

# Data Mining untuk Pemeliharaan Prediktif Mesin Produksi Berdasarkan Database Kerusakan Mesin Menggunakan Naïve Bayes Classifier

Fajar Martha Subqi dan Dyah Anggraini

Universitas Gunadarma

Jakarta, Indonesia

E-mail: fajar.martha@gmail.com, d\_anggraini@staff.gunadarma.ac.id

## Abstrak

*Preventive maintenance* adalah kegiatan pemeliharaan dan perawatan yang dilakukan untuk mencegah timbulnya kerusakan-kerusakan yang tidak terduga dan menemukan kondisi yang dapat menyebabkan fasilitas produksi mengalami kerusakan pada saat proses produksi dilakukan. Agar kegiatan *preventive maintenance* tidak mengganggu proses produksi, diperlukan penjadwalan mesin mana saja yang perlu dilakukan tindakan *preventive maintenance*. Dengan adanya aplikasi *data mining* saat ini, akan sangat membantu menemukan informasi yang tersembunyi pada *history* kerusakan mesin di dalam sebuah database dengan menggunakan algoritma tertentu. Penelitian ini akan membahas data kasus pada perusahaan *spare part* otomotif. Dimana data *history* kerusakan mesin yang akan diolah digunakan untuk menentukan mesin mana saja yang perlu dilakukan kegiatan *preventive maintenance* dengan menggunakan metode naïve bayes classifier. Data yang akan diolah diambil dari sepuluh besar mesin yang mengalami kerusakan, sample data diambil secara acak (*random sampling*) dengan menggunakan enam parameter meliputi nomor mesin, masalah, shift, durasi waktu, tipe *breakdown* dan tingkat resiko. Implementasi pengolahan *data mining* dengan menggunakan aplikasi rapid miner 9.6, dengan proporsi data 80/20 yaitu data training sebanyak 754 data kerusakan dan data testing 189 data kerusakan dari hasil pengujian didapat dengan nilai akurasi 93,65 % untuk menentukan mesin yang perlu dilakukan kegiatan *preventive maintenance*.

**Kata Kunci:** *preventive maintenance*, naïve bayes classifier, rapid miner

## Pendahuluan

Pemanfaatan teknologi infomasi dalam dunia industri sangatlah penting, untuk menunjang kesinambungan bisnis itu sendiri. Di era evolusi industri 4.0, banyak perusahaan baik skala besar maupun kecil sangat mengandalkan kemajuan teknologi informasi untuk menunjang akfititas bisnis perusahaan. Salah satu pemanfaatan teknologi infomasi adalah dalam hal pengolahan data.

Pengertian data menurut Vercellis [1] adalah sebuah representasi fakta yang tersusun secara terstruktur. Sedangkan menurut Wawan dan Munir [2] data adalah nilai yang merepresentasikan deskripsi dari suatu objek atau kejadian (*event*). Menurut Jhon J.Long [3] mengatakan bahwa data adalah suatu istilah majemuk dari fakta yang mengandung arti yang dihubungkan dengan kenyataan, simbol, gambar, angka, huruf yang menunjukkan suatu ide, objek, kondisi atau situasi. Jadi dapat disimpulkan bahwa data adalah fakta atau bagian dari fakta yang belum tersusun yang mempunyai arti yang dihubungkan dengan kenyataan, sim-

bol,gambar, kata, angka dan lain – lain.

Penelitian ini dilakukan di sebuah perusahaan otomotif yang memiliki kurang lebih 500 unit mesin produksi. Pihak operator produksi sangat bergantung kepada mesin – mesin tersebut untuk merubah bahan setengah jadi menjadi bahan jadi. Oleh karena itu perawatan terhadap mesin produksi sangat diperlukan untuk menghindari kerusakan yang dapat mengakibatkan *stop line* di area produksi. Saat ini perusahaan ini memiliki aplikasi MBA (*Machine Breakdown Action*) yaitu sebuah aplikasi web untuk mencatat dan memonitor mesin produksi yang mengalami kerusakan, berapa lama perbaikannya, siapa yang mengerjakannya dan lain –lain. Akan tetapi *record* data dari kerusakan mesin belum diolah untuk memprediksi mesin apa yang akan rusak, sehingga dapat dilakukan tindakan *preventive maintenance* untuk mengurangi potensi *stop line* diarea prosuksi yang mengakibatkan kerugian secara finansial. Berdasarkan permasalahan - permasalahan yang telah dikemukakan diatas, maka batasan masalah dari penelitian ini adalah :

1. Sistem informasi MBA (*Machine Breakdown Action*) mencatat semua *record* kerusakan mesin di 3 pabrik.
2. Untuk pengolahan *data mining* hanya akan dilakukan di perusahaan otomotif di kawasan industri MM2100.
3. *Data record* kerusakan mesin produksi akan diambil sample dari Maret 2019 sampai dengan Juli 2020.
4. Metode yang akan digunakan untuk *data mining* Naive Bayes Classifier.

Beberapa manfaat yang dapat diambil dari penelitian ini adalah sebagai berikut:

1. Dapat memberikan masukan untuk memprediksi mesin produksi yang akan mengalami kerusakan.
2. Meningkatkan produktivitas dan efisiensi pegawai *maintenance department*.
3. Melakukan penjadwalan untuk *preventive maintenance* mesin produksi.
4. Dapat memperkirakan dan menentukan spare part mesin yang akan digunakan untuk melakukan *preventive maintenance*.
5. Hasil penelitian ini diharapkan dapat digunakan sebagai bahan referensi atau rujukan bagi para pembaca, terutama mereka yang bermaksud melakukan penelitian dengan topik yang sama seperti yang saat ini dilakukan oleh peneliti.

## Knowledge Discovery in Database

*Knowledge Discovery in Database* (KDD) adalah sebuah proses pencarian atau penemuan pengetahuan (nilai tambah) di dalam sebuah database, karena *data mining* adalah suatu rangkaian proses [4].

## Data Mining

Merupakan sebuah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [5]. *Data mining* mempunyai fungsi yang penting untuk membantu mendapatkan informasi yang berguna serta meningkatkan pengetahuan bagi pengguna. Pada dasarnya, data mining mempunyai lima fungsi dasar yaitu:

1. Fungsi Prediksi (*prediction*)

Proses untuk menemukan pola tertentu dari suatu data dengan menggunakan beberapa variabel untuk memprediksikan variabel lain yang belum diketahui jenis atau nilainya.

2. Fungsi Deskripsi (*description*).

Proses untuk memahami lebih jauh tentang karakteristik penting dari data dalam suatu basis data.

3. Fungsi Klasifikasi (*classification*)

Klasifikasi merupakan suatu proses menemukan model yang menggambarkan dan membedakan *class* atau konsep dari suatu data. Proses yang digunakan untuk mendeskripsikan data yang penting serta dapat meramalkan kecenderungan data pada masa depan.

