

PEMODELAN PRODUKSI ENERGI PEMBANGKIT LISTRIK TENAGA SURYA 1000 WP DENGAN ALGORITMA *NAIVE BAYES*

ARIF FAHRIZAL¹, DANI RUSIRAWAN², LITA LIDYAWATI³

^{1,2,3}Institut Teknologi Nasional Bandung
Email: ariffahrizal1987@gmail.com

ABSTRAK

Dalam penelitian ini, akan diuraikan pemodelan karakteristik produksi energi Pembangkit Listrik Tenaga Surya (PLTS) 1000 *Watt peak* (Wp) dalam periode Juni 2020-Desember 2021, dengan menggunakan salah satu algoritma *machine learning*, yaitu *Naive Bayes*. Fitur yang digunakan dalam pemodelan adalah radiasi matahari, temperatur & pancaran sinar, yang dibagi menjadi lima label kelas yaitu *Very low*, *Low*, *Medium*, *High*, dan *Very High*. Perbandingan data latih dan uji dibagi menjadi 3 skenario, yaitu 90%:10%, 80%:20%, 75%:25%. Hasil pemodelan memperlihatkan bahwa skenario perbandingan 75%:25% memiliki nilai *mean absolute error* (MAE), *root mean squared error* (RMSE) dan *mean absolute percentage error* (MAPE) terkecil (dibanding skenario lain), masing-masing sebesar 0,14; 0,38; dan 6,52, yang artinya skenario ini memiliki akurasi relative lebih tinggi (dibanding skenario lain), jika dibandingkan terhadap karakteristik produksi energi referensi.

Kata kunci: modul fotovoltaik, *machine learning*, MAE, RMSE, MAPE

ABSTRACT

In this research, the modeling of energy production characteristics of 1000 Watt peak of the Solar Power Plant (SPP) will be elaborated for the period of June 2020-December 2021, using one of the machine learning algorithms, i.e., Naive Bayes. The features used are solar radiation, temperature & light beam. Each of feature is divided into five class labels, namely Very low, Low, Medium, High, and Very High. The comparison of training and test data is divided into 3 scenarios i.e., 90%: 10%, 80%: 20%, 75%: 25%. Based on modeling, it is found that a 75%;25% scenario shows the smallest of MAE, RMSE, and MAPE, i.e., 0.14, 0.38, and 6.52, respectively. It means that this scenario has the highest accuracy in this modeling (if compared to the actual's energy production characteristics).

Keywords: photovolataic module, *machine learning*, MAE, RMSE, MAPE

1. PENDAHULUAN

Kebutuhan energi dunia terus meningkat, dan berdasarkan laporan *Internasional Energy Agency* (IEA), kebutuhan energi dunia di tahun 2030 akan meningkat sebesar 45%. Di sisi lain, rata-rata kebutuhan energi di Indonesia meningkat sebesar 1,6 % per tahun **(Rudiyanto, dkk, 2019)**. Untuk meningkatkan pasokan energi berkelanjutan, pembangkit listrik berbasis energi terbarukan merupakan satu-satunya solusi **(Adibhatla & Kaushik, 2017)**. Potensi energi terbarukan di Indonesia sangat besar, namun belum tergarap secara maksimal **(Suharyati, dkk, 2019)**. Radiasi matahari adalah sumber energi yang paling tua dan modern sehingga dapat dijadikan sumber energi terbarukan **(Boyle, 2004)**. Produksi energi yang diperoleh instalasi *photovoltaic (PV)*/fotovoltaik (FV) dapat diestimasi melalui label klasifikasi melalui *Machine Learning* sehingga dapat digunakan untuk melatih berbagai skenario pengujian dengan algoritma klasifikasi *naïve bayes* dan *k-nearest neighbors*, dengan menggunakan fitur radiasi matahari harian, suhu harian dan pancaran sinar harian **(Ikhsan, 2020)**.

Sistem PLTS dapat mengubah langsung energi surya yang diterima oleh sel FV menjadi energi listrik berdasarkan efek fotoelektrik **(Kou, 2015)**. Beberapa sistem PLTS yang umum diantaranya adalah *stand alone*, *grid connected* dan hybrid **(Rusirawan D, 2012)**. Prinsip kerja sel mirip dengan instrumen semikonduktor dioda surya **(Zulkifli, 2011)**, **(Sari, 2019)**. Komponen-komponen elektron akan keluar dari lapisan anti refleksi sel surya sebagai akibat berinteraksi dengan cahaya, sehingga sel fotovoltaik menghasilkan arus searah (DC), yang dapat digunakan untuk peralatan DC, sedangkan untuk peralatan AC membutuhkan komponen lain selain modul surya **(Jos, 2017)**, **(Sari, 2019)**. Sementara itu, pembangkit listrik tenaga surya yang tidak mengubah energi surya menjadi energi listrik secara langsung disebut sistem tenaga surya terkonsentrasi (*concentrated solar power system*). Sistem ini bekerja dengan cara memusatkan energi matahari dari luar untuk ditangkap oleh suatu kaca dan diarahkan pada satu area kecil, dan energi matahari tersebut dapat diubah menjadi energi panas, yang selanjutnya dapat menggerakkan siklus turbin uap sebagai penggerak generator **(Bhutka, Gajjar, & Harinarayana, 2016)**. Memperhatikan karakteristik *I-V* (arus terhadap tegangan) dari FV, terlihat bahwa pada temperatur modul konstan, makin besar irradiasi akan berpengaruh besar terhadap perubahan arus hubung singkat (I_{sc}), tetapi efeknya akan kecil terhadap perubahan tegangan sirkuit terbuka (V_{oc}), sedangkan berdasarkan karakteristik *P-V* (daya terhadap tegangan) pada temperatur modul konstan, peningkatan irradiasi matahari akan memberikan pengaruh signifikan terhadap daya keluaran modul FV. Dari karakteristik ini terlihat bahwa daya akan memiliki puncak pada tegangan tertentu, dan jika tegangan tertentu telah dicapai, maka daya keluaran yang dihasilkan akan turun kembali **(Rusirawan & Farkas, 2014)**, **(Sari, 2019)**.

