

Analisis Algoritma *K-Medoids Clustering* Dalam Menentukan Pemesanan Hotel

Wawan Kurniawan¹; Achmad Rifai²; Windu Gata³; Deni Gunawan⁴

STMIK Nusa Mandiri/Program Studi Ilmu Komputer¹
14002315@nusamandiri.ac.id^[1]

STMIK Nusa Mandiri/Program Studi Teknik Informatika²
achmad.acf@nusamandiri.ac.id^[2]

STMIK Nusa Mandiri/Program Studi Ilmu Komputer³
windu@nusamandiri.ac.id^[3]

Universitas Bina Sarana Informatika/Program Studi Sistem Informasi⁴
deni.dee@bsi.ac.id^[4]

Abstract

At this time the hotel business is developing so rapidly, this is due to the emergence of many small, medium to star-rated hotels, so that the competition in the hotel world is getting tighter owned, then data owned by a hotel is one of the assets of a hotel. Therefore, hotels need to classify hotel bookings as a consideration in making decisions for marketing strategies to achieve the intended target. This study uses the K-Medoids method so that hotel booking patterns can be identified for hotel guests. K-Medoids is a partial clustering method that aims to find the K-cluser set among the data that best characterizes objects in a data set. Clustering or clustering is one of the tools in data mining that aims to group objects into clusters. From the results of research that has been carried out produced 3 clusters with the first cluster consisting of 2.343 transaction records, the second cluster consisting of 1833 transaction records, and the third cluster consisting of 4.530 transaction records. The highest number of hotel bookings is in the 3rd cluster with 4.530 transaction records, which can be used as a marketing barometer, determining customer targets and hotel promotion strategies in the future.

Keyword: Hotels Booking, Clustering, K-Medoids.

Abstract

Pada saat ini bisnis perhotelan berkembang begitu pesat, hal ini disebabkan karena banyak bermunculan hotel-hotel kecil, menengah sampai hotel berbintang dari satu sampai bintang lima sehingga persaingan dunia hotel semakin ketat. Pesatnya pertumbuhan hotel mendorong persaingan yang ketat sehingga banyak hotel melakukan inovasi baik dalam produk, layanan dan fasilitas yang dimiliki, kemudian data yang dimiliki suatu hotel merupakan salah satu aset dari suatu hotel tersebut. Oleh karena itu, hotel perlu mengelompokkan pemesanan hotel sebagai pertimbangan dalam mengambil keputusan untuk strategi pemasaran untuk mencapai target yang dituju. Penelitian ini menggunakan metode *K-Medoids* agar dapat diketahui pola pemesanan hotel bagi para tamu hotel. *K-Medoids* merupakan metode partisional *clustering* yang bertujuan untuk menemukan set k-cluster di antara data yang paling mencirikan objek dalam kumpulan data. *Clustering* atau klusterisasi adalah salah satu alat bantu pada data mining yang bertujuan mengelompokkan objek-objek ke dalam cluster-cluster. Dari hasil penelitian yang telah dilakukan menghasilkan 3 cluster dengan cluster pertama terdiri dari 2.343 record transaksi, cluster ke-2 terdiri dari 1833 record transaksi, dan cluster ke-3 terdiri dari 4.530 record transaksi. Jumlah pemesanan hotel terbanyak terdapat pada cluster ke-3 dengan 4.530 record transaksi, yang bisa dijadikan barometer pemasaran, penentuan target tamu dan strategi-strategi promosi pihak hotel di masa mendatang.

Kata Kunci: Pemesanan Hotel, *Clustering*, K-Medoids.

1. Pendahuluan

Pada saat ini bisnis perhotelan berkembang begitu pesat, hal ini disebabkan karena banyak bermunculan hotel-hotel kecil, menengah sampai hotel berbintang sehingga persaingan dunia hotel semakin ketat (Sihotang et al., 2018). Industri perhotelan memiliki karakteristik lain dari industri yang biasa kita kenal. dimana konsumen membeli jasa ini dalam jangka pendek, dipengaruhi oleh fisik atau produk hotel, strategi harga, promosi komunikasi dengan calon dan langganan, dan menetapkan lokasi dan saluran distribusi yang efektif sehingga keputusan konsumen untuk berkunjung bisa terus dilakukan (Sihotang et al., 2018). Pesatnya pertumbuhan hotel mendorong persaingan yang ketat sehingga banyak hotel melakukan inovasi baik dalam produk, layanan dan fasilitas yang dimiliki. Pada dasarnya hampir semua hotel memiliki fasilitas yang sama, yang membedakan salah satunya adalah kualitas layanan yang dimilikinya (Sihotang et al., 2018).

