

PENERAPAN ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI LAMA Pengerjaan Servis Berkala KENDARAAN PELANGGAN PT. NASMOCO MAJAPAHIT

Restu Andriar Okthavian¹, Heru Lestiawan . M.Kom.²

¹*Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang
email : 1111201106084@mhs.dinus.ac.id*

²*Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang
email : herul.at.dosen.dinus.ac.id*

Abstract PT. Nasmoco Majapahit is a company engaged in the automotive field. Workshop in which there are parts that serve the work of repairing the vehicle, be it a general periodic servicing and repair. At periodic servicing, Service Advisor provides time estimates hour delivery vehicle to the customer. Due to the ever-increasing customer expectations in PT. Nasmoco Majapahit feared disappointed if the estimated processing time is not appropriate periodic servicing. To overcome this would be to use data mining techniques to process data at regular servicing of vehicles. Data mining is the mining or the discovery of new information by looking for certain patterns or rules of a number of large amounts of data are expected to treat the condition. By utilizing the data service customers` vehicles, is expected to generate information about the speed of time required to perform periodic servicing of vehicles. The algorithm used is the naïve Bayes algorithm, so that information is displayed in the form of speed category. Category velocity obtained is influenced by mechanical, vehicles, miles periodic servicing and additional jobs.

Keys : data mining, naïve bayes algorithm, periodic servicing.

I. PENDAHULUAN

1.1 Latar Belakang

P.T. Nasmoco Majapahit Semarang merupakan salah satu cabang dari PT. New Ratna Motor yang bergerak dalam bidang otomotif. Terletak di Jl. Brigjend Sudiarto No. 426 Semarang. Perusahaan otomotif dibawah naungan Toyota Astra Motor ini dalam kegiatan organisasi memiliki salah satu prinsip yaitu mengutamakan kepuasan pelanggan. Kepuasan pelanggan diketahui setelah dilakukan follow up oleh petugas, baik pada proses penjualan dan pemeliharaan kendaraan. Untuk meningkatkan kepuasan pelanggan dan mempermudah operasional kerja, perusahaan telah menggunakan teknologi modern seperti website, media sosial, Blackberry Messenger, sistem antrian dan sistem

pembagi pekerjaan / NIS (Nasmoco Intregated System). Dengan teknologi tersebut kinerja perusahaan teroganisir dengan lancar. Mulai dari aktifitas pengkreditan kendaraan, customer satisfaction pada bagian penjualan serta alur pekerjaan pada bagian bengkel tercatat secara runtun oleh sistem yang ada.

Tetapi pada bagian after sales / bengkel masih terdapat sedikit kekurangan di ruang counter service. Di ruang counter service terdapat dua petugas yaitu service advisor dan control room. Sevice advisor mempunyai tugas menerima pelanggan sesuai nomor antrian kemudian membuat work order kendaraan, sedangkan tugas utama control room adalah membagikan work order yang telah dibuat service advisor kepada teknisi yang tersedia agar tercipta heijunka. Heijunka adalah kondisi dimana tercapainya kesamaan rata-rata pengerjaan jumlah kendaraan yang melakukan servis oleh teknisi. Secara keseluruhan proses penerimaan, pemeliharaan, dan pembayaran kendaraan yang melakukan servis telah tercatat

oleh sistem. Akan tetapi pengerjaan kendaraan yang melakukan servis berkala / servis ringan dan general repair / servis berat sering diluar estimasi jam penyerahan yang telah dijanjikan servis advisor kepada

customer. Estimasi jam penyerahan adalah lama waktu yang dibutuhkan dalam pengerjaan mobil / kendaraan untuk melakukan servis berkala.

Maka agar servis advisor dapat memberikan estimasi jam penyerahan tepat waktu kepada pelanggan akan dimanfaatkan dataset yang sudah tersimpan kemudian dianalisa dan diringkas kedalam informasi yang berguna, dimana informasi tersebut bisa digunakan untuk meningkatkan pendapatan dari efisiensi alur kerja pada counter service. Proses penggalian informasi baru dari data set yang ada dinamakan dengan data mining. Data Mining adalah ekstraksi pola yang menarik dari data dalam jumlah besar [1]. Salah satu teknik data mining adalah klasifikasi. Teknik klasifikasi yaitu membangun suatu model yang bisa mengklasifikasikan suatu objek berdasar atribut-atributnya. Kelas target sudah tersedia dalam data sebelumnya, sehingga fokusnya adalah bagaimana mempelajari data yang ada agar klasifikator bisa mengklasifikasikan sendiri. Contoh-contoh metode yang diberikan meliputi decision tree classifier, rule-based classifier, neural network, support vector machines, dan naive Bayes classifier [1]. Metode yang digunakan untuk penelitian ini adalah metode naive bayes. Klasifikasi dengan metode Naive Bayes memperlihatkan keakuratan yang tinggi dan kecepatan yang baik ketika di jalankan pada database yang besar [2]. Berdasarkan permasalahan yang terjadi, dengan memanfaatkan teknik data mining metode naive bayes, penulis ingin mengolah set data servis kendaraan PT. Nasmoco Majapahit Semarang dengan tujuan memperoleh informasi baru mengenai waktu yang dibutuhkan untuk melakukan servis berkala kendaraan.

