

Komparasi Akurasi Metode *Random Forest* Dan *Deep Learning* pada Reservasi Hotel

Saifudin

Universitas Bina Sarana Informatika

e-mail: saifudin.sfn@bsi.ac.id

Diterima	Direvisi	Disetujui
12-06-2023	06-07-2023	20-07-2023

Abstrak - Ketersediaan kamar tempat menginap (hotel) dikota besar akan berbeda dipengaruhi oleh akhir pekan dan juga hari libur nasional. Dengan kondisi seperti ini untuk sebaiknya sebelum melakukan perjalanan ke kota lain sebaiknya melakukan reservasi hotel terlebih dahulu yang bertujuan jika sudah sampai dikota yang dituju dapat langsung beristirahat dengan nyaman. Reservasi ini bisa dilakukan melalui website/online dan melalui telepon/offline. Pada era media sosial seperti sekarang ini untuk melakukan reservasi hotel sangat dimudahkan sekali, dan juga dibantu dari ulasan yang diberikan oleh orang yang pernah menginap di hotel tersebut. Pengalaman yang kurang dalam hal reservasi hotel dapat menyebabkan dibatakannya reservasi hotelnya. Pada saat melakukan reservasi hotel sebaiknya diteliti terlebih dahulu fasilitas-fasilitas yang disediakan oleh pihak hotel, lokasi hotel, dan juga yang harus dipertimbangkan adalah harga sewanya supaya dapat mengambil keputusan apakah menginap atau dibatalkan resevsi hotelnya. Penelitian ini menggunakan metode *Random Forest* dan *Deep Learning*. Dataset diambil dari laman Kaggle dengan judul *reservation hotel dataset*. Nilai *performance* tertinggi ada pada model dengan menggunakan metode *Random Forest* yakni mempunyai nilai akurasi 76,86%, presisi 83,45% dan *recall* 84,03%.

Kata Kunci: Komparasi, *Random Forest*, *Deep Learning*

Abstract - The availability of rooms to stay (hotels) in big cities will differ depending on weekends and also national holidays. With conditions like this, before traveling to other cities, you should make a hotel reservation in advance which aims if you have arrived at the destination city can immediately rest comfortably. This reservation can be made through the website / online and by phone / offline. In the era of social media like now, making hotel reservations is very easy, and also helped by reviews given by people who have stayed at the hotel. Lack of experience in terms of hotel reservations may result in cancellation of hotel reservations. When making a hotel reservation, you should first examine the facilities provided by the hotel, the location of the hotel, and also what must be considered is the rental price in order to make a decision whether to stay or cancel the hotel reservation. This research uses *Random Forest* and *Deep Learning* methods. The dataset is taken from the Kaggle page with the title *reservation hotel dataset*. The highest performance value is in the model using the *Random Forest* method, which has an accuracy value of 76.86%, precision of 83.45% and recall of 84.03 %.

Keywords: Comparisons, *Random Forest*, *Deep Learning*

PENDAHULUAN

Banyak orang melakukan perjalanan kota lain untuk melakukan wisata atau dikarenakan pekerjaan/perjalanan dinas maka diperlukan tempat untuk beristirahat atau menginap supaya kondisi kesehatan tetap terjaga. Ketersediaan kamar tempat menginap (hotel) dikota-kota besar akan berbeda-beda dipengaruhi oleh akhir pekan dan juga hari libur nasional. Dengan kondisi seperti ini untuk sebaiknya sebelum melakukan perjalanan ke kota lain sebaiknya melakukan reservasi hotel terlebih dahulu yang bertujuan jika sudah sampai dikota yang

dituju dapat langsung beristirahat dengan nyaman. Reservasi ini bisa dilakukan melalui website/online dan melalui telepon/offline. Pada era media sosial seperti sekarang ini untuk melakukan reservasi hotel sangat dimudahkan sekali, dan juga dibantu dari ulasan yang diberikan oleh orang yang pernah menginap di hotel tersebut. Pengalaman yang kurang dalam hal reservasi hotel dapat menyebabkan dibatakannya reservasi hotelnya. Pada saat melakukan reservasi hotel sebaiknya diteliti terlebih dahulu fasilitas-fasilitas yang disediakan oleh pihak hotel, lokasi hotel, dan juga yang harus



dipertimbangkan adalah harga sewanya supaya dapat mengambil keputusan apakah menginap atau dibatalkan resepsi hotelnya. Sebagian besar reservasi hotel batal dikarenakan pembatalan atau ketidakhadiran. Alasan pembatalan biasanya dikarenakan ada perubahan rencana, konflik penjadwalan.

