

Penerapan *Data Mining Clustering* Menggunakan Metode K-Means Dalam Pengelompokan Buku Perpustakaan Politeknik Negeri Balikpapan

Maria Ulfah
Teknik Elektro
Politeknik Negeri Balikpapan
Balikpapan, Indonesia
maria.ulfah@poltekba.ac.id

Andi Sri Irtwaty
Teknik Mesin
Politeknik Negeri Balikpapan
Balikpapan, Indonesia
Andi.sri@poltekba.ac.id

Abstrak—Perpustakaan Politeknik Negeri Balikpapan, setiap tahunnya mengadakan penambahan koleksi buku bacaan. Agar pengadaan buku tersebut sesuai dengan kebutuhan pengguna yakni mahasiswa maka diperlukan informasi koleksi buku yang dibutuhkan atau diminati. Untuk menjawab permasalahan ini, maka perlukan sistem clustering buku-buku yang ada dipergustakaan dengan melihat aspek frekuensi peminjaman. Sistem *Clustering* yang dibuat menggunakan metode K-Means dengan pemilihan 3 *Cluster* yakni, sangat diminati, diminati dan kurang diminati. Dari hasil pengolahan data melalui Aplikasi Rapid Miner dengan $K=3$, didapatkan hasil *cluster_0* (rendah) terdiri 82 judul buku dengan frekuensi peminjaman bukunya dalam kategori jarang dengan kata lain kurang diminati untuk dipinjam, *cluster_1* (sedang) terdiri dari 23 judul buku yang merupakan buku-buku dengan frekuensi peminjaman sedang dengan kata lain dikategorikan diminati untuk dipinjam, *cluster_2* (tinggi) terdiri dari 2 judul buku merupakan kelompok buku yang paling diminati yang terdiri dari 2 judul buku yakni Teknologi Beton dan buku Teori dan Praktik hotel *Front Office*. Hasil pengelompokan data buku-buku ini dapat menjadi masukan bagi pengelola perpustakaan dalam pengadaan koleksi buku berdasarkan frekuensi peminjaman buku.

Kata Kunci—*Clustering, K-Means, Centroid, Perpustakaan, Rapid Miner*

I. PENDAHULUAN

Keberadaan perpustakaan tidak dapat dipisahkan dari budaya manusia. Tinggi rendahnya peradaban suatu bangsa dapat dilihat dari kondisi perpustakaan yang dimiliki. Pada hakekatnya perpustakaan merupakan hasil budaya berupa lembaga yang mengumpulkan, menyimpan, mengatur baik berupa karya cetak maupun karya rekam sebagai sumber informasi dan belajar dari generasi ke generasi. Di Indonesia ada lima jenis perpustakaan dan kelima jenis perpustakaan ini yaitu perpustakaan nasional, perpustakaan umum, perpustakaan khusus dan perpustakaan perguruan tinggi.

Adapun koleksi perpustakaan perguruan tinggi diadakan melalui seleksi yang mengacu kepada kebutuhan program studi yang diselenggarakan dan diorganisasikan sedemikian rupa sehingga dapat menjamin efektivitas dan efisiensi layanan kepada kebutuhan *civitas academica*. Pada setiap *civitas academica* pun tidak dapat disamakan setiap kebutuhannya karena memiliki kebutuhan berbeda pada literturnya. Literatur yang digunakan mahasiswa setiap jurusan pun berbeda. Dengan adanya pengelompokan yang

terorganisir pada setiap penggunaannya dalam hal ini adalah peminjam dari berbagai jurusan maka dapat diketahui kelompok literatur apa saja yang paling sering dipinjam oleh mahasiswa [1].

Perpustakaan Politeknik Negeri Balikpapan bertujuan mendukung kebutuhan informasi, kondisi minat baca perpustakaan dapat dilihat dari persentase jumlah pengunjung, peminjam dan buku yang dipinjam. Data peminjaman buku dijadikan sebagai laporan. Pihak Perpustakaan setiap tahunnya melakukan pengadaan koleksi buku baru untuk diperbanyak sehingga harus mengetahui jenis koleksi buku prioritas untuk diperbanyak. Sehingga diperlukan adanya analisis pengolahan data transaksi peminjaman buku yang menjelaskan informasi sebaran intensitas peminjaman buku, informasi pengolahan tersebut membantu penambahan koleksi.

