

Komprasi Algoritma *Decision Tree*, *SVM* dan *ANN* untuk Reservasi Hotel

¹Dwi Hartanti*, ²Afu Ichsan Pradana, ³Sri Lestari

¹²³Universitas Duta Bangsa Surakarta, Indonesia

*dwhartanti@udb.ac.id

ABSTRAK

Hotel adalah tempat peristirahatan sementara yang menyediakan fasilitas akomodasi dan layanan lainnya seperti kamar tidur, kamar mandi, restoran, kolam renang, spa, dan pusat kebugaran yang dikelola oleh perusahaan dengan melakukan pembayaran sesuai ketentuan. *Machine learning* merupakan cabang dari kecerdasan buatan yang cara bekerjanya belajar dari pemikiran manusia dengan menggunakan algoritma matematika. Penelitian ini menggunakan *Decision Tree*, *SVM*, dan *ANN*. Tujuan penelitian membandingkan akurasi dari *Decision Tree*, *SVM*, dan *ANN* dalam identifikasi Reservasi Hotel. Hasil penelitian didapatkan Algoritma *Decision Tree* didapatkan nilai akurasi sebesar 85.7%, *Support Vector Machine (SVM)*, sebesar 82.71%, dan *Artificial Neural Network (ANN)* 80.48%.

Kata Kunci: *Machine Learning*, *Decision Tree*, *SVM*, *ANN*

Latar Belakang

Hotel adalah tempat peristirahatan sementara yang menyediakan fasilitas akomodasi dan layanan lainnya seperti kamar tidur, kamar mandi, restoran, kolam renang, spa, dan pusat kebugaran yang dikelola oleh perusahaan dengan melakukan pembayaran sesuai ketentuan. Hotel dikelompokkan dalam kategori berdasarkan fasilitas dan harga, seperti hotel bintang lima, empat bintang, tiga bintang. Proses reservasi hotel adalah bagian penting dari operasi hotel, yang memungkinkan tamu untuk memesan kamar sesuai dengan kebutuhan mereka. Metode reservasi yang paling populer saat ini adalah secara online (melalui situs web hotel/aplikasi) atau melalui telepon. Metode reservasi secara online ini juga mengakibatkan semakin besarnya faktor kemungkinan pembatalan reservasi, sehingga diperlukan adanya identifikasi reservasi hotel.

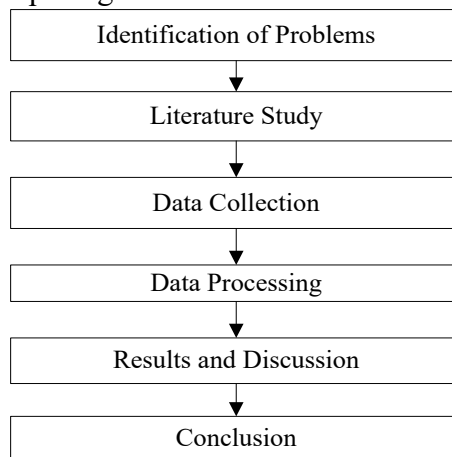
Kecerdasan Buatan merupakan ilmu yang ditujukan untuk pembuatan software dan hardware yang memiliki fungsi dalam untuk melakukan pemikiran seperti manusia. *Machine learning* merupakan cabang dari kecerdasan buatan yang cara bekerjanya belajar dari pemikiran manusia dengan menggunakan algoritma matematika.

Dalam identifikasi ini menggunakan 4 algoritma dalam Machine Learning yaitu *Decision Tree*, *Support Vector Machine (SVM)*, dan *Artificial Neural Network (ANN)*. Algoritma ini banyak digunakan untuk menyelesaikan permasalahan dan cocok untuk kasus penelitian terkait identifikasi reservasi hotel.

