

IMPLEMENTASI METODE RANDOM FOREST BERBASIS WEBSERVICES API PPoE UNTUK KLASIFIKASI PELANGGAN INTERNET DIAL UP

Fajar Ahmad Raharjo, Dian Asmarajati, Iman Ahmad Ihsannuddin

Teknik Informatika, Fakultas Teknik dan Ilmu Komputer

Email : ahmadraharjo40@gmail.com

Diterima : 31 Juli 2024 ; Disetujui : 25 Agustus 2024 ; Dipublikasikan : 31 Januari 2025

ABSTRAK

Internet merupakan kebutuhan primer bagi setiap orang untuk kegiatan sehari-hari baik berupa hiburan atau pekerjaan. Salah satu sumber mendapatkan internet ialah dari penyedia Internet Service Provider. Internet Service Provider atau ISP sekarang berlomba untuk melakukan layanan terbaik ke pelanggan. Penelitian ini dilakukan untuk menentukan klasifikasi pelanggan internet pada CV. Tabung Creative Digital sebagai ISP. Dalam menentukan klasifikasi, metode yang digunakan yaitu menggunakan Metode Random Forest. Dengan menggunakan Metode Random Forest membutuhkan atribut untuk melakukan perhitungan probabilitas, entropy dan gain. Dalam penelitian ini atribut yang digunakan yaitu profile, keadaan sosial dan nilai redaman. Klasifikasi dilakukan dengan cara merangkum dataset tersedia, kemudian dataset dirangkum berdasarkan klasifikasi training, kemudian menghitung entropy dari setiap atribut diperoleh nilai serta menghitung gain dari sub atribut kemudian diperoleh rule dari gain terbesar ke terkecil untuk melakukan klasifikasi. Selanjutnya klasifikasi diujikan dengan data riil dengan hasil uji system. Hasil dari klasifikasi tersebut kemudian diimplementasikan kedalam sistem dan menjadi sistem klasifikasi pelanggan internet serta diuji dengan blackbox supaya analisa dan perancangan sesuai sehingga dapat digunakan

Kata Kunci : *Random Forest, Pelanggan Internet, Gain, Entropy, Blackbox.*

ABSTRACT

The internet is a primary need for everyone for daily activities, whether in the form of entertainment or work. One source of getting internet is from an Internet Service Provider. Internet Service Providers or ISPs are now competing to provide the best service to customers. This research was conducted to determine the classification of internet customers at CV. Tabung Creative Digital as ISP. In determining the classification, the method used is the Random Forest Method. By using the Random Forest Method attribute is required to do probability, entropy and gain. In this research, the attributes used are profile, social situation and attenuation value. Classification is carried out by summarizing the available dataset, then the dataset is summarized based on the training classification, then calculating the entropy of each attribute obtained by value and calculating the gain of the sub-attributes then obtaining a rule from the largest gain to the smallest to carry out classification. Next, the classification is tested with real data with system test results. The results of the classification are then implemented into the system and become an internet customer classification system and tested using a black box so that the analysis and design are appropriate so that they can be used.

Keywords : *Random Forest, Internet Customers, Gain, Entropy, Blackbox.*

1. PENDAHULUAN

Penggunaan jaringan komputer sangat lazim ditemui di kehidupan sehari-hari bahkan di zaman sekarang internet merupakan kebutuhan primer manusia, dengan koneksi internet dapat memperoleh segala hal baik informasi, hiburan bahkan pekerjaan yang sangat menunjang kelangsungan hidup manusia [1]. Menurut lembaga riset MarkPlus Insight, angka jumlah pengguna Internet di Indonesia menembus 100 Juta jiwa ditahun 2015. Semakin banyak pengguna internet di kehidupan-masyarakat sehingga perlu adanya protokol dan metode untuk mengatur meminimalkan resiko masalah dilapangan sampai manajemen traffic internet yang beredar,hal ini berbanding lurus dengan semakin kompetitif tantangan bagi perusahaan penyedia internet (ISP) bereksplorasi mencari cara untuk melakukan diferensiasi dari penyedia lain meningkatkan ekstensif kualitas pelayanan [2]. Internet dial up merupakan koneksi yang umum digunakan oleh perusahaan penyedia internet atau ISP yang menggunakan kabel telepon atau kabel internet untuk media koneksi timbal balik dari sisi server ke sisi user menggunakan modem sesuai pemanggilan (dial) dari user menuju ke server (up) [3].