Data yang dimiliki suatu hotel merupakan salah satu aset dari suatu hotel tersebut. Dengan adanya kegiatan operasional sehari-hari akan semakin memperbanyak jumlah data yang dihasilkan. Atribut pertama diambil dari jenis dalam mendapatkan pola informasi, bisa kamar yang paling banyak dipesan. Hal ini dipertimbangkan karena *customer* yang menginap di hotel mempertimbangkan harga dari kamar. Atribut kedua adalah dari usia *customer* yang menginap, semakin muda usia *customer*, maka akan semakin besar kemungkinan memilih kamar standar. Kemudian, atribut ketiga diambil dari negara. Hal ini dikarenakan, gaya hidup masyarakat suatu negara pasti berbeda antara negara satu dengan negara yang lain. Seperti negara maju dan negara berkembang

Banyak orang yang memeriksa pendapat dari pembeli lain sebelum membeli produk untuk membuat pilihan yang tepat (Indrayuni et al., 2015). Kemudahan dalam penggunaan teknologi informasi khususnya internet membuat para pengguna internet dengan mudah mendapatkan informasi yang diinginkan (Taufik, 2018).

Terdapat beberapa penelitian yang pernah dilakukan mengenai metode *K-Medoids* diantaranya adalah penelitian yang

dilakukan oleh Nurliana pulungan Penerapan Algoritma *K-Medoids* untuk Mengelompokkan penduduk 15 Tahun keatas Menurut lapangan Pekerjaan Utama dengan hasil cluster rendah dan tinggi. Cluster rendah. Penelitian yang dilakukan Dyang Falila Pramesti (2017 Implementasi Metode *K-Medoids Clustering* Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas dengan hasil cluster dengan potensi tinggi dan kluster dengan potensi sedang. Penelitian yang dilakukan Wiwit Agus Triyanto (2015, Algoritma *K-Medoids* untuk Penentuan Strategi Pemasaran Produk dengan hasil 5 cluster, dimana cluster tertinggi terdapat pada cluster ke-lima yang memiliki kombinasi jumlah barang yang dibeli paling tinggi.

Berdasarkan paparan penelitian sebelumnya, penulis mencoba melakukan penelitian yang berjudul "Analisa pemesanan hotel dengan menggunakan metode *K-Medoids*" sebagai solusi agar memudahkan informasi customer yang berguna bagi hotel dalam mempersiapkan strategi-strategi kedepan

Dari alasan diatas, maka penelitian ini bertujuan untuk mendapatkan pola informasi customer, yang berguna bagi hotel dalam mempersiapkan strategi-strategi di masa depan, dengan data yang diambil adalah dari transaksi customer yang menginap. Pada penelitian ini, menggunakan teknik data mining dengan metode *K-medoids* untuk mengetahui pola transaksi hotel dari masing-masing negara. Data mining merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Sutinah et al., 2019).

Teknologi data mining sangat berguna untuk menemukan informasi yang sangat penting. Data mining dapat digunakan untuk mengelompokkan, memprediksi, mengestimasi, dan menentukan kaidah asosiasi dalam suatu data yang ada (Apriyana et al., 2016). Metode *K-medoid* digunakan untuk mengkluster tipe kamar yang dipesan pada suatu negara sesuai dengan segmen customer. Melihat perkembangan teknologi yang dipergunakan di setiap hotel untuk peningkatan kualitas pelayanan semakin ketat. Agar dapat bersaing dengan para

pebisnis jasa akomodasi atau pelayanan lainnya di bidang yang sama, sehingga hal ini pun harus mewajibkan para pengusaha hotel khususnya harus bisa memaksimalkan pemanfaat data, serta perlu lebih menganalisa data tersebut agar tidak hanya disimpan.