Penelitian yang dilakukan Moch Shandy Tsalasa Putra dkk menyimpulkan didapatkan hasil untuk nilai accuracy dengan menggunakan algoritma ANN sebesar 79.29% dan Logistic Regression 79.70% dalam kasus prediksi Pembatalan Hotel. Penelitian yang dilakukan oleh Antonius Yadi Kuntoro dkk menyimpulkan berdasarkan pengolahan data yang dilakukan didapatkan nilai untuk Algoritma SVM nilai akurasi sebesar 73.36% dan nilai *AUC* sebesar 0.794 sedangkan untuk algoritma Naïve Bayes nilai akurasi dan nilai *AUC* dari algoritma yaitu untuk SVM nilai akurasi sebesar 67.10% dan nilai *AUC* sebesar 0.573

Metodologi Penelitian

Alur penelitian yang dilakukan dalam penelitian yang dilakukan untuk Komparasi Algoritma *Decision Tree*, *SVM* dan *ANN* untuk Reservasi Hotel dapat dilihat pada gambar 1.



Keterangan

1. Melakukan identifikasi terkait dengan permasalahan terkait identifikasi reservasi hotel

2. Melakukan studi literature yang terkait dengan penelitian yang dilakukan melalui artikel, jurnal dan buku.
3. Melakukan pengumpulan data terkait reservasi hotel
4. Melakukan memproses data identifikasi reservasi hotel
5. Mendapatkan hasil akurasi dari pengolahan data dengan algoritma-algoritma yang telah ditentukan.
6. Melakukan pengambilan kesimpulan dari proses yang telah dilakukan.

Hasil dan Pembahasan

a. Data Preparation

Penelitian yang dilakukan menggunakan algoritma terkait identifikasi reservasi hotel menggunakan Dataset yang digunakan untuk data latih dalam program ini berjumlah 36276 baris, yang memiliki 14 Fitur atau Atribut mulai dari jumlah dewasa, jumlah anak anak, jumlah malam akhir pekan, jumlah malam minggu, ruang mobil parkir yang diperlukan, waktu terlama, tahun kedatangan, bulan kedatangan, tanggal kedatangan, tamu berulang, jumlah pembatalan sebelumnya, jumlah pemesanan sebelumnya tidak dibatalkan, harga rata per kamar, jumlah permintaan khusus.

Penelitian ini menggunakan class reservasi hotel dibatalkan atau tidak. Penelitian ini akan membandingkan akurasi dari algoritma Decision Tree, SVM, dan ANN.

```
In [3]: dataset.describe()
```

```
Out[3]:
```

	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	required_car_parking_space	lead_time	arrival_year	arrival_month	arrival
count	36275.000000	36275.000000	36275.000000	36275.000000	36275.000000	36275.000000	36275.000000	36275.000000	36275.000000
mean	1.844962	0.105279	0.810724	2.204300	0.030986	85.232557	2017.820427	7.423653	15.1
std	0.518715	0.402648	0.870644	1.410905	0.173281	85.930817	0.383836	3.069894	8.0
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	2017.000000	1.000000	1.0
25%	2.000000	0.000000	0.000000	1.000000	0.000000	17.000000	2018.000000	5.000000	8.0
50%	2.000000	0.000000	1.000000	2.000000	0.000000	57.000000	2018.000000	8.000000	16.0
75%	2.000000	0.000000	2.000000	3.000000	0.000000	126.000000	2018.000000	10.000000	23.0
max	4.000000	10.000000	7.000000	17.000000	1.000000	443.000000	2018.000000	12.000000	31.0

Gambar 2. Data Reservasi Hotel

b. Data Preprocessing

Pada tahapan data *preprocessing* ini melakukan persiapan data untuk data tersebut bisa digunakan seperti menganalisis data yang memiliki value kosong. Hasil dari data preprocessing tidak terdapat

data yang kosong. Hasil dari data *preprocessing* terlihat pada gambar 3.

```
In [4]: dataset.isnull().sum()
Out[4]: Booking_ID          0
        no_of_adults        0
        no_of_children      0
        no_of_weekend_nights 0
        no_of_week_nights   0
        type_of_meal_plan    0
        required_car_parking_space 0
        room_type_reserved   0
        lead_time            0
        arrival_year        0
        arrival_month       0
        arrival_date        0
        market_segment_type  0
        repeated_guest      0
        no_of_previous_cancellations 0
        no_of_previous_bookings_not_canceled 0
        avg_price_per_room   0
        no_of_special_requests 0
        booking_status       0
        dtype: int64
```

