

DOI: [10.46793/CIGRE37.A1.04](https://doi.org/10.46793/CIGRE37.A1.04)

A1.04

**KOMPARATIVNA ANALIZA REKURENTNIH I TRANSFORMER MODELA
MAŠINSKOG UČENJA ZA PREDIKCIJU PROIZVODNJE AKTIVNE SNAGE
VETROGENERATORA****COMPARATIVE ANALYSIS OF RECURRENT AND TRANSFORMER MACHINE
LEARNING MODELS FOR WIND TURBINE ACTIVE POWER GENERATION
FORECASTING****Luka Ivanović, Ilija Klasnić, Saša Milić***

Kratak sadržaj: Vetrogeneratori su inherentno intermitentni izvori energije koji pretežno zavise od meteoroloških uslova. Precizno predviđanje proizvodnje aktivne snage vetrogeneratora ključno je za optimizaciju rada modernih elektroenergetskih sistema bogatim obnovljivim izvorima energije, poboljšanje stabilnosti sistema, efikasnije balansiranje proizvodnje i potrošnje, optimizaciju tržišta električne energije i brojne druge aspekte. Razvojem i primenom naprednih metoda veštačke inteligencije i mašinskog učenja (ML) moguće je dodatno povećati tačnost predikcije, omogućavajući sistemu da se bolje prilagodi varijabilnoj prirodi vetroenergije. U radu istražujemo efikasnost dva pristupa mašinskog učenja za predikciju aktivne snage vetrogeneratora, fokusirajući se na ML modele zasnovane na rekurentnim neuralnim mrežama (LSTM i GRU) i na Transformer modele prilagođenije procesiranju i predikciji vremenskih serija. Transformer modeli su sve prisutniji u analizi vremenskih serija zbog sposobnosti da efikasno prepoznaju dugoročne zavisnosti u podacima primenom *self-attention* mehanizma, omogućavajući paralelnu obradu sekvenci i precizniju identifikaciju relevantnih obrazaca u podacima. Izbor arhitekture transformer modela i podešavanje njegovih parametara su dva uslovljena iterativna postupka. Validacija realizovanog ML modela se vrši na praktičnom open-source skupu podataka koji sadrži veliki broj ulaznih karakteristika vezanih za meteorološke i operativne parametre sistema. Selekcija najrelevantnijih obeležja vrši se na osnovu korelacije sa ciljanom promenljivom (u ovom slučaju aktivnom snagom), čime se smanjuje dimenzionalnost problema i poboljšava efikasnost modela. Eksperimentalna evaluacija uključuje analizu performansi modela pomoću standardnih metrika (RMSE i MAEI), pri čemu se paralelno analiziraju arhitektura transformera i skup vrednosti njegovih parametara. Dobijeni rezultati pružaju uvid u prednosti i ograničenja oba pristupa u više različitih scenarija sa kratkoročnom i dugoročnom predikcijom, sa ciljem poboljšanja šire strategija predikcije i optimizacije rada vetroelektrana.

* Luka Ivanović, Elektrotehnički Institut Nikola Tesla Beograd, luka.ivanovic@ieent.org
Ilija Klasnić, Elektrotehnički institut Nikola Tesla ad Beograd, ilija.klasnic@ieent.org
Saša Milić, Elektrotehnički institut Nikola Tesla, s-milic@ieent.org

Ključne reči: *Mašinsko učenje, Rekurentne neuralne mreže, Transformer modeli, Vremenskeserije*

Abstract: Wind turbines are inherently intermittent energy sources that predominantly depend on meteorological conditions. Accurate prediction of wind turbine active power generation is crucial for optimizing the operation of modern power systems rich in renewable energy sources, improving system stability, enhancing the efficiency of production and consumption balancing, optimizing the electricity market, and numerous other aspects. The development and application of advanced artificial intelligence and machine learning (ML) methods can further increase prediction accuracy, enabling the system to better adapt to the variable nature of wind energy. This paper investigates the efficiency of two machine learning approaches for predicting wind turbine active power, focusing on ML models based on recurrent neural networks (LSTM and GRU) and Transformer models adapted for processing and predicting time series data. Transformer models have recently gained prominence in time series analysis due to their ability to effectively recognize long-term dependencies in data through the self-attention mechanism, enabling parallel sequence processing and more accurate identification of relevant patterns in the data. The choice of transformer model architecture and parameter tuning is conditioned by two iterative processes. The validation of the developed ML model is performed using a practical open-source dataset that contains a large number of input features related to meteorological and operational system parameters. The selection of the most relevant features is based on their correlation with the target variable (in this case, active power), reducing the dimensionality of the problem and improving model efficiency. The experimental evaluation includes an analysis of model performance using standard metrics (RMSE and MAEI), while simultaneously examining the transformer architecture and its parameter set. The obtained results provide insight into the advantages and limitations of both approaches in various scenarios of short-term and long-term prediction, with the aim of improving broader prediction strategies and optimizing wind farm operations.

Key words: *windgenerator, machine learning, recurent neural network, transformer model, time series*

1 UVOD

Aktuelna energetska tranzicija, oslonjena na primenu obnovljivih izvora energije, je ključni faktor u borbi protiv klimatskih promena i smanjenju emisije štetnih gasova sa efektom staklene bašte [1,2]. Jedan od obnovljivih izvora su i vetroelektrane koje u svojoj proizvodnji koriste energiju vetra na ekološki prihvatljiv način. Prema [3] instalisana snaga vetroelektrana u svetu na kraju 2023. godine iznosi 1017390 MW i u poslednjih deset godina se povećala skoro 3 puta. Moderne vetroturbine mogu da rade pri manjim vrednostima brzine vetra i uključuju pametne tehnologije za praćenje performansi u realnom vremenu. Konstantni napretci u ovoj oblasti omogućavaju da energija vetra ostane konkurentna opcija u aktuelnoj energetskejoj tranziciji [4, 5].

Aktivna snaga vetroturbine zavisi od nekoliko faktora ali dominantan je brzina vetra. Međutim, kako je ova veličina po prirodi intermitentna i stohastička (sa velikim fluktuacijama) i zavisi od različitih faktora kao što su klimatski uslovi i godišnja doba, njeno predviđanje je prilično kompleksno [6].

Tradicionalno, sistemi za predikciju snage vetroturbina su se zasnivali na ne toliko preciznim statističkim modelima koji su uzimali u obzir istorijske podatke o brzini vetra, smeru vetra, atmosferskim uslovima (temperatura i vlažnost vazduha) i dr. U poslednje vreme se u ove svrhe sve više koriste različite tehnike mašinskog učenja zbog sposobnosti da efikasno prepoznaju dugoročne zavisnosti u podacima kroz *self-attention* mehanizam i preciznije identifikuju relevantne obrazce u podacima [7]-[10]. Nedostatak tačne predikcije brzine vetra, a samim tim i aktivne snage vetroturbine može dovesti do značajnih problema u elektroenergetskom sistemu koji je bogat vetroturbinama: nedovoljne rezerve aktivne snage, problemi sa stabilnošću sistema i dr. Iz tog razloga se uspešna integracija vetroturbina u moderne elektroenergetske sisteme čvrsto oslanja na preciznom predviđanju snage vetra.

