

Perbandingan Metode Klasifikasi Data Mining untuk Prediksi Tingkat Kepuasan Pelanggan Telkomsel Prabayar

Fattya Ariani¹, Andi Taufik²

¹STMIK Nusa Mandiri, fattya.fty@nusamandiri.ac.id, Jln. Jatiwaringin Raya No.2, Jakarta Timur, Indonesia

²STMIK Nusa Mandiri, a.taufik30@gmail.com, Jln. Jatiwaringin Raya No.2, Jakarta Timur, Indonesia

Informasi Makalah

Submit : 06 Okt 2020
Revisi : 03 Nov 2020
Diterima : 30 Nov 2020

Kata Kunci :

Algoritma C.45
Support Vector Machine
Naïve Bayes
Particle Swarm Optimization
Tingkat kepuasan
Data Mining
Klasifikasi

Abstrak

Saat ini Indonesia mempunyai lebih dari satu operator seluler yang aktif. Perkembangan teknologi yang membuat segala aktivitas masyarakat dilakukan secara *mobile* atau *online*. Karena kebutuhan komunikasi sangat tinggi maka perusahaan operator saling berlomba-lomba untuk mengambil hati para pelanggan dengan peningkatan layanannya. Telkomsel menjadi salah satu operator terbesar di Indonesia. Untuk memberikan pelayanan terbaik, telkomsel harus konsisten dan meningkatkan pelayanan. Ada beberapa metode yang dapat digunakan untuk menganalisa klasifikasi kepuasan pelanggan seperti *Support Vector Machine*, *Decission Tree* dan Metode *Naïve Bayes* dan PSO. Tetapi belum diketahui tingkat akurasi paling tinggi dalam mengklasifikasi kepuasan pelanggan. Untuk mengetahui tingkat akurasi tertinggi diantara metode tersebut, maka dilakukan penelitian perbandingan metode klasifikasi data mining untuk prediksi tingkat kepuasan pelanggan telkomsel prabayar. Penelitian dilakukan dengan penyebaran kuesioner dengan jumlah 500 responden. Dan variabel yang dinilai ada 4 yaitu harga, promosi, kualitas produk dan kualitas layanan. Hasil dari penelitian ini di dapatkan nilai akurasi algoritma C45 sebesar sebesar 96,50%, *Support Vector Machine* sebesar 89,66%, *Naïve Bayes* sebesar 89,88% dan metode Optimasi *Naïve Bayes* dengan pemilihan *fiture Particle Swarm Optimization* sebesar 95,85%. Jadi algoritma C.45 nilai akurasinya paling tinggi dibandingkan dengan metode lainnya. Sehingga menggunakan algoritma C.45 lebih baik dalam memprediksi tingkat kepuasan pelanggan telkomsel prabayar

Abstract

Currently, Indonesia has more than one active cellular operator. Technological developments that make all community activities carried out via mobile or online. Because the need for

Andi Taufik
Email: a.taufik30@gmail.com

communication is very high, operator companies are competing to win over the hearts of their customers by improving their services. Telkomsel is one of the largest operators in Indonesia. To provide the best service, Telkomsel must be consistent and improve service. There are several methods that can be used to analyze customer satisfaction classifications such as Naïve Bayes, Support Vector Machine, Decision Tree and Naïve Bayes and PSO methods. But it is not yet known the highest level of accuracy in classifying customer satisfaction. To determine the highest level of accuracy among these methods, a comparative study of data mining classification methods was conducted to predict the level of prepaid Telkomsel customer satisfaction. The research was conducted by distributing questionnaires with a total of 500 respondents. And there are 4 variables assessed, namely price, promotion, product quality and service quality. The results of this study obtained the accuracy value of the C45 algorithm of 96.50%, Support Vector Machine of 89.66%, Naïve Bayes of 89.88% and the Naïve Bayes Optimization method with the selection of Particle Swarm Optimization features of 95.85%. So the C.45 algorithm has the highest accuracy value compared to other methods. So using the C.45 algorithm is better at predicting the level of customer satisfaction of Telkomsel prepaid customers

1. Pendahuluan

Karena kebutuhan komunikasi sangat tinggi dan perkembangan penggunaan data akses internet makin dibutuhkan, terutama dimasa seperti sekarang ini dimasa pandemi maka perusahaan operator saling berlomba-lomba untuk mengambil hati para pelanggan dengan peningkatan layanannya. Salah satu operator seluler yang terbesar dan banyak digunakan di Indonesia adalah telkomsel. Untuk memberikan pelayanan terbaik kepada pelanggan telkomsel harus konsisten dan meningkatkan pelayanan.