Tahun 2013-2014 telah dilakukan penelitian mengenai produksi energi listrik harian FV dengan metode klasifikasi *naïve bayes*. Fitur yang digunakan adalah radiasi matahari, temperatur & pancaran sinar. Setiap fitur dibagi menjadi lima label kelas yaitu *Very low*, *Low*, *Medium*, *High*, dan *Very High* agar proses klasifikasi lebih sederhana. Hasilnya diperoleh prediksi *recall* 98,305 % dan akurasi 82,191% **(Bayindir, dkk, 2017)**. Selain itu, di Singapura telah dilakukan peramalan produksi energi FV jangka pendek menggunakan metode *Naïve Bayes Classifier* (NBC). Parameter input yang digunakan seperti daya sesaat, temperatur luar ruangan, temperatur panel, penyinaran di lokasi dan waktu harian untuk meramalkan energi yang akan dipanen oleh sistem FV dalam periode 15 menit berikutnya. Hasil perkiraan diklasifikasikan menjadi lima tipe yaitu *very low*, *low*, *medium*, *high* dan *very high*. Data uji historis dari sistem PV FV yang ada di Singapura digunakan untuk mengevaluasi akurasi metode peramalan NBC. Hasil perbandingan menunjukkan bahwa metode yang diusulkan mampu mencapai akurasi peramalan lebih dari 68 persen **(Quek, Woo, & Logenthiran, 2017)**.

Institut Teknologi Nasional Bandung (Itenas Bandung telah mengoperasikan PLTS jenis *grid-connected system* 1 kWp sejak Januari 2018. PLTS ini memiliki bagian utama modul FV dan inverter. Tujuan dari didirikannya PLTS ini selain digunakan sebagai penerangan pada sebagian koridor gedung 1 (Gedung kuliah dan studio Prodi Desain) lantai 4 digunakan juga untuk sarana penelitian, sarana edukasi bagi mahasiswa, pelajar serta umum. Gambar 1 menunjukkan Instalasi PLTS 1 kWp di Itenas (**Hidayat, Rusirawan, & Tanjung, 2019**).



Gambar 1. Instalasi FV PLTS Institut Teknologi Nasional Bandung

Dalam penelitian ini pemodelan produksi energi pembangkit listrik tenaga surya 1 kWp akan diuraikan dengan menggunakan algoritma *naive bayes*, sebagai metode klasifikasi dari *machine learning* dengan menggunakan beberapa skenario perbandingan data latih dan data uji. Modul FV PLTS yang dijadikan objek penelitian ini adalah modul JAM6-60-250, yang terletak di *roof-top* gedung 1, Institut Teknologi Nasional Bandung. Secara spesifik modul ini terletak pada ketinggian 714 meter diatas permukaan laut dengan letak geografis 6°90" lintang selatan dan 107°64" bujur timur.

2. METODE PENELITIAN

2.1 *Naive Bayes*

Klasifikasi *naive bayes* adalah suatu metode pengklasifikasian statistik berdasarkan pada teorema *bayes*. Dengan teorema ini peluang di waktu yang akan datang dapat diperkirakan (*forecasting*) berdasarkan pengalaman di waktu sebelumnya dengan asumsi bahwa setiap fitur penentu keputusan bersifat bebas (*independence*) sehingga keberadaan setiap fitur tidak ada kaitannya dengan keberadaan fitur yang lain. Dalam konteks *Naive* kondisi antar fitur diasumsikan saling bebas. Teorema *Bayes* ditunjukkan oleh persamaan (2.1) (**Ikhsan, 2020**).

$$P(A|B) = \frac{(P(B|A) \times P(A))}{P(B)} \quad (2.1)$$

Dimana:

- P(A|B) : Fitur B (*posterior probability*) mendasari probabilitas label A
- P(B|A) : Label A (*likelihood*) mendasari probabilitas fitur B
- P(A) : Probabilitas label A (*prior probability*)
- P(B) : Probabilitas fitur B (*evidence*)

Label yang memiliki *posterior probability* paling tinggi akan dijadikan hasil prediksi, dan dapat menentukan hasil prediksi label untuk suatu sampel data dengan suatu perbandingan. Pada penelitian ini digunakan lima fitur kelas dengan menggunakan tipe sangat rendah, rendah, sedang, tinggi, dan sangat tinggi. Sementara itu, parameter atau fitur yang digunakan adalah temperatur, pancaran, dan radiasi.

2.2 Confusion Matrix

Kinerja model klasifikasi (*classifier*) pada satu set data uji, nilai sebenarnya dapat diketahui dan digambarkan melalui *confusion matrix*. *Error matrix* atau *confusion matrix* adalah istilah yang sama pada *machine learning* dan khususnya pada klasifikasi statistik. Kinerja suatu algoritma dapat digambarkan pada tabel 2.1, *confusion matrix*. Hasil prediksi pada masalah klasifikasi, jumlah prediksi benar dan salah dirangkum dengan nilai-nilainya, serta dipecah kepada masing-masing label, ditampilkan di tabel 2.1, sehingga dapat diketahui jenis kesalahan yang dibuat oleh *classifier*.

