

PENERAPAN *K-MEANS* UNTUK MENGANALISIS PENGARUH CURAH HUJAN TERHADAP PRODUKSI LISTRIK (STUDI KASUS: PT. INDONESIA POWER)

Alifa Hanifah¹, Aji Primajaya², Agung Susilo Yuda Irawan³

^{1,2,3} Teknik Informatika, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang
¹alifa16030@student.unsika.ac.id, ²aji.primajaya@staff.unsika.ac.id, ³agung@unsika.ac.id

Abstrak

Listrik merupakan kebutuhan pokok bagi kehidupan masyarakat saat ini. Salah satu pembangkit tenaga listrik yaitu Pembangkit Listrik Tenaga Air (PLTA). Dalam sistemnya, PLTA memanfaatkan tenaga dari aliran/terjunan air, bendungan, atau saluran irigasi. Air merupakan energi utama dalam proses PLTA dan salah satu faktor dalam ketersediaan air adalah curah hujan sehingga karyawan PLTA lebih cenderung melihat curah hujan untuk menentukan target produksi listrik. *K-Means* merupakan metode *data clustering* non-hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*. Tujuan dari penelitian ini yaitu melakukan *cluster* terhadap produksi daya listrik dan curah hujan serta mengevaluasi hasil tersebut berhubungan atau tidak. Nilai yang digunakan dalam penelitian ini adalah nilai hasil produksi listrik (KWH) dan curah hujan (RR). Data yang digunakan sebanyak 1826 data produksi daya listrik dan data curah hujan. Pada penelitian ini ingin menganalisis pengaruh curah hujan terhadap produksi daya listrik dengan *K-Means Clustering*. Hasil dari penelitian ini adalah akurasi terhadap *clustering* produksi daya listrik dengan curah hujan.

Kata kunci : *Clustering*, Curah Hujan, *K-Means*, Produksi Daya Listrik

1. Pendahuluan

Energi listrik merupakan salah satu kebutuhan pokok masyarakat, memiliki atri penting dalam memajukan kesejahteraan umum, mencerdaskan kehidupan bangsa dan meningkatkan perekonomian (Maslichah, 2016). Dengan pesatnya pertumbuhan perindustrian, ekonomi dan teknologi informasi, dengan itu kebutuhan akan energi listrik semakin meningkat. Ketersediaan sumber energi pembangkit listrik dalam pemanfaatan sumber energi diatur dalam UU 30 Tahun 2009 Pasal 2 Tentang Ketenagalistrikan. Di Indonesia terdapat berbagai macam pembangkit listrik berdasarkan sumber energinya yaitu: Pembangkit Listrik Tenaga Air (PLTA), Pembangkit Listrik Tenaga Surya (PLTS), Pembangkit Listrik Tenaga Gas (PLTG), Pembangkit Listrik Tenaga Panas Bumi (PLTP), dan masih banyak lagi.

Menurut Badan Pusat Statistik, provinsi Jawa Barat menempati urutan ke 4 sebagai provinsi penghasil listrik atau tenaga listrik yang dibangkitkan dengan satuan *Giga Watt Hour* (GWh) terbanyak di Indonesia dalam 6 tahun yaitu pada tahun 2011, 2012, 2013, 2014, 2015, dan 2017 produksi listrik di provinsi Jawa Barat memiliki nilai yang fluktuatif atau memiliki penghasil listrik yang naik turun pada setiap tahunnya. Untuk nilai produksi listrik tiap tahunnya di provinsi Jawa Barat yaitu sebanyak 16089.53 GWh (*Giga Watt Hour*) pada tahun 2011, 17378.85 GWh (*Giga Watt Hour*) pada tahun 2012,

21615.84 GWh (*Giga Watt Hour*) pada tahun 2013, 20849.37 GWh (*Giga Watt Hour*) pada tahun 2014, 20617.92 GWh (*Giga Watt Hour*) pada tahun 2015, dan 37801,93 GWh (*Giga Watt Hour*) pada tahun 2017, dari 6 tahun tersebut hasil produksi listrik terbesar di provinsi Jawa Barat yaitu pada tahun 2017 sebesar 37801,93 GWh (*Giga Watt Hour*).