Industri telekomunikasi membantu pengguna telepon seluler agar dapat terhubung antara satu sama lain dengan cepat dan mudah. Keberhasilan industri telekomunikasi seluler tegantung pada infrastruktur untuk memenuhi jangkauan layanan, karena setiap telepon genggam harus dilengkapi dengan beberapa fasilitas agar dapat berfungsi dengan cepat dengan jangkauan yang luas. Pada negara-negara berkembang, Industri provider atau operator telekomunikasi (mobile network provider) mengalami perkembangan yang sangat pesat. Industri telekomunikasi menjadi persaingan bisnis yang banyak diminati dan sangat

menguntungkan dalam lingkup lingkungan bisnis.(Armin et al., 2017)

Operator telepon seluler atau operator nirkabel merupakan perusahaan yang menyediakan layanan untuk pengguna telepon seluler. Operator telepon seluler membuat kartu SIM dengan ukuran kecil untuk pelanggan dengan cara harus dimasukkan ke setiap ponsel untuk mendapatkan akses layanan tersebut. Operator telepon seluler yang populer sejak pada awal peluncurannya pada tanggal 26 Mei 1995 di Jakarta hingga kini adalah Telkomsel.

Telkomsel menjadi salah satu operator telepon seluler terbesar yang di miliki Indonesia dengan jumlah minat pelanggan terbanyak. Telkomsel juga memiliki kualitas jaringan yang bagus, tersedia pilihan bonus telepon dan sms, pilihan paket ineternet dan tersedianya customer service pada setiap Grapari yang tersebar di seluruh Indonesia. (Septi et al., 2020)

Kepuasan konsumen sebagai pengguna jasa adalah hal yang paling utama bagi sebuah pengelola jasa layanan. (Desyanti, 2018)

Kepuasan konsumen merupakan tingkat dimana anggapan terhadap produk sesuai dengan harapan para konsumen. Harapan konsumen umumnya merupakan prakiraan atau keyakinan konsumen tentang apa yang akan diterimanya bila telah membeli atau mengkonsumsi suatu produk. Seorang konsumen mungkin mengalami berbagai tingkat kepuasan, yaitu bila produk tidak sesuai dengan harapannya setelah dikonsumsi, maka konsumen tersebut akan merasa tidak puas. Namun bila terjadi sebaliknya yaitu produk sesuai dengan harapannya, maka konsumen akan merasa puas sehingga suatu saat akan mengkonsumsi kembali produk tersebut. (Shiddiq et al., 2018)

Berikut ini beberapa penelitian terdahulu yang telah melakukan penelitian berkaitan dengan analisa tingkat kepuasan pelanggan, yaitu :

Tahun 2016 Teguh Budi Santoso (Santoso, 2016) melakukan penelitian tentang analisa dan penerapan metode c4.5 untuk prediksi loyalitas pelanggan. Berdasarkan Hasil klasifikasi menggunakan algoritma C4.5 menunjukkan bahwa diperoleh akurasi mencapai 97.5%, yang menunjukkan bahwa algoritma C4.5 cocok digunakan untuk mengukur tingkat loyalitas pelanggan data seluler.

Tahun 2018 Sutriawaty Takalapeta (Takalapeta, 2018) melakukan penelitian tentang Penerapan Data Mining Untuk Menganalisis Kepuasan Konsumen Menggunakan Nilai gain tertinggi adalah variabel *tangibles* dengan nilai 0,378118. Sedangkan Metode Algoritma C4.5, pada sebuah resto dan hasil penelitian indikator dari variabel *tangibles* yang mempunyai nilai entropy tertinggi adalah indikator fasilitas dengan nilai 0,764205

Ahmad Shiddiq, Ratih Kumalasari Niswatin, dan Intan Nur Farida melakukan penelitian pada tahun 2018 tentang kepuasan konsumen di restoran dapur solo cabang

Kediri. Penelitian dilakukan dengan menggunakan algoritma C4.5. hasil dari penelitian tersebut adalah mengklasifikasikan konsumen merasa tidak puas apabila pelayanan dirasa kurang, fasilitas dirasa kurang dan pelayanan ramah, product dirasa kurang, fasilitas dirasa baik dan pelayanan dirasa ramah. Dan mengklasifikasikan konsumen merasa puas apabila pelayanan dirasa lumayan, product dirasa lumayan, fasilitas dirasa baik dan pelayanan dirasa ramah. (Shiddiq et al., 2018)

