

Model Klasifikasi Analisis Kepuasan Pengguna Perpustakaan Online Menggunakan K-Means dan Decision Tree

Khamim Surya Hadi Kusuma Al Atros¹, Abdul Robi Padri², Odi Nurdiawan^{3*}, Ahmad Faqih², Saeful Anwar²

¹ Program Studi Sistem Informasi, STMIK IKMI, Cirebon, Indonesia

² Program Studi Teknik Informatika, STMIK IKMI, Cirebon, Indonesia

³ Program Studi Manajemen Informatika, STMIK IKMI, Cirebon, Indonesia

Email : ¹khamimsuryahadi9@gmail.com ²abdulrobipadri06@gmail.com, ³odinurdiawan2020@gmail.com, ⁴ahmadfaqih367@gmail.com, ⁵saefulanwar419@gmail.com

Email Penulis Koresponden : odinurdiawan2020@gmail.com

Submitted 28-10-2021; Accepted 11-12-2021; Published 30-12-2021

Abstrak

Perpustakaan SD Riyadlul Muta'allimin adalah salah satu unit pendukung dalam kegiatan akademis sekolah. Masa pandemi seperti ini para siswa sangat sulit dalam mengakses perpustakaan disekolah. Oleh sebab itu pihak sekolah menerapkan Perpustakaan Online agar para siswa maupun guru dapat mengakses perpustakaan tidak perlu berbondong-bondong ke sekolah. Perpustakaan online yang sudah diterapkan perlu dilakukannya evaluasi kepuasan pengguna, karena ini adalah hal yang baru bagi SD Riyadlul Muta'allimin. Analisis kepuasan dapat dilakukan dengan metode K-Means dengan decision tree. K-Means digunakan untuk mengelompokkan data yang nantinya hasil kluster akan dirubah menjadi label. Selanjutnya diklasifikasikan menggunakan decision tree. Berdasarkan analisis tersebut, prediksi kluster 1 dengan true kluster 1 sebesar 48 item, true kluster 0 dan true kluster 0 tidak terdapat item, maka class precisionnya 100%.

Kata Kunci: Perpustakaan; K-Means; Decision Tree; Klusterisasi; Klasifikasi

Abstract

Riyadlul Muta'allimin Elementary School Library is one of the support units in the school's academic activities. During times like these, students are very difficult in accessing the library in school. Therefore, the school applies an Online Library, so those students and teachers who can access the library do not need to flock to school. Online libraries that have been implemented need to do an evaluation of user satisfaction because this is a new thing for Riyadlul Muta'allimin Elementary School. Satisfaction analysis can be done by the K-Means method with a decision tree. K-Means is used to group data that will be converted into labels. It is further classified using a decision tree. Based on the analysis, the prediction of cluster 1 with true cluster 1 of 48 items, true cluster 0 and true cluster 0 there are no items, then the class precision is 100%.

Keywords: Library; K-Means; Decision Tree; Clustering; Classification

1. PENDAHULUAN

Perpustakaan merupakan sistem informasi yang di dalamnya terdapat aktifitas pengumpulan, pengelolaan, pengawetan, pelestarian dan penyajian serta penyebaran informasi. Perpustakaan sebagaimana yang ada dan berkembang sekarang telah dipergunakan sebagai pusat informasi, sumber ilmu pengetahuan, penelitian, rekreasi, pelestarian khasanah budaya bangsa, serta memberikan layanan jasa lainnya. Berdasarkan Undang-Undang Perpustakaan No.43 Tahun 2007 pasal 23 mengamanatkan bahwa setiap sekolah / madrasah harus memiliki dan menyelenggarakan perpustakaan yang memenuhi standar nasional perpustakaan dengan memperhatikan standar nasional pendidikan [1].

Perpustakaan SD Riyadlul Muta'allimin merupakan menjadi sarana untuk mencari informasi dan sumber belajar siswa maupun guru. Beberapa siswa memanfaatkan perpustakaan untuk mencari referensi berkenaan dengan tugas yang diberikan oleh guru ataupun mencari referensi untuk menyelesaikan tugas. Tugas-tugas yang diberikan oleh guru tersebut, menjadikan para siswa termotivasi untuk mengunjungi perpustakaan guna memperoleh hasil belajar yang maksimal. Hanya saja pada masa pandemic, perpustakaan mengalami kemerosotan pengunjung. Oleh karena itu pihak sekolah mencari solusi agar siswa meskipun belajar dari rumah namun masih tetap dapat mendapatkan referensi belajar, yaitu dengan menerapkannya Perpustakaan digital/online. Karena perpustakaan digital tersebut masih terbilang baru, maka perlu dilakukannya analisis kepuasan user terkait perpustakaan digital tersebut.

