

dengan 105 kematian, dan tertinggi kedua terjadi di provinsi Jawa Tengah dengan jumlah kematian terjadi sebanyak 92. Berdasarkan informasi dari Kepala Bidang Pencegahan dan Pengendalian Penyakit Dinas Kesehatan (Dinkes) Jateng, Tatik Muharyati pada Rabu (11/3/2020) di Semarang menyatakan terdapat 17 kasus kematian yang disebabkan oleh DBD berasal dari 13 daerah di Jateng. Dan dalam 13 kasus tersebut dapat dirinci dengan 2 orang meninggal dunia yang ada di kota Semarang dan Kabupaten Banyumas. Adapun yang lainnya 1 atau 2 orang meninggal dunia di wilayah Kabupaten Cilacap, Banjarnegara, Purbalingga, Batang, Tegal, Kebumen, Karanganyar, Rembang, Jepara, Temanggung, dan [3]. Melihat demam berdarah adalah penyakit yang berhubungan dengan lingkungan yang menyebar relatif cepat. Gerakan pencegahan dipertimbangkan hal paling perlu dilakukan, oleh karena itu diperlukan sistem peringatan dini yang lebih efisien [4]. Salah satu yang dilakukan adalah dengan melakukan suatu Gerakan satu kesatuan yang menjadi program nasional, biasa disebut dengan adanya Pemberantasan Sarang Nyamuk atau PSN dimana PSN ini terfokus dengan 3M Plus aktifitas yang mendukung adanya pemberantasan pada sarang-sarang nyamuk sebagai sumber penyakit DBD. Dimana 3M Plus ini terdiri atas:

- 1) Menguras atau membersihkan wadah atau tempat yang biasanya sering dijadikan sebagai tempat penampungan air bersih/ tempat air menggenang didalamnya diantaranya seperti bak di kamar mandi, ember berisi air, tempayan air minum, penampung air lemari es, sela-sela dari ban bekas, wadah pot tanaman dan lain-lain.
- 2) Menutup dengan rapat wadah-wadah penampungan air contohnya seperti kendi, toren air, drum dan lain sebagainya.
- 3) Memanfaatkan kembali atau dengan kata lain mendaur ulang berbagai jenis barang bekas yang memiliki potensi sebagai tempat perkembangbiakan jentik nyamuk yang menyebabkan DBD.

Adapun untuk plus pada 3M adalah kegiatan-kegiatan yang dilakukan untuk pencegahan DBD lainnya, seperti :

- 1) Menaburkan bubuk abate di tempat/ wadah penampungan air yang sulit untuk dibersihkan
- 2) Memakai obat nyamuk atau anti nyamuk pada ruangan-ruangan tertentu.
- 3) Menggunakan kelambu saat tidur
- 4) Memelihara ikan di kolam/ akuarium sebagai pemakan jentik nyamuk
- 5) Menanam tanaman yang memiliki fungsi sebagai pengusir nyamuk
- 6) Mengatur cahaya dan sirkulasi udara melalui ventilasi dalam rumah
- 7) Menghindari kebiasaan menggantung baju atau pakaian lainnya yang selesai dipakai

karena biasanya menjadi tempat bersarang nyamuk, dan lain- lain.

Meningkatnya kasus DBD suatu wilayah, dapat dikarenakan beberapa faktor salah satunya adalah gerakan PSN yang tidak dilaksanakan dengan maksimal dan berkualitas. Contohnya ketika di lapangan, Desa Pandak, Kecamatan Baturaden, Kabupaten Banyumas yang paling banyak dijumpai penderita DBD, banyak terdapat tumpukan bambu. Dan, saat PSN, hal tersebut luput dari perhatian warga. Untuk itu pada saat kegiatan PSN tidak hanya berdasarkan kegiatan dilapangan saja namun juga terdokumentasi, salah bentuk usaha pendokumentasian gerakan PSN ini adalah yang dilakukan oleh Kelurahan Sumampir, Kabupaten Banyumas.

