

## BAB II

### TINJAUAN PUSTAKA

Penelitian yang dilakukan penulis berhubungan dengan penelitian sebelumnya. Penelitian tentang *clustering* menggunakan metode *K-means* pada evaluasi dosen di situs siatma UAJYs belum pernah dilakukan. Para peneliti terdahulu sudah banyak melakukan *clustering* menggunakan metode *K-Means*. Penelitian ini melakukan pengelompokan data evaluasi dosen UAJY menggunakan metode *K-means*. Hasil dari pengujian ini, diharapkan dapat dijadikan sebagai bahan untuk Universitas Atma Jaya Yogyakarta dalam mengembangkan proses pembelajaran dan juga dapat menemukan solusi dari kendala selama pembelajaran.

Penelitian tentang *clustering* menggunakan metode *K-means* dilakukan oleh Musthafawi, dkk tentang “Analisis Respon Emosi Marah Wanita Jawa Dengan Algoritma *K-Means Clustering*” pada tahun 2017. Pada proses pengujian *clustering* menggunakan metode *K-Means* yang dimana pengujian tersebut menggunakan 30 data yang sudah dipilih. Dalam menentukan *centroid* awal untuk melakukan proses pengujian mendapatkan hasil tingkat akurasi *clustering* yang pada *cluster* satu sebesar 100%, untuk *cluster* dua sebanyak 100%, dan yang terakhir pada *cluster* tiga sebanyak 100%. Berdasarkan pengujian *cluster* yang menghasilkan 3 kelompok memiliki tingkat akurasi pada sistem sebanyak 96,6%. Pada hasil penelitian ini dapat mengetahui faktor-faktor apa saja yang mempengaruhi respon emosi marah wanita Jawa dan dapat memberikan masukan bagi yang bersangkutan untuk mengembangkan strategi yang tepat dalam menangani masalah respon emosi marah wanita Jawa [5].

Penelitian selanjutnya yang dilakukan oleh Chandra, dkk penelitian yang berjudul “Pemanfaatan metode *K-Means Clustering* dalam penentuan penjurusan Siswa SMA” pada tahun 2014. Proses penelitian *clustering* menggunakan metode *K-Means* sehingga dengan menggunakan algoritma

tersebut dapat mengelompokan data siswa dimana data tersebut digunakan untuk membuat keputusan dalam menentukan penjurusan siswa yang berada pada tingkat SMA. Dari proses pengujian mendapatkan hasil yang paling baik pada proses *preprocessing* dalam *clustering* IPA dengan tingkat akurasi sebesar 0,905882 dengan ketepatan pada hasil dari pengujian sebesar 0,876923 dan untuk kesuaian hasil negatif terhadap actual negatif sebanyak 0,714285. Sebaliknya pada kelompok IPS menghasilkan akurasi 0,905882. Sehingga perbandingan hasil pada proses *clustering* terbaik membuktikan kalau tidak terdapat siswa yang masuk ke dalam kedua jurusan ataupun tidak masuk ke dalam kedua jurusan, hal tersebut disebabkan pada proses *clustering* memikirkan nilai dari kriteria akademik. *Intelligence Quotient* (IQ) serta minat siswa juga bisa pengaruhi hasil dalam proses *clustering* yang bisa menghindari kemungkinan siswa dapat diterima IPA dan IPS ataupun dapat ditolak semua jurusan [6].

Penelitian yang dilakukan oleh Aulia, dkk yang berjudul “Penerapan Algoritma *K-Means* dalam Proses *Clustering* Penilaian Kinerja ASN” pada tahun 2021, Penelitian ini bertujuan untuk dapat mengoptimalkan kinerja pegawai, sehingga hasil penelitian menampilkan kalau penerapan *data mining* dalam analisis evaluasi kinerja ASN untuk menetapkan besaran TPP di Sekretariat DPRD Pematangsiantar dapat membantu dalam mengoptimalkan kinerja ASN. Hasil proses penelitian *clustering* menggunakan metode *K-Means* mempunyai nilai yang sama, ialah menjadi tiga *cluster* yang pertama *cluster* yang paling sebanyak satu anggota, *cluster* yang sedang sebanyak sembilan anggota, serta untuk *cluster* yang rendah sebanyak 13 anggota. Penelitian ini diharapkan bisa menjadi masukan untuk lembaga yang terkait dan pegawai [4].

