

IDENTIFIKASI PROFIL KONSUMSI ENERGI LISTRIK UNTUK MENINGKATKAN PENDAPATAN DENGAN KLUSTERING

Mirza Hamdhani^{a,b}, Diana Purwitasari^{c,*}, Agus Budi Raharjo^c

^aMagister Manajemen Teknologi, Institut Teknologi Sepuluh Nopember

^bUnit Layanan Pelanggan Rappang, PT. PLN (Persero)

^cTeknik Informatika, Institut Teknologi Sepuluh Nopember

E-mail: mirzahamdhani@gmail.com, diana@if.its.ac.id, agus.budi@its.ac.id

Abstrak— Ketersediaan energi listrik pada sistem sulsel lebih dari cukup yakni 602 MW. Sejalan dengan surplusnya energi listrik, kantor pusat memberikan program program peningkatan penjualan kepada unit-unit layanan pelanggan untuk dijalankan. Program tersebut belum memberikan hasil yang baik untuk *Key performance indicator* penjualan tenaga listrik, karna bahwasannya program tersebut diberikan secara umum untuk seluruh unit layanan pelanggan tanpa memperhatikan kondisi pasar dan karakter pelanggan yang dimiliki unit layanan. Profil konsumsi energi listrik sangat penting untuk mendukung pengembangan strategi pemasaran yang dipersonalisasi agar tepat sasaran. Identifikasi profil konsumsi listrik dapat menunjukkan karakteristik pemakaian energi listrik tiap pelanggan. Pada penelitian ini *clustering* dilakukan permodelan melalui pengolahan data profil konsumsi listrik ditunjukkan dengan *variable daya*, pemakaian energi, penambahan pelanggan bulanan dari tahun 2019-2021. Selanjutnya dari hasil *clustering* tersebut menggali informasi karakteristik tiap klusternya untuk dijadikan informasi strategi pemasaran. Diharapkan dari penelitian ini mendapatkan model karakteristik profil konsumsi energi listrik. Hasil dari metode *sum of square error* mendapatkan $k=3$ dengan rasio 1,57. *Cluster_1* adalah pelanggan dengan kontribusi rupiah penjualan terendah yakni secara kumulatif hanya memberikan 18,5%. *Cluster_2* berkontribusi sedang secara kumulatif pada rupiah pendapatan yakni sebesar 34,86%. *Cluster_3* berkontribusi paling tinggi secara kumulatif pada rupiah pendapatan yakni sebesar 46,63%. Pada kelompok pelanggan yang berkontribusi terendah perlu dilakukan pemeriksaan persil pelanggan untuk memastikan pemanfaatan energi listrik dan mencurigai adanya pelanggaran penyaluran energi listrik. Pada kelompok pelanggan kontribusi sedang diberikan pendampingan dengan pengenalan alat elektronik dengan manfaatnya. Kemudian pada pelanggan kontribusi besar dapat diberikan layanan pendampingan dalam rangka menjaga loyalitas pelanggan, serta memberikan layanan informasi terkait tghian listrik. Teknik *k-means clustering* memberikan kemudahan identifikasi karakteristik pelanggan dan visualisasi yang baik untuk perusahaan menganalisa yang kemudian memberikan informasi rekomendasi kebijakan dan strategi peningkatan pendapatan.

Kata Kunci— *Clustering*, Energi Listrik, K-Means, Profil Kelistrikan, *Sum Of Square Error*.

Naskah Masuk : 04 Juli 2022

Naskah Direvisi : 25 Juli 2022

Naskah Diterima : 26 Juli 2022

*Corresponding Author : diana@if.its.ac.id



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

I. PENDAHULUAN

Sistem kelistrikan Sulawesi Selatan memiliki cadangan daya hingga 602 MW (data tahun 2021). Hal ini menyatakan bahwa kepastian ketersediaan energi lebih dari cukup untuk sistem Sulawesi Selatan. *Uncertainty* dan *disrupsi bisnis* membuat PT. PLN (Persero) memiliki banyak tantangan dari para kompetitor yakni *Independent Power Producer* (IPP) maupun perusahaan layanan kelistrikan mandiri (*captive power*). Hadirnya perusahaan-perusahaan tersebut memberikan adanya perubahan aturan-aturan dalam industri layanan kelistrikan yang didukung dengan Undang-Undang Dasar (UUD) Nomor 30 Tahun 2009. UUD tersebut ini membuka pintu kesempatan bagi perusahaan-perusahaan layanan kelistrikan untuk dapat menyediakan listrik secara mandiri. Dampak dari hadirnya UUD 30 tahun 2009 membuat persaingan layanan kelistrikan menjadi sangat terbuka dan sangat kompetitif [1].

Kondisi saat ini cadangan daya yang tersedia sangat besar, namun KPI penjualan tenaga listrik saat ini belum tercapai yakni hanya 78%. KPI penjualan merupakan nilai ukur terhadap energi yang terjual kepada pelanggan. Tentunya hal ini sangat [2] merugikan perusahaan. Upaya-upaya meningkatkan penjualan telah dilakukan dengan berbagai macam cara, yakni memasarkan melalui social media, iklan baliho, dengan menyampaikan promo-promo tambah daya. Anggaran biaya pemasaran ini sangat besar namun tidak memberikan pencapaian yang baik pada KPI. Berdasarkan hal ini perlu dilakukan analisa kembali dengan membedah data penjualan tenaga listrik per pelanggan untuk melihat karakteristik pola konsumsi pelanggan agar dapat memberikan informasi khusus terkait program ataupun layanan yang dapat diberikan ke pelanggan dalam rangka meningkatkan penjualan tenaga listrik.

Pendapatan PLN berasal dari penjualan tenaga listrik kepada para pelanggannya, baik sektor industri maupun pelanggan rumah tangga. Terdapat 2 jenis cara pengukuran konsumsi energi listrik PLN, yakni dengan meter pascabayar dan meter Prabayar. Pengukuran meter pascabayar dilakukan setiap bulan, sedangkan meter Prabayar berdasarkan kemampuan pelanggan untuk membeli token pulsa listrik.

Pada meter pascabayar, setiap meter pelanggan PLN akan dibaca untuk mendapatkan konsumsi energi listrik pelanggan. Perolehan konsumsi energi listrik didapatkan dengan 3 cara seperti yang ditunjukkan Gambar 1, yakni

dengan petugas baca meter, swacam atau yang lebih dikenal baca meter mandiri oleh pelanggan, serta pembacaan meter secara otomatis menggunakan pengiriman data melalui internet. Sedangkan meter Prabayar hanya di *record* oleh sistem AP2T berdasarkan setiap pembelian token listrik pelanggan.

Data penjualan pelanggan listrik meliputi daya terpasang dalam *watt*, jumlah pemakaian daya dalam satuan *watt hour*, rupiah penjualan tenaga listrik, durasi jam nyala pemakaian energi dalam sebulan dan golongan tarif yang dikenakan pada pelanggan [3]. Data data tersebut mempunyai karakteristik pola konsumsi tiap pelanggan listrik yang dapat memberikan informasi penjualan listrik kepada perusahaan [2].