Pembangkit Listrik Tenaga Air (PLTA) merupakan salah satu pembangkit listrik yang terdapat di provinsi Jawa Barat. Menurut Peraturan Menteri Energi dan Sumber Daya Mineral tentang Pemanfaatan Sumber Energi Terbarukan untuk Penyediaan Tenaga Listrik pada BAB 1 Pasal 1 Ayat 10 Pembangkit Listrik Tenaga Air yang selanjutnya disebut Tenaga Air adalah pembangkit listrik yang memanfaatkan tenaga dari aliran/terjunan air, waduk/bendungan, atau saluran irigasi yang pembangunannya bersifat multiguna.

Air merupakan energi utama dalam proses Pembangkit Listrik Tenaga Air dan salah satu faktor penentu dalam ketersediaan air adalah curah hujan. Curah hujan merupakan ketinggian air hujan yang terkumpul dalam tempat yang datar, tidak menguap, tidak meresap dan tidak mengalir. Curah hujan juga didefinisikan sebagai tinggi air (mm) yang diterima permukaan sebelum mengalami aliran permukaan, evaporasi dan peresapan ke dalam tanah. Curah hujan merupakan salah satu faktor dalam ketersediaan air dan air digunakan untuk menjalankan turbin dalam Pembangkit Listrik Tenaga Air. (Desvina & Ratnawati, 2014)

K-Means sendiri merupakan salah satu metode *data clustering* non-hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok. Metode ini mempartisi data ke dalam *cluster*/kelompok sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu *cluster* yang sama (Metisen & Sari, 2015). Kelebihan dari algoritma *K-Means* adalah mampu mengelompokkan objek besar dengan sangat cepat sehingga mempercepat proses pengelompokkan. Selain itu, pada algoritma *K-Means* proses iterasi akan berhenti dalam keadaan optimum (Poerwanto & Fa'rifah, 2019). Pada penelitian ini pengaruh curah hujan dan produksi daya listrik dapat diketahui dengan melihat hasil dari *clustering* pada kedua *dataset* menggunakan *K-Means clustering*.

Penelitian yang dilakukan oleh (Rohi & Tumbelaka, 2017) mengenai studi kinerja Pembangkit Listrik Tenaga Air (PLTA) menghasilkan kesimpulan bahwa debit air *inflow* sangat berpengaruh pada besarnya *capacity factor* dan daya.

Atas dasar itulah, penulis ingin melakukan penelitian tentang ANALISIS PENGARUH CURAH HUJAN TERHADAP PRODUKSI DAYA LISTRIK DENGAN *K-MEANS CLUSTERING* (Studi Kasus: PT. Indonesia Power)

2. Landasan Teori

2.1 Listrik

Energi listrik adalah kemampuan untuk atau menghasilkan usaha listrik (kemampuan yang diperlukan untuk memindahkan muatan dari satu titik ke titik yang lain), dilambangkan dengan *W* (Melipurbowo, 2016).

2.2 Pembangkit Listrik Tenaga Air (PLTA)

Pembangkit listrik Tenaga Air (PLTA) adalah pembangkit listrik yang memanfaatkan tenaga (aliran) air sebagai sumber penghasil energi. PLTA termasuk sumber energi terbarukan dan layak disebut *clean energy* karena ramah lingkungan (Rahmadi, Yusuf, & Priyatman, 2015).

2.3 Curah Hujan

Curah hujan merupakan ketinggian air hujan yang terkumpul dalam tempat yang datar, tidak menguap, tidak meresap, dan tidak mengalir. Satuan curah hujan selalu dinyatakan dalam satuan millimeter atau inchi namun untuk di Indonesia satuan curah hujan yang digunakan adalah satuan millimeter (mm) (Ajr & Dwirani, 2019).

2.4 Data Mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan (Putra & Wadisma, 2018).

2.5 Imputasi

Metode imputasi adalah pengisian nilai data nilai pada suatu survei. Imputasi juga merupakan suatu alternative dalam menangani permasalahan *missing data* dalam analisis statistik. Terdapat beberapa macam pendekatan untuk imputasi (Sartika, 2018).

2.6 Normalisasi Min-Max

Min-Max normalization atau normalisasi *min-max* merupakan metode normalisasi dengan melakukan transformasi linier terhadap data asli sehingga menghasilkan keseimbangan nilai perbandingan antar data saat sebelum dan sesudah proses (Nasution, Khotimah, & Chamidah, 2019).