Penelitian sebelumnya pernah dilakukan oleh cristin, dkk pada penelitiannya dengan judul Penerapan Algoritma K-Means Dalam Menentukan Tingkat Kepuasan Pembelajaran Online Pada Masa Pandemi Covid-19. Tujuan penelitian ini adalah melakukan analisis kepuasan terhadap model pembelajaran daring pada masa pandemi. Metode yang digunakan adalah K-Means. Hasil dari penelitian ini adalah kluster pertama memiliki 4 anggota kelompok, yang artinya kelompok ini memiliki tingkat kepuasan yang rendah. Kluster ke 2 memiliki 7 anggota yang artinya kelompok ini memiliki tingkat kepuasan yang sedang. Terakhir kluster ke 3 memiliki 9 anggota, yang artinya siswa dengan model pembelajaran daring ini memiliki tingkat kepuasan yang tinggi [2].

Penelitian sebelumnya yang pernah dilakukan oleh Windarto ,dkk menyatakan bahwa Daerah resapan air merupakan sebuah daerah tempat masuknya air dari permukaan tanah ke dalam zona jenuh air sehingga membentuk suatu aliran air di dalam tanah. Kurangnya daerah resapan air menjadi masalah saat ini, karena daerah resapan air berfungsi untuk menampung debit air hujan yang turun di daerah tersebut agar tidak terjadi banjir dan masalah kekeringan. Sumber data diambil dari situs Badan Pusat Statistik Indonesia. Penelitian ini bertujuan untuk mengelompokkan daerah resapan air menggunakan algoritma

datamining dengan metode K-Means. Data yang digunakan dalam penelitian ini adalah data persentase rumah tangga berdasarkan provinsi dan keberadaan area resapan air pada tahun 2017 yang terdiri dari 34 provinsi di Indonesia. Berdasarkan data tersebut dieproleh pengelompokan area resapan air berdasarkan provinsi menjadi 2 clusteryaitu cluster tinggi (C1) dan cluster rendahmenggunakan K-means clustering. Dari hasil penelitian diperoleh pengelompokan menjadi 16 provinsi dengan cluster tinggi (C1) dan 18 provinsi dengan clusterrendah (C2). Hal ini dapat menjadi masukan pada pemerintah untuk provinsi yang memiliki daerah resapan air rendah menjadi perhatian lebih berdasarkan cluster yang telah dilakukan[3].

Penelitian selanjutnya pernah dilakukan oleh Eki pada penelitiannya yang berjudul Analisis Kepuasan Pelanggan Menggunakan Algoritma C4 .5. Tujuan dari peneliti ini adalah menganalisis kepuasan pelanggan telkomsel Kota Semarang menggunakan decision tree. Metode yang digunakan adalah decision tree. Hasil dari lasifikasi tersebut menunjukkan menunjukan bahwa data 70% : 30% akurasi 87% dan data 90% : 10% akurasi meningkat menjadi 93% [4].

Penelitian selanjutnya juga pernah dilakukan oleh Sherly, dkk dengan judul penelitian Penerapan Algoritma C4.5 Untuk Prediksi Kualitas Pelayanan Terhadap Kepuasan Konsumen. Tujuan dari penelitian ini adalah untuk menganalisa kualitas pelayanan terhadap kepuasan konsumen Hinet Batam dengan menggunakan algoritma decision tree. Metode yang digunakan adalah decision tree. Hasil dari penelitian ini adalah nilai presisi prediksi puas sebesar 61.7% dan nilai presisi prediksi tidakpuas 38,3%. Hasil penelitian ini bisa dijadikan acuan untuk meningkatkan kualitas pelayanan serta untuk mempertahankan loyalitas konsumen dalam menggunakan produk paket data layanan internet dan meningkatkan penjualan[5].

