

Analisis Data Produksi Menggunakan Metode Decision Tree (C4.5) di PT. Baja Marga Kharisma Utama

Ali Rohman Zarkasi¹, Fahlepi Roma Doni²

Program Studi Informatika, Universitas Bina Sarana Informatika, Jl. Kamal Raya No.18, Jakarta 11730, Indonesia¹

Arz210499@gmail.com¹, fahlepi.fro@bsi.ac.id²

Abstrak – PT Baja Marga Kharisma Utama (BMKU), adalah perusahaan yang bergerak di bidang produksi dan distributor Besi Baja yang sudah berdiri selama 42 tahun. Kualitas bahan baku produksi merupakan salah satu komponen penting dalam menentukan hasil produk yang dibuat oleh suatu perusahaan. Saat ini PT. BMKU masih belum menerapkan penggunaan teknologi untuk mengolah dan menyimpan data produksi yang sifatnya permanen dan penting. Sehingga, hasil produksi tidak dapat dihitung secara tepat. Perbandingan jumlah masing-masing pipa dan hollow yang sesuai standar hanya dapat dilakukan secara perkiraan saja. Berdasarkan hal tersebut penelitian ini melakukan analisis data produksi untuk clustering hasil produksi baja pipa dan hollow dengan menggunakan metode Decision Tree (C.45). Dari hasil penelitian ini, dapat disimpulkan bahwa penerapan data mining dengan menggunakan metode decision tree sangat bermanfaat untuk mengetahui produksi meningkat atau tidaknya selama 3 bulan dan mengetahui mesin mana produksi terbanyak selama 3 bulan dan produksi terbanyak diperoleh pipa 3,0 mm dan untuk mesin produksi terbanyak diperoleh mesin ch-2.

Kata Kunci – Analisis Data, Decision Tree (C4.5), RapidMiner

I. PENDAHULUAN

Keberhasilan suatu perusahaan dalam memproduksi produk tidak hanya ditentukan oleh nilai sumber daya manusia, sarana dan prasarana yang tersedia, tetapi juga kualitas dari bahan baku hasil produksi yang digunakan. Kualitas bahan baku produksi merupakan komponen penting dalam menentukan hasil produk yang dibuat oleh suatu perusahaan.

PT Baja Marga Kharisma Utama (BMKU), Perusahaan ini bergelut di bidang produksi dan penyuplai Besi Baja yang telah berdiri selama 42 tahun. Kantor utama BMKU terletak di kapuk raya, Jakarta utara dan pabrik produksi terletak di Kamal Raya, Jakarta Utara dan beberapa cabang yg tersebar di beberapa wilayah di Indonesia.

Dalam mengekstraksi informasi dari kumpulan data data mining dapat digunakan yang kemudian data tersebut dipresentasikan agar dapat dipahami oleh manusia untuk meraih sebuah kepastian [1].

Dalam mengelompokan data yang memiliki kemiripan yang sama dapat menggunakan teknik *data mining clustering*. *Clustering*

adalah suatu prosedur analisis data yang memungkinkan pengelompokan objek-objek yang sifat dan bentuknya sedemikian rupa sehingga menciptakan kelompok-kelompok serupa di antara anggota-anggota kelompok yang sama [2].

Clustering memiliki manfaat dapat mengkategorikan data tanpa informasi sebelumnya dibandingkan teknik data mining lainnya yang tidak memiliki manfaat tersebut. Berdasarkan kesamaan yang dimiliki Dari berbagai jenis data *Clustering* memisahkan dan mengatur data dari tipe data yang berbeda ke dalam *cluster*. Pada Teknik clustering ada juga algoritma yang sering dipakai dalam penelitian yaitu Algoritma pohon keputusan (C.45).

Algoritma ini merupakan algoritma yang dapat membuat pohon keputusan. Algoritma pohon keputusan C4.5 mengunjungi setiap node keputusan secara rekursif, memilih pemisahan optimal hingga pemisahan lebih lanjut tidak lagi memungkinkan. Algoritma pohon keputusan C4.5 menggunakan konsep perolehan informasi atau penurunan entropi untuk memilih bagian yang optimal.

Data hasil produksi PT. BMKU di tahun 2020 total produksi mencapai 90.593.817 ton, dan tahun 2021 total produksi mencapai 78.052.727 ton, sedangkan tahun 2022 total produksi hanya 41.903.270 ton.

Saat ini PT. BMKU belum menerapkan penggunaan teknologi dalam proses pengolahan dan penyimpanan data yang persisten dan penting. Sehingga masalah yang diketahui adalah ketika hasil produksi yang dihasilkan tidak dapat dihitung secara tepat berapa perbandingan jumlah masing-masing pipa dan hollow diproduksi sesuai dengan standar kualitas yang ditentukan, sehingga hanya dapat dijadikan perkiraan. Untuk itu diperlukan pengelolaan yang tepat, hal ini dapat dilakukan melalui salah satu cara yaitu analisis data produksi untuk mengklasifikasikan hasil produksi baja pipa dan hollow dengan menggunakan metode Decision Tree (C.45).

