

## **Model Prediksi Penjadwalan Produksi Energi Terbarukan dengan Algoritma XGBoost dan Analisis Interpretatif Menggunakan SHAP**

**M. Safii<sup>1</sup>, Husain<sup>2</sup>, Ika Okta Kirana<sup>3</sup>, Sasha Aiko Leana<sup>4</sup>, Yuli Indahwati Gultom<sup>5</sup>**

<sup>1</sup>Informatika, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

<sup>3,5</sup>Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

<sup>4</sup>Teknik Informatika, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

<sup>2</sup>Sistem Informasi, Universitas Bumigora, Mataram, Indonesia

Email: <sup>1,\*</sup>m.safii@amiktunasbangsa.ac.id, <sup>2</sup>husain@universitasbumigora.ac.id, <sup>3</sup>ikaoktakirana@stikomtb.ac.id

<sup>4</sup>aikosasha556@gmail.com, <sup>5</sup>indahwatiyuli101@gmail.com

Email Penulis Korespondensi: m.safii@amiktunasbangsa.ac.id

### **Abstrak**

Penjadwalan produksi energi terbarukan adalah kegiatan untuk menyeimbangkan antara pasokan dan permintaan energi dalam siklus sistem energi berkelanjutan. Berbagai jenis energi terbarukan seperti hidro, angin, matahari, dan lainnya akan melalui pemodelan prediktif dari jadwal produksi menggunakan algoritma Extreme Gradient Boosting (XGBoost) yang dikombinasikan dengan pendekatan interpretabilitas model menggunakan SHapley Additive exPlanations (SHAP). Penelitian ini menggunakan data sekunder dengan parameter Tahun, Negara, Energi Surya, Energi Angin, Energi Hidro, Energi Terbarukan Lainnya, dan Total Energi Terbarukan. Pemodelan menunjukkan bahwa energi angin dan energi matahari memiliki prediksi produksi yang meningkat ketika nilai fitur tinggi dan energi angin memiliki efek negatif ketika nilai fitur rendah. Penelitian ini memiliki kontribusi yang signifikan terhadap faktor yang mempengaruhi penjadwalan dan juga berpeluang untuk penerapan sistem cerdas dalam pengambilan keputusan sektor energi. Hasil penelitian ini dapat menjadi dasar untuk merumuskan strategi manajemen energi berkelanjutan yang memiliki potensi untuk mengintegrasikan kecerdasan buatan dan transparansi model dalam kebijakan energi terbarukan.

**Kata Kunci:** Energi terbarukan, Extreme Gradient Boosting, Penjadwalan, Prediksi, SHapley Additive exPlanations

### **Abstract**

Renewable energy production scheduling is the activity of balancing energy supply and demand within the cycle of a sustainable energy system. Various types of renewable energy such as hydro, wind, solar, and others will go through predictive modeling of the production schedule using the Extreme Gradient Boosting (XGBoost) algorithm combined with a model interpretability approach using SHapley Additive exPlanations (SHAP). This research uses secondary data with the parameters of Year, Country, Solar Energy, Wind Energy, Hydro Energy, Other Renewable Energy, and Total Renewable Energy. The modeling shows that wind energy and solar energy have an increasing production prediction when the feature values are high, and wind energy has a negative effect when the feature values are low. This research has a significant contribution to the factors influencing scheduling and also has the potential for the implementation of intelligent systems in decision-making in the energy sector. The results of this study can serve as a foundation for formulating sustainable energy management strategies that have the potential to integrate artificial intelligence and model transparency in renewable energy policies.

**Keywords:** Renewable energy, Extreme Gradient Boosting, Scheduling, Prediction, SHapley Additive exPlanations (at least 5 words related to the research content separated by commas)