Penelitian selanjutnya juga pernah dilakukan oleh Sherly, dkk dengan judul penelitian Penerapan Algoritma C4.5 Untuk Prediksi Kualitas Pelayanan Terhadap Kepuasan Konsumen. Tujuan dari penelitian ini adalah untuk menganalisa kualitas pelayanan terhadap kepuasan konsumen Hinet Batam dengan menggunakan algoritma decision tree. Metode yang digunakan adalah decision tree. Hasil dari penelitian ini adalah nilai presisi prediksi puas sebesar 61.7% dan nilai presisi prediksi tidakpuas 38,3%. Hasil penelitian ini bisa dijadikan acuan untuk meningkatkan kualitas pelayanan serta untuk mempertahankan loyalitas konsumen dalam menggunakan produk paket data layanan internet dan meningkatkan penjualan[6]. Berikut tabel data pengunjung perpustakaan Raudari bulan September 2019 sampai Mei 2020 :

Tabel 1. Data Pengunjung Perpustakaan

Bulan	Pengunjung
September	47
Oktober	51
November	59
Desember	62
Januari	57
Februari	62
Maret	20
April	10
Mei	5

Berdasarkan data tersebut, dapat dilihat bahwa terjadi penurunan drastis pengunjung pada bulan maret sampai bulan setelahnya. Efek pandemi mengakibatkan kemrosotan jumlah pengunjung perpustakaan. Oleh sebab itu diperlukannya analisis kepuasan pengguna agar dapat meningkatkan lagi minat kunjungan ke perpustakaan sebagai literasi tambahan belajar siswa maupun guru. Selanjutnya mengevaluasi ke akuratan analisis yang dilakukan.

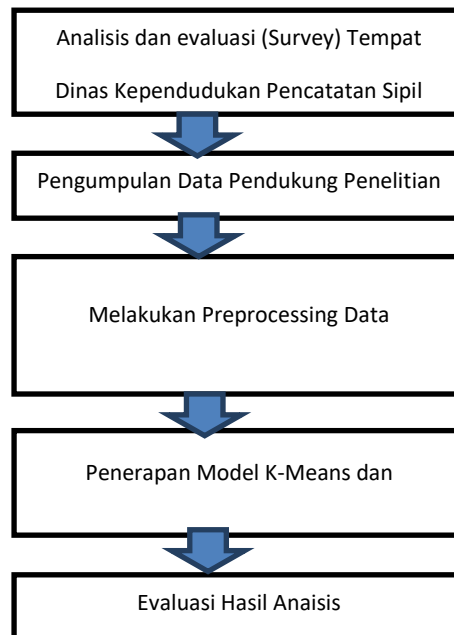
Solusi untuk mengatasi hal tersebut adalah dengan melakukannya analisis dengan penyebaran angket ke para siswa. Hal pertama yang dilakukan adalah mengklaster hasil angket menjadi 3 kelompok. Setelah terbentuknya 3 klaster, selanjutnya nama klaster dijadikan label untuk dapat diproses ke pemrosesan klasifikasi menggunakan Decission tree.

Metode analisis yang dapat digunakan berdasarkan pengalaman penelitian terdahulu yang sudah dipaparkan bahwa metode K-means dapat digunakan untuk mengelompokkan data kepuasan pengguna. Selanjutnya hasil pengklasteran dijadikan label yang nantinya akan diproses dengan teknik klasifikasi decision tree. Hasil klasifikasi akan dievaluasi performanya untuk bahan pertimbangan langkah selanjutnya.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Metode yang dapat digunakan berdasarkan pengalaman penelitian terdahulu yang sudah dipaparkan bahwa metode K-means dapat digunakan untuk mengelompokkan data kepuasan pengguna. Selanjutnya hasil pengklasteran dijadikan label yang nantinya akan diproses dengan teknik klasifikasi decision tree. Hasil klasifikasi akan dievaluasi performanya untuk bahan pertimbangan langkah selanjutnya.



Gambar 1. Alur Penelitian

Alur penelitian ini adalah yang pertama menganalisis dan survey tempat penelitian. Selanjutnya pengumpulan data pendukung yang digunakan untuk penelitian. Setelah data terkumpul, dilakukannya preprocessing data agar data dapat diproses. Setelah data normal, selanjutnya penerapan metode K-Means dan Decision tree. Terakhir adalah mengevaluasi hasil penelitian.

