

# PENERAPAN ALGORITMA *NAÏVE BAYES* UNTUK MEMPREDIKSI KEPUTUSAN CALON NASABAH DAN NASABAH TETAP BANK BRI SYARIAH MENERIMA PENAWARAN PROGRAM DEPOSITO BERJANGKA

Wahyu Nurjaya WK<sup>1</sup>, Yusrina Adani<sup>2</sup>

Program Studi Sistem Informasi, STMIK LPKIA Bandung

Jl. Soekarno Hatta No.456 Kota Bandung, 40266 Tlp. (022) 7564283, Fax.(022) 7564282

Email: <sup>1</sup>WahyuNWK@lpkia.ac.id, <sup>2</sup>yusrina@fellow.lpkia.ac.id

---

## ABSTRAK

Bank BRI Syariah memiliki banyak produk yang menarik untuk ditawarkan kepada calon nasabah maupun nasabah tetap berupa produk jangka panjang atau jangka pendek, yang menawarkan banyak keuntungan bagi nasabah itu sendiri. Salah satu produknya adalah Deposito berjangka yang merupakan produk investasi dengan menyimpan uang dan penarikannya hanya bisa dilakukan pada kurun waktu tertentu yang telah di janjikan oleh pihak bank dengan persetujuan nasabah. Dengan *telemarketing* yang baik oleh pihak bank maka diharapkan calon nasabah dan nasabah tetap mengetahui produk ini.

*Telemarketing* adalah salah satu cara dalam mempromosikan produk-produk atau jasa layanan yang ada di bank. Seorang *telemarketing* bank harus dapat membuat target nasabah, nasabah mana yang berpotensi untuk meningkatkan deposito dengan melihat data-data nasabah bank yang telah tersimpan dalam *database*. Dikarenakan *database* nasabah sangat besar, maka tidak mungkin untuk mencari pola prediksi calon nasabah atau nasabah tetap yang berminat untuk program Deposito dengan cara konvensional.

Berdasarkan hal tersebut, pengelolaan data yang sangat besar bisa diatasi dengan memanfaatkan *Data Mining* yaitu proses iteratif dan interaktif untuk menentukan pola atau model baru yang sempurna, bermanfaat dan dapat dimengerti dalam suatu *database* yang sangat besar. *Data Mining* berisi pencarian *trend* pola yang diinginkan dalam *database* besar untuk membantu pengambilan keputusan diwaktu yang akan datang. Dengan menggunakan *Data Mining* diharapkan dapat mengoptimasikan proses prediksi data nasabah oleh seorang *telemarketing*, sehingga dia mampu menawarkan deposito dengan target calon nasabah atau nasabah tetap yang tepat sasaran. Adapun Teknik Klasifikasi *Data Mining* menggunakan algoritma *Naïve Bayes*. *Naïve Bayes* bekerja sangat efektif saat diuji pada *dataset* yang besar untuk menentukan pola dimasa lalu dan mencari fungsi yang akan menjadi pola penilaian data dimasa yang akan datang. Untuk mencapai hasil yang diharapkan metode *CRISP-DM* (*Cross Industry Standard Process for Data Mining*) sangat cocok sebagai solusi, melalui proses *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation* dan *deployment*. Dengan ini hasil prediksi akan lebih akurat, sehingga untuk target *telemarketing* produk Deposito Bank BRI Syariah akan tepat sasaran.

**Kata Kunci :** *Data Mining, Deposito, Telemarketing, Algoritma Naive Baye, CRISP-DM*

---

## 1. PENDAHULUAN

Bank BRI Syariah salah satu diantara tiga bank syariah terbesar di Indonesia merupakan badan usaha yang mengelola dana dari nasabah yang mempunyai program atau produk-produk layanan yang banyak untuk ditawarkan kepada calon nasabah dan nasabah bank tersebut berdasarkan prinsip syariah islam. Dengan persaingan berbagai bank yang ada di

Indonesia dapat dilihat berbagai upaya bank untuk mendapatkan nasabah sebanyak-banyaknya dalam berbagai strategi pemasaran. Dengan memanfaatkan teknologi yang sudah canggih bank dapat menggunakan alat komunikasi jarak jauh seperti telepon atau media elektronik lainnya. Strategi pemasaran ini dilakukan oleh bank-bank untuk menawarkan produk-produknya kepada nasabah melalui media promosi salah satunya menggunakan strategi *Telemarketing*.