## **1. PENDAHULUAN**

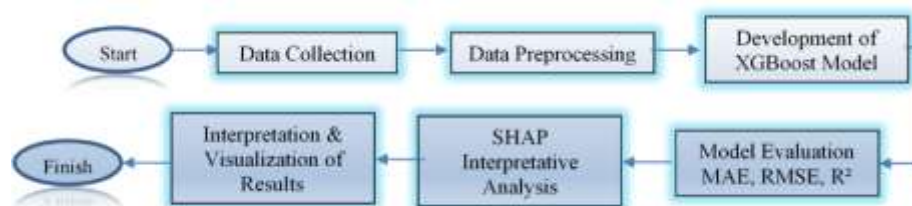
Pemanfaatan energi fosil telah menjadi ketergantungan bagi pengguna energi di seluruh dunia, yang mengakibatkan krisis energi global. Seiring perubahan iklim global mendorong negara-negara untuk memanfaatkan energi alternatif non-fosil seperti energi matahari, energi angin, dan lainnya. Secara historis, pembangunan ekonomi sangat berkorelasi dengan peningkatan konsumsi energi dan pertumbuhan emisi gas rumah kaca. Energi terbarukan dapat membantu memisahkan korelasi yang berkontribusi pada pembangunan berkelanjutan [1], [2], [3]. Perubahan besar dalam cara energi diproduksi adalah langkah yang dapat membantu mengurangi peningkatan suhu global akibat emisi gas rumah kaca. Pemerintah Indonesia bertujuan untuk mencapai target Emisi Nol Bersih atau netralitas karbon pada tahun 2060 sejalan dengan pengembangan sektor energi [4]. Tantangan dalam energi terbarukan adalah fluktuasi energi yang tergantung pada kondisi alam yang tidak dapat diprediksi. Hal ini menyebabkan kesulitan dalam proses penjadwalan produksi, yang mengarah pada ketidakseimbangan dalam efisiensi operasional. Berdasarkan kondisi di atas, dibutuhkan sistem prediksi dan penjadwalan untuk produksi energi terbarukan yang dapat dioptimalkan. Untuk mengatasi masalah ini, pendekatan pembelajaran mesin menggunakan algoritma XGBoost dapat memodelkan sistem prediksi dan penjadwalan produksi energi terbarukan [5], [6]. Algoritma XGBoost dengan menerapkan teknik oversampling dan penyesuaian parameter adalah metode klasifikasi yang cukup efektif untuk diterapkan dalam kasus klasifikasi [7], [8]. Keuntungan lain dari algoritma XGBoost adalah kemampuannya untuk memodelkan hubungan nonlinier, menangani data yang hilang, dan mengelola jumlah fitur yang besar dengan efisien [9], [10], [11]. Dalam penerapan model pembelajaran

mesin untuk menyelesaikan masalah, banyak studi yang tidak menunjukkan interpretabilitas, sehingga menyulitkan untuk memberikan pemahaman yang jelas kepada pengambil keputusan mengenai masalah tersebut. Dalam pemecahan masalah, tidak hanya prediksi yang akurat yang dibutuhkan, tetapi juga harus memberikan pemahaman tentang parameter yang memiliki pengaruh signifikan terhadap keputusan. Dengan demikian, pendekatan interpretatif dari SHapley Additive exPlanations (SHAP) dapat memvisualisasikan parameter yang berpengaruh signifikan terhadap hasil pemodelan[12], [13]. Integrasi XGBoost dan SHAP memiliki kemampuan prediktif yang baik dan transparansi seperti yang ditunjukkan dalam penelitian[14] XGBoost dapat menggeneralisasi dan mencapai akurasi (90,6%) dengan risiko overfitting yang lebih rendah, dan teknik XAI memfasilitasi interpretasi klinis, memberikan transparansi, dan membangun kepercayaan dalam keputusan model. Penelitian lain yang dilakukan oleh [15] kemampuan algoritma XGBoost untuk memperkirakan kategori kecelakaan maritim dengan nilai yang sesuai dan akurat, serta tingkat pengambilan kembali sebesar 97,14%. Faktor tim adalah elemen penting yang mempengaruhi risiko keselamatan dalam navigasi maritim, sementara menjaga peralatan dan fasilitas dalam kondisi optimal serta meningkatkan manajemen perusahaan pelayaran memiliki dampak positif terhadap keselamatan maritim. Penelitian lainnya oleh [16] menerapkan algoritma XGBoost sebagai model regresi dalam proses pelatihan dan prediksi, memperoleh MAPE sebesar 3,402% dan nilai R<sup>2</sup> yang hampir mencapai 1, menunjukkan bahwa model ini sangat sesuai. Ini berdasarkan pengembangan model XGBoost untuk menyelidiki secara mendalam faktor-faktor yang mempengaruhi dan menentukan pentingnya fitur-fitur tersebut. Penelitian sebelumnya yang membahas peramalan energi terbarukan menggunakan machine learning dan deep learning oleh [17] membandingkan model machine learning regresi linier klasik dengan model random forest, SVM, dan XGBoost menunjukkan bahwa model machine learning regresi linier klasik tidak dapat menangkap hubungan non-linear dan pola kompleks dalam data, sedangkan model random forest, SVM, dan XGBoost memiliki hasil yang lebih optimal dalam memprediksi energi terbarukan yang dapat menangani hubungan non-linear dan data yang kompleks.

Dalam penelitian sebelumnya belum dapat memvisualisasikan parameter yang mempengaruhi hasil prediksi. Dalam penelitian ini, model dapat memberikan prediksi yang optimal serta menjelaskan alasan dan cara fitur-fitur seperti jenis energi, lokasi, dan waktu dapat mempengaruhi hasil prediksi. Visualisasi SHAP menunjukkan bahwa fitur HydroEnergy, WindEnergy, SolarEnergy, dan OtherRenewableEnergy memiliki pengaruh terbesar pada hasil model. Diharapkan penelitian ini dapat membantu dalam pengambilan keputusan yang berbasis data dengan cara yang lebih jelas dan dapat beradaptasi dengan kompleksitas sistem energi saat ini.

