



Implementasi Algoritma C.45 Pada Data Pengolahan Limbah Kelapa Sawit

Wahyu Sri Rejeki

Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: milyani209@gmail.com

INFORMASI ARTIKEL

Article History

Received : Mar 11, 2021

Accepted : Nov 30, 2021

Published : Nov 30, 2021

KORESPONDENSI

Email: milyani209@gmail.com

A B S T R A K

Permasalahan yang dihadapi pada PT. PP. London Sumatera Indonesia Tbk, Bagerpang Palm Oil Mill, adalah masih terdapatnya kesalahan didalam pengelompokan limbah cair tersebut yang berdampak buruk serta mengakibatkan kerusakan lingkungan serta terganggunya aktifitas disekitar perusahaan. Salah satu cara untuk mengolah data tersebut adalah dengan menerapkan atau menggunakan teknik data mining. Dimana data mining merupakan proses analisa data yang besar menjadi informasi untuk menentukan pola dari kesimpulan data yang di analisa. Penerapan algoritma C.45 yang nanti nya akan menggali data yang sudah ada dan akan membentuk suatu pohon keputusan, dimana pohon keputusan inilah nanti nya yang akan membantu perusahaan agar mengetahui pengelompokan dan pengolahan limbah cair tersebut agar tidak berdampak buruk bagi sekitar perusahaan.

Kata Kunci: Data Mining; Algoritma C.45; Pohon Keputusan; Limbah

A B S T R A C T

The problems faced at PT. PP. London Sumatra Indonesia Tbk, Bagerpang Palm Oil Mill, is that there are still errors in the grouping of the liquid waste which have a bad impact and cause environmental damage and disrupt activities around the company. One way to process the data is to apply or use data mining techniques. Where data mining is the process of analyzing large data into information to determine the pattern of the conclusions of the data being analyzed. The application of the C.45 algorithm which later will explore existing data and will form a decision tree, where this decision tree will later help the company to know the grouping and processing of the liquid waste so as not to have a bad impact on the surrounding company.

Keywords: Data Mining; C.45 Algorithm; Decision Tree; Waste

1. PENDAHULUAN

PT.PP.London Sumatera Indonesia Tbk, Bagerpang *Palm Oil Mill*, merupakan suatu perusahaan perkebunan dan perdagangan yang berbasis di London. Perkebunan London-Sumatera kemudian lebih dikenal dengan nama (Lonsum), berkembang menjadi salah satu perusahaan perkebunan terkemuka di dunia, dengan lebih 100.000 hektar perkebunan kelapa sawit, karet, kakao dan teh di empat pulau terbesar di Indonesia. Lonsum juga dikenal sebagai produsen bibit kelapa sawit yang berkualitas, yang kini menjadi salah satu pendorong pertumbuhan perseroan. Lonsum juga merupakan penghasil minyak sawit terbesar kedua di Indonesia yang telah menerima sertifikat *Roundtable on Sustainable Palm Oil (RSPO)*, ditahun 2009 setelah pelaksanaan audit ahli independen atas perkebunan dan pabrik kelapa sawit di sumatera.

Namun PT.PP.London Sumatera Indonesia Tbk, Bagerpang *Palm Oil Mill* tidak berinteraksi langsung dengan konsumen maupun *supplier*, karena Lonsum merupakan anak perusahaan dari PT. Indofood. Pada lonsum sendiri tahapan pertama yang dilakukan untuk pengolahan kelapa sawit sebelum di proses menjadi limbah ialah penimbangan berat buah di stasiun penerimaan, kemudian dilakukan sortasi dan kemudian akan masuk kedalam stasiun perebusan, setelah itu masuk menuju stasiun penebah, selanjutnya akan dikirimkan pada stasiun kempa, selanjutnya akan menuju stasiun klarifikasi dan stasiun inti, kemudian hasilnya akan disimpan dan di ambil sampel untuk dilakukan pengecekan kadar FFA pada CPO dilaboratorium, kemudian hasil tersebut akan disimpan CPO dan Karna yang telah disimpan akan dikirim keperusahaan induk, apabila perusahaan telah menerima karna dan CPO maka proses pengolahan buah telah selesai. Kemudian sisa hasil dari proses pengolahan buah disebut sebagai limbah, yang dimana limbah tersebut harus diolah kembali agar tidak merusak ataupun mengganggu lingkungan juga aktifitas disekitar PT.PP.London Sumatera Indonesia Tbk, Bagerpang *Palm Oil Mill*. Salah satu limbah yang