Penelitian selanjutnya yang dilakukan oleh Sandra, dkk yang berjudul “*Clustering* Kualitas Kinerja Karyawan Pada Perusahaan Bahan Kimia Menggunakan Algoritma *K-Means*” pada tahun 2021, penelitian ini memberikan hasil yang positif bagi perusahaan. Dengan mengelompokan karyawan ke dalam kelompok yang memiliki kualitas kinerja yang sama

maka perusahaan dapat mengetahui bagian-bagian yang memerlukan peningkatan dan memberikan focus yang lebih pada peningkatan kualitas kinerja karyawan di bagian tersebut dan juga dapat membantu dalam proses pengambilan keputusan. Pada penelitian ini melakukan *clustering* pada evaluasi kinerja karyawan yang memiliki variabel nilai kualitas kerja, nilai tanggung jawab, nilai kerjasama, nilai kehadiran, dan nilai disiplin. Berdasarkan hasil yang didapat pada *clustering* ini memiliki tingkat akurasi sebesar 0,42% untuk *cluster* satu dengan kelompok sangat produktif yang memiliki 16 anggota, pada *cluster* dua memiliki tingkat akurasi sebesar 0,47% dengan anggota sebanyak 18, dan yang terakhir untuk *Cluster* ketiga memiliki tingkat akurasi sebesar 0,11% dengan anggota sebanyak empat. Pada penelitian ini memiliki batasan yang disebabkan oleh hanya menggunakan lima variabel [7].

Penelitian Dwitri, dkk yang berjudul berjudul “Penerapan Algoritma *K-means* dalam Menentukan Tingkat Penyebaran Pandemi *Covid-19* di Indonesia” pada tahun 2020, hasil dari penelitian ini akan dapat diketahui wilayah yang memiliki tingkat penyebaran yang lebih tinggi atau rendah, sehingga dapat dilakukan tindakan yang lebih cepat dan efektif. Selain itu, pada penelitian ini juga dapat membantu dalam mengidentifikasi faktor yang mempengaruhi tingkat penyebaran *covid-19* di Indonesia, sehingga selama pandemi dapat melakukan upaya yang lebih tepat. Penelitian ini dilakukan agar kondisi pandemi yang berada di Indonesia dikarenakan *covid-19* dapat mengurangi dalam penyebaran pandemi saat ini di Indonesia dengan menganalisa berdasarkan seluruh daerah di Indonesia yang terkena *covid-19*. Dari hasil pengelompokan yang sudah dilakukan terdapat daerah yang menjadi pusat penyebaran *covid-19* dan merupakan daerah dengan kasus meninggal terbanyak adalah berada di provinsi DKI Jakarta [8].

Pada yang dilakukan oleh Sugiono, dkk penelitian berjudul “Pengelompokan Perilaku Mahasiswa Pada Perkuliahan *E-Learning* dengan *K-Means Clustering*” pada tahun 2019, tujuan dari penelitian ini berfungsi untuk mengetahui perilaku mahasiswa ketika sedang melakukan perkuliahan

secara *E-Learning* dengan melakukan *clustering* dan perhitungan nilai jarak mahasiswa dengan riwayat aktivitas mahasiswa tersebut. Sehingga hasil pada penelitian ini berdampak bagi Universitas yang dapat mengetahui tingkat keterlibatan dan partisipasi mahasiswa dalam perkuliahan *E-Learning* dan juga dapat membantu dalam menentukan strategi peningkatan kualitas perkuliahan *E-Learning* yang dapat memberikan pengalaman belajar yang lebih baik bagi mahasiswa. Dalam proses *clustering* ini menggunakan metode *K-Means* yang mendapatkan hasil *cluster* sebanyak tiga, sehingga pada *cluster* satu memiliki anggota sebanyak 53, *cluster* dua memiliki anggota sebanyak 42, dan *cluster* tiga sebanyak 14 dan jika ditotal data yang diolah sebanyak 109 [9].

Penelitian selanjutnya yang dilakukan oleh Budiarto dengan topik “Penerapan Metode *K-Means* dalam Proses *Clustering* Penilaian Prestasi Kerja Pegawai Negeri Sipil” pada tahun 2017, penelitian ini memiliki tujuan untuk dapat menjamin objektivitas pembinaan PNS yang sudah tercatat dalam sistem prestasi kerja. Penelitian ini akan memberikan informasi mengenai tingkat prestasi kerja pegawai negeri sipil yang dapat digunakan sebagai pertimbangan dalam pemberian promosi atau hadiah bagi pegawai yang memiliki prestasi kerja yang baik dan juga memberikan bimbingan dan pelatihan bagi pegawai yang memiliki prestasi kerja yang kurang baik. Berdasarkan hasil pengujian yang menggunakan 100 data, membutuhkan sebanyak 10 iterasi untuk menyelesaikannya membutuhkan waktu 7 menit dan menghasilkan lima *cluster*, yaitu *cluster* dengan jumlah nilai sangat baik sebanyak 20 anggota, *cluster* dengan nilai baik sebanyak 31 anggota, *cluster* dengan nilai cukup sebanyak 10 anggota, dan untuk *cluster* dengan jumlah nilai buruk sebanyak 26 anggota [10].

