



Implementasi Algoritma J48 Data Mining Untuk Inovasi Bisnis Perhotelan Di Masa Pandemi Covid- 19 (Studi Kasus Hotel SNS Semarang)

¹Ana Wahyuni, ²Sri Anggraini
^{1,2}Universitas AKI Indonesia

Alamat Surat

Email: ana.wahyuni@unaki.ac.id , sri.anggraini@unaki.ac.id

Article History:

Diajukan: 27 Maret 2021; Direvisi: 15 April 2022; Diterima: 25 April 2022

ABSTRAK

Bisnis perhotelan merupakan salah satu bisnis yang terkena dampak signifikan terhadap pandemic covid-19. Pandemi ini mengharuskan mengurangi kegiatan di luar rumah dan tidak mengizinkan kegiatan *public* yang berpotensi menyebabkan kerumunan. Dampak tersebut yaitu sangat menurunnya tingkat okupansi dan *Mice* (*Meeting, Incentives, Convention & Exhibitions*) sehingga menurunkan profit hotel. Di lain pihak, biaya operasional hotel tetap tinggi sehingga banyak manajemen hotel yang terpaksa merumahkan (PHK) karyawannya. Meningkatnya PHK berarti mengancam kestabilan ekonomi dan keamanan bangsa. Hal ini menjadi latar belakang permasalahan sehingga dibutuhkan solusi yang inovatif. Permasalahan yang akan dibahas yaitu bagaimana mengimplementasikan algoritma J48 pada data mining sebagai solusi inovasi bisnis perhotelan dalam meningkatkan pendapatan, studi kasus di hotel SNS (*Sleep and Sleep*) Semarang. Tujuannya adalah mengetahui pola data pengunjung yang menginap di hotel. Metode yang digunakan algoritma J48. Hasil yang dicapai yaitu segmen pengunjung/ *customer* yang mengulang kedatangan kembali dan strategi inovasi yang sesuai dengan *customer*. Manfaat yang diperoleh yaitu dengan strategi inovasi yang tepat maka profil hotel dapat ditingkatkan sehingga operasional hotel tetap berjalan baik dan tidak mengurangi atau me-PHK karyawannya.

Kata kunci: Perhotelan, Data Mining, Algoritma J48

ABSTRACT

The hotel business is one of the businesses that has been significantly affected by the Covid-19 pandemic. This pandemic requires reducing activities outside the home and not allowing public activities that have the potential to cause crowds. This impact is the greatly reduced occupancy rate and Mice (Meeting, Incentives, Convention & Exhibitions), thereby reducing hotel profits. On the other hand, hotel operating costs remain high so that many hotel managers are forced to lay off their employees. The increase in layoffs means threatening the economic stability and security of the nation. This is the background of the problem so that innovative solutions are needed. The problems that will be discussed are how to implement the J48 algorithm in data mining as a solution to hotel business innovation in increasing revenue, a case study at the SNS (Sleep and Sleep) Semarang hotel. The goal is to find out the data patterns of visitors who stay at the hotel. The method used is the J48 algorithm. The results achieved are customer segments who repeat their return visits and innovation strategies according to customers. The benefit obtained is that with the right innovation strategy, the hotel profile can be improved so that hotel operations continue to run well.

Keywords: Hospitality, Data Mining, J48 Algorithm

1. PENDAHULUAN

Sumber keterangan ditunjuk dengan menuliskan di dalam kurung: nama akhir penulis dan tahun penerbitan. Disajikan secara sistematis sehingga didapatkan gambaran tentang dasar pembuatan makalah ini dan hasil yang diharapkan. Adanya pandemic covid-19 yang telah berlangsung lebih dari setahun menyebabkan banyak sektor terkena dampaknya termasuk kegiatan kerja praktek (KP) mahasiswa. Pada masa pandemic ini hanya mitra tertentu yang mengizinkan mahasiswa melakukan KP di tempat mitra. Salah satu mitra yaitu hotel *Sleep and Sleep* (SNS) yang menjadi unit bisnis kampus dan satu gedung dengan kampus Universitas AKI Semarang. Hal ini mendorong tim KP untuk menerapkan iptek yang diperoleh selama perkuliahan di bisnis perhotelan dengan studi kasus di hotel SNS Semarang.

Bisnis perhotelan merasakan dampak dari pandemic covid-19 yang sangat signifikan. Hal ini terjadi karena pemasukan dari okupansi dan Mice (*Meeting, Incentives, Convention & Exhibitions*) hotel menurun drastis padahal biaya operasional sangat tinggi. Walaupun pemasaran hotel telah menggunakan aplikasi *online* tetapi tingkat penambahan kasus covid-19 yang masih tinggi menjadikan banyak *customer* membatalkan kegiatan yang membutuhkan tempat (menginap ataupun *event*) di hotel. Sehingga banyak hotel yang dijual dan pihak manajemen terpaksa merumahkan karyawannya. Sektor perhotelan menyumbang peningkatan angka PHK yang menyebabkan ketidaksatabilan ekonomi dan keamanan bangsa. Hal ini tentu harus diberikan solusinya.

