

Penerapan Algoritma C4.5 untuk Klasifikasi Customer Churn pada Perusahaan Perbankan

Mohammad Aulia Riftiarraafi

Universitas Pembangunan Nasional “Veteran” Jawa Timur

Dira Ernawati

Universitas Pembangunan Nasional “Veteran” Jawa Timur

Jl. Rungkut Madya No.1, Gn. Anyar, Kec. Gn. Anyar, Kota SBY, Jawa Timur 60294

Korespondensi penulis: 20032010091@student.upnjatim.ac.id

Abstract. *In the fast-paced digital era, companies must continue to innovate to survive. Especially in banks, retaining customers is the main strategy to be able to project business continuity. The topic of customer retention pays more attention to studying customer patterns that have a tendency to survive. This is very difficult to do because the observed data is generally very complex and amounts to a very large amount. Therefore, computers are a technology that is very suitable to solve these problems. Decision tree is a method that is currently being widely used for research. Moreover, with various versions that continue to be developed such as the C4.5 algorithm, it is a guarantee that this algorithm is still relevant for use in today's industry. Even so, there is still no customer retention prediction research with banking industry case studies. The results of this research test, prove that using the C4.5 algorithm, successfully predict customer churn with an accuracy of 99.77%.*

Keywords: C4.5 Algorithm, Customer Churn, Decision Tree.

Abstrak. Di era digital yang serba cepat, Perusahaan haruslah terus berinovasi untuk dapat bertahan. Terkhusus pada bank, mempertahankan pelanggan menjadi strategi utama untuk bisa memproyeksikan keberlangsungan bisnis. Topik mengenai customer retention, memberikan perhatian lebih kepada mempelajari pola pelanggan yang memiliki kecenderungan untuk bisa bertahan. Hal ini sangatlah sulit untuk dilakukan karena data yang diamati umumnya sangatlah kompleks dan berjumlah sangat banyak. Oleh karena itu, computer menjadi teknologi yang sangat sesuai untuk menyelesaikan permasalahan tersebut. Decision tree adalah metode yang saat ini sedang marak digunakan untuk penelitian. Terlebih lagi dengan ragam versi yang terus dikembangkan seperti algoritma C4.5 menjadi jaminan bahwa algoritma ini masih tetap relevan untuk digunakan pada industri saat ini. Meskipun begitu, masih belum ada penelitian prediksi customer retention dengan studi kasus industri perbankan. Hasil uji penelitian ini, membuktikan bahwa menggunakan algoritma C4.5, berhasil memprediksi customer churn dengan akurasi sebesar 99.77%.

Kata kunci: Algoritma C4.5, Customer Churn, Decision Tree.

LATAR BELAKANG

Di era digital ini yang semuanya serba cepat, Perusahaan haruslah terus berinovasi untuk dapat bertahan. Terkhusus pada bank, customer retention menjadi focus utama untuk bisa memproyeksikan keberlangsungan bisnis. Hal ini menjadi sangat penting, karena bagi bank pelanggan tetaplah yang memberikan keuntungan terbesar. sehingga strategi untuk dapat mempertahankan pelanggan menjadi topik yang bagus untuk diteliti (Kassem et al., 2020; NGUYEN et al., 2020). Untuk menyelesaikan permasalahan tersebut, peneliti umumnya

menitik beratkan kepada Customer Relationship Management (Anees et al., 2020; Simanjuntak et al., 2020), Kualitas Pelayanan (NGUYEN et al., 2020; Slack et al., 2020), dan Feedback pelanggan (Larsson & Broström, 2020).

Topik mengenai customer retention, memberikan perhatian lebih kepada mempelajari pola pelanggan yang memiliki kecenderungan untuk bisa bertahan. Hal ini sangatlah sulit untuk dilakukan karena data yang diamati umumnya sangatlah kompleks dan berjumlah sangat banyak. Oleh karena itu, computer menjadi teknologi yang sangat sesuai untuk menyelesaikan permasalahan tersebut. Bahkan saat ini, Perusahaan dengan model B2B atau Business to Business mulai menggunakan kecerdasan buatan untuk membantu memetakan mana saja pelanggan mereka yang kemungkinan terbesar akan menetap (Kassem et al., 2020). Sayangnya permasalahan itu hanya diteliti pada model B2B saja, belum ada yang mengangkat topik mengenai aplikasi klasifikasi customer churn pada studi kasus bank karena bank tidak hanya menggunakan model B2B saja namun juga B2C atau Business to Customer. Padahal untuk mempertimbangkan churn pada skala yang lebih kecil yaitu pada customer adalah masalah yang lebih spesifik dan kompleks apabila dibandingkan dengan kerjasama bisnis. Oleh karena itu, dalam penelitian ini akan mengujikan data customer individu bank dengan menggunakan Decision Tree untuk memprediksi customer churn.