Penelitian Susilowati, dkk yang berjudul “Diagnosa Penyakit Kanker Payudara Menggunakan Metode *K-means Clustering*” Pada tahun 2019. Pada penelitian ini dilakukan *clustering* berujuan untuk dapat membantu dalam melakukan diagnosa pada pasien yang mengalami penyakit kanker payudara dengan kategori jenis penyakit jinak dan ganas dengan metode *K-Means*,

sehingga dapat membantu dalam mengelompokkan penyakit kanker payudara menjadi beberapa kelompok yang memiliki tingkat resiko yang sama, dapat membantu dalam menentukan tindakan dalam pengobatan yang tepat dan juga dapat membantu dalam faktor yang mempengaruhi terjadinya penyakit kanker payudara. Setelah melakukan pengelompokan mendapatkan hasil yang dapat menentukan penyakit yang dialami oleh pasien merupakan termasuk tergolong kanker payudara jinak atau ganas [11].

Penelitian yang dilakukan oleh Sugianto, dkk yang berjudul “Algoritma *K-means* untuk Pengelompokkan Penyakit Pasien pada Puskesmas Cigugur Tengah” pada tahun 2020. Penelitian ini bertujuan untuk dapat mempermudah dan membantu tenaga medis dan kesehatan dalam memberikan informasi kepada masyarakat. Dari hasil melakukan pengelompokan menggunakan metode *K-Means* mendapatkan informasi bahwa penyakit yang dialami oleh pasien di Puskesmas paling diderita adalah penyakit akut dengan nilai akurasi yang lumayan tinggi [12].

**Tabel 2. 1 Perbandingan Penelitian**

<b>Peneliti (Tahun)</b>	Chandra, dkk (2014) [6]	Musthafawi, dkk (2017) [5]	Sugiono, dkk (2019) [9]	Dwitri, dkk (2020) [8]	Aulia, dkk (2021) [4]	Penulis (2023)
<b>Judul</b>	Pemanfaatan metode <i>K-Means Clustering</i> dalam penentuan penjurusan Siswa SMA	Analisis Respon Emosi Marah Wanita Jawa Dengan Algoritma <i>K-Means Clustering</i>	Pengelompokan Perilaku Mahasiswa Pada Perkuliahan <i>E-Learning</i> dengan <i>K-Means Clustering</i>	Penerapan Algoritma <i>K-Means</i> Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 di Indonesia	Penerapan Algoritma <i>K-Means</i> dalam Proses <i>Clustering</i> Penilaian Kinerja ASN	<i>Clustering</i> Evaluasi Dosen Atma Jaya Yogyakarta Menggunakan Metode <i>K-Means</i>
<b>Metode</b>	<i>K-Means</i>	<i>K-Means</i>	<i>K-Means</i>	<i>K-Means</i>	<i>K-Means</i>	<i>K-Means</i>
<b>Objek Penelitian</b>	Siswa SMA	Wanita Jawa	Mahasiswa	Lokasi penyebaran pandemic Covid di Indonesia	ASN	Dosen

<b>Hasil dan Kesimpulan</b>	Hasil dari perbandingan <i>clustering</i> menunjukkan bahwa tidak ada siswa yang diterima di kedua jurusan yaitu IPA dan IPS atau siswa ditolak di keduanya	Terdapat 3 kelompok <i>cluster</i> berdasarkan kelompok responden tidak pemaarah, agak pemaarah, dan pemaarah. Pada kelompok responden agak pemaarah dan pemaarah terdapat aspek yang menjadi pemicu	Terdapat 3 kelompok <i>cluster</i> yaitu mahasiswa dengan jumlah aktivitas dan nilai tinggi, mahasiswa dengan aktivitas sedang dan nilai tinggi, dan mahasiswa dengan aktivitas sedikit dan nilai rendah	Terdapat 3 kelompok <i>cluster</i> yaitu penyebaran dalam jangka besar sebesar 5056 kasus, penyebaran dalam jangka menengah sebesar 4525 kasus, dan penyebaran jangka kecil sebesar 4043 kasus	Terdapat 3 kelompok cluster yaitu pegawai dalam pekerja termasuk sangat baik, baik, cukup, kurang atau buruk. Jumlah anggota tiap <i>cluster</i> sebanyak 20, 31, 10, 13, dan 26 pegawai	Terdapat 4 kelompok cluster yaitu kinerja dosen sangat baik, baik, cukup baik, kurang baik. Jumlah anggota tiap cluster sebanyak 7, 18, 9, dan 1 kelas kuliah.
<b>Tools</b>	<i>WEKA</i>	-	<i>WEKA</i>	<i>Rapid Miner</i>	<i>RapidMiner</i>	<i>RapidMiner</i>