dihasilkan adalah limbah cair yang dimana limbah cair tersebut dikelompokkan terhadap beberapa kategori. Hasil pengolahan limbah cair pada Lonsun akan dialirkan ke perkebunan kelapa sawit sebagai pupuk organik dan sebagai penyiraman kompos, maka limbah cair tersebut harus diolah terlebih dahulu. Tujuan utama dari pengolahan limbah cair adalah menyederhanakan senyawa yang kompleks, karena semakin tinggi kompleksitas suatu senyawa, maka semakin besar pula BOD nya atau kata lain kekompleksitasan senyawa berbanding dengan peningkatan BOD. Proses pengelompokan limbah cair tersebut dapat dilihat dari kandungan-kandungan yang terdapat pada limbah hasil sisa produksi, permasalahan yang dihadapi pada PT.PP.London Sumatera Indonesia Tbk, Bagerpang *Palm Oil Mill* adalah masih terdapatnya kesalahan didalam pengelompokan limbah cair tersebut yang berdampak buruk serta mengakibatkan kerusakan lingkungan serta terganggunya aktifitas disekitar PT.PP.London Sumatera Indonesia Tbk, Bagerpang *Palm Oil Mill*. Oleh karena itu dibutuhkan proses pengolahan data limbah pada PT.PP.London Sumatera Indonesia Tbk, Bagerpang *Palm Oil Mill* secara baik dan juga benar, salah satu cara untuk mengolah data tersebut adalah dengan menerapkan atau menggunakan teknik Data mining.

Data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data dengan melakukan penggalian pola-pola dari data menjadi informasi yang lebih berharga yang diperoleh dengan cara mengintraksi dan mengenali pola yang penting atau menarik dari data atau suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam database. Data mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengidentifikasi formasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan didalam database besar. Dengan menggunakan data mining diharapkan dapat memperoleh sebuah pengetahuan pada pengolahan limbah cair di PT.PP.London Sumatera Indonesia Tbk, Bagerpang *Palm Oil Mill* berdasarkan dari data-data pada basis data yang telah tersedia, hingga tidak terjadi lagi kesalahan pada pengelompokan limbah cair tersebut.

Algoritma C4.5 merupakan algoritma klasifikasi pohon keputusan yang banyak digunakan karena memiliki kelebihan utama dari algoritma lainnya. Kelebihan algoritma C4.5 dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisien dalam menangani atribut bertipe diskret dan dapat menangani atribut bertipe diskret dan numerik. Dalam mengkonstruksi pohon keputusan, algoritma C4.5 membaca seluruh sampel *data mining* dari *storage* dan memuatnya ke memori [1]. Berdasarkan penelitian terdahulu yang dilakukan oleh David Hartanto Kamagi yang berjudul Implementasi Data Mining dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan Mahasiswa mengatakan bahwasannya Data mining dengan algoritma C4.5 dapat diimplementasikan untuk mengetahui hal yang paling berpengaruh pada hasil [1].

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data Mining sebagai proses untuk mendapatkan informasi yang berguna dari gudang bisnis data yang besar. Istilah *data mining* terkadang disebut juga sebagai *knowledge discovery*. Salah satu teknik *data mining* yang dibuat untuk menelusuri data yang ada untuk membangun sebuah model, kemudian akan menggunakan model tersebut untuk dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan. Kebutuhan untuk prediksi ini juga dapat dilakukan dengan memanfaatkan teknik dalam data mining sebagai pengelompokan data[2].

Menurut Eko Prasetyo pada buku *Data Mining Konsep dan Aplikasi Menggunakan Matlab* mengatakan “Teknik *data mining* adalah bagaimana mencari data yang tersedia untuk menciptakan sebuah model, lalu memanfaatkan basis data yang tersimpan[2].

