

Klasifikasi Konsumsi Energi Listrik Dan Penggunaan Listrik Rumah Tangga Menggunakan Metode *Naïve Bayes*

Dea Destiani¹, Irma Santikarama²,

Fakultas Sains dan Informatika/Program Studi Sistem Informasi

Universitas Jenderal Achmad Yani

Cimahi, Indonesia

e-mail: ¹dea.destiani@lecture.unjani.ac.id, ²irma.santikarama@lecture.unjani.ac.id,

Correspondence : e-mail: ¹dea.destiani@lecturer.unjani.ac.id,

Diajukan: 15 Agustus 2024; Direvisi: 22 Agustus 2024; Diterima: 23 Agustus 2024

Abstrak

Energi listrik memiliki peranan penting bagi setiap kalangan masyarakat. Pentingnya peran listrik, tentu saja, berdampak pada permintaan listrik yang semakin besar, setiap peralatan elektronik yang berada di perumahan tidak dapat berfungsi jika tidak di hantarkan arus listrik. Atribut yang berpengaruh terhadap tingginya pemakaian listrik yaitu Global aktif power, global reaktif power, voltage dan global intensity. Atribut tersebut akan di update setiap satu menit sekali sehingga menghasilkan data set sebanyak 10.000 data. Metode *Naïve Bayes* Merupakan salah satu metode pemodelan yang digunakan dalam klasifikasi yang memiliki nilai accuracy, recall, precision yang tinggi. Penelitian ini pada pengujianya menggunakan k-fold validation. Dimana data set di bagi kedalam 10 bagian. 10 % data digunakan sebagai data uji dan 90 % digunakan untuk data latih . Hasil pengujian akan di evaluasi menggunakan confusion matrix sehingga dapat diperoleh nilai accuracy, recall dan precisionnya. Dan penelitian ini mengkombinasikan jumlah atribut. Pada penelitian ini nilai accuracy terbaik yaitu nilai 90 % Tujuan dari penelitian ini yaitu untuk mengetahui perbandingan accuracy, recall dan precision yang dihasilkan *Naïve Bayes* dalam menyelesaikan masalah prediksi konsumsi energi listrik di perumahan

Kata kunci: energy Listrik, global aktif power, global reaktif power, voltage, global intensity.

Abstract

Electrical energy plays an important role for all levels of society. The significance of electricity naturally impacts the increasing demand for it, as no electronic device in residential areas can function without the conduction of electric current. The attributes that influence high electricity consumption include global active power, global reactive power, voltage, and global intensity. These attributes are updated every minute, resulting in a dataset of 10,000 data points. The *Naïve Bayes* method is one of the modeling methods used in classification, known for its high accuracy, recall, and precision values. This study uses k-fold validation for testing, where the dataset is divided into 10 parts. Ten percent of the data is used as test data, while 90 percent is used for training data. The test results will be evaluated using a confusion matrix to obtain accuracy, recall, and precision values. This study also combines the number of attributes. The best accuracy obtained in this study is 90%. The purpose of this research is to compare the accuracy, recall, and precision produced by *Naïve Bayes* in solving the problem of predicting residential energy consumption

Keywords: electrical energy, global active power, global reactive power, voltage, global intensity.

1. Pendahuluan

Era modern ini, kebutuhan akan energi listrik terus meningkat seiring dengan pertumbuhan populasi dan perkembangan teknologi. Penggunaan energi listrik yang efisien dan berkelanjutan menjadi isu yang sangat penting, terutama di sektor rumah tangga, yang merupakan salah satu konsumen utama energi listrik. Klasifikasi dan analisis konsumsi energi listrik pada tingkat rumah tangga dapat memberikan wawasan penting untuk pengelolaan energi yang lebih baik dan untuk merancang kebijakan penghematan energi yang lebih efektif. [1]

Konsumsi energi listrik rumah tangga dapat bervariasi secara signifikan berdasarkan berbagai faktor, termasuk ukuran rumah, jumlah anggota keluarga, perangkat listrik yang digunakan, dan kebiasaan konsumsi. [2] Dengan data yang bervariasi dan kompleks ini, diperlukan metode yang efisien untuk