Handwritten form titled "HASIL PEMERIKSAAN JENTIK Aedes Aegypti DALAM RANGKA PSN 0SD TAHUN 2020". The form lists various household containers (e.g., Bak Kamar Mandi, Tempayan, Peralatan Berair) and their status (e.g., Terlihat, Tidak Terlihat). It includes a signature and date at the bottom.

Gambar 2. Contoh Form Hasil PSN per-masing-masing Kepala Keluarga di Kelurahan Sumampir Kabupaten Banyumas (Sumber: Kelurahan Sumampir)

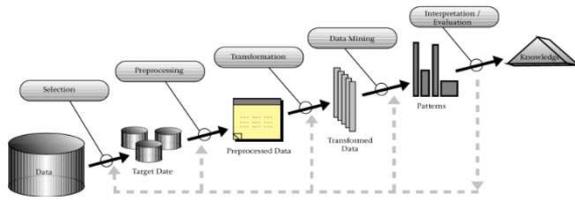
Dengan adanya dokumen tersebut diharapkan akan ada suatu dokumentasi yang bisa untuk diolah kedepannya, walaupun pada kenyataannya dokumen tersebut ternyata hingga sekarang ini belum ada pengolahan lebih lanjut, penelitian inipun akhirnya akan dilakukan dengan tujuan agar pemakai penelitian dapat melihat pengelompokan data penyebaran sarang jentik nyamuk berdasarkan jumlah container yang dimiliki oleh masing-masing kepala keluarga dan jumlah jentik nyamuk yang terlihat pada masing-masing container, selain itu pemakai juga dapat mengetahui keakuratan algoritma k-means dalam pengelompokan data penyebaran sarang jentik nyamuk di Kelurahan Sumampir. Hasil penelitian ini akan bermanfaat bagi masyarakat terutama untuk lebih memperhatikan penyebaran sarang jentik nyamuk yang ada pada barang-barang disekitar rumah.

METODE

A. Metode Penelitian

Adanya data yang terkumpul dalam jumlah banyak dan belum terolah, maka dapat digunakan sebagai bahan penelitian yang akan memberikan nilai/ manfaat lebih bagi data itu sendiri atau dalam hal ini umum disebut dengan data mining. Data

mining juga bisa diartikan dengan proses yang melibatkan suatu kecerdasan buatan, teknik statistik, matematika, dan machine learning yang berguna untuk mencari tahu lebih dalam dan menguraikan suatu informasi yang berguna dan pengetahuan terkait berdasarkan berbagai macam database yang besar [5].



Gambar 3. Data mining and Knowledge Discovery Process of Database
 (Sumber: Silwattananusarn & Tuamsuk in Sano [6])

- 1) **Selection**
 Selection digunakan untuk memilih variabel yang akan digunakan agar tidak terjadi kesamaan dan iterasi yang tidak diperlukan pada pengolahan data mining.
- 2) **Preprocessing**
 Pada tahap preprocessing memiliki dua tahap, yaitu:
 - a. Data Cleaning, menghapus data yang memang tidak digunakan seperti dalam menangani missing value, dan noise data yang menangani data-data yang tidak konsisten dan relevan.
 - b. Data Integration, dilakukan pada atribut yang mencari entitas yang unik.
- 3) **Transformation**
 Merubah data dengan bentuk ekstensi yang benar untuk pengolahan data mining, yang dikarenakan pada berbagai metode data mining membutuhkan format khusus agar dapat diproses dalam data mining.
- 4) **Data mining**
 Bagian yang paling penting dalam metode yang dipakai dalam menerima pengetahuan baru dari data yang telah diproses. Dalam penelitian ini menerapkan teknik cluster yakni menggunakan K-Means Clustering.
- 5) **Evaluation/ Interpretation**
 Memahami pola yang unik ke dalam dasar pengetahuan yang akan diidentifikasi. Dalam tahap ini, akan dihasilkan bentuk yang khas ataupun model prediksi yang dipakai untuk menilai kajian yang ada sudah memenuhi sasaran yang dicapai.
- 6) **Knowledge**
 Pola yang didapatkan nantinya akan disajikan kepada pengguna. Pada tahap ini, pengetahuan baru yang didapatkan sanggup dipahami oleh semua orang yang dijadikan acuan dalam pengambilan keputusan.