U radu su opisana dva pristupa mašinskog učenja za predikciju aktivne snage vetrogeneratora, fokusirajući se na ML modele zasnovane na rekurentnim neuralnim mrežama (*LSTM* i *GRU*) i na Transformer modelima. Validacija realizovanih ML modela je izvršena na praktičnom *open-source* skupu podataka jedne vetroturbine koji sadrži veliki broj ulaznih karakteristika vezanih za meteorološke i operativne parametre sistema [11]. Selekcija najrelevantnijih obeležja izvršena je na osnovu korelacije sa ciljanom promenljivom tj. aktivnom snagom vetrogeneratora, čime je smanjena dimenzionalnost problema i poboljšana efikasnost predloženih modela. Eksperimentalna evaluacija podrazumeva analizu performansi predloženih modela pomoću standardnih metrika (*RMSE* i *MAEI*). Cilj primene tehnika mašinskog učenja u predikciji proizvodnje aktivne snage vetrogeneratora je da se poboljšaju preciznost predikcije i pouzdanost same vetroelektrane.

Rad je strukturiran na sledeći način: U prvom poglavlju, **Uvod**, predstavljena je osnovna motivacija i ciljevi istraživanja. Drugo poglavlje, **Opis i obrada podataka**, obuhvata detaljan prikaz korišćenih podataka, proces njihove selekcije, metode obrade i način podele skupa podataka. U trećem poglavlju, **Rekurentne neuronske mreže**, dat je kratak pregled rekurentnih neuronskih mreža, uz objašnjenje potrebe za korišćenjem LSTM i GRU modela, kao i njihov opis. Četvrto poglavlje, **Transformer modeli**, sadrži opšti pregled transformer modela, uz detaljniji opis PatchTST modela koji je primenjen u ovom radu. U petom poglavlju, **Rezultati**, predstavljeni su dobijeni rezultati i izvršena je komparativna analiza modela. Konačno, u šestom poglavlju, **Zaključak**, sumirani su ključni zaključci istraživanja i predloženi pravci za budući rad na ovoj temi.

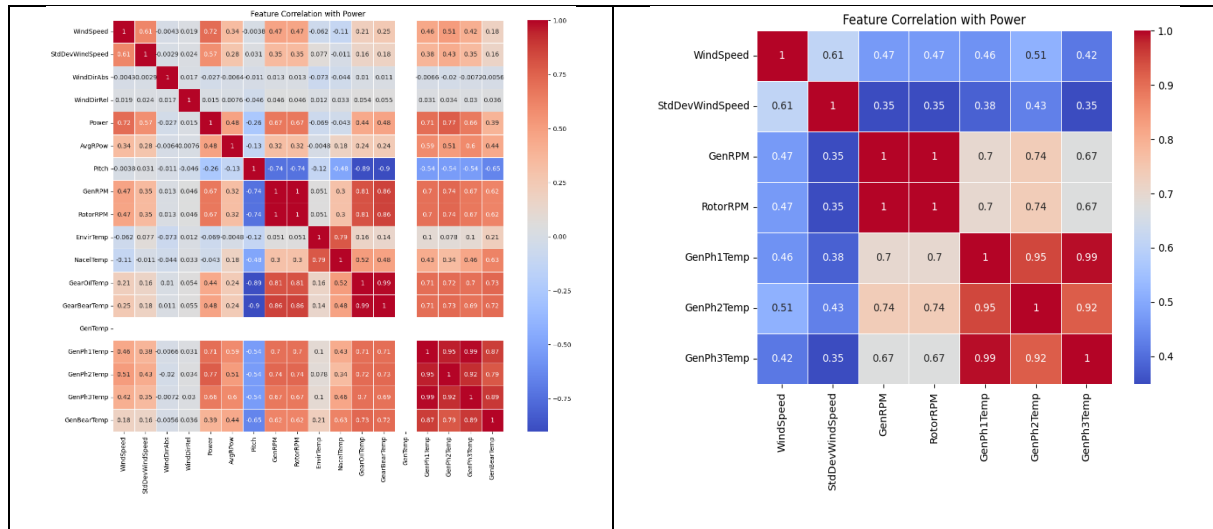
2 OPIS I OBRADA PODATAKA

Za predikciju aktivne snage vetrogeneratora korišćeni su arhivski podaci (merenja) prikupljeni na SCADA sistemu vetroelektrane sa desetominutnom rezolucijom za period od 14 godina. Na raspolaganju su bili sledeći podaci: brzina vetra, apsolutni smer vetra, relativni pravac gondole u odnosu na apsolutni smer vetra, aktivna snaga vetrogeneratora, reaktivna snaga vetrogeneratora, brzina obrtanja vetrogeneratora, temperatura ambijenta u okolini gondole, temperatura reduktora, temperature faznih namotaja i ležajeva vetrogeneratora.

2.1 Obrada podataka

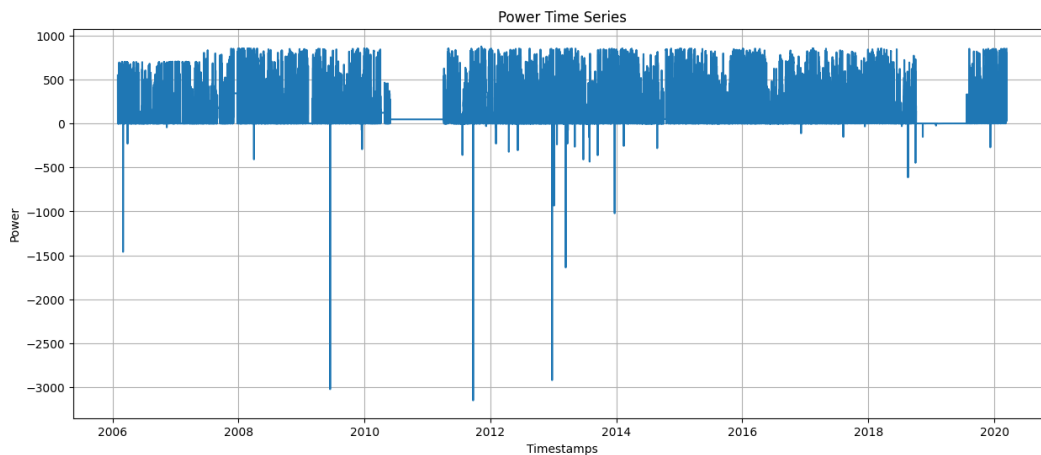
S obzirom na to da su prikupljeni podaci originalno arhivirani u intervalima od 10 minuta, izvršena je interpolacija kako bi se obezbedila satna rezolucija podataka. Za svaki sat, izračunata je srednja vrednost šest desetominutnih uzoraka i korišćena kao reprezentativna vrednost za taj vremenski korak. Ovaj pristup omogućava smanjenje šuma u podacima i poboljšava konzistentnost ulaznih podataka za modele vremenskih serija.

