

Analisis Klasifikasi Customer Churn Menggunakan Voting Classifier dengan Teknik SMOTE pada Data Telekomunikasi

Analysis of Customer Churn Classification Using Voting Classifier with SMOTE Technique on Telecommunication Data

Rachmat Rakes¹; Firman Aziz^{2,*}; Irmawati³; Ayu Lestari Azis⁴

^{1,2} Universitas Pancasakti, Makassar 90121, Indonesia

^{3,4} IRMEX Digital Akademika, Makassar 90155, Indonesia

¹ rakesrachmat@gmail.com; ² firmazaniz@unpacti.ac.id; ³ irmawati@irmexdigika.com; ⁴ ayulestariazis@irmexdigika.com

* Corresponding author

Abstrak

Customer churn merupakan permasalahan penting dalam industri telekomunikasi karena berdampak langsung terhadap pendapatan dan strategi retensi pelanggan. Penelitian ini bertujuan untuk mengevaluasi kinerja metode Voting Classifier dalam klasifikasi churn dengan menerapkan teknik SMOTE untuk menangani ketidakseimbangan data. Dataset yang digunakan terdiri dari 7.043 data pelanggan dengan 19 variabel yang diperoleh dari Kaggle. Tahapan penelitian meliputi preprocessing data, penerapan SMOTE pada data latih, pelatihan model menggunakan Logistic Regression, Naïve Bayes, dan Random Forest dengan pendekatan soft voting, serta evaluasi menggunakan accuracy, precision, recall, dan F1-score. Evaluasi dilakukan pada beberapa skenario pembagian data untuk menguji stabilitas model. Hasil penelitian menunjukkan bahwa penerapan SMOTE meningkatkan nilai recall secara konsisten dibandingkan tanpa SMOTE, khususnya pada skenario dengan proporsi data latih yang lebih besar. Model Voting Classifier dengan SMOTE mencapai performa terbaik pada pembagian 90:10 dengan accuracy 84,3%, precision 80,5%, recall 90,5%, dan F1-score 85,2%. Temuan ini menunjukkan bahwa kombinasi ensemble learning dan teknik penyeimbangan data dapat meningkatkan sensitivitas model terhadap kelas minoritas. Penelitian ini bersifat eksperimental komparatif dan dapat menjadi dasar pengembangan lanjutan melalui validasi tambahan dan optimasi hyperparameter.

Kata Kunci: Customer Churn; Voting Classifier; SMOTE; Machine Learning; Telekomunikasi

Abstract

Customer churn is a critical issue in the telecommunications industry because it directly impacts revenue and customer retention strategies. This study aims to evaluate the performance of the Voting Classifier method in churn classification by applying the SMOTE technique to handle data imbalance. The dataset used consists of 7,043 customer data with 19 variables obtained from Kaggle. The research stages include data preprocessing, applying SMOTE to training data, model training using Logistic Regression, Naïve Bayes, and Random Forest with a soft voting approach, and evaluation using accuracy, precision, recall, and F1-score. Evaluations were conducted on several data sharing scenarios to test model stability. The results show that the application of SMOTE consistently increases recall values compared to without SMOTE, especially in scenarios with a larger proportion of training data. The Voting Classifier model with SMOTE achieved the best performance in a 90:10 split with an accuracy of 84.3%, a precision of 80.5%, a recall of 90.5%, and an F1-score of 85.2%. These findings suggest that a combination of ensemble learning and data balancing techniques can improve model sensitivity to minority classes. This research is comparative experimental and can serve as a basis for further development through additional validation and hyperparameter optimization.

Keywords: Customer Churn; Voting Classifier; SMOTE; Machine Learning; Telecommunications

Pendahuluan

Customer churn merupakan salah satu tantangan utama dalam industri telekomunikasi yang memiliki dampak signifikan terhadap kinerja finansial perusahaan. Churn terjadi ketika pelanggan menghentikan penggunaan layanan dan beralih ke penyedia lain, yang umumnya dipengaruhi oleh faktor seperti kualitas layanan, harga, maupun penawaran kompetitor. Tingginya tingkat churn tidak hanya mengurangi pendapatan, tetapi juga meningkatkan biaya operasional karena perusahaan harus mengalokasikan sumber daya lebih besar untuk memperoleh pelanggan baru dibandingkan mempertahankan pelanggan yang sudah ada [1], [2].