2.7 Clustering

Clustering adalah pengelompokan dari *record*, observasi-observasi atau kasus-kasus ke kelas yang memiliki kemiripan pada objek-objeknya. *Cluster* adalah koleksi dari *record* yang mirip, dan tidak mirip dengan *record* dari *cluster* lain (Halim & Widodo, 2017).

2.8 Algoritma K-Means

K-Means merupakan salah satu metode data *clustering* non hierarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster* atau kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik yang berbeda di kelompokkan ke dalam kelompok yang lainnya (Metisen & Sari, 2015).

2.9 Metode Elbow

Metode *Elbow* merupakan suatu metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah *cluster* terbaik dengan cara melihat persentase hasil perbandingan antara jumlah *cluster* yang akan membentuk siku pada suatu titik (Merliana, Ernawati, & Santoso, 2015).

2.10 Weka

Weka merupakan rangkaian perangkat lunak pembelajaran mesin yang ditulis dalam bahasa Java, dikembangkan di Universitas Waikato, Selandia

Baru. Perangkat lunak ini memiliki banyak algoritma *machine learning* untuk keperluan *data mining*. Weka juga memiliki banyak *tool* untuk pengolahan data, mulai dari data *preprocessing*, *classification*, *association rules*, dan *visualization* (Faid, Jasri, & Rahmawati, 2019).

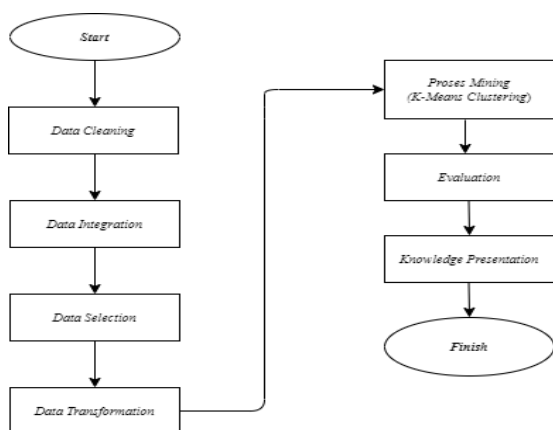
3. Metodologi Penelitian

3.1 Objek Penelitian

Objek dalam penelitian ini yaitu data curah hujan di Kabupaten Bogor dan data produksi listrik setiap harinya dari bulan Januari 2015 hingga bulan Desember 2019 pada Pembangkit Listrik Tenaga Air di PT. Indonesia Power.

3.2 Metode

Penelitian ini menggunakan metode *knowledge discovery in database*. merupakan suatu proses persimpangan dari beberapa disiplin ilmu seperti statistik, *database*, AI, visualisasi dan komputer parallel yang mempengaruhi pengetahuan (Ependi & Putra, 2019). Adapun alur penelitian dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

a. Data Cleaning

Data cleaning adalah membersihkan data dari data yang tidak konsisten dan bersifat noise karena secara umum data yang tersedia pada database memiliki isian-isian yang tidak sempurna karena hilang ataupun tidak valid.

b. Data Integration

Data Integration mengintegrasikan data dari berbagai sumber, dimana integrasi data ini dilakukan pada data-data yang memiliki atribut yang unik.

c. Data Selection

Data Selection merupakan data yang ada pada *database* yang sering kali tidak dipakai semuanya, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.

d. Data Transformation

Pada tahapan *data transformation* yaitu mentransformasikan data agar sesuai dengan kebutuhan *Data Mining*.

e. Data Mining

Pada proses *Data Mining* ini merupakan proses utama pada saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

f. Evaluation

Pada tahapan ini diharapkan adanya hasil berupa pola yang unik yang dapat digunakan sebagai model untuk memprediksi perilaku-perilaku bisnis yang sekiranya akan mempengaruhi sebuah kepurusan manajemen.

g. Knowledge

Tahapan terakhir ini merangkum dan menerjemahkan analisis dan merangkum hasil analisis yang telah didapat dengan bahasa yang semudah mungkin karena melibatkan orang lain pada manajemen yang kemungkinan besar awam dengan *Data Mining*.

