

PREDVIĐANJE OPTEREĆENJA SREDNJENAPONSKOG IZVODA UPOTREBOM VEŠTAČKIH NEURALNIH MREŽA I KLASTEROVANJA

THE FORECAST OF MEDIUM-VOLTAGE FEEDER LOAD USING NEURAL NETWORK AND CLUSTERING

Darko ŠOŠIĆ, Univerzitet u Beogradu, Elektrotehnički fakultet, Srbija
Mileta ŽARKOVIĆ, Univerzitet u Beogradu, Elektrotehnički fakultet, Srbija
Goran DOBRIĆ, Univerzitet u Beogradu, Elektrotehnički fakultet, Srbija

KRATAK SADRŽAJ

Ovaj rad predstavlja model za prognozu opterećenja izvoda srednjenaponske distributivne mreže. Pomoću ovog modela se vrši prognoza opterećenja izvoda sa 15-to minutnom rezolucijom. Za kreiranje modela korišćena je neuralna mreža koja je obučavana pomoću meteoroloških podataka koji su relevantni za analiziranu lokaciju, i istorijskih podataka o potrošnji električne energije na posmatranom izvodu. Da bi se smanjila greška prognoze određeni su karakteristični dijagrami opterećenja posmatranih izvoda. Prvu grupu sačinjavaju dijagrami opterećenja koji odgovaraju karakterističnom radnom i neradnom danu za svaki pojedinačni mesec. Druga grupa dijagrama koja je korišćena u obuci neuralne mreže je sačinjena za radni i neradni dan u okviru sezone, pri čemu je godina podeljena na tri sezone (zima, prelazni period i leto). Podela podataka na grupe za obuku, validaciju i testiranje je izvršena nakon klasterovanja *k-mean* metodom, gde je broj klastera odabiran na osnovu visine Dejvis Bouldinovog indeksa. Procena tačnosti prognoze je vršena upotrebom standardnih statističkih mera: MAPE (*Mean Absolute Percentage Error*), RMSE (*Root Mean Square Deviation*), CV-RMSE (*Coefficient of Variation of RMSE*). U ovom radu će biti razmatran jedan srednjenaponski izvod koji napaja potrošače sa centralnim grejanjem.

Ključne reči: prognoza opterećenja izvoda, neuralne mreže, srednjenaponska distributivna mreža, klasterovanje

ABSTRACT

This paper presents a model for forecasting the load of medium-voltage distribution network feeders. The model conducts load forecasting with a 15-minute resolution, employing a neural network. Meteorological data relevant to the analyzed location and data on previous electrical energy demand at the observed feeders were utilized for training the model. To enhance forecasting precision, characteristic load diagrams for the observed feeders were established. The first group comprises load diagrams corresponding to typical working and non-working days for each month. The second group of diagrams used in neural network training includes those for working and non-working days within each season, with the year divided into three seasons (winter, spring/autumn, and summer). Data division into training, validation, and testing groups occurred following clustering using the *k-means* method, with the number of clusters determined based on the Davis Bouldin index. Forecast accuracy was evaluated using standard statistical measures: MAPE (*Mean Absolute Percentage Error*), RMSE (*Root Mean Square Deviation*), CV-RMSE (*Coefficient of Variation of RMSE*). In this paper, one medium-voltage feeder supplying consumers with central heating will be considered.

Key words: Net load forecasts, Artificial Neural Network (ANN), Medium-voltage distribution network, *k-means* clustering

1. UVOD

Razvoj naprednih metoda za procenu stanja distributivnog sistema postaje sve značajniji zbog sve većeg prisustva distribuirane proizvodnje. Današnje distributivne mreže, posebno na nižim naponskim nivoima, često nisu opremljene dovoljnim brojem uređaja za daljinsko očitavanje, što zahteva inovativne pristupe za adekvatno praćenje i kontrolisanje stanja sistema. To stvara potrebu za naprednim metodama procene stanja distributivnog sistema na osnovu oskudnih i niskokvalitetnih merenja. Izrada modela mreže samo na osnovu podataka iz prošlosti može prouzrokovati ozbiljna odstupanja kada mreža pređe iz pasivne u aktivnu, odnosno kada se poveća udeo proizvođača u njoj [1]. Veštačke neuralne mreže (Artificial Neural Network - ANN) su se pokazale kao efikasan alat u proceni i predviđanju različitih aspekata rada distributivnih mreža. Na primer, u radu [2] testirane su tri različite neuralne mreže za procenu stacionarnih parametara pouzdanosti. Neuralne mreže su se koristile za procenu amplituda i faznih pomaka napona u čvorovima srednjenaponske distribucione mreže u kojima nisu bila dostupna merenja [3], dok su u [4] korišćene za kratkoročnu satnu prognozu opterećenja. Nadalje, u [5,6] različite vrste neuralnih mreža su korišćene za prognozu opterećenja srednjenaponskih izvoda u prisustvu distribuirane proizvodnje. U drugim istraživanjima neuralne mreže korišćene su za: određivanje redosleda prekidačkih operacija [7], procenu lokalnih cena prodaje električne energije [8], predviđanje graničnih vrednosti opterećenja na 30-to minutnom intervalu [9], procenu dnevne potrošnje domaćinstava koja su u kategoriji kupca-proizvođača [10], predviđanje potrošnje toplinskih pumpi na dnevnom nivou [11], predviđanje udela difuznog i reflektovanog zračenja na osnovu merenja ukupne iradijacije [12].