Menurut [4], Metode Point to Point Over Ethernet (PPoE) digunakan membangun jaringan VPN dimana koneksinya menggunakan point to point tunnel. PPoE dimulai dari inisiasi sesi sampai penetapan parameter PPP selalu membutuhkan metode dial up. Kedua unsur berkaitan satu sama lain, dial up metode pemanggilan sampai metode akses internet yang sangat berperan didalam koneksi PPoE dial-up. PPoE merupakan salah satu protokol jaringan yang sangat dikenal dikalangan publik untuk menciptakan sambungan antar komputer yang berlaku sebagai user dan server untuk kemudian disambungkan dengan autentikasi username dan password sekaligus sebagai pengatur traffic yang beredar. CV Tabung Creative Digital merupakan badan usaha yang bergerak di bidang Informasi dan Teknologi yang mulai beroperasi sejak 2020 silam, salah satu usaha yang dilakukan CV. Tabung Creative Digital meliputi jasa jual kembali akses internet. Didalam penelitian ini menggunakan algoritma random forest untuk mencari keputusan terbaik tindakan dilapangan pada CV. Tabung Creative

Digital ketika terjadi masalah sehingga pelayanan perusahaan dapat optimal ke pengguna dan efektif pembiayaan operasional, random forest memiliki prosedur yang sangat mudah untuk dipahami dan diimplementasikan dengan sistem PPoE dial-up.

Penelitian PPoE dial-up di CV. Tabung Creative Digital berkolaborasi dengan algoritma Random Forest (Three Tree) diharapkan bermanfaat membantu perusahaan dilapangan atau penggiat internet pada umumnya melakukan analisis sebelum melakukan tindakan perbaikan, monitoring bahkan manajemen internet yang digunakan khalayak umum. Beberapa aspek penunjang penelitian yaitu faktor koneksi, keadaan Ekonomi/Sosial pengguna bahkan faktor wilayah bisa menjadi pohon keputusan untuk menentukan Three tree dari algoritma Random Forest.

2. METODE

a. Metode Pengumpulan Data

Dalam proses pengumpulan data yang digunakan untuk penelitian,. metode yang digunakan yaitu :

1) Wawancara

Wawancara dilakukan bersama dengan narasumber yaitu Bapak Iman Ahmad Ihsanudin selaku pemilik dari CV Tabung Creative Digital. Wawancara dilakukan untuk memperoleh informasi – informasi serta data yang dibutuhkan dalam pembuatan sistem klasifikasi pelanggan internet.

2) Studi Literatur

Studi literatur dilakukan untuk mendapatkan konsep yang berkaitan dengan penerapan metode *Random Forest*, *Webservices* API PPoE dan Dial-up baik menelaah buku, jurnal maupun data di Internet.

b. Metode Pengembangan Sistem

Object Oriented Analysis Design (OOAD) dipilih dalam melakukan analisis dan desain sistem, selanjutnya Object Oriented Database dan Object Oriented Programming digunakan dalam mengimplementasikan hasil dari analisis menjadi sistem klasifikasi pelanggan internet.

c. Metode *Random Forest*

Random Forest merupakan algoritma klasifikasi suatu obyek atau entitas, dengan mempelajari data yang sudah ada kemudian ditentukan aturan sebelum ditentukan

menjadi kelompok atau klasifikasi baru dengan metode DecisionTree [5]. Jawaban pertama atau data awal akan berproyeksi mempengaruhi data berikutnya. Proses Random Forest sama dengan algoritma pada umumnya dimulai dari pengumpulan data, pengolahan data dan penentuan aturan, implementasi serta analisis, adapun implementasi secara detail sebagai berikut :

- 1) Penentuan Node
- 2) Pembentukan *Tree*
- 3) Penentuan *Rule* atau aturan
- 4) Penentuan Hasil *Vote*
- 5) Perhitungan *Vote*
- 6) Pengujian atau analisis

Persamaan umum dari perhitungan Random Forest adalah sebagai berikut :

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \tag{1}$$

Dimana :

- S : Himpunan Kasus
- N : Jumlah partisi S
- Pi : Proporsi dari Si terhadap S

3. HASIL DAN PEMBAHASAN

a. Tahapan *Random Forest*

Dalam melakukan prediksi, metode yang digunakan penulis yaitu menggunakan metode random forest, adapun prosedur yang ditempuh dalam perhitungan menentukan klasifikasi beserta hasil adalah sebagai berikut:

- 1) Menentukan atribut perhitungan dan dataset yang digunakan dalam melakukan klasifikasi. Atribut yang digunakan seperti yang tertera pada tabel 1.