Kesimpulan dari data mining menurut penulis yaitu serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data dengan melakukan penggalian pola-pola dari data dengan tujuan untuk memanipulasi data menjadi informasi.

Ada istilah lain yang mempunyai makna yang sama dengan *data mining* yaitu *Knowledge-Discovery in Database* (KDD). Memang *data mining* atau KDD bertujuan untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna. Ternyata data mining mempunyai empat akar bidang ilmu sebagai berikut [4]:

1. Statistik

Bidang ini merupakan akar paling tua, tanpa ada statistik maka data mining mungkin tidak ada. Dengan menggunakan statistik klasik ternyata data yang diolah dapat diringkas dalam apa yang umum dikenal sebagai *exploratory data analysis* ((EDA). EDA bertujuan sebagai proses identifikasi hubungan yang tersistematis antar variabel atau fitur ketika tidak memiliki informasi alami. Teknik EDA klasik yang digunakan dalam *data mining* di antaranya :

- a. Metode komputasional : *statistic deskriptif* (distribusi, *parameter statisticklasik* (*mean, median, rata – rata, varian dan sebagainya*), korelasi, tabel frekuensi, teknik eksplorasi *multivariate* (analisis *cluster*, analisis faktor, analisis komponen utama dan klasifikasi, analisis diskriminan, *classification tree*, analisis korespondensi), model *linear* atau *nonlinear* rlanjutan (*regresi linear* atau *nonlinear, time series* atau *forecasting* dan sebagainya).
- b. Visualisasi data :mengarah pada representasi informasi dalam bentuk visual dan dapat dipandang sebagai satu yang paling berguna. Pada saat yang sama, visualisasi data merupakan metode eksplorasi data yang

- atraktif. Teknik visualisasi yang paling umum yang dikenal adalah *histogram* semua jenis kolom, silinder, kerucut, piramida, lingkaran, batang dan sebagainya), kotak, *scatter*, *konturmatrik*, *icon* dan sebagainya.
2. Kecerdasan buatan atau *artificial intelligence* (AI)
Artificial Intelligence berbeda dengan statistik. Teorinya dibangun berdasarkan teknik *heuristic* sehingga *Artificial Intelligence* memiliki peran pada pengolahan informasi sesuai dengan model penalaran manusia. Salah satu cabang dari AI, yaitu pembelajaran mesin atau *machine learning*, merupakan ilmu penting direpresentasikan dalam pembangun *data mining*, menggunakan teknik di mana sistem komputer berjalan dengan pelatihan.
 3. Pengenalan Pola
Data mining juga merupakan bagian dari pengenalan pola, hanya saja pengolahan data berasal dari basis data. Data yang digunakan berasal dari basis data kemudian dilakukan pengolahan dimana proses pengolahan bukan pada bentuk relasi, tetapi pada bentuk normal. *Data mining* memiliki ciri tersendiri yakni melakukan pencarian pola asosiasi serta melakukan pencarian pola sekuensial.
 4. Sistem Basis Data
Sistem basis data pada *data mining* memiliki fungsi sebagai penyedia informasi data yang akan diproses untuk penggaliannya dengan memanfaatkan metode – metode yang sudah ada.

2.2 Pengelompokan Teknik – Teknik *Data Mining*

Pada proses pemecahan masalah dan pencarian pengetahuan baru terdapat beberapa pengelompokan berdasarkan tugas yang dilakukan [5]:

1. Estimasi
Dimanfaatkan terhadap estimasi pada data baru yang belum memiliki keputusan terhadap *histori* yang ada. Sebagai contoh, estimasi terhadap biaya pembangunan gedung diluar kota.
2. Asosiasi
Asosiasi dimanfaatkan pada proses pengenalan perilaku pada kejadian – kejadian tertentu atau hubungan asosiasi yang muncul dari setiap kejadian. Sebagai contoh bidang *market bisnis* pada swalayan, untuk menyusun tata letak barang yang akan dijual sesuai dengan produk – produk yang sering dibeli oleh pelanggan secara bersamaan.
3. Klasifikasi
Merupakan teknik yang melihat pada setiap kelakuan serta atribut pada kelompok yang telah terdefiniskan. Dengan teknik ini bisa menentukan klasifikasi terhadap data baru dengan melakukan manipulasi data yang telah terklasifikasi sebelumnya dan memanfaatkan hasilnya untuk menentukan klasifikasi yang baru. Sebagai contoh dilakukan pada Akademik dengan mengklasifikasikan siswa yang berhak mendapatkan beasiswa.
4. Klastering
Klastering merupakan teknik analisis kelompok yang berbeda dari basis data, klastering hampir sama dengan klasifikasi tetapi proses pengelompokan dilakukan pendefinisian dilakukan sebelum proses data mining dijalankan.
5. Prediksi
Prediksi umumnya dimanfaatkan untuk perkiraan ataupun *forecasting* dari suatu kejadian sebelum kejadian itu terjadi. Contohnya untuk memperkirakan cuaca apakah cerah, berawan, mendung atau hujan.

2.3 Algoritma C.45

Algoritma C4.5 adalah salah satu metode untuk membuat *decision tree* berdasarkan *training* data yang telah disediakan. Algoritma C4.5 merupakan pengembangan dari ID3. Beberapa pengembangan yang dilakukan pada C4.5 adalah sebagai antara lain bisa mengatasi *missing value*, bisa mengatasi *continiu* data, dan *pruning*[1].

Pada tahap algoritma C4.5 memiliki 2 prinsip kerja yaitu:

1. Pembuatan pohon keputusan. Tujuan dari algoritma penginduksi pohon keputusan adalah mengkonstruksi struktur data pohon yang dapat digunakan untuk memprediksi kelas dari sebuah kasus atau *record* baru yang belum memiliki kelas. C4.5 melakukan konstruksi pohon keputusan dengan metode *divide and conquer*. Pada awalnya hanya dibuat *node* akar dengan menerapkan algoritma *divide and conquer*. Algoritma ini memilih pemecahan kasus-kasus yang terbaik dengan menghitung dan membandingkan *gain ratio*, kemudian *node-node* yang terbentuk di *level* berikutnya, algoritma *divide and conquer* akan diterapkan lagi sampai terbentuk daun-daun.
2. Pembuatan aturan-aturan (*rule set*). Aturan- aturan yang terbentuk dari pohon keputusan akan membentuk suatu kondisi dalam bentuk *if- then*.

Aturan-aturan ini didapat dengan cara menelusuri pohon keputusan dari akar sampai daun. Setiap *node* dan syarat percabangan akan membentuk suatu kondisi atau suatu *if*, sedangkan untuk nilai-nilai yang terdapat pada daun akan membentuk. Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

2.4 Pilih atribut sebagai akar.

1. Buat cabang untuk tiap-tiap nilai.
2. Bagi kasus dalam cabang.
3. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk melihat atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus seperti tertera dalam persamaan 1 berikut :

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy}(S_i) \dots \dots \dots (3.1)$$

Keterangan:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

|S_i| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

Sementara itu, penghitungan nilai *entropi* dapat dilihat pada persamaan 2 berikut :

$$\text{Entropy}(S) = \sum_{i=1}^n - p_i * \log_2 p_i \dots \dots \dots (3.2)$$

Keterangan:

S : himpunan kasus

A : fitur

n : jumlah partisi

S p_i : proporsi dari S_i terhadap S

Pada saat pembangunan pohon keputusan, banyaknya cabang mungkin mencerminkan adanya *noise* atau *outlier* pada *training* data. Pemangkasan pohon dapat dilakukan untuk mengenali dan menghapus cabang-cabang tersebut. Pohon yang dipangkas akan menjadi lebih kecil dan lebih mudah dipahami, pohon semacam itu biasanya juga menjadi lebih cepat dan lebih baik dalam melakukan klasifikasi[5].

