

Klasifikasi Customer Churn pada Telekomunikasi Industri Untuk Retensi Pelanggan Menggunakan Algoritma C4.5

Stevan Desena Damanik, Muhammad Ihsan Jambak*

Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Sriwijaya, Palembang, Indonesia

Email: ¹09031282025093@student.unsri.ac.id, ^{2,*}jambak@unsri.ac.id

Email Penulis Korespondensi: jambak@unsri.ac.id

Abstrak—Ketidaktahuan perusahaan telekomunikasi mengenai alasan dan ciri-ciri customer churn membuat perusahaan telekomunikasi mengalami kerugian yang besar. Hal tersebut menjadikan customer churn sebagai masalah besar bagi perusahaan telekomunikasi. Penelitian ini menggunakan data mining dengan teknik klasifikasi sebagai solusi untuk menganalisis karakteristik customer churn. Pada penelitian ini akan menggunakan Rapid Miner dan algoritma C4.5 untuk melakukan proses data mining. Tujuan dari penelitian ini yaitu mengetahui apa saja karakteristik customer churn sehingga perusahaan dapat membuat kebijakan yang dapat mempertahankan customer serta meningkatkan retensi pelanggan. Penelitian ini didasarkan pada CRISP-DM. Data yang digunakan diambil dari kaggle.com dengan 21 atribut dan 7034 baris data dan akan dilakukan data preparation. Dari hasil penelitian diketahui terdapat 5 atribut yang memiliki pengaruh cukup besar terhadap customer churn yaitu contract, InternetService, TotalChares, tenure, PaperlessBilling, MultipleLines, StreamingMovies. Dan dari hasil penelitian ini memiliki tingkat akurasi sebesar 79.53%.

Kata Kunci: Data Mining; Karakteristik Customer Churn; Algoritma C4.5; Perusahaan Telekomunikasi; Rapid Miner

Abstract—Ignorance of telecommunications companies regarding the reasons and characteristics of customer churn causes telecommunications companies to suffer huge losses. This makes customer churn a big problem for telecommunications companies. This study uses data mining with classification techniques as a solution to analyze customer churn characteristics. This research will use Rapid Miner and the C4.5 algorithm to carry out the data mining process. The purpose of this research is to find out what are the characteristics of customer churn so that companies can make policies that can retain customers and increase customer retention. This research is based on CRISP-DM. Data taken from kaggle.com with 21 attributes and 7034 rows of data and data preparation will be carried out. From the research results it is known that there are 5 attributes that have a considerable influence on customer churn, namely contracts, InternetService, TotalChares, tenure, PaperlessBilling, MultipleLines, StreamingMovies. And from the results of this study has an accuracy rate of 79.53%.

Keywords: Data Mining; Characteristics of Customer Churn; C4.5 Algorithm; Telecommunication Companies; Rapid Miner

1. PENDAHULUAN

Salah satu hal fundamental di negara maju adalah bidang telekomunikasi. Seiring berkembangnya zaman, teknologi pun mengalami perkembangan yang sangat signifikan. Hal ini membuat telekomunikasi menjadi hal yang sangat penting di zaman ini. Pentingnya telekomunikasi membuat banyak perusahaan bermunculan yang mengakibatkan persaingan yang ketat di industri ini. Untuk memenangkan kompetisi banyak perusahaan melakukan inovasi demi mempertahankan dan menarik customer baru untuk menggunakan layanan mereka. Persaingan ini membuat ada pihak yang harus menelan kerugian karena kehilangan customer (churn) dan menerima keuntungan dari mendapatkan customer [1]–[5]

Dalam kompetisi ini perusahaan telekomunikasi akan fokus untuk mempertahankan customer lama dan menarik customer baru. Jika dibandingkan dari kedua hal tersebut, diketahui bahwa lebih sulit untuk mencari pelanggan baru daripada mempertahankan pelanggan lama karena mencari pelanggan baru akan memakan biaya yang lebih besar. Hal ini berhubungan dengan loyalitas, pelanggan lama cenderung akan lebih loyal untuk menggunakan layanan yang disediakan karena mereka sudah mempercayai layanan yang mereka gunakan. Presentase kerugian yang dialami oleh telekomunikasi di Eropa sebesar 8% hingga 12%. Dan biaya churn rate sebesar 500 euro. Sedangkan di Amerika Serikat memiliki churn rate sekitar 2-3% per bulan. Biaya untuk mendapatkan pelanggan baru sekitar 400 dolar sedangkan biaya untuk mempertahankan pelanggan lama kurang dari 400 dolar. Total kerugian yang diakibatkan churn mencapai 6,3 juta dolar pada industri [6]. Hal tersebut membuat pelanggan lama harus diberikan layanan yang lebih baik untuk mempertahankan mereka agar tidak churn.