Clustering termasuk dalam satu metode data mining [7] dan clustering sudah menjadi bagian dari instrumen yang valid guna memecahkan permasalahan yang kompleks dalam ilmu komputer dan statistik. Clustering merupakan tahap pengelompokan titik-titik data menjadi 2 kelompok atau lebih sehingga titik-titik tersebut yang ada dalam kelompok sama akan lebih mirip dibandingkan kelompok titik yang berbeda, hanya menurut informasi yang ada dalam poin data [8].

K-Means termasuk dalam salah satu metode pengelompokan data non-hierarki (sekatan) yang bisa membagi data menjadi dua kelompok atau lebih. Metode itu akan membagi data ke dalam kelompok dimana data yang memiliki karakteristik sama akan dimasukkan dalam satu kelompok yang sama, sedangkan data yang memiliki karakteristik tidak sama akan dimasukkan dalam kelompok lainnya. Tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang akan di atur dalam proses pengelompokan, biasanya diusahakan untuk meminimalkan variasi dalam satu kelompok dan memaksimalkan variasi dari antar kelompok.

Proses dasar algoritma k-means dapat dilihat di bawah ini [7]:

- 1) Menentukan jumlah cluster yang akan dibuat dan menetapkan pusat cluster k.
- 2) Menggunakan jarak euclidean lalu hitung tiap data ke pusat cluster.

$$d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \dots \dots \dots (1)$$

- 3) Kelompokkan data pada cluster dengan jarak paling pendek dengan persamaan

$$M_{in} \sum_k^k = d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \dots \dots \dots (2)$$

- 4) Hitung pusat cluster baru dengan menggunakan persamaan

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \dots \dots \dots (3)$$

Dengan :
 $x_{ij} \in$ Kluster ke - k
 p = banyaknya anggota kluster ke - k

Ulangi tahap 2) hingga 4) sehingga tidak terdapat lagi data yang dapat berpindah ke cluster yang lain.

B. Penelitian Terkait

1. Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru [8]. Tujuan dari penelitian ini adalah untuk mengelompokkan data obat-obatan pada Rumah Sakit Umum Daerah Pekanbaru yang bisa dipakai sebagai acuan dalam pengambilan keputusan pada perencanaan dan pengendalian pasokan medis pada rumah sakit tersebut. Hasil dari penelitian tersebut menghasilkan kelompok obat dengan pemakaian sedikit, rata-rata permintaan obat tiap tahunnya kurang dari

18000 buah, kelompok obat dengan pemakaian sedang rata-rata permintaan obat tiap tahunnya antara 18000–70000 buah, sedangkan kelompok obat dengan pemakaian tinggi rata – rata permintaan obat tiap tahunnya diatas 70000 buah.

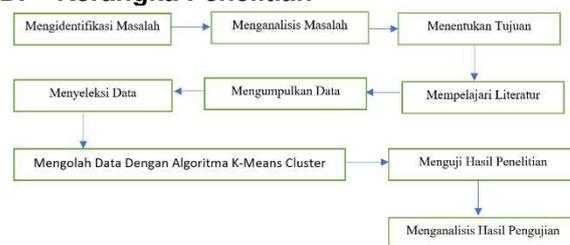
- Implementasi Algoritma K-Means Untuk Pengelompokan Penyakit Pasien Pada Puskesmas Kajen Pekalongan [9], bertujuan mendapatkan informasi tentang pengelompokan penyakit “AKUT” dan “TIDAK AKUT” yang biasanya banyak diderita oleh pasien pada Puskesmas Kajen Pekalongan dan didapat kesimpulan bahwa inialisasi jumlah cluster sebanyak 2 buah sesuai dengan mendefinisikan nilai k untuk jumlah cluster akut sebanyak 376 item, cluster tidak akut sebanyak 624 item dengan total jumlah data adalah 1000 item.
- Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (DBD) Berdasarkan Provinsi [10], mendapatkan hasil 3 cluster yaitu cluster dengan tingkat tinggi (C1), cluster dengan tingkat sedang (C2) dan cluster dengan tingkat rendah (C3), dimana 4 provinsi dengan cluster tingkat tinggi untuk penderita kasus DBD yakni Sumatera Utara, Jawa Barat, Jawa Tengah dan Jawa Timur, 13 provinsi dengan cluster tingkat sedang dan 17 provinsi lainnya termasuk dalam cluster tingkat rendah.

