

Pemanfaatan Metode SMOTE dan PSO Untuk Mengoptimalkan Tingkat Akurasi Klasifikasi Kepuasan Pelanggan

Arif Muqtadir¹, Dini Silvi Purnia²

¹Sistem Informasi, Universitas Nusa Mandiri, Jakarta Barat, Indonesia

²Sistem Informasi Kampus Kota Tasikmalaya, Universitas Bina Sarana Informatika, Tasikmalaya, Indonesia
e-mail: muqtadir971@gmail.com¹, dini.dlv@nusamandiri.ac.id²

ABSTRAK

Penyedia layanan telekomunikasi Indihome Witel Jakarta Selatan mempunyai keunggulan dan kekurangan yang dapat memberikan kepuasan dan ketidakpuasan yang membuat pelanggan berpindah ke penyedia layanan internet lain. Pengetahuan tentang klasifikasi konsumen yang puas dan tidak puas sangat dibutuhkan dalam mengambil keputusan. Tujuan dari penelitian ini adalah untuk menganalisa algoritma klasifikasi data mining mana yang menghasilkan nilai akurasi yang paling tinggi untuk klasifikasi kepuasan pelanggan Indihome Jakarta Selatan antara algoritma Naïve Bayes, Support Vector Machine, Decision Tree. Dalam mengukur atau memprediksi tingkat kepuasan pelanggan dapat menggunakan algoritma klasifikasi data mining yaitu Naïve Bayes, Support Vector Machine, Decision Tree dengan penambahan fitur Metode Synthetic Minority Over-sampling Technique (SMOTE) dan teknik Particle Swarm Optimization (PSO). SMOTE berguna untuk menyeimbangkan data dan PSO untuk mengoptimalkan hasil nilai pada algoritma klasifikasi. Penelitian ini menghasilkan penggunaan Algoritma SVM + SMOTE + PSO menghasilkan Accuracy 98.57%, precision tertinggi 98.51%, recall tertinggi 98.51% dan AUC tertinggi 0.995 yang termasuk dalam excellent classification. Penggunaan SMOTE dan PSO pada algoritma klasifikasi terbukti berpengaruh meningkatkan accuracy algoritma SVM namun tidak secara signifikan.

Katakunci: data mining, kepuasan pelanggan, layanan indihome

ABSTRACTS

The South Jakarta Indihome Witel telecommunication service provider has advantages and disadvantages that can provide satisfaction and dissatisfaction that make customers switch to other internet service providers. Knowledge of the classification of satisfied and dissatisfied consumers is needed in making decisions. The purpose of this study is to analyze which data mining classification algorithm produces the highest accuracy value for the classification of South Jakarta Indihome customer satisfaction between the Naïve Bayes algorithm, Support Vector Machine, and Decision Tree. In measuring or predicting the level of customer satisfaction, data mining classification algorithms can be used, namely Naïve Bayes, Support Vector Machine, Decision Tree with the addition of the Synthetic Minority Over-sampling Technique (SMOTE) method if the data is unbalanced and the Particle Swarm Optimization (PSO) technique. SMOTE is useful for balancing data and PSO to optimize the value results in the classification algorithm. This research resulted in the use of the SVM + SMOTE + PSO algorithm resulting in an accuracy of 98.57%, the highest precision of 98.51%, the highest recall of 98.51% and the highest AUC of 0.995 which is included in the excellent classification. The use of SMOTE and PSO in the classification algorithm has proven to have an effect on increasing the accuracy of the SVM algorithm but not significantly.

Keywords: customer satisfaction, data mining, indihome service



1. PENDAHULUAN

Pada era globalisasi seperti saat ini kebutuhan fundamental bagi masyarakat adalah kebutuhan teknologi informasi. Kesadaran masyarakat akan teknologi sudah semakin tinggi. Di negara kita, Indonesia presentase pengguna internet mencapai 56% atau sekitar 150 juta jiwa dan presentase pengguna mobile sebesar 53% atau sekitar 142,8 juta jiwa (Kominfo, 2019). Adanya teknologi tidak jauh dengan penyedia layanan telekomunikasi, salah satunya adalah Indihome. Di sisi pelanggan, IndiHome terus menanamkan momentum positif, di tahun 2019, pelanggan pelanggan indihom naik sebesar 37,2% yang berarti pelanggan bertambah dari 1,9 juta pelanggan menjadi 7 juta pelanggan. Hal tersebut memperlihatkan begitu besarnya permintaan layanan internet dengan kualitas tinggi dan fasilitas yang menarik pada masyarakat Indonesia (Telkom, 2018). Meningkatnya intensitas persaingan dan jumlah kompetitor sehingga menuntut perusahaan untuk selalu memperhatikan kebutuhan dan keinginan konsumen. Indihome bukanlah satu-satunya penyedia layanan internet, ada beberapa penyedia layanan internet lain yang menyediakan produk serupa. Hal tersebut menyebabkan persaingan untuk memperoleh pelanggan dan sebaliknya menimbulkan kebingungan konsumen dalam memilih jasa layanan internet (Telkom, 2018). Besaran data tersebut mencerminkan keseluruhan data pengguna selama 2019, sedangkan pada tahun 2015 di Indihome wilayah Jakarta Selatan mempunyai 34.856 pelanggan dan mengalami penurunan pada tahun 2016 sehingga jumlah hanya mencapai 20.708. Pada saat itu sejumlah 14.148 dengan presentase penurunan sebesar 40,5% hal ini dikarenakan pelanggan berhenti berlangganan Indihome (Putra & Widiyanesti, 2018). Salah satu penurunan presentase data pelanggan berhubungan kualitas Indihome yang menyebabkan ketidakpuasan pelanggan. Data keluhan pelanggan terhadap layanan indihome beragam, pada Maret sampai dengan Desember 2016 terjadi peningkatan sebesar 1058 data keluhan (Putra & Widiyanesti, 2018). Meningkatnya keluhan pelanggan harus diatasi supaya tidak banyak pelanggan indihome yang beralih dengan layanan provider lain.

Berdasarkan penelitian terhadap kualitas pelayanan dan nilai pelanggan berbasis digital pada media sosial Facebook terhadap kepuasan

pelanggan pada produk IndiHome di Kota Bandung, meningkatnya jumlah keluhan yang ada pada media sosial maupun customer service mencerminkan masih adanya masalah yang dialami oleh pelanggan Indihome. Dari hasil survei menunjukkan bahwa pengalaman pelanggan terhadap layanan Indihome pada tahun 2018 meningkat dari 18% menjadi 25% (Syahputra & Sampurno, 2019). Kepuasan konsumen sebagai pengguna jasa merupakan hal yang vital terhadap sebuah pengelola jasa layanan (Desyanti, 2019). Setiap penyedia layanan telekomunikasi mempunyai keunggulan yang dapat memberikan kepuasan dan juga terdapat kekurangan yang membuat pelanggan merasa tidak puas dan akhirnya berpindah ke penyedia layanan internet lain. Keunggulan dan kekurangan produk layanan internet harus diperhatikan untuk menghindari penurunan layanan (Rofiqoh et al., 2017).