T. Yudi Hadiwandura melakukan penelitian tentang Perbandingan Kinerja Model Klasifikasi *Decision Tree*, *Bayesian Classifier*, *Instance Base*, *Linear Function Base*, *Rule Base* pada 4 Dataset Berbeda. Hasil dari penelitian tersebut adalah kelas model *Linear Function* dan *Decision Tree* membutuhkan waktu training yang relatif lebih lama dibanding model lainnya, namun membutuhkan waktu *testing* yang relatif lebih cepat dibanding model lainnya. Secara umum kelas model *Decision Tree* lebih unggul dibanding kelas model lainnya. (Hadiwandura, 2019)

Nur Lutfiyana melakukan penelitian pada tahun 2018 tentang prediksi hasil layanan kemudahan Donasi zakat dan program dengan penerapan algoritma C4.5 berbasis PSO. Hasil penelitian untuk nilai akurasi yang didapat pada model algoritma C4.5 adalah 83,91% lebih baik jika dibandingkan dengan model algoritma C4.5 berbasis PSO adalah 83,49%.(Lutfiyana, 2018)

Julia Ferlin, Fitra Abdurrachman Bachtiar, dan Alfi Nur Rusydi melakukan penelitian pada tahun 2019 dengan judul Klasifikasi Customer Intent untuk Mengetahui Tingkat Kepuasan Pelanggan menggunakan Metode *Support Vector Machine* pada Restoran Bakso President. Penelitian ini menghasilkan nilai *Accuracy model* sebesar 80% dengan rata – rata nilai *Precision* sebesar 66%, *Recall* sebesar

43,7%, dan *F1-score* sebesar 47%.(Ferlin, 2019)

Yusran Timur Samuel, Kemala Dewi melakukan penelitian tahun 2019 dengan judul, Penggunaan Metode *Naïve Bayes* Dalam Mengukur Tingkat Kepuasan Pengguna Terhadap *Online System* Universitas Advent Indonesia. Hasil pengujian menunjukkan bahwa klasifikasi kepuasan pengguna menggunakan metode *naïve bayes* untuk pengujian pertama dengan *full data training* mendapatkan nilai akurasi sebesar 81.3%, pengujian kedua dengan 80% *data training* 20% *data testing* mendapatkan nilai akurasi 80%, pengujian ketiga dengan *cross-validation* mendapatkan nilai akurasi sebesar 78.7 dan pengujian keempat dengan 66% *data training* dan 33% *data testing* mendapatkan nilai akurasi 68.6%(Samuel & Dewi, 2019)

Tahun 2018 Agus Darmawan, Nunu Kustian dan Wanti Rahayu meakukan penelitian dengan judul Implementasi Data Mining Menggunakan Model SVM Untuk Prediksi Kepuasan Pengunjung Taman Tabebuya. Hasil dari penelitian melakukan evaluasi dan validasi dapat disimpulkan bahwa algoritma SVM memiliki akurasi dan performa secara rata-rata yaitu sebesar 86,00% dan nilai AUC sebesar 0.947. (Darmawan et al., 2018)

Dari penelitian sebelumnya dapat dilihat bahwa terdapat beberapa metode untuk menganalisa kepuasan pelanggan. Tetapi belum diketahui tingkat akurasi paling tinggi dalam mengklasifikasi kepuasan pelanggan.

Tujuan dari penelitian ini adalah untuk menganalisa algoritma mana yang mendapatkan nilai akurasi yang paling tinggi untuk klasifikasi kepuasan pelanggan telkomsel antara algoritma *Naïve Bayes*, *Support Vector Machine*, *Decission Tree* dan *Metode Naïve Bayes* dan PSO.

2. Metode Penelitian

Metode penelitian yang dilakukan adalah metode penelitian eksperimen dengan tahapan sebagai berikut :

1. Pengumpulan Data

Pengumpulan data menggunakan penyebaran kuesioner kepada para pelanggan telkomsel. Pada penelitian ini kuesioner dibagikan kepada 500 responden. Dalam kuesioner terdapat 4 variabel (Harga, Promosi, Kualitas Produk, dan kualitas layanan) dan 14 atribut. Data kuesioner tersebut digunakan sebagai data awal untuk mengetahui jumlah puas dan tidak puas pada data kepuasan pelanggan telkomsel yang secara rinci di jabarkan dalam Tabel 1.