Berlandaskan latar belakang tersebut, maka penelitian ini dirumuskan sebagaimana menganalisis data produksi menggunakan metode Pohon Keputusan (C4.5) di PT. Baja Marga Kharisma Utama.

Penulisan ini bertujuan untuk menganalisis data produksi menggunakan metode decision tree (C4.5) di PT. Baja Marga Kharisma Utama.

Memberikan kemudahan bagi kinerja karyawan dalam menganalisis data produksi di PT. Baja Marga Kharisma Utama (BMKU) menggunakan metode decision tree (C4.5) sehingga lebih efektif dan efisien.

II. TINJAUAN PUSTAKA

A. Definisi Data

Definisi Data adalah bahan baru yang butuh diolah agar dapat memberikan keterangan atau informasi, baik kuantitatif maupun kualitatif, yang memberikan fakta sedemikian rupa peneliti dapat memperoleh manfaat atau peneliti juga dapat memperoleh wawasan tentang suatu kondisi atau situasi. Sedangkan informasi adalah kumpulan data yang telah diolah untuk melakukan analisis untuk digunakan oleh pihak-pihak yang memerlukannya [3].

Data juga dapat diartikan sebagai informasi tentang peristiwa atau fakta yang dirumuskan dalam suatu kelompok lembaga tertentu, tidak

bersifat acak, dan menunjukkan jumlah, tindakan, peristiwa, kegiatan dan operasi yang tidak dimilikinya dan dampak langsung bagi penggunaannya [4]

B. Jenis Data

Menurut pandangan Muri jenis data sederhana dapat dibedakan menjadi empat macam yaitu diantaranya sebagai berikut.

a. Data Nominal.

Dalam pengukuran data, data nominal merupakan alat ukur paling sederhana. Skala nominal ini mempunyai fungsi terbatas, yaitu membedakan dan mengidentifikasi. Data nominal berisi data klasifikasi atau mengkategorikan berdasarkan atribut atau simbol lain secara sempurna dan luas. Ciri-ciri data skala nominal adalah:

- 1) Tidak ada bilangan pecahan pada hasil penghitungan,
- 2) Hanya label yang tertera pada angka,
- 3) Tidak ada urutan (peringkat),
- 4) Ukuran baru yang tidak ada,
- 5) Nol mutlak yang tidak dimiliki,
- 6) Statistik non parametrik adalah uji statistic yang akan digunakan.

Contoh :Jenis kelamin: 1. Pria 2. Wanita.

b. Data Ordinal.

Skala ordinal adalah pengukurannya yang tidak hanya menentukan kategori melainkan juga memberi kelas pada konstruk yang akan diukur. Variabel penelitian ini terdapat banyak konsep yang tidak juga dapat diberi atribut atau diklasifikasikan secara lengkap, tetapi juga berkaitan satu sama lain. Hubungan dicirikan oleh derajat, keteraturan, atau keteraturan pada tingkat yang berbeda. Atau karena ingin mengetahui sifat variabel yang akurat, bahwa pengukuran ini lebih tepat untuk kondisi tersebut. Ciri-ciri dari data tersebut adalah sebagai berikut:

- a. Mereka saling eksklusif,
- b. Data memiliki aturan logis,
- c. Skalanya ditentukan oleh jumlah karakter khusus.

Skala ordinal ini kian unggul dibandingkan skala nominal lalu sering juga disebut skala nilai. Hal ini disebabkan akibat pada skala ordinal, simbol-simbol pengukuran selain membuktikan selisih, juga membuktikan keteraturan atau tahapan benda yang diukur

dalam sifat tertentu. Misalnya tahapan kepuasan seseorang kepada produk. Bisa kita beri angka 5 = sangat puas, 4 = puas, 3 = kurang puas, 2 = tidak puas, dan 1 = sangat tidak puas.

C. Data Interval

Data yang diambil dengan pengukuran ketika jarak antara dua titik pada skala diketahui. Berlainan dengan skala nominal dan ordinal, skala interval mempunyai satuan pengukuran khusus sehingga jaraknya tetap. Berikut adalah ciri-ciri dari data interval:

- 1) Mereka saling eksklusif,
- 2) Logis,
- 3) Skala tersebut ditentukan berdasarkan banyaknya ciri-ciri yang dimiliki,
- 4) Adanya kelainan karakteristik yang serupa tercermin dari kelainan jumlah yang sama yang diberikan antar bilangan,
- 5) Angka nol hanya mewakili satu titik pada skala.