Tabel 2.1 Confusion Matrix

		Fitur Perkiraan	
		Positive (P)	Negative (N)
Fitur Sebenarnya	P	True Positive (TP)	False Negative (FN)
	N	False Positive (FP)	True Negative (TN)

Dimana:

True Positive : Data positif dan diperkirakan benar

True Negative : Data negatif dan diperkirakan benar

False Positive : Data negatif namun diperkirakan sebagai data positif (error tipe 1)

False Negative: Data positif namun diperkirakan sebagai data negatif (error tipe 2)

Tabel 2.2 memperlihatkan *confusion matrix* lebih kompleks dibandingkan dengan tabel 2.1 karena menggunakan 5 fitur klasifikasi.

Tabel 2.2 Confusion Matrix 5 × 5

		Fitur Perkiraan				
		Fitur A	Fitur B	Fitur C	Fitur D	Fitur E
Fitur Sebenarnya	Fitur A	AA	AB	AC	AD	AE
	Fitur B	BA	BB	BC	BD	BE
	Fitur C	CA	CB	CC	CD	CE
	Fitur D	DA	DB	DC	DD	DE
	Fitur E	EA	EB	EC	ED	EE

2.3 Penilaian Model

a. Akurasi

Jumlah perkiraan benar dibagi dengan keseluruhan jumlah data yaitu akurasi. Persamaan (2.2) dapat digunakan untuk menghitung akurasi

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+ FN} \quad (2.2)$$

b. Presisi

Presisi yaitu perbandingan banyaknya perkiraan yang benar-benar positif dari semua kelas positif yang diperkirakan dengan benar. Nilai presisi dapat dihitung menggunakan persamaan (2.3).

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (2.3)$$

c. Recall

Recall atau *sensitivity* yaitu banyaknya perkiraan yang benar dari semua kelas positif. Nilai *recall* dapat dihitung dengan persamaan (2.4).

$$Recall = \frac{TP}{TP+FN} \quad (2.4)$$

d. F-Score

Dengan diperbanyak prediksi sampel pada kelas positif akan mengakibatkan *FP* juga bertambah, dan *recall* bisa meningkat secara maksimal. Selanjutnya, presisi akan dibuat turun karena tujuan utama dari evaluasi model menggunakan presisi adalah mengurangi jumlah *FP*. Sementara itu, untuk *FN* yang mempengaruhi *recall*, berlaku juga sebaliknya. Model yang seimbang membutuhkan suatu cara. Mencari nilai tengah dari presisi dan *recall* yang merupakan rata-rata harmonik dari keduanya dapat menggunakan *F-Score*, sehingga *F-Score* bisa dihitung dengan menggunakan persamaan (2.5).

$$F - Score = 2 \times \frac{(\text{Presisi} \times \text{Recall})}{\text{Presisi} + \text{Recall}} \quad (2.5)$$

2.4 Evaluasi Model

Setelah melakukan *modeling*, model tersebut perlu dievaluasi untuk menentukan seberapa besar *error* yang dihasilkan dibandingkan dengan kondisi actual/referensi. Penentuan kecocokan model dapat dilakukan dengan beberapa cara, diantaranya dengan menghitung nilai RMSE (*Root Mean Squared Error*), MAPE (*Mean Absolute Percentage Error*) dan MAE (*Mean Absolute Error*) (AI Muluk, 2021).

a. Mean Absolute Error (MAE)

MAE mengukur variabilitas *error* hasil prediksi dari model yang dibuat (Montgomery, Jennings, & Kulahci, 2015). Model yang dibuat diharapkan memiliki variabilitas *error* kecil. MAE dapat dirumuskan melalui persamaan (2.6).

$$MAE = \frac{1}{N} \sum_{i=1}^n |y_i - f_i| \quad (2.6)$$

Dimana:

- y_i = *Expected value* (Data Actual)
- f_i = *Observed value* (Data Forecasting)
- N = Panjang sample (Jumlah Data)

b. Root Mean Squared Error (RMSE)

Seperti MAE, RMSE digunakan untuk melihat variabilitas *error* dari nilai prediksi dengan mengkuadratkan selisih dari nilai sebenarnya. Namun karena nilai kuadrat, maka akan menghasilkan besaran *error* yang dikuadratkan. Untuk memperkecilnya dilakukan persamaan matematika dengan mengakarkan hasil MSE. Nilai RMSE ini dapat dihitung dengan persamaan (2.7).

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (y_i - f_i)^2} \quad (2.7)$$

Dimana:

- y_i = *Expected value* (Data actual)
- f_i = *Observed value* (Data forecasting)
- N = Panjang sample (Jumlah Data)

c. Mean Absolute Percentage Error (MAPE)

MAPE bisa difungsikan untuk menghitung nilai *error* hasil prediksi relatif terhadap nilai aslinya. Perhitungan MAPE dapat dirumuskan seperti pada persamaan (2.8).

$$MAPE = \frac{1}{N} \sum_{i=1}^n \left| \frac{y_i - f_i}{y_i} \right| \times 100\% \quad (2.8)$$

Dimana:

y_i = *Expected value (Data actual)*

f_i = *Observed value (Data forecasting)*

N = Panjang sample (Jumlah data)

3. HASIL DAN PEMBAHASAN

3.1 Analisa Data Keseluruhan

Data yang digunakan untuk pemodelan diambil dari tanggal 1 Juni 2020 - tanggal 31 Desember 2021, dengan jumlah 579 sampel data. Parameter masukan yang digunakan ada 3 yaitu temperatur (°C), pancaran (*hour*) dan radiasi matahari (kWh/m sementara parameter luaran adalah data analisa yang digunakan, yaitu 1 label klasifikasi data Energi (kWh).