Penelitian yang berupa analisis penerapan algoritma C4.5 untuk mengukur tingkat Kepuasan pelanggan Indihome Pada Kota Pematangsiantar dilakukan untuk mengetahui tingkat kepuasan pelanggan dengan menggunakan metode Algoritma C.45. Data yang digunakan yaitu hasil dari kuesioner pelanggan Indihome di Kota Pematangsiantar. Penelitian ini menghasilkan nilai akurasi sebesar 94,00% label puas dan 50% label tidak puas (Bastian et al., 2021). Penelitian terhadap perbandingan metode klasifikasi data mining untuk memprediksi tingkat kepuasan pelanggan telkomsel prabayar fattya dilakukan untuk mengetahui tingkat akurasi tertinggi diantara metode tersebut, maka dilakukan penelitian perbandingan metode klasifikasi data mining untuk prediksi tingkat kepuasan pelanggan telkomsel prabayar. Penelitian ini menghasilkan nilai akurasi pada algoritma C45 sebesar 96,50%, SVM sebesar 89,66%, dan NB sebesar 89,88% dan metode NB + PSO sebesar 95,85%. Jadi algoritma C.45 nilai akurasinya paling tinggi dibandingkan algoritma lainnya (Ariani & Taufik, 2020). Perbandingan algoritma svm dengan naïve bayes untuk mengukur kepuasan pelanggan indihome juga telah dilakukan (Rohanah et al., 2021) yang menghasilkan algoritma svm lebih baik dari algoritma naïve bayes. Penelitian lain menambahkan fitur Backward Elimination pada klasifikasi kepuasan pelanggan menggunakan algoritme k-NN dan Naïve Bayes dengan peningkatan sebesar 33.74% pada algoritma K-NN, namun dengan algoritma naïve bayes mendapatkan nilai

akurasi 99.04% tanpa optimasi fitur Backward Elimination (Yunitasari et al., 2021). Sedangkan (Somantri & Apriliani, 2018) menerapkan Feature Selection dengan information Gain (IG) (SVM-IG) memiliki tingkat akurasi lebih baik.

Berdasarkan pembahasan yang telah dipaparkan, dapat dilihat permasalahan untuk mengetahui tingkat kepuasan pelanggan itu penting. Dalam mengukur atau memprediksi tingkat kepuasan pelanggan dapat menggunakan algoritma klasifikasi data mininig yaitu *Naive Bayes*, *Support Vector Machine*, *Decision Tree* dengan penambahan fitur Metode *Synthetic Minority Over-sampling Technique* (SMOTE) apabila data tidak seimbang dan teknik *Particle Swarm Optimization* (PSO). Oleh karena itu, penelitian ini perlu dilakukan untuk memprediksi tingkat kepuasan pelanggan terhadap layanan Indihome Jakarta Selatan menggunakan metode klasifikasi data mining.

2. METODE PENELITIAN

Penelitian ialah suatu proses terhadap investigasi sistematis yang mempelajari dengan bermacam bahan serta sumber selaku membangun dalam suatu kenyataan serta menarik kesimpulan tersebut. Tujuan tersebut yakni selaku mendapatkan ataupun menciptakan suatu informasi selaku iktidak serta mempunyai tujuan tersebut. Riset ataupun studi biasanya dideskripsikan selaku sesuatu proses investigasi yang dicoba dengan aktif, tekun serta sistematis, yang bertujuan buat menciptakan data pada suatu topik tertentu. Tata cara riset yang digunakan dalam riset ini merupakan dengan memakai tata cara eksperimen. Tata cara yang dicoba buat bisa memperoleh data yang mendalam tentang memprediksi tingkatan kepuasan pelanggan bisa memakai algoritma klasifikasi informasi mininig ialah *Naive Bayes*, *Support Vector Machine*, *Decision Tree*. Gambar 1 menjelaskan model yang diusulkan dalam penelitian ini dengan memakai tata cara CRISP- DM.

Dalam mendesain metode penelitian eksperimen menggunakan metode penelitian standar yang digunakan pada data mining yaitu CRISP-DM yaitu terdiri dari enam tahap, namun penelitian ini hanya melakukan lima tahap dengan langkah-langkahnya adalah *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, dan *Evaluation*.

Pada tahap *Business Understanding* bertujuan untuk memahami bidang masalah,

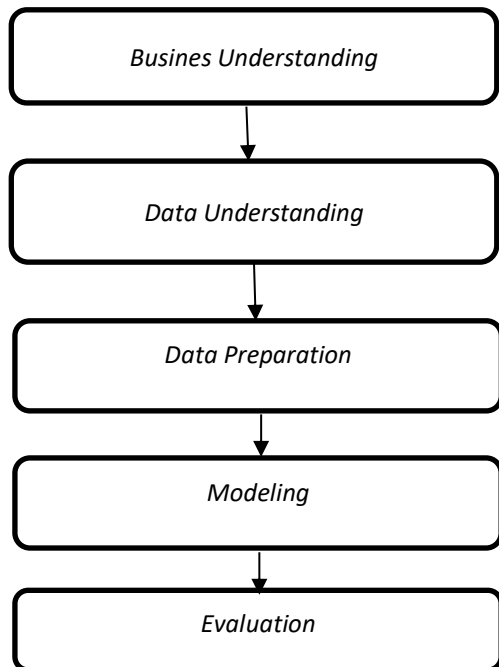
menghasilkan solusi yang tepat, dan mengungkapkan faktor penting yang berpengaruh pada hasil penelitian. Dalam memprediksi tingkat kepuasan pelanggan terhadap layanan Indihome Jakarta Selatan penulis melakukan pemahaman bisnis berjalan pada PT Telkom Witel Jakarta Selatan dengan menganalisa kebutuhan apa saja yang diperlukan

Pada tahap *Data Understanding* bertujuan untuk mengumpulkan, mengidentifikasi, dan memahami data yang dimiliki. Data tersebut juga harus dapat diverifikasi kebenarannya. Data yang akan digunakan dalam penelitian ini adalah data yang dihasilkan dari kuesioner yang telah diisi oleh pelanggan Indihome Witel Jakarta Selatan.

Tahap *Data Preparation* yaitu dengan mempersiapkan data yang dihasilkan dari kuesioner yang telah diisi oleh pelanggan Indihome Witel Jakarta Selatan. Dalam proses ini melakukan cleaning data yang tidak dibutuhkan sehingga menjadi data yang clean (bersih), ada beberapa atribut data yang tidak digunakan yaitu waktu pengisian data (timestamp), nama, jenis kelamin dan pekerjaan. Adapun atribut yang akan digunakan yaitu kualitas produk, harga, kualitas jasa, pelayanan petugas, fasilitas dan tanggapan.

Dalam tahap *Modelling* ini akan dilakukan teknik pengklasifikasian data yang paling akurat. Untuk membandingkan atau mengkomparasi, pada penelitian ini akan menggunakan algoritma *Naive Bayes*, *Support Vector Machine*, *Decision Tree* dengan penambahan fitur Metode *Synthetic Minority Over-sampling Technique* (SMOTE) apabila data tidak seimbang dan teknik *Particle Swarm Optimization* (PSO).

Model yang terbentuk akan diuji menggunakan *Confusion Matrix* yang akan mengetahui tingkat akurasi. *Confusion Matrix* akan menggambarkan hasil akurasi. Akurasi akan dihitung dari seluruh hasil prediksi yang benar (baik prediksi positif dan negatif) dibandingkan dengan seluruh data testing. Semakin tinggi nilai akurasi, semakin baik pula model yang dihasilkan. Pengujian juga diukur dengan menggunakan *ROC Curve*. *ROC Curve* akan menggambarkan kelas positif dalam bentuk kurva. Pengujian dilakukan dengan menghitung nilai *Area Under Curve* (AUC), semakin tinggi nilai AUC dan *ROC Curve*, maka semakin baik pula model klasifikasi yang terbentuk.



