

Analisis Faktor Keterlambatan Pemasangan Meter Baru Menggunakan Metode Decision Tree Pada PT PLN Unit Layanan Pelanggan Daerah Pringgabaya Lombok Timur

Yahya^{1*}, Arnita Sandi²

¹Program Studi Sistem Informasi, Universitas Hamzanwadi

²Program Studi Teknik Komputer, Universitas Hamzanwadi

*ayhay7078@gmail.com

Abstrak

PT. PLN (Persero) ULP Pringgabaya adalah salah satu unit yang didirikan oleh PT PLN agar mampu memberikan pelayanan secara langsung kepada masyarakat pada wilayah dengan cakupan lebih kecil. Akan tetapi penggunaan energi listrik yang banyak dilakukan oleh masyarakat menyebabkan peningkatan permintaan pemasangan meter baru. Penelitian ini bertujuan untuk menganalisis penyebab keterlambatan pemasangan meter baru. Untuk dapat mengklasifikasikan penyebab paling tinggi dalam keterlambatan pemasangan, digunakan metode Decision Tree. Jumlah data set yang digunakan sebanyak 700 data dengan atribut, seperti jarak lokasi dengan jaringan PLN (Meter), tegangan listrik (Volt), kontak pelanggan, SLO (Sertifikat Laik Operasi), Instalasi dan Status Pemasangan. Dari hasil perhitungan, perolehan meliputi Jarak ≤ 37 M dari jaringan PLN, tegangan ≥ 220 Volt, kontak yang tidak aktif, SLO yang lama terbit, dan Instalasi yang tidak terpasang. Pengujian dilakukan sebanyak 10 kali pengujian menggunakan *cross validation* dengan hasil accuracy 100% dengan nilai AUC = 1.00 yang berarti menggunakan algoritma decision tree dapat dimanfaatkan dalam menentukan klasifikasi penyebab keterlambatan pemasangan meter baru PT PLN ULP (Unit Layanan Pelanggan) Pringgabaya. Sehingga dapat disimpulkan bahwa penyebab keterlambatan pemasangan meter baru PT PLN ULP (Unit Layanan Pelanggan) Pringgabaya bukan disebabkan oleh kriteria pelayanan melainkan dipengaruhi oleh jarak, tegangan, kontak pelanggan baru.

Kata kunci : *Decision Tree*, Klasifikasi, Pelayanan

Abstract

PT. PLN (Persero) ULP Pringgabaya is one of the units established by PT PLN to be able to provide direct services to the community in areas with smaller coverage. However, the use of electricity by the community has increased the demand for the installation of new meters. This study aims to analyze the causes of delays in the installation of new meters. In order to classify the highest causes of installation delays, the Decision Tree method is used. The number of data sets used is 700 data with attributes, such as the distance of the location to the PLN network (Meter), electric voltage (Volt), customer contact, SLO (Certificate of Operational Worthiness), Installation and Installation Status. From the calculation results, the acquisition includes Distance ≤ 37 M from the PLN network, voltage ≥ 220 Volts, inactive contacts, old SLOs issued, and uninstalled installations. Testing was carried out 10 times using cross validation with 100% accuracy results with an AUC value of 1.00, which means that the use of the decision tree algorithm can be utilized in determining the classification of the causes of delays in the installation of new meters at PT PLN ULP (Customer Service Unit) Pringgabaya. So it can be concluded

that the cause of the delay in the installation of new meters at PT PLN ULP (Customer Service Unit) Pringgabaya is not caused by service criteria but is influenced by distance, voltage, and new customer contacts.

Keywords: Decision Tree, Classification, Service

1. Pendahuluan

Energi listrik menjadi sumber utama dalam pembuatan alat-alat penunjang kehidupan manusia. Melonjaknya penggunaan listrik dalam kehidupan sehari-hari menjadikan tuntutan kepada PT PLN untuk dapat memberikan pelayanan kepada masyarakat secara tepat waktu. Pada PT PLN ULP Pringgabaya banyaknya permintaan pemasangan meter baru menjadi salah satu permasalahan yang sulit diatasi. Keterlambatan pemasangan menjadi dampak yang harus dihadapi sehingga menjadikan permasalahan ini salah satu pengaduan yang paling sering diajukan dalam sistem Aplikasi PLN Mobile.