II. METODE YANG DIUSULKAN

Beberapa kasus berikut merupakan acuan literatur dalam penelitian ini :

- a. *Decision Tree and Naive Bayes Algorithm for Classification and Generation of Actionable Knowledge for Direct Marketing*
- b. *Graduation Prediction of Gunadarma University Students Using Naive Bayes Algorithm and C4.5 Algorithm*

Dari beberapa jurnal yang telah dibaca akan digunakan sebagai acuan literature yang memiliki kesamaan konsep data mining untuk memprediksi suatu permasalahan. Maka penelitian ini akan menggunakan teknik klasifikasi metode naive bayes untuk mengolah *data set* servis berkala PT. Nasmoco Majapahit Semarang. Tujuan dari pengolahan data adalah memperoleh informasi baru untuk memprediksi waktu yang dibutuhkan dalam pengerjaan servis berkala

kendaraan di bawah km 80.000.

A. Metode Pengumpulan Data

1. Jenis Data

Jenis data yang akan digunakan untuk penelitian ini adalah data kuantitatif. Data kuantitatif merupakan data yang dapat dideskripsikan secara numerik dalam hal objek, variabel, dan nilai [17]. Data yang digunakan dalam penelitian ini berupa data servis berkala PT Majapahit Semarang bulan Mei sampai dengan September tahun 2014. Dari data tersebut hanya diambil data yang berkaitan dengan lama servis berkala berdasarkan tipe kendaraan dibawah 80.000 km.

2. Sumber Data

Sumber data yang digunakan dalam penelitian ini ada dua, yaitu data primer dan data sekunder.

a) Data Primer

Data Primer adalah data yang berasal dari sumber asli atau pertama. Pada penelitian ini, data primer adalah data pelanggan servis kendaraan dibawah 80.000 km PT Majapahit Semarang pada bulan Mei sampai dengan September tahun 2014.

b) Data Sekunder

Data sekunder merupakan hasil pengolahan lanjut dari data primer yang ditampilkan dalam bentuk lain dan digunakan untuk mendukung data primer. Dalam penelitian ini data sekunder diperoleh dari Small Group Activity (SGA), laporan kepuasan pelanggan, buku, dan literature. Peneliti menggunakan data sekunder ini untuk memperkuat data primer dan melengkapi informasi yang telah dikumpulkan melalui observasi langsung.

3. Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini merupakan teknik pengumpulan observasi dan dokumentasi. Data pelanggan PT. Nasmoco Majapahit Semarang pada bulan Mei sampai dengan September 2014 akan dianalisa

untuk memperoleh sebuah informasi baru yang akan digunakan untuk memprediksi waktu yang dibutuhkan untuk melakukan servis berkala dibawah 80.000 km dengan teknik klasifikasi menggunakan algoritma naive bayes.

B. Analisa Data Mining

Dalam penelitian ini akan dicari atribut yang memiliki hubungan antara waktu yang dibutuhkan untuk melakukan servis dengan data pelanggan PT. Nasmoco Majapahit Semarang. Tidak semua data pelanggan dicari hubungannya dengan waktu yang dibutuhkan untuk melakukan servis berkala, hanya beberapa atribut yang berguna dan sebarannya tidak terlalu acak. Karena data yang terlalu acak akan membuat proses mining memakan waktu lama dan tingkat hubungannya pun rendah. Data customer yang akan

dicari hubungannya meliputi :

a). Hubungan atribut tipe kendaraan dengan jumlah waktu yang dibutuhkan.

Atribut tipe kendaraan akan dicari hubungannya dengan waktu yang dibutuhkan untuk melakukan servis berkala dibawah km 80.000. Atribut tipe kendaraan merupakan salah satu atribut yang mempengaruhi total waktu dikarenakan perbedaan tingkat kesulitan pada saat pengerjaan job servis.

b). Hubungan atribut pekerjaan dengan jumlah waktu yang dibutuhkan.

Hubungan pekerjaan dengan jumlah waktu yang dibutuhkan bermanfaat untuk mengetahui kilometer servis berkala yang membutuhkan waktu lebih lama atau lebih cepat.

c). Hubungan atribut additional job dan mekanik dengan jumlah waktu yang dibutuhkan.

Hubungan additional job dan mekanik berpengaruh pada jumlah waktu yang dibutuhkan untuk melakukan servis berkala. Additional job berpengaruh karena pengerjaan harus sesuai dengan antrian, sedangkan untuk mekanik berpengaruh karena setiap mekanik memiliki alur pengerjaan yang berbeda.

