Journal of Computer Science and Technology
Volume 3 ; Nomor 2 ; May 2025 ; Page 97-103
DOI : https://doi.org/10.59435/jocstec.v3i2.463
https://jurnal.padangtekno.com/index.php/jocstec
E-ISSN : 2985-4318
P-ISSN : 2985-5772

Pengelompokan Data Nilai Siswa Madrasah Ta'hiliyah Menggunakan Metode K-Means Clustering

Fahrillah[1*], Zaehol Fatah[2]

[1,2] Sistem Informasi, Universitas Ibrahimy
[1*]fahrillah74@gmail.com, [2]zaeholfatah@gmail.com

Abstrak

Data mining, atau penambangan data merupakan proses pengumpulan dan pengolahan data untuk mengekstrak informasi penting. Tahapan dalam proses data mining berguna untuk mencari sebuah pola tertentu dari data penilaian yang sangat banyak. Tujuan ini yaitu mengetahui dan membentuk cluster data siswa berdasarkan nilai sehingga menjadi sebuah cluster, sehingga hasil cluster siswa dapat menjadi acuan dalam meningkatkan nilai siswa dalam proses pembelajaran selanjutnya. Hasil evaluasi dan penilaian terhadap siswa dilakukan oleh tenaga pengajar atau guru dalam melakukan penilaian selama proses pembelajaran. Dalam proses pembelajaran terdapat 2 kategori penilaian yaitu nilai siswa UTS serta UAS. Hasil pengelompokan data nilai siswa menggunakan metode K-Means clustering menunjukan bahwa berdasarkan hasil cluster data siswa dalam satu semester, maka didapatkan cluster 0 berjumlah 7 siswa, cluster 1 berjumlah 3. Hasil pengujian menggunakan rapid miner maka terdapat 7 siswa yang memiliki nilai dengan rata-rata yang baik dan terdapat 3 siswa dengan rata-rata nilai yang kurang baik.

Kata Kunci : Clustering, Data Mining, K-Means, Pengelompokan

Abstract

Data mining, or data mining is the process of collecting and processing data to extract important information. The stages in the data mining process are useful for finding a particular pattern from a large amount of assessment data. This goal is to find out and form student data clusters based on grades so that they become a cluster, so that the results of student clusters can be a reference in improving student grades in the next learning process. The results of the evaluation and assessment of students are carried out by teaching staff or teachers in conducting assessments during the learning process. In the learning process there are 2 assessment categories, namely UTS and UAS student grades. The results of grouping student grade data using the K-Means clustering method show that based on the results of student data clusters in one semester, cluster 0 is obtained with 7 students, cluster 1 is 3. The results of testing using rapid miner show that there are 7 students who have grades with a good average and there are 3 students with a poor average grade..

Keyword : *Clustering, Data Mining, K-Means, Grouping*

PENDAHULUAN

Di era digital yang semakin berkembang, pengelolaan dan pemrosesan data menjadi salah satu aspek penting dalam berbagai bidang, termasuk dalam dunia pendidikan. Data yang terkumpul, seperti nilai siswa, dapat memberikan gambaran yang jelas mengenai pencapaian, kekuatan, dan kelemahan masing-masing individu dalam proses pembelajaran. Di Madrasah Ta'hiliyah, yang merupakan lembaga pendidikan dengan basis agama, pemanfaatan data untuk meningkatkan kualitas

pendidikan sangatlah krusial. Salah satu metode yang dapat digunakan untuk mengelompokkan data nilai siswa adalah metode *clustering*, khususnya K-Means Clustering.[1]