## 2. METODOLOGI PENELITIAN

Penelitian ini menggunakan metode kuantitatif-prediktif yang berbasis pada pembelajaran mesin untuk mengembangkan model prediktif yang akurat dan mudah dipahami untuk jadwal produksi energi terbarukan. Metodologi dirancang secara sistematis mulai dari pengumpulan data, pra-pemrosesan, pembangunan model menggunakan algoritma XGBoost, hingga analisis interpretatif dengan SHAP (SHapley Additive exPlanations). Diagram alir dari keseluruhan metodologi digambarkan sebagai berikut:



Gambar 1. Diagram Alir Penelitian

Gambar di atas merupakan alur kegiatan penelitian yang dimulai dari proses pengumpulan data, pra-pemrosesan data, membangun model dengan XGBOOST, mengevaluasi model, menginterpretasi dengan SHAP dan visualisasi. Berikut penjelasan alur penelitian tersebut:

### 2.1 Desain dan Pendekatan Penelitian

Penelitian ini adalah studi eksperimental yang fokus pada supervised learning regression dengan hasil model yang memprediksi total jumlah produksi energi terbarukan yang diukur dalam GWh. Prediksi ini didasarkan pada beberapa fitur input, termasuk jenis energi (seperti hidro, surya, angin, dan lainnya), dimensi spasial (dalam bentuk negara), dan dimensi temporal (tahun). Algoritma yang digunakan adalah Extreme Gradient Boosting (XGBoost), sementara SHAP digunakan untuk interpretasi model terkait setiap fitur.

### 2.2 Data Penelitian

Data penelitian berasal dari data sekunder yang terdiri dari 240 record data dengan 7 variabel yaitu Year, Country, Solar energy, Wind Energy, Hydro Energy, Other Renewable energy, and Total Renewable energy. Seperti yang ditunjukkan pada tabel 1 berikut:

Tabel 1. Data Penelitian

| No         | Year  | Country | Solar Energy               | Wind Energy                | Hydro Energy               | Other Renewable Energy     | Total Renewable Energy     |
|------------|-------|---------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| 1          | 2000  | USA     | 4.370.861.0<br>69.626.260  | 1.435.928.59<br>8.332.890  | 15.443.897.0<br>10.793.800 | 3.193.963.17<br>8.886.660  | 37.368.007.2<br>42.635.700 |
| 2          | 2001  | USA     | 24.041.677.<br>639.819.200 | 40.279.287.6<br>43.706.300 | 3.987.421.40<br>6.859.390  | 4.397.792.65<br>5.987.200  | 14.817.310.5<br>91.199.100 |
| 3          | 2002  | USA     | 6.410.035.1<br>05.688.870  | 11.204.943.5<br>11.348.500 | 33.499.364.0<br>30.286.400 | 48.645.943.3<br>47.289.700 | 2.582.950.93<br>5.479.500  |
| 4          | 2003  | USA     | 8.491.983.7<br>67.203.790  | 476.040.843.<br>881.759    | 6.091.024.44<br>2.520.710  | 13.253.202.9<br>43.404.500 | 2.066.873.69<br>4.288.250  |
| 5          | 2004  | USA     | 3.738.180.1<br>86.635.830  | 8.821.833.61<br>1.219.090  | 10.343.065.3<br>16.915.900 | 18.105.311.3<br>08.911.800 | 2.471.361.02<br>4.566.200  |
| .....<br>. | ..... | .....   | .....                      | .....                      | .....                      | .....                      | .....                      |
| 24<br>0    | 2023  | Japan   | 5.670.259.3<br>47.545.530  | 2.610.557.56<br>8.075.710  | 5.826.817.26<br>8.853.340  | 3.821.151.27<br>3.918.660  | 1.792.878.54<br>5.839.320  |

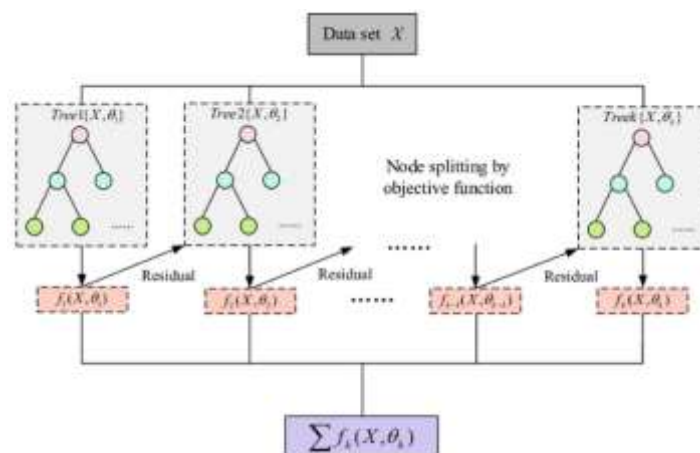
Berdasarkan data di atas merupakan kumpulan informasi yang terdiri dari data tahunan tentang produksi energi terbarukan dari berbagai negara dari tahun 2000 hingga 2023. Kumpulan informasi ini mencakup rincian tentang produsen energi dari matahari, angin, air, dan sumber terbarukan lainnya. Kumpulan informasi ini bertujuan untuk memberikan pemahaman tentang perkembangan dan distribusi produksi energi terbarukan di seluruh dunia.