Sa povećanom izgradnjom solarnih sistema kod dosadašnjih potrošača neminovno dolazi do promena koje sve više zabrinjavaju operatore distributivnog sistema. U jutarnjim časovima će se javiti povećani upliv energije iz distribuirane proizvodnje koji neće moći lokalno da se potroši, što će dalje prouzrokovati postojanje kontra tokova u primarnoj srednjenaponskoj distributivnoj mreži. S druge strane, ni suprotni efekat koji se javlja uveče, kada ponestane resursa, nije zanemarljiv posebno zbog očekivanog porasta opterećenja u večernjim časovima kada se poveća broj električnih vozila. Stoga je neophodno napraviti odgovarajući model koji bi uverljivo simulirao ponašanje srednjenaponskog izvoda u odgovarajućim situacijama.

Kratkoročno predviđanje (od jednog do nekoliko sati unapred) korisno je za planiranje ekonomskog rasporeda opterećenja i sigurnosti rada na tržištu električne energije [13], ali i za optimizaciju rada skladišta električne energije [14]. Predviđanje unapred za jednu nedelju (srednjoročno) pogodno je za odluke o angažovanju većih proizvodnih jedinica, i upravljanju rezervama [15].

U [16], predviđanje je definisano kao izračunavanje ili predviđanje budućeg događaja ili uslova na osnovu analize istorijskih podataka, događaja ili posmatranja. Jedan vremenski trenutak zahteva pažljivu analizu kada se govori o algoritmima za predviđanje. Tu se prvenstveno misli na izbor ulaznih podataka koji će činiti model, a koji moraju da imaju visok stepen korelacije sa izlaznom veličinom modela. Međutim, veoma je teško tačno predvideti električno opterećenje, jer ono zavisi od različitih promenljivih, kao što su društveni faktori, ekonomske situacije, nepravilna ponašanja i drugi nelinearni faktori [17].

U ovom radu razmatraće se kratkoročna prognoza opterećenja na nivou izvoda. Razlog za posmatranje celog izvoda je činjenica da se opterećenje pojedinačnih domaćinstava veoma često i stohastički menja, tako da je izuzetno teško bez dodatnih senzora i merenja predvideti izgled takvog dijagrama opterećenja. Sa druge strane, pošto na izvodu postoji veliki broj potrošača sličnog stepena elektrifikacije moguće je izvršiti agregaciju opterećenja i na taj način kreirati relativno glatku krivu koju je moguće predvideti. Sa napretkom tehnologije postoji verovatnoća da će i predviđanje ponašanja pojedinačnih potrošača biti moguće u realnom vremenu, što dalje otvara mogućnosti za bolje planiranje rada distributivnog sistema, ali to je još uvek u budućnosti. Što se tiče kratkoročnog predviđanja opterećenja izvoda, njega je moguće vršiti na osnovu dostupnih podataka iz bliske i dalje prošlosti. Osnovne prednosti ovog rada su:

- binarna segmentacija i filtriranje podataka,
- primena *K-mean* metode za optimalno klasterovanje podataka u svrhu lakšeg obučavanja ANN,
- odabir optimalnih ulaza za obuku ANN.

Nakon uvodnih razmatranja, u drugom delu rada su opisani alati koji su korišćeni za filtriranje podataka. U trećem poglavlju rada opisana je metoda za odabir broja klastera. U četvrtom delu navedena je korišćena metrika za procenu rezultata sprovedene prognoze, dok je u petom poglavlju dat osnovni opis genetičkog algoritma i neuralnih mreža. Šesto poglavlje predstavlja rezultate primenjenih metoda i komparaciju grešaka.

2. FILTRIRANJE PODATAKA

Podaci koji se mere u realnom vremenu često su zaprljani šumom i velikim odstupanjima. Odstupanja se odnose na grubu grešku merenja [19], odnosno mogu se definisati kao podaci koji su značajno neusaglašeni sa preostalim podacima. U zavisnosti od konteksta, proporcija grubih grešaka u podacima je između 0,1% do 10% [20]. Ovde će biti opisane dve metode koje se koriste za filtriranje takvih odstupanja. Da bi se primenila metoda

prepoznavanja odstupanja potrebno je prvo posmatranu veličinu podeliti na segmente. Za tu svrhu će se koristiti „pohlepna“ binarna segmentacija.

Prvo je potrebno objasniti usvojenu notaciju obeležavanja signala. Za razmatrani signal $y = \{y_t\}_{t=1}^T$, odsečak signala čije su granice odbirci a i b , $\{y_t\}_{t=a+1}^b$ ($1 \leq a < b \leq T$) može se jednostavnije napisati kao $y_{a..b}$. Neka je sa $\tau = \{t_1, \dots, t_k\}$ obeležen skup tačaka koje predstavljaju indekse odbiraka, odnosno mogu uzimati vrednosti iz opsega $[1, T]$. Kardinalni broj ovog skupa je u stvari broj članova skupa, odnosno $|\tau| = K$. Ovde se smatra da su dostupna i dva fiktivna indeksa $t_0 = 0$ i $t_{k+1} = T$.

