doi: 10.46793/EEE26-3.14M

Originalni naučni rad
UDK 662.613:004.8

Predikcija toplotne moći uglja termoelektrane „Kostolac B" pomoću modela mašinskog učenja

Aleksandar Milićević, Srđan Belošević, Ivan Tomanović, Nenad Crnomarković, Milić Erić, Zoran Marković, Nikola Živković, Ivan Lazović

Laboratorija za termotehniku i energetiku, Institut za nuklearne nauke „Vinča" - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu

Rezime - Toplotna moć uglja predstavlja jedan od ključnih parametara za upravljanje procesom sagorevanja i energetskom efikasnošću termoelektrana. Njeno tačno određivanje od suštinskog je značaja za kontrolu potrošnje goriva, optimizaciju rada kotlova i smanjenje negativnih uticaja na životnu sredinu. Iako se toplotna moć uglja tradicionalno određuje laboratorijskim merenjima, ova procedura je vremenski zahtevna za kontinuirano praćenje kvaliteta goriva u realnim uslovima rada termoelektrane. U kontekstu rastućih zahteva za digitalizacijom energetskih sistema i razvojem prediktivnih metoda zasnovanih na bazama podataka, cilj ovog rada je razvoj modela mašinskog učenja za pouzdanu predikciju donje toplotne moći lignita korišćenog u termoelektrani „Kostolac B", na osnovu rezultata laboratorijskih merenja. Baza podataka sa rezultatima tehničke i elementarne analize lignita i donje toplotnoj moći uglja, formirana je na osnovu eksperimentalnih merenja u akreditovanom Odeljenju za karakterizaciju goriva u okviru Laboratorije za termoenergetiku i energetiku Instituta za nuklearne nauke "Vinča". U radu su razvijena tri metode mašinskog učenja na formiranoj bazi: *Linear Regression*, kao bazni, linearni model, *Random Forest* kao robustan ansambl-model zasnovan na agregaciji podataka tokom treninga modela (bagging) i *Gradient Boosting* kao napredna metoda zasnovana na sekvencijalnoj nadogradnji tzv. stabala odlučivanja (decision tree). Za predikciju toplotne moći na osnovu tehničke analize uglja, model linearne regresije je pokazao stabilne i pouzdane rezultate, dok složeniji nelinearni pristupi ne donose dodatnu korist. Pri predikciji toplotne moći zasnovanoj na elementarnoj analizi uglja, modeli linearne regresije i *Gradient Boosting* su ostvarili slične tačnosti na test skupu, pri čemu nelinearni pristupi omogućavaju modeliranje složenijih hemijskih odnosa. Rezultati studije pokazuju da izbor modela zavisi od složenosti dostupnih ulaznih podataka i specifičnih zahteva primene. Primena ovih modela omogućava bržu procenu kvaliteta uglja i podržava efikasnije operativno planiranje, čime se smanjuje potreba za laboratorijskim analizama.

Ključne reči - lignit, toplotna moć, laboratorijska merenja, modeli mašinskog učenja, termoelektrana

I UVOD

Termoelektrane na ugalj imaju značajnu ulogu u elektroenergetskom sistemu Republike Srbije, pri čemu lignit predstavlja dominantno gorivo. Efikasnost i pouzdanost rada termoelektrana zavise od kvaliteta uglja, koji je kod lignita izraženo promenljiv usled složenog i nehomogenog sastava goriva. Ova varijabilnost direktno utiče na sagorevanje goriva, emisije i ukupne performanse sistema [1].

Jedan od najvažnijih pokazatelja kvaliteta uglja jeste njegova toplotna moć, koja predstavlja osnovni parametar za termičke proračune i upravljanje radom kotlova. Toplotna moć se definiše kao količina toplote oslobođena pri potpunom sagorevanju goriva i značajno utiče na potrošnju goriva, stepen korisnosti kotlova, stabilnost procesa sagorevanja i emisije zagađujućih materija u atmosferu. Kod čvrstih fosilnih goriva, a naročito lignita, toplotna moć pokazuje značajnu varijabilnost usled promenljivog sadržaja vlage, pepela i sagorivih komponenti [2].