Untuk mengetahui Buku apa yang paling diminati, dilakukan teknik *cluster* menggunakan metode K-Means. Algoritma K-Means merupakan algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah *K Cluster* yang sudah ditetapkan di awal” [2].

Oleh karena itu penulis dalam penelitian ini mengambil judul “Penerapan *Data Mining Clustering* Menggunakan Metode K-Means Dalam Pengelompokan Buku Perpustakaan Politeknik Negeri Balikpapan” untuk membantu mengelompokkan data buku pinjaman yang diminati dalam mendukung Pengelola Perpustakaan Politeknik Negeri Balikpapan dalam pengadaan koleksi buku.

II. TINJAUAN PUSTAKA

A. Data Mining

Data mining adalah suatu istilah yang digunakan untuk untuk menguraikan penemuan pengetahuan di dalam database. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakrit dari berbagai database besar [3].

Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual [4]. Dimana hasil dari proses penggalian tersebut akan membentuk pola-pola dari kumpulan data, yang sering disebut dengan

pengenalan pola (*pattern recognition*). Pengenalan pola merupakan bagian dari *data mining*. *Data mining* sering juga disebut *Knowledge Discovery in Database (KDD)*, yaitu kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [5]

B. Clustering

Pada dasarnya *clustering* merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*Similarity*) antara satu dengan data yang lain. *Clustering* merupakan merupakan salah satu metode *data mining* yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak menemukan target *output*. Dalam *data mining* ada dua jenis metode *clustering* yang digunakan dalam pengelompokan data, yaitu *hierarchical clustering* dan *non hierarchical clustering* [5]. *Hierarchical clustering* adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua. Demikian proses seterusnya sehingga cluster akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah *cluster*. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut [5]

Berbeda dengan metode *hierarchical clustering*, metode *non-hierarchical clustering* justru dimulai dengan menentukan terlebih dahulu jumlah *cluster* yang diinginkan (dua *cluster*, tiga *cluster*, atau lain sebagainya). Setelah jumlah *cluster* diketahui, baru proses *cluster* dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *K-Means Clustering* [5], [6]

Algoritma K-Means merupakan algoritma yang relatif sederhana untuk mengklasifikasikan atau mengelompokkan sejumlah besar objek dengan atribut tertentu kedalam kelompok-kelompok *cluster* sebanyak K. Pada algoritma k-Means, jumlah *cluster* K sudah ditentukan terlebih dahulu. K-Means adalah algoritma *clustering* untuk data mining yang diciptakan tahun 70-an dan berguna untuk melakukan *clustering* secara *unsupervised learning* (pembelajaran yang tidak terawasi) dalam suatu kumpulan data berdasarkan parameter tertentu. K-Means adalah sebuah algoritma untuk mengklasifikasikan atau mengelompokkan objek-objek (dalam hal ini data) berdasarkan parameter tertentu kedalam sejumlah group, sehingga berjalan lebih cepat dari pada *hierarchical clustering* (jika kecil) dengan jumlah variabel yang besar dan menghasilkan *cluster* yang lebih rapat [7]. K-Means memiliki properti: selalu ada K *cluster*, paling tidak memiliki satu data dalam tiap *cluster*, *cluster* ini merupakan non-hierarki dan tidak akan terjadi *Overlap*, dan setiap member dari sebuah cluster berdekatan di *cluster* terhadap *cluster* lainnya karena kedekatan tidak selalu melibatkan pusat dari *cluster* itu.