Metode *clustering* dapat mengelompokkan populasi pelanggan ke dalam kelompok-kelompok tertentu dengan tingkat kemiripan dari variabel input yang diberikan. Misalnya, mengelompokkan pelanggan ke dalam beberapa cluster dengan pola karakteristik yang serupa atau sifat kuat dari variabel input tertentu.. Metode pendekatan *clustering* menggunakan karakteristik beban energi listrik dapat digunakan untuk menyimpulkan prediktor dasar hipotesis tertentu [4]. Pada penelitian Amin Rajabi mengatakan untuk mendapatkan manfaat bagi berbagai pemangku kepentingan dalam sistem tenaga listrik dia meneliti pola konsumsi listrik dari beban pemakaian rumah tangga untuk mengetahui perbedaan pola penggunaan energi dalam sebulan dari data konsumsi listrik hariannya yang tren masa depan dan aplikasi utama pola konsumsi pengelompokan diuraikan untuk menginformasikan praktisi industri dan peneliti akademis untuk mengoptimalkan penggunaan dan efektivitas operasional *smart meter* dan segmentasi pelanggan berdasarkan pola konsumsi tenaga listrik sangat penting untuk mendukung pengembangan strategi pemasaran yang dipersonalisasi dan tepat sasaran serta peningkatan efisiensi energi [5]. Selanjutnya penelitian S.Yilmaz menggunakan pendekatan *clustering* konsumsi energi listrik rumah tangga dengan metode K-Means, memberikan informasi implikasi pola permintaan energi pelanggan listrik yang dapat dijadikan informasi pembuat kebijakan tarif berdasarkan waktu penggunaan [6].

Studi kasus penelitian ini dilaksanakan dengan metode *clustering* K-means. Data uji menggunakan data pelanggan PLN Unit Layanan Pelanggan Rappang dari tahun 2019-2021 dengan variabel uji yakni daya terpasang, pemakaian energi, rupiah penjualan tenaga listrik, dan Jam nyala agar dapat memberikan interpretasi karakteristik dalam identifikasi profil konsumsi energi listrik.

II. TINJAUAN PUSTAKA

Saat ini PT. PLN (Persero) memiliki pesaing pada bidang layanan kelistrikan. Pesaing tersebut merupakan *independent power producer* (IPP) dan Perusahaan yang melakukan bisnis di bidang yang sama, yaitu layanan kelistrikan mandiri (*captive power*). PT. PLN (Persero) harus berbenah dalam melakukan strategi penjualan yang efisien dan tepat sasaran agar dapat memenuhi kebutuhan energi dan menghasilkan *revenue* yang baik. Pada penelitian ini total data yang akan digunakan sebanyak 13.510 pelanggan. Setiap pelanggan tersebut memiliki masing-masing karakteristik konsumsi energi listrik dan tidak semua pelanggan dapat di sama ratakan teknik pemasarannya. Metode pengelompokan

pelanggan dapat berfokus pada pembagian target pasar menjadi kelompok-kelompok yang dapat dikelola sesuai dengan karakteristik konsumsi energi listrik. Metode pengelompokan bertujuan agar dapat mengembangkan strategi bisnis yang efektif dan tepat sasaran berdasarkan informasi kelompok tersebut [7]. Pembagian kelompok ataupun segmentasi ini dilakukan dengan membagi basis pelanggan yang ada menjadi kelompok yang dapat dikelola berdasarkan karakteristik umum pelanggan listrik [8].

Penelitian sebelumnya pada Table 1 memberikan perspektif cara untuk menganalisa profil konsumsi energi listrik serta perencanaan management dalam melayani pelanggan pada PT. PLN (Persero) ULP Rappang. Data uji dalam penelitian terdahulu tersebut menggunakan data pemanfaatan beban energi listrik dalam bentuk harian dengan deret waktu time series. Pada penelitian ini data yang digunakan yakni data penggunaan energi listrik bulanan pada pelanggan pascabayar, daya terpasang, rupiah pendapatan, serta jam nyala energi listrik pelanggan. Selanjutnya pemrosesan data dengan permodelan *clustering* dan kesiapan utilitas trafo distribusi yang melayani tiap pelanggan yang akan memberikan informasi bagaimana bentuk profil konsumsi untuk peningkatan strategi pemasaran, serta perencanaan persiapan distribusi penyaluran listrik ke pelanggan dari akurasi tertinggi dari model *cluster*.

A. Pemasaran

Pemasaran yang telah dilakukan memerlukan kajian analisis yang lebih dalam agar usaha pemasaran yang dilaksanakan tepat sasaran, efektif dan efisien. Layanan promo tambah daya listrik harus tepat sasaran kepada pelanggan yang memiliki tren peningkatan penggunaan, kemudian diupayakan dapat dilayani sesegera mungkin dan tanpa penambahan asset ini lah yang disebut tepat sasaran, efektif dan efisien [9]. Karena kurangnya analisis data profil konsumsi listrik, pengambilan keputusan pada skema program penambahan daya tidak sejalan dengan peningkatan penjualan yang tepat sasaran, efektif dan efisien [10].

Analisis pemasaran adalah cara mempelajari daya tarik pasar khusus dalam industri tertentu yang di dasari sajian data data informasi konsumsi pasar terhadap produk. Faktor faktor analisis pasar yang paling umum digunakan SWOT (*Strengths Weakness Opportunities dan Threats*) [11]. Data konsumsi energi listrik pelanggan yang dimiliki pada *database* PT PLN (Persero) akan digunakan untuk melihat karakteristik pola konsumsi yang akan dihubungkan dengan keputusan perusahaan dalam melaksanakan program *marketing*.

B. Metode Customer Relationship Management

Pada era saat ini pelanggan tidak lagi hanya dipandang sebagai *buyer* melainkan lebih dari pada hal itu Pengelolaan pelanggan saat ini telah berubah pola menjadi *Customer Relationship Management* (CRM). CRM strategi bisnis untuk memahami pelanggan, mengenali kebutuhan mereka, hingga mempengaruhi perilaku pelanggan menggunakan komunikasi antara pelanggan dengan perusahaan. Hal ini bertujuan untuk meningkatkan potensi penjualan, loyalitas pelanggan, dan profitabilitas pelanggan. Metode CRM menyediakan struktur pola untuk membangun hubungan yang kuat antara pelanggan dan perusahaan. Membina hubungan baik kepada pelanggan memerlukan informasi data dari pelanggan terkait keluhannya, kepuasannya, hingga

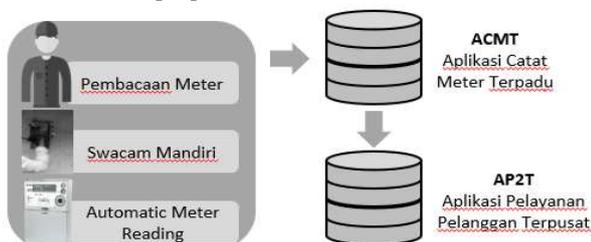
produk yang diinginkan. Menurut Abbott dan Julie terdapat 3 komponen utama pada CRM yakni *Operational, Analytical, Collaborative CRM* [12].