Oleh karena itu skala interval yaitu skala yang memberitahu jarak antara satu titik data dengan titik data lainnya lalu diberi bobot yang bobotnya sama.

a. Data Rasio

Data skala rasio adalah data yang dilakukan dengan pengukuran apabila jarak antara dua titik pada skala ini dan mempunyai angka nol mutlak. Pada tipe data ini, relasi tersebut mempunyai derajat dimensi terbesar dan memiliki nilai mutlak nol. Jika pada skala ini titik nolnya sembarang dan tiada membagi atau mengalikan, maka pada skala rasio keempatnya bisa dilakukan. Seluruh ciri-ciri skala nominal, ordinal, dan interval juga ditemukan pada skala rasio. Ciri-ciri skala rasio adalah:

- 1) Mereka dipisahkan satu sama lain berdasarkan golongan data,
- 2) Mempunyai aturan logistik untuk golongan data,
- 3) Skala tersebut ditentukan berdasarkan banyaknya ciri-ciri tertentu yang dimiliki,
- 4) Kelainan karakteristik yang sama tercermin dalam kelainan serupa dalam total yang diberikan antar golongan,
- 5) Angka nol melambangkan suatu titik pada skala yang menunjukkan tidak adanya suatu sifat (memiliki nilai mutlak nol).

D. Decision Tree

Pohon keputusan adalah struktur data yang

terbagi dari node dan edge. Node dalam pohon ini terbagi dalam beberapa bagian yaitu node akar, cabang/internal, dan daun [5].

Pohon keputusan adalah metode klasifikasi untuk beberapa jumlah kelas terbatas, di mana ikatan internal dan ikatan akar diberi label dengan sebutan atribut, tepi dikasih label dengan nilai karakter, dan ikatan daun diberi label dengan tingkatan yang beragam [6].

Decision Tree C4.5 menggambarkan salah satu metode prediksi dan klasifikasi yang sangat kuat dan sering digunakan dalam analisis data. Algoritma C4.5 membangun pohon keputusan berdasarkan data pelatihan, yang direpresentasikan dalam database dalam bentuk observasi atau catatan (tupel). Setiap kasus berisi nilai karakter tingkatan. Setiap karakter dapat berisikan data diskrit atau kontinu[7]

Terdapat beberapa model dan algoritma untuk sebarang kasus pada metode decision tree salah satunya ialah klasifikasi menggunakan Decision Tree. Berdasarkan hal yang telah diberitahu sebelumnya pada data mining terbagi beberapa langkah yaitu: pembersihan data ialah untuk menghapus noise dan data yang bertentangan, Beberapa sumber data yang dapat dikombinasikan disebut integrasi data, pemilihan data merupakan data yang penting dengan kepentingan yang diambil dari basis data, data yang dievolusi ke dalam bentuk data yang searah untuk mining merupakan proses transformasi data, data mining merupakan cara esensial dimana cara cerdas ini diaplikasikan untuk mengekstrak pola data, untuk mengidentifikasi pola yang benar-benar menarik yang mewakili pengetahuan yaitu evaluasi pola [8].

Sebagai algoritma C4.5 untuk membangun pohon keputusan yaitu sebagai berikut:

1. Memilih atribut sebagai akar.
2. Membentuk cabang untuk masing-masing nilai.
3. Membuat pembagian kasus dalam cabang.
 4. Ulangi proses untuk tiap-tiap cabang sampai semua kasus tiap cabang memiliki kelompok yang sama.

E. RapidMiner

RapidMiner adalah platform perangkat lunak data ilmiah yang dikembangkan oleh

perusahaan dengan nama yang sama. Rapidminer menyediakan lingkungan terpadu untuk pembelajaran mesin, pembelajaran mendalam, penambangan data, dan analisis prediktif. Rapidminer dapat digunakan sebagai aplikasi dalam lingkungan bisnis dan komersial, juga dapat digunakan untuk penelitian, pendidikan, pelatihan, pembuatan prototipe cepat, dan pengembangan aplikasi, serta mendukung semua tahapan proses pembelajaran mesin, termasuk persiapan data, visualisasi hasil, verifikasi dan optimalisasi [9].

F. Penelitian Terkait

Dari pengujian yang dilakukan menggunakan data training uji sebanyak 1003 record dengan 5 variabel, diketahui bahwa produksi yang tercapai sebanyak 709 dan produksi yang tidak tercapai sebanyak 192. Dengan hasil akurasi sebesar 99.9002 %. [10]

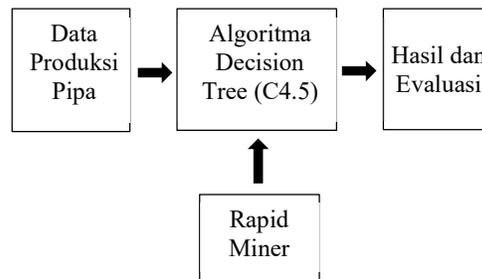
Metode Decision Tree terbukti membantu stokist herbal produk HWI di Benai dalam mengetahui produk mana yang laris dan tidak laris sehingga Stokist HWI di Benai bisa memasok barang lebih banyak dan dapat meningkatkan penjualan dan juga dapat mempertahankan pelanggan. [11]

Algoritma C4.5 telah terbukti dapat digunakan untuk mengklasifikasikan data kepuasan pelanggan. Pohon keputusan yang dihasilkan memberikan dua aturan yang menjadi pedoman penting dalam menilai kepuasan pelanggan, yaitu rasa dan kemurnian. [12]

III. ANALISA DAN PERANCANGAN SISTEM

A. Kerangka Penelitian

Pada penelitian ini untuk memecahkan permasalahan yang ada digunakan kerangka penelitian sebagai gambaran bagaimana penelitian ini akan dilakukan. Berikut merupakan kerangka penelitian pada penelitian ini.



