

SISTEM REKOMENDASI KELAYAKAN *LINK* SCADA MENGGUNAKAN *DECISION TREE C4.5*

Taufiq Dwi Cahyono¹, Rizki Eka Pratama²

^{1,2}Fakultas Teknologi Informasi dan Industri, Universitas Stikubank

^{1,2}Jurusan Teknik Elektro, Universitas Semarang

e-mail: ¹taufiq_dc@edu.unisbank.ac.id, ²rizki.takiata1@gmail.com

ABSTRAK

SCADA 20 KV merupakan layanan pengatur beban di Gardu Induk (GI), Infrastruktur jaringan SCADA 20 KV terdiri dari beberapa komponen yang sangat rentan terhadap gangguan, yang dapat mengakibatkan hilangnya kontrol pada suatu GI. Oleh karena itu, sebuah tautan cadangan telah dibuat dengan tujuan untuk menyediakan cadangan jika terjadi gangguan pada tautan utama. Tautan cadangan ini akan menggantikan tautan utama yang mengalami gangguan, dan sebaliknya. Pentingnya sebuah tautan SCADA yang dapat diandalkan memunculkan kebutuhan akan sistem rekomendasi guna menilai kelayakan tautan cadangan SCADA 20 KV. Algoritma Decision Tree C4.5 digunakan untuk menyusun rekomendasi ini. Sistem hasil keputusan yang dihasilkan sangat terperinci dan dapat diuraikan melalui representasi pohon keputusan. Hasil rekomendasi akan menunjukkan apakah tautan tersebut layak atau tidak layak untuk digunakan.

Kata Kunci: Klasifikasi, decision tree, SCADA, gardu induk

1. PENDAHULUAN

Dalam era industri 4.0, keberlanjutan operasional dan efisiensi sistem kendali otomatis menjadi krusial untuk menjaga kelancaran proses produksi. Sistem Kontrol Supervisory and Data Acquisition (SCADA) memainkan peran sentral dalam memastikan integritas, kinerja, dan keamanan operasional suatu fasilitas. Dengan semakin kompleksnya jaringan dan perangkat yang terhubung dalam suatu sistem SCADA, pentingnya memiliki mekanisme yang efektif untuk mengelola dan merekomendasikan kelayakan link SCADA menjadi semakin mendesak.

Pada suatu tautan SCADA 20 KV, terdapat beberapa komponen yang dapat memengaruhi validitas suatu tautan cadangan, yang harus dipertimbangkan dalam mengevaluasi kelayakannya. Oleh karena itu, dalam penelitian ini, metode yang diterapkan menggunakan teknik data mining dipilih. Klasifikasi merupakan teknik data mining yang mampu memetakan target [1], [2]. Penelitian ini bertujuan memberikan rekomendasi mengenai kelayakan tautan cadangan pada layanan SCADA 20 KV berdasarkan atribut yang telah ditentukan, dengan menerapkan algoritma tertentu. Algoritma yang digunakan dalam penelitian ini adalah Decision Tree C4.5. Keunggulan algoritma C4.5 terletak pada kemudahannya dalam penggunaan dan kemampuannya menghasilkan pohon keputusan yang sederhana. Dengan menggunakan algoritma C4.5, peneliti berharap bahwa proses penelitian ini akan menghasilkan analisis dengan tingkat akurasi yang tinggi dan rekomendasi yang tepat, berdasarkan parameter dari atribut yang telah ditetapkan [3].

Dalam penelitian ini, peneliti menggunakan metode yang menerapkan Data Mining. Klasifikasi, sebagai cara atau teknik dalam data mining, digunakan untuk menentukan rekomendasi kelayakan tautan cadangan pada layanan SCADA 20 KV berdasarkan atribut yang telah ditetapkan, dengan menerapkan algoritma. Algoritma yang diterapkan dalam penelitian ini adalah Decision Tree C4.5. Pemilihan algoritma C4.5 dipilih karena kemudahannya dalam penggunaan dan kemampuannya menghasilkan pohon keputusan yang simpel [4]. Dengan menggunakan algoritma C4.5, peneliti berharap bahwa proses penelitian ini akan memberikan analisis dengan tingkat akurasi yang tinggi dan rekomendasi yang tepat, sesuai dengan parameter dari atribut yang telah ditentukan.