III. METODOLOGI PENELITIAN

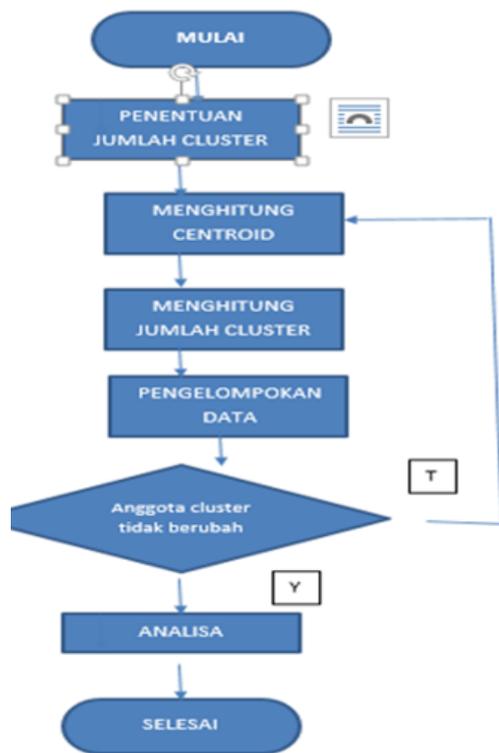
Gambar 1 merupakan gambaran metode pelaksanaan pada penelitian ini.



Gambar 1. Alur Penelitian

- 1) Langkah awal dalam penelitian ini adalah melakukan perencanaan penelitian. Ada empat kegiatan dalam perencanaan yaitu penentuan tujuan penelitian, identifikasi masalah, penentuan batasan masalah dan studi pustaka. Dalam tahap pengumpulan data dilakukan dengan dua cara yaitu:
 - a) Studi pustaka Pada kegiatan ini peneliti mencari dan mempelajari buku, *ebook* dan jurnal untuk mendukung penulisan yang berkaitan dengan topik yang diangkat. Dari studi pustaka yang di pelajari di peroleh teori dan metode dalam penyelesaian masalah penelitian yang berkaitan dengan topik yang diangkat.
 - b) Observasi dan wawancara, Dari hasil wawancara diperoleh data wawancara yang nantinya akan digunakan untuk penulisan laporan dan dari hasil observasi diperoleh data dokumentasi dan data peminjaman buku perpustakaan.
- 2) Tahap *Preprocessing*: Dalam tahap *preprocessing* kegiatan yang dilakukan adalah mengetik kembali data yang didapatkan kedalam microsoft excel untuk mencatat semua transaksi peminjaman buku setelah data di salin pada tahap ini juga akan di lakukan *cleaning* data, yaitu penghapusan data yang tidak jelas penulisannya atau data yang tidak bisa dibaca. Proses *cleaning* ini dilakukan agar mendapatkan hasil perhitungan yang benar.
- 3) Tahap Pengolahan Data: Mengolah data dengan *software* rapidminer dan juga mengolah data berdasarkan identifikasi masalah pada penelitian. Menggunakan metode K-Means, dan melalui penerapan algoritma k-means *clustering* diharapkan dalam mengolah data ini mendapatkan hasil yang baik untuk mengelompokkan buku berdasarkan frekuensi peminjaman. hasil belajar dan menghasilkan suatu pengetahuan baru. Mengolah data dengan *software* Rapidminer
- 4) Tahap Analisa: Setelah semua data terkumpul selanjutnya dilakukan tahap analisa.
- 5) Tahap Dokumentasi: Proses dokumentasi dilakukan dari awal penelitian hingga akhir penelitian. Hasil akhir dari proses dokumentasi berupa laporan kemajuan dan Laporan Akhir penelitian.

Gambar 2 merupakan diagram alir penelitian ini, dapat digunakan untuk lebih memahami arah yang akan dituju.



Gambar 2. Diagram Alir Penelitian

IV. HASIL DAN PEMBAHASAN

A. Dataset

Dataset yang digunakan adalah data peminjaman buku pada Perpustakaan Politeknik Negeri Balikpapan dari kurun waktu 2019- Juni 2022. Banyaknya jenis buku yang terdata sejumlah 107 judul dengan total frekuensi peminjaman 231, terlihat pada Tabel 1 yang berisi *record* judul buku dan frekuensi peminjaman setiap tahun.

