

## Prediksi *Churn* Pelanggan *Multinational Bank* Menggunakan Algoritma *Machine Learning*

M Ainur Syawaludin<sup>1</sup>, Rifki Hidayat<sup>2</sup>, Nurmalitasari<sup>3</sup>

<sup>1,2,3</sup>Universitas Duta Bangsa

Jl. Bhayangkara No.55 Tipes, Surakarta, Indonesia

e-mail: <sup>1</sup>mainurs2112@gmail.com, <sup>2</sup>ripkihidayat@gmail.com, <sup>3</sup>nurmalitasari@udb.ac.id

Artikel Info : Diterima : 08-07-2024 | Direvisi : 07-10-2024 | Disetujui : 01-12-2024

**Abstrak** - Dalam menghadapi persaingan pasar yang ketat, prediksi *churn* pelanggan menjadi penting bagi perusahaan perbankan untuk mempertahankan loyalitas pelanggan. Penelitian ini mengaplikasikan algoritma *machine learning* meliputi *Naive Bayes*, *Decision Tree*, dan *Random Forest* untuk prediksi *churn* pelanggan pada ABC *Multinational Bank*. Data yang digunakan adalah *dataset* publik yang diambil dari *Kaggle* yang mencakup informasi 10.000 nasabah bank. Proses penelitian melibatkan beberapa tahapan yaitu pengumpulan data, *preprocessing*, pemodelan, prediksi, dan evaluasi. Hasil evaluasi memperlihatkan bahwa model *Random Forest* memberikan performa terbaik dengan akurasi 85% dan AUC 0.83. *Naive Bayes* dan *Decision Tree* masing-masing memiliki akurasi 82% dan 77%. Kesimpulan menunjukkan bahwa *Random Forest* lebih unggul dalam memprediksi *churn* pelanggan dibandingkan dua algoritma lainnya, sehingga dapat digunakan untuk strategi pemasaran yang lebih efektif dalam industri perbankan.

**Kata Kunci** : Prediksi *Churn* Pelanggan, *Machine Learning*

**Abstracts** - In facing intense market competition, predicting customer churn is important for banking companies to maintain customer loyalty. This research applies machine learning algorithms, namely *Naive Bayes*, *Decision Tree*, and *Random Forest*, to predict customer churn at ABC *Multinational Bank*. The data used is a public dataset taken from *Kaggle*, which includes information on 10,000 bank customers. The research process involves several stages, namely data collection, *preprocessing*, modeling, prediction, and evaluation. The evaluation results show that the *Random Forest* model provides the best performance with an accuracy of 85% and an AUC of 0.83. *Naive Bayes* and *Decision Tree* have 82% and 77% accuracy, respectively. The conclusion shows that *Random Forest* is superior in predicting customer churn compared to the other two algorithms so that it can be used for more effective marketing strategies in the banking industry.

**Keywords** : Prediction, Customer Churn, Machine Learning

### PENDAHULUAN

Perkembangan kemajuan teknologi dan persaingan pasar yang ketat, menyebabkan bisnis saat ini diharuskan untuk mengalihkan perhatian perusahaan dari produk ke pelanggan. Hal ini karena pelanggan dianggap sebagai penguasa pasar yang sebenarnya. *Churn* pelanggan yang didefinisikan sebagai kecenderungan pelanggan untuk berhenti berbisnis dengan suatu perusahaan dalam suatu periode tertentu, telah menjadi masalah yang signifikan bagi ABC *Multinational Bank* (Rizki Kurniawan et al., 2023). Perbankan mengalami krisis loyalitas pelanggan dikarenakan persaingan yang ketat dan meningkatnya pilihan pelanggan. Akibatnya, bank dipaksa untuk menjaga hubungan pelanggan dalam jangka panjang. Sangat penting bagi sebuah bisnis untuk memprediksi berapa banyak pelanggan yang meninggalkan perusahaannya. Hal ini dikarenakan mengakuisisi pelanggan baru sering kali lebih mahal daripada mempertahankan pelanggan lama. Memprediksi kemungkinan terjadinya *churn* pelanggan akan menjadi panduan bagi bank dalam mengembangkan strategi pemasaran yang berbeda untuk basis nasabah yang berbeda sambil memenuhi kebutuhan mereka. Bank dapat mengembangkan kampanye pemasaran proaktif untuk menjaga agar nasabah yang sudah ada tidak pergi.

