

Penerapan Clustering dalam Data Science Untuk Mengembangkan Keterampilan Analitik di SMK Media Informatika

Application of Clustering in Data Science to Develop Analytical Skills at Media Informatics Vocational Schools

Eliyani¹, Muhammad Rifqi², Saruni Dwiasnati^{3*}

^{1,2,3}Fakultas Ilmu Komputer, Teknik Informatika, Universitas Mercu Buana

Correspondence: eliyani@mercubuana.ac.id, m.rifky@mercubuana.ac.id, saruni.dwiasnati@mercubuana.ac.id

Abstrak

Pasar kerja saat ini dan masa depan memerlukan tenaga kerja yang terampil dalam data science, termasuk teknik analitik seperti clustering. Dengan berkembangnya Industri 4.0, keterampilan dalam analisis data menjadi semakin penting, dan pendidikan harus menyesuaikan diri untuk memenuhi kebutuhan ini. Meningkatkan literasi data di kalangan siswa SMK sangat penting untuk mempersiapkan mereka menghadapi dunia yang semakin didominasi oleh data. Keterampilan analitik membantu siswa dalam memecahkan masalah kompleks dan membuat keputusan berdasarkan data. Keterampilan dalam data science mendorong inovasi dan kreativitas di berbagai bidang. Keterlibatan siswa agar dapat menggunakan alat dan keterampilan untuk menganalisis dan memahami dunia di sekitar mereka melalui data. Clustering adalah sebuah teknik dalam data science dan statistik yang digunakan untuk mengelompokkan objek-objek atau data point yang memiliki karakteristik serupa ke dalam kelompok atau cluster.

Kata kunci: Clustering, Data Science, SMK

Abstract

The current and future job market requires a workforce skilled in data science, including analytical techniques such as clustering. With the development of Industry 4.0, skills in data analysis are becoming increasingly important, and education must adapt to meet this need. Increasing data literacy among vocational school students is very important to prepare them to face a world increasingly dominated by data. Analytical skills help students in solving complex problems and making data-driven decisions. Skills in data science encourage innovation and creativity in various fields. Involvement of students so they can use tools and skills to analyze and understand the world around them through data. Clustering is a technique in data science and statistics that is used to group objects or data points that have similar characteristics into groups or clusters.

Keywords: Clustering, Data Science, Vocational school

1. PENDAHULUAN

Teknologi di era saat ini semakin berkembang dan memberikan berbagai dampak dalam kehidupan. Salah satu contoh perkembangan teknologi tersebut yaitu Artificial Intelligence (AI). Artificial intelligence adalah program komputer yang dapat meniru kecerdasan manusia, seperti pengambilan keputusan, penalaran, dan karakteristik manusia yang lain [1]. Perkembangan teknologi informasi dan komputer yang semakin pesat membawa dampak yang sangat besar dalam aktivitas kehidupan. Pada saat ini komputer banyak digunakan oleh masyarakat, baik lingkungan sekolah, kantor maupun umum. Pada era digital yang terus berkembang, keterampilan analitik telah menjadi kompetensi yang sangat penting dalam berbagai bidang, termasuk pendidikan, bisnis, dan industri. Kemampuan untuk menganalisis data, mengidentifikasi pola, dan mengambil keputusan berbasis data tidak hanya menjadi nilai tambah tetapi juga kebutuhan yang mendasar. Salah satu teknik yang signifikan dalam analisis data adalah clustering, sebuah metode dalam data science yang digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik. Dunia pendidikan di Indonesia telah mengalami banyak perubahan. Dimulai dengan pembelajaran yang berpusat pada guru (teacher-centered learning), berkembang menjadi pembelajaran yang berpusat pada peserta didik (students-centered learning). Pengabdian masyarakat adalah penerapan bidang pendidikan dan ekonomi untuk menyelesaikan masalah di masyarakat baik warga kota atau daerah maupun masyarakat industri dan pemerintah. Pengabdian masyarakat ini dilakukan dalam rangka menghasilkan karya monumental bagi masyarakat sekitar.

