**SISTEM SEGMENTASI KELUHAN PELANGGAN MENGGUNAKAN METODE**

**K-MEDOIDS DI PT PDAM TIRTAWENING BANDUNG**

Diajukan Sebagai Salah Satu Syarat

Matakuliah Tugas Akhir 1

Jenjang Strata 1 pada Program Studi Informatika

Universitas Jenderal Achmad Yani

Oleh:

**Febry Ramadhan**

**3411161046**

|  |  |  |
| --- | --- | --- |
| Pembimbing 1 |  | Pembimbing 2 |
| **Yulison Herry Chrisnanto** |  |  |
| **NID. 4121 668 63** |  |  |

****

**JURUSAN INFORMATIKA**

**FAKULTAS SAINS DAN INFORMATIKA**

**UNIVERSITAS JENDERAL ACHMAD YANI**

**CIMAHI**

**2019**

1. **Latar Belakang**

Data mining adalah proses analitik yang dirancang untuk memeriksa sejumlah data yang besar dalam mencari suatu pengetahuan tersembunyi yang berharga dan konsisten [1]. Selain itu data mining merupakan proses pencarian pola-pola yang menarik dan tersembunyi (hidden patten) dari suatu kumpulan data yang berukuran besar yang tersimpan dalam suatu basis data, data warehouse, atau tempat penyimpanan data lainnya [2].

Pada penelitian ini *data mining* akan digunakan untuk mengelola data pengaduan atau keluhan pelanggan di PT PDAM Tirtawening. Dimana PT PDAM Tirtawening ini merupakan perusahaan yang bergerak di bidang pengelolaan air minum dan pengelolaan sarana air limbah di daerah, untuk meningkatkan kesejahteraan masyarakat yang mencakup aspek ekonomi, sosial, kesehatan dan pelayanan umum.

Dimana jumlah orang yang berlangganan sangat lah banyak dan tersebar di banyak wilayah yang terdiri dari 174.583 yang berlangganan air minum dan 117.505 yang berlangganan mengenai pengaturan air limbah, yang mana terdiri dari beberapa golongan meliputi pelanggan golongan sosial, niaga, non niaga, industri, dan golongan pelanggan kelompok khusus. Jumlah pelanggan yang sebanyak itu tentu akan menghasilkan keluhan yang sangat banyak dan beragam, namun begitu sangat dimungkinkan jika sebuah wilayah mempunyai kecenderungan keluhan yang sama. Untuk mengetahui kecenderungan keluhan pada suatu wilayah tertentu jika tidak menggunakan bantuan komputasi tentunya akan sangat sulit untuk mendapatkan hasil yang akurat

Dari data keluhan tersebut tanpa bantuan komputasi tetap dapat dilakukan analisis terhadap data tersebut, tetapi dalam proses analisis dari data keluhan tersebut akan memerlukan waktu yang tidak sedikit, maka diperlukannya bantuan komputasi untuk dapat melakukan analisis data keluhan menjadi lebih mudah, cepat, dan efisien.

Selain *data mining* adapun pengelompokan data dari keluhan perlu dilakukan proses pengelompokkan data untuk memudahkan dalam pengelolaan data keluhan. Dalam penelitian ini untuk dapat menghasilkan hasil yang lebih baik, maka proses mengelolaan data keluhan PT PDAM Tirtawening dilakukan menggunakan algoritma K-Medoids, dimana algortima K-Medoids sendiri yaitu metode yang digunakan untuk pengelompokan data sehingga informasi jenis keluhan yang lebih banyak di suatu daerah tertentu akan diketahui dengan waktu yang singkat.

Ada beberapa teknik untuk melakukan klasterisasi, diantaranya *K-Means*, *K-Medoids* dan lainnya. Teknik *K-Means* diperkenalkan oleh MacQueenm (1967), sedangkan teknik *K-Medoids* (*Partitioning Around Medoids*) diperkenalkan oleh Kaufman dan Rousseeuw (1990). Berdasarkan penelitian yang dilakukan oleh (Aishwaya) didapati kesimpulan bahwa algoritma *K-Means* efisien untuk kumpulan data yang lebih kecil dan *K-Medoids* nampaknya berperforma lebih baik untuk dataset besar[3].

