

## **BAB II**

### **TINJAUAN PUSTAKA DAN DASAR TEORI**

#### **2.1. Tinjauan Pustaka**

Tinjauan pustaka yang digunakan dalam penelitian ini didapat dari penelitian yang dilakukan oleh Andik (2018), dalam penelitiannya mengimplementasikan sebuah data pelanggan PLN Kota Lhokseumawe yang digunakan sebagai bahan klasifikasi metode K-NN untuk penggolongan subsidi dan non subsidi berdasarkan data variabel. Variabel yang digunakan antara lain, pendapatan keluarga secara keseluruhan (K1), pekerjaan orang tua (K2), dan luas rumah (K3). Penggolongan subsidi terdapat 2 golongan tarif yaitu 450 VA dan 900 VA, Kemudian untuk penggolongan non subsidi yaitu 1352 VA. Tujuan penelitian yang dilakukan, untuk memudahkan pelanggan PLN Kota Lhokseumawe mengetahui besaran daya listrik yang akan digunakan.

Cici Astria (2019), dalam penelitiannya melakukan penerapan k-medoids pada rumah tangga yang memiliki sumber penerangan listrik PLN berdasarkan provinsi. Tujuan dari penelitiannya yaitu untuk mengelompokkan pelanggan listrik golongan rumah tangga berdasarkan provinsinya, sehingga dengan metode k-medoids dapat membantu pemerataan listrik ke provinsi-provinsi yang masih kurang dalam penerangannya. Jumlah klaster yang digunakan berjumlah 2 klaster yaitu tingkat tinggi (C1) dan tingkat rendah (C2). Dan hasil yang didapat menyimpulkan bahwa pengelompokan listrik golongan rumah tangga

diterapkan dengan baik, terdiri dari 34 provinsi, 23 provinsi termasuk klaster tingkat tinggi dan 11 provinsi termasuk klaster tingkat rendah.

Afian Nugroho (2020), dalam penelitiannya mengimplementasikan data dari PLN Unit Pelaksana Pengatur Distribusi (UP2D) Kota Semarang menggunakan metode Decision Tree C4.5 untuk mengklasifikasikan data sebagai penentu subsidi listrik untuk masyarakat kurang mampu. Variabel yang digunakan untuk syarat pemberian subsidi yaitu pekerjaan, penghasilan, aset kendaraan, jumlah tanggungan keluarga, kepemilikan rumah, jenis lantai dan sumber air minum. Tujuan penelitian ini, di dasari pada kebijakan subsidi listrik yaitu subsidi harga, dan penyimpangan subsidi listrik yang dinikmati oleh golongan rumah tangga yang mampu dari segi ekonomi, sehingga menyebabkan listrik subsidi menjadi tidak tepat sasaran. Dan dengan metode Decision Tree C4.5 dapat menampilkan bahan keputusan untuk menentukan pelanggan yang berhak mendapatkan subsidi listrik.

Sukma Sindi (2020), dalam penelitiannya melakukan analisa klaster pengelompokan penyebaran covid-19 di Indonesia dengan data yang diperoleh dari Kementerian Kesehatan Indonesia, data yang diperoleh tersebut dikelola menggunakan metode k-medoids dengan 3 klaster. Atribut data yang digunakan meliputi data kasus positif, sembuh dan meninggal. Pengolahan data tersebut dilakukan menggunakan software rapidminer dengan menginputkan data excel yang didapat dari kementerian kesehatan. Sehingga hasil analisa tersebut menampilkan hasil dari seluruh data berjumlah 34 data record terbagi

menjadi klaster pertama 1 *record*, klaster kedua 2 *record* dan klaster ketiga 31 *record*.

