

Nama : Febianus Felix W

NIM : 00000072737

Pilihan Industri : Banking

Company Name : Bank Uhuy

# Optimisasi Kampanye Pemasaran Langsung di Sektor Perbankan dengan CRISP-DM Menggunakan Decision Tree, Random Forest, Neural Network, dan SVM

Febianus Felix Widisulistyo<sup>1</sup>

<sup>1</sup>Information System, Faculty of Engineering & Information  
Universitas Multimedia Nusantara, Tangerang Banten 15810, Indonesia.  
[febianus.felix@student.umn.ac.id](mailto:febianus.felix@student.umn.ac.id)<sup>1</sup>

**Abstract** — *This research evaluates and enhances the effectiveness of Bank Uhuy's direct marketing campaign using the CRISP-DM methodology. Starting with business understanding, the study covers the comprehension of Bank Uhuy's marketing campaign objectives, including understanding customer preferences, market trends, and the dynamics of the banking industry. Utilizing SAS Visual Analytics, various classification techniques such as Decision Tree, Neural Network, Random Forest, and Support Vector Machine were employed to analyze customer data. The dataset includes customer characteristics and interaction history with Bank Uhuy, including age, occupation, marital status, education, balance, home loans, and previous campaign outcomes. Data preparation involved checking for missing values, handling outliers, and feature selection to simplify the analysis process. Model evaluation revealed that Neural Network outperformed Decision Tree, Random Forest, and Support Vector Machine, showing the lowest Misclassification Rate and the highest F1-Score, indicating an optimal balance of precision and recall. The implementation phase involved integrating the Neural Network model into Bank Uhuy's operational environment for automated and accurate campaign outcome predictions. This implementation aims to improve marketing effectiveness, customer conversion, loyalty, and competitive position in the banking industry.*

**Keywords**—*Direct marketing campaigns, Banking industry, CRISP-DM methodology, Neural Network, Customer data analysis*

**Abstrak** — *Penelitian ini mengevaluasi dan meningkatkan efektivitas kampanye pemasaran langsung Bank Uhuy menggunakan metodologi CRISP-DM. Dimulai dengan pemahaman bisnis, penelitian ini mencakup pemahaman tujuan kampanye pemasaran Bank Uhuy, termasuk memahami preferensi pelanggan, tren pasar, dan dinamika industri perbankan. Dengan memanfaatkan SAS Visual Analytics, berbagai teknik klasifikasi seperti Decision Tree, Neural Network, Random Forest, dan Support Vector Machine digunakan untuk menganalisis data pelanggan. Dataset mencakup karakteristik pelanggan dan riwayat interaksi dengan Bank Uhuy, termasuk usia, pekerjaan, status pernikahan, pendidikan, saldo, pinjaman rumah, dan hasil kampanye sebelumnya. Pra-pemrosesan data melibatkan pemeriksaan nilai yang hilang, penanganan pencilaan, dan pemilihan fitur untuk menyederhanakan proses analisis. Evaluasi model mengungkapkan bahwa Neural Network mengungguli Decision Tree, Random Forest, dan Support Vector Machine, menunjukkan Misclassification Rate terendah dan F1-Score tertinggi, menandakan keseimbangan presisi dan recall yang optimal. Tahap implementasi melibatkan integrasi model*

*Neural Network ke dalam lingkungan operasional Bank Uhuy untuk prediksi hasil kampanye yang otomatis dan akurat. Implementasi ini bertujuan untuk meningkatkan efektivitas pemasaran, konversi pelanggan, loyalitas, dan posisi bersaing di industri perbankan.*

**Kata kunci**—*Kampanye pemasaran langsung, Industri perbankan, Metodologi CRISP-DM, Jaringan Saraf Tiruan, Analisis data pelanggan*

## A. I. LATAR BELAKANG DAN PEMAHAMAN BISNIS

Pemasaran langsung merupakan salah satu strategi yang umum digunakan oleh bank untuk menjangkau calon pelanggan dan meningkatkan basis pelanggan mereka [1]. Melalui strategi ini, komunikasi langsung dengan calon pelanggan dilakukan melalui berbagai saluran, seperti telepon, email, dan media sosial, dengan tujuan menarik minat mereka terhadap produk dan layanan yang ditawarkan [1], [2]. Berbagai kampanye pemasaran langsung telah dijalankan oleh Bank Uhuy, sebagai salah satu institusi keuangan, untuk menarik pelanggan baru dan mempertahankan pelanggan yang sudah ada. Kampanye pemasaran ini dirancang dengan berbagai pendekatan kreatif dan inovatif untuk menarik perhatian pelanggan dan meningkatkan loyalitas mereka terhadap bank.

Namun, untuk terus meningkatkan efektivitas kampanye-kampanye tersebut, evaluasi yang mendalam terhadap hasil kampanye pemasaran sebelumnya perlu dilakukan oleh Bank Uhuy. Evaluasi ini dianggap sangat penting karena wawasan berharga tentang apa yang berhasil dan apa yang tidak dalam strategi pemasaran yang telah dijalankan dapat diberikan [3] [4]. Dengan hasil dari kampanye sebelumnya yang dievaluasi, faktor-faktor yang mempengaruhi keberhasilan kampanye dapat diidentifikasi oleh Bank Uhuy, seperti segmen pelanggan yang merespons dengan baik, jenis pesan yang paling efektif, dan saluran komunikasi yang paling efisien. Informasi ini sangat berharga untuk digunakan dalam perancangan strategi pemasaran di masa mendatang yang lebih terarah dan efektif [4].

Evaluasi ini juga bertujuan untuk mengoptimalkan strategi pemasaran di masa mendatang. Dengan pola dan tren dari data kampanye sebelumnya yang dipahami, keputusan yang lebih informasional dalam merancang kampanye baru

[5] dapat dibuat oleh Bank Uhuy. Selain itu, dengan saluran komunikasi yang paling efektif yang diketahui, sumber daya pemasaran dapat dialokasikan lebih efisien oleh bank, sehingga hasil yang lebih baik dengan biaya yang lebih rendah dapat dicapai [6].

Dalam konteks ini, analisis data yang komprehensif dan penggunaan metode klasifikasi dapat memberikan wawasan yang mendalam tentang efektivitas kampanye pemasaran. Dengan data kampanye sebelumnya yang dimanfaatkan, pola perilaku pelanggan dan variabel yang berkontribusi terhadap keberhasilan kampanye dapat diidentifikasi oleh Bank Uhuy. Ini termasuk analisis variabel demografis, perilaku pelanggan, dan respons mereka terhadap berbagai elemen kampanye. Dengan menggunakan metode *Decision Tree*, *Random Forest*, *Neural Network*, dan *SVM*, serta membandingkan keempat model tersebut, Bank Uhuy dapat menentukan pendekatan yang paling efektif. Dengan metodologi yang tepat, seperti *CRISP-DM* (*Cross-Industry Standard Process for Data Mining*) yang digunakan, proses evaluasi dapat dipastikan dilakukan secara sistematis dan terstruktur, yang pada akhirnya akan membantu dalam mengoptimalkan strategi pemasaran mereka dan mencapai tujuan bisnis yang telah ditetapkan.[7]

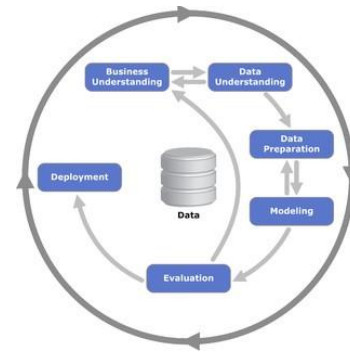
Oleh karena itu, evaluasi hasil kampanye pemasaran dipandang sebagai langkah yang krusial bagi Bank Uhuy dalam upaya untuk meningkatkan efektivitas kampanye pemasaran dan mencapai tujuan bisnis jangka panjang mereka. Dengan evaluasi yang mendalam dan pendekatan analisis data yang tepat yang dilakukan, adaptasi terhadap perubahan pasar dan kebutuhan pelanggan dapat terus dilakukan oleh Bank Uhuy, serta daya saing dalam industri perbankan dapat ditingkatkan [8].

