



PENERAPAN ALGORITMA ID3 MELALUI APLIKASI ORANGE UNTUK PREDIKSI AKURASI AKREDITASI SEKOLAH DASAR DI DEPOK

Febrianto¹, Hasbi Yallah², Muhammad Darwis³, Retno Hendrowati⁴

¹²³⁴Program Studi Teknik Informatika, Universitas Paramadina

Jl. Raya Mabes Hankam No.Kav 9, Setu, Kec. Cipayung, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta 16680

¹febrianto@students.paramadina.ac.id,

²hasbi.yallah@students.paramadina.ac.id,³muhammaddarwis.@paramadina.ac.id,

⁴retno.hendrowati@paramadina.ac.id

Abstract

This research aims to explore the application of the ID3 algorithm using the Orange application to predict the accuracy of elementary school accreditation in the Depok City area. The focus of the research is focused on increasing school accreditation scores, in predicting the results of accreditation scores as an indicator of educational quality. By analyzing relevant data regarding supporting factors and barriers in relation to primary school accreditation, this research evaluates the performance of ID3 as a predictive tool. The research methodology involves the stages of data collection, processing and implementation of ID3 via the Orange platform. The results of this research can provide in-depth insight into ID3's ability to predict the accuracy of elementary school accreditation in Depok.

Keywords : Accreditation, Elementary School, ID3 Algorithm, Prediction, Education Quality

Abstrak

Penelitian ini bertujuan untuk mengeksplorasi penerapan Algoritma ID3 dengan menggunakan aplikasi Orange untuk memprediksi akurasi akreditasi sekolah dasar di wilayah Kota Depok. Fokus penelitian difokuskan pada peningkatan nilai akreditasi sekolah, dalam meramalkan hasil nilai akreditasi sebagai indikator kualitas pendidikan. Dengan menganalisis data yang relevan mengenai faktor-faktor pendukung dan hambatan dalam kaitannya dengan akreditasi sekolah dasar, penelitian ini mengevaluasi kinerja ID3 sebagai alat prediktif. Metodologi penelitian melibatkan tahapan pengumpulan data, pemrosesan, dan implementasi ID3 melalui platform Orange. Hasil penelitian ini dapat memberikan wawasan mendalam mengenai kemampuan ID3 dalam memprediksi akurasi akreditasi sekolah dasar di Depok.

Kata kunci : Akreditasi, Sekolah Dasar, Algoritma ID3, Prediksi, Kualitas Pendidikan

1. PENDAHULUAN

Pentingnya akreditasi dalam meningkatkan mutu pendidikan sekolah menjadi suatu hal yang tak terbantahkan. Akreditasi bukan sekadar indikator kualitas, melainkan juga sebagai panduan untuk perbaikan dan pengembangan. Dalam upaya meningkatkan mutu pendidikan tingkat dasar, pemanfaatan data dari Kementerian Pendidikan dan

Kebudayaan (Kemendikbud) merupakan langkah strategis. Menurut ketentuan yang tercantum dalam Keputusan Menteri Pendidikan dan Kebudayaan Nomor 59 Tahun 2012 (Pasal 1Ayat 2), Komisi Akreditasi Nasional Sekolah/Madrasah bertugas sebagai lembaga penilaian yang independen dalam mengevaluasi dan menetapkan kecocokan program serta unit Pendidikan yang menilai dan menetapkan



kelayakan program dan satuan pendidikan jenjang pendidikan dasar dan menengah jalur formal dengan mengacu pada standar nasional Pendidikan. Dataset tersebut bukan hanya menjadi dasar kebijakan, tetapi juga membantu mengidentifikasi tren dan mengoptimalkan proses akreditasi dalam memberikan nilai yang dapat berdampak positif yang lebih berpengaruh besar bagi dunia pendidikan dasar di Indonesia terutama di kota Depok.[1][2]

Akreditasi memiliki peran utama dalam menilai sejauh mana sekolah memenuhi standar yang telah ditetapkan oleh otoritas pendidikan. Pengukuran ini melibatkan berbagai aspek, termasuk kualitas pengajaran, fasilitas fisik, keberlanjutan sekolah, dan faktor lain yang berkontribusi pada mutu pendidikan. Data yang diperoleh dari Kemendikbud memberikan gambaran menyeluruh tentang kondisi dan kinerja sekolah-sekolah di berbagai wilayah, memberikan evaluasi yang lebih akurat dan dasar pengambilan keputusan yang lebih akurat.[3][4]

Dalam konteks ini, penelitian ini bertujuan untuk memanfaatkan data akreditasi dari Kemendikbud dengan menggunakan pendekatan data mining dan ilmu data. Algoritma pemrosesan data, seperti Algoritma Iterative Dichotomiser 3 (ID3), akan digunakan untuk menganalisis dan memahami pola-pola dalam proses akreditasi sekolah. Dengan menerapkan teknologi ini, diharapkan penelitian ini dapat memberikan kontribusi yang substansial dalam merumuskan strategi pengembangan, memberikan rekomendasi perbaikan, dan secara keseluruhan meningkatkan kualitas pendidikan dasar di Indonesia.[5][6]

Dari banyak metode yang tersedia, penelitian ini memilih metode ID3 sebagai alternatif karena aturan yang dihasilkan melalui Decision Tree mudah dipahami dibandingkan dengan metode klasifikasi lainnya. Melalui pendekatan ini, penelitian ini tidak hanya memfokuskan pada akreditasi sebagai alat evaluasi, tetapi juga sebagai instrumen proaktif untuk peningkatan mutu pendidikan. Dengan memahami dan memanfaatkan data dari Kemendikbud, diharapkan penelitian ini dapat memberikan wawasan yang mendalam, mendukung pengambilan keputusan berbasis bukti, dan mengoptimalkan pemanfaatan akreditasi sebagai alat strategis untuk mencapai mutu pendidikan yang lebih tinggi di nasional.[7]

2. TINJAUAN PUSTAKA

Penelitian ini bertujuan untuk meningkatkan akurasi prediksi Reformasi Akreditasi Sekolah/Madrasah Pendekatan Model Prediksi dengan menggunakan metode c4.5 untuk klasifikasi potensi akademik siswa. Langkah-langkah preprocessing data dilakukan dengan mengisi nilai yang hilang dan normalisasi data sebelum menerapkan algoritma c4.5. Analisis data kemudian dilakukan menggunakan pohon keputusan.