Gambar 1 Kerangka Penelitian

B. Objek Penelitian

Pada penelitian ini objek yang akan diambil adalah di PT Baja Marga Kharisma Utama (BMKU) yang terletak di Jalan kapuk kamal nomor 17, Jakarta Utara dan berkantor di Jalan kapuk raya nomor 26 serta memiliki cabang di wilayah Tangerang, Bekasi dan Bandung. Dengan mengambil data produksi pipa di PT Baja Marga Kharisma Utama (BMKU) bulan Oktober 2023 - Desember 2023.

C. Sumber Data

Dalam metode penelitian ini ada beberapa metodologi dalam pengumpulan data yaitu sebagai berikut :

1. Wawancara

Dalam pengumpulan data untuk penelitian yaitu metode yang digunakan ialah wawancara. Guna mendapatkan informasi atau data yang penting dengan eksplorasi yang dilakukan peneliti pada proses ini melakukan sesi tanya jawab dengan narasumber.

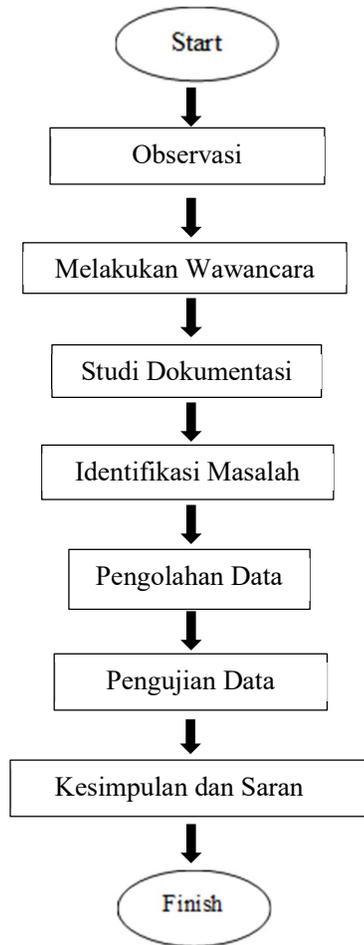
2. Pengamatan

Dalam konteks ilmiah, observasi, dapat diartikan sebagai proses pengamatan yang dilakukan secara sistematis dan disengaja. Proses ini menjadi salah satu cara pengumpulan data serta mengamati dan mencatat gejala atau peristiwa yang sedang terjadi menggunakan indra manusia, terutama mata. Dalam penelitian ini, penulis melakukan observasi terhadap subjek data untuk mencari hubungan antara data produksi pipa dengan kinerja mesin yang digunakan dalam proses produksi pipa.

3. Studi dokumentasi

Pada tahap ini proses yang dilakukan yaitu mengumpulkan dokumen-dokumen penting, seperti data produksi pipa dari setiap mesin yang digunakan, serta gambaran struktur Perusahaan.

D. Langkah Penelitian



Gambar 2 Langkah Penelitian

1. Observasi

Penulis melakukan peninjauan langsung pada proses produksi pipa yang dihasilkan untuk dapat memahami bagaimana proses pembuatan produksi pipa berlangsung.

2. Melakukan Wawancara

Penulis melakukan wawancara dengan Admin Produksi untuk mencapai pemahaman yang lebih mendalam mengenai permasalahan dan kendala yang dihadapi.

3. Studi Dokumentasi

Penulis mengumpulkan dokumen-dokumen penting dari data yang akan diteliti, seperti data produksi pipa dari setiap mesin yang digunakan, serta gambaran struktur perusahaan secara detail.

4. Identifikasi Masalah

Penulis menganalisis hasil wawancara tentang kendala dan permasalahan yang dialami perusahaan serta menganalisis dokumentasi data produksi yang ada.

5. Pengolahan Data

Penulis melakukan pengolahan data berlandaskan data yang penulis terima dari para sumber.

6. Pengujian Data

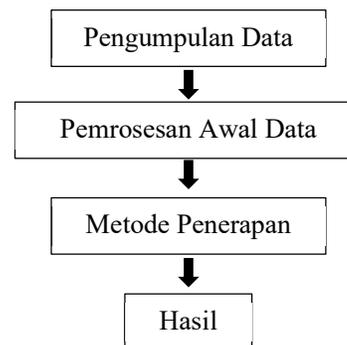
Pada tahap ini, penulis melakukan pengujian data menggunakan Aplikasi Perangkat Lunak RapidMiner agar data yang diperoleh dan diolah dapat memperkuat hasil dari perhitungan Pohon Keputusan C4.5 secara akurat.