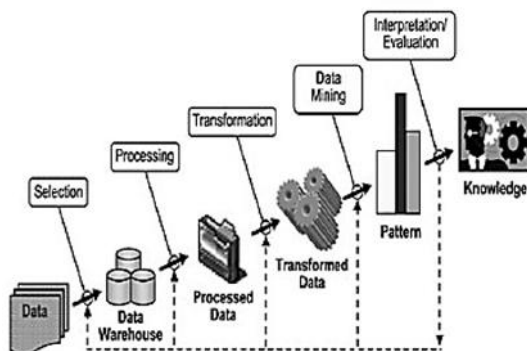
4. Fungsi Asosiasi (*association*)

Proses ini digunakan untuk menemukan suatu hubungan yang terdapat pada nilai atribut yang muncul dalam suatu waktu dari sekumpulan data.

5. *Forecasting*

Memperkirakan atau meramalkan nilai pada masa yang akan datang berdasarkan pola – pola dengan sekumpulan data yang besar.

Tahapan yang dilakukan pada proses *data mining* diawali dari seleksi data dari data sumber ke data target, tahap *preprocessing* untuk memperbaiki kualitas data, transformasi *data mining* serta tahap interpretasi dan evaluasi yang menghasilkan *output* berupa pengetahuan baru yang diharapkan memberikan kontribusi yang lebih baik. Secara detail dijelaskan sebagai berikut [6]:



Gambar 1: Tahapan data mining

1. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi di KDD dimulai. Data hasil seleksi yang digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. *Pre – Processing/ Cleaning*

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning*

pada data yang menjadi fokus KDD. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data.

### 3. Transformation

Pengkodean adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses pengkodean dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

### 4. Data Mining

Proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

### 5. Interpretation/ Evaluation

Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

### 6. Knowledge

Merupakan tahap terakhir dimana pengetahuan yang telah ditemukan secara visual ditampilkan kepada *user*. Tahap penting ini menggunakan teknik visualisasi untuk membantu *user* dalam mengerti dan menginterpretasikan hasil dari *data mining*.

## Naïve Bayes

Sebuah metoda pembelajaran mesin yang memanfaatkan perhitungan probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. Ciri dari naïve bayes adalah asumsi independensi yang kuat (naïve).

Dari sisi performa naïve bayes memiliki kecepatan dan keakuratan yang tinggi bila di implementasikan dalam data yang ukurannya besar. Metode ini sering digunakan dalam menyelesaikan masalah dalam bidang mesin pembelajaran karena metode ini dikenal memiliki tingkat akurasi yang tinggi dengan perhitungan sederhana. Teorema bayes merupakan dasar aturan dari naïve bayes *classifier* berikut teorema bayes akan disajikan

pada persamaan. Bentuk umum teorema bayes adalah sebagai berikut:

$$P(H/X) = \frac{P(X/H)P(H)}{P(X)}$$

Dimana:

$X$  = Data dengan kelas yang belum diketahui

$H$  = Hipotesa data  $X$  merupakan suatu kelas spesifik

$P(H|X)$  = Probabilitas hipotesis  $H$  berdasarkan kondisi  $X$  (*posterior probability*)

$P(H)$  = Probabilitas hipotesis  $H$  (*prior probability*)

Klasifikasi Naive Bayes untuk menghitung probabilitas suatu peristiwa dalam langkah-langkah berikut:

1. Hitung probabilitas sebelumnya untuk label kelas yang diberikan.
2. Temukan probabilitas peluang dengan setiap atribut untuk setiap kelas.
3. Masukkan nilai ini dalam Formula Bayes dan hitung probabilitas posterior.
4. Lihat kelas mana yang memiliki probabilitas lebih tinggi, mengingat input milik kelas probabilitas lebih tinggi

## Rapidminer

Merupakan sebuah *platform* perangkat lunak ilmu data yang dikembangkan oleh sebuah perusahaan dengan nama yang sama yang menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam, penambahan teks, dan analisis prediktif. Hal ini digunakan untuk bisnis dan komersial, juga untuk penelitian, pendidikan, pelatihan, rapid *prototyping*, dan pengembangan aplikasi serta mendukung semua langkah dalam proses pembelajaran mesin termasuk persiapan data, hasil visualisasi, validasi model, dan optimasi.



Gambar 2: Aplikasi rapid miner studio

## Metode Penelitian

Penelitian ini dilakukan di perusahaan otomotif yang terletak di Kawasan Industri MM2100 Cikarang Kabupaten Bekasi Provinsi Jawa Barat. Penelitian ini dilakukan pada bulan Mei 2020 hingga bulan Juli 2020. Pengumpulan data pada penelitian ini menggunakan beberapa metode. Adapun metode yang digunakan sebagai berikut:

### 1. Studi literatur

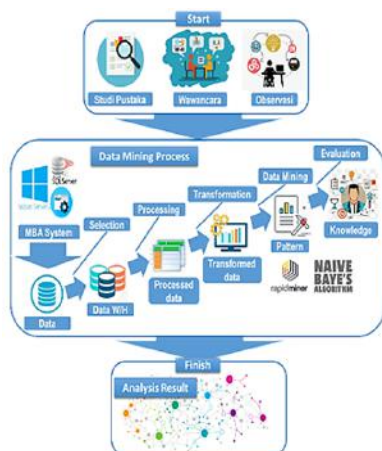
Penulis melakukan penelitian dengan membaca dan mempelajari jurnal, buku literatur yang dapat dijadikan pendukung dalam penyelesaian penulisan ini. Penulis juga mempelajari proses bisnis yang ada pada perusahaan otomotif dalam melakukan aktifitas pencatatan *record* kerusakan mesin pada aplikasi MBA (*Machine Breakdown Action*). Dimana data yang diolah dalam penelitian ini merupakan data primer. Data primer adalah data yang dikumpulkan dan diolah sendiri oleh peneliti langsung dari subjek atau objek penelitian.

### 2. Wawancara

Wawancara ini dilakukan di kantor *maintenance department* di perusahaan otomotif dengan karyawan yang memiliki wewenang dalam aplikasi MBA sekaligus berdiskusi singkat untuk memudahkan pengambilan keputusan dari sistem yang akan dikembangkan.

### 3. Observasi

Observasi dilakukan untuk mempelajari yang berjalan pada perusahaan otomotif, dengan memperhatikan beberapa tahapan dalam proses pencatatan record kerusakan mesin.



Gambar 3: Tahapan penelitian

Beberapa penelitian telah banyak dilakukan dengan menggunakan teknik *data mining* untuk menggali berbagai informasi dari sebuah data

kerusakan mesin produksi. Penelitian tersebut membahas tentang topik yang terkait dengan penelitian penulis.

1. Perancangan Penjadwalan *Preventive Maintenance* pada PT. Artha Prima Sukses Makmur tahun 2015 [7].
2. Analisa Perancangan Jadwal *Preventive dan Predictive Maintenance* pada Mesin Kapal di Daerah Limbangan Indramayu ” tahun 2017 [8].
3. Perbandingan Performansi Teknik Klasifikasi *Breakdown* Mesin pada Proses Produksi Pembuatan *Battery* Mobil” tahun 2019 [9].
4. Analisa Data Gangguan Kerusakan Mesin Produksi Menggunakan Teknik *Association Rules*” tahun 2019 [10].
5. Sistem Pakar Diagnosa Penyakit Batu Ginjal Menggunakan Teorema Bayes” tahun 2016 [11].