2. TINJAUAN PUSTAKA

Klasifikasi dalam data mining adalah suatu proses di mana data dibagi ke dalam kategori atau kelas berdasarkan atribut-atribut tertentu. Tujuan utama klasifikasi adalah untuk mengelompokkan data ke dalam kategori yang telah ditentukan sebelumnya, sehingga dapat memprediksi kelas atau label dari data yang baru atau belum dikenal [1], [5].

Proses klasifikasi melibatkan penggunaan algoritma pembelajaran mesin yang memanfaatkan pola atau karakteristik dari data pelatihan untuk membangun model klasifikasi. Model ini kemudian dapat digunakan untuk

mengklasifikasikan data baru ke dalam kategori yang sesuai. Algoritma C4.5 yang di gunakan untuk menentukan tingkat kelayakan sebagai target [6], [7]

Implementasi Algoritma C4.5 dalam penentuan beasiswa bertujuan agar mahasiswa berprestasi yang akan mendapatkan beasiswa lebih tepat sasaran , selain ipk yang bagus penghasilan orang tua juga berpengaruh dalam penentuan mahasiswa yang akan menerima beasiswa tersebut.

Dalam pengembangan sistem rekomendasi pemilihan jurusan untuk calon siswa Sekolah Menengah, disajikan model menggunakan decision tree dengan pendekatan penelitian dan pengembangan (RND). Variabel input yang dimasukkan dalam penelitian ini meliputi minat, bakat, prestasi akademis, nilai ujian nasional, dan jenis kelamin. Proses analisis data penelitian dilakukan dengan membandingkan hasil output sistem, menunjukkan kemampuan sistem memberikan peringkat seajar sebagai data uji. Hasil pengujian sistem menunjukkan bahwa sistem dapat memberikan rekomendasi yang akurat sebesar 83,33% dari 48 data uji [3].

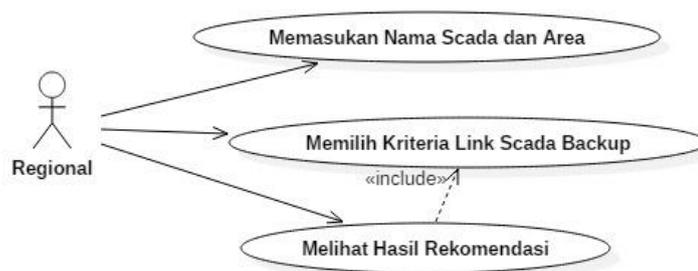
Pentingnya mendeteksi gangguan autisme pada perkembangan anak menjadi fokus utama, dengan mayoritas orang tua masih kurang memahami kondisi autisme yang mungkin dialami oleh anak-anak mereka, serta kurang familiar dengan cara penanganannya. Konsultasi dengan dokter atau pakar seringkali memerlukan biaya yang signifikan, sehingga diperlukan cara atau metode yang dapat membantu orang tua dalam mendeteksi dini gangguan autisme pada anak-anak mereka. Salah satu metode yang digunakan adalah algoritma C4.5 dalam data mining untuk memprediksi tingkat akurasi, dan hasil penelitian ini menunjukkan tingkat akurasi sebesar 72%.

Decision tree Merupakan alat yang dapat mendukung keputusan menggunakan model keputusan yang berbentuk seperti akar pohon .Decision tree dapat memetakan berbagai alternatif yang memungkinkan untuk mengatasi suatu masalah, dan terdapat juga faktor-faktor kemungkinan yang dapat mempengaruhi alternative tersebut beserta estimasi akhirnya jika memiliki alternative yang ada. Decision tree merupakan salah satu metode yang bisa digunakan untuk menampilkan algoritma dimana hanya berisi pertanyaan control yang bersyarat [8], [9], [10].