2. Metode Penelitian

Berdasarkan pendahuluan yang telah dijabarkan, penelitian ini melalui beberapa tahap, berikut tahapan kerangka kerja dalam penyusunan penelitian:



Gambar 1. Metode Penelitian

A. Identifikasi Masalah

Dalam penelitian ini, Langkah awal yang dilakukan adalah mengidentifikasi masalah yang bertujuan untuk mengetahui berapa banyak jumlah pengelompokan pada masalah kasus pemesanan hotel menggunakan metode *K-Medoids Clustering*. Dan mencari akar masal dengan table data yang kita punya dengan table yang kita miliki tersebut data yang banyak dapunya potensi besar sebagai bahan strategi pemasaran di masa mendatang dan belum diolah dengan baik menjadi pokok permasalahan ini.

Dalam penelitian ini, identifikasi masalah bertujuan untuk mengetahui seberapa banyak pengelompokan

pemesanan hotel menggunakan metode *K-Medoids Clustering*.

B. Pengumpulan Data

Data yang digunakan oleh penulis adalah kumpulan data pemesan hotel di beberapa negara, Sumber data di peroleh dari repository dari situs <https://www.kaggle.com>

C. Pengolahan Data Awal

Pengolahan Data dengan melakukan pemilihan table terlebih dahulu, lalu dilanjut menentukan Atribut untuk melakukan clustering

D. Penerapan Algoritma K-Medoid

Penerapan atau implementasi algoritma *K-Medoids* Pada penelitian kali dengan menentukan cluster (K). dimana dari data yang ada menjadi 3 Cluster(k). Penentuan k = 3 ini dilakukan agar lebih mudah dalam mendapatkan hasil

3. Hasil Dan Pembahasan

3.1 Hasil

Metode K-medoids tidak menentukan nilai rata-rata dari objek yang ada dalam cluster sebagai titik acuan, namun menggunakan medoids (median), yang merupakan objek yang paling terletak dipusat sebuah cluster (Keatas et al., 2019). Algoritma K-medoids atau sering dikenal pula dengan PAM (*Partitioning Around Medoids*) menggunakan metode partisi *clustering* untuk mengelompokkan sekumpulan n objek menjadi sejumlah k cluster (Mustofa & Suasana, 2018).

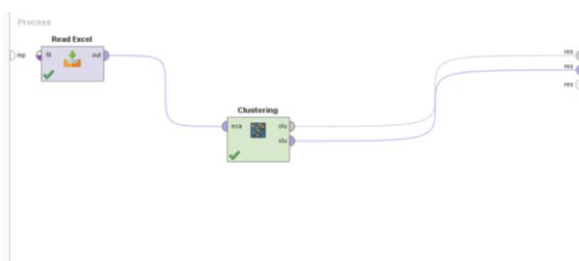
Jumlah Data Hotel adalah 8.705 transaksi *booking*, yang terdiri dari beberapa negara, segmen pasar, *customer* dan tipe kamar.

1	country	market_segment	reserved_room_type	assigned_room_type	booking_changes	deposit_type
2	PRT	Direct	C	C		3 No Deposit
3	PRT	Direct	C	C		4 No Deposit
4	GBR	Direct	A	C		0 No Deposit
5	GBR	Corporate	A	A		0 No Deposit
6	GBR	Online TA	A	A		0 No Deposit
7	GBR	Online TA	A	A		0 No Deposit
8	PRT	Direct	C	C		0 No Deposit
9	PRT	Direct	C	C		0 No Deposit
10	PRT	Online TA	A	A		0 No Deposit
11	PRT	Offline TA/TO	D	D		0 No Deposit
12	PRT	Online TA	E	E		0 No Deposit
13	PRT	Online TA	D	D		0 No Deposit
14	USA	Online TA	D	E		0 No Deposit
15	ESP	Online TA	G	G		1 No Deposit
16	PRT	Online TA	E	E		0 No Deposit
17	IRL	Online TA	D	E		0 No Deposit
18	PRT	Offline TA/TO	E	E		0 No Deposit
19	IRL	Online TA	A	E		0 No Deposit
20	FRA	Corporate	A	G		0 No Deposit
21	GBR	Direct	G	G		0 No Deposit
22	GBR	Online TA	F	F		0 No Deposit

