



Analisis Data Penjualan Tiket Pesawat Ke Jepang Menggunakan *Classic Machine Learning* Pada PT. TTD

Syarifuddin

Program Studi Teknik Informatika S-2, Universitas Pamulang, Tangerang Selatan, Banten

Email: syarif.unpam@gmail.com

ABSTRACT

Airline ticket sales to Japan is an important topic in the ever-evolving travel industry. The unpredictable demand for airline tickets is a common challenge in the aviation industry due to many complex and variable factors. This research analyzes the data on airline ticket sales to Japan using classic machine learning approaches. The data used for the research are the airline ticket sales to Japan in 2022 and 2023 for Japan Airlines and All Nippon Airways. Classical methods such as classification were applied to identify factors influencing the ticket sales patterns. The collected data was cleaned and organized to obtain a dataset for use in Machine Learning with the K-Nearest Neighbors, Naïve Bayes, and Decision Tree algorithms. After evaluating the models created using these three algorithms, the evaluation results with prediction models, model test & score, and confusion matrix, showed that the K-Nearest Neighbor algorithm achieved the highest values compared to the Naïve Bayes and Decision Tree algorithms with an accuracy of 99.5% (model predictions evaluation) & 98.9% (model test & score evaluation). The majority of ticket sales to Japan were for JAL flights, economy class tickets, and spring season being the most popular choices. The conclusion from this data is that Japan Airlines holds a strong market share in ticket sales to Japan.

Keywords: Data Mining, Classification, Airline Ticket Sales, KNN, Naïve Bayes, Decision Tree.

ABSTRAK

Penjualan tiket pesawat ke Jepang adalah topik penting dalam industri perjalanan yang terus berkembang. Permintaan tiket pesawat yang sulit diprediksi adalah tantangan umum dalam industri penerbangan karena dipengaruhi oleh banyak faktor yang kompleks dan berubah-ubah. Dalam penelitian ini dilakukan analisis data terhadap pola penjualan tiket pesawat ke Jepang menggunakan pendekatan *classic machine learning*. Data yang digunakan untuk penelitian adalah data penjualan tiket pesawat khusus ke Jepang tahun 2022 dan 2023 dengan tipe pesawat *Japan Airlines* dan *All Nippon Airways*. Metode klasik seperti klasifikasi diterapkan untuk mengidentifikasi faktor-faktor yang mempengaruhi pola penjualan tiket. Data yang sudah dikumpulkan kemudian dilakukan pembersihan dan perapihan sehingga mendapat dataset yang akan digunakan pada *Machine Learning* dengan menggunakan Algoritma *K-Nearest Neighbors*, *Naïve Bayes* dan *Decision Tree*. Setelah dilakukan evaluasi model yang dibuat menggunakan ke tiga algoritma tersebut, Hasil evaluasi dengan model *prediction*, model *test & score* dan *confusion matrix*, Pada Algoritma *K-Nearest Neighbor* mendapatkan nilai tertinggi dibandingkan dengan algoritma *Naïve Bayes* dan *Decision Tree* dengan nilai *accuracy* sebesar 99,5% (*evaluasi model predictions*) & 98,9% (*evaluasi model test & score*) dan Penjualan tiket ke Jepang terbanyak menggunakan pesawat JAL, kelas tiket kelas ekonomi dan musim *spring* merupakan pilihan terbanyak. kesimpulan dari data tersebut adalah bahwa *Japan Airlines* memegang pangsa pasar yang kuat dalam penjualan tiket ke Jepang.

Kata Kunci : Data Mining, Klasifikasi, Penjualan tiket pesawat, *KNN*, *Naïve Bayes*, *Deision Tree*.

1. PENDAHULUAN

PT. TTD (*tours & travel*) yang sudah berdiri sejak tahun 1972 merupakan agen *tour & travel* yang menawarkan paket perjalanan dan penjualan tiket pesawat khususnya untuk perusahaan Jepang yang ada di Indonesia dengan tujuan ke negara Jepang. Sebanyak 1120 perusahaan perjalanan wisata berdasarkan informasi dari statistik jasa perjalanan wisata *tours & travel service statistics 2011* membuat perusahaan saling bersaing untuk membuat penawaran yang terbaik kepada *customer* sehingga membuat para pengelola ingin menunjukkan strategi pemasaran yang lebih baik. Untuk itu maka para pengelola harus mencermati strategi penjualan dan melakukan analisa berdasarkan data-data penjualan yang tersimpan di *database*. PT. TTD tidak hanya melayani penjualan tiket pesawat ke Jepang saja tetapi ke beberapa negara dan antar pulau di Indonesia dengan sistem *Business to Business* (B2B) dengan perusahaan-perusahaan yang ada di Indonesia khususnya perusahaan Jepang.