Sumber data untuk penelitian diunduh dari aplikasi MBA (*Machine Breakdown Action*) berupa data kerusakan mesin produksi. Dari keseluruhan data kerusakan mesin tersebut, untuk penelitian ini tidak semua data kerusakan mesin akan digunakan hanya sepuluh besar mesin produksi yang sering mengalami kerusakan, *sample* data diambil secara acak.

Tabel 1: Sepuluh besar mesin produksi

Machine Number	Data Training	Data Testing
ME-7214-13-012	68	15
ME-7214-19-01	58	30
ME-7215-15-0019	65	15
ME-7215-15-007	60	15
ME-7217-12-012	87	22
ME-7217-15-001	70	15
ME-7217-15-009	141	21
ME-7217-15-014	60	15
ME-7252-08-010	50	15
ME-7252-15-003	95	26
<b>Total</b>	<b>754</b>	<b>189</b>

Dengan perbandingan data training dan data testing 80/20, yang menghasilkan data training sebanyak 754 *record* dan data testing sebanyak 189 *record*. Dari data tersebut hanya diambil beberapa *attribute* untuk pengolahan data mining yaitu:

1. *Machine Number*
2. *Problem*
3. *Shift*
4. *Durasi*
5. *Breakdown Type*

6. Tingkat Resiko

Proses data mining terbagi dalam beberapa tahap yaitu :

1. Data selection

Pada tahap ini data akan diseleksi dengan cara melihat kecenderungan data/kesesuaian data topik/ judul penelitian yang akan diteliti, adapun data yang akan diteliti diperoleh dari database aplikasi MBA (*Machine Breakdown Action*).

Tabel 2: Data kerusakan mesin produksi

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
ME-2017-15-009	issue	status	issuor name	category	machine number	spesifikasi	produksi	stopper	broken	broken	broken	broken	broken	broken	broken	broken	broken	broken	broken	broken	broken

Pada Tabel 2 dapat kita lihat terdapat data yang tidak sesuai, maka dari itu perlu dilakukan seleksi.

2. Data preprocessing/cleaning

Dalam tahapan *preprocessing* yaitu menghilangkan *noise* data yang tidak konsisten, data *error* atau data kosong. Perlu dilakukan proses penghapusan data pada attribute *machine number* dikarenakan terdapat data yang kosong atau tidak diisi ketika proses perbaikan mesin.

Tabel 3: Data cleaning attribute machine number

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
ME-2017-15-009	issue	status	issuor name	category	machine number	spesifikasi	produksi	stopper	broken	broken	broken	broken	broken	broken	broken	broken

3. Data Transformation

Data akan berubah atau bersatu menjadi bentuk yang tepat untuk menambang dengan ringkasan performa atau operasi agresi. Adapun dalam tahap

ini data akan diubah menjadi bentuk yang sesuai untuk proses data mining.

Dalam penelitian ini akan dilakukan uji coba secara teoritis dan menggunakan *software* Rapid-Miner, maka data yang telah melalui proses sebelumnya akan ditransporasikan agar sesuai dengan algoritma yang dipakai yaitu naïve bayes *classifier*. Dalam tahapan ini *attribute* yang akan dipakai akan diberi label mengikuti kondisi data-data pada *attribute* tersebut.

- Klasifikasi *attribute problem*

Berdasarkan hasil dataset yang siap diolah terdapat 227 jenis problem dari kerusakan mesin.

- Klasifikasi *attribute breakdown type*

*Breakdown type* Dikelompokkan menjadi 6 kategori : *Electric, Electronic, Hydraulic, Lain-lain, Mechanic* dan *Pneumatic*.

- Klasifikasi *attribute tingkat resiko*

Tingkat resiko dikelompokkan menjadi 3 kategori yaitu : Rendah, Sedang dan Tinggi.

- Label *preventive maintenance*

Untuk label *preventive maintenance* didasarkan 2 keputusan “YA” dan “TIDAK”.

Tabel 4: Dataset format

Machine Number	Problem	Shift	Durasi (min)	Breakdown Type	Tingkat Resiko	Preventive Maintenance
ME-2217-15-009	Hole punch	1	4	Mechanic	Rendah	TIDAK
ME-2217-15-009	Hole punch	2	10	Electric	Rendah	TIDAK
ME-2217-15-009	Hose	1	15	Pneumatic	Rendah	TIDAK
ME-2217-15-009	Hose	2	26	Pneumatic	Rendah	TIDAK
ME-2217-15-009	Hose	2	33	Pneumatic	Rendah	TIDAK
7252-NEW	Screw roller	2	73	Mechanic	Sedang	YA
7252-NEW	Breaker line up	2	18	electric	Rendah	TIDAK
7252-NEW	Config parameter	2	47	Mechanic	Rendah	TIDAK
7252-NEW	Block cutter	1	110	Mechanic	Sedang	YA
7252-NEW	Plate spring	1	27	Mechanic	Rendah	TIDAK
ME-2217-12-012	Tool screw	1	12	Mechanic	Rendah	TIDAK
ME-2217-12-012	Config parameter	2	87	Mechanic	Sedang	YA
ME-2217-12-012	Trap oring	1	17	Mechanic	Rendah	TIDAK
ME-2217-12-012	Trap kembang	2	2	Mechanic	Rendah	TIDAK
ME-2217-12-012	Auto switch cylinder	1	24	Electric	Rendah	TIDAK
ME-2217-15-001	Conveyor chain	2	19	Mechanic	Rendah	TIDAK
ME-2217-15-001	Roller	1	187	Mechanic	Tinggi	YA
ME-2217-15-001	Bearing sproket	1	29	Mechanic	Rendah	TIDAK
ME-2217-15-001	Conveyor chain	2	20	Mechanic	Rendah	TIDAK
ME-2217-15-001	Bearing roller	1	135	Mechanic	Tinggi	YA
ME-2217-15-001	Conveyor belt	2	18	Mechanic	Rendah	TIDAK
ME-2217-15-001	Conveyor chain	1	30	Mechanic	Rendah	TIDAK
ME-2217-15-001	Conveyor belt	2	83	Mechanic	Sedang	YA
ME-2217-15-001	Conveyor chain	1	19	Electric	Rendah	TIDAK
ME-2217-15-001	Roller	2	187	Pneumatic	Tinggi	YA
ME-2217-15-001	Bearing sproket	1	29	Mechanic	Rendah	TIDAK
ME-2214-13-012	Guide shammer	1	15	lain-lain	Rendah	TIDAK
ME-2214-13-012	V-belt	1	153	Mechanic	Tinggi	YA
ME-2214-13-012	Alarm press	1	15	Electric	Rendah	TIDAK
ME-2214-13-012	Socket power	2	97	Electric	Sedang	YA
ME-2214-13-012	Regulator helium	1	279	Pneumatic	Tinggi	YA
ME-2217-13-014	label sensor	1	81	Electric	Rendah	TIDAK
ME-2217-13-014	Pusher resin	1	13	Mechanic	Rendah	TIDAK
ME-2217-13-014	Sensor fiber optic	2	28	Electric	Rendah	TIDAK