22	GBR	Online TA	F	F	0	No Deposit
23	PRT	Direct	A	A	1	No Deposit
24	PRT	Direct	A	A	1	No Deposit
25	PRT	Direct	D	D	1	No Deposit
26	GBR	Offline TA/TO	D	I	0	No Deposit
27	PRT	Offline TA/TO	D	D	0	No Deposit
28	IRL	Offline TA/TO	D	D	0	No Deposit
29	PRT	Online TA	E	E	0	No Deposit
30	PRT	Online TA	A	A	0	No Deposit
31	PRT	Online TA	D	D	0	No Deposit
32	NULL	Direct	A	A	2	No Deposit
33	GBR	Offline TA/TO	D	D	0	No Deposit
34	PRT	Direct	E	E	0	No Deposit
35	IRL	Offline TA/TO	A	C	0	No Deposit
36	PRT	Online TA	D	D	0	No Deposit
37	PRT	Online TA	D	D	0	No Deposit
38	ESP	Online TA	A	C	0	No Deposit
39	PRT	Online TA	D	D	0	No Deposit
40	PRT	Online TA	D	D	0	No Deposit
41	ROU	Direct	E	E	0	No Deposit
42	PRT	Online TA	G	G	0	No Deposit
43	IRL	Offline TA/TO	D	D	0	No Deposit
44	ESP	Direct	F	F	0	No Deposit
45	ROU	Direct	E	E	0	No Deposit
46	PRT	Online TA	A	A	0	No Deposit
47	PRT	Online TA	G	G	0	No Deposit
48	ESP	Offline TA/TO	A	A	0	No Deposit

Gambar 3. Data Hotel

Dari data yang ada kita *Clustering* Data dengan *K-Medoids*. Dalam *K-Medoids* kita tentukan menjadi 3 Cluster(k). Penentuan k = 3 ini dilakukan agar lebih mudah dalam mendapatkan hasilnya terkait keterbatasan support kemampuan PC. Hasil dari proses *clustering* tersebut akan diekspor dalam bentuk file terpisah sesuai hasil *clustering*. Proses *clustering* data ditunjukkan pada gambar di bawah



Gambar 4. Proses Clustering Data

Setelah *clustering* pada *Rapid miner* selesai didapat 3 cluster, yakni cluster 0, Cluste 1 dan cluster 2. Kemudian disorting atau disurutkan dan dipisahkan menjadi 3 file yang kemudian dilakukan proses asosiasi menggunakan *Frequent Pattern-Growth (FP-Growth)*. Hasil yg diperoleh dari clustering menggunakan *K-Medoids* adalah : Cluster 0: 2342 items, Cluster 1: 1833 items, Cluster 2: 4530 items

3.2 Pembahasan

a. fData Set

Pada penelitian ini, dataset yang digunakan oleh penulis adalah kumpulan data pemesan hotel di beberapa negara, Sumber data di peroleh dari repository dari situs <https://www.kaggle.com>.

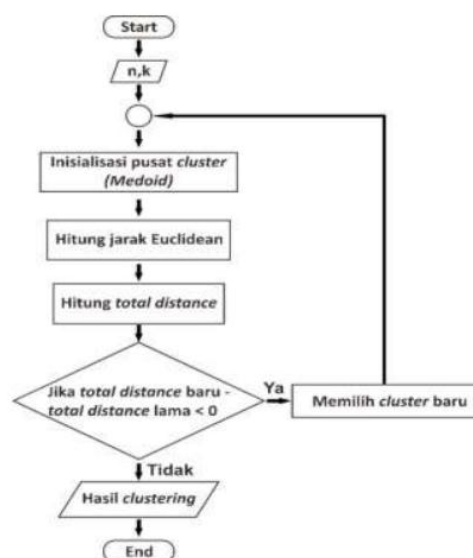
Atribut yang diambil untuk model *K-Medoids* sebanyak 3 atribut yaitu : tipe kamar, negara dan segmen pasar. Melihat dari kebiasaan pemesanan tiap negara

b. Clustering

Clustering atau klusterisasi adalah salah satu alat bantu pada data mining yang bertujuan mengelompokkan objek-objek ke dalam cluster-cluster (Silitonga et al., 2019). Prinsip dasar dari *clustering* adalah memaksimalkan kesamaan antar anggota satu cluster dan meminimalisir kesamaan antar anggota cluster yang berbeda (Simamora et al., 2017).