Pemasangan Meter Baru pada PT. PLN ULP, masih banyak kendala dan hambatan yang membuat pelayanan menjadi terganggu, menurut informasi yang diperoleh dari bagian pelayanan, hambatan-hambatan yang dimaksud antara lain : 1. Di beberapa wilayah di Pringgabaya, jarak jangkauan yang sulit untuk dijangkau, 2, Kapasitas Voltase yang masih kurang memadai. 3, Adanya kesulitan komunikasi dengan pelanggan baru.

1. Berdasarkan informasi di atas, penulis berusaha menguraikan dan menganalisis

pelayanan menggunakan metode decision tree untuk mengolah data, dimana decision tree adalah metode yang tepat untuk pengambilan keputusan dalam permasalahan ini. Penggunaan decision tree juga dimaksudkan agar dapat diketahui faktor apa saja yang paling mempengaruhi permasalahan sehingga dapat diketahui prioritas yang perlu didahulukan dalam mengatasinya.

2. Tinjauan Pustaka

2.1. Penelitian Terkait

Dalam melakukan penelitian ini, penulis menggunakan beberapa acuan pada penelitian sebelumnya sebagai berikut :

1. Penelitian yang dilakukan oleh Biaq Andriskha Candra Permana, Muhammad Sadali, dan Ramli Ahmad dalam jurnal yang berjudul “Penerapan Model Decision Tree Menggunakan Python Untuk Prediksi Faktor Dominan Penyebab Penyakit Stroke”. Penelitian ini dibuat untuk mempermudah dalam melihat gejala yang paling beresiko hingga yang memiliki resiko terendah karena penyakit stroke[1].
2. Penelitian yang dilakukan oleh Yahya, Nurhidayati dan Andi Suherman dalam

jurnal yang berjudul “Prediksi Tingkat Kesehatan Masyarakat Berdasarkan Penggunaan Alat Kontrasepsi Menggunakan Algoritma Random Forest” Penelitian ini dibuat untuk mengetahui bahwa alat kontrasepsi yang paling banyak dan efisien untuk digunakan di wilayah Kecamatan Suralaga adalah Pil KB dan Suntik KB[2].

3. Penelitian yang dilakukan oleh Reni Kurniah, Dadang Yunika Surya Putra dan Elviza Diana dalam jurnal yang berjudul “Penerapan Data Mining Decision Tree Algoritma C4.5 Untuk Mengetahui Tingkat Kepuasan Mahasiswa Terhadap Layanan Akademik Dan Kemahasiswaan” Penelitian ini dibuat untuk keperluan evaluasi kebijakan perguruan tinggi dalam meningkatkan kualitas pelayanan pada bidang akademik dan kemahasiswaan[3].
4. Penelitian yang dilakukan oleh Mahpuz, Amri Muliawan Nur dan L M Samsu dalam jurnal yang berjudul “Penerapan Algoritma C4.5 Dalam Mengklasifikasi Status Gizi Balita Pada Posyandu Desa Dames Damai Kabupaten Lombok Timur”. Penelitian ini dibuat untuk memberikan informasi kepada masyarakat umum mengenai jadwal kegiatan dan jenis layanan yang terdapat pada poloklinik[4].

5. Penelitian yang dilakukan oleh Yupi Kuspandi Putra, Hamzan Ahmadi dan Suhartini dalam jurnal yang berjudul “Komparasi Algoritma C4.5 Dan C4.5 Berbasis Pso Untuk Prediksi Jumlah Penggunaan BBM Perbulan Pada Kantor Dinas Lingkungan Hidup Dan Kebersihan Kabupaten Lombok Timur”. Penelitian ini dilakukan untuk memprediksi jumlah penggunaan Bahan Bakar Minyak (BBM) perbulan pada Kantor Dinas Lingkungan Hidup dan Kebersihan Kabupaten Lombok Timur[5].

2.2. Landasan Teori

2.1. Pengertian Data Mining

Data mining adalah proses menganalisa data dari perspektif yang berbeda dan menyimpulkannya menjadi informasi-informasi penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya. Secara teknis, data mining dapat disebut sebagai proses untuk menemukan korelasi atau pola dari ratusan atau ribuan field dari sebuah relasional database yang besar. Pengelompokan Data Mining dibagi menjadi beberapa kelompok yaitu[6]:

1. Deskripsi

Deskripsi merupakan cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data yang dimiliki.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variable target estimasi lebih ke arah numerik daripada ke arah kategori. Model yang dibangun menggunakan record lengkap yang menyediakan nilai variable target sebagai nilai prediksi.

