemar's Test P-Value : 1.0000

      Sensitivity : 0.750
      Specificity : 1.000
      Pos Pred Value : 1.000
      Neg Pred Value : 0.500
      Prevalence : 0.800
      Detection Rate : 0.600
      Detection Prevalence : 0.600
      Balanced Accuracy : 0.875

      'Positive' Class : Berhasil
```

Gambar 12. Pengukuran Performance Metode SVM.

Gambar 12. menunjukkan confusion matrix dan tingkat akurasi metode SVM dengan bahasa pemrograman R pada RStudio, dari output tersebut dapat nyatakan bahwa berdasarkan confusion matrix prediksi yang dilakukan metode SVM terdapat 1 prediksi yang salah dimana hasil diprediksi gagal tetapi instance actual dinyatakan berhasil, tingkat akurasinya adalah 0.8 dan class positif adalah berhasil.



Gambar 13. Kurva ROC Imunoterapi dengan SVM.

4. SIMPULAN

Berdasarkan hasil pengujian metode *Naive bayes* dan *Support Vector Machine* (SVM) terhadap prediksi keberhasilan pengobatan imunoterapi untuk penyakit kutil menggunakan bahasa pemrograman R pada RStudio, maka dapat diambil kesimpulan bahwa hasil pengujian menggunakan dapat memprediksi class untuk semua instance secara tepat dengan nilai akurasi 1, sedangkan metode SVM masih terdapat prediksi yang tidak tepat dengan tingkat akurasi 0.8. Hasil pengujian tersebut menunjukkan bahwa metode *Naive bayes* merupakan metode yang lebih unggul dibandingkan dengan metode *Support Vector Machine* (SVM).

DAFTAR PUSTAKA

- [1] F. Khozeimeh *et al.*, "Intralesional immunotherapy compared to cryotherapy in the treatment of warts," *Int. J. Dermatol.*, vol. 56, no. 4, pp. 1-5, 2017.
- [2] F. Khozeimeh, R. Alizadehsani, M. Roshanzamir, A. Khosravi, P. Layegh, and S. Nahavandi, "An expert system for selecting wart treatment method," *Comput. Biol. Med.*, vol. 81, no. August 2016, pp. 167-175, 2017.
- [3] H. Amalia and E. Evicienna, "Komparasi metode Data Mining Untuk Penentuan Proses Persalinan Ibu Melahirkan," *J. Sist. Inf. (Journal Inf. Syst.*, vol. 13, pp. 103-109, 2017.
- [4] N. Saputra, T. B. Adji, and A. E. Permanasari, "Analisis Sentimen Data Presiden Jokowi dengan Preprocessing Normalisasi dan Stemming menggunakan Metode *Naive bayes* dan SVM," *J. Din. Inform.*, vol. 5, no. 1, 2015.
- [5] F. A. Novianti and S. W. Purnami, "Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan *Support Vector Machine* (SVM) Berdasarkan Hasil Mamografi," *J. Sains dan Seni ITS*, vol. 1, no. 1, pp. D147--D152, 2012.
- [6] M. Ridwan, H. Suyono, and M. Sarosa, "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma *Naive bayes* Classifier," *Eeccis*, vol. 7, no. 1, pp. 59-64, 2013.
- [7] K. Hastuti, "Analisis komparasi algoritma klasifikasi data mining untuk prediksi mahasiswa non aktif," *Semin. Nas. Teknol. Inf. Komun. Terap.*, vol. Juni, no. Semantik, pp. 241-249, 2012.
- [8] M. R. Faisal, *Seri Belajar Data Science: Klasifikasi dengan Bahasa Pemrograman R*. Banjarmasin: Indonesia .NET Developer Community, 2016.
- [9] M. R. Faisal, "Seri Belajar Pemrograman : Pengenalan Bahasa Pemrograman R," no. April. Indonesia .NET Developer Community, Banjarmasin, p. 147, 2016.
- [10] W. Budiharto and R. N. Rachmawati, *Pengantar Praktis Pemrograman R Untuk Ilmu Komputer*, 1st ed. Jakarta: Halaman Moeka, 2013.
- [11] S. Mrinalini, N. S. Abinayalakshmi, and C. Vinoth Kumar, "Wavelet feature based SVM and *NAIVE BAYES* classification of glaucomatous images using PCA and Gabor filter," *Proc. 10th Int. Conf. Intell. Syst. Control. ISCO 2016*, 2016.
- [12] G. Karthick, "Comparative Performance Analysis of *Naive bayes* and SVM classifier for Oral X-ray images," vol. 17, pp. 6-10.
- [13] A. Jananto, "Algoritma *Naive bayes* untuk Mencari Perkiraan Waktu Studi Mahasiswa," *Teknol. Inf. Din.*, vol. 18, no. 1, pp. 9-16, 2013.