Algoritma C4.5 adalah sebuah model yang dapat membentuk pohon keputusan dari data yang telah ada sebelumnya. Ini termasuk dalam kategori algoritma decision tree dan merupakan pengembangan dari algoritma ID3. Dalam penerapannya, algoritma C4.5 memproses komponen pemrosesan dengan memberikan data untuk dipelajari. Proses pembelajaran ini melibatkan pengolahan dataset uji yang disebut sebagai dataset test. Hasil dari pemrosesan dataset test ini menghasilkan beberapa kelompok data yang dikelompokkan ke dalam beberapa kelas.

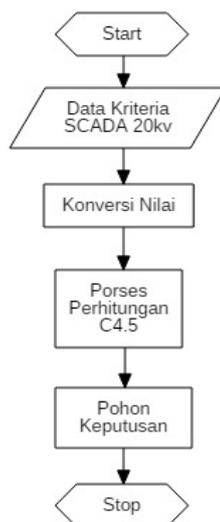
3. METODE PENELITIAN

Implementasi *Decision Tree* C4.5 pada sistem rekomendasi kelayakan link backup SCADA 20kv diharapkan dapat membantu untuk menentukan kelayakan link yang akan di bangun maupun yang sudah ada, dengan kriteria yang sudah di tentukan. Pada sistem kali ini regional atau admin di sebuah kantor cabang akan menginputkan data yang di dapat dari lapangan dan engineer, setelah itu data akan di inputkan kedalam sistem rekomendasi yang hasilnya akan di data dan menjadi *improvement* bagi perusahaan. Berikut adalah usecase analisa sistem :



Gambar 1. Analisa Sistem

Terdapat beberapa langkah yang harus dilalui untuk menghasilkan pohon keputusan, dimulai dari proses pengolahan data mentah, konversi nilai atribut, perhitungan menggunakan algoritma C4.5 dan pada akhirnya, pohon keputusan dapat diperoleh. Sebuah flowchart telah disusun untuk memvisualisasikan rangkaian proses yang diterapkan dalam perhitungan Decision Tree C4.5.



Gambar 2. Flowchart Perhitungan

Tahap setelah data dimasukkan kedalam sistem rekomendasi maka data akan di hitung menggunakan algoritma c4.5. rumus untuk menghitung entropy pada algoritma C4.5 ditunjukkan pada rumus 1 dan perhitungan gain akan di tunjukan pada persamaan (2) :

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

Keterangan :

S = Himpunan (dataset) kasus k = banyaknya partisi S

Pi = probabilitaas yang didapat dari Sum(Ya) atau Sum(Tidak) dibagi total kasus

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

Keterangan :

S = Himpunan

A = Atribut

n = Jumlah partisi atribut A

| Si | = Jumlah kasus pada partisi ke-i

| S | = Jumlah kasus dalam S

Beberapa atributt atau komponen dalam sistem rekomendasi kelayakan scada 20kv sebagai berikut:

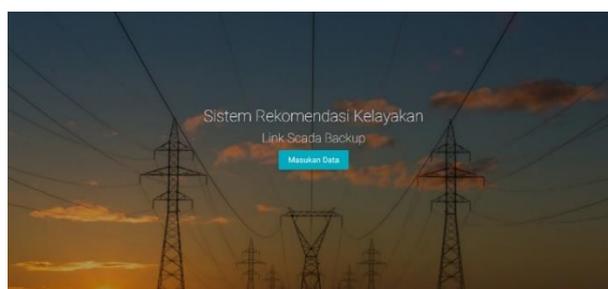
- a. Nama Scada
Nama scada merupakan nama dari sebuah site scada 20kv.
- b. Area
Area merupakan wilayah dimana secada 20kv tersebut di kelompokkan.
- c. Backup Power Layer 2
Backup power layer 2 merupakan cadangan daya yang akan memberikan daya ke perangkat ketika ada PLN *off*.
- d. Layer 3
Layer 3 merupakan gateway dari beberapa layanan yang melewati layer 2 (fisik).
- e. Jalur Kabel
Jalur kabel merupakan wujud fisik dari jalur akses user ke server berupa serat optic.
- f. Konverter
Konverter merupakan perubahan sinyal data ke sinyal cahaya atau sebaliknya agar perangkat user dan server dapat berkomunikasi.