Decision tree adalah metode yang saat ini sedang marak digunakan untuk penelitian. Terlebih lagi dengan ragam versi yang terus dikembangkan menjadi jaminan bahwa algoritma ini masih tetap relevan untuk digunakan pada industry saat ini. Diantara semua varian decision tree yang ada, satu yang paling menonjol yakni Decision Tree C4.5 Varian ini memiliki keunggulan untuk memudahkan pengembang memodifikasi konfigurasi untuk bisa mencapai parameter terbaik (Harahap et al., 2022). Selain itu varian ini sudah diterapkan pada penelitian dengan studi kasus di berbagai bidang seperti BUMN (Muttaqien et al., 2021; Sinaga et al., 2021), Manufaktur (Apriliyawan & Lestari, n.d.), dan Telekomunikasi (Desena Damanik & Ihsan Jambak, 2023; Gu et al., 2021). Meskipun begitu, masih belum ada penelitian yang mengangkat studi kasus industry perbankan menggunakan algoritma C4.5.

KAJIAN TEORITIS

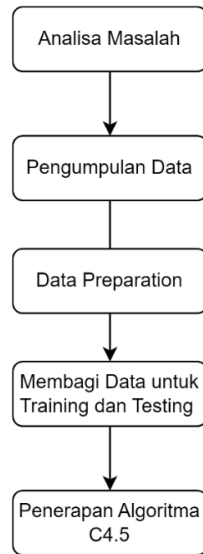
Dalam 5 tahun terakhir, memprediksi customer churn rate berdasarkan data customer menjadi topik yang hangat untuk diteliti. Seiring dengan waktu, data yang digunakan pun juga semakin besar dan memiliki atribut yang beraneka ragam. Untuk menyelesaikan hal ini,

banyak sekali metode yang dapat digunakan. Mulai dari feature selection (Alkitbi et al., 2021; Kassem et al., 2020) hingga klasifikasi (AL-Shatnwai & Faris, 2020; Kelvin et al., 2020; Miryam Clementine & Arum, 2022; Suhandha et al., 2022). Meskipun begitu, perlu diperhatikan bahwa data dengan dimensi yang besar juga memiliki pemrosesan waktu yang semakin besar pula. Oleh karena itu, penelitian mengenai topik ini seringkali didominasi oleh metode yang bukan berbasis deeplearning seperti FP-Growth (Kelvin et al., 2020), XGBoost (AL-Shatnwai & Faris, 2020), Random Forest (Suhanda et al., 2022), dan Decision Tree (Höppner et al., 2020; Saleh & Saha, 2023). Decision tree dipilih karena dianggap memiliki waktu pemrosesan tercepat, serta masih bisa mendapatkan akurasi yang baik (Kassem et al., 2020).

Decision Tree yang populer diterapkan dalam 5 tahun terakhir adalah Varian Algoritma C4.5. Karena pada dasarnya Algoritma C4.5 memudahkan pengembang untuk memodifikasi konfigurasinya, hal ini menyebabkan maraknya riset menggunakan algoritma ini. Diantaranya yaitu pengaplikasian algoritma (Desena Damanik & Ihsan Jambak, 2023; Kim & Lee, 2022), Data Mining (Muttaqien et al., 2021; Sinaga et al., 2021), Big Data (Gu et al., 2021), dan Integrasi dengan algoritma lain seperti Naïve Bayes (Apriliyawan & Lestari, n.d.) maupun K-Means (Harahap et al., 2022). Terlebih lagi, algoritma ini juga sudah dapat dengan mudah dimodifikasi melalui aplikasi RapidMiner (Desena Damanik & Ihsan Jambak, 2023).

METODE PENELITIAN

Metodologi yang digunakan dalam penelitian ini adalah metode CRISP-DM sebagai alur penelitian, menerapkan business dan data understanding untuk menganalisa dan memahami masalah yang ada. Oleh karena mendeteksi churn adalah suatu proses yang kompleks oleh karena itu dalam penelitian ini akan menggunakan Decision Tree C4.5 sebagai pilihan metode. Berikut merupakan gambaran kerangka pikir yang dilakukan dalam penelitian ini.