Neka se razmatra multi-varijabilni nestacionarni slučajni proces $y = \{y_1, \dots, y_T\}$ koji uzima vrednosti u R^d ($d \geq 1$) i ima T odbiraka. Usvaja se pretpostavka da je signal y parcijalno nestacionaran, odnosno postoje tačke promena pri kojima se neke karakteristike (medijana, srednja vrednost, varijansa) drastično menjaju. Detekcija tačaka promene sastoji se u proceni indeksa t^*_k , pri čemu važi $t^*_1 < t^*_2 < \dots < t^*_K$. U zavisnosti od problema koji se razmatra ukupan broj tačaka promene, K , može ili ne mora biti poznat. Pošto u praksi ovaj broj najčešće nije poznat potrebno ga je proceniti.

Određivanje tačaka promene se može definisati kao problem izbora modela, što podrazumeva izbor najbolje moguće segmentacije τ u skladu sa kriterijumskom funkcijom $V(\tau, y)$ koji treba minimizovati. U ovom radu usvojena je kriterijumska funkcija koja predstavlja zbir funkcija troškova svakog segmenta, odnosno:

$$V(\tau, y) = \sum_{k=0}^K c(y_{t_k..t_{k+1}}), \quad (1)$$

$$c(y_{a..b}) = \sum_{t=a+1}^b |y_t - \hat{y}_{a..b}|,$$

pri čemu je $\hat{y}_{a..b}$ medijan odsečka signala. Binarna podela počinje podelom celog signala na dva segmenta nalaženjem prve tačke podele. Nakon toga proces se nastavlja sa novonastalim segmentima. Da bi se sprečila prevelika podela potrebno je uvesti neko ograničenje. Jedno od ograničenja je broj odbiraka u segmentu. Naime pošto se u ovom radu razmatra promena energije usvojeno je da jedan segment mora da sadrži najmanje jedan ceo dan. Druga provera je zadata sledećim uslovom [18]:

$$c(y_{1..r}) + c(y_{r+1..n}) + 4 \log n < c(y_{1..n}) \quad (2)$$

gde je $r \in [1, n]$ broj odbirka u posmatranom segmentu. U slučaju da nije zadovoljen bilo koji od uslova posmatrani segment se neće deliti. Predloženi algoritam je prikazan pomoću pseudo koda.

Inicijalizacija: signal $\{y_t\}_{t=1}^T$, funkcija troškova $c(\cdot)$, kriterijum zaustavljanja, $L = \{\}$.

repeat

$k = |L|$

$t_0 = 0$ i $t_{k+1} = T$

if $k > 0$ **then**

 Presložiti L u rastućem poretku

end if

 Inicijalizovati niz G dimenzija $(k+1 \times 1)$

for $i = 0, \dots, k$ **do**

$$G(i) = c(y_{t_i..y_{t_{i+1}}}) - \min_{t_i < t < t_{i+1}} \left(c(y_{t_i..y_t}) + c(y_{t..y_{t_{i+1}}}) \right)$$

end for

$\hat{i} = \text{argmax}_i G(i)$

$$t = \text{argmin}_{t_i < t < t_{i+1}} \left(c(y_{t_i..t}) + c(y_{t..t_{i+1}}) \right)$$

$L = L \cup \{t\}$

until kriterijum zaustavljanja

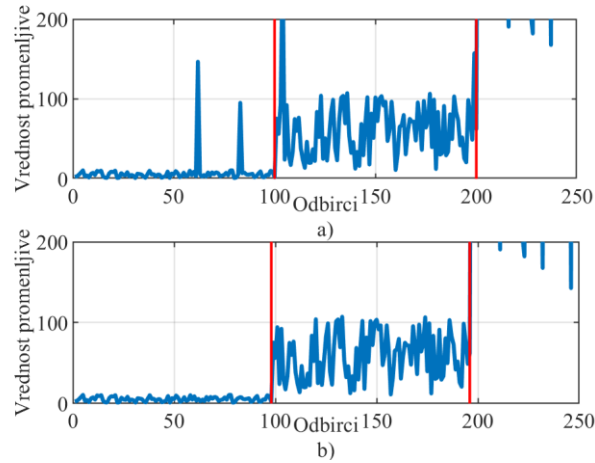
Izlaz: L predstavlja skup tačaka promene

Sada je potrebno razmotriti metode za prepoznavanje odstupanja. Generalno gledano postoje dva standardna pristupa za postupanje s odstupanjima u problemima regresije: dijagnostički pristup i robusna regresija. Dijagnostički pristup identifikuje i uklanja odstupanja iz podataka pre nego što prilagodi model očišćenim podacima, dok robusni pristup prvo prilagođava model celom skupu podataka, a zatim identifikuje izuzetke kao one podatke koji imaju velike ostatke.

Tukejevo pravilo odnosi se na robusnu metodu primenjenu za identifikaciju odstupanja putem *boxplot* vizuelizacije. Potencijalna odstupanja se obeležavaju na osnovu gornjih i donjih granica koje su povezane s karakteristikama skupa merenja. Pomoću prve (q_1) i treće granice (q_3) određuje se širina centralnog dela podataka, odnosno dobija se odgovarajući raspon od $IR = q_3 - q_1$ [21]. Na ovaj način su definisani podaci od interesa, dok se donja i gornja granica određuju pomoću sledećeg izraza:

$$\begin{aligned} Lb &= q_1 - 1.5IR, \\ Ub &= q_3 + 1.5IR. \end{aligned} \quad (3)$$

Vrednost 1,5 se može menjati, a detaljnije objašnjenje se može potražiti u literaturi [22]. Za vrednosti q_1 i q_3 u ovom radu su odabrane vrednosti od 5% i 95%. Primena segmentacije i detekcije odstupanja je ilustrovano na Slici 1.