Analisis pemodelan produksi energi dengan metode algoritma *naive bayes* berupa grafik disertai dengan penilaian akurasi, presisi, *recall* atau *sensitivity* dan *f-score* sedangkan untuk evaluasi model dilakukan melalui nilai RMSE (*Root Mean Squared Error*), MAPE (*Mean Absolute Percentage Error*) dan MAE (*Mean Absolute Error*) dengan menggunakan 3 (tiga) skenario perbandingan data latih dan data uji, yaitu skenario pertama 90%:10%, skenario kedua 80%:20%, dan skenario ketiga 75%:25%. Sebelum mengevaluasi seluruh skenario, terlebih dahulu dibuat pembagian nilai *range* atau jarak pada masing-masing parameter seperti terlihat pada tabel 3.1-3.4.

Tabel 3.1. Pembagian Jarak Pada Energi

Kelas	Kategori Klasifikasi	Nilai Jarak	
		Minimal	Maksimal
Kelas 1	<i>Very low</i>	0	1,152
Kelas 2	<i>Low</i>	1,153	2,305
Kelas 3	<i>Medium</i>	2,306	3,458
Kelas 4	<i>High</i>	3,459	4,611
Kelas 5	<i>Very High</i>	4,612	5,764

Tabel 3.2. Pembagian Jarak Pada Suhu

Kelas	Kategori Klasifikasi	Nilai Jarak	
		Minimal	Maksimal
Kelas 1	<i>Very low</i>	0	4,73
Kelas 2	<i>Low</i>	4,731	9,461
Kelas 3	<i>Medium</i>	9,462	14,192
Kelas 4	<i>High</i>	14,193	18,923
Kelas 5	<i>Very High</i>	18,924	23,654

Tabel 3.3. Pembagian Jarak Pada Pancaran

Kelas	Kategori Klasifikasi	Nilai Jarak	
		Minimal	Maksimal
Kelas 1	<i>Very low</i>	0	1,96
Kelas 2	<i>Low</i>	1,961	3,921
Kelas 3	<i>Medium</i>	3,922	5,882
Kelas 4	<i>High</i>	5,883	7,843
Kelas 5	<i>Very High</i>	7,844	9,804

Tabel 3.4. Pembagian Jarak Pada Radiasi

Kelas	Kategori Klasifikasi	Nilai Jarak	
		Minimal	Maksimal
Kelas 1	<i>Very low</i>	0	1,37
Kelas 2	<i>Low</i>	1,371	2,741
Kelas 3	<i>Medium</i>	2,742	4,112
Kelas 4	<i>High</i>	4,113	5,483
Kelas 5	<i>Very High</i>	5,484	6,854

Setelah dibuat pembagian *range* atau jarak pada masing-masing parameter, kemudian dibuat pemodelan dengan algoritma *naive bayes* yang ditunjukkan pada tabel 3.5, untuk mendapatkan peluang data pada masing-masing parameter berdasarkan 5 label kelas yang dikategorikan sebagai *very low, low, medium, high, very high*.

Tabel 3.5. Perhitungan Data Peluang Metode Algoritma *Naive Bayes*

P(Very Low/Low/Medium/High/Very High)	1%	7%	37%	39%	16%
Jumlah	100%				
P(Suhu=↓ ...	<i>Very Low</i>	<i>Low</i>	<i>Medium</i>	<i>High</i>	<i>Very High</i>
<i>Very Low</i>	0%	0%	0%	0%	0%
<i>Low</i>	0%	0%	0%	0%	0%
<i>Medium</i>	0%	0%	0%	0%	0%
<i>High</i>	0%	0%	0%	0%	2%
<i>Very High</i>	100%	100%	100%	100%	98%
Jumlah	100%	100%	100%	100%	100%
P(Pancaran=↓ ...	<i>Very Low</i>	<i>Low</i>	<i>Medium</i>	<i>High</i>	<i>Very High</i>
<i>Very Low</i>	33%	16%	19%	12%	12%
<i>Low</i>	50%	28%	19%	17%	7%
<i>Medium</i>	17%	23%	31%	25%	25%
<i>High</i>	0%	28%	25%	29%	23%
<i>Very High</i>	0%	5%	6%	16%	34%
Jumlah	100%	100%	100%	100%	100%
P(Radiasi=↓ ...	<i>Very Low</i>	<i>Low</i>	<i>Medium</i>	<i>High</i>	<i>Very High</i>
<i>Very Low</i>	0%	9%	0%	0%	0%
<i>Low</i>	17%	26%	6%	1%	0%
<i>Medium</i>	33%	35%	24%	8%	3%
<i>High</i>	50%	30%	59%	64%	21%
<i>Very High</i>	0%	0%	11%	26%	76%
Jumlah	100%	100%	100%	100%	100%

3.2 Analisa Data Metode Algoritma *Naïve Bayes*

3.2.1 Analisa Data Metode Algoritma *Naïve Bayes* Skenario 1

Analisis data metode algoritma *naïve bayes* skenario 1 menggunakan data latih 521 sampel dan data uji 58 sampel. Dari analisis akan diperoleh data berupa data prediksi berdasarkan label klas berupa kategori *very low, low, medium, high, very high*. Setelah itu kemudian dibuatkan suatu data dengan sistem *confusion table* atau *confusion matrix* seperti diperlihatkan pada tabel 3.6. Setelah data *confusion matrix* diperoleh, maka bisa dihitung nilai-nilai akurasi, presisi, *recall* atau *sensitivity* dan *f-score*.