Gambar 1. Kerangka Penelitian

Pada penelitian ini digunakan dataset yang diperoleh dari Kaggle dengan link URL: <https://www.kaggle.com/datasets/radheshyamkollipara/bank-customer-churn>. Dalam dataset uji penelitian, terdapat 10000 data yang masing-masing memiliki 16 atribut. Atribut-atribut ini merupakan faktor yang mempengaruhi customer churn pada perusahaan perbankan. Meskipun begitu bobot pengaruh atribut tersebut dapat mempengaruhi customer churn itulah yang akan dilatihkan pada model Decision tree.

Atribut	Keterangan
CustomerId	Id Pelanggan
Surname	Nama nasabah bang
CreditScore	Skor kredit ekonomi nasabah
Geography	Lokasi tinggal nasabah
Gender	Jenis kelamin nasabah
Age	Umur nasabah
Tenure	Lama tahun menjadi nasabah
Balance	Saldo yang dimiliki nasabah
NumOfProducts	Jumlah produk yang telah dibeli nasabah melalui bank
HasCrCard	Kepemilikan kartu kredit
IsActiveMember	Nasabah yang aktif mengikuti program bank
EstimatedSalary	Gaji nasabah
Exited	Keputusan nasabah untuk meninggalkan bank atau tidak
Complain	Apakah nasabah mengajukan komplain
Satisfaction Score	Skor kepuasan nasabah terhadap bank
Card Type	Tipe kartu yang dimiliki nasabah

Gambar 2. Deskripsi Atribut Dataset

Langkah selanjutnya adalah data preparation, meliputi menyusun, mengatur, menggabungkan dan mengumpulkan data sehingga data dapat digunakan pada proses machine learning. Data preparation adalah langkah untuk mempersiapkan data yang dimiliki agar sesuai

dengan model yang akan digunakan. Berikut adalah langkah yang dilakukan pada tahap data preparation :

- a. Drop kolom yang tidak digunakan.

Terdapat 1 kolom yang tidak digunakan yaitu kolom CustomerId. Kolom tersebut akan dihapus dari dataset yang akan digunakan

- b. Handling missing value.

Dalam dataset pelanggan perusahaan perbankan tidak terdapat missing value atau Null, Maka langkah ini dapat dilewati

- c. Transformasi data.

Pada kolom HasCrCard, IsActiveMember, Exited, dan Complain memiliki nilai 1 dan 0. Hal ini tidak sesuai dengan jenis data yang dapat diinputkan ke algoritma C4.5 sehingga perlu dilakukan transformasi data. Kolom-kolom tersebut akan ditransformasi sehingga kolom yang memiliki nilai 0 menjadi “No” dan kolom yang memiliki nilai 1 akan menjadi “Yes”

CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts
15634602	Hargrave	600 - 850	France	Female	30 - 60		2 0 - 100 k	1
15647311	Hill	600 - 850	Spain	Female	30 - 60		1 0 - 100 k	1
15619304	Onio	350 - 600	France	Female	30 - 60		8 100 - 250 k	3
15701354	Boni	600 - 850	France	Female	30 - 60		1 0 - 100 k	2
15737888	Mitchell	600 - 850	Spain	Female	30 - 60		2 100 - 250 k	1
15574012	Chu	600 - 850	Spain	Male	30 - 60		8 100 - 250 k	2
15592531	Bartlett	600 - 850	France	Male	30 - 60		7 0 - 100 k	2
15656148	Obinna	350 - 600	Germany	Female	30 kebawah		4 100 - 250 k	4
15792365	He	350 - 600	France	Male	30 - 60		4 100 - 250 k	2