*Telemarketing* adalah salah satu cara dalam mempromosikan produk-produk atau jasa layanan yang ada di bank, salah satu produk yang ditawarkan bank yaitu deposito. Deposito adalah produk investasi dengan menyimpan uang dan penarikannya hanya bisa dilakukan dengan kurun waktu tertentu yang telah di janjikan oleh pihak bank dengan persetujuan nasabah. Seorang *telemarketing* bank harus dapat membuat target nasabah, nasabah mana yang berpotensi untuk meningkatkan deposito dengan melihat data-data nasabah bank.

Adapun Permasalahan berdasarkan latar belakang yang telah diuraikan diantara lain adalah :

1. Bagaimana penerapan *Data Mining* dengan algoritma *Naïve Bayes* dan metode *CRISP-DM* dapat menentukan pola untuk memprediksi keputusan calon nasabah atau nasabah tetap yang layak membuka deposito di Bank BRI Syariah, sehingga memudahkan seorang *Telemarketing* dalam mencapai target.
2. Bagaimana hasil prediksi dan tingkat akurasi terhadap banyaknya data-data nasabah yang ada dengan menggunakan klasifikasi algoritma *Naïve Bayes*.

Berdasarkan permasalahan yang ada diatas maka perlu membatasi ruang lingkup dari permasalahan tersebut. Adapun permasalahan yang akan dibahas meliputi :

1. Proses *Data Mining* yang digunakan adalah klasifikasi algoritma *Naïve Bayes* dengan metode *CRISP-DM*.
2. Pokok pembahasan hanya fokus membahas tentang prediksi keputusan calon nasabah dan nasabah tetap yang siap menerima atau tidak penawaran program Deposito berdasarkan hasil *telemarketing*.
3. Di asumsikan data yang akan di analisa telah melalui proses *cleaning* data.
4. *Dataset* yang akan digunakan untuk penerapan data *training* sudah tersedia dan data *testing* yang akan diterapkan terdiri dari 10 atribut dan 1 data terget sebagai *output*.
5. Hanya menampilkan perbandingan hasil data prediksi dan akurasi data untuk memprediksi keputusan dalam menentukan calon nasabah dan nasabah tetap yang berminat atau tidak terhadap produk Deposito berdasarkan hasil *telemarketing*.

Adapun tujuan dari perancangan sistem yang baru adalah sebagai berikut :

1. Menghasilkan pola penerapan *Data Mining* untuk menentukan keputusan calon nasabah dan nasabah tetap yang berminat atau tidak terhadap produk Deposito berdasarkan hasil *telemarketing* yang memiliki standar baik, agar dapat mendukung dan memudahkan pihak bank bagian *marketing* dalam mengetahui dan mengambil keputusan yang tepat sasaran.
2. Menghasilkan perbandingan data prediksi yang memiliki akurasi tinggi, sehingga produk Deposito Bank BRI Syariah dapat diterima oleh calon nasabah maupun nasabah tetap.

## 2. DASAR TEORI

### II.1. *Data Mining*

*Data Mining* adalah proses iteratif dan interaktif untuk menentukan pola atau model baru yang sempurna, bermanfaat dan dapat dimengerti dalam suatu *database* yang sangat besar. *Data Mining* berisi pencarian *trend* pola yang diinginkan dalam *database* besar untuk membantu pengambilan keputusan di waktu yang akan datang. [1]

### II.2. *Klasifikasi*

Klasifikasi merupakan proses pembelajaran suatu fungsi tujuan (target) yang menetapkan tiap himpunan atribut  $x$  ke satu dari label kelas  $y$  yang di definisikan sebelumnya. Dalam mencapai tujuan tersebut, proses klasifikasi membentuk suatu model yang mampu membedakan data kedalam kelas-kelas yang berbeda berdasarkan aturan atau fungsi tertentu. Model itu sendiri bisa berupa aturan berupa pohon keputusan atau formula matematis.

### II.3. *Deposito*

Deposito menurut Undang-undang RI No. 7 tahun 1992 Bab I pasal 1 butir 8: "Deposito didefinisikan sebagai simpanan yang penarikannya hanya dapat dilakukan pada waktu tertentu menurut perjanjian antara penyimpan dengan bank yang bersangkutan".

