

Penerapan Metode *Machine Learning* Dalam Memprediksi Keberhasilan Panggilan *Telemarketing* Menjual Produk Bank

Ahmad Fauzi¹, Fanny Fatma Wati², Indah Sulistyowati³, Muhammad Faittullah Akbar⁴,
Eka Rahmawati⁵, Ratna Kurnia Sari⁶

^{1,2,3,4}Prodi Sistem Informasi Akuntansi, Universitas Bina Sarana Informatika

^{5,6}Prodi Sistem Informasi, Universitas Bina Sarana Informatika

¹Ahmad.fzx@bsi.ac.id, ²fanny.ffw@bsi.ac.id, ³indahsulistyowati1409@gmail.com,

⁴Muhammad.mtl@bsi.ac.id, ⁵eka.eat@bsi.ac.id, ⁶ratna.rus@bsi.ac.id

Abstract: Competition between banks can be seen from the various attempts by banks to find customers through various marketing activities in order to get as many customers as possible. In the past, business actors offered goods or services to consumers in a face-to-face manner, now by utilizing existing and sophisticated technology, they can use long-distance communication tools such as telephone and fax, as well as other electronic media. To make it easier to manage customer data, a data calcification is needed. Machine Learning Algorithms can be used to predict or classify data. One of the algorithms in Machine Learning is the Naive Bayes method. Naive Bayes is a simple probabilistic classification that calculates a set of probabilities by summing the frequency and value combinations from a given dataset. This research will predict a successful Telemarketing call in selling Bank products to customers. The Naive Bayes algorithm and the Backward Elimination feature selection can increase the accuracy value in predicting the success of telemarketing in selling bank products well, as evidenced by the accuracy value generated by Naive Bayes of 83.04%, then after being applied with the selection of the backward elimination feature it increases by 6.41. % to 89.45%.

Keywords: Telemarketing, Machine Learning, Naive Bayes

Abstrak: Persaingan antar bank dapat dilihat dari berbagai upaya bank dalam mencari nasabah dengan berbagai kegiatan pemasaran agar mendapat nasabah sebanyak-banyaknya. Dahulu para pelaku usaha menawarkan barang atau jasa kepada konsumen dengan cara bertatap muka langsung, sekarang dengan memanfaatkan teknologi yang ada dan canggih bisa menggunakan alat komunikasi jarak jauh seperti telepon dan fax, serta media elektronik lainnya. Untuk mempermudah mengelola data nasabah maka dibutuhkan sebuah pengklasifikasian data. Algoritma *Machine Learning* dapat digunakan dalam memprediksi atau mengklasifikasikan sebuah data. Salah satu algoritma dalam *Machine Learning* adalah metode *Naive Bayes*. *Naive Bayes* merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Pada penelitian ini akan memprediksi sebuah keberhasilan panggilan *Telemarketing* dalam menjula produk Bank kepada para nasabah. Algoritma *Naive Bayes* dan seleksi fitur *Backward Elimination* mampu meningkatkan nilai akurasi dalam memprediksi keberhasilan *telemarketing* dalam menjual produk bank dengan baik, dibuktikan dengan nilai akurasi yang dihasilkan *naive bayes* sebesar 83,04 %, kemudian setelah diterapkan dengan seleksi fitur *backward elimination* meningkat sebesar 6,41% menjadi 89,45%.

Kata kunci: Telemarketing, Machine Learning, Naive Bayes



This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2019 by author and IJSE-Indonesian Journal on Software Engineering.

A. PENDAHULUAN

Bank merupakan peranan yang sangat penting dalam menggerakkan roda perekonomian nasional. Sebagaimana umumnya negara berkembang, sumber pembiayaan dunia usaha di Indonesia masih didominasi oleh penyaluran deposito perbankan yang diharapkan dapat mendorong pertumbuhan ekonomi (Marcos & Hidayah, 2014).

Sulitnya mengetahui keputusan calon nasabah, *telemarketing* untuk melakukan deposito pada bank, menyebabkan bank selalu menghadapi ancaman krisis keuangan. Oleh karena itu bank ditekan untuk meningkatkan persyaratan modal dengan menawarkan deposito jangka panjang pada masyarakat.

Untuk mengetahui keberhasilan *telemarketing* dalam menjual produk bank, dapat dilakukan dengan penerapan dan pemanfaatan teknik *Machine Learning*. Teknik *Machine Learning* merupakan suatu proses pendukung pengambilan keputusan dimana kita mencari pola informasi dalam data (Marcos & Hidayah, 2014).

Salah satu algoritma dalam *Machine Learning* adalah metode *Naive Bayes*. *Naive Bayes* merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Keuntungan penggunaan *Naive Bayes* yaitu bahwa metode ini hanya membutuhkan jumlah data pelatihan (*Training Data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian (Saleh, 2015).