Tabel 3.6. Data *Confusion Matrix* Skenario 1

Actual ↓	Prediksi				
	<i>Very Low</i>	<i>Low</i>	<i>Medium</i>	<i>High</i>	<i>Very High</i>
<i>Very Low</i>	0	0	0	0	0
<i>Low</i>	0	5	0	2	0
<i>Medium</i>	0	0	0	0	0
<i>High</i>	0	0	0	8	1
<i>Very High</i>	0	0	0	0	2

Dari perhitungan data *confusion matrix* skenario 1 yang ditunjukkan pada tabel 3.6, diperoleh beberapa data sebagai parameter penilaian model produksi energi, yaitu:

1. Menghitung Akurasi

Diperoleh data-data sebagai berikut:

Data *Very Low* (*True Positif*) = 0, Data *Low* (*True Positif*) = 5, Data *Medium* (*True Positif*) = 0, Data *High* (*True Positif*) = 8, Data *Very High* (*True Positif*) = 2

Total Data Set = 18

Mencari Akurasi = (*True Positif* / Total Data Set)

$$\begin{aligned} \text{Akurasi} &= (0 + 5 + 0 + 8 + 2) / 18 \\ &= 0,83 \times 100\% = 83\% \end{aligned}$$

2. Menghitung Presisi

Diperoleh data-data sebagai berikut:

Data *Very Low* (*True Positif*) = 0 dan (*False Positif*) = 0

Data *Low* (*True Positif*) = 5 dan (*False Positif*) = 0

Data *Medium* (*True Positif*) = 0 dan (*False Positif*) = 0

Data *High* (*True Positif*) = 8 dan (*False Positif*) = 2

Data *Very High* (*True Positif*) = 2 dan (*False Positif*) = 1

Mencari Presisi = *True Positif* / (*True Positif* + *False Positif*)

$$\begin{aligned} \text{Presisi} &= P(\textit{Very Low}) + P(\textit{Low}) + P(\textit{Medium}) + P(\textit{High}) + P(\textit{Very High}) / \text{Jumlah Klas} \\ &= (0 + 1 + 0 + 0,8 + 0,67) / 5 \\ &= 0,49 \times 100\% = 49\% \end{aligned}$$

3. Menghitung *Recall* atau *Sensitivity*

Diperoleh data-data sebagai berikut:

Data *Very Low* (*True Positif*) = 0 dan (*False Negatif*) = 0

Data *Low* (*True Positif*) = 5 dan (*False Negatif*) = 2

Data *Medium* (*True Positif*) = 0 dan (*False Negatif*) = 0

Data *High* (*True Positif*) = 8 dan (*False Negatif*) = 1

Data *Very High* (*True Positif*) = 2 dan (*False Negatif*) = 0
 Mencari *Recall* atau *Sensitivity* = $\frac{\text{True Positif}}{\text{True Positif} + \text{False Negatif}}$
 $\text{Recall} = \frac{P(\text{Very Low}) + P(\text{Low}) + P(\text{Medium}) + P(\text{High}) + P(\text{Very High})}{\text{Jumlah Klas}}$
 $= \frac{(0 + 0,72 + 0 + 0,89 + 1)}{5}$
 $= 0,52 \times 100\% = 52\%$

4. Menghitung *F-Score*

$F - \text{Score} = \frac{2(\text{Presisi} \times \text{Recall})}{\text{Presisi} + \text{Recall}}$
 $F - \text{Score} = \frac{2(0,49 \times 0,52)}{0,49 + 0,52}$
 $F - \text{Score} = 0,507 \times 100\% = 51\%$

3.2.2 Analisa Data Metode Algoritma *Naive Bayes* Skenario 2

Analisis data metode algoritma *naive bayes* skenario 2 menggunakan data latih 463 sampel dan data uji 116 sampel. Dari analisis akan diperoleh data berupa data prediksi berdasarkan label klas berupa kategori *very low, low, medium, high, very high*. Setelah itu dibuat data dengan sistem *confusion table* atau *confusion matrix* seperti diperlihatkan pada tabel 3.7. Setelah data *confusion matrix* diperoleh maka dapat dihitung nilai-nilai akurasi, presisi, *recall* atau *sensitivity* dan *f-score*.

Tabel 3.7. Data *Confusion Matrix* Skenario 2

Actual ↓	Prediksi				
	<i>Very Low</i>	<i>Low</i>	<i>Medium</i>	<i>High</i>	<i>Very High</i>
<i>Very Low</i>	0	0	0	0	0
<i>Low</i>	0	6	0	2	0
<i>Medium</i>	0	0	0	0	0
<i>High</i>	0	0	0	27	1
<i>Very High</i>	0	0	0	1	6

Dari perhitungan data *confusion matrix* pada tabel 3.7 diperoleh beberapa data sebagai parameter penilaian model produksi energi, yaitu:

1. Menghitung Akurasi

Diperoleh data-data sebagai berikut:
 Data *Very Low* (*True Positif*) = 0, Data *Low* (*True Positif*) = 6, Data *Medium* (*True Positif*) = 0
 Data *High* (*True Positif*) = 27, Data *Very High* (*True Positif*) = 6
 Total Data Set = 43
 Mencari Akurasi = $\frac{\text{True Positif}}{\text{Total Data Set}}$
 Akurasi = $\frac{(0 + 6 + 0 + 27 + 6)}{43}$
 $= 0,91 \times 100\% = 91\%$

2. Menghitung Presisi

Diperoleh data-data sebagai berikut:
 Data *Very Low* (*True Positif*) = 0 dan (*False Positif*) = 0
 Data *Low* (*True Positif*) = 6 dan (*False Positif*) = 0
 Data *Medium* (*True Positif*) = 0 dan (*False Positif*) = 0

Data *High* (True Positif) = 27 dan (False Positif) = 3
 Data *Very High* (True Positif) = 6 dan (False Positif) = 1
 Mencari Presisi = True Positif / (True Positif + False Positif)
 Presisi = P(*Very Low*) + P(*Low*) + P(*Medium*) + P(*High*) + P(*Very High*) / Jumlah Klas
 = (0 + 1 + 0 + 0,9 + 0,86) / 5
 = 0,55 x 100% = 55%