3. Prediksi

Prediksi menerka sebuah nilai yang belum diketahui dan juga memperkirakan nilai untuk masa mendatang.

4. Klasifikasi

Dalam klasifikasi terdapat target variable kategori, misal penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu tinggi, sedang, dan rendah.

5. Pengklasteran

Merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

6. Asosiasi

Asosiasi bertugas menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

2.2. Tahap-Tahapan Data Mining

Karena data mining adalah suatu rangkaian proses maka dibagi menjadi beberapa tahap antara lain[7]:

- Pembersihan data: untuk membuang data yang tidak konsisten dan noise.
- Integrasi data: untuk menggabungkan data dari beberapa sumber.
- Transformasi data: untuk mengubah data menjadi bentuk yang sesuai untuk di mining.
- Aplikasi teknik data mining.
- Evaluasi pola yang ditemukan: untuk menemukan informasi yang menarik atau pun bernilai.
- Presentasi pengetahuan dengan teknik visualisasi

2.3. Python

Kepopuleran penggunaan Python menempatkannya menjadi bahasa pemrograman yang mulai banyak dipelajari oleh mahasiswa terutama mahasiswa di kampus yang berbasis IT, guna menyelesaikan tugas kuliah, tugas akhir maupun tugas penelitian, untuk dapat menyelesaikan berbagai tugas pemrograman, seseorang perlu memahami algoritma, karena pada dasarnya program computer adalah implementasi dari algoritma[8]. .

1. Google Colaboratory

Colaboratory, atau “Colab” merupakan produk dari Google Research. Colab memungkinkan siapa saja menulis dan mengeksekusi kode python arbitrer melalui browser, dan sangat cocok untuk machine learning, analisis data,

serta pendidikan. Secara lebih teknis, Colab merupakan layanan notebook Jupyter yang dihosting dan dapat digunakan tanpa penyiapan, serta menyediakan akses gratis ke resource komputasi termasuk GPU. Resource Colab tidak dijamin dan sifatnya terbatas, serta batas penggunaannya terkadang berfluktuasi. Hal ini diperlukan agar Colab dapat menyediakan resource secara gratis. Pengguna yang ingin memiliki akses lebih andal ke resource yang lebih baik dapat menggunakan Colab Pro. Memperkenalkan Colab Pro merupakan Langkah pertama yang Google ambil untuk melayani pengguna yang ingin melakukan lebih banyak hal di Colab. Tujuan jangka panjang pihak Google adalah untuk terus menyediakan versi gratis Colab, dan disaat yang bersamaan berkembang secara berkelanjutan untuk memenuhi kebutuhan pengguna Google[9]

2. Jupiter Notebook

Jupiter Notebook adalah aplikasi open source sebagai tempat untuk membuat dan berbagi dokumen python. Untuk penginstalannya cukup menggunakan perintah pip. Sedangkan untuk mengakses Jupyter Notebook dengan mengetikkan "jupyter notebook" di Command Prompt[10].

Proyek dan komunitas yang tujuannya adalah untuk "mengembangkn perangkat lunak sumber terbuka, standar terbuka, dan layanan untuk komputasi interaktif di puluhan bahasa

pemrograman". Itu dipisahkan dari IPython pada tahun 2014 oleh Fernando Pérez. Nama Project Jupyter adalah referensi ke tiga bahasa pemrograman inti yang didukung oleh Jupyter, yaitu Julia, Python dan R, dan juga penghormatan kepada notebook Galileo yang merekam penemuan bulan Jupiter. Project Jupyter telah mengembangkan dan mendukung produk komputasi interaktif Jupyter Notebook, JupyterHub, dan JupyterLab[11].

3. Pycharm

Pycharm adalah Integrated Development Environment (IDE) yang berfokus pada pengembangan project dengan Bahasa pemrograman python. Pycharm di kembangkan oleh JetBrains, Pycharm merupakan IDE cross-platform yang artinya dapat dijalankan pada berbagai sistem informasi seperti Windows, Linux, dan Mac. Terdapat beberapa library yang digunakan dengan mengupgrade pip seperti scipy, pillow, matplotlib dan opencv contrib python. Python dapat digunakan untuk berbagai keperluan pengembangan perangkat lunak dan dapat berjalan di berbagai macam sistem operasi karena sifatnya yang multiplatform. Pada kebanyakan sistem operasi linux, bahasa pemrograman ini menjadi standarisasi untuk disertakan dalam paket distribusinya. Seperti halnya bahasa pemrograman dinamis, python seringkali digunakan sebagai bahasa skrip[12]..