Reservasi adalah proses, pembuatan, pemesanan (tempat, barang, dll) kepada orang lain. Selain itu, semakin banyak lokasi menarik yang dapat digunakan sebagai lokasi acara khusus. Umumnya, lokasi yang relatif menarik ini jauh dari perkotaan. Lokasi yang jauh dari perkotaan sering diminati banyak orang karena memiliki pemandangan yang indah. Namun, di balik banyaknya peminat, informasi yang tersedia masih minim. Beberapa telah menggunakan media sosial sebagai media promosi. Namun tetap saja, informasi yang bisa didapatkan melalui akun media sosial tidak lengkap. Begitu juga bagi pelanggan yang ingin melakukan reservasi harus mengunjungi lokasi untuk mendapatkan informasi lengkap dan membuat perjanjian reservasi (Rahmatya dkk., 2020).

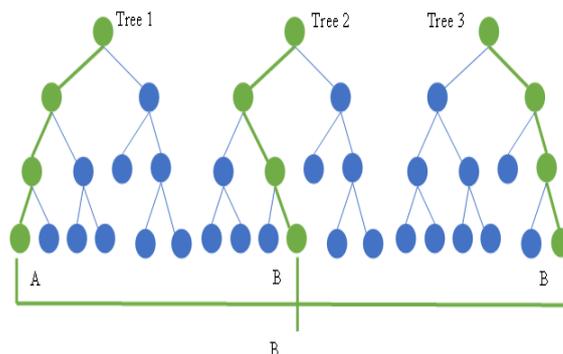
Menurut Monaghan dalam (Wardani, 2021) Reservasi merupakan suatu proses klerikal ataupun klerikal atau elektronik dimana produk ekspedisi semacam tiket pesawat, kamar hotel, kamar pada kapal pesiar ada buat dipakai serta pada kesimpulannya dibeli oleh orang secara khusus. Kata reservasi dalam dunia pariwisata lebih dikenal dengan istilah *booking*.

RANDOM FOREST

Menurut Han dalam (Khoirunissa dkk., 2021) *Random Forest* dimulai dengan teknik dasar penambangan data, pohon keputusan. Pada pohon keputusan dimasukkan input di bagian atas (*root*) kemudian turun ke bawah (*leaf*) untuk menentukan data termasuk kelas apa. Hutan Acak adalah klasifikasi yang terdiri dari kumpulan klasifikasi pohon terstruktur di mana setiap pohon melempar unit suara untuk kelas yang paling populer. Dengan kata lain, *Random Forest* terdiri dari sekumpulan pohon keputusan, di mana kumpulan pohon keputusan digunakan untuk mengklasifikasikan data ke kelas.

Menurut (Pramoedyo, 2022) *Random Forest* adalah pengembangan pohon keputusan dengan menggunakan banyak pohon keputusan dimana setiap pohon keputusan telah dilatih menggunakan sampel individu dan setiap atribut dibagi menjadi pohon yang elit antara satu set atribut acak dan setiap atribut dibagi menjadi pohon yang dipilih antara subset atribut acak dan setiap atribut dibagi menjadi pohon yang dipilih antara satu set atribut acak dan setiap atribut *Random Forest* mungkin merupakan

teknik organisasi yang mengumpulkan variabel *freelance* masih sebagai data sampel, yang mengarah ke pohon klasifikasi dengan berbagai ukuran dan bentuk.