4. Hasil Penelitian

4.1 Data Cleaning

Pada tahapan *data cleaning* dilakukan pembersihan data dari *missing value* dan *outlier* agar kualitas data menjadi semakin baik untuk diolah. Pada tahapan ini akan dilakukan *cleaning* pada *dataset* Hasil Produksi Listrik dan *dataset* Curah Hujan. Pada tahap ini dilakukan proses *cleaning* pada seluruh atribut yang memiliki *missing value* data tersebut akan diganti menjadi data hasil imputasi *mean*.

4.2 Data Integration

Data Intergration merupakan tahapan mengintegrasikan seluruh data, kemudian digabungkan menjadi satu *dataset*. Dalam *dataset* Hasil Produksi Listrik PT. Indonesia Power Sub Unit PLTA Kracak dan *dataset* Curah Hujan di Bogor berasal dari satu sumber yang sama, namun *dataset* tersebut terdiri dari beberapa *file* yang berbeda, sesuai dengan bulan dan tahun pencatatan datanya.

4.3 Data Selection

Proses *pre processing* dilakukan untuk memperoleh citra yang selanjutnya dapat *Data Selection* merupakan tahap pemilihan data dari sekumpulan data yang berada pada tabel, tidak semua data dipakai, hanya yang relevan pada *database* sesuai dengan kebutuhan, karena tidak semua data cocok untuk digunakan. Pada penelitian ini data Hasil Produksi Listrik didapat dari PT. Indonesia Power Sub Unit PLTA Kracak dan data Curah hujan didapat

dari Data Online Pusat Database Badan Meteorologi, Klimatologi, dan Geofisika (BMKG).

4.4 Data Transformation

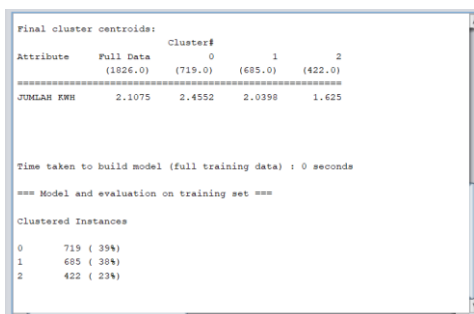
Tahap transformasi adalah tahap mengubah atau mentransformasikan bentuk *dataset* agar sesuai dengan kebutuhan *data mining* dan dapat diolah menggunakan metode *K-Means Clustering*. Pada tahap transformasi dilakukan normalisasi terhadap data dengan pengskalaan nilai atribut dari data sehingga bisa jatuh pada *range* tersebut dengan metode normalisasi *Min-Max* dengan melakukan transformasi linier terhadap data asli. Metode ini dapat menggunakan rumus sebagai berikut :

$$Newdata = \frac{(data - min) \times (newmax - newmin)}{(max - min) + newmin}$$

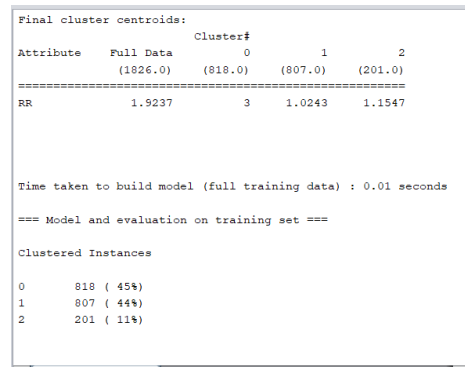
- Newdata = Data hasil normalisasi
- Min = Nilai minimum dari data per kolom
- Max = Nilai maksimum dari data per kolom
- Newmin = Batas minimum yang kita berikan
- Newmax = batas maksimum yang kita berikan

4.5 Data Mining

Setelah melakukan tahap transformasi data menjadi data yang sesuai untuk penerapan *data mining* yang akan dilakukan pemilihan teknik algoritma untuk menemukan pola atau informasi dalam pengelompokkan Curah Hujan dan Jumlah Hasil Produksi Listrik. Pada penelitian ini, menggunakan metode *Clustering* dengan algoritma *K-Means* dengan proses *clustering* menggunakan $K=3$ dengan *seed* 10.