Proses lengkap analisa data mining pada penelitian ini menggunakan tahapan CRISP-DM (Cross-Industry Standard Process for Data Mining) dengan enam fase berikut ini :

1) Fase Pemahaman Bisnis

Tujuan dilakukannya penelitian ini adalah untuk mengetahui waktu yang dibutuhkan untuk melakukan servis berkala dibawah km 80.000 pada PT Nasmoco Majapahit Semarang. Data set pelanggan pada bulan Mei sampai dengan September tahun 2014 merupakan data yang akan dianalisa untuk memperoleh informasi yang dibutuhkan. Data set akan diolah dengan cara data mining menggunakan teknik klasifikasi metode naïve bayes. Dari algoritma terpilih akan diimplementasikan ke GUI (Graphical User Interface) menggunakan engine Java untuk memudahkan dalam melihat prediksi jumlah waktu yang dibutuhkan dalam melakukan servis berkala.

2) Fase Pemahaman Data (Data Understanding Phase)

Fase ini merupakan pengumpulan data set yang diperoleh dari PT. Nasmoco Majapahit Semarang berupa record berjumlah 3688 dari bulan Mei sampai dengan bulan September. Selanjutnya data yang didapatkan akan dianalisa atributnya untuk mendapatkan pengetahuan awal. Atribut dari data set pelanggan yang melakukan servis berkala dapat dilihat pada tabel 2.2 :

Tabel 3.2.2 : Pemahaman Data

Atribut	Keterangan
Nomor Polisi	Merupakan nomor seri dari kepolisisan untuk

	kendaraan.
Nomor <i>Work Order</i>	Nomor WO merupakan nomor yang tertera pada kertas job pekerjaan.
Nomor <i>Invoice</i>	Nomor <i>invoice</i> adalah nomor yang muncul setelah <i>service advisor</i> melakukan semua pekerjaan dan menutup <i>work order</i> di dalam program NIS.
<i>Customer</i>	Atribut <i>customer</i> merupakan nama pelanggan yang melakukana servis kendaraan.
Jam	Jam merupakan waktu dimana customer mengambil nomor antrian.
Serah	Atribut serah menunjukkan waktu dimana SA menutup semua proses didalam program NIS dikarenakan kendaraan akan diserahkan kepada <i>customer</i> .
Nomor Rangka	Nomor rangka kendaraan merupakan nomor seri yang tertera pada bodi kendaraan

	(<i>chasis</i>).
Model Kendaraan	Model kendaraan merupakan nama kendaraan. Contoh : Avanza, Innova, Yaris.
Tunggu Servis	Merupakan atribut yang menunjukkan apakah kendaraan ditunggu atau ditinggal dalam proses pengerjaan.
<i>Service Advisor</i>	Merupakan seseorang yang berinteraksi langsung dengan pelanggan. Salah satu tugasnya adalah membuat <i>work order</i> .
<i>Mechanic</i>	Seseorang yang melakukan / mengerjakan servis kendaraan.
Kelompok	Atribut kelompok menunjukkan <i>group</i> yang mengerjakan <i>work order</i> .
Pekerjaan	Atribut ini menunjukkan tipe pekerjaan. Tipe pekerjaan ada 3 : Servis Berkala Internal, Servis berkala Eksternal, dan <i>General Repair</i> .
<i>Rate</i>	Rate adalah waktu yang digunakan

	untuk pengerjaan servis tidak meliputi <i>washing</i> .
<i>Actual Rate</i>	Atribut ini menunjukkan waktu yang sebenarnya dalam pengerjaan servis tidak meliputi <i>washing</i> .
Status	Merupakan atribut yang menunjukkan kendaraan ditunggu pelanggan atau ditinggal dalam proses pengerjaan servis.

Dari atribut-atribut tersebut hanya 4 atribut yang akan digunakan untuk diolah di dalam fase selanjutnya. Pemilihan atribut didasarkan keterkaitan nilai didalamnya. Atribut yang digunakan adalah Model Kendaraan, Pekerjaan, Mekanik dan *Additional Job*.

3) Fase Pengolahan Data (Data Preparation Phase)

Dari atribut yang digunakan akan diambil record yang sesuai batasan. Data yang berjumlah 3688 diambil pada bulan Juli sampai dengan bulan September setelah diseleksi sesuai kebutuhan hanya terdapat 2059 record. Kriteria data yang dibutuhkan mencakup :

- Tipe kendaraan merupakan tipe avanza, inova atau sedan.
- Kilometer kendaraan ≤ 80.000 .

Maka dari kriteria data yang ditentukan contoh data sampel dapat dilihat pada tabel 2.3 :

Tabel 3.3.3 : Fase Pengolahan Data.

No Polisi	Pekerjaan	Model Kendaraan	Jam	Serah
H-9127-PQ	SBE 40.000	Avanza	09.38	10.55
H-8861-XX	SBE 30.000	Innova	14.05	16.43
H-9822-XX	SBE 50.000	Agya	13.03	15.49
H-8686-QM	SBE 10.000	Yaris	07.35	10.21
H-8631-XX	SBE 60.000	Vios	11.36	14.65

4) Fase Permodelan (Modelling Phase)

Fase ini akan memilih dan mengaplikasikan teknik permodelan yang sesuai untuk mengoptimalkan hasil. Pada penelitian ini menggunakan teknik klasifikasi data mining algoritma naive bayes.