7. Kesimpulan dan Saran.

Tahapan proses terakhir ini yang dilakukan berisi penelitian ini adalah pencabutan kesimpulan dan saran. Temuan tersebut diperoleh dari penelitian penerapan pohon keputusan C4.5 pada data produksi pipa di PT Bajamarga Kharisma Utama. Selain itu, perbaikan dan penyempurnaan disarankan untuk penelitian selanjutnya.

E. Metode Pengolahan dan Analisis Data

Metode analisis yang akan digunakan adalah klasifikasi data. Klasifikasi data yaitu proses mendapatkan beberapa pola atau fungsi yang mendefinisikan dan mengasingkan kelas data satu sama lain sehingga dapat digunakan untuk melakukan prediksi pada data yang tidak memiliki kelas data tertentu.



Gambar 3 Tahap Penelitian

1. Pengumpulan Data

Data yang digunakan pada penelitian ini adalah data produksi pipa dan yang dihasilkan oleh PT. BMKU. Data ini merupakan data

pokok yang diperoleh langsung dari PT. BMKU. Data yang dikumpulkan berjumlah 20 data, dimana data ini merupakan data produksi pipa dan diproduksi oleh PT. BMKU akan diterima pada tahun 2023. Data tersebut berisi atribut mesin, bahan mentah, nama barang, ukuran (mm), jumlah produksi (1 bulan), jumlah produksi (2 bulan), jumlah produksi (3 bulan), ketentuan, total produksi dan keterangan.

2. Pemrosesan Awal Data

Pemrosesan awal data adalah proses menghilangkan data duplikat, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan pada data. Selain itu, proses ini juga mencakup pengayaan, yaitu proses memperkaya data yang ada dengan data lain yang relevan. Untuk memperoleh data yang berkualitas, digunakan beberapa metode pemrosesan awal data:

a. Validasi Data

Hal ini dilakukan untuk mengenali dan menghilangkan data ganjil (outlier/noise), data tidak konsisten, dan data tidak lengkap (nilai yang hilang). Data persediaan dan penjualan barang yang dijadikan dataset akan diidentifikasi data ganjil, data yang bertentangan dan data yang tidak lengkap.

b. Pembersihan Data

Pembersihan data ini mencakup beberapa data lain yaitu; menghapus data duplikat, mengamati data yang tidak konsisten, dan membenahi kesalahan data. Dari 20 data yang terkumpul akan dianalisis untuk mengetahui apakah terdapat data yang tidak konsisten atau tidak relevan dan mungkin melanggar pola aturan algoritma yang dihasilkan.

3. Metode Penerapan

Pohon keputusan adalah struktur pohon yang setiap simpulnya mewakili atribut yang diuji. Setiap cabang mewakili subdivisi hasil pengujian, dan simpul daun mewakili kelompok kelas tertentu. Node tingkat teratas dari pohon keputusan adalah akar, yang biasanya merupakan atribut yang memiliki dampak terbesar pada kelas tertentu. Secara umum, pohon keputusan menerapkan strategi top-down untuk mencari solusi. Pada proses klasifikasi, nilai atribut akan diperiksa dengan memeriksa jalur dari node akar ke node daun

(leaf) dan kemudian akan ditentukan kelas baru.

4. Hasil

Setelah hasil klasifikasi diperoleh, hasilnya kemudian dievaluasi menggunakan validasi silang (confusion matriks) untuk melihat akurasi, presisi, dan recall yang diberikan oleh model yang diusulkan.

IV. IMPLEMENTASI DAN PEMBAHASAN

A. Sumber Data

Peneliti mendapatkan data dan informasi yang dibutuhkan sesuai kebutuhan peneliti, sumber data produksi ini di dapat pada masa penelitian selama 3 bulan pada tahun 2023 yang diperoleh melalui observasi dan wawancara secara langsung kepada admin produksi di PT. Baja Marga Kharisma Utama.

B. Karakteristik Data

Data yang digunakan peneliti berupa data hasil produksi barang yang dilakukan dalam 3 bulan di tahun 2023 di PT. Baja Marga Kharisma Utama. Data diperoleh melalui observasi dan wawancara kepada admin produksi dan studi pustaka. Jumlah data yang digunakan adalah 20 data, data ini di dapat dari hasil wawanca kepada admin produksi selama 3 bulan di tahun 2023. Data ini akan diproses menggunakan software rapid miner dengan metode algoritma decision tree (C4.5) untuk mendapatkan hasil produksi terbanyak.