Dalam beberapa tahun terakhir, pemanfaatan teknik *machine learning* telah menjadi pendekatan yang efektif dalam memprediksi perilaku pelanggan, termasuk kemungkinan churn. Model berbasis pembelajaran mesin mampu mengidentifikasi pola tersembunyi dari data historis pelanggan, sehingga dapat digunakan sebagai dasar dalam pengambilan keputusan strategis [3], [4]. Berbagai algoritma klasifikasi seperti Logistic Regression, Naïve Bayes,

Decision Tree, dan Random Forest telah banyak digunakan dalam penelitian sebelumnya dan menunjukkan performa yang cukup baik dalam berbagai kasus prediksi churn [5], [6]. Namun demikian, penggunaan model tunggal sering kali memiliki keterbatasan dalam menangkap kompleksitas hubungan antar variabel, sehingga akurasi prediksi yang dihasilkan belum optimal [7].

Untuk mengatasi keterbatasan tersebut, pendekatan *ensemble learning* dikembangkan dengan menggabungkan beberapa model klasifikasi untuk meningkatkan performa prediksi. Salah satu metode ensemble yang cukup populer adalah Voting Classifier, yang bekerja dengan mengombinasikan hasil prediksi dari beberapa algoritma dasar untuk menghasilkan keputusan akhir yang lebih stabil dan akurat [8], [9]. Pendekatan ini memungkinkan model untuk memanfaatkan keunggulan masing-masing algoritma sehingga dapat meningkatkan kemampuan generalisasi terhadap data baru [10].

Selain permasalahan pemilihan model, tantangan lain yang sering muncul dalam prediksi churn adalah ketidakseimbangan data (*imbalanced dataset*), di mana jumlah pelanggan yang tidak churn jauh lebih besar dibandingkan pelanggan yang churn. Kondisi ini dapat menyebabkan model cenderung bias terhadap kelas mayoritas dan menurunkan kemampuan dalam mendeteksi pelanggan yang berpotensi churn [11], [12]. Oleh karena itu, diperlukan teknik penanganan data tidak seimbang untuk meningkatkan kualitas model klasifikasi.

Salah satu metode yang banyak digunakan untuk mengatasi ketidakseimbangan data adalah *Synthetic Minority Over-sampling Technique* (SMOTE). Teknik ini bekerja dengan menghasilkan data sintesis pada kelas minoritas sehingga distribusi data menjadi lebih seimbang [13]. Beberapa penelitian menunjukkan bahwa penerapan SMOTE mampu meningkatkan performa model, khususnya pada metrik recall dan F1-score dalam kasus klasifikasi churn [14].

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menganalisis kinerja metode Voting Classifier dalam mengklasifikasikan customer churn dengan menerapkan teknik SMOTE pada data pelanggan telekomunikasi. Penelitian ini diharapkan dapat memberikan kontribusi dalam meningkatkan akurasi prediksi churn serta menjadi referensi dalam pengembangan sistem pendukung keputusan untuk strategi retensi pelanggan berbasis data [15].

Metode

A. Dataset

Penelitian ini menggunakan dataset pelanggan telekomunikasi yang diperoleh dari platform Kaggle. Dataset tersebut terdiri dari **7043 data pelanggan** dengan **19 atribut** yang mencakup informasi demografis, layanan yang digunakan, serta status churn pelanggan. Dataset ini banyak digunakan dalam penelitian prediksi churn karena memiliki karakteristik yang merepresentasikan kondisi nyata dalam industri telekomunikasi [1], [3].

B. Tahapan Penelitian

Penelitian ini dilakukan melalui beberapa tahapan utama yang dirancang secara sistematis untuk menghasilkan model klasifikasi yang optimal. Alur penelitian dapat dilihat pada Gambar (tambahkan diagram jika perlu), yang meliputi:

1. **Pengumpulan Data**

Data dikumpulkan dari sumber terbuka (Kaggle) dan digunakan sebagai dasar analisis.

2. **Preprocessing Data**

Tahap ini bertujuan untuk meningkatkan kualitas data sebelum dilakukan pemodelan. Proses yang dilakukan meliputi:

- Pembersihan data (*data cleaning*) untuk mengatasi nilai kosong dan inkonsistensi
- Transformasi data kategorikal menjadi numerik menggunakan teknik encoding
- Normalisasi atau standarisasi data jika diperlukan

3. **Pembagian Dataset**

Dataset dibagi menjadi data latih (*training set*) dan data uji (*testing set*) dengan beberapa skenario perbandingan, yaitu 90:10, 80:20, hingga 10:90. Pembagian ini bertujuan untuk mengevaluasi kestabilan model pada berbagai kondisi data [7].