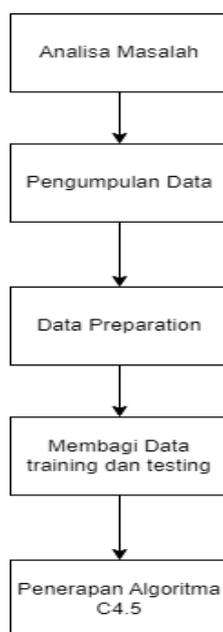
Churn merupakan masalah yang besar bagi perusahaan telekomunikasi dan tidak dapat dihindari. Hal tersebut terus terjadi dan sangat merugikan perusahaan telekomunikasi karena ketidaktahuan perusahaan telekomunikasi mengenai ciri-ciri customer churn. Hal tersebut bukanlah hal yang mudah untuk diatasi. Untuk mengatasi hal tersebut perusahaan telekomunikasi perlu mengetahui alasan dan ciri-ciri dari customer yang akan churn, sehingga kita dapat meningkatkan pelayanan dan kepuasan mereka. Dalam mempertahankan dan menarik customer baru tidak dapat dilakukan secara acak, ada beberapa faktor yang harus diperhatikan dalam menarik customer secara efektif. Oleh karena itu perlu diketahui faktor penyebab churn dalam upaya peningkatan retensi pelanggan di perusahaan telekomunikasi.

Analisis karakteristik sangat diperlukan untuk mengetahui penyebab customer churn. Perusahaan telekomunikasi perlu menerapkan ini dalam perusahaan mereka untuk meningkatkan retensi pelanggan [7]. Dengan adanya analisis yang baik tentunya akan menghasilkan insight yang baik sehingga perusahaan dapat membuat keputusan terbaik mengenai pelanggan. Terdapat banyak cara untuk menganalisis karakteristik pelanggan. Salah satunya yaitu data mining. Data mining dapat digunakan untuk menemukan karakteristik customer churn berdasarkan data yang dimiliki. Dengan adanya data mining knowledge yang dihasilkan pun akan memiliki value yang baik.

Rules-base learning, neural networks dan decision tree merupakan algoritma yang dapat memprediksi customer churn dengan cukup baik (Kavitha et al., 2020). Terdapat penelitian terdahulu yang sudah melakukan klasifikasi mengenai customer churn yaitu “Penerapan Algoritma C4.5 Untuk Prediksi Churn Rate Pengguna Jasa Telekomunikasi” penelitian ini dari penelitian tersebut menghasilkan nilai accuracy sebesar 87.03%, precision 87.5% dan recall sebesar 96%. Berdasarkan penelitian terdahulu Penulis memutuskan untuk menggunakan Algoritma C4.5 untuk mengetahui karakteristik customer yang akan churn dengan tujuan meningkatkan retensi pelanggan di perusahaan telekomunikasi. Untuk memecahkan masalah tersebut Penulis membuat model dengan teknik klasifikasi data mining yaitu algoritma C4.5 untuk mengetahui karakteristik dari pelanggan yang churn. Penelitian ini memiliki tujuan untuk mendapatkan karakteristik customer churn sehingga perusahaan di industri telekomunikasi dapat membuat kebijakan yang tepat mengenai masalah customer churn. Dengan adanya keputusan yang efektif dalam mengatasi customer churn penulis berharap akan meningkatkan retensi pengguna dan mengurangi kerugian yang disebabkan oleh churn.