Tabel 1. Variabel dan atribut

| Variabel | Atribut |
|------------------|---|
| Harga | Harga kartu perdana Telkomsel Prabayar terjangkau |
| | Tarif SMS lebih murah dibandingkan dengan operator lain |
| | Tarif telepon lebih murah dibandingkan dengan operator lain |
| Promosi | Harga paket internet yang ditawarkan lebih murah dibandingkan dengan operator lain |
| | Sering memberikan gratis atau bonus SMS dan telepon |
| | Memberikan gratis atau bonus Internet |
| Kualitas Produk | Selalu mengadakan undian seperti telkomsel point |
| | Selama Pemakaian kualitas SIM Card tahan lama |
| | Telkomsel mempunyai sinyal yang kuat serta jangkauan yang luas dan dapat mempermudah komunikasi |
| Kualitas Layanan | Koneksi internet yang diterima sangat cepat. |
| | Sinyal selalu stabil dan jarang mengalami gangguan. |
| | Tersedianya grapari Telkomsel yang mudah ditemukan sebagai pusat layanan pelanggan. |
| | Layanan Customer Service Telkomsel dapat membantu saya ketika terjadi keluhan terkait produk Telkomsel. |

Tersedia layanan aplikasi MY
 Telkomsel.

Tabel 1 merupakan variabel dan atribut yang digunakan dalam penelitian ini.

Pengukuran data responden dalam penelitian ini menggunakan Skala *Linkert*. Skala *Linkert* merupakan skala yang digunakan untuk mengukur variabel sikap, pendapat, dan persepsi seseorang atau sekelompok mengenai fenomena sosial yang disajikan pada Tabel 2 (Puspa & Sudibya, 2016).

Tabel 2. Skala *linkert*

| No. | Simbol | Skala jawaban | Nilai |
|-----|--------|---------------------|-------|
| 1. | STS | Sangat Tidak Setuju | 1 |
| 2. | TS | Tidak Setuju | 2 |
| 3. | CS | Cukup Setuju | 3 |
| 4. | S | Setuju | 4 |
| 5. | SS | Sangat Setuju | 5 |

Tabel 2 merupakan skala *linkert* yang terdiri dari 5 simbol yaitu STS, TS, CS, S, dan SS.

2. Metode yang Digunakan

Untuk mengetahui metode klasifikasi data mining yang paling akurat pada kepuasan pelanggan telkomsel. Metode yang digunakan yaitu, C4.5, *Support Vector Machine*, *Naïves Bayes* dan metode *Naïve Bayes* dengan pemilihan fitur *Particle Swarm Optimization*.

a. Algoritma C.45

Metode pohon keputusan digunakan untuk mengubah fakta yang sangat besar, mengeksplorasi data, dan menemukan hubungan tersembunyi antara sejumlah variabel input dengan variabel target.

Tahapan dalam membuat sebuah pohon keputusan dengan algoritma C4.5. (Kusrini & Luthfi, 2009) yaitu:

- 1) Mempersiapkan data *training*, dapat diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan dalam kelas-kelas tertentu.

- 2) Menentukan akar dari pohon dengan menghitung nilai *gain* yang tertinggi dari masing-masing atribut atau berdasarkan nilai *index entropy* terendah. Sebelumnya dihitung terlebih dahulu nilai *index entropy*, dengan rumus:

$$Entropy(S) = \sum_{i=1}^n (-pi) * (pi) \quad (1)$$

Keterangan :

S: himpunan kasus

n : jumlah partisi S

pi : proporsi dari S_i terhadap S

- 3) Hitung nilai *gain* dengan rumus:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$

Keterangan:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

|S_i| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

- 4) Ulangi langkah ke-2 hingga semua semua tupel terpartisi.
- 5) Proses partisi pohon keputusan akan berhenti saat:
 - a) Semua tupel dalam node N mendapat kelas yang sama.
 - b) Tidak ada atribut di dalam tupel yang dipartisi lagi.
 - c) Tidak ada tupel di dalam cabang yang kosong

b. *Support Vector Machine* (SVM)

Support Vector Machine merupakan *machine learning* yang termasuk dalam model *supervised learning* atau pembelajaran dengan pengawasan yang berhubungan dengan analisis data dan pengenalan pola Fungsi keputusan klasifikasi sign (f(x)):

$$f(x) = wx + b \quad (3)$$

Atau

$$fx = \sum_i^m = 1 a_i a y_i K(x, x_i) + b \quad (4)$$

Keterangan:

N : Banyaknya data

n : dimensi data atau banyaknya fitur

Ld : *Dualitas Lagrange Multiplier*

α_i : nilai bobot setiap titik data

C: nilai konstanta

M : jumlah *support vector*/titik data yang memiliki $\alpha_i > 0$

K(x,x_i) : fungsi kernel

c. *Naïves Bayes*

Dalam *Bayes* (terutama *Naïve Bayes*), maksud independensi yang kuat pada fitur adalah bahwa sebuah data tidak berkaitan dengan ada atau tidaknya fitur lain dalam data yang sama. Prediksi *Bayes* didasarkan pada teorema *Bayes* dengan formula umum sebagai berikut :

$$P(E) = \frac{P(HP(H))xP(H)}{P(E)} \quad (5)$$

Keterangan

P(H|E) : Probabilitas akhir bersyarat (*conditional probability*) suatu hipotesis H terjadi jika diberikan bukti (*evidence*) E terjadi.

