

Implementasi Data Mining Penentuan Daya Pelanggan Baru untuk Klasifikasi Subsidi dan Non Subsidi di Wilayah PLN Kota Lhokseumawe

Andik Bintoro¹, Safwandi²

Program Studi Teknik Elektro, Program Studi Teknik Informatika

Universitas Malikussaleh

Jl. Cot Tgk Nie-Reulet, Aceh Utara

email : andik@unimal.ac.id¹, safwandihst@gmail.com²

ABSTRAK

Penerapan model data mining dalam klasifikasi K-Nearest Neighbor (*KNN*) dalam penentuan Daya Pelanggan Baru untuk Klasifikasi Subsidi dan Non Subsidi di Wilayah PLN Kota Lhokseumawe sangatlah perlu untuk dilakukan. Hal ini untuk mengetahui pelanggan PLN dalam klasifikasi pengelompokan berdasarkan dengan variable yang digunakan. Pelanggan baru PLN yang ingin mengetahui berapa klasifikasi pengelompokan listrik untuk pasang baru dapat menggunakan aplikasi ini dalam penentuan besaran daya yang digunakan untuk pasang baru. Variable yang digunakan adalah untuk K1 adalah pendapatan keluarga secara keseluruhan, K2 adalah pekerjaan orang tua, K3 adalah pekerjaan orang tua, K4 adalah luas rumah. Dari variable ini akan dimasukkan kedalam sistem dengan sub kriteria yang akan dimasukkan. Adapun Golongan tarif daya listrik subsidi terbagi menjadi dua yaitu Tarif R-1/450 VA untuk penggolongan K1 dan golongan R-1/900 VA Subsidi Tarif R-1/586 kWh untuk penggolongan K2, sedangkan Golongan K3 non subsidi Tarif R-1/1352 VA untuk rumah tangga mampu. Tujuan dari penelitian ini adalah memudahkan pelanggan baru dalam penentuan golongan yang diberikan oleh pihak PLN kota lhokseumawe dan pelanggan mengetahui besaran yang akan digunakan. Sehingga pelanggan kota lhokseumawe dapat melihat besaran jumlah daya yang diberikan. Model klasifikasi *KNN* menggunakan data training untuk di uji coba dengan data pengujian untuk pelanggan baru dan hasilnya dari nilai terdekat dari data training yang telah diujikan. Dalam penelitian ini data training yang digunakan adalah 30 data training, hal ini berguna untuk nilai pengujian lebih akurat dan hasil yang diharapkan menjadi lebih baik. Untuk pengujian V1 adalah dengan nilai 3, untuk kriteria V2 adalah 3, dan nilai variabel V=3 adalah 2 dan nilai variabel terakhir V4 adalah 1. Selanjutnya untuk hasil pengujian dari pelanggan baru adalah termasuk kedalam golongan subsidi R-1/900 VA Subsidi Tarif R-1/586 kWh yang telah dilihat dari hasil training. Hasil dari aplikasi ini dapat menampilkan besaran daya yang digunakan untuk pelanggan baru berdasarkan nilai pengujian yang dilakukan.

Kata kunci : *Klasifikasi, Daya Listrik, KNN*

1.Pendahuluan

Penentuan daya pelanggan baru dapat dilihat dengan menggunakan model data mining. Salah satu bagian dari data mining dapat menggunakan dengan model klasifikasi K-Nearest Neighbor (*KNN*). Model klasifikasi ini dapat digunakan dalam penentuan daya pelanggan baru untuk klasifikasi subsidi dan non subsidi di wilayah kota lhokseumawe. Sehingga pelanggan yang ingin pasang baru dapat langsung mengetahui berapa besaran golongan yang diberikan. Teknik data mining yang bisa dapat digunakan untuk pengolahan data menjadi sumber informasi biasanya menggunakan klasifikasi. Teknik data mining pada umumnya digunakan oleh instansi, perusahaan dan pelanggan baru dan dapat menyebarkan informasi yang dapat mendukung pimpinan dalam mengambil tindakan dalam penentuan pengelompokkan daya listrik yang digunakan.

Hal ini sangat berpengaruh dalam melihat pengelompokkan pelanggan baru serta sesuai dengan daya yang diberikan, akan tetapi apabila pelanggan baru juga dapat melihat dan daya yang diberikan yang tidak sesuai dalam kriteria yang dimasukkan, pelanggan langsung dapat menanyakan kepada petugas PLN kota lhokseumawe yang disesuaikan dengan kebutuhan dari masing-masing pelanggan baru.