2.2 Langkah Proses Model K-Means dan Decision Tree

K-Means adalah salah satu metode data clustering non hierarki yang prosesnya membagi data ke dalam bentuk satu atau lebih cluster atau kelompok sehingga data yang memiliki karakteristik yang persis dikelompokkan ke dalam satu cluster yang serupa dan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam kelompok lainnya. K-Means adalah metode clustering berbasis jarak yang membagi data ke dalam sejumlah cluster dan algoritma ini hanya bekerja pada atribut yang bersifat numeric[7]. Langkah-langkah melakukan clustering dengan menggunakan metode K-Means adalah sebagai berikut:

- Menentukan nilai K, jumlah cluster
- Menentukan nilai centroid, dalam menentukan nilai centroid untuk awal iterasi, nilai awal centroid dilakukan secara acak. Sedangkan nilai centroid ketika proses iterasi berlangsung, maka digunakan rumus :

$$\bar{v}_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj} \quad (1)$$

Keterangan:

\bar{v}_{ij} = data cluster ke- i kolom j

x_{kj} = data ke-k kolom ke-j

N_i = banyaknya anggota kluster ke- i

- Menghitung jarak antara titik centroid dengan titik tiap objeknya, dengan rumus Euclidean distance space tidak jarang digunakan dalam perhitungan jarak, hal ini dikarenakan hasil yang diperoleh adalah jarak terpendek antara dua titik yang diperhitungkan. Adapun persamaannya adalah sebagai berikut:

$$D_L(x_2 - x_1) = \sqrt{\sum_{j=1}^p (x_{2j} - x_{1j})^2} \quad (2)$$

Keterangan

$D_L(x_2 - x_1)$ = Jarak objek antara objek i dan j

P = Dimensi data

x_{2j} = Koordinat dari obyek i pada dimensi k

x_{1j} = Koordinat dari obyek j pada dimensi k

- Pengelompokkan objek untuk menentukan anggota cluster adalah dengan mencari jarak minimum objek.

Kembali ke tahap kedua, lakukan perulangan hingga nilai centroid yang dihasilkan tetap atau tidak berubah dan anggota cluster tidak berpindah ke cluster lainnya [8]. Salah satu contoh pengolahan data mining adalah dengan Decision Tree yaitu salah satu metode klasifikasi yang paling populer karena mudah untuk diinterpretasi oleh manusia. Decision Tree adalah struktur flowchart yang menyerupai Tree (pohon), dimana setiap simpul internal menandakan suatu tes pada atribut, setiap cabang merepresentasikan hasil tes, dan simpul daun merepresentasikan kelas atau distribusi kelas. Alur pada Decision Tree di telusuri dari simpul akar ke simpul daun yang memegang prediksi [9].

Nilai Gain adalah Information Gain yang digunakan untuk mencari satu variabel/ atribut dari dataset (S) untuk dijadikan root/ node dan branch node, yaitu satu atribut yang mempunyai nilai gain tertinggi. Untuk pencarian nilai gain (information Gain) dapat dilakukan dengan konsep Entropy, Gini Index dan Classification Error. Nilai gain terbesar yang di dapat dari atribut pada dataset (data pembelajaran) pertama kali digunakan untuk mencari atribut yang layak menjadi root (akar) pohon keputusan (Decision Tree). Selanjut proses akan diulang untuk mencari atribut yang menjadi cabang hingga menemukan leaf yang merupakan label kelas. Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy. Entropi adalah formula untuk menghitung homogenitas atribut (A) dari sebuah sample data (S). Untuk menghitung nilai entropy digunakan rumus : [10]

$$\text{Entropy (S)} = \sum_{i=1}^n - p_i * \log_2 p_i \quad (3)$$

Keterangan:

S = Himpunan kasus

n = Jumlah partisi S

p_i = Proporsi S_i terhadap S

Setelah itu mencari nilai information gain. Information Gain merupakan kriteria yang paling populer untuk pemilihan atribut. Information gain dapat dihitung dari output data atau variabel dependen y yang dikelompokkan berdasarkan atribut A, dinotasikan dengan gain (y,A). Gain (y,A) dari atribut A relatif terhadap output data y adalah:

$$\text{Gain (S,A)} = \text{Entropy (S)} - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy (S}_i\text{)} \quad (4)$$