P (E|H) : Probabilitas sebuah bukti E terjadi akan mempengaruhi hipotesis H.

P (H) : Probabilitas awal (priori) hipotesis H terjadi tanpa memandang bukti apapun.

P (E) : Probabilitas awal (priori) bukti E terjadi tanpa memandang hiotesis/bukti yang lain

d. *Particle Swarm Optimization*.

Particle Swarm Optimization dirumuskan pertama kali oleh Edward dan Kennedy pada tahun 1995. Proses pemikiran dibalik

algoritma ini terinspirasi dari perilaku sosial hewan. Seperti burung yang berkelompok atau sekelompok ikan.(Taufik, 2018)

3. Pengujian Metode

Pengujian pada model yang dilakukan dengan menggunakan RapidMiner untuk mengolah data. Model diuji untuk melihat hasil yang akan dimanfaatkan untuk mengambil keputusan hasil penelitian.

Tahapan dalam melakukan pengujian kepuasan pelanggan sebagai berikut:

- a) Menyiapkan dataset untuk eksperimen yang sudah diketahui classnya Mendesain arsitektur algoritma klasifikasi C4.5, *Support Vector Machine*, *Naïves Bayes* dan metode *Naïve Bayes* dengan pemilihan fitur *Particle Swarm Optimization*.
- b) Melakukan *training* dan *testing* terhadap algoritma C4.5, *Support Vector Machine*, *Naïves Bayes* dan metode *Naïve Bayes* dengan pemilihan fitur *Particle Swarm Optimization*, kemudian mencatat hasil *accuracy* dan AUC

4. Evaluasi dan Validasi Data

Validasi dilakukan menggunakan validation tertinggi dalam tiap algoritma. Sedangkan pengukuran akurasi diukur dengan confusion matrix dan kurva ROC (Receiver Operating Characteristics) untuk mengukur nilai AUC.

3. Hasil dan Pembahasan

Setelah perhitung data kuesioner, didapatkan jumlah nilai PUAS terdapat 366 data dan jumlah TIDAK PUAS berjumlah 145 data, dan total keseluruhan kasus terdapat 500.

A. Hasil pengujian menggunakan Algoritma C.45

Pengujian Algoritma C.45 menggunakan teknik pengujian dengan merubah nilai

validasi nya dari 1 sampai dengan 10, berikut tabel 3 ini pengujiannya untuk mendapatkan nilai akurasi yang paling tinggi maka peneliti melakukan

Tabel 3. Hasil pengujian algoritma C.45

| Validation | Accuracy | AUC |
|------------|---------------|--------------|
| 1 | 94,00% | 0.685 |
| 2 | 94,00% | 0.685 |
| 3 | 95,00% | 0.887 |
| 4 | 94,90% | 0.794 |
| 5 | 96,04% | 0.950 |
| 6 | 93,96% | 0.682 |
| 7 | 96,10% | 0.959 |
| 8 | 95,95% | 0.952 |
| 9 | 96,50% | 0.960 |
| 10 | 95,00% | 0.980 |

Berdasarkan hasil pengujian menggunakan Algoritma C45 pada Tabel 3 dengan merubah nilai validasi 1-10 menggunakan Rapid Miner untuk mendapatkan nilai akurasi tertinggi terdapat pada nilai validasi 9 sebesar 96,50 % dan nilai AUC nya 0.960

B. Hasil pengujian menggunakan *Support Vector Machine*

Pada Pengujian Menggunakan Algoritma *Support Vector Machine* (SVM) untuk mendapatkan hasil yang terbaik peneliti mencoba merubah beberapa parameter agar mendapatkan nilai akurasi yang tinggi, pengujian pertama yang dilakukan dengan Nilai C 0.0 , Nilai *Epsilon* 0.0 dan merubah nilai Validitas 1-10, berikut Tabel 4 pengujian yang telah dilakukan hanya menampilkan nilai akurasi yang tertinggi yaitu dengan nilai validasi dari 6 sampai 10