4. Tensorflow

TensorFlow adalah pustaka perangkat lunak sumber terbuka dan gratis untuk pembelajaran mesin. TensorFlow dapat digunakan dalam berbagai tugas tetapi memiliki fokus khusus pada pelatihan dan inferensi jaringan neural dalam. Tensorflow adalah pustaka matematika simbolis berdasarkan dataflow dan pemrograman. Saat ini, TensorFlow merupakan pustaka pembelajaran mesin paling terkenal di dunia[13]

Dari semua tools yang tertera di atas, yang digunakan adalah Google Colaboratory, karena tools ini mempunyai fitur yang sederhana yang lebih mudah untuk digunakan dalam proses pengolahan data yang digunakan.

2.4. Metode Klasifikasi

Klasifikasi merupakan teknik penambahan data yang memetakan data ke dalam kelompok atau kelas yang telah ditentukan. Ini adalah metode pembelajaran terawasi yang membutuhkan data pelatihan berlabel untuk menghasilkan aturan untuk mengklasifikasikan data uji ke dalam kelompok atau kelas yang telah ditentukan. Metode klasifikasi mengacu pada pembentukan kelompok data dengan menerapkan algoritma dikenal ke gudang data di bawah pemeriksaan. Metode ini berguna untuk proses bisnis yang membutuhkan informasi kategori seperti pemasaran atau penjualan[14]. Secara umum metode klasifikasi dibagi menjadi dua macam yaitu klasifikasi supervised dan

unsupervised. Metode klasifikasi supervised adalah metode learning melalui bantuan dari supervisor (learning melalui contoh). Sedangkan metode klasifikasi unsupervised merupakan kebalikan dari metode supervised yakni learning tanpa supervisor (learning melalui observasi). Klasifikasi unsupervised dikenal juga sebagai clustering[15]

2.5. Algoritma Decision Tree

Pohon keputusan merupakan salah satu metode klasifikasi yang populer karena dapat dengan mudah diinterpretasi oleh manusia. Pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan[16]. Algoritma ini merupakan algoritma yang populer digunakan dan memiliki tingkatan akurasi yang lebih tinggi. Algoritma C4.5 merupakan pengembangan dari Algoritma ID3, ID3 sendiri dikembangkan oleh J.Ross Quinlan. Dalam prosedur algoritma ID3, inputannya berupa sampel training, label training dan atribut.

Berikut kelebihan metode pohon keputusan[17] : 1. Daerah pengambilan keputusan yang sebelumnya kompleks dan sangat global, dapat diubah menjadi lebih simpel dan spesifik. 2. Eliminasi perhitungan-perhitungan yang tidak diperlukan, karena ketika menggunakan metode pohon keputusan maka sample diuji hanya

berdasarkan kriteria atau kelas tertentu. 3. Fleksibel untuk memilih fitur dari node internal yang berbeda, fitur yang terpilih akan membedakan suatu kriteria dibandingkan kriteria yang lain dalam node yang sama. Kefleksibelan metode pohon keputusan ini meningkatkan kualitas keputusan yang dihasilkan jika dibandingkan ketika menggunakan metode penghitungan satu tahap yang lebih konvensional. 4. Dalam analisis multivarian, dengan kriteria dan kelas yang jumlahnya sangat banyak, seorang penguji biasanya perlu mengestimasi baik itu distribusi dimensi tinggi ataupun parameter tertentu dari distribusi kelas tersebut. Metode pohon keputusan dapat menghindari munculnya permasalahan ini dengan menggunakan kriteria yang jumlahnya lebih sedikit pada setiap node internal tanpa banyak mengurangi kualitas keputusan yang dihasilkan. Kekurangan pada pohon keputusan[18]: 1. Terjadi overlapping, terutama ketika kelas-kelas dan kriteria yang digunakan jumlahnya sangat banyak. Hal tersebut juga dapat menyebabkan meningkatnya waktu pengambilan keputusan dan jumlah memori yang diperlukan. 2. Pengakumulasi jumlah kesalahan dari setiap tingkat dalam sebuah pohon keputusan yang besar. 3. Kesulitan dalam mendesain pohon keputusan yang optimal. 4. Hasil kualitas keputusan yang didapatkan dari metode pohon keputusan sangat

tergantungan pada bagaimana pohon tersebut didesain.

