

PRIMENA RANDOM FOREST ALGORTIMA U PROGNOZI BRZINE VETRA I PROFITA VETROGENERATORA**APPLICATION OF THE RANDOM FOREST ALGORITHM IN PREDICTING WIND SPEED AND WIND FARM PROFIT**

Nataša Gavrilović, *Fakultet tehničkih nauka, Novi Sad*

Oblast – ELEKTROTEHNIKA I RAČUNARSTVO

Kratak sadržaj – U ovom radu predstavljen je sistem za predikciju brzine vetra i profita vetrogeneratora korišćenjem tehnologija integracije aplikacija, uz primenu mašinskog učenja. Sistem koristi podatke o brzini vetra i proizvedenoj energiji kako bi precizno procenio budući profit vetroparkova. Implementacija uključuje metode za prikupljanje podataka, njihovu analizu i primenu Random forest algoritma, omogućavajući korisnicima da predvide efikasnost proizvodnje. Rad doprinosi razvoju inovativnih rešenja za praćenje i procenu obnovljivih resursa, čime se olakšava prelazak na održiviju energetska budućnost.

Ključne reči: Mašinsko učenje, Random Forest algoritam, Analiza podataka, Predikcija brzine vetra i profita

Abstract – This paper presents a system for predicting the wind speed and profit of wind turbines using integration technologies, along with the applications of machine learning. The system utilizes data on wind speed and energy production to accurately estimate future profit of wind farms. The implementation includes methods of data collection, analysis, and the use of random forest algorithm, enabling users to forecast production efficiency. The paper contributes to the development of innovative solutions for monitoring and assessing renewable resources, facilitating the transition to a more sustainable energy future

Keywords: Machine Learning, Random Forest algorithm, Data Analysis, Wind speed and profit prediction

1. UVOD

Uzeći u obzir sve veći značaj obnovljivih izvora energije u savremenom svetu, naročito vetroenergije, potreba za optimizacijom proizvodnje i upravljanjem vetroparkovima postaje ključna za postizanje energetske efikasnosti i održivosti. Vetrogeneratori predstavljaju važan izvor čiste energije, ali njihova profitabilnost zavisi od niza faktora. Ovaj rad se bavi razvojem naprednog sistema za predikciju brzine vetra i profita vetrogeneratora, koristeći tehnologije integracije uz primenu algoritma mašinskog učenja.

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bio dr Aleksandar Bošković, docent

Glavni cilj rada je izgradnja sistema koji može prikupljati podatke o brzini vetra, proizvodnji energije i drugim relevantnim parametrima, analizirati ih i koristiti za predikciju brzine vetra i budućeg profita vetroparkova. Pored algoritma za predikciju, rad se bavi i izazovima u prikupljanju i obradi podataka sa različitih izvora, integraciji podataka između različitih aplikacija i implementaciji rešenja koje je prilagodljivo i lako skalabilno.

2. TEORIJSKE OSNOVE I DEFINICIJE

U ovom poglavlju biće predstavljene teorijske osnove i definicije najbitnijih koncepata i tehnologija koje su korišćene u izradi sistema za predikciju profita vetrogeneratora. Ovo uključuje osnovne principe obnovljivih izvora energije, mašinskog učenja kao i integraciju između različitih softverskih komponenti.

