

Prediksi Tingkat Kepenuhan Penumpang Pesawat dari Bandara Hang Nadim Batam

You Ari Faeni
BPS Kabupaten OKU Timur
Provinsi Sumatera Selatan
Martapura, Indonesia
you@bps.go.id

Nyanwar Eko Pribadi
Badan Pengusahaan
Batam
Batam, Indonesia
nyanwar@bpbatam.go.id

Jackson Bobby Romano Daba
BPS Kabupaten Bengkayang
Provinsi Kalimantan Barat
Bengkayang, Indonesia
jackson@bps.go.id

Abstract— Batam is a very strategic area located in the Indonesia Malaysia Singapore Growth Triangle region. Transportation is an important factor in the economic growth of Batam City, especially air transportation. Various types of data are collected, but only a few are analyzed so that they can generate new knowledge. This research will analyze flight data from Hang Nadim Batam airport to predict the fullness of aircraft from Batam to various destinations. The results showed that the support vector machine (SVM) model showed a higher level of accuracy than decision tree in predicting the fullness of the aircraft.

Keywords— support vector machine, decision tree, predictive analytic, data mining

I. PENDAHULUAN

Pembangunan Batam diharapkan mampu menjadi kawasan berdaya saing di dalam perkembangan global sehingga dapat memberikan nilai tambah bagi Indonesia. Batam Rempang Galang atau disingkat Barelang di desain agar menjadi salah satu kawasan baru yang maju. Posisi Batam merupakan lokasi yang strategis, berada di jalur pelayaran internasional paling ramai ke dua di dunia setelah Selat Dover di Inggris. Pulau Batam juga berada di kawasan Segitiga Pertumbuhan Indonesia Malaysia Singapura (Indonesian-Malaysia-Singapore Growth Triangle, IMS-GT). Wilayah Batam tidak perlu dipagar karena sudah terpisah dengan sendirinya dari daratan lainnya oleh lautan. Dengan topografi seperti itu, semua akses dan kegiatan yang ada di pulau tersebut menjadi sangat mudah diawasi dan diarahkan.

Posisi strategis Batam itulah yang mendorong pemerintah menetapkan Batam untuk dikembangkan sebagai kawasan industri (industrial zone) berteknologi tinggi untuk Indonesia. Batam diharapkan mampu menjadi pionir gerakan industrialisasi nasional. Bahkan, Batam tidak hanya menjadi kawasan industri yang berdaya saing, tetapi juga menjadi pusat perdagangan bebas di kawasan Asia Pasifik.

Untuk dapat memanfaatkan peluang mata rantai perdagangan dunia tersebut, pemerintah pusat melalui BP Batam telah membangun sarana infrastruktur yang memadai seperti pelabuhan transshipment yang terintegrasi dalam FTZ Batam dan juga bandara internasional. Bandara internasional Hang Nadim Memberikan layanan pesawat (datang, berangkat, lokal), penumpang (datang, berangkat, transit), dan cargo (bongkar, muat, transit). Dengan Fasilitas Pelayanan antara lain : Operator Penerbangan, Operator Lapangan, Pos, Kurir, dan Kargo, Operator Bagasi, Transportasi Darat, Asuransi, Tour & Travel, Bank dan Tempat Penukaran Uang, Executive Lounge, Pemerintah dan Pelayanan Umum, Katering, Agen Iklan Dalam dan Luar Ruang, Koneksi Internet berbasis Wifi, Cafe, Restoran, Toko. Dan Fasilitas Bandara berupa: Runway, Apron, Terminal, Aircraft, Terminal Capacity, Freight Storage, Fuel Storage.

Dalam rangka meningkatkan daya saing bandara hang nadim, maka perlu dilakukan sebuah langkah – langkah strategis dalam melakukan transformasi digital sehingga mampu meningkatkan layanan – layanan bandara Hang Nadim . salah satunya adalah menjadikan data sebagai dasar analisa dalam pengambilan keputusan baik dalam rangka meningkatkan efektifitas dan efisiensi operasional bandara maupun inovasi layanan bandara. Dalam makalah ini, kami melakukan kegiatan analisa data penumpang bandara berasal dari data