U zavisnosti od toga da li se uzima u obzir toplota kondenzacije vodene pare nastale sagorevanjem vodonika i isparavanjem vlage, razlikuju se gornja i donja toplotna moć. Donja toplotna moć (DTM) se u praksi najčešće koristi u energetskim proračunima i analizi rada termoelektrana. Kod lignita, koji se odlikuje visokim sadržajem vlage i pepela, razlika između gornje i donje toplotne moći je posebno izražena, što dodatno naglašava značaj njenog pouzdanog određivanja u realnim uslovima rada postrojenja [3]. Toplotna moć uglja se standardno određuje eksperimentalno u laboratorijskim uslovima, najčešće primenom kalorimetrijske bombe, u skladu sa međunarodnim standardom ISO 1928:2020 [4].

Iako ove metode obezbeđuju visoku tačnost i pouzdanost rezultata, njihova primena je ograničena u kontekstu brzog i kontinuiranog praćenja kvaliteta uglja. Ograničenja se prvenstveno odnose na vremensku zahtevnost ispitivanja, troškove analize i nemogućnost neposredne primene u operativnom upravljanju procesima sagorevanja. Dodatni problem u eksploataciji lignita predstavlja njegova nehomogenost, koja proizlazi iz geoloških karakteristika ležišta, različitih zona otkopavanja, kao i mešanja goriva tokom transporta i skladištenja. Promene u sadržaju vlage, pepela i sagorivih komponenti uzrokuju fluktuacije toplotne moći i mogu dovesti do nestabilnosti sagorevanja u kotlovima. Zbog toga se u savremenoj energetici javlja potreba za razvojem alternativnih metoda koje omogućavaju brzu, pouzdanu i dovoljno tačnu procenu energetske vrednosti goriva, uz smanjenje oslanjanja na česta laboratorijska ispitivanja [5].

Razvoj digitalnih tehnologija i metoda mašinskog učenja omogućio je značajan napredak u analizi eksperimentalnih podataka u energetici. Dosadašnja istraživanja pokazala su da

modeli mašinskog učenja mogu ostvariti visoku tačnost u predikciji toplotne moći uglja na osnovu rezultata tehničke i elementarne analize [6]. Prednost ovih modela ogleda se u njihovoj sposobnosti da identifikuju složene i nelinearne zavisnosti između ulaznih parametara i energetskih karakteristika goriva [7]. U tom kontekstu, posebno se ističu modeli nadgledanog učenja [8], koji omogućavaju direktno povezivanje laboratorijskih ulaznih karakteristika sa poznatim vrednostima donje toplotne moći. Ipak, dosadašnja istraživanja pokazuju da je efikasnost modela mašinskog učenja u velikoj meri zavisna od lokalnih karakteristika uglja, što naglašava potrebu za razvojem modela prilagođenih konkretnim termoelektranama [9].

Laboratorija za termoenergetiku i energetiku Instituta za nuklearne nauke „Vinča“ ima dugogodišnju saradnju sa AD „Elektroprivreda Srbije“ na realizaciji različitih projekata termotehničkih ispitivanja u domaćim termoelektranama. U okviru periodično-povremenih ispitivanja emisija zagađujućih materija u vazduh, sprovode se i detaljna određivanja fizičko-hemijskog sastava goriva u skladu sa standardima, uključujući rezultate tehničke i elementarne analize uglja, u akreditovanom odeljenju Laboratorije.