*Machine learning* adalah sub bidang ilmu komputer dan kecerdasan buatan (AI) yang meniru cara manusia belajar dengan menggunakan data dan algoritma, yang secara bertahap menjadi lebih akurat (Kurniawan et al., 2023) (Diana et al., 2023). *Machine learning* sebagian besar digunakan untuk memprediksi



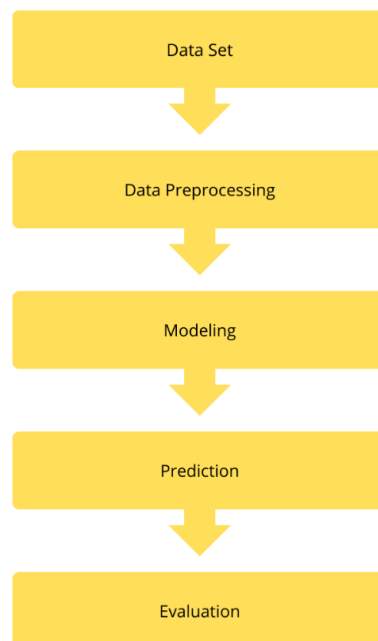
*churn* pelanggan untuk membantu perusahaan memprediksi *churn* pelanggan dengan lebih baik dan menyediakan strategi pemasaran yang ditargetkan dan efektif untuk kemungkinan *churn* (Hon et al., 2023).

Studi tentang prediksi *churn* pelanggan yang dilakukan oleh Miryam Clementine & Arum (2022) menggunakan klasifikasi *Naive Bayes* menunjukkan bahwa *Naive Bayes* memiliki nilai akurasi 85.71% di atas ID3. Pada penelitian sebelumnya yang dilakukan oleh Rahmawati & Larasati (2021) memprediksi persetujuan bank untuk kredit nasabah menggunakan metode *Decision Tree* memperoleh nilai kinerja yang lebih tinggi daripada dengan metode ANN, kinerja akurasi yang diperoleh yaitu 99,36%. Selain itu, penelitian sebelumnya yang dilakukan oleh Husein & Harahap (2021) menggunakan *machine learning* untuk prediksi kemungkinan *churn* pelanggan menggunakan dataset bank seperti Regresi Logistik (LR), *Decision Tree* (DT), *K-Nearest Neighbor* (KNN), dan *Random Forest* (RF). Temuan penelitian menunjukkan bahwa, dari semua model yang digunakan pendekatan *Random Forest* memiliki kinerja terbaik.

Meskipun telah ada penelitian yang mengaplikasikan algoritma ini dalam konteks prediksi *churn* pelanggan, masih terdapat ruang untuk pengembangan lebih lanjut terutama dalam konteks industri perbankan. Penelitian sebelumnya telah menunjukkan bahwa ketiga model ini dapat memberikan hasil yang memuaskan dalam memprediksi *churn* pelanggan. Sehingga, tujuan dari penelitian ini untuk membandingkan performa terbaik model *Naive Bayes*, *Decision Tree*, dan *Random Forest* dalam prediksi *churn* pelanggan pada ABC Multinational Bank.

## METODE PENELITIAN

Penelitian ini menerapkan algoritma machine learning meliputi *Naive Bayes*, *Decision Tree*, dan *Random Forest* untuk prediksi *churn* pelanggan bank. Tahapan metode penelitian terdiri dari 5 tahap meliputi pengumpulan dataset, data *preprocessing*, modeling, prediksi dan evaluasi. Pada penelitian ini terdapat 5 tahapan, karena penelitian ini menggunakan data-data publik (Hasanah, 2023). Pada Gambar 1 menunjukkan tahapan penelitian.



Sumber: Hasil Penelitian (2024)  
Gambar 1. Tahapan Penelitian

Tahapan pertama yaitu penjelasan dataset yang akan digunakan penelitian ini. Setelah itu masuk pada tahapan data *preprocessing* yang akan dilakukan pembersihan data seperti menghapus variabel fitur yang tidak relevan. Tahap ketiga modeling, dataset akan dibagi menjadi 2 data yaitu data latih dan data uji yang melalui pemodelan *Naive Bayes*, *Decision Tree*, dan *Random Forest*.

*Naive Bayes* adalah metode yang mengelompokkan masalah ke dalam kelas-kelas berdasarkan kesamaan dan perbedaan ciri-ciri menggunakan statistik untuk memprediksi suatu kelas (Lestari et al., 2020). Persamaan teorema *Bayes* sebagai berikut.