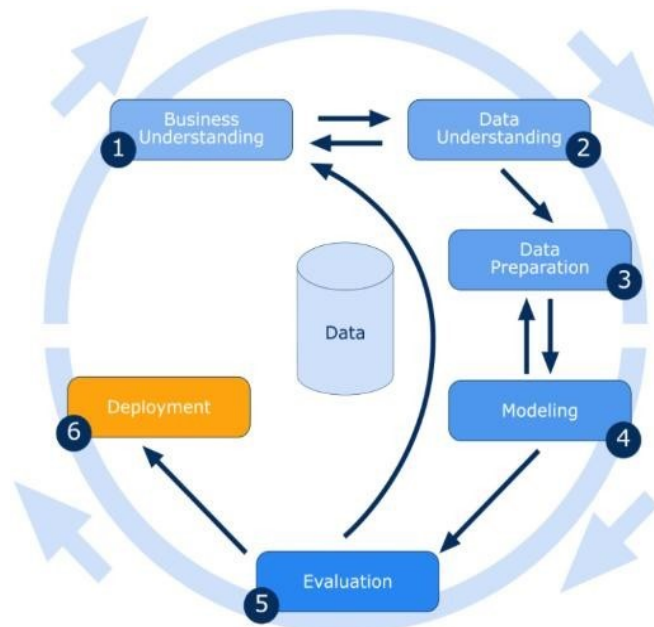
manifest yang disediakan oleh maskapai. Dari data penumpang tersebut, dilakukan prediksi tingkat kepenuhan pesawat pada hari dan bulan tertentu sehingga dapat dijadikan informasi manajemen bandara dalam meningkatkan efektifitas dan efisiensi operasional bandara hang Nadim.

II. TINJAUAN PUSTAKA

Untuk mendapatkan solusi penelitian, maka diperlukan analisa data manifest bandara Hang Nadim Batam melalui proses data mining yang dikenal dengan framework Cross Industry Standard Process Model for Data Mining (CRISP-DM). Selain itu untuk melakukan Predictive Analytics menggunakan 2 metode yaitu Support Vector Machine (SVM), dan Decision Tree

A. Data Mining

Data Mining adalah proses menganalisis pola data untuk dikategorisasi menjadi informasi yang berguna, yang dikumpulkan dari berbagai sumber data seperti data warehouse dan data informasi yang digunakan untuk memotong biaya pengeluaran dan mendapatkan penghasilan. Data Mining digunakan untuk memfasilitasi kinerja organisasi dalam membuat keputusan berdasarkan historis data.



Gambar 1. Proses Data Mining (CRISP-DM)

Proses pada Data Mining yang dikenal dengan CRISP-DM :

1. Business Understanding

Fase ini untuk memahami tujuan bisnis dengan jelas dan mencari tahu apa kebutuhan bisnis. Dari tujuan bisnis dan situasi saat ini, buat tujuan Data Mining untuk mencapai tujuan bisnis dalam situasi saat ini.

2. Data Understanding

Fase ini dimulai dengan pengumpulan data awal, yang dikumpulkan dari sumber data yang tersedia, untuk membantu membiasakan diri dengan data. Beberapa kegiatan penting harus dilakukan termasuk memuat data dan integrasi data agar pengumpulan data berhasil. Data perlu dieksplorasi dengan menangani pertanyaan Data Mining, yang dapat diatasi dengan menggunakan kueri, pelaporan, dan visualisasi.

3. Data Preparation

Hasil dari fase persiapan data adalah set data akhir. Setelah sumber data yang tersedia diidentifikasi, mereka perlu dipilih, dibersihkan, dibangun dan diformat ke dalam formulir yang diinginkan. Tugas eksplorasi data pada kedalaman yang lebih besar dapat dilakukan selama fase ini untuk memperhatikan pola berdasarkan pemahaman bisnis.

4. Modeling

Teknik pemodelan harus dipilih untuk digunakan untuk set data yang disiapkan. Model perlu dinilai secara hati-hati yang melibatkan pemangku kepentingan untuk memastikan bahwa model yang dibuat dipenuhi inisiatif bisnis

5. Evaluation

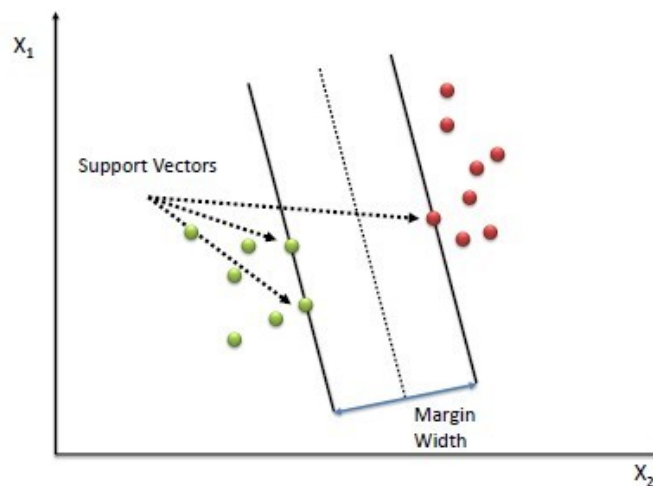
Pada fase evaluasi, hasil model harus dievaluasi dalam konteks tujuan bisnis pada fase pertama. Dalam fase ini, persyaratan bisnis baru dapat dinaikkan karena pola baru yang telah ditemukan dalam hasil model atau dari faktor lain. Mendapatkan pemahaman bisnis adalah proses berulang dalam Data Mining. Keputusan akhir harus dibuat dalam langkah ini untuk pindah ke fase Deployment.