Pendidikan merupakan unsur terpenting dalam membentuk karakter dan budaya bangsa, terlebih bangsa yang memiliki masyarakat yang majemuk seperti halnya Indonesia. Pendidikan berfungsi untuk meningkatkan kemampuan manusia, dalam membentuk karakter serta peradaban bangsa yang bermartabat dalam rangka mencerdaskan kehidupan bangsa.[2] Sebagai bagian dari proses evaluasi pendidikan, penilaian terhadap hasil belajar siswa menjadi hal yang sangat penting untuk dilakukan. Salah satu cara untuk mempermudah dalam menganalisis dan mengelompokkan data nilai siswa adalah dengan menggunakan metode statistik dan algoritma tertentu, seperti K-Means Clustering. Metode K-Means adalah salah satu teknik dalam data mining yang digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik atau atribut tertentu, dalam hal ini adalah nilai siswa. Salah satu cara untuk memanfaatkan data nilai siswa secara efektif adalah dengan mengelompokkan siswa berdasarkan kesamaan kinerja akademik. Dalam hal ini, metode K-Means Clustering dapat digunakan untuk membagi siswa ke dalam beberapa kelompok berdasarkan pola-pola nilai yang diperoleh, baik itu nilai ujian, keterampilan, atau sikap.

Clustering adalah salah satu teknik dalam *machine learning* dan *data mining* yang digunakan untuk mengelompokkan sekumpulan objek atau data yang memiliki kemiripan atau kesamaan dalam karakteristik tertentu ke dalam kelompok-kelompok yang disebut kluster.[3] Metode K-Means Clustering adalah salah satu teknik dalam data mining yang digunakan untuk mengelompokkan data ke dalam sejumlah grup atau kluster berdasarkan kesamaan karakteristik yang dimiliki oleh data tersebut. Dalam konteks pendidikan, khususnya untuk nilai siswa, metode ini dapat digunakan untuk mengelompokkan siswa berdasarkan nilai yang diperoleh dalam ujian atau penilaian lainnya.[4]

Penelitian ini bertujuan untuk menerapkan metode K-Means Clustering pada data nilai siswa dengan tujuan utama untuk mengidentifikasi pola kinerja akademis siswa. Melalui penerapan K-Means Clustering, penelitian ini diharapkan dapat membagi siswa ke dalam beberapa kelompok berdasarkan kemiripan nilai sehingga pola-pola dalam kinerja akademis siswa dapat terlihat dengan jelas. Dengan begitu, pengelola pendidikan dapat mengidentifikasi kelompok siswa yang memiliki performa baik, sedang, atau rendah dalam bidang akademis.[5]

METODOLOGI PENELITIAN

Metode Pengambilan Data

Pada bagian ini merupakan cara yang dilakukan untuk mengumpulkan data terkait dengan penelitian yang dilakukan. Data yang dikumpulkan adalah data kuantitatif, yaitu nilai akademis siswa dari beberapa mata pelajaran. Data nilai diperoleh dari laporan hasil belajar semester siswa Maddrasah Ta'hiliyah Ibrahmy. Setelah data terkumpul, kemudian data di seleksi dan akan digunakan metode K-Means Clustering untuk mengelompokkan siswa berdasarkan nilai yang diperoleh menggunakan aplikasi Rapidminer. Berikut adalah data nilai siswa pada gambar di bawah ini:

Tabel 1. Data nilai siwa

NO	NAMA	UTS	UAS
1	AHMAD BAIHAQI	75	70
2	AHMAD HIDAYAT	60	50
3	ANDRIA PAHLOVI	70	60
4	ARDIANSYAH	50	75
5	FITRAH MAULANA ZACKY	40	65

6	KHAIRUL FAHMI TAMIMI	70	70
7	M. IMRON ROSADI	75	65
8	M. SANDI SOPIAN	65	70
9	MAULANA KASTURI	50	75
10	MOH. DAFID HALID	70	75