Slika 1 – a) Primer binarne segmentacije, b) Detekcije odstupanja

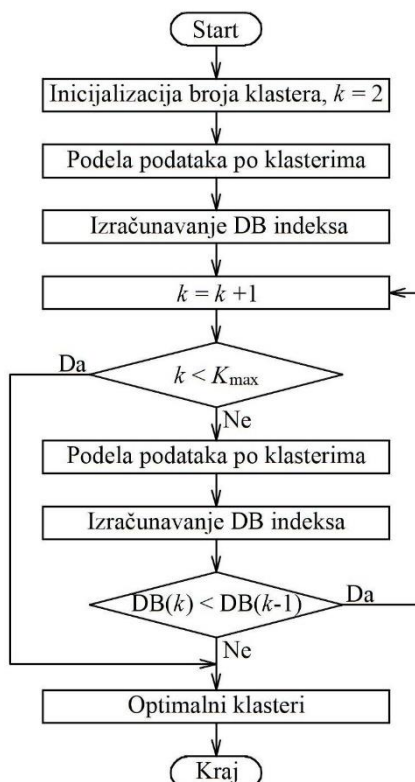
3. ODABIR BROJA KLASTERA

Prilikom obučavanja neuralnih mreža uvek je bolje dodatno obraditi ulazne podatke. Zbog činjenice da se radi sa većim brojem ulaznih podataka ne postoji praktičan način prikazivanja svih podataka da bi se uočila neka zajednička karakteristika. Jedna od mogućnosti koja se najčešće primenjuje je klasterovanje podataka, odnosno podela podataka u slične grupe. U ovom radu za klasterovanje je odabrana K -mean metoda. Međutim, K -mean metoda zahteva zadavanje ukupnog broja klastera na koje će se podeliti ulazni podaci. Optimalan broj klastera zavisi od mnogo čega i u praksi ga je teško unapred zadati, zato je u ovom radu primenjena adaptivni način klasterovanja koji je predložen u [23]. U ovoj metodi ukupan broj klastera se određuje na osnovu visine Dejvis-Bouldinog indeksa koji se određuje pomoću sledećeg izraza:

$$\begin{aligned} S(i) &= \frac{1}{|C_i|} \sum_{z_j \in C_i} \|z_j - A_i\|, \\ DB &= \frac{1}{k} \sum_{i=1}^k \max_{j=1:k, j \neq i} \frac{S(i) + S(j)}{\|A_i - A_j\|}. \end{aligned} \quad (4)$$

U ovom izrazu C_i predstavlja i -ti klaster, z_j je vektor u i -tom klasteru, dok je A_i centroid i -tog klastera. Na osnovu svega navedenog $S(i)$ predstavlja meru raštrkanosti unutar posmatranog klastera. Pomoću DB indeksa uvažava se međusobna raznolikost susednih klastera, ali i sličnost unutrašnjih članova svakog pojedinačnog klastera. Za optimalan broj klastera se proglašava onaj broj klastera za koji se dobija najniža vrednost DB indeksa. Upravo na osnovu ovoga je kreiran adaptivni algoritam odabira broja klastera [24], koji je prikazan na Slici 2.

Pošto se na ovaj način ulazni podaci dele u optimalan broj grupa pretpostavlja se da će se time olakšati i ubrzati rad, odnosno proces obuke, neuralnih mreža.



Slika 2 – Adaptivni algoritam odabira broja klastera

4. GREŠKE PROGNOZE

U ovom radu koristile su se tri metrike, srednja apsolutna procentualna greška (MAPE), srednje kvadratno odstupanje i koeficijent varijacije srednjeg kvadratnog odstupanja (CVRMSE). MAPE se koristi za procenu relativne devijacije rezultata prognoze od stvarnih vrednosti, RMSE meri kolika je koncentracija podataka oko najbolje prilagođene krive, dok CVRMSE opisuje prilagodljivost modela prognoze. Što je niža vrednost CVRMSE, manji su ostaci u odnosu na prognoziranu vrednost, što sugeriše fleksibilniju prilagodljivost modela. Ove tri metrike se računaju na sledeći način:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|P_i^F - P_i^R|}{P_i^R} \cdot 100\%, \quad (5)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i^F - P_i^R)^2}, \quad (6)$$

$$CVRMSE = \frac{\sqrt{\sum_{i=1}^N (P_i^F - P_i^R)^2}}{\sum_{i=1}^N P_i^R / N} \quad (7)$$

gde P_i^F predstavlja prognozirano opterećenje izvoda, a P_i^R stvarno opterećenje izvoda.