Hasil penelitian menunjukkan bahwa penggunaan preprocessing data berhasil meningkatkan akurasi prediksi dari 80% menjadi 88%. Temuan ini menekankan pentingnya kualitas data dalam analisis prediktif.

Referensi: [8]

Penelitian ini memperkenalkan beberapa aspek baru yang belum dieksplorasi dalam penelitian sebelumnya. Pertama, fokus penelitian ini adalah pada prediksi akurasi akreditasi sekolah dasar di Depok, yang merupakan area yang belum banyak diteliti. Kedua, penelitian ini memanfaatkan aplikasi Orange, yang memiliki antarmuka grafis lebih intuitif dan efisien dibandingkan perangkat lunak lain seperti Weka atau RapidMiner yang digunakan dalam penelitian terdahulu. Ketiga, selain menggunakan algoritma ID3, penelitian ini juga mengevaluasi dampak preprocessing data terhadap peningkatan akurasi prediksi. Melalui pendekatan ini, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam meningkatkan kualitas analisis dan prediksi akreditasi sekolah dasar di Indonesia.

3. METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Untuk menyelesaikan penelitian ini penulis menerapkan metode klasifikasi atau pengelompokan dengan metode Iterated Dichotomous 3 (ID3). Tujuan dari klasifikasi tersebut adalah bertujuan agar dapat memprediksi akreditasi sekolah dengan mempertimbangkan berbagai faktor yang



menjadi kriteria untuk memperoleh nilai tertinggi Metode tersebut melibatkan beberapa tahapan yang dimulai dengan tinjauan literatur, identifikasi dataset, dan persiapan data, termasuk validasi data, transformasi data, reduksi data, dan eksplorasi data. Pada tahapan eksplorasi data, aplikasi Orange digunakan untuk membantu penulis dalam melakukan analisis hasil mining data berdasarkan dataset yang ada. Bukan hanya itu, aplikasi tersebut juga untuk menguji hasil klasifikasi dalam penelitian dengan fitur yang tersedia dengan mudah. Peneliti melakukan analisis dan pengujian data yang diperoleh melalui sebuah Alamat link <https://dapo.kemdikbud.go.id/sp/2/026600>. [9] [10]

3.2 Akreditasi

Akreditasi sekolah adalah proses evaluasi formal oleh lembaga yang berwenang seperti Dinas Pendidikan untuk menilai kualitas pendidikan yang diberikan kepada sekolah. Proses ini memastikan bahwa sekolah memenuhi standar yang ditetapkan mencakup aspek pendidikan, infrastruktur, sumber daya manusia, dan pengelolaan. Akreditasi ini penting karena memberikan jaminan kepada siswa, orang tua, dan masyarakat bahwa sekolah tersebut memenuhi kriteria kualitas yang ditetapkan oleh otoritas pendidikan atau badan akreditasi. [11][12]

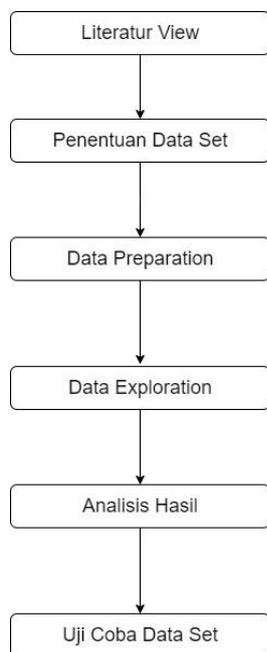
Score Akreditasi Sekolah 60-100 tersebut maksudnya kurang lebih sebagai berikut:

- a. Score 91-100 "Akreditasi A", maksudnya sekolah dengan nilai Score memenuhi setandar kualitas yang ditetapkan. Mencakup nilai kualitas kurikulum, kualifikasi guru, fasilitas sekolah, manajemen sekolah, dan hasil pendidikan
- b. Score 81-90 "Akreditasi B", maksudnya sekolah dengan nilai Score memenuhi sebagian besar standar kualitas, tetapi mungkin ada beberapa area yang memerlukan perbaikan. Meskipun tidak setinggi akreditasi A, sekolah dengan akreditasi B masih dianggap baik dan

mampu memberikan pendidikan yang memadai kepada siswanya.

- c. Score 71-80 "Akreditasi C", Akreditasi ini diberikan kepada sekolah yang memenuhi standar minimal. Artinya, sekolah tersebut memiliki kekurangan dalam beberapa area yang penting. Sekolah dengan akreditasi C perlu melakukan beberapa perbaikan untuk memastikan bahwa mereka memberikan pendidikan yang memadai dan memenuhi standar yang lebih tinggi..
- d. Score 61-70 "Akreditasi D", pada score ini, Ini adalah tingkat akreditasi terendah. Sekolah dengan akreditasi D sering kali jauh di bawah standar dalam berbagai aspek penting. Mereka perlu melakukan perbaikan yang signifikan dalam banyak area, seperti kualitas pengajaran, fasilitas, atau manajemen. Dalam beberapa sistem, sekolah yang mendapat akreditasi D mungkin berisiko kehilangan izin operasionalnya jika tidak melakukan perbaikan.

Berdasarkan hal rentang score diatas, penulis memberikan definisi sesuai dengan yang di terapkan oleh Pemerintah dalam nilai akreditasi untuk sekolah sd. terhadap nilai akreditasi tersebut yaitu 91-100 dengan defines "Akreditasi A", 81-90 dengan definisi "Akreditasi B" 71-80 dengan definisi "Akreditasi C". dan 61-70 definisi "Akreditasi D" Definisi tersebut nantinya yang penulis gunakan dalam dataset untuk membuat kalsifikasi dan pohon keputusan dalam mencari akreditasi. Hal ini tentunya agar proses dan hasil penelitian nantinya mudah untuk dibaca dan dipahami. [13]



Gambar 1. Flow dan Tahapan Penelitian

3.3 Literatur view

Penulis menggunakan beberapa jurnal akademis sebagai dokumen pendukung penelitian ini diantaranya adalah, Analisis dan Penerapan Algoritma Iterative Dichotomiser 3 (ID3) dalam , Pendekatan Metode Pohon Keputusan Menggunakan Algoritma ID3 Untuk memprediksi keberhasilan sekolah menghadapi UN.[14], Analisa Prediksi Efek Kerusakan Gempa Dari Magnitudo (Skala Richter) Dengan Metode Algoritma Id3 Menggunakan Aplikasi Data Mining Orange.[15], Penerapan Metode ID3 terhadap Implementasi FIS Sugeno pada Algoritma C4.5 Berbasis Particle Swarm Optimization (PSO) untuk Prediksi Prestasi Siswa.[16], Selain itu, penulis juga menggunakan jurnal yang dikhususkan untuk data mining.

