

Pengembangan Aplikasi Berbasis Web dengan Rshiny untuk Data Klasifikasi Menggunakan Metode Naive Bayes

by Junior Macheda

Submission date: 26-Nov-2021 01:13PM (UTC+0700)

Submission ID: 1712958355

File name: naive_bayes_kolokium.pdf (600.52K)

Word count: 2540

Character count: 16143

Pengembangan Aplikasi Berbasis Web dengan Rshiny untuk Data Klasifikasi Menggunakan Metode *Naive Bayes*

Abstrak—Melonjaknya jumlah data menjelaskan bahwa data adalah salah satu hal yang akan terus berkembang, beriringan dengan berkembangnya teknologi. Di era saat ini, data membawa banyak sekali manfaat jika pemilikinya dapat menggunakannya dengan benar. Dalam setiap data pasti terdapat sebuah informasi yang berharga namun, untuk mengolah data tersebut butuh pengetahuan yang cukup mengenai *data science*. *Data science* mencakup berbagai macam ilmu pengetahuan, salah satunya adalah *Machine Learning*. Dalam pengolahan data, salah satu metode perhitungan data dalam *Machine Learning* adalah *Naive Bayes*, yang merupakan teknik *Supervised Learning*. *Naive Bayes* merupakan *classifier* terbaik dibandingkan *classifier* lainnya. Klasifikasi data merupakan hal yang cukup penting dalam pengolahan data, agar data lebih terstruktur. Penelitian ini mengembangkan sebuah aplikasi berbasis web dengan *R Shiny*, salah satu *framework* dari *R Studio*. Diharapkan aplikasi ini dapat berguna bagi para pengguna yang ingin mengklasifikasikan datanya menggunakan metode *Naive Bayes*.

Keywords—*Data Science, Machine Learning, Data Classification, Naive Bayes, R, R Shiny*

I. PENDAHULUAN

Perkembangan teknologi informasi yang tiada henti memaksa manusia harus cepat beradaptasi jika tidak ingin ketinggalan zaman. Teknologi Informasi (TI) memberikan banyak sekali manfaat di berbagai hal, terutama di bidang ilmu pengetahuan. Manusia sangat mudah menemukan informasi, pengetahuan, dan wawasan hanya dalam hitungan menit bahkan detik. Menurut B. Uno dan Nina Lamatenggo, “Teknologi informasi adalah suatu teknologi yang digunakan untuk mengolah data. Pengolahan itu termasuk memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk menghasilkan informasi yang berkualitas, yaitu informasi yang relevan, akurat, dan tepat waktu” [1]. TI mengandung ribuan bahkan jutaan data. Data tersebut diolah agar mendapatkan sebuah informasi yang diinginkan dan dapat dipahami oleh banyak orang. Data-data seperti inilah yang dikategorisasikan sebagai data-data ilmiah. Data ilmiah ini disebut

sebagai *data science*

Data science merupakan sebuah ilmu komputasi yang digunakan untuk memperoleh informasi berharga dari sebuah data. *Data science* memerlukan keterampilan dalam ilmu komputer, matematika dan statistika, dan individu yang mempunyai pengetahuan mendalam terhadap suatu bidang. *Data Science* merupakan ilmu mencakup berbagai macam ilmu pengetahuan. Salah satu ilmu tersebut adalah *Machine Learning*

Machine learning termasuk salah satu ilmu *artificial intelligence* yang mempunyai tujuan untuk mengembangkan sebuah mesin atau komputer agar dapat mempelajari sesuatu dengan sendirinya tanpa arahan dari pengguna. Terdapat 2 teknik dasar *Machine Learning* yaitu *Supervised* dan *Unsupervised*. *Supervised* adalah sebuah teknik pembelajaran mesin yang bersumber dari informasi yang sudah tersedia. Teknik ini menghasilkan informasi yang nantinya dapat dimanfaatkan guna memprediksi suatu kejadian yang akan terjadi di waktu yang akan datang. Sedangkan, *Unsupervised* tidak mempunyai informasi yang dijadikan acuan sebelumnya. *Unsupervised* membantu menemukan pola yang tersembunyi pada data [2].