2.1. Obnovljivi izvori energije i vetrogeneratori

Obnovljivi izvori zauzimaju sve važnije mesto u savremenom društvu, jer pružaju mogućnost proizvodnje energije bez štetnog uticaja na životnu sredinu. Među njima, energija vetra se ističe kao jedan od najperspektivnijih i ekološki najprihvatljivijih izvora, jer ne proizvodi emisije štetnih gasova niti zagađuje okolinu. Vetro turbine, koje koriste kinetičku energiju vetra za generisanje električne energije, predstavljaju ključnu tehnologiju u prelazu sa fosilnih goriva na čiste izvore energije. Ova tehnologija ne samo ga smanjuje zavisnost od fosilnih goriva, već doprinosi i energetske sigurnosti i održivosti. Vetro turbine funkcionišu na osnovu jednostavnog, ali efektivnog principa. One pretvaraju kinetičku energiju vetra u mehaničku energiju, a zatim u električnu energiju. Kada vetar duva kroz elise vetro turbine, one se okreću i pokreću rotor, deo vetro turbine koji sakuplja energiju vetra. Rotor je povezan sa glavnim vratilom, koje prenose rotaciono kretanje na generator. Generator potom pretvara ovu mehaničku energiju u električnu energiju, koja se može koristiti za napajanje domova, fabrika i drugih potrošača. Jedan od ključnih aspekata koji utiče na efikasnost proizvodnje energije vetro turbine je brzina i stabilnost vetra. Projektovane su da optimalno rade unutar specifičnih opsega brzina vetra. U uslovima slabog vetra, produktivnost vetroelektrana može biti smanjena, dok se u slovima jakih vetrova njihova proizvodnja drastično povećava. Osim tehničkih prednosti, energija vetra donosi

i ekonomske koristi. Sa smanjenjem troškova tehnologije i povećanjem kapaciteta proizvodnje, vetroelektrane postaju sve konkurentnije u poređenju sa tradicionalnim izvorima energije. Međutim, osiguranje stabilne i predvidive proizvodnje energije iz vetra zahteva strateško planiranje i integraciju sa drugim izvorima energije u mreži. Energija vetra se pokazuje kao ključni deo budućeg energetskog sistema. Njena sposobnost da doprinese smanjenju zavisnosti od fosilnih goriva, poboljšanju energetske sigurnosti i smanjenju emisija stakleničkih gasova čini je neizostavnim resurskom u borbi protiv klimatskih promena. Uzeći u obzir sve veći fokus na obnovljive izvore energije, energija vetra će nastaviti da igra centralnu ulogu u globalnoj tranziciji ka održivijem i čistijem energetskom sistemu [1].

2.2. Mašinsko učenje i Random forest algoritam

Mašinsko učenje trenutno predstavlja dominantnu formu veštačke inteligencije. Algoritmi mašinskog učenja su osnova širokog spektra aplikacija koje imaju stvarni uticaj na pojedince, kompanije i vlade. Jedan od ključnih izazova u razvoju sistema zasnovanih na mašinskom učenju je osigurati da tehnički napredna rešenja imaju pozitivan društveni uticaj [2]. Random forest algoritam je metoda statičkog učenja koja se koristi u mnogim oblastima primene, ne samo zbog odličnih performansi u predviđanju, već i zbog svoje fleksibilnosti, uz nekoliko ograničenja u pogledu prirode podataka. Prilagođen je i za nadzirane probleme klasifikacije i za probleme regresije. Omogućava razmatranje kvalitativnih i kvantitativnih objašnjavajućih varijabli zajedno, bez potrebe za prethodnom obradom. Takođe, mogu se koristiti za obradu standardnih podataka kod kojih je broj posmatranja veći od broja varijabli, a veoma je efikasan i u slučaju visoke dimenzionalnosti, kada je broj varijabli znatno veći u odnosu na broj posmatranja. Kao rezultat toga, danas su među preferiranim metodama u alatu statističara i drugih naučnika koji rade sa podacima [3].

2.3. Prikupljanje i analiza podataka

Efikasno predviđanje profita vetrogeneratora zahteva prikupljanje preciznih podataka o brzini vetra, vremenskim uslovima i trenutnoj proizvodnji električne energije. U okviru sistema razvijenog u ovom radu, podaci se prikupljanju za praćenje vremenskih uslova, te se potom obrađuju kako bi se predvidela proizvodnja energije. Nakon prikupljanja, ovi podaci prolaze kroz proces obrade kako bi se identifikovale promene u vremenskim obrascima koje mogu uticati na proizvodnju energije. Prikupljeni podaci o vremenskim uslovima, koriste se za predikciju količine energije koja će biti proizvedena u budućem vremenu, što je ključno za predviđanje profita vetrogeneratora.