Keterangan:

S = Himpunan kasus

A = Fitur

n = Jumlah partisi atribut A

$|S_i|$ = Proporsi S_i terhadap S

$|S|$ = jumlah kasus dalam S

3. HASIL DAN PEMBAHASAN

Dalam melakukan penelitian ini menggunakan data dari angket yang disebar ke para siswa beserta guru di SD Riyadlul Muta'allimin. Adapun pertanyaan yang ada di dalam kuisioner adalah sebagai berikut :

Tabel 2. Butir Pertanyaan Kuisioner

No	Indikator	Pertanyaan
1	Antarmuka	1. Apakah Perpustakaan digital mudah digunakan? (1.1) 2. Apakah Perpustakaan digital mempermudah anda untk mencari materi yang dibutuhkan? (1.2) 3. Apakah Perpustakaan digital sangat <i>user-friendly</i> (1.3)
2	Personalisasi	1. Apakah Perpustakaan digital dapat membantu meningkatkan kemajuan belajar anda? (2.1) 2. Apakah Perpustakaan digital membantu anda mencari materi yang dibutuhkan? (2.2)
3	Support	1. Apakah user guide yang disediakan dapat membantu anda? (3.1) 2. Apakah pelatihan yang dilakukan sudah sangat jelas? (3.2) 3. Jika terjadi kendala, apakah ditanggapi dengan cepat? (3.3)
4	Kemudahan akses	1. Apakah Perpustakaan digital dapat diakses dengan mudah? (4.1) 2. Apakah materi yang anda cari mudah ditemukan? (4.2)

Siswa dan guru yang mengisi form angket akan diberi bantuan penilaian dengan teknik skala likert, adapun aturan skala likert yang digunakan adalah:

Tabel 3. Skala Likert

Angka	Deskripsi
5	Sangat Puas
4	Puas
3	Cukup Puas
2	Tidak Puas
1	Sangat Tidak Puas

Setelah kuisioner sudah tersebar, selanjutnya seleksi data. Data yang digunakan sebagai sampel 189 data. Agar tidak ada lagi data yang mengandung noise atau missing value dan data dinyatakan normal, data terlebih dahulu harus melalui

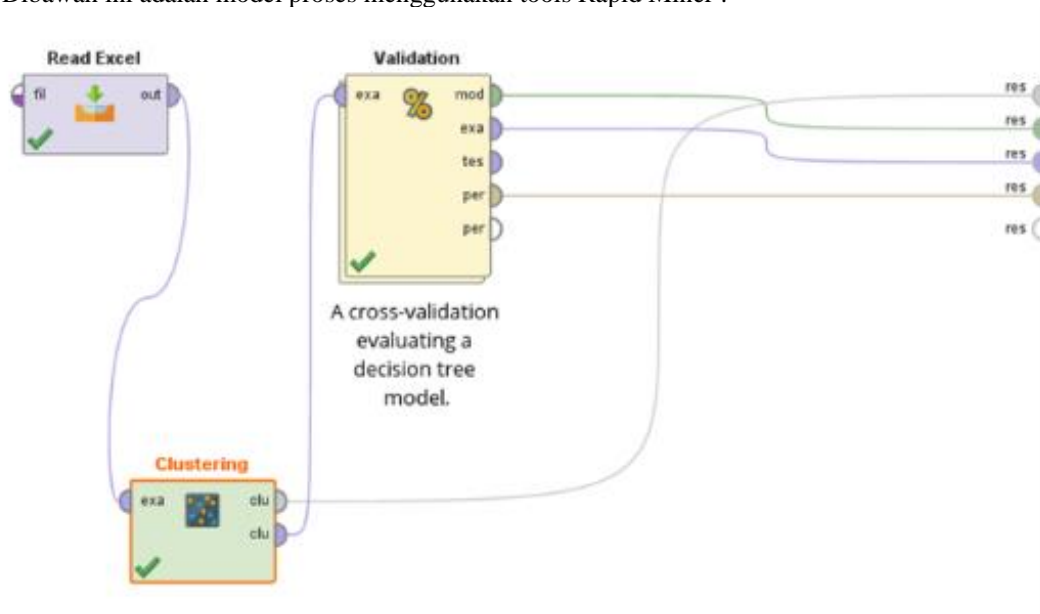
preprocessing data. Selanjutnya Syarat agar data dapat diproses ialah data harus dalam keadaan normal. Bebas dari missing value, dan bebas dari redudansi data.