TABEL I. Dataset Penelitian

No	Judul Buku	2019	2020	2021	2022	Total
1	101 Amazing Public Relation Ideas				1	1
2	9 Days Umratan	1				1
3	Akuntansi Dasar 1 Dan 2				2	2
4	Akuntansi Manajerial Suatu Orientasi Praktis				2	2
5	Akuntansi Perbankan Syariah				2	2
6	Akuntansi Perbankan Syariah Teori Dan Praktik Kontemporer Berdasarkan PAPSII 2013 Edisi Kedua			2	1	3
7	Akuntansi Syariah				1	1
8	Akuntansi Perusahaan Manufaktur				2	2
9	Algoritma & Pemrograman Dengan C++ Edisi Kedua		1		1	2
10	All About Corporate Valuation				1	1
11	Anak Rantau		1			1

TABEL I. Dataset Penelitian (Lanjutan)

No	Judul Buku	2019	2020	2021	2022	Total
12	Analisa Laporan Keuangan Edisi Keempat		2		4	6
13	Analisis Kritis Atas Laporan Keuangan				7	7
14	Aneka Kue		1			1
15	Aplikasi Logika Fuzzy Edisi Kedua				1	1
16	Arduino Belajar Cepat Dan Pemrograman		1		2	3
17	Aroma Karsa		1			1
18	Auditing Buku 1 Edisi Keenam				5	5
19	Bahan Produk Bakery		1			1
20	Belajar Jaringan Komputer Berbasis Mikrotik OS	1				1
21	Budidaya Unggul Lele Phyton			1		1
22	Build Up Your English Reading Skill				1	1
23	Catatan Harian Anne Frank		1			1
24	Dahsyatnya Bisnis Hotel Di Indonesia				1	1
25	Dasar - Dasar Pengolahan Makanan	1				1
26	Dasar - Dasar Perbankan Edisi Revisi 2014				6	6
27	Drainase Perkotaan				3	3
28	Ekonomi Manajerial Dengan Pendekatan Matematis Edisi Revisi		1			1
29	Ekonomi Moneter				1	1
30	Emotional Intelligence	1	1			2
31	Ensiklopedia Mini - Hotel				2	2
32	Food And Beverage Service Operational				2	2
33	Front Office Management				5	5
34	Hidrologi				1	1
35	Hidrologi Jilid 1				5	5
36	Hidrologi untuk Pengairan				1	1
37	Hotel Courtesy				1	1
38	Housekeeping Hotel				1	1
39	Housekeeping Hotel Edisi Kedua				2	2
40	Hujan		2			2
41	Investasi Dan Pasar Modal Indonesia				3	3
42	Jurus Sukses Beternak Lele Sangkuriang			1		1
43	Kamus Populer Istilah Komputer dan Informatika		2			2
44	Khasiat Buah Dan Sayur				1	1
45	Kitab Khasiat Buah Dan Sayur				1	1

TABEL I. *Dataset* Penelitian (Lanjutan)

No	Judul Buku	2019	2020	2021	2022	Total
46	Kitab Kue Super Yummy		1			1
47	Konfigurasi Wireless Routerboard Mikrotik	1				1
48	Lintang				1	1
49	Manajemen Divisi Kamar				2	2
50	Manajemen Keuangan (Finance Management)				2	2
51	Manajemen Keuangan Modern				2	2
52	Manajemen Keuangan Sebagai Dasar Pengambilan Keputusan Bisnis				2	2
53	Manajemen Keuangan Teori Konsep & Aplikasi Edisi Revisi				2	2
54	Manajemen Kinerja Edisi Kelima				1	1
55	Manajemen Pemasaran Bank				5	5
56	Manajemen Pemasaran Jilid 1 Edisi 12				4	4
57	Manajemen Penyelenggaraan Hotel				2	2
58	Manajemen Perkreditan				1	1
59	Manajemen Risiko				2	2
60	Mengelola Kualitas Layanan Perbankan (Cover Baru)				3	3
61	Metode Penelitian Kuantitatif				2	2
62	Metode Penelitian Pendidikan - Pendekatan Kuantitatif, Kualitatif Dan R&D				4	4
63	Metodologi Penelitian Bisnis & Ekonomi				6	6
64	Metodologi Penelitian Bisnis Untuk Akuntansi & Manajemen Edisi Pertama				4	4
65	Metodologi Penelitian Dan Statistika				3	3
66	Mikrotik Kung Fu Kitab 1		1			1
67	Mikrotik Untuk Pemula		1			1
68	Modul Pemrograman Web (HTML, PHP, & MySQL) Edisi Ketiga				1	1
69	Napoleon Hill's Keys To Success				1	1
70	Nikmatnya Bangun Pagi, Tahajud, Subuh dan Dhuha				1	1
71	Panduan Praktis Dasar Analisa Laporan Keuangan				4	4
72	Paradigma, Metodologi & Aplikasi Ekonomi Syari'ah				1	1