2. METODOLOGI PENELITIAN

Kerangka kerja penelitian merupakan tahapan prosedur yang dilakukan para peneliti untuk melakukan penelitian [12], [13]. Pada penelitian ini penulis mengadopsi metode CRISP-DM sebagai alur penelitian. Berikut adalah tahapan prosedur yang dilakukan dalam penelitian ini:



Gambar 1. Kerangka Penelitian

Proses pemahaman terhadap masalah yang ada disebut analisa masalah. Analisa masalah termasuk tahapan business understanding pada CRISP-DM. Pada tahap analisa masalah, masalah akan di analisa dan dipelajari lebih lanjut untuk menentuekan data yang sesuai dengan masalah [12]. Berdasarkan permasalahan yang telah ditemukan akan dilakukan analisa untuk memahami masalah yang ada. Permasalahan yang ada ketidkadmampuan perusahaan telekomunikasi untuk mengetahui penyebab customer churn. Penulis melakukan penelian untuk mengetahui perilaku atau karakteristik pelanggan yang akan churn. Pada tahap ini juga penulis akan menentukan data apa saja yang dapat memecahkan masalah tersebut.

Dalam konteks penelitian yang dilakukan diperlukan dataset yang memiliki informasi mengenai pelanggan telekomunikasi. Terdapat berbagai macam sumber dataset yang dapat diambil untuk menjadi dataset penelitian. Penulis memilih untuk mengambil dataset dari situs kaggle.com <https://www.kaggle.com/datasets/blastchar/telco-customer-churn> yang terdiri dari 21 atribut yang terlihat di tabel 1. Atribut-atribut ini merupakan faktor yang mempengaruhi customer churn. Dataset yang dimiliki memiliki 7033 data (yang dapat dilihat di tabel 2) dan memiliki 1 label yaitu churn yang memiliki 2 kelas yaitu ‘yes’ sebanyak 1869 dan ‘no’ sebanyak 5164.

Tabel 1. Informasi Dataset

Atribut	Keterangan
customerID	Id pelanggan
gender	Jenis kelamin pelanggan
SeniorCitizen	Apakah pelanggan lansia (1 : ya, 0 : tidak)
Partner	Apakah pelanggan memiliki pasangan
Dependents	Apakah customer memiliki tanggungan

tenure	Jangka waktu pelanggan telah berlangganan (dalam bulan)
PhoneService	Apakah customer mengambil layanan telepon
MultipleLines	Layanan multiplelines yang dipilih customer (Yes, No, No Phone Service)
InternetService	Pilihan internet service provider oleh pelanggan (DSL, Fiber Optic, No)
OnlineSecurity	Apakah Pelanggan memilih layanan online security(Yes, No, No internet service),
OnlineBackup	Apakah pelanggan memilih layanan onlineBackup (Yes, No, No internet service)
DeviceProtection	Apakah customer memilih layanan device protection (Yes, No, No internet service)
TechSupport	Apakah customer memilih layanan tech support (Yes, No, No internet service)
StreamingTV	Apakah customer memilih streaming tv (Yes, No, No internet service)
StreamingMovies	Apakah customer memilih layanan streaming movies (Yes, No, No internet service)
Contract	Contract yang dipilih pelanggan (Month-to-month, One year, Two year)
PaperlessBilling	Apakah customer memilih paperless billing (Yes, No)
PaymentMethod	Metode pembayaran (Eletronic check, Mailed check, Bank transfet automatic), credit card (automatic))
MonthlyCharges	Pembayaran perbulan
TotalCharges	Total pembayaran
Churn	Customer melakukan churn atau tidak

Tabel 2. Dataset Pelanggan telekomunikasi

customerID	gender	SeniorCitizen	Partner	Dependents	Tenure	PhoneService	...	Churn
7590-VHVEG	Female	0	Yes	No	1	No	...	No
5575-GNVDE	Male	0	No	No	34	Yes	...	No
3668-QPYBK	Male	0	No	No	2	Yes	...	Yes
7795-CFOCW	Male	0	No	No	45	No	...	No
9237-HQITU	Female	0	No	No	2	Yes	...	Yes
9305-CDSKC	Female	0	No	No	8	Yes	...	Yes
...
3186-AJIEK	Male	0	No	No	66	Yes	...	No

Langka selanjutnya adalah data preparation adalah proses penyiapan data meliputi menyusun, mengatu, menggabungkan dan mengumpulkan sehingga dapat digunakan pada machine learning. Data preparation seringkali mengumpulkan dari data internal dan eksternal yang berbeda[8]–[10]. Data preparation adalah langkah untuk mempersiapkan data yang dimiliki agar sesuai dengan model yang akan digunakan. Berikut adalah langkah pada tahap data preparation :

a. Drop kolom yang tidak digunakan.