Dengan menggunakan algoritma *Naive Bayes* dapat mengatasi masalah dalam mengambil keputusan klien mana yang akan dihubungi terlebih dahulu dan berminat deposito. Berdasarkan pemaparan latar belakang permasalahan, penulis memutuskan menggunakan dataset dengan judul *Bank Marketing Data Set* yang diunggah di *Portuguese banking institution Machine Learning Repository*. Maka dalam penelitian ini, data yang didapatkan akan diolah menggunakan teknik klasifikasi *data mining* algoritma *Naive Bayes*.

B. TINJAUAN PUSTAKA

1. Bank

Bank adalah bank umum sebagaimana dimaksud dalam Undang-Undang No 10 Tahun 1998 Pasal 1 ayat 2 Bank adalah badan usaha yang menghimpun dana dari masyarakat dalam bentuk simpanan dan menyalurkan kepada masyarakat dalam kredit atau bentuk-bentuk lainnya dalam rangka meningkatkan taraf hidup rakyat. Bank dikenal sebagai lembaga keuangan yang kegiatannya utamanya menerima simpanan giro, tabungan dan deposito. Kemudian bank juga dikenal sebagai tempat untuk menukar uang, memindahkan uang atau menerima segala bentuk pembayaran listrik, telepon, air, pajak, uang kuliah dan pembayaran lainnya (Memahami Bisnis Bank Syariah & Indonesia, 2014).

Bank merupakan peranan yang sangat penting dalam menggerakkan roda perekonomian nasional. Sebagaimana umumnya negara berkembang, sumber pembiayaan dunia usaha di Indonesia masih didominasi oleh penyaluran deposito perbankan yang diharapkan dapat mendorong pertumbuhan ekonomi (Marcos & Hidayah, 2014).

Dengan persaingan antar bank dapat dilihat dari berbagai upaya bank dalam mencari nasabah dengan berbagai kegiatan pemasaran agar mendapat nasabah sebanyak-banyaknya. Dahulu para pelaku usaha menawarkan barang atau jasa kepada konsumen dengan cara bertatap muka langsung, sekarang dengan memanfaatkan teknologi yang ada dan canggih bisa menggunakan alat komunikasi jarak jauh seperti telepon dan fax, serta media elektronik lainnya. Strategi *telemarketing* saat ini yang dilakukan oleh bank-bank untuk menawarkan produknya kepada konsumen (Silalahi, Prihatiningsih, P, Merugikan, & Bank, 2018) .

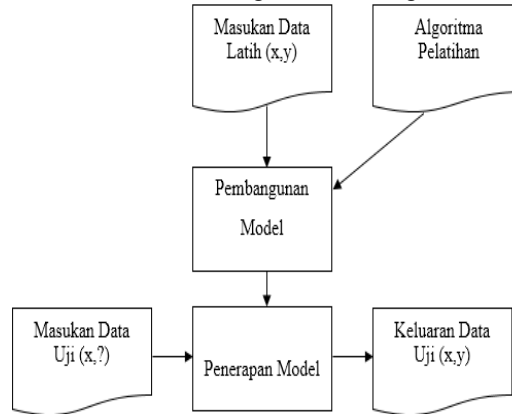
2. Telemarketing

Telemarketing adalah penggunaan telepon untuk menjual produk secara langsung kepada konsumen dan pelanggan bisnis. *Telemarketing* adalah penggunaan telepon dan pusat panggilan untuk menarik prospek, menjual produk kepada pelanggan, dan memberikan layanan kepada pelanggan (Monareh, Dh, & Nuralam, 2018). Dapat disimpulkan bahwa *Telemarketing* adalah penggunaan telepon dan pusat data untuk menjual produk secara langsung dan memberikan layanan kepada pelanggan dengan tujuan mendapatkan laba. Dengan pemanfaatan teknologi komunikasi melalui *Telemarketing* sebagai salah satu bentuk dominan dari pemasaran langsung

diharapkan dapat memberikan kemudahan kepada pelanggan untuk mengakses informasi yang dibutuhkan dalam tahapan Proses Keputusan Pembelian.

3. Klasifikasi

Menurut Muflikhah et al (2018:9) “Klasifikasi merupakan peranan dalam data yang menggunakan metode pendekatan prediktif, jika terdapat semumpulan *record* (*training set*) dimana setiap *record* terdiri dari sekumpulan atribut dan satu atribut merupakan kelas, digunakan untuk menentukan model dari atribut kelas sebagai suatu fungsi nilai dari atribut lain”.