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (1)$$

Dimana:

- $P(A|B)$  = Probabilitas bersyarat A yang diberikah oleh B
- $P(B|A)$  = Probabilitas bersyarat B yang diberikah oleh A
- $P(A)$  = Probabilitas independen A
- $P(B)$  = Probabilitas independen B

Algoritma *Decision Tree* merupakan teknik klasifikasi yang direpresentasikan dalam bentuk struktur pohon, menyerupai diagram alir pada diagram alur. Setiap *node* internal dalam struktur ini sesuai dengan pengujian yang dilakukan pada atribut, sementara setiap cabang menandakan hasil pengujian. Selain itu, simpul atau daun di dalam pohon mewakili kelas atau distribusi kelas tertentu (Marutho, 2019). Untuk menghitung entropi sampel  $s$  dapat dinyatakan menggunakan rumus sebagai berikut.

$$Entropy(S) = \sum_i^c -p_i \log_2 P_i \quad (2)$$

Dimana:

- $c$  = Jumlah kelas klasifikasi.
- $p_i$  = Jumlah proporsi peluang untuk kelas  $i$ .

*Random Forest* adalah teknik dalam bidang pembelajaran ansambel, secara khusus berfokus pada penggabungan atau *bagging*, yang melibatkan pembuatan beberapa pohon yang berasal dari data sampel. Pohon-pohon ini secara kolektif membentuk pohon tunggal di mana proses pelatihan tidak bergantung pada pohon sebelumnya. Pada akhirnya, Keputusan akhir dibuat berdasarkan mayoritas suara dari masing-masing pohon. *Random Forest* memberikan pengukuran setiap variabel atau *feature* yang terpenting untuk digunakan sebagai *predictor* agar menghasilkan akurasi yang tinggi. Algoritma *Random Forest* dapat diterapkan pada data dengan dimensi yang tinggi (Satria et al., 2023).

Tahap selanjutnya yaitu prediksi antara nilai aktual dan nilai prediksi. Setelah itu, hasil akhir dari proses *machine learning* ini akan dievaluasi melalui pemanfaatan analisis AUC, CA, F1, *Precision*, *Recall*, *Confusion Matrix*, dan ROC.

## HASIL DAN PEMBAHASAN

### 1. Dataset

Penelitian ini menggunakan data yang terdiri dari data *churn* nasabah bank bersumber dari *website Kaggle*, yang secara khusus diidentifikasi sebagai *Bank Customer Churn Dataset* yang tersedia di URL berikut: <https://www.kaggle.com/datasets/gauravtopre/bank-customer-churn-dataset>. Dataset ini berisi informasi 10.000 nasabah bank dari negara-negara Eropa yang dapat dilihat pada Gambar 2. Dengan fitur target menjadi variabel biner yang menunjukkan apakah nasabah telah meninggalkan bank atau masih menjadi nasabah. Fitur target akan menjadi nilai 1 ketika nasabah telah meninggalkan bank selama suatu periode. Ketika nasabah tidak meninggalkan bank, pada fitur target nilai 0 akan ditampilkan. Dataset berisi dua belas fitur yang berasal dari data pelanggan dan transaksi yang diproses pelanggan. Tabel 1 memberikan rincian lebih lanjut tentang fitur-fitur ini.

	customer_id	credit_score	country	gender	age	tenure	balance	products_number	credit_card	active_member	estimated_salary	churn
0	15634602	619	France	Female	42	2	0.00	1	1	1	101348.88	1
1	15647311	608	Spain	Female	41	1	83807.86	1	0	1	112542.58	0
2	15619304	502	France	Female	42	8	159660.80	3	1	0	113931.57	1
3	15701354	699	France	Female	39	1	0.00	2	0	0	93826.63	0
4	15737888	850	Spain	Female	43	2	125510.82	1	1	1	79084.10	0
...	...	...	...	...	...	...	...	...	...	...	...	...
9995	15606229	771	France	Male	39	5	0.00	2	1	0	96270.64	0
9996	15569892	516	France	Male	35	10	57369.61	1	1	1	101699.77	0
9997	15584532	709	France	Female	36	7	0.00	1	0	1	42085.58	1
9998	15682355	772	Germany	Male	42	3	75075.31	2	1	0	92888.52	1
9999	15628319	792	France	Female	28	4	130142.79	1	1	0	38190.78	0

Sumber: Kaggle (2022)  
 Gambar 2. Dataset

Tabel 1. Deskripsi Dataset

Nama Fitur	Deskripsi Fitur
<i>customer_id</i>	ID unik untuk identifikasi nasabah bank
<i>credit_score</i>	Skor kredit pelanggan