2.5 Limbah

Limbah adalah buangan yang dihasilkan dari suatu proses produksi baik industri maupun domestik (rumah tangga)[6]. Dimana masyarakat bermukim, disanalah berbagai jenis limbah akan dihasilkan, seperti sampah, air kakus (*black water*), dan ada air buangan dari berbagai aktivitas domestik lainnya. Limbah juga merupakan semua material sisa atau buangan yang berasal dari proses teknologi maupun dari proses alam dimana kehadirannya tidak bermanfaat bagi lingkungan dan tidak memiliki nilai ekonomis. Pada dasarnya berbagai jenis limbah dihasilkan oleh kegiatan manusia, baik itu kegiatan *industry* maupun domestik (rumah tangga) dan berdampak buruk terhadap lingkungan dan juga bagi kesehatan manusia. Menurut Peraturan Pemerintah Nomor 101 tahun 2014, limbah adalah sisa suatu usaha atau kegiatan, berdasarkan dari wujud limbah yang dihasilkan, limbah dibagi menjadi tiga yaitu limbah padat, limbah cair dan gas dengan penjelasan sebagai berikut:

1. Limbah padat adalah limbah yang berwujud padat. Limbah padat bersifat kering, tidak dapat berpindah kecuali ada yang memindahkannya. Limbah padat ini misalnya, sisa makanan, sayuran, potongan kayu, sobekan kertas, sampah, plastik, dan logam
2. Limbah cair adalah limbah yang berwujud cair. Limbah cair terlarut dalam air, selalu berpindah, dan tidak pernah diam. Contoh limbah cair adalah air bekas mencuci pakaian, air bekas pencelupan warna pakaian, dan sebagainya.
3. Limbah gas adalah limbah zat (zat buangan) yang berwujud gas. Limbah gas dapat dilihat dalam bentuk asap. Limbah gas selalu bergerak sehingga penyebarannya sangat luas. Contoh limbah gas adalah gas pembuangan kendaraan bermotor. Pembuatan bahan bakar minyak juga menghasilkan gas buangan yang berbahaya bagi lingkungan.

Limbah memiliki ciri-ciri tertentu yang membedakannya dengan bend lainnya. Adapun beberapa karakteristik limbah adalah sebagai berikut:

1. Berukuran Mikro, limbah memiliki ukuran kecil atau partikel-partikel kecil yang masih dapat dilihat oleh mata manusia.
2. Bersifat dinamis, limbah selalu bergerak sesuai dengan lingkungan sekitarnya. Misalnya, ketika limbah masuk kesungai maka limbah tersebut akan mengikuti arah aliran sungai tersebut.
3. Penyebarannya berdampak luas, dampak yang ditimbulkan oleh limbah pada lingkungan dan manusia efeknya beragam. Ketika kontaminasi limbah sudah berat maka akan menyebabkan kerusakan bagi lingkungan dan manusia.
4. Dampak jangka panjang, limbah dapat menimbulkan dampak yang cukup lama di wilayah yang terkontaminasi. Sehingga dibutuhkan waktu yang cukup lama untuk mengembalikan kondisi wilayah tersebut.

3. HASIL DAN PENELITIAN

Analisa berfungsi untuk mengetahui dan mengamati permasalahan yang sedang terjadi, kemudian merumuskan masalah dan mencari cara solusi untuk memecahkan permasalahan yang ada. Untuk mengetahui data pengolahan limbah agar lebih akurat dan tidak terjadi kesalahan pada PT. PP. London Sumatera Indonesia Tbk, Bagerpang *Palm Oil Mill*, yaitu dengan menggunakan teknik Data Mining yang diharapkan dapat memperoleh sebuah pengetahuan pada pengolahan limbah cair di Lonsum yang dimana data-data di ambil berdasarkan dari basis data yang telah tersedia, hingga tidak terjadi lagi kesalahan pada pengelompokan limbah cair.

Tahapan pertama yang harus dilakukan agar limbah tersebut tidak mengganggu dan merusak lingkungan juga aktifitas disekitar perusahaan yaitu dengan mengolah kembali limbah pertama dengan cara melakukan penyaringan, kemudian proses pengolahan awal, setelah proses pengolahan awal maka akan dilakukan pengendapan, dan setelah

itu pengapungan yang berfungsi untuk menyingkirkan lemak atau minyak kepermukaan air. Dimana limbah yang dihasilkan adalah limbah cair, kemudian akan dikelompokkan terhadap beberapa kategori dan dapat dilihat dari kandungan-kandungan yang terdapat pada limbah hasil sisa produksi, yang kemudian hasil pengolahan limbah cair pada Lonsom akan dialirkan ke perkebunan kelapa sawit sebagai pupuk organik dan sebagai penyiraman kompos.