Tabel 4. Hasil pengujian SVM nilai C 0.0 dan epsilon 0.0

| Validasi | C | Epsilon | Accuracy | AUC |
|----------|------------|------------|---------------|--------------|
| 6 | 0.0 | 0.0 | 65.11% | 0,970 |
| 7 | 0.0 | 0.0 | 67.84% | 0.950 |
| 8 | 0.0 | 0.0 | 66.90% | 0,969 |

| | | | | |
|----|-----|-----|--------|-------|
| 9 | 0.0 | 0.0 | 60.05% | 0.975 |
| 10 | 0.0 | 0.0 | 66.00% | 0,973 |

Hasil Pengujian dengan menggunakan algoritma *Support Vector Machine* masih sangat rendah nilai akurasi tertinggi pada validasi 7 dengan nilai akurasi 67,84% dan nilai AUC 0.950

Pengujian kedua dilakukan untuk mencari nilai akurasi yang besar menggunakan Algoritma *Support Vector Machine* (SVM) dengan nilai parameter C 0.1, nilai Epsilon 0.0 dan merubah nilai Validasi 1 sampai 10 yang disajikan pada Tabel 5.

Tabel 5. Hasil pengujian SVM nilai C 0.1 dan epsilon 0.0

| Validasi | C | Epsilon | Accuracy | AUC |
|----------|------------|------------|---------------|--------------|
| 6 | 0.1 | 0.0 | 85.80% | 0,970 |
| 7 | 0.1 | 0.0 | 85.74% | 0.986 |
| 8 | 0.1 | 0.0 | 88.90% | 0,976 |
| 9 | 0.1 | 0.0 | 87.69% | 0.975 |
| 10 | 0.1 | 0.0 | 88.40% | 0,960 |

Pada Tabel 5 hasil pengujian kedua menggunakan *Support Vector Machine* yang telah dilakukan mengalami peningkatan pada nilai akurasi dengan nilai validasi 8, nilai parameter C 0.1, dan nilai Epsilon 0.0 sebesar 88,90%

Pengujian Ketiga Menggunakan Algoritma *Support Vector Machine* (SVM) dilakukan untuk mencari nilai akurasi lebih baik lagi dengan parameter Nilai C 1.0 , nilai Epsilon 0.0 dan merubah nilai validasi 1 sampai dengan 10

Tabel 6. Hasil pengujian SVM nilai C 1.0 dan epsilon 0.0

| Validasi | C | Epsilon | Accuracy | AUC |
|----------|------------|------------|---------------|--------------|
| 6 | 1.0 | 0.0 | 86.80% | 0,980 |
| 7 | 1.0 | 0.0 | 87.52% | 0.977 |
| 8 | 1.0 | 0.0 | 89.66% | 0,968 |
| 9 | 1.0 | 0.0 | 86.90% | 0.981 |
| 10 | 1.0 | 0.0 | 89.40% | 0,945 |

Pada tabel 6 pengujian ketiga menggunakan algoritma *Support Vector Machine* mengalami peningkatan pada nilai akurasi dengan nilai Validasi 8 Nilai C 1.0 dan nilai Epsilision 0.0 sebesar 89.66%

C. Hasil pengujian menggunakan *Naïves Bayes*

Untuk mendapatkan nilai akurasi yang baik pengujian algoritma *Naïve Bayes* menggunakan teknik *Fold Cross Validation* atau yang disebut dengan merubah nilai validasinya 1 sampai dengan 10 berikut tabel hasil pengujian nya.

Tabel 7. Hasil pengujian *Naïves Bayes*

| Validation | Accuracy | AUC |
|------------|---------------|--------------|
| 1 | 85,00% | 0.682 |
| 2 | 85,00% | 0.682 |
| 3 | 80,70% | 0.500 |
| 4 | 86,00% | 0.500 |
| 5 | 84,00% | 0.550 |
| 6 | 89,88% | 0.500 |
| 7 | 87,07% | 0.550 |
| 8 | 86,90% | 0.500 |
| 9 | 83,89% | 0.500 |
| 10 | 84,96% | 0.500 |

Pada tabel 7 hasil pengujian algoritma *Naïve Bayes* Menggunakan *Rapidminer*, nilai akurasi tertinggi pada nilai validasi 6 sebesar 89,88% dan nilai AUC 0.500

D. Hasil pengujian menggunakan *Naïve Bayes* dengan pemilihan fitur *Particle Swarm Optimization*

Dalam Pengujian metode *Naïve Bayes* dengan pemilihan *future particle Swarm Optimization* peneliti melakukan perubahan nilai parameter untuk mendapatkan nilai akurasi tertinggi

Pengujian pertama yang telah dilakukan memberikan nilai *Population Size* 10, nilai *maximum of Generation* 80, nilai *ineteria Weight* 1.0 bernilai tetap dan merubah nilai