6. Deployment

Pengetahuan atau informasi, yang diperoleh melalui proses Data Mining, perlu disajikan sedemikian rupa sehingga pemangku kepentingan dapat menggunakannya saat mereka menginginkannya. Berdasarkan persyaratan bisnis, fase deployment bisa sesederhana membuat laporan atau serumit proses Data Mining yang berulang di seluruh organisasi. Dalam fase deployment, rencana deployment, pemeliharaan, dan pemantauan harus dibuat untuk implementasi dan juga dukungan di masa mendatang. Dari sudut pandang proyek, laporan akhir proyek perlu meringkas pengalaman proyek dan meninjau proyek untuk melihat apa yang perlu untuk meningkatkan pelajaran yang telah dipelajari.

B. Support Vector Machine

SVM atau Support Vector Machine adalah model linier untuk masalah klasifikasi dan regresi. Itu dapat memecahkan masalah linear dan non-linear dan bekerja dengan baik untuk banyak masalah praktis. Support Vector Machine (SVM) melakukan klasifikasi dengan menemukan hyperplane yang memaksimalkan margin antara dua kelas. Vektor (kasus) yang menentukan hyperplane adalah vektor dukungan.



Gambar 2. Support Vector Machine

Algoritma pada Support Vector Machine sebagai berikut :

1. Mendefinisikan hyperplane optimal dengan cara memaksimalkan margin
2. Memperluas definisi di atas untuk masalah yang dapat dipisah secara non-linear dengan melihat data yang tidak terklasifikasi.

3. Memetakan data ke vector machine di mana lebih mudah untuk diklasifikasikan dengan permukaan keputusan linier dengan merumuskan kembali masalah sehingga data dipetakan secara implisit ke vector machine.

C. Decision Tree

Metodologi Decision Tree adalah metode Data Mining yang umum digunakan untuk membangun sistem klasifikasi berdasarkan beberapa kovariat atau untuk mengembangkan algoritma prediksi untuk variabel target. Metode ini mengklasifikasikan populasi menjadi segmen seperti cabang yang membangun tree terbalik dengan simpul akar, simpul internal, dan simpul daun. Algoritma ini non-parametrik dan dapat secara efisien menangani kumpulan data yang besar dan rumit tanpa memaksakan struktur parametrik yang rumit. Ketika ukuran sampel cukup besar, data studi dapat dibagi menjadi set data pelatihan (training) dan validasi (testing). Menggunakan dataset training untuk membangun model Decision Tree dan dataset testing untuk memutuskan ukuran tree yang sesuai yang diperlukan untuk mencapai model akhir yang optimal. Penggunaan Umum metode Decision Tree :

1. Pemilihan variable

Pemilihan variabel bertahap dalam analisis regresi, metode Decision Tree dapat digunakan untuk memilih variabel input paling relevan yang harus digunakan untuk membentuk model pohon keputusan, yang selanjutnya dapat digunakan untuk merumuskan hipotesis dan menginformasikan penelitian selanjutnya

2. Menilai kepentingan relatif variabel

Setelah satu set variabel yang relevan diidentifikasi, peneliti mungkin ingin tahu variabel mana yang memainkan peran utama. Secara umum, variabel penting dihitung berdasarkan pengurangan akurasi model (atau dalam kemurnian simpul di pohon) ketika variabel dihapus. Dalam sebagian besar keadaan, semakin banyak catatan variabel berpengaruh, semakin besar pentingnya variabel.

3. Penanganan nilai yang hilang

Metode umum dalam menangani data yang hilang adalah untuk mengecualikan kasus dengan nilai yang hilang; ini tidak efisien dan berisiko menimbulkan bias dalam analisis. Analisis Decision Tree dapat menangani data yang hilang dengan dua cara: itu dapat mengklasifikasikan nilai yang hilang sebagai kategori terpisah yang dapat dianalisis dengan kategori lain atau menggunakan model Decision Tree yang dibangun dengan mengatur variabel dengan banyak nilai yang hilang sebagai variabel target untuk membuat prediksi dan mengganti yang hilang ini dengan nilai prediksi.

4. Prediksi.

Ini adalah salah satu penggunaan paling penting dari model Decision Tree. Menggunakan model tree yang berasal dari data historis, mudah untuk memprediksi hasil untuk catatan di masa mendatang.