### 2.3 Pra-pemrosesan Data

Informasi yang diperoleh sebelumnya melalui tahap persiapan adalah proses pengolahan data, yang melibatkan penghapusan entri yang memiliki nilai yang hilang dan outlier yang tidak sesuai, dilanjutkan dengan pengolahan fitur kategori yang melibatkan negara-negara yang telah dilabeli menggunakan teknik One-Hot Encoding agar dapat diproses oleh model XGBoost.

### 2.4 Pengembangan Model XGBoost

Model prediksi dikembangkan menggunakan XGBoost, sebuah metode peningkatan gradien ensemble berbasis pohon keputusan yang telah terbukti efektif dalam menangani masalah prediksi[18]. Diagram berikut menunjukkan langkah-langkah terstruktur dalam pembentukan model XGBoost, dimulai dari pengumpulan data hingga evaluasi dan pemahaman hasil model. Diagram alir model dapat dilihat pada gambar berikut:



Gambar 2. Diagram Alir XGBOOST

Dapat dijelaskan bahwa gambar di atas adalah arsitektur dasar dari algoritma Gradient Boosting Decision Tree yang umum digunakan dalam model XGBoost. Ini diimplementasikan sebagai[19]:

$$l = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k)$$

(1)

$$\Omega(fk) = \gamma T_k + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (2)$$

Dimana  $l(y_i, \hat{y}_i)$ ,  $y_i$  dan  $\hat{y}_i$  adalah total kerugian pelatihan, target pelatihan, dan nilai prediksi, masing-masing. Selain itu,  $\Omega$ ,  $\lambda$  dan  $\gamma$  adalah istilah untuk regularisasi, koefisien, dan parameter masing-masing. Proses pelatihan model dilakukan dengan menggunakan teknik pencarian yang efisien untuk mengevaluasi potensi titik pembagian, memilih opsi terbaik untuk setiap tingkat pohon. Proses pelatihan model dilakukan pada tahap-tahap berikut:

- Pembagian data: data dibagi menjadi set pelatihan (80%) dan set pengujian (20%)
- Pemilihan hyperparameter: dilakukan menggunakan teknik Grid Search dan 5-fold Cross-Validation pada parameter.
- Pelatihan model: menggunakan pustaka xgboost di lingkungan Python.
- Evaluasi kinerja: menggunakan metrik Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan R-squared ( $R^2$  Score).

## 2.5 Analisis Interpretatif Menggunakan SHAP

Setelah proses pelatihan model XGBoost selesai, analisis kontribusi dari fitur-fitur dilakukan menggunakan metode SHAP. SHAP menunjukkan seberapa banyak setiap fitur memberikan kontribusi secara individu terhadap hasil yang diperoleh oleh model, mengacu pada teori permainan Shapley dalam permainan kooperatif. Analisis ini dilakukan pada dua level: Interpretasi Global untuk mengidentifikasi fitur mana yang memiliki dampak paling besar pada prediksi keseluruhan model, dan Interpretasi Lokal untuk menjelaskan prediksi spesifik untuk setiap data masukan. Kerangka kerja untuk pendekatan interpretabilitas ini dapat diuraikan sebagai berikut[20]:

$$g(\mathbf{z}) = \varphi_0 + \sum_{i=1}^M \varphi_i z_i \quad (3)$$

Dimana  $g$  dan  $\mathbf{z}$  adalah fungsi penjelas dan kumpulan parameter prediktor biner yang disederhanakan.  $M$  dan  $\varphi_i$  menunjukkan jumlah fitur dan koefisien dari masing-masing model penjelas. Proses perhitungan untuk ukuran pemahaman ini menggunakan pendekatan permutasi berdasarkan prinsip-prinsip teori permainan kolaboratif. Ini mengukur pengaruh variabel input dengan menganalisis semua kombinasi fitur yang mungkin. Algoritma ini menilai signifikansi setiap fitur dengan membandingkan hasil prediksi ketika fitur tersebut ada dan ketika tidak ada, melalui semua variasi fitur yang mungkin.

## 3. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penerapan dan evaluasi model ramalan untuk penjadwalan produksi energi terbarukan menggunakan algoritma XGBoost, serta analisis mendalam mengenai model tersebut menggunakan metode SHAP (SHapley Additive exPlanations). Diskusi mencakup penjelasan tentang dataset, langkah-langkah dalam pra-proses data, hasil dari pelatihan dan pengujian model, penjelasan mengenai fitur yang mempengaruhi hasil prediksi, serta diskusi tentang temuan penting yang diperoleh. Analisis dilakukan untuk mengevaluasi seberapa efektif model dalam memberikan prediksi yang akurat dan memberikan pemahaman yang lebih dalam tentang faktor-faktor yang mempengaruhi produksi energi terbarukan.