Gambar 1. Tahapan Penelitian Menggunakan CRISPM-DM

Pengumpulan data juga dilakukan untuk mengambil data terkait faktor-faktor yang mempengaruhi kepuasan pelayanan Indihome Jakarta Selatan. Sebelum melakukan pengumpulan data yang lebih mendalam, peneliti melakukan penelusuran terlebih dahulu untuk memperoleh gambaran umum tentang situasi teliti sebagai sumber data kunci (*key person*) yaitu semua pihak yang terlibat dalam memberikan pelayanan pada Indihome yang sedang bekerja pada saat itu. Melalui penelusuran ini, peneliti dapat memperoleh gambaran umum tentang pelayanan Indihome khususnya yang terkait dengan kepuasan pelanggan terhadap layanan Indihome. Adapun metode untuk mengumpulkan data dilakukan melalui observasi, wawancara dan angket (kuesioner).

Observasi dilakukan dengan melakukan pengamatan terhadap pelayanan indohome di Witel Jakarta Selatan. Adapun pelayanan yang diamati adalah customer service/petugas yang memberikan pelayanan, konsumen/pelanggan Indihome, fasilitas dan hubungan timbal baik antar bebrapa pihak yang terkait. Sedangkan wawancara dilakukan terhadap pihak terkait yaitu terhadap customer service/petugas, pelanggan, dan staff PT.Telkom Witel Jakarta Selatan untuk memperoleh data yang berhubungan dengan materi pembahasan tingkat kepuasan pelanggan terhadap layanan

Indihome. Pengumpulan data dengan kuesioner dilakukan dengan memberikan atau menyebarkan daftar pertanyaan kepada responden, sehingga pelanggan Indihome memberikan respon atas daftar pertanyaan tersebut. Dalam hal ini peneliti menyebarkan kuesioner kepada beberapa responden.

Instrument dalam riset ini memakai angket ataupun kuesioner, sebagai bahan dalam pengumpulan informasi yang digunakan untuk pengolahan informasi. Jumlah persoalan yang digunakan untuk mengenali mutu kepuasan pelanggan terhadap pelayanan Indihome pada Witel Jakarta Selatan ialah 6 butir, butir soal persoalan digunakan untuk mengenali seberapa tingkatan mutu kinerja serta harapan terhadap kepuasan pelanggan Indihome. Instrumen riset memegang peranan berarti sebab berperan mengatakan kenyataan jadi informasi, sehingga bila instrumen riset memiliki mutu yang mencukupi hingga informasi yang diperoleh hendak cocok dengan kenyataan ataupun kondisi sebetulnya dilapangan. Ada pula instrumen riset yang digunakan dalam riset ini merupakan kuesioner. Wujud kuesioner yang digunakan merupakan kuesioner tertutup. Kuesioner tertutup ialah responden cuma memilah jawaban yang sudah ada yang ialah sesuai buat memastikan jawaban terhadap mutu pelayanan Indihome. Tabel 1 menunjukkan indikator dari atribut penelitian yang digunakan sebagai pertanyaan kuesioner.

Pertanyaan pertanyaan pada angket tertutup dibuat dengan skala *Likert* dengan skor 1 sampai 4. Skor ini digunakan peneliti karena lebih sederhana dan memiliki nilai tengah yang digunakan untuk menjelaskan keragu-raguan atau netral dalam memilih jawaban. Skala *Likert* ini lazim digunakan di Indonesia untuk mengukur sikap, pendapat, dan persepsi orang. Penilaian skor pada Pertanyaan dan jawaban tingkat kepuasan pelanggan dapat dilihat pada tabel 2.

Populasi yang digunakan untuk memenuhi penelitian ini yaitu pelanggan Indihome Witel Jakarta Selatan. Pelanggan tersebut dijadikan sebagai responden dalam penelitian ini. *Survey* yang telah dilakukan demi pengumpulan data penelitian dengan cara menyebarkan kuesioner kepada para responden. Berdasarkan populasi yang sekian banyak jumlah pelanggan Indihome Witel Jakarta Selatan, adapun jumlah responden yang mengisi kuesione adalah 120 pelanggan.

Tabel 1. Indikator Dari Atribut Penelitian Yang Digunakan

No.	Atribut
1	Nama
2	Jenis Kelamin
3	Pekerjaan
4	Kualitas Produk Indohome : IndiHome mempunyai kualitas produk yang baik sehingga mudah digunakan, kualitas jaringan IndiHome tidak pernah bermasalah, mempunyai kemudahan dalam mengakses internet dan jarang mengalami hambatan dalam penggunaannya
5	Harga: Indihome menawarkan biaya yang terjangkau dan wajar
6	Kualitas Jasa: Jasa yang ditawarkan beragam yaitu menyediakan paket yang sesuai dengan Kebutuhan Pengguna, layanan telepon yang spesial, dan bisa membayar di mana saja
7	Pelayanan Petugas: Petugas memberikan pelayanan baik terhadap pelanggan, petugas melayani dengan ramah, dan memberikan jaminan kecepatan dan ketepatan waktu pelayanan dengan maksimal.
8	Fasilitas: Fasilitas yang disediakan sudah memadai.
9	Tanggapan: Silahkan pilih tanggapan secara umum terhadap kualitas produk, harga, kualitas jasa, pelayanan petugas, fasilitas.

Tabel 2. Skala *Likert* Tingkat Kepuasan Pelanggan

Pilihan Jawaban	Skor
Sangat Tidak Puas	1
Tidak Puas	2
Puas	3
Sangat Puas	4

Metode yang digunakan untuk memprediksi tingkat kepuasan pelanggan terhadap layanan Indihome Jakarta Selatan dengan metode klasifikasi data mining yaitu *Naive Bayes*, *Decision Tree*, dan *Support Vector Machine*. Pemakaian teorema Bayes pada algoritma *Naive Bayes* ialah dengan

mengombinasikan prior probability serta probabilitas bersyarat dalam suatu rumus yang dapat digunakan buat menghitung probabilitas masing- masing klasifikasi yang bisa jadi Model *independence* ini menciptakan pemecahan yang terbaik. Efektifitas tata cara *Naive Bayes* serta perbandingan empiris lebih jauh, dengan hasil yang sama. Klasifikasi Bayes didasarkan pada teorema Bayes, diambil dari nama seseorang pakar matematika yang pula menteri Prebysterian Inggris, Thomas Bayes (1702-1761),[14] yaitu :

$$\frac{p(x|y) = p(y|x)p(x)}{p(y)}$$

Keterangan:

y : data dengan kelas yang belum diketahui

x : hipotesis data y merupakan suatu kelas spesifik

P(x|y) : probabilitas hipotesis x berdasarkan kondisi y

P(x) : probabilitas hipotesis x

P(y|x) : probabilitas y berdasarkan kondisi pada hipotesis x

P(y) : probabilitas dari y

Decision Tree (DT) adalah representasi sederhana untuk mengklasifikasikan suatu contoh. Proses dalam *Decision Tree* yaitu mengubah bentuk data (tabel) menjadi model pohon (tree) kemudian mengubah model pohon tersebut menjadi aturan (*rule*). Algoritma yang digunakan adalah algoritma ID3 [15]. Algoritma ID3 menggunakan konsep dari Entropy dan Information Gain. Untuk nilai Entropy dapat ditemukan dengan menggunakan rumus:

$$Entropy(S) = \sum_{j=1}^k -p_j \log_2 p_j$$

Keterangan:

S adalah ruang (data) sample yang digunakan untuk training

K adalah banyaknya partisi pada S

p_j adalah probabilitas yang didapat dari Sum(Ya) dibagi dengan total *sample*

$$Gain(A) = Entropy(S) - \sum_{i=1}^k \frac{|S_i|}{|S|} Entropy(S_i)$$

Keterangan:

S adalah ruang (data) sample yang digunakan untuk training

A adalah atribut

$|S_i|$ adalah jumlah sample untuk nilai V

$|S|$ adalah jumlah seluruh sample data

Entropy (S_i) adalah entropy untuk sample – sample yang memiliki nilai i

Support Vector Machine (SVM) adalah suatu teknik yang relatif baru untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. SVM masuk kelas supervised learning, dimana dalam implementasinya perlu adanya tahap pelatihan menggunakan sequential training SVM dan disusun tahap pengujian. Konsep klasifikasi dengan SVM adalah mencari *hyperplane* terbaik yang berfungsi sebagai pemisah dua kelas data. SVM mampu bekerja pada dataset yang berdimensi tinggi dengan menggunakan kernel trik. SVM hanya menggunakan beberapa titik data terpilih yang berkontribusi (support vector) untuk membentuk model yang akan digunakan dalam proses klasifikasi [6].

$$f(x) = w \cdot x + b$$

Atau

$$f(x) = \sum_{i=1}^m a_i y_i k(x, x_i) + b$$

Keterangan:

w : parameter *hyperplane* yang dicari (garis yang tegak lurus antara garis

hyperplane dan titik support vector

x : titik data masukan *Support Vector Machine*

a_i : nilai bobot setiap titik data

$k(x, x_i)$: fungsi kernel

b : parameter *hyperplane* yang dicari (nilai bias)

Untuk penelitian ini menggunakan kernel polynomial. Persamaannya:

$$K(x, y) = (x \cdot y + c)^d$$

Keterangan:

$K(x, y)$: Nilai kernel dari data x dan data y

x : fitur data 1

y : fitur data 2

c : nilai konstanta

d : nilai *degree*

3. HASIL DAN PEMBAHASAN

Penelitian dilakukan dengan menggunakan laptop dengan processor Intel Core i5-2520M CPU @2.50GHz dengan RAM 4GB, dan system operasi Windos 7 Professional 64bit. Sedangkan dalam melakukan eksperimen didukung dengan tools Rapidminer Versi 9.9. Penelitian ini menggunakan dataset hasil kuesioner yang telah diisi oleh pelanggan Indihome Witel Jakarta Selatan tentang kepuasan pelanggan terhadap Indihome dengan 120 data yang terdiri dari 6 atribut dan 2 kelas.

dataset ini berkaitan dengan respon atau penilaian kualitas Indihome di Witel Jakarta Selatan.

Tahap implementasi menggunakan tahapan CRISP-DM yaitu terdiri dari enam tahap, namun penelitian ini hanya melakukan lima tahap dengan langkah-langkahnya adalah *Business Understanding, Data Understanding, Data Preparation, Modelling, dan Evaluation*. Berikut menjelaskan tentang implementasi tahapan CRISP-DM yang digunakan pada penelitian sebagai berikut:

3.1. Pemahaman Bisnis (*Business Understanding*)

Pada tahap *Business Understanding* bertujuan untuk memahami bidang masalah, menghasilkan solusi yang tepat, dan mengungkapkan faktor penting yang berpengaruh pada hasil penelitian. Dalam memprediksi tingkat kepuasan pelanggan terhadap layanan Indihome Jakarta Selatan penulis melakukan pemahaman bisnis berjalan pada PT Telkom Witel Jakarta Selatan dengan menganalisa kebutuhan apa saja yang diperlukan.

Business Understanding juga dilakukan untuk memahami dan mendapatkan algoritma terbaik yang akan digunakan untuk mengolah data dalam memprediksitingkat kepuasan pelanggan Indihome Witel Jakarta Selatan. Adapun algoritma yang di dapat yaitu *Support Vector Machine* (SVM), *Decision Tree* (DT) dan *Naïve Bayes* (NB) dengan menggunakan teknik SMOTE untuk menyeimbangkan data dan *Optimasi Particle Swarm Optimization* (PSO) untuk mengoptimalkan hasil.

3.2. Pemahaman Data (*Data Understanding*)

Pada tahap *Data Understanding* bertujuan untuk mengumpulkan, mengidentifikasi, dan memahami data yang dimiliki. Data tersebut juga harus dapat diverifikasi kebenarannya. Data yang akan digunakan dalam penelitian ini adalah data yang dihasilkan dari kuesioner yang telah diisi oleh pelanggan Indihome Witel Jakarta Selatan.

Data yang digunakan akan diolah dari hasil kuesioner yang diberikan kepada pelanggan Indihome pada Witel Jakarta Selatan. Sampel yang digunakan adalah 120 responden. Setelah data yang diperlukan diperoleh, kemudian ditentukan variabel – variabel yang akan digunakan dalam penelitian. Sedangkan variabel responnya adalah kepuasan, yaitu

pelanggan Indihome yang Puas dan Tidak Puas terhadap kualitas pelayanan terhadap Indihome Witel Jakarta Selatan. Setiap aspek terdiri dari beberapa pertanyaan yang diajukan kepada pelanggan indihome.

3.3. Persiapan Data (*Data Preparation*)

Pada tahap *preprocessing* data yang pertama yaitu menghilangkan (*cleansing*) atribut-atribut yang tidak digunakan untuk mengolah data seperti No, Nama, Jenis Kelamin dan Pekerjaan, dan Setelah menghilangkan atribut data yang tidak dipakai kemudian dilakukan *preprocessing* mengubah jenis data numerical menjadi nominal untuk atribut A1, A2, A3, A4 dan A5, sehingga dapat dilihat data *cleansing* atau data yang sudah siap untuk diolah pada Tabel 3.

Tabel 3. Data Kepuasan Pelanggan Indihome yang telah di *cleansing*

No	A1	A2	A3	A4	A5	AH
1	3	3	4	4	3	Puas
2	3	2	3	3	3	Puas
3	4	3	4	4	4	Puas
4	3	3	3	3	3	Puas
5	3	4	3	4	3	Puas
6	4	4	4	4	4	Puas
7	3	3	3	3	3	Puas
8	3	2	3	4	3	Puas
9	3	3	3	4	4	Puas
10	3	3	3	3	4	Puas
11	4	3	4	4	4	Puas
12	4	3	3	4	3	Tidak Puas
13	2	3	3	2	3	Puas
14	3	4	4	4	4	Puas
15	3	3	3	3	3	Puas
..
..
117	3	2	2	2	3	Tidak Puas
118	4	3	3	3	4	Puas
119	4	2	3	3	4	Puas
120	4	2	3	3	4	Puas

3.3.1. *Preprocessing* Data Numerical Menjadi Nominal

Tahap *preprocessing* data numerical menjadi nominal yang diubah adalah atribut A1, A2, A3, A4 dan A5 dapat dilihat pada Gambar 2. Gambar 2 menunjukkan bahwa *type data* telah diubah dari numerik menjadi nominal di Atribut A1, A2, A3, A4, dan A5. Data yang telah didapat menunjukkan nilai Tidak Puas berjumlah 53 dan nilai Puas 67 dengan keseluruhan data berjumlah 120 data.