$$Gain(Profile) = 0,987138 - \left(\left(\frac{6}{30} \times 1 \right) + \left(\frac{8}{30} \times 0,811278124 \right) + \left(\frac{8}{30} \times 0,811278124 \right) + \left(\frac{8}{30} \times 0,811278124 \right) \right) = 0.138115275$$

Tabel 1. Atribut Perhitungan

Atribut Perhitungan		
Nama Atribut	Keterangan	
Profile	± 5 Mbps	High Priority
	± 10 Mbps	Medium Priority
	± 15 Mbps	Low Priority
	± 20 Mbps	Low Priority
Pekerjaan/Keadaan Ekonomi	Wiraswasta	High Priority
	Pedagang	Medium Priority
	Pegawai Negeri Sipil	Low Priority
	Guru/Dosen	Medium Priority
Jenis Wilayah/Redaman	Petani	High Priority
	>-27 dBM	High Priority
	>=-27 dBM && <-8 dBM	Medium Priority
	>=-17 dBM && <-8 dBM	Low Priority

Dataset yang digunakan seperti pada tabel 2 berikut

Tabel 2 Dataset

Dataset				
No	Profile (Mbps)	Keadaan Sosial	Redaman/Jenis Wilayah	Prioritas
1	5	Wiraswasta	-10	Ya
2	10	Pedagang	-20	Tidak
3	20	Petani	-8	Ya
4	15	PNS	-30	Ya
5	10	Guru/Dosen	-15	Tidak
6	15	Pedagang	-25	Tidak
7	20	Wiraswasta	-8	Ya
8	20	Petani	-27	Ya
9	5	Pedagang	-8	Tidak
10	15	Wiraswasta	-17	Tidak
11	10	Pedagang	-12	Tidak
12	20	Petani	-30	Ya
13	15	Pedagang	-10	Tidak
14	5	Wiraswasta	-27	Ya
15	15	PNS	-8	Tidak
16	10	Guru/Dosen	-22	Tidak
17	20	Pedagang	-17	Tidak
18	5	Petani	-20	Ya
19	15	Wiraswasta	-12	Ya
20	10	Pedagang	-30	Ya
21	20	Petani	-10	Ya
22	15	PNS	-25	Tidak
23	10	Guru/Dosen	-18	Tidak
24	5	Pedagang	-27	Tidak
25	20	Wiraswasta	-15	Ya
26	10	Guru/Dosen	-20	Tidak
27	15	Guru/Dosen	-20	Tidak
28	10	Petani	-18	Ya
29	5	Guru/Dosen	-15	Tidak
30	20	Guru/Dosen	-15	Tidak

- 2) Dataset yang didapatkan kemudian diambil sebagai data training dalam melakukan prediksi klasifikasi pelanggan sesuai yang tertera pada tabel diatas.
- 3) Setelah data training didapatkan, selanjutnya menghitung class probabilitas, Prosedur pertama yang ditempuh meringkas kasus dan menghitung jumlah probabilitas.

Tabel 3 Menghitung Probabilitas

Probabilitas	Jml Kasus	Prioritas	
		Tidak	Ya
Total	30	17	13
Profile			
± 5 Mbps	6	3	3
± 10 Mbps	8	6	2
± 15 Mbps	8	6	2
± 20 Mbps	8	2	6
Keadaan Sosial			
Wiraswasta	6	1	5
Pedagang	8	7	1
PNS	3	2	1
Guru/Dosen	7	7	0
Petani	6	0	6
Jenis Wilayah/Nilai redaman			
>-27 dBM	3	0	3
> -8 dBM && <-27 dBM	27	17	10
> -8 dBM && <-17 dBM	15	9	6

4) Menghitung jumlah entropy total dan turunannya.

$$Entropy(\text{Total}) = \left(-\frac{17}{30} \times \log_2\left(\frac{17}{30}\right)\right) - \left(\frac{13}{30} \times \log_2\left(\frac{13}{30}\right)\right)$$

$$Entropy(\text{Total}) = \left(-\frac{17}{30} \times -0,81943\right) - \left(\frac{13}{30} \times -1,20645\right)$$

$$Entropy(\text{Total}) = 0,464342 + 0,522795$$

$$Entropy(\text{Total}) = 0,987138$$

Tabel 4 Data Entropy

Probabilitas	Jml Kasus	Prioritas		ENTROPY
		Tidak	Ya	
Total	30	17	13	0.987137774
Profile				
± 5 Mbps	6	3	3	1
± 10 Mbps	8	6	2	0.811278124
± 15 Mbps	8	6	2	0.811278124
± 20 Mbps	8	2	6	0.811278124
Keadaan Sosial				
Wiraswasta	6	1	5	0.650022422
Pedagang	8	7	1	0.543564443
PNS	3	2	1	0.918295834
Guru/Dosen	7	7	0	0
Petani	6	0	6	0
Jenis Wilayah/Nilai redaman				
>-27 dBM	3	0	3	0
> -8 dBM && <-27 dBM	27	17	10	0.950956048
> -8 dBM && <-17 dBM	15	9	6	0.970950594