Teknik dasar *Machine Learning* yang digunakan pada penelitian ini adalah *supervised learning*, dengan algoritma *Naive Bayes classifier*. *Naive Bayes classifier* menggunakan metode probabilitas statistik yang memprediksi probabilitas di masa yang akan datang, berdasarkan informasi di masa sebelumnya. *Naive Bayes classifier* merupakan model classifier terbaik Hal tersebut dibuktikan oleh D. Xhemali, C. J. Hinde, and R. G. Stone, dalam jurnal mereka yang berjudul “Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages” [3]. Pada jurnal tersebut dapat menyatakan *Naive Bayes* mendapatkan tingkat *accuracy* paling tinggi dibandingkan *classifier* lain. Hal tersebut menjadi salah satu pertimbangan penelitian ini mengaplikasikan metode *Naive Bayes* sebagai metode klasifikasinya.

Visualisasi data akan ditampilkan pada Shiny yang merupakan salah satu paket dalam aplikasi R untuk membangun aplikasi berbasis web. Shiny dapat memberikan kemudahan kepada penggunanya dalam pembuatan aplikasi berbasis web, terutama untuk menampilkan visualisasi data dengan halaman website yang interaktif. *Server* pada *Shiny* menggunakan pemrograman yang reaktif sehingga memungkinkan penggunaannya menggunakan logika, dan algoritma apapun dalam pengerjaannya [4]. Karena fleksibilitas dan kemudahan *Shiny* dalam pembuatan aplikasi berbasis web terutama dalam memvisualisasikan data, penggunaan *Shiny* untuk penelitian ini adalah pilihan terbaik demi mendapatkan hasil yang diharapkan.

10

II. TINJAUAN PUSTAKA

A. Naive Bayes Classifier

Naive Bayes adalah sebuah pengklasifikasi yang menggunakan metode statistik dan juga probabilitas, yaitu memprediksi probabilitas atau peluang di waktu yang akan datang berdasarkan data yang sudah ada [5]. *Naive Bayes* adalah pengklasifikasi terbaik dibandingkan dengan pengklasifikasi lainnya karena memiliki nilai akurasi terbaik.

B. Data Classification

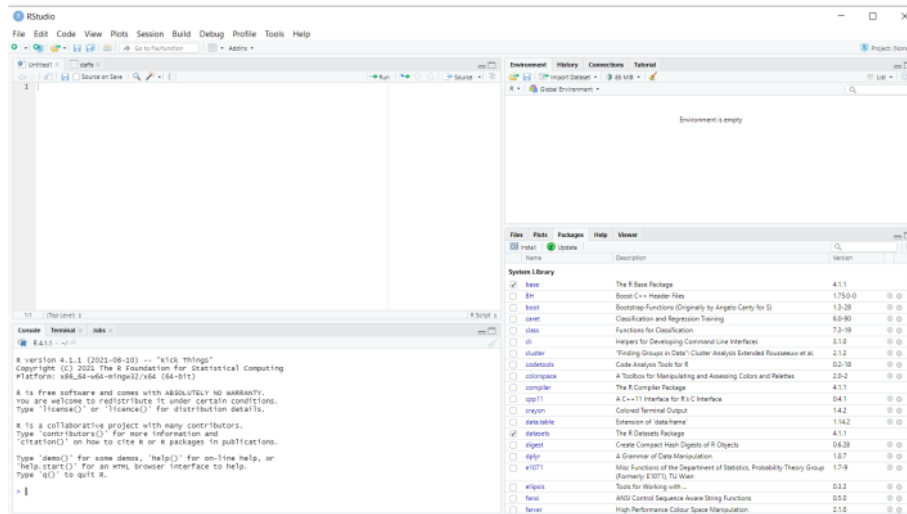
Data Classification atau klasifikasi data adalah sebuah metode pengelompokan berdasarkan karakteristik dari masing-masing objek. Klasifikasi mempunyai fungsi untuk mengklasifikasikan atau mengelompokkan data yang belum terstruktur menjadi kelas-kelas data yang menunjukkan isi dari data tersebut [6].

C. Bahasa Pemrograman R

R adalah bahasa pemrograman yang digunakan untuk komputasi, analisis data dan juga menampilkan grafik untuk statistika. R mempunyai banyak sekali *package* yang dapat dimanfaatkan sesuai dengan kebutuhan. R salah satu bahasa pemrograman yang interaktif sehingga pengguna dapat langsung melihat hasil perhitungannya [7].