C. K-Means Clustering

Machine Learning merupakan teknik pengolahan data yang dapat mempelajari pola bentuk data yang tersedia. *Machine learning* dapat melakukan tugas-tugas tertentu sesuai dengan apa yang dipelajari. *Machine Learning* didefinisikan sebagai aplikasi komputer yang dijalankan menggunakan algoritma matematika [13]. Metode *Clustering* merupakan bagian daripada *machine learning*, bekerja dengan cara mengelompokkan data kedalam beberapa *cluster* berdasarkan unsur kesamaannya. Algoritma *clustering* disebut juga dengan metode *unsupervised learning*, yakni tanpa membutuhkan target. *Clustering* adalah pengolahan informasi data yang dikelompokkan menjadi beberapa cluster atau kelompok sehingga kesamaan data dalam cluster dimaksimalkan dan kesamaan data di setiap cluster diminimalkan. *Clustering* membagi sekumpulan objek data ke dalam kelompok-kelompok bagian yang disebut cluster. [14].

Pada pengelolaan bisnis saat ini *clustering* menjadi *tools* manajemen perusahaan dalam mempelajari data data pelanggan. *Clustering* membagi sekumpulan objek data ke dalam kelompok-kelompok bagian yang disebut cluster. Pada penelitian Q. Zhao, dkk *clustering* memberikan informasi menyeluruh serta kemampuan dalam menganalisa perilaku konsumsi listrik yang kemudian dapat memverifikasi efektivitas model *cluster* yang dikerjakan [15]. Demikian pula menghubungkan perilaku konsumsi energi listrik tinggi ataupun rendah dengan kombinasi tertentu dari daya terpasang, konsumsi energi listrik rupiah pendapatan penjualan energi listrik [4].

Algoritma k-means membutuhkan k parameter input dan membagi n objek menjadi k cluster untuk kesamaan antar anggota dalam cluster. Sedangkan pada tingkat kemiripan *cluster* satu dengan *cluster* lainnya haruslah rendah. Kesamaan anggota dalam suatu klaster diukur dari kedekatan objek dengan rata-rata dalam *cluster* tersebut. Atau, kadang-kadang disebut sebagai besarnya jarak objek dari pusat *centroid cluster*. [16].



Gambar, 1. Diagram Alir Perolehan Pengukuran meter Pascabayar

D. Metode Elbow

Metode *elbow* adalah teknik dalam menentukan jumlah *cluster* yang optimal dengan mengukur rasio jumlah *cluster* yang membentuk sudut pada suatu titik nilai k terhadap titik nilai k lainnya. Ukuran dari setiap perhitungan nilai k terhadap nilai k lainnya ditunjukkan dalam grafik yang selanjutnya membandingkan rasio penurunan *sum square error* (SSE). Nilai k dengan nilai k yang selanjutnya apabila menunjukkan signifikansi penurunan paling besar, maka nilai k tersebut menjadi k terbaik. Ini juga dapat dikonfirmasi

dengan membandingkan nilai rasio SSE dengan nilai maksimum. [17]. Untuk menghitung metode *elbow* :

1. Inisiasi nilai k dari proses algoritma looping k-means.
2. Berdasarkan hasil SSE, nilai K mengalami penurunan yang cukup signifikan.
3. Bandingkan nilai rasio hitung hasil SSE dari setiap nilai k.
4. Nilai hitung rasio SSE tertinggi dipilih menjadi pengelompokan sejumlah k tersebut.

TABEL I
PENELITIAN TERDAHULU

No.	Deskripsi Jurnal	Pembahasan
1	Judul Penelitian : <i>Comparison of clustering approaches for domestic electricity load profile characterisation - Implications for demand side management</i> Tahun Publikasi : 2019 Nama Peneliti : S. Yilmaz	Hasil Penelitian : Analisis kluster profil permintaan harian dalam rumah tangga dilakukan, menghasilkan kluster yang dapat menunjukkan bahwa setiap kluster profil berbeda secara signifikan dari profil rata-rata.
2	Judul Penelitian : <i>Clustering of residential electricity customers using load time series</i> Tahun Publikasi : 2019 Nama Peneliti : Omidh Motlagh	Hasil Penelitian : Pengelompokan pelanggan listrik mendukung segmentasi dan pengelolaan pasar yang efektif. Tantangan utamanya adalah penerapan proses yang tepat untuk mengurangi dimensi ekstrim dari deret waktu dalam pembuatan cluster.
3	Judul Penelitian Pengaruh Bauran Pemasaran Terhadap Peningkatan Penjualan Tenaga Listrik Pada PT PLN (Persero) Unit Pelaksana Pelayanan Pinrang. Tahun Publikasi : 2021 Nama Peneliti : Andi Muh. Armin Yusfin	Hasil Penelitian : Menemukan bahwa produk, harga, promosi, dan tindak lanjut implementasi lapangan berpengaruh positif dan signifikan terhadap peningkatan penjualan tenaga listrik berdasarkan hasil uji regresi.

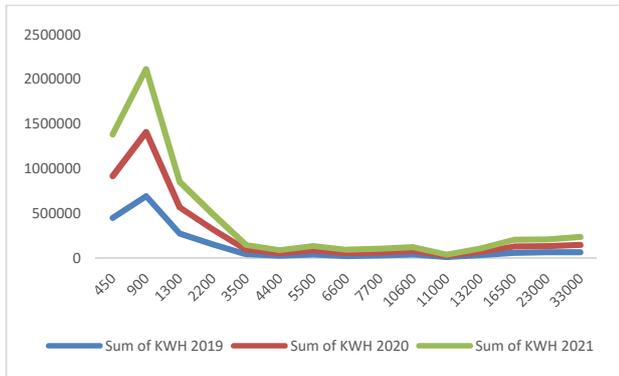
III. METODE PENELITIAN

Penelitian ini menggunakan CRISP-DM, model proses CRISP-DM memberikan gambaran tentang siklus hidup proyek *data mining*. CRISP-DM memiliki 6 tahapan yaitu Pemahaman terhadap bisnis yang dihadapi, Penjelasan sumber data, *Data Preparation, Modeling, Evaluation, dan Deployment* [18]. Penelitian akan menggunakan *Knime Analytical Tools* pemodelan *machine learning*. Model uji pada penelitian ini menggunakan k-means *clustering*, yang bertujuan untuk membentuk kelompok data konsumsi tenaga listrik dengan data set ditunjukkan pada Tabel III. K-means *algorithm* mampu menangani data konsumsi pelanggan listrik dengan mendapatkan kesamaan karakteristik pelanggan yang kemudian memberikan informasi pada perusahaan terkait perencanaan strategis perusahaan berdasarkan hasil *clustering* [17].

A. Bisnis Understanding

Situasi yang dihadapi yakni belum tercapainya kinerja penjualan tenaga listrik dan program program peningkatan penjualan yang belum tepat sasaran. Pada penelitian ini dimaksudkan untuk melakukan identifikasi profil karakteristik konsumsi energi pelanggan pada unit layanan pelanggan rapping, agar dalam menjalankan upaya peningkatan penjualan dapat didasari pada karakteristik kelompok pelanggan.