Proizvodnja električne energije vetrogeneratora direktno zavisi od brzine vetra. Međutim, kako bi se tačno izračunala količina proizvedene energije, koristi se formula koja uzima u obzir nekoliko ključnih parametara: minimalnu brzinu vetra za pokretanje vetrogeneratora V_{min} , brzinu na kojoj generator radi sa punim kapacitetom V_{full} i maksimalnu brzinu vetra nakon koje generator prestaje sa radom V_{max} . Parametar P_n predstavlja nominalnu snagu generatora, tj. maksimalnu količinu energije koju može proizvesti pod idealnim uslovima. Kada se razmatra snaga vetrogeneratora, uzimaju se u obzir

različite brzine vetra i prilagođeni režimi rada kako bi se obezbedila stabilna i bezbedna proizvodnja energije. Formula za izračunavanje snage vetrogeneratora izgleda ovako:

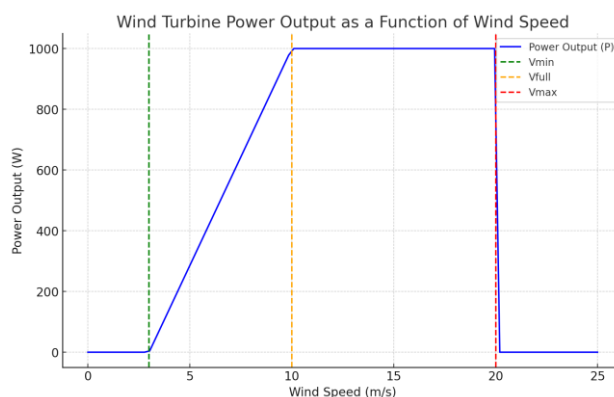
- Ako je brzina vetra manja od V_{min} , vetrogenerator ne proizvodi energiju jer je brzina vetra nedovoljna za pokretanje elisa. Proizvedena snaga je nula.
- Ako je brzina vetra između V_{min} i V_{full} , u ovom opsegu brzine vetra, snaga vetrogeneratora se postepeno povećava u skladu sa rastom brzine vetra. Formula za izračunavanje snage glasi:

$$P = \frac{(V - V_{min})}{(V_{full} - V_{min})} \times P_n$$

gde su P proizvedena snaga pri brzini vetra, V je trenutna brzina vetra, V_{min} minimalna brzina vetra potrebna za pokretanje vetrogeneratora, V_{full} brzina vetra pri kojoj vetrogenerator dostiže pun kapacitet i P_n nominalna snaga vetrogeneratora. Ova formula omogućava linearan rast snage sa povećanjem brzine vetra, dok se ne dostigne maksimalni kapacitet vetrogeneratora.

- Ako je brzina vetra između V_{full} i V_{max} , generator radi sa punim kapacitetom, odnosno proizvodi maksimalnu moguću snagu P_n bez obzira na dodatno povećanje brzine vetra.
- Ako je brzina vetra veća od V_{max} , vetrogenerator se automatski isključuje kako bi se izbegla potencijalna oštećenja zbog prejakih vetrova. U ovom slučaju, proizvedena snaga je ponovo nula.

Ova formula omogućava tačno izračunavanje proizvedene snage vetrogeneratora na osnovu promena brzine vetra.



Slika 1. Zavisnost snage vetrogeneratora od brzine vetra

Prikupljeni podaci o proizvodnji energije se koriste kao ulaz u model, koji predviđa budući profit vetrogeneratora. Model uzima u obzir istorijske podatke o proizvodnji, promene u vremenskim uslovima i predviđene cene energije na tržištu kako bi se procenio očekivani profit u različitim scenarijima. Predikcioni model time pomaže u planiranju i donošenju odluka o daljim investicijama, optimizaciji rada vetrogeneratora i izračunavanje potencijalnih ušteda ili dobiti.