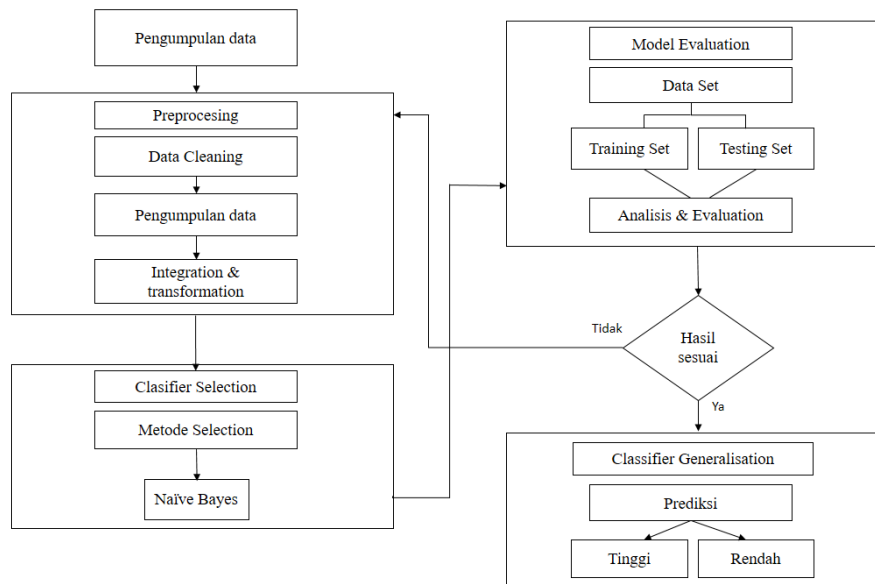
mengidentifikasi pola dan klasifikasi konsumsi energi. [3] Salah satu metode yang dapat digunakan untuk tujuan ini adalah *Naïve Bayes*.

Penelitian sebelumnya dengan terkait peningkatan prediksi kebutuhan permintaan energi listrik menggunakan Model *Neural Network* (NN) dan *Support Vector Machine* (SVM) digunakan untuk memprediksi kebutuhan permintaan listrik. Namun, kedua model ini memiliki kelemahan, yaitu kesulitan dalam menentukan nilai parameter yang digunakan, sehingga berdampak pada tingkat akurasi yang dihasilkan. [3]

Penelitian ini bertujuan untuk menerapkan metode *naïve bayes* dalam klasifikasi konsumsi energi listrik dan penggunaan listrik rumah tangga. Penerapan diharapkan mampu untuk memprediksi besarnya penggunaan listrik tiap rumah tangga agar penggunaan listrik lebih hemat. Penggunaan energi listrik bisa mengalami kenaikan ataupun penurunan, Maka dibutuhkan prediksi sehingga data penggunaan listrik lebih efektif dalam penggunaannya dan membantu pemerintah dalam menyediakan energi daya tampung energi listrik. Dalam hal ini dibutuhkan perbandingan metode menggunakan *Naïve Bayes* untuk prediksi konsumsi energi listrik.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode klasifikasi untuk menganalisis pola konsumsi energi listrik rumah tangga. Fokus utama adalah penerapan metode *Naïve Bayes* untuk mengklasifikasikan data konsumsi energi berdasarkan berbagai fitur yang relevan. [4]



Gambar 1 Kerangka Penelitian

2.1. Atribut yang digunakan

1. *Global Aktif Power* (Daya Aktif Global):

Daya aktif adalah daya yang digunakan oleh perangkat listrik untuk melakukan pekerjaan nyata seperti menyalakan lampu atau menggerakkan motor. Daya ini diukur dalam satuan watt (W) atau kilowatt (kW). Ini adalah daya yang benar-benar bermanfaat dalam penggunaan listrik sehari-hari.

2. *Global Re-Aktif Power* (Daya Reaktif Global):

Daya reaktif adalah daya yang diperlukan untuk membentuk medan magnet dalam peralatan listrik seperti motor dan transformator. Daya ini tidak menghasilkan pekerjaan nyata, tetapi tetap penting untuk pengoperasian peralatan tertentu. Ini diukur dalam volt-ampere reaktif (VAR).

3. *Voltage* (Tegangan):

Tegangan adalah perbedaan potensial listrik antara dua titik dalam rangkaian listrik. Tegangan inilah yang mendorong arus listrik mengalir melalui rangkaian. Satuan pengukuran tegangan adalah volt (V).