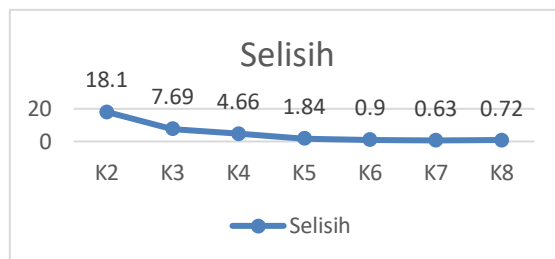
Gambar 2. Hasil *cluster* Produksi Listrik Tahun 2015-2019



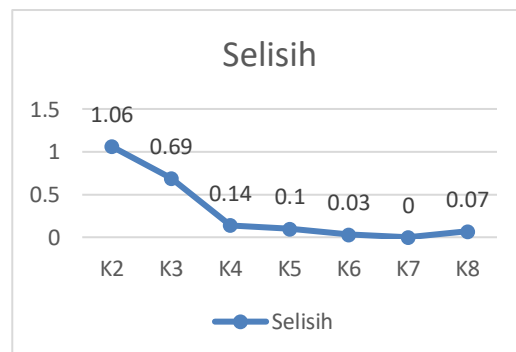
Gambar 3. Hasil *Cluster* Curah Hujan Tahun 2015-2019

4.6 Evaluation

Pada penelitian ini menggunakan metode *Elbow* untuk menghasilkan informasi dalam menentukan jumlah *cluster* terbaik dengan cara melihat persentase hasil perbandingan antara jumlah *cluster* yang akan membentuk siku pada suatu titik. Untuk mendapatkan perbandingannya dilakukan pengujian data menggunakan *Sum of Square Error* untuk menentukan data menjadi data yang sesuai. Pada penelitian sebelumnya dilakukan pengujian terhadap *cluster* dari $K=2$ sampai dengan $K=8$ untuk menentukan jumlah k terbaik dan menghasilkan $K=3$ sebagai *cluster* terbaik (Merliana et al., 2015).



Gambar 4. Grafik *Sum of Square Error* Hasil Produksi Listrik



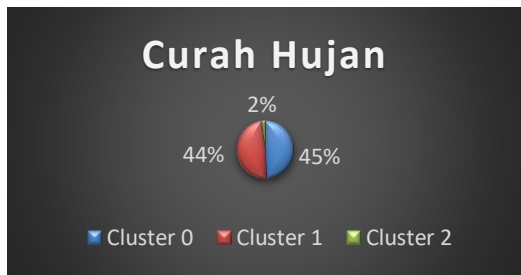
Gambar 5. Grafik *Sum of Square Error* Curah Hujan

Berdasarkan hasil dari grafik *Sum of Square Error* tersebut didapat nilai yang mengalami penurunan disertai perubahan *cluster*. Analisis dalam penentuan jumlah K optimal menggunakan metode

Elbow dengan perhitungan SSE berdasarkan hasil percobaan tersebut dengan melihat penurunan selisih nilai *Sum of Square Error* yang signifikan dan berbentuk siku diperoleh jumlah *cluster* optimal 3.

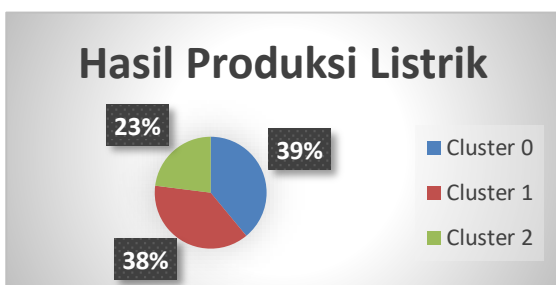
4.7 Knowledge

Berdasarkan hasil dari evaluasi *cluster* yang telah terbentuk melalui proses evaluasi dengan metode *Sum of Error* (SSE) maka didapat pengetahuan baru sebagai berikut :