5) Setelah nilai entropy diketahui maka mencari nilai gain dari setiap entropy untuk menentukan rule.

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

$$Gain(S, A) = Entropy(\text{Total}) - \sum_{i=1}^n \frac{|Profile|}{|Total|} \times Entropy(\text{Profile}),$$

6) Menentukan nilai gain terbesar untuk menentukan pelanggan prioritas.

Tabel 5 Data Gain

Probabilitas	Jml Kasus	Prioritas		ENTROPY	GAIN
		Tidak	Ya		
Total	30	17	13	0.987137774	
Profile					0.138115275
± 5 Mbps	6	3	3	1	
± 10 Mbps	8	6	2	0.811278124	
± 15 Mbps	8	6	2	0.811278124	
± 20 Mbps	8	2	6	0.811278124	
Keadaan Sosial					0.620353188
Wiraswasta	6	1	5	0.650022422	
Pedagang	8	7	1	0.543564443	
PNS	3	2	1	0.918295834	
Guru/Dosen	7	7	0	0	
Petani	6	0	6	0	
Jenis Wilayah/Nilai redaman					0.354197966
>-27 dBM	3	0	3	0	
> -8 dBM && <-27 dBM	27	17	10	0.950956048	
> -8 dBM && <-17 dBM	15	9	6	0.970950594	

7) Setelah semua probabilitas dan rule setiap atribut dihitung, selanjutnya melakukan prediksi dengan menggunakan data testing sebagai berikut:

Tabel 6 Data testing

Profile (Mbps)	Keadaan Sosial	Jenis Wilayah/Nilai Redaman
10	Wiraswasta	-16

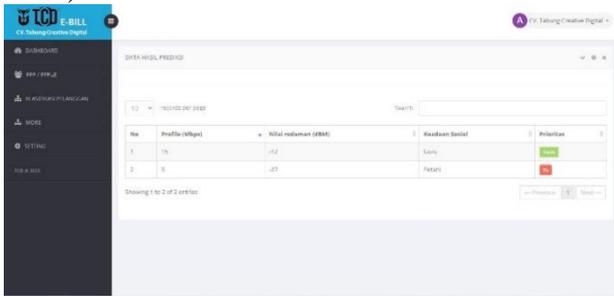
Berdasarkan data testing pada tabel 6 maka didapatkan perhitungan untuk setiap nilai atributnya sebagai berikut.

- DecisionTree berdasarkan nilai gain terbesar diperoleh Keadaan Sosial, Profile dan Nilai redaman.
- Berdasarkan data testing tabel diperoleh Profile 10 Mbps, Keadaan Sosial Wiraswasta dan nilai redaman -16 dBm.
- Keadaan Sosial wiraswasta bernilai > 0 maka bernilai belum pasti dibandingkan dengan gain selanjutnya.
- Profile 10 Mbps bernilai > 0 maka bernilai belum pasti dibandingkan dengan gain selanjutnya.
- Nilai Redaman -16 maka bernilai > 0 maka bernilai Tidak untuk pelanggan prioritas
- Kesimpulan data testing pada tabel 1.6 merupakan pelanggan dengan status "Tidak" Prioritas.

b. Implementasi Sistem

Berikut merupakan implementasi dari sistem yang dibuat

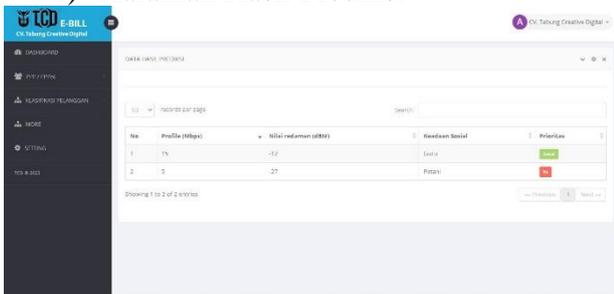
1) Halaman Proses Prediksi



Gambar 1. Halaman Prediksi

Pada halaman ini, terdapat step by step atau prosedur didalam proses perhitungan yang dimulai dari perhitungan *Entropy* dan *Gain* sampai penentuan *rule* serta implementasi data testing.