S obzirom na veliki broj ulaznih podataka od interesa je bilo utvrditi međusobnu korelaciju u cilju odabira najrelevantnijih podataka za predikciju aktivne snage. Na slici 1a prikazana je korelaciona matrica koja sadrži koeficijente korelacije između svih ulaznih podataka, na osnovu koje su određeni najrelevantniji za predikciju aktivne snage, a koji su prikazani na slici 1b. Kriterijum za odabir najrelevantnijih podataka je da apsolutna vrednost koeficijenta bude veća od 0.5.



Slika 1: a) korelaciona matrica za sve ulazne podatke b) korelaciona matrica za najrelevantnije ulazne podatke za predikciju aktivne snage vetrogeneratora

Prikupljeni podaci pokrivaju period od 2006. do 2020. godine i obuhvataju arhivske vrednosti radnih parametara vetrogeneratora i meteoroloških uslova. Pre početka modeliranja, izvršena je temeljna analiza podataka kako bi se osigurala njihova validnost i relevantnost za zadatak predikcije aktivne snage vetroturbine. Na slici 2 prikazane su vrednosti aktivne snage (izlazne varijable čija vrednost se predviđa):

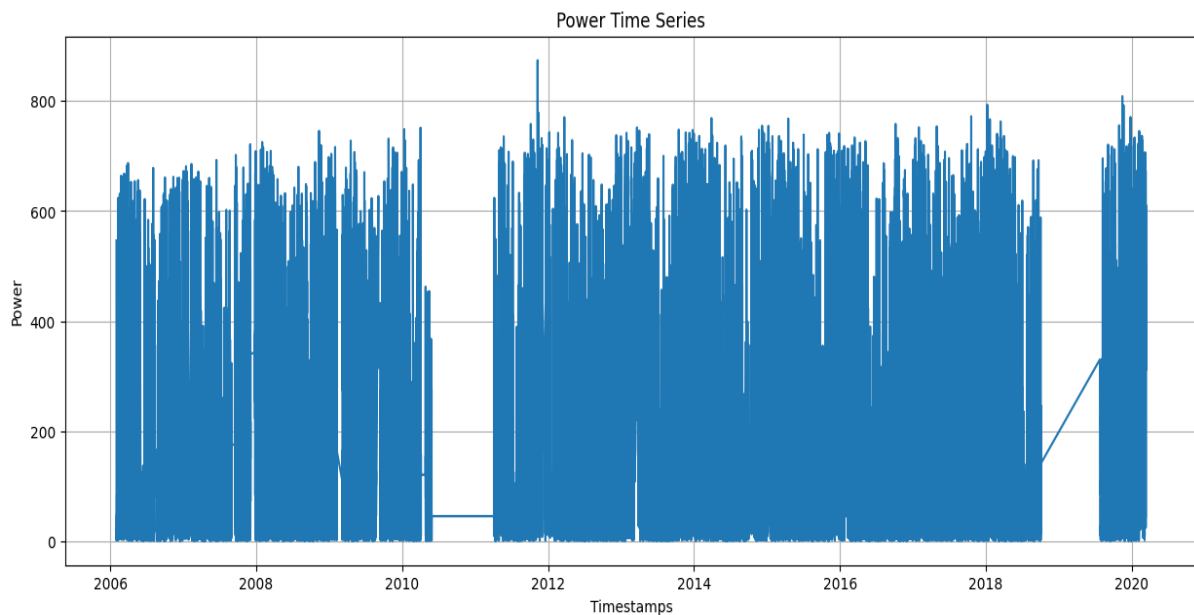


Slika 2: Vrednosti aktivne snage za period od 2006. do 2020. godine

Pošto je uočen veliki broj odstupajućih podataka (*outlier-a*), podaci su isfiltrirani i iskorišćena je *z-score* metoda kao tehnika koja se koristi za identifikaciju outlier-a u skupu podataka (2.1):

$$Z = \frac{x - \mu}{\sigma} \quad (2.1)$$

gde su: x vrednost odbirka, μ je srednja vrednost skupa podataka a σ standardna devijacija skupa podataka. Izabrana vrednost praga je $Z > 2$, odnosno *outlier* je onaj odbirak čija je vrednost *z-score*-a veća za 2 standardne devijacije od srednje vrednosti. Rezultati filtriranja ulaznih podataka primenom *z-score* metoda su prikazani na slici 3.



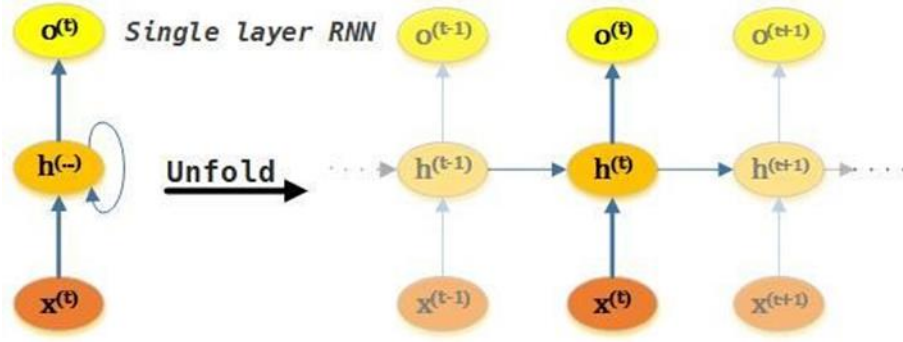
Slika 3: Vrednosti aktivne snage bez outlier-a

Zbog neadekvatnih vrednosti, podaci iz 2011. i 2019. godine su isključeni iz skupa, ali je struktura vremenske serije očuvana uprkos uklanjanju te podsekvence. Nakon filtriranja podataka, primenjen je postupak normalizacije kako bi se obezbedila ujednačenost različitih obeležja. Korišćen je MinMaxScaler metod, kojim su sve vrednosti skalirane u opsegu $[0,1]$. Ovaj postupak omogućava bržu i stabilniju konvergenciju modela tokom procesa treniranja.

Na kraju, podaci su podeljeni na trening i test skup u odnosu 80:20 u korist trening skupa. Trening skup je korišćen za obučavanje modela, dok je test skup korišćen za evaluaciju generalizacione sposobnosti modela i poređenje performansi različitih arhitektura (PatchTST, LSTM i GRU). Pri selekciji trening i test skupa vođeno je računa da se očuva vremenska struktura podataka, kako bi se obezbedilo realno testiranje modela u kontekstu vremenske predikcije.