Dengan metode yang digunakan adalah klasifikasi seperti Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes.[17], Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS.[18]

3.4 Penentuan dataset sekolah SD

Dalam menyelesaikan permasalahan yang ada, penulis mencoba mengambil dataset yang diperoleh dengan mengakses <https://dapo.kemdikbud.go.id/sp/2/026600>. dari link di atas peneliti mencoba mengambil data untuk penelitian akreditasi sekolah sd di kota depok mengenai akreditasi sekolah pada preode tahun 2022 sampai 2023. dari data akreditasi sekolah sd yang diperoleh ini, penulis mencoba melakukan cek data sekolah sd agar tidak ada data yang sama, setelah itu data untuk selanjutnya dibawa ke tahap persiapan untuk selanjutnya di bawa ke tahap preparation.

Jumlah dataset sekolah sd yang diperoleh sebanyak 422 data, dimana didalamnya berisikan informasi rinci mengenai akreditasi sekolah seperti jumlah guru, ruang kelas, fasilitas perpustakaan, fasilitas laboratorium, dan sebagainya. Untuk penelitian ini, penulis memilih beberapa atribut yang dianggap relevan sebagai bahan pertimbangan untuk menentukan tingkat akreditasi sekolah. Atribut yang dipertimbangkan mencakup jumlah guru, jumlah ruang kelas, fasilitas perpustakaan, fasilitas laboratorium, dan lain-lainnya dari beberapa atribut yang ada dalam data sekolah sd tersebut, diharapkan dapat terbentuk pola yang dapat membantu dalam memprediksi tingkat akreditasi sekolah. Langkah ini bertujuan untuk memberikan insight kepada pihak lembaga pendidikan agar dapat melakukan perbaikan dan peningkatan kualitas Pendidikan.

3.5 Data preparation

Langkah selanjutnya melibatkan beberapa langkah persiapan data. Fase ini terdiri dari tiga bagian penting: validasi dataset, transformasi dataset, dan reduksi dataset. Untuk memudahkan penelitian pada tahap ini, penulis memanfaatkan fungsionalitas yang tersedia pada aplikasi Orange. Di bawah ini rincian ketiga tahapan.

3.6 Dataset validation

peneliti memverifikasi dataset untuk memastikan dataset sekolah dasar yang diperoleh dalam kondisi baik tanpa adanya nilai



yang hilang. Nilai yang hilang atau bermasalah mencakup kumpulan dataset yang tidak lengkap, outlier (dataset tidak normal), adalah nilai atau angka yang terdapat di dalam dataset yang tidak konsisten. Tahap validasi dataset ini diharapkan dapat menghasilkan kumpulan dataset dimana tidak ada nilai yang hilang didalam dataset. Artinya, pada fase selanjutnya keadaannya normal.

3.7 Data transformation

Langkah selanjutnya adalah transformasi dataset. Hal ini dimaksudkan untuk menjaga dan menjamin akuratan dataset sekolah dasar. Pada tahap ini peneliti menerapkan teknik outlier dengan menggunakan fitur “Metode Deteksi Outlier” yang tersedia pada aplikasi Orange. Proses transformasi data ini diharapkan dapat menghasilkan dataset yang lebih akurat dan dapat dipindahkan ke tahap selanjutnya.

Reduksi data dilakukan penulis untuk mengambil sampel data dari dataset sekolah dasar yang ada. Tujuannya adalah untuk memastikan bahwa kumpulan data yang dipilih benar-benar siap untuk klasifikasi. Dari total 422 data penulis menyaringnya hingga tersisa 80% dari data, atau sebanyak 388 data. Dengan demikian, dari tahapan ini, akan dihasilkan data sampel yang siap untuk diproses ke tahap selanjutnya.

3.8 Data Exploration

Pada fase ini merupakan bagian penting dari proses klasifikasi, dataset yang telah disiapkan sebelumnya oleh peneliti dan akan diolah secara menyeluruh dengan menghitung nilai dan menganalisis hasilnya. Penelitian ini menggunakan teknik pengolahan data dan komputasi dengan menggunakan berbagai fitur yang tersedia pada aplikasi Orange. Fitur yang digunakan oleh penulis termasuk, namun tidak terbatas pada, fitur outlier dan model pohon. Ini divisualisasikan menggunakan penampil pohon untuk memudahkan pemahaman hasil klasifikasi.

3.9 Analisis Hasil Klasifikasi

Dalam tahapan ini peneliti mencoba menganalisis hasil klasifikasi yang diperoleh dari perhitungan cepat dengan menggunakan aplikasi Orange. Dari hasil yang di peroleh penulis dapat mendeskripsikan beberapa atribut atau turunan mana yang paling cocok untuk klasifikasi atau pemisahan. Selain itu, tahap ini juga memungkinkan Anda melihat dan menganalisis pola-pola yang terbentuk dari hasil klasifikasi menggunakan aplikasi Orange.

3.10 Uji Coba Dataset

Dalam menguji kumpulan data, Peneliti menerapkan beberapa metode validasi silang 5 kali lipat yang membagi kumpulan dataset secara acak atau random menjadi 10 subset independen. Dataset dapat prediksi atau diperoleh menggunakan aplikasi Orange kemudian dataset yang telah di peroleh akan dihitung dan dianalisis menggunakan beberapa teknik pengujian seperti Tree dan Naive Bayes. Dalam hal ini Penulis secara manual menghitung nilai ketelitian dan keakuratan setiap lipatan dan membandingkan kedua hasilnya. Selain itu, penulis juga menyajikan matriks konfusi untuk memastikan kebenaran algoritma yang diperoleh dari dataset. Bagus ya, nilai matriks konfusi harus sesuai dengan akurasi rata-rata atau akurasi lipat dalam validasi silang 5 kali lipat.