Tabel 4. Hasil Angket yang sudah disebar

Responden	Pertanyaan									
	1.1	1.2	1.3	2.1	2.2	3.1	3.2	3.3	4.1	4.2
1	3	2	2	3	3	3	3	3	3	3
2	2	4	5	2	3	4	4	3	3	2
3	4	5	2	5	4	3	5	4	4	2
4	4	5	5	5	4	5	4	5	4	4
5	4	5	5	5	5	5	5	5	4	5
6	2	3	2	2	4	4	2	4	3	3
7	5	5	5	4	5	4	5	4	5	4
8	2	4	5	2	3	4	4	3	3	2
9	3	2	2	3	3	3	3	3	3	3
10	5	5	4	5	5	5	4	5	5	5
11	4	5	5	5	4	5	4	5	4	4
12	2	2	3	3	2	2	2	2	2	3
13	5	5	4	5	5	5	4	5	5	5
14	4	5	5	5	4	5	4	5	4	4
15	2	3	3	3	3	2	2	2	2	3
....

Langkah selanjutnya adalah menerapkan teknik K-Means dan Decision tree dalam menganalisis kepuasan pengguna perpustakaan online. Adapun langkah pertama yaitu data harus melalui proses K-Means terlebih dahulu, dan selanjutnya nama klaster dijadikan label untuk proses decision tree. Langkah- Langkah perhitungan K-Means ialah menentukan Jumlah cluster data. Selanjutnya tentukan titik pusat cluster. Setelah itu menghitung jarak obyek dengan centroid. Kemudian kelompokkan obyek. Jika kelompok data hasil perhitungan baru sama dengan hasil perhitungan kelompok data baru maka selesailah perhitungannya.

Model proses yang digunakan adalah Algoritma K-Means dan Decision tree, dengan menggunakan tools Rapid Miner . Rapid Miner merupakan software platform yang digunakan mengolah berbagai data mining. Sebelum mengolah sebuah data mining dengan tools Rapid Miner , langkah pertama yang harus dilakukan yaitu menyusun atau membuat model proses dengan operator yang sesuai dengan teknik data mining yang digunakan dan operator yang telah tersedia pada tools Rapid Miner. Dibawah ini adalah model proses menggunakan tools Rapid Miner :



Gambar 2. Model Proses

Setelah model proses sudah dibuat, langkah selanjutnya adalah mengeksekusi model proses menggunakan rapid miner. Berikut adalah hasil dari pemrosesan K-Means :

Cluster Model

```
Cluster 0: 51 items
Cluster 1: 52 items
Cluster 2: 47 items
Total number of items: 150
```

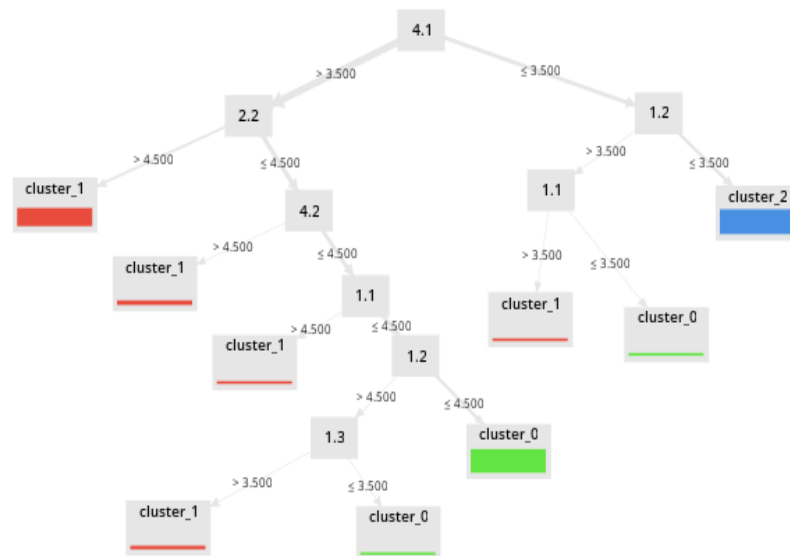
Gambar 3. *Cluster Model*

Agar hasil klaster dapat dibaca dengan mudah. Maka nama klaster di ubah sesuai dengan rata-ratanya.