Terdapat 1 kolom yang tidak digunakan dari 21 kolom yang ada. Kolom tersebut yaitu kolom costumerID. Kolom tersebut akan dihapus dari tabel dataset yang akan digunakan.

b. *Handling Missing value.*

Terdapat beberapa kolom yang missing value atau NULL pada dataset ini. Hal ini dapat terjadi karena kesalahan teknis atau memang kolom tersebut tidak memiliki nilai. Untuk mengatasi masalah ini digunakan nilai rata-rata dari 1 kolom dari tempat missing value berada.

c. Transformasi data.

Pada kolom SeniorCitizen memiliki nilai 1 dan 0. Hal ini tidak sesuai dengan jenis data yang dapat diinput ke algoritma C4.5 sehingga perlu dilakukan transformasi data. Pada kolom SeniorCitizen kolom yang memiliki nilai nol akan ditransformasi menjadi “No” dan kolom yang bernilai 1 akan ditransformasi menjadi “Yes”.

Tabel 3. Dataset setelah Preparation

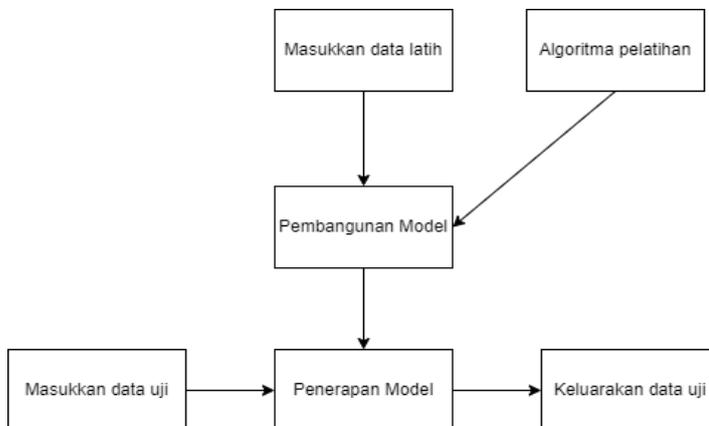
gender	SeniorCitizen	Partner	Dependents	Tenure	PhoneService	...	Churn
Female	No	Yes	No	1	No	...	No
Male	No	No	No	34	Yes	...	No
Male	No	No	No	2	Yes	...	Yes
Male	No	No	No	45	No	...	No
Female	No	No	No	2	Yes	...	Yes
Female	No	No	No	8	Yes	...	Yes
...
Male	No	No	No	66	Yes	...	No

Terdapat istilah training dan testing dalam klasifikasi pada machine learning. Pada penelitian ini akan menggunakan linear sampling sebagai type split dengan ratio sebesar 0.7 untuk data training dan 0.3 untuk data testing. Pada machine learning training merupakan proses membangun model, lalu testing sendiri merupakan proses menguji model yang telah dibuat pada tahap training. Secara umum, dataset dibagi menjadi dua jenis, yaitu:

a. Training set yaitu data yang akan dipakai untuk membangun model

b. Validation set yaitu data yang dipakai untuk mengoptimalkan model

Tahapan dalam menemukan model untuk membedakan sebuah kelas atau konsep data disebut dengan klasifikasi. Klasifikasi memiliki banyak algoritma salah satunya yaitu algoritma C4.5. Model didapat dari proses menganalisis data training yang digunakan untuk memprediksi label kelas dari objek yang label kelasnya tidak diketahui [8]. Kerangka kerja klasifikasi dijelaskan pada gambar dibawah ini. Pada kerangka kerja terdapat data latih yang digunakan untuk membangun model, kemudian model yang di bangun dari data latih akan digunakan untuk memprediksi kelas sehingga data uji diketahui kelas yang seharusnya. Kerangka kerja dibawah ini meliputi 2 proses, yaitu deduksi dan induksi. Induksi merupakan proses pelatihan dengan menggunakan data latih yang dimiliki, sedangkan deduksi merupakan penerapan model dengan data uji untuk melakukan prediksi kelas [4], [11].