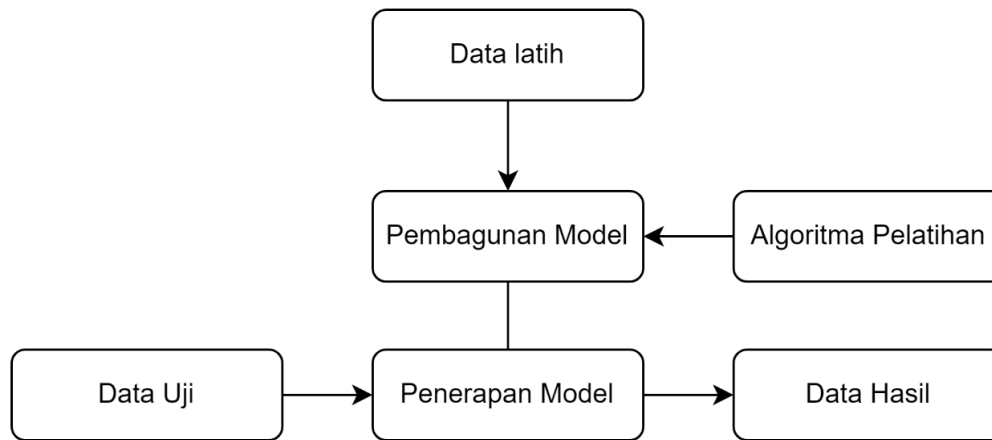
Gambar 3. Dataset setelah Preparation

HasCrCard	IsActiveMember	EstimatedSalary	Exited	Complain	Satisfaction Score	Card Type
Yes	Yes	100 - 200 k	Yes	Yes		2 DIAMOND
No	Yes	100 - 200 k	No	Yes		3 DIAMOND
Yes	No	100 - 200 k	Yes	Yes		3 DIAMOND
No	No	0 - 100 k	No	No		5 GOLD
Yes	Yes	0 - 100 k	No	No		5 GOLD
Yes	No	100 - 200 k	Yes	Yes		5 DIAMOND
Yes	Yes	0 - 100 k	No	No		2 SILVER
Yes	No	100 - 200 k	Yes	Yes		2 DIAMOND
No	Yes	0 - 100 k	No	No		3 GOLD

Gambar 4. Lanjutan Dataset setelah Preparation

Pada penelitian ini akan menggunakan RapidMiner sebagai alat bantu machine learning. Tahapan dalam menentukan 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 menganalisa training yang digunakan untuk

memprediksi label kelas dari objek yang label kelasnya tidak diketahui. Kerangka kerja klasifikasi dijelaskan pada gambar dibawah ini.



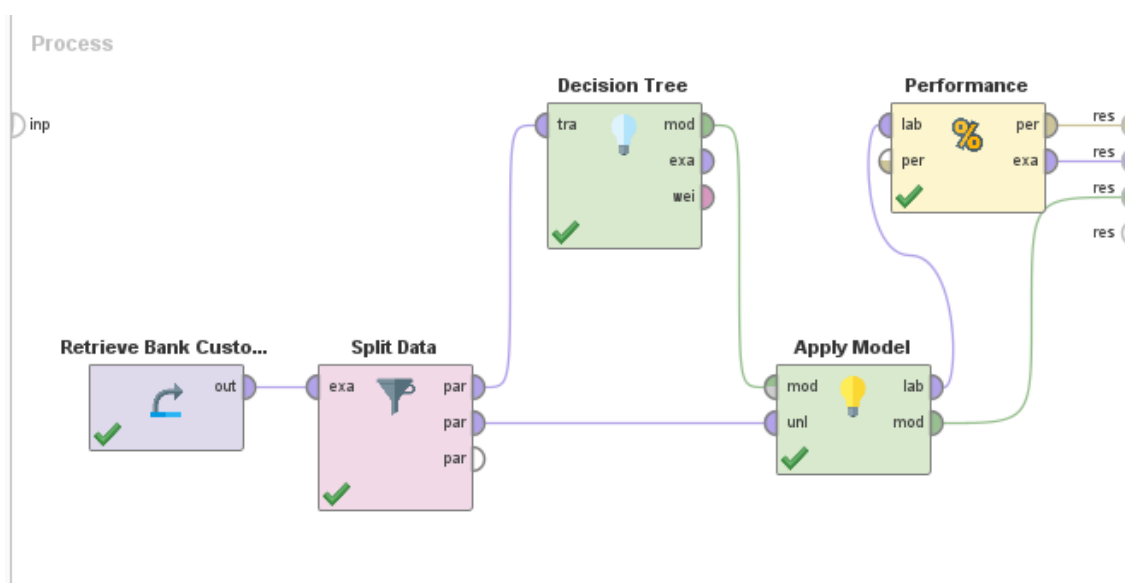
Gambar 5. Kerangka Kerja Klasifikasi

Algoritma pembentuk decision tree yaitu algoritma C4.5. Algoritma C4.5 merupakan algoritma induk decision tree yaitu ID3. Input yang digunakan dalam algoritma ID3 berupa sample 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 menggunakan algoritma C4.5 (Desena Damanik & Ihsan Jambak, 2023).