Gambar 1. Proses Pekerjaan Klasifikasi
 Sumber: (Prasetyo, 2014)

Model yang sudah dibangun pada saat pelatihan kemudian dapat digunakan untuk memprediksi label kelas dari data baru yang belum diketahui label kelasnya.

4. Backward Elimination

Menurut Han, J. Kember dalam Hermawanti (2015:44) “*Backward Elimination* menghilangkan atribut-atribut yang tidak relevan. Sedangkan menurut Noori, dkk, dalam Hermawanti (2015:44) *Backward Elimination* didasarkan pada model regresi linear. Langkah-langkah *Backward Elimination* adalah:

- a. Mulai semua variabel pada model F-statistik parsial dihitung setiap variabel pada model
- b. Menentukan variabel dengan F-statistik parsial terkecil dan menguji F_{min} .
- c. Jika F_{min} tidak signifikan, dalam kasus ini, variabel dihilangkan dari model.
- d. Pada sisi lain, variabel dengan F-statistik terkecil adalah variabel indikator. Bagaimanapun, p-value diasosiasikan dengan F_{min} tidak cukup membenarkan model yang tidak inklusi (non-inclusion) menurut kriteria (lebih dari bit). Maka dari itu, prosedur menghasilkan dan melaporkan model sebagai berikut:

$$Y = \beta_0 + \beta_1(\text{single epithelial cell size}) + \beta_2(\text{normal nucleoli}) + \beta_3(\text{marginal adhesion}) + \epsilon$$
- e. Menghitung F-test parsial

5. Algoritma Naive Bayes

Menurut Wengking (2018:44) “*Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai *Teorema Bayes*”. Sedangkan menurut Saleh (2015:208) “*Naive Bayes* merupakan sebuah pengklasifikasian *probabilistic* sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan”.

Naive bayes dapat digunakan untuk berbagai macam keperluan antara lain untuk klasifikasi dokumen, deteksi spam atau filtering spam, dan masalah klasifikasi lainnya (Pratiwi, 2016). *The naive bayes classifier* bekerja sangat baik dibanding dengan *model classifier* lainnya. Hal ini dibuktikan oleh Xhemali, Hinde dan Stone dalam jurnalnya “*Naive bayes vs. Decision Trees vs. Neural Network in the Classification of Training Web Pages*” mengatakan bahwa “*Naive bayes classifier* memiliki tingkat akurasi yang lebih baik dibanding *model classifier* lainnya” (Wengking, 2018).

Untuk sebuah klasifikasi *Bayes* sederhana yang lebih dikenal sebagai *Naïve Bayesian Classifier* dapat diasumsikan bahwa efek dari suatu nilai atribut sebuah kelas yang diberikan adalah bebas dari atribut-atribut lain. (Fauzi, 2019).

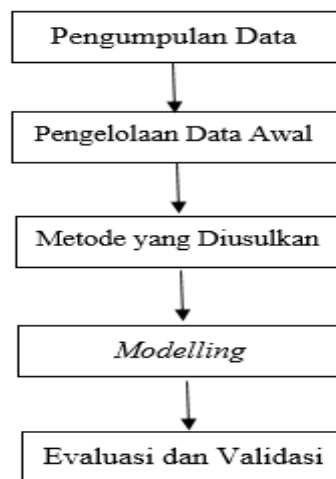
6. Data Mining

Data mining atau penambangan data dapat didefinisikan sebagai proses seleksi, eksplorasi, dan pemodelan dari sejumlah besar data untuk menemukan pola atau kecenderungan yang biasanya tidak disadari keberadaannya. *Data mining* dapat dikatakan sebagai proses mengekstrak pengetahuan dari sejumlah besar data yang tersedia. Pengetahuan yang dihasilkan dari proses data *mining* harus baru, mudah dimengerti, dan bermanfaat. Dalam data *mining*, data disimpan secara elektronik dan diproses secara otomatis oleh komputer menggunakan teknik dan perhitungan tertentu (Pramadhani & Setiadi, 2014).

C. METODE PENELITIAN

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah mengambil *dataset* yang diperoleh dari *Portuguese banking institution Machine Learning Repository* dengan judul *Bank Marketing Data Set*. *Dataset* disediakan oleh Sergio Moro, Paulo Cortez, dan Paulo Rita dari Institusi Bank Portugis, dan data tersebut disumbangkan untuk *UCI Repository* pada tahun 2012. *Dataset* meliputi 4521 instansi dan 17 atribut.

Sampel yang digunakan untuk pengujian terhadap model yang dihasilkan atau dapat dikatakan sebagai data *testing*, sedangkan data *training* berfungsi sebagai bahan pelatihan suatu model. *Bank marketing dataset* memiliki total *record* 4521, kemudian data tersebut akan dibagi menjadi dua bagian yaitu sebagai data *training* sejumlah 3165 dan data *testing* sejumlah 1356.