Salah satu potensi dari hotel yaitu adanya data pengunjung yang menginap (*customer*). Data tersebut dapat diolah menggunakan teknik data mining. Data mining merupakan serangkaian proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mendapatkan pengetahuan, informasi atau pola dari kumpulan data. (Witten 2011) Data mining menganalisa data yang telah ada pada *database* untuk tujuan memecahkan suatu masalah. Data mining mempunyai fungsi minor yaitu deskripsi, estimasi dan prediksi. Sedangkan fungsi mayor data mining yaitu *clasification, clustering* dan *association* (Susanto 2010).

Salah satu teknik data mining yaitu menggunakan algoritma J48. Algoritma J48 merupakan algoritma yang populer digunakan karena mudah untuk diinterpretasi dan dapat membantu dalam menentukan keputusan. (Kade et al. 2016) Algoritma J48 dapat mengolah data diskrit dan kontinu, dapat menangani nilai atribut yang hilang, menghasilkan aturan-aturan yang mudah diinterpretasikan dan tercepat diantara algoritma-algoritma yang menggunakan memori utama di komputer. (Hastuti 2012) Salah satu *tool* untuk menerapkan algoritma J48 yaitu aplikasi Weka.

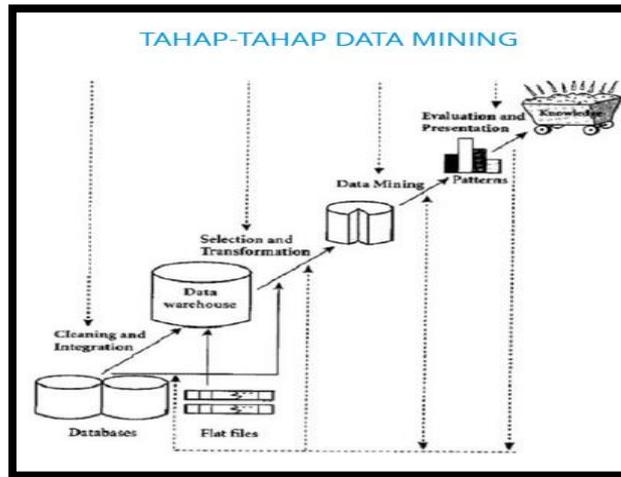
WEKA merupakan singkatan dari *Waikato Environment for Knowledge Analysis*, yang dibuat di Universitas Waikato, New Zealand. WEKA mampu menyelesaikan masalah-masalah data mining, khususnya klasifikasi yang mendasari pendekatan-pendekatan *machine learning*. (Dimov 2007) Algoritma J48 merupakan implementasi algoritma C4.5 pada Weka. Menurut Ferry, 2016 teknik data mining menggunakan algoritma J48 dapat digunakan untuk menentukan pola customer yang berguna untuk strategi manajemen hotel.

Hotel SNS Semarang memerlukan inovasi manajemen dalam meningkatkan profitnya. Salah satu inovasi yang dapat dilakukan yaitu merumuskan strategi manajemen hotel. Strategi tersebut memerlukan *database* pengunjung/ *customer* dan analisisnya. Sehingga permasalahan yang mendesak yaitu bagaimana menerapkan algoritma J48 yang bertujuan untuk mencari pola pengunjung dan memprediksi kedatangan kembali pengunjung. Pola pengunjung tersebut bermanfaat untuk pihak manajemen hotel dalam merumuskan strategi bisnis sehingga dapat meningkatkan pendapatan hotel dan meningkatkan kinerja karyawannya.

2. METODE

Penelitian ini mengambil populasi transaksi di hotel SNS Semarang selama tahun 2019. Dari sejumlah besar *database*, setelah data transaksi dilakukan proses *cleaning/* pembersihan data didapat sejumlah 5248 data.

Teknik pengumpulan data dilakukan dengan wawancara dan dokumentasi. Wawancara dilakukan dengan bagian administrasi hotel tentang paket menginap, unit pelayanan hotel dan struktur manajemen hotel. Dokumentasi dilakukan terhadap data transaksi selama tahun 2019. Metode penelitian yang digunakan yaitu metode data mining menggunakan algoritma J 48.