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

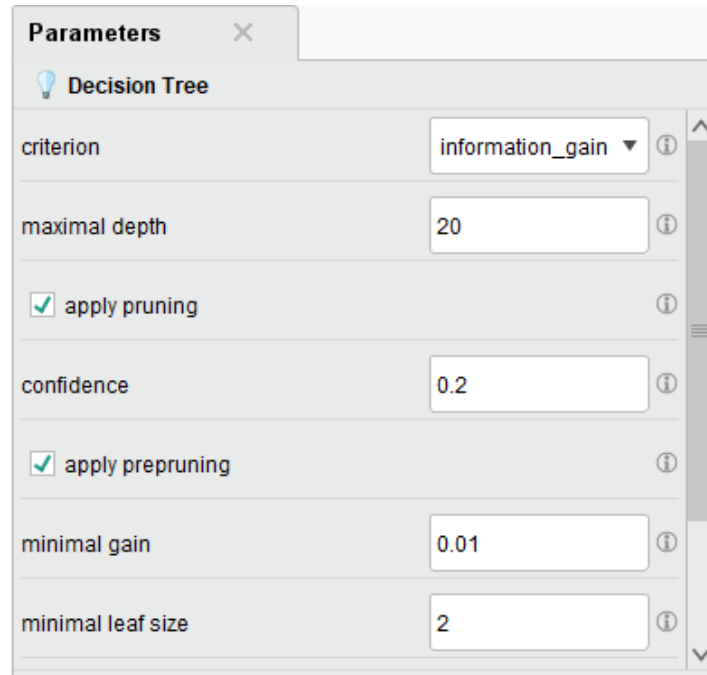
Penerapan algoritma C4.5 atau proses modelling merupakan tahapan untuk membuat model prediktif, yaitu untuk mengklasifikasi dan membuat akar pohon keputusan algoritma C4.5 dari dataset perusahaan perbankan. Pada tahap ini dilakukan dengan menggunakan

machine learning untuk mendapatkan insight dari data untuk mencapai tujuan penelitian. Berikut adalah proses yang telah dibuat pada RapidMiner.



Gambar 6. Alur Melatih Model Decision Tree

Pada Gambar 6 merupakan proses modelling atau penerapan algoritma C4.5 pada tools RapidMiner. Diawali dengan menginput dataset, lalu dataset tersebut akan dibagi menjadi data training dan data testing dengan rasio sebesar 0.7 dan 0.3. Setelah split data langkah selanjutnya yaitu membangun model decision tree, model decision tree yang digunakan menggunakan parameter yang telah disesuaikan seperti yang dapat dilihat pada gambar 7. Sedangkan data testing akan diinput ke module apply model untuk mencoba model yang telah dibangun. Langkah terakhir yaitu melihat hasil atau performa yang telah dihasilkan oleh model yang telah dibangun.



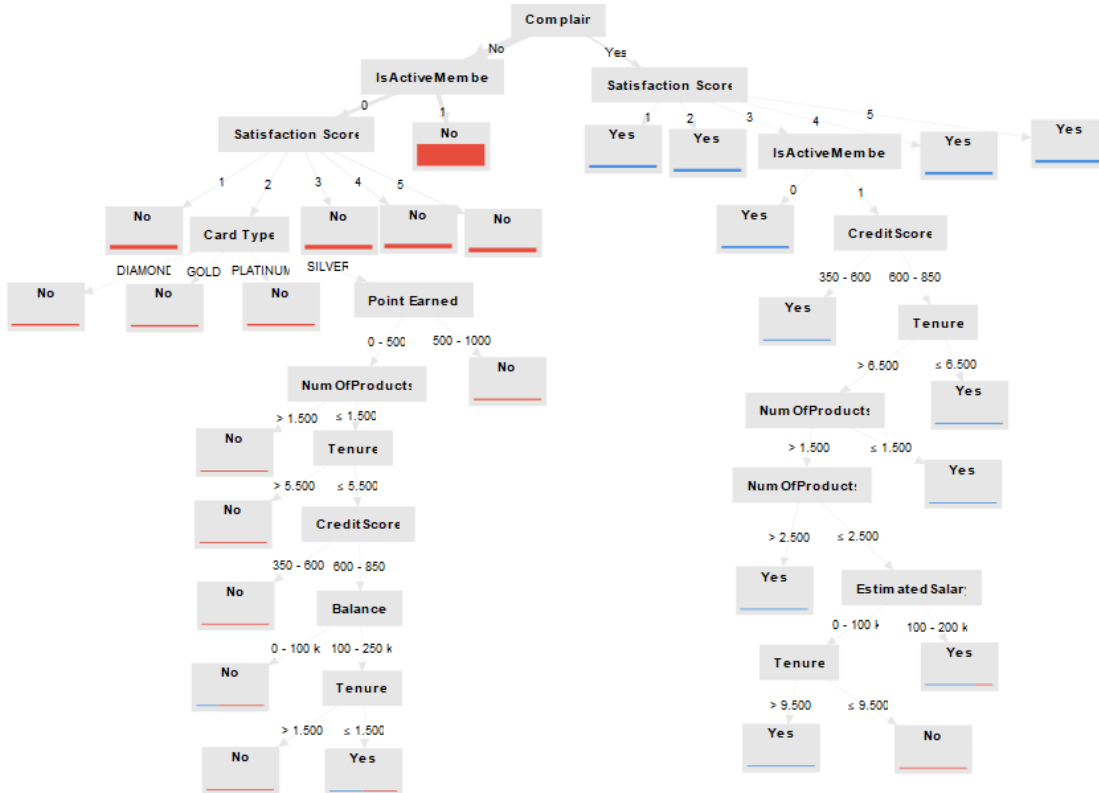
Gambar 7. Konfigurasi Decision Tree Pada RapidMiner