Salah satu cara untuk mengolah data tersebut adalah dengan menerapkan atau menggunakan teknik *data mining*. Dimana *data mining* merupakan proses analisa data yang besar menjadi informasi untuk menentukan pola dari kesimpulan data yang di analisa. Penerapan algoritma C.45 yang nanti nya akan menggali data yang sudah ada dan akan membentuk suatu pohon keputusan, dimana pohon keputusan inilah nanti nya yang akan membantu perusahaan agar mengetahui pengelompokan dan pengolahan limbah cair tersebut agar tidak berdampak buruk bagi sekitar perusahaan.

3.1 Analisa Data Mining Menggunakan Algoritma C.45

Sistem yang digunakan dalam klasifikasi penyebab pengolahan limbah cair yaitu dengan menggunakan data awal *numeric* dan *nonnumeric* yang akan dibagi melalui atribut-atribut untuk memudahkan analisa berikutnya, kemudian setelah semua data yang dibagi melalui atributnya maka akan dilakukan proses klasifikasi dengan membuat pohon keputusan sebagai *output*. Proses pengambilan keputusan untuk mengklasifikasi penyebab pengelompokan pengolahan limbah yang menjadi keputusan adalah kandungan-kandungan yang terdapat didalam limbah sisa hasil produksi.

Atribut merupakan kolom data yang menunjukan fungsi nya sebagai pembentuk karekteristik yang mempunyai banyak nilai, katagori, maupun kode untuk setiap basis data. Berikut yang merupakan keterangan atribut dalam pengolahan limbah cair yaitu :

1. Suhu
Merupakan ukuran derajat panas atau dingin nya suatu benda, suhu mempunyai 4 skala yaitu *celcius*, *fahrenheit*, *reamur*, dan *kelvin*.
2. Ph
Merupakan derajat keasaman yang digunakan untuk menyatakan tingkat keasamaan atau kebebasan yang dimiliki oleh suatu larutan.
3. BOD-5(*BiochemicalOxygenDemand,5days*)
BOD-5 menunjukan jumlah oksigen (mg O₂) yang dikomsumsi mikroba aerobik saat menguraikan organik terurai dalam waktu 5 hari pada 1 liter limbah cair.
4. COD (*Chemical Oxygen Demand*)
Merupakan senyawa organik yang tidak terurai, seperti selisih antara nilai COD dan nilai BOD 5 dari suatu limbah cair dianggap menunjukan jumlah senyawa organik yang tidak sesuai.
5. TSS (*Total suspended solid* atau pedatan *tersuspensi*)
Merupakan residu yang tertahan oleh saringan dengan ukuran partikel maksimal 2pm atau lebih besar dari ukuran partikel koloid, contohnya lumpur, logam, tanah liat, iksida, sulfide, bakteri dan jamur.
6. NH-N3
Merupakan senyawa kompleks ionik, yang terdiri atas ion positif (*cation*) dan ion negative (*anion*).
7. PO-P4
Merupakan persamaan dan pemyetaraan reaksi redoks.

Berikut adalah tabel data pengelohan limbah pada PT.PP.London Sumatera Indonesia Tbk, Bagerpang *Palm Oil Mil*.