5) Fase Evaluasi (Evaluation Phase)

Pada fase ini akan mengevaluasi model yang digunakan dalam fase permodelan serta menetapkan apakah model tadi sudah sesuai dengan tujuan pada fase awal. Pada penelitian ini akan dilakukan 2 proses evaluasi, yakni pengujian dan perbandingan. Pengujian akan dilakukan dengan unit kendaraan secara langsung dan untuk perbandingan akan membandingkan hasil setelah pengimplementasian dengan data awal apakah sesuai dengan tujuan yang diinginkan.

6) Fase Penyebaran (Deployment Phase)

Setelah tahap evaluasi dimana menilai secara detail hasil dari permodelan dilakukan pengimplementasian dari keseluruhan model yang telah dirancang. Selain itu juga dilakukan penyesuaian dari model dengan sistem yang akan dibangun sehingga dapat menghasilkan suatu hasil yang sesuai dengan target diawal tahap CRISP-DM.

C. Metode Pengembangan Sistem Waterfall

Pengembangan sistem yang akan digunakan pada penelitian ini adalah metode *System Development Life Cycle (SDLC)* model *waterfall*. Adapun tahapannya adalah sebagai berikut :

a. Requirement Analysis

Requirement Analysis atau analisa kebutuhan sistem merupakan langkah awal pada metode pengembangan sistem waterfall. Pada tahap ini dianalisa mengenai kebutuhan sistem yang akan dikembangkan. Sistem data mining ini dirancang untuk menampilkan daftar history pelanggan yang melakukan servis berkala, menampilkan daftar SA, dan memberikan informasi tentang prediksi kategori waktu, dan melakukan input data pelanggan.

b. System Design

System Design bertujuan menentukan spesifikasi detail dari komponen-komponen sistem informasi dan produk-produk informasi yang sesuai dengan hasil tahap analisis. Desain sistem data mining ini akan digambarkan menggunakan UML, yaitu menggunakan use case diagram dan class diagram. Pada use case digambarkan interaksi antara actor dan sistem. Sedangkan untuk class diagram adalah untuk menggambarkan alur desain program yang akan dibuat. Setelah use case dan class diagram dibuat tahap selanjutnya adalah desain interface sesuai dengan kebutuhan sistem.

c. Implementastion

Pada langkah ini adalah proses dilakukannya pemrograman sesuai dengan desain yang telah dibuat. Proses pengimplementasian harus sesuai dengan use case

dan class diagram yang dirancang agar hasil dari sistem data mining ini memenuhi kebutuhan. Proses penulisan program data mining ini akan menggunakan Bahasa pemrograman java dan menggunakan tools Java NetBeans.

d. Integration & Testing

Setelah proses penulisan program selesai maka tahap selanjutnya adalah pengujian sistem data mining. Pengujian disini bertujuan untuk mengetahui apakah sistem yang dibuat sudah memenuhi kebutuhan atau masih perlu perbaikan agar memndapatkan program yang sempurna untuk dijalankan. Program data mining yang telah dibuat akan diuji menggunakan metode Black Box Testing.

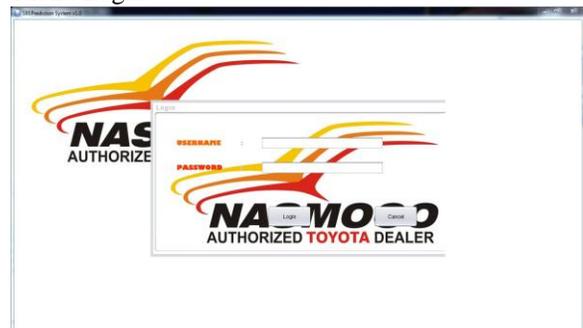
e. Operation & Maintenance

Tahap perawatan terhadap sistem data mining akan dilakukan secara rutin untuk mengetahui apakah sistem perlu dikembangkan agar dapat memenuhi kebutuhan yang semakin bertambah.

III. IMPLEMENTASI

Implementasi aplikasi data mining ini terdiri dari beberapa interface yaitu Menu Login, Menu Utama, Menu Service Advisor, Menu History Service, Menu About, dan Menu Input Data.

1. Menu Login



Gambar 3.1 : User Interface Login

Pada halaman ini user / admin memasukkan akun yang telah terdaftar agar bisa masuk ke halaman utama.

2. Menu Utama



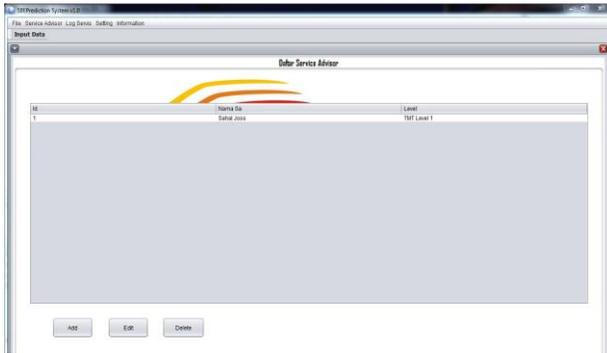
Gambar 3.2 : User Interface Menu Utama.