Perkembangan sistem transportasi hadir untuk mendukung kinerja orang-orang di dunia untuk meningkatkan kualitas hidup mereka. Tidak bisa dihindari lagi, bahwa pada masa ini, sistem transportasi merupakan salah satu kebutuhan yang sangat penting dan bahkan menjadi salah satu faktor pengendali kehidupan manusia. Di masa ini, jumlah penduduk Indonesia meningkat sangat tinggi, jumlah penduduk yang tinggi, menjadi faktor penyebab meningkatnya mobilitas masyarakat dalam kehidupan sehari-hari. Tingginya mobilitas masyarakat membuat sistem transportasi diperlukan untuk mendukung penuh dalam mobilitas tersebut. Kebutuhan sistem transportasi yang tinggi, dijadikan oleh banyak orang di masa kini, sebagai salah satu peluang usaha yang menguntungkan. Untuk itu mereka membuat suatu usaha untuk mempermudah masyarakat dalam melakukan mobilitas, dengan cepat, dan mudah. Usaha-usaha itu sangat berkembang pesat di masa ini, yaitu usaha di bidang *travel*. Dalam dunia industri *travel*, banyak sekali data yang dikelola dan pengelolaan tersebut dilakukan hampir setiap hari [1].

Pada penelitian yang pernah dilakukan oleh (Suci & Samsudin, 2022) dengan judul Algoritma *K-Nearest Neighbors* dan *Synthetic Minority Oversampling Technique* dalam prediksi pemesanan tiket pesawat pada CV. Trend Pelangi Mandiri. Data yang digunakan 5711 (80% pelatihan, 20% Pengujian), data pemesanan tiket pesawat tahun 2019, 2020, dan 2021 dengan hasil yang didapatkan adalah Akurasi SMOTE (*Synthetic Minority*

Oversampling Technique) dengan K-NN lebih rendah yaitu sebesar 79,65%, K-NN tanpa menggunakan SMOTE yaitu sebesar 97,81% K-NN dengan menggunakan SMOTE presisi naik 18,00%, *recall* naik 28,00%, *F1-Score* naik 27,00%.

Pada penelitian-penelitian yang sudah dilakukan masih sedikit yang membahas mengenai analisa penjualan tiket pesawat terutama tujuan kejepang oleh karena itu Penulis melakukan penelitian dengan pengelompokan tiket yang banyak dan sedikit dipesan menggunakan *Classic Machine Learning* dengan algoritma *K-Nearest Neighbors (KNN)*, *Naïve Bayes* & *Decision Tree* untuk melakukan klasifikasi data sesuai dengan banyak atau sedikitnya permintaan tiket, sehingga dapat memprediksi ketersediaan tiket yang akan dipesan sesuai dengan data penjualan sebelumnya menggunakan metode klasifikasi seperti yang dijelaskan diatas. Dengan hasil yang didapat PT. TTD dapat melakukan persiapan jika ada pemesanan dalam jumlah besar sesuai dengan waktu pemesanan sehingga dapat mengantisipasi ketersediaan tiket pesawat atau membuat strategi penjualan melihat dari hasil analisa dimana bisa meningkatkan penjualan disaat penjualan sedang sedikit dan dapat memberikan promo-promo yang menarik, sehingga perusahaan akan mendapatkan pemasukan yang lebih banyak dari penjualan tiket pesawat. Selain untuk melakukan klasifikasi data penelitian ini juga dilakukan untuk mengukur nilai akurasi dengan menggunakan 3 algoritma berbeda untuk melakukan klasifikasi, Ketiga algoritma tersebut adalah *K-Nearest Neighbors (KNN)*, *Naïve Bayes* & *Decision Tree* sehingga akan didapat nilai akurasi mana yang lebih tinggi dalam melakukan klasifikasi data. Sehingga hasil yang didapatkan lebih akurat dan maksimal.