Gambar 2. Hasil *preprocessing* data numerical menjadi nominal

3.3.2. *Preprocessing* Data Menggunakan *Synthetic Minority Over-sampling Technique* (SMOTE)

Metode *Synthetic Minority Over-sampling Technique* (SMOTE) merupakan metode yang diterapkan dalam rangka menangani data yang tidak seimbang. Teknik ini mensintesis sampel baru dari kelas minoritas untuk menyeimbangkan dataset dengan cara sampling ulang sampel kelas minoritas. Berdasarkan data yang diperoleh dari hasil kuesioner terdapat 53 data dengan class Tidak Puas dan 67 data dengan class Puas. Hal ini menunjukkan ketidak seimbangan data antara dua class yaitu Tidak Puas dan Puas dengan selisih 14 data, maka dari ini dalam penelitian ini akan menggunakan Teknik SMOTE untuk menyeimbangkan data, hasil dari Teknik SMOTE dapat dilihat pada Gambar 3.



Gambar 3. Hasil dari data menggunakan Teknik SMOTE

Pada Gambar 3. dapat dilihat bahwa dengan menggunakan Teknik SMOTE dapat membuat data menjadi seimbang (*balance*), dengan jumlah class Tidak Puas sejumlah 67 dan class Puas juga berjumlah 67. Adapun jumlah data setelah menggunakan teknik SMOTE dapat dilihat pada Tabel 4.

Tabel 4. jumlah data setelah menggunakan teknik SMOTE

Atribut	Sangat Tidak Puas	Tidak Puas	Puas	Sangat Puas
A1	0	36	62	36
A2	20	57	44	13
A3	4	35	60	35
A4	3	38	61	32
A5	5	50	47	32

Tabel 4. dapat dilihat bahwa setelah melakukan teknik SMOTE terjadi perbedaan jumlah data peratribut berdasarkan nilai Sangat Tidak Puas, Tidak Puas, Puas dan Sangat Puas sehingga menghasilkan data yang seimbang antara nilai Puas dan tidak Puas.

3.4. Modelling

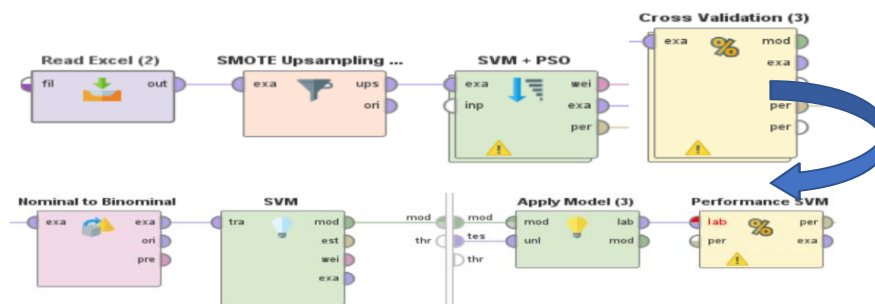
Modelling merupakan fase pemilihan teknik mining dengan menentukan algoritma yang digunakan. Dalam tahap *Modelling* ini dilakukan teknik pengklasifikasian data dengan membandingkan tiga algoritma yaitu *Support Vector Machine (SVM)*, *Decision Tree (DT)* dan

Naïve Bayes (NB) menggunakan *Cross Validation*, *SMOTE* dan *Particle Swarm Optimization (PSO)*.

3.4.1. Metode *Support Vector Machine (SVM)* dengan *SMOTE* dan *Optimasi Particle Swarm Optimization*

Desain proses pengujian model SVM dengan *Cross Validation*, *SMOTE* dan *PSO* yang digunakan dengan pada aplikasi Rapid Miner dapat dilihat pada Gambar 4.

Pada Gambar 4. yaitu *Modelling* metode model SVM dengan *PSO*, *Cross Validation* dan *SMOTE* diawali dengan *Read Excel* untuk mengambil data kepuasan pelanggan Indihome kemudian dihubungkan dengan operator *SMOTE* dengan nilai number of neighbours 5 untuk menyeimbangkan data *PSO* yang merupakan salah satu metode yang digunakan untuk optimasi multirespon pada proses manufaktur kemudian disambungkan dengan *node Cross Validation* dan setelah mengubah data class menjadi binominal selanjutnya selanjutnya dihubungkan dengan *node SVM*, *Apply Model* dan *Performance*.



Gambar 4. *Modelling Support Vector Machine (SVM)* dengan *SMOTE* dan *Optimasi Particle Swarm Optimization*

3.4.2. Metode *Decision Tree (DT)* dengan *SMOTE* dan *Optimasi Particle Swarm Optimization*

Desain proses pengujian model DT dengan *Cross Validation*, *SMOTE* dan *PSO* yang digunakan dengan pada aplikasi Rapid Miner dapat dilihat pada gambar 5. *Modelling* metode model DT dengan *PSO*, *Cross Validation* dan *SMOTE* diawali dengan *Read Excel* untuk mengambil data kepuasan pelanggan Indihome kemudian dihubungkan dengan operator *SMOTE* dengan nilai number of neighbours 5 untuk menyeimbangkan data *PSO* yang merupakan salah satu metode yang digunakan untuk optimasi multirespon pada proses manufaktur kemudian disambungkan dengan

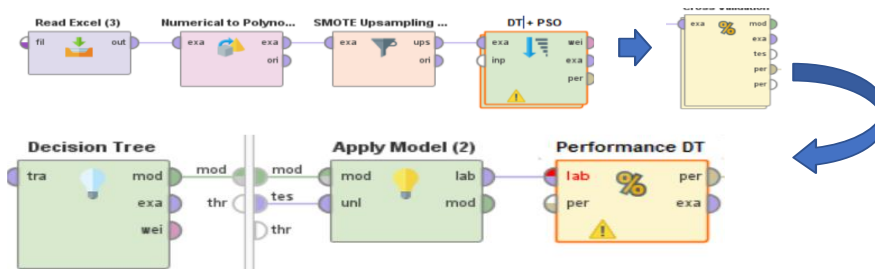
node Cross Validation dan selanjutnya selanjutnya dihubungkan dengan *node DT*, *Apply Model* dan *Performance*.

3.4.3. Metode *Naïve Bayes (NB)* dengan *SMOTE* dan *Optimasi Particle Swarm Optimization*

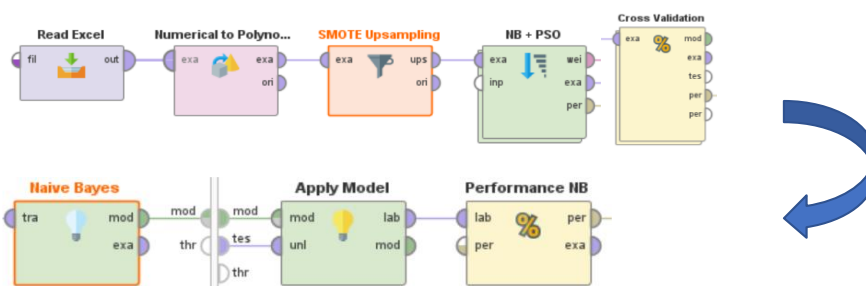
Desain proses pengujian model NB dengan *Cross Validation*, *SMOTE* dan *PSO* yang digunakan dengan pada aplikasi Rapid Miner dapat dilihat pada Gambar 6. *Modelling* metode model NB dengan *PSO*, *Cross Validation* dan *SMOTE* diawali dengan *Read Excel* untuk mengambil data kepuasan pelanggan Indihome kemudian dihubungkan dengan operator *SMOTE* untuk menyeimbangkan data *PSO* yang

merupakan salah satu metode yang digunakan untuk optimasi multirespon pada proses manufaktur kemudian disambungkan dengan

node Cross Validation dan selanjutnya selanjutnya dihubungkan dengan *node* NB, Apply Model dan *Performance*.