Halaman utama merupakan interface yang menampilkan menu fungsional berupa File, Service Advisor, Log service, Setting, Information dan toolbar

Input Data.

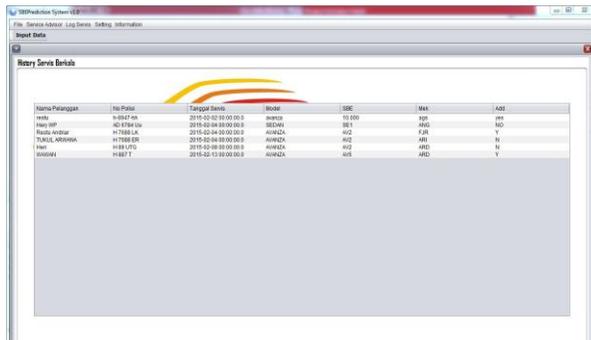
3. Menu Service Advisor

Daftar SA merupakan tampilan dari item menu Expand pada menu Service Advisor. User dapat melihat list SA, memanipulasi data pada daftar berupa tambah SA, Edit SA, dan Delete SA dengan cara memilih button yang disediakan.



Gambar 3.3 : User Interface Daftar SA.

4. Menu History



Gambar 3.4 : User Interface History Service.

History Service adalah *interface* dari item menu Unit Masuk pada menu Log Service. Pada menu ini user dapat melihat daftar pelanggan yang telah melakukan servis berkala dan dapat melakukan pencarian data pelanggan.

5. Menu About



Gambar 3.5 : User Interface About

Pada halaman about dijelaskan mengenai versi dan cv dari developer perangkat lunak data mining.



6. Menu Input Data

Gambar 3.6 : User Interface Input Data

Pada halaman Input Data user dapat memasukkan data pelanggan yang melakukan servis berkala dan mendapatkan hasil prediksi kategori speed dari data yang di input.

IV. HASIL & PEMBAHASAN

A. Hasil Penelitian

Hasil penelitian adalah terciptanya sistem untuk memprediksi waktu yang dibutuhkan dalam proses servis berkala. Sistem terdiri dari beberapa *interface* yaitu : *Login* , *Main Menu*, *Input Data*, *Service Advisor*, *History Service*, *User Setting* dan *About*. Setiap antar muka mempunyai fungsi yang berbeda. Pada implementasi program dan *interface* pengguna dapat memperoleh informasi berupa daftar *Service Advisor*, daftar pelanggan yang pernah melakukan service berkala, daftar SA dan mendapatkan kategori waktu yang dibutuhkan untuk melakukan servis berkala menggunakan algoritma naïve bayes.

B. Perhitungan Algoritma Naïve Bayes

Setelah proses implementasi perhitungan naïve bayes pada sistem adalah sebagai berikut :

1. Diketahui pada data training yang berjumlah 2.059, jumlah masing-masing class adalah :

- Kelas "FAST" berjumlah
- Kelas "MEDIUM" berjumlah
- Kelas "LOW" berjumlah

Maka proses penghitungan dengan algoritma naïve bayes adalah menghitung $P(C_i)$ untuk setiap atribut, dalam kasus dataset pada penelitian ini yaitu atribut tahun kelulusan yang terdiri dari 3 kelas yaitu kelas waktu servis cepat dinyatakan "FAST", kelas waktu servis sedang dinyatakan dengan "MEDIUM", dan kelas waktu servis rendah dinyatakan dengan "LOW". Kemudian hitung $P(X|C_i)$, $i=1,2,3$ untuk setiap kelas. Setelah itu bandingkan $P(X|C_i)$,

dicari nilai yang terbesar maka nilai yang keluar merupakan output. Prosesnya adalah sebagai berikut :

2. Menghitung jumlah kelas dari tahun lulus berdasarkan klasifikasi yang terbentuk (prior probability) :

a). C1 (Class Speed = "FAST") = jumlah "FAST" pada kolom Speed = 763. Maka jumlah "FAST" dibagi dengan total data Speed = $763 / 2059 = 0,370568$.

b). C2 (Class Speed = "MEDIUM") = jumlah "MEDIUM" pada kolom Speed = 893. Maka jumlah "MEDIUM" dibagi dengan total data Speed = $893 / 2059 = 0,433706$.

c). C3 (Class Speed = "LOW") = jumlah "LOW" pada kolom Speed = 403. Maka jumlah "LOW" dibagi dengan total data Speed = $403 / 2059 = 0,195726$.