5. Manipulasi data

Terlalu banyak kategori dari satu variabel kategorikal atau data kontinu adalah umum dalam penelitian. Dalam keadaan ini, model Decision Tree dapat membantu dalam memutuskan bagaimana cara terbaik menentukan variabel kategori menjadi jumlah kategori yang lebih mudah dikelola atau bagaimana membagi variabel dalam rentang angka tertentu.

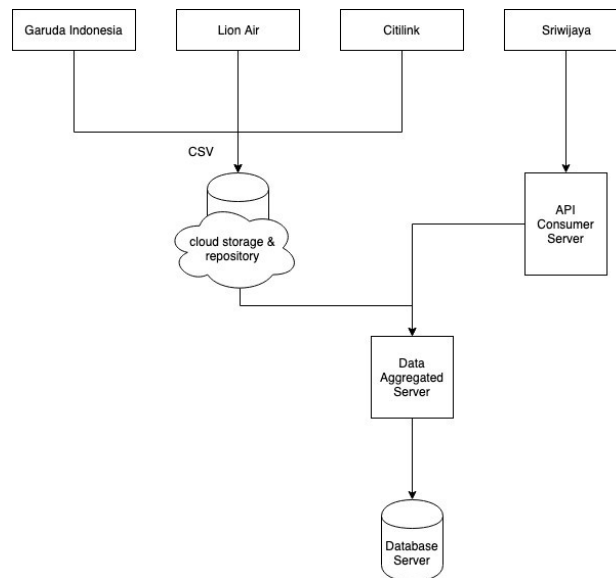
III. DATA PRE-PROCESSING

A. Pengumpulan Data

Bagian ini memaparkan poin-poin pengumpulan data, data manifest dikumpulkan melalui 2 metode yaitu :

1. upload dokumen csv harian oleh maskapai ke cloud storage yang disiapkan oleh pihak bandara Hand Nadim Batam.
2. konsumsi API manifest yang telah disediakan pihak maskapai oleh server API consumer bandara Hang Nadim Batam.

data yang tersimpan dilakukan proses konversi, transformasi dan agregasi sehingga data yang diterima memiliki format data sesuai standar aplikasi, database dan operasional bandara Hang nadim Batam. terlampir diagram pengumpulan data manifest penerbangan bandara hang nadim Batam.



Gambar 3. Proses Pengumpulan Data

B. Cleaning Data

Data yang digunakan merupakan data log dari aktivitas pesawat pada bandara Hang Nadim, Batam. Data berasal dari perusahaan penerbangan dan dikumpulkan oleh BP Batam. Data semula dalam database oracle dan kemudian dikonversi menjadi bentuk excel (TBAR_MANIFEST_AIRCRAFT.xls). Data berisi informasi nama maskapai, nomor penerbangan, tanggal departure, tanggal arrival, jenis pesawat, kapasitas pesawat, jumlah penumpang (dewasa, anak dan bayi), jumlah bagasi, jumlah kargo dsb. Data terdiri dari 14218 record yang berisi manifest pesawat dari bulan oktober 2018-april 2019.

Tahap selanjutnya adalah data preparation yang terdiri dari proses pemilihan variable yang akan dipakai dan menghilangkan data anomaly serta missing data. pemilihan variable disesuaikan dengan tujuan analisis. Dalam makalah ini dipilih beberapa variabel yang akan digunakan antara lain:

1. Nama maskapai
2. Asal bandara
3. Tujuan bandara
4. Waktu keberangkatan
5. Kapasitas maksimal pesawat
6. Jumlah penumpang

A	B	C	D	E	O	P	Q	R	S
AIRLINE_NAME	FROM_PORT	TO_PORT	ETD	SEAT_MAX	ADULT	CHILD	INFANT	HOUR	MINUTE
CITILINK	SUB	BTH	11/24/2018 6:20	180	155	3	1	6	20
CITILINK	BTH	CGK	11/24/2018 14:15	180	173	5	1	14	15
CITILINK	BTH	CGK	11/24/2018 18:20	180	162	16	4	18	20
CITILINK	CGK	BTH	11/24/2018 12:05	180	165	6	2	12	5
LION AIR	BTH	KNO	10/4/2018	215	207	2	3	0	0
LION AIR	BTH	SUB	10/4/2018 14:55	215	206	5	3	14	55
LION AIR	BTH	PNK	10/4/2018 12:00	189	87	1	0	12	0
CITILINK	CGK	BTH	11/24/2018 16:00	180	152	7	3	16	0
CITILINK	BTH	KNO	11/24/2018 10:20	180	163	14	6	10	20
CITILINK	BTH	KNO	11/24/2018 17:30	180	169	7	5	17	30
CITILINK	KNO	BTH	11/24/2018 13:00	180	172	8	6	13	0
CITILINK	KNO	BTH	11/24/2018 15:20	180	168	11	6	15	20
CITILINK	BTH	PDG	11/24/2018 16:20	180	174	6	3	16	20
CITILINK	PDG	BTH	11/24/2018 18:15	180	163	16	8	18	15
CITILINK	BTH	PKU	11/24/2018 13:30	180	171	8	1	13	30
CITILINK	PKU	BTH	11/24/2018 15:00	180	169	8	5	15	0
CITILINK	PNK	BTH	11/24/2018 8:15	180	89	2	1	8	15
CITILINK	SUB	BTH	11/24/2018 11:55	180	164	5	5	11	55
CITILINK	BTH	CGK	11/24/2018 9:10	180	172	8	1	9	10