Gambar 6. Hasil Presentasi *Cluster* pada Dataset Curah Hujan

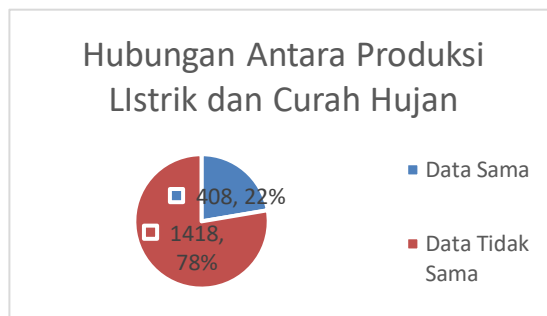
Pada diagram diatas, dapat dilihat bahwa *Cluster* 2 memiliki anggota paling sedikit. Berdasarkan tabel hasil *clustering*, dapat dilihat hasil produksi listrik yang terasuk ke dalam *Cluster* 0 sebagian besar memiliki frekuensi tinggi (curah hujan tinggi), *cluster* 1 memiliki frekuensi curah hujan yang rendah dan *Cluster* 2 memiliki sebagian frekuensi curah hujan sedang. Oleh karena itu, curah hujan *cluster* 0 dapat dikategorika tinggi, *cluster* 2 dapat dikategorikan sedang dan *cluster* 1 dapat dikategorikan rendah.



Gambar 7. Hasil Presentasi *Cluster* pada Dataset Curah Hujan

Pada diagram diatas, dapat dilihat bahwa *Cluster* 0 memiliki anggota paling banyak. Berdasarkan tabel hasil *clustering*, dapat dilihat hasil produksi listrik yang termasuk ke dalam *Cluster* 2 sebagian besar memiliki frekuensi hasil produksi terendah (hasil produksi listriknya rendah), *Cluster* 1 sebagian besar memiliki frekuensi hasil produksi listrik sedang dan *Cluster* 0 memiliki sebagian besar frekuensi hasil produksi listrik tinggi. Oleh karena itu, hasil produksi listrik *Cluster* 0 dapat dikategorikan tinggi, *Cluster* 1 dapat dikategorikan sedang, dan *Cluster* 2 dikategorikan rendah.

Berdasarkan hasil *cluster* terhadap data Hasil Produksi Listrik yang didapat dari PT. Indonesia Power dan data Curah Hujan yang didapat dari Data Online Pusat Database Badan Meteorologi, Klimatologi, dan Geofisika (BMKG. Dalam 5 tahun terakhir atau pada tahun 2015 hingga 2019 data hasil *clustering* menunjukkan hasil data kesamaan dengan melihat data hasil *cluster* antara hasil produksi listrik dan curah hujan adalah seperti pada gambar di bawah ini.



Gambar 8. Grafik Perbandingan Produksi Listrik dan Curah Hujan

Dari grafik diatas didapatkan hasil produksi listrik dan curah hujan hanya sebanyak 22% atau sebanyak 408 memiliki data hasil *clustering* yang sama dan hasil produksi listrik dan hujan yang tidak sama adalah 78% atau sebanyak 1418 data. Pada pengamatan yang telah penulis lakukan di PT. Indonesia Power Sub Unit PLTA Kracak ada faktor lain yang mempengaruhi hasil produksi listrik yaitu mesin yang berfungsi dengan baik atau tidak.

5. Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilaksanakan maka didapatkan kesimpulan sebagai berikut:

1. Hasil Produksi Listrik di PT Indonesia Power dan Curah Hujan BMKG dengan menggunakan algoritma *K-Means Clustering* dalam proses pengolahan datanya dan menggunakan Metode *Elbow* dengan perhitungan *Sum Square of Error* dalam proses evaluasi *dataset* menghasilkan 3 *cluster*. Secara keseluruhan dalam tahapan *data mining* pada penelitian ini dibantu dengan *tools* WEKA 3.8.4.
2. Secara keseluruhan hasil dari evaluasi menggunakan Metode *Elbow* dengan perhitungan *Sum Square of Error* menunjukkan bahwa hasil dari *cluster* Hasil Produksi Listrik pada PT. Indonesia Power dan *cluster* Curah Hujan menunjukkan hasil yang 22% memiliki data hasil *cluster* yang sesuai dan 78% memiliki data yang tidak sesuai atau sebanyak 1826 data hasil produksi listrik hanya sebanyak 408 data yang sesuai dengan curah hujan sisanya sebanyak 1418 data hasil produksi listrik tidak

sesuai dengan curah hujan. Hasilnya ditunjukkan pada grafik yang telah dijelaskan pada tahapan *Knowledge*. Hal tersebut menentukan bahwa curah hujan sebagai salah satu faktor yang mempengaruhi debit air tidak berpengaruh terhadap hasil dari produksi daya listrik namun ada faktor lain yang perlu dimasukkan dalam pembuatan model karena akan mempengaruhi hasil dari *clustering*.