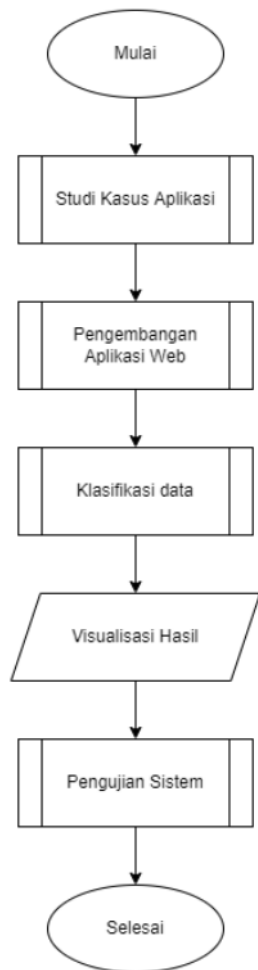
D. R Shiny

R Shiny adalah sebuah *tool* dari R yang dapat memfasilitasi penggunaannya dalam pembuatan *user interface* berbasis web. Shiny menggunakan bahasa pemrograman R yang memungkinkan penggunaannya dapat membuat web yang interaktif. Hal ini diharapkan dapat mempermudah user dalam mengoperasikannya



Gambar 1. Tampilan aplikasi R Studio

III. METODE PERANCANGAN SISTEM



Gambar 2. Flowchart perancangan sistem

A. Studi Kasus Aplikasi

Perancangan sistem ini diawali dengan melakukan studi kasus terhadap aplikasi serupa yang sudah digunakan oleh banyak orang. Aplikasi berbasis *Machine learning* dan sains data yang dijadikan pembandingan pada perancangan sistem ini adalah aplikasi “Waikato Environment for

Knowledge Analysis” atau yang lebih sering didengar dengan “WEKA” [8].