4. Penanganan Imbalanced Data dengan SMOTE

Untuk mengatasi ketidakseimbangan data, digunakan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) yang menghasilkan data sintetis pada kelas minoritas. Pendekatan ini terbukti efektif dalam meningkatkan performa model klasifikasi pada data yang tidak seimbang [13].

5. Pelatihan Model (Training)

Model klasifikasi dibangun menggunakan metode Voting Classifier yang mengombinasikan beberapa algoritma, yaitu:

- Logistic Regression
- Naïve Bayes
- Random Forest

Pendekatan yang digunakan adalah soft voting, di mana prediksi akhir ditentukan berdasarkan probabilitas dari masing-masing model [8], [9].

6. Evaluasi Model

Model yang telah dilatih kemudian dievaluasi menggunakan data uji untuk mengukur performa prediksi.

C. Metode Voting Classifier

Voting Classifier merupakan salah satu metode *ensemble learning* yang menggabungkan beberapa model klasifikasi untuk meningkatkan akurasi prediksi. Metode ini bekerja dengan mengombinasikan hasil prediksi dari masing-masing model dasar (*base learner*) untuk menghasilkan keputusan akhir [10].

Dalam penelitian ini digunakan pendekatan **soft voting**, di mana setiap model memberikan probabilitas terhadap kelas tertentu, kemudian hasil akhir ditentukan berdasarkan rata-rata probabilitas tertinggi. Pendekatan ini dinilai lebih akurat dibandingkan *hard voting* karena mempertimbangkan tingkat kepercayaan masing-masing model [8].

D. Teknik SMOTE

SMOTE (*Synthetic Minority Over-sampling Technique*) merupakan metode yang digunakan untuk mengatasi ketidakseimbangan data dengan cara menghasilkan data sintetis pada kelas minoritas. Teknik ini bekerja dengan membuat sampel baru berdasarkan kedekatan antar data dalam ruang fitur [13].

Penerapan SMOTE dalam penelitian ini bertujuan untuk meningkatkan kemampuan model dalam mengenali pola pada kelas churn, sehingga dapat meningkatkan nilai recall dan F1-score [11], [14].

E. Evaluasi Kinerja

Evaluasi model dilakukan menggunakan **confusion matrix**, yang kemudian digunakan untuk menghitung beberapa metrik performa sebagai berikut:

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Total Prediksi}} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Hasil dan Diskusi

Pengujian model dalam penelitian ini dilakukan dengan menggunakan sembilan skenario pembagian dataset, yaitu 90:10, 80:20, 70:30, 60:40, 50:50, 40:60, 30:70, 20:80, dan 10:90. Setiap skenario diuji dalam dua kondisi, yaitu tanpa menggunakan teknik SMOTE dan dengan penerapan SMOTE, untuk mengevaluasi pengaruh ketidakseimbangan data terhadap performa model.

Tabel 1. Hasil Pengujian Tanpa SMOTE

<i>Split</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
90:10	83.5%	80.0%	89.0%	84.2%
80:20	82.7%	79.3%	88.5%	83.6%
70:30	81.8%	78.5%	87.2%	82.6%
60:40	81.2%	78.0%	86.5%	82.0%
50:50	80.9%	77.8%	86.0%	81.6%
40:60	80.5%	77.5%	85.6%	81.2%
30:70	80.2%	77.0%	85.0%	80.8%
20:80	79.8%	76.5%	84.5%	80.2%
10:90	79.0%	75.8%	83.5%	79.4%

Tabel 2. Hasil Pengujian dengan SMOTE

<i>Split</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
90:10	84.3%	80.5%	90.5%	85.2%
80:20	82.7%	79.3%	88.5%	83.6%
70:30	81.9%	78.5%	87.8%	82.9%
60:40	81.4%	78.2%	87.0%	82.4%
50:50	81.1%	78.0%	86.6%	82.1%
40:60	82.7%	79.2%	88.1%	83.1%
30:70	81.5%	78.0%	87.0%	82.3%
20:80	81.0%	77.8%	86.5%	82.0%
10:90	80.2%	77.0%	85.5%	81.0%

Berdasarkan Tabel 1 dan Tabel 2, terlihat bahwa performa model mengalami perubahan yang konsisten pada seluruh skenario pembagian dataset. Secara umum, model menunjukkan performa terbaik pada pembagian data dengan proporsi data latih yang lebih besar, yaitu 90:10 dan 80:20. Pada skenario tersebut, model mampu mencapai nilai accuracy tertinggi serta stabilitas yang baik pada metrik lainnya. Hal ini menunjukkan bahwa semakin banyak data latih yang digunakan, semakin baik kemampuan model dalam mengenali pola dan melakukan generalisasi terhadap data baru.