B. Pemahaman Sumber Data



Gambar. 2. Komulatif Penjualan Tenaga Listrik 2019-2021

Jenis data dalam penelitian ini merupakan data primer PT. PLN (Persero) ULP Rapping. yang didapat dan dikumpulkan langsung dari data penjualan tenaga listrik. Data yang dikumpulkan dari tahun 2019-2021 atau selama 3 tahun dengan total data sebanyak 13.866 Pelanggan.

Berdasarkan Gambar 2 diatas, dapat terlihat bahwasannya secara kumulatif data penjualan tenaga listrik per daya tersambung terjadi peningkatan penjualan disetiap daya pelanggan. Namun hal ini belum terpastikan populasi pelanggan yang berkontribusi meningkatkan penjualan dan populasi pelanggan yang justru menurun pemakaian energinya. Komposisi populasi pelanggan terbanyak yakni pada daya 450 VA, 900 VA, 1300 VA, 2200 VA, dan 3500 VA. Selebihnya merupakan pelanggan dengan daya yang besar namun poulasinya sedikit, sehingga kontribusi terhadap penjualan tidak begitu signifikan. Berikutnya yakni skema diagram alir penelitian yang ditunjukkan Gambar 3. Model penelitian ini menggunakan *unsupervised learning* dengan teknik *k-means clustering*.

C. Data Preparation

Teknik yang diterapkan pada database untuk menghilangkan nilai yang hilang, *noise*, dan informasi data yang tidak konsisten [19]. *Preprocessing* digunakan untuk pengolahan data yang tidak lengkap dan tidak konsisten seringkali mengakibatkan hasil *data mining* yang tidak akurat. *Preprocessing* data dalam penelitian ini antara lain:

1. Melakukan pemilahan data pada data penjualan tenaga listrik pelanggan menggunakan fitur "Value filter". Data yang digunakan hanya yang memiliki keterkaitan dengan tujuan mendapatkan *cluster*.
2. *Imput Data* - Dalam penarikan data penjualan lampau yakni 2019 dan 2020 terdapat beberapa *error* pada besaran jam nyala. Pada data penjualan tenaga listrik untuk pelanggan yang baru jadi pelanggan, missal di juni tahun 2020 baru menjadi pelanggan sehingga data di 2019 dan januari-mei 2020 tidak ada data. *Imputing* data tersebut menggunakan nilai mean.

3. *Missing Value* - Rekam data dari system tidak sempurna artinya ada data penjualan energi dan jam nyala pelanggan yang tidak terekam. Menghilangkan data bila terdapat lebih dari 10 data pada kolom kwh jual tenaga listrik, rupiah pendapatan tenaga listrik, dan jam nyala pelanggan maka data tersebut dihilangkan (keseluruhan kolom yakni 108 kolom).
4. *Numeric outlier* untuk melakukan seleksi pelanggan yang terdapat penyimpangan sangat tinggi pada data jam nyala dan energi pakai.
5. Transformasi data - Keseluruhan data rekam dari penjualan tenaga listrik tidak seluruhnya digunakan sehingga penulis membuang kolom-kolom data tersebut dengan hanya menyisakan kolom id pelanggan, nama pelanggan, tarif, daya, kwh penjualan, rupiah pendapatan, dan jam nyala. Untuk data fitur yang digunakan dari 108 kolom, penulis melakukan *average* data penjualan KWH, rupiah pendapatan, dan jam nyala pelanggan menjadi per tahun rekam data tersebut.
6. *Normalizer* - pada Table II untuk mengubah data dengan skala minimal dan maksimal karna gap yang terlalu jauh, untuk memudahkan pada tahapan proses pemodelan. menggunakan fitur *normalizer* pada skala nilai data variabel rupiah pendapatan tenaga listrik, energi pakai (kWh), dan jam nyala.

TABEL II
STRUKTUR DATA PENJUALAN TENAGA LISTRIK

No.	Nama Field	Type	Keterangan
1	Nama	Character	Nama Pelanggan
2	Alamat	Character	Alamat Pelanggan
3	Tarif	Character	Jenis Pengenaan Tarif
4	Daya	Numerical	Daya Kontrak Pelanggan
5	Energi Pakai	Numerical	Pemakaian Energi
6	Rupiah Jual	Numerical	Rupiah Pendapatan
7	Jam Nyala	Numerical	Jam Durasi Pemakaian

D. Data Modeling

Dalam proses clustering data dijelaskan mengenai hasil pentuan nilai K dengan metode elbow dan proses *clustering* menggunakan metode K-Means serta uji performa *cluster* sebagai berikut :

1. Menentukan besarnya nilai k dengan metode elbow. Proses ini menggunakan dengan cara memilih nilai k yang mengalami penurunan nilai *sum of square error* (SSE) terbesar dan ditunjukkan pada pembentukan sudut dalam grafik antara nilai k pertama terhadap nilai k setelahnya. Besarnya jumlah k mempengaruhi nilai SSE. Berikut perhitungan SSE:

$$SSE = \sum_{k=1}^k \sum_{x_i \in S_k} \|X_i - C_k\|^2 \quad (1)$$

Keterangan :

k sebagai jumlah *cluster*.

X_i adalah banyaknya data.

C_k adalah jumlah *cluster* pada cluster ke k

2. Mengukur nilai *silhouette* terbaik dari pengujian. Nilai *silhouette* diukur dengan menghitung semua obyek pada *cluster*.

$$SIL(c) = SIL(k) \frac{1}{|k|} \sum_i^k SIL(c_i) \quad (2)$$

Keterangan :

$SIL(k)$ adalah nilai *silhouette* semua *cluster*.

$|k|$ adalah jumlah *cluster* k .

$SIL(c_i)$ adalah rata-rata nilai *silhouette*

TABEL III
STRUKTUR DATA SET

No	Nama Field	Type	Keterangan	Kriteria	Ket.
1	Daya	Numeric	Daya Kontrak	Fitur 1	Tunggal
2	Kwh Jan-Des 2019	Numeric	Kwh Jual Per Pelanggan	Fitur 2	Average
3	Kwh Jan-Des 2020	Numeric	Kwh Jual per Pelanggan	Fitur 3	Average
4	Kwh Jan-Des 2021	Numeric	Kwh Jual perPelanggan	Fitur 4	Average
5	RP Jan-Des 2019	Numeric	Rupiah Pendapatan Per Pelanggan	Fitur 5	Average
6	RP Jan-Des 2020	Numeric	Rupiah Pendapatan Per Pelanggan	Fitur 6	Average
7	RP Jan-Des 2021	Numeric	Rupiah Pendapatan Per pelanggan	Fitur 7	Average
8	Jam Nyala Jan-Des 2019	Numeric	Jam Pemakaian Energi Per Pelanggan	Fitur 8	Average
9	Jam Nyala Jan-Des 2020	Numeric	Jam Pemakaian Energi Per Pelanggan	Fitur 9	Average
10	Jam Nyala Jan-Des 2021	Numerical	Jam Pemakaian Energi Per Pelanggan	Fitur 10	Average