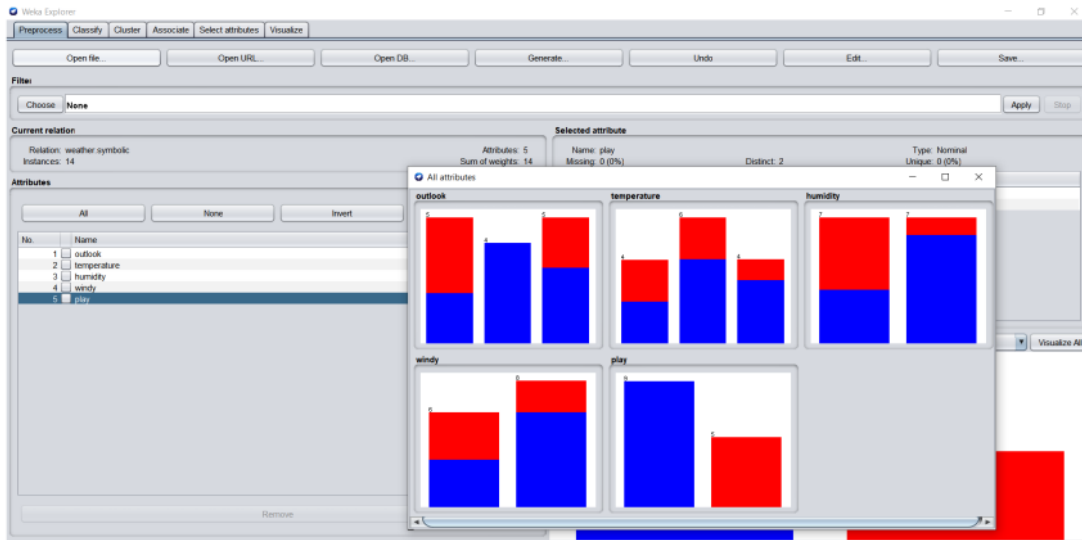


Gambar 3. Tampilan aplikasi WEKA

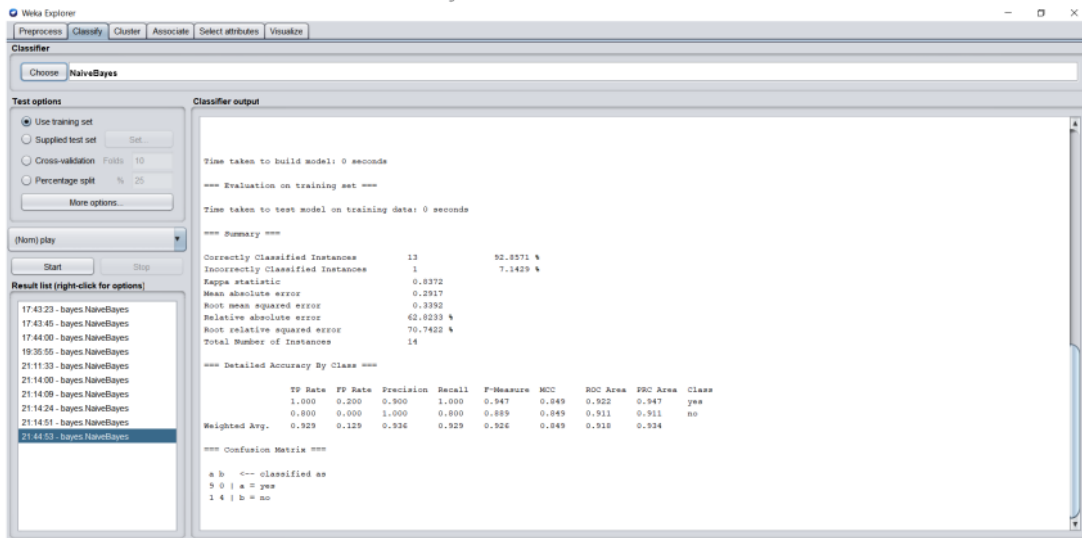
WEKA adalah sebuah *tools machine learning* berbasis Java, yang dapat memenuhi kebutuhan *data mining* dengan menggunakan berbagai macam algoritma *machine learning*. Penggunaanya dapat menggali informasi dari dataset yang mereka miliki dan mereka akan diberikan fleksibilitas dalam memilih antara algoritma *supervised* atau *unsupervised learning*. Pengaplikasian *supervised learning* dengan metode *naive bayes* dilakukan untuk menjadi bahan komparasi. Berikut adalah hasil dari penggunaan WEKA menggunakan metode *naive bayes* dengan menggunakan dataset yang telah disediakan disediakan oleh Aplikasi WEKA.

id	outlook	temperatura	humidity	windy	play
1	sunny	hot	high	FALSE	no
2	sunny	hot	high	TRUE	no
3	overcast	hot	high	FALSE	yes
4	rainy	mid	high	FALSE	yes
5	rainy	cool	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
7	overcast	cool	normal	TRUE	yes
8	sunny	mid	high	FALSE	no
9	sunny	cool	normal	FALSE	yes
10	rainy	mid	normal	FALSE	yes
11	sunny	mid	normal	TRUE	yes
12	overcast	mid	high	TRUE	yes
13	overcast	hot	normal	FALSE	yes
14	rainy	mid	high	TRUE	no

Gambar 4. Data weather nominal yang digunakan pada aplikasi WEKA



Gambar 5. Penjabaran variabel dari data weather:nominal.



Gambar 6. Hasil klasifikasi menggunakan aplikasi WEKA dengan metode naive bayes

B. Pengembangan Aplikasi

Pada tahap pengembangan aplikasi web, penelitian ini menggunakan salah satu paket dari R yaitu R Shiny. Pada proses pembuatannya, dibutuhkan 3 jenis script yaitu UI, server, dan shiny. Berikut adalah penjelasan dari masing-masing script.

a. UI

UI atau User Interface adalah script yang digunakan sebagai tampilan utama dari aplikasi ini. Keindahan tampilan web ini bergantung pada *code*

UI. Script ini menggunakan bahasa R dan juga HTML dalam pengkodeannya. Seperti pembuatan web pada umumnya, script ini memberikan kebebasan kepada penggunanya untuk mengatur user interface sesuai dengan kebutuhan dan kreativitas mereka.

b. Server

Server menjadi penghubung antara 2 script lainnya yaitu UI dan Bayes. Script ini membuat UI menjadi source yang harus ditampilkan.