Tabel 3. Penamaan Klaster

Klaster	Nama Klaster
0	Tidak Puas
1	Cukup
2	Puas

Penamaan klaster ini nantinya akan dijadikan sebagai label untuk pemrosesan model Decission Tree. Berikut hasil pemrosesan Decission tree yang data setnya berasal dari hasil pemrosesan K-Means:



Gambar 4. Hasil Decission Tree

Berdasarkan pohon keputusan tersebut, node tertinggi dimiliki oleh atribut 4.1, sedangkan turunan node 4.1 yang memiliki node terbanyak adalah 2.2 dengan klasifikasi atribut 4.1 memiliki nilai lebih dari 3,5. Sedangkan akar sebelahnya hanya memiliki 2 akar dan 3 node saja.

Gambar tersebut menggambarkan hasil klasifikasi yang dibentuk pohon keputusan. Adapun performa yang dihasilkan adalah :

Tabel 4. Performance Vector

	true cluster_2	true cluster_0	true cluster_1	class precision
pred. cluster_2	47	3	2	90.38%
pred. cluster_0	0	48	2	96.00%
pred. cluster_1	0	0	48	100.00%
class recall	100.00%	94.12%	92.31%	

Berdasarkan tabel tersebut, prediksi kluster 1 dengan true kluster 1 sebesar 48 item, true kluster 0 dan true kluster 0 tidak terdapat item, maka class precissionnya 100%. Prediksi kluster 2 dan true kluster 2 sebanyak 47, true kluster 0 sebanyak 3, dan true kluster 1 sebanyak 2 item dengan class precission sebesar 90,38%. Serta prediksi kluster 0 dengan true kluster 0 terdapat 48 item, dan true kluster 1 sebanyak 2 item dengan class precission sebesar 96%. Sehingga akurasi dari pemodelan ini menjadi 95,33%, dengan kata lain pembelajaran mesin dengan ketepatan klasifikasinya sebesar 95,33%.

3.1 Implementasi

Model proses ini adalah gabungan dari 2 model, yaitu model K-Means dan model Decision Tree. Sehingga penelitian ini bukanlah penelitian untuk membandingkan 2 metode. Berdasarkan hasil penelitian yang dilakukan untuk analisis kepuasan pengguna perpustakaan digital menggunakan algoritma K-Means dan decision tree, dapat disimpulkan bahwa dengan dataset berjumlah 150 data menghasilkan persentase akurasi sebesar 95,33%.

Berdasarkan dari hasil penelitian yang dilakukan di SD Riyadlul Muta'allimin, maka kesimpulan yang didapat adalah penggabungan Algoritma K-Means dan Decision tree merupakan suatu penggabungan metode yang relevan untuk digunakan dalam analisis indeks kepuasan pengguna perpustakaan digital di era pandemi ini. Dengan menerapkan model Penggabungan Algoritma K-Means dan Decision Tree yang diproses melalui rapidminer, maka pihak sekolah dapat dengan mudah menganalisis indeks kepuasan pengguna perpustakaan digital di SD Riyadlul Muta'allimin.

Kelompok data klaster 0 adalah data yang memiliki asumsi pengguna yang tidak puas dengan perpustakaan digital sekolah. Sedangkan klaster 1 merupakan kelompok data yang memiliki tingkat kepuasan yang cukup. Selanjutnya klaster 2 adalah data yang memiliki tingkat kepuasan yang paling tinggi karena mereka puas dengan adanya perpustakaan online ini.