Beberapa kriteria yang akan dimasukan oleh admin yang akan menyeleksi kelayakan link SCADA 20Kv dan datanya akan di olah oleh pusata untuk improvement.

4. HASIL DAN PEMBAHASAN

4.1 Halaman Utama

Pada bagian pembahasan akan dijelaskan tahapan dalam menjalankan aplikasi program yang sudah di buat, tahapan yang akan dijelaskan adalah mengenai masuk ke menu memasukan data, pengisian data, dan sampai ke hasil rekomendasi. Pada langkah awal aplikasi sistem rekomendasi kelayakan SCADA 20Kv akan menuju ke menu utama, yang di dalamnya memiliki tombol untuk memasukan data.



Gambar 3. Halaman Utama

4.2 Halaman Pengisian Data

Pada halaman pengisian data atau form didalamnya terdapat beberapa form yaitu nama scada yang merupakan nama dari sercada, area merupakan area dari secada, backup power merupakan cadangan daya yang terdapat pada ruang server yang membackup power ketika ada PLN off, jalur kabel merupakan jalur kabel optical yang menghubungkan antara 2 perangkat dari sisi POP dan latmile (sisi user), layer 3 atau UPE merupakan gateway dari perangkat layer 2 yang terdapat di POP, konverter merupakan alat untuk membantu mengubah sinyal optical agar dapat di baca oleh perangkat, dan ada tombol proses untuk melakukan pengolahan data dan akan menuju ke halaman rekomendasi atau hasil.

Gambar 4. Halaman Pengisian Data Pada Form

4.3 Halaman Hasil Rekomendasi

Pada halaman hasil rekomendasi, data yang sudah di masukan sebelumnya akan di olah menggunakan algoritma decision tree C4.5 yang akan menghasilkan rekomendasi kelayakan link backup secada 20kv, pada halaman hasil rekomendasi berisi nama scada, area scada, urutan nomor, parameter yang sudah di input, keterangan yang sudah di masukan, hasil dari rekomendasi algoritma decision tree C4.5

No	Parameter	Keterangan
1	Backup Power	Genesi
2	Layer 3	Tidak Sama Dengan Main
3	Jalur Kabel	Tidak Satu Jalur Dengan Main
4	Konverter	SPP10G

Hasil Rekomendasi Link Backup
Berdasarkan Parameter Yang Sudah Anda Masukan, Hasil Rekomendasi Link Backup Layak Untuk Di Cerdikan

Gambar 5. Halaman Hasil Rekomendasi

4.4 Perhitungan Manual Decision Tree C4.5

Pada proses perhitungan manual membutuhkan data yang harus di olah, berikut data dan parameter yang akan di hitung:

Tabel 1. Konversi Nilai

Nilai	Peringkat Keterangan
Menggunakan Genset Layer 3 Tidak Sama Dengan Main Jalur Kabel Tidak Satu Jalur Dengan Main Konverter SFP10G	Ya
Menggunakan Baterai Layer 3 Satu Kaki Dari Main Jalur Kabel Melewati Jalur Main Konverter SFP1G	Mungkin
Tidak Menggunakan Backup Layer 3 Sama Dengan Main Jalur Kabel Satu Jalur Dengan Main Konverter Telways	Tidak

Tabel 2. Data Scada

No	Scada Backup	Backup Power Layer 2 (CPE/Switch)	Layer 3 (UPE/Router)	Jalur Kabel	Konverter
1	SCADA A	Genset	Satu kaki dari main	Melewati Jalur Main	SFP10G
2	SCADA B	Baterai	Tidak sama dengan main	Tidak Satu Jalur Dengan Main	SFP1G
3	SCADA C	Baterai	Satu kaki dari main	Melewati Jalur Main	SFP1G
4	SCADA D	Baterai	Tidak sama dengan main	Tidak Satu Jalur Dengan Main	SFP1G
5	SCADA E	Genset	Tidak sama dengan main	Tidak Satu Jalur Dengan Main	SFP10G

Setelah menyelesaikan pengelompokan data ke dalam beberapa atribut, tindakan berikutnya melibatkan transformasi nilai parameter agar memiliki bobot individualnya.