Gambar 2. Tahap Penelitian

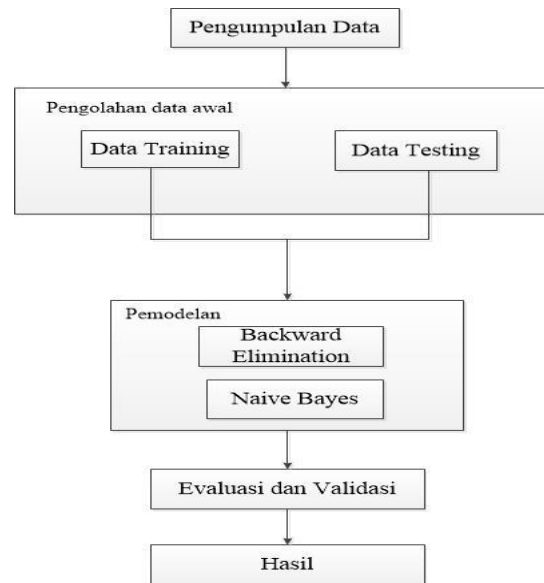
Berdasarkan hasil tahapan-tahapan gambar diatas, rincian penjelasan tahapan-tahapan yang akan disajikan dalam penelitian ini yaitu, sebagai berikut :

1. Pengumpulan data
Penulis mengumpulkan data sebagai bahan informasi yang harus dicari. Dalam penelitian ini penulis melakukan pengumpulan data melalui observasi dan data yang digunakan adalah data sekunder. kemudian mengintegrasikan semua data, termasuk variabel yang diperlukan dalam proses.
2. Pengolahan data awal
Pada tahap ini mencakup semua kegiatan untuk mempersiapkan data yang akan dimasukkan ke dalam alat pemodelan, dimana data tersebut merupakan pengolahan dari data mentah awal. Dalam melakukan pengolahan ini dilakukan beberapa tahapan sehingga pada akhirnya didapatkan data yang dapat digunakan pada tahap berikutnya.
3. Metode yang diusulkan
Pada tahap ini data dianalisis, dan setiap variabel-variabel yang berhubungan dikelompokkan. Setelah data selesai dianalisis selanjutnya diterapkan model yang sesuai

dengan jenis data. Data dibagi menjadi dua bagian yaitu, data latih (*training data*) dan data uji (*testing data*) sehingga dapat di proses kedalam model.

4. *Modelling*
Pada tahap ini dilakukan proses pengolahan data *mining* dengan menggunakan algoritma *naïve bayes*. Hal ini bertujuan untuk memperoleh nilai akurasi yang ingin dicapai dalam penerapan data *mining*.
5. Evaluasi dan validasi
Tahap evaluasi merupakan tahapan dimana dilakukan interpretasi terhadap hasil data mining yang telah dihasilkan pada tahap sebelumnya. Dalam proses ini dilakukan secara mendalam agar hasil pada tahapan sebelumnya sesuai dengan tujuan yang ingin dicapai

Dalam penelitian ini metode yang diusulkan adalah metode klasifikasi data mining algoritma *Naive Bayes*. Pengujian model menggunakan *split Data* dengan pembagian 70% data *training* dan 30% data *testing*. Hasil klasifikasi akan diukur menggunakan tingkat akurasi.



Gambar 3. Desain penelitian

D. HASIL DAN PEMBAHASAN

1. Hasil Eksperimen menggunakan algoritma *Naive Bayes*

Pada tahap ini adalah proses eksperimen tahap pertama yaitu pengujian model *Naive Bayes* menggunakan *Software Rapidminer. Validation*. Setelah dilakukan uji model maka diperoleh hasil seperti berikut:

Tabel 1. Hasil Akurasi Model Algoritma *Naive Bayes*

Model (Algoritma)	Akurasi
<i>Naives Bayes</i>	83,04%

	true no	true yes	class precision
pred. no	1069	99	91.52%
pred. yes	131	57	30.32%
class recall	89.08%	36.54%	

Gambar 4. Hasil *Confusion Matrix Naive Bayes*

Eksperimen ini menggunakan data sebanyak 4521 *record*. Berdasarkan *confusion matrix* terlihat bahwa 1069 *record* diprediksi *no* sebagai kelompok data *no* dan sebanyak 99 *record* diprediksi *no* sebagai kelompok data *yes*. Selanjutnya terlihat bahwa 131 *record* diprediksi *yes* sebagai kelompok data *no* dan sebanyak 57 *record* di diprediksi *yes* sebagai kelompok data *yes*.