Batasan masalah penelitian ini adalah terfokus pada analisis dampak kampanye pemasaran terhadap konversi pelanggan, dengan penekanan pada data kampanye pemasaran sebelumnya serta pembatasan lingkup penelitian pada periode dan wilayah tertentu. Rumusan masalahnya meliputi pertanyaan tentang tingkat keberhasilan kampanye pemasaran langsung dalam mengubah prospek menjadi pelanggan aktif, bagaimana tahapan pemahaman bisnis dan eksplorasi data dapat membantu mengidentifikasi kebutuhan dan tantangan bisnis Bank Uhuy, serta seberapa efektif model klasifikasi dalam memprediksi hasil kampanye pemasaran berdasarkan data yang tersedia.

Tujuan penelitian ini adalah mengevaluasi tingkat konversi prospek menjadi pelanggan aktif sebagai indikator keberhasilan kampanye pemasaran langsung, menganalisis proses pemahaman bisnis untuk memahami tujuan dan tantangan utama Bank Uhuy dalam kampanye pemasaran, dan mengukur efektivitas model klasifikasi dengan menghitung *Misclassification Rate*, *F1-score*, dan *Confusion Matrix*. Kegunaannya adalah untuk mengidentifikasi faktor-faktor yang mempengaruhi efektivitas kampanye pemasaran, seperti jenis pesan, saluran komunikasi, dan segmentasi pelanggan, serta memberikan pemahaman tentang metodologi analisis data, seperti *CRISP-DM*, dalam evaluasi hasil kampanye pemasaran.

## B. II. TINJAUAN TEORITIS

### C. CRISP DM



Gambar 1. Proses CRISP-DM [10]

*CRISP-DM* (*Cross-Industry Standard Process for Data Mining*) adalah sebuah metode yang telah diakui secara luas dalam industri untuk mengelola dan mengimplementasikan proyek-proyek data mining [7]. Metode ini telah terbukti efektif dalam memandu proses analisis data dari tahap awal hingga tahap akhir, memungkinkan para profesional untuk mengorganisir dan mengelola proyek-proyek data mining dengan lebih terstruktur dan sistematis [9]. *CRISP-DM* meliputi enam tahap: pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penyebaran. Tahap pemahaman bisnis menyoroti tujuan dan tantangan bisnis, sementara tahap pemahaman data berkaitan dengan pemahaman data yang akan dianalisis. Tahap persiapan data melibatkan pembersihan dan transformasi data, sementara tahap pemodelan melibatkan pengembangan dan evaluasi model. Evaluasi melibatkan penilaian kinerja model dan validasi hasil, sementara tahap penyebaran fokus pada implementasi solusi ke dalam lingkungan produksi.

Dengan adopsi *CRISP-DM*, organisasi dapat mengurangi risiko dan meningkatkan peluang kesuksesan dalam proyek-proyek data mining. Oleh karena itu, *CRISP-DM* tetap menjadi salah satu panduan yang paling diandalkan dan digunakan secara luas dalam industri untuk mengelola proyek-proyek data mining dengan efektif dan efisien [10].

### D. Decision Tree

*Decision Tree* merupakan salah satu algoritma klasifikasi yang umum digunakan dalam analisis data pemasaran. Prosesnya terdiri dari membagi data menjadi subset yang semakin kecil berdasarkan serangkaian keputusan atau aturan, dengan setiap node mewakili keputusan berdasarkan fitur-fitur data tertentu dan cabang-cabangnya merepresentasikan kemungkinan hasil dari keputusan tersebut. Kelebihanannya terletak pada kemampuannya memberikan interpretasi yang mudah dipahami oleh manusia dan dapat menangani data kategori dan numerik, serta mengatasi masalah overfitting dengan teknik pruning [11]. Rumus dasar untuk membangun pohon keputusan dalam algoritma *Decision Tree* adalah

$$Entropy = -\sum_{i=1}^n p_i \log_2(p_i)$$

$n$  adalah jumlah kelas atau label dalam data

$p_i$  adalah proporsi sampel yang termasuk dalam kelas atau label ke- $i$

$\log_2$  adalah logaritma basis 2

Entropy mengukur ketidakpastian dalam data; semakin rendah entropi, semakin teratur dan terstruktur data tersebut.

dengan parameter yang tepat dan teknik pruning yang efektif, pohon keputusan tetap menjadi alat yang efektif

dalam analisis data pemasaran. Dengan menggunakan pohon keputusan, Bank Uhuy dapat mengidentifikasi faktor-faktor yang mempengaruhi keberhasilan kampanye pemasaran dan merancang strategi pemasaran yang lebih efektif.

#### E. Random Forest

Random Forest merupakan teknik populer dalam machine learning yang mengembangkan konsep pohon keputusan untuk pemodelan klasifikasi [11]. Dibandingkan dengan pohon keputusan tunggal, Random Forest mengatasi overfitting dengan cara membuat sejumlah pohon keputusan secara acak berdasarkan sampel data yang diambil dengan penggantian. Hasil prediksi dari setiap pohon keputusan dikombinasikan secara ensemble untuk menghasilkan prediksi akhir yang lebih umum dan stabil [12].

Random Forest memiliki keunggulan dalam mengatasi overfitting dan menangani data besar serta kompleks. Dengan memanfaatkan banyak pohon keputusan dan teknik pengambilan sampel acak, model yang dihasilkan lebih umum dan stabil. Selain itu, Random Forest juga dapat secara otomatis melakukan seleksi fitur, membuatnya populer dalam analisis data pemasaran [11], [12]. Bagi Bank Uhuy, Random Forest dapat digunakan untuk mengidentifikasi faktor-faktor penting dalam kampanye pemasaran, melakukan segmentasi pelanggan, dan memprediksi hasil kampanye baru, meningkatkan efektivitas strategi pemasaran dan pencapaian tujuan bisnis.

#### F. Neural Network

Jaringan saraf tiruan (*Neural Network*) merupakan teknik pembelajaran mesin yang mampu mengenali pola-pola kompleks dalam data pemasaran [13]. Dengan memanfaatkan arsitektur terinspirasi dari otak manusia, jaringan saraf dapat mengklasifikasikan calon pelanggan berdasarkan kemungkinan konversi mereka menjadi pelanggan aktif. Model jaringan saraf dilatih menggunakan data kampanye pemasaran sebelumnya, yang menghubungkan fitur input seperti demografi pelanggan, jenis pesan, dan saluran komunikasi dengan output yang diinginkan, misalnya, konversi atau tidak. Rumus inti dalam jaringan saraf adalah:

$$a_j^{(l)} = f(\sum_{i=1}^n w_{ij}^{(l)} a_i^{(l-1)} + b_j^{(l)})$$

$a_j^{(l)}$  adalah Output neuron ke-j pada layer ke-l.

$f$  adalah Fungsi aktivasi untuk memperoleh output dari neuron.

$w_{ij}^{(l)}$  adalah Bobot yang menghubungkan neuron ke-i dari layer ke-(l-1) dengan neuron ke-j pada layer ke-l.

$a_i^{(l-1)}$  adalah Output neuron ke-i pada layer ke-(l-1).

$b_j^{(l)}$  adalah Bias dari neuron ke-j pada layer ke-l.