2.6. Pelayanan

Dalam kamus besar bahasa Indonesia dinyatakan pengertian pelayanan bahwa pelayanan adalah suatu usaha untuk membantu menyiapkan (mengurus) apa yang diperlukan orang lain. Oleh karena itu, pelayanan berfungsi sebagai sebuah sistem yang menyediakan apa yang dibutuhkan oleh masyarakat. Istilah publik, yang berasal dari Bahasa Inggris (public), terdapat beberapa pengertian yang memiliki variasi arti dalam Bahasa Indonesia, yaitu umum, masyarakat dan negara. Konsep pelayanan publik diturunkan dari makna public service yang berarti: "berbagai aktivitas yang bertujuan memenuhi kebutuhan masyarakat akan barang dan jasa", atau pelayanan umum yang diartikan sebagai segala bentuk kegiatan pelayanan yang dilaksanakan oleh instansi pemerintah pusat, daerah dan lingkungan BUMN/BUMD dalam bentuk barang dan/atau jasa baik dalam rangka pelaksanaan perundang-undangan[19]

3. Metode Penelitian

Metode pengumpulan Data yang dilakukan dalam penelitian ini antara lain: (1) Observasi yaitu dengan melakukan pengamatan secara langsung pada lokasi penelitian, sehingga diperoleh data yang sebenarnya. (2) Wawancara yaitu dengan melakukan

wawancara menggunakan teknik wawancara tidak terstruktur, jadi penulis menanyakan beberapa pertanyaan langsung ke pelanggan. (3) Dokumentasi yaitu dengan melakukan pengumpulan data dengan cara mengumpulkan data yang tertulis (*print*)

4. Hasil dan Pembahasan

4.1. Datasets

Datasets yang digunakan adalah datasets pemasangan meter baru PT PLN ULP (Unit Layanan Pelanggan) Pringabaya

NAMA	ALAMAT	JAKA K LOKASI QABET REU	TEGAN GAN ARUS	KONTAK PELANG GAN	SLO	INSTAL ASI	STATUS PEMASAN GAN
SUDIANTI	DS DUSUN TELAGA TAMPAT	31	222	TIDAK AKTIF	TER BIT	TERPAS ANG	TERLAMBAT
HURNI	DS BENYER DAYA TELAGA	32	224	TIDAK AKTIF	TER BIT	TERPAS ANG	TERLAMBAT
PATIMAH	DS PADAK SELATAN OADAK	34	223	AKTIF	BEL UM TER BIT	BELUM TERPAS ANG	TERLAMBAT
ASMAWATI	DS LENDANG BELU SELAPAR ANG	37	224	AKTIF	BEL UM TER BIT	BELUM TERPAS ANG	TERLAMBAT
BUSAIRI	DS DUN QUBUZ TIMUK	58	191	AKTIF	TER BIT	TERPAS ANG	TIDAK TERLAMBAT
IM PADIL FAISAL	DS DUSUN BUENE BEBIDAS	56	218	AKTIF	TER BIT	TERPAS ANG	TIDAK TERLAMBAT
SAMIRIN	DS BENTENG KERUMUT	15	225	TIDAK AKTIF	TER BIT	TERPAS ANG	TIDAK TERLAMBAT
WIDIYAT UL JANNAH	DS KP BATEAN LABUHAN	20	227	AKTIF	BEL UM TER BIT	TERPAS ANG	TERLAMBAT
IRWANSYAH	DS DASAN BARU KALIAGA	34	180	AKTIF	TER BIT	TERPAS ANG	TIDAK TERLAMBAT
MURHIDAYATI	DS DASAN TENGGE	27	221	TIDAK AKTIF	TER BIT	TERPAS ANG	TERLAMBAT

Gambar 1 : Datasets

4.2. Pembahasan

1) *Splitting* Data

Splitting data berfungsi untuk membagi data menjadi data pelatihan dan data pengujian. Data pelatihan berjumlah 560 sampel (80%) sedangkan data pengujian berjumlah 140 sampel (20%). Untuk menghasilkan nilai yang konsisten pada saat pengujian, dibuatkan jumlah fitur yang sama pada data pelatihan dan data pengujian berjumlah 587 fitur