3. Menghitung jumlah kasus yang sama pada setiap atribut dari kelas Speed (FAST / MEDIUM / LOW) berdasarkan data testing yang dipilih. Semisal inputan yang diberikan berupa Unit = "AVANZA", Sbe = "AV1", Mek = "RST" dan Adt = "Yes" maka penghitungan datanya adalah :

a). Menghitung jumlah kasus yang sama pada setiap atribut dari kelas Speed (Fast / Medium / Low) berdasarkan data testing.

$$P(\text{Unit} = \text{"AVANZA"} \mid \text{Class Speed} = \text{"FAST"}) = 461 / 763 = 0,6041$$

$$P(\text{Unit} = \text{"AVANZA"} \mid \text{Class Speed} = \text{"MEDIUM"}) = 445 / 893 = 0,4983$$

$$P(\text{Unit} = \text{"AVANZA"} \mid \text{Class Speed} = \text{"LOW"}) = 196 / 403 = 0,4863$$

$$P(\text{Sbe} = \text{"AV1"} \mid \text{Class Speed} = \text{"FAST"}) = 81 / 763 = 0,1061$$

$$P(\text{Sbe} = \text{"AV1"} \mid \text{Class Speed} = \text{"MEDIUM"}) = 90 / 893 = 0,1007$$

$$P(\text{Sbe} = \text{"AV1"} \mid \text{Class Speed} = \text{"LOW"}) = 40 / 403 = 0,0992$$

$$P(\text{Mek} = \text{"RST"} \mid \text{Class Speed} = \text{"FAST"}) = 39 / 763 = 0,0511$$

$$P(\text{Mek} = \text{"RST"} \mid \text{Class Speed} = \text{"MEDIUM"}) = 34 / 893 = 0,0380$$

$$P(\text{Mek} = \text{"RST"} \mid \text{Class Speed} = \text{"LOW"}) = 18 / 403 = 0,0446$$

$$P(\text{Adt} = \text{"YES"} \mid \text{Class Speed} = \text{"FAST"}) = 351 / 763 = 0,4600$$

$$P(\text{Adt} = \text{"YES"} \mid \text{Class Speed} = \text{"MEDIUM"}) = 412 / 893 = 0,4613$$

$$P(\text{Adt} = \text{"YES"} \mid \text{Class Speed} = \text{"LOW"}) = 155 / 403 = 0,3846$$

b). Setelah penghitungan kelas, hitung perkalian semua atribut sesuai dengan kelas :

$$\begin{aligned} \text{Kelas} &= \text{"FAST"} \\ &= P(\text{Unit}=\text{"FAST"}) \\ & \cdot P(\text{Sbe}=\text{"FAST"}) \cdot P(\text{Mek}=\text{"FAST"}) \cdot P(\text{Adt}=\text{"FAST"}) \\ &= 0.6041 * 0.1061 * 0.0511 * 0.4600 \\ &= 0.0015 \end{aligned}$$

$$\begin{aligned} \text{Kelas} &= \text{"MEDIUM"} \\ &= P(\text{Unit}=\text{"MEDIUM"}) \cdot P(\text{Sbe}=\text{"MEDIUM"}) \cdot P(\text{Mek}=\text{"MEDIUM"}) \cdot P(\text{Adt}=\text{"MEDIUM"}) \\ &= 0,4983 * 0,1007 * 0,0380 * 0,4613 \\ &= 0.000879 \end{aligned}$$

$$\begin{aligned} \text{Kelas} &= \text{"LOW"} \\ &= P(\text{Unit}=\text{"LOW"}) \cdot P(\text{Sbe}=\text{"LOW"}) \cdot P(\text{Mek}=\text{"LOW"}) \cdot P(\text{Adt}=\text{"LOW"}) \\ &= 0,4863 * 0,0992 * 0,0446 * 0,3846 \\ &= 0,000827 \end{aligned}$$

c). Hasil dari perkalian atribut dikalikan dengan prior probability.

$$\begin{aligned} P(C_i \mid \text{Class Speed} = \text{"FAST"}) & \cdot P(X \mid \text{Class Speed} = \text{"FAST"}) \\ &= 0.370568 * 0,0015 = 0.000555 \end{aligned}$$

$$\begin{aligned} P(C_i \mid \text{Class Speed} = \text{"MEDIUM"}) & \cdot P(X \mid \text{Class Speed} = \text{"MEDIUM"}) \\ &= 0,433706 * 0.000879 = 0.000381 \end{aligned}$$

$$\begin{aligned} P(C_i \mid \text{Class Speed} = \text{"LOW"}) & \cdot P(X \mid \text{Class Speed} = \text{"LOW"}) \\ &= 0,195726 * 0,000827 = 0.000161 \end{aligned}$$

d). Penentuan hasil dengan membandingkan hasil perkalian setiap kelas dicari nilai terbesar.

$$\begin{aligned} P(C_i \mid \text{Class Speed} = \text{"FAST"}) & \cdot P(X \mid \text{Class Speed} = \text{"FAST"}) \parallel P(C_i \mid \text{Class Speed} = \text{"MEDIUM"}) \cdot P(X \mid \text{Class Speed} = \text{"MEDIUM"}) \parallel P(C_i \mid \text{Class Speed} = \text{"LOW"}) \cdot P(X \mid \text{Class Speed} = \text{"LOW"}) \\ &= 0.000555 \parallel 0.000381 \parallel 0.000161 \end{aligned}$$

Kesimpulan yang diperoleh dari contoh kasus diatas adalah kelas Speed = "FAST"

C. Pengujian

Selanjutnya akan digunakan teknik pengujian Black Box Testing terhadap sistem atau aplikasi data mining dan akan dilakukan penghitungan akurasi algoritma naïve bayes yang telah diterapkan menggunakan confusion matrix.