<i>country</i>	Negara asal pelanggan
<i>gender</i>	Jenis kelamin laki-laki atau perempuan
<i>age</i>	Usia pelanggan
<i>tenure</i>	Jumlah tahun nasabah telah menjadi pemegang rekening Bank ABC
<i>balance</i>	Saldo rekening nasabah bank
<i>products_number</i>	Jumlah produk bank yang digunakan nasabah (rekening Tabungan, <i>mobile banking</i> , <i>internet banking</i> , dll)
<i>credit_card</i>	Tipe data bilangan biner untuk mengetahui apakah nasabah memegang kartu kredit di bank atau tidak
<i>active_member</i>	Tipe data bilangan biner untuk mengetahui apakah nasabah merupakan anggota aktif bank atau bukan
<i>estimated_salary</i>	Perkiraan gaji pelanggan dalam dolar
<i>churn</i>	Tipe data bilangan biner 1 jika nasabah telah meninggalkan bank selama suatu periode dan 0 jika nasabah dipertahankan

Sumber: Hasil Penelitian (2024)

### 2. Data Preprocessing

Pada tahap ini, sebelum melakukan implementasi model data harus diproses ke dalam format yang dapat ditangani algoritma *machine learning*. Ada beberapa fitur dalam dataset nasabah bank yang tidak diperlukan untuk prediksi penelitian ini. Oleh karena itu, data akan dibersihkan dengan menghapus fitur yang tidak relevan. Fitur tersebut meliputi *customer\_id*, *country*, dan *gender*. Tabel 2 menunjukkan hasil *preprocessing data* dan fitur ini akan digunakan dalam penelitian untuk mengevaluasi prediksi *churn* pelanggan.

Tabel 2. Deskripsi Hasil *Preprocessing* Dataset

Nama Fitur	Min	Max	Mean	STD
credit score	350	850	650,5288	96,6532
age	18	92	38,9218	10,4878
tenure	0	10	5,0128	2,8921
balance	0	250898,09	76485,8892	62397,4052
products number	1	4	1,5302	0,5816
credit card	0	1	0,7055	0,4558
active member	0	1	0,5151	0,4997
estimated salary	11,58	199992,48	100090,2398	57510,4928
churn	0	1	0,2037	0,4027

Sumber: Hasil Penelitian (2024)

### 3. Modeling dan Prediction

Pada penelitian ini, perbandingan akan dilakukan di antara tiga model meliputi *Naive Bayes*, *Decision Tree*, dan *Random Forest*. Proses yang melibatkan pengembangan, pelatihan, dan evaluasi data dijalankan menggunakan bahasa pemrograman *Python* dengan bantuan *JupyterLab* dan *library Scikit-learn*. Dataset akan menjalani partisi menjadi 2 sub set data terpisah yaitu data latih dan data uji, dengan rasio *split* 7:3 yang bertujuan untuk mencapai akurasi optimal. Partisi data dapat diamati dalam Tabel 3.

Tabel 3. Partisi Dataset

Komposisi Data	Data Latih	Data Uji	Jumlah Keseluruhan Data
7 : 3	7000	3000	10000

Sumber: Hasil Penelitian (2024)

Pada penelitian ini, konsep *churn* dilambangkan oleh nilai biner 1, sedangkan *non-churn* dilambangkan dengan nilai biner 0 di seluruh dataset. Tabel 4 menampilkan contoh data yang menunjukkan deskripsi *churn* dan *non-churn*.