### II.4. *Telemarketing*

*Telemarketing* adalah suatu saluran pemasaran yang menggunakan sarana telekomunikasi untuk berhubungan dengan pelanggan ataupun calon pelanggan dalam mengomunikasikan ataupun mempromosikan produk yang dimiliki oleh perusahaan. [3]

### II.5. *Algoritma Naïve Bayes*

Algoritma *Naïve Bayes* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naïve Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes*, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. Teorema tersebut dikombinasikan dengan *Naïve* dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naïve Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. [4]

Persamaan dari teorema *Bayes* adalah :

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

#### **Keterangan :**

X : Data dengan class yang belum diketahui

H : Hipotesis data  $X$  merupakan suatu

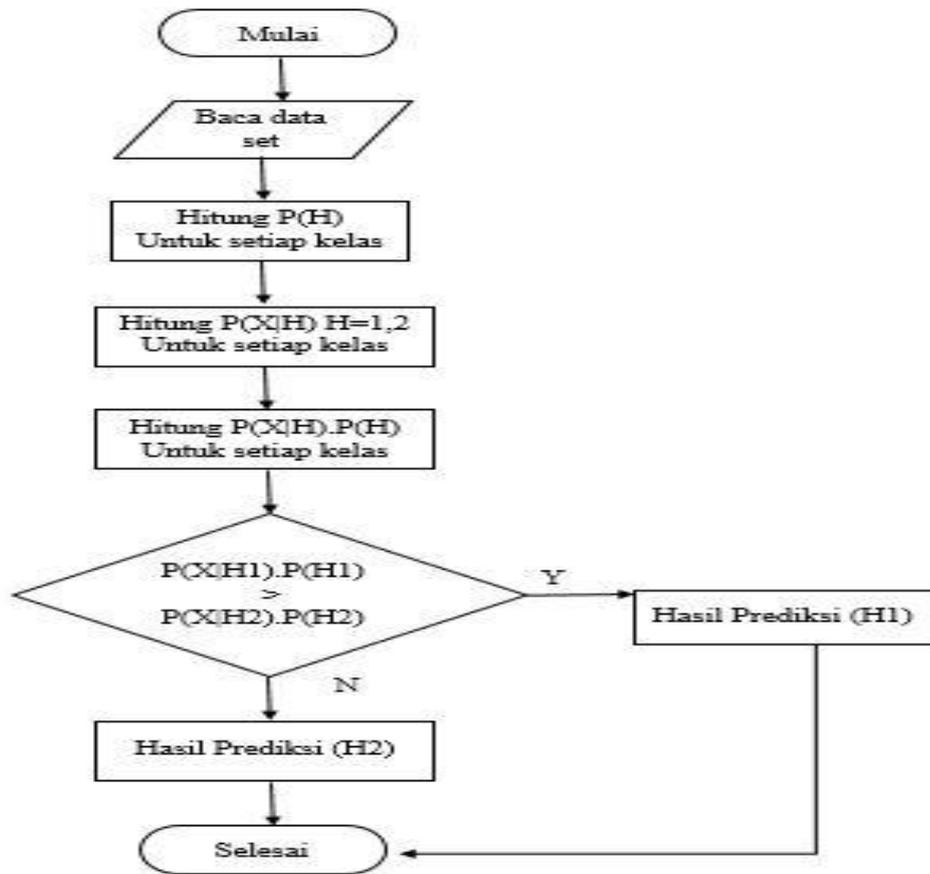
$P(H|X)$  : Probabilitas hipotesis  $H$  berdasar class spesifik kondisi  $X$  (posteriori probability)

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

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

$P(X)$  : Probabilitas  $X$

Adapun alur dari metode *Naïve Bayes* adalah sebagai berikut :