Data klasifikasi *KNN* dapat disimpan dalam suatu database untuk data training, sehingga pemrosesan dapat dilakukan dengan cepat. Pelanggan baru PT PLN kota lhokseumawe langsung dapat melihat hasil pemrosesan dengan *KNN* tanpa harus ada kendala teknis pada saat melakukan data pengujian. Selanjutnya apabila terjadi kesalahan yang dilakukan pada saat mengolah data klasifikasi tersebut. Sistem langsung dapat menangani apabila ada pengujian eror dan pihak dari pelanggan yang mengatasi tidak dapat dirugikan. Hal ini sangat membantu pelanggan lama dan baru dalam melihat hasil dari model *KNN* tersebut. Kriteria yang digunakan dalam penentuan besaran daya yang diberikan pada implementasi data mining untuk pelanggan baru untuk klasifikasi subsidi dan non subsidi di wilayah pln kota lhokseumawe meliputi K1 adalah pendapatan keluarga secara keseluruhan, K2 adalah pekerjaan orang tua, K3 adalah jumlah ruang, K4 adalah luas rumah

. Dari variable ini akan dimasukkan kedalam sistem dengan sub kriteria yang akan dimasukkan. Adapun Golongan tarif daya listrik

subsidi terbagi menjadi dua yaitu Tarif R-1/450 VA untuk penggolongan K1 dan golongan R-1/900 VA Subsidi Tarif R-1/586 kWh untuk penggolongan K2, sedangkan Golongan K3 non subsidi Tarif R-1/1352 VA untuk rumah tangga mampu., selanjutnya alternatif penggolongan daya yang diberikan pada calon nasabah terbagi menjadi 3 siklus yaitu alternatif besaran dayayang diberikan.

2. Tinjauan Pustaka

2.1 Data Mining

Data mining dapat diartikan untuk menemukan pengetahuan yang tersembunyi di dalam database. Data mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam database besar. (Turban et al, 2005).

Data Mining juga dapat mengeksplorasi basis data untuk menemukan pola-pola yang ada apada aturan dan mencari informasi untuk dapat mengklasifikasi pelanggan baru dalam penentuan klasifikasi penggolongan daya listrik.

Data mining adalah sebuah istilah yang digunakan untuk menggambarkan penemuan ilmu pengetahuan dalam bidang database, sebuah bidang analisis informasi yang mencari pola tersembunyi dalam sekelompok data yang dapat digunakan untuk memprediksi perilaku masa depan (Turban et al, 2007:202).

Data mining adalah proses menemukan hubung dalam data yang tidak diketahui oleh pengguna dan menyajikannya dengan cara yang dapat dipahami hingga hubungan tersebut dapat menjadi dasar pengambilan keputusan (McLeod, 2007).

2.1 Algoritma K-Nearest Neighbour

Algoritma *K-Nearest Neighbor* merupakan salah satu metode yang digunakan dalam pengklasifikasian. Prinsip kerja *K-Nearest Neighbor* (KNN) adalah mencari jarak terdekat antara data yang akan dievaluasi dengan *K* tetangga (*neighbor*) terdekatnya dalam data pelatihan. Teknik ini termasuk dalam kelompok klasifikasi *nonparametric*. Di sini kita tidak memperhatikan distribusi dari data yang ingin kita kelompokkan. Teknik ini sangat sederhana dan mudah diimplementasikan. Mirip dengan teknik

klustering, kita mengelompokkan suatu data baru berdasarkan jarak data baru itu ke beberapa data/tetangga (*neighbor*) terdekat (Kusrini, 2009).

Algoritma k-NN menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari sampel uji yang baru. Jarak yang digunakan adalah jarak Euclidean Distance (Han J and Kamber, 2011). Jarak Euclidean adalah jarak yang paling umum digunakan pada data numeric.

Algoritma *K-Nearest Neighbor* adalah metode yang sering digunakan dalam pengklasifikasian untuk nasabah baru. Model *K-Nearest Neighbor* (KNN) merupakan mencari jarak terdekat antara data *K* tetangga (*neighbor*) dengan yang akan dievaluasi dengan berdasarkan nilai terdekatnya dalam data pelatihan. Teknik data mining sangat mudah dan penerapannya juga mudah misalnya mengelompokkan suatu data baru berdasarkan jarak data baru itu ke beberapa data/tetangga (*neighbor*) terdekat (Kusrini dan Luthfi, 2009).