Sumber: (Leonardo dkk., 2020)

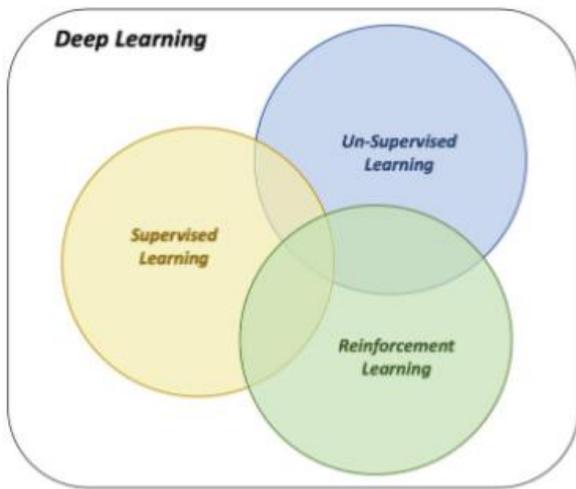
Gambar 1. Proses pengujian *Random Forest*

DEEP LEARNING

Deep learning adalah sejenis metode pembelajaran mesin berdasarkan jaringan saraf. Ini mengacu pada mode berpikir otak hewan, dan dapat mengenali fitur secara otomatis tanpa pengawasan apa pun. Dapat mengolah *big data* secara akurat, sehingga menghemat banyak sumber daya manusia (Yunfei Lai, 2019)

Deep learning adalah salah satu cabang ilmu pendidikan mesin yang memakai basis jaringan saraf tiruan (JST), ataupun dapat dikatakan kalau *deep learning* ini merupakan pengembangan dari JST. Dengan memakai *deep learning*, dimana suatu *personal computer* (PC) sanggup menekuni klasifikasi secara langsung lewat foto ataupun suara (Ilahiyah & Nilogiri, 2018).

Beberapa teknik dalam *deep learning* dapat dikategorikan menjadi *supervised*, *semi-supervised*, dan *unsupervised*. Kategori lain seperti *Reinforcement Learning* atau *Deep Reinforcement Learning*, seringkali dikategorikan menjadi *semi-supervised* atau *unsupervised* (Diponegoro dkk., 2021).



Sumber: (Diponegoro dkk., 2021)

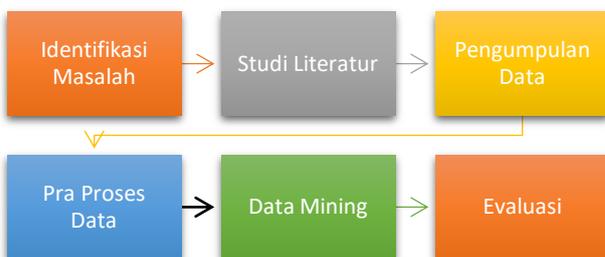
Gambar 2. Kategori dalam Deep Learning

METODE PENELITIAN

Pengumpulan data

Pada penelitian ini menggunakan data yang berasal dari dataset reservasi hotel yang diunduh dari site <https://www.kaggle.com/datasets>.

Dalam melakukan penelitian, satu langkah yang harus diperhatikan adalah tahapan penelitian. Tahap penelitian ini merupakan suatu proses untuk mendapatkan informasi dari data yang diolah atau diidentifikasi secara sistematis, logis dan ilmiah. Proses penelitian ini memiliki beberapa tahapan yaitu sebagai berikut:



Sumber: (Astuti dkk., 2022)

Gambar 3. Tahapan Penelitian

TEKNIK SAMPLING

Beberapa teknik yang digunakan untuk penentuan sample salah satunya adalah teknik simple random sampling. Data penelitian sendiri didapat dari laman kaggle *datasets* sejumlah 36.275 baris data, dan akan diambil sampel sebagai data *training* dengan menggunakan teknik *simple random sampling* berdasarkan teori slovin, diharapkan mampu memberikan jawaban yang akurat terhadap

jumlah populasi tanpa memperhatikan tingkat anggota populasi yang dipilih untuk menjadi sampel.

Menurut Sugiyono dalam (Febriyanti & Widiyanto, 2017) mengatakan sederhana(simpler) disebabkan cara pengambilan responden ilustrasi dari populasi diambil secara acak tanpa mencermati strata yang terdapat dalam populasi itu.