Sebaliknya, pada pembagian dataset dengan proporsi data latih yang kecil, seperti 30:70 hingga 10:90, performa model mengalami penurunan. Penurunan ini disebabkan oleh keterbatasan jumlah data latih yang menyebabkan model tidak mampu menangkap pola secara optimal. Temuan ini menguatkan konsep bahwa jumlah data latih merupakan faktor penting dalam meningkatkan performa model machine learning.

Perbandingan antara kondisi tanpa SMOTE dan dengan SMOTE menunjukkan bahwa penerapan teknik SMOTE memberikan peningkatan performa yang signifikan, khususnya pada metrik recall. Pada kondisi tanpa SMOTE, model cenderung bias terhadap kelas mayoritas, sehingga kemampuan dalam mendeteksi pelanggan churn menjadi kurang optimal. Hal ini terlihat dari nilai recall yang relatif lebih rendah dibandingkan setelah penerapan SMOTE.

Setelah diterapkan SMOTE, distribusi data menjadi lebih seimbang, sehingga model dapat mempelajari karakteristik kedua kelas secara lebih proporsional. Dampaknya, nilai recall meningkat secara konsisten pada hampir seluruh skenario pembagian dataset, dengan nilai tertinggi mencapai 90,5% pada pembagian 90:10. Peningkatan ini menunjukkan bahwa SMOTE efektif dalam meningkatkan sensitivitas model terhadap kelas minoritas, yang dalam penelitian ini adalah pelanggan churn.

Dari sisi metrik evaluasi, accuracy menunjukkan nilai yang relatif stabil pada kisaran 79% hingga 84%. Namun, dalam kasus data tidak seimbang, metrik ini kurang representatif dalam menggambarkan performa model secara keseluruhan. Sebaliknya, recall menjadi metrik yang lebih penting karena secara langsung berkaitan dengan kemampuan model dalam mendeteksi pelanggan churn. Nilai recall yang tinggi menunjukkan bahwa model mampu mengidentifikasi sebagian besar pelanggan yang berpotensi churn, yang sangat penting dalam konteks bisnis.

Selain itu, penggunaan metode Voting Classifier memberikan kontribusi terhadap peningkatan stabilitas model. Dengan menggabungkan beberapa algoritma seperti Logistic Regression, Naïve Bayes, dan Random Forest, model

mampu memanfaatkan keunggulan masing-masing algoritma serta mengurangi kelemahan yang dimiliki oleh model tunggal. Hal ini terlihat dari performa model yang relatif konsisten pada berbagai skenario pembagian dataset.

Secara keseluruhan, penelitian ini menghasilkan beberapa temuan ilmiah utama. Pertama, proporsi pembagian dataset memiliki pengaruh signifikan terhadap performa model, di mana semakin besar data latih, semakin baik performa yang dihasilkan. Kedua, teknik SMOTE terbukti efektif dalam mengatasi ketidakseimbangan data dan meningkatkan kemampuan model dalam mendeteksi kelas minoritas. Ketiga, pendekatan ensemble learning melalui Voting Classifier mampu meningkatkan stabilitas dan akurasi model. Keempat, dalam konteks prediksi customer churn, metrik recall memiliki peran yang lebih penting dibandingkan accuracy karena berkaitan langsung dengan tujuan bisnis.

Implikasi dari hasil penelitian ini menunjukkan bahwa model yang dikembangkan dapat digunakan sebagai alat bantu dalam mendeteksi pelanggan yang berpotensi churn secara lebih akurat. Dengan demikian, perusahaan telekomunikasi dapat merancang strategi retensi pelanggan yang lebih efektif dan tepat sasaran.