3 REKURENTNE NEURONSKE MREŽE

Rekurentne neuronske mreže (RNN) su arhitekture neuronskih mreža dizajnirane za obradu sekvencijalnih podataka tako što zadržavaju informacije iz prethodnih vremenskih koraka kroz rekurentne veze. RNN se "odmotava" kroz vremenske korake, pri čemu se isti parametri primenjuju na svakom koraku. Standardne veze se prenose kroz slojeve, dok rekurentne veze omogućavaju prenos informacija između susednih vremenskih koraka. Kako je prikazano na slici 4, RNN se može posmatrati kao neuronska mreža s direktnim prosleđivanjem, ali sa zajedničkim parametrima kroz vreme:



Slika 4: „Odmotana“ rekurentna neuronska mreža

Svaka skrivena jedinica (h) u RNN prima dva ulaza: preaktivaciju iz ulaznog sloja i aktivaciju istog skrivenog sloja iz prethodnog vremenskog koraka $t-1$. Na početku, kada je $t = 0$, skrivene jedinice se inicijalizuju na nule ili male nasumične vrednosti. Za korake gde je $t > 0$, skrivene jedinice primaju vrednost trenutnog ulaza $x(t)$, kao i vrednosti skrivenih jedinica iz prethodnog koraka $h(t-1)$. Vrednost skrivene jedinice u trenutku t se računa na sledeći način:

$$\mathbf{h}^{(t)} = \sigma_h(\mathbf{W}_{xh}\mathbf{x}^{(t)} + \mathbf{W}_{hh}\mathbf{h}^{(t-1)} + \mathbf{b}_h) \quad (3.1)$$

gde je $\mathbf{W}_{xh}$ težinska matrica između ulaza $\mathbf{x}^{(t)}$ i skrivene jedinice h , $\mathbf{W}_{hh}$ je težinska matrica rekurentne veze, a $\mathbf{b}_h$ je pomak odnosno bias.

Nakon što se izračunaju aktivacije skrivenih jedinica u vremenskom trenutku t , aktivacije izlaznih jedinica će biti izračunate na sledeći način:

$$\mathbf{o}^{(t)} = \phi_o(\mathbf{W}_{ho}\mathbf{h}^{(t)} + \mathbf{b}_o) \quad (3.2)$$

gde je $\mathbf{W}_{ho}$ težinska matrica između skrivene jedinice h i izlaznog sloja o , a $\mathbf{b}_o$ je pomak odnosno bias.

Gubitak u vremenskom trenutku t se označava sa $l(\mathbf{o}^{(t)}, \mathbf{y}^{(t)})$, a objektna funkcija L čiji minimum se traži za vremenski period T je predstavljena sledećom formulom:

$$L = \frac{1}{T} \sum_{t=1}^T l(\mathbf{o}^{(t)}, \mathbf{y}^{(t)}) \quad (3.3)$$

Objektna funkcija L zavisi od težinskih matrica $\mathbf{W}_{xh}$ i $\mathbf{W}_{hh}$ vezane za skrivene jedinice $\mathbf{h}^{(1)}, \dots, \mathbf{h}^{(T)}$. Da bismo izračunali gradijente u odnosu na ove parametre, $\partial L / \partial \mathbf{W}_{xh} \in \mathbb{R}^{h \times d}$ i $\partial L / \partial \mathbf{W}_{hh} \in \mathbb{R}^{h \times h}$ primenjujemo lančano pravilo i dobijamo sledeće izraze:

$$\frac{\partial L}{\partial \mathbf{W}_{xh}} = \sum_{t=1}^T \frac{\partial L}{\partial \mathbf{h}^{(t)}} \times \frac{\partial \mathbf{h}^{(t)}}{\partial \mathbf{W}_{xh}} = \sum_{t=1}^T \frac{\partial L}{\partial \mathbf{h}^{(t)}} \mathbf{x}^{(t)T} \quad (3.4)$$

$$\frac{\partial L}{\partial \mathbf{W}_{hh}} = \sum_{t=1}^T \frac{\partial L}{\partial \mathbf{h}^{(t)}} \times \frac{\partial \mathbf{h}^{(t)}}{\partial \mathbf{W}_{hh}} = \sum_{t=1}^T \frac{\partial L}{\partial \mathbf{h}^{(t)}} \mathbf{h}^{(t-1)T} \quad (3.5)$$

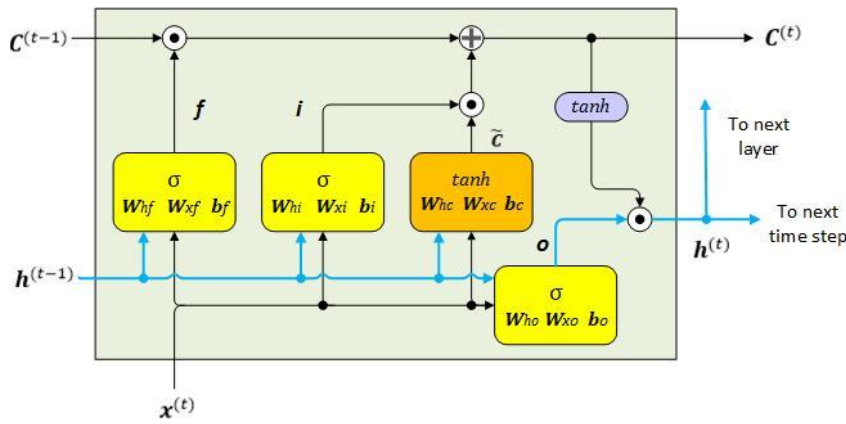
Problem nestajućih gradijenata u RNN-u nastaje usled ponovljenog množenja gradijenata kroz vremenske korake, što može dovesti do njihovog zanemarljivog uticaja pri ažuriranju težina, posebno kod dugih sekvenci ili malih inicijalnih težina (jednačine 3.4 i 3.5). Ovo usporava učenje i može dovesti do gubitka bitnih informacija.

Sa druge strane, problem eksplodirajućih gradijenata nastaje kada parcijalni izvodi postanu preveliki tokom propagacije unazad zbog velikih vrednosti članova težinskih matrica, što sprečava konvergenciju i može izazvati oscilacije RNN oko lokalnog minimuma.

Zbog toga su dizajnirane sledeće dve arhitekture, LSTM i GRU.

3.1 LSTM (Long Short-Term Memory)

LSTM je specijalna vrsta RNN-a dizajnirana za praćenje dugoročnih zavisnosti u sekvencijalnim podacima. Ključna inovacija LSTM-a je mehanizam kapija (*input*, *forget* i *output gate*) koji omogućava mreži da efikasno zadrži ili odbaci informacije tokom dužih sekvenci. LSTM ćelija se sastoji od memorijskih blokova koji održavaju stanje kroz vreme i regulišu protok informacija pomoću nelinearnih kapija. Arhitektura LSTM bloka je prikazana na slici 5.



Slika 5: Arhitektura LSTM bloka

Stanje ćelije $C^{(t)}$ se ažurira pomoću prethodnog stanja $C^{(t-1)}$, pri čemu kapije odlučuju koje informacije će biti sačuvane ili odbačene.