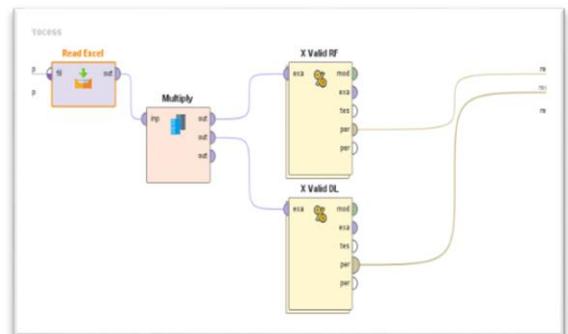
Nilai besaran sampel yang diambil menurut teori slovin dengan memakai rumus sebagai berikut:

$$n = \frac{N}{1+N(e)^2} \dots\dots\dots(1)$$

Berdasarkan rumus *simple random simpling* berdasarkan slovin dengan mengambil nilai kritis (e) sebesar 7%, maka penggunaan sampel dalam penelitian ini adalah sejumlah 203 baris data.

HASIL DAN PEMBAHASAN

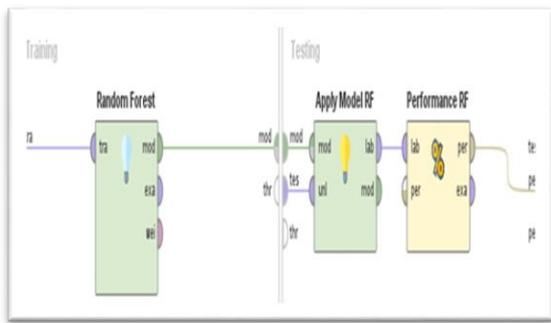
Dalam penelitian ini, pemodelan dilakukan dengan menggunakan aplikasi Rapid Mincer. Data training sebanyak 203 baris data yang diperoleh kemudian diolah untuk mendapatkan nilai prediksi akurasi dan presisi dari masing-masing metode.



Sumber: Proses di Rapidminer

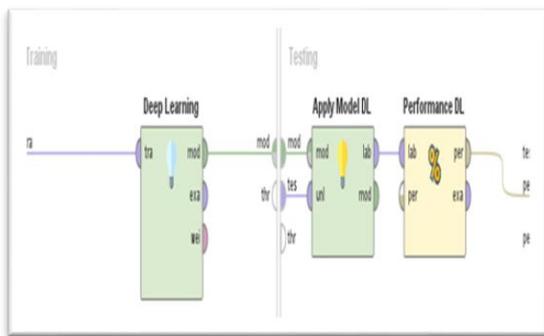
Gambar 4. Proses pengujian metode *Random Forest* dan *Deep Learning*

Model yang akan dipakai dalam proses pengujian model perbandingan dalam penelitian ini adalah algoritma *Random forest* dan *Deep Learning*. Model yang terbaik nantinya yang akan dipilih dari masing-masing pengujian. Pengukuran terhadap model algoritmik diperlukan untuk menentukan model dan hasil yang terbaik.



Sumber: Proses di Rapidminer

Gambar 5. Proses dan Pengujian Metode *Random Forest*



Sumber: Proses di Rapidminer

Gambar 6. Proses dan Pengujian metode *Deep Learning*

Pengujian Model *Random Forest* Dan *Deep Learning*

Dalam pengujian data training pada penelitian ini menggunakan dataset yang memiliki 19 atribut. Setelah dilakukan pengujian terhadap masing-masing model, hasil nilai akhir dari masing-masing pengujian dari masing-masing metode memiliki nilai akurasi rata-rata 76.86% untuk metode *Random Forest* dan 72.45% untuk model *Deep Learning*.

Evaluasi Model

Evaluasi model yang digunakan merupakan matriks penilaian universal digunakan buat mengukur performa masing-masing model yang dibanding Matriks penilaian yang digunakan, ialah akurasi, presisi, sensitivitas (*recall*), spesifisitas, serta nilai *F1-score* (Saputro & Sari, 2019).