Data mining merupakan proses mencari informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi [2]. Data mining terbagi dalam beberapa task antara lain : asosiasi, klasifikasi, clustering dan sequence pattern. Data mining adalah analisis yang memeriksa kumpulan data untuk menemukan hubungan yang tidak terduga dan merangkum data dengan cara yang berbeda dengan cara yang berbeda dari sebelumnya, yang mudah dipahami dan berguna bagi pemilik data [3]. Data mining adalah proses menemukan hubungan baru yang bermakna, pola dan kebiasaan dengan menyortir sebagian besar data yang disimpan dalam media penyimpanan menggunakan teknologi pengenalan pola, seperti teknik statistik dan matematika [4]. Clustering adalah proses pengelompokan data data menjadi beberapa kelompok sehingga objek di dalam satu kelompok memiliki banyak kemiripan serta memiliki banyak perbedaan dengan objek dikelompok lain [5]. Salah satu cara untuk meningkatkan wawasan tentang AI di kalangan siswa SMK adalah melalui kegiatan workshop; Sebagaimana yang dilakukan untuk bidang lainnya [6] [7].

Salah satu domain dari AI adalah Data Science. Data Science juga telah banyak digunakan dalam kehidupan kita sehari-hari. Contoh penggunaan Data Science seperti peramalan cuaca [8], di mana cuaca yang diramal atau diprediksi

menggunakan data dari cuaca di hari-hari sebelumnya dengan tingkat akurasi yang tinggi. Contoh lain seperti penggunaan sistem rekomendasi pada Youtube, sistem rekomendasi penempatan barang di swalayan, dan banyak lagi.

SMK Media Informatika adalah sekolah SMK pertama di Jakarta Selatan yang bergerak dalam bidang Teknologi Informasi. SMK Media Informatika juga memiliki tag line “Sekolah Berbasis Project” dimana siswa di persiapkan untuk menghadapi dunia usaha dan dunia kerja melalui pembelajaran berbasis project. Saat ini SMK Media Informatika memiliki total 10 Laboratorium dan Ruang Praktek terdiri dari 3 Lab DKV, 2 Lab Broadcasting, 2 Lab PPLG, dan 3 Lab TJKT, dengan jumlah masing-masing 40 komputer dalam tiap Lab. Sehingga jumlah seluruhnya 400 unit komputer. Selain itu tersedia juga 1 Studio Siaran dan 1 Lab Bahasa. Didukung dengan perlengkapan lab lainnya seperti proyektor, printer, scanner dan ruangan yang semuanya dilengkapi dengan fasilitas AC.