Tabel 3. Data Scada Dikonversi

No	Scada Backup	Backup Power Layer 2 (CPE/Switch)	Layer 3 (UPE/Router)	Jalur Kabel	Konverter
1	SCADA A	Layak	Mungkin	Mungkin	Layak
2	SCADA B	Mungkin	Layak	Layak	Mungkin
3	SCADA C	Mungkin	Mungkin	Mungkin	Mungkin
4	SCADA D	Mungkin	Layak	Layak	Mungkin
5	SCADA E	Layak	Layak	Layak	Layak

Setelah data SCADA telah diberi bobot, langkah selanjutnya adalah melakukan perhitungan nilai entropi untuk setiap data, seperti yang dijelaskan dalam perhitungan berikut:

$$\text{Total Kasus} = -(70/85) * \log_2 (70/85) + -(15/85) * \log_2 (15/85) = 0,672294817$$

Nilai Backup Power Layer 2 Layak 2
 $= (-26/28) * \log_2 (26/28) + (-2/28) * \log_2 (2/28) = 0,371232327$
 Nilai Backup Power Layer 2 Mungkin
 $= (-41/47) * \log_2 (41/47) + (-6/47) * \log_2 (6/47) = 0,550984443$
 Nilai Backup Power Layer 2 Tidak Layak
 $= (-3/10) * \log_2 (3/10) + (-7/10) * \log_2 (7/10) = 0,881290899$
 Nilai Layer 3 Layak
 $= (-52/54) * \log_2 (52/54) + (-2/54) * \log_2 (2/54) = 0,228538144$
 Nilai Layer 3 Mungkin
 $= (-17/28) * \log_2 (17/28) + (-11/28) * \log_2 (11/28) = 0,966618633$
 Nilai Layer 3 Tidak Layak
 $= (-1/3) * \log_2 (1/3) + (-2/3) * \log_2 (2/3) = 0,918295834$
 Nilai Jalur Kabel Layak
 $= (-51/54) * \log_2 (51/54) + (-2/54) * \log_2 (2/54) = 0,309543429$
 Nilai Jalur Kabel Mungkin
 $= (-18/29) * \log_2 (18/29) + (-11/29) * \log_2 (11/29) = 0,957553484$
 Nilai Jalur Kabel Tidak Layak
 $= (-1/2) * \log_2 (1/2) + (-1/2) * \log_2 (1/2) = 1,000000000$
 Nilai Konverter Layak
 $= (-30/32) * \log_2 (30/32) + (-2/32) * \log_2 (2/32) = 0,337290067$
 Nilai Konverter Mungkin
 $= (-39/49) * \log_2 (39/49) + (-10/49) * \log_2 (10/49) = 0,730016630$
 Nilai Konverter Tidak Layak
 $= (-1/4) * \log_2 (1/4) + (-3/4) * \log_2 (3/4) = 0,811278124$

Langkah selanjutnya adalah menghitung gain pada setiap atribut data :