4. HASIL DAN PEMBAHASAN

Peneliti mencoba membuat struktur rancangan klasifikasi dataset dimulai dari dataset, tahap desain, dan implementasi data mining. Hal ini menciptakan pola-pola tertentu yang nantinya bisa digunakan dalam memprediksi kemungkinan sertifikasi, sehingga menimbulkan banyak minat dan menghasilkan nilai yang tinggi. Fase-fase tersebut meliputi mengidentifikasi kumpulan data, menyiapkan data, mengeksplorasi data dengan Orange, serta menganalisis dan memvalidasi hasil komputasi.



4.1 Dataset sekolah SD depok

Pada fase ini yang dilakukan oleh peneliti adalah dengan pendefinisian dataset sekolah dasar, seperti yang terlihat pada Gambar 2. Dataset ini berisi data sekolah dasar yang terletak di wilayah Kota Depok dari tahun 2023 sampai 2024 semester genap. Data tersebut kemudian diunggah ke dalam Platform oranye yang memfasilitasi pemrosesan data seperti persiapan data. Dataset ini tidak ada missing valuenya karena datanya sudah dibersihkan terlebih dahulu. Oleh karena itu, kemungkinan hilangnya nilai sangat kecil. Dari tabel data terlihat bahwa penulis memilih setidaknya sembilan atribut dari kumpulan data dan selanjutnya mengolahnya menjadi pola. Atribut AKREDITASI pada tabel tersebut dikenal dengan sebutan kelas dengan nilai 'A', 'B', 'C', 'D' yang ditetapkan oleh penulis berdasarkan pengertian akreditasi di atas. Empat atribut sisanya kini sudah lengkap.

- a. JUMLAH siswa Pria dan Wanita (PD), berisi data jenis genre siswa dan merupakan atribut biner. Nilainya adalah Rombel, Guru, Pegawai dan Ruang kelas.
- b. Rombe merupakan aribut biner yang berisi data Kumpulan siswa yang terdaftar ditiap kelas dalam satuan Pendidikan. Nilainya adalah PD (genre).
- c. GURU yaitu data jumlah pengajar yang ada pada satuan pendidik di dalam sekolah atau madrasah.
- d. PEGAWAI merupakan data yang berisikan orang yang bekerja di dalam lingkungan satuan pendidik di dalam sekolah.

	Akreditasi	PD	Rombel	Guru	Pegawai
1	A	587.000	18	20	2
2	A	376.000	12	12	2
3	A	383.000	12	14	3
4	A	472.000	15	15	6
5	A	377.000	12	12	3
6	A	550.000	17	18	3
7	B	373.000	11	14	3
8	A	292.000	13	26	6
9	A	310.000	15	27	3
10	A	237.000	10	11	2
11	A	208.000	8	15	0
12	A	567.000	22	54	6
13	A	565.000	21	59	14
14	B	38.000	6	8	2
15	A	318.000	12	21	4
16	B	83.000	6	6	3
17	A	95.000	8	6	2
18	A	820.000	24	25	5
19	A	663.000	20	19	4
20	A	574.000	18	19	5
21	A	363.000	12	12	4
22	B	316.000	12	12	4

Gambar 2. Data Set Sekolah SD Depok

4.2 Dataset preparation

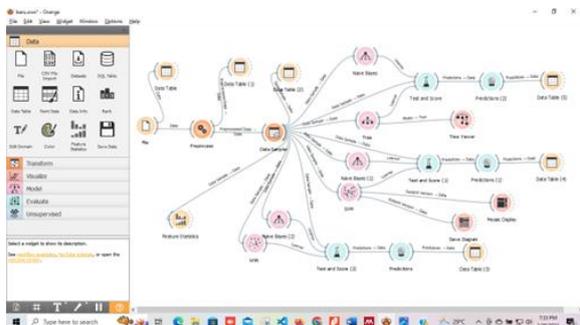
Pada bagian ini peneliti mempersiapkan dataset, yang mencakup validasi dataset, transformasi dataset, dan reduksi dataset. Penulis menggunakan aplikasi Orange untuk melakukan semua tahap persiapan ini guna memudahkan dan meningkatkan efektivitas pada proses dan pengolahan dataset.

a. Data validation

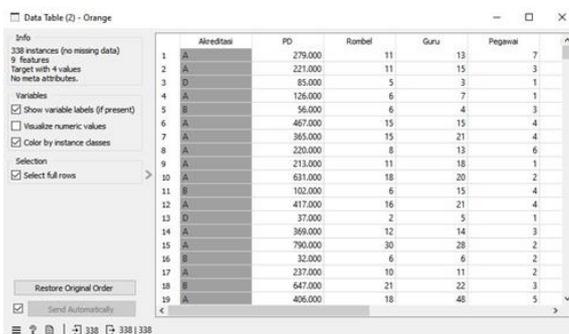
Sesuai dengan penjelasan sebelumnya, dataset yang digunakan oleh penulis sudah terbukti berkualitas baik tanpa adanya nilai yang hilang. Namun, untuk meningkatkan validitas data, peneliti menggunakan beberapa fitur preprocessing yang tersedia didalam aplikasi Orange untuk memastikan kembali. Selain itu, fitur Data Table memungkinkan penulis untuk meninjau data dan hasil validasinya. Hasilnya menunjukkan bahwa tidak ada nilai yang hilang dalam dataset, sehingga jumlah dataset tetap sebanyak 422. Bila dalam proses ini tidak ada nilai yang hilang maka peneliti bisa melanjutkan ke tahap selanjutnya.

b. Data transformation

Dalam proses transformasi data, peneliti memanfaatkan fitur Outlier yang ada di aplikasi Orange. Outlier adalah data yang tidak wajar atau tidak normal dalam suatu kumpulan dataset dan harus diperbaiki sebelum melanjutkan ke tahap berikutnya. Penulis kemudian menampilkan hasil proses pada data inlier. yang menunjukkan dataset normal atau semua data kecuali yang dianggap sebagai outlier. Proses transformasi menggunakan fitur Outlier dalam Orange ditunjukkan pada Gambar 3.