### 3.1 Hasil Pelatihan dan Evaluasi Model XGBoost

Pada tahap ini, model prediksi penjadwalan produksi energi terbarukan dibangun menggunakan algoritma XGBoost Regressor. Implementasi kode menggunakan python dapat dilihat pada gambar berikut:

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Inisialisasi model
model = xgb.XGBRegressor(objective='reg:squarederror', n_estimators=100, max_depth=5, learning_rate=0.1)

# Training
model.fit(X_train, y_train)

# Prediksi
y_pred = model.predict(X_test)

# Evaluasi
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

print(f"Mean Squared Error: {mse}")
print(f"R2 Score: {r2}")
```

Gambar 3. Pelatihan dan Evaluasi Model XGBoost

Dataset dibagi menjadi dua bagian, 80% untuk data pelatihan dan 20% untuk data pengujian menggunakan fungsi `train_test_split`, dan mengatur parameter `random_state=42` untuk memastikan hasil yang konsisten. Pembagian data ini bertujuan agar model dapat dilatih dan diuji secara objektif tanpa kebocoran informasi. Model XGBoost diatur dengan parameter `objective='reg:squarederror'`, `n_estimators=100`, `max_depth=5`, dan `learning_rate=0.1`. Pemilihan parameter ini didasarkan pada eksperimen awal untuk mencapai keseimbangan antara tingkat kompleksitas model dan akurasi hasil prediksi. Selanjutnya, model dilatih dengan data pelatihan ( $X_{\text{train}}$ ,  $y_{\text{train}}$ ) dan menghasilkan model prediksi yang akan diuji pada data yang sebelumnya tidak diketahui ( $X_{\text{test}}$ ). Hasil prediksi model kemudian dinilai menggunakan dua metrik utama, yaitu Mean Squared Error (MSE) dan skor R-squared ( $R^2$ ). Hasil Mean Squared Error (MSE) dan skor R-squared ( $R^2$ ) dapat dilihat pada gambar berikut:

Mean Squared Error: 23738.02404501845  
R2 Score: 0.9439083751481915

Gambar 4. Hasil Mean Squared Error (MSE) dan skor R-squared ( $R^2$ )

Nilai MSE yang tercatat sebesar 23.738,02 menggambarkan rata-rata kuadrat kesalahan dalam memprediksi model dibandingkan dengan nilai aktual. Di sisi lain, nilai  $R^2$  yang mencapai 0,9439 menunjukkan bahwa model ini dapat menjelaskan hampir 94,39% variasi dari variabel yang ditargetkan. Secara keseluruhan,  $R^2$  yang mendekati 1 menunjukkan bahwa model memiliki kemampuan prediksi yang sangat baik, sementara nilai MSE yang cukup rendah menunjukkan bahwa kesalahan dalam prediksi masih dalam batas yang wajar. Temuan ini menunjukkan bahwa algoritma XGBoost dapat memahami pola dalam data dengan baik dan memberikan hasil prediksi yang akurat mengenai penjadwalan produksi energi terbarukan

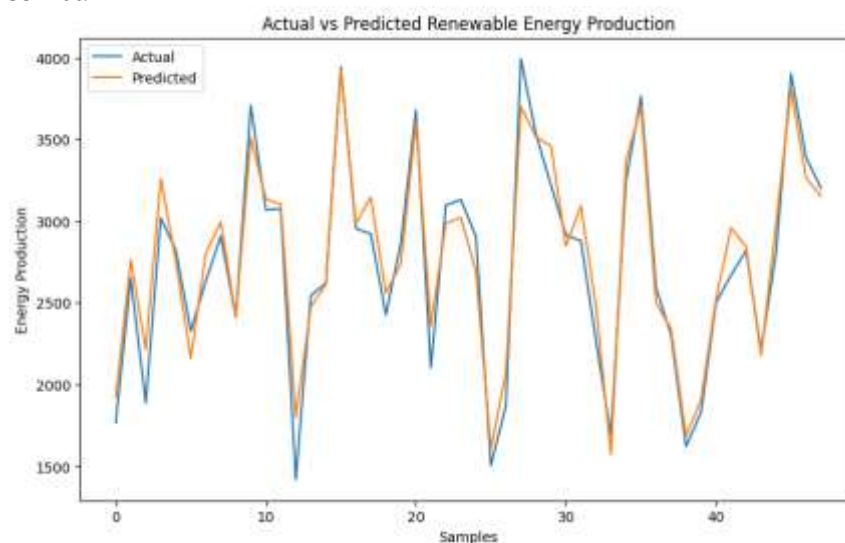
### 3.2 Visualisasi Perbandingan Antara Nilai Aktual dan Nilai Prediksi

Untuk menilai kemampuan model dalam mencocokkan pola data aktual, dibuat perbandingan visual antara nilai aktual ( $y_{\text{test}}$ ) dan nilai prediksi ( $y_{\text{pred}}$ ) melalui grafik garis seperti yang ditunjukkan dengan kode pada gambar berikut:

```
plt.figure(figsize=(10, 6))
plt.plot(y_test.values, label='Actual')
plt.plot(y_pred, label='Predicted')
plt.legend()
plt.title("Actual vs Predicted Renewable Energy Production")
plt.xlabel("Samples")
plt.ylabel("Energy Production")
plt.show()
```

Gambar 5. Skrip Perbandingan Nilai Aktual dan Prediksi

Script diatas merupakan script python untuk membuat visualisasi data dalam bentuk plot garis yang biasanya digunakan dalam prediksi machine learning. Model akan dianggap tepat jika garis prediksi mendekati garis aktual seperti yang dapat dilihat pada gambar berikut:



Gambar 6. Visualisasi Perbandingan antara Nilai Aktual dan Nilai Prediksi

Gambar yang dihasilkan menunjukkan dua garis, masing-masing mewakili nilai aktual dan nilai yang diprediksi dari produksi energi terbarukan. Sumbu horizontal menunjukkan urutan sampel dalam data uji, sedangkan sumbu vertikal menunjukkan besaran nilai produksi energi. Garis yang mewakili nilai aktual diberi label "Aktual", sementara garis yang diprediksi oleh model diberi label "Diprediksi". Hasil visualisasi menunjukkan bahwa kurva yang dihasilkan oleh model XGBoost sangat mirip dengan nilai aktual, menunjukkan bahwa model ini berhasil mencerminkan perubahan dalam produksi energi dengan cukup akurat. Ini juga mendukung hasil evaluasi kuantitatif yang dilakukan sebelumnya, di mana model mencapai nilai  $R^2$  sebesar 0,9439, yang menunjukkan tingkat akurasi yang tinggi dalam menjelaskan variasi data. Dengan demikian, grafik ini menunjukkan bahwa model ini tidak hanya memiliki kemampuan prediktif yang baik secara numerik tetapi juga secara visual menunjukkan kesamaan pola antara hasil yang diprediksi dan kenyataan, yang sangat penting untuk pengambilan keputusan berbasis data dalam sistem penjadwalan produksi energi terbarukan.

### 3.3. Interpretasi Model SHAP

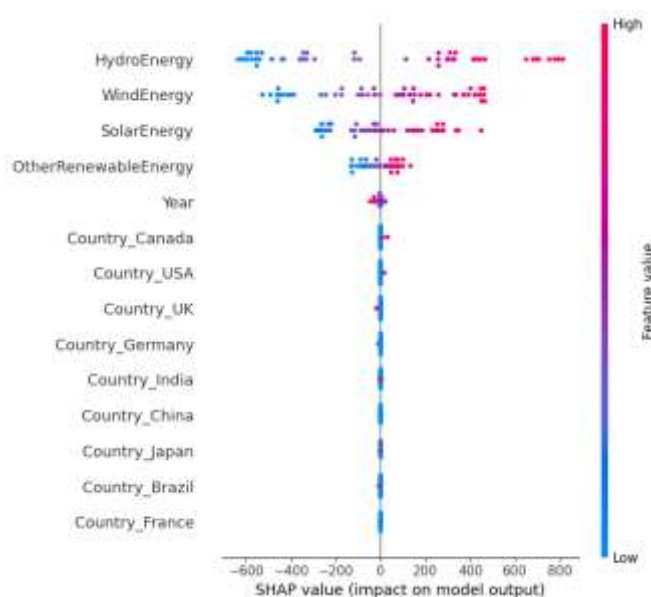
Untuk meningkatkan pemahaman mengenai model prediksi yang telah dibuat, penelitian ini menerapkan pendekatan SHAP (SHapley Additive exPlanations) untuk menjelaskan peran masing-masing fitur dalam hasil prediksi. SHAP adalah metode interpretasi model yang berbasis pada teori permainan kooperatif, yang secara matematis memastikan distribusi yang adil dari kontribusi semua fitur terhadap hasil akhir[21]. Dalam penerapannya, objek TreeExplainer dari pustaka shap digunakan untuk menghitung nilai SHAP dari model yang telah dibuat. Objek ini dirancang khusus untuk model yang menggunakan pohon keputusan. Skrip berikut digunakan dalam proses interpretasi:

```
explainer = shap.TreeExplainer(model)
shap_values = explainer.shap_values(X_test.values)