Idham Kholid Ramadhani (Diusulkan), dalam penelitian yang diajukan, yaitu mengimplementasikan data pelanggan PLN Kebumen, sebagai bahan pertimbangan untuk penentuan daya baru dan mengklasterkannya menjadi 2 klaster yaitu subsidi dan non subsidi, penelitian ini menggunakan metode K-Medoids. Kemudian hasil klaster yang didapat digunakan untuk melakukan prediksi data pelanggan baru dengan mengimplementasi metode K-NN. Penelitian yang akan dilakukan ini, merupakan pengembangan dari penelitian yang dilakukan oleh Andik, karena melakukan klasifikasi hanya menentukan kelas berdasarkan pola, sehingga perlu meminimalkan kemiripan datanya supaya lebih tepat dan akurat.

Tabel 1.1 Tinjauan Pustaka

No	Peneliti	Judul	Metode	Objek Data
1	Andik Bintoro, Safwandi (2018)	Implementasi Data Mining Penentuan Daya Pelanggan Baru untuk Klasifikasi Subsidi dan Non Subsidi di Wilayah PLN Kota Lhokseumawe	Algoritma K-Nearest Neighbour	PLN Kota Lhokseumawe

Tabel 1.1 Tinjauan Pustaka

No	Peneliti	Judul	Metode	Objek Data
2	Cici Astria, Agus Perdana Windarto, Dedy Hartama (2019)	Penerapan K- Medoid Pada Rumah Tangga Yang Memiliki Sumber Penerangan Listrik PLN Berdasarkan Provinsi	K-Medoids	Situs Badan Pusat Statistik Indonesia yaitu data Persentase Rumah Tangga Menurut Provinsi dan Perlakuan, 2013-2014.
3	Afian Nugroho, Aji Supriyanto (2020)	Implementasi Metode Decision Tree C4.5 Dalam Pemberian Subsidi Listrik Kepada Masyarakat	Decision Tree C4.5	PLN Unit Pelaksana Pengatur Distribusi (UP2D) Kota Semarang
4	Sukma Sindi, Weni Ratnasari Orktapia Ningse, Irma Agustika Sihombing, P.P.P.A.N.W Fikrul Ilmi R.H.Zer, Dedy Hartama (2020)	Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid- 19 Di Indonesia	K-Medoids	Kementerian Kesehatan Indonesia

Tabel 1.1 Tinjauan Pustaka

No	Peneliti	Judul	Metode	Objek Data
5	Idham Kholid Ramadhani (Diusulkan)	Implementasi Data Mining Untuk Penentuan Kelompok Pelanggan Listrik Subsidi Atau Non Subsidi	K-Medoids	PT. PLN Kebumen

## 2.2. Dasar Teori

### 2.2.1. Data Mining

Dalam bentuk penafsiran, *data mining* merupakan proses menemukan pengetahuan tersembunyi di basis data. *Data mining* adalah semi-proses otomatis dengan teknik statistik, matematika, kecerdasan buatan dan *machine learning*. Implementasi data mining biasanya dilakukan sebagai proses menemukan hubungan dalam data yang tidak diketahui dan kemudian menyajikannya dengan cara yang dapat dipahami oleh semua orang misalnya dalam gambar diagram ataupun grafik, sehingga dapat menjadi dasar dalam pengambilan keputusan. (Bintoro & Safwandi, 2018).

Dalam melakukan penggalian informasi, istilah *data mining* dan *knowledge discovery in database* (KKD) sering digunakan secara bergantian. Kedua istilah tersebut, memiliki konsep yang berbeda, tetapi saling berkaitan. (Yunita, 2018). Proses KKD secara garis besar sebagai berikut:

a. *Data Selection*

Dalam proses ini, sekumpulan data di seleksi berdasarkan kebutuhan, sehingga hasil seleksi data tersebut dapat digunakan dalam proses *data mining* yang disimpan dalam satu berkas, terpisah dari basis data.

b. *Pre-Processing / Cleaning*

Proses ini dilakukan sebagai pembersihan data dari hasil seleksi data yang sebelumnya dilakukan. Proses pembersihan data mencakup duplikasi data, data inkonsisten, dan memperbaiki data kesalahan cetak (*tipografi*).