Gambar 5. Modelling Naive Bayes (NB) dengan SMOTE dan Optimasi Particle Swarm Optimization



Gambar 6. Modelling Naive Bayes (NB) dengan SMOTE dan Optimasi Particle Swarm Optimization

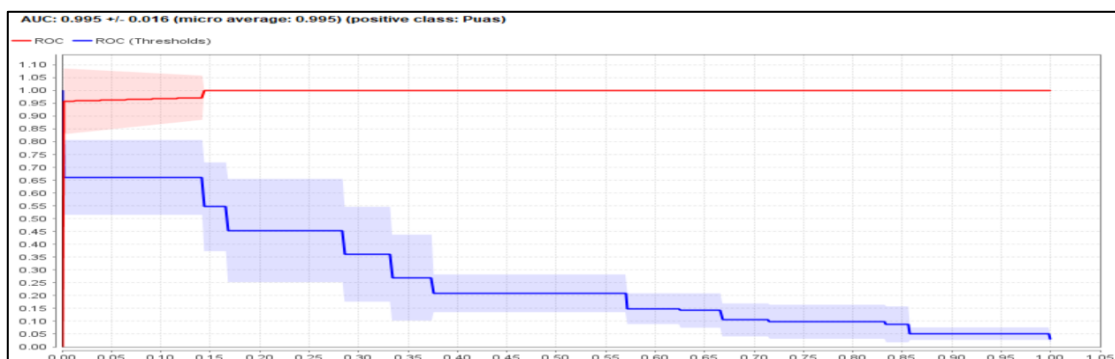
3.5. Evaluation

3.5.1. Metode Support Vector Machine (SVM) dengan SMOTE dan Optimasi Particle Swarm Optimization

Kurva ROC dan *Confusion Matrix* dari algoritma SVM dengan SMOTE dan PSO dapat dilihat pada Gambar 7. Kurva ROC SVM dengan SMOTE dan PSO dengan nilai AUC (*Area Under Curve*) yang dihasilkan dari gambar 7 sebesar 0,995 dimana diagnosa hasilnya adalah *excellent classification*. *Confusion Matrix* memberikan penilaian *performance* klasifikasi berdasarkan objek benar dan salah. *Confusion Matrix* berisi informasi aktual dan prediksi pada sistem klasifikasi. Gambar 8 menampilkan *Confusion*

Matrix menggunakan algoritma NB dengan SMOTE dan PSO Pada *Confusion Matrix* menunjukkan nilai *accuracy*, *precision* dan *recall*.

Berdasarkan gambar 8 dapat dilihat bahwa nilai *accuracy* yang diperoleh yaitu 98,57%, *precision* 98,51%, *recall* 98,51% dan AUC sebesar 0,995. Akurasi yang diperoleh yaitu 98,57% dari 134 data kepuasan pelanggan Indihome. Class data Puas yang sesuai prediksi yaitu 66 data. Class data Tidak Puas yang termasuk ke dalam prediksi Puas yaitu 1 data. Data Puas yang termasuk kedalam prediksi Tidak Puas yaitu 1 data dan data Tidak Puas yang sesuai prediksi yaitu 66 data.



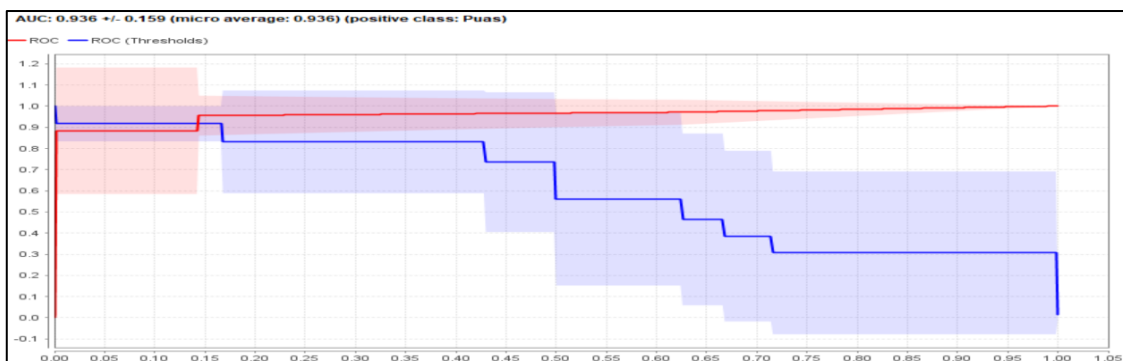
Gambar 7. ROC Support Vector Machine (SVM) dengan SMOTE dan Optimasi Particle Swarm Optimization

accuracy: 98.57% +/- 4.52% (micro average: 98.51%)			
	true Puas	true Tidak Puas	class precision
pred. Puas	66	1	98.51%
pred. Tidak Puas	1	66	98.51%
class recall	98.51%	98.51%	

Gambar 8. Confusion Matrix menggunakan algoritma SVM dengan SMOTE dan Optimasi Particle Swarm Optimization

3.5.2. Metode *Decision Tree* (DT) dengan SMOTE dan *Optimasi Particle Swarm Optimization* Kurva ROC dan *Confusion Matrix* dari algoritma DT dengan SMOTE dan PSO dapat dilihat pada Gambar 9. Kurva ROC DT dengan SMOTE dan PSO dengan nilai AUC (*Area Under Curve*) yang dihasilkan dari gambar 9 sebesar 0,936 dimana diagnosa hasilnya adalah *excellent classification*. *Confusion Matrix* memberikan penilaian *performance* klasifikasi berdasarkan objek benar dan salah. *Confusion Matrix* berisi informasi aktual dan prediksi pada sistem klasifikasi. Gambar 10 menampilkan *Confusion Matrix* menggunakan algoritma DT dengan SMOTE dan PSO Pada *Confusion Matrix* menunjukkan nilai *accuracy*, *precision* dan *recall*. Berdasarkan gambar 10 dapat dilihat bahwa nilai *accuracy* yang diperoleh yaitu 97,14%, *precision* 98,46%, *recall* 95,52% dan AUC sebesar 0,936. Akurasi yang diperoleh yaitu 97,14% dari 134 data kepuasan pelanggan Indihome. Class data Puas yang sesuai prediksi yaitu 64 data. Class data Tidak Puas yang termasuk ke dalam prediksi Puas yaitu 3 data. Data Puas yang termasuk kedalam prediksi Tidak Puas yaitu 1 data dan data Tidak Puas yang sesuai prediksi yaitu 66 data.