Dalam gambar 7 tertera parameter yang digunakan dalam memodelkan decision tree pada rapidminer, konfigurasi decision tree yang digunakan adalah criterion yang mengatur kriteria training decision tree, maximal depth yang mengatur Tingkat kedalaman leaf node dari tree yang dapat dicapai, dan minimal leaf size yang menandakan bahwa tree hanya boleh memiliki dua percabangan karena studi kasus yang digunakan adalah Klasifikasi Biner yang hanya bisa mendeteksi dua label.

HASIL DAN PEMBAHASAN

Berdasarkan hasil decision tree, dapat dipahami bahwa atribut Complain merupakan root dari decision tree yang berarti Complain merupakan atribut yang perlu diperhatikan pertama kali untuk melihat karakteristik customer churn. Gambar dibawah merupakan jabaran yang terbentuk dari decision tree pada Gambar 8.

Penerapan Algoritma C4.5 untuk Klasifikasi Customer Churn pada Perusahaan Perbankan



Gambar 8. Hasil Pemetaan Decision Tree

```

Tree
Complain = No
  IsActiveMember = 0
    Satisfaction Score = 1: No {Yes=1, No=481}
    Satisfaction Score = 2
      Card Type = DIAMOND: No {Yes=0, No=118}
      Card Type = GOLD: No {Yes=0, No=120}
      Card Type = PLATINUM: No {Yes=0, No=122}
      Card Type = SILVER
        Point Earned = 0 - 500
          NumOfProducts > 1.500: No {Yes=0, No=22}
          NumOfProducts ≤ 1.500
            Tenure > 5.500: No {Yes=0, No=8}
            Tenure ≤ 5.500
              CreditScore = 350 - 600: No {Yes=0, No=2}
              CreditScore = 600 - 850
                Gender = Female
                  Balance = 0 - 100 k : Yes {Yes=1, No=1}
                  Balance = 100 - 250 k: No {Yes=0, No=4}
                Gender = Male: No {Yes=1, No=2}
              Point Earned = 500 - 1000: No {Yes=0, No=78}
            Satisfaction Score = 3: No {Yes=0, No=502}
            Satisfaction Score = 4: No {Yes=1, No=502}
            Satisfaction Score = 5: No {Yes=0, No=510}
          IsActiveMember = 1: No {Yes=0, No=3093}
    Complain = Yes
      Satisfaction Score = 1: Yes {Yes=276, No=1}
      Satisfaction Score = 2: Yes {Yes=297, No=1}
      Satisfaction Score = 3
        IsActiveMember = 0: Yes {Yes=186, No=1}
        IsActiveMember = 1
          CreditScore = 350 - 600: Yes {Yes=33, No=0}
          CreditScore = 600 - 850
            Tenure > 6.500
              NumOfProducts > 1.500
                NumOfProducts > 2.500: Yes {Yes=2, No=0}
                NumOfProducts ≤ 2.500
                  Gender = Female: Yes {Yes=4, No=1}
                  Gender = Male: No {Yes=1, No=2}
              NumOfProducts ≤ 1.500: Yes {Yes=12, No=0}
            Tenure ≤ 6.500: Yes {Yes=41, No=1}
          Satisfaction Score = 4: Yes {Yes=276, No=0}
          Satisfaction Score = 5: Yes {Yes=295, No=1}
  