2) Halaman Hasil Prediksi



Gambar 2 Halaman Hasil Prediksi

Pada halaman ini memuat hasil prediksi dari data pelanggan yang telah di uji coba atau di *update* status prioritasnya.

4. PENTUTUP

a. Kesimpulan

Kesimpulan yang dapat diambil dari laporan “Implementasi Metode Random Forest Berbasis Webservices API PPOE untuk klasifikasi Pelanggan Internet Dial Up” adalah sebagai berikut :

1. Klasifikasi pelanggan internet berbasis Webservices API PPOE menggunakan metode random forest dengan tahapan menentukan dataset, menentukan data testing, menghitung probabilitas, menghitung nilai entropy dari setiap atribut, menghitung gain, menentukan decision tree, menghitung probabilitas antara data actual dengan data prediksi serta mengambil kesimpulan dari implementasi klasifikasi dengan random forest.

2. Dalam melakukan perhitungan klasifikasi, atribut yang dipakai yaitu profile, keadaan sosial dan nilai redaman.
3. Proses prediksi dengan menghitung probabilitas setiap nilai atribut, probabilitas 5 Mbps, 10 Mbps, 15 Mbps, 20 Mbps untuk atribut profile, atribut keadaan sosial meliputi Wiraswasta, Guru/Dosen, Petani, Pedagang dan PNS, serta untuk Nilai redaman <-27 dBm, > -8 dBm && <=-27 dBm dan > -8 dBM && <-17 dBM.
4. Pengujian menggunakan black box menunjukkan bahwa sistem implementasi sudah sesuai dengan hasil perancangan.

b. Saran

Dalam melakukan penelitian dan pembuatan Sistem Implementasi Klasifikasi Pelanggan Internet berbasis Webservices API PPOE Menggunakan Random Forest penulis menyadari masih adanya kekurangan, baik dalam penulisan laporan maupun dalam pembuatan implementasi. Adapun saran yang diharapkan peneliti untuk penelitian kedepan yaitu :

1. Terdapat metode lain dalam melakukan prediksi klasifikasi lain sehingga hasil klasifikasi lebih akurat .Klasifikasi Pelanggan Internet dial up menggunakan metode *random forest* dapat ditambahkan lagi metode prediksi lainnya sehingga dapat lebih berkembang dan tidak mengacu pada 1 prediksi saja.
2. Terdapat data atribut bersifat manual untuk proyeksi kedepan pengambilan bisa dilaksanakan secara riil dengan OLT atau modular yang terhubung, adapun yang dimaksud nilai atribut redaman.
3. Penambahan dataset untuk training model sehingga data yang diperoleh lebih matang.

5. DAFTAR PUSTAKA

[1] FAUZI, Aditya Ahmad, et al. Pemanfaatan Teknologi Informasi di Berbagai Sektor

- Pada Masa Society 5.0. PT. Sonpedia Publishing Indonesia, 2023.
- [2] FEBRIAN, Ahmad Dandi; DARMAWAN, Redi. Implementasi Jaringan Komputer Berbasis Virtual LAN untuk Layanan Iconnet VIP Pada Jaringan MPLS (Multi Protocol Label Switching): Studi Kasus di PT Indonesia Comnets Plus. *Scientia Sacra: Jurnal Sains, Teknologi dan Masyarakat*, 2022, 2.3: 466-480.
- [3] ISKANDAR, Eggy. MEMBANGUN RT/RW NET SEBAGAI PORTAL INFORMASI DAN LAYANAN INTERNET MASUK DESA SUKAHURIP. 2021. PhD Thesis. Universitas Siliwangi.
- [4] WARMAN, Indra; HANAFAI, Ahmad. Analisa Perbandingan kinerja Generic routing encapsulation (GRE) tunnel dengan point to point protocol over ethernet (PPPoE) tunnel mikrotik routers. *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang*, 2019, 7.1: 58-66.
- [5] PAMUJI, Fandi Yulian; RAMADHAN, Viry Puspaning. Komparasi Algoritma Random Forest dan Decision Tree untuk Memprediksi Keberhasilan Immunotherapy. *Jurnal Teknologi dan Manajemen Informatika*, 2021, 7.1: 46-50.