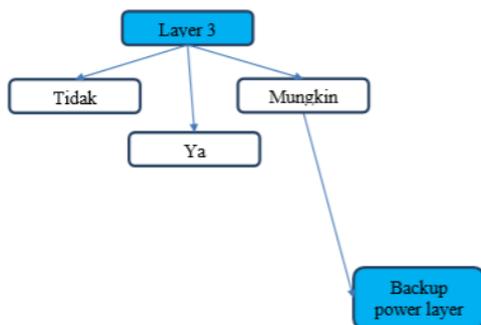
Nilai Backup Power Layer 2
 $= 0,672295 - ((28/85 * 0,371232327) + (47/85 * 0,550984443) + (10/85 * 0,881290899)) = 0,141663$
 Nilai Layer 3
 $= 0,672295 - ((54/85 * 0,228538144) + (28/85 * 0,966618633) + (3/85 * 0,918295834)) = 0,176280$
 (paling tinggi)
 Nilai Jalur Kabel
 $= 0,672295 - ((54/85 * 0,309543429) + (29/85 * 0,957553484) + (2/85 * 1,000000000)) = 0,125420$
 Nilai Konverter
 $= 0,672295 - ((32/85 * 0,337290067) + (49/85 * 0,730016630) + (4/85 * 0,811278124)) = 0,086304$

Setelah perhitungan selesai nilai paling tinggi akan menentukan akar pohon pertama, berikut gambar perhitungan manual entropy dan gain pertama :

Tabel Perhitungan Node Akar	JUMLAH	Layak	Tidak Layak	Entropy	Gain
Total	85	70	15	0,672294817	
Backup Power Layer 2					0,141663
Ya	28	26	2	0,371232327	
Mungkin	47	41	6	0,550984443	
Tidak	10	3	7	0,881290899	
Layer 3 Upe/ Router					0,176280
Ya	54	52	2	0,228538144	
Mungkin	28	17	11	0,966618633	
Tidak	3	1	2	0,918295834	
Jalur Kabel					0,125420
Ya	54	51	3	0,309543429	
Mungkin	29	18	11	0,957553484	
Tidak	2	1	1	1,000000000	
Konverter					0,086304
Ya	32	30	2	0,337290067	
Mungkin	49	39	10	0,730016630	
Tidak	4	1	3	0,811278124	

Gambar 6. Perhitungan Entropy dan Gain

Dari perhitungan yang manual tersebut di dapatkan akar pertama seperti pada Gambar 9:

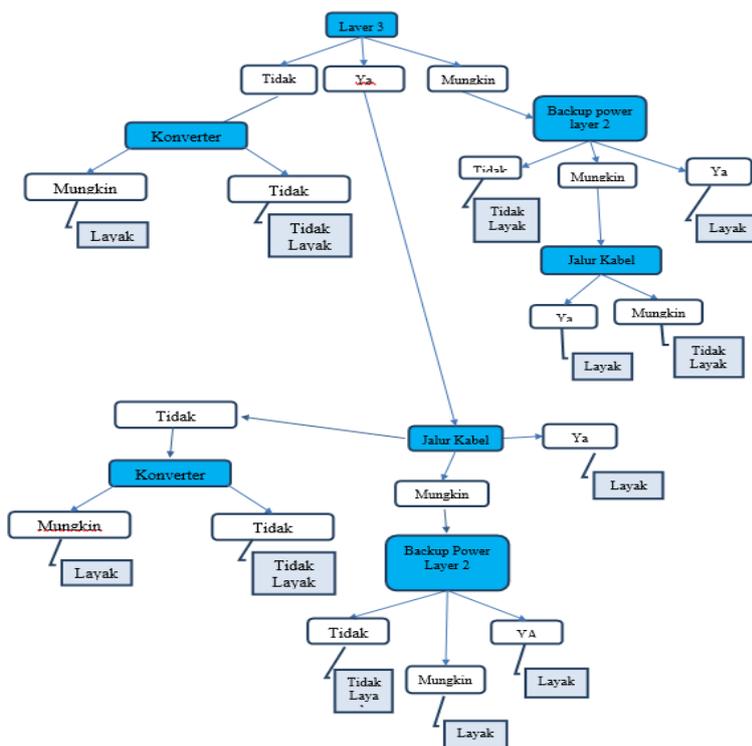


Gambar 7. Akar Pertama Dari Pehitungan Manual

Perhitungan untuk mencari gain atau node akan dilakukan sampai tidak dapat dihitung lagi, berikut perhitungan akhir gain pada G:

Tabel Perhitungan Node Akar	JUMLAH	Layak	Tidak Layak	Entropy	Gain
Total	3	1	2	0,528320834	
Backup Power Layer 2					0,251629
Ya	0	0	0	0,000000000	
Mungkin	2	1	1	0,000000000	
Tidak	1	0	1	0,000000000	
Jalur Kabel					0,251629
Ya	2	1	1	0,000000000	
Mungkin	1	0	1	0,000000000	
Tidak	0	0	0	0,000000000	
Konverter					0,918296
Ya	1	1	0	0,000000000	
Mungkin	0	0	0	0,000000000	
Tidak	0	0	1	0,000000000	

Gambar 8. Perhitungan Akhir



Gambar 9. Akar Terakhir

Dari perhitungan pada Gambar 7 akan dihasilkan pohon keputusan seperti pada Gambar 9, perhitungan sudah dilakukan sampai tidak dapat dihitung lagi dari semua perhitungan tersebut setelah dilakukan pengetesan menggunakan rapidminer menghasilkan akurasi sebesar 84%. Berikut adalah rule yang di dapatkan dari pohon keputusan :

- IF Layer Tiga Satu Kaki Dari Main and Backup Power Layer 2 Genset Then Layak
- IF Layer Tiga Satu Kaki Dari Main and Backup Power Layer 2 Baterai and Jalur Kabel Melewati Main Then Tidak Layak
- IF Layer Tiga Satu Kaki Dari Main and Backup Power Layer 2 Baterai and Jalur Kabel Tidak Melewati Main Then Layak
- IF Layer Tiga Satu Kaki Dari Main and Backup Power Layer 2 Tidak Ada Then Tidak Layak
- IF Layer Tiga Tidak Sama Dengan Main and Jalur Kabel Tidak Satu Jalur Then Layak
- IF Layer Tiga Tidak Sama Dengan Main and Jalur Kabel Melewati Main and Backup Power Layer 2 Genset Then Layak
- IF Layer Tiga Tidak Sama Dengan Main and Jalur Kabel Melewati Main and Backup Power Layer 2 Baterai Then Layak
- IF Layer Tiga Tidak Sama Dengan Main and Jalur Kabel Melewati Main and Backup Power Layer 2 Tidak Ada Then Tidak Layak
- IF Layer Tiga Tidak Sama Dengan Main and Jalur Kabel Satu Jalur and Konverter SFP1G Then Layak
- IF Layer Tiga Tidak Sama Dengan Main and Jalur Kabel Satu Jalur and Konverter Telways Then Tidak Layak
- IF Layer Sama Dengan Main and Konverter SFP10G Then Layak
- IF Layer Sama Dengan Main and Konverter Telways Then Tidak Layak

Dari rule yang sudah didapat akan di ujikan dengan hasil dari sistem. Dan hasil dari pengujian berhasil atau sesuai dengan rule yang dihasilkan. Berikut data sampling yang di compare dari rule yang di dapat dari pohon keputusan dengan sistem yang ada:

Tabel 4. Data Sampling Hasil Compare dari Pohon Keputusan

No	Scada Backup	Backup Power Layer 2 (CPE/Switch)	Layer 3 (UPE/Router)	Jalur Kabel	Konverter	Hasil Dari Rule	Hasil Dari Sistem	Hasil Pengujian
1	SCADA A	Layak	Mungkin	Mungkin	Layak	Layak	Layak	Berhasil
2	SCADA B	Mungkin	Layak	Layak	Mungkin	Layak	Layak	Berhasil
3	SCADA C	Mungkin	Mungkin	Mungkin	Mungkin	Tidak layak	Tidak layak	Berhasil
4	SCADA D	Mungkin	Layak	Layak	Mungkin	Layak	Layak	Berhasil
5	SCADA E	Layak	Layak	Layak	Layak	Layak	Layak	Berhasil
6	SCADA F	Layak	Layak	Mungkin	Mungkin	Layak	Layak	Berhasil
7	SCADA G	Tidak Layak	Tidak Layak	Layak	Tidak Layak	Tidak layak	Tidak layak	Berhasil