c. Bayes

Bayes berisi seluruh code pemodelan Naive Bayes. Bisa dikatakan script ini menjadi penentu berhasil atau tidaknya perhitungan dan prediksi dari aplikasi ini. Jika code bayes berhasil, maka hasil dari pemodelan tersebut akan ditampilkan oleh UI yang dihubungkan oleh server. Penamaan script ini bebas sesuai dengan keinginan kita

C. Klasifikasi Data

Proses klasifikasi data merupakan hal yang paling utama pada penelitian ini, karena proses klasifikasi yang akan menentukan apakah aplikasi ini bisa bekerja atau tidak. Pada tahap ini akan dibuat code menggunakan bahasa pemrograman R untuk melakukan klasifikasi dengan algoritma Naive Bayes. Hasil dari klasifikasi tersebut yang akan dimanfaatkan oleh pengguna.

D. Visualisasi Data

Hasil dari data yang telah diklasifikasi akan divisualisasikan dengan menggunakan R Shiny. Beberapa informasi yang akan divisualisasikan mulai dari prediksi, accuracy, specificity, sensitivity dan sebagainya. Tahap ini adalah hal yang dibutuhkan banyak pengguna karena mereka dapat melihat hasil klasifikasi dari dataset yang mereka gunakan.

17 E. Pengujian Sistem

Pada tahap pengujian, sistem akan diuji coba apakah sistem dapat berjalan sesuai rencana, dan dapat memberikan hasil yang diharapkan atau tidak. Tahap ini harus dilakukan secara berulang kali demi mendapatkan hasil yang sempurna. Pengujian dimulai dari penggunaan berbagai macam dataset dengan format yang berbeda hingga hasil visualisasinya.

IV. IMPLEMENTASI SISTEM

]Tampilan Aplikasi

Tampilan pada aplikasi ini dirancang agar user friendly dan langsung pada intinya. Tampilan aplikasi menggambarkan judul, penjelasan mengenai

aplikasi dan navbar yang berisi upload dataset dan juga hasil modeling dari Naive Bayes.



Gambar 7. Tampilan dari aplikasi klasifikasi Naive Bayes

A. Dataset

Sistem akan menjabarkan informasi mengenai dataset yang digunakan pada aplikasi ini. Apa nama dataset yang digunakan, berapa jumlah datanya, apakah data tersebut numeric, logical atau character, apa saja atribut yang terdapat pada dataset tersebut dan lain sebagainya.

```
> glimpse(daffan)
#> # A tibble: 14 x 8
#>   outlook      temperature      humidity      windy      day      outlook_prob
#>   <chr>          <dbl>          <dbl>      <dbl>      <chr>      <dbl>
#> 1 sunny         85.1           86           1         sunny      0.33
#> 2 sunny         73           70           1         sunny      0.41
#> 3 sunny         81.5           75           1         sunny      0.37
#> 4 sunny         68.3           71           1         sunny      0.29
#> 5 overcast      63           71           1         overcast   0.09
#> 6 overcast      71           76           1         overcast   0.15
#> 7 overcast      69           80           1         overcast   0.12
#> 8 overcast      77           76           1         overcast   0.17
#> 9 rainy         68.3           81           1         rainy       0.21
#> 10 rainy         71           83           1         rainy       0.24
#> 11 rainy         81.5           76           0         rainy       0.27
#> 12 rainy         68.3           80           0         rainy       0.22
#> 13 sunny         81.5           71           0         sunny      0.37
#> 14 sunny         68.3           71           0         sunny      0.29
```

Gambar 8. Penjabaran informasi mengenai dataset yang digunakan

B. Perhitungan Naive Bayes

Perhitungan pada Naive Bayes melakukan splitting data atau membagi data menjadi data latih (data training) dan juga data testing. Pembagian tersebut bertujuan agar sistem dapat mempelajari dataset terlebih dahulu lalu akan dilakukan pengujian pada data tersebut untuk menentukan akurasi. Pada sistem ini splitting data dilakukan sebanyak 80%.

id	outlook	temperature	humidity	windy	play
5	rainy	cool	normal	FALSE	yes
4	rainy	mild	high	FALSE	yes
9	sunny	cool	normal	FALSE	yes
4.1	rainy	mild	high	FALSE	yes
13	overcast	hot	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
2	2 sunny	hot	high	TRUE	no
9.1	sunny	cool	normal	FALSE	yes
2.1	sunny	hot	high	TRUE	no
1	sunny	hot	high	FALSE	no
10	rainy	mild	normal	FALSE	yes