Table 1: Data Overview

Selanjutnya dilakukan cleaning data dengan menghilangkan missing value dan data pencilan. Beberapa informasi hilang yang masih bisa dicari nilainya seperti kapasitas maksimal, nama maskapai dilakukan imputasi data. selanjutnya dilakukan cleaning data, dari 14218 record ditemukan:

- 198 blank data pada data penumpang
- 6 data outlier

Data yang sudah clean kemudian dilakukan pengolahan awal data dengan membuat variable baru dari variable yang sudah ada antara lain: pembuatan variable hari dan jumlah penumpang. Hari digenerate dari ETD(estimated time departure) dengan rumus “=text(ETD,”dddd”)”. Jumlah penumpang dihitung dari penjumlahan penumpang dewasa dan anak.

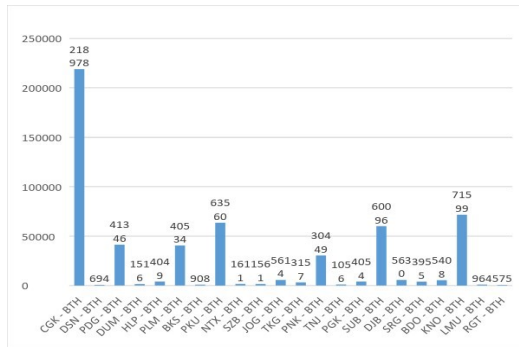
ETD	Hari	SEAT_MAX	ADULT	CHILD	INFANT	Jumlah penumpang I
11/24/2018 6:20	Saturday	180	155	3	1	158
11/24/2018 14:15	Saturday	180	173	5	1	178
11/24/2018 18:20	Saturday	180	162	16	4	178
11/24/2018 12:05	Saturday	180	165	6	2	171
10/4/2018	Thursday	215	207	2	3	209
10/4/2018 14:55	Thursday	215	206	5	3	211
10/4/2018 12:00	Thursday	189	87	1	0	88
11/24/2018 16:00	Saturday	180	152	7	3	159
11/24/2018 10:20	Saturday	180	163	14	6	177
11/24/2018 17:30	Saturday	180	169	7	5	176
11/24/2018 13:00	Saturday	180	172	8	6	180
11/24/2018 15:20	Saturday	180	168	11	6	179
11/24/2018 16:20	Saturday	180	174	6	3	180
11/24/2018 18:15	Saturday	180	163	16	8	179
11/24/2018 13:30	Saturday	180	171	8	1	179
11/24/2018 15:00	Saturday	180	169	8	5	177
11/24/2018 8:15	Saturday	180	89	2	1	91
11/24/2018 11:55	Saturday	180	164	5	5	169
11/24/2018 9:10	Saturday	180	172	8	1	180
11/24/2018 14:40	Saturday	180	157	10	3	167

Table 2: Pembuatan Variabel Baru

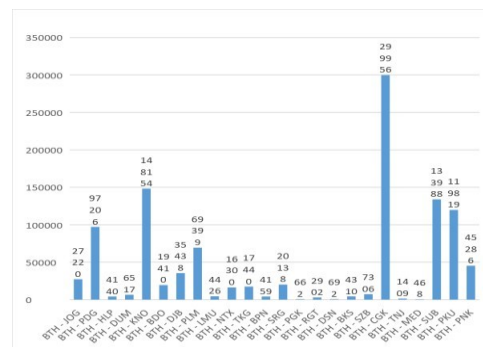
C. Visualisasi Data

Untuk dapat lebih memahami data maka akan di tampilkn visualisasi terhadap data setelah di lakukan pengolahan seperti summary, counting ataupun grouping. Visualisasi yang ditampilkan antara lain :