Forget gate (f_t) određuje koje informacije iz prethodnog stanja treba zaboraviti:

$$f_t = \sigma(W_{xf}x^{(t)} + W_{hf}h^{(t-1)} + b_f) \quad (3.6)$$

Input gate (i_t) i vrednost $\tilde{C}_t$ ažuriraju stanje ćelije:

$$i_t = \sigma(W_{xi}x^{(t)} + W_{hi}h^{(t-1)} + b_i) \quad (3.7)$$

$$\tilde{C}_t = \tanh(W_{xc}x^{(t)} + W_{hc}h^{(t-1)} + b_c) \quad (3.8)$$

Novo stanje ćelije C u trenutku t se računa na sledeći način:

$$C^{(t)} = (C^{(t-1)} \odot f_t) \oplus (i_t \odot \tilde{C}_t) \quad (3.9)$$

gde znakovi $\odot$ i $\oplus$ označavaju element-po-element množenje i element-po-element sabiranje, respektivno.

Output gate (o_t) određuje izlazne vrednosti:

$$o_t = \sigma(W_{xo}x^{(t)} + W_{ho}h^{(t-1)} + b_o) \quad (3.10)$$

Skriveno stanje h se računa kao:

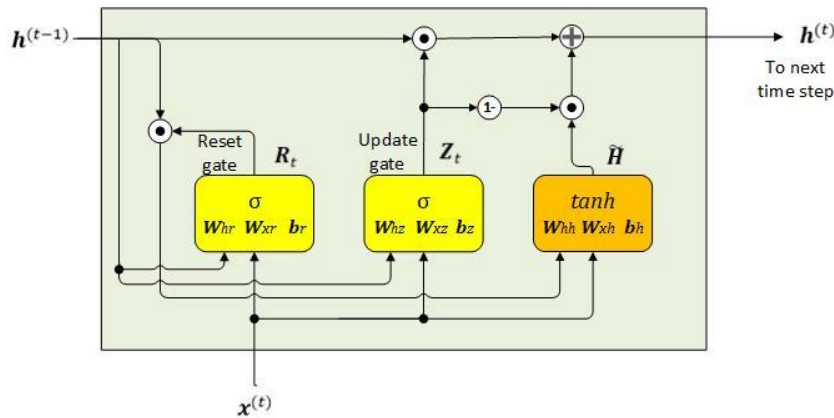
$$h^{(t)} = o_t \odot \tanh(C^{(t)}) \quad (3.11)$$

LSTM arhitektura omogućava stabilnije treniranje i bolju sposobnost praćenja dugoročnih zavisnosti u podacima.

3.2 GRU (Gated Recurrent Unit)

Pošto propagacija unazad u LSTM-u zahteva veliki broj parametara, vreme računanja odnosno vreme obučavanja modela je dugo. U [12] Cho i saradnici su predložili GRU, koji ima manje *gate*-ova od LSTM-a kako bi se skratilo vreme obučavanja. Iako je arhitektura GRU-a pojednostavljena, njegoa efikasnost je uporediva sa LSTM-om.

Slično LSTM-u, GRU koristi *gate*-ove za praćenje dugoročnih zavisnosti i rešavanje problema nestajućih i eksplodirajućih gradijenata. Na slici 6 prikazana je arhitektura GRU bloka.



Slika 6: Arhitektura GRU bloka

GRU jedinica je definisana sledećim jednačinama, gde $\tilde{H}_t$ označava kandidata skrivenog stanja:

$$R_t = \sigma(W_{xr}x^{(t)} + W_{hr}h^{(t-1)} + b_r) \quad (3.12)$$

$$Z_t = \sigma(W_{xz}x^{(t)} + W_{hz}h^{(t-1)} + b_z) \quad (3.13)$$

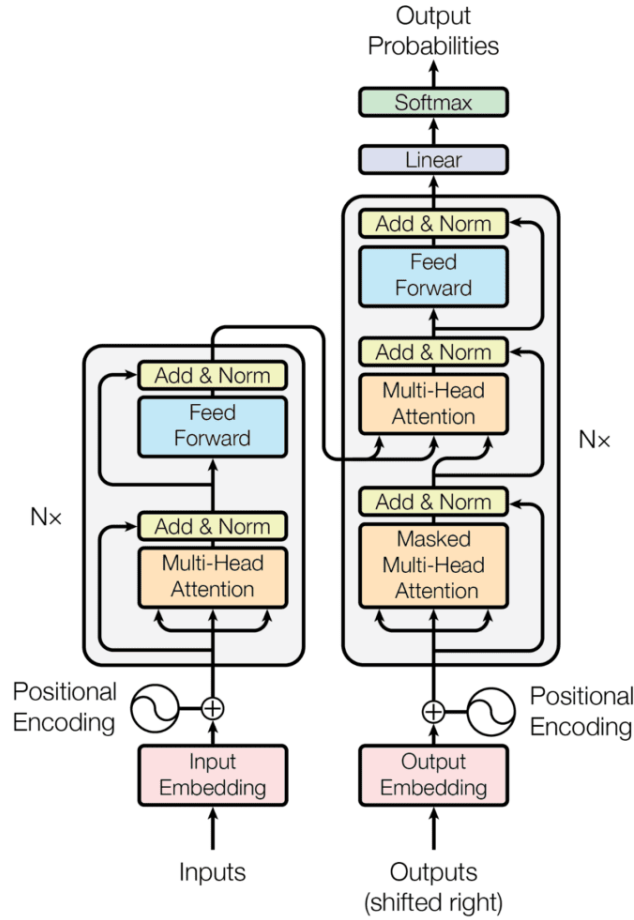
$$\tilde{H}_t = \tanh(W_{xh}x^{(t)} + W_{hh}(R_t \odot h^{(t-1)}) + b_h) \quad (3.14)$$

$$h^{(t)} = (h^{(t-1)} \odot Z_t) \oplus (1 - Z_t) \odot \tilde{H}_t \quad (3.15)$$

4 TRANSFORMER MODELI

Arhitektura Transformera, predstavljena u [13], zasniva se na enkoder i dekoder blokovima (Slika 7).

Iako su rani Transformer modeli eliminisali potrebu za rekurentnim i konvolucionim slojevima, oslanjajući se isključivo na mehanizme pažnje (*self-attention*) za modelovanje veza unutar sekvence, to ne važi za sve njihove varijante.



Slika 7: Arhitektura Transformer modela

U novijim arhitekturama namenjenim obradi vremenskih serija, često se integrišu LSTM slojevi kako bi se efikasnije uhvatile kratkoročne zavisnosti unutar podsekvenci. Takođe, tehnika tokenizacije igra ključnu ulogu u prilagođavanju principa obrade jezika specifičnim zahtevima segmentacije vremenskih podataka. Ovo omogućava brže obučavanje modela, veću paralelizaciju i bolju skalabilnost.

Blok enkodera se sastoji od više identičnih slojeva, koji uključuju mehanizam pažnje sa više glava (*multi-head attention*) i *Feedforward* neuronsku mrežu (FNN). Dekoder je sličan enkoderu, ali sadrži dodatni sloj pažnje između enkodera i dekodera (*encoder-decoder attention layer*), koji omogućava pristup izlazima enkodera.