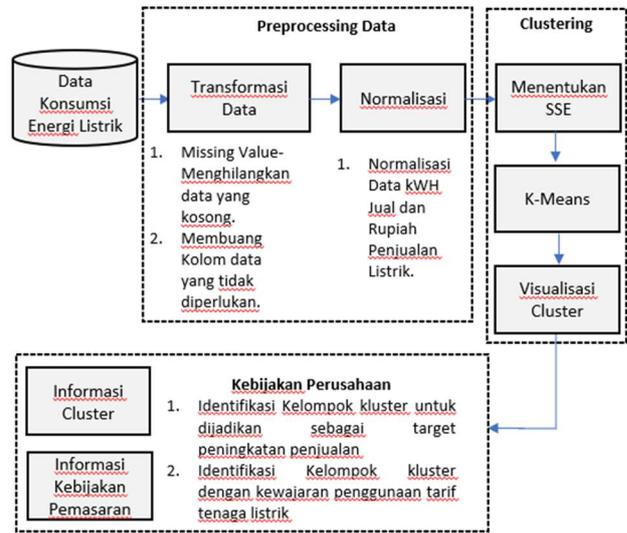
E. Deploy Clustering

Pemodelan Machine Learning dilakukan dengan menggunakan tool *KNIME Analytic*. Algoritma yang akan digunakan yakni *K-Means Clustering*. Teknik analisis data menggunakan algoritma *K-Means* yang ditunjukkan pada gambar 3, dapat mengidentifikasi pola konsumsi listrik pada tiap kelompok kluster yang terbentuk [6].

Variabel yang digunakan dalam metode *k-means* yakni daya terpasang, penjualan tenaga listrik, energi pakai (kWh), dan Jam Nyala pelanggan. Keempat variabel ini digunakan untuk mendapatkan *cluster* identifikasi kelompok karakteristik pelanggan. Tahapan yang dilakukan untuk melakukan pemodelan klaster, yakni :

1. *K-means* menggunakan *random initialization* dengan iterasi sebanyak 99 kali.
2. *Visualisasi* hasil *cluster* dari model dengan menggunakan *scatter plot*

Model kluster digunakan untuk mendapatkan informasi karakteristik pelanggan pada tiap *cluster* yang kemudian dianalisa untuk rekomendasi sasaran strategi eksekusi program pemasaran oleh perusahaan sehingga program pemasaran dapat efektif dan efisien.

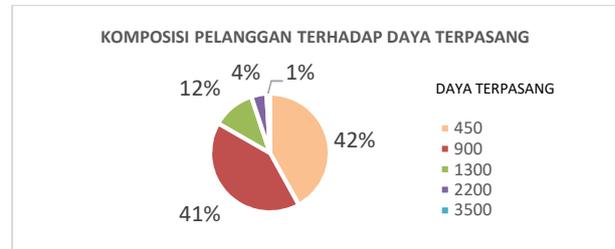


Gambar. 3. Alir Penelitian

IV. HASIL DAN PEMBAHASAN

Simulasi model pada penelitian ini dilakukan dengan menggunakan aplikasi *knime analytical*. Simulasi model *clustering* dilakukan dengan tujuan mendapatkan pengelompokan pelanggan yang akan diidentifikasi untuk dianalisa.

A. Data Penjualan Tenaga Listrik



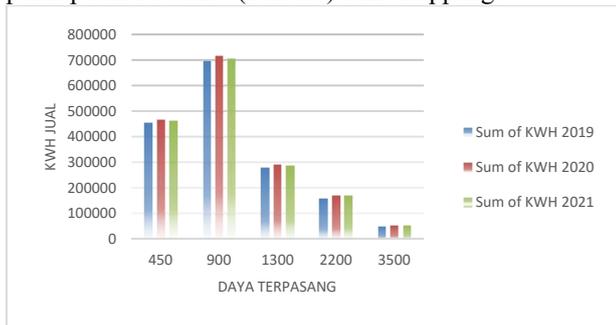
Gambar. 4. Komposisi pelanggan Terhadap Daya Terpasang

Data merupakan rekaman penjualan yang berisikan semua informasi transaksi yang tersusun sebagai bahan laporan dan analisa. Laporan data penjualan tenaga listrik merupakan laporan bulanan yang dikumpulkan selama 3 tahun. Informasi yang terdapat dalam laporan berupa data transaksi yang dihasilkan setiap bulan untuk pelanggan listrik pasca bayar..

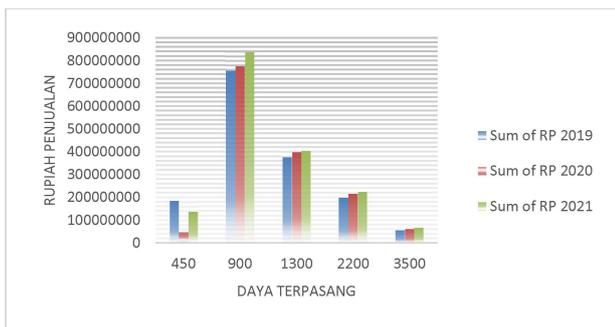
Pada Gambar 4 data pelanggan listrik PT. PLN (Persero) ULP Rappang sebanyak 13.510 pelanggan dengan sebaran daya 450 VA, 900 VA, 1300 VA, 2200 VA, dan 3500 VA. Daya terpasang pelanggan yang dipilih pada penelitian ini merupakan peminatan permintaan daya pada umumnya. Populasi terbesar yakni pada pelanggan daya terpasang 450 VA sebanyak 42% yakni 5.672 pelanggan, daya terpasang 900 VA sebanyak 41% yakni 5.583 pelanggan, daya terpasang 1300 VA sebanyak 12% yakni 1.570 pelanggan, daya terpasang 2200 VA sebanyak 4% yakni 553 pelanggan, dan daya terpasang 3500 VA sebanyak 1% yakni 129 pelanggan.

Penjualan energi listrik berdasarkan Gambar 5 dapat dilihat secara kumulatif selama 3 tahun, kWh jual tertinggi pada pelanggan daya 900 VA dengan puncak tertinggi terjadi pada tahun 2020. Seluruh pelanggan berkontribusi pada penjualan energi listrik pada PT. PLN (Persero) ULP Rappang. Namun secara karakteristik pemakaian energi listrik tiap pelanggan berbeda-beda, ada yang pasif dan ada yang tumbuh. Identifikasi menggunakan metode *clustering* digunakan untuk mengelompokkan pelanggan yang kemudian menganalisa setiap *cluster* yang terbentuk untuk dapat mengidentifikasi karakteristik pelanggan.

Rupiah penjualan energi listrik berdasarkan Gambar 6 dapat dilihat secara kumulatif selama 3 tahun, rupiah penjualan tertinggi pada pelanggan daya 900 VA dengan puncak tertinggi terjadi pada tahun 2021. Hal ini berbanding terbalik dengan pemakaian energi listrik yang terlihat pada Gambar 5. Demikian pula yang terjadi pada pelanggan daya 450 VA, secara pemakaian energi pada Gambar 5 lebih tinggi dari pada daya 1300 VA, 2200 VA, dan 3500 VA. Namun pada rupiah penjualan cenderung lebih rendah dari pada daya 1300 VA, 2200 VA, dan 3500 VA. Seluruh pelanggan berkontribusi pada rupiah penjualan energi listrik pada PT. PLN (Persero) ULP Rappang. Namun secara karakteristik rupiah pemakaian energi listrik tiap pelanggan berbeda-beda, ada yang pasif dan ada yang tumbuh. Rupiah penjualan dan kWh jual tenaga listrik merupakan komponen penentu untuk pendapatan PT. PLN (Persero) ULP Rappang.