2. METODE

Jenis penelitian yang digunakan adalah dengan model eksperimen. Penelitian ini berfokus menggunakan metode klasifikasi dengan Algoritma *K-Nearest Neighbors (KNN)*, *Naïve Bayes* & *Decision Tree*. Pada PT TTD dengan data yang digunakan adalah data penjualan tiket pesawat internasional ke Jepang tahun 2022 & 2023 dengan data sebanyak 7670 data dan dibagi menjadi 80% (6136) sebagai *training* dan 20% (1534) sebagai *testing* untuk digunakan dalam menentukan klasifikasi *tren* permintaan tiket pesawat

NoInvoice	Airline	Month	Destination	Class	Season	Price	Request
IN0122010001	ALL NIPPON AIRWAYS	JAN	Haneda	Economy	Winter	Mid Price	Low
IN0122010002	JAPAN AIRLINES CO LTD	FEB	Narita	Economy	Winter	Mid Price	High
IN0122010003	JAPAN AIRLINES CO LTD	FEB	Narita	Economy	Winter	Low Price	High
IN0122010004	JAPAN AIRLINES CO LTD	FEB	Narita	Economy	Winter	Low Price	High
IN0122010005	JAPAN AIRLINES CO LTD	APR	Narita	Economy	Spring	Mid Price	High
IN0122010006	JAPAN AIRLINES CO LTD	MAR	Narita	Economy	Winter	Low Price	High
IN0122010007	JAPAN AIRLINES CO LTD	MAR	Narita	Economy	Winter	Low Price	High
IN0122010008	JAPAN AIRLINES CO LTD	MAR	Narita	Economy	Winter	Low Price	High
IN0122010009	JAPAN AIRLINES CO LTD	MAR	Narita	Economy	Winter	Low Price	High
IN0122010010	ALL NIPPON AIRWAYS	MAR	Haneda	Economy	Winter	Mid Price	High
IN0122010011	JAPAN AIRLINES CO LTD	MAR	Narita	Economy	Winter	Low Price	High
IN0122010012	JAPAN AIRLINES CO LTD	MAR	Narita	Economy	Winter	Low Price	High
IN0122010013	JAPAN AIRLINES CO LTD	MAR	Narita	Economy	Winter	Low Price	High
IN0122010014	JAPAN AIRLINES CO LTD	MAR	Narita	Economy	Winter	Low Price	High
IN0122010015	JAPAN AIRLINES CO LTD	MAR	Narita	Economy	Winter	Low Price	High

Gambar 1. Contoh Data Set

2.1 Pengklasifikasian Data

2.1.1 Kategori Season

Langkah selanjutnya adalah melakukan pengklasifikasian atribut *Season* kedalam kelompok untuk mempermudah pembentukan model, pengelompokan atribut season sebagai berikut :

Tabel 1. Kategori *Season*

Bulan	Kategori
Desember – Maret	Winter
April - Mei	Spring
Juni -Agustus	Summer
September – November	Autumn

2.1.2 Kategori *Price*

Kemudian melakukan pengklasifikasian nilai atribut *Price* kedalam kelompok untuk mempermudah pembentukan model, Pengelompokan nilai atribut adalah sebagai berikut:

Tabel 2. Kategori *Price*

Price	Kategori
0 sd Rp 15.000.000	Low Price
Rp. 15.000.000 sd Rp. 25.000.000	Mid Price
➤ Rp. 25.000.000	High Price

2.1.3 Kategori Permintaan Tiket

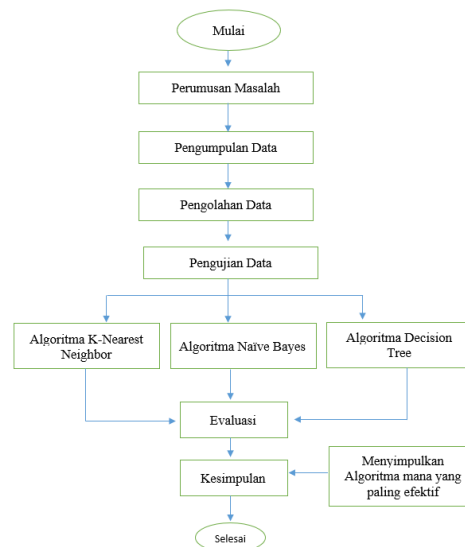
Kemudian melakukan pengklasifikasian nilai atribut *request* (permintaan tiket) kedalam kelompok menjadi seperti dibawah ini :