5. VEŠTAČKA INTELIGENCIJA

Veštačka inteligencija (*Artificial Intelligence* - AI) predstavlja vid efikasnog korišćenja računarskih sistema i velikog broja podataka u svrhu uočavanja zavisnosti različitih parametara na vremenske serije različitih podataka u elektroenergetici. Različiti AI algoritmi mogu da predstavljaju veoma koristan alat u distributivnim elektroenergetskim sistemima kada inženjeri nisu u mogućnosti da sagledaju obimne količine podataka. Takvi alati se odnose i na predikciju vremenskih serija podataka, pa je u ovom radu AI iskorišćena za predikciju potrošnje srednjenaponskog izvoda. U ovom radu genetički algoritam se koristi za određivanje optimalnih ulaza za obuku ANN. Algoritam na optimalan način formira što bolju bazu ulaznih podataka koja se kasnije koristi za što uspešnije obučavanje ANN. Na taj način ekspertima u oblasti je olakšan odabir i priprema relevantne baze podataka.

5.1 Genetički algoritam

Genetički algoritam je vrsta evolutivnog algoritma inspirisanog procesima prirodne selekcije i genetskog nasleđivanja. Koristi se za rešavanje optimizacionih problema, posebno u situacijama kada je tradicionalne algoritme teško ili nemoguće primeniti, ili kada je dovoljno pronaći približno optimalno rešenje.

Osnovna ideja genetičkog algoritma je simulacija procesa evolucije. Algoritam počinje sa populacijom jedinki, koje predstavljaju potencijalna rešenja problema. Svaka jedinka je predstavljena nizom parametara ili genetskim kodom koji se mogu menjati tokom evolucije.

Proces evolucije se odvija kroz iterativno ponavljanje sledećih koraka:

1. **Selekcija:** Odabir roditelja iz trenutne populacije za reprodukciju, u ovom slučaju je korišteno pravilo točka ruleta. Kod ovog pravila verovatnoća selekcije zavisi od prilagođenosti jedinki, tj. koliko je dobro svaka jedinka prilagođena ciljnoj funkciji.
2. **Reprodukcija:** Generisanje potomaka odabranih roditelja. To se obično postiže kombinovanjem genetskih informacija roditelja, što može uključivati ukrštanje (kombinovanje delova genetskog koda roditelja) i mutaciju (slučajno menjanje genetskog koda sa veoma malom verovatnoćom).
3. **Evaluacija:** Ocena prilagođenosti svakog potomka, odnosno koliko dobro rešenje odgovara ciljnoj funkciji.
4. **Selekcija za preživljavanje:** Odabir jedinki koje će preživeti u sledećoj generaciji.

Ovi koraci se ponavljaju kroz niz generacija sve dok se ne ispuni određeni uslov zaustavljanja, kao što je dostizanje maksimalnog broja generacija ili postizanje zadovoljavajućeg rešenja.

5.2 Neuralne mreže

Neuralne mreže su računarski modeli inspirisani strukturom ljudskog mozga, koji se koriste za obradu podataka i rešavanje različitih problema. Jedna od najčešće korišćenih arhitektura neuralnih mreža je *feedforward* neuralna mreža sa *backpropagation* algoritmom.

U ovoj arhitekturi, podaci se propuštaju kroz mrežu od ulaza do izlaza bez ikakvih povratnih veza. Svaki neuron u mreži prima ulazne podatke, obrađuje ih i prosleđuje ih dalje sledećem sloju neurona. Konačni izlaz mreže se koristi za rešavanje zadatog problema, u ovom slučaju procena opterećenja srednjenaponskog kabla.

Backpropagation je tehnika koja se koristi za treniranje neuralnih mreža. Ona se sastoji od iterativnog procesa prilagođavanja težina veza između neurona kako bi se minimizovala greška između stvarnih i predviđenih izlaza. Ovaj proces se ponavlja kroz više iteracija sve dok se ne postigne željeni nivo tačnosti.

U ovom konkretnom slučaju, neuralna mreža sa *feedforward* arhitekturom i *backpropagation* algoritmom koristi se za procenu opterećenja srednjenaponskog kabla. Ulazni podaci uključuju datum (dan u nedelji, nedelja u godini, mesec, sat, četvrtina sata), meteorološke podatke (minimalna, maksimalna i srednja dnevna temperatura) i istorijske podatke o opterećenju kabla. Neuralna mreža obrađuje ove podatke i generiše procenu opterećenja kabla kao izlaz. Ovaj model može biti koristan za predviđanje opterećenja kabla u različitim vremenskim uslovima i pomaže u planiranju i upravljanju elektroenergetskim sistemima.

Da bi neuralne mreže radile potrebno je definisati određeni broj parametara. Za aktivacionu funkciju neurona u skrivenom sloju odabrana je tangens hiperbolička funkcija kako bi se poboljšala aproksimacija nelinearnih zavisnosti, dok je za izlazni sloj odabrana linearna funkcija. U cilju ubrzanog smanjivanja srednje kvadratne greške korištena je Levenberg–Marquardt metoda. Odlučeno je da neuralna mreža ima jedan skriveni sloj pošto je on dovoljan za aproksimaciju funkcija skoro svakog oblika, dok bi se povećanjem broja skrivenih slojeva model previše usmerio na postojeće podatke i ne bi bio u stanju da prevaziđe nedostatak nekog parametra ili veće odstupanje pojedinih ulaznih veličina. Da bi se odredio broj neurona u skrivenom sloju korišten je sledeći izraz [25]:

$$N_n = \frac{2}{3} N_i + N_o, \quad (8)$$

gde je sa N_n označen broj neurona, sa N_i i N_o broj ulaza i izlaza. Da bi se ubrzao proces konvergencije obuke neuralne mreže ulazni podaci su podeljeni u tri grupe. Prvu grupu su sačinjavali podaci za obuku (60%) na osnovu kojih su podešavani parametri neuralne mreže. Drugu grupu su sačinjavali podaci za validaciju (20%). Ovi podaci nisu korišteni za obuku mreže nego samo kao kriterijum zaustavljanja. Naime kada ukupna greška procene prestane da opada i počne da raste potrebno je zaustaviti proces obuke. Poslednju grupu sačinjavaju podaci za testiranje (20%) na osnovu kojih se daje konačna ocena kreiranog modela.