Gambar. 5. Kumulatif kWh Jual Terhadap Daya Terpasang

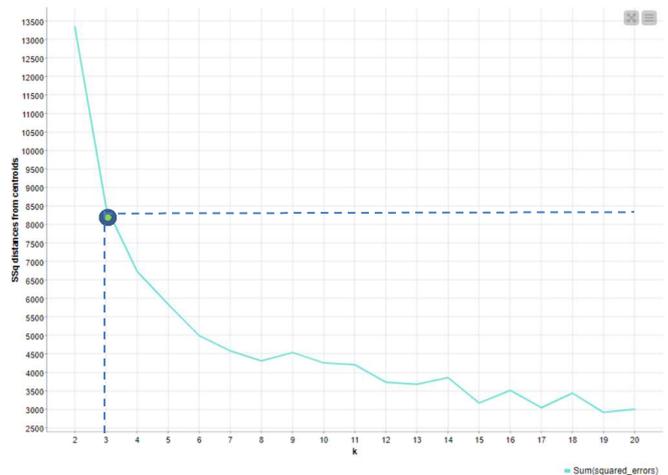


Gambar. 6. Kumulatif Rupiah Penjualan Energi Listrik Terhadap Daya Terpasang

B. Menentukan Nilai k dengan Metode Elbow

Nilai k ditentukan dengan menggunakan metode *elbow*. Visualisasi grafik perhitungan metode elbow ditampilkan dalam grafik 2D pada Gambar 7 Nilai sumbu x mewakili nilai k dalam rentang 2-20, dan sumbu y mewakili jumlah nilai *sum square error* (SSE) yang dihasilkan oleh setiap nilai k yang dihitung. Hasil dari metode elbow dapat dijelaskan dengan fakta bahwa nilai k yang dipilih adalah tingkat penurunan maksimum nilai SSE. [20]. Dapat dilihat pada plot

nilai k terhadap nilai SSE, ketika nilai k meningkat, nilai SSE menurun. Berdasarkan Gambar 7, nilai pada k = 3.



Gambar. 7. Metode Elbow

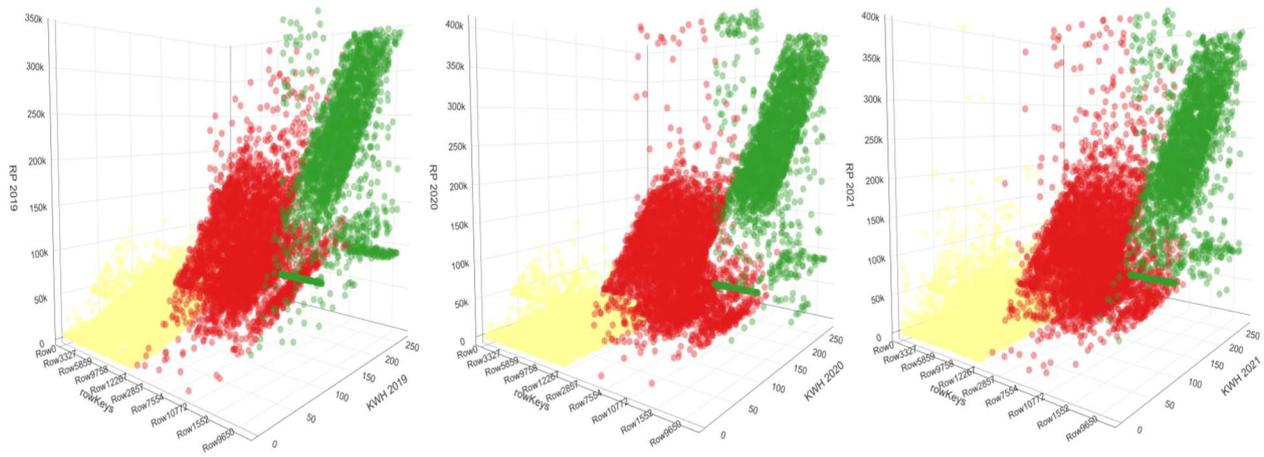
Pada Tabel III terlihat bahwa K=3 memiliki penurunan nilai rasio SSE terbesar ditunjukkan pada Tabel IV, dimana nilai rasio penurunan terbesar berada pada nilai k = 3 sehingga nilai SSE yang didapatkan adalah 8488,13.

TABEL IV
RASIO PENURUNAN SSE

No.	k	Sum(squared_errors)	Rasio
1	2	13342,09075	
2	3	8488,136243	1,571852
3	4	6733,291745	1,260622
4	5	5842,355291	1,152496
5	6	4989,376785	1,170959
6	7	4580,754754	1,089204
7	8	4309,848239	1,062858
8	9	4535,666407	0,950213
9	10	4257,163343	1,06542
10	11	4206,683278	1,012
11	12	3732,940701	1,126909
12	13	3676,123219	1,015456
13	14	3856,752861	0,953165
14	15	3170,548443	1,216431
15	16	3512,260049	0,902709
16	17	3043,037022	1,154196
17	18	3437,161143	0,885334
18	19	2915,936555	1,17875
19	20	3001,473853	0,971502

C. Hasil Cluster Pelanggan dengan K-means

Hasil dari proses *clustering* menghasilkan empat kelompok dari 13.510 pelanggan. Metode *clustering k-means* menghitung jarak semua data dari titik-titik dalam *cluster*. Data hasil uji berada pada kelompok *cluster* yang memiliki jarak terkecil terhadap titik *centroid*, sehingga membentuk *cluster*. Setiap *cluster* memiliki karakteristik yang berbeda-beda [21]. Keempat variabel uji memberikan informasi karakteristik pemakaian energi listrik dan potensi pendapatan penjualan tenaga listrik. Pada Table V dari hasil *clustering* populasi pelanggan paling banyak pada *cluster_1* yakni 6.047 pelanggan dengan *mean silhouette* 0,668. Selanjutnya hasil *cluster_3* sedikit lebih kecil dari pada *cluster_1* yakni sebanyak 4.918 pelanggan dengan *mean silhouette* 0,261 dan populasi paling kecil berada pada *cluster_2* sebanyak 2.545 dengan *mean silhouette* 0,216.