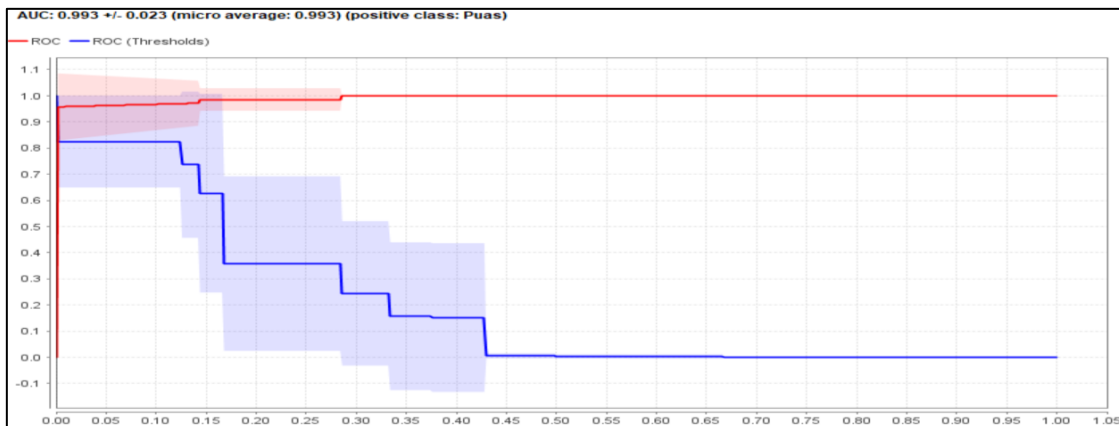
3.5.3. Metode *Naïve Bayes* (NB) dengan SMOTE dan *Optimasi Particle Swarm Optimization* Kurva ROC dan *Confusion Matrix* dari algoritma NB dengan SMOTE dan PSO dapat dilihat pada gambar 11. Kurva ROC NB dengan SMOTE dan PSO dengan nilai AUC (*Area Under Curve*) yang dihasilkan dari gambar 11 sebesar 0,993 dimana diagnosa hasilnya adalah *excellent classification*. *Confusion Matrix* memberikan penilaian *performance* klasifikasi berdasarkan objek benar dan salah. *Confusion Matrix* berisi informasi aktual dan prediksi pada sistem klasifikasi. Gambar 12 menampilkan *Confusion Matrix* menggunakan algoritma NB dengan SMOTE dan PSO Pada *Confusion Matrix* menunjukkan nilai *accuracy*, *precision* dan *recall*. Berdasarkan gambar 12 dapat dilihat bahwa nilai *accuracy* yang diperoleh yaitu 97,32%, *precision* 97,06%, *recall* 98,51% dan AUC sebesar 0,993. Akurasi yang diperoleh yaitu 97,32% dari 134 data kepuasan pelanggan Indihome. Class data Puas yang sesuai prediksi yaitu 65 data. Class data Tidak Puas yang termasuk ke dalam prediksi Puas yaitu 2 data. Data Puas yang termasuk kedalam prediksi Tidak Puas yaitu 1 data dan data Tidak Puas yang sesuai prediksi yaitu 66 data.



Gambar 9. ROC Decision Tree (DT) dengan SMOTE dan Optimasi Particle Swarm Optimization

accuracy: 97.14% +/- 6.90% (micro average: 97.01%)			
	true Puas	true Tidak Puas	class precision
pred. Puas	64	1	98.46%
pred. Tidak Puas	3	66	95.65%
class recall	95.52%	98.51%	

Gambar 10. Confusion Matrix menggunakan algoritma DT dengan SMOTE dan Optimasi Particle Swarm Optimization



Gambar 11. ROC Naïve Bayes (NB) dengan SMOTE dan Optimasi Particle Swarm Optimization

precision: 97.32% +/- 5.66% (micro average: 97.06%) (positive class: Puas)			
	true Tidak Puas	true Puas	class precision
pred. Tidak Puas	65	1	98.48%
pred. Puas	2	66	97.06%
class recall	97.01%	98.51%	

Gambar 12. Confusion Matrix menggunakan algoritma NB dengan SMOTE dan Optimasi Particle Swarm Optimization

3.5.4. Hasil Evaluasi

Perbandingan hasil *accuracy*, *precision*, *recall* dan AUC Algoritma SVM, DT dan NB dengan

Cross Validation, SMOTE dan PSO dapat dilihat pada tabel 5.

Tabel 5. Perbandingan Hasil antar Algoritma

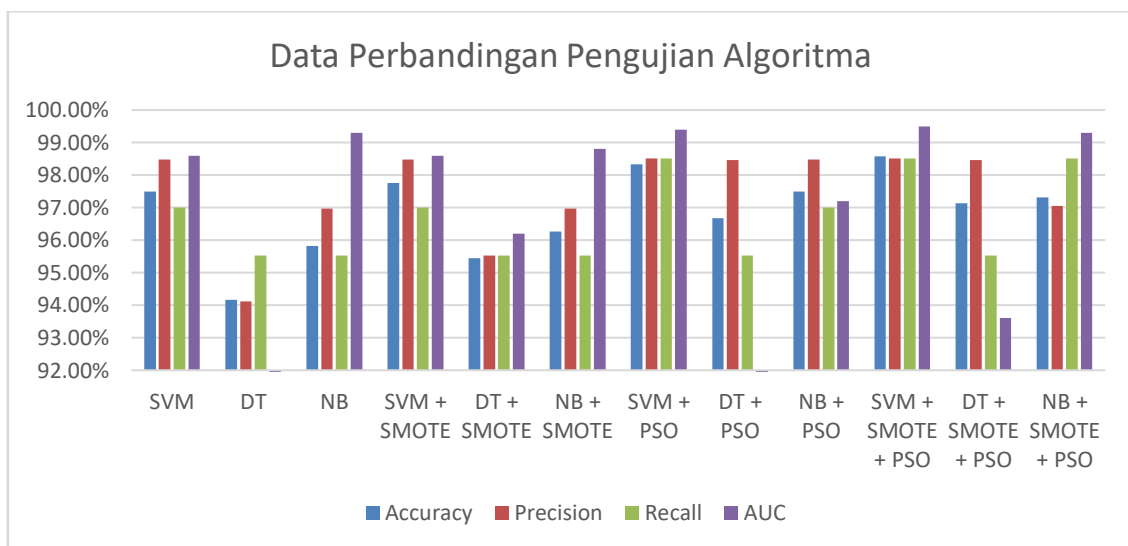
Algoritma	Accuracy	Precision	Recall	AUC
SVM	97.50%	98.48%	97.01%	0.986
DT	94.17%	94.12%	95.52%	0.859
NB	95.83%	96.97%	95.52%	0.993
SVM + SMOTE	97.75%	98.48%	97.01%	0.986
DT + SMOTE	95.44%	95.52%	95.52%	0.962
NB + SMOTE	96.26%	96.97%	95.52%	0.988
SVM + PSO	98.33%	98.51%	98.51%	0.994
DT + PSO	96.67%	98.46%	95.52%	0.906
NB + PSO	97.50%	98.48%	97.01%	0.972
SVM + SMOTE + PSO	98.57%	98.51%	98.51%	0.995
DT + SMOTE + PSO	97.14%	98.46%	95.52%	0.936
NB + SMOTE + PSO	97.32%	97.06%	98.51%	0.993

Dari tabel 5 dapat dilihat bahwa untuk memprediksi tingkat kepuasan pelanggan terhadap layanan Indihome Jakarta Selatan menggunakan algoritma SVM mendapatkan hasil akurasi 97,50%, precision 98,48%, recall 97,01% dan nilai AUC 0,986. Pengolahan data menggunakan algoritma DT mendapatkan hasil akurasi 94,17%, precision 94,12%, recall 95,52%

dan nilai AUC 0.859. Kemudian pengolahan data menggunakan algoritma NB mendapatkan hasil akurasi 95,83%, precision 96,97%, recall 95,52% dan nilai AUC 0.993. Sedangkan pengolahan data menggunakan algoritma SVM+SMOTE mendapatkan hasil akurasi 97,75%, precision 96,97%, recall 95,52% dan nilai AUC 0.986. Pengolahan data menggunakan algoritma