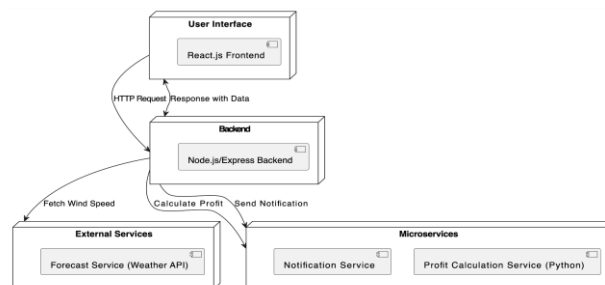
3. ARHITEKTURA SISTEMA

Arhitektura sistema za predikciju profita vetroparkova zasnovana je na integraciji između *Node.js* i *Python*

aplikacija, što omogućava efikasno prikupljanje podataka, analizu i predikciju profita. Sistem je dizajniran kao modularana i skalabilna arhitektura koja omogućava laku integraciju novih komponenti i funkcionalnosti. Glavne komponente sistema su:

1. Servis za prikupljanje i obradu podataka: Ovaj servis je odgovoran za prikupljanje podataka o brzini vetra i trenutnoj proizvodnji energije. Aplikacija se povezuje sa bazom podataka i čuva sve dobijene podatke za dalju obradu. Uz to, implementirana je funkcija za izračunavanje trenutne proizvedene snage koristeći formulu koja uzima u obzir brzinu vetra i karakteristike generatora, kao što su minimalna, puna i maksimalna brzina rada.
2. Servis za predikciju profita: Servis je implementiran kao *Python* aplikacija koja koristi algoritme mašinskog učenja. On funkcioniše kao ključni deo sistema za predviđanje budućih prihoda vetrogeneratora. Koristeći model zasnovan na Random Forest algoritmu, servis vrši predikciju profita na osnovu podataka o brzini vetra, trenutno proizvedenoj energiji i drugih parametara koji utiču na rad vetrogeneratora. Kada glavni servis prikupi podatke, kao što su brzina vetra i trenutno proizvedena energija, on ih prosleđuje servisu za predikciju putem HTTP zahteva. Ovi podaci se unose u model mašinskog učenja, koji je prethodno obučen na velikom skupu istorijskih podataka o vremenskim uslovima, brzinama vetra i prihodima, čime se obezbeđuje visoka tačnost predikcija. Model kontinuirano uči iz novih podataka, omogućavajući adaptaciju na promene u vremenskim obrascima. Algoritam Random Forest analizira različite faktore koji utiču na promene u brzinama vetra i generisanu energiju u različitim vremenskim periodima i, na osnovu tog odnosa, predviđa očekivani prihod u budućnosti. Kada servis završi obradu, povratne informacije o predviđenom profitu šalju se nazad glavnom servisu, gde se podaci automatski skladište i prikazuju korisnicima u interfejsu.
3. Baza podataka: Sistem koristi *NoSql* bazu podataka *MongoDB* za skladištenje podataka o vetroparkovima, uključujući informacije o njihovoj lokaciji, tipu generatora, trenutnoj i ukupnoj proizvedenoj snazi, kao i istoriju profita. Svaki podatak o vetroparku povezan je sa korisnikom koji upravlja tim sistemom, čime omogućava lako upravljanje podacima na nivou korisnika.
4. Interfejs za upravljanje i analizu podataka: Korisnički interfejs omogućava operaterima da u realnom vremenu prate stanje svojih vetroparkova, brzinu vetra i proizvodnju energije, kao i da dobiju predikciju profita. Interfejs je implementiran kao *React* aplikacija koja komunicira sa glavnim servisom preko HTTP API zahteva. Korisnici mogu da biraju određeni vetropark na mapi i zatraže predikciju profita za određen vremenski period, nakon čega se rezultati prikazuju direktno u aplikaciji.

5. Servis za slanje notifikacija: Ovaj servis je zadužen za automatsko obaveštavanje korisnika o važnim događajima i promenama na vetroparkovima. Kada se registruje novi korisnik ili kada dođe do značajnih promena u podacima u brzini vetra ili proizvedenoj energiji, sistem šalje notifikaciju putem email-a koristeći *SendGrid*. Glavna aplikacija šalje poruke preko *RabbitMQ* servera, koji upravlja redovima i obezbeđuje pouzdanu i asinhronu isporuku obaveštenja.