6. REZULTATI

Kod kreiranja modela srednjenaponskog izvoda potrebno je odabrati ulazne podatke koji će i u budućnosti biti dostupni. Ponekad je slučaj da postoji veliki broj različitih podataka koji bi mogli da se koriste za obuku. Međutim, nije potrebno uvažavati sve trenutno dostupne veličine pogotovo ako željena veličina, u ovom slučaju

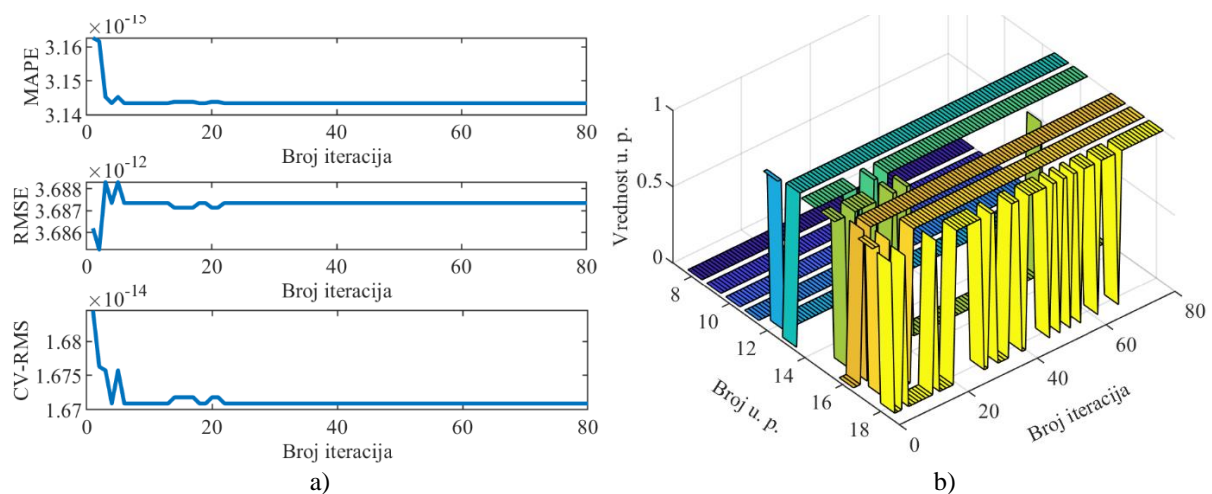
opterećenje izvoda, malo ili ni malo ne zavise od posmatrane ulazne veličine. Često puta se u praksi pokazalo da relativno jednostavno rešenje može da daje odlične rezultate. U literaturi je moguće pronaći radove koji su predviđali vršno opterećenje izvoda na osnovu meteoroloških podataka (temperatura, relativna vlažnost vazduha) [26], podataka iz bliske i dalje prošlosti [9], ili upotrebom karakterističnog dana [6]. U Tabeli 1 prikazani su podaci sa kojima je sprovedena opisana procedura, pri čemu je na raspolaganju bilo 96300 parova podataka.

Tabela 1 – Raspoloživi podaci

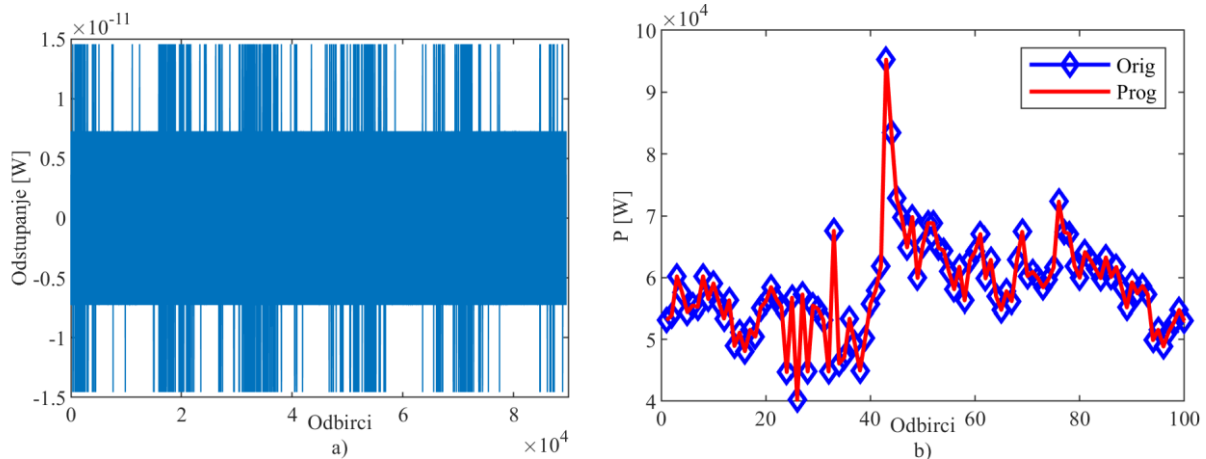
Broj ulaza	Opis	Broj ulaza	Opis
1	Redni broj meseca	10	P_{i-45m}
2	Redni broj dana u mesecu	11	P_{i-7d}
3	Sat u danu	12	P_{i-14d}
4	Broj minuta	13	P_{i-21d}
5	Redni broj nedelje	14	Srednja dnevna temperatura
6	Redni broj dana u nedelji	15	Maksimalna dnevna temperatura
7	Aktivna sanga (P_i)	16	Minimalna dnevna temperatura
8	Snaga iz prethodnog intervala (P_{i-15m})	17	Karakterističan dan u mesecu
9	P_{i-30m}	18	Karakterističan dan u sezoni