`((560, 587), (140, 587))`

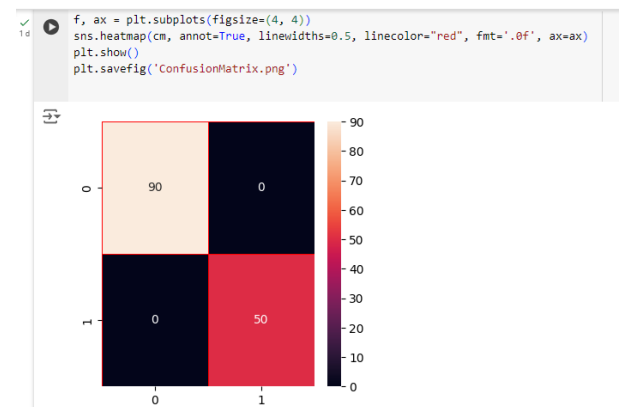
`[ ] y_train.shape , y_test.shape`

`((560,), (140,))`

Gambar 2 : *Splitting* Data

2) Confusion Matrix

Confusion matrix digunakan untuk mengukur bagaimana kinerja dari model klasifikasi dengan cara memberikan gambaran rinci tentang hasil prediksi yang dilakukan oleh model tersebut.



Gambar 3 : *Confusion Matrix*

Dari gambar di atas diperoleh confusion matrik sebagai berikut :

	True puas	True tidak puas
Pred. puas	90	0
Pred. Tidak puas	0	50

Memiliki 4 komponen yang terdiri dari:

1. True Positive (TP) : 50

Ada 50 kasus dimana model memprediksi Positif dengan benar (Terlambat)

2. True Negative (TN) : 90

Ada 90 kasus yang memprediksi negative dengan benar (Tidak Terlambat)

3. False Positive (FP) : 0

Tidak ada kesalahan dalam kasus model ini

4. False Negative : 0

Tidak ada kesalahan dalam kasus model ini

Dalam perhitungan yang menghasilkan Accuracy:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \dots\dots\dots(1)$$

$$\text{Accuracy} = \frac{50 + 90}{50 + 90 + 0 + 0} = \frac{140}{140} = 1.0$$

Tingkat accuracy : Accuracy = 1.0 × 100% = 100%

3) K-Fold Validation

Uji coba pada pengolahan data sering kali menggunakan metode K-Fold Validation dengan jumlah fold yang bervariasi dari 2 hingga 10. Menurut para ahli, percobaan yang ekstensif dan pembuktian teoritis menunjukkan bahwa penggunaan 10 K-Fold Validation adalah pilihan terbaik untuk mendapatkan hasil yang akurat dan dapat diandalkan.



Gambar 4 : K-Fold Validation

4) AUC

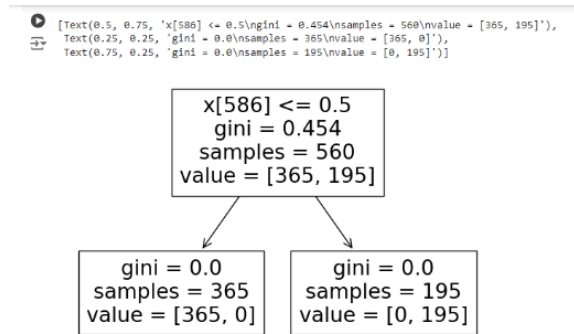
AUC (*Area Under the Curve*) digunakan untuk mengukur kinerja model dalam menganalisis prediktif, khususnya dalam mengevaluasi model klasifikasi.



Gambar 5 : AUC

5) Decision Tree

- Kriteria



Gambar 6 : Kriteria Decision Tree

1. $X[586] \leq 0.5$ \gini = 0.454 \sampil = 560 \value = [365,195]

a. $X[586] \leq 0.5$: jika nilai fitur ini kurang atau sama dengan 0.5, maka akan bercabang ke sebelah kiri. Sedangkan jika nilai fitur ini lebih dari 0.5, maka akan bercabang ke kanan.

b. Gini = 0.454 : menunjukkan nilai gini yang berjumlah 0.454

c. Sampel = 560 : memiliki sampel sebanyak 560.

d. Value = [365,195] : nilai sampel kelas pertama berjumlah 365 dan nilai sampel kelas kedua berjumlah 195.

2. Gini = 0.0 \sampil = 365 \value = [365,0]

a. Gini = 0.0 : menunjukkan nilai pada sampel ini bernilai sama.

b. Sampel = 560 : memiliki sampel sebanyak 560.

c. Value = [365,0] : nilai sampel kelas pertama berjumlah 365 dan nilai sampel kelas kedua berjumlah 195.

