

BAB II. LANDASAN TEORI

2.1 Studi Literatur

Studi literatur merupakan bagian yang akan membahas tentang penyelesaian masalah yang akan memberikan jalan keluarnya. Dalam hal ini akan dikemukakan beberapa teori-teori yang berkaitan dengan masalah yang diangkat.

2.1.1 Penelitian Terdahulu

Penelitian (Asroni & Adrian, 2015) yang berjudul “Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang” meneliti suatu kasus yang ada di Kampus UMM Magelang yaitu menentukan peserta yang akan dikirimkan untuk perlombaan berdasarkan data yang terdapat pada *data warehouse* Universitas Muhammadiyah Magelang. Dengan menerapkan metode *K-Means* dalam melakukan pengelompokan peserta yang cocok sesuai dengan kriteria yang ada untuk dijadikan peserta lomba diharapkan dapat memberikan rekomendasi. Hasil yang didapatkan berdasarkan penelitian, metode ini dapat digunakan untuk mengelompokkan mahasiswa berdasarkan IPK dan beberapa atribut mata kuliah.

Penelitian (Rosadi & dkk, 2020) tentang pemilihan program prioritas dana desa dengan menerapkan metode *K-Means Clustering* dalam kasusnya. Permasalahan yang terjadi pada penelitian ini yaitu, kurang efektif, profesional dan efisien terhadap pengelolaan dan pemangku kebijakan di desa. Selain itu, dalam menentukan program kerja desa sering muncul kecurigaan-kecurigaan masyarakat terhadap pemerintah desa. Masyarakat beranggapan bahwa pemerintah desa tidak adil dalam melakukan rumusan program kerja. Untuk mencapai efektifitas dalam pengelolaan keuangan desa maka dari itu peneliti menerapkan metode tersebut di dalam sistem agar dapat pembagian program kerja menjadi tiga tingkat prioritas, tingkat kebutuhan tinggi sedang, rendah dan tinggi. Hasil yang didapatkan pada penelitian, metode ini cocok digunakan untuk menghasilkan keputusan yang dijadikan acuan atau bahan pertimbangan dalam merancang program kerja desa.

Penerapan *Clustering* pada Aplikasi Pendeteksi Kemiripan Dokumen Teks Bahasa Indonesia. Teknik *clustering* ini diterapkan di sebuah sistem agar mudah

digunakan dan dapat mempersingkat waktu pengukuran kesamaan. Aplikasi ini digunakan untuk melakukan *check plagiarisme* agar mengurangi tindakan *plagiarisme* yang dapat merugikan orang pertama yang mempunyai pikiran dalam menuliskan karya tulis. Peneliti menyarankan bagi pengembang aplikasi agar menggunakan teknik lainnya agar hasil yang didapatkan lebih akurat (Kambey, Sangkey, & Jacobus, 2020).

Penelitian yang berjudul “Penerapan Data Mining pada Presentasi Penerimaan Imunisasi Anak-Anak menurut Provinsi Menggunakan *K-Means Clustering*” yang dilakukan oleh (Amri, Hartama, & Windarto, 2020) menggunakan metode *K-Means* untuk melakukan pengelompokan wilayah dalam kasus kurangnya pemberian imunisasi terhadap beberapa wilayah karena kondisi yang jauh dari daerah perkotaan. Berdasarkan penelitian yang dilakukan menyatakan bahwa hasil yang didapatkan metode ini cocok diterapkan dalam ilmu kesehatan salah satunya terhadap kasus ini.

Penelitian yang dilakukan (Fauzia, Dermawan, & Padilah, 2020) yang berjudul “Penerapan *K-Means Clustering* pada Penyakit Infeksi Saluran Pernapasan Akut (ISPA) di Kabupaten Karawang” meneliti tentang situasi dalam perlakuan Dinas Kesehatan Karawang terhadap seluruh puskesmas sama, tidak ada sama sekali penelitian khusus terhadap salah satu puskesmas yang berada di Karawang. Maka dari itu, Dinas Kesehatan Karawang harus bertindak tegas dalam melakukan penanganan yang efektif untuk penanganan penyakit Infeksi Saluran Pernapasan Akut. Untuk mendapatkan perhatian khusus dari Dinas Kesehatan Karawang harus menggunakan metode untuk mendapatkan daerah persebaran penyakit ISPA di Kabupaten Karawang. Sehingga, dapat dilakukan penanganan yang efektif dan dapat membantu dalam mengambil suatu kebijakan terhadap penyakit ISPA di Kabupaten Karawang. Penerapan metode *K-Means* dalam melakukan *clustering* mendapatkan hasil yang optimal dan lebih akurat dalam pengelompokan data.