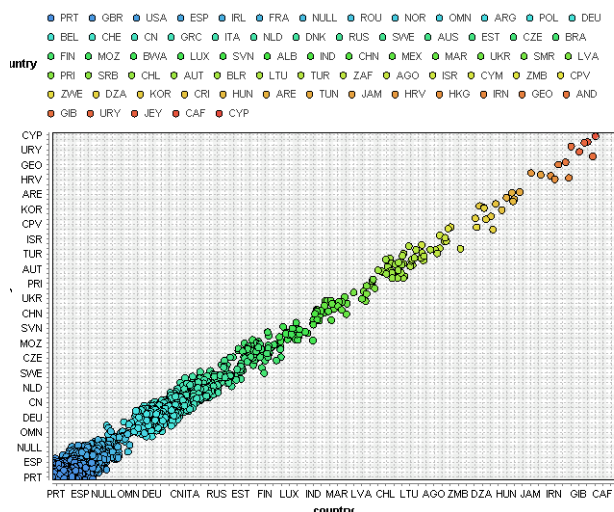
c. Algoritma K-Medoids

K-medoids tidak menentukan nilai rata-rata dari objek dalam cluster sebagai titik acuan, tapi menggunakan medoids (median), yang merupakan objek yang paling terletak dipusat sebuah cluster(Keatas et al., 2019). Algoritma *K-medoids* atau dikenal pula dengan *PAM (Partitioning Around Medoids)* menggunakan metode partisi *clustering* untuk mengelompokkan sekumpulan n objek menjadi sejumlah k cluster(Mustofa & Suasana, 2018).

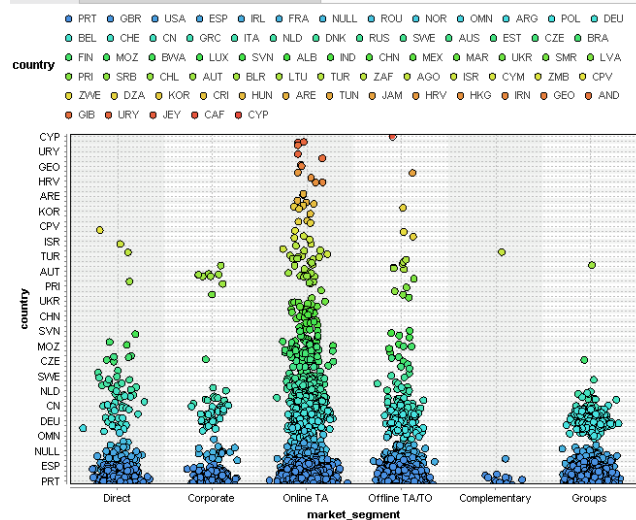
Gambar 5. Diagram Alur Proses *K-Medoids*

Pengelompokan dengan metode *K-Medoids*, lalu analisa *cluter* dengan menggunakan *cluter* model.

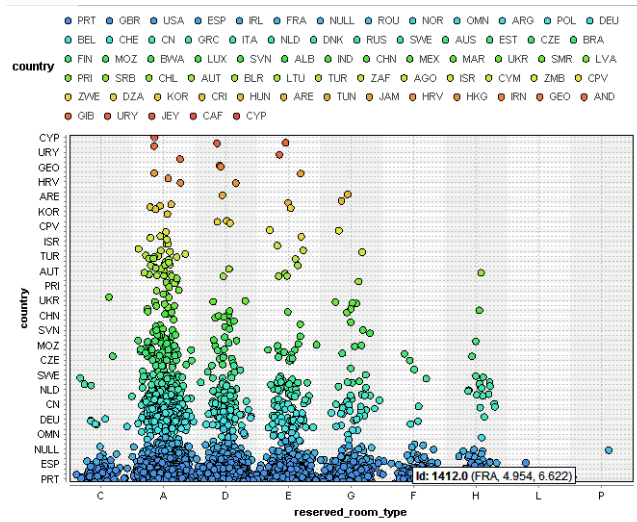
Data yang telah diimport akan dilakukan proses *clustering* dengan membagi data tersebut menjadi 3 cluster (k). Berikut hasil cluster dalam bentuk *scatter slot* :