Gambar 9. Data training

id	outlook	temperature	humidity	windy	play
3	overcast	hot	high	FALSE	yes
7	overcast	cool	normal	TRUE	yes
8	sunny	mild	high	FALSE	no
11	sunny	mild	normal	TRUE	yes
12	overcast	mild	high	TRUE	yes
14	rainy	mild	high	TRUE	no

Gambar 10. Data testing

C. Visualisasi Naive Bayes

Hasil dari perhitungan *Naive Bayes* tersebut akan divisualisasikan agar pengguna dapat lebih mudah memahami hasil dari klasifikasi yang dilakukan pada *dataset* mereka. Beberapa informasi yang akan divisualisasikan mengenai pembagian kelas, akurasi, specificity, sensitivity, prediksi dan sebagainya.

```

confusion matrix and statistics

predict no yes
no 2 3
yes 0 2

Accuracy : 0.5714
ROC AUC : 0.1861, 0.8013
No Information Rate : 0.7143
P-value (acc > NRI) : 0.8917

Kappa : 0.2759
McNemar's Test P-value : 0.3482

Sensitivity : 1.0000
Specificity : 0.4000
Pos Pred Value : 0.4000
Neg Pred Value : 1.0000
Prevalence : 0.2857
Detection Rate : 0.2857
Detection Prevalence : 0.7143
Balanced Accuracy : 0.7000

"positive" class : no

```

Gambar 11. Hasil klasifikasi pada dataset yang digunakan. Visualisasi dalam bentuk *shiny* masih dalam tahap pengembangan

V. PEMBAHASAN

A. Dataset

Dataset yang digunakan pada sistem ini akan dijabarkan terlebih dahulu. Penjabaran ini mempunyai tujuan agar pengguna dapat lebih mudah memahami *dataset* yang mereka gunakan dan apakah *dataset* yang mereka input sudah sesuai atau belum.

Akan diberikan informasi mengenai nama *dataset* yang digunakan, jumlah kolom dan juga baris, apa saja tipe data dari *dataset* dan akan diberikan *summary* dari *dataset* tersebut. Untuk menjabarkan *dataset* ini digunakan *library* “*skimr*”, “*dplyr*”.

Library “*dplyr*” berfungsi untuk menampilkan dimensi dari *dataset* yang digunakan. Pada *library* ini digunakan perintah “*glimpse*” untuk menampilkan data secara vertika. Selanjutnya *library* “*skimr*”, berfungsi memberikan informasi dari masing-masing data variabel, dan menampilkan informasi mengenai data *value* yang hilang.

Terdapat beberapa syarat *dataset* agar dapat digunakan untuk melakukan klasifikasi menggunakan metode *Naive Bayes classifier*. Data yang akan diinput harus sudah siap untuk dianalisis. Data dikatakan dapat dianalisis jika data sudah rapi dan juga lengkap. Apakah data tersebut secara penulisan sudah sistematis dan apakah terdapat baris atau kolom yang belum terisi. Selanjutnya adalah memastikan data yang digunakan telah melewati proses *data cleaning*. *Data cleaning* adalah proses menghapus data-data yang terduplikat, penulisan data yang salah atau tipografi. Menghapus berbagai macam anomali yang mungkin masih dapat ditemukan pada data tersebut dengan dipenuhinya syarat tersebut, proses klasifikasi data akan jauh lebih mudah dan cepat.

B. Perhitungan Naive Bayes

Pada perhitungan *Naive Bayes*, dilakukan proses *data splitting* atau memisahkan *dataset* menjadi *Data* latih dan juga *Data Testing*.