Pada label status akan dijadikan sebagai proses untuk melakukan proses eksekusi pada program rapid miner sehingga harus digolongkan menjadi 2 kriteria yaitu YA dan TIDAK. Pada *attribute* dengan tingkat resiko rendah akan diberi label *preventive maintenance* YA dan TIDAK tergantung dengan ketentuan dari divisi *maintenance*. Jika kerusakan mesin dengan durasi lebih dari 60 menit dan kerusakan mesin berulang dengan kerusakan yang sama lebih dari 5 kali maka akan diberi label YA, jika tidak memenuhi ketentuan tersebut maka akan diberi label TIDAK. Sedangkan pada *attribute* dengan tingkat resiko sedang dan tinggi akan diberi label *preventive maintenance*

YA. Setelah hasil proses klasifikasi tersebut, maka langkah selanjutnya agar dapat diolah dengan menggunakan *software data mining* yaitu Rapid-Miner. Data yang sudah dilakukan pada proses tahapan diatas ini adalah data yang sudah siap atau sudah bersih dari data *redundancy* dan data yang *loss* (kosong). Pada Tabel 4 merupakan dataset yang siap diolah untuk dilakukan proses *data mining*.

#### 4. Data Mining

Perhitungan *data mining* dengan menggunakan naïve bayes classifier:

1. Menghitung jumlah kelas/label

Dari total jumlah data set 754 terdapat 365 label “YA” dan 389 label “TIDAK”. Jadi:

$$P(H=YA) = \text{jumlah label YA} / \text{jumlah data} = 365/754 = 0.4841$$

$$P(H=TIDAK) = \text{jumlah label TIDAK} / \text{jumlah data} = 389/754 = 0.5159$$

Pada tabel 3.5 dapat dilihat nilai probabilitas dari kelas.

Tabel 5: Probabilitas kelas/label

Probabilitas	Preventive Maintenance	
	YA	TIDAK
	0,4841	0,5159

Probabilitas *Preventive Maintenance* YA TIDAK 0,4841 0,5159 Menghitung jumlah kasus yang sama dengan *class* yang sama, terdapat 6 attribute yang harus dicari nilai probabilitasnya. PM (*Preventive Maintenance*)

2. Menghitung probabilitas *attribute machine number* untuk menghitung nilai probabilitas *machine number* adalah sebagai berikut:.

Perhitungan manual *machine number*:

$$\text{Jumlah kerusakan mesin ME-7217-15-009} = 141$$

$$\text{Jumlah Preventive Maintenance "YA"} = 78$$

$$\text{Jumlah Preventive Maintenance "TIDAK"} = 63$$

$$P(\text{Machine number} = M_x | PM=YA)$$

$$P(\text{ME-7217-15-009} = YA)$$

$$= (78/141) * (141/754) / 0.4841$$

$$= (0.5532 * 0.1870) / 0.4841$$

$$= 0.2137$$

$$P(\text{Machine number} = M_x | PM=TIDAK)$$

$$P(\text{ME-7217-15-009} = TIDAK)$$

$$= (63/141) * (141/754) / 0.5159$$

$$= (0.4468 * 0.1870) / 0.5159$$

$$= 0.1620$$

Tabel 6: Probabilitas attribute Machine Number

Machine Number	Preventive Maintenance	
	YA	TIDAK
ME-7217-15-009	0.2137	0.1620
ME-7214-13-012	0.1014	0.0797
ME-7214-19-01	0.0438	0.1080
ME-7215-15-0019	0.0740	0.0977
ME-7215-15-007	0.0658	0.0925
ME-7217-12-012	0.1178	0.1131
ME-7217-15-001	0.0685	0.1157
ME-7217-15-014	0.0658	0.0925
ME-7252-08-010	0.0521	0.0797
ME-7252-15-003	0.1973	0.0591

3. Menghitung probabilitas *attribute problem*

Untuk menghitung nilai probabilitas problem adalah sebagai berikut:.

Perhitungan manual problem:

$$\text{Jumlah problem adjust coolant} = 3$$

$$\text{Jumlah Preventive Maintenance "YA"} = 0$$

$$\text{Jumlah Preventive Maintenance "TIDAK"} = 3$$

$$P(\text{Adjust coolant} = P_x | PM=YA)$$

$$P(\text{Adjust coolant} = YA) = 0/0.4841 = 0$$

$$P(\text{Adjust coolant} = P_x | PM=TIDAK)$$

$$P(\text{Adjust coolant} = TIDAK)$$

$$= (3/754) / 0.5159$$

$$= (0.0040) / 0.5159$$

$$= 0.0077$$

Tabel 7: Probabilitas attribute problem

Problem	Preventive Maintenance	
	YA	TIDAK
Adjust coolant	0,0000	0,0077
Air & Gas flow	0,0027	0,0000
Air fitting	0,0027	0,0000
Alarm chamber	0,0000	0,0026
Alarm press	0,0356	0,0000

4. Menghitung probabilitas *attribute shift*

Untuk menghitung nilai probabilitas *shift* adalah sebagai berikut:.

Perhitungan manual shift:

$$\text{Jumlah shift 1} = 428$$

$$\text{Jumlah Preventive Maintenance "YA"} = 199$$

$$\text{Jumlah Preventive Maintenance "TIDAK"} = 229$$

$$P(\text{Shift 1} = S_x | PM=YA)$$

$$P(\text{Shift 1} = YA)$$

$$= (199/428) * (428/754) / 0.4841$$

$$= (0.4650 * 0.5676) / 0.4841$$

$$\begin{aligned}
 &= 0.5452 \\
 P(\text{Shift 1} = S_x | \text{PM} = \text{TIDAK}) \\
 P(\text{Shift 1} = \text{TIDAK}) \\
 &= (299/428) * (428/754) / 0.5159 \\
 &= (0.5350 * 0.5676) / 0.5159 \\
 &= 0.5587
 \end{aligned}$$

Tabel 8: Probabilitas attribute shift

Shift	Preventive Maintenance	
	YA	TIDAK
1	0,5452	0,5887
2	0,4548	0,4113

5. Menghitung probabilitas nilai *attribute* durasi

Untuk menghitung nilai probabilitas durasi adalah sebagai berikut:

Perhitungan manual durasi:

Jumlah durasi 2 menit = 1

Jumlah *Preventive Maintenance* “YA” = 0

Jumlah *Preventive Maintenance* “TIDAK” = 1

$P(\text{Durasi 2} = D_x | \text{PM} = \text{YA})$

$P(\text{Durasi 2} = \text{YA})$

$$= 0/0.4841$$

$$= 0$$

$P(\text{Durasi 2} = D_x | \text{PM} = \text{TIDAK})$

$P(\text{Durasi 2} = \text{TIDAK})$

$$= (1/754) / 0.5159$$

$$= (0.0013) / 0.5159$$

$$= 0,0026$$

Tabel 9: Probabilitas attribute durasi

Durasi	Preventive Maintenance	
	YA	TIDAK
2	0,0000	0,0026
3	0,0082	0,0051
4	0,0055	0,0154
5	0,0137	0,0231
6	0,0110	0,0231

6. Menghitung nilai probabilitas *attribute breakdown type*

Untuk menghitung nilai probabilitas *breakdown type* adalah sebagai berikut :

Perhitungan manual *breakdown type electric*:

Jumlah *breakdown type electric* = 260

Jumlah *Preventive Maintenance* “YA” = 110

Jumlah *Preventive Maintenance* “TIDAK” = 150

$P(\text{Breakdown type} = B_x | \text{PM} = \text{YA})$

$P(\text{breakdown type electric} = \text{YA})$

$$= (110/260) * (260/754) / 0.4841$$

$$= (0.4231 * 0.3448) / 0.4841$$

$$= 0.3014$$

$P(\text{Breakdown type} = B_x | \text{PM} = \text{TIDAK})$

$P(\text{breakdown type electric} = \text{TIDAK})$

$$= (150/260) * (260/754) / 0.5159$$

$$= (0.5769 * 0.3448) / 0.5159$$

$$= 0.3856$$

Tabel 10: Probabilitas *attribute breakdown type*

Breakdown	Preventive Maintenance	
	YA	TIDAK
Electric	0,3014	0,3856
Electronic	0,0082	0,0129
Hydraulic	0,0027	0,0154
lain-lain	0,0877	0,0925
Mechanic	0,5096	0,4139
Pneumatic	0,0904	0,0797

7. Menghitung nilai probabilitas *attribute* tingkat resiko

Untuk menghitung nilai probabilitas tingkat resiko adalah sebagai berikut.

Perhitungan manual tingkat resiko rendah:

Jumlah tingkat resiko rendah = 607

Jumlah *Preventive Maintenance* “YA” = 219

Jumlah *Preventive Maintenance* “TIDAK” = 338

$P(\text{Tingkat resiko} = T_x | \text{PM} = \text{YA})$

$P(\text{tingkat resiko rendah} = \text{YA})$

$$= (219/607) * (607/754) / 0.4841$$

$$= (0.3608 * 0.8050) / 0.4841$$

$$= 0.6000$$

$P(\text{Tingkat resiko} = T_x | \text{PM} = \text{TIDAK})$

$P(\text{tingkat resiko rendah} = \text{TIDAK})$

$$= (338/607) * (607/754) / 0.5159$$

$$= (0.6392 * 0.8050) / 0.5159$$

$$= 0.9974$$

Tabel 11: Probabilitas *attribute* tingkat resiko

Resiko	Preventive Maintenance	
	YA	TIDAK
Rendah	0,6000	0,9974
Sedang	0,3178	0,0026
Tinggi	0,0822	0,0000



8. Evaluation

Untuk melakukan pengujian terhadap metode naïve bayes dengan cara mengkalikan semua hasil variable “YA” & ”TIDAK”, untuk menentukan hasil prediksi bandingkan  $P(Preventive\ maintenance = YA)$  dan  $P(Preventive\ maintenance = TIDAK)$ .

$$\begin{aligned}
 &P(Preventive\ maintenance = YA) \\
 &= P(Machine\ Number = YA) * P(Problem = YA) \\
 & * P(Shift = YA) * P(Durasi\ (min) = YA) \\
 & * P(Breakdown\ Type = YA) \\
 & * P(Tingkat\ Resiko = YA) * P(YA) \\
 &P(Preventive\ maintenance = TIDAK) \\
 &= P(Machine\ Number = TIDAK) \\
 & * P(Problem = TIDAK) \\
 & * P(Shift = TIDAK) * P(Durasi\ (min) = TIDAK) \\
 & * P(Breakdown\ Type = TIDAK) \\
 & * P(Tingkat\ Resiko = TIDAK) * P(TIDAK)
 \end{aligned}$$

Berikut hasil pengujian dengan menggunakan data testing sebanyak 189 data.

Tabel 12: Hasil pengujian manual

Machine Number	Problem	Shift	Durasi	Breakdown Type	Tingkat Resiko	Preventive Maintenance	Class Prediction	Preventive Maintenance YA	TIDAK
ME-2217-12-012	Oilshak oring	2	4	lain-lain	Rendah	TIDAK	TIDAK	0,00000000	0,00000000
ME-2217-12-012	Lifter oring	2	115	lain-lain	Sedang	YA	YA	0,00000016	0,00000000
ME-2217-12-012	Sensor gripper	1	26	Electric	Rendah	TIDAK	TIDAK	0,00000000	0,00001000
ME-2217-12-012	Home jesso	1	13	lain-lain	Rendah	TIDAK	TIDAK	0,00000000	0,00000020
ME-2217-12-012	Sensor tras	2	18	Electric	Rendah	TIDAK	TIDAK	0,00000000	0,00001048
ME-2217-12-012	Kabel sensor	2	15	Electric	Rendah	YA	TIDAK	0,00000035	0,00001040
ME-2217-12-012	Sensor oring	1	8	Mechanic	Rendah	TIDAK	TIDAK	0,00000012	0,00000147
ME-2217-12-012	Sensor oring	1	26	Mechanic	Rendah	TIDAK	TIDAK	0,00000019	0,00000321
ME-2217-12-012	Sensor oring	1	6	lain-lain	Rendah	TIDAK	TIDAK	0,00000018	0,00000020
ME-2217-12-012	Coflak parameter	1	20	Pneumatic	Rendah	YA	YA	0,00000495	0,00000154
ME-2217-12-012	Kabel sensor	2	30	Electric	Rendah	YA	TIDAK	0,00000035	0,00001040
ME-2217-12-012	Lifter oring	2	13	lain-lain	Rendah	TIDAK	TIDAK	0,00000053	0,00000007
ME-2217-15-001	Coversep chain	2	30	Mechanic	Rendah	YA	YA	0,00000059	0,00000054
ME-2217-15-001	Roller	1	23	Mechanic	Rendah	YA	YA	0,00000018	0,00000000
ME-2217-15-001	Roller	1	28	Mechanic	Rendah	YA	YA	0,00000249	0,00000178
ME-2217-15-001	Relay	1	14	Electric	Rendah	TIDAK	TIDAK	0,00000000	0,00000007
ME-2217-15-001	Coversep bath	2	110	Mechanic	Sedang	YA	YA	0,00000018	0,00000000
ME-2217-15-001	Sensor radio loading	1	20	Electric	Rendah	TIDAK	TIDAK	0,00000000	0,00000134
ME-2217-15-001	Sensor filter	1	12	Electric	Rendah	YA	TIDAK	0,00000060	0,00000404
ME-2217-15-001	Sensor filter	2	42	Electric	Rendah	YA	TIDAK	0,00000073	0,00000008
ME-2217-15-001	Sensor filter	2	16	lain-lain	Rendah	YA	TIDAK	0,00000007	0,00000000
ME-2217-15-001	Sensor filter optic	1	30	Electric	Rendah	TIDAK	TIDAK	0,00000000	0,00000568

Berdasarkan hasil pengujian diatas dapat dilihat persentase untuk tingkat akurasi prediksi yang tepat sebesar 93,65% dan sedangkan prediksi yang tidak tepat sebesar 6,35 %. Dimana dari 189 data kerusakan mesin, ada sebanyak 177 data mesin yang harus dilakukan preventive maintenance berhasil diprediksi dengan tepat dan sebanyak 12 data tidak dapat diprediksi dengan tepat.