C. Skalabilitas Data

Skalabilitas dalam data mining mengacu pada kemampuan algoritma data mining untuk menangani data dalam jumlah besar secara efisien dan efektif. Artinya algoritma harus mampu memproses data secara tepat waktu, tanpa mengorbankan kualitas hasilnya. Dengan kata lain, algoritma data mining yang scalable harus mampu menangani peningkatan jumlah data tanpa memerlukan peningkatan sumber daya komputasi yang signifikan. Hal ini penting karena jumlah data yang tersedia untuk analisis berkembang pesat, dan kemampuan memproses data tersebut dengan cepat dan akurat sangat penting untuk mengambil keputusan yang tepat. Untuk menentukan skalabilitasnya kita dapat menghitung entropy dan gainnya sebagai berikut:

TABEL 1
TOTAL PRODUKSI

Mesin	Nama Barang	Ukuran (mm)	Jumlah Produksi (1 bulan)	Jumlah Produksi (2 bulan)	Jumlah Produksi (3 bulan)	Ketentuan	Total Produksi	Ket
CH-1	pipa	0,5	2939	5878	8817	2000	17634	meningkat
CH-2	hollow	0,6	1993	3986	5979	2000	11958	meningkat
CH-3	pipa	0,7	657	1314	1971	2000	3942	meningkat
CH-4	hollow	0,8	2422	4844	7266	2000	14532	meningkat
CH-5	hollow	0,9	1557	3114	4671	2000	9342	meningkat
CH-6	hollow	1	2688	5376	8064	2000	16128	meningkat
CH-7	hollow	1,1	980	1960	2940	2000	5880	meningkat
CH-8	hollow	1,2	1762	3524	5286	2000	10572	meningkat
CH-9	hollow	1,3	5581	11162	1743	2000	3486	meningkat
CH-10	hollow	1,4	365	730	1095	2000	2190	meningkat
TW-1	hollow	1,5	430	860	1290	2000	2580	meningkat
TW-2	pipa	1,6	285	570	855	2000	1710	menurun
TW-3	hollow	1,7	356	712	1068	2000	2136	menurun
TW-4	pipa	1,8	389	778	1167	2000	2334	meningkat
TW-5	hollow	1,9	283	566	849	2000	1698	menurun
TW-6	pipa	2	621	1242	1863	2000	3726	meningkat
TW-7	pipa	2,3	1116	2232	3348	2000	6696	meningkat
TW-8	pipa	2,5	556	1112	1668	2000	3336	meningkat
TW-9	pipa	2,8	339	678	1017	2000	2034	meningkat
TW-10	pipa	3	534	1068	1602	2000	3204	meningkat

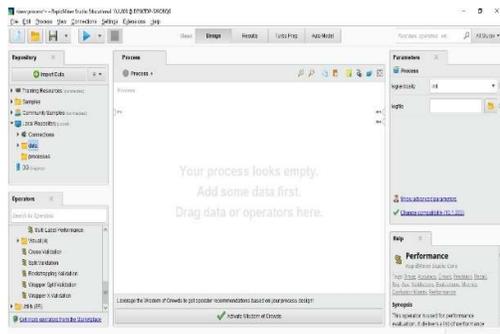
Untuk di tabel ini ada 8 atribut yaitu ada mesin, nama barang, ukuran (milimeter), produksi (1 bulan), produksi (2 bulan), Produksi (3 bulan), Ketentuan, total produksi, dan keterangan. Untuk mencari total produksi menggunakan rumus excel yaitu =SUM(produksi1bulan;produksi2bulan;produksi3bulan) maka akan muncul hasil total produksi setiap mesin seperti tabel diatas.

V. PENGUJIAN

A. Pemrosesan Data

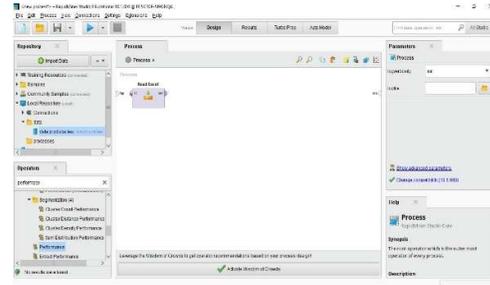
Langkah-langkah dalam mengimplementasi algoritma *Decision Tree (C4.5)* menggunakan software *RapidMiner* adalah sebagai berikut:

1. Buka aplikasi rapidminer yang telah diunduh dan akan muncul tampilan seperti dibawah ini.



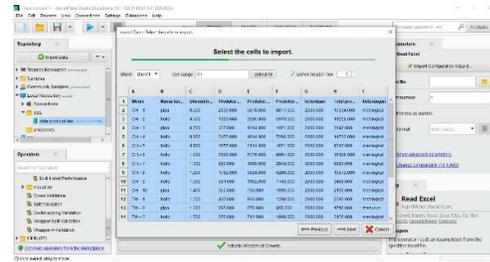
Gambar 4 Tampilan RapidMiner

2. Kemudian pilih operator data excel untuk meletakkan data yang akan kita uji.



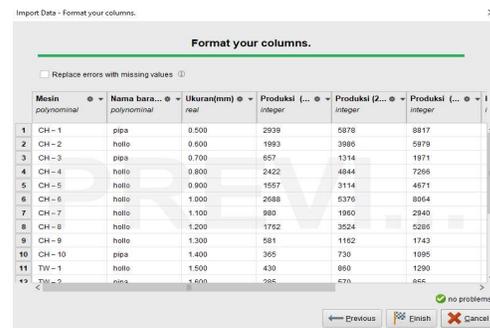
Gambar 5 Operator Pada RapidMiner

3. Pilih data yang yang akan dilakukan pengujian menggunakan rapidminer.



Gambar 6 Data Yang Akan Di Import

4. Sesuaikan data dengan format yang akan dilakukan pengujian.



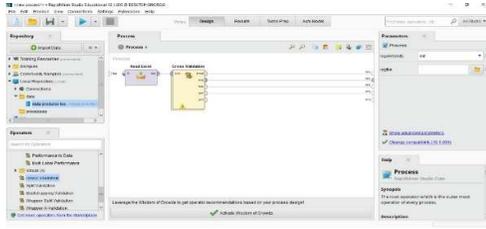
Gambar 7 Penyesuaian Data Yang Di Import