c. *Transformation*

Proses tahap ini bisa disebut sebagai *coding* dimana data yang dipilih ditransformasikan ke nilai yang sesuai untuk proses *data mining*, proses ini merupakan proses kreatif dalam mencari pola informasi pada basis data.

d. *Data Mining*

Tahap ini merupakan proses mencari pola atau informasi menarik yang terdapat dalam data yang terpilih dengan menggunakan metode tertentu, pada *data mining* terdapat variasi dalam teknik, metode atau algoritma, sehingga metode yang digunakan tergantung pada tujuan implementasi atau penelitian yang dilakukan.

e. *Interpretation/evaluation*

Pada proses *data mining* menghasilkan sebuah pola informasi sehingga diperlukan visualisasi data ke dalam bentuk yang mudah dimengerti dan dipahami oleh semua orang atau pihak yang terkait.

Pada metode *data mining*, dibagi menjadi 6 metode berdasarkan pekerjaan atau tugas yang dilakukan., di antaranya deskripsi, estimasi, prediksi, klasifikasi, pengklasteran dan asosiasi. (Wahdi, 2018).

a. Deskripsi

Pada metode deskripsi, peneliti dan analisis menjabarkan secara sederhana mengenai kemungkinan penjelasan terhadap pola dan kecenderungan yang terdapat dalam sebuah data. Metode deskripsi biasanya digunakan dalam analisis yang berkarakteristik tentang suatu kondisi, objek, sejarah ataupun atribut-atribut yang dapat dijelaskan dengan data maupun kata-kata. (Purba, Hasibuan, & Hatmi, 2019).

b. Estimasi

Metode estimasi hampir sama seperti metode klasifikasi, perbedaannya terdapat pada variabel target, pada metode estimasi variabel target cenderung ke arah numerik sedangkan klasifikasi ke arah kategori.

c. Prediksi

Metode prediksi dilakukan untuk menganalisa atau memperhitungkan keadaan di masa mendatang. Dalam metode ini, teknik yang digunakan sama seperti klasifikasi dan estimasi, yang membedakan

metode prediksi dengan klasifikasi dan estimasi yaitu pada nilai dari hasilnya, karena nilai tersebut masih dalam bentuk kemungkinan.

d. Klasifikasi

Metode klasifikasi merupakan proses mengidentifikasi objek ke dalam sebuah kelas berdasarkan variabel yang diamati dengan tujuan mengklasifikasikan kelas yang belum diketahui variabel kategorinya, sehingga variabel kategori tersebut menjadi target utamanya. Contoh dalam penelitian ini yaitu, klasifikasi pelanggan baru ke dalam kategori subsidi atau non subsidi.

e. Pengklasteran

Pengklasteran adalah pengelompokan suatu data *record*, pengamatan, dan bentuk kelas objek yang memiliki kemiripan dan dalam satu klaster memiliki tingkat kemiripan yang maksimum sedangkan data antar klaster memiliki kemiripan yang minimum.

f. Asosiasi

Metode asosiasi merupakan proses *data mining* dengan tugas menemukan aturan asosiatif yaitu “*if...then...*” atau “jika...maka...” yang muncul dalam satu waktu. Metode ini biasanya disebut “*Market Basket Analysis*” karena sering digunakan dalam bisnis untuk menentukan kombinasi diantara himpunan item-item. Sehingga didapat hasil kebiasaan konsumen membeli produk.



### 2.2.2. Klasterisasi K-Medoids

Metode *clustering* mengidentifikasi objek yang memiliki kesamaan karakteristik tertentu, dan kemudian menggunakan karakteristik tersebut sebagai “vektor karakteristik” atau “centroid”. (Yunita, 2018). Peran *clustering* dalam menganalisa data yaitu membagi data ke dalam beberapa kluster sesuai dengan kemiripannya.