Tabel 3. Kategori Permintaan Tiket

Permintaan	Kategori
0 sd 100 tiket	Low
>100 tiket	High

2.2 Perancangan Penelitian

Rancangan penelitian dirumuskan dengan tujuan adanya arah yang jelas dan target yang hendak dicapai dalam penelitian. Jika tujuan penelitian jelas dan terumuskan dengan baik, maka penelitian dan pemecahan masalah akan berjalan dengan baik pula

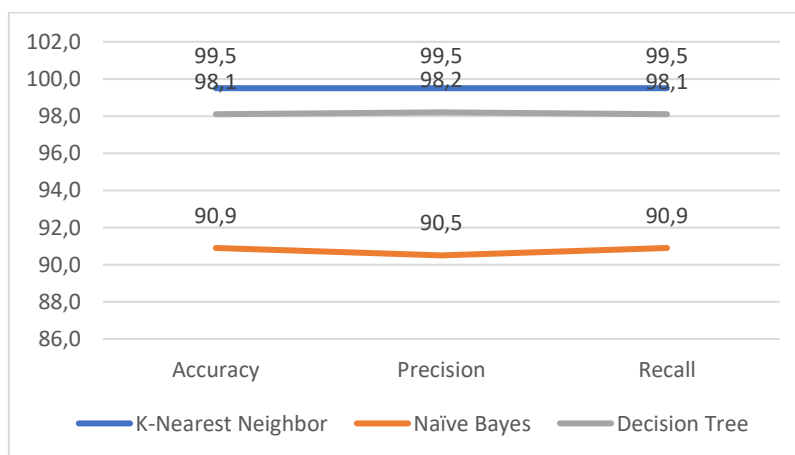


Gambar 2. Perancangan Penelitian

Dari **Gambar 1** seperti diatas secara umum dapat dideskripsikan perancangan penelitian bahwa pada tahap awal akan dilakukan pengumpulan data yang dibutuhkan dalam penelitian ini, kemudian akan dilakukan pengolahan data dengan menggunakan aplikasi pengolahan data agar lebih mudah diproses pada saat pengujian data. Setelah dilakukan pengolahan data kemudian data tersebut dilakukan pengujian data dengan menggunakan tiga algoritma yang telah dipilih diantaranya Algoritma *K-Nearest Neighbors (KNN)* , *Naïve Bayes & Decision Tree*. Pada tahap evaluasi, jika hasil mengalami ketidak sesuaian dengan yang diharapkan maka program dievaluasi Kembali dan dilakukan iterasi Kembali pada tahap pengujian, pengumpulan data, maupun pengolahan data berdasarkan hasil evaluasi mana yang perlu dikaji ulang. Dari implementasi tiga algoritma tersebut, akan didapatkan hasil kesimpulan berdasarkan hasil evaluasi untuk memperoleh akurasi yang lebih tinggi dari pengujian tersebut dan akan diimplementasikan pada aplikasi yang akan dibuat.