Gambar 3. Proses Outlier Dalam Orange



	Akreditasi	PD	Rombel	Guru	Pegawai
1	A	279.000	11	13	7
2	A	221.000	11	15	3
3	D	85.000	5	3	1
4	A	126.000	6	7	1
5	B	56.000	6	4	3
6	A	467.000	15	15	4
7	A	365.000	15	21	4
8	A	220.000	8	13	6
9	A	213.000	11	18	1
10	A	631.000	18	20	2
11	B	102.000	6	15	4
12	A	417.000	16	21	4
13	D	37.000	2	5	1
14	A	368.000	12	14	3
15	A	790.000	30	28	2
16	B	32.000	6	6	2
17	A	237.000	10	11	2
18	B	647.000	21	22	3
19	A	406.000	18	48	5

Gambar 4. Data Sampling Inliers

c. Data reduction

Data reduction ini pada dasarnya adalah pengambilan sampling terhadap data yang ada terutama jika row nya sangat banyak. Selain untuk mengefisienkan algoritma, hal ini dilakukan agar proses yang akan berjalan nantinya tidak terlalu berat. Meskipun beberapa row data direduksi pada tahap ini, namun kualitas data yang dihasilkan tetap sama sehingga tetap memenuhi syarat penelitian.

Dalam aplikasi Orange, untuk melakukan tahap ini, dapat digunakan fitur "Data Sampler" yang dihubungkan dari fitur Preprocess. Dalam penjelasan sebelumnya dari dataset penulis memperoleh dan menggunakan pada tahap ini adalah sekitar 80% dari total data atau sebanyak 422 data. Teknik yang penulis gunakan adalah Random Sampling dimana datanya akan dipilih secara acak. Dari output "Data Sampler" kemudian penulis menggunakan fitur Outlier sehingga nantinya diperoleh data inlier yakni data selain data dari data outlier. Data tersebut kemudian diolah lebih lanjut sehingga didapatkan total 388 data dari seluruh data yang ada. Data inliers yang didapatkan sebagai outputnya seperti terlihat pada Gambar 4. Data Inliers tersebutlah adalah dataset yang diproses lebih lanjut pada tahap berikutnya. sehingga pada tahap ini, dataset sudah sepenuhnya siap untuk diolah dan dieksplorasi lebih lanjut.

d. Data exploration

Pada tahap ini proses pengolahan dataset dan eksplorasi menggunakan Aplikasi Orange, penulis memberikan perhatian khusus pada fungsi-fungsi yang ada dan melakukan perhitungan statistik. Selain itu, penulis berusaha untuk menentukan kekuatan jaringan atau hubungan antar atribut. Pada tahapan ini peneliti saat ini adalah melakukan uji coba menggunakan teknik Analisis Univariat. Meskipun pada tahap ini juga bisa dilakukan pengujian dengan teknik Analisis Bivariat dan Analisis Multivariat, Namun, karena keterbatasan dataset yang digunakan, penulis hanya menggunakan analisis univariat. yang dilakukan dengan menyelidiki properti dari setiap atribut. Teknik ini diimplementasikan melalui fitur "Feature Statistics" yang tersedia di Orange. Rincian visualisasi dapat dilihat pada Gambar 5.

Dari Gambar 5 tersebut, dari pengolahan dataset secara umum dapat diperoleh. Nilai rata-rata yang dihasilkan oleh mean, modal, dan median dari nilai tersebut menunjukkan bahwa peneliti mendapatkan nilai pada variabel PD (mungkin Personal Development) yang cukup besar dengan mean sebesar 370.55336. Oleh karena itu, variabel PD mungkin merupakan faktor yang mempengaruhi akreditasi dalam sekolah dasar.

Selanjutnya, penelitian dilanjutkan dengan pemrosesan dataset menggunakan pendekatan klasifikasi dengan tujuan mengkategorikan dataset sekolah berdasarkan akreditasinya, yaitu "A", "B", "C", dan "D". Oleh karena itu peneliti melakukan klasifikasi dataset menggunakan aplikasi Orange, penulis juga mencoba

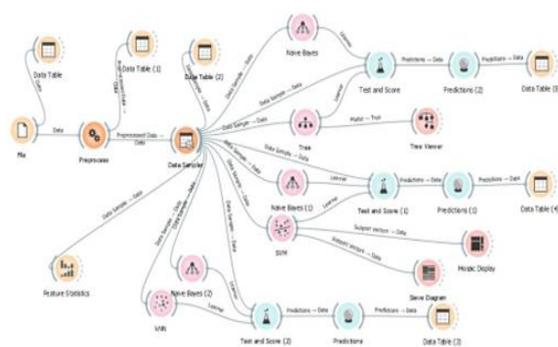


menggunakan fitur "Tree" dan peneliti mencoba menampilkan grafik pohon keputusannya dengan fitur "Tree Viewer" seperti yang terlihat pada Gambar 6.

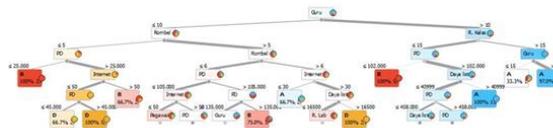
Aplikasi Orange memudahkan penulis melihat hasil pohon keputusan melalui fitur "Tree Viewer". Hasil yang diperoleh dengan pohon keputusan kemudian dianalisis oleh penulis. Untuk lebih jelasnya, penulis cukup mengklik fitur "Tree Viewer" yang ada pada kanvas. seperti yang terlihat pada Gambar 7. Grafik pohon keputusan yang dihasilkan cukup kompleks karena menggunakan 5 atribut kategorikal dengan nilai 3-4 kategori. Meskipun demikian, percabang pohon yang dihasilkan dalam pohon keputusan akan tetap dapat dibaca dan dianalisis dengan baik.