shap.summary_plot(shap_values, X_test, feature_names=X_test.columns)
```

Gambar 7. Interpretabilitas model SHAP.

Baris pertama menjelaskan penjelas menggunakan model yang telah dilatih. Selanjutnya, nilai shap dihitung menggunakan data uji  $X_{test}$ , yang menghasilkan matriks nilai SHAP yang menunjukkan seberapa banyak setiap fitur berkontribusi terhadap prediksi untuk setiap instance. Akhirnya, `shap.summary_plot()` digunakan untuk membuat visualisasi plot ringkasan, yaitu plot beeswarm yang menggambarkan distribusi nilai SHAP dari semua fitur. Warna pada grafik menunjukkan nilai sebenarnya dari fitur (biru untuk nilai rendah, merah untuk nilai tinggi), sementara sumbu horizontal mewakili besarnya kontribusi (baik positif maupun negatif) terhadap prediksi. Hasil script diatas dapat dilihat pada gambar berikut:



Gambar 8. Visualisasi Interpretasi Model Menggunakan SHAP.

Dari analisis grafik ringkasan SHAP, dapat dilihat bahwa fitur yang memiliki pengaruh paling besar terhadap hasil prediksi adalah HydroEnergy, diikuti oleh WindEnergy, SolarEnergy, dan OtherRenewableEnergy. Keempat fitur ini mencerminkan total kapasitas produksi energi terbarukan berdasarkan jenis dan menunjukkan pengaruh yang kuat

terhadap hasil yang diprediksi oleh model. Distribusi titik pada sumbu X mencerminkan nilai SHAP untuk setiap fitur, di mana nilai positif menunjukkan kontribusi dalam meningkatkan total prediksi energi, sedangkan nilai negatif menunjukkan kontribusi dalam mengurangi prediksi. Warna titik menunjukkan nilai aktual dari fitur, di mana merah menunjukkan nilai tinggi dan biru menunjukkan nilai rendah.

#### 4. KESIMPULAN

Hasil dari studi ini menunjukkan bahwa algoritma XGBoost dapat secara efektif memahami pola dalam data dan memberikan hasil prediksi yang akurat dengan nilai MSE tercatat sebesar 23.738,02, yang merepresentasikan kesalahan kuadrat rata-rata dalam memprediksi model dibandingkan dengan nilai aktual. Nilai  $R^2$  sebesar 0,9439 menunjukkan bahwa model ini dapat menjelaskan hampir 94,39% variasi dalam variabel target. Dari analisis grafik ringkasan SHAP, fitur-fitur yang memiliki pengaruh terbesar terhadap hasil prediksi adalah HydroEnergy, WindEnergy, SolarEnergy, dan OtherRenewableEnergy. Keempat fitur ini menunjukkan pengaruh yang kuat terhadap hasil yang diprediksi oleh model. Kontribusi dari penelitian ini memiliki potensi untuk diterapkan secara nasional dan regional sebagai bagian dari sistem pendukung keputusan di sektor energi dengan menggabungkan model XGBoost dan SHAP untuk menghasilkan prediksi yang akurat dan mudah dipahami. Implikasi dari pemanfaatan energi terbarukan di masa depan dapat membantu memisahkan korelasi yang berkontribusi pada pembangunan berkelanjutan. Energi terbarukan dipengaruhi oleh kondisi alam yang tidak dapat diprediksi, dan penelitian ini dilakukan berdasarkan kondisi alam saat ini. Karena data yang fluktuatif, hasil penelitian ini mungkin berdampak pada permintaan energi terbarukan di masa depan. Tindak lanjut dari hasil penelitian ini, dengan perkembangan teknologi informasi dan ketidakpastian kondisi alam, harus melibatkan pengujian ulang dengan algoritma lain untuk mencapai hasil yang lebih optimal terkait sumber energi terbarukan yang tersedia.

#### UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya penelitian ini, khususnya **Kementerian Pendidikan Tinggi, Sains, dan Teknologi. Direktorat Jenderal Riset dan Pengembangan**, Yayasan STIKOM Tunas Bangsa yang telah memberikan bantuan sarana dan prasarana dan tim peneliti yang bersama-sama memberikan ide.

#### DAFTAR PUSTAKA

- [1] P. A. Østergaard, N. Duic, Y. Noorollahi, and S. Kalogirou, "Renewable energy for sustainable development," *Renew. Energy*, vol. 199, pp. 1145–1152, 2022, doi: <https://doi.org/10.1016/j.renene.2022.09.065>.
- [2] V. Agbakwuru, P. Obidi, O. Salihu, and C. Ogwu, "The role of renewable energy in achieving sustainable development goals," *Int. J. Eng. Res. Updat.*, vol. 7, pp. 13–27, Nov. 2024, doi: [10.53430/ijeru.2024.7.2.0046](https://doi.org/10.53430/ijeru.2024.7.2.0046).
- [3] T. Güney, "Renewable Energy Consumption and Sustainable Development: A Panel Cointegration Approach," *J. Knowl. Econ.*, vol. 15, no. 1, pp. 1286–1301, 2024, doi: [10.1007/s13132-023-01107-0](https://doi.org/10.1007/s13132-023-01107-0).
- [4] H. Prasodjo, "Green diplomacy as an effort by the Indonesian government in Realizing Net Zero Emission (NZE) in the year 2060," 2023, pp. 184–190. doi: [10.1201/9781003360483-21](https://doi.org/10.1201/9781003360483-21).