Penelitian berjudul “Implementasi Algoritma *K-Medoids* dan *K-Means* untuk Mengelompokkan Wilayah Sebaran Cacat pada Anak” meneliti permasalahan tentang dampak penolakan di sebuah lingkungan dan masalah sosial lainnya yang menimpa seseorang yang menderita cacat, sehingga mendorong dari berbagai pihak untuk

melakukan proses pencegahan sedini mungkin. Pertumbuhan dan pengembangan anak juga dapat terganggu saat usia dibawah 18 tahun. Maka dari itu, di ambil langkah awal yang dapat dilakukan yaitu dengan memperkirakan wilayah mana saja yang menderita cacat tertinggi. Dengan memanfaatkan persebaran data penderita cacat, maka akan dilakukan proses pengelompokan sesuai dengan informasi yang dimiliki oleh data, sehingga dapat diketahui wilayah-wilayah yang memiliki persebaran penyandang cacat tertinggi. Penelitian kali ini yang dilakukan oleh (Marlina, Putri, Fernando, & dkk, 2018) menggunakan teknik *clustering* untuk menentukan wilayah tersebut. Teknik *clustering* dengan memanfaatkan gabungan dua metode yaitu *K-Medoids* dan *K-Means* memberikan hasil bahwa metode *K-Medoids* mampu melakukan pengelompokan data persebaran anak cacat yang memiliki nilai validitas 0.5009 sedangkan metode *K-Means* mendapatkan nilai validitas sebesar 0.1443. Berdasarkan hasil dari validitas tersebut, maka menunjukkan bahwa metode *K-Medoids* lebih baik dalam melakukan pengelompokan data sebaran anak cacat daripada menggunakan metode *K-Means*.

Penelitian (Asminatun, Wakhidah, & Putri, 2020) yang berjudul “Penerapan Metode K-Medoids Untuk Pengelompokan Kondisi Jalan di Kota Semarang” penelitian yang menggunakan metode K-Medoids dalam penelitiannya dari aplikasi roadroid didapatkan data sejumlah 638 dengan memiliki attribute antara lain nama jalan, prosentase kondisi jalan baik, sedang, rusak ringan dan juga rusak berat. Data yang dikelompokkan oleh peneliti tersebut berhasil dikelompokkan menjadi 4 cluster yang mana dapat menunjukkan tingkat frekuensi jalan kondisi baik, sedang, rusak ringan dan juga berat. Berdasarkan hasil dari pengujian kualitas *cluster* dengan menggunakan *silhouette efficient* bahwa jumlah *cluster* 2 merupakan *cluster* yang kualitasnya paling baik, dikarenakan hasilnya paling mendekati angka

1. Untuk analisis pengelompokan data dari kondisi jalan rusak lebih baik menggunakan *cluster*
2. Sehingga data yang dihasilkan dari penerapan metode tersebut yaitu K-Medoid akan semakin mendekati ketepatan pengelompokan data kondisi jalan.

Pada penelitian selanjutnya tentang “Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia”.

Penelitian yang dilakukan oleh (Sindi & et al., 2020) mendapatkan data yang ditampung pastinya banyak sekali, apalagi seluruh wilayah Indonesia. Berdasarkan data yang diperoleh, peneliti dapat melihat dari pola-pola penentuan pengelompokan dari penyebaran penyakit *covid-19* yang dilakukan berdasarkan nilai test. Berdasarkan pengelompokan dalam penelitian tersebut menunjukkan hasil bahwa masyarakat yang berasal dari berbagai wilayah Indonesia yang memiliki ciri-ciri antara lain suhu badan diatas 36,9°C disertai demam dan juga batuk yang sampai berkelanjutan, menunjukkan bahwa salah satu ciri-ciri dari gejala penyakit *covid-19*.