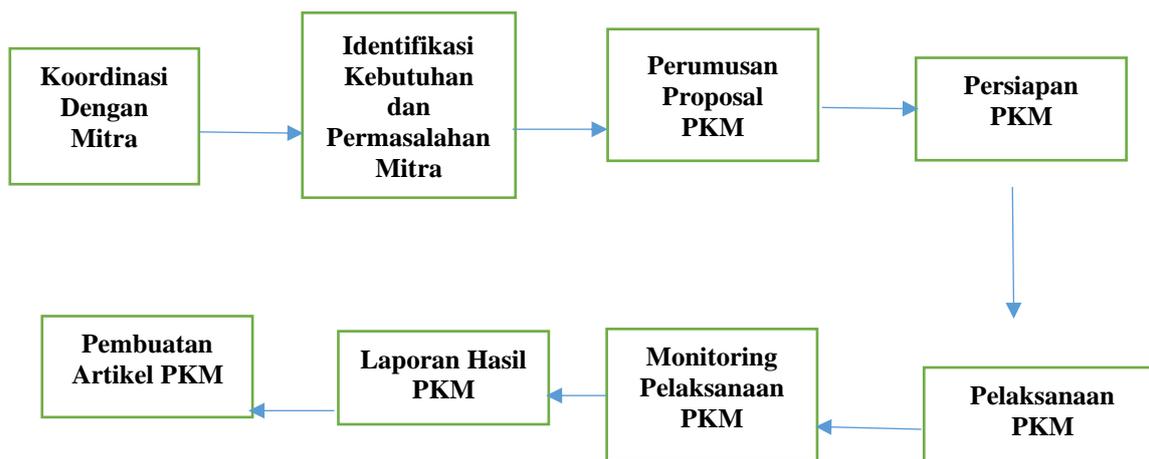
Gambar 1. Gedung SMK Media Informatika

Berbagai keuntungan lain juga akan didapatkan apabila suatu perusahaan menggunakan teknologi informasi dalam proses pengolahan dan analisis data mereka, seperti waktu pengolahan data dan pengambilan informasi yang lebih cepat sehingga pengambilan keputusan pada perusahaan tersebut juga menjadi lebih cepat. Peran yang dapat diberikan oleh aplikasi teknologi informasi ini adalah mendapatkan informasi untuk kehidupan pribadi seperti informasi tentang kesehatan, hobi, rekreasi, dan Rohani [9]. Hasil analisis yang diperoleh dengan menggunakan visualisasi dalam bentuk grafik manajemen sangat cepat dan mengoptimalkan pengolahan data sehingga mengetahui perkembangan keadaan database akademik [10]. Penelitian yang dilakukan oleh [11] menyatakan bahwa penulis akan memprediksi penjualan sepeda yang paling laris di provinsi tersebut dengan tujuan agar para pelaku usaha atau perusahaan penjual sepeda di wilayah

tersebut dapat lebih mengoptimalkan stok sepeda yang dimilikinya agar sesuai dengan target pasar konsumennya.

2. METODE

Langkah-langkah dalam melaksanakan solusi yang ditawarkan untuk mengatasi permasalahan mitra dari judul Penerapan Clustering dalam Data Science untuk Mengembangkan Keterampilan Analitik di SMK Media Informatika, sebagai berikut:



Gambar 2. Tahapan Pengabdian

Proses pelaksanaan Kegiatan PkM dapat dimulai dengan melakukan koordinasi dengan mitra dan mengidentifikasi kebutuhan dan permasalahan yang mitra hadapi. Kemudian, perumusan dari proposal PkM dibuat dan persiapan dilakukan dengan melakukan koordinasi antara mitra dan panitia yang terlibat. Setelah proses persiapan, kegiatan dilaksanakan sesuai dengan jadwal yang telah ditentukan oleh kedua belah pihak. Setelah kegiatan selesai, dilakukan proses monitoring dari pelaksanaan PKM, lalu membuat laporan dari hasil PKM yang dilakukan dan membuat artikel dari judul yang diambil.

A. Partisipasi Mitra dalam Pelaksanaan Program

Berikut adalah beberapa cara di mana mitra dapat berpartisipasi aktif:

1. Menyediakan ruang kelas atau laboratorium komputer yang memadai untuk pelatihan dan workshop.
2. Mengikutsertakan guru dalam pelatihan dan workshop untuk meningkatkan kapasitas mereka dalam mengajar materi terkait.
3. Mendorong dan memotivasi siswa untuk aktif berpartisipasi dan memanfaatkan program ini sepenuhnya.

- B. Evaluasi pelaksanaan program dan keberlanjutan program di Mitra
1. Mengumpulkan feedback dari siswa, guru, dan staf sekolah tentang pelatihan dan materi yang disampaikan.
 2. Melakukan survei atau wawancara untuk menilai tingkat pemahaman dan keterampilan siswa setelah mengikuti program.
 3. Mengadakan diskusi dengan pihak sekolah untuk membahas hasil evaluasi dan rencana tindak lanjut.
- C. Peran dan tugas dari masing-masing anggota tim sesuai dengan kompetensinya dan penugasan mahasiswa.

3. HASIL DAN PEMBAHASAN

Kegiatan yang akan dikemas kedalam tema besar oleh tim, yaitu “*Menggali Potensi Data Science dan Cybersecurity Untuk Siswa SMK Media Informatika*”, dalam pelaksanaannya dibagi menjadi lima materi pelatihan.