Setelah dilatih, *Neural Network* mampu memberikan prediksi terhadap hasil kampanye pemasaran baru, sementara kemampuannya dalam menangani data yang besar dan tidak terstruktur menjadikannya alat yang efektif dalam analisis data pemasaran yang kompleks. Dengan demikian, model jaringan saraf dapat memberikan wawasan yang mendalam tentang pola perilaku pelanggan dan faktor-faktor yang mempengaruhi keberhasilan kampanye, membantu Bank Uhuy merancang strategi pemasaran yang lebih efektif dan efisien.

#### G. Support Vector Machine

Algoritma machine learning yang dikenal sebagai Support Vector Machine (SVM) digunakan untuk menangani permasalahan klasifikasi dengan menciptakan hyperplane terbaik yang memisahkan berbagai kelas data dalam ruang fitur [14]. Keunggulan utama SVM terletak pada kemampuannya menangani dataset dengan dimensi tinggi dan kompleks, termasuk yang tidak linier, serta menghasilkan model klasifikasi yang konsisten dan dapat diandalkan dengan penentuan *hyperplane* yang berorientasi pada margin maksimum [15]. Dalam konteks analisis data pemasaran, SVM memiliki potensi besar untuk mengklasifikasikan pelanggan berdasarkan berbagai fitur seperti demografi, perilaku pembelian, atau preferensi produk, memungkinkan Bank Uhuy untuk mengoptimalkan segmentasi pelanggan dan merancang strategi pemasaran yang lebih efektif. Rumus yang digunakan dalam penelitian ini adalah

$$\min_{w,b} \frac{1}{2} ||w||^2 + C \sum_{i=1}^n \varepsilon_i$$

$w$  adalah vector bobot

$\varepsilon_i$  adalah variabel slack

$C$  adalah parameter penalti yang mengontrol trade-off antara margin maksimum dan jumlah kesalahan klasifikasi

#### H. Misclassification Rate

*Misclassification rate*, sebuah metrik evaluasi dalam analisis klasifikasi, dihitung dengan membagi jumlah data yang salah diklasifikasikan oleh model dengan total data yang dievaluasi. Signifikansinya terletak pada kemampuannya untuk memberikan gambaran tentang performa model klasifikasi dalam memprediksi kelas yang benar [16]. Dalam analisis pemasaran, misclassification rate digunakan untuk mengevaluasi efektivitas model dalam memprediksi hasil kampanye, memungkinkan area perbaikan diidentifikasi dan strategi pemasaran yang lebih efektif dirancang oleh Bank Uhuy.

#### I. F1 Score

*F1 score* merupakan metrik evaluasi umum dalam analisis klasifikasi yang menggabungkan presisi dan recall. Metrik ini memberikan gambaran menyeluruh tentang kinerja model klasifikasi dengan mempertimbangkan kedua aspek tersebut. *F1 score* dihitung sebagai rata-rata harmonis dari presisi dan recall, mencerminkan keseimbangan antara keduanya [9].

Dalam analisis data pemasaran, *F1 score* digunakan untuk mengevaluasi performa model klasifikasi dalam memprediksi hasil kampanye. Dengan mempertimbangkan presisi dan recall, *F1 score* memberikan gambaran komprehensif tentang efektivitas model dalam mengklasifikasikan data. Dengan demikian, Bank Uhuy dapat memanfaatkan *F1 score* untuk mengevaluasi dan meningkatkan model, serta merancang strategi pemasaran yang lebih efektif guna mencapai tujuan bisnis.

#### J. Confusion Matrix

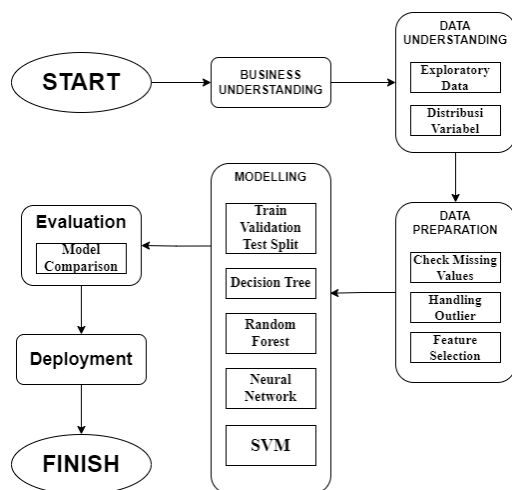
		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Gambar 2 Confusion Matrix [18]

Confusion matrix adalah tabel evaluasi kinerja model klasifikasi yang menunjukkan prediksi model terhadap data aktual. Terdiri dari empat sel: true positive (TP), true negative (TN), false positive (FP), dan false negative (FN), confusion matrix memungkinkan perhitungan berbagai metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Melalui analisis confusion matrix, peneliti dapat mengidentifikasi kelemahan dan kelebihan model serta memahami di mana model cenderung membuat kesalahan atau menunjukkan kekuatan. Dengan demikian, confusion matrix menjadi alat penting dalam evaluasi dan peningkatan model klasifikasi, memberikan wawasan yang berharga untuk pengembangan lebih lanjut [17].

### K. III. METODOLOGI PENELITIAN

Dalam penelitian ini, metodologi yang digunakan melibatkan penerapan metode CRISP-DM (Cross-Industry Standard Process for Data Mining) untuk melakukan evaluasi hasil kampanye pemasaran Bank Uhuy.



Gambar 2. Proses Penelitian

Diagram proses tersebut mencerminkan langkah-langkah utama dalam mengintegrasikan data dan melakukan analisis CRISP DM (Cross-Industry Standard Process for Data Mining). Berikut adalah penjelasan untuk setiap tahapan dalam diagram tersebut:

#### A. Business Understanding

Tahap awal dalam proses penelitian, yaitu Business Understanding, melibatkan pemahaman yang mendalam terhadap tujuan utama Bank Uhuy untuk meningkatkan efektivitas kampanye pemasaran mereka. Hal ini mencakup

pemahaman komprehensif tentang target pasar, tujuan spesifik dari setiap kampanye pemasaran, serta tantangan yang dihadapi dalam mencapai tujuan tersebut. Dalam konteks ini, wawasan yang jelas tentang preferensi pelanggan, tren pasar, dan dinamika industri perbankan yang dapat mempengaruhi hasil kampanye pemasaran perlu diperoleh oleh Bank Uhuy.

Dalam tahap ini, Bank Uhuy perlu memahami evaluasi kampanye pemasaran sebelumnya untuk mengidentifikasi faktor-faktor yang mempengaruhi keberhasilan atau kegagalan, seperti jenis pesan yang efektif, saluran komunikasi yang responsif, dan segmen pelanggan yang responsif. Dengan memahami pengalaman masa lalu, Bank Uhuy dapat membuat keputusan yang lebih informasional untuk merencanakan dan melaksanakan kampanye pemasaran di masa mendatang.

Dalam penelitian ini, alat yang digunakan adalah SAS Visual Analytics. Analisis dilakukan menggunakan beberapa teknik klasifikasi, termasuk *Decision Tree*, *Neural Network*, *Random Forest*, dan *Support Vector Machine*. Hasil dari analisis ini dievaluasi menggunakan metrik *Misclassification Rate* dan *F1 Score*, yang akan memberikan gambaran tentang akurasi dan kinerja model dalam mengklasifikasikan data.