4. KESIMPULAN

Berdasarkan dari hasil penelitian yang dilakukan di SD Riyadlul Muta'allimin maka kesimpulan yang didapat adalah Analisis Indeks Kepuasan pengguna perpustakaan digital dapat dilakukan dengan menggunakan K-Means dan Decision tree di SD Riyadlul Muta'allimin. Hasil evaluasi dari analisis Kepuasan pengguna perpustakaan digital menggunakan K-Means dan decision tree Di SD Riyadlul Muta'allimin didapati dengan dataset berjumlah 150 data menghasilkan persentase akurasi sebesar 95,33%. Klaster terbaik berdasarkan hasil proses data mining adalah Klaster 2. Agar selanjutnya penelitian ini dapat dikembangkan lagi kedepannya, berikut adalah saran dari peneliti yang dijadikan bahan literasi penelitian yang akan datang. Agar hasil penelitian ini lebih baik lagi, dapat menggunakan jumlah data yang lebih banyak lagi dan melakukan pengujian dengan dataset yang lebih lengkap atributnya sehingga model yang sudah dihasilkan akan lebih teruji lagi kedepannya. Menggunakan algoritma lain dalam pengelompokan data. Untuk mendapatkan nilai dbi yang lebih baik lagi dari penelitian ini, dapat melakukan optimasi dengan teknik optimasi lainnya.

REFERENCES

- [1] R. Setyaningrum and Y. Mulyanto, "Pengembangan Sistem Informasi Perpustakaan Pada Universitas Teknologi Sumbawa Berbasis Android," *J. Tambora*, vol. 3, no. 2, pp. 26–37, 2019.
- [2] K. D. R. Sianipar, S. Wanti Siahaan, M. Siregar, and P. R. Fikrul Ilmi Zer, "Penerapan Algoritma K-Means Dalam Menentukan Tingkat Kepuasan Pembelajaran Online Pada Masa Pandemi Covid-19," *J. Teknol. Inf.*, vol. 4, no. 1, pp. 101–105, 2020.
- [3] D. A. Silitonga, A. P. Windarto, and D. Hartama, "Penerapan Metode K-Medoid pada Pengelompokan Rumah Tangga Dalam Perlakuan Memilah Sampah Menurut Provinsi," *Semin. Nas. Sains Teknol. Inf. SENSASI 2019 ISBN*, pp. 313–318, 2019.
- [4] R. M. Eki, "Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5," *Fak. Ilmu Komput.*, vol. 3, no. 2, pp. 1–14, 2017.
- [5] S. M. Putri and S. A. Arnomo, "Penerapan Algoritma C4.5 Untuk Prediksi Kualitas Pelayanan Terhadap Kepuasan Konsumen (Studi Kasus : Hinet Batam)," *J. Inf. Syst. Res.*, vol. 1, no. 2, pp. 70–76, 2020.
- [6] A. Sarker, S. M. Shamim, M. Shahiduz, Z. M. Rahman, M. Shahiduz Zama, and M. Rahman, "Employee's Performance Analysis and Prediction using K-Means Clustering & Decision Tree Algorithm Mawlana Bhashani Science and Technology University Employee's Performance Analysis and Prediction using K-Means Clustering & Decision Tree Algorithm," *Type Double Blind Peer Rev. Int. Res. J. Softw. Data Eng. Glob. J. Comput. Sci. Technol. C*, vol. 18, no. 1, 2018.
- [7] S. Kusumadewi and I. Ruspita, "Penerapan Metode Inferensi Backward Chaining Dalam Sistem Pakar Pemilihan Obat Untuk Pasien Gigi Dengan Penyakit Sistemik," *J. Inform. dan Komputer) Akreditasi KEMENRISTEKDIKTI*, vol. 3, no. 2, pp. 92–100, 2020, doi: 10.33387/jiko.
- [8] F. Wang, Q. Wang, F. Nie, Z. Li, W. Yu, and F. Ren, "A linear multivariate binary decision tree classifier based on K-means splitting," *Pattern Recognit.*, vol. 107, no. June, p. 107521, 2020, doi: 10.1016/j.patcog.2020.107521.
- [9] J. Eska, "Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4.5," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 2, 2018, doi: 10.31227/osf.io/x6svc.
- [10] P. Kasih, "Pemodelan Data Mining Decision Tree Dengan Classification Error Untuk Seleksi Calon Anggota Tim Paduan Suara," *Innov. Res. Informatics*, vol. 1, no. 2, pp. 63–69, 2019, doi: 10.37058/innovatics.v1i2.918.