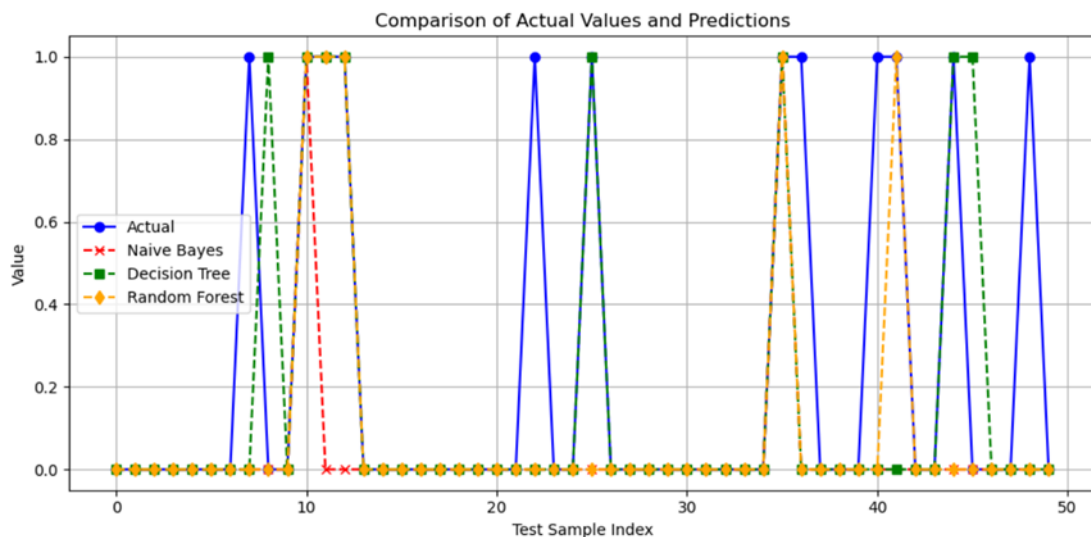
Tabel 4. Deskripsi Label

Label	Jumlah Data	Persentase
<i>churn</i>	2037	20,37%
<i>non-churn</i>	7963	79,63%

Sumber: Hasil Penelitian (2024)

Karena penelitian ini terdapat data yang sangat banyak. Maka, pada penerapan algoritma *Naïve Bayes* cukup menggunakan model *GaussianNB*. Selain itu, penerapan algoritma *Decision Tree* juga cukup menggunakan model *DecicisonTreeClassifier*. Demikian pula, pada penerapan algoritma *Random Forest* cukup menggunakan model *RandomForestClassifier* yang sudah tersedia pada *library Scikit-learn*.

Pada tahap prediksi ini akan menunjukkan grafik perbandingan antara nilai aktual dan nilai prediksi setiap sampel uji pada masing-masing model pengujian. Tahap ini menggunakan 50 data sampel uji untuk memprediksi. Berikut pada Gambar 3 merupakan hasil perbandingan antara nilai aktual dan nilai prediksi dari semua model.

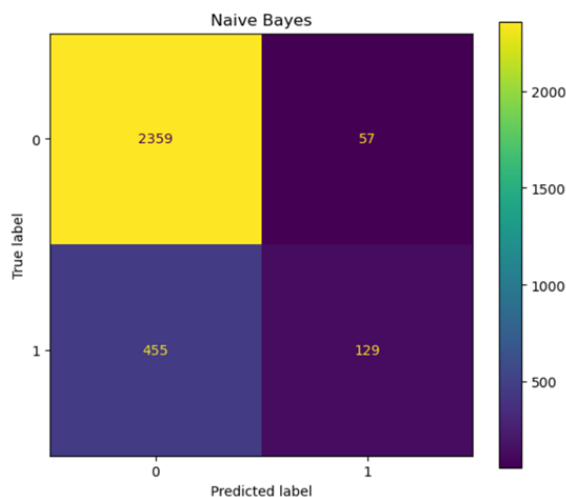


Sumber: Hasil Penelitian (2024)  
 Gambar 3. Grafik Perbandingan Nilai Aktual dan Prediksi

Dapat dilihat pada perbandingan gambar di atas sebagian besar prediksi mendekati nilai aktual, meskipun masih terdapat beberapa perbedaan di beberapa titik. Untuk kesalahan prediksi terdapat pada titik-titik di mana tanda setiap metode tidak bertepatan dengan tanda lingkaran.

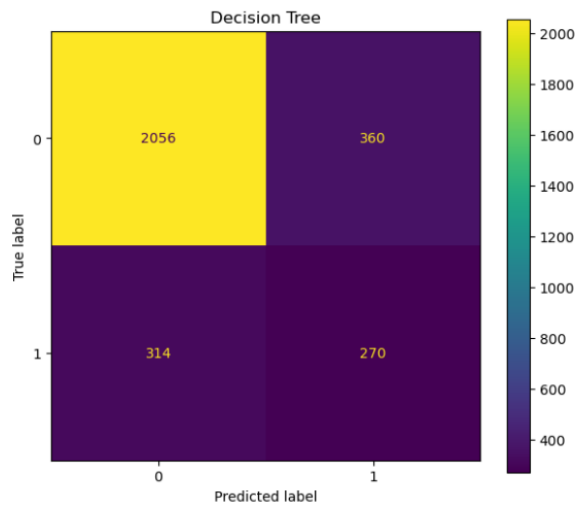
#### 4. Evaluation

Tahap evaluasi penelitian ini melibatkan penggunaan berbagai metrik seperti AUC, CA, F1, *Precision*, *Recall*, *Confusion Matrix*, dan analisis ROC. AUC digunakan untuk menggambarkan probabilitas dalam penelitian ini. CA berperan dalam menentukan keakuratan dataset yang sedang dipertimbangkan. *Precision* menunjukkan keakuratan data dengan mengizinkan hanya dua kejadian yaitu, 1 dan 0. *Recall* digunakan untuk mengukur rasio dalam dataset. Metrik F1 melibatkan perbandingan antara *recall* dan *precision*.