DT+SMOTE mendapatkan hasil akurasi 95.44%, precision 95,52%, recall 95,52% dan nilai AUC 0.962. Data yang diolah dengan menggunakan algoritma NB+SMOTE mendapatkan hasil akurasi 96.26%, precision 96,97%, recall 95,52% dan nilai AUC 0.988. Penggunaan PSO dengan algoritma SVM mendapatkan hasil akurasi 98.33%, precision 98,51%, recall 98,51% dan nilai AUC 0.994. Kemudian penggunaan PSO dengan algoritma DT mendapatkan hasil akurasi 96.67%, precision 98,46%, recall 95,52% dan nilai AUC 0.906. Sedangkan penggunaan PSO dengan algoritma NB mendapatkan hasil akurasi 97.50%, precision 98,48%, recall 97,01% dan nilai AUC

0.972. Pengolahan algoritma SVM dikolaborasikan dengan Teknik SMOTE untuk menyeimbangkan data dan teknik PSO mendapatkan hasil akurasi 98.57%, precision 98,51%, recall 98,51% dan nilai AUC 0.995. Algoritma DT dikolaborasikan dengan Teknik SMOTE untuk menyeimbangkan data dan teknik PSO mendapatkan hasil akurasi 97.14%, precision 98,46%, recall 95,52% dan nilai AUC 0.936. Sedangkan algoritma NB dikolaborasikan dengan Teknik SMOTE untuk menyeimbangkan data dan teknik PSO mendapatkan hasil akurasi 97.32%, precision 97,06%, recall 98,51% dan nilai AUC 0.993.



Gambar 13. Grafik Perbandingan Pengujian Algoritma

Pada gambar 13 dapat dilihat bahwa pengolahan data untuk memprediksi tingkat kepuasan pelanggan Indihome Witel Jakarta Selatan memperoleh nilai akurasi tertinggi dengan menggunakan algoritma SVM dikolaborasikan dengan Teknik SMOTE untuk menyeimbangkan data dan teknik PSO untuk mengoptimalkan hasil pengolahan. Nilai yang didapat dari algoritma SVM + SMOTE + PSO dengan akurasi tertinggi 98.57%, precision tertinggi 98.51%, recall tertinggi 98.51% dan AUC tertinggi 0.995 sehingga termasuk dalam *excellent classification*.

4. KESIMPULAN

Penyedia layanan telekomunikasi Indihome Witel Jakarta Selatan mempunyai keunggulan dan kekurangan yang dapat memberikan kepuasan dan ketidakpuasan yang membuat pelanggan berpindah ke penyedia layanan internet lain. Pengetahuan tentang klasifikasi konsumen yang puas dan tidak puas

sangat dibutuhkan dalam mengambil keputusan. Dalam mengukur atau memprediksi tingkat kepuasan pelanggan dapat menggunakan algoritma klasifikasi data mining yaitu *Naive Bayes*, *Support Vector Machine*, *Decision Tree* dengan penambahan fitur Metode *Synthetic Minority Over-sampling Technique* (SMOTE) apabila data tidak seimbang dan teknik *Particle Swarm Optimization* (PSO). Berdasarkan hasil penelitian dapat ditarik kesimpulan bahwa prediksi tingkat kepuasan pelanggan terhadap layanan Indihome Jakarta Selatan dengan metode klasifikasi data mining Algoritma algoritma SVM + SMOTE + PSO merupakan metode yang terbaik dalam penelitian ini dibandingkan dengan NB dan DT baik yang menggunakan SMOTE dan PSO. Algoritma SVM dengan SVM + SMOTE + PSO menghasilkan Accuracy 98.57%, precision tertinggi 98.51%, recall tertinggi 98.51% dan AUC tertinggi 0.995 yang termasuk dalam *excellent classification*.

Usulan penggunaan SMOTE yang berguna untuk menyeimbangkan data dan PSO untuk mengoptimalkan hasil nilai pada algoritma klasifikasi terbukti berpengaruh meningkatkan accuracy algoritma SVM namun tidak secara signifikan. Sehingga dapat disimpulkan bahwa penerapan optimasi yang terbaik pada model ini adalah algoritma SVM dengan SMOTE dan PSO dapat memberikan solusi terhadap permasalahan untuk memprediksi tingkat kepuasan pelanggan terhadap layanan Indihome Jakarta Selatan.

5. REFERENSI

- Ariani, F., & Taufik, A. (2020). Perbandingan Metode Klasifikasi Data Mining untuk Prediksi Tingkat Kepuasan Pelanggan Telkomsel Prabayar. *SATIN – Sains Dan Teknologi Informasi*, 16(1), 1–6. <https://doi.org/10.33372/stn.v6i2.666>
- Bastian, Y., Tambunan, H. S., Saputra, W., Bangsa, S. T., Utara, S., Bangsa, A. T., & Utara, S. (2021). Analisis Penerapan Algoritma C4.5 Dalam Mengukur Tingkat Kepuasan Pelanggan Indihome Pada Kota Pematangsiantar. *KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)*, 2(1), 62–69.
- Desyanti, D.-. (2019). Penerapan Data Mining Algoritma C4.5 untuk Mengetahui Tingkat Kepuasan Konsumen di Hotel Grand Zuri Dumai. *SATIN - Sains Dan Teknologi Informasi*, 4(2), 36. <https://doi.org/10.33372/stn.v4i2.403>
- Kominfo, A. (2019). *Penggunaan Internet di Indonesia*. Aptika Kominfo. <https://aptika.kominfo.go.id/2019/08/penggunaan-internet-di-indonesia/>
- Putra, H. E., & Widiyanesti, S. (2018). Analisis Masalah Kualitas Layanan PT . Telekomunikasi Indonesia Witel Jakarta Selatan Pada Produk Indihome (Menggunakan Metode Pendekatan Six Sigma) Problem Analysis Service Quality Pt . Telekomunikasi Indonesia Witel South Jakarta On Indihome Product (Us. *E-Proceeding of Management*, 5(1), 195–202.
- Rofiqoh, U., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 1(12), 1725–1732.
- Rohanah, A., Rianti, D. L., Sari, B. N., Informatika, T., & Karawang, U. S. (2021). Perbandingan Naïve Bayes dan Support Vector Machine untuk Klasifikasi Ulasan Pelanggan Indihome. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)*, 6(1), 23–30. <https://doi.org/10.30998/STRING.V6I1.9232>
- Somantri, O., & Apriliani, D. (2018). Support Vector Machine Berbasis Feature Selection Untuk Sentiment Analysis Kepuasan Pelanggan Terhadap Pelayanan Warung dan Restoran Kuliner Kota Tegal. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 5(5), 537–548. <https://doi.org/10.25126/JTIK.201855867>
- Syahputra, R. D., & Sampurno, W. (2019). Analisis Kualitas Pelayanan Dan Nilai Pelanggan Berbasis Digital Pada Media Sosial Facebook Terhadap Kepuasan Pelanggan (Studi Pada Produk Indihome Di Plaza Telkom Lembong Kota Bandung Tahun 2019). *EProceedings of Applied Science*, 5(3), 1841.
- Telkom. (2018). *Laporan Tahunan Telkom Indonesia Tahun 2018*. 01(01), 1689–1699.
- Yunitasari, Siti Hopipah, H., & Mayasari, R. (2021). Optimasi Backward Elimination untuk Klasifikasi Kepuasan Pelanggan Menggunakan Algoritme k-nearest neighbor (k-NN) and Naive Bayes. *Technomedia Journal*, 6(1 Agustus), 99–110. <https://doi.org/10.33050/TMJ.V6I1.1531>