Pengelompokan

pengelompokan (atau klasifikasi) adalah suatu cara untuk mengelompokkan data atau objek berdasarkan karakteristik atau sifat-sifat tertentu yang dimilikinya. Pengelompokan ini bertujuan untuk memudahkan dalam menganalisis data secara lebih sederhana dan terstruktur. [6]
Penjelasan mendalam tentang pengelompokan dalam konteks analisis data. Pengelompokan di sini merujuk pada teknik statistik yang digunakan untuk mengelompokkan data ke dalam kelompok-kelompok yang disebut *cluster*. Setiap kelompok memiliki tingkat kemiripan tinggi antara anggota dalam satu kelompok dan perbedaan yang besar dibandingkan dengan kelompok lain.[7]

K-Means

K-Means adalah salah satu algoritma yang paling populer digunakan dalam analisis klaster. K-Means adalah teknik partisional clustering yang membagi data menjadi sejumlah k klaster (kelompok), di mana setiap data point (titik data) dimasukkan ke dalam klaster yang memiliki pusat (centroid) terdekat.[8]
K-Means adalah algoritma partisi, yang berarti membagi data menjadi sejumlah kelompok yang lebih kecil, di mana setiap kelompok memiliki data dengan karakteristik yang lebih serupa satu sama lain dibandingkan dengan kelompok lain.[9]

Data Mining

Data Mining adalah proses menggunakan teknik-teknik pembelajaran mesin untuk mengambil pola dan informasi yang berguna dari dataset besar yang tidak terstruktur atau yang memiliki kompleksitas tinggi. Hal ini melibatkan pengumpulan data, pemilihan fitur, pelatihan model, dan evaluasi model untuk menemukan pola atau membuat prediksi yang akurat.[10]
Data Mining muncul Tahun 1960-an - 1970-an Konsep dasar analisis data muncul, dengan fokus pada statistik dan analisis data. Pengembangan teknik dasar, seperti analisis regresi dan analisis multivariat. Tahun 1990-an Istilah "data mining" mulai dikenal secara luas. Penelitian dan pengembangan mulai berfokus pada teknik seperti clustering dan asosiasi. Penemuan pengetahuan dari basis data (Knowledge Discovery in Databases - KDD) menjadi fokus utama.[11]