Pada implementasi ini, metode *clustering* yang digunakan adalah algoritma K-Medoids atau PAM (*Partitioning Around Medoids*) merupakan sebuah metode partisi kluster yang bertujuan untuk mengelompokkan sekumpulan data menjadi sejumlah k kluster dan menggunakan objek pada *dataset* untuk mewakili sebuah kluster. Medoid merupakan objek yang letaknya terpusat di dalam suatu kluster sehingga *robust* (tidak terpengaruh) terhadap *outlier*. *Cluster* dibangun dengan menghitung kedekatan yang dimiliki antara medoids dengan objek non medoids. (Mustofa & Suasana, 2018). Dengan metode ini hasil yang didapat berbentuk dataset yang kemudian akan digunakan dalam melakukan klasifikasi pelanggan.

Langkah-langkah dalam pengklusteran menggunakan K-Medoids sebagai berikut (Hidayati, Rizmayanti, Dewi, & Fatmasari, 2020):

- 1) Menginisialisasi pusat kluster sebanyak K (jumlah kluster).
- 2) Mengalokasikan setiap data ke kluster terdekat menggunakan persamaan ukuran jarak Euclidian Distance. Rumus:

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (\text{Rumus 2.1})$$

Keterangan:

$d_{ij}$  = jarak antara objek i dengan j

$x_{ik}$  = nilai objek i pada variabel ke-k

$x_{jk}$  = nilai objek j pada variabel ke-k

$p$  = banyaknya variabel yang diamati

- 3) Memilih secara acak objek pada masing-masing kluster, bertujuan sebagai kandidat medoid baru.
- 4) Menghitung jarak setiap objek pada masing-masing kluster dengan kandidat medoid baru.
- 5) Menghitung total simpangan, dengan menghitung nilai total jarak baru dikurangi dengan nilai total jarak lama. Rumus:

$$S = b - a \quad (\text{Rumus 2.2})$$

Keterangan:

$S$  = Simpangan baku

$b$  = Total jarak baru

$a$  = Total jarak lama

Jika  $S < 0$ , maka menukar objek dengan data untuk membentuk sekumpulan k baru sebagai medoid.

- 6) Mengulangi langkah 3 sampai 5 sehingga medoid tidak mengalami perubahan dan hasil tersebut menjadi kluster untuk masing-masing anggota klusternya.

### 2.2.3. *K-Nearest Neighbor (K-NN)*

Metode k-nn sering digunakan untuk memprediksi atau objek baru yang belum diketahui kelas atau kategorinya. Dalam implementasi ini, prediksi data yang digunakan adalah algoritma K-NN (*K-Nearest Neighbor*) dalam perhitungannya didasarkan pada jarak yang paling dekat dengan objek yang akan diprediksinya. Selanjutnya mayoritas k tetangga terdekat menjadi dasar dalam memutuskan kategori dari sampel. Pada algoritma K-NN terdapat 3 jenis jarak yaitu Euclidean, Minkowski, Manhattan. Ketiga jarak tersebut memiliki nilai yang baik, diantaranya 84.47% (untuk metode *euclidean distance*), 83.85% (untuk metode *manhattan distance*), dan 83.85% (untuk metode *minkowski*). (Nishom, 2019).

Dalam implementasi ini, jarak yang digunakan adalah Euclidean karena memiliki nilai akurasi yang baik dan mudah dalam penerapannya.

Pada implementasi klasifikasi K-NN ini menggunakan jarak *Euclidean Distance*, maka langkah-langkah yang dilakukan dalam metode klasifikasi ini seperti berikut, (Dinata, Akbara, & Hasdyna, 2020):

- 1) Menentukan parameter K.
- 2) Menghitung jarak antara data yang akan dievaluasi (data baru) dengan semua data latih.