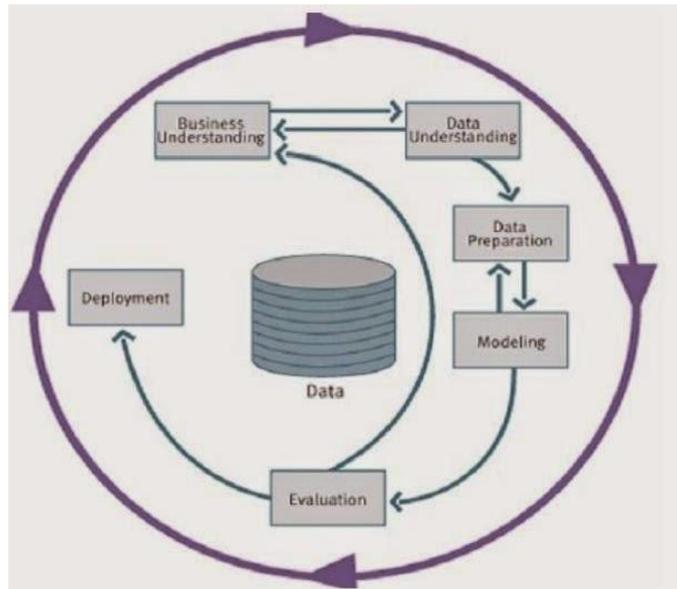


**Gambar 1.** Alur Algoritma Naïve Bayes

## II.6. Metode Analisis Data Mining

Dalam metode analisis data ini digunakan metode *Cross-Industry Standart Process for Data Mining (CRISP-DM)* karena proses ini salah satu tujuannya untuk menemukan pola yang menarik dan bermakna dalam *Data Mining*. [5]

Berikut gambar dan enam fase tersebut adalah:



**Gambar 2.** Proses Data Mining CRISP-DM

Berdasar pada model referensi yang telah ditampilkan pada gambar 2, berikut merupakan penjelasannya:

1. *Business Understanding*, adalah pemahaman tentang substansi dari kegiatan *Data Mining* yang akan dilakukan, kebutuhan dari perspektif bisnis. Kejadiannya antara lain menentukan sasaran atau tujuan bisnis, memahami situasi bisnis, menerjemahkan tujuan bisnis kedalam tujuan *Data Mining*.
2. *Data Understanding*, adalah pengumpulan data, mempelajari data untuk dapat memahami data yang akan digunakan dalam penelitian, mengidentifikasi masalah yang berkaitan dengan data.
3. *Data Preparation*, pada tahap ini struktur basis data akan dipersiapkan sehingga mempermudah proses mining.
4. *Modeling Phase*, adalah fase menentukan teknik *Data Mining* yang digunakan, menentukan tools *Data Mining*, algoritma *Data Mining*, menentukan parameter dengan nilai yang optimal.
5. *Evaluation Phase*, adalah fase interpretasi terhadap hasil *Data Mining* yang ditunjukkan dalam proses pemodelan yang terdapat pada fase sebelumnya. Evaluasi dilakukan secara mendalam dengan tujuan menyesuaikan model yang didapat agar sesuai dengan sasaran yang ingin dicapai dalam fase pertama.
6. *Deployment Phase*, atau fase penyebaran adalah fase penyusunan laporan atau presentasi dari pengetahuan yang didapat dari evaluasi pada proses *Data Mining*.

### III. ANALISIS DAN PERANCANGAN

1. Dalam penelitian ini mengumpulkan data-data dan menggunakan metode analisis *Data Mining* yaitu proses *CRISP-DM* seperti pada Gambar II.2 juga untuk penerapan *Data Mining* menggunakan metode klasifikasi algoritma *Naïve Bayes*.
2. Dan untuk perancangan sistem menggunakan konsep *Object Oriented* dan kebutuhan sistem digambarkan dengan *usecase* diagram seperti pada Gambar 3.

### III.1 Pemahaman Bisnis

Pemahaman bisnis (*business understanding*), tahap pertama dalam proses *CRISP-DM* yang juga dapat disebut sebagai tahap pemahaman bisnis (penelitian).

1. Menentukan Tujuan Bisnis

Tujuan bisnis berdasarkan pengolahan data-data nasabah *telemarketing* diantara lain untuk memprediksi keputusan nasabah telemarketing dalam menawarkan deposito di bank.

2. Menentukan Strategi Awal *Data Mining*

Strategi awal dalam menerapkan tujuan dilakukannya *Data Mining* adalah menggunakan data-data yang didapatkan dari penelitian tersebut diperoleh langsung dari <http://archive.uci.edu/> data tersebut diambil dari bagian marketing sebuah bank Portugal, dan telah didonasi pada tanggal 14 Februari 2012 oleh S.Moro, P.Cortez. Rita. Data-data tersebut akan digunakan dan akan diterapkan untuk mendukung dalam proses bisnis dalam menawarkan deposito di bank.