Gambar 1. Tahapan Data Mining (Sumber: Sis.Binus.Ac.Id)

Tahapan penelitian ini secara garis besar meliputi pengumpulan data, pengolahan data, pengujian, dan hasil analisis. Tahap pertama yaitu pengumpulan data dilakukan dengan merekap data transaksi selama tahun 2019. Kemudian dilakukan juga integrasi data transaksi yang dilakukan *on the spot* dan melalui aplikasi *online*. Proses selanjutnya yaitu pembersihan data. Data yang dibutuhkan sebagai atribut yaitu *customer* yang mengulang menginap dan tingkat pendidikannya. Sehingga data yang tidak ada atribut tersebut atau memiliki isian-isian yang tidak lengkap juga data yang tidak valid maka data akan dihapus/ dibersihkan. Data yang sudah diseleksi tersebut dalam format *Microsoft Excel 2010*, selanjutnya dilakukan transformasi data. Data dalam format *excel* ditransformasi dalam format **.scv (Comma Separated Values)* dan format **.arff (arffViewer)* untuk dilakukan proses mining oleh aplikasi Weka.

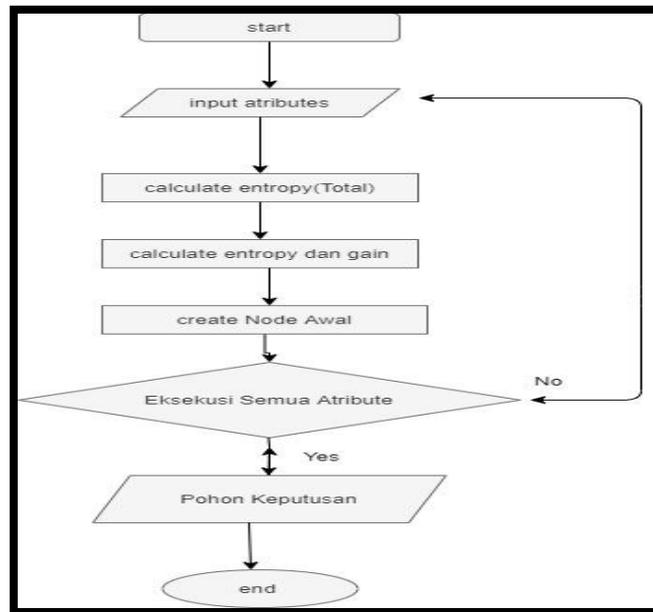
Tahap kedua yaitu pengolahan data menggunakan algoritma J48 dan aplikasi Weka. Algoritma J48 atau C4.5 (pada Weka) dilakukan dengan memilih atribut. Atribut akar dipilih berdasar nilai gain tertinggi (membutuhkan nilai perhitungan entropy). (Marcos et al. 2014)

Weka mudah digunakan dan diterapkan pada beberapa tingkatan yang berbeda. Weka mempunyai *tools* untuk *pre-processing* data, klasifikasi, regresi, *clustering*, aturan asosiasi, dan visualisasi. (Bouckaert 2008) Metodologi mengambil keputusan data Mining pada WEKA, salah satunya adalah pohon keputusan (*decision tree*). (Sujana 2010) Pohon keputusan adalah sebuah *flowchart* seperti struktur *tree*, dimana tiap internal *node* menunjukkan sebuah *test* pada sebuah atribut, setiap cabang menunjukkan hasil dari test dan *leaf node* menunjukkan kelas. Pohon keputusan dimulai dengan sebuah *root node* (titik awal) yang dipakai oleh *user* untuk mengambil tindakan. Dari *node root* ini, *user* memecahkan sesuai dengan algoritma *decision tree*. Hasil akhirnya adalah sebuah pohon keputusan dengan setiap cabangnya menunjukkan kemungkinan skenario dari keputusan yang diambil serta hasilnya. Hasil pengujian dan hasil analisisnya dari *output* aplikasi Weka kemudian digunakan untuk merumuskan strategi manajemen hotel untuk meningkatkan pemasukan hotel.

3. HASIL DAN PEMBAHASAN

Algoritma J48 merupakan implementasi algoritma C4.5 pada Weka. Algoritma C4.5 dapat menangani data numerik (kontinyu) dan diskret. Dalam algoritma C4.5 digunakan *information gain* untuk memilih atribut yang akan digunakan untuk pemisahan obyek. Pemilihan atribut yang akan menempati suatu simpul dilakukan dengan melakukan perhitungan entropi informasi (*information*

entropy) dan mencari nilai yang paling minimum. Proses C4.5, dengan memilih atribut sebagai akar (root), buat cabang untuk tiap nilai, bagi kasus ke dalam cabang, ulangi proses untuk tiap cabang hingga semua cabang memiliki kelas yang sama. Atribut akar dipilih berdasar nilai gain tertinggi yang membutuhkan nilai perhitungan entropy. Entropy adalah ukuran dari teori informasi yang dapat mengetahui karakteristik dari impurity, dan homogeneity dari kumpulan data. Entropy mengukur jumlah dari informasi yang ada pada atribut. Proses mining dijelaskan pada flowchart gambar 2 berikut:



Gambar 2.Flowchat Proses Mining

Analisis dalam penelitian ini menggunakan Pohon Keputusan (Decision Tree). Pohon keputusan ini dimulai dengan penentuan atribut yang digunakan, kemudian menghitung entropy total dan gain total dari setiap atributnya, kemudian menuliskan proses perhitungan entropy dan gain total (create node awal), selanjutnya mengeksekusi semua atribut yang akan digunakan, selanjutnya menghasilkan pohon keputusan dan proses selesai.