Tabel 1. Tema Kegiatan PKM

No	Kegiatan
<i>Menggali Potensi Data Science dan Cybersecurity Untuk Siswa SMK Media Informatika</i>	
1	Materi Pengenalan Phyton
2	Materi Pengenalan Metode Clustering
3	Materi Pengenalan Metode Klasifikasi
4	Materi Pengenalan Algoritma Naive Bayes
5	Materi Proses Pre-Processing
6	Materi Implementasi menggunakan Naive Bayes
7	Materi Visualisasi menggunakan Tableau
8	Materi Cybersecurity
9	Materi Pengenalan Algoritma Linear Regesi
10	Materi Pengenalan Algoritma K-Means & Implementasi

Pelaksanaan kegiatan dilakukan secara luring/offline pada SMK Media Informatika, kegiatan tersebut dibatasi hanya untuk 30 peserta.

Kegiatan PKM dilaksanakan dua tahap, yaitu tahap koordinasi dan tahap pelaksanaan. Pada tahap koordinasi, tim berkoordinasi dengan pihak mitra, untuk membahas permasalahan yang tengah dihadapi saat ini.

Berdasarkan hasil koordinasi disepakati kegiatan PKM ditujukan kepada Kepala Sekolah SMK Media Informatika dan disepakati juga untuk tanggal kegiatan, yaitu hari Kamis, Tanggal 29 Februari 2024.

Pada tahap pelaksanaan kegiatan PKM kegiatan dilakukan pada 1(satu) hari dimulai dari jam 09:00 – 15:30 WIB. Jadi untuk kegiatan PKM terbagi menjadi 10 materi yang disampaikan dalam 1(satu) hari, sesi kegiatan tersebut dapat terlihat pada tabel berikut:

Tabel 2. Jadwal Kegiatan Program kegiatan PPM yang Dilaksanakan

Hari, Tanggal	Jam	Aktivitas
Kamis 29 Februari 2024	09:00 - 09:20 WIB	• Pembukaan Oleh MC & Doa
		• Sambutan Kepala Sekolah
		• Sambutan Perwakilan Tim PKM
	09:20 - 09:50 WIB	Materi Pengenalan Phyton
	09:50 - 10:20 WIB	Materi Pengenalan Metode Clustering
	10:20 - 10:50 WIB	Materi Pengenalan Metode Klasifikasi
	10:50 - 11:20 WIB	Materi Pengenalan Algoritma Naive Bayes
	11:20 - 11:50 WIB	Materi Proses Pre-Processing
	11:50 - 12:20 WIB	Materi Implementasi menggunakan Naive Bayes
	12:20 - 13:00 WIB	Ishoma
13:00 - 13:30 WIB	Materi Visualisasi menggunakan Tableau	
13:30 - 14:00 WIB	Materi Cybersecurity	
14:00 - 14:30 WIB	Materi Pengenalan Algoritma Linear Regesi	
14:30 - 15:00 WIB	Materi Pengenalan Algoritma K-Means & Implementasi	
15:00 - 15:30 WIB	Quiz	

Untuk pelaksanaan kegiatan kelompok kami mendapat giliran pada jam 13:30 - 14:00 WIB. Fasilitas kegiatan berupa proyektor, ruang lab komputer yang berada dilingkungan SMK Media Informatika.



Gambar 3. Foto Kegiatan PKM



Gambar 4. Foto Penyerahan Piagam PKM



Gambar 5. Foto Bersama Setelah Kegiatan PKM

Potensi dan kemampuan penyerapan dari materi yang dipaparkan, terlihat sangat baik, para peserta mengerti apa yang telah dipaparkan oleh narasumber.

Keberhasilan pelaksanaan program pengabdian masyarakat ini dapat dilihat dari tolak ukur sebagai berikut:

Tabel 3. Tolak Ukur Keberhasilan Pelatihan

Respons positif dari peserta PKM	Respons peserta PKM diukur dengan hasil evaluasi yang di isi oleh para peserta PKM.
----------------------------------	---

Sebelum melakukan pelatihan dilakukan pre test dengan jumlah responden 30 orang, hasil dari pre test tersebut terlihat pada gambar 6, Mengerti Sekali 13,3%, Mengerti 60%, Biasa 23,3% dan Tidak Mengerti 3,4%.