4. *Intensity* (Arus Listrik):

Arus listrik adalah aliran muatan listrik melalui konduktor dalam rangkaian listrik. Arus ini diukur dalam satuan ampere (A). Semakin besar arus yang mengalir, semakin besar pula jumlah muatan listrik yang bergerak melalui suatu titik dalam waktu tertentu.

2.2. Pengumpulan Data

Pengumpulan data merupakan faktor penting demi keberhasilan penelitian. Hal ini berkaitan dengan bagaimana cara mengumpulkan data, siapa sumbernya, dan apa alat yang digunakan [5]. Pada penelitian ini sumber data didapatkan dari hasil observasi disuatu rumah dengan bantuan PLN Majalengka.

2.3. Data Cleaning

Data cleaning merupakan adalah proses menghapus data yang tidak digunakan. Penghapusan data dengan menghapus nilai yang hilang, *duplikasi data*, memulihkan dan memperbaiki kesalahan data. [6] Proses pembersihan data dilakukan secara manual dengan bantuan microsoft excel. Data biaya, sub metering 1, submetering 2 dan sub metering 3 dihapus dikarenakan biaya tidak termasuk kedalam katagori data yang diperlukan dan sebagian data ada yang kosong.

2.4. Data Selection

Data Selection merupakan proses pemilihan data dari kumpulan data operasional yang ada sebelum memasukkan data atau penggalian informasi. [7] Data yang dipilih akan digunakan untuk proses penambangan data yang disimpan dalam data terpisah dari basis data operasional Data selection yang dipilih yaitu data atribut tanggal, waktu, *global active power* (kWh), *global reactive power* (kWh), *voltage* (V), *global intensity* (A) sebanyak 10.000 data.

2.5. Data Transformation

Transformasi data adalah proses mengubah data dari satu format ke format lain sesuai kebutuhan yang bertujuan untuk mempermudah proses pembacaan data pada program maupun tool yang digunakan. Berikut adalah tampilan data hasil dari *transformation*. [1]

Tabel 1 Data Transformation

Date	Time	Global active_ power	Global Reactive power	Voltage	Global intensity	Prediksi
17/12/2022	12:00:00	1.356	0.058	240.89	5.6	tinggi
17/12/2022	12:01:00	1.022	0.007	240.51	5.8	rendah
17/12/2022	12:02:00	1.014	0.007	240.68	5.8	rendah
17/12/2022	12:03:00	1.372	0.084	239.68	5.6	tinggi
17/12/2022	12:04:00	1.378	0.086	240.4	5.6	tinggi
17/12/2022	12:05:00	1.378	0.086	240.08	5.6	tinggi
17/12/2022	12:06:00	1.376	0.086	239.63	5.6	tinggi
17/12/2022	12:07:00	1.368	0.084	239.11	5.6	tinggi
17/12/2022	12:08:00	1.376	0.086	239.9	5.6	tinggi
17/12/2022	12:09:00	1.378	0.086	240.22	5.6	tinggi
17/12/2022	12:10:00	1.378	0.088	240.47	5.6	tinggi

Pada penelitian ini akan dilakukan pemodelan menggunakan metode *Naïve Bayes* untuk pengolahan data konsumsi energi listrik. Berikut perhitungan Metode *Naïve Bayes*

2.6. K-Fold Cross Validation

Penelitian ini menggunakan *K-fold cross validation* dengan menggunakan 10.000 data yang akan digunakan sebagai data test dan dan train. [6] Dimana k yang digunakan pada penelitian ini yaitu K-10. Artinya proses pelatihan dan pengujian sebanyak 10 kali. Sehingga dari 10.000 data yang digunakan dibagi menjadi 10, maka setiap 1x testing menggunakan 1000 data. [7]Pembagian data untuk dijadikan *data training dan testing* dilakukan sebanyak sepuluh kali pengujian. Selain jumlah datanya, penelitian ini menggunakan 10 kali pengujian dengan melibatkan 10 kombinasi atribut sehingga dapat terlihat atribut mana yang akan menghasilkan akurasi terbaik dalam klasifikasi. Pengujian menggunakan *k-fold cross validation*. [4]

2.7. Evaluasi Pengujian

Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada. Pada penelitian ini untuk pengujiannya menggunakan *confusion matrix* sehingga dapat menghitung *accuracy, precision dan recall*.