Penelitian selanjutnya mengenai “Penerapan Metode K-Medoids Clustering Pada Penanganan Kasus Diare Di Indonesia” yang dilakukan oleh (Hardiyanti, Tambunan, & Saragih, 2019) mendapatkan kesimpulan bahwa penerapan data mining yang menggunakan algoritma *K-Medoids* untuk pengelompokan penanganan kasus diare menurut provinsi. Data uji yang diperoleh peneliti disini yaitu sebanyak 34 provinsi dengan menggunakan 2 *cluster*. *Cluster* yang pertama yaitu (nilai rendah) sebanyak 31, *cluster* kedua yaitu (nilai tinggi) sebanyak 3. Hasil yang diperoleh berdasarkan penelitian tersebut yang diperoleh dalam melakukan meng-*cluster* jumlah penanganan kasus diare yang ada di Indonesia pada tahun 2017 menggunakan metode ini adalah jumlah provinsi yang termasuk kedalam *cluster* pertama (nilai rendah) yang memiliki jumlah 31 provinsi sebagai berikut Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat dan Papua. Sedangkan, yang termasuk kedalam *cluster* kedua yaitu (nilai tinggi) terdapat 3 provinsi, yaitu Jawa Barat, Jawa Tengah dan Jawa Timur. Hasil yang diperoleh oleh peneliti tersebut sama dengan hasil yang telah di implementasikan ke dalam software *rapidminer* versi 5.3.

Penelitian selanjutnya yaitu penelitian milik (Defiyanti, Jajuli, & W , 2017). Penelitian ini melakukan pengelompokan data mahasiswa penerima beasiswa

berguna untuk menentukan mahasiswa yang berhak, dipertimbangkan atay tidak berhak. Dengan pengelompokan mahasiswa penerima beasiswa ini dapat memudahkan pihak tata usaha dalam menentukan penerima beasiswa khususnya beasiswa BBM. Pengelompokan tersebut dalam dilakukan dengan teknik *clustering* berbasis partisi atau dengan algoritma *K-Medoids*. Data-data yang didapat untuk dilakukan pengelompokan terdiri dari atribut SKS, IPK, Tanggungan Orang Tua dan Jumlah Penghasilan Orang Tua. Dari data-data yang didapat memiliki nilai yang beragam dan memiliki rentang satu dengan lainnya yang berjauhan. Dengan itu, pada penelitian ini terbentuklah tiga buah pengelompokan untuk data pengajuan beasiswa BBM yang terdiri dari kelompok penerima, kelompok dipertimbangkan menerima dan kelompok penerima beasiswa. Berdasarkan hasil evaluasi yang telah dilakukan dengan menghitung nilai *Cubic Clustering Criterion* (CCC) didapat bahwa dataset dengan kodefikasi keseluruhan data menempati predikat terbaik dalam keseragaman dalam pengelompokan dengan nilai 2,245 yang menempati predikat baik. Hal ini disebabkan oleh keseragaman data dari setiap atribut dan juga didukung dengan nilai *RMS Std Deviation* yang kecil.

Penelitian yang dilakukan oleh (Oktarina, 2020) bertujuan untuk mengkaji penggerombolan data besar menggunakan metode *K-Means* dan *K-Medoids*. Hasil kajian simulasi menunjukkan bahwa secara umum tidak bisa dibedakan antara kedua metode tersebut. Penggerombolan untuk data Twitter dilakukan berdasarkan pelabelan sentiment yang dilakukan secara manual yang bersifat objektif dengan kriteria pelabelan sebagai berikut, label pertama menunjukkan *tweet* yang berpihak pada Jokowi, label kedua menunjukkan *tweet* yang berpihak pada Prabowo dan label *tweet* ketiga yang menunjukkan opini lainnya. Metode *K-Medoid* disini menutupi kekurangan *K-Means* yaitu *noise*. Cara untuk mengatasi *noise* tersebut yaitu dengan melakukan pra pemrosesan yang dilakukan yaitu mentransformasi data tidak terstruktur menjadi data terstruktur sehingga analisis lebih lanjut dapat dilakukan. Dari hasil yang diperoleh terlihat bahwa ada gerombol yang memihak pada dua label. Hal ini terlihat dari tidak ada persentase yang mayoritas. Sedangkan gerombol yang menghasilkan nilai persentase yang lebih dari 50% pada pelabelan menunjukkan bahwa gerombol tersebut memihak pada label tersebut.

2.2 Dasar Teori

2.2.1 Wanita Jawa

Negara Kesatuan Republik Indonesia ini merupakan negara yang kaya akan suku dan budaya. Salah satu suku terbesar di Indonesia yaitu suku Jawa. Suku Jawa disini memiliki ciri khas pada wanita yang memiliki sifat dasar penurut, setia, lemah lembut, menjunjung tinggi nilai keluarga, mampu mengerti dan memahami orang lain serta sopan. Hal tersebut, dapat dilihat dari bagaimana sikap wanita tersebut menghargai suami, tidak banyak menuntut dan juga mematuhi suami. Seorang wanita jawa disini sangat pintar dalam mengatur situasi dan kondisi di segala kondisi apapun, mereka akan tetap menghormati dan menghargai suami meskipun kelakuan dari suami sudah keterlaluan, Bukan karena suami saja yang dapat mempengaruhi kemarahan seseorang wanita jawa. Tetapi, juga dapat dipengaruhi oleh lingkungan sekitarnya, yaitu anak, orang tua, saudara bahkan di lingkup kerja. Maka dari itu wanita jawa disini dikenal dengan wanita hebat dan kuat. Karena, bagi orang jawa khususnya seorang wanita, tidak sepatasnya menunjukkan emosi secara berlebih apalagi hingga menunjukkan konflik (Musthafawi, 2017).