TABEL I. *Dataset* Penelitian (Lanjutan)

No	Judul Buku	2019	2020	2021	2022	Total
73	Pedoman Pengerjaan Beton				3	3
74	Penerapan Soft Computing Dengan Matlab Edisi Revisi				1	1
75	Pengantar Akuntansi Edisi IFRS				1	1
76	Pengantar Akuntansi Lengkap Dengan Kumpulan Soal Dan Solusinya				1	1
77	Pengantar Bisnis Respon Terhadap Dinamika Global Ed.2				1	1
78	Pengantar Falsafah Ekonomi Dan Keuangan Syariah				1	1
79	Pengantar Ilmu Bisnis				4	4
80	Pengantar Metodologi Penelitian				1	1
81	Pengantar Statistika Sosial				3	3
82	Perempuan Kedua				1	1
83	Perencanaan dan Pemeliharaan Sistem Plambing				1	1
84	Perencanaan Struktur Baja Dengan Metode LRFD			1		1
85	Perilaku Konsumen Di Era Internet Implikasinya Pada Strategi Pemasaran				2	2
86	Praktikum Akuntansi Keuangan Manual Kasus Perusahaan Dagang Ed. Revisi				2	2
87	Rahasia Sukses Bisnis Dan Budi Daya Lele Unggul			1		1
88	Ranah 3 Warna		1			1
89	Rancang Bangun 3D dengan AutoCad 2012				1	1
90	Rangkaian Listrik		1			1
91	Santri Cangkir		2			2
92	Smart Traders Not Gamblers				3	3
93	Soal & Jawab Ekonomika Untuk Manajer		1			1
94	Statistik Deskriptif Untuk Ekonomi				2	2
95	Statistika Ekonomi & Bisnis				1	1
96	Sukses Beternak Lele Dumbo & Lele Lokal			1		1
97	Teknik Dan Prosedur Divisi Kamar Pada Bidang Hotel				3	3
98	Teknik Perhitungan Debit Rencana Bangunan Air				2	2
99	Teknologi Beton				14	14

TABEL I. *Dataset* Penelitian (Lanjutan)

No	Judul Buku	2019	2020	2021	2022	Total
100	Teknologi Perbankan				2	2
101	Teori Dan Praktik Hotel Front Office				13	13
102	The Greats On Leadership			1		1
103	The Maxwell Daily Reader	1		1		2
104	Ubur - Ubur Lembur				1	1
105	UMKM Aspek Hukum Dan Manajemen Pemasaran Produk				1	1
106	Who Am I				1	1
107	Who Moved My Cheese		2			2

B. Preprocessing Data

Pada tahapan ini, menggunakan *tools* RapidMiner dengan metode K-Means. Tampilan *New Process* ialah untuk mempersiapkan lembar kerja pada RapidMiner. melakukan impor data yang akan diuji berisi format .xls atau .xlsx. selanjutnya ialah model untuk melakukan impor file Microsoft Excel. Untuk memasukkan data yang akan dijalankan perlu melalui klik kanan -> Insert Operator -> Data Access -> File -> Read Excel seperti Gambar 3.