2.8. Confusion Matrix

Confusion Matrix mengenai hasil klasifikasi informasi perbandingan dilakukan oleh sistem dengan hasil klasifikasi yang berisi tabel yang berisi informasi tentang klasifikasi aktual dan prediktif untuk

setiap algoritma klasifikasi yang disediakan. Hasil dari perhitungan data tersebut menyatakan tingkat *accuracy*, *precision* dan *recall* dari *Naïve Bayes*, [8]

Tabel 2 *Confusion Matrix*

		Actual class	
		Positif	Negatif
Predicted Class	Positif	TP = 2 (True Positive)	FP= 1 (False Positive)
	Negatif	FN= 1 (False Negative)	TN =2 (True Negative)

Setelah didapat *true positives*, *false positives*, *true negatives* dan *false negatives*, selanjutnya hitung nilai *accuracy*, *precision* dan *recall*

Accuracy adalah nilai yang mengidentifikasi persentase dari jumlah data dalam data uji yang disetujui dengan benar oleh pengklasifikasi.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \text{ [4]}$$

Precision adalah nilai kepastian dari persentase data yang berlabel positif benar pada saat nilai sebenarnya.

$$Precision = \frac{TP}{TP+FP} \times 100\% \text{ [9]}$$

Recall adalah kelengkapan nilai-nilai data positif diberi label sebagai positif.

$$Recall = \frac{TP}{TP+FN} * 100\% \text{ [10]}$$

3. Hasil dan Pembahasan

3.1. Analisa Data

Analisa data konsumsi energi listrik digunakan pengguna untuk mengetahui perubahan konsumsi listrik di perumahan. Data yang dipergunakan pada penelitian ini merupakan data konsumsi energi listrik yang di *update* setiap menit. Jumlah data yang dipergunakan sebanyak 10.000 data. Data ini terdapat 4 atribut yang digunakan dalam proses klasifikasi konsumsi energi listrik. Adapun keterangan dari data

Tabel 3 Atribut Data

No	Atribut	Type Data
1	<i>Global_aktif_power</i>	Bilangan real
2	<i>Global_re-aktif_power</i>	Bilangan real
3	<i>Voltage</i>	Bilangan real
4	<i>Global_intensity</i>	Bilangan real
5	<i>Prediksi</i>	Label

Berdasarkan data dan atribut yang diperoleh, maka data tersebut dipisahkan menjadi dua jenis data yaitu :

- a. Data Pelatihan atau data pelatihan adalah data yang sudah ada sebelumnya berdasarkan fakta yang telah terjadi (kelas yang telah ditentukan).
- b. Data Pengujian adalah data yang berkelas / berlabel digunakan untuk perhitungan model klasifikasi yang dibuat.

3.2. Klasifikasi

Pada penelitian ini klasifikasi dilakukan dengan menggunakan metode *Naïve Bayes*.. Klasifikasi dilakukan untuk mengetahui metode yang memiliki nilai akurasi terbaik untuk melakukan klasifikasi penggunaan energi listrik.

Hasil uji pada penelitian ini dilakukan menggunakan Metode *Naïve Bayes* yang bertujuan untuk mendapatkan nilai akurasi, recall dan precision Pengujian data tersebut diolah menggunakan *tools* R Studio. Hasil pengujian menggunakan Metode *Naïve Bayes*.

3.3. Perhitungan Confusion matrix

Confusion Matrix mengenai hasil klasifikasi informasi perbandingan dilakukan oleh sistem dengan hasil klasifikasi yang berisi tabel yang berisi informasi tentang klasifikasi aktual dan prediktif untuk setiap algoritma klasifikasi yang disediakan. Misalnya, *confusion matrix* untuk model klasifikasi digunakan dalam set data dengan 1000 entri. Klasifikasi 190 entri positif dengan benar, dan 700 entri negatif dengan benar. Namun, ada data yang mengklasifikasikan 80 entri positif sebagai negatif, dan 30 entri negatif sebagai positif.