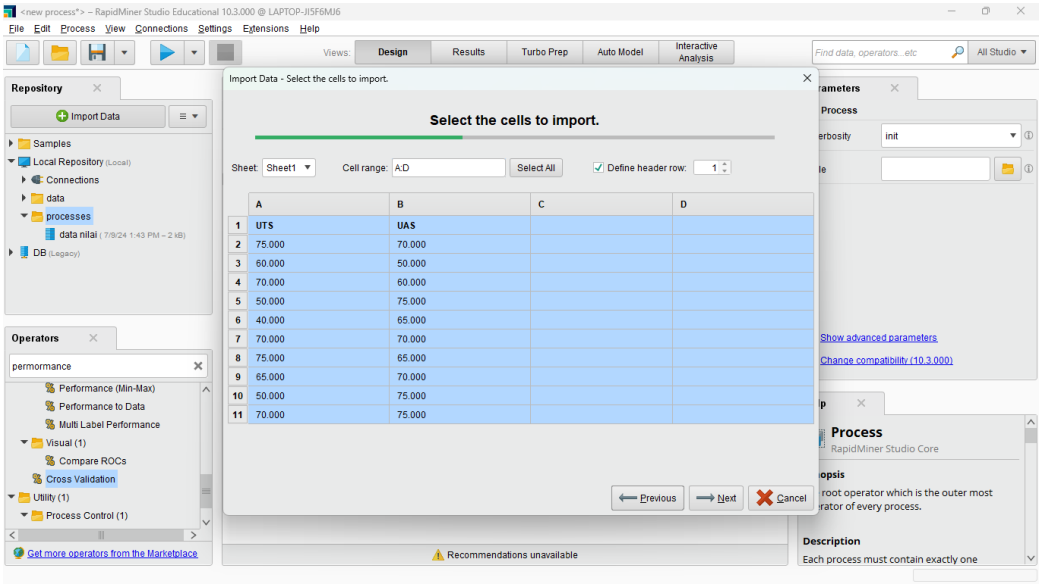
Clustering

Clustering adalah teknik dalam analisis data yang digunakan untuk mengelompokkan sekumpulan objek atau data sehingga objek dalam satu kelompok (cluster) memiliki kemiripan satu sama lain, sementara objek di kelompok yang berbeda memiliki perbedaan yang signifikan. Clustering sering digunakan dalam berbagai bidang, termasuk statistik, machine learning, dan analisis data.[12]
Clustering adalah teknik analisis data yang digunakan untuk mengelompokkan objek berdasarkan kesamaan atau jarak antar siswa. Tujuan utamanya adalah untuk memisahkan data ke dalam grup atau cluster sehingga objek dalam satu cluster lebih mirip satu sama lain dibandingkan dengan objek di cluster lain.[13]

HASIL DAN PEMBAHASAN

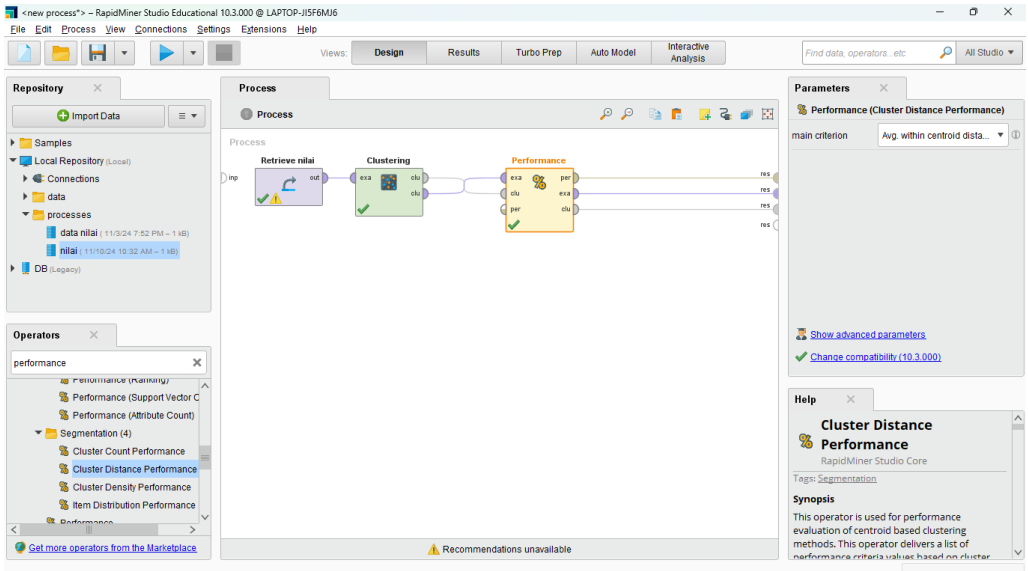
Penerapan Data Mining Menggunakan Rapidminer

Tahap awal ini proses pengolahan data di mulai dengan menginput dataset pada Rapidminer, dataset diatur dengan Nilai Uts Dan Uas memiliki tipe data interger seperti di bawah ini:



Gambar.1 Dataset Nilai

Setelah proses input dataset, proses selanjutnya akan memilih modul dengan menentukan nilai K awal ialah 2 dengan measure types menggunakan BregmenDivergences dikarenakan data yang diolah semuanya merupakan data numeric, kemudian menggunakan maximum perpindahan titik centroid maximum 100 kali sebelum dilakukan cluster harus menambah pengukuran performance untuk melihat nilai pembagian cluster yang akan terbagi menjadi 2.