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \quad (\text{Rumus 2.3})$$

Keterangan:

$x_1$  = Data uji

$x_2$  = Data sampel

$i$  = Variabel data

$d$  = jarak

$p$  = Jumlah data latih

- 3) Mengurutkan jarak yang terbentuk (nilai terkecil ke nilai yang terbesar).
- 4) Menentukan jarak terdekat sampai urutan  $K$ .
- 5) Memasangkan kelas yang bersesuaian.
- 6) Mencari jumlah kelas dari tetangga yang terdekat dan menetapkan kelas tersebut menjadi kelas data baru atau data yang akan dievaluasi.

#### **2.2.4. Daya Listrik Rumah Tangga**

Daya listrik merupakan laju hantaran yang dikeluarkan atau daya untuk melakukan usaha serangkaian listrik. Manfaat penggunaan daya listrik dalam kehidupan sehari-hari atau rumah tangga antara lain:

- a. Sebagai penerangan, misalnya penggunaan lampu di rumah, di jalan, di taman, dan lain-lain.
- b. Sumber energi, misalnya penggunaan alat elektronik seperti kulkas, televisi, laptop, dan lain-lain.
- c. Menghasilkan panas, misalnya penggunaan kompor listrik, setrika, dan penanak nasi.

Daya listrik golongan rumah tangga berkisaran 450 VA, 900 VA, 1300 VA, dan 2200 VA, daya listrik digunakan menyesuaikan dengan kebutuhan rumah tangga, secara umum daya yang digunakan 450 VA dan 900 VA, untuk

daya 1300 VA sampai 2200 VA biasanya memiliki kebutuhan khusus, seperti memiliki warung, bengkel, atau jasa lainnya. (Putra & Mukhaiyar, 2020), (PT PLN (Persero), 2022).

### **2.2.5. Subsidi Listrik**

Subsidi merupakan bentuk sebuah bantuan dari pemerintah baik berupa uang atau barang. Tujuannya untuk menjaga kestabilan harga, menutupi kerugian yang diderita perusahaan dan lain-lain. Data perincian subsidi mencakup subsidi bahan bakar, pupuk, dan salah satunya yaitu listrik. (Badan Pusat Statistik, 2022).

Pelanggan yang memperoleh bantuan dari pemerintah dikategorikan sebagai pelanggan subsidi, dan di luar itu dikategorikan sebagai pelanggan non subsidi. Sehingga pelanggan yang memperoleh subsidi listrik dapat membayar tagihan listrik lebih rendah dari pelanggan non subsidi, dan selisih antara pelanggan subsidi dengan non subsidi akan ditanggung oleh pemerintah yang kemudian dibayarkan ke pihak PLN. (Siaran Pers, 2022).

### **2.2.6. Framework**

Framework merupakan kerangka kerja dalam pengembangan webiste, sehingga memudahkan web developer dalam penulisan kode dan terstruktur. Fungsi penggunaan framework antara lain, (Setiawan, 2021):

- a. Kode program lebih terstruktur.
- b. Meningkatkan keamanan.

- c. Hemat waktu.
- d. *Maintenance* lebih mudah.

Dalam pengembangan website, penggunaan framework terdapat beberapa jenis dan memiliki fungsi masing-masing, yaitu:

- a. Framework CSS

Framework jenis ini merupakan kerangka yang digunakan dalam mengatur *layout* dan tampilan pada HTML sehingga menjadi lebih menarik, contohnya Bootstrap, Semantic UI, Materialize, UIKit, dan lain-lain.

- b. Framework JavaScript

Framework jenis ini merupakan kerangka yang berjalan pada sisi *front-end* yang digunakan dalam pembuatan *interface* menjadi lebih *user-friendly*, contohnya AngularJS, VueJS, ReactJS, NodeJS dan lain-lain.

- c. Framework PHP

Framework jenis ini merupakan kerangka yang berjalan pada sisi server atau *back-end* dan digunakan sebagai komunikasi dari sisi server tersebut, contohnya CodeIgniter, Laravel, Symfony dan lain sebagainya.