#### B. Data Understanding

Tahap Data Understanding bertujuan untuk memperoleh pemahaman yang komprehensif tentang data yang akan digunakan dalam analisis kampanye pemasaran Bank Uhuy. Proses ini mencakup beberapa langkah penting:

##### 1) Exploratory Data:

Columns	Rows	Size
17	11.2 K	2.8 MB

Gambar 3. total Kolom dan Baris

data yang diteliti memiliki 11.162 baris dan 17 kolom yang mana berisi informasi tentang direct marketing bank beserta faktor pendukung yang mencakup :

TABLE I. TABEL INFORMASI DATASET

Column	Deskripsi Data	
	Data Type	Penjelasan
Age	Double	Umur Pelanggan
Balance	double	Saldo Rekening
Campaign	double	Jumlah kontak yang dilakukan selama kampanye ini dan untuk klien ini.
Contact	string	Komunikasi kontak ('cellular', 'telephone')

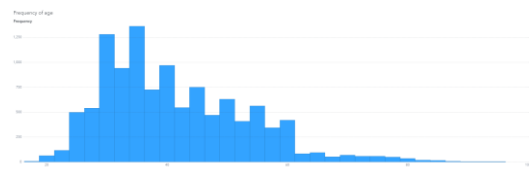
Day	double	Hari terakhir kontak (1,2,3....31)
Default	boolean	Apakah memiliki kredit bermasalah? ('no', 'yes', 'unknown')
Deposit	boolean	Apakah klien telah berlangganan deposito berjangka? ('ya', 'tidak')
Duration	double	Durasi kampanye
Education	string	Pendidikan Terakhir Pelanggan
housing	boolean	Apakah memiliki pinjaman perumahan? ('no', 'yes', 'unknown')
job	string	Jenis Pekerjaan Pelanggan ('admin.', 'bluecollar', 'entrepreneur', 'housemaid', 'management', 'retired', 'self-employed', 'services', 'student', 'technician', 'unemployed', 'unknown')
loan	boolean	Apakah memiliki pinjaman pribadi? ('no', 'yes', 'unknown')
marital	string	Status Pernikahan ('divorced', 'married', 'single', 'unknown')
month	string	Bulan terakhir kontak dalam setahun ( 'jan', 'feb', 'mar', ..., 'nov', 'dec')
pdays	double	Jumlah hari yang berlalu setelah klien terakhir dihubungi dari kampanye sebelumnya.
poutcome	string	Hasil dari kampanye pemasaran sebelumnya ('failure', 'nonexistent', 'success')
previous	double	Jumlah kontak yang dilakukan sebelum kampanye ini dan untuk klien ini.

Dataset yang digunakan dalam penelitian ini terdiri dari beberapa variabel yang menggambarkan karakteristik pelanggan dan riwayat interaksi mereka dengan Bank Uhu. Variabel-variabel ini mencakup umur pelanggan (*Age*) dengan tipe data *double*, jenis pekerjaan (*Job*), status pernikahan (*Marital*), pendidikan terakhir (*Education*), status kredit bermasalah (*Default*) dengan tipe data *varchar* yang merepresentasikan *boolean*, saldo rekening (*Balance*) dengan tipe data *double*, kredit rumah (*Housing*) dengan tipe data *varchar* yang merepresentasikan *boolean*, kepemilikan pinjaman (*Loan*) dengan tipe data *varchar* yang merepresentasikan *boolean*, jenis kontak komunikasi (*Contact*), bulan terakhir kontak (*Month*), hari terakhir kontak (*Day*) dengan tipe data *double*, durasi penawaran (*Duration*) dengan tipe data *double*, jumlah kontak kampanye (*Campaign*) dengan tipe data *double*, jumlah hari setelah kontak terakhir dari kampanye sebelumnya (*Pdays*) dengan tipe data *double*, jumlah kontak sebelum kampanye saat ini (*Previous*) dengan tipe data *double*, hasil dari kampanye pemasaran sebelumnya (*Poutcome*), dan langganan deposito (*Deposit*) dengan tipe data *varchar* yang merepresentasikan *boolean*.

## 2) Distribusi Variabel:

Distribusi dilakukan untuk menganalisis pola dan karakteristik data dalam dataset yang relevan, dengan beberapa variabel yang dianggap penting disoroti sebagai perwakilan analisis distribusi variabel, tanpa semua variabel dibahas dalam Pembahasan, namun, fokus diberikan pada variabel kunci yang dianggap penting untuk memperoleh pemahaman yang mendalam tentang dataset yang digunakan.

### • Variabel Age:



Gambar 4. Distribusi Variabel Age

Dari gambar di atas, teramati bahwa distribusi variabel *Age* menunjukkan kecondongan ke kanan (*Skewed-Right*), yang tercermin dari kecenderungan bar condong ke arah kiri. Median data tersebut terlihat berada pada angka 39, dengan nilai mean sebesar 41.23, dan modus sebesar 31. Dalam konteks Bank Uhu, distribusi variabel *Age* yang condong ke kanan menunjukkan bahwa sebagian besar nasabah atau calon nasabah cenderung berusia muda.

### • Variabel Balance:

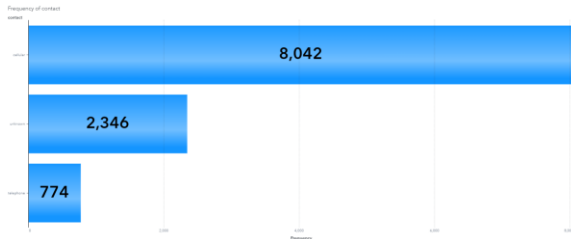


Gambar 5. Distribusi Variabel Marital



Dari gambar di atas, teramati bahwa distribusi variabel *Balance* menunjukkan kecondongan ke kanan (*Skewed-Right*), yang tercermin dari kecenderungan bar condong ke arah kiri. Median data tersebut terlihat berada pada angka 550, dengan nilai mean sebesar 1,528.54. distribusi variabel *Balance* yang *Skewed-Right* menunjukkan bahwa sebagian besar nasabah cenderung memiliki saldo rekening yang relatif rendah, dengan hanya sedikit yang memiliki saldo yang lebih tinggi.

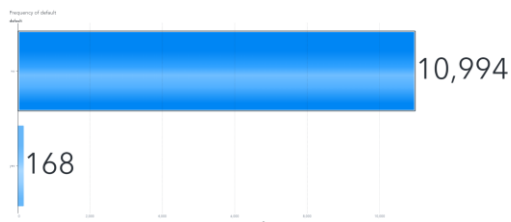
- **Variabel Contact:**



Gambar 6. Distribusi Variabel *Marital*

Dari gambar di atas, dapat diamati bahwa variabel kontak memiliki tiga nilai yang berbeda, yaitu *"Cellular"*, *"Unknown"*, dan *"Telephone"*, dengan masing-masing memiliki jumlah yang berbeda, yaitu 8.042 untuk *"Cellular"*, 2.346 untuk *"Unknown"*, dan 774 untuk *"Telephone"*. Nilai *"Unknown"* dalam variabel ini dapat mewakili berbagai jenis kontak, seperti melalui email atau kontak langsung.

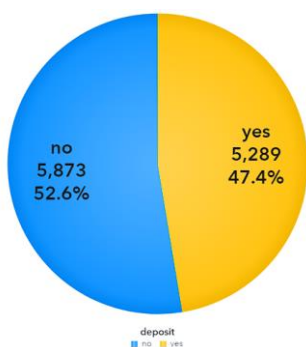
- **Variabel Default:**



Gambar 7. Distribusi Variabel *Default*

Dari gambar di atas, dapat diamati bahwa variabel *"Default"* memiliki dua nilai yang berbeda, yaitu *"yes"* dan *"no"*, dengan masing-masing memiliki jumlah yang berbeda, yaitu 10.994 untuk *"yes"*, dan 168 untuk *"no"*. Variasi nilai *"Default"* menunjukkan mayoritas pelanggan tidak memiliki catatan tunggakan, namun jumlah yang signifikan memiliki catatan tunggakan.