- [5] S. Hakkal and A. A. Lahcen, "XGBoost To Enhance Learner Performance Prediction," *Comput. Educ. Artif. Intell.*, vol. 7, p. 100254, 2024, doi: <https://doi.org/10.1016/j.caeai.2024.100254>.
- [6] Q. Wu *et al.*, "Application of shale TOC prediction model using the XGBoost machine learning algorithm: a case study of the Qiongzhusi Formation in central Sichuan Basin," *Carbonates and Evaporites*, vol. 40, no. 1, p. 8, 2024, doi: [10.1007/s13146-024-01042-4](https://doi.org/10.1007/s13146-024-01042-4).
- [7] Z. Ali, Z. Abduljabbar, H. Tahir, A. Sallow, and S. Almufti, "Exploring the Power of eXtreme Gradient Boosting Algorithm in Machine Learning: a Review," *Acad. J. Nawroz Univ.*, vol. 12, pp. 320–334, May 2023, doi: [10.25007/ajnu.v12n2a1612](https://doi.org/10.25007/ajnu.v12n2a1612).
- [8] X. Y. Liew, N. Hameed, and J. Clos, "An investigation of XGBoost-based algorithm for breast cancer classification," *Mach. Learn. with Appl.*, vol. 6, p. 100154, 2021, doi: <https://doi.org/10.1016/j.mlwa.2021.100154>.
- [9] N. B. Shaik, K. Jongkittinarukorn, and K. Bingi, "XGBoost based enhanced predictive model for handling missing input parameters: A case study on gas turbine," *Case Stud. Chem. Environ. Eng.*, vol. 10, p. 100775, 2024, doi: <https://doi.org/10.1016/j.csee.2024.100775>.
- [10] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, "A comparative analysis of gradient boosting algorithms," *Artif. Intell. Rev.*, vol. 54, no. 3, pp. 1937–1967, 2021, doi: [10.1007/s10462-020-09896-5](https://doi.org/10.1007/s10462-020-09896-5).
- [11] S. Fatima, A. Hussain, S. Amir, S. H. Ahmed, and S. Aslam, "XGBoost and Random Forest Algorithms: An in Depth Analysis," *Pakistan J. Sci. Res.*, vol. 3, pp. 26–31, Oct. 2023, doi: [10.57041/pjosr.v3i1.946](https://doi.org/10.57041/pjosr.v3i1.946).
- [12] D. Mane, "Unlocking Machine Learning Model Decisions: A Comparative Analysis of LIME and SHAP for Enhanced Interpretability," *J. Electr. Syst.*, vol. 20, pp. 598–613, Apr. 2024, doi: [10.52783/jes.1480](https://doi.org/10.52783/jes.1480).
- [13] Y. Nohara, K. Matsumoto, H. Soejima, and N. Nakashima, "Explanation of machine learning models using shapley additive explanation and application for real data in hospital," *Comput. Methods Programs Biomed.*, vol. 214, p. 106584, 2022, doi: <https://doi.org/10.1016/j.cmpb.2021.106584>.
- [14] S. Manoharan, "A Study On Evaluating The Impact Of Hybrid Work Model On Employee Performance: Adaptability, Sustainability And Organizational Success In The It Sector In Bangalore." 2024.
- [15] C. Zhang, X. Zou, and C. Lin, "Fusing XGBoost and SHAP Models for Maritime Accident Prediction and Causality Interpretability Analysis," *Journal of Marine Science and Engineering*, vol. 10, no. 8. 2022. doi: [10.3390/jmse10081154](https://doi.org/10.3390/jmse10081154).

- [16] D. Ying, P. Hua, and M. Hao, "Research and Application of SMOTE-Based Method with XGBoost Regression Prediction," in *2023 IEEE International Conference on Image Processing and Computer Applications (ICIPCA)*, 2023, pp. 1737–1740. doi: 10.1109/ICIPCA59209.2023.10257809.
- [17] N. E. Benti, M. D. Chaka, and A. G. Semie, "Forecasting Renewable Energy Generation with Machine Learning and Deep Learning: Current Advances and Future Prospects," *Sustainability*, vol. 15, no. 9. 2023. doi: 10.3390/su15097087.
- [18] D. Kumar, S. K. Sood, and K. S. Rawat, "Early health prediction framework using XGBoost ensemble algorithm in intelligent environment," *Artif. Intell. Rev.*, vol. 56, no. 1, pp. 1591–1615, 2023, doi: 10.1007/s10462-023-10565-6.
- [19] M. Alizamir *et al.*, "An interpretable XGBoost-SHAP machine learning model for reliable prediction of mechanical properties in waste foundry sand-based eco-friendly concrete," *Results Eng.*, vol. 25, p. 104307, 2025, doi: <https://doi.org/10.1016/j.rineng.2025.104307>.
- [20] K. K. P. M. Kannangara, W. Zhou, Z. Ding, and Z. Hong, "Investigation of feature contribution to shield tunneling-induced settlement using Shapley additive explanations method," *J. Rock Mech. Geotech. Eng.*, vol. 14, no. 4, pp. 1052–1063, 2022, doi: <https://doi.org/10.1016/j.jrmge.2022.01.002>.
- [21] M. Li, H. Sun, Y. Huang, and H. Chen, "Shapley value: from cooperative game to explainable artificial intelligence," *Auton. Intell. Syst.*, vol. 4, no. 1, p. 2, 2024, doi: 10.1007/s43684-023-00060-8.