U ovom radu su na osnovu prikupljenih eksperimentalnih podataka razvijeni modeli nadgledanog mašinskog učenja za predikciju donje toplotne moći lignita koji se koristi u termoelektrani „Kostolac B“. Kao ulazne promenljive korišćeni su rezultati tehničke i elementarne analize uglja, dok je donja toplotna moć predstavljala izlaznu veličinu modela. Razvijeni su i upoređeni sledeći modeli: Linear Regression kao bazni linearni model, Random Forest kao ansambl-metoda zasnovana na bagging pristupu i Gradient Boosting kao napredna metoda sekvencijalne nadogradnje stabala odlučivanja. Cilj rada je procena tačnosti i pouzdanosti ovih modela u predikciji toplotne moći lignita, kao i analiza njihove primenljivosti u realnim radnim uslovima termoelektrane.

II OPIS TERMOELEKTRANE „KOSTOLAC B” I ANALIZA UZORAKA LIGNITA

Termoelektrana „Kostolac B“ (TEKO B) nalazi se oko 5 km istočno od termoelektrane „Kostolac A“ i približno 5 km od ušća reke Mlave u Dunav, na desnoj obali reke Mlave. Lokacija termoelektrane smeštena je između naselja Drmno, Bradarac, Kličevac, sela Kostolac i grada Kostolca. Proizvodnja električne energije u TEKO B zasniva se na transformaciji hemijske energije uglja u toplotnu energiju, koja se oslobađa sagorevanjem u kotlovskim postrojenjima. U turbinskim postrojenjima toplotna energija sadržana u radnom medijumu (vodenoj pari) pretvara se u mehaničku energiju, da bi se na kraju mehanička energija u generatoru transformisala u električnu energiju.

Na lokaciji TEKO B, blok B1 pušten je u pogon 1988. godine, dok je blok B2 uveden u rad tri godine kasnije. Oba bloka imaju nazivnu snagu od 348,5 MW i nalaze se u kontinuiranoj eksploataciji. Kotlovi u termoelektrani, tipa Sulzer, namenjeni su sagorevanju sprašenog uglja u letu.

Nakon uzorkovanja uglja u blokovima B1 i B2, uzorci se analiziraju u Odeljenju za karakterizaciju goriva, koje se bavi fizičko-hemijskim i fizičko-mehaničkim ispitivanjima uglja, koksa i pepela uglja, u okviru Laboratorije za termotehniku i energetiku, akreditovane prema standardu ISO 17025. Tehnička i elementarna analiza uglja određuju se pomoću termogravimetrijskog analizatora LECO TGA701 i elementarnog analizatora LECO 628s, redosledno (slika 1). Donja toplotna moć uglja se određuje korišćenjem kalorimetrijske bombe.

Periodična i povremena ispitivanja emisija zagađujućih materija u vazduh na blokovima B1 i B2 izvršena su u periodu od 2015. do 2025. godine. Dobijene vrednosti tehničke i elementarne analize nekoliko uzoraka lignita, zajedno sa odgovarajućim vrednostima donje toplotne moći, prikazane su u Tabeli 1.



Slika 1. Aparatura za elementarnu (LECO 628s**)** i tehničku **(**LECO TGA701) analizu uglja

Tabela 1. Hemijski sastav (tehnička i elementarna analiza) nekoliko uzoraka lignita termoelektrane „Kostolac B“

	2015-B2-2	2018-B1-1	2021-B2-4	2022-B1-3	2025-B2-3
Volatili [%]*	24,26	24,07	23,57	23,18	24,72
Vlaga [%]*	41,96	40,71	39,45	39,94	37,81
Pepeo [%]*	18,02	19,49	22,07	21,35	21,48
Fiksni ugljenik [%]*	15,76	15,73	14,91	15,53	15,99
Ugljenik [%]*	26,51	26,81	25,07	24,09	26,70
Vodonik [%]*	2,31	2,41	2,24	2,30	2,39
Kiseonik [%]*	10,15	9,49	10,13	11,34	10,55
Azot [%]*	0,42	0,37	0,37	0,32	0,40
Sagorljivi sumpor [%]*	0,63	0,72	0,67	0,66	0,67
Donja toplotna moć [kJ/kg]*	9037	9021	8728	8573	9339