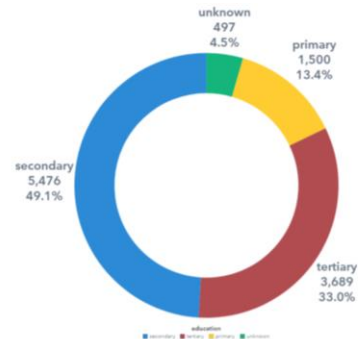
- **Variabel Deposit:**



Gambar 8. Distribusi Variabel *Deposit*

Dari gambar tersebut, terlihat bahwa variabel *"Deposit"* menunjukkan dua nilai yang berbeda, yaitu *"yes"* dan *"no"*, dengan jumlah masing-masing sebesar 5,873 (52.6%) untuk *"yes"*, dan 5,289 (47.4%) untuk *"no"*. Variasi nilai *"Deposit"* menunjukkan tingkat penetrasi produk deposit yang baik di Bank Uhuy

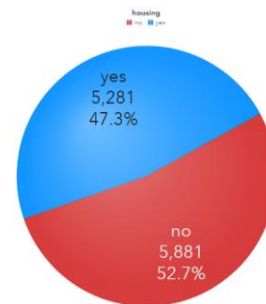
- **Variabel Education:**



Gambar 9. Distribusi Variabel *Education*

Dari gambar di atas, dapat diamati bahwa variabel *"Education"* menampilkan empat nilai yang berbeda, yakni *"Primary"*, *"Secondary"*, *"Tertiary"*, dan *"Unknown"*, dengan jumlah masing-masing sebesar 1,500 (13.4%) untuk *"Primary"*, 5,476 (49.1%) untuk *"Secondary"*, 3,689 (33%) untuk *"Tertiary"*, dan 497 (4.5%) untuk *"Unknown"*. Distribusi nilai dalam variabel *"Education"* menunjukkan mayoritas pelanggan atau prospek Bank Uhuy memiliki tingkat pendidikan sekunder atau tersier, memungkinkan bank untuk menyusun strategi pemasaran yang lebih sesuai dengan profil pendidikan mereka.

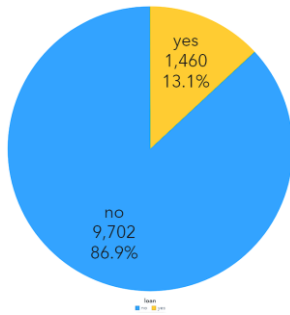
- **Variabel Housing:**



Gambar 10. Distribusi Variabel *Housing*

Dari gambar tersebut, terlihat bahwa variabel *"housing"* menunjukkan dua nilai yang berbeda, yaitu *"yes"* dan *"no"*, dengan jumlah masing-masing sebesar 5,281 (47.3%) untuk *"yes"*, dan 5,881 (52.7%) untuk *"no"*. Variasi nilai dalam variabel *"housing"* menunjukkan bahwa sebagian besar pelanggan atau prospek Bank Uhuy tidak memiliki pinjaman rumah.

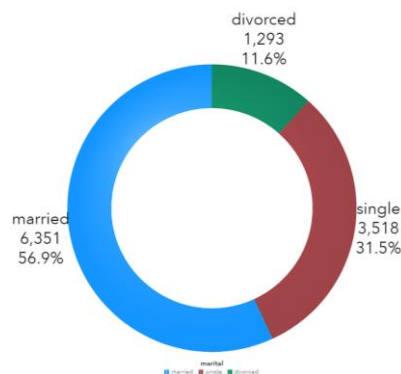
- **Variabel Loan:**



Gambar 11. Distribusi Variabel *Loan*

Dari gambar tersebut, terlihat bahwa variabel "*Loan*" menunjukkan dua nilai yang berbeda, yaitu "*yes*" dan "*no*", dengan jumlah masing-masing sebesar 1,460 (13.1%) untuk "*yes*", dan 9,702 (86.9%) untuk "*no*". Variasi nilai dalam variabel "*Loan*" menunjukkan bahwa sebagian besar pelanggan atau prospek Bank UhuY tidak memiliki pinjaman.

- *Variabel Marital*:



Gambar 12. Distribusi Variabel *Marital*

Dari gambar di atas, dapat diamati bahwa variabel "*Marital*" menampilkan empat nilai yang berbeda, yakni "*Married*", "*Single*", dan "*Divorced*", dengan jumlah masing-masing sebesar 6,351 (56.9%) untuk "*Married*", 3,518 (31.5%) untuk "*Single*", dan 1,293 (11.6%) untuk "*Divorced*". Variasi nilai dalam variabel "*Marital*" menunjukkan distribusi status pernikahan pelanggan atau prospek di Bank UhuY, dengan mayoritas merupakan pasangan yang menikah.

### C. Data Preparation

Tahap Data Preparation bertujuan untuk Menyiapkan dan mengolah data yang akan digunakan dalam analisis kampanye pemasaran Bank UhuY. Proses ini melibatkan beberapa tahapan esensial:

#### 1) Check Missing Values

Pada langkah ini, pengecekan *Missing Values* dilakukan terhadap seluruh variabel dalam dataset ini. Pengecekan ini bertujuan untuk memastikan kualitas dan kelengkapan data yang akan digunakan dalam analisis. Dengan identifikasi dan penanganan data yang hilang, diharapkan bias atau kesalahan dalam hasil analisis dapat diminimalisir.

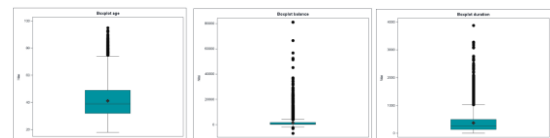
Variable	N	N Miss
age	11162	0
balance	11162	0
day	11162	0
duration	11162	0
campaign	11162	0
pdays	11162	0
previous	11162	0

Gambar 13. Cek *Missing Values*

Pada gambar di atas, terlihat bahwa tidak ada missing values pada dataset ini, sehingga data siap digunakan untuk analisis lebih lanjut tanpa memerlukan penanganan tambahan terhadap data yang hilang.

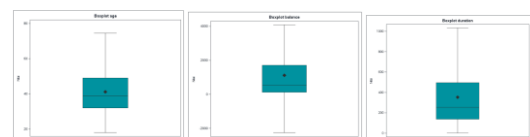
#### 2) Handling Outlier

Pada langkah ini, pengecekan outlier dilakukan dengan tujuan untuk diidentifikasi dan ditangani data yang memiliki nilai ekstrim. Proses ini penting karena outlier dapat mempengaruhi hasil analisis secara signifikan dan menyebabkan bias dalam model prediktif.



Gambar 14. Sebelum outlier ditangani

Pada gambar di atas, tampak bahwa terdapat outlier yang signifikan pada variabel *age*, *balance*, dan *duration*. Penanganan outlier dilakukan menggunakan metode IQR, dengan hasil akhir berupa pemindahan outlier ke nilai mean. Langkah ini dilakukan untuk mengurangi dampak negatif dari nilai ekstrim tersebut, sehingga analisis dapat menghasilkan model yang lebih akurat dan bebas bias.