3. Menghitung Recall atau Sensitivity

Diperoleh data-data sebagai berikut:
 Data *Very Low* (True Positif) = 0 dan (False Negatif) = 0
 Data *Low* (True Positif) = 6 dan (False Negatif) = 2
 Data *Medium* (True Positif) = 0 dan (False Negatif) = 0
 Data *High* (True Positif) = 27 dan (False Negatif) = 1
 Data *Very High* (True Positif) = 6 dan (False Negatif) = 1
 Mencari Recall atau Sensitivity = True Positif / (True Positif + False Negatif)
 Recall = P(*Very Low*) + P(*Low*) + P(*Medium*) + P(*High*) + P(*Very High*) / Jumlah Klas
 = (0 + 0,75 + 0 + 0,96 + 0,86) / 5
 = 0,51 x 100% = 51%

4. Menghitung F-Score

$$F - Score = \frac{2(Presisi \times Recall)}{Presisi + Recall}$$

$$F - Score = \frac{2(0,55 \times 0,51)}{0,55 + 0,51}$$

F - Score = 0,53 x 100% = 53 %

3.2.3 Analisa Data Metode Algoritma Naïve Bayes Skenario 3

Analisis data metode algoritma *naïve bayes* skenario 3 menggunakan data latih 434 sampel dan data uji 145 sampel. Dari analisis akan diperoleh data berupa data prediksi berdasarkan label klas berupa kategori *very low*, *low*, *medium*, *high*, *very high*. Setelah itu dibuat data dengan sistem *confusion table* atau *confusion matrix* seperti diperlihatkan pada tabel 3.8. Setelah data *confusion matrix* diperoleh, maka dapat dihitung nilai-nilai akurasi, presisi, *recall* atau *sensitivity* dan *f-score*.

Tabel 3.8. Data Confusion Matrix Skenario 3

Actual ↓	Prediksi				
	<i>Very Low</i>	<i>Low</i>	<i>Medium</i>	<i>High</i>	<i>Very High</i>
<i>Very Low</i>	0	0	0	0	0
<i>Low</i>	0	7	0	2	0
<i>Medium</i>	0	0	0	0	0
<i>High</i>	0	0	0	40	1
<i>Very High</i>	0	0	0	1	13

Dari perhitungan data *confusion matrix* yang ditunjukkan pada tabel 3.8 diperoleh beberapa data yang bisa dijadikan parameter penilaian model produksi energi, diantaranya yaitu:

1. Menghitung Akurasi

Diperoleh data-data sebagai berikut:

Data *Very Low* (*True Positif*) = 0, Data *Low* (*True Positif*) = 7, Data *Medium* (*True Positif*) = 0, Data *High* (*True Positif*) = 40, Data *Very High* (*True Positif*) = 13

Total Data Set = 64

Mencari Akurasi = (*True Positif* / Total Data Set)

$$\begin{aligned} \text{Akurasi} &= (0 + 7 + 0 + 40 + 13) / 64 \\ &= 0,94 \times 100\% = 94\% \end{aligned}$$

2. Menghitung Presisi

Diperoleh data-data sebagai berikut:

Data *Very Low* (*True Positif*) = 0 dan (*False Positif*) = 0

Data *Low* (*True Positif*) = 7 dan (*False Positif*) = 0

Data *Medium* (*True Positif*) = 0 dan (*False Positif*) = 0

Data *High* (*True Positif*) = 40 dan (*False Positif*) = 3

Data *Very High* (*True Positif*) = 13 dan (*False Positif*) = 1

Mencari Presisi = *True Positif* / (*True Positif* + *False Positif*)

$$\begin{aligned} \text{Presisi} &= P(\text{Very Low}) + P(\text{Low}) + P(\text{Medium}) + P(\text{High}) + P(\text{Very High}) / \text{Jumlah Klas} \\ &= (0 + 1 + 0 + 0,93 + 0,93) / 5 \\ &= 0,57 \times 100\% = 57\% \end{aligned}$$

3. Menghitung Recall atau Sensitivity

Diperoleh data-data sebagai berikut:

Data *Very Low* (*True Positif*) = 0 dan (*False Negatif*) = 0

Data *Low* (*True Positif*) = 7 dan (*False Negatif*) = 2

Data *Medium* (*True Positif*) = 0 dan (*False Negatif*) = 0

Data *High* (*True Positif*) = 40 dan (*False Negatif*) = 1

Data *Very High* (*True Positif*) = 13 dan (*False Negatif*) = 1

Mencari Recall atau Sensitivity = *True Positif* / (*True Positif* + *False Negatif*)

$$\begin{aligned} \text{Recall} &= P(\text{Very Low}) + P(\text{Low}) + P(\text{Medium}) + P(\text{High}) + P(\text{Very High}) / \text{Jumlah Klas} \\ &= (0 + 0,78 + 0 + 0,96 + 0,93) / 5 \\ &= 0,54 \times 100\% = 54\% \end{aligned}$$

4. Menghitung F-Score

$$F - \text{Score} = \frac{2(\text{Presisi} \times \text{Recall})}{\text{Presisi} + \text{Recall}}$$