Slika 2. Arhitektura sistema za predikciju profita

4. PREDIKCIJA PROFITA I RANDOM FOREST ALGORITAM

Servis za predikciju profita predstavlja važan deo sistema za optimizaciju rada vetrogeneratora i upravljanje vetroparkovima. Server koristi napredne tehnike mašinskog učenja, posebno algoritam Random Forest, kako bi omogućio precizno predviđanje budućih prihoda vetroparkova na osnovu istorijskih podataka o brzini vetra i ostvarenim profitima. Implementacija servisa realizovana je kao odvojena Python aplikacija. Model analizira složene odnose između brzine vetra i ostvarenih prihoda, omogućavajući izradu predikcija zasnovanih na više faktora. Ulazni podaci koje servis zahteva uključuju *wind_speed_data*, niz podataka o istorijskim brzinama vetra koji su prikupljeni tokom određenog perioda, i *actual_profits*, niz stvarnih podataka o profitima koji su ostvareni u istom vremenskom periodu kao i brzine vetra. Izlazni podaci predstavljaju *predicted_profit*, niz predviđenih profita koji se izračunavaju na osnovu podataka o brzini vetra primenom Random Forest algoritma. On funkcioniše tako što koristi veliki broj pojedinačnih stabala odluke i kombinuje njihove rezultate kako bi postigao stabilniju i precizniju predikciju. Svako stablo se trenira na slučajnom podskupu podataka, čime se model bolje prilagođava složenim i nelinearnim obrascima u podacima. Na osnovu prošlih podataka o brzini vetra i profitima, Random Forest model procenjuje uticaj različitih varijabli na profit, omogućavajući efikasne predikcije budućih prihoda. Ovaj pristup omogućava modelu da, na osnovu prethodnih podataka, precizno proceni kako se promena brzine vetra odražava na promenu u profitu. On koristi ansambl stablo odluka gde svako stablo pravi svoje procene, a krajnja predikcija nastaje kombinacijom predikcija svih stabala.

5. IMPLEMENTACIJA RANDOM FOREST

Implementacija funkcije za predikciju koristi biblioteke za mašinsko učenje iz *Python* okruženja, kao što su *numpy* i *scikit-learn*, specijalizovane za rad sa podacima i modele mašinskog učenja. Funkcija prolazi kroz nekoliko koraka. Prvi korak je učitavanje i priprema podataka kroz funkciju

`load_and_prepare_data`, koji učitava CSV fajlove za 2022, 2023 i 2024 godinu, koji sadrže informacije o vremenskim parametrima kao što su prosečna, minimalna i maksimalna temperatura, pravac vetra, brzina vetra, pritisak i druge varijable. Potom filtrira nulte vrednosti i interpolira nedostajuće podatke. Zatim se podaci transformišu u odgovarajući format za treniranje modela, zadržavajući samo relevantne vremenske parametre. Sledeći korak je treniranje Random Forest modela, koje se odvija u funkciji `train_random_forest_model`. Da bi se pronašli optimalni hiperparametri, koristi se `GridSearchCV`, koji pretraćuje kroz napred definisani skup parametara, uključujući broj stabala `n_estimators`, maksimalnu dubinu `max_depth` i druge.

```
X = data[['tavg', 'tmin', 'tmax', 'wdir']]
y = data['wspd']

param_grid = {
    'n_estimators': [100, 200],
    'max_depth': [10, 20]}

rf_model = RandomForestRegressor(random_state=42)

grid_search = GridSearchCV(
    estimator=rf_model,
    param_grid=param_grid,
    cv=3,
    scoring='neg_mean_squared_error',
    n_jobs=-1)

grid_search.fit(X, y)

best_model = grid_search.best_estimator_
```