* sa dostavnom vlagom

III MODELI MAŠINSKOG UČENJA

Linearna regresija predstavlja jedan od fundamentalnih modela mašinskog učenja, koji se koristi za modelovanje linearne zavisnosti između skupa ulaznih promenljivih i zavisne (ciljne) varijable [10]. Osnovna pretpostavka ovog modela jeste da se ciljna varijabla može izraziti kao linearna kombinacija ulaznih promenljivih i odgovarajućih regresionih koeficijenata, uz dodatak slučajne greške. Parametri modela se najčešće procenjuju metodom najmanjih kvadrata, čime se minimizuje suma kvadrata odstupanja između stvarnih i predviđenih vrednosti. Zbog svoje jednostavne matematičke formulacije i lake interpretabilnosti, linearna regresija omogućava uvid u uticaj pojedinačnih ulaznih promenljivih na izlaznu varijablu. Međutim, njena primena je ograničena u slučajevima kada su odnosi između promenljivih nelinearni ili kada postoje složene interakcije koje model ne može adekvatno da opiše. Uprkos tim ograničenjima, često se koristi kao bazni model i polazna tačka za poređenje sa složenijim metodama mašinskog učenja.

Random Forest je nelinearni ansambl-model zasnovan na kombinovanju velikog broja stabala odlučivanja, sa ciljem poboljšanja tačnosti i stabilnosti predikcije [11]. Ovaj model koristi tzv. tehniku *bootstrap aggregating* (*bagging*), pri čemu se svako stablo trenira na nasumično odabranom podskupu trening podataka, dobijenom uzorkovanjem sa vraćanjem [12]. Dodatno, pri svakoj podeli čvora (*node*) razmatra se samo slučajan podskup ulaznih obeležja, čime se smanjuje korelacija između pojedinačnih stabala i povećava raznovrsnost modela. Konačna predikcija dobija se agregacijom rezultata svih stabala, najčešće izračunavanjem prosečne vrednosti njihovih predikcija u regresionim problemima. Random Forest modeli su poznati po svojoj robusnosti, otpornosti na šum u podacima i smanjenoj sklonosti ka tzv. preprilagođavanju u poređenju sa pojedinačnim stablima odlučivanja. Zahvaljujući ovim karakteristikama, Random Forest se često koristi u problemima sa velikim brojem promenljivih i kompleksnom strukturom podataka.

Gradient Boosting predstavlja naprednu ansambl-metodu koja se zasniva na sekvencijalnoj izgradnji modela, gde se svaki novi model trenira tako da koriguje greške prethodnih modela [13]. Najčešće se kao bazni učenici koriste stabla odlučivanja ograničene dubine, dok se proces učenja zasniva na minimizaciji izabrane funkcije gubitka primenom gradijentnog spusta u prostoru funkcija. Za razliku od Random Forest metode, gde se stabla treniraju nezavisno jedno od drugog, Gradient Boosting uvodi zavisnost između uzastopnih modela, čime se omogućava precizno modelovanje složenih nelinearnih odnosa i interakcija između obeležja. Ovaj pristup često rezultuje visokom tačnošću predikcije, ali istovremeno zahteva pažljivo podešavanje hiperparametara, kao što su stopa učenja, broj stabala i njihova dubina, kako bi se izbeglo preprilagođavanje. Zbog toga se Gradient Boosting najčešće kombinuje sa tehnikama validacije i regularizacije u praktičnim primenama.

Evaluacija performansi modela je ključna za procenu njihove sposobnosti predikcije i generalizacije na testnoj bazi podataka. U ovom radu, tačnost modela je procenjena korišćenjem korena srednje kvadratne greške (RMSE) i srednje apsolutne procentualne greške (MAPE). Srednja apsolutna procentualna greška meri prosečno procentualno odstupanje predviđenih vrednosti od stvarnih vrednosti ciljne varijable. Ova metrika omogućava intuitivnu interpretaciju rezultata, jer grešku izražava u procentima, nezavisno od skale podataka. Niže vrednosti MAPE predstavlja bolju tačnost modela, a metrika je posebno pogodna za poređenje performansi različitih modela [14]. RMSE predstavlja standardnu devijaciju reziduala, odnosno prosečnu udaljenost predviđenih vrednosti od stvarnih. Niža vrednost RMSE podrazumeva veću preciznost predikcije, a ova metrika je posebno korisna kada se želi naglasiti apsolutna greška u istim jedinicama kao ciljne varijable [10]. Kombinacija RMSE i MAPE omogućava potpuniji uvid u performanse modela, jer MAPE pruža relativnu meru greške izraženu u procentima, dok RMSE kvantifikuje apsolutnu preciznost predikcije.