5. Pilih operator cross validation pada kolom pencarian.



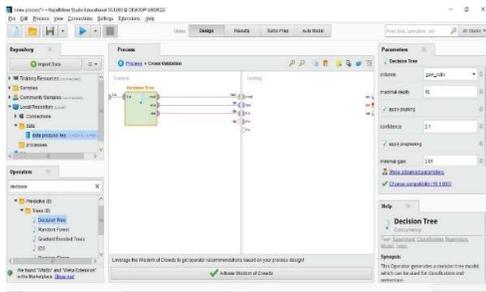
Gambar 8 Operator Cross Validation

6. Kemudian hubungkan cross validation dengan data excel yang sudah dipindahkan terlebih dahulu.



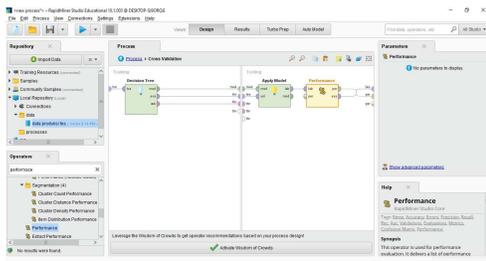
Gambar 9 Operator Validation Terhubung

7. Kemudian Klik 2x operator cross validation maka kita akan ke proses selanjutnya. Lalu pilih operator decision tree di bagian training dan hubungkan dari training ke testing.



Gambar 10 Menghubungkan Training Ke Testing

8. Lalu di bagian testing kita masukan operator apply model dan performance dibagian kolom pencarian, setelah itu hubungkan setiap operatornya seperti gambar ini.



Gambar 11 Menghubungkan Apply dan Performance

B. Output Pemrosesan Data

Setelah melakukan running data dengan melakukan langkah langkah diatas, maka hasil pengujian data yang didapatkan yaitu sebagai berikut:

1. Hasil keputusan produksi yang meningkat dan menurun pada beberapa mesin.



Gambar 12 Hasil Pada Setiap Mesin

2. Deskripsi Hasil Analisis

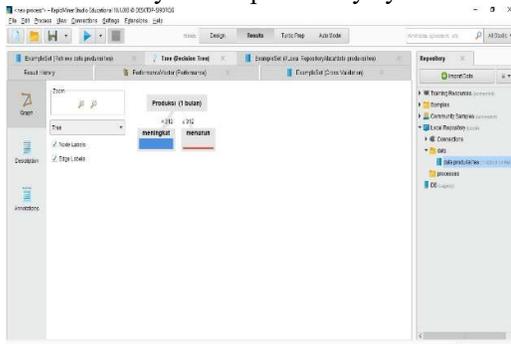
- Mesin = CH – 1 : meningkat {meningkat=1, menurun=0}
- Mesin = CH – 10: meningkat {meningkat=1, menurun=0}
- Mesin = CH – 2 : meningkat {meningkat=1, menurun=0}
- Mesin = CH – 3 : meningkat {meningkat=1, menurun=0}
- Mesin = CH – 4 : meningkat {meningkat=1, menurun=0}
- Mesin = CH – 5 : meningkat {meningkat=1, menurun=0}
- Mesin = CH – 6 : meningkat {meningkat=1, menurun=0}
- Mesin = CH – 7 : meningkat {meningkat=1, menurun=0}
- Mesin = CH – 8 : meningkat {meningkat=1, menurun=0}
- Mesin = CH – 9 : meningkat {meningkat=1, menurun=0}
- Mesin = TW – 10: meningkat {meningkat=1, menurun=0}
- Mesin = TW – 1 : meningkat {meningkat=1, menurun=0}
- Mesin = TW – 2 : menurun {meningkat=0, menurun=1}
- Mesin = TW – 3 : meningkat {meningkat=1, menurun=0}
- Mesin = TW – 4 : meningkat {meningkat=1, menurun=0}
- Mesin = TW – 5 : menurun {meningkat=0, menurun=1}
- Mesin = TW – 6 : meningkat {meningkat=1, menurun=0}