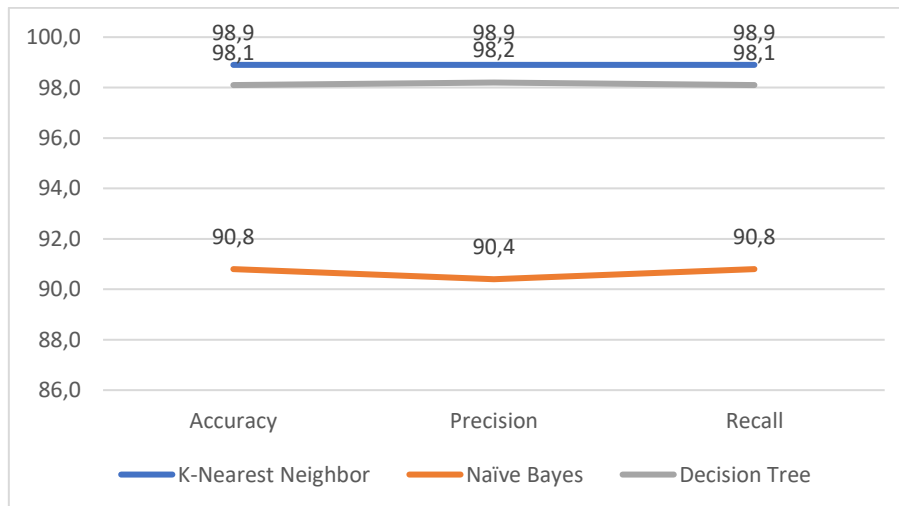
3. HASIL DAN PEMBAHASAN

Berdasarkan hasil penelitian yang dilakukan pada PT TTD dengan menggunakan metode klasifikasi dari jumlah data yang dilakukan pengujian sebanyak 80% (6136 dataset) untuk *training* dan 20% (1534 dataset) sebagai *testing*. Dilakukan evaluasi model dengan model *predictions* dan Model *testing & score*. Kemudian dilakukan evaluasi dengan menggunakan *confusion matrix* untuk kedua model tersebut. Kita akan mendapatkan perbandingan hasil nilai evaluasi yang dilakukan dengan menggunakan algoritma *K-Nearest Neighbor*, *Naïve Bayes* dan *Decesion Tree*. Berikut ini adalah grafik hasil dari evaluasi model yang sudah dilakukan :



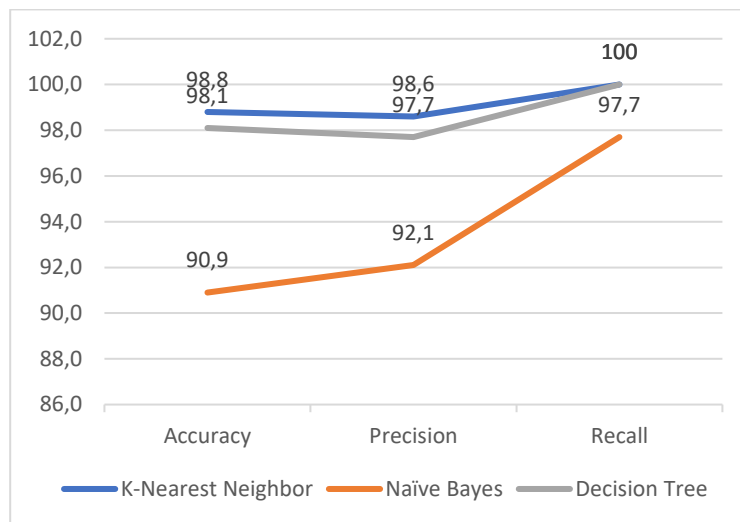
Gambar 2. Grafik Hasil Evaluasi Model *Predictions*

Berdasarkan grafik pada Gambar 3. dapat kita lihat bahwa hasil dari evaluasi menggunakan model *predictions* dengan menggunakan algoritma yang berbeda mendapatkan hasil yang berbeda juga, sehingga kita dapat melihat algoritma mana yang memiliki nilai yang lebih tinggi dibandingkan dengan algoritma lainnya. Berdasarkan Gambar 2. algoritma yang memiliki nilai tertinggi adalah *K-nearest Neighbor* dengan nilai akurasi tertinggi sebesar 99.3% dibandingkan dengan nilai akurasi menggunakan algoritma *Naïve bayes* sebesar 90.9% dan algoritma *Descision Tree* sebesar 98.1%. Sehingga jika menggunakan algoritma *K-nearest Neighbor* dalam melakukan klasifikasi dataset memiliki tingkat ke akuratan sebesar 99,3%.