2. Hasil Eksperimen menggunakan algoritma *Naive Bayes* dengan *Backward Elimination*

Pada tahap ini adalah proses eksperimen tahap kedua yaitu pengujian model *Naive Bayes* dengan *Backward Elimination* menggunakan *Software Rapidminer*. *Bank marketing dataset* dibagi menjadi beberapa ukuran pembagian data training dan data *testing*. *Dataset* yang telah disiapkan untuk implementasi proses uji model kemudian diuji pada algoritma *Naive Bayes* dengan *Backward Elimination* menggunakan metode *Split Validation*. Setelah dilakukan uji model maka diperoleh hasil seperti berikut:

Tabel 2. Hasil Akurasi Model Algoritma *Naive Bayes* dengan *Backward Elimination*

Model (Algoritma)	Akurasi
<i>Naive Bayes+Backward Elimination</i>	89,45%

	true no	true yes	class precision
pred. no	1170	113	91.19%
pred. yes	30	43	58.90%
class recall	97.50%	27.56%	

Gambar 5. Hasil *Confusion Matrix Naive Bayes* dengan *Backward Elimination*

Eksperimen ini menggunakan data sebanyak 4521 *record*. Berdasarkan *confusion matrix* terlihat bahwa 1170 *record* diprediksi *no* sebagai kelompok data *no* dan sebanyak 113 *record* diprediksi *no* sebagai kelompok data *yes*. Selanjutnya terlihat bahwa 30 *record* diprediksi *yes* sebagai kelompok data *yes* dan sebanyak 43 *record* di diprediksi *yes* sebagai kelompok data *no*.

3. Perbandingan Algoritma Naive Bayes dengan Algoritma Naive Bayes Menggunakan Backward Elimination

Berdasarkan eksperimen yang sudah dilakukan optimasi *Backward Elimination* mampu meningkatkan akurasi hingga 89,45% dari model *Naive Bayes* tanpa seleksi fitur yaitu 83,04%. Algoritma *Naive Bayes* menggunakan seleksi fitur *Backward Elimination* mampu meningkatkan akurasi sebesar 6,41% dari hasil yang diperoleh tanpa menggunakan seleksi *Backward Elimination*.

Tabel 3. Perbandingan Naive Bayes dengan Naive Bayes+Backward Elimination

No	Model (Algoritma)	Akurasi
1	<i>Naive Bayes</i>	83,04%
2	<i>Naive Bayes+Backward Elimination</i>	89,45%

4. Perhitungan Algoritma Naive Bayes

Pada penelitian ini kriteria atau atribut yang digunakan dalam memprediksi keberhasilan panggilan *telemarketing* dalam menjual produk bank meliputi *age, job, marital, education, default, balance, housing, loan, contact, day, month, campaign, pdays, previous, poutcome, dan y*. Nilai Kriteria ini akan diimplementasikan menggunakan algoritma *Naive Bayes*.

a. Perhitungan Manual Data Training

Tabel 4. Perhitungan Manual Data Training

age	job	marital	education	De- fault	Bal- ance	Hou- sing	loan	contact	day	month	Cam- paign	pdays	Pre- vious	poutcome	y
30	unemployed	married	primary	no	1787	no	no	cellular	19	oct	1	-1	0	unknown	no
33	services	married	secondary	no	4789	yes	yes	cellular	11	may	1	339	4	failure	no
35	management	single	tertiary	no	1350	yes	no	cellular	16	apr	1	330	1	failure	no
30	management	married	tertiary	no	1476	yes	yes	unknown	3	jun	4	-1	0	unknown	no
59	blue-collar	married	secondary	no	0	yes	no	unknown	5	may	1	-1	0	unknown	no
35	management	single	tertiary	no	747	no	no	cellular	23	feb	2	176	3	failure	no
36	self- employed	married	tertiary	no	307	yes	no	cellular	14	may	1	330	2	other	no
39	technician	married	secondary	no	147	yes	no	cellular	6	may	2	-1	0	unknown	no
41	entrepreneur	married	tertiary	no	221	yes	no	unknown	14	may	2	-1	0	unknown	no
43	services	married	primary	no	-88	yes	yes	cellular	17	apr	1	147	2	failure	no
39	services	married	secondary	no	9374	yes	no	unknown	20	may	1	-1	0	unknown	no
43	admin	married	secondary	no	264	yes	no	cellular	17	apr	2	-1	0	unknown	no
36	technician	married	tertiary	no	1109	no	no	cellular	13	aug	2	-1	0	unknown	no
20	student	single	secondary	no	502	no	no	cellular	30	apr	1	-1	0	unknown	yes
31	blue-collar	married	secondary	no	360	yes	yes	cellular	29	jan	1	241	1	failure	no
40	management	married	tertiary	no	194	no	yes	cellular	29	aug	2	-1	0	unknown	no
56	technician	married	secondary	no	4073	no	no	cellular	27	aug	5	-1	0	unknown	no
37	admin	single	tertiary	no	2317	yes	no	cellular	20	apr	1	152	2	failure	no
25	blue-collar	single	primary	no	-221	yes	no	unknown	23	may	1	-1	0	unknown	no
31	services	married	secondary	no	132	no	no	cellular	7	jul	1	152	1	other	no
38	management	divorced	unknown	no	0	yes	no	cellular	18	nov	2	-1	0	unknown	no
42	management	divorced	tertiary	no	16	no	no	cellular	19	nov	3	-1	0	unknown	no
44	services	single	secondary	no	106	no	no	unknown	12	jun	2	-1	0	unknown	no
44	entrepreneur	married	secondary	no	93	no	no	cellular	7	jul	2	-1	0	unknown	?
26	housemaid	married	tertiary	no	543	no	no	cellular	30	jan	3	-1	0	unknown	no