### **2.2.7. Laravel**

Laravel adalah sebuah framework PHP yang dirilis di bawah lisensi MIT dan dibangun dengan konsep MVC (*Model View Controller*). Konsep MVC menjadikan file program menjadi lebih terstruktur dan lebih mudah dimodifikasi isinya serta mempermudah developer membangun aplikasi sejenis. Penerapan website dengan konsep *Model View Controller* untuk

memudahkan pengelolaan website dan pengembangan secara modular. Konsep MVC dipisah-pisahkan menjadi tiga bagian yang terpisah menjadi saling terhubung. Bagian itu adalah *Model, View, dan Controller*. Berikut dasar-dasar laravel, (Mediana & Nurhidayat, 2018):

a. *Artisan*

*Artisan* adalah perintah yang dijalankan pada terminal selama melakukan pengembangan dan pembuatan sistem atau aplikasi, seperti *php artisan serve*, bertujuan untuk membuka web yang dibuat tanpa menggunakan web lokal.

b. *Routing*

*Routing* adalah proses mengatur lalu lintas file berdasarkan *request* user, sehingga halaman-halaman web terbuka berdasarkan keinginan user.

c. *Controller*

*Controller* adalah proses yang bertugas mengambil *request*, melakukan inisialisasi, dan memanggil model untuk dikirimkan pada view, dan *request* yang dibuat pada *controller* kemudian di lempar melalui *routing* supaya mendapatkan *request user* yang diinginkan.

d. *View (Blade Templating)*

*Blade* merupakan tema bawaan laravel yang digunakan untuk memisahkan elemen-elemen kodingan supaya tidak ditulis secara berulang, misalnya pemisahan kode *navbar, header* dan *footer*.

e. *Middleware*

*Middleware* adalah penengah antara *input request* dengan *controller*, sehingga dengan penggunaan *middleware*, aplikasi menjadi terjamin keamanannya karena untuk mengakses aplikasi diperlukan *request* seperti login dan autentikasi.

f. *Session*

*Session* adalah cara yang digunakan dalam penyimpanan pada server secara sementara, contohnya proses login dan pengecekan login.

### **2.2.8. Bootstrap**

Bootstrap merupakan sebuah *library* framework CSS yang telah dibuat khusus untuk mengembangkan *front-end* sebuah website. Bootstrap juga dikenal sebagai salah satu framework CSS, HTML, Javascript yang begitu populer di kalangan website developer atau pengembang website. Bootstrap digunakan untuk mengembangkan website agar lebih responsive. Dengan adanya bootstrap tersebut tentu saja membuat halaman website dapat menyesuaikan dengan ukuran monitor device. Baik jika diakses lewat ponsel, tablet ataupun desktop. (Jeklin Harefa, 2019).

### **2.2.9. Python**

Bahasa pemrograman python merupakan salah satu bahasa pemrograman populer dan banyak digunakan dalam berbagai bidang, termasuk analisis data. Python menawarkan kemudahan penggunaan dan fleksibilitas dalam penelitian atau pengembangan. Pada analisis data, python berproses sebagai *mining data*



dalam berbagai hal berdasarkan suatu metode yang diimplementasikan pada dataset sehingga mampu menghasilkan suatu informasi, semakin kompleks metode yang digunakan semakin mendalam wawasan yang diperoleh dari data, namun dibutuhkan juga sumber daya manusia dan teknologi yang canggih. Dengan demikian, python memberikan alat dan teknik yang memudahkan proses *data mining* secara efektif dan efisien. (Lo, et al., 2023).

#### **2.2.10. MySQL**

MySQL merupakan singkatan dari *My Structured Query Language*. SQL merupakan bahasa terstruktur yang khusus digunakan untuk mengolah *database*. MySQL merupakan sistem manajemen *database* yang bersifat relasional, yaitu data yang dikelola dalam *database* akan diletakkan pada beberapa tabel yang terpisah sehingga manipulasi atau proses pengorganisasian data akan jauh lebih cepat (Nirsal, Rusmala, & Syafriadi, 2020).