

Analiza podataka i primjena algoritama mašinskog učenja na podacima o proizvodnji solarne energije

Opis problema

Projekat se fokusira na analizu podataka o proizvodnji solarne energije na području Čačka u Srbiji, sa ciljem identifikacije ključnih faktora koji utiču na proizvodnju solarne energije. Glavni cilj je kreirati modele koji mogu precizno predviđati proizvodnju električne energije po satu.

Ovo rešenje omogućava:

- Optimizaciju proizvodnje energije.
- Bolje planiranje potrošnje energije.
- Efikasnije korišćenje solarne energije u pametnim gradovima.

Razlog za odabir problema proizilazi iz značaja solarne energije kao obnovljivog izvora, njene uloge u smanjenju emisije štetnih gasova i postizanju održivog razvoja, budućnost leži u obnovljivim izvorima energije.

Ovo rješenje se može integrisati u distribuirane sisteme pametnih gradova, gdje bi više lokalnih čvorova (npr. solarne elektrane) doprinijelo centralizovanom upravljanju proizvodnjom i potrošnjom energije.

Skup podataka

Opis podataka:

Podaci korišćeni u projektu preuzeti su sa platforme [AI4EU](#). Skup podataka sadrži sledeće karakteristike:

- AirTemperature (°C): Temperatura vazduha koja utiče na efikasnost solarnih panela.
- CloudOpacity (%): Količina oblačnosti na nebu izražena u procentima. Veće vrijednosti neprozirnosti oblaka znače više oblačnosti, što smanjuje količinu sunčeve svjetlosti koja dopire do solarnih panela. Niže vrijednosti ukazuju na jasnije nebo sa više direktne sunčeve svjetlosti.
- DHI (kW/m²): Difuzno horizontalno zračenje (DHI) je sunčevo zračenje primljeno na horizontalnu površinu sa cijelog neba. Uključuje i direktnu i raspršenu sunčevu svjetlost. Mjeri se u kilovatima po kvadratnom metru (kW/m²).
- DNI (kW/m²): Direct Normal Irradiance (DNI) predstavlja sunčevo zračenje primljeno direktno od sunca na površinu okomitu na sunčeve zrake. Mjeri se u kilovatima po kvadratnom metru (kW/m²).

- EBH (kW/m^2): Extraterrestrial Horizontal Irradiance (EBH) je količina sunčevog zračenja koja bi se primila na površini Zemlje da ne postoji atmosfera. Mjeri se u kilovatima po kvadratnom metru (kW/m^2).
- GHI (kW/m^2): Global Horizontal Irradiance (GHI) je ukupno sunčevo zračenje primljeno na horizontalnu površinu, uključujući direktne i difuzne komponente. Mjeri se u kilovatima po kvadratnom metru (kW/m^2).
- Produkcija – Lokacija 1 (kWh): Ova varijabla predstavlja proizvodnju električne energije iz solarnih panela na lokaciji 1 u kilovat-satima (kWh). Daje se kao satni prosjek, koji označava količinu proizvedene električne energije tokom svakog sata.
- Produkcija – Lokacija 2 (kWh): Ova varijabla predstavlja proizvodnju električne energije iz solarnih panela na lokaciji 2 u kilovat-satima (kWh). Daje se kao satni prosjek, koji označava količinu proizvedene električne energije tokom svakog sata.
- Produkcija – Lokacija 3 (kWh): Ova varijabla predstavlja proizvodnju električne energije iz solarnih panela na lokaciji 3 u kilovat-satima (kWh). Daje se kao satni prosjek, koji označava količinu proizvedene električne energije tokom svakog sata.

Izvori podataka: [AI4EU - Solar Energy Production Dataset](#).

Algoritmi i modeli

Za rješavanje problema korišćeni su sledeći algoritmi mašinskog učenja:

1. Random Forest Regressor:
 - Ensemble metoda sa više stabala odluka za predviđanje kontinuiranih vrednosti.
 - Robustna je i može modelovati nelinearne odnose.
2. Pipeline sa StandardScaler:
 - Normalizacija podataka za efikasno treniranje modela.

Evaluacija modela uključuje:

- MAE (Mean Absolute Error): Prosečna apsolutna greška.
 - MSE (Mean Squared Error): Prosečna kvadratna greška.
 - R^2 Score: Koeficijent determinacije.
-

Tehnologije i alati

Projekat koristi sledeće tehnologije:

- Python: Glavni programski jezik.

- Pandas: Manipulacija i analiza podataka.
 - NumPy: Numeričke operacije.
 - Scikit-learn: Implementacija algoritama mašinskog učenja.
 - Plotly: Vizualizacija podataka.
 - Flask: Backend framework.
 - HTML/CSS/Bootstrap: Frontend za aplikaciju.
 - Jinja2: Templating engine.
 - Visual Studio Code: Razvojno okruženje.
-

Implementacija

1. Preprocesiranje podataka
 - Uklanjanje nedostajućih vrednosti.
 - Grupisanje podataka po satnom intervalu.
 2. Razvoj modela
 - Trenirani su zasebni modeli za svaku lokaciju i ukupnu proizvodnju.
 3. Evaluacija i vizualizacija
 - Prikaz grafova za predikcije i uticaj ulaznih parametara na proizvodnju.
 4. Razvoj korisničkog interfejsa
 - Dinamički frontend sa grafičkim prikazima rezultata.
 - Responzivni dizajn za prikaz podataka na svim uređajima.
-

Ciljevi projekta

Primarni ciljevi:

- Analiza podataka i identifikacija ključnih faktora.
- Razvoj modela za predikciju proizvodnje energije.
- Implementacija aplikacije sa korisničkim interfejsom.

Sekundarni ciljevi:

- Vizualizacija uticaja ulaznih parametara kroz interaktivne grafove.

- Unapređenje korisničkog iskustva kroz moderni dizajn aplikacije.
-

Zaključak

Projekat pruža alat za analizu i predikciju solarne energije sa sledećim benefitima:

- Bolje razumevanje faktora koji utiču na proizvodnju.
- Optimizacija energetskeg sistema.
- Vizualizacija rezultata i intuitivno korisničko iskustvo.

Budući pravci razvoja:

- Proširenje na druge geografske lokacije.
- Integracija sa podacima u realnom vremenu.
- Primena naprednijih algoritama kao što su duboke neuronske mreže.

Model omogućava buduću integraciju u distribuirane energetske mreže pametnih gradova, čime se doprinosi efikasnijem upravljanju obnovljivim izvorima energije na većem geografskom nivou.

Reference

1. [AI4EU - Solar Energy Production Dataset](#)
2. [Scikit-learn Documentation](#)
3. [Flask Documentation](#)
4. [Plotly Documentation](#)
5. [MAE \(Mean Absolute Error\), MSE \(Mean Squared Error\), R² Score](#)