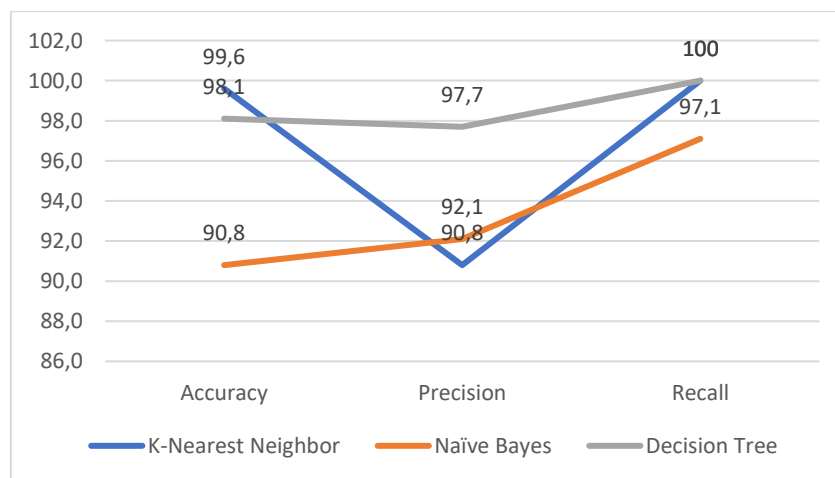
Gambar 4. Grafik Hasil Evaluasi Model *Test & Score*

Pada Gambar 4. kita mendapatkan hasil nilai tertinggi untuk *Accuracy*, *Precision* & *Recall* adalah menggunakan algoritma *K-Nearest Neighbor* sebesar *Accuracy* (98,9%), *Precision* (98,9%), *Recall* (98,9%). Sehingga memiliki keakuratan dalam melakukan klasifikasi dataset sebesar 98,9 %. Sama seperti hasil evaluasi dengan model *prediction* menyatakan bahwa algoritma *K-Nearest Neighbor* adalah yang memiliki nilai akurasi terbaik.



Gambar 5. Grafik Hasil Evaluasi *Confusion Matrix* dengan Model *Predictions*

Pada Gambar 5. kita dapat melihat hasil dari evaluasi *confusion matrix* dengan *model predictions* mendapatkan nilai tertinggi adalah menggunakan algoritma *K-Nearest Neighbor* dengan nilai *Accuracy* sebesar 98,8% .



Gambar 6. Grafik Hasil Evaluasi *Confusion Matrix* dengan Model *Test & Score*

Pada Gambar 6. kita dapat melihat hasil dari evaluasi *confusion matrix* dengan *model predictions* mendapatkan nilai tertinggi adalah menggunakan algoritma *K-Nearest Neighbor* dengan nilai *Accuracy* sebesar 98,8% . Dari percobaan yang sudah dilakukan dengan menggunakan 2 model yaitu *Model Predictions* dan *Model Test & Score* dan dengan *Confusion Matrix* untuk mencari nilai *Accuracy*, *Precision* & *Recall*. Berdasarkan

semua percobaan yang sudah dilakukan diatas algoritma yang paling tinggi nilainya adalah *K-Nearest Neighbor* karena mendapatkan nilai tertinggi dibandingkan dengan algoritman lainnya. Hasil evaluasi dengan model *prediction* dan model *test & score* , Algoritma *K-Nearest Neighbor* mendapatkan nilai tertinggi dibandingkan dengan algoritma *Naïve Bayes* dan *Decision Tree* dengan nilai *accuracy* sebesar 99,5% (*evaluasi model predictions*) & 98,9% (*evaluasi model test & score*). Hasil evaluasi *Confusion Matrix* dengan model *prediction* dan model *test & score*, Algoritma *K-Nearest Neighbor* mendapatkan nilai tertinggi, dengan nilai *accuracy* sebesar 98,8%, *precision* 98,6%, *recall* 100% (*model predictions*) & *accuracy* 99,6%, *precision* 99,6 % , *recall* 100% (*model test & score*). Dengan ini algoritma *K-Nearest Neighbor* dapat digunakan oleh PT. TTD dalam melakukan klasifikasi terhadap permintaan tiket apakah termasuk dalam kategori *High* atau *Low* , Sehingga dapat mengantisipasi ketersediaan tiket yang akan dipesan.