Mesin = TW - 7 : meningkat {meningkat=1, menurun=0}

Mesin = TW - 8 : meningkat {meningkat=1, menurun=0}

Mesin = TW - 9 : meningkat {meningkat=1, menurun=0}

3. Hasil pohon keputusan produksi terbanyak setiap bulan nya yaitu:



Gambar 13 Total Hasil 3 bulan

4. Deskripsi decision tree nya yaitu :

Produksi (3 bulan) > 312: meningkat {meningkat=18, menurun=0}

Produksi (3 bulan) ≤ 312: menurun {meningkat=0, menurun=2}

Setelah dilakukan proses running terbentuklah pohon keputusan dan juga description rules. Pembentukan pohon keputusan dan description rules diperoleh setelah menghubungkan operator *cross validation*, *decision tree*, *apply model*, dan *performance*. Dan untuk accuracynya yaitu pred.meningkat sebesar 90% dan untuk pred.menurunnya sebesar 10%. Dan produksi pipa terbanyak diperoleh pipa 3,0mm.

VI. KESIMPULAN

Pada penelitian ini penulis melakukan analisis data mining di PT. Baja Marga Kharisma Utama menggunakan data produksi untuk mencari produksi terbanyak disetiap mesinnya dan produksi bulannya. Penelitian ini menggunakan software rapidminer untuk membantu pengolahan data menghasil output menjadi lebih mudah dan hasil penelitian ini menggunakan metode decision tree (C4.5). Berdasarkan hasil penelitian, dapat disimpulkan bahwa penerapan data mining

dengan menggunakan metode decision tree sangat bermanfaat untuk mengetahui produksi meningkat atau tidaknya selama 3 bulan dan mengetahui mesin mana produksi terbanyak selama 3 bulan dan produksi terbanyak diperoleh pipa 3,0 mm dan untuk mesin produksi terbanyak diperoleh mesin ch-2.

Berisikan Kesimpulan dari hasil penelitian tersebut berdasarkan hasil implementasi sistem yang didapatkan. Serta pengembangan yang memungkinkan untuk dilakukan di masa yang akan datang.

REFERENSI

- [1] B. Ginting and F. Riandari, "Implementasi Metode K-Means Clustering Dalam Pengelompokan Bibit Tanaman Kopi Arabika," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 3, no. 2, 2020.
- [2] D. Hedyati and I. M. Suartana, "Penerapan Principal Component Analysis (PCA) Untuk Reduksi Dimensi Pada Proses Clustering Data Produksi Pertanian Di Kabupaten Bojonegoro."
- [3] F. Gunadi and S. R. Widiyanto, *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) Perbandingan Data Warehouse Cloud Computing Menggunakan Konvensional Berbasis Kriptografi*.
- [4] N. Nestary, "PERANCANGAN SISTEM INFORMASI PENJUALAN PADA TOKO STOCK POINT LILY BERBASIS PHP MYSQL," 2020.
- [5] R. Rusliyawati and A. Wantoro, "Decision support system model using FIS Mamdani for determining tire pressure," *Jurnal Teknologi dan Sistem Komputer*, vol. 9, no. 1, pp. 56–63, Jan. 2021, doi: 10.14710/jtsiskom.2020.13776.
- [6] A. A. Aldino and H. Sulistiani, "DECISION TREE C4.5 ALGORITHM FOR TUITION AID GRANT PROGRAM CLASSIFICATION (CASE STUDY: DEPARTMENT OF INFORMATION SYSTEM, UNIVERSITAS TEKNOKRAT INDONESIA)," 2020.

- [7] A. Erfina, N. Destria Arianti, P. Studi Sistem Informasi, and U. Nusa Putra Sukabumi Jl Raya Cibolang No, "PENERAPAN METODE DATA MINING TERHADAP DATA TRANSAKSI PENJUALAN MENGGUNAKAN ALGORITMA APRIORI (STUDI KASUS: TOKO FASENTRIO FANCY)," vol. 2, no. 3, pp. 2715–1875, 2020.
- [8] A. Kurnia and A. Haidar Mirza, "Penerapan Decision Tree Data Mining Pada Produksi Kelapa Sawit PT Hindoli Di Sungai Lilin Kabupaten Musi Banyuasin," 2020. [Online]. Available: <http://journal.jis-institute.org/index.php/jpsii/index>
- [9] R. Nofitri and N. Irawati, "INTEGRASI METODE NEIVE BAYES DAN SOFTWARE RAPIDMINER DALAM ANALISIS HASIL USAHA PERUSAHAAN DAGANG," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 6, no. 1, pp. 35–42, Dec. 2019, doi: 10.33330/jurteksi.v6i1.393.
- [10] N. P. Setyadi, "Penerapan Data Mining Untuk Prediksi Hasil Produksi Karet Menggunakan Algoritma Decision Tree C4.5," 2022.
- [11] R. Sulastri, "Identifikasi Tingkat Penjualan Produk Herbal HWI Menggunakan Algoritma C4.5," *Jurnal Informatika Ekonomi Bisnis*, pp. 84–89, Sep. 2022, doi: 10.37034/infv4i3.141.
- [12] E. E. Azwanti Nurul, "Analisa Kepuasan Konsumen Menggunakan Algoritma C4.5," 2020.