Pada perhitungan prediksi keberhasilan panggilan *telemarketing* dalam menjual produk bank data yang digunakan 4521 *record* data yang dibagi menjadi 70% data *training* dan 30% data *testing*. Sehingga untuk data *training* yang digunakan 3165 data dan data *testing* 1356. Berdasarkan pembagian data *training* apabila diberikan inputan baru menggunakan algoritma *Naive Bayes* maka langkah-langkahnya sebagai berikut:

b. Data Input

Tabel 5. Data Testing

Age	44
Job	Entrepreneur
Marital	Married
Education	Secondary
Default	No
Balance	93
Housing	No
Loan	No
Contact	Cellular
Day	7
Month	Jul
Campaign	2
Pdays	-1
Previous	0
Poutcome	unknown
Y	???

Pada data klasifikasi tersebut dengan $P(C_i)$ merupakan *class* target, kemudian akan ditentukan *class* atribut yang digunakan dengan ketentuan:

$C1 = (\text{Yes})$

$C2 = (\text{No})$

$X1 = (\text{Age}="44")$

$X2 = (\text{Job}="Entrepreneur")$

$X3 = (\text{Marital}="Married")$

$X4 = (\text{Education}="Secondary")$

$X5 = (\text{Default}="No")$

$X6 = (\text{Balance}="93")$

$X7 = (\text{Housing}="No")$

$X8 = (\text{Loan}="No")$

$X9 = (\text{Contact}="Cellular")$

$X10 = (\text{Day}="7")$

$X11 = (\text{Month}="Jul")$

$X12 = (\text{Campaign}="2")$

$X13 = (\text{Pdays}="-1")$

$X14 = (\text{Previous}="0")$

$X15 = (\text{Previous}="0")$

c. Mengitung jumlah probabilitas prior untuk *class*/label pertama

$$1. P(C1) = P(\text{Yes}) = \frac{366}{3165} = 0.11563981$$

$$2. P(C2) = P(\text{No}) = \frac{2799}{3165} = 0.88436019$$

d. Menghitung probabilitas bersyarat untuk setiap kelas $P(X|C_i)$, $i = \text{Yes, No}$ dan untuk setiap atribut pada sampel data masukan.

$$1. P(X_1|C_1) = P(\text{Age}="44"|\text{Yes}) = \frac{8}{366} = 0.021857923$$

$$2. P(X_1|C_2) = P(\text{Age}="44"|\text{No}) = \frac{64}{2799} = 0.022865309$$

$$3. P(X_2|C_1) = P(\text{Job}="Entrepreneur"|\text{Yes}) = \frac{11}{366} = 0.030054645$$

$$4. P(X_2|C_2) = P(\text{Job}="Entrepreneur"|\text{No}) = \frac{112}{2799} = 0.040014291$$

$$5. P(X_3|C_1) = P(\text{Marital}="Married"|\text{Yes}) = \frac{193}{366} = 0.527322404$$

$$6. P(X_3|C_2) = P(\text{Marital}="Married"|\text{No}) = \frac{1765}{2799} = 0.630582351$$