Gambar 6. Pre Test Sebelum Pelatihan

Setelah melakukan pemaparan materi dilakukan post test dengan jumlah responden 28 orang, hasil dari post test terlihat pada gambar 7, Mengerti Sekali 14,3%, Mengerti 64,3%, Biasa 21,4% terlihat untuk peserta pelatihan yang tidak mengerti tidak ada lagi.



Gambar 7. Post Test Setelah Pelatihan

Secara keseluruhan materi yang disampaikan oleh pemateri dapat dimengerti dan dipahami oleh peserta kegiatan PKM tersebut. Kegiatan PKM ini menambah pengetahuan bagi peserta kegiatan yang terdiri dari siswa yang berada dilingkungan SMK Media Informatika.

4. KESIMPULAN

Penerapan metode clustering seperti K-Means, Hierarchical Clustering, DBSCAN, dan Gaussian Mixture Models (GMM) membantu siswa memahami struktur dan pola dalam data. Dengan mengelompokkan data ke dalam cluster yang bermakna, siswa dapat lebih mudah mengidentifikasi karakteristik dan hubungan antara berbagai atribut data.

DAFTAR PUSTAKA

- [1]. Harkut, Dinesh G, Kashmira Kasat, and Vaishnavi D Harkut. 2019. "Artificial Intelligence." In Chapter, IntechOpen, 1-5. <https://www.intechopen.com/chapters/66147>.
- [2]. Sowmya, R., Suneetha, K.,R " Data Mining with Big Data". International Conference on Intelligent Systems and Control (ISCO) Vol 10, pp.1109, Februari 2017.
- [3]. D. Diy, "Analisis Data Mining Untuk Memprediksi Lama Perawatan Pasien Covid-19 Bianglala Informatika," vol. 10, no. 1, pp. 21-29, 2022.
- [4]. S. Hendrian, "Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan," Fakt. Exacta, vol. 11, no. 3, pp. 266-274, 2018, doi: 10.30998/faktorexacta.v11i3.2777
- [5]. Alam Jusia, P., Muhammad Irfan, F., & Dinamika Bangsa Jambi Jl Jend Sudirman Thehok Jambi, S. (2019). Clustering Data Untuk Rekomendasi Penentuan Jurusan Perguruan Tinggi Menggunakan Metode K-Means. Jurnal IKRA-ITH Informatika, 3(3), 75.
- [6]. Slamet, S., Pratikno, H., dan Maulana Y. M., "Workshop JARKOM Berbasis CISCO dan MIKROTIK untuk Persiapan Uji Kompetensi Keahlian (UKK) bagi Guru dan Murid di SMK KARTIKA 1 Surabaya."Share: Journal of Service Learning, 2021, 7(1), 1-7.
- [7]. Margasari, N, Sholikhah, Z., Andhini, M. M., Fitrianna, H., "Peningkatan Literasi Digital untuk Membentuk Jiwa Student-preneurship pada Siswa Sekolah Menengah Atas (SMA) di Yogyakarta," Darma Sabha Cendekia, 2020, 2(2), 32-39.
- [8]. Hiran, Kamal Kant, Deepak Khazanchi, Ajay Kumar Vyas, and Sanjeevikumar Padmanaban. 2021. Machine Learning for Sustainable Development. eds. Kamal Kant Hiran, Deepak Khazanchi, Ajay Kumar Vyas, and Sanjeevikumar Padmanaban. De Gruyter
- [9]. Cecep Abdul Cholik. Perkembangan Teknologi Informasi Komunikasi / ICT Dalam Berbagai Bidang. Jurnal Fakultas Teknik. e-ISSN:2746-220X. p-ISSN: 2746-1209 Vol. 2 No. 2 Mei 2021.

- [10]. Dedy Hartama. 2018. Analisa Visualisasi Data Akademik Menggunakan Tableau Big Data. Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK). Volume (3). Juli 2018, pp 46-55
- [11]. Anggriawan, I., & Gunawan, W. (2022). Implementation of Data Mining Using K-Means Algorithm for Bicycle Sales Prediction. ILKOM Jurnal Ilmiah, 14(3), 284–293. <https://doi.org/10.33096/ilkom.v14i3.1291.284-293>