Tabel 4. Rekap Hasil Evaluasi Mode

Hasil Evaluasi Model Predictions				
No	Algoritma	Accuracy	Precision	Recall
1	<i>K-Nearest Neighbor</i>	99,5%	99,5%	99,5%
2	<i>Naïve Bayes</i>	90,9%	90,5%	90,9%
3	<i>Decision Tree</i>	98,1%	98,2%	98,1%
Hasil Evaluasi Model Test & Score				
1	<i>K-Nearest Neighbor</i>	98,9%	98,9%	98,9%
2	<i>Naïve Bayes</i>	90,8%	90,4%	90,8%
3	<i>Decision Tree</i>	98,1%	98,2%	98,1%
Hasil Evaluasi Model Predictions dengan Confusion Matrix				
1	<i>K-Nearest Neighbor</i>	98,8%	98,6%	100%
2	<i>Naïve Bayes</i>	90,9%	92,1%	97,2%
3	<i>Decision Tree</i>	98,1%	97,7%	100%

Hasil Evaluasi Model <i>Test & Score</i> dengan <i>Confusion Matrix</i>				
1	<i>K-Nearest Neighbor</i>	99,6%	99,6%	100%
2	<i>Naïve Bayes</i>	90,8%	92,1%	97,0%
3	<i>Decision Tree</i>	98,1%	97,7%	100%

4. KESIMPULAN

Dari hasil penelitian untuk melakukan klasifikasi data penjualan tiket tahun 2022 & 2023 dengan menggunakan algoritma *K-Nearest Neighbor*, *Naïve Bayes*, *Decision Tree* (C4.5) Pada PT.TTD penulis simpulkan sebagai berikut : Algoritma *K-Nearest Neighbor* (KNN) adalah pilihan terbaik untuk dataset ini berdasarkan akurasi yang diperoleh pada kedua tahap evaluasi. Dengan nilai akurasi 99,5% pada model *prediction* dan 98,9% pada model *test & score*, KNN secara signifikan lebih unggul dibandingkan dengan *Naïve Bayes* dan *Decision Tree* dalam hal prediksi dan generalisasi data. Dari data tersebut, dapat disimpulkan bahwa algoritma *K-Nearest Neighbor* (KNN) secara konsisten memberikan performa yang lebih baik dibandingkan dengan algoritma *Naïve Bayes* dan *Decision Tree* dalam kasus ini. Akurasi yang tinggi pada kedua jenis evaluasi (model *prediction* dan model *test & score*) menunjukkan bahwa KNN efektif dalam melakukan prediksi dan generalisasi terhadap data yang diberikan. Hasil evaluasi *Confusion Matrix* dengan model *prediction* dan model *test & score*, Model *prediction* memiliki akurasi sebesar 98,8%, sedangkan model *test & score* memiliki akurasi sebesar 99,6%. Akurasi mengukur seberapa banyak prediksi yang benar dari total prediksi yang dilakukan oleh model. Model *prediction* memiliki presisi sebesar 98,6%, dan model *test & score* memiliki presisi sebesar 99,6%. Presisi mengukur seberapa tepat model dalam memprediksi positif. Kedua model, baik model *prediction* maupun model *test & score*, memiliki *recall* sebesar 100%. *Recall* mengukur seberapa baik model dalam menemukan semua kasus positif.

5. DAFTAR PUSTAKA

- [1] A'yuniyah, Q., Elvira, W., Nazira, N., Ambarani, I., Intan, S. F., & Ramadhani, D. (2023). Analisa Algoritma Naïve Bayes Classifier (NBC) Untuk Prediksi Penjualan Alat Kesehatan. *IJRSE: Indonesian Journal of Informatic Research and Software*