Validasi dari 1 sampai dengan 10, berikut tabel 8 pengujian yang telah dilakukan yang ditampilkan dengan nilai akurasi tertinggi yaitu pada nilai validasi 6-10

Tabel 8. Hasil pengujian *Naïves Bayes* pemilihan fitur PSO

| Validasi | Populasi Size (P) | Maximum Number Of Generation | Ineteria Weight (w) | Accuracy |
|----------|-------------------|------------------------------|---------------------|---------------|
| 6 | 10 | 80 | 1.0 | 90.94% |
| 7 | 10 | 80 | 1.0 | 95.77% |
| 8 | 10 | 80 | 1.0 | 93.95% |
| 9 | 10 | 80 | 1.0 | 94.87% |
| 10 | 10 | 80 | 1.0 | 95.30% |

Hasil pengujian menggunakan metode *naïve bayes* dan PSO pada Tabel 8 dengan merubah nilai validasi 1-10 menghasilkan nilai akurasi tertinggi pada validasi 7 sebesar 95,77%

Pengujian kedua metode *Naïve Bayes* dengan pemilihan *future particle Swarm Optimization* dilakukan dengan merubah nilai *Population Size* 1-10, sedangkan nilai *Validasi 7* dan nilai *Maximum Number Of Generation* 80 bernilai tetap untuk mendapatkan nilai akurasi lebih baik lagi yang disajikan pada Tabel 9 dibawah ini.

Tabel 9. Hasil pengujian *Naïves Bayes* pemilihan fitur PSO

| Validasi | Populasi Size (P) | Maximum Number Of Generation | Ineteria Weight (w) | Accuracy |
|----------|-------------------|------------------------------|---------------------|---------------|
| 7 | 6 | 80 | 1.0 | 93.35% |
| 7 | 7 | 80 | 1.0 | 95.85% |
| 7 | 8 | 80 | 1.0 | 94.75% |
| 7 | 9 | 80 | 1.0 | 93.80% |
| 7 | 10 | 80 | 1.0 | 94.45% |

Pada Tabel 9 hasil pengujian kedua metode Naïve Bayes dengan pemilihan *fiture particle Swarm Optimization* dengan merubah nilai *population size* 1-10 mengalami peningkatan pada nilai akurasi terdapat pada *population size* 7 sebesar 95,85%

Hasil Perbandingan Metode pada tingkat kepuasan pelanggan Telkomsel Prabayar adalah sebagai berikut pada Tabel 10.

Tabel 10. Hasil perbandingan metode

| Metode | Accuracy | AUC |
|-------------|----------|-------|
| C45 | 96,50% | 0.960 |
| SVM | 89.66% | 0,968 |
| Naïve Bayes | 89,88% | 0.500 |
| NB Dan PSO | 95.85% | 0.500 |

Pada Tabel 10 adalah hasil perbandingan metode menggunakan aplikasi rapid miner dengan nilai *Confusion Matrix* maupun *ROC Curve*, nilai akurasi yang bisa didapatkan menggunakan algoritma C.45 sebesar 96,50% dan AUC 0.960, sedangkan untuk nilai akurasi menggunakan Algoritma *Support Vector Machine* sebesar 89,66% dan AUC 0,968, Akurasi menggunakan Algoritma *Naïve Bayes* sebesar 89,88% dan auc 0.500 dan nilai akurasi dengan metode Optimasi *Naïve Bayes* dengan pemilihan *fiture Particle Swarm Optimization* sebesar 95,85% dan AUC 0.500

Pada Hasil pengujian perbandingan metode yang telah dilakukan nilai akurasi dan auc menggunakan algoritma C.45 lebih baik dibandingkan dengan 3 metodenya

4. Simpulan

Simpulan dari penelitian ini adalah hasil perhitungan dengan bantuan software RapidMiner studio 9.7 dengan perbandingan empat metode yang berbeda yaitu menggunakan algoritma C.45, *Support Vector Machine*, *Naïve Bayes* dan *Optimasi Naïve Bayes* dengan *Particle Swarm Optimization* untuk mengetahui metode mana

yang lebih baik dalam memprediksi tingkat kepuasan pelanggan telkomsel prabayar. Nilai akurasi algoritma C45 sebesar sebesar 96,50%, sedangkan untuk nilai akurasi menggunakan Algoritma *Support Vector Machine* sebesar 89,66%, Akurasi menggunakan Algoritma *Naïve Bayes* sebesar 89,88% dan nilai akurasi dengan metode Optimasi *Naïve Bayes* dengan pemilihan *fiture Particle Swarm Optimization* sebesar 95,85%

Berdasarkan hasil pengolahan perbandingan metode C45, *Support Vector Machine*, *Naïve Bayes* dan *Optimasi Naïve Bayes* dengan pemilihan *fitur Particle Swarm Optimization* dapat disimpulkan bahwa menggunakan algoritma C.45 Lebih baik dalam memprediksi tingkat kepuasan pelanggan telkomsel prabayar dan Hasil penelitian ini dapat digunakan dalam pengambilan keputusan peningkatan produk dan layanan operator seluler untuk kedepannya.