Tabel 1. Data Limbah 3 BulanTerakhir

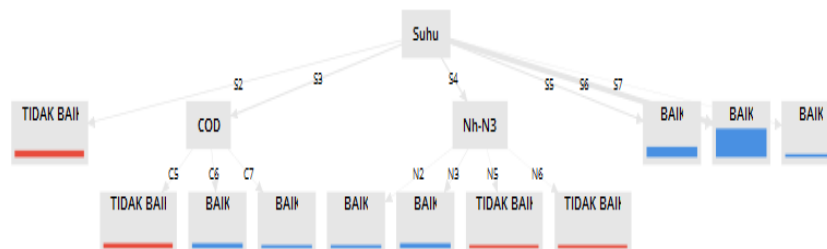
Tanggal	Suhu	Ph	BOD-5	COD	TSS	NH-N3	PO-P4	Keterangan
01-01-19	28°C	7.54	8	90	5	1	3.17	BAIK
03-01-19	30°C	7.55	15	50	52	18	2.91	BAIK
05-01-19	25°C	7.56	10	80	19	6	0.925	BAIK
08-01-19	15°C	7.57	11	64	20	9	2.85	BAIK
11-01-19	18°C	7.58	6	66	40	7	0.256	BAIK
13-01-19	25°C	7.59	5	75	19	12	3.17	BAIK
16-01-19	30°C	8.1	8	80	10	19	2.85	BAIK
19-01-19	33°C	8.11	9	35	15	15	0.827	BAIK
21-01-19	15°C	8.12	10	60	55	13	0.834	BAIK
23-01-19	10°C	8.13	20	80	33	11	3.17	TIDAK BAIK
26-01-19	25°C	8.14	15	90	35	15	0.929	BAIK
29-01-19	30°C	8.15	25	89	40	15	0.245	BAIK
03-02-19	30°C	8.16	9	55	50	12	2.87	BAIK
06-02-19	27°C	8.17	15	69	30	22	2.85	BAIK
09-02-19	12°C	8.18	5	50	55	25	2.69	TIDAK BAIK
12-02-19	10°C	8.19	10	35	30	33	2.66	TIDAK BAIK
15-02-19	18°C	8.2	11	80	22	12	3.23	BAIK
18-02-19	22°C	8.21	14	55	45	17	3.35	BAIK

Tanggal	Suhu	Ph	BOD-5	COD	TSS	NH-N3	PO-P4	Keterangan
21-02-19	30°C	6.3	12	70	25	22	2.13	BAIK
24-02-19	20°C	8.55	17	75	44	29	3.17	TIDAK BAIK
27-02-19	28°C	6.6	19	89	55	33	3.16	BAIK
30-02-19	18°C	8.13	22	90	11	25	3.16	TIDAK BAIK
03-03-19	15°C	8.14	30	50	17	12	3.1	TIDAK BAIK
06-03-19	10°C	8.15	29	60	19	11	0.321	TIDAK BAIK
09-03-19	25°C	8.16	28	70	50	1	0.825	BAIK
11-03-19	18°C	8.17	12	80	45	15	0.865	BAIK
13-03-19	27°C	8.18	15	55	34	17	0.833	BAIK
16-03-19	30°C	8.19	11	90	39	23	0.825	BAIK
19-03-19	15°C	8.2	10	60	44	28	0.825	BAIK
21-03-19	29°C	8.21	5	40	66	33	0.885	BAIK
23-03-19	30°C	7.54	15	55	60	36	0.825	BAIK
26-03-19	28°C	7.55	18	50	62	12	0.875	BAIK
29-03-19	30°C	7.56	20	90	12	21	0.825	BAIK
31-03-19	30°C	7.57	30	80	10	18	0.825	BAIK

Sebelum menerapkan data pada algoritma C.45 yang digunakan perlu kiranya dilakukan tahapan preprosesing yang berguna untuk menyesuaikan data terhadap karakteristik terhadap algoritma yang digunakan. Adapun preprosesing data dapat dilihat pada tabel dibawah ini:

3.2 Hasil Pengujian

Hasil akhir atau *output* dari aplikasi *RapidMiner* adalah berupa pohon keputusan atau *decision tree* atau *id3*. Dari pohon keputusan inilah akan dihasilkan *rules* atau aturan-aturan yang dapat membantu dalam mengambil keputusan. Pohon keputusan akhir dapat dilihat pada gambar 1 berikut :



Gambar 1. Pohon Keputusan Hasil Pengujian

Dari hasil pohon keputusan tersebut, didapatkan rule aturan sebagai berikut :