Self-attention mehanizam računa indekse pažnje (*attention scores*) za sve pozicije u ulaznoj sekvenci omogućavajući modelu da prepozna globalne zavisnosti. Za svaki element ulazne sekvence, model generiše tri vektora: Q - upitni (*Query*) vektor, K - ključ (*Key*) vektor i V - vrednost (*Value*) vektor. Formula za računanje indeksa pažnje:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) V \quad (4.1)$$

gde d_k predstavlja dimenziju vektora K.

Poziciono kodiranje (*Positional encoding*) daje informaciju o redosledu elemenata u sekvenci, jer *self-attention* mehanizam nema vremensku strukturu. *Feedforward* slojevi povećavaju nelinearnost i dubinu modela, čime poboljšavaju sposobnost modelovanja složenih nelinearnih sekvenci. *Multi-head attention* omogućava da model istovremeno analizira različite aspekte interakcije elemenata unutar sekvence. Konačni indeks pažnje se dobija spajanjem i linearnom transformacijom izlaza svih glava.

4.1 PatchTST

PatchTST (*Patch Time-Series Transformer*) [14] je napredni model za analizu vremenskih serija čija je glavna inovacija vremensko *patch*-ovanje (*temporal patching*), gde se duge vremenske sekvence dele na manje vremenske "parčiće" (*patches*), slično kao kod analize slika. Ključne karakteristike PatchTST modela su: segmentacija vremenskih serija (*patching layer*) čiji je zadatak da podeli ulaznu vremensku seriju na *patch*-eve fiksne dužine, linearno ugrađivanje (*linear embedding*) linearno projektuje svaki *patch* u vektorski prostor, standardni transformatorski enkoder i izlazni sloj koji projektuje transformisane *patch*-eve u željene predviđene vrednosti.

5 REZULTATI

Kao što je pokazano u poglavlju 2, modeli su dizajnirani da prihvataju 7 različitih ulaznih karakteristika, uključujući brzinu vetra, standardnu devijaciju brzine vetra, broj obrtaja generatora, temperaturu generatora i rotora. Svaka od ovih karakteristika pruža važne informacije za preciznu predikciju aktivne snage. Za svoje obučavanje, modeli koriste sekvence dužine 360 koraka unazad (petnaest prethodnih dana) kako bi predvideo snagu u sledeća 24 sata (24 koraka unapred). Optimizator koji se koristi je *Adam* sa vrlo malim početnim stepenom učenja (*learning rate*) od $1e^{-5}$. Ovako podešeni parametar omogućava stabilno učenje modela kroz više epoha. Takođe, koristi se *ReduceLROnPlateau scheduler*, koji smanjuje stopu učenja kada model ne pokazuje značajan napredak u minimalizaciji gubitka.

Modeli su trenirani kroz **20 epoha** korišćenjem *Mean Squared Error* (MSE) za ocenu kvaliteta treniranja i procene funkcije gubitka. Uz pomoć *DataLoader-a*, podaci su podeljeni na trening i test skupove, a tokom treninga su se računali različiti metrički pokazatelji kao što su *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), *Mean Absolute Error Index* (MAEI) što je omogućilo detaljno praćenje performansi modela.

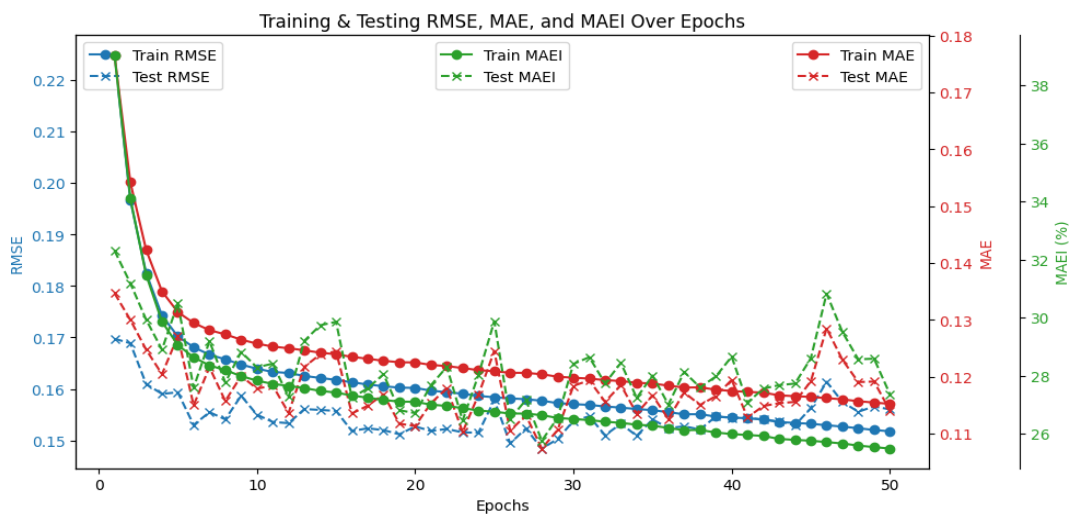
PatchTST model se sastoji od nekoliko ključnih komponenti:

1. *Patch Embedding* (1D Konvolucija): Model koristi 1D konvolucioni sloj kako bi izvukao informaciju iz ulaznih vremenskih serija. Ovaj sloj koristi filtere (kernel veličine 8 sa pomakom od 8), što omogućava efikasno smanjenje dimenzionalnosti podataka kroz lokalne interakcije. Konvolucija igra ključnu ulogu u ekstrakciji prostornih i vremenskih obrazaca, čime model postaje efikasan u obradama vremenskih serija.
2. *Transformer Encoder*: Nakon što su podaci prošli kroz konvolucioni sloj, rezultat se prosleđuje kroz *Transformer Encoder*. Ovaj deo modela koristi višeslojnu arhitekturu zasnovanu na Self-Attention mehanizmima, koji omogućavaju modelu da procesuiru celu vremensku seriju, uzimajući u obzir sve pozicije unutar nje.

Korišćenjem *positional encoding*-a, model prepoznaje i povezuje obrasce koji se ponavljaju kroz vreme, što mu omogućava da identifikuje relevantne informacije sa svih vremenskih tačaka za preciznu predikciju.