Tabel 4 Confusion Matrix

		Predicted Class	
		Positif	Negatif
Actual class	Positif	TP = 190 (True Positive)	FN=30 (False Negative)
	Negatif	FP=80 (False Positive)	TN =700 (True Negative)

Setelah didapat *true positives*, *false positives*, *true negatives* dan *false negatives*, selanjutnya hitung nilai *accuracy*, *precision* dan *recall*.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

$$Accuracy = \frac{190 + 700}{190 + 700 + 80 + 30} \times 100\%$$

$$Accuracy = \frac{890}{1000} \times 100\% = 0.89\%$$

Maka nilai *Accuracy* yang dihasilkan bernilai 0.89%

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

$$Precision = \frac{190}{190 + 80} \times 100\%$$

$$Precision = \frac{190}{270} \times 100\% = 0.70\%$$

Maka nilai *precision* yang dihasilkan bernilai 0.70%

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

$$Recall = \frac{190}{190 + 30} \times 100\%$$

$$Recall = \frac{19}{22} \times 100\% = 0.90\%$$

Maka nilai *recall* yang dihasilkan bernilai 0.90%

Hasil dari perhitungan data tersebut menyatakan tingkat *accuracy*, *precision* dan *recall* dari algoritma *Naïve Bayes*

Tabel 5 Hasil Pengujian

Pengujian	Atribut yang digunakan	Accuracy	Recall	Precision
1	<i>Global_aktif_power, global_re-aktif_power, voltage, global_intensity</i>	62,34%	59,21%	86.80%
2	<i>Global_aktif_power</i>	46.10%	48.29%	60.34%

3	<i>Global_re-aktif_power</i>	90.70%	89.62%	95.70%
4	<i>Voltage</i>	50.40 %	38.85 %	71.50%
5	<i>Global_intensity</i>	64.00%	84.82%	67.37%
6	<i>Global_aktif_power, global_re-aktif_power</i>	84.20%	83.43%	91.35%
7	<i>Voltage, global_intensity</i>	55.20%	52.32%	70.71%
8	<i>Global_aktif_power, global_intensity</i>	49.40%	57.89%	61.51%
9	<i>Global_aktif_power, voltage</i>	41.00%	29.87%	58.48%
10	<i>Global_re-aktif_power, voltage</i>	85.38 %	80.34 %	96.46 %
11	<i>Global_re-aktif_power, global_intensity</i>	90.69 %	90.25 %	95.10 %
12	<i>Global_re-aktif_power, voltage, global_intensity</i>	85.98 %	83.12 %	94.54 %
13	<i>Global_aktif_power, voltage, global_intensity</i>	47.14 %	44.42 %	62.93 %
14	<i>Global_aktif_power, global_re-aktif_power, global_intensity</i>	81.38 %	81.57 %	88.72 %
15	<i>Global_aktif_power, global_re-aktif_power, voltage</i>	77.80 %	72.44 %	91.40 %

Analisis hasil dari penelitian ini merupakan rangkuman hasil percobaan dari setiap metode dan mengetahui atribut mana yang berpengaruh terhadap tingginya accuracy, recall dan precision menggunakan Metode *Naïve Bayes*, untuk mengetahui klasifikasi konsumsi energi listrik di perumahan. Pengujian dilakukan 15 kali pengujian dengan menggunakan metode *Naïve Bayes*, menggunakan kombinasi atribut yang berbeda,

Berdasarkan pengujian yang telah dilakukan, nilai accuracy tertinggi yaitu accuracy yang pada pengujian ke 3, pengujian ke 3 dengan nilai accuracy sebesar 90,70 %. dimana pada pengujian ke 2 menggunakan atribut *global_re-aktif_power*, Sedangkan accuracy terendah yaitu pada pengujian ke 9 dengan nilai accuracy sebesar 41.00%, dengan kombinasi atribut *global_aktif_power, voltage*.

Berdasarkan pengujian yang telah dilakukan, nilai Recall tertinggi yaitu Recall yang pada pengujian ke 11, pengujian ke 11 dengan nilai recall sebesar 90,50%. dimana pada pengujian ke 11 menggunakan atribut *global_re-aktif_power, global_intensity*, Sedangkan recall terendah yaitu pada pengujian ke 9 dengan nilai recall sebesar 29,87%, dengan kombinasi atribut *global_aktif_power, voltage*.

Berdasarkan pengujian yang telah dilakukan, nilai precision tertinggi yaitu precision yang pada pengujian ke 10, pengujian ke 10 dengan nilai precision sebesar 96,46 %. dimana pada pengujian ke 10 menggunakan atribut *global_re-aktif_power, voltage* Sedangkan precision terendah yaitu pada pengujian ke 9 dengan nilai precision sebesar 58.48%, dengan kombinasi atribut *global_aktif_power, voltage*.