Listing 1. Kod treniranja Random Forest modela u Python-u

Kada je model testiran na parametrima maksimalna dubina 15, maksimalne karakteristike `sqrt`, minimalni broj uzoraka u listu 1, minimalni broj uzoraka za podelu 2 i broj stabala 200, postigao je *Mean Squared Error* (MSE) 1.1600050874260937 i *Mean Score* od 0.9740540595776065 što ukazuje na visoku tačnost modela. *Mean Score* modela odražava koliko dobro model objašnjava varijaciju u podacima, dok MSE pokazuje prosečnu kvadratnu grešku između predikcija i stvarnih vrednosti. Nakon što je model treniran i optimizovan, sledeći korak je predikcija buduće brzine vetra. Funkcija `predict_future_wind_speed` kreira vremensku seriju budućih datuma između `start_date` i `end_date` i zatim popunjava podatke korišćenjem poslednjeg poznatog reda iz `all_data`, što simulira buduće vremenske uslove. Random forest model koristi ove podatke za predikciju buduće brzine vetra za svaki dan u zadatom vremenskom periodu. Ova funkcionalnost omogućava da se na osnovu prethodnih vremenskih parametara predvidi kako će se brzina vetra menjati, što je ključna informacija za izračunavanje potencijalnog profita vetrogeneratora.

Za kraj, API omogućava korisnicima da postave zahteve za predikciju buduće brzine vetra, kao i profita korišćenjem ovog modela. Kada korisnik pošalje POST zahtev sa početnim i krajnjim datumom, API koristi `predict_future_wind_speed` za kreiranje predikcija i vraća

rezultate u JSON formatu. Sa visokim Mean Score-om od 0.974, model pokazuje visok nivo preciznosti, što potvrđuje njegovu pouzdanost u predikciji brzine vetra i budućih prihoda vetroparkova.

6. ZAKLJUČAK

Ovaj rad osvetljava značaj razvoja sistema za predikciju brzine vetra i profita vetrogeneratora kao ključni korak ka unapređenju energetske efikasnosti i ekonomičnosti obnovljivih izvora energije. Kroz integraciju tehnologija mašinskog učenja, prikupljanja podataka i interfejsa za upravljanje, omogućeno je preciznije planiranje proizvodnje i efikasnija alokacija resursa u vetroparkovima. Jedan od glavnih doprinosa sistema razvijenog u ovom radu jeste njegova sposobnost da koristi istorijske podatke o brzini vetra i energetskoj proizvodnji za preciznu predikciju profita. Sistem nudi fleksibilnost i modularnu arhitekturu, što olakšava buduće nadogradnje i proširenja. Integracija između aplikacija pokazala je kako različite tehnologije mogu efikasno da sarađuju kako bi se obezbedila preciznost i pouzdanost sistema. Rezultati predikcija profita dobijeni primenom Random forest algoritma ukazuju na mogućnost korišćenja algoritama mašinskog učenja u realnim poslovnim procesima. Rad potvrđuje da je razvoj sistema za predikciju profita važan korak ka efikasnijem korišćenju obnovljivih izvora energije, smanjenju troškova i unapređenju održivosti, čime doprinosi globalnim naporima u borbi protiv klimatskih promena i prelasku na čistu energiju.

7. LITERATURA

- [1] Wei tong, "Fundamentals of Wind Energy", *WIT Transactions on State of the Art in Science and Engineering*, Vol 44, 2010.
- [2] Fuwei Li, Lifeng Lai, Shuguang Cui, "Machine Learning Algorithms", 2022
- [3] Robin Genuer, Jean-Michel Poggi, "Random Forests with R", Springer Nature Switzerland, 2020.

Kratka biografija:



Nataša Gavrilović rođena je 8. novembra 1999. godine u Zrenjaninu. Fakultet tehničkih nauka u Novom Sadu, studentski program Primenjeno softversko inženjerstvo upisala je 2018. godine. Nakon završenih osnovnih studija, 2023. godine upisala je master akademske studije iz iste oblasti.
Kontakt:
gavrilovicnataša123@gmail.com