Algoritma *K-Nearest Neighbor* yang diambil adalah Nilai *k* dimana nilai *k* tersebut terbaik dalam algoritma ini yang semuanya tergantung pada data training. Secara umum, nilai *k* yang tinggi akan mengurangi efek *noise* pada klasifikasi. Kasus khusus dimana klasifikasi diprediksikan berdasarkan training data yang paling dekat disebut algoritma *K-Nearest Neighbor*.

Algoritma *K-Nearest Neighbor* sebagai berikut :

Rumus KNN:

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

Keterangan:

x_1 = Data Uji

x_2 = Data Sampel

i = Variabel Data

d = Jarak

p = Dimensi Data

2.2 BASIS DATA

Basis data adalah "mekanisme yang digunakan untuk menyimpan informasi atau data. basis data sebagai kumpulan data, umumnya alah satu komponen yang paling penting di dalam sistem informasi karena

merupakan dasar dalam menyediakan informasi bagi pemakai. Tujuan dari **desain Database** yaitu untuk menentukan data-data yang dibutuhkan dalam sistem , sehingga informasi yang nantinya akan dihasilkan dapat terpenuhi dengan baik." (Abdul Kadir, 2008)".

Database adalah kumpulan dari item data yang saling berhubungan satu sama lain, yang diorganisasikan berdasarkan sebuah skema atau struktur tertentu, tersimpan di hardware computer, dan harus menggunakan software untuk melakukan manipulasi tertentu.

4.3 Klasifikasi

Klasifikasi merupakan pengelompokan data atau objek baru berdasarkan variabel yang diamati dengan tujuan untuk memprediksi/klasifikasi suatu objek baru yang masih belum diketahui kelas atau kategorinya

K-Nearest Neighbor (KNN) merupakan model yang sering digunakan pada klasifikasi data. Algoritma ini digunakan untuk mengklasifikasikan terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. KNN merupakan suatu metode yang menggunakan algoritma *supervised* dengan hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada KNN. Tujuan dari algoritma ini ialah mengklasifikasikan objek baru berdasarkan atribut dan *training sample* (Larose, 2006).

Prinsip umum dari algoritma ini adalah menemukan k data training untuk menentukan *k-nearest neighbor* berdasarkan ukuran jarak. Selanjutnya mayoritas dari k tetangga terdekat akan menjadi dasar untuk memutuskan kategori dari sample.

Klasifikasi merupakan proses mengidentifikasi obyek ke dalam sebuah kelas, grup, atau ketegori berdasarkan prosedur, karakteristik & definisi yang telah ditentukan sebelumnya (KNN dilakukan dengan mencari kelompok k objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data *testing* Han dan Kamber. 2006).

3. Metodologi Penelitian

3.1 Tahapan-Tahapan Penelitian

3.1.1 Jenis Data

Penelitian ini penggunaan data primer yaitu data observasi langsung di PT PLN Lhokseumawe dan wawancara langsung dengan pelanggan baru. Sample-sample variabel adalah adalah dua yaitu Tarif R-1/450 VA untuk penggolongan K1 dan golongan R-1/900 VA Subsidi Tarif R-1/586 kWh untuk penggolongan K2, sedangkan Golongan K3 non subsidi Tarif

R-1/1352 VA untuk rumah tangga mampu. Kemudian diproses menggunakan Algoritma K-Nearest Neighbour untuk pasang baru dan penambahan daya.

3.1.2 Analisis Pengumpulan Data

Metode Analisis data penelitian ini menggunakan analisis data yaitu prosedur pencatatan untuk data yang diteliti pada PT PLN Lhokseumawe mengenai pengumpulan informasi terperinci mengidentifikasi masalah membuat evaluasi dalam menentukan variabel penentuan pemberian daya listrik dan jenis variabel termasuk dalam klasifikasi Algoritma K-Nearest Neighbour.

3.1.3 Analisis Sumber Data

Data diperlukan untuk menghasilkan informasi yang baik, karena informasi pada dasarnya merupakan hasil dari pengolahan data yang diinputkan pada sistem.

1. Data internal

Merupakan data yang berasal dari dalam organisasi untuk mendukung peramalan dan data mining klasifikasi, Data penilaian masing-masing adalah data kebutuhan dari masing-masing kebutuhan stok listrik daerah untuk pelanggan lama dan pemasangan baru.

2. Data eksternal

Merupakan data yang berasal dari luar organisasi atau perusahaan namun tetap memiliki pengaruh dalam menciptakan data mining yang baik. Adapun beberapa data eksternal yang mempengaruhi pengambilan keputusan pada peramalan dan klasifikasi data mining.