Sebelum menemukan gain, terlebih dahulu menghitung entropy keseluruhan dan entropy dari setiap atribut seperti menggunakan rumus berikut:

$$\text{Entropy}(S) = - \sum p_i * \log_2 p_i$$

Keterangan:

S : himpunan kasus

Pi : proporsi dari Si terhadap S.

Setelah entropy diperoleh selanjutnya menentukan gain dari setiap atribut, dan mencari gain yang tertinggi untuk dijadikan akar dari cabang pohon keputusan seperti rumus dibawah ini:

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum \frac{s_i}{s} * \text{Entropy}(S_i)$$

Keterangan:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

|Si| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

Hasil perhitungan dengan aplikasi *excel* diberikan pada tabel 1. Kolom total adalah total data masing-masing tingkat pendidikan. Kolom Ya adalah total *customer* yang mengulang menginap di hotel menurut tingkat pendidikan. Kolom Tidak adalah total *customer* yang tidak mengulang menginap di hotel. Kolom *entropy*, *gain*, *split info* dan *gain ration* adalah hasil proses mining.

Tabel 1. Data Hasil *Entropy*, *Gain*, *Split Info*, dan *Gain Ration*

Keterangan	Total	Ya	Tidak	Entropy	Gain	Split Info	Gain Ration
SMP							
SEDERAJAT	1549	772	777	0.999992484	0.681474511	0.519608796	1.311514578
MAHASISWA	1941	1546	395	0.728879699	0.707052419	0.530730205	1.332225701
SMA							
SEDERAJAT	1758	777	981	0.990264775	0.644908749	0.528549885	1.220147364
Total	5248	3095	2153	0.976632353			

Perhitungan mencari nilai *gain ratio* tertinggi yaitu atribut tingkat pendidikan mahasiswa. Nilai *gain ratio* yang paling tinggi yang akan menjadi akar pertama.

Proses mendapatkan nilai *entropy* untuk masing-masing atribut, pertama adalah mencari nilai *entropy* untuk keseluruhan data (*class* = "*tested_positive*" dan *class* = "*tested_negative*") terlebih dahulu.

Jumlah *class* target (*class*) = 2 ("*tested_positive*" dan "*tested_negative*")

Dengan perhitungan yang sama dilakukan terhadap tiap atribut dengan berdasar pengelompokan jumlah kasus pada tiap atribut dan subset atribut di dalamnya. Berdasarkan rumus perhitungan *entropy* diperoleh hasil sebagai berikut:

a. *Entropy Class* SMP SEDERAJAT

Jumlah *tested_positive* = 772

Jumlah *tested_negative* = 777

Total keseluruhan = 1549

$$\text{Entropy} = \sum - p_i * \log_2 p_i$$

$$I(772,777) = ((-772/1549) * \log_2 (772/1549)) + ((-777/1549) * \log_2 (777/1549)) = 0,999992484$$

Kemudian hitung nilai *gain* pada tiap atribut. Nilai *gain* pada atribut SMP SEDERAJAT dapat dihitung sebagai berikut ini:

$$\begin{aligned} \text{Gain (SMP SEDERAJAT)} &= (0,813518 - ((1549/5248) * 0,999992484)) \\ &= 0,681474 \end{aligned}$$

Untuk menghitung *gain ratio* dibutuhkan nilai *split info*. Untuk menghitung *split info* sebagai berikut:

$$\begin{aligned} \text{Split info} &= - \sum_{i=1}^n \frac{\text{Jumlah}}{\text{Total}} * \log_2 \frac{\text{jumlah}}{\text{Total}} \\ &= - \left(\frac{1549}{5248} \right) * \text{LOG}_2 \left(\frac{1549}{5248} \right) \\ &= 0,519608796 \end{aligned}$$

$$\text{GainRation (A)} = \left(\frac{\text{Gain (A)}}{\text{Split info (A)}} \right) = \frac{0,681474511}{0,519608796} = 1,311514578$$

b. Entropy Class MAHASISWA

Jumlah *tested_positive* = 1546Jumlah *tested_negative* = 395

Total keseluruhan = 1941

$$\text{Entropy} = \sum - p_i * \log_2 p_i$$

$$I(1546,395) = ((1546/1758) * \log_2(1546/1758)) + ((-395/1758) * \log_2(395/1758)) \\ = 0,72887969$$