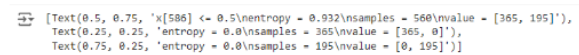
3. Gini = 0.0 \sampil = 365 \value = [0,195]

a. Gini = 0.0 : menunjukkan nilai pada sampel ini bernilai sama.

d. Sampel = 195 : memiliki sampel sebanyak 560.

e. Value = [0,195] : tidak ada sampel dari kelas pertama dan terdapat 195 sampel pada kelas kedua.

- Klasifikasi



Gambar 7 : Klasifikasi

1. $X[586] \leq 0.5$ \entropy = 0.932 \sampil = 560 \value = [365,195]

a. $X[586] \leq 0.5$: nilai fitur ke-586 kurang dari atau sama dengan 0.5

b. Nilai entropy yang berjumlah 0.932. nilai entropy adalah nilai tertinggi yang dihasilkan dalam perhitungan data .

c. Sampel = 560. Sampe yang digunakan berjumlah 560 sampel.

d. Value = [365,195] : dari 560 sampel, 365 sampel dalam kelas pertama dan 195 sampel dalam kelas kedua.

2. Entropy = 0.0\sampel = 365\value = [365,0]

- Entropy 0.0 : menunjukkan bahwa nilai pada sampel ini semuanya bernilai sama
- Sampel = 365 : memiliki sampel sebanyak 365
- Value [365,0] : 365 sampel berasal dari kelas pertama dan tidak ada sampel pada kelas kedua.

3. Entropy = 0.0\sampel = 195\value = [0, 195]

- Entropy = 0.0 : menunjukkan bahwa nilai pada sampel ini semuanya bernilai sama
- Sampel = 195 : memiliki sampel sebanyak 195
- Value = [0,195] : tidak ada sampel dari kelas pertama dan terdapat 195 sampel pada kelas kedua.

5. Kesimpulan

Pada hasil perhitungan yang dilakukan didapatkan nilai tertinggi diperoleh pada perhitungan ke-5 pada K-Fold Validation yang mencapai 100% dan hasil AUC memperoleh nilai sebesar 1.00 yang artinya hasil yang didapatkan sangat bagus. Klasifikasi data mining pada nilai Area Under ROC (AUC) masuk ke dalam bagian Excellent Classification yang berarti nilai yang diperoleh berada pada

bagian 0.90 – 1.00 atau sangat bagus. Dari penelitian yang dilakukan, Decision Tree dapat dimanfaatkan dalam menentukan klasifikasi penyebab keterlambatan pemasangan meter baru PT PLN ULP (Unit Layanan Pelanggan) Pringgabaya. Di mana hasil akurasi yang didapatkan mencapai 100% yang artinya hasil yang didapatkan sangat bagus. Selain itu, dapat disimpulkan bahwa penyebab keterlambatan pemasangan meter baru PT PLN ULP (Unit Layanan Pelanggan) Pringgabaya yang paling mempengaruhi adalah mulai dari : jarak jangkauan, kapasitas egangan, kontak, SLO dan Instalasi. Bukan disebabkan oleh kriteria pelayanan.

Daftar Pustaka

- [1] A. M. Nur, M. F. Wazdi, B. Harianto, and M. F. Zaini, "Implementation of Naive Bayes Algorithm in Analyzing Acceptance of Poor Student Assistance," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Jul. 2020. doi: 10.1088/1742-6596/1539/1/012018.
- [2] B. Andriska, C. Permana, M. Sadali, and R. Ahmad, "Penerapan Model Decision Tree Menggunakan Python Untuk Prediksi Faktor Dominan Penyebab Penyakit Stroke," *Jurnal Informatika dan Teknologi*, vol. 7, no. 1, p. 23, 2024, doi: 10.29408/jit.v7i1.23232.
- [3] Y. Yahya, N. Nurhidayati, and A. Suherman, "Prediksi Tingkat Kesehatan Masyarakat Berdasarkan Penggunaan Alat Kontrasepsi Menggunakan Algoritma Random Forest," *Infotek : Jurnal Informatika dan Teknologi*, vol. 7, no. 1, pp.