5.2 Saran

Dari hasil penelitian yang telah dilakukan, maka saran yang diberikan untuk referensi penelitian selanjutnya yaitu:

1. Diharapkan pada penelitian selanjutnya dapat menggunakan scenario lain untuk *clustering* yang lebih baik untuk proses *clustering* untuk mengetahui pengaruh curah hujan terhadap produksi listrik.
2. Penelitian ini hanya dilakukan di satu Pembangkit Listrik Tenaga Air (PLTA), diharapkan pada penelitian selanjutnya dilakukan pada lebih dari satu Pembangkit Listrik tenaga Air (PLTA)

Daftar Pustaka:

- Ajr, E. Q., & Dwirani, F. (2019). Menentukan stasiun hujan dan curah hujan dengan metode polygon thiessen daerah kabupaten lebak. *Jurnal Lingkungan Dan Sipil*, 2(2), 139–146.
- Desvina, A. P., & Ratnawati. (2014). Penerapan model vector autoregressive (VAR) untuk peramalan curah hujan kota pekanbaru. *Jurnal Sains, Teknologi Dan Industri*, 11(2), 151–159.
- Ependi, U., & Putra, A. (2019). Solusi prediksi persediaan barang dengan menggunakan algoritma apriori (studi kasus: regional part depo auto 2000 Palembang). *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 5(2), 139–145.
<https://doi.org/10.26418/jp.v5i2.32648>
- Faid, M., Jasri, M., & Rahmawati, T. (2019). Perbandingan kinerja tool data mining weka dan rapidminer dalam algoritma klasifikasi. *Teknika*, 8(1), 11–16.
<https://doi.org/10.34148/teknika.v8i1.95>
- Halim, N. N., & Widodo, E. (2017). Clustering dampak gempa bumi di indonesia menggunakan kohonen self organizing maps. *Prosiding SI MaNIS (Seminar Nasional Integrasi Matematika Dan Nilai Islami)*, 1(1), 188–194.
- Maslichah, K. (2016). Hemat energi listrik studi kasus di badan diklat provinsi banten. *Jurnal Lingkar Widya Swara*, 3(1), 47–52. Retrieved from http://juliwi.com/published/E0301/Juliwi0301_47-52.pdf
- Merliana, N. P. E., Ernawati, & Santoso, A. J. (2015). Analisa penentuan jumlah cluster terbaik pada metode k-means clustering. *Prosiding Seminar Nasional Multi Disiplin Ilmu & Call For Paper UNISBANK (SENDI_U)*, 978–979.
- Metisen, B. M., & Sari, H. L. (2015). Analisis clustering menggunakan metode k-means dalam pengelompokan penjualan produk pada swalayan fadhila. *Jurnal Media Infotama*, 11(2), 110–118.
- Nasution, D. A., Khotimah, H. H., & Chamidah, N. (2019). Perbandingan normalisasi data untuk klasifikasi wine menggunakan algoritma k-nn. *Computer Engineering, Science and System Journal*, 4(1), 78.
<https://doi.org/10.24114/cess.v4i1.11458>
- Poerwanto, B., & Fa'rifah, R. Y. (2019). Algoritma k-means dalam mengelompokkan kecamatan di tana luwu berdasarkan produktifitas hasil pertanian. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.
<https://doi.org/10.1017/CBO9781107415324.004>
- Putra, R. R., & Wadisma, C. (2018). Implementasi data mining pemilihan pelanggan potensial menggunakan algoritma k-means. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.
<https://doi.org/10.1017/CBO9781107415324.004>
- Rahmadi, J., Yusuf, I., & Priyatman, H. (2015). Studi kelayakan pemanfaatan pembangkit listrik kincir air terapung desa ella hilir kecamatan ella hilir kabupaten melawi. *Jurnal ELKHA*, 7(1), 11–18.
- Rohi, D., & Tumbelaka, H. H. (2017). Studi kinerja pembangkit listrik tenaga air (plta) di daerah aliran sungai (das) brantas. *Jurnal Teknik Elektro*, 10(1), 17–23.
<https://doi.org/10.9744/jte.10.1.17-23>
- Sartika, E. (2018). Analisis metode k nearest neighbor imputation (knni) untuk mengatasi data hilang pada estimasi data survey. *TEDC*, 12(3), 219–227.