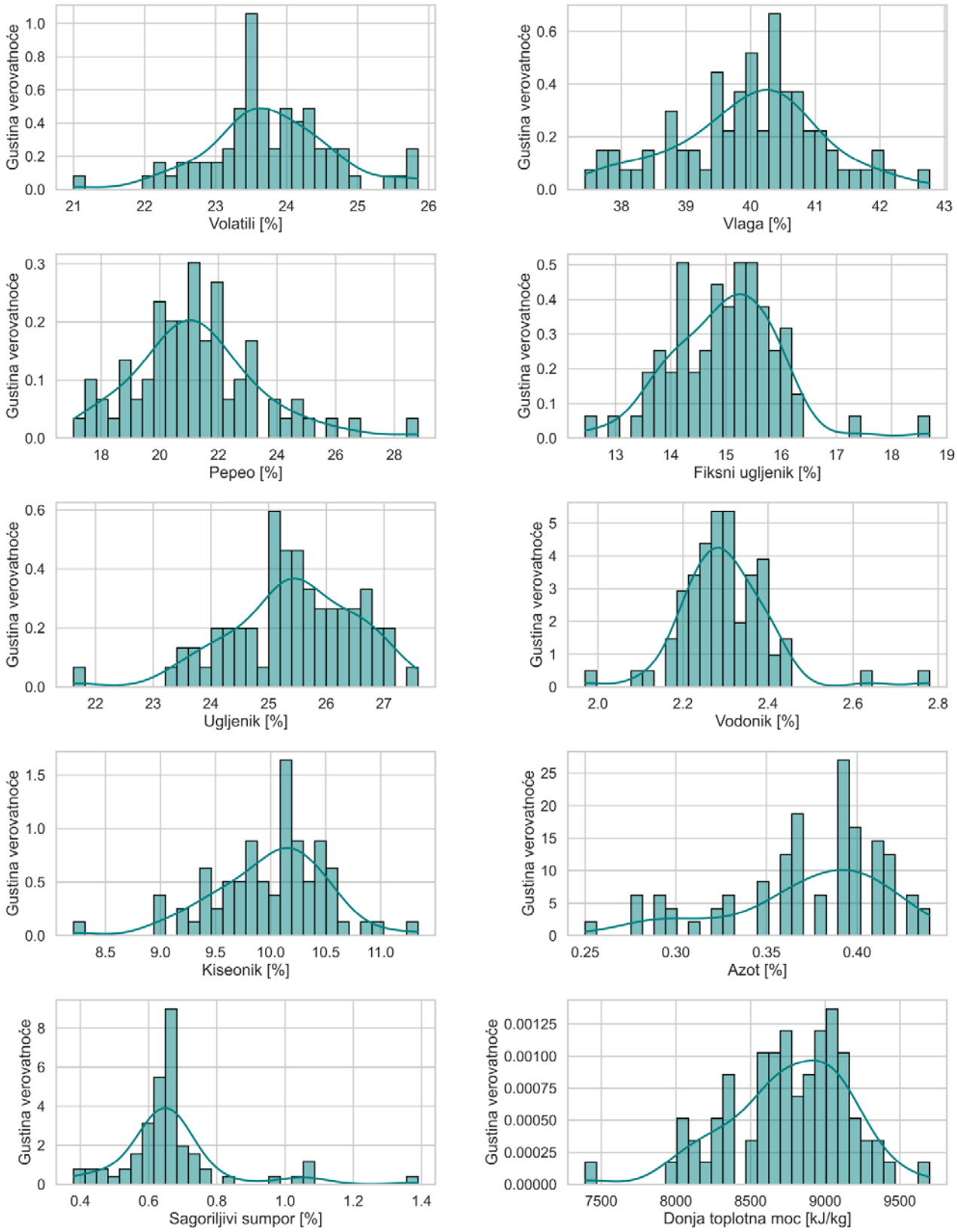
IV STATISTIČKA ANALIZA HEMIJSKOG SASTAVA I DONJE TOPLOTNE MOĆI UGLJA

Pre razvoja i evaluacije prediktivnih modela, sprovedena je deskriptivna statistička analiza hemijskog sastava i donje toplotne moći uzoraka uglja, kako bi se sagledale raspodele podataka i stepen njihove varijabilnosti. Na Slici 2 prikazane su raspodele hemijskog sastava i donje toplotne moći analiziranih uzoraka uglja kroz histograme i krive gustine verovatnoće. Na *y*-osi prikazana je gustina verovatnoće, što omogućava poređenje raspodela različitih parametara nezavisno od njihove apsolutne skale. Površina ispod svake krive normalizovana je na 1, što znači da predstavlja verovatnoću da slučajan uzorak ima vrednost unutar posmatranog intervala. Binovi u histogramima služe za vizuelizaciju raspodele, dok kriva gustine verovatnoće daje kontinualnu procenu raspodele.

Analiza rezultata ukazuje na relativno mali stepen varijabilnosti većine komponenti uglja. Sadržaj volatila iznosi u proseku 23,77 %, uz standardnu devijaciju od 0,88 %. Sličnu homogenost pokazuju i sadržaj vlage (39,98 % ± 1,12 %), ugljenika (25,44 % ± 1,09 %) i vodonika (2,30 % ± 0,11 %). Takođe, sadržaj kiseonika (10,00 % ± 0,51 %) i azota (0,37 % ± 0,04 %) imaju relativno male varijacije, što ukazuje na stabilnost ovih komponenti među uzorcima.

Sadržaji pepela (21,26 % ± 2,14 %) i fiksnog ugljenika (14,99 %