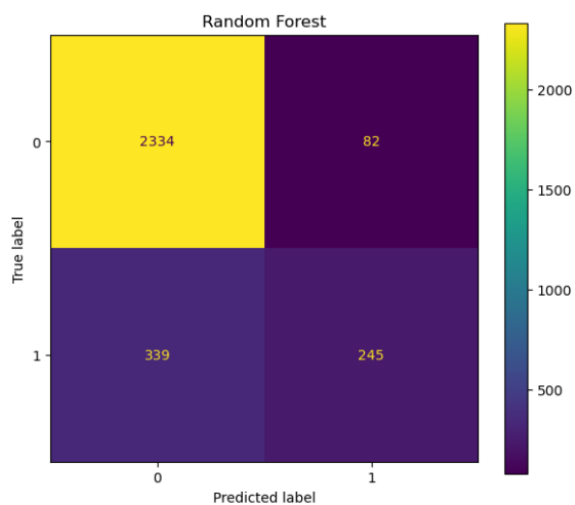
Sumber: Hasil Penelitian (2024)  
 Gambar 4. *Confusion Matrix Naive Bayes*

Seperti yang ditunjukkan pada Gambar 4 algoritma *Naive Bayes* memiliki 2359 + 129 prediksi akurat dan 57 + 455 prediksi salah yang terdiri dari total 3000 data pengujian. Tingkat akurasi algoritma dapat ditentukan dengan cara berikut  $2488/3000 * 100\% = 82\%$ .



Sumber: Hasil Penelitian (2024)  
 Gambar 5. *Confusion Matrix Decision Tree*

Seperti yang ditunjukkan pada Gambar 5 algoritma *Decision Tree* memiliki 2056 + 270 prediksi akurat dan 360 + 314 prediksi salah yang terdiri dari total 3000 data pengujian. Tingkat akurasi algoritma dapat ditentukan dengan cara berikut  $2326/3000 * 100\% = 77\%$ .



Sumber: Hasil Penelitian (2024)  
 Gambar 6. *Confusion Matrix Random Forest*

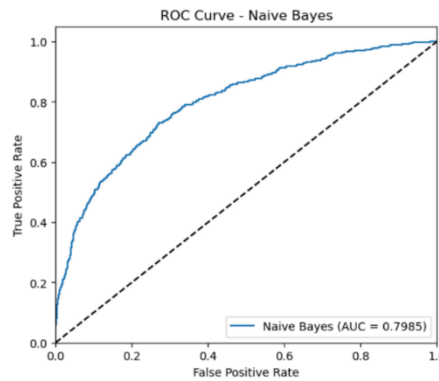
Seperti yang ditunjukkan pada Gambar 6 algoritma *Random Forest* memiliki 2334 + 245 prediksi akurat dan 82 + 339 prediksi salah yang terdiri dari total 3000 data pengujian. Tingkat akurasi algoritma dapat ditentukan dengan cara berikut  $2579/3000 * 100\% = 85\%$ .

Tabel 5. Kinerja Algoritma

Model	AUC	CA	F1	Precision	Recall
<i>Naive Bayes</i>	0,79	0,82	0,33	0,69	0,22
<i>Decision Tree</i>	0,65	0,77	0,44	0,42	0,46
<i>Random Forest</i>	0,83	0,85	0,53	0,74	0,41

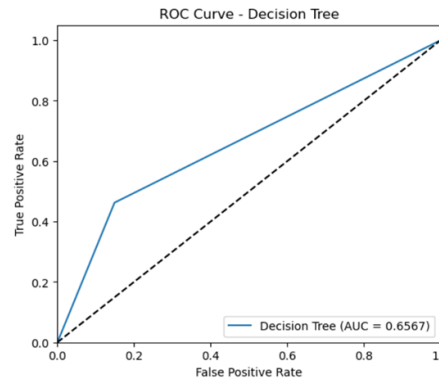
Sumber: Hasil Penelitian (2024)

Tabel 5 menggambarkan bahwa algoritma *Random Forest* menunjukkan akurasi yang lebih tinggi dalam memprediksi dataset churn pelanggan. Secara khusus, nilai AUC dicatat pada 0,83; CA pada 0,85; F1 pada 0,53; *Precision* pada 0,74; dan *Recall* pada 0,41.