3.2. Model yang Digunakan

Identifikasi variabel/kriteria yang akan digunakan dalam penentuan klasifikasi jenis pengelompokan daya listrik bagi pasang baru. Model yang digunakan dalam proses klasifikasi memprediksi pengelompokan kebutuhan daya listrik pasang baru menggunakan model K Nearest Neighbors.

3.3 Analisis Data

Data yang digunakan untuk implementasi perangkat lunak ini dibagi menjadi tiga bagian utama yaitu sebagai berikut : (1) Data Masukan(Input) yaitu input untuk melakukan proses pengambilan variable yang akan dimasukkan kedalam sistem. Selanjutnya proses Algoritma K-Nearest Neighbour adalah untuk menentukan hasil klasifikasi jarak terdekat

ketetangaan. berdasarkan klasifikasinya dalam pemberian daya yang diperlukan.

4.HASIL DAN PEMBAHASAN

4.1 Implementasi Sistem

1. Tampilan Menu Utama

Adapun tampilan menu utama dalam implementasi data mining penentuan daya pelanggan baru untuk klasifikasi subsidi dan non subsidi di wilayah PLN kota lhokseumawe adalah sebagai berikut :



Gambar 4.1 Tampilan Menu Utama

2.Input Data Training

Adapun tampilan menu input data training dalam implementasi data mining penentuan daya pelanggan baru untuk klasifikasi subsidi dan non subsidi di wilayah PLN kota lhokseumawe adalah sebagai berikut :

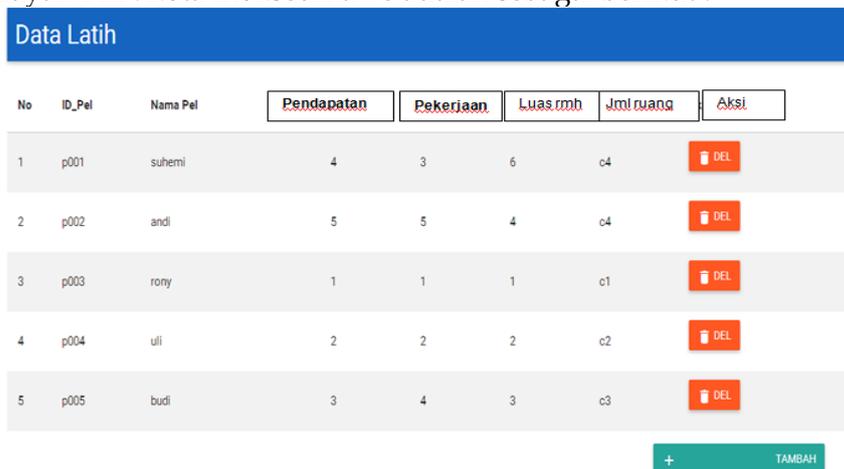
[Tambah Data Training](#)
===== Tambah Data Training =====
account_circle Input ID_Pelanggan
account_circle Input Nama
looks_one Input V1
looks_two Input V2
looks_3 Input V3
looks_4 Input Kelas

[Batal](#)

Gambar 4.2 Input Data Pelanggan

3. Data Latih

Adapun tampilan menu data latih dalam implementasi data mining penentuan daya pelanggan baru untuk klasifikasi subsidi dan non subsidi di wilayah PLN kota lhokseumawe adalah sebagai berikut :



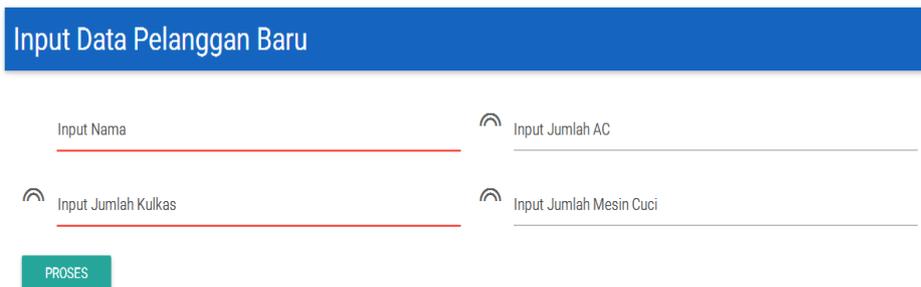
No	ID_Pel	Nama Pel	Pendapatan	Pekerjaan	Luas.rmh	Jml.ruang	Aksi
1	p001	suhermi	4	3	6	c4	DEL
2	p002	andi	5	5	4	c4	DEL
3	p003	rony	1	1	1	c1	DEL
4	p004	uli	2	2	2	c2	DEL
5	p005	budi	3	4	3	c3	DEL