2.2.2 Respon Emosi

Berbagai jenis respon dalam emosi antara lain yaitu respon emosi positif (emosi yang menyenangkan) dan respon emosi negative (emosi yang tidak menyenangkan). Respon emosi yang positif merupakan emosi yang dapat menimbulkan perasaan positif pada orang yang mengalaminya yakni dampak yang menyenangkan dan menenangkan. Sedangkan emosi yang respon yang menimbulkan perasaan negatif terhadap orang yang mengalaminya, dampaknya antara lain marah, benci, takut, dan sebagainya.

Emosi merupakan pengalaman yang efektif disertai dengan penyesuaian batin secara menyeluruh yang aman keadaan mental dan fisiologi seseorang sedang mengalami kondisi yang meluap-luap dan juga dapat terlihat dengan tingkah laku yang nyata. Emosi juga merupakan salah satu dari aspek yang memiliki pengaruh besar terhadap sikap manusia. Prinsip dari emosi itu sendiri yaitu menggambarkan perasaan dari seseorang dalam menghadapi berbagai macam situasi yang berbeda-beda. Wanita cenderung lebih emosional dan lebih penuh perasaan daripada laki-

laki dikarenakan seorang laki-laki lebih dominan rasional dan menggunakan logika. (Rosiani, Permatasari, & Yunhasnawa, 2018).

Untuk respon emosi orang jawa tentunya berbeda, melihat pulau jawa terdapat tiga provinsi, antara lain Jawa Timur, Jawa Tengah dan Jawa Barat. Pembawaan diri wanita jawa selalu terlihat lebih tenang dan dapat terkendali emosionalnya (Rosiani, Permatasari, & Yunhasnawa, 2018). Respon emosi dari orang jawa tengah itu sendiri cenderung lebih tidak terlihat, dikarenakan karakteristik dari sifat sikap wanita Jawa Tengah yang memang mampu menyembunyikan situasi emosi ketika marah. Sedangkan wanita Jawa Timur, cenderung kurang bisa menutupi emosi, itu semua dapat terlihat dari nada wanita Jawa Timur saat berbicara.

Berdasarkan berbagai macam penyebab dari kemarahan seorang wanita jawa tentunya berbeda-beda di setiap individunya. Itu semua tergantung dari segi kondisi dan lingkungan sekitar. Berikut beberapa aspek yang mempengaruhi kemarahan terhadap wanita jawa, aspek pertama suami, aspek kedua anak, aspek ketiga orang tua, aspek keempat saudara dan aspek pekerjaan (Musthafawi, 2017)

Berdasarkan uraian dari beberapa aspek diatas, maka dapat ditarik sebuah kesimpulan bahwa kemarahan wanita jawa disini, dapat disebabkan oleh gangguan untuk pencapaian suatu tujuan dan juga melakukan beberapa aktivitas. Selain itu, kemarahan sosok wanita jawa juga dapat dipicu oleh ucapan dari orang lain yang dapat menyakiti hati seorang wanita jawa, ancaman fisik dan perilaku dari seseorang lain yang dapat menimbulkan *psychological insultment* (Musthafawi, 2017)

2.2.3 Metode K-Means Clustering

K-Means merupakan suatu algoritma yang digunakan dalam pengelompokan secara partisi yang memisahkan data kedalam kelompok yang berbeda-beda. Pada penelitian yang sebelumnya dilakukan oleh (Musthafawi, Rosiani, & Yunhasnawa, 2017) dengan menggunakan metode ini sebagai algoritma untuk menentukan respon emosi wanita jawa. Metode *K-Means Clustering* ini merupakan metode yang mana diminta untuk menentukan jumlah klaster nya terlebih dahulu dan ditetapkan pusat klaster nya, lalu dilakukan proses perhitungan nilai rata-rata dari setiap data