### III.2 Pemahaman Data

Setelah pemahaman terhadap bisnis, aktivitas utama dari proses *Data Mining* berikutnya adalah mengidentifikasi data yang relevan dari database yang ada. Dataset tersebut terdiri dari 10 atribut predictor dan 1 atribut label. Atribut prediktor adalah atribut yang dijadikan penentu nasabah yang hampir pasti membuka rekening deposito dan atribut label yaitu atribut yang dijadikan hasil learning. Atribut prediktor yang digunakan pada penelitian ini adalah *age, job, marital, education, balance, housing, loan, contact, previous, dan poutcome*. Sedangkan atribut hasil atau label adalah *y*. Dataset yang didapatkan dari berupa dokumen excel sejumlah 4521 record.

### III.3 Pengelolaan Data

Adapun sumber data utama yang digunakan dalam penelitian ini adalah dataset berupa dokumen excel. Namun tidak semua data dan atribut dapat digunakan sehingga harus melewati tahap persiapan *Data Mining* terlebih dahulu. Untuk mendapatkan data yang berkualitas maka dapat dilakukan teknik-teknik sebagai berikut, terdapat tiga proses yaitu:

1. *Data selection* yaitu memilih data yang akan digunakan dalam proses *Data Mining*. Dalam proses ini dilakukan juga pemilihan atribut-atribut yang disesuaikan dengan proses *Data Mining*.
2. *Data cleaning* yaitu memastikan kualitas data yang telah dipilih pada tahap *data selection*, pada tahap ini masalah yang harus dihadapi adalah apabila terdapat noisy data dan missing values. Proses pembersihan data atau cleaning dilakukan untuk menemukan anomali data yang mungkin saja masih terdapat pada data tersebut.
3. *Data transformation* yaitu pengelompokkan atribut-atribut atau field yang telah terpilih kedalam basis data baru untuk bahan *Data Mining*. Dalam hal ini menggunakan rumus distribusi frekuensi kelompok terhadap beberapa attribute yaitu *age, balance* dan *previous* agar data-data tersebut dapat membantu proses peneceapan algoritma tersebut.

### III.4 Pemodelan Data

Pemodelan adalah fase yang secara langsung melibatkan teknik *Data Mining* yaitu dengan melakukan pemilihan teknik *Data Mining* dan menentukan algoritma yang akan digunakan sesuai proses bisnis yang berjalan. Dalam penelitian ini menggunakan metode klasifikasi algoritma *Naïve Bayes*. Proses dalam penerapan aplikasi ini akan menggunakan *yii framework* untuk menjalankan dan menemukan pola algoritma tersebut dengan

dilakukannya alur perhitungan algoritma *Naïve Bayes* sesuai gambar 1 agar hasil yang diharapkan untuk membandingkan prediksi nasabah lebih akurat dengan menggunakan dataset yang telah di persiapkan berupa dokumen excel tersebut.

### III.5 Evaluasi Data

Dalam tahap ini dilakukan validasi dan keakuratan hasil akurasi serta efisiensi algoritma *Naïve Bayes* pada model tersebut dengan menggunakan beberapa teknik yaitu *yii framework* untuk penerapan aplikasi dan membantu untuk mengukur akurasi pada model yang digunakan dengan menggunakan adataset yang telah di tentukan dari awal fase. Pada fase ini terdapat dua proses yaitu:

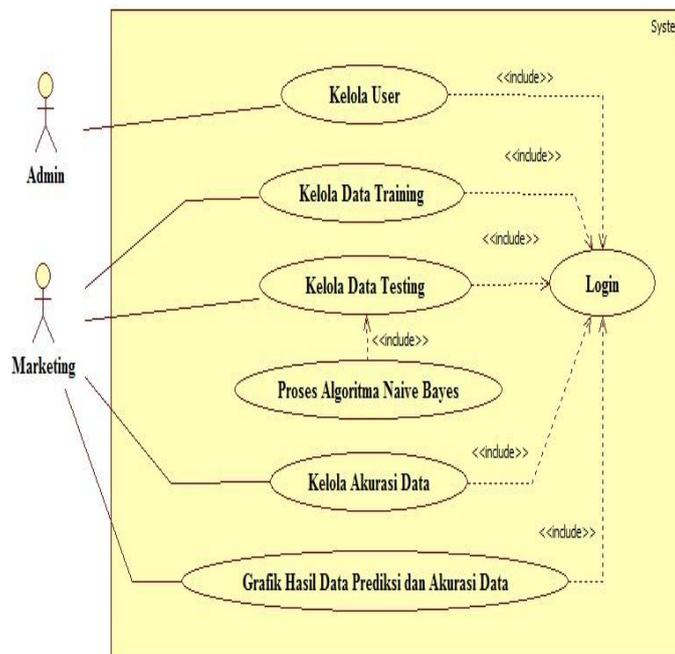
1. *Evaluate result* yaitu merangkum hasil penilaian dari segi kriteria keberhasilan bisnis, termasuk pernyataan akhir mengenai apakah penelitian sudah memenuhi tujuan bisnis. Setelah dilakukan evaluasi terhadap penerapan *Data Mining* tersebut hasilnya mendapatkan nilai perbandingan prediksi dan ingkat akurasi yang terstandarisasi dengan baik.
2. *Determine next step* yaitu memberikan keputusan apakah teknik pemodelan yang dipergunakan dapat dijadikan standar dalam menentukan tujuan penelitian atau kembali ke fase awal.

### III.6 Penyebaran Data

1. Pada *fase* ini terdapat dua proses yaitu: *Deployment plan* yaitu menjelaskan gambaran mengenai rencana terhadap pembuatan laporan yang akan dibuat.
2. *Produce final report* yaitu memberikan visualisasi dari laporan yang telah dibuat berdasarkan pada *deployment plan*.

### III.7 Perancangan Sistem

metode analisis *Data Mining* yaitu proses *CRISP-DM* seperti pada Gambar 2 untuk perancangan sistem menggunakan konsep *Object Oriented* dan kebutuhan sistem digambarkan dengan usecase diagram seperti pada Gambar 3.



Gambar 3. Usecase Diagram

### III.8 Use Case Diagram

Dalam menganalisa kebutuhan perangkat lunak sesuai dengan notasi *use case diagram* pada Gambar 3 ada 2 (dua) aktor yang akan terlibat dalam Perangkat Lunak yaitu *Administrator* dan *Marketing* Semua aktor sangat berperan membantu proses analisa yang akan dilakukan oleh perangkat lunak agar dapat berjalan dengan baik.

Berdasarkan Gambar 3 maka usecase dibagi menjadi sebagai berikut :

- a. *Use case* Mengelola Users akan melayani Aktor *Administrator*.
- b. *Use case* Mengelola Data *Training*, Data *Testing* dan Data Akurasi Data dan Menampilkan Grafik Perbandingan dan akurasi data akan melayani Aktor *Marketing*.