1. **Rumusan Masalah**

PT PDAM Tirtawening menerima ribuan keluhan setiap tahun nya, dimana keluhan

Tersebut terdiri dari beberapa jenis keluhan yang berbeda dengan alamat yang berbeda pula, dari keberagaman jenis keluhan tersebut tentu ada sebuah kecenderungan jika disebuah wilayah mempunya jenis keluhan yang realatif seragam. Untuk mengetahui kecenderungan jenis keluhan disebuah wilayah tertentu jika dilakukan analisis terhadap data tersebut tanpa menggunakan bantuan komputasi, proses dari analisis akan memerlukan waktu yang tidak sedikit mengingat data yang ditampilkan bersifat tabular sehingga memelukan ketelitian untuk mengecek satu persatu dari data keluhan tersebut. Sehingga untuk pengelolaan data keluhan tersebut diperlukannya sistem komputasi untuk mengelola datanya. Serta untuk mengetahui jenis keluhan seperti apa yang banyak dikeluhkan oleh pelanggan di sebuah wilayah dapat menggunakan metode K-Medoids. Dari hasil analisis data keluhan dapat dimanfaatkan sebagai bahan referensi untuk cara menanggulangi masalah tersbut agar tidak terjadi lagi dimasa yang akan datang atau setidaknya meminimalisir risiko keluhan tersebut terjadi kembali, karena dari hasil identifikasi sebaran keluhan dan jenis keluhannya dapat diketahui kecenderungan masalah yang sering terjadi di suatu wilayah, sehingga perusahaan dapat mencegah atau setidaknya meminimalisir masalah tersebut untuk terjadi lagi di kemudian hari.

# **Batasan Masalah**

Dalam melaksanakan penelitian ini diperlukan suatu batasan – batasan agar lebih fokus serta tidak penyimpang dari yang telah direncanakan, batasan – batasan tersebut yaitu sistem hanya dapat diakses secara offline. Sehingga dalam penggunaan sistem ini hanya dapat dilakukan di daerah lingkungan perusahaan saja, tidak dapat dilakukan di luar daerah perusahaan.

Serta dalam penelitian ini diberikannya batasan daerah terkait tempat tersebarnya pelanggan. Dalam penelitian ini hanya akan meneliti di daerah Bandung, oleh karena itu data keluhan yang akan dijadikan sebagai penelitian merupakan data yang berasal dari pelanggan di wilayah Bandung saja. Sehingga dalam penelitian ini tidak akan melibatkan pelanggan yang tidak terletak di kota Bandung.

# **Tujuan Penelitian**

Tujuan dari penelitian ini yaitu membangun sistem yang dapat mengidentifikasi kecenderungan masalah yang sering terjadi disebuah wilayah tertentu, dengan mengetahui jenis dan letak keluhan yang tersebar dari banyak pelanggan yang ditujukan pada PT PDAM Tirtawening. Serta untuk mempermudah dalam melakukan analisis dari data keluhan di PT PDAM Tirtawening.

# **Keluaran dan Manfaat**

Keluaran dari penelitian ini yaitu sebuah sistem yang dapat mengidentifikasi sebaran keluhan, serta memberikan kemudahan terhadap analisis data keluhan yang masuk ke PT PDAM Tirtawening.

Manfaat dari penelitian ini yaitu, untuk mempermudah dalam melakukan analisis data terhadap data keluhan pelanggan, sehingga untuk mengetahui informasi yang dibutuhkan dapat dilakukan dengan waktu yang lebih singkat, serta untuk menganalisis data keluhan supaya tidak memerlukan ketelitian yang sangat detail, serta dengan bantuan sistem data yang ditampilkan akan dikelompokan sesuai dengan jenis keluhan dan dimana masalah yang dikeluhkan tersebut terjadi. Sehingga dalam sistem ini akan menampilkan suatu kelompok – kelompok yang mewakili suatu kondisi dimana kondisi tersebut merupakan suatu kecenderungan keluhan yang terjadi pada sebuah daerah di daerah kota Bandung.