Suhu = S2: TIDAK BAIK {BAIK=0, TIDAK BAIK=3}

Suhu = S3

| COD = C5: TIDAK BAIK {BAIK=0, TIDAK BAIK=2}

| COD = C6: BAIK {BAIK=2, TIDAK BAIK=0}

| COD = C7: BAIK {BAIK=1, TIDAK BAIK=0}

Suhu = S4

| Nh-N3 = N2: BAIK {BAIK=1, TIDAK BAIK=0}

| Nh-N3 = N3: BAIK {BAIK=2, TIDAK BAIK=0}

| Nh-N3 = N5: TIDAK BAIK {BAIK=0, TIDAK BAIK=1}

| Nh-N3 = N6: TIDAK BAIK {BAIK=0, TIDAK BAIK=1}

Suhu = S5: BAIK {BAIK=5, TIDAK BAIK=0}

Suhu = S6: BAIK {BAIK=15, TIDAK BAIK=0}

Suhu = S7: BAIK {BAIK=1, TIDAK BAIK=0}

Selanjutnya setelah didapatkan rule dan juga pohon keputusan, kemudian menghitung tingkat akurasi yang dari proses klasifikasi. Adapun tingkat akurasi dapat dilihat pada gambar 5.9 berikut :

accuracy: 100.00%			
	true BAIK	true TIDAK BAIK	class precision
pred. BAIK	27	0	100.00%
pred. TIDAK BAIK	0	7	100.00%
class recall	100.00%	100.00%	

Gambar 2. Hasil Tingkat Akurasi

Pada gambar 2 telah dilakukan pengujian data pengolahan limbah kelapa sawit dengan menggunakan aplikasi *Rapid Miner*, dan didapatkan bahwa tingkat akurasi dari pohon keputusan sebesar 100%.

4. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, disimpulkan dengan memanfaatkan data mining dapat mengetahui pengelompokan pengolahan limbah kelapa sawit yang baik untuk dijadikan pupuk organik pada PT.PP.London Sumatera Indonesia.Tbk. Algoritma C.45 dapat membantu untuk membuat suatu keputusan dalam mengidentifikasi suatu permasalahan dan melihat hubungan antara faktor-faktor yang mempengaruhi suatu masalah, dan dapat membantu untuk memecahkan suatu masalah. Aplikasi *RapidMiner Classification Decision Tree* digunakan sebagai aplikasi pendukung keputusan dan pengujian atas hasil yang didapatkan secara manual, yang menghasilkan sebuah pohon keputusan. Dari pohon keputusan inilah akan menghasilkan sebuah aturan-aturan yang dapat membantu pihak PT.PP.London Sumatera Indonesia.Tbk dalam data pengolahan limbah kelapa sawit serta mudah dipahami oleh pengguna aplikasi.

REFERENCES

- [1] S. D. N. S. Hikmah Widya., "Data Mining Untuk Memprediksi Jenis Transaksi Nasabah Pada Koperasi Simpan Pinjam Dengan Algoritma C.45," Media Informatika Budidarma, vol. 1, no. 2, pp. 32-37, Juni 2017.
- [2] E. Prasetyo, Data Mining Konsep Dan Aplikasi Menggunakan Matlab, Yogyakarta: CV. Andi Offset, 2012.
- [3] E. Prasetyo, Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab, Yogyakarta: CV. Andi Offset, 2014.
- [4] D. Nofriansyah, Buku Algoritma Data Mining dan Pengujian, 2017.
- [5] D. H. Kamagi and S. Hansun, "Implementasi Data Mining dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan Mahasiswa," ULTMATICS, vol. VI, no. 1, pp. 15-20, 2014.
- [6] P. Limbah, Agustus 2018. [Online]. Available: <http://googleweblight.com>.
- [7] B. S. D. Oetomo, Perencanaan dan Pengembangan Sistem Informasi, Yogyakarta: CV. Andi Offset, 2002.
- [8] A. Preantoko, Jago Microsoft Excel 2010, Jakarta: Kawan Pustaka, 2010.
- [9] D. A. C, D. A. Baskoro, L. Ambarwati and I. W. S. Wicaksana, Belajar Data Mining Dengan Rapid Miner, 2013.