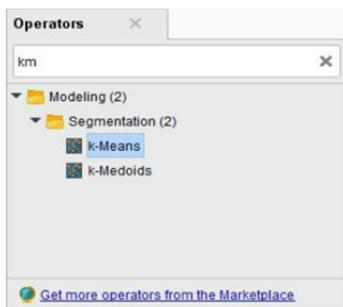


Gambar 3. Impor Data

Dari *dataset* yang ada kemudian dilakukan *preprocessing* data, salah satunya yaitu mengubah nama atribut judul buku menjadi kode unik (id) dengan tipe data adalah polynominal. Untuk atribut Total peminjaman tiap buku setiap tahunnya tipe data adalah *integer*. Hasil *preprocessing* data dengan menggunakan Rapid Miner, tidak terdapat *missing value* untuk kedua atribut.

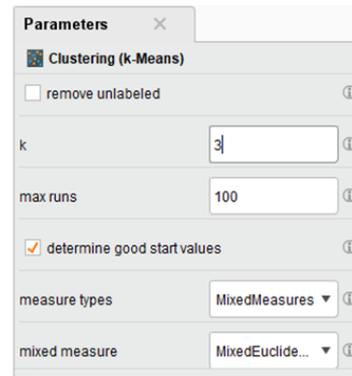
C. Modelling dan Evaluasi

Pada tahapan ini, menambahkan operator K-Means. melalui klik kanan -> *Insert Operator* -> Modeling -> Segmentation -> K-Means.



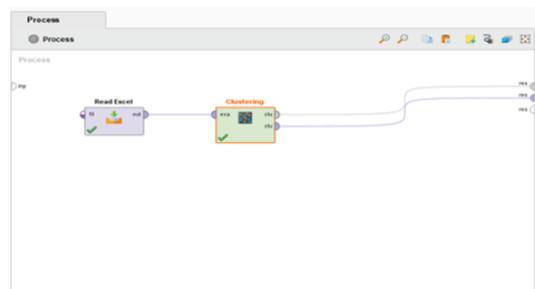
Gambar 4. Penambahan Operator K-Means

Berikutnya mengerjakan pengaturan pada menu Parameter *Clustering K-Means*, dilakukan pengaturan nilai k, dimana k mejadi nilai yang akan digunakan untuk penentuan jumlah *cluster* yang akan dibuat. Disini jumlah *cluster* yang akan dibuat yakni sebanyak 3 *cluster* (rendah, sedang, tinggi) yang terlihat pada Gambar 5.



Gambar 5. Jumlah Cluster

Bentuk pemodelan menjadi seperti berikut:



Gambar 6. Pemodelan Clustering dengan Operator K-Means

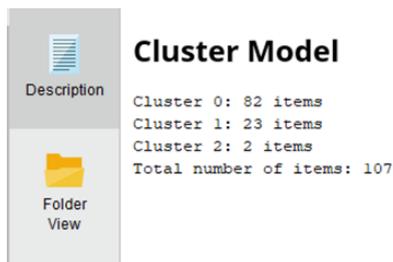
Dalam *ExampleSet* dapat dilihat beberapa hasil *cluster* yang terdapat pada *Data View* yang merupakan tampilan hasil *cluster* data yang telah di masukkan.

Row No.	Judul Buku	cluster	Total
1	101 Amazing...	cluster_0	1
2	9 Days Unrat...	cluster_0	1
3	Akuntansi Da...	cluster_0	2
4	Akuntansi Ma...	cluster_0	2
5	Akuntansi Pe...	cluster_0	2
6	Akuntansi Pe...	cluster_1	3
7	Akuntansi Sy...	cluster_0	1
8	Akuntansi Per...	cluster_0	2
9	Algoritma & P...	cluster_0	2
10	All About Cor...	cluster_0	1
11	Anak Rantau	cluster_0	1
12	Analisa Lapo...	cluster_1	6
13	Analisa Kritis...	cluster_1	7
14	Aneka Kue	cluster_0	1
15	Aplikasi Logi...	cluster_0	1