Gambar 2. Kerangka Kerja Klasifikasi

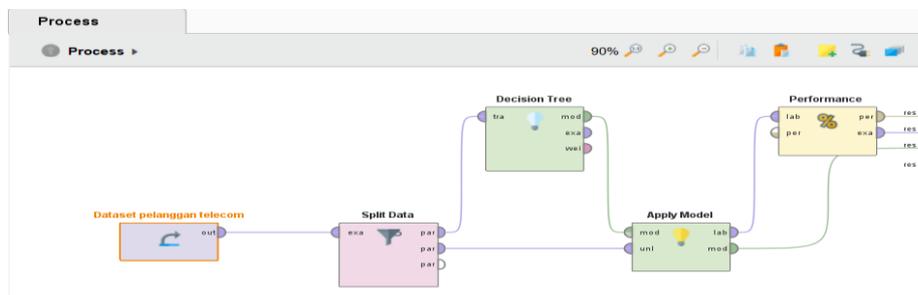
Algoritma pembentuk decision tree yaitu algoritma C4,5. Algoritma C4.5 merupakan algoritma induksi decision tree yaitu ID3. Input yang digunakan dalam algoritma ID3 berupa sampel training, atribut dan label training. Hasil pengembangan dari algoritma ID3 adalah algoritma C4.5. Dasar dari algoritma ID3 adalah membuat pohon keputusan dengan memperhatikan pemilihan atribut dengan prioritas tertinggi.. Kemudian secara berulang cabang-cabang pohon yang ada akan diperluas sehingga seluruh pohon terbentuk. Berikut adalah langkah pembuat keputusan dengan menggunakan algoritma C4.5[4][14]:

- Memilih atribut sebagai akar
- Membuat cabang
- Pembagian setiap kasus yang ada dalam cabang
- Pengulangan proses setiap cabang hingga setiap cabang memiliki kelas yang sama.

Penerapan algoritma C4.5 atau proses modelling merupakan tahapan untuk membuat model prediktif, yaitu untuk mengklasifikasi dan membuat akar pohon keputusan algoritma C4.5 dari dataset telekomunikasi industri. Alasan menggunakan algoritma C4.5 karena dapat memberikan karakteristik customer churn yang jelas, sehingga dapat membantu perusahaan telekomunikasi dalam memahami karteristik customer churn. Pada tahap ini dilakukan dengan menggunakan *machine learning* untuk menggali pengetahuan dari data yang ada untuk mencapai tujuan penelitian. Dalam penerapan algoritma C4.5 akan menggunakan Rapid Miner sebagai tools data mining.

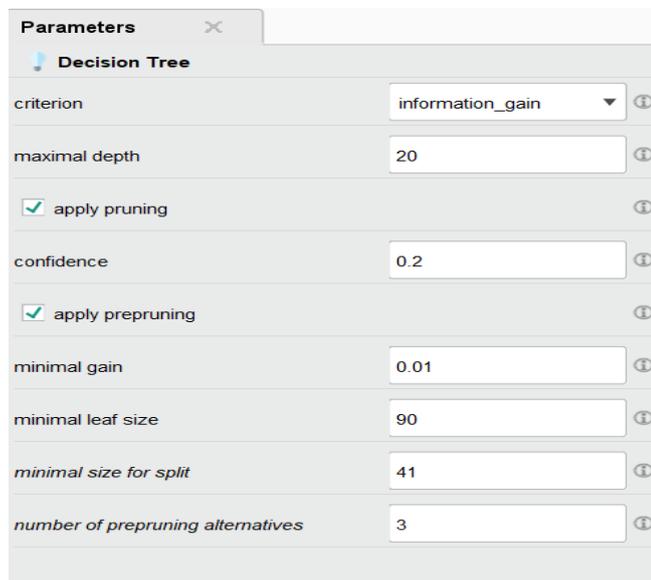
3. HASIL DAN PEMBAHASAN

Penerapan algoritma C4.5 atau proses modelling merupakan tahapan untuk membuat model prediktif, yaitu untuk mengklasifikasi dan membuat akar pohon keputusan algoritma C4.5 dari dataset telekomunikasi industri. Pada tahap ini dilakukan dengan menggunakan *machine learning* untuk mendapatkan insight dari data untuk mencapai tujuan penelitian. Untuk melakukan klasifikasi pada data pelanggan pada perusahaan telekomunikasi akan menggunakan algoritma C4.5. Berikut adalah proses yang telah dibuat pada Rapid Miner:



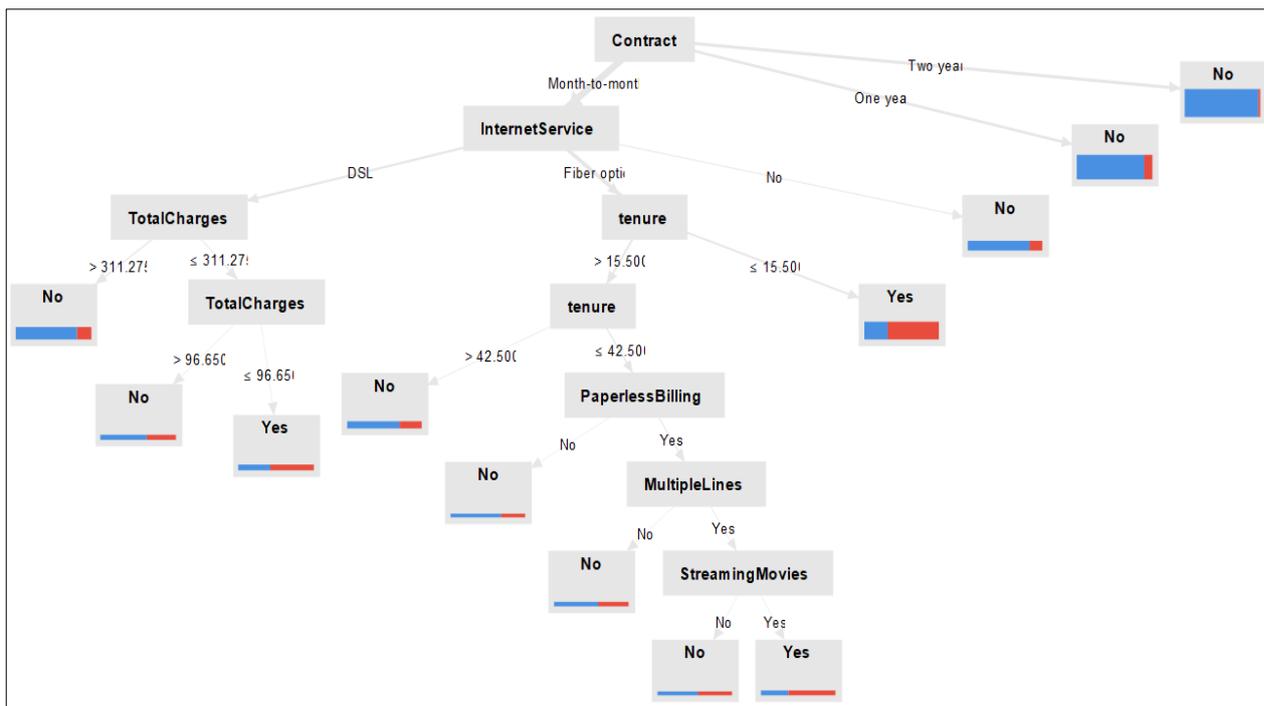
Gambar 3. Proses Modeling

Pada Gambar 3. Merupakan proses modeling atau penerapan algoritma C4.5 pada tools Rapid Miner. Diawali dengan dataset pelanggan telekomunikasi yang dimiliki, lalu dataset tersebut akan dibagi menjadi 2 yaitu data training dan data testing dengan sebuah module split data. Type split yang dipilih adalah linear sampling dengan ratio sebesar 0.7 untuk data training dan 0.3 untuk data testing. Setelah split data langkah selanjutnya yaitu membangun model decision tree. Data training akan diinput ke decision tree dan akan menghasilkan sebuah model yang akan di apply di module apply model. Model decision tree yang digunakan telah di sesuaikan lagi parameternya seperti yang dapat dilihat pada gambar 3. Sedangkan data testing akan diinput ke module apply model untuk mencoba model yang telah dibangun. Langkah terakhir yaitu melihat performa yang dihasilkan oleh model yang telah dibangun.