Name	atribut	Mean	Mode	Median	Dispersion	Min.	Max.	Missing
PD		370.55336	58.000	364.000	0.60750	1.036	974.000	0 (0 %)
Rombel		13.55	6	12	0.46	2	32	0 (0 %)
Guru		17.91	12	15	0.58	1	66	0 (0 %)
Pegawai		3.16	2	3	0.65	0	19	0 (0 %)
R. Kelas		11.39	6	10	0.51	4	39	0 (0 %)
R. Lab		0.83	0	1	1.11	0	5	0 (0 %)
R. Pcpus		0.92	1	1	0.43	0	2	0 (0 %)
Internet Mb		115.53	100	100	1.08	0	500	0 (0 %)

Gambar 5. Hasil Fitur "Feature Statistics" Pada Orange



Gambar 6. Proses Fitur "Tree" dan "Tree Viewer" Pada Orange



Gambar 7. Grafik Hasil Klasifikasi "Tree" Terhadap Data Sekolah SD

4.3 Analisis hasil klasifikasi

Dari hasil yang diperoleh dari pohon keputusan yang digunakan pada penelitian ini ditemukan bahwa atribut AKREDITASI merupakan simpul akar. Jika nilai atribut guru di atas 15, maka sekolah tersebut dijamin mendapat akreditasi A. Jika skor atribut guru lebih besar dari 10, maka sekolah tersebut dianggap memiliki Akreditasi B. Namun, jika nilai atribut Guru kurang dari atau sama dengan 10, maka sekolah tersebut akan mendapatkan akreditasi D. Dari pohon keputusan tersebut pula, dapat dijelaskan bahwa dari data set yang ada, dapat dibuatkan suatu pengelompokan atau classification. Lebih jauh.

Dari hasil pengelompokan tersebut, dapat dijadikan sebagai suatu aturan (rules) atau prosedur atau pola untuk memprediksi akreditasi, yang kemudian dapat menjadi tolak ukur dalam memilih sekolah bagi orang tua atau murid. Akreditasi A seringkali dijadikan pertimbangan utama bagi orang tua dalam memilih sekolah sd untuk anak-anak mereka. Berikut ini beberapa rules atau pola yang bisa diambil:

Pola menghitung akreditasi A sekolah dengan menggunakan pohon keputusan:

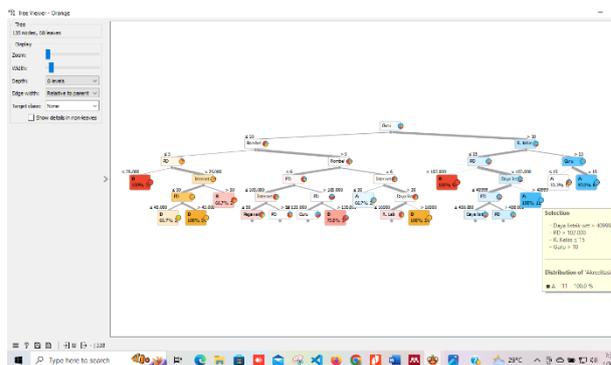
Pohon keputusan memberikan beberapa nilai dalam perhitungan terhadap nilai akreditasi sekolah dengan banyak Pola tersebut antara lain sebagai berikut.

- 1) Dari dataset yang ada, ditemukan bahwa jika jumlah Daya Listrik (wtt) > 40999, PD > 102.000, R.kelas > 15, dan Guru > 10, maka dapat dipastikan sekolah tersebut akan mendapatkan akreditasi A.
- 2) Jika nilai Internet (Mb) > 50, PD > 25.000, nilai R.kelas ≤ 5, dan Guru ≤ 10, maka dapat dipastikan sekolah tersebut termasuk dalam akreditasi B

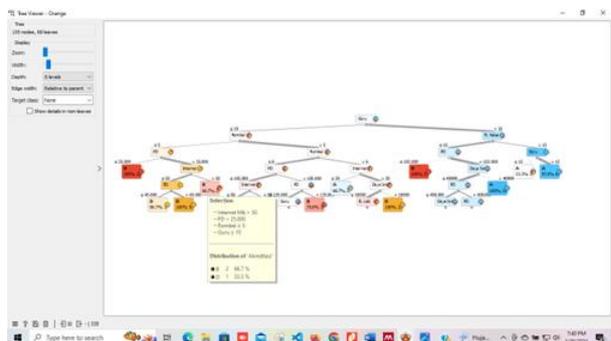


- 3) Jika nilai Daya Listrik (wtt) > 16500, nilai Internet (Mb) > 30, Rombel > 6, dan Guru ≤ 10, maka dapat dipastikan sekolah tersebut termasuk dalam akreditasi D.

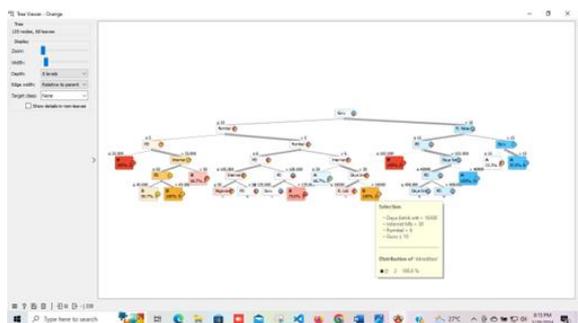
Data di atas bisa di lihat dari gambar 8,9 dan 10



Gambar 8. Nilai Perhitungan Akreditasi A



Gambar 9. Nilai Perhitungan Akreditasi B

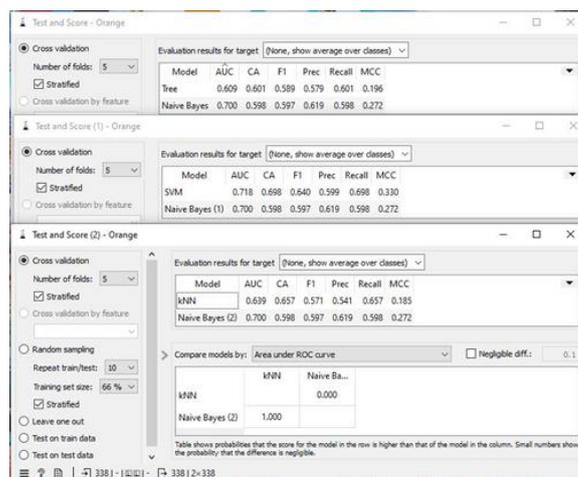


Gambar 10. Nilai Perhitungan Akreditasi D

3.4 Uji coba data set

Uji coba dilakukan dengan menggunakan 3 metode yaitu KNN, Tree, dan SVM, dimana himpunan data set dibagi secara acak menjadi 5 bagian. Untuk memastikan proses pengujian

lebih akurat Pada tahap ini, penulis menggunakan teknik pengujian dengan metode Tree dan Naïve Bayes. Untuk mempermudah, penulis dalam memanfaatkan fitur 'Test & Score' yang telah tersedia didalam aplikasi Orange. Dalam fitur 'Test & Score', peneliti mencoba menggunakan metode cross validation, seperti yang ditunjukkan oleh Gambar 11.