- Engineering Journal*, 3(2), 119–126.
- [2] Fadillah, A. P. (2015). Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ). *Jurnal Teknik Informatika Dan Sistem Informasi*, 1(3), 260–270. <https://doi.org/10.28932/jutisi.v1i3.406>
- [3] Heliyanti Susana. (2022). Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet. *Jurnal Riset Sistem Informasi Dan Teknologi Informasi (JURSISTEKNI)*, 4(1), 1–8. <https://doi.org/10.52005/jursistekni.v4i1.96>
- [4] Irmayani, W. (2021). Visualisasi Data Pada Data Mining Menggunakan Metode Klasifikasi. *Jurnal Khatulistiwa Informatika*, IX(I), 68–72.
- [5] Juwita, Safii, M., & Efendi Damanik, B. (2022). Algoritma Naive Bayes Untuk Memprediksi Penjualan Pada Toko VJCakes Pematang Siantar. *Journal of Machine Learning and Artificial Intelligence*, 1(4), 337–346. <https://doi.org/10.55123/jomlai.v1i4.1674>
- [6] Leonardi, M., Emilda, R., Katrin, I., & Yulianto, A. (2021). Prediksi Penjualan Produk Rokok Pada PT. Indomarco Prismatama Menggunakan Algoritma C4.5. *Paradigma - Jurnal Komputer Dan Informatika*, 23(2), 182–190. <https://doi.org/10.31294/p.v23i2.11151>
- [7] Nopianti, Y., Ruliansyah, R., & Fadilah, E. (2017). Sistem Informasi Pemesanan Tiket Travel Berbasis Web (Studi Kasus: Lantra Wisata Travel PO. Sejahtera). *Jusifo (Jurnal Sistem Informasi)*, 3(2), 143–150. <https://doi.org/10.19109/jusifo.v3i2.3869>
- [8] Permadi, M., Febriani, F., & Panggabean, M. J. (2023). *Application of the K-Nearest Neighbor (KNN) Algorithm for the Classification of Indihome Products Sales in the Riau Region Penerapan Algoritma K-Nearest Neighbor (KNN) untuk Klasifikasi Penjualan Produk Indihome Daerah Riau*. 105–111.
- [9] Putra, V., Tri Pranoto, G., & Eko Putra, F. (2023). Klasifikasi Kebutuhan Sparepart Dengan Algoritma K-Nearest Neighbor Untuk Meningkatkan Penjualan Sparepart. *Bulletin of Information Technology (BIT)*, 4(2), 287–293.
- [10] Rismala, R., Ali, I., & Rizki Rinaldi, A. (2023). Penerapan Metode K-Nearest Neighbor Untuk Prediksi Penjualan Sepeda Motor Terlaris. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 585–590. <https://doi.org/10.36040/jati.v7i1.6419>

- [11] Robianto, Sitorus, S. H., & Ristian, U. (2021). Penerapan Metode Decision Tree Untuk Mengklasifikasikan Mutu Buah Jeruk Berdasarkan Fitur Warna Dan Ukuran. *Jurnal Komputer Dan Aplikasi*, 09(01), 76–86.
- [12] Suci, W., & Samsudin, S. (2022). Algoritma K-Nearest Neighbors dan Synthetic Minority Oversampling Technique dalam Prediksi Pemesanan Tiket Pesawat. *Jurnal Media Informatika Budidarma*, 6(3), 1775. <https://doi.org/10.30865/mib.v6i3.4374>
- [13] Suhanda, Y., Kurniati, I., & Norma, S. (2020). Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik. *Jurnal Teknologi Informatika Dan Komputer*, 6(2), 12–20. <https://doi.org/10.37012/jtik.v6i2.299>
- [14] Suliztia, M. L., & Fauzan, A. (2019). Comparing Naive Bayes, K-Nearest Neighbor, and Neural Network Classification Methods of Seat Load Factor in Lombok Outbound Flights. *Jurnal Matematika, Statistika Dan Komputasi*, 16(2), 187. <https://doi.org/10.20956/jmsk.v16i2.7864>
- [15] Suryadewiansyah, M. K., & Tju, T. E. E. (2022). Naïve Bayes dan Confusion Matrix untuk Efisiensi Analisa Intrusion Detection System Alert. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 8(2), 81–88. <https://doi.org/10.25077/teknosi.v8i2.2022.81-88>
- [16] Wardhana, A. W., Patimah, E., Shafarindu, A. I., Siahaan, Y. M., Haekal, B. V., & Prasvita, D. S. (2021). Klasifikasi Data Penjualan pada Supermarket dengan Metode Decision Tree. *Senamika*, 2(1), 660–667. <https://conference.upnvj.ac.id/index.php/senamika/article/view/1389>
- [17] Winantu, A., & Khatimah, C. (2023). Perbandingan Metode Klasifikasi Naive Bayes Dan K-Nearest Neighbor Dalam Memprediksi Prestasi Siswa. *INTEK : Jurnal Informatika Dan Teknologi Informasi*, 6(1), 58–64. <https://doi.org/10.37729/intek.v6i1.3006>