

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Pengelompokan data penjualan merupakan salah satu teknik yang dibutuhkan sebagai penunjang dalam meningkatkan produktifitas penjualan bibit kentang, terlebih di daerah yang menjadi salah satu penghasil komoditas pertanian kentang terbesar. Hal ini juga menjadi salah satu pemicu bertambahnya konsumen yang membutuhkan bibit kentang. Selain itu, mengetahui karakteristik dari penjualan bibit kentang merupakan hal yang bermanfaat dalam meningkatkan penjualan(1). Pengolah data penjualan bibit kentang untuk menghasilkan informasi dibutuhkan teknik data mining. Data mining adalah studi yang mengmpulkan, membersihkan, mengolah, menganalisis, dan memperoleh informasi baru dari data(4). Data mning dapat memperoleh informasi dengan cara melakukan prses klasifikasi,prediksi,perkiraan dan asumsi dari sekumulan data yang berukuran besar(5).Salah Teknik analisis data mining yang paling sering digunakan dalam mengelompokkan data penjualan adalah clustering (2). DBSCAN (Densy-based Spatial Cluster of Aplications with Noise) algoritma clustering yang didsarkan pada kepadatan (density) data (6) dan hasil pengelompokkannya akan menghasilkan 3 macam status dari setiap data, yaitu inti (*core*), batas (*border*), dan (*noise*)(7).

Penelitian (Ni Mde Anindya Santika Devi¹, I Ketut Gede Darma Putra²,) menjelaskan bahwa melalui metode DBSCAN dapat membantu dalam pengelompokan data dan menghasilkan penentuan pembeli potensial berdasarkan dari dataset penjualan(8). (Fitri Indriyani & Eni irfani, 2019) Menjelaskan bahwa clustering dikembangkan untuk menghasilkan kinerja pengelompokan yang baik serta memiliki kemampuan memperkirakan volume dari penjualan(9). Algoritma DBSCAN memiliki cara pengelompokan data yang berbeda dibandingkan clustering yang lain sehingga menghasilkan data yang akurat dengan melakukan kinerja penerapan yang luas pada pengelompokan data(10). (Betha Nurina Sari,& Aji Primajaya , 2019) menjelaskan clustering DBSCAN pada data penjualan untuk mengetahui karakteristik dari setiap konsumen yang melakukan pembelian terhadap satu barang(11). Clustering DBSCAN merupakan algoritma yang efisien

dalam mengelompokkan dataset yang besar(12). DBSCAN telah banyak digunakan dalam bidang lain karena kesederhanaan dan efek pengelompokan yang baik dan akurat (13).

Penelitian ini menerapkan metode algoritma DBSCAN dalam melakukan pengelompokan bibit kentang berdasarkan dataset penjualan bibit kentang untuk mengetahui penjualan bibit kentang jenis apa yang banyak dibutuhkan oleh petani sebagai konsumen bibit kentang dan mengetahui jenis pembayaran secara tunai atau kredit yang banyak digunakan konsumen dalam mengoptimalkan penjualan bibit kentang.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, maka yang menjadi bahan penelitian adalah bagaimana cara meningkatkan produktifitas penjualan bibit kentang dengan algoritma DBSCAN dalam melakukan proses clustering data.

1.3. Tujuan dan Manfaat

1.3.1. Tujuan

Adapun tujuan yang dilakukan dalam penelitian ini adalah untuk membantu meningkatkan produktifitas penjualan bibit kentang.

1.3.2. Manfaat

Adapun manfaat yang diperoleh dengan selesainya penelitian ini:

1. Mengetahui kinerja dari Metode DBSCAN.
2. Menjadi bahan referensi untuk pembelajaran dan penelitian metode *clustering* dalam melakukan pengelompokkan data penjualan.

1.4. Keterbaruan

Beberapa penelitian yang pernah dilakukan sebelumnya menggunakan DBSCAN sebagai berikut :

Penelitian (Sharma Dr.A.K.,2017) menguji nilai performa kinerja dari EPS DBSCAN dan E-DBSCAN, menghasilkan nilai performa paling baik pada nilai EPS : 1.33 dan E-DBSCAN : 0,9 yang memiliki tingkat akurasi yang paling tinggi(10). (JEBARIS.,2019) meneliti tentang kinerja dari algoritma DBSCAN dan menunjukkan performa kinerja

yang paling unggul berdasarkan nilai AF-DSCAN sebesar 2.562, FN-DBSCAN: 2.185 dan AF-DBSCAN: 2.562.(12).

Penelitian (Zhu,Q.X,2020) meneliti perbandingan K-Means dan DBSCAN dalam kecepatan penggelompokkannya yang menunjukkan algoritma k-means memiliki waktu eksekusi lebih cepat dibandingkan DBSCAN dengan waktu komputasi K-Means sebesar 0.278 dan DBSCAN sebesar 1.956 namun indeks rand rendah menandakan kualitas pengelompokan yang buruk berdasarkan nilainya dengan perbandingan K-Means: 0.28 dan DBSCAN: 1.00, (13).

Penelitian (Safitri.D,Wuryandari.T,R.R,2017) menerapkan algoritma DBSCAN dalam mengelompokkan daerah yang memproduksi padi pada provinsi jawa tengah dari 35 kabupaten/kota yang ada di Jawa Tengah dan hasil penelitian menunjukkan cluster 3 merupakan kelompok daerah penghasil padi tertinggi dengan 2 noise pada cluster data yaitu Kabupaten Cilacap dan Kabupaten Wonogiri(14).

Penelitian (Isnarwaty,D P,Irhamah,2019) meneliti berdasarkan nilai silhouette coefficient menunjukkan metode DBSCAN lebih produktif dari K-Means dalam mengelompokkan tweet yang ditujukan kepada layanan ekspedisi JNE, J&T, dan Pos Indonesia karena menghasilkan silhouette coefficient yang lebih tinggi(15).

Penelitian (M,P.Dewi,C.Siam,Wijayanti,AuliaN,Raeni,2020) membandingkan algoritma K-Means dan DBSCAN, mendapatkan hasil bahwa DBSCAN lebih baik dalam hal pengelompokan data yang dilihat dari nilai koefisien *silhouette*(16).

(Soni.D,2017) melakukan penelitian untuk meningkatkan tingkat akurasi algoritma DBSCAN berdasarkan algoritma back propagation dalam menghitung jarak euclidean secara dinamis.

Penelitian (Li X,Zhang P,Zhu G.2019) bertujuan untuk mengidentifikasi kelompok para pengguna layanan publik menggunakan algoritma DBSCAN berdasarkan kepadatan data

yang ada pada pengguna layanan public bahwa hasil identifikasi cluster penumpang bermanfaat untuk optimalisasi fasilitas pelayanan, organisasi penumpang dan pembinaan, evakuasi arus penumpang yang tidak normal dan hasil penelitian ini menjadi masukan dalam menentukan sikap dalam keadaan darurat yang terjadi pada layanan publik(18).

Penelitian (Lazarevic A,Xu X,Fiez T, Obradovic b,2020) menerapkan information gain untuk mengevaluasi algoritma DBSCAN dalam mengintegrasikan data agar setiap cluster mendapatkan kelompok pada setiap daerah masing masing (19).

Penelitian (Li H,a,Zhang A,Pei X,2019) menerapkan algoritma DBSCAN dan BP Neural Network untuk mengoptimasi titik pengukuran sensor suhu alat mesin CNS DAN hasil penelitian mendapatkan metode koefisien korelasi pearson mengurangi titik pengukuran suhu dari 16 hingga 5. (20).