Gambar 4. Parameter Decision Tree

Dibawah ini merupakan pohon keputusan dari proses data mining yang telah dilakukan:



Gambar 5. Pohon Keputusan Data Pelanggan Telekomunikasi

Berdasarkan dari pohon keputusan yang dihasilkan, dapat dipahami bahwa 7 atribut memiliki pengaruh besar terhadap customer churn yaitu contract, InternetService, TotalChares, tenure, PaperlessBilling, MultipleLines, StreamingMovies. Atribut contract merupakan root dari pohon keputusan yang berarti contract merupakan variabel atau atribut yang perlu diperhatikan pertama kali untuk melihat karakteristik customer churn. Berikut dibawah ini merupakan aturan atau rule yang terbentuk dari pohon keputusan diatas :

Tree

```

Contract = Month-to-month
| InternetService = DSL
| | TotalCharges > 311.275: No {No=399, Yes=93}
| | TotalCharges ≤ 311.275
| | | TotalCharges > 96.650: No {No=102, Yes=64}
| | | TotalCharges ≤ 96.650: Yes {No=84, Yes=114}
| InternetService = Fiber optic
| | tenure > 15.500
| | | tenure > 42.500: No {No=181, Yes=76}
| | | tenure ≤ 42.500
| | | | PaperlessBilling = No: No {No=73, Yes=35}
| | | | PaperlessBilling = Yes
| | | | | MultipleLines = No: No {No=84, Yes=58}
| | | | | MultipleLines = Yes
| | | | | StreamingMovies = No: No {No=62, Yes=52}
| | | | | StreamingMovies = Yes: Yes {No=58, Yes=99}
| | tenure ≤ 15.500: Yes {No=227, Yes=496}
| InternetService = No: No {No=308, Yes=63}
Contract = One year: No {No=910, Yes=111}
Contract = Two year: No {No=1146, Yes=28}
    
```

Dari rule yang telah didapatkan dapat dipahami bahwa contract merupakan variabel yang perlu diperhatikan dari seorang customer. Customer dengan contract one year dan two year memiliki frekuensi ‘NO’ lebih tinggi. Sedangkan month-to-month perlu memperhatikan internet service, total charge, tenure, paperlessbilling, multiplelines dan streaming movies. Masing masing variabel sangat mempengaruhi customer churn, nilai frekuensi nya dapat dilihat pada rule di atas. Dari proses membangun model lalu menguji model dengan data testing yang telah disiapkan menghasilkan confusion matrix sebagai berikut :

Tabel 4. Confusion Matrix

	true No	true Yes	class precision
pred. No	1374	276	83.27%
pred. Yes	156	304	66.09%
class recall	89.80%	52.41%	

Accuracy: 79.53%

Berdasarkan hasil prediksi yang sudah tertera pada confusion matrix diatas menghasilkan nilai sebagai berikut[15]:

- a. Akurasi
Rasio dari prediksi benar dengan keseluruhan data disebut akurasi.
- b. Class Precision
Rasio dari nilai prediksi benar positif lalu dibandingkan dengan nilai dari seluruh prediksi positif disebut class precision
- c. Class recall
Rasio dari prediksi benar positif dibandingkan dengan nilai dari keseluruhan benar positif disebut class recall.

4. KESIMPULAN

Berdasarkan hasil penelitian yang diperoleh dengan menerapkan metode klasifikasi menggunakan algoritma C4.5 pada permasalahan customer churn pada perusahaan di industri telekomunikasi didapati beberapa karakteristik yang customer churn. Karakteristik tersebut yaitu contract, InternetService, TotalCharges, tenure, PaperlessBilling MultipleLines, StreamingMovies. Karakteristik contract merupakan karakteristik pertama yang harus dilihat dan diperhatikan pada customer churn karena pada penelitian ini karakteristik tersebut menjadi akar dari pohon keputusan. Dari karakteristik yang telah didapatkan tersebut dapat membantu perusahaan telekomunikasi untuk mengambil keputusan kedepannya dalam konteks mempertahankan customer yang ada, seperti memperhatikan customer yang contractnya month-to-month agar customer tersebut dapat bertahan lebih lama dalam menggunakan layanan yang ada.

Mengingat angka akurasi yang dihasilkan di output menggunakan algoritma C4.5 sebesar 79,53 maka penulis menyarankan penelitian selanjutnya untuk melakukan optimalisasi algoritma C4.5 menggunakan data yang ada untuk mendapatkan hasil yang optimal. Sehingga dapat mengetahui prediksi dari seorang customer akan churn atau tidak dengan lebih baik lagi.