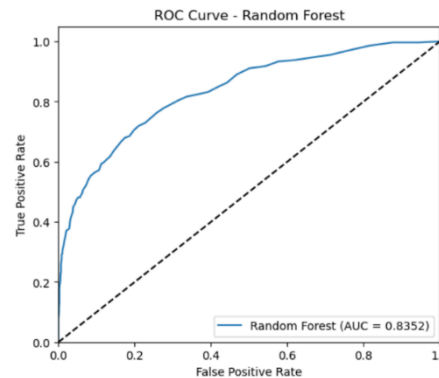
Sumber: Hasil Penelitian (2024)  
Gambar 7. ROC Naive Bayes

Dapat dilihat pada Gambar 7 ROC untuk *Naive Bayes* terbukti bahwa nilai sensitivitas (*True Positive Rate*) dan spesifisitas (*False Positive Rate*) tidak melebihi 1. Nilai AUC terhitung 0,79 menunjukkan tingkat pengklasifikasi berkualitas cukup baik.



Sumber: Hasil Penelitian (2024)  
Gambar 8. ROC Decision Tree

Dapat dilihat pada Gambar 8 ROC untuk *Decision Tree* terbukti bahwa nilai sensitivitas (*True Positive Rate*) dan spesifisitas (*False Positive Rate*) tidak melebihi 1. Nilai AUC terhitung 0,65 menunjukkan tingkat pengklasifikasi berkualitas tergolong lemah.



Sumber: Hasil Penelitian (2024)  
Gambar 9. ROC Random Forest

Dapat dilihat pada Gambar 9 ROC untuk *Random Forest* terbukti bahwa nilai sensitivitas (*True Positive Rate*) dan spesifisitas (*False Positive Rate*) tidak melebihi 1. AUC terhitung bernilai 0,83 menunjukkan tingkat kinerja pengklasifikasi yang sangat baik. Hal ini memperlihatkan bahwa algoritma *Random Forest* mempunyai kemampuan klasifikasi yang efektif untuk mengatasi masalah yang terkait dengan prediksi churn pelanggan.

Dari analisa kurva ROC pada Gambar 7, Gambar 8, dan Gambar 9 dapat dilihat kinerja dari masing-masing algoritma *machine learning*. Kedekatan kurva ke batas kiri dan atas ruang ROC secara langsung berkorelasi dengan tingkat akurasi pengklasifikasi. Setelah melihat hasil analisis ROC, terlihat jelas bahwa model *Random Forest* lebih akurat daripada model *Naive Bayes* dan *Decision Tree* dalam hal akurasi

pengklasifikasi. Mengamati kurva *Random Forest* yang rata-rata menuju sumbu Y menunjukkan tingkat akurasi pengklasifikasi yang optimal.

## KESIMPULAN

Berdasarkan hasil evaluasi, algoritma *Random Forest* menunjukkan performa terbaik dalam memprediksi *churn* pelanggan bank dengan tingkat akurasi sebesar 85%, AUC 0.83, F1 0.53, *Precision* 0.74, dan *Recall* 0.41. Algoritma ini menunjukkan kemampuan klasifikasi yang lebih unggul, yang juga didukung oleh analisis kurva ROC yang menunjukkan bahwa *Random Forest* memiliki kurva yang lebih mendekati sumbu Y dan batas kiri ruang ROC, menandakan keakuratan tinggi. Sebaliknya, algoritma *Naive Bayes* meskipun memiliki akurasi yang cukup tinggi sebesar 82% dan AUC 0.79, menunjukkan kelemahan pada nilai *recall* yang rendah sebesar 0.22, sehingga kurang efektif dalam mendeteksi *churn* pelanggan secara keseluruhan. Algoritma *Decision Tree* menunjukkan performa paling rendah dengan akurasi sebesar 77%, AUC 0.65, F1 0.44, *Precision* 0.42, dan *Recall* 0.46, menunjukkan bahwa algoritma ini kurang optimal untuk dataset ini dibandingkan dua algoritma lainnya.

Keseluruhan, prediksi *churn* pelanggan menggunakan *machine learning*, khususnya dengan algoritma *Random Forest*, sangat bermanfaat bagi industri perbankan untuk mengembangkan strategi pemasaran yang lebih efektif dan proaktif. Dengan mengetahui pelanggan yang berpotensi *churn*, bank dapat melakukan intervensi yang tepat waktu untuk mempertahankan pelanggan dan mengurangi biaya akuisisi pelanggan baru.