Hasil dan Pembahasan

Dalam penelitian ini dilakukan pemodelan Naïve Bayes menggunakan aplikasi Rapid Miner 9.6 dengan menggunakan data training sebanyak 754 record dan data testing sebanyak 189 record. Berikut adalah spesifikasi perangkat keras dan perangkat lunak yang akan digunakan:

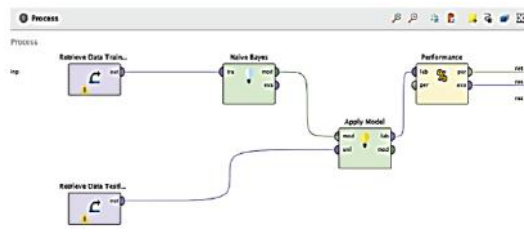
- Spesifikasi Perangkat Keras
  1. Processor intel core i5 2,3 Ghz Memory 8 GB

2. Hard disk 256 GB
3. Graphics intel iris 640 1,5 GB
4. Display mode (2560 x 1600)

• Spesifikasi Perangkat Lunak

1. Operating System MacOS Catalina version 10.15.5
2. RapidMiner 9.6

Berikut merupakan untuk pemodelan naïve bayes pada aplikasi rapidminer 9.6.



Gambar 4: Proses pemodelan data naïve bayes

Berikut merupakan hasil akurasi prediksi yang tepat sebesar 93,65% yang dapat dilihat pada Gambar 5, sedangkan prediksi yang tidak tepat sebesar 6,35 %. Dimana dari 189 data testing terdapat sebanyak 177 data mesin yang harus dilakukan preventive maintenance berhasil diprediksi dengan tepat dan sebanyak 12 data tidak dapat diprediksi dengan tepat.

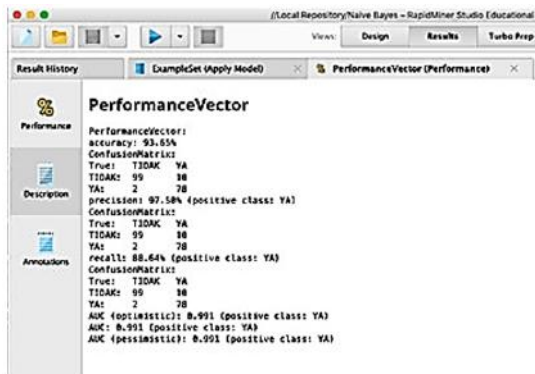
	true TIDAK	true YA	class prediction
pred. TIDAK	99	10	90,83%
pred. YA	2	78	97,50%
class recall	98,02%	88,64%	

Gambar 5: Akurasi Naïve bayes

Precision adalah jumlah data yang true positive (jumlah data positif yang dikenali secara benar sebagai positif) dibagi dengan jumlah data yang dikenali sebagai positif). Dari hasil pengujian nilai precision yaitu 97,50%.

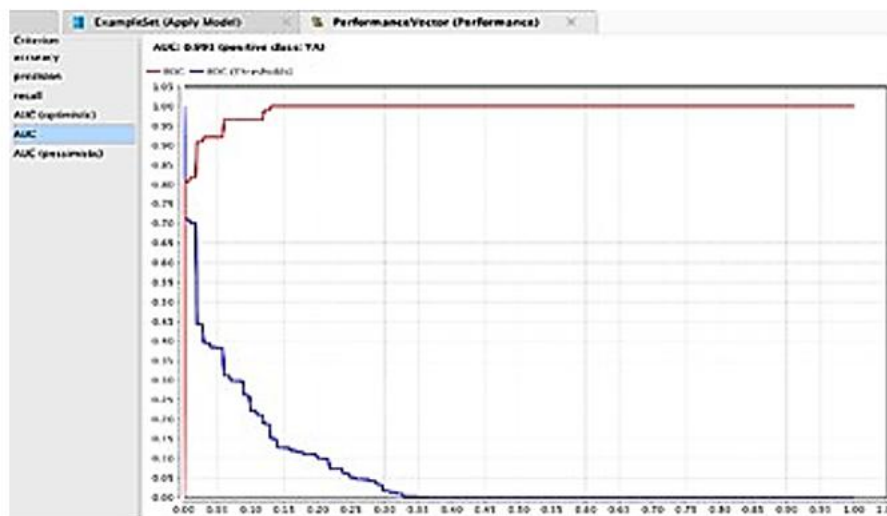
Recall adalah jumlah data yang true positive dibagi dengan jumlah data yang sebenarnya positif (true positive + true negative). Dari hasil pengujian nilai recall yaitu 88,64%.