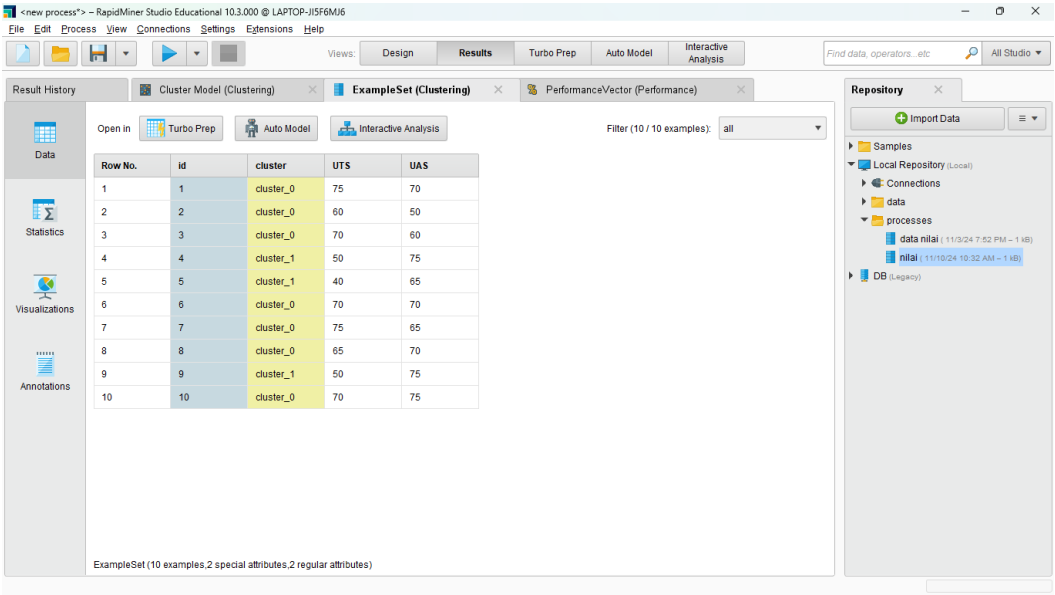


Gambar.2 Proses Data Mining

Hasil Algoritma K-Means

Data View

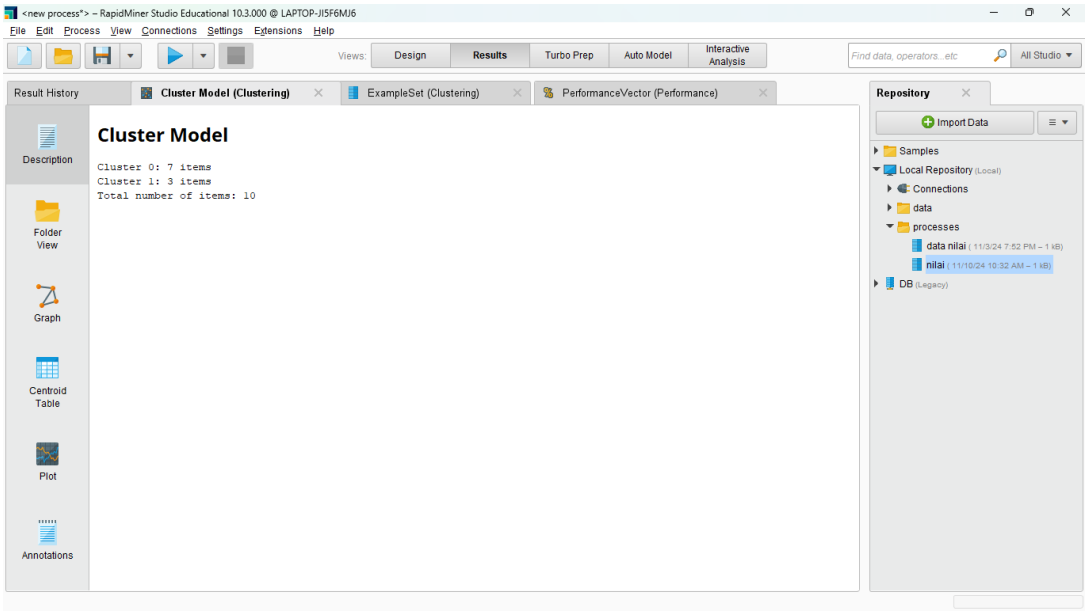
Setelah tahap proses K-Means Clustering menggunakan aplikasi Rapidminer Studio maka proses berkut ini adalah menampilkan hasil dari penerapan tersebut, pada gambar di bawah ini.