5. Referensi

- Armin, D., Mulyadi, & Fauziah. (2017). Evaluasi Kualitas Layanan terhadap Operator Telekomunikasi: Tinjauan Perspektif Pelanggan. *Jurnal Optimasi Sistem Industri*, 16(2), 106.
- Darmawan, A., Kustian, N., & Rahayu, W. (2018). Implementasi Data Mining Menggunakan Model SVM untuk Prediksi Kepuasan Pengunjung Taman Tabebuaya. *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 2(3), 299. <https://doi.org/10.30998/string.v2i3.2439>
- Desyanti. (2018). Penerapan Data Mining Algoritma C4.5 untuk Mengetahui Tingkat Kepuasan Konsumen di Hotel Grand Zuri Dumai. *SATIN- Sains dan Teknologi*, 4(2).
- Ferlin, J. (2019). *Klasifikasi Customer Intent Melalui Text Mining Untuk Mengetahui Tingkat Kepuasan Pelanggan Berdasarkan Komentar*

- Pelanggan Menggunakan Metode Support Vector Machine Pada Restoran Bakso President.* 3(10), 9867–9875.
- Hadiwandura, T. Y. (2019). SATIN – Sains dan Teknologi Informasi Perbandingan Kinerja Model Klasifikasi Decision Tree , Bayesian Classifier , Instance Base , Linear Function Base , Rule Base pada 4 Dataset Berbeda. *SATIN - Sains dan Teknologi Informasi*, 5(1), 70–78.
- Kusrini, & Luthfi, E. taufiq. (2009). *Algoritma Data Mining*. Cv. Andi Offset.
- Lutfiyana, N. (2018). Penerapan Algoritma C4.5 Berbasis Particle Swarm Optmization Untuk Prediksi Hasil Layanan Kemudaha Donasi Zakat Dan Program. *Jurnal Pilar Nusa Mandiri*, 14(1), 103–110.
- Puspa, C., & Sudibya, I. G. A. (2016). Pengaruh Kepemimpinan Transformasional Terhadap Psychological Empowerment Pada PT . PLN (PERSERO) Distribusi Bali Fakultas Ekonomi dan Bisnis Universitas Udayana (Unud), Bali , Indonesia Hasil observasi langsung di perusahaan , permasalahan yang terja. 5(8), 5143–5171.
- Samuel, Y. T., & Dewi, K. (2019). Penggunaan Metode NAÏVE BAYES Dalam Mengukur Tingkat Kepuasan Pengguna Terhadap Online System Universitas Advent Indonesia. *TeIKa*, 9(02), 147–153. <https://doi.org/10.36342/teika.v9i02.2162>
- Santoso, T. B. (2016). Analisa Dan Penerapan Metode C4.5 Untuk Prediksi Loyalitas Pelanggan. *Jurnal Ilmiah Fakultas Teknik LIMIT'S*, 10(1).
- Septi, N., Widodo, J., & Zulianto, M. (2020). Pengaruh Kualitas Produk Terhadap Kepuasan Konsumen Kartu Prabayar simPATI Telkomsel (Studi Kasus Pada Mahasiswa Pendidikan Ekonomi Fakultas Keguruan dan Ilmu Pendidikan Universitas Jember Angkatan 2015-2017). *JPE- Jurnal Pendidikan Ekonomi*, 14(1), 166–170.
- Shiddiq, A., Niswatin, R. K., & Farida, I. N. (2018). Analisa Kepuasan Konsumen Menggunakan Klasifikasi Decision Tree Di Restoran Dapur Solo (Cabang Kediri). *Generation Journal*, 2(1), 9–18.
- Takalapeta, S. (2018). Penerapan Data Mining Untuk Menganalisis Kepuasan Konsumen Menggunakan Metode Algoritma C4.5. *JIMP - Jurnal Informatika Merdeka Pasuruan*, 3(3), 34–38.
- Taufik, A. (2018). Komparasi Algoritma Klasifikasi Text Mining Untuk Klasifikasi Review Hotel. *jurnal teknik komputer AMIK BSI*, 14(2), 112–118. <https://doi.org/10.31294/jtk.v4i2.3461>