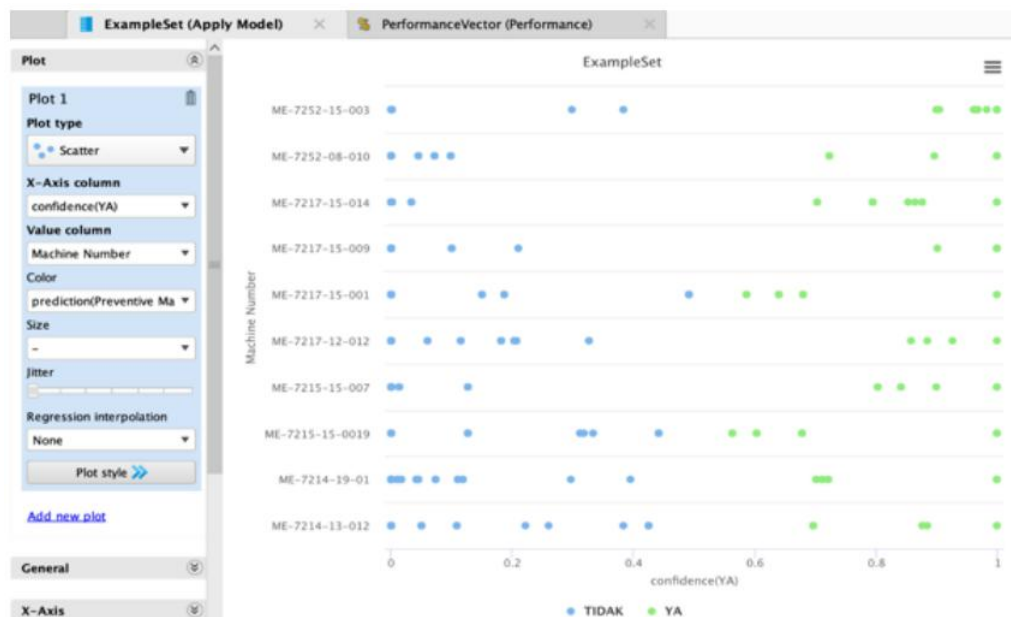


Gambar 6: Performance Vector Naïve bayes

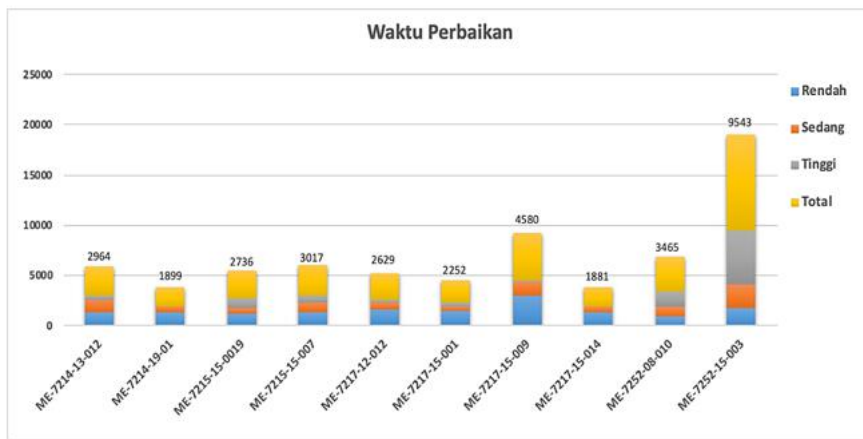
Selain *confusion matrix*, kurva *Receiver Operating Characteristic* (ROC) dihasilkan oleh rapid miner. Kurva *Receiver Operating Characteristic* (ROC) digunakan untuk mengekspresikan data *confusion matrix*. Garis horizontal mewakili nilai *false positives* (FP) dan garis vertikal mewakili nilai *true positives* (TP). Pada gambar dibawah ini dapat diketahui bahwa nilai *Area Under Curve* (AUC) model algoritma naïve bayes adalah 0,991, lihat Gambar 7.



Gambar 7: Kurva AUC (Area Under Curve)



Gambar 8: Chart prediction Rapid Miner



Gambar 9: Tingkat resiko kerusakan mesin

Pada Rapid Miner terdapat tampilan grafik sebaran data hasil prediksi perhitungan. Untuk titik warna biru merupakan hasil dari nilai confidence (TIDAK), sedangkan titik berwarna hijau merupakan nilai confidence (YA), lihat Gambar 8..

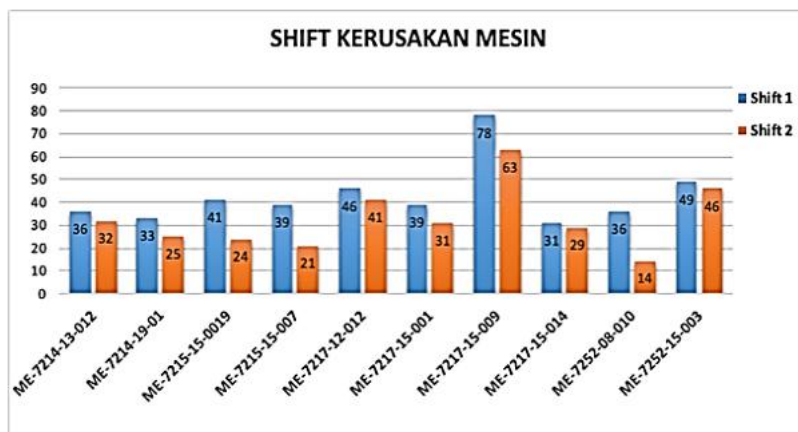
Dari penelitian ini data kerusakan mesin produksi yang terjadi pada *maintenance department*, pada Gambar 9 dapat disimpulkan bahwa tingkat kerusakan dengan resiko rendah mendominasi kerusakan mesin produksi dengan persentase 80,50%. Dimana mesin dengan nomor ME-7217-15-009 paling sering mengalami kerusakan dengan persentase 16,18%, sedangkan mesin yang paling sedikit mengalami kerusakan dengan nomor ME-7252-08-010 dengan presentase 6,63%. Berdasarkan dari hasil analisa maka dapat ditentukan bahwa mesin ME-7217-15-009 perlu dilakukan *preventive maintenance* terlebih dahulu.

Perusahaan ini menetapkan dua jenis pembagian waktu kerja yaitu *shift 1* (siang) dan *shift 2* (malam). Pada Gambar 10 kerusakan mesin banyak terjadi pada *shift 1* dengan persentase 56,76%. Dikarenakan tingkat resiko kerusakan mesin sering terjadi pada *shift 1*, sebaiknya tindakan *preventive*

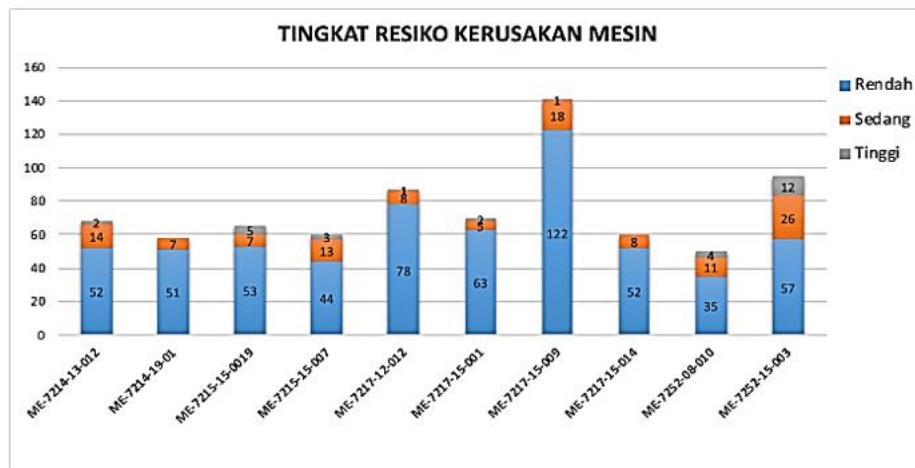
*maintenance* dapat dilakukan pada *shift 2* agar tidak mengganggu proses produksi pada *shift 1*.

Pada Gambar 11 terlihat bahwa tipe kerusakan mesin banyak terjadi pada bagian *mechanic* dengan persentase 46,02%. Dimana mesin dengan nomor ME-7252-15-003 paling sering mengalami kerusakan dengan persentase 17,29 %, sedangkan mesin yang paling sedikit mengalami kerusakan dengan nomor ME-7214-13-012 dengan presentase 3,17%. Berdasarkan hasil analisa maka dapat ditentukan bahwa mesin ME-7252-15-003 perlu dilakukan *preventive maintenance* terlebih dahulu, untuk mengurangi resiko kerusakan mesin.