```

Gambar 9. Pseudocode Cabang Decision Tree

Dari rule yang dapat dilihat pada Gambar 9, dapat dipahami bahwa Complain merupakan variabel yang perlu diperhatikan untuk mengantisipasi customer churn. Nasabah yang mengajukan complain bisa dipastikan akan churn, dengan pertimbangan satisfaction score mereka dengan bank, apakah mereka member aktif bank, credit score, tenure, dan jumlah produk yang telah mereka beli melalui bank. Dari proses pembangunan model lalu menguji model dengan data testing yang telah disiapkan menghasilkan confusion matrix sebagai berikut :

accuracy: 99.77%

	true Yes	true No	class precision
pred. Yes	610	6	99.03%
pred. No	1	2383	99.96%
class recall	99.84%	99.75%	

Gambar 10. Confusion Matrix Hasil Pengujian Decision Tree

Berdasarkan hasil prediksi yang tertera pada confusion matrix pada Gambar 10, didapatkan informasi sebagai berikut: Akurasi dari proses uji decision tree menghasilkan akurasi sebesar 99.77%, dengan precision dari prediksi churn sebesar 99.03% dan precision dari prediksi not-churn sebesar 99.96%, serta Recall label churn sebesar 99.84% dan recall label not-churn sebesar 99.75%.

KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang diperoleh dengan menerapkan metode klasifikasi menggunakan algoritma C4.5 pada permasalahan customer churn pada perusahaan perbankan didapatkan beberapa karakteristik customer churn. Karakteristik tersebut yaitu Complain, Satisfaction Score, IsActiveMember, Tenure, dan NumOfProducts. Karakteristik Complain merupakan karakteristik pertama yang harus diperhatikan pada customer churn karena pada penelitian ini karakteristik tersebut menjadi akar dari decision tree. Karakteristik tersebut dapat membantu perusahaan perbankan untuk mengambil keputusan kedepannya dalam konteks mempertahankan nasabah yang ada, seperti memperhatikan layanan yang diberikan agar nasabah dapat bertahan lebih lama. Hasil akurasi decision tree dengan algoritma C4.5 yang dilatihkan menggunakan data customer perbankan, menghasilkan akurasi sebesar 99.77%. Hal ini berarti model dapat memprediksi customer churn dengan kemungkinan 99.77% benar.

Dengan memperhatikan keterbatasan yang dimiliki penelitian ini, yakni pada keterbatasan akses mengenai data perbankan yang lebih besar dan kurang seimbang antar kedua kelas tersebut. Akan lebih baik untuk melakukan penelitian dengan menggunakan Algoritma C4.5 dengan menggunakan data yang lebih besar. Selain itu terdapat juga topik untuk menyeimbangkan data antara kedua kelas tersebut dengan data augmentation.

DAFTAR REFERENSI

- Alkitbi, S. S., Alshurideh, M., Al Kurdi, B., & Salloum, S. A. (2021). *Factors Affect Customer Retention: A Systematic Review* (pp. 656–667). https://doi.org/10.1007/978-3-030-58669-0_59
- AL-Shatnwai, A. M., & Faris, M. (2020). Predicting Customer Retention using XGBoost and Balancing Methods. *International Journal of Advanced Computer Science and Applications*, 11(7). <https://doi.org/10.14569/IJACSA.2020.0110785>
- Anees, R. T., Nordin, N. A., Anjum, T., Cavaliere, L. P. L., & Heidler, P. (2020). Evaluating the Impact of Customer Relationship Management (CRM) Strategies on Customer Retention (A Study of Fast Food Chains in Pakistan). *Business Management and Strategy*, 11(2), 117. <https://doi.org/10.5296/bms.v11i2.17934>
- Apriliyawan, B., & Lestari, S. (n.d.). *Performance Analysis Of C4.5 And Naïve Bayes Algorithm On Customer Relationship Management (CRM) In Jatimas Furniture Ltd.*
- Desena Damanik, S., & Ihsan Jambak, M. (2023). Klasifikasi Customer Churn pada Telekomunikasi Industri Untuk Retensi Pelanggan Menggunakan Algoritma C4.5. *Kajian Ilmiah Informatika Dan Komputer*, 3(6), 1303–1309. <https://doi.org/10.30865/klik.v3i6.829>
- Gu, Y., Palaoag, T. D., & Dela Cruz, J. S. (2021). Comparison of Main Algorithms in Big Data Analysis of Telecom Customer Retention. *IOP Conference Series: Materials Science and Engineering*, 1077(1), 012045. <https://doi.org/10.1088/1757-899X/1077/1/012045>
- Höppner, S., Stripling, E., Baesens, B., Broucke, S. vanden, & Verdonck, T. (2020). Profit driven decision trees for churn prediction. *European Journal of Operational Research*, 284(3), 920–933. <https://doi.org/10.1016/j.ejor.2018.11.072>
- Kassem, E. A. el, Ali, S., Mostafa, A., & Kamal, F. (2020). Customer Churn Prediction Model and Identifying Features to Increase Customer Retention based on User Generated Content. *International Journal of Advanced Computer Science and Applications*, 11(5). <https://doi.org/10.14569/IJACSA.2020.0110567>
- Kelvin, Cindy, Charles, Peter Leonardo, D., & Yennimar. (2020). Customer Churn's Analysis In Telecommunications Company Using Fp-Growth Algorithm. *Jurnal Mantik*, 4(2), 1285–1291. <https://doi.org/https://doi.org/10.35335/mantik.Vol4.2020.933.pp1285-1290>