1. Black Box Testing

Teknik yang digunakan dalam pengujian Black Box antara lain :

a. Digunakan untuk menguji fungsi-fungsi khusus dari perangkat lunak yang dirancang.

b. Kebenaran perangkat lunak yang diuji hanya dilihat berdasarkan keluaran yang dihasilkan dari data atau kondisi masukan yang diberikan untuk fungsi yang ada tanpa melihat bagaimana proses untuk mendapatkan keluaran tersebut dan bagaimana hasil dari proses mining.

c. Dari keluaran yang dihasilkan, kemampuan program dalam memenuhi kebutuhan pemakai dapat diukur sekaligus dapat diketahui kesalahan-kesalahannya.

Hasil uji dianggap sukses jika pada tabel pengujian, hasil

yang didapat sesuai dengan kriteria evaluasi hasil dan hasil yang diharapkan. Tabel hasil pengujian dapat dilihat pada

tabel 4.c.1

Tabel 4.c.1 : Hasil Uji Aplikasi *Data Mining*.

Deskripsi	Prosedur Pengujian	Keluaran yang Diharapkan	Hasil yang Didapat	Simpulan
Fungsi login	Menekan tombol ok pada form login	Tampilan menu utama	Menu utama dapat ditampilkan	Diterima
Fungsi Exit pada item menu File	Menekan item menu exit	Program data mining akan tertutup	Program tertutup	Diterima
Fungsi Expand pada item menu SA	Menekan item menu expand	Menampilkan daftar Sa yang tersedia	Daftar SA dapat ditampilkan	Diterima
Fungsi Add, Edit dan Delete pada interface Expand	Menekan tombol Add, Edit dan Delete	Daftar Sa dan database akan diperbaharui	Database dapat diperbaharui	Diterima
Fungsi History Service pada item menu Log Service	Menekan item menu unit masuk	Menampilkan semua record pelanggan	Record pelanggan servis berkala dapat ditampilkan	Diterima
Fungsi User Setting pada item menu Setting	Menekan item menu user setting	Menampilkan daftar user yang dapat menggunakan program	Daftar user dapat ditampilkan	Diterima

Fungsi Add, Edit dan Delete pada interface User Setting	Menekan tombol Add, Edit dan Delete	Daftar user dan database akan diperbaharui	Database dapat diperbaharui	Diterima
Fungsi About pada item menu <i>Information</i>	Menekan item menu about	Ditampilkan tentang data developer data mining sistem	Data / interface about dapat ditampilkan	Diterima
Fungsi input data	Menekan input data pada toolbar	Akan muncul interface input data untuk mengisi data pelanggan	Interface Input data dapat ditampilkan	Diterima
Fungsi Prediksi pada interface Input Data	Menekan tombol prediksi	Data pelanggan akan tersimpan ke database dan didapatkan kategori speed pengerjaan	Record pelanggan dapat disimpan ke database dan didapatkan kategori speed	Diterima

D. Pengujian Algoritma Naïve Bayes

Pengujian ini bertujuan untuk mengetahui unjuk kerja dari algoritma NBC dalam mengklasifikasikan data ke dalam kelas yang telah ditentukan. Pada uji coba ini, diberikan data latih untuk membentuk tabel probabilitas. Langkah selanjutnya akan diberikan data uji untuk menguji tabel probabilitas yang sudah terbentuk. Unjuk kerja diperoleh dengan memberikan nilai pada confusion matrix untuk menghitung nilai precision, recall, dan accuracy dari hasil pengujian.

a. Percobaan ke-1 diperoleh tabel nilai *precision*, *recall*, dan *accuracy*.

$$\text{Precision: } P = 0 / (0 + 0) \times 100\% = 0 \%$$

$$\text{Recall: } R = 10 / (10 + 0) \times 100\% = 100 \%$$

$$\text{Accuracy: } A = (10 + 0) / 10 \times 100\% = 100 \%$$

b. Percobaan ke-2 diperoleh tabel nilai *precision*, *recall*, dan *accuracy*.

$$\text{Precision: } P = 1 / (0 + 1) \times 100\% = 100 \%$$

$$\text{Recall: } R = 9 / (9 + 1) \times 100\% = 90 \%$$

$$\text{Accuracy: } A = (9 + 1) / 11 \times 100\% = 90,90 \%$$

c. Percobaan ke-3 diperoleh tabel nilai *precision*, *recall*, dan *accuracy*.

$$\text{Precision: } P = 1 / (0 + 1) \times 100\% = 100 \%$$

$$\text{Recall: } R = 9 / (9 + 1) \times 100\% = 90\%$$

$$\text{Accuracy: } A = (9 + 1) / 11 \times 100\% = 90,90\%$$

Berdasarkan pengujian black box di atas, dapat disimpulkan bahwa Aplikasi Data Mining dapat digunakan untuk mengetahui hasil hubungan model kendaraan, mekanik, km servis berkala dan additional job berupa kategori speed pengerjaan. Pada pengujian algoritma naïve bayes menggunakan confusion matrix didapatkan rata-rata precision sebesar 66.66 %, recall sebesar 93.33 %, dan accuracy sebesar 93.93 %.