7. $P(X_4|C_1) = P(\text{Education}=\text{"Secondary"}|\text{"Yes"}) = \frac{174}{366} = 0.475409836$
8. $P(X_4|C_2) = P(\text{Education}=\text{"Secondary"}|\text{"No"}) = \frac{1428}{2799} = 0.510182208$
9. $P(X_5|C_1) = P(\text{Default}=\text{"No"}|\text{"Yes"}) = \frac{361}{366} = 0.986338798$
10. $P(X_5|C_2) = P(\text{Default}=\text{"No"}|\text{"No"}) = \frac{2753}{2799} = 0.98356555$
11. $P(X_6|C_1) = P(\text{Balance}=\text{"93"}|\text{"Yes"}) = \frac{0}{366} = 0$
12. $P(X_6|C_2) = P(\text{Balance}=\text{"93"}|\text{"No"}) = \frac{4}{2799} = 0.001429082$
13. $P(X_7|C_1) = P(\text{Housing}=\text{"No"}|\text{"Yes"}) = \frac{210}{366} = 0.573770492$
14. $P(X_7|C_2) = P(\text{Housing}=\text{"No"}|\text{"No"}) = \frac{1167}{2799} = 0.41693462$
15. $P(X_8|C_1) = P(\text{Loan}=\text{"No"}|\text{"Yes"}) = \frac{334}{366} = 0.912568306$
16. $P(X_8|C_2) = P(\text{Loan}=\text{"No"}|\text{"No"}) = \frac{2339}{2799} = 0.835655591$
17. $P(X_9|C_1) = P(\text{Contact}=\text{"Cellular"}|\text{"Yes"}) = \frac{297}{366} = 0.81147541$
18. $P(X_9|C_2) = P(\text{Contact}=\text{"Cellular"}|\text{"No"}) = \frac{1766}{2799} = 0.630939621$
19. $P(X_{10}|C_1) = P(\text{Day}=\text{"7"}|\text{"Yes"}) = \frac{7}{366} = 0.019125683$
20. $P(X_{10}|C_2) = P(\text{Day}=\text{"7"}|\text{"No"}) = \frac{124}{2799} = 0.44301536$
21. $P(X_{11}|C_1) = P(\text{Month}=\text{"Jul"}|\text{"Yes"}) = \frac{44}{366} = 0.120218579$
22. $P(X_{11}|C_2) = P(\text{Month}=\text{"Jul"}|\text{"No"}) = \frac{465}{2799} = 0.166130761$
23. $P(X_{12}|C_1) = P(\text{Campaign}=\text{"2"}|\text{"Yes"}) = \frac{101}{366} = 0.275956284$
24. $P(X_{12}|C_2) = P(\text{Campaign}=\text{"2"}|\text{"No"}) = \frac{101}{2799} = 0.036084316$
25. $P(X_{13}|C_1) = P(\text{Pdays}=\text{"-1"}|\text{"Yes"}) = \frac{239}{366} = 0.653005464$
26. $P(X_{13}|C_2) = P(\text{Pdays}=\text{"-1"}|\text{"No"}) = \frac{2368}{2799} = 0.846016434$
27. $P(X_{14}|C_1) = P(\text{Previous}=\text{"0"}|\text{"Yes"}) = \frac{239}{366} = 0.653005464$
28. $P(X_{14}|C_2) = P(\text{Previous}=\text{"0"}|\text{"No"}) = \frac{2368}{2799} = 0.846016434$
29. $P(X_{15}|C_1) = P(\text{Poutcome}=\text{"Unknown"}|\text{"Yes"}) = \frac{236}{366} = 0.653005464$
30. $P(X_{15}|C_2) = P(\text{Poutcome}=\text{"Unknown"}|\text{"No"}) = \frac{2368}{2799} = 0.846016434$