5. KESIMPULAN

Kesimpulan setelah selesai perancangan dan pembuatan sistem rekomendasi kelayakan scada 20kv dengan menggunakan algoritma Decision Tree C4.5 adalah Dapat menghasilkan sistem rekomendasi scada 20kv dengan menggunakan pohon keputusan dari algoritma C4.5. Rancangan atau bahkan link scada 20kv backup yang sudah jadi dapat di uji kelayakanya menggunakan sistem ini berdasarkan beberapa parameter yang sudah ditentukan. Dari data yang di dapat belum maksimal karena masih ada beberapa kriteria yang belum mendapat hasil rekomendasi setelah di proses menggunakan algoritma C4.5.

DAFTAR PUSTAKA

- [1] M. Sujatha and C. D. Jaidhar, "Classification of Soil Fertility using Machine Learning-based Classifier," in *ICSCCC 2021 - International Conference on Secure Cyber Computing and Communications*, Institute of Electrical and Electronics Engineers Inc., May 2021, pp. 138–143. doi: 10.1109/ICSCCC51823.2021.9478169.
- [2] S. Bahassine, A. Madani, M. Al-Sarem, and M. Kissi, "Feature selection using an improved Chi-square for Arabic text classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 32, no. 2, pp. 225–231, Feb. 2020, doi: 10.1016/j.jksuci.2018.05.010.
- [3] G. N. Awaludin *et al.*, "Comparison of Decision Tree C4.5 Algorithm with K-Nearest Neighbor (KNN) Algorithm in Hadith Classification," in *6th International Conference on Computing, Engineering, and Design, ICCED 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020. doi: 10.1109/ICCED51276.2020.9415796.
- [4] C. Deng and Z. H. Ma, "Research on C4.5 Algorithm Optimization for User Churn," in *2021 IEEE International Conference on Computer Science, Artificial Intelligence and Electronic Engineering, CSAIEE 2021*, Institute of Electrical and Electronics Engineers Inc., Aug. 2021, pp. 75–79. doi: 10.1109/CSAIEE54046.2021.9543367.
- [5] K. Djouzi, K. Beghdad-Bey, and A. Amamra, "A new adaptive sampling algorithm for big data classification," *J Comput Sci*, vol. 61, May 2022, doi: 10.1016/j.jocs.2022.101653.
- [6] F. Aaboub, H. Chamlal, and T. Ouaderhman, "Analysis of the prediction performance of decision tree-based algorithms," in *2023 International Conference on Decision Aid Sciences and Applications, DASA 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 7–11. doi: 10.1109/DASA59624.2023.10286809.
- [7] B. Zhang, "Tactical Decision System of Table Tennis Match based on C4.5 Decision Tree," in *Proceedings - 2021 13th International Conference on Measuring Technology and Mechatronics Automation, ICMTMA 2021*, Institute of Electrical and Electronics Engineers Inc., Jan. 2021, pp. 632–635. doi: 10.1109/ICMTMA52658.2021.00146.
- [8] C. J. D. Albia and D. Origines, "Selecting of Pet Adopters using C4.5 Decision Tree Model Algorithm," in *Proceedings - 2022 2nd International Conference in Information and Computing Research, iCORE 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 30–35. doi: 10.1109/iCORE58172.2022.00025.
- [9] P. Svefors *et al.*, "Relative importance of prenatal and postnatal determinants of stunting: Data mining approaches to the MINIMat cohort, Bangladesh," *BMJ Open*, vol. 9, no. 8, Aug. 2019, doi: 10.1136/bmjopen-2018-025154.
- [10] A. F. A., "Penerapan Data Mining Untuk Mengetahui Tingkat Kekuatan Beton Yang Dihasilkan Dengan Metode Estimasi Menggunakan Linear Regression," pp. 1–12, 2009.