Gambar 11. Test & Score Cross Validation Pada Orange

Dari gambar tersebut, dapat terlihat bahwa pengujian dengan metode Naïve Bayes menunjukkan penurunan nilai pada beberapa metrik, terutama pada nilai Akurasi dan skor F1. Dalam pengujian menggunakan metode Tree, diperoleh akurasi sebesar 0,601 (61,0%), presisi sebesar 0,579 (57,9%), dan recall sebesar 0,601 (61,0%), dan F1 sebesar 0,589 (58,9%).

Sedangkan pada teknik uji coba dengan Naïve Bayes, terjadi penurunan akurasi menjadi 0,598 (59,8%), namun terjadi peningkatan nilai presisi menjadi 0,619 (61,9%). Untuk nilai recall mengalami penurunan menjadi 0,598 (59,8%), tetapi nilai F1 mengalami peningkatan menjadi 0,597 (59,7%). Hal ini mungkin disebabkan oleh jumlah data set yang digunakan dalam penelitian yang relatif sedikit, sehingga dalam teori, metode SVM mungkin lebih cocok untuk menguji Test & Score dari data set tersebut.

Dari Gambar 11, peneliti dapat menyimpulkan bahwa Setelah mengevaluasi dataset yang terdiri dari 422 data, metode SVM terbukti lebih baik dibandingkan dengan K-NN dan Decision Tree. Pengolahan data dengan metode SVM menunjukkan nilai tertinggi untuk akurasi sebesar 0,698 (69,8%), nilai F1 sebesar



0,640 (64,0%), presisi sebesar 0,599 (59,9%), dan recall sebesar 0,698 (69,8%). Meskipun demikian, hasil pengujian seperti yang ditunjukkan pada Gambar 11 sebelumnya menunjukkan bahwa perbedaan antara teknik pengujian Tree, SVM, K-NN, dan Naive Bayes hampir sama.

Selanjutnya presisi, recall, dan F1 semuanya menunjukkan hasil yang wajar, menunjukkan bahwa algoritma klasifikasi bekerja dengan baik. Namun, jika melihat area di bawah kurva ROC (AUC), yang mengukur performa model algoritmik, kami menemukan bahwa pendekatan Naive Bayes jauh lebih baik. Model yang baik memiliki nilai AUC mendekati 1,0; dan dalam pengujian ini, teknik Naive Bayes memperoleh nilai AUC sebesar 0,700 (70%), lebih baik dibandingkan dengan teknik Tree yang hanya mencapai nilai AUC sebesar 0,609 (60,9%).

Peneliti sangat puas dengan metode yang digunakan untuk pengujian ini, karena menunjukkan bahwa algoritma telah berfungsi dengan baik. Dengan demikian, klasifikasi ID3 (Tree) yang digunakan untuk memprediksi akreditasi sekolah yang kemungkinan akan mendapatkan skor akreditasi A dapat dijadikan acuan oleh peneliti dalam pengambilan keputusan terkait akreditasi sekolah dasar

5. KESIMPULAN

Penelitian ini telah menggunakan data set sekolah dasar di Kota Depok, dengan atribut dalam data set sebagai root dan diikuti oleh node-node lainnya seperti PD (Personal Development), Guru, Pegawai, Rombel, serta atribut pendukung lainnya seperti R.kelas, R.Perpustakaan, R.Laboratorium, Internet (Mb), dan Listrik (Watt). Hasil penelitian ini menghasilkan pola yang dapat digunakan oleh sekolah untuk memprediksi kemungkinan tinggi atau rendahnya akreditasi sekolah dasar. Dengan menerapkan algoritma ID3 yang dikembangkan dalam penelitian ini, peneliti dapat dengan mudah mengevaluasi kemungkinan akreditasi sekolah, apakah akan memiliki nilai A, B, atau D. Uji coba penelitian ini menunjukkan hasil yang baik, sehingga sekolah dapat mengimplementasikan teknik klasifikasi ID3 menggunakan data sekolah dasar yang telah ada di situs resmi Kementerian Pendidikan dan

Kebudayaan (Kemdikbud) untuk hasil yang lebih akurat. Hal ini penting karena hasil pohon keputusan yang dihasilkan dapat membantu dalam proses pengambilan keputusan akreditasi, yang akan menentukan hasil nilai akreditasi dalam sekolah. Dengan menggunakan, teknik atau metode klasifikasi ID3 dapat menjadi alat ukur dalam mendapatkan nilai akreditasi sekolah Sd. yang berguna untuk membantu pihak sekolah dan Dinas Pendidikan kota Depok dalam pengambilan keputusan terkait akreditasi sekolah. Dan memudahkan calon orang tua murid dalam melihat Akreditasi sekolah yang akan dituju.