C. Kerangka Pemikiran

Tabel 1. Kerangka Pemikiran Penelitian

Permasalahan		
Data penyebaran sarang jentik nyamuk di Kelurahan Sumampir hanya menjadi tumpukan statistik yang belum digali informasinya.		
Tujuan		
Mengelompokkan data penyebaran sarang jentik nyamuk		
Mengetahui keakuratan algoritma k-means dalam pengelompokan data penyebaran sarang jentik nyamuk di Kelurahan Sumampir		
Eksperimen		
Inputan	Metode	Implementasi
Data penyebaran sarang jentik nyamuk	K-means	Rapid minner
Hasil		
Menghasilkan sistem yang mampu mengelompokkan data penyebaran sarang jentik nyamuk di Kelurahan Sumampir dengan menggunakan algoritma k-means		
Manfaat		
Membantu petugas PSN dalam mengetahui tempat mana saja yang menjadi peringkat pertama sebagai sarang jentik nyamuk, dan juga bisa menjadi bahan acuan dalam memonitoring dan pemberian penyuluhan kesehatan kepada masyarakat sekitar		

D. Kerangka Penelitian



Gambar 4. Kerangka Penelitian

HASIL DAN PEMBAHASAN

Dalam penelitian ini, sumber data diambil dari Petugas PSN Kelurahan Sumampir, Kabupaten Banyumas dimana data tersebut merupakan data yang diambil dari hasil pemeriksaan jentik Aedes Aegypti yang ada pada countainer pada masing-masing Kepala Keluarga di RT.01/RW.08 Kelurahan Sumampir, Kabupaten Banyumas. Dalam melakukan pengolahan data, data yang diperoleh akan diclustering berdasarkan 2 kriteria penilaian yakni 1). Jumlah Coutainer 2). Jumlah keberadaan jentik nyamuk seperti yang ditunjukkan pada table berikut:

Tabel 2. Tabel Rekapitulasi Hasil Pemeriksaan Jentik *Aedes Aegypti* di wilayah RT.01/08 Kelurahan Sumampir Kabupaten Banyumas

Countainer	Jumlah Countainer	Jumlah Keberadaan jentik nyamuk
Bak Kamar Mandi	397	12
Penampung Bawah Kulkas	251	16
Penampung Dispenser	219	44
Barang Bekas	206	2
Akuarium	129	1
Vas Bunga	81	5
Pot Bunga	477	8
Lain-lain	6	0

Setelah diakumulasikan maka akan didapatkan ada nilai dari setiap variabel. Data tersebut masuk ke dalam tahapan clustering dan diolah dengan K-means berdasarkan cluster yang sudah ditentukan.

Cendroid Data

Pada penerapan algoritma K-means akan dihasilkan nilai titik tengah atau centroid dari data. Proses menghitung centroid data dilakukan dengan mengambil nilai terbesar (maksimum) untuk cluster dengan tingkat tinggi (C1), nilai rerata (average) untuk cluster dengan tingkat sedang (C2) dan nilai terkecil (minimum) untuk cluster dengan tingkat rendah (C3). Berikut merupakan nilai centroid awal untuk iterasi 1 yang dapat dilihat pada tabel berikut:

Tabel 3. Tabel Cendroid Data Awal (Iterasi 1)

Centroid		
Max (C1)	487	10
Average (C2)	252.750	11
Min (C3)	72	1,667

Clustering Data

Proses clustering dengan menggunakan nilai centroid awal yang terdapat pada tabel 3, akan memperoleh hasil pengelompokan pada iterasi 1 yang dipaparkan pada tabel di bawah ini:

Tabel 4. Tabel Pengelompokan Data Iterasi

Countainer	C1	C2	C3
Bak Kamar Mandi	1		
Penampung Bawah Kulkas		1	
Penampung Dispenser		1	
Barang Bekas		1	
Pecahan Botol/ Air Kemasan		1	
Akuarium			1
Vas Bunga			1
Pot Bunga	1		
Lain-lain			1
Hasil	2	4	3

Berdasarkan tabel 4, diperoleh pengelompokan untuk iterasi 1: Cluster tingkat tinggi (C1) terdiri dari Proses iterasi akan terus berlangsung sampai hasil iterasi yang terakhir sama dengan hasil iterasi sebelumnya. Proses nilai titik tengah atau centroid akan menyesuaikan sesuai dengan iterasi yang ada. Dalam hal ini proses akan dilanjutkan dengan menggunakan aplikasi RapidMiner.