Što se tiče karakterističnih dana oni su određivano posebno za radni, a posebno za neradni dan. Ovi dani su određivani na osnovu najmanjeg kvadratnog odstupanja odabranog dana od svih ostalih dana iz posmatranog skupa. Prilikom obuke neuralnih mreža sedmi podatak je korišćen kao izlaz, tako da je te podatke trebalo odstraniti sa ulaza. Zbog velikog broja ulaznih promenljivih, kao i bojazni da optimizacioni algoritam neće dati dobro rešenje u prihvatljivom vremenskom intervalu, odlučeno je da se prvih 6 veličina smatra uvek dostupnim. Razlog za ovo leži u jednostavnoj činjenici da će kalendarski i hronološki podaci uvek biti dostupni. Ostali podaci koji su navedeni u Tabeli 1 su korišteni kao upravljačke promenljive u optimizacionom procesu.

Nakon kreiranja potencijalnog rešenja, svaka veličina kojoj je dodeljena vrednost 0 je izostavljena iz procesa obuke. Prvo je za definisani skup ulaznih veličina određivan broj klastera, nakon čega je iz svakog klastera uziman unapred definisani procenat jedinki za obuku, validaciju i testiranje. Upravo zbog ove činjenice može se desiti da za dva identična vektora rešenja budu kreirani drugačiji klasteri, a samim tim i neuralne mreže. U ovom radu je korištena jednokriterijumska optimizacija. Za kriterijumsku funkciju odabrana je srednja apsolutna procentualna greška. Sa Slike 3 a) se može primetiti da se ova greška (MAPE) postepeno smanjuje, pri čemu zbog osobine genetičkog algoritma dolazi do trenutnog gubitka najboljeg rešenja oko dvadesete iteracije. Poređenja radi na dodatnim dijagramima su prikazane i promene druge dve greške najboljeg rešenja sa protokom iteracija. Posmatranjem samo ovih krivih moglo bi se pretpostaviti da je do zaustavljanja algoritma moglo doći i u nekoj ranijoj iteraciji, i da nije bilo potrebno forsirati proračun do maksimalnog dozvoljenog broja iteracija koji je u ovom slučaju iznosio 80. Međutim, ako bi se pogledala promena upravljačkih promenljivih najboljeg rešenja, Slika 3 b), došlo bi se do zaključka da je dolazilo do intenzivnih promena i da bi zaustavljanje u nekoj ranijoj iteraciji doveo do lošijeg rešenja. Odstupanje prognozirane veličine od stvarne vrednosti je prikazano na Slici 4 a), dok je na Slici 4 b) radi lakšeg uočavanja prikazana stvarna i prognozirana vrednost opterećenja na manjem vremenskom intervalu. Kao što se može primetiti zbog male greške došlo je do idealnog preklapanja krivih, i one se mogu razlikovati jedino pomoću prikazanih markera.



Slika 3 – a) Promena greški, b) Promena najboljeg rešenja po iteracijama



Slika 4 – a) Odstupanje prognozirane vrednosti od stvarne vrednosti, b) Stvarna i prognozirana vrednost

Prema prikazanim rezultatima opterećenje voda se može pored kalendarskih i hronoloških podataka vršiti pomoću karakterističnog dana na mesečnom i sezonskom nivou, najmanje i srednje dnevne temperature i pomoću krive opterećenja od pre tri nedelje. Stoga, u skrivenom sloju je bilo ukupno 8 neurona.

7. ZAKLJUČAK

U istraživanju je razvijen efikasan model za predviđanje opterećenja srednjenaponskog izvoda distributivnog sistema. Korišćenjem ograničenog broja ulaznih promenljivih, uključujući kalendarske i hronološke podatke, kao i karakteristične dane i minimalne i srednje dnevne temperature, postignuti su zadovoljavajući rezultati u predviđanju opterećenja. Za odabir ulaznih veličina neuralne mreže korišten je genetički algoritam. Pokazalo se da u slučaju realnog izvoda koji napaja potrošače koji imaju centralno grejanje odstupanje prognozirane vrednosti od stvarne vrednosti ne prelazi 1%, što ukazuje na njegovu promenljivost u praksi za planiranje i upravljanje distributivnim sistemom. U svrhu formiranja alata moguće je grafički interfejs dizajnirati gde bi se korisniku omogućilo učitavanje dostupnih podataka i primena razvijene metode. Dalji razvoj modela predstavlja primenu za prognozu različitih vremenskih serija podataka, distribuirane proizvodnje energije i drugih.