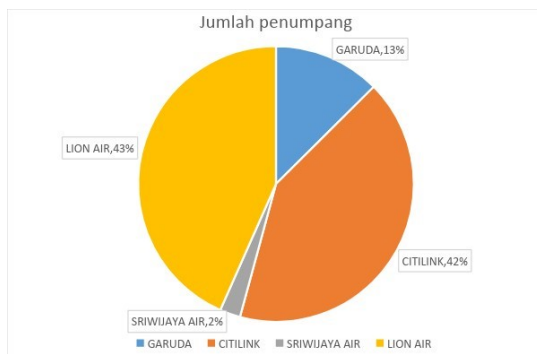
1. Jumlah penumpang selama 6 bulan terakhir menuju bandara Hang Nadim Batam
2. Jumlah penumpang selama 6 bulan terakhir dari bandara Hang Nadim Batam
3. Persentase Jumlah penumpang berdasarkan maskapai dari bandara Hang Nadim Batam
4. Jumlah penerbangan setiap maskapai dari bari bandara Hang Nadim Batam
5. Rata – rata jumlah penumpang setiap maskapai dari bandara Hang Nadim Batam



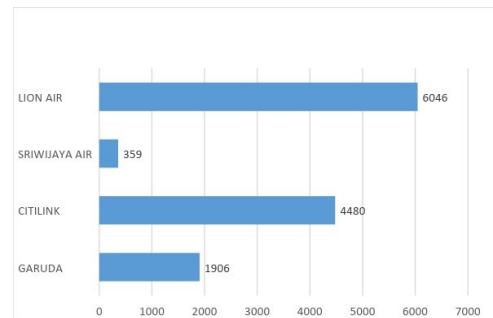
Gambar 4. Jumlah penumpang selama 6 bulan terakhir menuju bandara Hang Nadim Batam



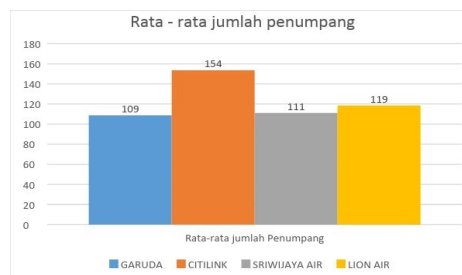
Gambar 5. Jumlah penumpang selama 6 bulan terakhir dari bandara Hang Nadim Batam



Gambar 6. Persentase Jumlah penumpang berdasarkan maskapai dari bandara Hang Nadim Batam



Gambar 7. Jumlah penerbangan setiap maskapai dari Bandara Hang Nadim Batam



Gambar 8. Rata - rata jumlah penumpang setiap maskapai dari bandara Hang Nadim Batam

D. Data Pre-Modeling

Pada bagian ini dilakukan proses pengolahan data melalui beberapa tahapan sehingga didapatkan hasil pengujian berupa prediksi tingkat kepenuhan pesawat.

a. IV.1. Pre-Modelling Task

Dalam tahap ini data pada variable mulai dikategorikan sesuai dengan kebutuhan. Data waktu dikategorikan menjadi 3 kategori :

- kategori 1 (pagi: pukul 04.00 - 11.59),
- kategori 2 (siang: pukul 12.00 - 17.59),
- kategori 3 (malam: pukul 18.00 - 03.59)

Data hari diubah menjadi integer 1- 7 berurutan dari minggu-sabtu. Maskapai diubah menjadi integer dengan nilai 1=citilink, 2=lion air, 3=garuda, dan 4=sriwijaya air. Data tujuan diklasifikasikan menjadi 4 antara lain:

- Kategori 1 (tujuan CGK)
- Kategori 2 (tujuan jawa, selain Jakarta)
- Kategori 3 (tujuan sumatera)
- Kategori 4 (tujuan selain jawa, sumatera dan Jakarta)

Selanjutnya untuk variable yang akan diestimasi adalah tingkat kepenuhan pesawat yang didapatkan dari hasil bagi dari jumlah penumpang dengan kapasitas pesawat. Pesawat dianggap penuh apabila hasil bagi jumlah penumpang dengan kapasitas pesawat lebih dari 90%. Hasil dari data sebelum dilakukan pemodelan adalah sebagai berikut:

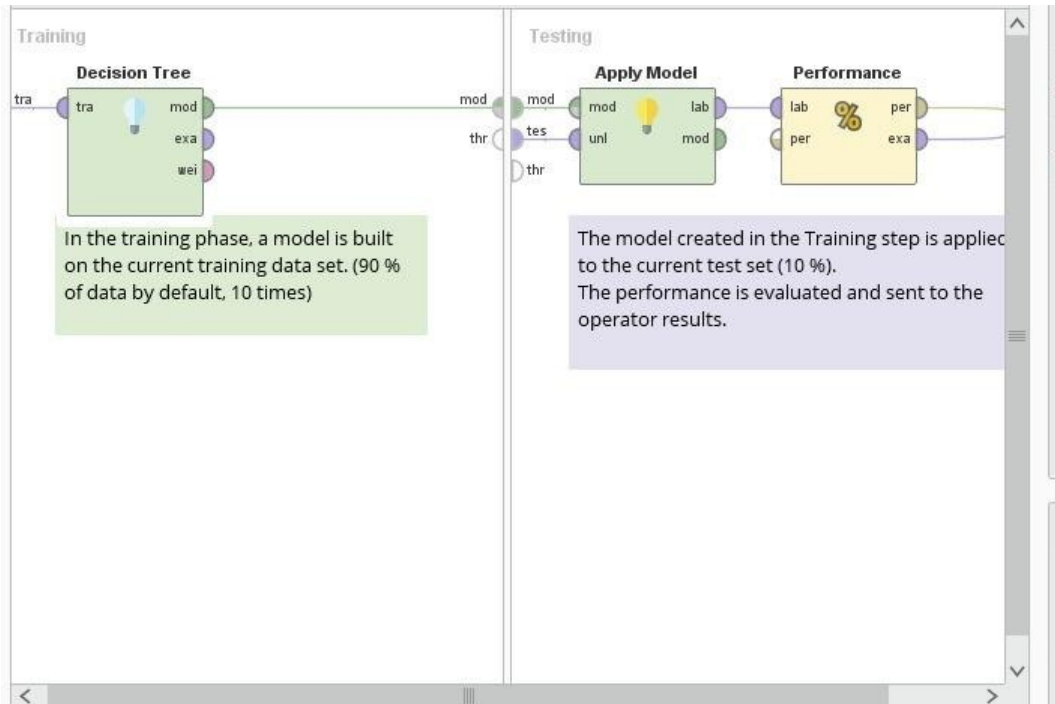
penuh	maskapai	hari	bulan	waktu	TUJUAN
yes	1	7	11	2	3
yes	1	7	11	3	3
yes	2	5	10	3	1
yes	2	5	10	2	1
no	2	5	10	2	4
yes	1	7	11	1	1
yes	1	7	11	2	1
yes	1	7	11	2	3
yes	1	7	11	2	3
yes	1	7	11	1	1
yes	1	7	11	2	1
no	1	3	11	1	3
yes	4	4	10	2	3
yes	4	4	10	2	4
yes	1	3	11	1	3
no	1	3	11	2	3
yes	1	3	11	3	1
yes	1	3	11	2	3
yes	1	3	11	2	3
no	1	3	11	2	4

Table 3: Hasil data premodeling task sebelum pembuatan model

IV. EKSPERIMEN DAN HASIL

A. Data Modeling

Data modeling menggunakan 2 metode yaitu support vector machine(SVM) dan decision tree. Data diolah menggunakan software R dan rapid miner. Data tingkat kepenuhan pesawat akan diprediksi menggunakan beberapa predictor anatara lain maskapai, hari, bulan, waktu dan tujuan. Software rapid miner menggunakan metode decision tree sedangkan software R menggunakan metode svm.



Gambar 9. Decision Tree dengan Rapid Miner

Decision Tree dibuat menggunakan software rapid miner dimulai dengan melakukan import data penumpang kemudian membuat aturan validasi. Data displit menjadi 2 80% data training dan 20% data uji. Setelah itu dibuat model menggunakan decision tree dan kemudian dilihat bagaimana performancinya.

```
#SVM_pesawat
#variabel tujuan digabung 4 (cgk) dan 8 (hlp), 4 tidak ada
rm(list = ls())
#Data penumpang
penum=read.csv("penumpang.csv")
head(penum, 5)
attach(penum)

#split data menjadi 80 persen data train
smp_size <- floor(0.80 * nrow(penum))
set.seed(10^6)
train_id <- sample(1:nrow(penum), size = smp_size)
trainset <- penum[train_id, ]
testset <- penum[-train_id, ]

#split testset
x=subset(testset,select = -penuh)
y=testset$penuh

#SVM awal
library(e1071)

#Memasukkan C dan gamma optimal
svm_model_after_tune<-svm(x,y,data=penum, kernel="radial", cost=10, gamma=.5)
summary(svm_model_after_tune)

#Running prediksi dengan cost dan gamma optimal
pred2<-predict(svm_model_after_tune,x)

#Melihat hasil confusion matrix result dari prediksi
table(pred2,y)
```

Gambar 10. Support Vector Machine dengan R

Model SVM dibangun menggunakan software R dengan membagi data menjadi 2 80% data training dan sisanya untuk data testing. Library yang digunakan adalah e1071. Svm menggunakan kernel radial dengan cost 10 dan gamma=5.