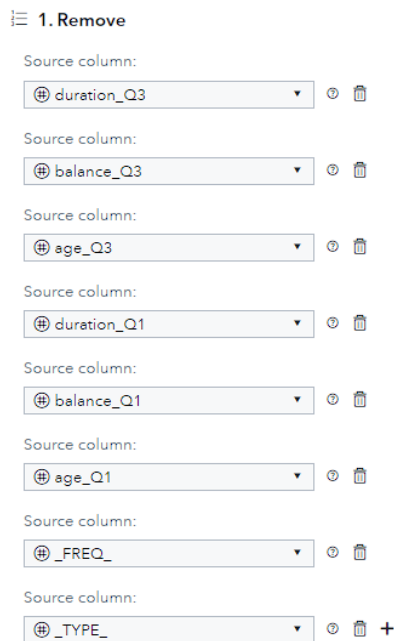


Gambar 15. Setelah penanganan outlier

Setelah outlier ditangani, dapat terlihat bahwa distribusi data menjadi lebih terkonsentrasi di sekitar nilai mean atau median, dengan rentang nilai yang lebih konsisten dan tidak ada nilai ekstrim yang mencolok.

#### 3) Feature Selection

Setelah menangani pemeriksaan Missing Values dan Outlier, terdapat kolom tambahan yang dimasukkan ke dalam dataset yang dibutuhkan untuk langkah sebelumnya. Namun, untuk langkah berikutnya dalam analisis, kolom-kolom tersebut tidak lagi diperlukan. Oleh karena itu, diperlukan pemilihan fitur atau variabel yang relevan untuk dipertahankan dalam proses analisis berikutnya. Proses ini bertujuan untuk menyaring variabel yang paling berpengaruh atau relevan terhadap tujuan analisis yang ditetapkan, sehingga dapat meningkatkan efisiensi dan akurasi model yang dibangun.



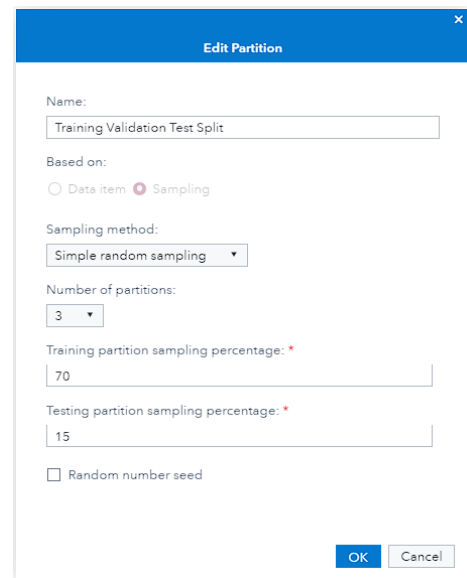
Gambar 16. Menghapus fitur yang tidak penting

Penghapusan kolom "Duration\_Q3", "Balance\_Q3", "Age\_Q3", "Duration\_Q1", "Balance\_Q1", "Age\_Q1", "\_FREQ\_", dan "\_Type\_" dari dataset menghasilkan pengurangan dimensi dataset, yang pada gilirannya mengurangi kompleksitas analisis data. Dengan mengurangi jumlah variabel yang tidak relevan atau tidak diperlukan, proses analisis menjadi lebih efisien dan fokus pada fitur-fitur yang lebih signifikan untuk mencapai tujuan analisis yang ditetapkan. Hal ini juga dapat meningkatkan kinerja model yang dibangun dengan mengurangi kelebihan variabilitas yang tidak memberikan kontribusi signifikan terhadap hasil analisis.

#### D. Modelling

Pada tahap pemodelan untuk membangun dan mengevaluasi model prediktif deposit pelanggan atas *direct marketing* di Bank Uhu. Variabel response yang digunakan adalah 'Deposit', dengan 'yes', dan 'No', sementara variabel predictor meliputi 'Contact', 'default', 'education', 'housing', 'job', 'loan', 'marital', 'month', 'poutcome', 'age', 'balance', 'campaign', 'day', 'duration', 'previous', dan 'pdays'. Empat algoritma yang diuji adalah *Random Forest*, *Decision Tree*, *Support Vector Machine (SVM)*, dan *Neural Network*.

Pada tahap ini, dilakukan pemisahan dataset menjadi tiga bagian: data pelatihan (*training data*), data validasi (*validation data*), dan data pengujian (*test data*) dengan tujuan untuk memastikan bahwa model yang dibangun dapat diuji dan divalidasi dengan baik, sehingga performanya dapat diukur dengan lebih akurat.



Gambar 17. Train Validation Test Split

Dalam gambar terlihat bahwa pemisahan dilakukan ke dalam tiga bagian, dengan proporsi 70% untuk data training, 15% untuk data testing, dan 15% untuk data validation. Hal ini membantu dalam penyesuaian parameter oleh mesin dan pembelajaran pola-pola dalam data training, diikuti oleh pengujian kinerja model pada data testing, serta validasi hasilnya pada data validation.

#### E. Evaluation

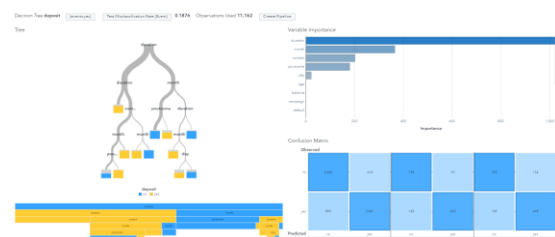
Tahap selanjutnya, evaluasi dilakukan menggunakan metrik kinerja seperti tingkat *misclassification Rate*, *Confusion Matrix*, dan *F1-score*. Hasil analisis dan prediksi disajikan dalam bentuk perbandingan model dan dasbor interaktif di *SAS Visual Analytics*, membantu dalam mengenali pola deposit dan faktor yang mempengaruhinya. Dengan pendekatan ini, peneliti dapat memastikan pemilihan model terbaik untuk digunakan dalam prediksi klasifikasi hasil *direct marketing*.

#### F. Deployment

Pola *Direct Marketing* yang telah dibangun dan divalidasi dengan predictor *Deposit* menggunakan *SAS Visual Analytics* akan diimplementasikan ke dalam *Marketing*. Hasil prediksi *Direct Marketing* akan digunakan oleh tim manajemen untuk menganalisa pola marketing, seperti melakukan penawaran saat tertentu atau dengan melihat data pelanggan yang berpotensi.

## L. IV. HASIL DAN PEMBAHASAN

### A. Decision Tree



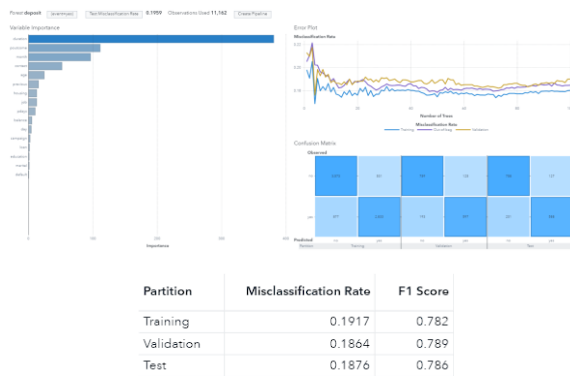
Partition	Misclassification Rate	F1 Score
Training	0.1911	0.789
Validation	0.1864	0.796
Test	0.1906	0.790



Gambar 18. Modelling *Decision Tree*

Pada pengujian dengan model *Decision Tree*, yang menghasilkan *Misclassification Rate* sebesar 0.1965 dan *F1-Score* sebesar 0.785 pada data test, terdapat hasil yang konsisten dengan *confusion matrix*. *Misclassification Rate* yang rendah, ditunjukkan oleh jumlah *TP* dan *TN* yang tinggi, menandakan bahwa model jarang melakukan kesalahan dalam prediksi, seperti yang terlihat pada data training dengan *TP* sebesar 3541 dan *TN* sebesar 2807. Selain itu, *F1-Score* yang tinggi menunjukkan keseimbangan yang baik antara *precision* dan *recall*, sesuai dengan data *validation* yang menunjukkan *TP* sebesar 766 dan *TN* sebesar 576. Dengan hasil tersebut, model *Decision Tree* mempertahankan performa yang konsisten dan mampu mengklasifikasikan data dengan baik, sebagaimana yang tercermin dalam hasil pengujian tes dengan *TP* sebesar 759 dan *TN* sebesar 582.