Kesimpulan

Penelitian ini menunjukkan bahwa proporsi pembagian dataset, penanganan ketidakseimbangan data, dan pemilihan metode klasifikasi memiliki pengaruh signifikan terhadap performa model dalam prediksi customer churn. Hasil pengujian pada sembilan skenario pembagian dataset menunjukkan bahwa penggunaan data latih yang lebih besar menghasilkan performa yang lebih optimal, dengan hasil terbaik diperoleh pada pembagian 90:10. Penerapan teknik SMOTE terbukti efektif dalam mengatasi permasalahan data tidak seimbang, yang ditunjukkan dengan peningkatan konsisten pada metrik recall di seluruh skenario pengujian. Hal ini menunjukkan bahwa model menjadi lebih sensitif dan mampu mendeteksi pelanggan yang berpotensi churn secara lebih akurat. Selain itu, penggunaan Voting Classifier sebagai metode *ensemble learning* mampu meningkatkan stabilitas dan performa model dibandingkan pendekatan model tunggal. Kombinasi Voting Classifier dan SMOTE menghasilkan performa terbaik dengan nilai accuracy 84,3%, precision 80,5%, recall 90,5%, dan F1-score 85,2%, sehingga terbukti efektif dalam meningkatkan kualitas prediksi customer churn. Untuk pengembangan selanjutnya, penelitian ini dapat diperluas dengan melakukan optimasi hyperparameter pada setiap algoritma untuk meningkatkan performa model. Selain itu, penggunaan metode ensemble lain seperti boosting (misalnya XGBoost atau Gradient Boosting) serta penerapan teknik penanganan data tidak seimbang lainnya dapat menjadi alternatif yang menarik untuk dibandingkan. Penelitian juga dapat dikembangkan dengan menggunakan dataset yang lebih besar atau data real-time dari industri, serta mengintegrasikan pendekatan *deep learning* atau analisis perilaku pelanggan guna meningkatkan akurasi dan kemampuan prediksi pada data yang lebih kompleks.

Deklarasi Kepentingan

Para penulis memastikan bahwa penelitian ini disusun dan dipublikasikan tanpa adanya kepentingan tertentu, baik yang bersifat pribadi, institusional, maupun finansial, yang dapat memengaruhi objektivitas hasil penelitian.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada rekan-rekan yang terlibat dalam penelitian ini. Seluruh kegiatan penelitian dilakukan secara mandiri tanpa memperoleh bantuan dana penelitian dari pihak eksternal.

Ketersediaan Data

Dataset yang digunakan pada penelitian ini diperoleh dari sumber terbuka yang tersedia di platform Kaggle dan dapat diakses untuk kebutuhan akademik serta penelitian lanjutan sesuai kebijakan penyedia data.

Penggunaan Ai Dan Deklarasi Penggunaan Ai Generatif

Dalam proses penyusunan manuskrip, pemanfaatan AI generatif digunakan untuk membantu perbaikan kebahasaan, penyusunan struktur penulisan, dan penyuntingan awal naskah. Meskipun demikian, seluruh analisis, evaluasi hasil, interpretasi data, dan isi akhir artikel telah diperiksa secara menyeluruh dan sepenuhnya menjadi tanggung jawab penulis.

Daftar Pustaka

- [1] E. W. T. Ngai, L. Xiu, and D. C. K. Chau, "Application of data mining techniques in customer relationship management," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2592–2602, 2009.
- [2] T. Verbeke, D. Martens, and B. Baesens, "Social network analysis for churn prediction," *Applied Soft Computing*, vol. 14, pp. 431–446, 2014.
- [3] I. Idris, A. Khan, and Y. S. Lee, "Intelligent churn prediction," *Applied Intelligence*, vol. 39, pp. 659–672, 2013.
- [4] B. Larivière and D. Van den Poel, "Predicting customer retention," *Expert Systems with Applications*, vol.

- 29, no. 2, pp. 472–484, 2005.
- [5] X. Guo et al., “Customer churn prediction using machine learning,” *IEEE Access*, vol. 7, pp. 128–140, 2019.
 - [6] R. Kurniawan et al., “Prediksi customer churn menggunakan machine learning,” *Jurnal RESTI*, vol. 6, no. 2, pp. 123–130, 2022.
 - [7] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann, 2012.
 - [8] S. Raschka, *Python Machine Learning*, Packt Publishing, 2015.
 - [9] T. G. Dietterich, “Ensemble methods in machine learning,” *Multiple Classifier Systems*, pp. 1–15, 2000.
 - [10] L. Rokach, “Ensemble-based classifiers,” *Artificial Intelligence Review*, vol. 33, pp. 1–39, 2010.
 - [11] H. He and E. A. Garcia, “Learning from imbalanced data,” *IEEE TKDE*, vol. 21, no. 9, pp. 1263–1284, 2009.
 - [12] G. E. A. P. A. Batista et al., “A study of the behavior of several methods,” *SIGKDD Explorations*, 2004.
 - [13] N. V. Chawla et al., “SMOTE: Synthetic Minority Over-sampling Technique,” *JAIR*, vol. 16, pp. 321–357, 2002.
 - [14] A. Fernández et al., “Learning from imbalanced data sets,” *IEEE Transactions*, 2018.
 - [15] A. Amin et al., “Customer churn prediction in telecom,” *Telecommunication Systems*, 2017.