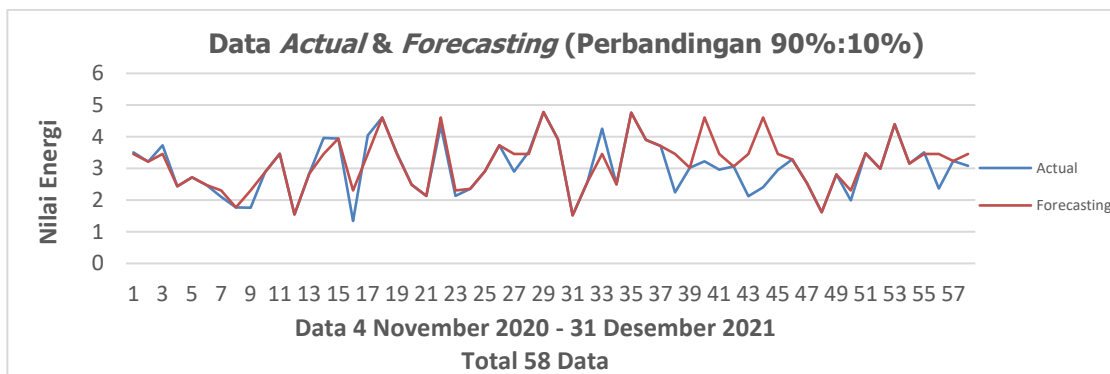
$$F - \text{Score} = \frac{2(0,57 \times 0,54)}{0,57 + 0,54}$$

$$F - \text{Score} = 0,55 \times 100\% = 55\%$$

3.3 Pemodelan Produksi Energi

3.3.1 Pemodelan Produksi Energi Skenario 1

Pemodelan produksi energi skenario 1 yaitu pemodelan berupa hasil grafik dengan perbandingan 90%:10% dengan data latih 521 sampel dan data uji 58 sampel. Hasil pemodelan skenario 1 diperlihatkan pada gambar 3.1.

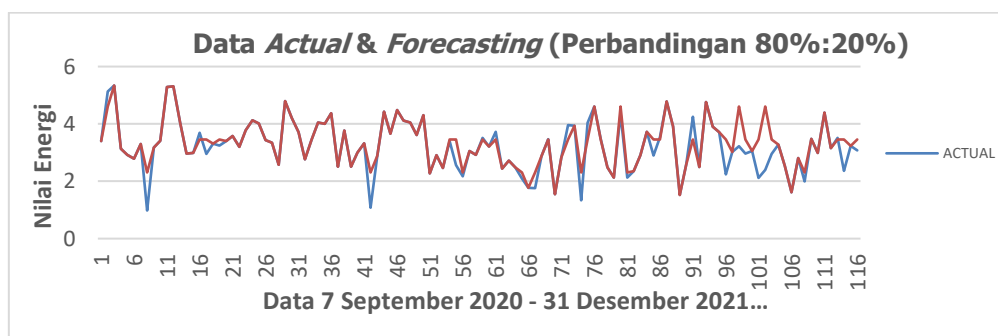


Gambar 3.1 Hasil Grafik Data Actual & Forecasting Skenario 1

Dari analisis sebelumnya, untuk pemodelan skenario 1 diperoleh nilai-nilai akurasi 83%, presisi 49%, *recall* atau *sensitivity* 52%, dan *f-score* 51%.

3.3.2 Pemodelan Produksi Energi Skenario 2

Pemodelan produksi energi skenario 2 yaitu pemodelan berupa hasil grafik dengan perbandingan 80%:20% dengan data latih 463 sampel dan data uji 116 sampel. Hasil pemodelan skenario 2 diperlihatkan pada gambar 3.2.

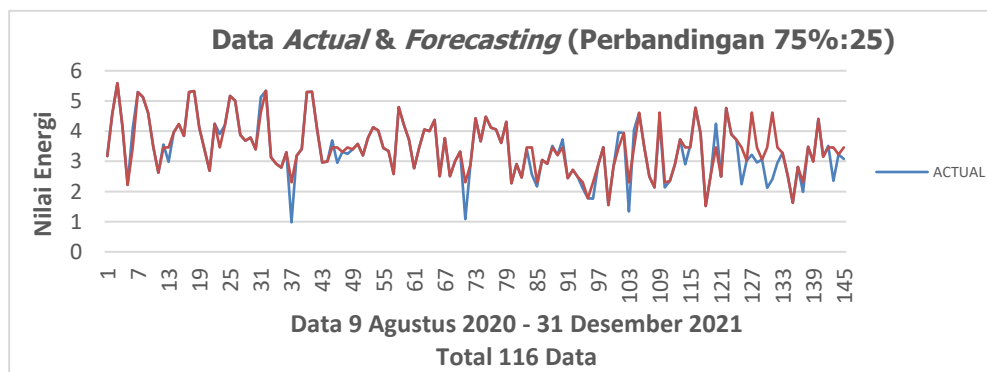


Gambar 3.2. Hasil Grafik Data Actual & Forecasting Skenario 2

Dari analisis sebelumnya, diperoleh nilai-nilai akurasi 91%, presisi 55%, *recall* atau *sensitivity* 51%, dan nilai *f-score* 53%.

3.3.3 Pemodelan Produksi Energi Skenario 3

Pemodelan produksi energi skenario 3 yaitu pemodelan berupa hasil grafik dengan perbandingan 75%:25% dengan data latih 434 sampel dan data uji 145 sampel. Hasil pemodelan skenario 3 diperlihatkan pada gambar 3.3.



Gambar 3.3 Hasil Grafik Data Actual & Forecasting Skenario 3

Dari analisis sebelumnya, diperoleh nilai-nilai akurasi 94%, presisi 57%, *recall* atau *sensitivity* 54%, dan nilai *f-score* 55%.

3.4 Evaluasi Pemodelan

Evaluasi pemodelan ini akan mencoba menilai kesalahan/*error* hasil pemodelan setiap skenario terhadap data produksi energi harian aktual. Nilai *error* seluruh skenario, yang dinyatakan dalam nilai MAE, RMSE dan MAPE ditunjukkan pada tabel 3.9. Berdasarkan tabel 3.9, dapat dilihat model skenario yang terbaik untuk memodelkan hasil produksi energi PLTS 1 kWp di ITENAS.