Gambar. 8. 3D Scatter Plot Hasil Clustering Populasi Pelanggan Terhadap KWH Jual dan Rupiah Pendapatan Tahun 2019-2021

TABEL V
HASIL CLUSTERING

No.	Cluster	Pelanggan	Mean Silhouette
1	cluster_1	6.047	0,668
2	cluster_2	2.545	0,261
3	cluster_3	4.918	0,216
	Grand Total	13.510	0,294

D. Cluster Analisis

Pada *cluster_1* dapat dilihat pada Gambar 8 sebaran warna hijau menunjukkan sebaran populasi pelanggan terhadap kwh pendapatan tahun 2019 serta terhadap rupiah pendapatan tahun 2019. Representasi hasil *cluster_1* sebaran pelanggan terhadap terhadap kwh penjualan menunjukkan rentang yang cenderung sedang dan rentang terhadap rupiah pendapatan rentangnya lebih kecil terhadap *cluster_2* dan *cluster_3* yang artinya memiliki populasi pelanggan yang cukup banyak namun kontribusi rupiah penjualannya rendah.

cluster_2 dilihat pada Gambar 8 sebaran warna merah menunjukkan sebaran populasi pelanggan terhadap kwh pendapatan tahun 2020 serta terhadap rupiah pendapatan tahun 2020. Representasi hasil *cluster_2* sebaran pelanggan terhadap terhadap kwh penjualan menunjukkan rentang yang cenderung lebih kecil dan rentang terhadap rupiah pendapatan cukup lebar terhadap *cluster_1* dan *cluster_3* yang artinya memiliki populasi pelanggan yang sedikit dan kontribusi rupiah penjualannya lebih besar dari pada *cluster_1*.

cluster_3 dapat dilihat pada Gambar 8 sebaran warna merah menunjukkan sebaran populasi pelanggan terhadap kwh pendapatan tahun 2021 serta terhadap rupiah pendapatan tahun 2021. Representasi hasil *cluster_3* sebaran pelanggan terhadap terhadap kwh penjualan menunjukkan rentang yang cenderung paling besar dan rentang terhadap rupiah pendapatan paling lebar terhadap *cluster_1* dan *cluster_2* yang artinya memiliki populasi pelanggan berada diantara *cluster_1* dan *cluster_2* namun memiliki kwh jual dan pendapatan yang paling besar.

Berikut hasil Analisa yang disimpulkan karakteristik pelanggan tiap *cluster* :

1. Pelanggan *cluster_1* memiliki paling banyak populasi pelanggan yakni 6.047 pelanggan dengan memberikan kontribusi kwh jual dan jam nyala pakai pada kategori

tinggi namun kontribusi terendah untuk rupiah pendapatan, yakni 18,5%.

2. Pelanggan *cluster_2* memiliki 2.545 pelanggan dengan memberikan kontribusi kwh jual dan jam nyala pakai pada kategori rendah namun memberikan kontribusi sedang untuk rupiah pendapatan, yakni 34,9%.
3. Pelanggan *cluster_3* memiliki 4.917 pelanggan dengan memberikan kontribusi paling besar untuk kategori rupiah pendapatan, yakni 46,63%, namun secara kwh jual dan jam nyala pakai berkontribusi sedang.

TABEL VI
KONTRIBUSI TIAP CLUSTER

No.	Kriteria	Custer 1	Cluster 2	Cluster 3
1	Pelanggan	6.047	2.545	4.917
2	Rp. Penjualan	18,5%	34,9%	46,6%
3	kWh Jual	42,9%	21,0%	36,1%
4	JN Pakai	42,9%	21,0%	37,2%

E. Implikasi Bisnis

Berdasarkan pada Tabel VI diatas maka sebagai manajemen perusahaan, dapat menerapkan strategi pelaksanaan kegiatan pemasaran dan pengawasan untuk menjaga revenue dan meningkatkan pendapatan, berdasarkan diskusi dengan bidang niaga terkait hasil *clustering*, sebagai berikut :

1. Kinerja penjualan tenaga listrik diukur dari rupiah pendapatan dengan inisiasi dari kWh jual dan jam nyala pakai pelanggan. Hasil *clustering* pada Tabel V produktifitas pelanggan didapatkan pada kelompok pelanggan *cluster_3*.
2. Pelanggan *Cluster_1* perlu dilakukannya pemeriksaan penggunaan energi listrik pada persil pelanggan untuk memastikannya tidak ada anomali pada pengukuran meter. Kemudian melakukan pengecekan kesesuaian penggunaan tarif dipelanggan dikarenakan secara kWh jual dan jam nyala pakai tinggi.
3. Pelanggan *Cluster 2* dapat dilakukan sosialisasi terkait kenyamanan penggunaan peralatan listrik di era saat ini, menawarkan promo bundling tambah daya peralatan elektronik secara agresif. Hal ini agar terjadi peningkatan pemakaian energi listrik yang berdampak pada peningkatan revenue.

4. Pelanggan *Cluster_3* diberikan layanan-layanan ekstra dalam rangka membangun *customer relationship*. Berikan kunjungan khusus untuk mengetahui testimoni dan harapan yang diinginkan. Tawarkan informasi terkait peralatan listrik yang terbaru terkait kenyamanan dan kemudahan beraktifitas dengan peralatan elektronik modern. Loyalitas pada *cluster_3* sangat perlu dijaga terkait kontribusinya yang mendominasi, serta berikan informasi tagihan listrik agar menjaga revenue yang didapat tidak menjadi piutang.

V. KESIMPULAN

Berdasarkan hasil *clustering* menggunakan metode k-means kemudian disimpulkan sebagai berikut :

1. Penggunaan metode *clustering* dalam *data mining* sangat membantu perusahaan dalam mengidentifikasi karakteristik pelanggan. Data pelanggan dikelompokkan kedalam masing-masing *cluster*. Teknik *clustering* k-means yang diterapkan di atas menghasilkan tiga *cluster* yang berbeda berdasarkan pemilihan beberapa variabel antara lain daya terpasang, penjualan kWh 2019-2021, pendapatan rupiah 2019-2021, dan jam nyala 2019-2021. Dengan pembagian *cluster* ini, perusahaan mendapatkan informasi untuk membuat strategi bisnis yang tepat untuk dijalankan pada setiap informasi karakteristik masing-masing *cluster*.
2. Hasil *clustering* yang telah dilakukan memberikan informasi bahwa kelompok pelanggan yang berkontribusi besar dalam kwh jual dan rupiah pendapatan yakni pelanggan *cluster_3* sebesar 46,63%. Di susul selanjutnya pelanggan *cluster_2* kontribusi 34,86% dan paling rendah pelanggan *cluster_1* dengan kontribusi 18,5%.
3. Visualisasi *clustering* yang terbentuk memberikan gambaran persebaran data pada tiap *cluster*. Hal ini memungkinkan perusahaan untuk dengan mudah memetakan hasil pengelompokan yang ditampilkan. Perusahaan mendapatkan distribusi visual jumlah, persentase, dan karakteristik distribusi pelanggan di setiap area *cluster*.