V. PENUTUP

A. Kesimpulan

Kesimpulan yang dapat diambil dalam pengerjaan tugas

akhir ini adalah :

1. Aplikasi Data Mining ini dapat digunakan untuk menampilkan prediksi speed pengerjaan servis berkala kendaraan. Informasi yang ditampilkan berupa daftar service advisor, daftar unit kendaraan yang pernah servis berkala, daftar user aplikasi, serta mendapatkan kategori speed pengerjaan berdasarkan unit, km servis berkala, mekanik dan additional job. Hasil dari proses data mining ini dapat digunakan service advisor dan control room sebagai pertimbangan dalam mengambil keputusan lebih lanjut tentang faktor yang mempengaruhi waktu pengerjaan servis berkala.

2. Dari sistem yang telah dijalankan didapatkan hasil akurasi naïve bayes sebesar 93.93%, recall sebesar 93.33% , dan precision sebesar 66.66%.

B. Saran

Untuk pengembangan Aplikasi Data Mining lebih lanjut, dapat menggunakan teknik klasifikasi dengan algoritma Naïve Bayes yang lain. Seperti algoritma C.4.5, metode Nearest Neighbor dan Decision Tree. Selain dikembangkan dengan algoritma lain data testing yang digunakan sebagai penghitungan, sebaiknya selalu diperbaharui secara rutin agar estimasi waktu yang didapatkan diperoleh dari hasil penghitungan data terbaru.

REFERENCES

- [1] J. Han and M. Kamber, *Data Mining: Concepts and Techniques* Second Edition, San Francisco: Morgan Kaufmann Publishers, 2006.
- [1] [2] R. C. Nagendra K V, "Customer Behaviour Analysis Using CBA," *IJECCCE*, pp. 65-68, 2012.
- [2] [3] N. Purwokerto, "<http://toyotapurwokerto.blogspot.com>," Nasmoco Purwokerto, 07 2010. [Online]. Available: <http://toyotapurwokerto.blogspot.com/2010/07/servis-berkala.html>. [Accessed 24 Oktober 2014].
- [3] [4] M. Durairaj and C. Vijitha, "Educational Data Mining for Prediction of Student Performance Using Clustering Algorithm," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 0975-9646, pp. 5987-5991, 2014.
- [4] [5] B. Santosa, *Data Mining. Teknik Pemanfaatan Data untuk Keperluan Bisnis*, 1 ed., Yogyakarta: Graha Ilmu, 2007.
- [5] [6] L. T. Daniel, *Data Mining Methods and Models*, John Wiley & Sons, Inc Publication, 2006.
- [6] [7] Bustami, "Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi," *Jurnal Penelitian Teknik Informatika*, p. 129.
- [7] [8] R. S. Pressman, *Software Engineering edition 2*, Mc Graw Hill, 2001.
- [8] [9] Isom, "Pengertian Netbeans," 06 Juni 2010. [Online]. Available: <http://www.isomwebs.net/2012/09/pengertian-netbeans/>. [Accessed 07 July 2014].
- [9] [10] L. Thomson and L. Welling, *PHP and MySQL Web Development*, Indiana-USA: SAMS, 2001.
- [10] [11] J. Peter, "xampp," 4 Juni 2014. [Online]. Available: www.ruangkecil.or.id. [Accessed 4 Juni 2014].
- [11] [12] A. Ramadhan and H. Saputra, *SQL Server 2000 dan Visual Basic*, Jakarta: PT. Elex Media Komputindo, 2005.
- [12] [13] C. B. B. Agarwad, *Software Engineering & Testing*, Boston, 2010.
- [13] [14] T. Sriwahyuni, "Implementasi Perancangan Sistem Informasi Ekspedisi Paket Pada PT. Pos Indonesia," *Jurnal Teknologi Informasi & Pendidikan*, vol. 4, p. 45, 2011.
- [14] [15] M. Bramer, *Principles of Data Mining*. London: Springer, ISBN-10: 1-84628-765-0, 2007.
- [15] [16] T. W. Diana, S. T and L. Ardytha, "PREDIKSI HASIL PEMILU LEGISLATIF DKI JAKARTA MENGGUNAKAN NAÏVE BAYES DENGAN ALGORITMA GENETIKA SEBAGAI FITUR SELEKSI," pp. 5-6, 2012.
- [16] [17] J. J. Hox and H. R. Boeije, "Primary vs Secondary," in *Encyclopedia of Social Measurement*, Utrecht, Elsevier, 2005, p. 593.