Dilihat dari beberapa percobaan dengan kombinasi atribut dalam prediksi konsumsi energi listrik di perumahan. Metode *Naïve Bayes* memiliki nilai *accuracy* yang tinggi tetapi tidak mencapai 100%, disebabkan oleh untuk perhitungan *Naïve Bayes* kurang bagus akurasi ketika data yang di uji tidak ada dalam *data training*. Dan untuk perhitungan *Naïve Bayes* lebih efisien digunakan untuk jumlah data yang sedikit. Metode *Naïve Bayes* menganggap atribut atau kategori yang digunakan dalam klasifikasi tidak memiliki ketergantungan satu sama lain.

4. Kesimpulan

Penelitian ini bertujuan untuk memprediksi konsumsi energi listrik dengan metode *Naïve Bayes*. Berdasarkan hasil pengujian dalam penelitian ini dapat disimpulkan hal-hal sebagai berikut:

1. Nilai *accuracy* tertinggi yaitu *accuracy* yang pada pengujian ke 3, pengujian ke 3 dengan nilai *accuracy* sebesar 90,70 %. dimana pada pengujian ke 2 menggunakan atribut *global_re-aktif_power*,
2. Nilai *Recall* tertinggi yaitu *Recall* yang pada pengujian ke 11, pengujian ke 11 dengan nilai recall sebesar 90,50%. dimana pada pengujian ke 11 menggunakan atribut *global_re-aktif_power, global_intensity*,
3. Nilai *precision* tertinggi yaitu *precision* yang pada pengujian ke 10, pengujian ke 10 dengan nilai *precision* sebesar 96,46 %. dimana pada pengujian ke 10 menggunakan atribut *global_re-aktif_power, voltage*

Daftar Pustaka

- [1] A. Silviani E Rumagit, “Prediksi Pemakaian Listrik Menggunakan Jaringan Syaraf Tiruan dan ARIMA Di Wilayah Sulluttenggo,” p. 2017.
- [2] D. T. T. R. Prita Ayuningtyas, “Prediksi Beban Listrik Pada Pt.Pln (Persero) Menggunakan Regresi Intervaldengan Neural Fuzzy,” *Jurnal Coding, Sistem Komputer UNTAN*, Vol. %1 dari %2Volume 04, No.1, pp. 1-10, 2016.
- [3] C. S. Oman Somantri, “Algoritme Genetika untuk Peningkatan Prediksi Kebutuhan Permintaan Energi Listrik,” *JNTETI*, Vol. %1 dari %2Vol. 5, No. 2,, Mei 2016.
- [4] A. Saleh, “Implementasi Metode Klasifikasi *Naïve Bayes* Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga,” *Citec Journal*, , pp. Vol. 2, No. 3,, Mei 2015 – Juli 2015.
- [5] A. Z. A.-G. S S A K Javeed Nizami, “Forecasting electric energy consumption using neural networks,” *butterworth heinemann*, vol. Vol. 23, p. 1097 I 104, 2020.
- [6] M. K. M. D. A. G. A. M. Ramadani Dwisatya, “Prediksi Beban Listrik Jangka Pendek Berbasis Algoritma Feed Forward Back Propagation Dengan Mempertimbangkan Variasi Tipe Hari,” *e-Proceeding of Engineering*, Vol. %1 dari %2 Vol.2, No.3 , p. 7315, Desember 2015.
- [7] I. G. L. C. L. ., Xinfu Song, “Electricity Consumption Prediction for Xinjiang Electric Energy Replacement,” *Hindawi*, p. 11, 2019.
- [8] A. K. Z. R. Hadapiningradja Kusumodestoni, “Prediksi Kecepatan Angin Menggunakan Model *Neural Network* Untuk Mengetahui Besar Daya Listrik Yang Dihasilkan,” *Jurnal DISPROTEK*, vol. Vol 6, pp. 53-59, 2015.
- [9] R. R. Ruliah S, “Prediksi Pemakaian Listrik Dengan Pendekatan Back Propagation,” *JUTISI* , Vol. %1 dari %2Vol. 3, No. 1, , p. 465 – 526, April 2014 .
- [10] J.-Y. K. a. S.-B. Cho, “Electric Energy Consumption Prediction by Deep Learning with State Explainable Autoencoder,” *Energies*, 2019,.