# **Tinjauan Pustaka**

## **Data Mining**

Data *Mining* (Witten, 2011) didefinisikan sebagai proses penemuan pola dalam data. Berdasarkan tugasnya, *data mining* dikelompokkan menjadi deskripsi, estimasi, prediksi, klasifikasi, *clustering* dan asosiasi (Larose, 2005) [4]. Serta data mining merupakan proses pencarian pola atau informasi dalam data terpilih dengan menggunakan teknik atau metode tertentu [5].

* 1. **Clustering**

*Clustering* atau pengelompokan data merupakan kegiatan mempertimbangkan sebuah pendekatan penting untuk mencari kesamaan dalam data dan menempatkan data yang sama ke dalam kelompok-kelompok. *Clustering* dianggap sebagai *unsupervised learning* yang paling penting dimana setiap masalah pada jenis ini berkaitan untuk menemukan struktur dalam kumpulan data yang tidak berlabel. Sebuah pengelompokkan *cluster* membagi kumpulan data kedalam sejumlah kelompok dimana kesamaan pada sebuah kelompok tertentu lebih besar daripada kelompok lainnya [6]. Selain itu Clustering adalah proses dari pembentukan segmen – segmen data dengan cara mengukur kemiripan sebuah data dengan data lainnya [7]. Tujuan dari data *Clustering* ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses *Clustering*, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu *cluster* dan memaksimalisasikan variasi antar *cluster* [8].

* 1. **K-Medoids**

Algoritma PAM sering dikenal dengan algoritma k-Medoids adalah sebuah algoritma yang merepresentasikan cluster yang dibentuk menggunakan medoids. Proses pembentukan cluster dimulai dengan menentukan k objek dari dataset secara acak sebagai medoid, selanjutnya hitung cost setiap objek non-medoid dengan k objek, cost terkecil setiap objek non-medoid terhadap medoid akan masuk dalam cluster dimana medoid tersebut berada. Secara iterative ditentukan kembali objek baru secara acak, lalu proses perhitungan cost dilakukan kembali. Apabila total cost yang dihasilkan lebih kecil dari cost setiap objek dengan medoid lama, maka objek baru tersebut dapat menjadi medoid baru. Iterasi berakhir hingga tidak ada perubahan terhadap cost yang dihasilkan oleh medoid baru [9]. Teknik K-Medoid klastering : algoritma ini adalah sebuah metode pengelompokan canggih yang mirip dengan K-Means. Kedua algoritma ini membagi database menjadi beberapa bagian dan mencoba untuk meminimalkan jarak antara pusat setiap cluster dan masing-masing titik data berlabel. K-medoid memiliki dua fitur penting sedangkan K-means menghasilkan hasil yang lebih baik dalam hal segmentasi. Pertama, K-medoid menginisialkan komponen matriks dengan data input yang disebut titik data, sementara itu bukan batasan untuk K-means. Kedua, Teknik K-medoid memilih pusat data sebagai pusatnya yang disebut medoid, tetapi yang lain menggunakan rata-rata dari titik data sebagai pusat. [10].

# **Metode Penelitian**

Metode penelitian menjelaskan langkah-langkah yang akan digunakan serta perancangan dalam melakukan implementasi metode K-Medoids clustering untuk identifikasi persebaran pelanggan yang melakukan keluhan.

**7.1 Analisis Sistem Berjalan**

**7.2 Identifikasi dan Pengumpulan Data**

**7.3 Perancangan Perangkat Lunak dan Pembuatan Sistem**

**7.4 Pengujian dan Evaluasi**

* 1. **Pelaporan dan Publikasi Ilmiah**

# **Jadwal Penelitian**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Kegiatan | Bulan ke-1 | | | | Bulan ke-2 | | | | Bulan ke-3 | | | | Bulan ke-4 | | | |
| 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| 1. | Analis Sistem Berjalan |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2. | Identifikasi dan Pengumpulan Data |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3. | Perancangan Perangkat Lunak dan Pembuatan Sistem |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4. | Pengujian dan Evaluasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5. | Pelaporan dan Publikasi Ilmiah |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

1. References