DAFTAR PUSTAKA

- [1] A. Yusuf, T. Abduh and H. Abubakar, "Pengaruh Bauran Pemasaran Terhadap Peningkatan Penjualan Tenaga Listrik Pada PT PLN (PERSERO) Unit Pelaksana Pelayanan Pelanggan Pinrang," *Indonesian Journal of Business and Management*, vol. 3, no. 2460-3767, pp. 115-120, 2021.
- [2] Charmicael, Gross, Hanna, Rhodes and Green, "The Demand Response Technology Cluster: Accelerating UK residential consumer engagement with time-of-use tariffs, electric vehicles and smart meters via digital comparison tools," *Science Direct-Renewable and Sustainable Energy Reviews*, vol. 139, p. 110701, 2021.
- [3] S. Ramos, J. Soare, S. S. Cembranel, I. Tavares, Z. Foroozandeh, Z. Vale and R. Fernandes, "Data mining techniques for electricity customer characterization," *ScienceDirect-Intelligent Systems*, vol. 186, pp. 475-488, 2021.
- [4] O. Motlagh, A. Berry and L. O'neil, "Clustering of residential electricity customers using load time series," *ScienceDirect- Applied Energy*, vol. 237, pp. 11-24, 2019.
- [5] A. Rajabi, M. Eskandari, M. Jabbari, L. Li, J. Zhang and P. Siano, "A comparative study of clustering techniques for electrical load pattern segmentation," *Science Direct-Renewable and Sustainable Energy Reviews*, vol. 120, p. 109628, 2020.
- [6] S. Yilmaz, J. Chambers and M. Patel, "Comparison of clustering approaches for domestic electricity load," *Science Direct-Energy*, vol. 180, pp. 665-677, 2019.
- [7] D. Jaiswal, V. Kaushal, P. K. Singh and A. Biswas, "Green market segmentation and consumer profiling: a cluster approach to an emerging consumer market," *Emerald Insight*, vol. 28, pp. pp. 792-812, 2021.
- [8] J. Alhilman, M. R. M. Wiyono and Marina, "Predicting And Clustering Customer to Improve Customer Loyalty and Company Profit," *IEEE*, 2014.
- [9] E. A. Darko, F. Donkor, S. Adarkwah and E. Kyei, "Management of accounts receivables in utility companies: A focus on Electricity Company of Ghana," *International Journal of Academic Research in Business and Social Sciences*, pp. 486-518, 2016.
- [10] D. Drosos, G. L. Kyriakopoulos, G. Arabatzis and N. Tsotsolas, "Evaluating Customer Satisfaction in Energy Markets Using a Multicriteria Method: The Case of Electricity Market in Greece," *MDPI-Sustainability*, vol. 12, p. 3862, 2020.
- [11] E. Kapassa, M. Touloupou and M. Themistocleous, "Local Electricity and Flexibility Markets: SWOT Analysis and Recommendations," in *IEEE-International Conference on Smart and Sustainable Technologies*, Bol and Split, Croatia, 2021.
- [12] "The effect of digital marketing capabilities on business performance enhancement: Mediating the role of customer relationship management (CRM)," *Growing Science-International Journal of Data and Network Science*, vol. 6, pp. 295-304, 2022.
- [13] A. Roihan, P. A. Sunarya and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review Paper," *Indonesian Journal on Computer and Information Technology*, vol. 5, pp. 75-82, 2020.
- [14] R. Afthoni, M. Hamdhani, Ardianto, A. F. K and H. Patria, "Pemanfaatan Algoritma Machine Learning untuk Segmentasi Pelanggan Berbasis Data Konsumsi Listrik di PT PLNXYZ," in *Seminar Nasional Teknik dan Manajemen Industri dan Call for Paper(SENTEKMI 2021)*, Bandung, 2021.
- [15] Q. Zhao, H. Li, X. Wang, T. Pu and J. Wang, "Analysis of users' electricity consumption behavior based on ensemble clustering," *KeAi-Global Energy Interconnection*, vol. 2, pp. 479-488, 2019.
- [16] R. S. Rochman, W. E. Y. Retnani and O. Juwita, "Rancang Bangun Aplikasi Analisis Indeks Kepuasan Pelanggan pada PT. PLN (Persero) Area Jember dengan Menggunakan Pendekatan Metode Servqual dan K-Means Clustering," *Sainstek 2020*, vol. 3, no. 2339-0069, pp. 96-101, 2020.
- [17] L. Czetany, V. Vamos, M. Horvath, Z. Szalay, A. Mota, Z. Deme and T. Csoknyai, "Development of electricity consumption profiles of residential buildings based on smart meter data clustering," *Science Direct - Energy & Buildings*, vol. 252, p. 111376, 2021.
- [18] C. Schoer, F. Kruse and J. M. Gomez, "A Systematic Literatur Review on Applying CRISP-DM Process Model," *Science Direct*, vol. 181, pp. 526-534, 2021.
- [19] S. Ramosa, J. Soares, S. Cembranel, F. Inês Tavares, Z. Vale and R. Piara, "Data mining techniques for electricity customer characterization," *Science Direct*, vol. 186, pp. 475-488, 2021.
- [20] R. Shanker, R. Singh and M. B., "Segmentation of Tumor and Edema Based on K-mean clustering and hierarchical centroid shape descriptor," *IEEE*, pp. 1105-1109, 2017.
- [21] Y. Asri, D. Kuswardani, E. Yosrita and F. H. Wullur, "Clusterization of customer energy usage to detect power shrinkage in an effort to increase the efficiency of electric energy consumption," *Indonesian Journal of Electrical Engineering and Computer*, vol. 22, no. 2502-4752, pp. p. 10-17, 2021.
- [22] R. Menteri ESDM and PLN, *Rencana Usaha Penyediaan Tenaga Listrik (RUPTL) PT PLN (PERSERO) 2021-2030*, JAKARTA: ESDM dan PT. PLN (Persero), 2021.
- [23] S. Alasadi and W. Bhaya, "Review Of Data Preprocessing Techniques in Data Mining," *ACADEMIA*, vol. 12, no. 1816-949x, pp. 4102-4107, 2017.
- [24] Yasirli, Amanda, Fatmawati, Setiani and Rani, "Analysis Clustering of Electricity Usage Profile Using K-Means Algorithm," *IOP Science*, vol. 105, 2016.

- [25] K. Li, Z. Ma, D. Robinson, W. Lin and Z. Li, "A data-driven strategy to forecast next-day electricity usage and peak electricity demand of a building portfolio using cluster analysis, Cubist regression models and Particle Swarm Optimization," *ScienceDirect-Journal of Cleaner Production*, vol. 273, pp. 115-123, 2020.
- [26] K. Zhou, S. Yang and Z. Shao, "Household monthly electricity consumption pattern mining: A fuzzy clustering-based model and a case study," *Science Direct-Journal Of Cleaner Production*, vol. 141, no. 2016, pp. 900-908, 2016.
- [27] A. R. Arsyah and E. Listiani, "Efektifitas Personal selling Pada komunikasi pemasaran keliling Listrik Prabayar (Listrik Pintar) PT. PLN (Persero) Di Wilayah Cijaura Bandung," *e-Proceeding Of Management*, vol. 2, no. 2355-9357, p. 2176, 2015.
- [28] S. Balasubramanian and P. Balachandra, "Characterising electricity demand through load curve clustering: A," *Science Direct-Computers and Chemical Engineering*, vol. 150, p. 107, 2021.
- [29] C. Liu, X. Wang, Y. Huang, Y. Liu, R. Li, Y. Li and J. Liu, "A Moving Shape-based Robust Fuzzy K-modes Clustering Algorithm for Electricity Profiles," *ScienceDirect-Electric Power Systems Research*, vol. 187, p. 106425, 2020.