Kemudian hitung nilai *gain* pada tiap atribut. Nilai *gain* pada atribut MAHASISWA dapat dihitung sebagai berikut ini:

$$\text{Gain (Class MAHASISWA)} = (0,813518 - ((1941/5248) * 0,72887969)) \\ = 0,707052419$$

Untuk menghitung *gain ratio* dibutuhkan nilai *split info*. Untuk menghitung *split info* sebagai berikut:

$$\text{Split info} = - \sum_{i=1}^n \frac{\text{Jumlah}}{\text{Total}} * \log_2 \frac{\text{jumlah}}{\text{Total}} \\ = - \left(\frac{1941}{5248} \right) * \text{LOG}_2 \left(\frac{1941}{5248} \right) \\ = \mathbf{0,530787}$$

$$\text{GainRation (A)} = \left(\frac{\text{Gain (A)}}{\text{Split info (A)}} \right) \\ = \frac{0,707052419}{0,530730205} \\ = 1,332225701$$

c. Atribut SMA SEDERAJAT

Jumlah *tested positive* = 777Jumlah *tested negative* = 981

Total keseluruhan = 1748

$$\text{Entropy} = \sum - p_i * \log_2 p_i$$

$$I(777,981) = \left(\left(- \frac{777}{1758} \right) \log_2 \left(\frac{777}{1758} \right) \right) + \left(\left(- \frac{981}{1758} \right) \log_2 \left(\frac{981}{1758} \right) \right) \\ = 0,990264775$$

Kemudian hitung nilai *gain* pada tiap atribut. Nilai *gain* pada atribut SMA SEDERAJAT dapat dihitung sebagai berikut ini:

$$\text{Gain (Class SMA SEDERAJAT)} = \text{Entropy (s)} - \frac{si}{s} \text{entropy (si)} \\ = (0,813518 - \left(\left(\frac{1758}{5248} \right) * 0,990264775 \right)) \\ = 0,644908749$$

Untuk menghitung *gain ratio* dibutuhkan nilai *split info*. Untuk menghitung *split info* sebagai berikut:

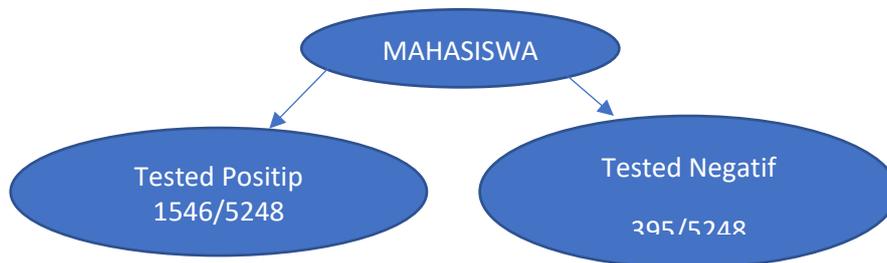
$$\text{Split info} = - \sum_{i=1}^n \frac{\text{Jumlah}}{\text{Total}} * \log_2 \frac{\text{jumlah}}{\text{Total}}$$

$$= -\left(\frac{1758}{5248}\right) * \text{LOG}_2\left(\frac{1758}{5248}\right)$$

$$= 0,5285499$$

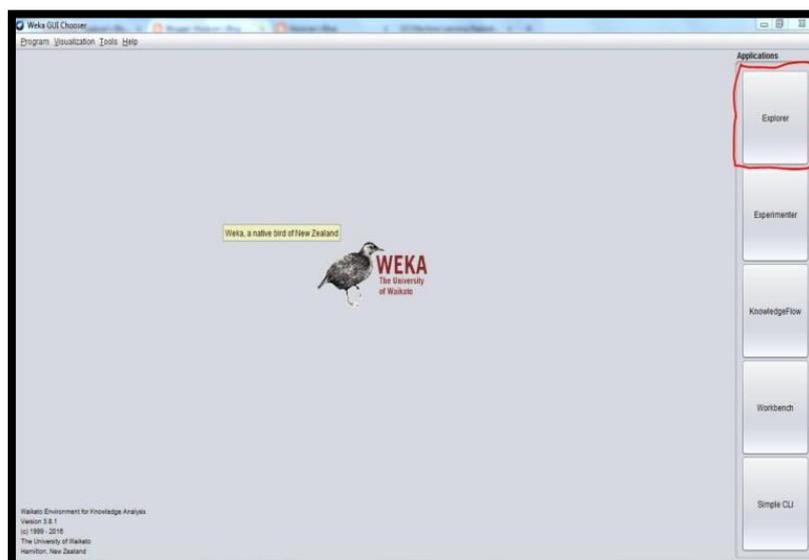
$$\text{GainRation (A)} = \frac{\text{Gain (A)}}{\text{Split info (A)}} = \frac{0,644908749}{0,5285499} = 1,220147364$$