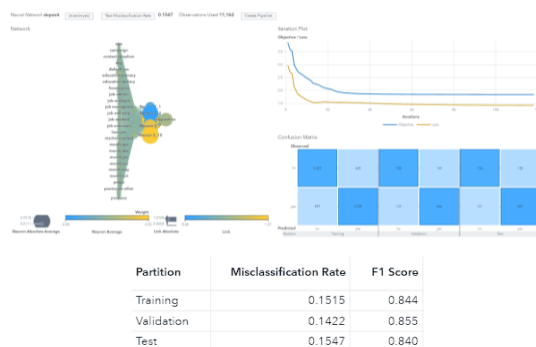
## B. Random Forest



Gambar 19. Modelling *Random Forest*

Pada pengujian dengan model *Random Forest*, yang menghasilkan *Misclassification Rate* sebesar 0.1959 dan *F1-Score* sebesar 0.780 pada data test, terdapat korelasi yang konsisten dengan hasil *confusion matrix*. *Misclassification Rate* yang rendah, yang tercermin dari jumlah *TP* dan *TN* yang tinggi, menunjukkan bahwa model jarang melakukan kesalahan dalam prediksi, seperti yang terlihat pada data training dengan *TP* sebesar 3694 dan *TN* sebesar 2616. Selain itu, *F1-Score* yang tinggi menunjukkan keseimbangan yang baik antara *precision* dan *recall*, sebagaimana yang tercermin dalam data *validation* yang menunjukkan *TP* sebesar 792 dan *TN* sebesar 533. Dengan hasil tersebut, model *Random Forest* mempertahankan performa yang konsisten dan mampu mengklasifikasikan data dengan baik, seperti yang diamati dalam hasil pengujian tes dengan *TP* sebesar 788 dan *TN* sebesar 544.

## C. Neural Network



Gambar 20. Modelling *Neural Network*

Pada pengujian dengan model *Neural Network*, diperoleh *Misclassification Rate* sebesar 0.1750 dan *F1-Score* sebesar 0.813 pada data test. Hasil ini konsisten dengan analisis dari *confusion matrix*, di mana *Misclassification Rate* yang rendah tercermin dari jumlah *TP* dan *TN* yang tinggi, seperti pada data training dengan *TP* sebesar 3435 dan *TN* sebesar 3205. *F1-Score* yang tinggi juga menunjukkan keseimbangan yang baik antara *precision* dan *recall*, sebagaimana yang diamati pada data *validation* dengan *TP* sebesar 738 dan *TN* sebesar 664. Dengan demikian, model *Neural Network* menunjukkan performa yang baik dan konsisten, sebagaimana terlihat dalam hasil pengujian tes dengan *TP* sebesar 728 dan *TN* sebesar 687.

## D. Support Vector Machine



Gambar 21. Modelling *Support Vector Machine*

Pada pengujian dengan model *Support Vector Machine* (SVM), didapatkan *Misclassification Rate* sebesar 0.1798 dan *F1-Score* sebesar 0.806 pada data test. Hasil tersebut mencerminkan performa yang baik dari model, yang diperkuat oleh analisis *confusion matrix*. *Misclassification Rate* yang rendah, seperti terlihat pada data training dengan jumlah *FP* dan *FN* yang relatif kecil (masing-masing 667 dan 628), menunjukkan bahwa model jarang membuat kesalahan dalam prediksi. Hal ini juga tercermin dalam *F1-Score* yang tinggi, mengindikasikan keseimbangan yang baik antara *precision* dan *recall*. Pada data *validation*, ditemukan *Misclassification Rate* sebesar 0.1786 dan *F1-Score* sebesar 0.810, menunjukkan konsistensi performa model selama proses validasi. Hasil pengujian tes juga menegaskan kualitas model SVM, dengan *TP* sebesar 655 dan *TN* sebesar 734. Dengan demikian, model SVM menunjukkan kinerja yang baik dan konsisten dalam mengklasifikasikan data.

## E. Perbandingan Model



Gambar 23. Deskripsi Model Comparison

Dalam pengujian yang melibatkan empat model utama, yaitu Decision Tree, Random Forest, Neural Network, dan Support Vector Machine (SVM), dilakukan analisis menyeluruh terhadap kinerja masing-masing model. Misclassification Rate sebesar 0.1965 dan F1-Score 0.785 dihasilkan oleh Decision Tree, menunjukkan performa yang konsisten dengan hasil confusion matrix yang menunjukkan jaranganya kesalahan prediksi. Random Forest juga menunjukkan performa yang baik dengan Misclassification Rate 0.1959 dan F1-Score 0.780, serta memiliki keseimbangan yang baik antara precision dan recall. Namun, Neural Network menonjol sebagai model terbaik dengan Misclassification Rate terendah sebesar 0.1750 dan F1-Score tertinggi sebesar 0.813. Hasil ini konsisten dengan analisis confusion matrix yang menunjukkan performa konsisten dan tingginya akurasi prediksi. Model SVM juga menunjukkan performa yang baik dengan Misclassification Rate sebesar 0.1798 dan F1-Score sebesar 0.806, serta konsistensi dalam hasil confusion matrix.

Secara keseluruhan, Neural Network adalah model yang paling unggul dalam mengklasifikasikan data, diikuti oleh SVM, Decision Tree, dan Random Forest secara berurutan. Evaluasi ini memberikan panduan yang solid dalam pemilihan model yang dapat diandalkan untuk aplikasi praktis, terutama yang membutuhkan tingkat akurasi tinggi dalam pengklasifikasian data.

#### F. Pembahasan

Dalam tahap evaluasi, empat model utama - Decision Tree, Random Forest, Neural Network, dan Support Vector Machine (SVM) - telah dianalisis secara menyeluruh untuk menilai kemampuan mereka dalam mengklasifikasikan data. Dengan menggunakan metrik seperti Misclassification Rate dan F1-Score, pemahaman yang mendalam tentang kinerja masing-masing model diperoleh. Hasilnya menunjukkan bahwa Neural Network memiliki performa paling baik, dengan Misclassification Rate terendah dan F1-Score tertinggi, menunjukkan keseimbangan yang optimal antara precision dan recall. Evaluasi ini memberikan panduan yang solid untuk pemilihan model yang andal dalam aplikasi praktis yang membutuhkan tingkat akurasi tinggi dalam pengklasifikasian data.

Setelah membandingkan kinerja model, langkah selanjutnya adalah menerapkan model terbaik dalam lingkungan produksi atau operasional. Model Neural Network terpilih untuk diimplementasikan oleh Bank Uhuy karena kinerjanya yang superior. Integrasi model ini dengan sistem informasi Bank Uhuy akan memungkinkan prediksi hasil kampanye pemasaran secara otomatis, membantu bank membuat keputusan yang lebih informatif dalam merancang strategi pemasaran yang efektif. Pengembangan alat atau platform khusus akan memungkinkan tim pemasaran dan manajemen Bank Uhuy untuk mengakses dan memanfaatkan hasil prediksi dengan mudah, mendukung pengambilan keputusan strategis. Dengan demikian, implementasi model Neural Network akan membantu Bank Uhuy meningkatkan efektivitas kampanye pemasaran mereka, mencapai tujuan bisnis yang ditetapkan, dan tetap bersaing dalam industri perbankan.