c. Kalikan probabilitas $P(X|C_i)P(C_i)$ untuk setiap kelas C_1 dan C_2

1. $P(X|\text{"Yes"}) P(\text{"Yes"})$
 $= P(\text{Age}=\text{"44"}|\text{"Yes"}) * P(\text{Job}=\text{"Entrepreneur"}|\text{"Yes"}) * P(\text{Marital}=\text{"Married"}|\text{"Yes"}) * P(\text{Education}=\text{"Secondary"}|\text{"Yes"}) * P(\text{Default}=\text{"No"}|\text{"Yes"}) * P(\text{Balance}=\text{"93"}|\text{"Yes"}) * P(\text{Housing}=\text{"No"}|\text{"Yes"}) * P(\text{Loan}=\text{"No"}|\text{"Yes"}) * P(\text{Contact}=\text{"Cellular"}|\text{"Yes"}) * P(\text{Day}=\text{"7"}|\text{"Yes"}) * P(\text{Month}=\text{"Jul"}|\text{"Yes"}) * P(\text{Campaign}=\text{"2"}|\text{"Yes"}) * P(\text{Pdays}=\text{"-1"}|\text{"Yes"}) * P(\text{Previous}=\text{"0"}|\text{"Yes"}) * P(\text{Poutcome}=\text{"Unknown"}|\text{"Yes"})$
 $= 0.021857923 * 0.030054645 * 0.527322404 * 0.475409836 * 0.986338798 * 0 * 0.573770492 * 0.912568306 * 0.81147541 * 0.019125683 * 0.120218579 * 0.275956284 * 0.653005464 * 0.653005464 * 0.653005464 = 0$
2. $P(X|\text{"No"}) P(\text{"No"})$
 $= P(\text{Age}=\text{"44"}|\text{"No"}) * P(\text{Job}=\text{"Entrepreneur"}|\text{"No"}) * P(\text{Marital}=\text{"Married"}|\text{"No"}) * P(\text{Education}=\text{"Secondary"}|\text{"No"}) * P(\text{Default}=\text{"No"}|\text{"No"}) * P(\text{Balance}=\text{"93"}|\text{"No"}) * P(\text{Housing}=\text{"No"}|\text{"No"}) * P(\text{Loan}=\text{"No"}|\text{"No"}) * P(\text{Contact}=\text{"Cellular"}|\text{"No"}) * P(\text{Day}=\text{"7"}|\text{"No"}) * P(\text{Month}=\text{"Jul"}|\text{"No"}) * P(\text{Campaign}=\text{"2"}|\text{"No"}) * P(\text{Pdays}=\text{"-1"}|\text{"No"}) * P(\text{Previous}=\text{"0"}|\text{"No"}) * P(\text{Poutcome}=\text{"Unknown"}|\text{"No"})$
 $= 0.022865309 * 0.040014291 * 0.630582351 * 0.510182208 * 0.983565559 * 0.001429082 * 0.41693462 * 0.835655591 * 0.630939621 * 0.44301536 * 0.166130761 * 0.036084316 * 0.846016434 * 0.846016434 = 1.4626E-11$

Bandingkan hasil probabilitas posterior *class Yes dan No* $P(X|C_1)P(C_1)$

Dari hasil perhitungan *class* diatas yang mempunyai nilai probabilitas tertinggi yaitu *class* $P|No$ sehingga dapat disimpulkan bahwa kategori prediksi keberhasilan *telemarketing* dalam

menjual produk bank tersebut termasuk *class* “NO”. Perbandingan ketiga *class* dapat dilihat pada tabel dibawah ini:

Tabel 6. Hasil Iterasi

CLASS	HASIL
YES	0
NO	0,00000014626

E. KESIMPULAN

1. Algoritma *naïve bayes* dengan seleksi fitur *backward elimination* dapat memprediksi keberhasilan *telemarketing* dalam menjual produk bank.
2. Algoritma *Naive Bayes* dan seleksi fitur *Backward Elimination* mampu meningkatkan nilai akurasi dalam memprediksi keberhasilan *telemarketing* dalam menjual produk bank dengan baik, dibuktikan dengan nilai akurasi yang dihasilkan *naive bayes* sebesar 83,04 %, kemudian setelah diterapkan dengan seleksi fitur *backward elimination* meningkat sebesar 6,41% menjadi 89,45%.

REFERENSI

- Hermawanti, L. (2015). Mendiagnosis, Untuk Kanker, Penyakit, 11, 42–45.
- Indonesia, I. B. (2014). *Memahami Bisnis Bank Syariah*. Gramedia Pustaka Utama. Retrieved from https://books.google.co.id/books?id=jqFLDwAAQBAJ&dq=memahami+bisnis+bank+syaria h&hl=id&source=gbs_navlinks_s
- Monareh, J. A., Dh, A. F., & Nuralam, I. P. (2018). PENGARUH TELEMARKETING TERHADAP KEPUTUSAN PEMBELIAN (Survei Online pada Pelanggan Produk Multiguna Astra Credit Companies Priority di PT Astra Sedaya Finance). *Jurnal Administrasi Bisnis*, 58(2).
- Moro, S., Cortez, P., & Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62, 22–31. <https://doi.org/10.1016/j.dss.2014.03.001>
- Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*. (Nikodemus WK, Ed.) (1st ed.). Yogyakarta: C.V ANDI OFFSET.
- Pratiwi, H. (2016). *Buku Ajar Sistem Pendukung Keputusan*. Retrieved from <https://books.google.co.id/books?id=8HB5DwAAQBAJ&dq=buku+ajar+sistem+pendukung +keputusan&hl=id&sa=X&ved=0ahUKEwjBjer02dbjAhUyhuYKHxEWBwIQ6AEILTAB>
- Saleh, A. (2015). Implementasi Metode Klasifikasi Naïve Bayes dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *Citec Journal*, 2(3), 207–217.
- Fauzi, A. (2019). *Sentimen Analisis Berinternet Pada Media Sosial dengan Menggunakan Algoritma Bayes*. *Jurnal Informatika*. <https://doi.org/10.24076/CITEC.2015V2I3.49> (Fauzi, 2019)