Gambar.3 Hasil Cluster

Cluster Model

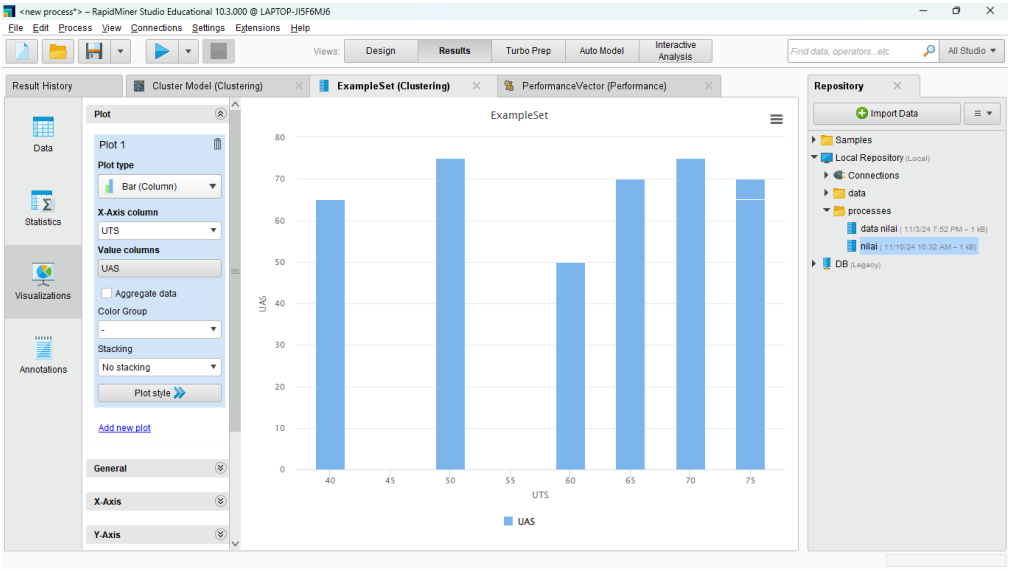
Pada gambar di bawah ini adalah hasil pengujian dataset yang berjumlah 10 item menggunakan Rapidminer sehingga ada 2 cluster. Pada cluster 0 terdapat 7 item dan cluster 1 terdapat 3 item.