+ TAMBAH

Gambar 4.3 Data Latih

4. Input Data Pelanggan Baru

Adapun tampilan input data pelanggan baru dalam implementasi data mining penentuan daya pelanggan baru untuk klasifikasi subsidi dan non subsidi di wilayah PLN kota lhokseumawe adalah sebagai berikut :



Input Data Pelanggan Baru

Input Nama

Input Jumlah AC

Input Jumlah Kulkas

Input Jumlah Mesin Cuci

PROSES

Gambar 4.4 Input Data Pelanggan Baru

5. Input Data Training

Adapun tampilan input data training dalam implementasi data mining penentuan daya pelanggan baru untuk klasifikasi subsidi dan non subsidi di wilayah PLN kota lhokseumawe adalah sebagai berikut :

Nilai K adalah sebesar 3

No	id_pel	nama	V1	V2	V3	Kelas	Jarak
1	p005	budi	3	4	3	c3	1.41421
2	p004	uli	2	2	2	c2	1.41421
3	p003	rony	1	1	1	c1	3

Gambar 4.5 Input Data Training

6. Perhitungan K-NN

Adapun tampilan perhitungan KNN dalam implementasi data mining penentuan daya pelanggan baru untuk klasifikasi subsidi dan non subsidi di wilayah PLN kota lhokseumawe adalah sebagai berikut :

Perhitungan K-NN

Menghitung jarak antar kasus

No	id_pel	nama	V1	V2	V3	Kelas	Jarak
1	p001	suhemi	4	3	6	c4	3.60555
2	p002	andi	5	5	4	c4	3.74166
3	p003	rony	1	1	1	c1	3
4	p004	uli	2	2	2	c2	1.41421
5	p005	budi	3	4	3	c3	1.41421

Gambar 4.6 Input Data Pelanggan Baru

5. Kesimpulan dan Saran

5.1 Kesimpulan

Kesimpulan Implementasi Data Mining Penentuan Daya Pelanggan Baru untuk Klasifikasi Subsidi dan Non Subsidi di Wilayah PLN Kota Lhokseumawe adalah sebagai berikut:

1. Dengan adanya penentuan daya pelanggan baru untuk klasifikasi subsidi dan non subsidi dapat memudahkan pelanggan baru dalam melihat klasifikasi besaran daya yang diberikan.
2. Adanya penerapan sistem yang dapat di rekomendasikan kepada pimpinan PT PLN Lhokseumawe dan sistem dapat menampilkan

besaran daya dalam penentuan daya pelanggan baru untuk klasifikasi subsidi dan non subsidi di wilayah pln kota lhokseumawe.

5.2 Saran

1. Dalam penentuan klasifikasi kebutuhan daya listrik untuk masing-masing daerah di kota lhokseumawe lebih baiknya menggunakan penggabungan model cluster
2. Program selanjutnya lebih baik digunakan cluster untuk masing-masing daerah yang ada di kota lhokseumawe.

Daftar Pustaka

- D. T. Larose, DATA MINING METHODS AND MODELS, New Jersey: John Wiley & Sons, Inc, 2006.
- Larose, D.T. (2005). Discovering Knowledge in Data. New Jersey :John Willey & Sons Inc.
- Kusrini, M.Kom. 2007. Strategi Perancangan dan Pengelolaan Basis Data: CV. Andi Offset.
- Kusrini, Luthfi, Emha Taufiq, Algoritma Data Mining, Penerbit ANDI, Yogyakarta, 2009. Xu, R., & Wunsch II, D. C. (2009). Clustering. Kanada: IEEE Press.
- McLeod, Jr.R. dan G.P. Schell. 2007. Management Information System. 10th ed. Pearson Education, Inc.
- Han, J., Kamber M., Pei, J., 2011, Data Mining : Technique and Concepts 3rd ed., San Fransisco, CA, USA : Morgan Kaufmann.
- Han, J., and Kamber M., 2006, Data Mining: Concepts and Techniques 2e, Morgan Kaufmann Publishers , San Francisco.
- Turban, E., Aronson, J. E., Liang, T. P., & Sharda, R.(2007). Decision support and business intelligencesystems (Eighth ed.). Pearson Education
- Turban, E, 2005, Decision Support Systems and Intelligent Systems Edisi Bahasa Indonesia Jilid 1, Andi, Yogyakarta.