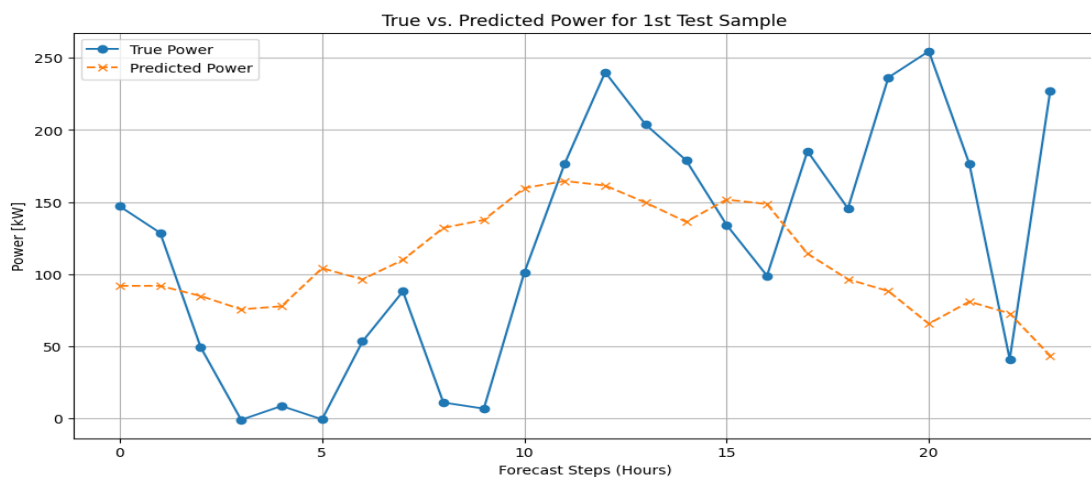
3. *Linearni Sloj za Projekciju*: Na kraju, izlaz iz Transformer enkodera se povezuje sa linearnim slojem za projekciju koji transformiše izlaz u traženi oblik predikcije. Ovaj sloj omogućava modelu da generiše tačne prognoze za buduće vrednosti, u ovom slučaju za 24 vremenska koraka unapred odnosno 24 sata unapred.
4. *Aktivacija*: Za aktivaciju izlaza koristi se *ReLU* funkcija, koja osigurava da izlazni podaci budu nenegativni. To je ključno za problem predikcije snage, jer fizička vrednost snage ne može biti negativna.

Na slici 8 predstavljene su metrike obučavanja PatchTST modela.



Slika 8: Metrike obučavanja PatchTST modela

Na slici 9 prikazane su stvarne izmerene vrednosti aktivne snage i predikcija PatchTST modela.



Slika 9: Izmerene i prediktivne vrednosti PatchTST modela za aktivnu snagu

Na slici 8 prikazane su promene metrika tokom procesa obučavanja kroz epohe. Analizom vrednosti metrika na trening i test skupu može se zaključiti da nije došlo ni do podobučavanja a ni do preobučavanja modela. Dodatno, slika 9 potvrđuje da PatchTST model efikasno generalizuje.

U okviru ovog istraživanja korišćeni su i LSTM i GRU modeli za predikciju aktivne snage vetrogeneratora.

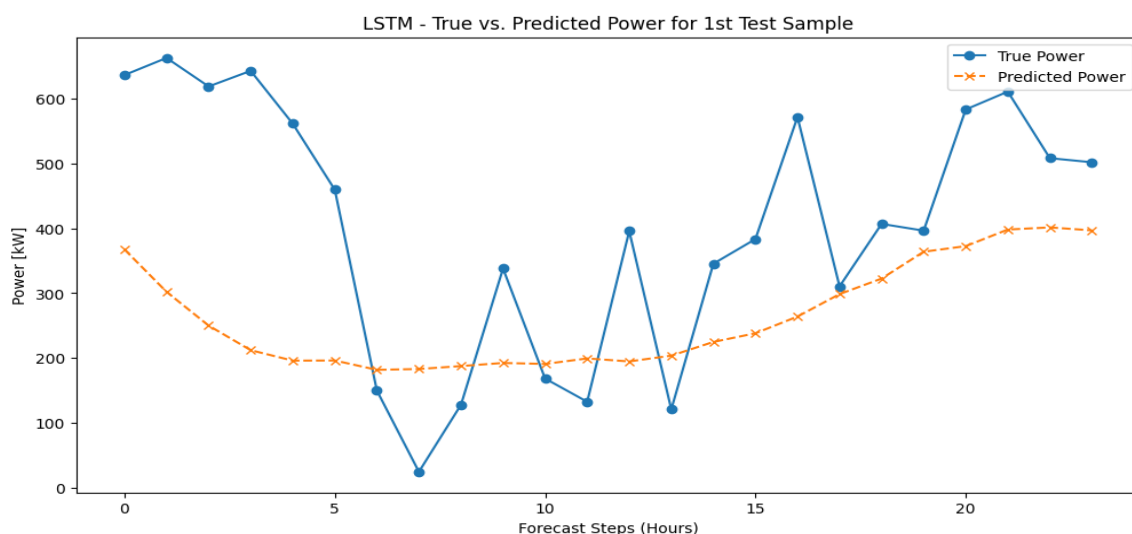
LSTM model korišćen u ovom radu sastoji se od nekoliko ključnih komponenti:

- *Ulazni sloj*: Prilagođen za rad sa sekvencijalnim podacima, gde svaki uzorak obuhvata vremenski prozor od 360 koraka unazad.
- *LSTM slojevi*: Model sadrži dva LSTM sloja sa *hidden_size* od 128 jedinica, koji omogućavaju učenje dugoročnih zavisnosti u podacima.
- *Dropout slojevi*: Kako bi se smanjio rizik od preobučavanja, između LSTM slojeva dodati su dropout slojevi sa stopom od 0.2.
- *Potpuno povezani sloj (Fully Connected Layer)*: Na izlazu, linearni sloj transformiše podatke u vektor dužine 24 za predikciju aktivne snage 24 koraka unapred (24 sata unapred).

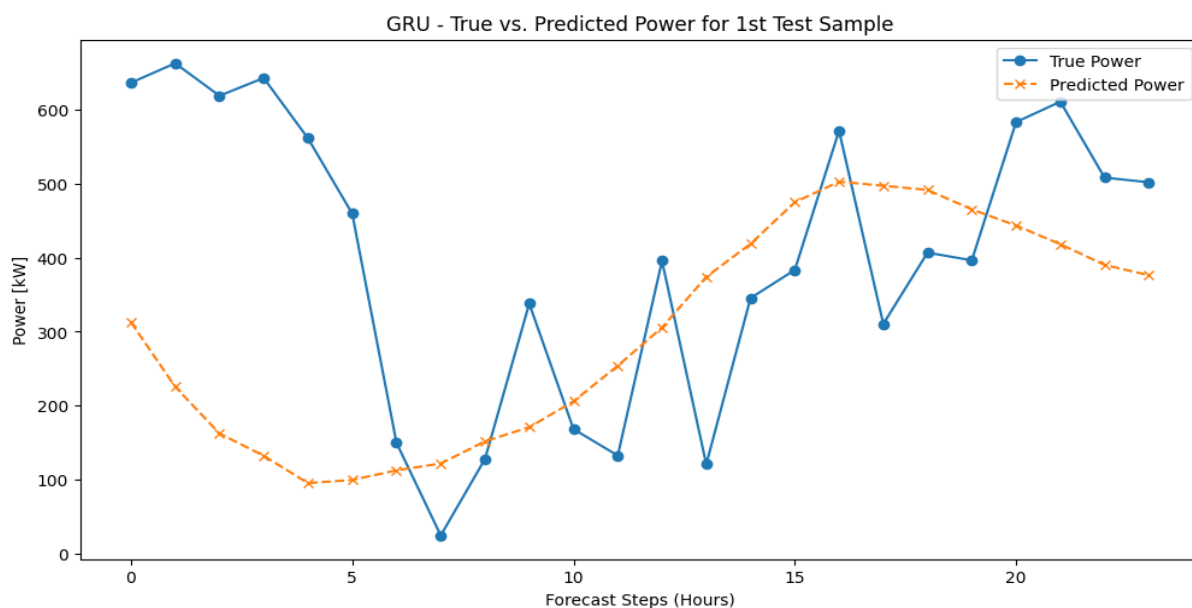
GRU model koristi sličnu arhitekturu kao LSTM, ali sa određenim razlikama:

- *GRU slojevi*: Dva GRU sloja sa *hidden_size* od 128 jedinica.
- *Manji broj parametara*: GRU koristi pojednostavljenu strukturu sa dve vrste vrata (*reset* i *update*), što smanjuje broj parametara i omogućava bržu obuku u poređenju sa LSTM-om.
- *Dropout slojevi*: Implementiran je dropout od 0.2 radi poboljšanja generalizacije modela.
- *Izlazni sloj*: Kao i kod LSTM modela, izlaz je vektor dimenzije 24, koji predviđa aktivnu snagu za sledećih 24 vremenskih koraka.

Na slikama 10 i 11 prikazane su stvarne izmerene vrednosti aktivne snage i predikcije LSTM i GRU modela, respektivno.



Slika 10: Izmerene i prediktivne vrednosti LSTM modela za aktivnu snagu



Slika 11: Izmerene i prediktivne vrednosti GRU modela za aktivnu snagu

Slike 10 i 11 ukazuju na to da modeli zasnovani na RNN takođe postižu dobru sposobnost generalizacije, pri čemu GRU model ostvaruje preciznije rezultate u kratkoročnom predviđanju aktivne snage vetrogeneratora. U Tabeli 1 prikazane su vrednosti metrika na test skupu za sva 3 modela:

Tabela 1: Vrednosti metrika na test skupu sva tri modela

Model	<i>RMSE</i>	<i>MAE</i>	<i>MAEI</i>
PatchTST	0.1545	0.1172	27.82%
GRU	0.2194	0.1676	31.81%
LSTM	0.2197	0.1722	31.97%

6 ZAKLJUČAK

U ovom radu analizirana je kratkoročna predikcija aktivne snage vetrogeneratora korišćenjem tri različita modela mašinskog učenja za predikciju vremenskih serija: PatchTST, LSTM i GRU. Svi modeli su pokazali sposobnost tačne predikcije, pri čemu je PatchTST ostvario najbolje rezultate, nadmašujući ostale modele u pogledu preciznosti i sposobnosti generalizacije. GRU je takođe pokazao dobre performanse, dok je LSTM, iako efikasan, imao nešto slabije rezultate u poređenju s prethodna dva modela. Transformer modeli, za razliku od rekurentnih modela poput LSTM i GRU, pokazuju bolje sposobnosti u predikciji na dužim vremenskim horizontima. Zbog svoje arhitekture, koja omogućava bolje učenje globalnih zavisnosti u podacima, Transformer modeli su efikasniji u obradama dužih sekvenci i mogu pružiti preciznije predikcije, dok rekurentni modeli imaju tendenciju da se oslanjaju na prošle informacije u vremenu, što može ograničiti njihovu tačnost na duže staze. Značaj ovih modela ogleda se u sposobnosti da povezuju kompleksne, nelinearne zavisnosti između mnoštvo varijabli različitog tipa u cilju predikcije unapred definisane veličine.

Dalji pravci istraživanja usmeriće se na evaluaciju modela za dugoročne predikcije, što bi omogućilo bolje planiranje i optimizaciju rada vetrogeneratora.

Takođe, razmatraće se mogućnosti integracije ovih modela u postojeće energetske sisteme, gde bi mogli igrati savetodavnu ulogu u donošenju operativnih odluka.

ZAHVALNICA

Ovaj rad je podržalo Ministarstvo nauke, tehnološkog razvoja i inovacija Republike Srbije kroz Ugovor o realizaciji i finansiranju naučnoistraživačkog rada NIO u 2025. godini.

7 LITERATURA

- [1] IRENA Global Energy Transformation: A Roadmap to 2050 report, April 2019.
- [2] Ivanović, L., Milić, S. D., Sokolović, Ž., & Rakić, A. (2024). Komparativna analiza dubokih neuronskih mreža i algoritama sa pojačanjem gradijenta u dugoročnom predviđanju snage vetra. *Zbornik radova, Elektrotehnički institut "Nikola Tesla"*, 34, 15-36. <https://doi.org/10.5937/zeint34-51258>
- [3] IRENA (2024), Renewable energy statistics 2024, International Renewable Energy Agency, Abu Dhabi.
- [4] Clifton-Smith, M., D. Wood, and A. Wright. "Optimising wind turbine design for operation in low wind speed environments." *Wind Energy Systems*. Woodhead Publishing, 2011. 366-e388.
- [5] Bošnjaković, Mladen, et al. "Wind turbine technology trends." *Applied Sciences* 12.17 (2022): 8653.
- [6] Zhang, H.; Yue, D.; Dou, C.; Li, K.; Hancke, G. Two-Step Wind Power Prediction Approach with Improved Complementary Ensemble Empirical Mode Decomposition and Reinforcement Learning. *IEEE Syst. J.* **2022**, 16, 2545.
- [7] Chu, X.; Bai, W.; Sun, Y.; Li, W.; Liu, C.; Song, H. A Machine Learning-Based Method for Wind Fields Forecasting Utilizing GNSS Radio Occultation Data. *IEEE Access* **2022**, 10, 30258.
- [8] Karaman, Ö.A. Prediction of Wind Power with Machine Learning Models. *Appl. Sci.* **2023**, 13, 11455. <https://doi.org/10.3390/app132011455>
- [9] Liu, T.; Huang, Z.; Tian, L.; Zhu, Y.; Wang, H.; Feng, S. Enhancing Wind Turbine Power Forecast via Convolutional Neural Network. *Electronics* **2021**, 10, 261.
- [10] Solas, M.; Cepeda, N.; Viegas, J.L. Convolutional Neural Network for Short-Term Wind Power Forecasting. In Proceedings of the 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), Bucharest, Romania, 29 September–2 October 2019; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2019; pp. 1–5.
- [11] Byrne, Raymond; MacArtain, Paul (2022), "Vestas V52 Wind Turbine, 10-minute SCADA Data, 2006-2020 - Dundalk Institute of Technology, Ireland", Mendeley Data, V2, doi: 10.17632/tm988rs48k.2

- [12] J. Chung, et al., "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," 2015. [Online]. Available: <https://doi.org/10.48550/arXiv.1412.3555>
- [13] A. Vaswani, et al., "Attention Is All You Need," In the Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS), 2017, Long Beach, CA, USA. <https://doi.org/10.48550/arXiv.1706.03762>
- [14] Y. Nie, N. H. Nguyen, P. Sinthong, J. Kalagnanam, "A time series is worth 64 words: Long-term forecasting with transformers," arXiv:2211.14730, 2022. <https://doi.org/10.48550/arXiv.2211.14730>