Dapat diketahui bahwa atribut dengan *gain ratio* tertinggi dipilih sebagai atribut test 1,332225701 (Atribut MAHASISWA) sehingga prediksi/kemungkinan terbesar yang kembali menginap adalah tamu dengan status/ tingkat pendidikan Mahasiswa. Dari hasil tersebut dapat digambarkan pohon keputusan berikut:

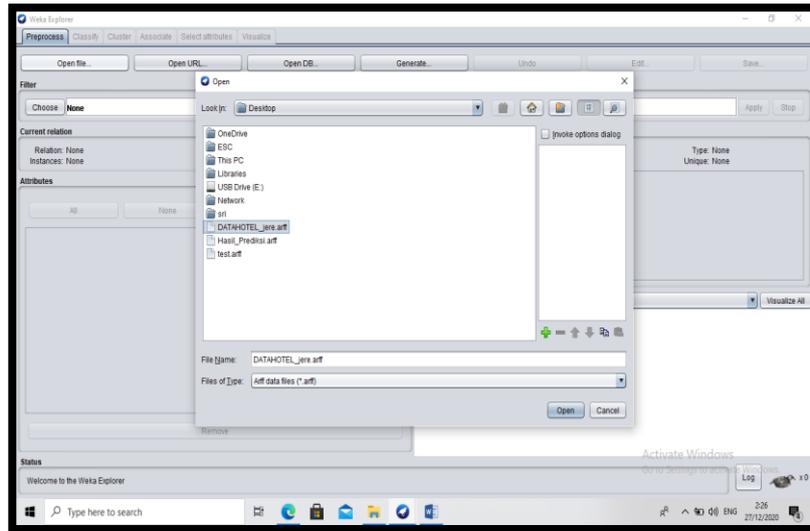


Gambar 3. Pohon Keputusan

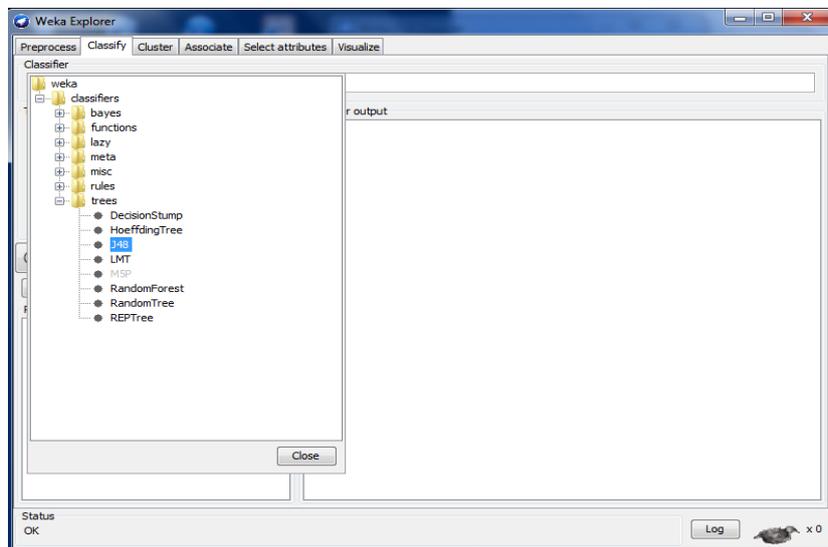
Dalam Penelitian ini digunakan juga perhitungan menggunakan Weka. Data yang diproses dan diolah dengan menggunakan *Microsoft Excel 2007*, kemudian file ditransformasi dalam bentuk *.scv (*Comma Separate Values*) dan format *.arff supaya dapat dengan mudah dibaca oleh aplikasi Weka. berikut ini implementasi perhitungan menggunakan Aplikasi WEKA 2016.



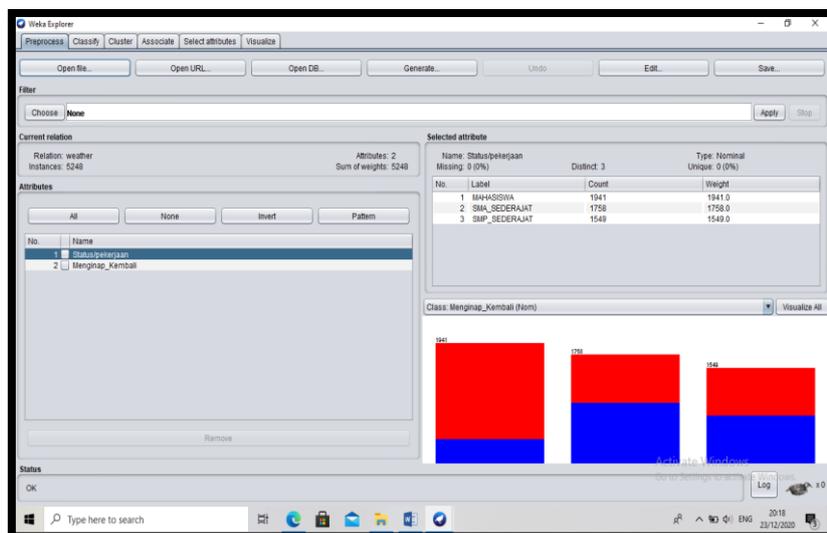
Gambar 4. Tampilan Perhitungan Menggunakan Weka (2016)