LITERATURA

- [1] Pertl M, Heussen K, Gehrke O, Rezkalla M, 2016, "Voltage Estimation in Active Distribution Grids Using Neural Networks", 2016 IEEE Power and Energy Society General Meeting (PESGM), Boston, USA
- [2] Weng G, Zhu S, Gong Y, Ma T, Xie F, Fang M, 2017, "Research on Power Quality Prediction for DG Integrated Smart Grid Based on Neural Network", 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China
- [3] Hong G, Kim Y-S, 2020, "Supervised Learning Approach for State Estimation of Unmeasured Points of Distribution Network", IEEE Access, Vol. 8, pp. 113918 – 113931
- [4] Faiazy M, Ebtehaj M, 2013, "Short Term Load Prediction of a Distribution Network based on an Artificial Intelligent Method", 22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013), Stockholm
- [5] Tziolis G, Spanias C, Theodoride M, Theocharides S, Lopez-Lorente J, Livera A, Makrides G, Georghiou GE, 2023, "Short-term electric net load forecasting for solar-integrated distribution systems based on Bayesian neural networks and statistical post-processing", Energy, Vol 271, , 127018
- [6] Chu Y, Pedro H, Kaur A, Kleissl J, Coimbra C, 2017, "Net load forecasts for solar-integrated operational grid feeders", Solar Energy, Vol. 158, pp. 236-246
- [7] Išlić M, Sučić S, Havelka J, Marušić A, 2020, "0" Sustainable Energy, Grids and Networks, Volume 22, 100331
- [8] Liu B, Dong J, Lian J, Kuruganti T, Wang X, Li F, 2023, "Enhanced deep neural networks with transfer learning for distribution LMP considering load and PV uncertainties", Electrical Power and Energy Systems, Vol. 147, 108780
- [9] Berrisch J, Narajewski M, Ziel F, 2023, "High-resolution peak demand estimation using generalized additive models and deep neural networks", Energy and AI, vol. 13, 100236
- [10] Kobylinski P, Wierzbowski M, Piotrowski K, 2020, "High-resolution net load forecasting for micro-neighbourhoods with high penetration of renewable energy sources", Electrical Power and Energy Systems, Vol. 117, 105635

- [11] Cao Z, Han X, Lyons W, O'Rourke F, 2021, "Energy management optimisation using a combined Long Short-Term Memory recurrent neural network – Particle Swarm Optimisation model", *Journal of Cleaner Production*, Vol. 326, 129246
- [12] Schlager E, Feichtinger G, Gursch H, 2023, "Development and comparison of local solar split models on the example of Central Europe", *Energy and AI*, Vol. 12, 100226
- [13] Kakhki IN, Taherian H, Aghaebrahimi MR, 2013, "Short-Term Price Forecasting Under High Penetration of Wind Generation Units in Smart Grid Environment", 3rd International Conference on Computer and Knowledge Engineering (ICCKE 2013), Ferdowsi University of Mashhad.
- [14] Sepasi S, Reihani E, Howlader A, Roose L, Matsuura M, 2017, "Very short term load forecasting of a distribution system with high PV penetration", *Renewable Energy*, Vol. 106, pp. 142–148
- [15] Chang WY, 2014, "A Literature Review of Wind Forecasting Methods", *Journal of Power and Energy Engineering*, pp.161 – 168
- [16] Asare-Bediako B, Kling WL, Ribeiro PF, 2013, "Day-Ahead Residential Load Forecasting with Artificial Neural Networks using Smart Meter Data", *Power Tech (POWERTECH)*, Grenoble
- [17] Wang J, Li L, Niu D, Tan Z, 2012, "An annual load forecasting model based on support vector regression with differential evolution algorithm", *Journal of Applied Energy*, vol. 95, pp 65 – 70
- [18] Huyghues-Beaufond N, Tindemans S, Falugi P, Sun M, Strbac G, 2020, "Robust and automatic data cleansing method for short-term load forecasting of distribution feeders", *Applied Energy*, Vol. 261, 114405
- [19] Aggarwal CC, 2013, "Outlier analysis", Springer
- [20] Hampel FR, 1974, "The influence curve and its role in robust estimation", *Journal of the American Statistical Association*, vol 69 (346), pp. 383–93
- [21] Hampel FR, 1971, "A general qualitative definition of robustness". *Ann Math Stat*, vol 42(6), pp.1887–96
- [22] Hoaglin DC, Iglewicz B, Tukey JW, 1986, "Performance of some resistant rules for outlier labeling", *Journal of the American Statistical Association*, vol 81(396), pp. 991–9
- [23] Ye C, Ding Y, Wang P, et al, 2019, "A data-driven bottom-up approach for spatial and temporal electric load forecasting", *IEEE Trans Power Syst*, vol 34(3), pp. 1966–79
- [24] Zhou B, Meng Y, Huang W, Wang H, Deng L, Huang S, Wei J, 2021, "Multi-energy net load forecasting for integrated local energy systems with heterogeneous prosumers", *Electrical Power and Energy Systems*, Vol. 126, 106542
- [25] Karsoliya S, 2012, "Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture", *Int J Eng Trends Technol*, vol. 3(6)
- [26] Ahmed M, Abdelrazek S, Kamalasan S, Enslin J, Fenimore T, 2016, "Weather Forecasting Based Intelligent Distribution Feeder Load Prediction", 2016 IEEE Power and Energy Society General Meeting (PESGM), Boston, USA