Gambar 6. Cluster pada atribut Country



Gambar 7. Cluster pada atribut market segment



Gambar 8. Cluster pada atribut reserved room type

Selanjutnya agar dapat lebih dipahami dari dari setiap atribut dan *cluster*, berikut hasil *clustering* data hotel

ExampleSet (8705 examples, 2 special attributes, 3 regular attributes) Filter (8,705 / 8,705 examples):

Row No.	id	cluster	country	market_seg...	reserved_ro...
1	1	cluster_2	PRT	Direct	C
2	2	cluster_2	PRT	Direct	C
3	3	cluster_0	GBR	Direct	A
4	4	cluster_0	GBR	Corporate	A
5	5	cluster_0	GBR	Online TA	A
6	6	cluster_0	GBR	Online TA	A
7	7	cluster_2	PRT	Direct	C
8	8	cluster_2	PRT	Direct	C
9	9	cluster_2	PRT	Online TA	A
10	10	cluster_0	PRT	Offline TA/TO	D
11	11	cluster_1	PRT	Online TA	E
12	12	cluster_0	PRT	Online TA	D
13	13	cluster_0	USA	Online TA	D
14	14	cluster_1	ESP	Online TA	G
15	15	cluster_1	PRT	Online TA	E
16	16	cluster_1	IRL	Online TA	D

4. Kesimpulan

Dari hasil penelitian yang sudah dilakukan bisa disimpulkan sebagai berikut:

1. Algoritma K-Medoids dapat diimplementasikan untuk Menentukan Pemesanan Hotel sebagai strategi pemasaran dengan tahapan-tahapan perhitungan K-Medoids *Clustering*.
2. Analisis penentuan pemesanan hotel diperoleh hasil dimana pemesanan tinggi ditunjukkan pada cluster 2 dengan 4530 items/pemesanan.
3. terbanyak.Strategi pemasaran dapat dilakukan dengan melakukan promosi pada cluster ke-2 yang memiliki item/pemesanan yang paling tinggi.

Referensi

- Apriyana, K. F., Karang, I. G., Putra, K., & Indrawan, G. (2016). *Teknik Data Mining untuk Mendapatkan Pola Transaksi Hotel Bendesa dengan Algoritma J48. Senapati*.
- Indayuni, E., Wahyudi, M., Informasi, S., Selatan, J., Komputer, I., & Selatan, J. (2015). *PENERAPAN CHARACTER N-GRAM UNTUK SENTIMENT ANALYSIS REVIEW HOTEL MENGGUNAKAN ALGORITMA NAIVE BAYES*. 83–88.
- Keatas, T., Lapangan, M., Utama, P., Pulungan, N., & Suhendro, D. (2019). *PENERAPAN ALGORITMA K-MEDOIDS UNTUK MENGELOMPOKKAN PENDUDUK*. 3, 329–334. <https://doi.org/10.30865/komik.v3i1.1609>
- Mustofa, Z., & Suasana, I. S. (2018). *ALGORITMA CLUSTERING K-MEDOIDS PADA E-GOVERNMENT BIDANG INFORMATION AND COMMUNICATION*. 9, 1–10.
- Sihotang, B. K., Wanto, A., Informatika, T., Studi, P., & Informatika, T. (2018). *Analisis Jaringan Syaraf Tiruan Dalam Memprediksi Jumlah Tamu Pada Hotel Non Bintang*. 17(4), 333–346.
- Silitonga, D. A., Windarto, A. P., & Hartama, D. (2019). *Penerapan Metode K-Medoid pada Pengelompokan Rumah Tangga Dalam Perlakuan Memilah Sampah Menurut Provinsi*. 313–318.
- Simamora, D. A. S., Furqon, M. T., & Priyambadha, B. (2017). *Clustering Data Kejadian Tsunami Yang Disebabkan Oleh Gempa Bumi Dengan Menggunakan Algoritma K-Medoids*. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(8), 635–640.
- Sutinah, E., Agustina, N., & Asmoro, R. A. (2019). *Data Mining Untuk Klasifikasi Tamu Hotel Dengan Algoritma Apriori*. 7(1), 69–78.
- Taufik, A. (2018). *Komparasi Algoritma Text Mining Untuk Klasifikasi Review Hotel*. IV(2). <https://doi.org/10.31294/jtk.v4i2.3461>