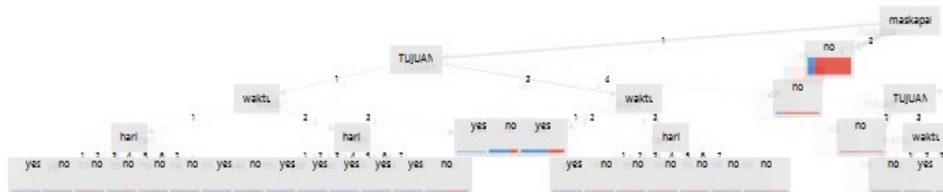
dilihat yaitu nilai asli yes tapi diprediksi model tidak sebanyak 238, nilai asli no tapi diprediksi yes sebanyak 98

Tabel 1. Hasil Akurasi SVM menggunakan aplikasi R

	True No	True Yes	Persentase
Pred No	1198	238	83,42%
Pred Yes	98	311	76,04%
Persentase	92,43%	56,67%	81.78%

C. Evaluasi Model

Model decision Tree diuji menggunakan data asli dengan parameter predictor(maskapai,hari,bulan,waktu dan tujuan) semua bernilai 1. Hasil decision Tree menunjukkan akan bernilai yes pada tingkat kepenuhan pesawat. Setelah dilihat dengan data asli maka hasil ditemukan dari 16 data 14 bernilai yes dan 2 bernilai no.



penuh	maskapai	hari	bulan	waktu	TUJUAN
yes	1	1	1	1	1
no	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1
no	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1
yes	1	1	1	1	1

Gambar 13. Evaluasi Decision Tree

Untuk model SVM langsung diuji menggunakan data asli dimasukan dalam syntax R. dari 4 jenis data uji dihasilkan 3 no dan 1 yes. Setelah dilihat pada table awal ditemukan bahwa hasil dari model sudah sesuai dengan data asli. Hasil uji dari svm dapat dilihat dari gambar dibawah ini:

```
> maskapai=c(1, 2, 3, 4)
> hari=c(7, 5, 2, 3)
> bulan=c(12, 10, 2, 3)
> waktu=c(2, 3, 1, 2)
> tujuan=c(8, 5, 8, 9)
> cek=data.frame(maskapai, hari, bulan, waktu, tujuan)
> cek
  maskapai hari bulan waktu tujuan
1         1   7   12     2       8
2         2   5   10     3       5
3         3   2    2     1       8
4         4   3    3     2       9
> prediksi<-predict(svm_model_after_tune, cek)
> prediksi
  1  2  3  4
yes no no no
Levels: no yes
> |
```

penyakit	maskapai	hari	bulan	waktu	tujuan
yes	1	7	12	2	8
no	1	7	12	2	15
yes	1	7	12	2	17
no	1	7	12	2	19
yes	1	7	12	2	15
yes	1	7	12	2	15
yes	1	7	12	2	8
no	1	7	12	2	19

Gambar 14. Syntax R Untuk Pengujian Model

V. KESIMPULAN

Dari penjelasan diatas dapat disimpulkan bahwa dalam membuat suatu model prediktif diperlukan beberapa tahapan dan asumsi yang harus dipenuhi. Hasil menunjukkan bahwa dengan menggunakan metode SVM akurasi data menjadi 81,7%. Sedangkan untuk penggunaan metode Decision Tree akurasi data menjadi 79,36%. Kasus yang dibahas pada makalah ini untuk pembuatan model tingkat kepenuhan pesawat, model SVM memiliki tingkat akurasi yang lebih tinggi jika dibandingkan dengan metode Decision Tree.

Manfaat dari hasil penelitian ini antara lain; Pihak bandara dalam mengalokasikan jumlah pegawai yang menunggu pada tiap2 terminal penerbangan; Pihak rekanan penjual bandara dapat mengalokasikan waktu buka dan tutup; Pihak maskapai dapat mengkaji ulang jadwal dan rute penerbangan.

REFERENCES

- [1] Abbott, D., 2014. Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst. Hoboken, New Jersey: Wiley.
- [2] Cichosz, P., 2015. Data Mining Algorithm: Explained Using R. Wiley.
- [3] Lantz, B., 2015. Machine Learning With R. 2nd ed. Birmingham: Packt Publishing Ltd
- [4] Wessler, M., 2014. Predictive Analytics For Dummies. Hoboken, New Jersey: Wiley.
- [5] Rachmat, Arif. 2019. *Survei Penerapan Model Machine Learning Dalam Bidang Keamanan Informasi*. Bandung: Jurnal Sistem Cerdas 2019 Volume 02 No 01 ISSN : 2622-8254 Hal :47 -60.