Untuk meningkatkan akurasi prediksi dan pemahaman terhadap faktor-faktor yang mempengaruhi *churn*, penggunaan dataset yang lebih besar dan beragam dari berbagai wilayah dan segmentasi pasar sangat disarankan untuk penelitian selanjutnya. Selain itu, penambahan fitur seperti riwayat interaksi pelanggan dan aktivitas online dapat memberikan *insight* yang lebih mendalam. Menguji algoritma lain seperti *Gradient Boosting*, *XGBoost*, dan model *deep learning* juga perlu dipertimbangkan untuk mengevaluasi apakah mereka mampu memberikan performa yang lebih baik.

## REFERENSI

- Diana, R., Warni, H., & Sutabri, T. (2023). PENGGUNAAN TEKNOLOGI MACHINE LEARNING UNTUK PELAYANAN MONITORING KEGIATAN BELAJAR MENGAJAR PADA SMK BINA SRIWIJAYA PALEMBANG. *JUTEKIN (Jurnal Teknik Informatika)*, 11(1). <https://doi.org/10.51530/jutekin.v11i1.709>
- Hasanah, H. (2023). Perbandingan Tingkat Akurasi Algoritma Support Vector Machines (SVM) dan C45 dalam Prediksi Penyakit Jantung. In *Universitas Nusantara PGRI Kediri. Kediri (Vol. 2)*.
- Hon, H. S., Wah Khaw, K., Chew, X., & Wong, W. P. (2023). PREDICTION OF CUSTOMER CHURN FOR ABC MULTISTATE BANK USING MACHINE LEARNING ALGORITHMS. *Malaysian Journal of Computing*, 8(2), 1602–1619. <https://doi.org/10.24191/mjoc.v8i2.21393>
- Husein, A. M., & Harahap, M. (2021). Pendekatan Data Science untuk Menemukan Churn Pelanggan pada Sector Perbankan dengan Machine Learning. *Data Sciences Indonesia (DSI)*, 1(1), 8–13. <https://doi.org/10.47709/dsi.v1i1.1169>
- Kurniawan, I., Cahya, D., Buani, P., Apriliah, W., Saputra, R. A., Bina, U., Informatika, S., Mandiri, U. N., Pusat, J., & Korespondensi, P. (2023). IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK MENENTUKAN IMPLEMENTATION OF RANDOM FOREST ALGORITHM FOR DETERMINING RECIPIENTS OF RASKIN. 10(2), 421–428. <https://doi.org/10.25126/jtiik.202396225>
- Lestari, S., Badrul, M., Studi Sistem Informasi, P., & Nusa Mandiri Jakarta, S. (2020). IMPLEMENTASI KLASIFIKASI NAIVE BAYES UNTUK PREDIKSI KELAYAKAN PEMBERIAN PINJAMAN PADA KOPERASI ANUGERAH BINTANG CEMERLANG. 7(1).
- Marutho, D. (2019). PERBANDINGAN METODE NAIVE BAYES, KNN, DECISION TREE PADA LAPORAN WATER LEVEL JAKARTA.
- Miryam Clementine, & Arum. (2022). Prediksi Churn Nasabah Bank Menggunakan Klasifikasi Naive Bayes dan ID3. *Jurnal Processor*, 17(1), 9–18. <https://doi.org/10.33998/processor.2022.17.1.1170>
- Rahmawati, P., & Larasati, A. (2021). Comparing Decision Tree and Artificial Neural Network Model in Predicting Bank Approval on Customer Credit.
- Rizki Kurniawan, M., Nurul Sabrina, P., Ilyas Teknik Informatika, R., Jenderal Achmad Yani Jl Terusan Jend Sudirman, U., Cimahi Sel, K., Cimahi, K., & Barat, J. (2023). PREDIKSI CUSTOMER CHURN PADA PERUSAHAAN TELEKOMUNIKASI MENGGUNAKAN ALGORITMA C4.5 BERBASIS PARTICLE SWARM OPTIMIZATION. In *Jurnal Mahasiswa Teknik Informatika (Vol. 7, Issue 5)*.



Satria, A., Badri, R. M., & Safitri, I. (2023). Prediksi Hasil Panen Tanaman Pangan Sumatera dengan Metode Machine Learning. *Digital Transformation Technology*, 3(2), 389–398.  
<https://doi.org/10.47709/digitech.v3i2.2852>