Evaluasi terhadap kinerja model *machine learning* dalam riset ini memakai metode confusion matrix. Jumlah prediksi yang benar serta salah dirangkum dengan nilai hitungan serta dipecah oleh tiap- tiap kelas. Dalam confusion matrix buat mengukur akurasi klasifikasi, *precision*, *recall*, serta *F1-score* dengan memakai Persamaan(2) hingga(5) (NOR dkk., 2022).

$$accuracy = \frac{TP+}{TP+TN+FP+F} \dots\dots\dots (2)$$

$$precision = \frac{TP}{TP+} \dots\dots\dots(3)$$

$$Recall = \frac{TP}{TP+} \dots\dots\dots(4)$$

$$F1 - Score = \frac{2 \times precision \times Recall}{Precision+Recall} \dots\dots\dots(5)$$

Tabel 1. Perbandingan kinerja metode *Random Forest* dan *deep learning*

No	Algoritma	Akurasi	Presisi	Recall	F1 Score
1	<i>Random Forest</i>	76,86%	83,45%	84,03%	83,73%
2	<i>Deep Learning</i>	72,45%	74,44%	93,06%	82,71%

Sumber: Perhitungan Akurasi, presisi, *recall* dan *F1-score*

Berdasarkan tabel 1, pengujian untuk memprediksi reservasi hotel menggunakan metode *Random Forest* dan *deep learning* dihasilkan nilai akurasi tertinggi terdapat pada metode *Random Forest* dengan nilai akurasi 76,86%.

KESIMPULAN

Metode *Random Forest* dan *Deep Learning* dapat digunakan untuk memprediksi akurasi reservasi hotel dengan nilai akurasi diatas 76%. Berdasarkan pengujian data, metode *Random Forest* mempunyai nilai akurasi, presisi dan *recall* yang tertinggi dibandingkan dengan metode *deep learning* yaitu dengan nilai akurasi sebesar 76,86%, presisi 83,45% *recall* 84,03%.

DAFTAR PUSTAKA

Astuti, N., Utamajaya, J. N., & Pratama, A. (2022). Penerapan Data Mining Pada Penjualan Produk Digital Konter Leppangeng Cell Menggunakan Metode K-Means Clustering. *JURIKOM*, 9(3), 754–760.

Diponegoro, M. H., Kusumawardani, S. S., & Hidayah, I. (2021). Tinjauan Pustaka Sistematis: Implementasi Metode Deep Learning pada Prediksi Kinerja Murid. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 10(2), 131–138.

Febriyanti, F., & Widiyanto, K. (2017). Pengukuran Penggunaan E-Learning Pada Smk–Smak Bogor Dengan Tecnology Acceptance Model

- (Tam). *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*, 2(2), 32–38.
- Ilahiyah, S., & Nilogiri, A. (2018). Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network. *Jurnal Sistem dan Teknologi Informasi (Justindo)*, 3(2).
- Khoirunissa, H. A., Widyaningrum, A. R., & Maharani, A. P. A. (2021). Comparison of Random Forest, Logistic Regression, and Multilayer Perceptron Methods on Classification of Bank Customer Account Closure. *Indonesian Journal of Applied Statistics*, 4(1), 14–20.
- Leonardo, R., Pratama, J., & Chrisnatalis. (2020). Perbandingan Metode Random Forest Dan Naïve Bayes Dalam Prediksi Keberhasilan Klien Telemarketing. *Jutikomp*, 3(2), 455–459.
- NOR, S., MUSLIM, M. A., & ASWIN, M. (2022). Pengenalan Pola Dasar Angka berdasarkan Gerakan Tangan menggunakan Machine Learning. *ELKOMIKA*, 10(3), 596–608.
- Pramoedyo, H. (2022). Comparison Of Random Forest And Naïve Bayes Methods For Classifying And Forecasting Soil Texture In The Area Around Das Kalikonto, East Java. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 16(4), 1411–1422.
- Rahmatya, M. D., Wicaksono, M. F., Sari, D. P., & Mubarak, M. N. (2020). Design of Reservation Information System. *The 1st Paris Van Java International Seminar on Computer, Science, Engineering, and Technology (PVJ_ISComSET)*, 879(3).
- Saputro, I. W., & Sari, B. W. (2019). Uji Performa Algoritma Naïve Bayes untuk Prediksi Masa Studi Mahasiswa. *Citec Journal*, 6(1), 1–11.
- Wardani, D. M. (2021). Hotel Reservation Policy Pada Masa Pandemi: Refund, Reschedule Atau Cancel Di Labuan Bajo. *Pariwisata*, 8(1), 63–72.
- Yunfei Lai. (2019). A Comparison of Traditional Machine Learning and Deep Learning in Image Recognition. *Journal of Physics: Conference Series International Conference on Electrical, Mechanical and Computer Engineering, Guizhou, China*, 1314(3).