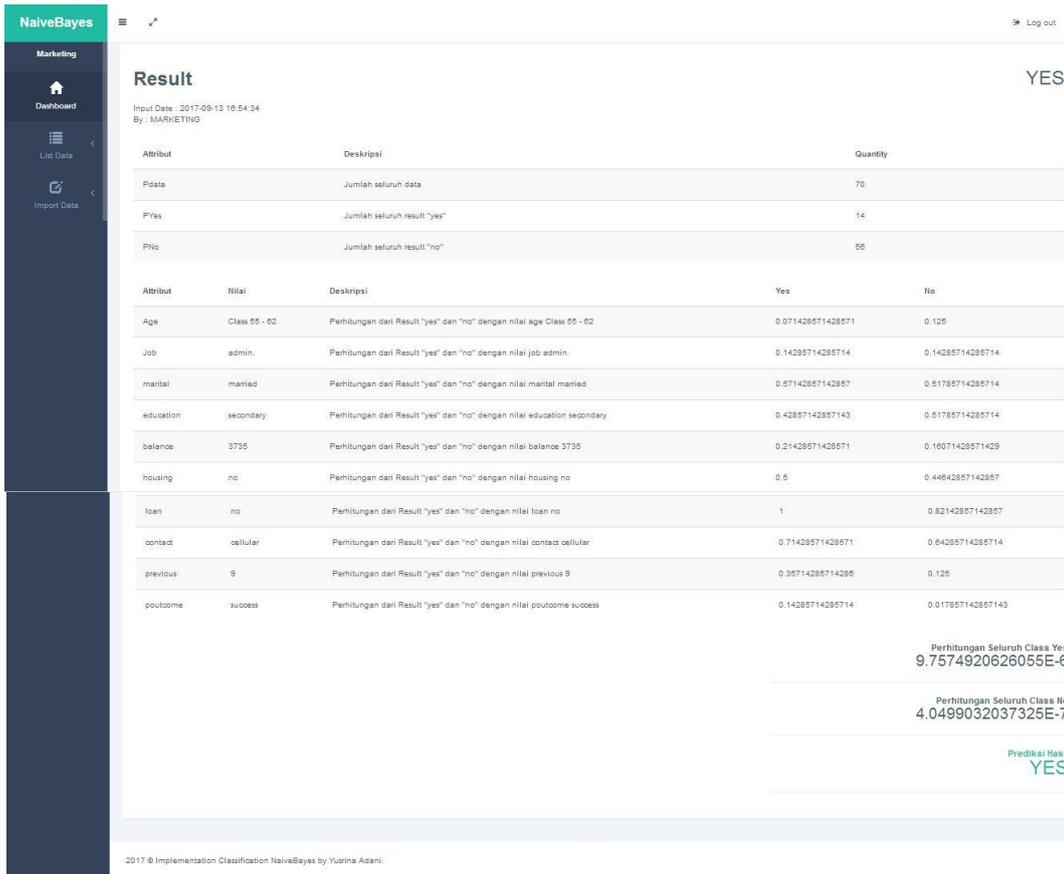
## IV. IMPLEMENTASI

Pada tahap pengujian analisa data menggunakan data *dummy* yang sudah dijelaskan pada tahap pemahaman data dalam analisis *Data Mining CRISP-DM*. Dengan ketentuan data yang diambil untuk penerapan tersebut berjumlah 102 *record*. Dari data *training* sebanyak 70 *record*. dan data *testing* sebanyak 32 *record*.

1. Pertama untuk pengujian memprediksi data nasabah yang akan membuka deposito tersebut diambil 1 *record* dengan kriteria 10 atribut sebagai predictor dan 1 label sebagai hasil atau outpunya dan nanti hasil dari pengujian 1 data *testing* ini didapatkan sesuai dengan perhitungan rumus algoritma *Naïve Bayes* pada aturan klasifikasi yang dijelaskan pada gambar 1 yang di proses dengan melihat data data *training* sebanyak 70 record untuk menghasilkan hasil dari semple 1 data *testing* . Dan sample data nasabahnya untuk pengujian seperti pada tabel 1. Dan hasil penerapan pola tersebut dapat tampilan dan di proses menggunakan aplikasi yang akan dijelaskan pada Gambar 4.

**Tabel 1** *Sample Data Testing*

Age	job	marital	education	balance	housing	loan	contact	previous	poutcome	y
60	admin.	married	secondary	3735	no	no	cellular	9	success	?



**Gambar 4.** Penerapan pola aplikasi hasil perhitungan

2. Kemudian kedua untuk hasil tingkat akurasi data dapat diambil dan dihitung dari data *testing* yang sudah ada dibandingkan dengan data *testing* menggunakan pola algoritma *Naive Bayes* yang masing-masing tiap data nya sebanyak 32 *record*, dengan menggunakan rumus perhitungan akurasi data yang dijelaskan pada Gambar 5 dan Gambar 6 sebagai berikut :

Dari hasil ini, dapat dihitung nilai akurasi nya.

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{banyaknya prediksi yang benar}}{\text{total banyaknya prediksi}} \times 100\% \\
 \text{Akurasi} &= \frac{21}{32} \times 100\% \\
 &= 0,65625 \times 100\% \\
 &= 65.625 \Rightarrow 66\% \text{ (Data Akurat)}
 \end{aligned}$$

**Gambar 5** Nilai Akurasi Akurat

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{banyaknya prediksi yang salah}}{\text{total banyaknya prediksi}} \times 100\% \\
 \text{Akurasi} &= \frac{11}{32} \times 100\% \\
 &= 0,34375 \times 100\% \\
 &= 34,375 \Rightarrow 34\% \text{ (Data Tidak Akurat)}
 \end{aligned}$$