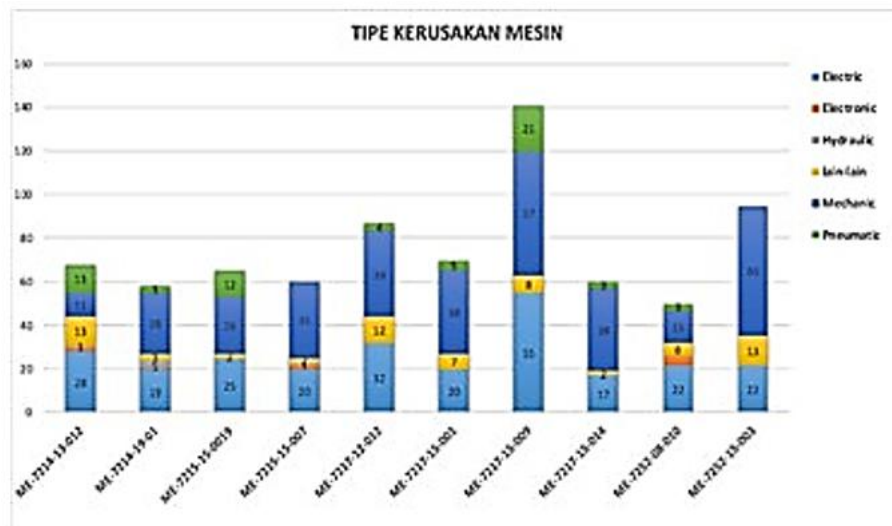
Pada gambar 12 terlihat bahwa total waktu perbaikan kerusakan mesin yang membutuhkan waktu perbaikan yang paling lama adalah mesin dengan nomor ME-7252-15-003 dengan presentase 27,29%, sedangkan mesin yang paling sedikit membutuhkan waktu perbaikan dengan nomor mesin ME-7217-15-014 dengan presentase 5,38%. Berdasarkan hasil analisa maka untuk mesin ME-7252-15-003 harus dilakukan *improvement process* untuk perbaikan kerusakan mesin. Agar proses perbaikan dapat dilakukan lebih cepat.



Gambar 10: Shift terjadi kerusakan mesin



Gambar 11: Tipe kerusakan mesins



Gambar 12: Waktu Perbaikan

## Penutup

Berdasarkan penelitian yang telah dilakukan mulai dari tahap awal hingga proses pengujian algoritma dapat disimpulkan bahwa dengan metode algoritma naïve bayes berdasarkan data dari *history* kerusakan mesin dengan jumlah data training 754 *record* dan data testing 189 *record*, metode ini dapat memprediksi mesin mana saja yang perlu dilakukan *preventive maintenance* dengan tingkat akurasi 93,65%. Sehingga dari hasil penelitian ini dapat membantu pihak *maintenance department* sebagai acuan untuk menyusun jadwal untuk melakukan kegiatan *preventive maintenance* terhadap mesin-mesin tersebut untuk mengurangi *breakdown time* yang dapat mengganggu proses produksi, serta pihak *maintenance department* dapat mempersiapkan dan memperkirakan *spare part* apa saja yang digunakan untuk proses tersebut. Sehingga tidak terjadi penumpukan *spare part* di gudang *maintenance department* dan menjaga *stock level spare part*

agar tidak *over stock* sehingga dapat mengakibatkan *dead stock*. Serta dapat meningkatkan produktivitas kerja divisi *maintenance* karena dapat memprediksi mesin apa saja yang akan mengalami kerusakan.

Untuk penelitian selanjutnya diharapkan hasil dari proses *data mining* tersebut dapat ditampilkan detail *schedule preventive maintenance* pada aplikasi MBA (*Machine Breakdown Action*) untuk memudahkan proses monitoring mesin mana saja yang sedang dalam proses *preventive maintenance*. Serta aplikasi MBA dapat terintegrasi dengan *spare part* sistem agar dapat diketahui ketersediaan *spare part* dan total biaya perbaikan mesin yang mengalami kerusakan, untuk memudahkan *management* dalam mengambil suatu keputusan.

## Daftar Pustaka

- [1] Carlo Vercellis, "Data Warehousing", in Business Intelligence, Onlinelibrary.wiley.com, <https://doi.org/10.1002/9780470753866.ch3>, 2009.
- [2] Wawan dan Munir, "Pengantar Teknologi Informasi, Sistem Informasi", Bandung, Universitas Pendidikan Indonesia, 2006.
- [3] Jhon J. Longkutoy, "Pengenalan Komputer", Mutiara Sumber Widya, Jakarta, 1996.
- [4] F. Sari dan D. Saro, "Implementasi Algoritma C4.5 Dalam Menentukan Lokasi Prioritas Penyuluhan Program Keluarga Berencana di Kecamatan Dumai Timur", Jurnal Penelitian Pos dan Informatika, Vol 8, No. 1, 63 - 76, 2018.
- [5] E. Turban, "Decision Support Systems and Intelligent Systems", Edisi Bahasa Indonesia Jilid 1, Andi, Yogyakarta, 2005.
- [6] U. Fayyad, G. Piatetsky-Shapiro & P. Smyth, "From Data Mining to Knowledge Discovery in Databases", AI Magazine, 17(3), 37. <https://doi.org/10.1609/aimag.v17i3.1230>, 1996.
- [7] Y. Phaharsi, I. K. Sriwana dan D.M. Sari, "Perancangan Penjadwalan Preventive Maintenance Pada PT Artha Prima Sukses Makmur", Jurnal Ilmiah Teknik Industri, 14(1), 59–65. <http://journals.ums.ac.id/index.php/jiti/article>," J. Ilm. Tek. Ind., vol. 14, no. 1, pp. 59–65, 2015.
- [8] Y. N. Rohmat, Rachmatullah, D. Canra, and Suliono, "Analisis Perancangan Jadwal Preventive Dan Predictive Maintenance Pada Mesin Kapal Di Daerah", Pros. Semin. Nas. Teknol. dan Rekayasa, pp. 1–7, 2017.
- [9] I. A. Marie, L. Hakim, D. Sugiarto, and W. Septiani, "Perbandingan Performansi Teknik Klasifikasi Breakdown Mesin pada Proses Produksi Pembuatan Battery Mobil", J. Ilm. Tek. Ind., vol. 18, no. 1, pp. 33–41, 2019, doi: 10.23917/jiti.v18i1.7232.
- [10] I. A. Marie, L. Hakim, D. Sugiarto, and W. Septiani, "Analisis Data Gangguan Kerusakan Mesin Produksi Menggunakan Teknik Association Rules", J. Ilm. Tek. Ind., vol. 7, no. 1, pp. 43–52, 2019, doi: 10.24912/jitiuntar.v7i1.5035.
- [11] I. Russari, "Sistem Pakar Diagnosa Penyakit Batu Ginjal Menggunakan Teorema Bayes", J. Ris. Komput., vol. 3, pp. 18–22, 2016.