#### M. V. KESIMPULAN

##### A. Kesimpulan

Pada penelitian ini, dilakukan evaluasi terhadap hasil kampanye pemasaran Bank Uhuy menggunakan pendekatan metodologi CRISP-DM. Berbagai model machine learning, termasuk Decision Tree, Random Forest, Neural Network, dan Support Vector Machine, telah diuji untuk memprediksi hasil kampanye pemasaran berdasarkan berbagai faktor seperti kontak pelanggan, kondisi keuangan, dan sejarah interaksi sebelumnya.

Dari hasil evaluasi, dapat disimpulkan bahwa model Neural Network menunjukkan kinerja yang paling baik dalam mengklasifikasikan data, dengan Misclassification Rate terendah dan F1-Score tertinggi dibandingkan dengan model lainnya. Hal ini menunjukkan bahwa pendekatan machine learning, terutama Neural Network, dapat menjadi alat yang efektif dalam meningkatkan efektivitas kampanye pemasaran Bank Uhuy.

Selain itu, penerapan model ini dalam lingkungan produksi atau operasional dapat membantu Bank Uhuy dalam mengambil keputusan yang lebih informatif dan merancang strategi pemasaran yang lebih efektif. Integrasi model ini dengan sistem informasi Bank Uhuy akan memungkinkan pembuatan prediksi secara otomatis berdasarkan data pelanggan yang tersedia, sehingga memperkuat proses pengambilan keputusan strategis.

#### B. Saran

Dalam konteks saran untuk Bank Uhuy, disarankan agar terus dilakukan pemantauan dan evaluasi terhadap kinerja model Neural Network setelah implementasi. Hal ini dianggap penting untuk memastikan bahwa model tetap efektif dalam menghadapi perubahan lingkungan bisnis dan respons pelanggan yang mungkin terjadi. Selain itu, pertimbangan untuk pengembangan model yang lebih kompleks dan canggih di masa depan dapat dipertimbangkan oleh Bank Uhuy guna meningkatkan tingkat akurasi dan prediktabilitas kampanye pemasaran mereka.

Untuk penelitian lebih lanjut, direkomendasikan untuk dilibatkannya analisis yang lebih mendalam tentang faktor-faktor tambahan yang dapat mempengaruhi efektivitas kampanye pemasaran, seperti preferensi pelanggan yang lebih spesifik, tren pasar yang berkembang, dan faktor-faktor eksternal lainnya. Selain itu, penelitian tentang penggunaan teknik machine learning yang lebih lanjut dalam konteks pemasaran perbankan dapat memberikan wawasan yang lebih luas tentang bagaimana model-model tersebut dapat diterapkan secara optimal untuk mencapai tujuan bisnis yang diinginkan.

#### REFERENCES

- [1] A. M. Rafli, "Penjualan Langsung (Direct Selling), Tugas, dan Contohnya," *Mekari Jurnal*, Dec. 12, 2023. <https://www.jurnal.id/blog/pengertian-penjualan-langsung-direct-selling-contohnya-sbc/>
- [2] M. Rosyida, "Direct Marketing adalah Strategi Mudah Untuk Meraih Pelanggan Dalam Bisnis - DomaiNesia," *DomaiNesia*, Feb. 25, 2024. <https://www.domainesia.com/berita/direct-marketing-adalah/#:~:text=Direct%20marketing%20adalah%20strategi%20pemasaran,%2C%20sosial%20media%2C%20dan%20seba> gainya.
- [3] D. Amalia, "Riset Pemasaran: Fungsi, Metode dan Langkah Penyusunan," *Mekari Jurnal*, Oct. 13, 2023. <https://www.jurnal.id/blog/fungsi-langkah-penyusunan-riset-pemasaran-atau-riset-pasar-adalah/>
- [4] M. G. Haque-Fawzi, A. S. Iskandar, H. Erlangga, Nurjaya, and D. Sunarsi, *STRATEGI PEMASARAN Konsep, Teori dan Implementasi*. Pascal Books, 2022.

- [5] A. Asia, "Mengenal Strategi Pemasaran Berbasis Data untuk Pengambilan Keputusan Bisnis," ADA, Apr. 23, 2024. <https://www.ada-asia.com/insights/id-strategi-pemasaran-berbasis-data>
- [6] J. Jatmiko, "Komunikasi pemasaran sebagai strategi memperluas pasar," KOMUNIKOLOGI : Jurnal Ilmiah Ilmu Komunikasi, vol. 9, no. 2, Jan. 2012, [Online]. Available: <https://ejurnal.esaunggul.ac.id/index.php/Kom/article/download/T016/944>
- [7] T. Mauritsius and F. Binsar, "Cross-Industry Standard Process for Data Mining (CRISP-DM)," MMSI BINUS University, Sep. 18, 2020. <https://mmsi.binus.ac.id/2020/09/18/cross-industry-standard-process-for-data-mining-crisp-dm/>
- [8] I. Y. Musyawarah and D. Idayanti, "Forecasting: Jurnal ilmiah Ilmu Manajemen," Analisis Strategi Pemasaran Untuk Meningkatkan Penjualan Pada Usaha Ibu Bagas Di Kecamatan Mamuju, vol. 1, no. 1, 2022, [Online]. Available: <https://stiemmamuju.e-journal.id/FJIIM/article/download/83/40>
- [9] R. Maulid, "Teknik Analisis data CRISP-DM dalam data mining," DQLab.id, Nov. 07, 2022. <https://dqlab.id/teknik-analisis-data-crisp-dm-dalam-data-mining>
- [10] I. K. Sukesha, "CRISP DM Sebagai Salah Satu Standard untuk Menghasilkan Data Driven Decision Making yang Berkualitas," DJKN KEMENKEU, Jun. 22, 2022. <https://www.djkn.kemenkeu.go.id/artikel/baca/15134/CRISP-DM-Sebagai-Salah-Satu-Standard-untuk-Menghasilkan-Data-Driven-Decision-Making-yang-Berkualitas.html>
- [11] D. Feby, "Apa itu decision tree di machine learning model?," DQLab, May 16, 2023. <https://dqlab.id/apa-itu-decision-tree-di-machine-learning-model>
- [12] S. Misra and H. Li, "Noninvasive fracture characterization based on the classification of sonic wave travel times," in Elsevier eBooks, 2020, pp. 243–287. doi: 10.1016/b978-0-12-817736-5.00009-0.
- [13] N. Anisa, "Mengenal artificial neural network," School of Information Systems, Feb. 14, 2022. <https://sis.binus.ac.id/2022/02/14/mengenal-artificial-neural-network/>
- [14] GeeksforGeeks, "Support Vector Machine (SVM) algorithm," GeeksforGeeks, Jun. 10, 2023. <https://www.geeksforgeeks.org/support-vector-machine-algorithm/>
- [15] P. A. Octaviani, Y. Wilandari, and D. Ispriyanti, "PENERAPAN METODE KLASIFIKASI SUPPORT VECTOR MACHINE (SVM) PADA DATA AKREDITASI SEKOLAH DASAR (SD) DI KABUPATEN MAGELANG," Jurnal Gaussian, vol. 3, no. 4, pp. 811–820, Oct. 2014, doi: 10.14710/j.gauss.v3i4.8092.
- [16] Z. Bobbitt, "Misclassification rate in Machine Learning: definition & example," Statology, Mar. 25, 2022. <https://www.statology.org/misclassification-rate/>
- [17] M. S. ANGGREANY, "Confusion Matrix," School of Computer Science, Nov. 01, 2020. <https://socs.binus.ac.id/2020/11/01/confusion-matrix/>
- [18] S. Narkhede, "Understanding Confusion Matrix - towards Data science," Medium, Jun. 15, 2021. [Online]. Available: <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>