Gambar 3. Hasil data *preprocessing*

c. Training Data dan Penerapan Algoritma

Melakukan pembagian data training dan data testing untuk identifikasi reservasi hotel. Pembagian data training sebesar 80% dan data testing sebesar 20%. Hasil Pembagian data training dan data testing terlihat pada gambar 4.

```
In [10]: # membagi data training 80% dan testing 20%
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y,test_size=0.2, random_state=0)
```

Gambar 4. Pembagian Data Training dan Data Testing

1) *Decision Tree*

Dalam implementasi identifikasi reservasi hotel untuk Algoritma *Decision Tree* didapatkan nilai akurasi sebesar 85.7%. Hasil akurasi pada Gambar 5.

```
In [11]: # Decision Tree Classifier
from sklearn.tree import DecisionTreeClassifier

tree_model = DecisionTreeClassifier()
tree_model = tree_model.fit(x_train, y_train)

In [12]: from sklearn.metrics import accuracy_score
y_pred = tree_model.predict(x_test)
acc_score = round(accuracy_score(y_pred, y_test), 3)
print('The Accuracy for Decision Tree is: ',acc_score*100, '%')

The Accuracy for Decision Tree is:  85.7 %
```

Gambar 5. Hasil akurasi *Decision Tree*

2) *Support Vector Machine (SVM)*

Dalam implementasi identifikasi reservasi hotel untuk Algoritma *Support Vector Machine (SVM)* didapatkan nilai akurasi sebesar 82.71%. Hasil akurasi pada Gambar 6.

```
In [14]: # Metode Support Vector Machine (SVM)
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score

svm = SVC()
svm.fit(x_train, y_train)
y_predsvm = svm.predict(x_test)
acc_svm = accuracy_score(y_predsvm, y_test)
print('The accuracy for SVM is :', acc_svm*100, '%')

The accuracy for SVM is : 82.19159200551344 %
```

Gambar 6. Hasil akurasi *Support Vector Machine (SVM)*

3) *Artificial Neural Network*

Dalam implementasi identifikasi reservasi hotel untuk Algoritma *Artificial Neural Network (ANN)* didapatkan nilai akurasi sebesar 80.48%. Hasil akurasi pada gambar 7.

```
In [16]: # Menerapkan Metode Artificial Neural Network
from sklearn.neural_network import MLPClassifier
clf = MLPClassifier(activation='relu', solver='lbfgs', max_iter=500, alpha=1e-5, hidden_layer_sizes=(5,3))
clf.fit(x_train, y_train)

from sklearn.metrics import accuracy_score
y_model = clf.predict(x_train)
print('The accuracy Neural Network is : ', accuracy_score(y_train, y_model)*100, '%')

The accuracy Neural Network is : 80.48587181254308 %
```

Gambar 7. Hasil akurasi *Artificial Neural Network (ANN)*

Berdasarkan hasil pengolahan didapatkan nilai akurasi untuk Algoritma *Decision Tree* didapatkan nilai akurasi sebesar 85.7%., *Support Vector Machine (SVM)* didapatkan nilai akurasi sebesar 82.71%, *Artificial Neural Network (ANN)* didapatkan nilai akurasi sebesar 80.48%.

Kesimpulan

Berdasarkan hasil akurasi yang telah di olah dengan menggunakan tiga algoritma didapatkan hasil bahwa Algoritma *Decision Tree* didapatkan nilai akurasi sebesar 85.7%, *Support Vector Machine (SVM)* didapatkan nilai akurasi sebesar 82.71%, *Artificial Neural Network (ANN)* didapatkan nilai akurasi sebesar 80.48%. Untuk identifikasi reservasi hotel algoritma *Decision Tree* memiliki nilai

akurasi tertinggi dibandingkan dengan *Artificial Neural Network* (ANN) dan *Support Vector Machine* (SVM).