Tabel 3.9. Evaluasi Model

Pemodelan Skenario	Nilai MAE	Nilai RMSE	Nilai MAPE
1 (90% : 10%)	0,24	0,51	9,78
2 (80% : 20%)	0,17	0,41	7,75
3 (75% : 25%)	0,14	0,38	6,52

4. KESIMPULAN

Beberapa kesimpulan penting yang dapat diambil dari penelitian ini adalah:

Skenario 3 memiliki empat parameter penilaian model terbesar dibandingkan dengan skenario lainnya. Skenario 3 memiliki akurasi 94%, presisi 57%, *recall* atau *sensitivity* 54% dan *f-score* 55%. Sementara itu, skenario 1 memiliki akurasi 83%, presisi 49%, *recall* atau *sensitivity* 52% dan *f-score* 51%, dan skenario 2 memiliki akurasi 91%, presisi 55%, *recall* atau *sensitivity* 51% dan *f-score* 53%.

Skenario 3 memiliki tiga parameter evaluasi model terkecil dibandingkan dengan skenario lainnya. Skenario 3 memiliki MAE (*Mean Absolute Error*) yaitu 0,14, RMSE (*Root Mean Squared Error*) 0,38, dan MAPE (*Mean Absolute Percentage Error*) 6,52. Sementara itu, skenario 1 memiliki MAE (*Mean Absolute Error*) 0,24, RMSE (*Root Mean Squared Error*) 0,51, dan MAPE (*Mean Absolute Percentage Error*) 9,78, dan skenario 2 memiliki MAE (*Mean Absolute Error*) 0,17, RMSE (*Root Mean Squared Error*) 0,41, MAPE (*Mean Absolute Percentage Error*) 7,75.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh tim Tugas Akhir/Tesis dalam bidang fotovoltaik, atas berbagai diskusinya selama proses penelitian.

DAFTAR RUJUKAN

- Adibhatla, S., & Kaushik, S. C. (2017). Energy, exergy and economic (3E) analysis of integrated solar direct steam generation combined cycle power plant . *Sustainable Energy Technologies and Assessments* 20 , 88-97.
- Quek, Y. T., Woo, W. L., & Logenthiran, T. (2017). A Naïve Bayes Classification Approach for Short-Term Forecast of a Photovoltaic System. *Sustainable Energy and*

- Environmental Sciences* (hal. 1-6). Singapore: School of Engineering, Republic Polytechnic.
- Al Muluk, M. D. (2021). *Pemodelan Produksi Energi Pembangkit Listrik 1 kWp Menggunakan Sarima dan FTS*. Bandung: Tugas Akhir Program Sarjana Teknik Mesin, Institut Teknologi Nasional.
- Bhutka, J., Gajjar, J., & Harinarayana, T. (2016). Modelling of Solar Thermal Power Plant Using Parabolic Trough Collector. *Journal of Power and Energy Engineering*, 9-25.
- Boyle, G. (2004). *Renewable Energy: Power for Sustainable Future Second Edition*. New York: Oxford University Press Inc.
- Hidayat, F., Rusirawan, D., & Tanjung, I. R. (2019). Evaluasi Kinerja PLTS 1000 Wp di Itenas Bandung. *ELKOMIKA*, 7(1), 195 - 208.
- Ikhsan, M. (2020). *Peramalan Energi Photovoltaic Dengan Menggunakan Algoritma Naive Bayes Classifier Dan K-Nearest Neighbors*. Yogyakarta : Tugas Akhir Program Sarjana Teknik Elektro, Universitas Islam Indonesia.
- Jos, B. (2017). *Solar Energy Materi Kuliah Energi Terbarukan*. Semarang: Universitas Diponegoro.
- Kou, W. d. (2015). A direct Phase-coordinates Approach to Fault Ride Through of Unbalanced Faults in Large-scale Photovoltaic Power Systems. *Electric Power Components and Systems*, 902-913.
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction Time Series Analysis and Forecasting (2nd ed.)*. Hoboken New Jersey USA: John Wiley & Sons, Ltd.
- Ramazan, B., Yesilbudak, M., Colak, M., & Genc, N. (2017). A Novel Application of Naïve Bayes Classifier in Photovoltaic Energy Prediction. *16th IEEE International Conference on Machine Learning and Applications* (hal. 523-527). Turkey: ICMLA.
- Rudiyanto, B., Wardani, T. A., Anwar, S., Jamali, L. A., Wibowo, K. M., Pambudi, N. A., & Prasetyo, T. (2019). Energy and Exergy Analysis of Steam Power Plant in Paiton,. *Earth and Environmental Science* (hal. 1-11). IOP Conference Series.
- Rusirawan, D. (2012). *Energetic Modelling of Photovoltaic Modules in Grid- Connected Systems*. Godollo: Doctoral Dissertation, Szent Istvan University, Hungary.
- Rusirawan, D., & Farkas, I. (2014). Identification of Model Parameters of The Photovoltaic Solar Cells. *Energy Procedia* 57, 39-46.
- Sari, N. (2019). *Pemodelan Modul Fotovoltaik Model Dioda Tunggal dan Dioda Ganda dengan Visual Basic for Application (VBA) Microsoft Excel*. Bandung: Tugas Akhir Program Sarjana Teknik Mesin, Institut Teknologi Nasional.
- Suharyati., Pambudi, S. H., Wibowo, J. L., & Pratiwi, N. I. (2019). *Outlook Energi Indonesia (OEI) 2019*. Jakarta: Sekretariat Jenderal Dewan Energi Nasional.
- Zulkifli. (2011). Perencanaan Pemanfaatan Energi Surya sebagai Energi Alternatif. *Semantic Scholar*.