REFERENCES

- [1] M. Syarif and W. Nugraha, "MWMOTE DALAM MENGATASI KETIDAKSEIMBANGAN KELAS PADA PREDIKSI CHURN MENGGUNAKAN KLASIFIKASI C4.5," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, Feb. 2023.
- [2] Anita, A. Wicaksono, and T. N. Padilah, "Pengaruh Jumlah Record Dataset Terhadap Algoritma Klasifikasi Berdasarkan Data Customer Churn," *Jurnal Ilmiah Informatika*, vol. 6, no. 1, pp. 1–10, Jun. 2021, doi: 10.35316/jimi.v6i1.1223.
- [3] H. N. Irmanda, R. Astriratma, and S. Afrizal, "PERBANDINGAN METODE JARINGAN SYARAF TIRUAN DAN POHON KEPUTUSAN UNTUK PREDIKSI CHURN," *JSI: Jurnal Sistem Informasi (E-Journal)*, vol. 11, no. 2, 2019, [Online]. Available: <http://ejournal.unsri.ac.id/index.php/jsi/index>
- [4] Y. Tri Utami, D. Asiah Shofiana, Y. Heningtyas, J. Ilmu Komputer FMIPA Universitas Lampung Jl Soemantri Brodjonegoro No, and B. Lampung, "Penerapan Algoritma C4.5 Untuk Prediksi Churn Rate Pengguna Jasa Telekomunikasi," 2020.
- [5] A. Wicaksono, "Uji Performa Teknik Klasifikasi untuk Memprediksi Customer Churn," *Bianglala Informatika*, vol. 9, no. 1, 2021.
- [6] R. Govindaraju et al., "PERANCANGAN SISTEM PREDIKSI CHURN PELANGGAN PT. TELEKOMUNIKASI SELULER DENGAN MEMANFAATKAN PROSES DATA MINING."
- [7] R. Pratama, M. I. Herdiansyah, D. Syamsuar, and A. Syazili, "Prediksi Customer Retention Perusahaan Asuransi Menggunakan Machine Learning," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 1, pp. 96–104, Mar. 2023, doi: 10.32736/sisfokom.v12i1.1507.
- [8] F. A. Pambudi, A. P. Windarto, M. Fauzan, J. T. Hardinata, and R. Winanjaya, "Analisis Klasifikasi C4.5 Pada Pola Pembayaran Sepeda Motor Adira Cabang Pematangsiantar," *Kajian Ilmiah Informatika dan Komputer*, vol. 1, no. 4, pp. 160–169, 2021, [Online]. Available: <https://djournals.com/klik>
- [9] Y. Indah, "Prediksi Tingkat Kepuasan Pelayanan Online Menggunakan Metode Algoritma C4.5," *Jurnal Informatika Ekonomi Bisnis*, pp. 59–64, Jun. 2022, doi: 10.37034/inf.v4i2.99.
- [10] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *Jurnal Edik informatika*.
- [11] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "ANALISIS DATA MINING DATA NETFLIX MENGGUNAKAN APLIKASI RAPID MINER," *JBASE - Journal of Business and Audit Information Systems*, vol. 4, no. 1, Apr. 2021, doi: 10.30813/jbase.v4i1.2729.
- [12] I. Latief, A. Subekti, and W. Gata, "PREDIKSI TINGKAT PELANGGAN CHURN PADA PERUSAHAAN TELEKOMUNIKASI DENGAN ALGORITMA ADABOOST," *Jurnal Informatika*, vol. 21, 2021.
- [13] L. Navia, "Klasifikasi Nasabah Menggunakan Algoritma C4.5 Sebagai Dasar Pemberian Kredit," *JURNAL INOVTEK POLBENG*, vol. 1, no. 2, 2016.
- [14] K. Aidi Saputra, J. Tata Hardinata, M. Ridwan Lubis, S. Retno Andani, and I. Syahputra Saragih, "KLIK: Kajian Ilmiah Informatika dan Komputer Klasifikasi Algoritma C4.5 Dalam Penerapan Tingkat Kepuasan Siswa Terhadap Media Pembelajaran Online," *Media Online*, vol. 1, no. 3, pp. 113–118, 2020, [Online]. Available: <https://djournals.com/klik>
- [15] D. Marlina and M. Bakri, "PENERAPAN DATA MINING UNTUK MEMPREDIKSI TRANSAKSI NASABAH DENGAN ALGORITMA C4.5," *Jurnal Teknologi dan Sistem Informasi (JTSI)*, vol. 2, Mar. 2021.