- Kim, S., & Lee, H. (2022). Customer Churn Prediction in Influencer Commerce: An Application of Decision Trees. *Procedia Computer Science*, 199, 1332–1339. <https://doi.org/10.1016/j.procs.2022.01.169>
- Larsson, A., & Broström, E. (2020). Ensuring customer retention: insurers' perception of customer loyalty. *Marketing Intelligence & Planning*, 38(2), 151–166. <https://doi.org/10.1108/MIP-02-2019-0106>
- Miryam Clementine, & Arum. (2022). Prediksi Churn Nasabah Bank Menggunakan Klasifikasi Naïve Bayes dan ID3. *Jurnal Processor*, 17(1), 9–18. <https://doi.org/10.33998/processor.2022.17.1.1170>
- Muttaqien, R., Galih, M., & Pramuntadi, A. (2021). Implementation of Data Mining Using C4.5 Algorithm for Predicting Customer Loyalty of PT. Pegadaian (Persero) Pati Area Office. In *International Journal of Computer and Information System (IJCIS) Peer Reviewed-International Journal* (Vol. 02, Issue 03). <https://ijcis.net/index.php/ijcis/indexJournalIJCIShomepage-https://ijcis.net/index.php/ijcis/index>
- NGUYEN, D. T., PHAM, V. T., TRAN, D. M., & PHAM, D. B. T. (2020). Impact of Service Quality, Customer Satisfaction and Switching Costs on Customer Loyalty. *The Journal of Asian Finance, Economics and Business*, 7(8), 395–405. <https://doi.org/10.13106/jafeb.2020.vol7.no8.395>
- Harahap, N. H. S., Amirullah, A., Saputro, M. B., & Tamaroh, I. A. (2022). Classification of potential customers using C4.5 and k-means algorithms to determine customer service priorities to maintain loyalty. *Journal of Soft Computing Exploration*, 3(2). <https://doi.org/10.52465/josce.v3i2.89>
- Saleh, S., & Saha, S. (2023). Customer retention and churn prediction in the telecommunication industry: a case study on a Danish university. *SN Applied Sciences*, 5(7), 173. <https://doi.org/10.1007/s42452-023-05389-6>
- Simanjuntak, M., Putri, N. E., Yuliati, L. N., & Sabri, M. F. (2020). Enhancing customer retention using customer relationship management approach in car loan bussiness. *Cogent Business & Management*, 7(1), 1738200. <https://doi.org/10.1080/23311975.2020.1738200>
- Sinaga, T. H., Wanto, A., Gunawan, I., Sumarno, S., & Nasution, Z. M. (2021). Implementation of Data Mining Using C4.5 Algorithm on Customer Satisfaction in Tirta Lihou PDAM. *Journal of Computer Networks, Architecture, and High-Performance Computing*, 3(1), 9–20. <https://doi.org/10.47709/cnahpc.v3i1.923>
- Slack, N., Singh, G., & Sharma, S. (2020). The effect of supermarket service quality dimensions and customer satisfaction on customer loyalty and disloyalty dimensions.

International Journal of Quality and Service Sciences, 12(3), 297–318.
<https://doi.org/10.1108/IJQSS-10-2019-0114>

Suhanda, Y., Nurlaela, L., Kurniati, I., Dharmalau, A., & Rosita, I. (2022). Predictive Analysis of Customer Retention Using the Random Forest Algorithm. *TIERS Information Technology Journal*, 3(1), 35–47. <https://doi.org/10.38043/tiers.v3i1.3616>