Penerapan aplikasi RapidMiner:

Data view: merupakan lembar kerja yang terdapat pada Result Perspective untuk menampilkan dan menjelaskan data yang sudah diolah secara lengkap berdasarkan dengan klasternya dari example set (read excel).

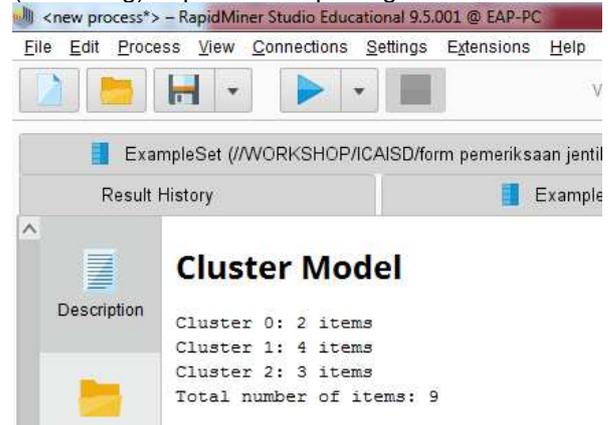
Open in Turbo Prep Auto Model

Row No.	id	cluster	nama	A1	B1
1	1	cluster_0	Bak Kamar M...	397	12
2	2	cluster_1	Tempayan	251	16
3	3	cluster_1	Pecahan Bo...	219	1
4	4	cluster_1	Barang Bekas	206	2
5	5	cluster_1	Kulkas / Disp...	335	25
6	6	cluster_2	Tandon Air	129	0
7	7	cluster_2	Vas Bunga	81	5
8	8	cluster_0	Pot Bunga	577	8
9	9	cluster_2	Lain-lain	6	0

Gambar 5. Gambaran Data View

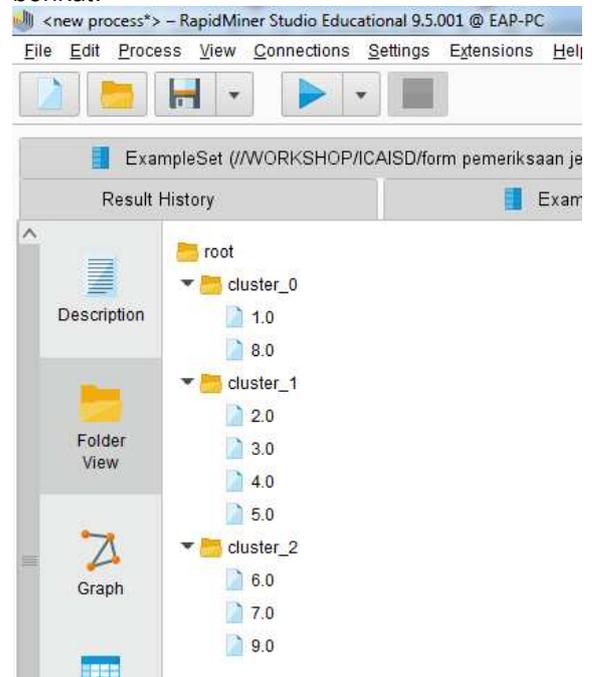
Flot view : adalah area lembar kerja yang ada dalam Result Perspective dan berguna untuk menampilkan data yang sudah diolah secara keseluruhan dan lengkap dengan klasternya dari example set (read excel) dalam bentuk Scatter diagram.

Text view : merupakan area lembar kerja yang digunakan untuk menampilkan database yang sudah diolah secara lengkap dengan clusternya. Tampilan dari cluster model (clustering) dapat terlihat pada gambar berikut.