Gambar 5. Proses file converter ke Weka (2016)



Gambar 6. Tampilan Memilih Metode J48



Gambar 7. Tampilan Proses Perhitungan Metode J48

```

Time taken to build model: 0.03 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      3280           62.5 %
Incorrectly Classified Instances    1968           37.5 %
Kappa statistic                    0.2691
Mean absolute error                 0.4329
Root mean squared error             0.4654
Relative absolute error             89.4533 %
Root relative squared error         94.6105 %
Total Number of Instances          5248

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0.741   0.456   0.531     0.741   0.619     0.284   0.658    0.519    Tidak
                0.544   0.259   0.751     0.544   0.631     0.284   0.658    0.716    Ya
Weighted Avg.   0.625   0.340   0.661     0.625   0.626     0.284   0.658    0.635

=== Confusion Matrix ===

  a    b  <-- classified as
1596  557 |  a = Tidak
1411 1684 |  b = Ya

```

Gambar 8. Hasil Hitung Algoritma J48, Hasil Keluaran WEKA

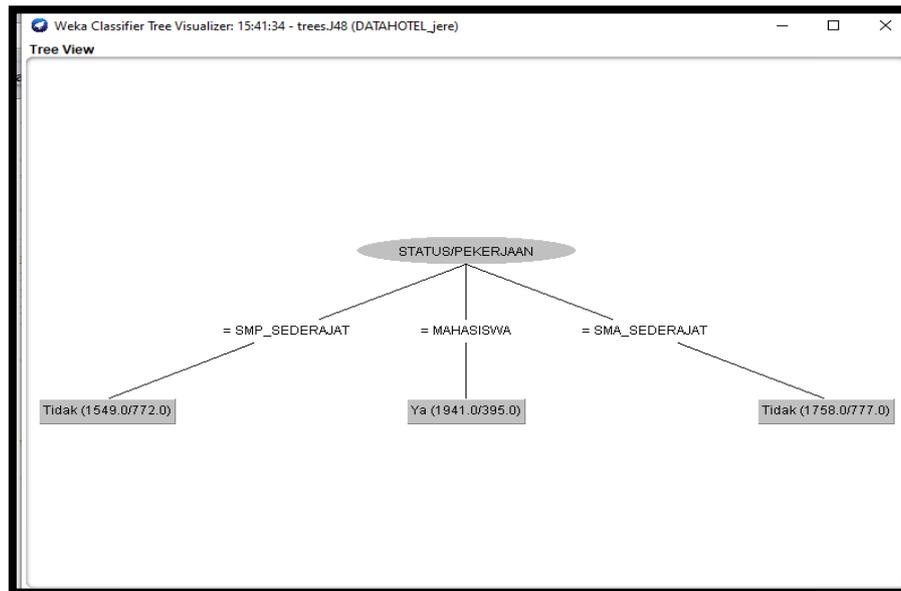
Hasil yang diperoleh dari proses perhitungan prediksi tamu yang datang kembali menginap dengan menggunakan Weka pada “*use training set*”, terlihat bahwa 67 Data Tamu Menginap kembali Hotel *Sleep and Sleep* menghasilkan hasil yang baik.

Hal ini dapat diketahui dari evaluasi indeks yang ditunjukkan pada area *Classifier output*, meliputi: banyaknya tingkat kebenaran (*Correctly Classified Instances*), sebanyak = 52 data yaitu 77. 6119 % sedangkan banyaknya tingkat kesalahan (*Incorrectly Classified Instances*), sebanyak = 15 data yaitu 22.3881 % dari keseluruhan data, *Correctly Classified Instances* dan *Incorrectly Classified Instances* ini tergantung kepada konsistensi hubungan sebab akibat antara data masukan dan data keluaran, sehingga semakin rendah data yang terklasifikasi dengan benar maka semakin rendah data yang terklasifikasi dengan benar. *Kappa statistic* sebesar 0. 5012, *kappa* adalah ukuran koreksi antara klasifikasi dan *class* yang benar, sedangkan hasil *Test option 10 folds Cross validation* menghasilkan *Mean absolute error* sebesar 0.3326 & *Root mean squared error* sebesar 0.4147. dari hasil diatas menyatakan 0% yang sama artinya akurasi diperoleh 100% menunjukkan *classifier* yang terbentuk dapat mengklasifikasikan semua data pada kelas masing-masing secara tepat.