- 185–194, Jan. 2024, doi: 10.29408/jit.v7i1.24321.
- [4] M. Terhadap Layanan Akademik Dan Kemahasiswaan, R. Kurniah, D. Yunika Surya Putra, E. Diana, P. Studi Informatika, and U. Prof. Dr. Hazairin, “Penerapan Data Mining Decision Tree Algoritma C4.5 Untuk Mengetahui Tingkat Kepuasan,” *Jurnal Informatika dan Teknologi*, vol. 5, no. 2, p. 316, 2022, doi: 10.29408/jit.v5i2.5910.
- [5] D. Kabupaten, L. Timur Mahfuz, A. Muliawan Nur, and L. M. Samsu, “Penerapan Algoritma C4.5 Dalam Mengklasifikasi Status Gizi Balita Pada Posyandu Desa Dames,” *Jurnal Informatika dan Teknologi*, vol. 5, no. 1, p. 72, 2022, doi: 10.29408/jit.v5i1.4414.
- [6] H. A. S. Yupi Kuspani Putra, “Komparasi Algoritma C4.5 Dan C4.5 Berbasis Pso Untuk Prediksi Jumlah Penggunaan BBM Perbulan Pada Kantor Dinas Lingkungan Hidup Dan Kebersihan Kabupaten Lombok Timur,” 2019.
- [7] “Penerapan Data Mining dalam Mengelompokkan Jumlah Kematian Penderita COVID-19 Berdasarkan Negara di Benua Asia”, doi: 10.31294/p.v2i1i2.
- [8] G. G. I. (2021) Gunawan, “Analisis Perilaku Pengguna Internet Dengan Data Mining Metode K-Means,” 2021.
- [9] M. Romzi and B. Kurniawan, “Pembelajaran Pemrograman Python Dengan Pendekatan Logika Algoritma,” 2020.
- [10] G. I. E. Soen, M. Marlina, and R. Renny, “Implementasi Cloud Computing dengan Google Colaboratory pada Aplikasi Pengolah Data Zoom Participants,” *JITU : Journal Informatic Technology And Communication*, vol. 6, no. 1, pp. 24–30, Jun. 2022, doi: 10.36596/jitu.v6i1.781.
- [11] E. Nofiyanti and E. M. Oki Nur Haryanto, “Analisis Sentimen terhadap Penanggulangan Bencana di Indonesia,” *Jurnal Ilmiah SINUS*, vol. 19, no. 2, p. 17, Jul. 2021, doi: 10.30646/sinus.v19i2.563.
- [12] L. N. Fadhila, I. Kadek, and D. Nuryana, “Teks Ringkas Otomatis pada Portal Berita CNN Indonesia Menggunakan Algoritma Textrank,” 2024.
- [13] A. Thariq and R. Y. Bakti, “Sistem Deteksi Masker dengan Metode Haar Cascade pada Era New Normal COVID-19,” *Jurnal Sistem dan Teknologi Informasi (Justin)*, vol. 9, no. 2, p. 241, Apr. 2021, doi: 10.26418/justin.v9i2.44309.
- [14] L. Rahma, H. Syaputra, A. H. Mirza, and S. D. Purnamasari, “Objek Deteksi Makanan Khas Palembang Menggunakan Algoritma YOLO (You Only Look Once),” 2021.
- [15] S. Bahri and A. Lubis, “METODE KLASIFIKASI DECISION TREE UNTUK MEMPREDIKSI JUARA ENGLISH PREMIER LEAGUE,” vol. 2, no. 1, 2020.
- [16] K. Algoritma Nonparametrik untuk Klasifikasi Citra Wajah Berdasarkan Suku di, S. Hartono, H. Sujaini, and A. Perwitasari, “JEPIN (Jurnal Edukasi dan Penelitian Informatika)”.
- [17] P. B. N. Setio, D. R. S. Saputro, and B. Winarno, “PRISMA, Prosiding Seminar Nasional Matematika Klasifikasi dengan Pohon Keputusan Berbasis Algoritme C4.5,” vol. 3, pp. 64–71, 2020, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [18] I. Teknik, P. Peminatan, J. Siswa, and M. Algoritma, “Indra Gunawan 3) , Iin Parlina 4) , Irawan 5) 1) 3) 4) 5) Teknik Informatika,” vol. 2, no. 2, pp. 1–5, 2021, [Online]. Available:



<http://creativecommons.org/licenses/by/4.0>
/

- [19] Y. Yahya, N. Nurhidayati, and A. Suherman, "Prediksi Tingkat Kesehatan Masyarakat Berdasarkan Penggunaan Alat Kontrasepsi Menggunakan Algoritma Random Forest," *Infotek : Jurnal Informatika dan Teknologi*, vol. 7, no. 1, pp. 185–194, Jan. 2024, doi: 10.29408/jit.v7i1.24321.