yang diperoleh. Jarak data dengan klaster tentunya berbeda-beda, maka dari itu dilakukanlah proses perhitungan jarak data ke klaster dengan menggunakan *Euclidean Distance*. Lakukan proses perhitungan jarak tadi secara menerus hingga banyaknya data. Melakukan alokasikan semua data ke dalam klaster terdekat. Caranya yaitu dengan melihat dari nilainya. Jika nilai jarak klaster 1 lebih kecil daripada jarak ke klaster lainnya maka langsung masuk ke klaster 1. Melakukan penentuan pusat *cluster* baru berdasarkan klaster pertama hingga banyaknya *cluster*, dengan cara nilai hasil dibagi dengan banyak hasil. Lakukan proses perhitungan jarak pusat klaster baru, caranya sama seperti tadi dengan menggunakan *Euclidean Distance*, proses perhitungan jarak tadi secara menerus hingga banyaknya data dan alokasikan semua data ke dalam klaster terdekat. Caranya yaitu dengan melihat dari nilainya. Setelah itu, melakukan perbandingan antara hasil klaster pertama dan kedua, jika hasil yang diperoleh antara pertama dan kedua tidak bergeser atau berubah, maka klaster sudah benar.

2.2.4 Metode K-Medoid

K-Medoid merupakan salah satu metode yang mana digunakan ke dalam *clustering*. Metode ini mirip dengan metode *K-Means*, yang membedakan disini antara *K-Means* dan *K-Medoid* adalah metode *K-Medoid* menggunakan object sebagai perwakilan pusat *cluster* untuk setiap *cluster*. Sedangkan *K-Means* dia menggunakan nilai rata-rata (*Mean*) sebagai pusat cluster. *K-Medoid* memiliki kelebihan untuk mengatasi kelemahan pada algoritma *K-Means* yang sensitif terhadap *noise* dan *outlier*. Berikut, langkah-langkah metode ini :

1. Cleaning data : menghapus data-data yang tidak relevan / tidak valid.
2. Transformasi data : mengubah data text menjadi numeric.
3. Normalisasi data : Mencari minimum dan maximum dari setiap data yang sudah di numeric tadi. Rumusnya normalisasi data :
$$=(\text{data} - \text{minimum keseluruhan data}) / (\text{maximum keseluruhan data} - \text{minimum keseluruhan data})$$
.
4. Tentukan jumlah *cluster* yang diinginkan. Misal $K = 2$. Pilih secara acak sebanyak K dari n data.

5. Lalu ambil acak medoid nya. Misal object ambil data 18 dan 19.
6. Hitung jarak masing-masing object dengan object sementara dengan metode euclidean distance. Hasilnya (cost 1). Untuk cari cost 2 sama.
7. Tandai jarak terdekat object ke medoid, lalu hitung totalnya. Misal :
 $= \text{minimum}(\text{jarak awal dan jarak akhir})$ ":". = kedekatan. Mencari total
 $= \text{sum}(\text{jumlah kedekatan})$.
8. Penentuan anggota cluster terhadap medoid sementara. $= \text{if}(\text{cost1} - (\text{cost 1} : \text{cost 2}); 1; 2)$. Maka akan didapatkan anggota cluster.
9. Lakukan iterasi medoid. (Ulangi langkah 4-8).
10. Hitung total simpangan. Misal : selisih antara total cost medoid baru dengan total cost medoid lama.
11. Jika mendapatkan hasil percobaan pengurangan sudah lebih dari 0 maka hasil clustering mencari medoid berhenti.

2.2.5 PHP

PHP kepanjangan dari *Hypertext Preprocessor* merupakan salah satu bahasa pemrograman *open source* yang sangat cocok atau di khususkan untuk pengembangan *web* dan dapat ditanamkan pada sebuah skripsi HTML. Bahasa PHP dapat dikatakan menggambarkan beberapa bahasa pemrograman, diantaranya yaitu : *C*, *Java*, dan *Perl* serta yang mudah untuk dipelajari.

PHP juga merupakan bahasa *scripting server – side*, dimana pemrosesan datanya dilakukan pada sisi *server*. Secara singkat, *server*lah yang akan menerjemahkan skrip program, yang kemudian hasilnya akan dikirim kepada *client* yang melakukan permintaan.

Adapun pengertian PHP lain yaitu akronim dari *Hypertext Preprocessor*, yakni suatu bahasa pemrograman berbasis kode-kode (*script*) yang digunakan untuk mengolah suatu data dan mengirimkannya kembali ke *web browser* menjadi kode HTML (Firman, Wowor, & Najoran , 2016)