Gambar 6. Gambaran Text View

Folder view: merupakan sheet untuk menampilkan database yang telah diolah secara keseluruhan lengkap dengan clusternya. Tampilan folder view dapat dilihat pada gambar berikut:



Gambar 7. Gambaran Folder View

KESIMPULAN

Dari hasil penelitian yang telah dilakukan, maka penulis dapat menarik beberapa kesimpulan, diantaranya sebagai berikut :

1. Pengujian yang dilakukan dalam penelitian ini, iterasi clustering pada data penyakit DBD (Demam Berdarah Dengue) terjadi sebanyak 6 kali iterasi.
2. Data tersebut diolah menggunakan Rapidminer untuk menentukan nilai centroid dalam 3 cluster yaitu cluster dengan tingkat

- tinggi (C1), cluster dengan tingkat sedang (C2) dan cluster dengan tingkat rendah (C3).
3. Dari 9 data jumlah keberadaan jentik nyamuk berdasarkan container yang dijadikan sarang dapat diketahui, 2 container cluster tingkat tinggi untuk jumlah keberadaan jentik nyamuk tertinggi yakni bak mandi dan penampung dispenser, 4 container cluster tingkat sedang dan 3 container lainnya termasuk pada cluster tingkat rendah.
 4. Dari analisa cluster diatas, perlu dilakukan lagi penelitian lanjutan supaya clusterisasi data jentik nyamuk dapat dilakukan secara lebih valid dengan menetapkan nilai centroid terbaik.

DAFTAR PUSTAKA

- [1] D. Kobayashi *et al.*, "Dengue Virus Infection in *Aedes albopictus* during the 2014 Autochthonous Dengue Outbreak in Tokyo Metropolis, Japan," *Am. J. Trop. Med. Hyg.*, vol. 98, no. 5, pp. 1460–1468, 2018, doi: 10.4269/ajtmh.17-0954.
- [2] Y. H. Lai, "The climatic factors affecting dengue fever outbreaks in southern Taiwan : an application of symbolic data analysis," *Biomed. Eng. Online*, vol. 17, no. S2, pp. 1–14, 2018, doi: 10.1186/s12938-018-0575-4.
- [3] Y. A. Indrayani and T. Wahyudi, "Info Datin-Situasi Demam Berdarah Di Indonesia Tahun 2017-," *Pusat Data dan Informasi Kementerian Kesehatan RI*. p. 7, 2018.
- [4] T. W. Kesetyaningsih, S. Andarini, Sudarto, and H. Pramodyo, "Determination Of Environmental Factors Affecting Dengue Incidence In Sleman District, Yogyakarta, Indonesia," *Ajid*, vol. 12, no. 3, pp. 13–25, 2018.
- [5] F. Gorunescu, *Data Mining: Concepts, Models and Techniques (Vol.12)*. Springer-Verlag Berlin Heidelberg: Intelligent Systems References Library, 2013.
- [6] A. V. D. Sano and H. Nindito, "Application Of K-Means Algorithm For Cluster Analysis On Poverty Of Provinces In Indonesia," *ComTech*, vol. 7, no. 2, pp. 141–150, 2016.
- [7] A. K. Wardhani, "K-Means Algorithm Implementation For Clustering Of Patients Disease In Kajen Clinic Of Pekalongan)," *J. Transform.*, vol. 14, no. 1, pp. 30–37, 2016.
- [8] M. H. Adiya and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru," *J. Nas. Teknol. dan Sist. Inf.*, vol. 05, no. 01, pp. 17–24, 2019.
- [9] A. K. Wardhani, "Implementasi Algoritma K-Means untuk Pengelompokan Penyakit Pasien pada Puskesmas Kajen Pekalongan," *J. Transform.*, vol. 14, no. 1, pp. 30–37, 2016.
- [10] K. Fatmawati and A. P. Windarto, "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (DBD) Berdasarkan Provinsi," *Comput. Eng. Sci. Syst. J.*, vol. 3, no. 2, p. 173, 2018, doi: 10.24114/cess.v3i2.9661.