Gambar 7. *Data Example Set* / Hasil *cluster* pada *Data View*

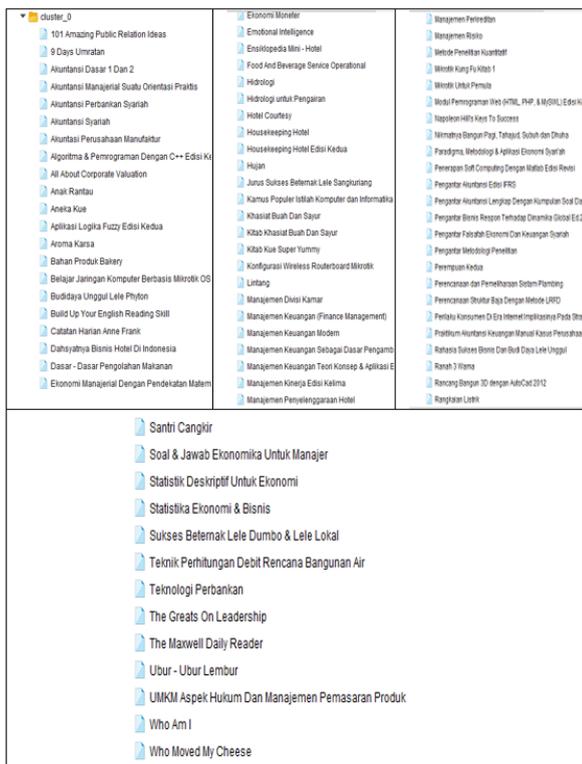
Pada Gambar 7 hasil data terurai menjadi 3 *cluster* yaitu *cluster 0*, *cluster 1*, *cluster 2* dengan setiap *cluster*

mengantongi hasil pengelompokkan data *cluster_0* terdiri 82 judul buku, *cluster_1* terdiri dari 23 judul buku, *cluster_2* terdiri dari 2 judul buku.



Gambar 8. Cluster Model

Berikut anggota dari *Cluster 0*



Gambar 9. Daftar Anggota Cluster 0



Gambar 10. Daftar Anggota Cluster 1



Gambar 11. Daftar Anggota Cluster 2

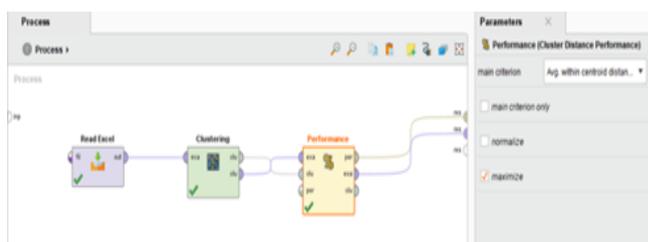
Dari hasil klusterisasi dengan $K=3$, terlihat bahwa pada *Cluster 2* (tinggi) merupakan kelompok buku yang paling diminati yang terdiri dari 2 judul buku yakni Teknologi Beton dan buku Teori dan Praktik hotel *Front Office* judul-judul buku yang paling banyak diminati untuk dipinjam selama kurun waktu 2019 hingga Juni 2022 sebanyak 14 dan 13 kali frekuensi peminjaman.

Cluster 1 (sedang) terdiri dari 23 judul buku yang merupakan buku-buku dengan frekuensi peminjaman sedang dengan kata lain dikategorikan diminati untuk dipinjam. Sedangkan *Cluster 0* (rendah) terdiri dari 82 judul buku dengan frekuensi peminjaman buku dalam kategori jarang dengan kata lain kurang diminati untuk dipinjam.

D. Performance

Pada tahapan ini, menambahkan operator K-Means. melalui klik kanan -> *Insert Operator* -> *Modeling* -> *Segmentation* -> K-Means.

Untuk *performance* klusterisasi dengan $K=3$ menggunakan nilai *Indeks Davies Bouldin* (DBI). Evaluasi dengan menggunakan *Davies Bouldin Index* ini memiliki skema evaluasi dari *internal cluster*, dimana baik atau tidaknya hasil *cluster* dilihat dari kuantitas dan kedekatan antar data hasil *cluster*.