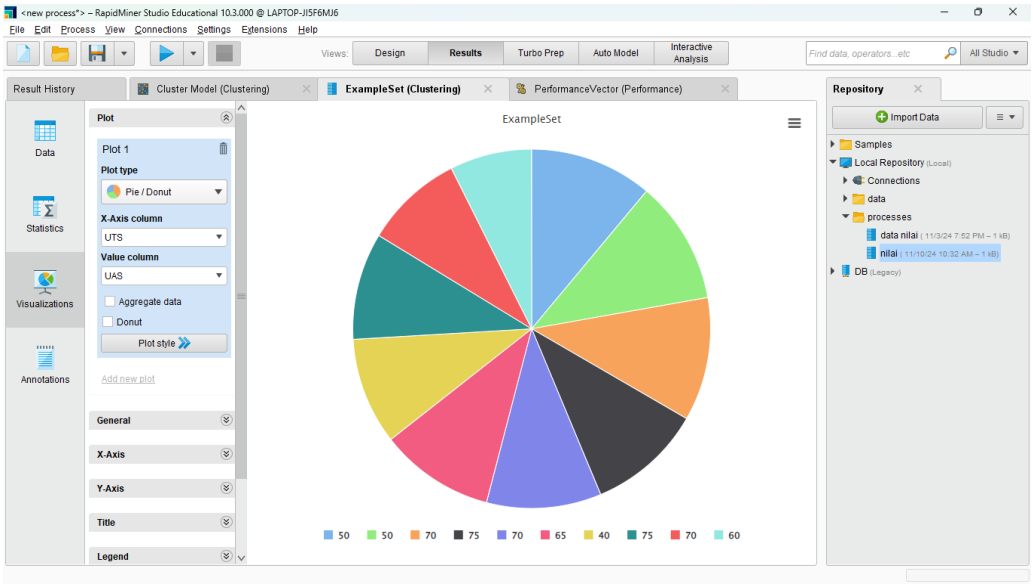
Gambar.4 Cluster Model

Visualization Hasil Clustering

Gambar di bawah ini merupakan tampilan grafik hasil pengelompokan data nilai siswa pada setiap nilai uas dan uts. Berikut ini tampilannya.



Gambar.5 Tampilan Bar (colomn)



Gambar.6 Tampilan Pie/Donut

KESIMPULAN

Berdasarkan Hasil perhitungan nilai akhir siswa menggunakan K-means dapat disimpulkan bahwa dari nilai siswa terdapat 2 cluster. Maka didapatkan cluster 0 berjumlah 7 siswa, cluster 1 berjumlah 3 siswa, dalam pembagian 2 penilain dan menunjukkan bahwa hanya terdapat 7 siswa yang memiliki nilai dengan rata-rata yang baik dan terdapat 3 siswa dengan nilai rata-rata yang kurang baik hal ini dapat menjadi kendala kedepan untuk nilai selanjutnya sehingga diperlukan strategi yang harus di lakukan untuk menurunkan nilai pada cluster 1 agar tidak menjadi kendala pada nilai selanjutnya.

UCAPAN TERIMA KASIH

Syukur Alhamdulillah, kami ucapkan kehadirat Allah SWT. Yang telah melimpahkan rahmat taufiq dan hidayah-Nya kepada kami, sehingga kami dapat menyelesaikan jurnal ini. Terimakasih sebesar-besarnya kepada dosen pembimbing yang telah meluangkan waktunya untuk membimbing saya dengan penuh kesabaran. Dan juga terimakasih kepada teman-teman yang selalu mengingatkan saya akan terselesainya jurnal ini.

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya penelitian ini.

DAFTAR PUSTAKA

[1] Manning, Raghavan, & Schütze, (2008).*dengan buku Introduction to Information Retrieval*. Cambridge University Press.
[2] moh. Syahrul Iskandar, Zaehol Fatah, "Implementasi Metode Algoritma K-Means Clustering Untuk Menentukan Penerima Program Indonesia Pintar (PIP)" Volume 2 ; Nomor 10 ; November 2024 ; Page 1-8
[3] Tan, Steinbach, & Kumar, (2006). *Introduction to Data Mining*. Pearson Addison-Wesley.
[4] Jain, (2010). *Data Clustering: 50 Years Beyond K-Means*. *Pattern Recognition Letters*, 31(8).
[5] Han, Kamber, & Pei. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
[6] Sudjana, (2010). *Statistik Deskriptif*. Bandung: Tarsito..
[7] Everitt, (2011). *Cluster Analysis*. 5th Edition. Chichester: Wiley.
[8] Han, Kamber, & Pei. (2011). *Data Mining: Concepts and Techniques* (3rd Edition). San Francisco: Morgan Kaufmann Publishers.
[9] Alpaydin. (2014). *Introduction to Machine Learning* (3rd Edition). Cambridge, MA: MIT Press.
[10] Ng, A. (2018). *Machine Learning Yearning*. Stanford: Self-published.
[11] Pang-Ning Tan, Michael Steinbach, & Vipin Kumar **(2006)**. *Introduction to Data Mining*. Pearson.
[12] "Pattern Recognition and Machine Learning" oleh Christopher Bishop.
[13] "Data Mining: Concepts and Techniques" oleh Jiawei Han, Micheline Kamber, dan Jian Pei.