DAFTAR PUSTAKA

- Antonio, N., Almeida, A. de, & Nunes, L. (2017). Predicting hotel booking cancellations to decrease uncertainty and increase revenue. *Tourism & Management Studies*, 13(2), 25–39. <https://doi.org/10.18089/tms.2017.13203>
- Antonio, N., de Almeida, A., & Nunes, L. (2019). Hotel booking demand datasets. *Data in Brief*, 22, 41–49. <https://doi.org/10.1016/j.dib.2018.11.126>
- Diyah, I., Arifah, C., Nusantara, U. B., & Surabaya, U. N. (2022). Job Replacementartificial Intelligencedi Industri Jasa: Tinjauan Pustaka Sistematis. *Jurnal Ilmu Manajemen*, 10(3), 911–929.
- Handayani, F., Sari Kusuma, K., Leoni Asbudi, H., Guines Purnasiwi, R., Kusuma, R., Sunyoto, A., & Mega Pradnya, W. (2021). *JEPIN (Jurnal Edukasi dan Penelitian Informatika) Komparasi Support Vector Machine, Logistic Regression Dan Artificial Neural Network dalam Prediksi Penyakit Jantung*. 7(3), 329–334.
- Kuntoro, A. Y., Asra, T., Sistem, S., Fakultas, I., Informatika, T., Mandiri, U. N., Timur, J., Studi, P., Komputer, T., Teknik, F., Bina, U., Informatika, S., Studi, P., Perangkat, R., Fakultas, L., Informatika, T., Bina, U., & Informatika, S. (2022). Klasifikasi Keluhan Pengguna Kai Access Untuk Pemesanan. *JIKA (Jurnal Informatika) Universitas Muhammadiyah Tangerang*, 161–169.
- Lee, M. (2018). Modeling and forecasting hotel room demand based on advance booking information. *Tourism Management*, 66, 62–71. <https://doi.org/10.1016/j.tourman.2017.11.004>
- Nawassyarif, Yuli Santika, N. D. S. (2022). RANCANG BANGUN APLIKASI RESERVATION HOTEL BERBASIS WEB (STUDI KASUS HOTEL TAMBORA SUMBAWA). *JINTEKS (Jurnal Informatika Teknologi Dan Sains)*, 4(2), 87–93.

- Putra, M. S. T., & Azhar, Y. (2021). Perbandingan Model Logistic Regression dan Artificial Neural Network pada Prediksi Pembatalan Hotel. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 6(1), 29–37. <https://doi.org/10.14421/jiska.2021.61-04>
- Ramdan, D. S. (2014). Komparasi Antara Algoritma Support Vector Machine Dan Artificial Neural Network Untuk Pemodelan Data Time Series. *Tedc*, 8(2), 154–159.
- Roihan, A., Sunarya, P. A., & Rafika, A. S. (2020). Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper. *IJCIT (Indonesian Journal on Computer and Information Technology)*, 5(1), 75–82. <https://doi.org/10.31294/ijcit.v5i1.7951>
- Strandberg, R., & Laas, J. (2019). *A comparison between Neural networks, Lasso regularized Logistic regression, and Gradient boosted trees in modeling binary sales*. <https://www.diva-portal.org/smash/get/diva2:1319871/FULLTEXT02>
- Tuhenay, D., & Mailoa, E. (2021). *Perbandingan Klasifikasi Bahasa Menggunakan Metode Naïve Bayes Classifier (Nbc) Dan Support Vector Machine (Svm) Comparison of Language Classification Using Naive Bayes Classifier (Nbc) and Support Vector Machine (Svm) Method*. 4(2), 105–111. <https://doi.org/10.33387/jiko>
- Zein, A. (2021). Kecerdasan Buatan Dalam Hal Otomatisasi Layanan. *Jurnal Ilmu Komputer JIK*, 4(2), 16–25. <https://jurnal.pranataindonesia.ac.id/index.php/jik/article/download/96/49>