Data latih digunakan untuk melatih metode *Naive Bayes*, untuk mempelajari *dataset* yang digunakan. Teorema Bayes digunakan pada saat data dipelajari. Berikut adalah rumus dari *Naive Bayes*.

$$P(X|Y) = \frac{P(Y|X) \times P(X)}{P(Y)}$$

- $P(X|Y)$ = Kemungkinan bersyarat. X terjadi jika Y benar
- $P(Y|X)$ = Kemungkinan bersyarat. Y terjadi jika X benar
- $P(X)$ = Kemungkinan X tanpa syarat

- $P(Y)$ = Kemungkinan Y tanpa syarat
- Kondisi X dan Y tidak boleh sama

Data testing digunakan untuk menguji apakah metode yang digunakan memberikan nilai atau akurasi yang baik dalam mempelajari *Data Training*. Hasil dari data testing tersebut yang nantinya akan divisualisasikan. Analoginya seperti seorang siswa diberikan latihan soal oleh gurunya, lalu di akhir akan diberikan ujian. Nilai ujian itulah yang akan menjadi nilai akurasi.

Jika nilai akurasi rendah terdapat beberapa kemungkinan, yaitu kesalahan pada *dataset* yang digunakan. Bisa jadi terjadi kesalahan pada struktur *dataset*, seperti adanya *data missing*, *data outlier*, atau *data imbalance* (data tidak seimbang). Kemungkinan kedua adalah kesalahan pada pemilihan metode *classifier*. Ada kemungkinan *dataset* tersebut tidak cocok digunakan untuk *classifier Naive Bayes*.

Package yang digunakan pada perhitungan *Naive Bayes* ini adalah *package* "e1071". *Package* "e1071" digunakan karena di dalamnya terdapat metode *Naive Bayes*.

C. Visualisasi Naive Bayes

Setelah dilakukan perhitungan dengan algoritma *Naive Bayes*, maka hasil dari perhitungan klasifikasi itu akan divisualisasikan melalui *framework* R Shiny.

Pada visualisasi *Naive Bayes*, *package* yang digunakan adalah "caret". "Caret" merupakan sebuah *package* yang digunakan untuk menentukan apakah metode *Naive Bayes* yang digunakan baik atau tidak dilihat dari nilai *accuracy*, *sensitivity* dan *specificity*. Informasi yang akan ditampilkan adalah *accuracy*, *precision*, *sensitivity*, *specificity*, *F1-Score*, *AUC*, dan *prediction*. Berikut adalah penjelasan dari masing-masing penilaian.

True Positive (TP) = Seberapa banyak data yang bernilai positif dan modelnya juga memprediksi positif.

True Negative (TN) = Seberapa banyak data yang bernilai negatif dan modelnya juga memprediksi negatif.

False Positive (FP) = Seberapa banyak data yang bernilai negatif dan modelnya juga memprediksi positif.

False Negative (FN) = Seberapa banyak data yang bernilai positif dan modelnya juga memprediksi negatif.

1. Accuracy

Accuracy adalah total keseluruhan dari seberapa sering model *Naive Bayes* bernilai benar. Rumus *Accuracy* adalah :

$$\frac{TP + TN}{Total}$$

2. Precision

Precision adalah seberapa sering sebuah prediksi bernilai benar jika model memprediksi positif. Rumus dari *precision* adalah :

$$\frac{TP}{FN + TP}$$

3. Sensitivity

Sensitivity merupakan sebuah kebenaran prediksi dari data seluruh data positif. Rumus dari *Sensitivity* adalah :

$$\frac{TP}{TP + FN}$$

4. Specificity

Specificity merupakan sebuah kebenaran prediksi dari data seluruh data negatif. Rumus dari *Specificity* adalah :

$$\frac{TN}{TN + FP}$$

5. F1-Score

F1-Score adalah perhitungan rata-rata dari *precision* dan juga *recall* atau *sensitivity*. Rumus dari *F1-Score* adalah :

$$2 \times \frac{Precision \times Sensitivity}{Precision + Sensitivity}$$

6. AUC

AUC atau *Area Under the Curve* adalah sebuah kurva yang menunjukkan *probability* dengan variabel *Sensitivity* dan variabel *Specificity*. Semakin tinggi nilai *AUC* maka semakin baik model klasifikasinya.

7. Prediction

Prediction akan menampilkan hasil prediksi dari berbagai perhitungan yang telah dilakukan oleh *Naive Bayes classifier*.