DAFTAR PUSTAKA:

- [1] D. Iskamto, Jeli Nata Liyas, Elida Gultom, P. B. Ansori, Y. Harwina, and T. Hendra, "Pelaksanaan Proses Akreditasi Sekolah untuk menjaga kualitas Pendidikan Sekolah/Madrasah," *J. Pengabd. Masy. Akad.*, vol. 1, no. 2, pp. 46–51, 2022, doi: 10.54099/jpma.v1i2.132.
- [2] I. Martinelli and N. Khairiah, "Sosialisasi Urgensi Akreditasi Sekolah Bagi Masyarakat dalam Memperoleh Layanan Pendidikan yang Unggul," *Wahana Inov.*, vol. 9, no. 2, pp. 60–67, 2020.
- [3] Z. Nurhasanah, H. Asyari, and S. Ratnaningsih, "Analisis Fungsi Akreditasi Sekolah Dalam Peningkatan Produktivitas Sekolah," *J. Manaj. Pendidik.*, vol. 10, no. 02, pp. 118–124, 2022, [Online]. Available: <https://journal.unpak.ac.id/index.php/jmp>
- [4] V. Trysia, "A Case Study Of Teacher Perception And Commitment On Accreditation At School XYZ In Depok , West Java Studi Kasus Persepsi Dan Komitmen Guru Terhadap Akreditasi Sekolah XYZ Di Depok , Jawa Barat," *J. Pendidik.*, vol. 19, no. 1, pp. 31–42, 2018.
- [5] Nengah Widya Utami and I Wayan Budi Suryawan, "Implementasi Data Mining Untuk Mengklasifikasikan Produk Pada Sebuah Supermarket Menggunakan Algoritma Id3 Pada Orange," *Smart Techno (Smart Technol. Informatics Technopreneurship)*, vol. 3, no. 1, pp. 33–36, 2021, doi: 10.59356/smart-techno.v3i1.33.



- [6] D. Irmayanti, Y. Muhyidin, and D. A. Nurjaman, "Prediksi Mahasiswa Berpotensi Drop Out Dengan Metode Iteratif Dichotomiser 3 (ID3)," *J. Teknol. Inf.*, vol. 5, no. 2, pp. 103–113, 2021, doi: 10.36294/jurti.v5i2.2054.
- [7] M. Darwis, M. Firmasnyah, I. S. Sukma, and N. Ahady, "Implementasi algoritma ID3 untuk memprediksi kemungkinan film yang memiliki skor penilaian tinggi Implementation of ID3 algorithm to predict the possibility of films with high rating scores," 2020, doi: 10.14710/jtsiskom.x.x.202x.xx-xx.
- [8] N. F. A. Firdaus, D. M. Midyanti, and R. Hidayati, "PENERAPAN METODE C4. 5 UNTUK KLASIFIKASI POTENSI AKADEMIK SISWA (Studi kasus: SMP DDI Sultan Syarif Abdurachman Pontianak)," *Coding J. Komput. dan ...*, vol. 11, no. 01, 2023, [Online]. Available: <https://jurnal.untan.ac.id/index.php/jcskommipa/article/view/54266%0Ahttps://jurnal.untan.ac.id/index.php/jcskommipa/article/download/54266/75676597155>
- [9] F. Dwi Meliani Achmad, Budanis, Slamet, "Klasifikasi Data Karyawan Untuk Menentukan Jadwal Kerja Menggunakan Metode Decision Tree," *J. IPTEK*, vol. 16, no. 1, pp. 18–23, 2012, [Online]. Available: <http://jurnal.itats.ac.id/wp-content/uploads/2013/06/3.-BUDANIS-FINAL-hal-17-23.pdf>
- [10] W. N. Purba, D. Situmorang, Y. Alfani, D. Hutabarat, and W. Anggiono, "Implementasi Data Mining Dengan Metode Pohon Keputusan Algoritma Id3 Untuk Memprediksi Penjualan Pada Cv. Mitra Baja Cemerlang," *J. TEKINKOM*, vol. 2, pp. 110–115, 2019.
- [11] T. A. Yoga Siswa and Naufal Azmi Verdikha, "Komparasi Algoritma Klasifikasi Untuk Menentukan Evaluasi Kinerja Terbaik Pada Status Akreditasi Sekolah/Madrasah Kalimantan Timur Berdasarkan Iasp 2020," *J. Inform. Teknol. dan Sains*, vol. 4, no. 3, pp. 185–192, 2022, doi: 10.51401/jinteks.v4i3.1807.
- [12] S. Irawan, U. Tagela, and Y. Windrawanto, "Hubungan akreditasi sekolah dan supervisi oleh kepala sekolah dengan kualitas sekolah," *J. Akuntabilitas Manaj. Pendidik.*, vol. 8, no. 2, pp. 165–174, 2020, doi: 10.21831/jamp.v8i2.33905.
- [13] A. Fadlil, I. Riadi, and Y. Mulyana, "Penerapan Algoritma K-Means pada Pengelompokan Data Pendaftar Bantuan Biaya Pendidikan," *J. Teknol. Inform. dan Komput.*, vol. 8, no. 2, pp. 352–366, 2022, doi: 10.37012/jtik.v8i2.1261.
- [14] Y. Angraini, S. Fauziah, and J. L. Putra, "Analisis Kinerja Algoritma C4.5 Dan Naive Bayes Dalam Memprediksi Keberhasilan Sekolah Menghadapi Un," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 5, no. 2, pp. 285–290, 2020, doi: 10.33480/jitk.v5i2.1233.
- [15] L. Irawan, L. H. Hasibuan, and F. Fauzi, "Analisa Prediksi Efek Kerusakan Gempa Dari Magnitudo (Skala Richter) Dengan Metode Algoritma Id3 Menggunakan Aplikasi Data Mining Orange," *J. Teknol. Inf. J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 14, no. 2, pp. 189–201, 2020, doi: 10.47111/jti.v14i2.1079.
- [16] M. A. Yulianto, "Implementasi FIS Sugeno pada Algoritma C4. 5 Berbasis Particle Swarm Optimization (PSO) untuk Prediksi Prestasi Siswam," *JOAIIA J. Artif. Intell. Innov. Appl.*, vol. 1, no. 1, pp. 12–22, 2020, [Online]. Available: <http://www.openjournal.unpam.ac.id/index.php/JOAIIA/article/view/4272>
- [17] H. Hozairi, A. Anwari, and S. Alim, "Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes," *Netw. Eng. Res. Oper.*, vol. 6, no. 2, p. 133, 2021, doi: 10.21107/nero.v6i2.237.
- [18] R. Puspita and A. Widodo, "Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS," *J. Inform. Univ. Pamulang*, vol. 5, no. 4, p. 646, 2021, doi: 10.32493/informatika.v5i4.7622.