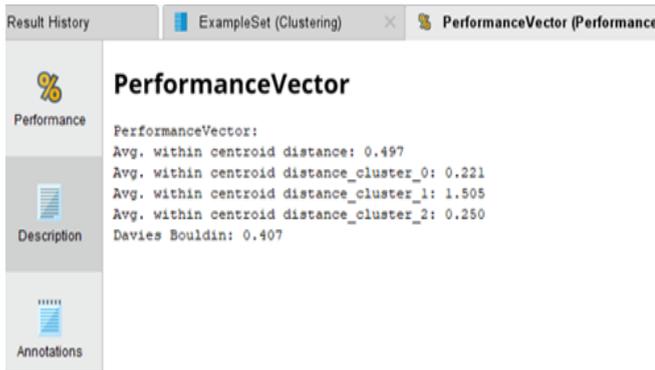
Gambar 12. Model Performance

Dari hasil *performance* didapatkan nilai *index Davies Bouldin* sebesar 0,407 seperti Gambar 13.

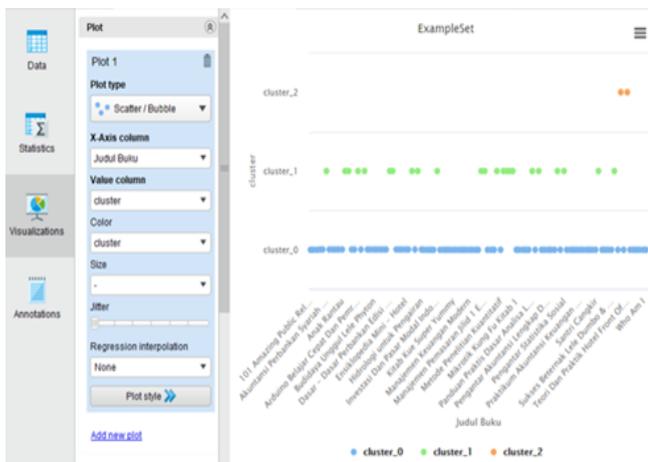


Gambar 13. Nilai Index Davis Bouldin

Dan untuk nilai *centroid* masing-masing cluster ditampilkan pada Gambar 14.



Gambar 14. Nilai *Centroid* dari setiap *Cluster*



Gambar 15. Plot *Cluster*

Gambar 15. memperlihatkan *plot cluster* secara *scattering* untuk $K=3$ (*cluster 0*, *cluster 1*, *cluster 2*)

V. KESIMPULAN

Dari hasil pengolahan data melalui Aplikasi Rapid Miner dengan $K=3$, didapatkan hasil *cluster 0* (rendah) terdiri 82 judul buku dengan frekuensi peminjaman bukunya dalam kategori jarang dengan kata lain kurang diminati untuk dipinjam, *cluster 1* (sedang) terdiri dari 23 judul buku yang merupakan buku-buku dengan frekuensi peminjaman sedang dengan kata lain dikategorikan diminati untuk dipinjam, *cluster 2* (tinggi) terdiri dari 2 judul buku merupakan kelompok buku yang paling diminati yang terdiri dari 2 judul buku yakni Teknologi Beton dan buku Teori dan Praktik *hotel Front Office*. Hasil performansi didapatkan nilai *index Davies Bouldin* sebesar 0,407. Pengelompokan data buku-buku ini dapat menjadi masukan bagi pengelola perpustakaan dalam pengadaan koleksi buku berdasarkan frekuensi peminjaman buku.

REFERENSI

- [1] Intan fitri andyni. (2013) Pengelompokan peminjam buku Dengan metode k-means Di perpustakaan pusat upn “veteran” jawa timur
- [2] Parlina, Iin. (2018). Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center. CESS (Journal of Computer Engineering System and Science), 3(1), 87–93.

- [3] Kusrini dan Lutfi, E.T. 2009. Alogoritma Data Mining. Yogyakarta: Andi Offset
- [4] Pramudiono, (2006), Apa Itu Data Mining?, [online], diakses tanggal 15 Maret 2022
- [5] Santosa, Budi. (2007). Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis. Yogyakarta: Graha Ilmu.
- [6] Sulaiman. (2020). Analisis pola belanja konsumen menggunakan algoritma k-means dan apriori pada haura swalayan. UIN Sultan Syarif Kasim Riau
- [7] Arai, K., dan Barakbah, A. R. (2007). Hierarchical k-means: an algorithm for centroids initialization for k-means. Reports of the Faculty of Science and Engineering, 36(1), 25–31.