VI. KESIMPULAN

Berdasarkan penjelasan mengenai penelitian dan sistem, maka dapat disimpulkan bahwa :

1. Aplikasi *classifier* yang dikembangkan menggunakan bahasa pemrograman R dan *tool* R Shiny sebagai wadah untuk pembuatan *user interface*.
2. Metode klasifikasi yang digunakan pada aplikasi ini adalah algoritma *Naive Bayes*, karena merupakan *classifier* terbaik dibandingkan *classifier* lain.
3. Aplikasi yang berbasis web ini memberikan kemudahan pada pengguna dalam aksesnya karena tidak perlu mendownload aplikasi seperti penggunaan aplikasi WEKA pada studi kasus sebelumnya yang masih berbasis java.
4. Aplikasi ini bertujuan untuk menggali informasi dari dataset pengguna. Dengan menggunakan algoritma *Naive Bayes* diharapkan pengguna bisa memperoleh informasi dan prediksi dari dataset yang mereka miliki.

VII. DAFTAR PUSTAKA

- [1] H. B. Uno and N. Lamatenggo, *Teknologi Komunikasi dan Informasi Pembelajaran*, 1st ed. Jakarta: Bumi Aksara, 2010.
- [2] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.
- [3] D. Xhemali, C. J. Hinde, and R. G. Stone, "Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages," *Int. J. Comput. Sci.*, vol. 4, no. 1, pp. 16–23, 2009, [Online]. Available: <http://cogprints.org/6708/>.
- [4] A. T. Utomo, A. S. Ahmar, and M. K. Aidid, "Pengembangan Paket R untuk Analisis Time Series Dengan Graphical User Interface (GUI)," pp. 1–6, 2017.
- [5] D. Nofriansyah, K. Erwansyah, and M. Ramadhan, "Penerapan Data Mining dengan Algoritma Naive Bayes Clasifier untuk Mengetahui Minat Beli Pelanggan terhadap Kartu Internet XL (Studi Kasus di CV. Sumber Utama Telekomunikasi)," *J. Saindikom*, vol. 15, no. 2, pp. 81–92, 2016.
- [6] A. Indriani, "Klasifikasi Data Forum dengan menggunakan Metode Naive Bayes Classifier," *Semin. Nas. Apl. Teknol. Inf.*

Yogyakarta, vol. 1, no. 1, pp. 21–2014, 2014, [Online]. Available: www.bluefame.com.

- [7] D. Animasi, S. Berbasis, and I. G. A. A. Yudistira, "Pemrograman R," vol. 1, pp. 2–7, 2013.
- [8] D. Purnamasari, J. Henharta, Y. P. Sasmita, F. Ihsani, and I. W. S. Wicaksana, "Machine Learning 'Get Easy Using WEKA,'" *Dapur Buku*, pp. 1–40, 2013, [Online]. Available: www.DapurBuku.com.

Pengembangan Aplikasi Berbasis Web dengan Rshiny untuk Data Klasifikasi Menggunakan Metode Naive Bayes

ORIGINALITY REPORT

10%

SIMILARITY INDEX

10%

INTERNET SOURCES

4%

PUBLICATIONS

0%

STUDENT PAPERS

PRIMARY SOURCES

1	e-journal.uajy.ac.id Internet Source	1%
2	ojs.stmikpringsewu.ac.id Internet Source	1%
3	jurnal.upnyk.ac.id Internet Source	1%
4	eprints.umm.ac.id Internet Source	1%
5	repository.radenintan.ac.id Internet Source	1%
6	ejournal.upbatam.ac.id Internet Source	1%
7	strix.ciens.ucv.ve Internet Source	1%
8	ejournal.bsi.ac.id Internet Source	<1%
9	lib.unnes.ac.id Internet Source	<1%

10	ejurnal.itats.ac.id Internet Source	<1 %
11	123dok.com Internet Source	<1 %
12	etheses.uin-malang.ac.id Internet Source	<1 %
13	journal.upgris.ac.id Internet Source	<1 %
14	mujidinahmad.files.wordpress.com Internet Source	<1 %
15	dspace.uii.ac.id Internet Source	<1 %
16	ejournal.nusamandiri.ac.id Internet Source	<1 %
17	scholar.unand.ac.id Internet Source	<1 %
18	sistemasi.ftik.unisi.ac.id Internet Source	<1 %
19	www.ejurnal.stmik-budidarma.ac.id Internet Source	<1 %

Exclude quotes On

Exclude matches Off

Exclude bibliography On