Selanjutnya berdasarkan Uji Test *Validation Option 10 folds* menghasilkan *Relative absolute error* sebesar 70. 8159 %, dan *Root relative squared error* sebesar 85. 5824 %, berdasarkan Tingkat hubungan antara variabel.

Dari hasil perhitungan keluaran WEKA 2016 Aplikasi versi 3.9.4, untuk selanjutnya hasil tersebut memperoleh Hasil Prediksi Tamu yang mengulang kembali kedatangan dan Tamu yang Tidak Mengulang Kedatangannya.

Kemudian penelitian dilanjutkan dengan pengujian *Dissicion Tree* untuk mengetahui Pola keputusan *customer* yang kembali menginap. Berikut ini tampilan *Visualisasi Tree (disscion tree)* menggunakan Weka Explorer (2016).



Gambar 9. Tampilan Pohon Keputusan

Berdasarkan hasil dari algoritma J48 menggunakan WEKA diperoleh pola keputusan yang memiliki informasi mengenai transaksi hotel untuk mengulang menginap kembali ke Hotel *Sleep And Sleep* Semarang menurut tingkat pendidikan terbesar yaitu tingkat pendidikan mahasiswa.

Hasil mining tersebut menunjukkan porsi market terbesar *customer* hotel adalah dengan tingkat pendidikan mahasiswa. Hal ini menjadi pertimbangan manajemen hotel dalam merumuskan kebijakannya untuk meningkatkan pendapatan hotel. Mengingat masa pandemi covid-19 masih berlangsung dan okupansi hotel menurun drastis dan pihak hotel tidak menerima untuk tempat isolasi mandiri bagi *customer* yang terindikasi covid-19. Jadi strategi manajemen dari tingkat okupansi pada masa pandemi ini belum dapat ditingkatkan secara signifikan. Sehingga ditentukan sektor pemasukan hotel dari bidang penjualan makanan/ minuman (*resto*) dan *souvenir*. Terkait dengan market terbesar pada tingkat mahasiswa, maka dirumuskan strategi pemasaran makanan/ minuman (*resto*) dan *souvenir* yang sesuai dengan gaya bahasa dan gaya hidup mahasiswa.

4. SIMPULAN DAN SARAN

Teknik data mining dimanfaatkan untuk mendapat pengetahuan dari data transaksi *customer* yang kembali menginap di Hotel *Sleep And Sleep* Semarang selama 1 tahun terakhir. Hasil pengujian pada algoritma J48 di dapat akurasi pengujian sebesar 62,5% dan hasil pohon keputusan memberikan informasi bahwa tingkat pendidikan mahasiswa merupakan *customer* terbanyak yang memilih untuk menginap kembali. manajemen hotel dalam merumuskan kebijakannya untuk meningkatkan pendapatan hotel. Mengingat masa pandemi covid-19 masih berlangsung dan okupansi hotel menurun, sehingga ditentukan sektor pemasukan hotel dari bidang penjualan makanan/ minuman (*resto*) dan *souvenir*. Terkait dengan market terbesar pada tingkat mahasiswa, maka perlu dikembangkan sistem pemasaran makanan/ minuman (*resto*) dan *souvenir online* ataupun *offline* yang sesuai dengan gaya bahasa dan gaya hidup mahasiswa.

5. DAFTAR PUSTAKA

Binus University, 2021, *Data Mining, website*, <https://sis.binus.ac.id/2019/04/04/data-mining-2/>, diakses pada tanggal 17 Maret 2021

- Bouckaert, R., 2008. WEKA manual for version 3-6-0 WEKA Manual for Version 3-6-0. In University of Waikato.
- Dimov, R., 2007. *Weka: Practical machine learning tools and techniques with Java implementations*, University of Saarland.
- Hastuti, K., 2012. Analisis komparasi algoritma klasifikasi data mining untuk prediksi mahasiswa non aktif. *Prosiding Semantik*, 2012, pp.241–249.
- Kade, F. et al., 2016. Teknik Data Mining untuk Mendapatkan Pola Transaksi Hotel Bendesa dengan Algoritma J48. , (Prosiding Senapati).
- Marcos, H. et al., 2014. Implementasi Data Mining Untuk Klasifikasi Nasabah Kredit Bank “ X ” Menggunakan Classification Rule. *Prosiding Seminar Nasional Teknologi Informasi dan Multimedia 2014*, (February), pp.1–7.
- Sujana, 2010. Aplikasi mining data mahasiswa dengan metode klasifikasi decision tree. *Prosiding Snati 2010*.
- Susanto, S., 2010. Pengantar Data Mining. *Andi Yogyakarta*.
- Witten, I., 2011. *Data Mining Practical Machine Learning Tool & Techniques*, Morgan Kaufmann Publisher USA.