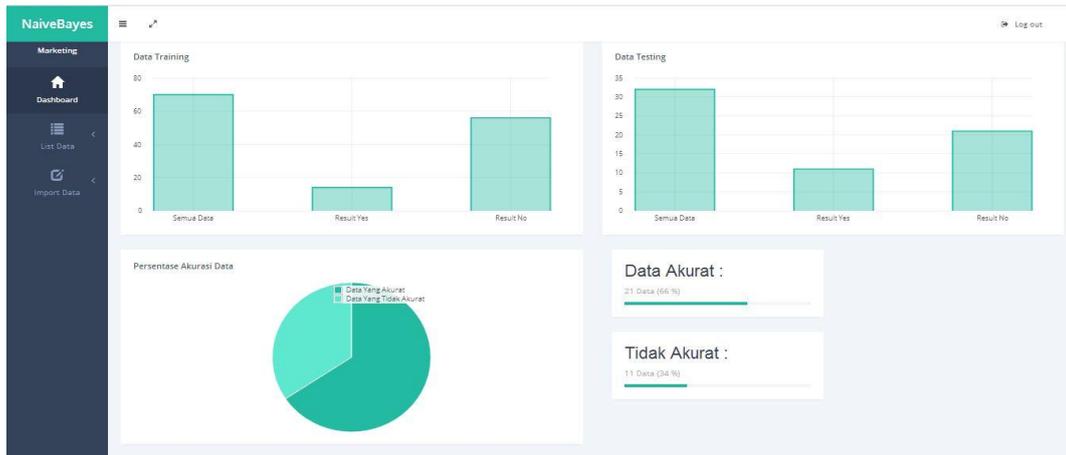
**Gambar 6** Nilai Akurasi Akurat

## V. KESIMPULAN

Berdasarkan hasil analisa, perancangan, implementasi, dan pengujian sesuai dengan proses analisis *Data Mining* menggunakan metode *CRISP-DM* maka dapat ditarik kesimpulan bahwa penerepan algoritma *Naïve Bayes* dapat menampilkan pola dengan menghasilkan perbandingan data prediksi dan akurasi data dalam menawarkan deposito yang terstadarisasi dengan baik, yang dapat mendukung dan membantu pihak bank bagian *marketing* untuk mengetahui keputusan calon nasabah dan nasabah tetap yang berminat atau tidak terhadap produk Deposito berjangka sesuai dengan hasil *telemarketing*. Pada pengujian ini menggunakan pengelolaan *dataset* yang sudah melalui tahapan *data cleaning* terlebih dahulu, sehingga memudahkan dalam proses perhitungan algoritma *Naïve Bayes* yang akan menghasilkan data-data yang lebih akurat.

Dengan ketentuan data yang diambil untuk penerapan tersebut berjumlah 102 *record*. Dari data *training* sebanyak 70 *record*. dan data *testing* sebanyak 32 *record*. Dapat dilihat pada Gambar 7 yaitu Grafik Hasil Prediksi dan Akurasi Data.

1. Maka hasil prediksi nasabah yang membuka deposito sebanyak 11 nasabah (*yes*) dan nasabah yang tidak membuka deposito sebanyak 21 nasabah (*no*).
2. Dan hasil tingkat akurasi data lebih dari 50%, hasilnya dibandingkan dari data *testing* yang sudah ada dengan data *testing* menggunakan pola algoritma *Naïve Bayes* sebanyak 32 *record*, maka hasil data yang akurat sebesar 66% dan hasil data yang tidak akurat 34%. Artinya rata-rata calon nasabah dan nasabah tetap dapat menerima penawaran produk Deposito berjangka dari Bank BRI Syariah.



**Gambar 7** Hasil Grafik Prediksi dan Akurasi data

## Daftar Pustaka

- [1] F. A. Hermawati, *Data Mining*, Yogyakarta: CV ANDI OFFSET, 2013.
- [2] B. Santosa, *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*, Yogyakarta: Graha Ilmu, 2007.
- [3] D. M. A Fauzi, *Aplikasi Excel untuk telemarketing*, Jakarta: Elex Media, 2007.
- [4] S. Kumusadewi, *Artificial Intelligence*, Yogyakarta: Graha Ilmu, 2003.
- [5] D. T. Larose, *Discovering Knowledge in Data an Introduction to Data Mining*, John Wiley and Sons, 2005.