± 0,97 %) ispoljavaju nešto veću varijabilnost, što može biti posledica prisustva različitih mineralnih ostataka u uglju. Najveću varijabilnost među hemijskim komponentama ima sagorljivi sumpor (0,67 % ± 0,16 %), pri čemu su uočljivi i pojedinačni uzorci sa vrednostima značajno različitim od prosečne, što može imati značaj za procenu emisije SO_2.



Slika 2. Histogramske raspodele i krive gustine verovatnoće hemijskog sastava i donje toplotne moći uzoraka lignita

Donja toplotna moć ima prosečnu vrednost od 8759 kJ/kg, uz standardnu devijaciju od 403,9 kJ/kg. Iako su apsolutne vrednosti velike, relativna varijabilnost je mala, što potvrđuje visok stepen homogenosti energetskog sadržaja analiziranih uzoraka uglja.

Sve raspodele, osim raspodele sagorljivog sumpora, približno su simetrične, što se ogleda u sličnim vrednostima srednje vrednosti i medijana. Kombinacija deskriptivne statistike i vizualizacije gustine verovatnoće na Slici 2 potvrđuje da su hemijski sastav i energetski sadržaj uglja u analiziranim uzorcima uglavnom relativno homogeni. Ipak, pojedine komponente, kao što su pepeo, fiksni ugljenik i sagorljivi sumpor, ispoljavaju veći stepen varijabilnosti među uzorcima

V PREDIKCIJA DONJE TOPLOTNE MOĆI UGLJA NA OSNOVU PARAMETARA TEHNIČKE ANALIZE

U ovom poglavlju izvršena je predikcija donje toplotne moći uglja na osnovu parametara tehničke analize, tj. sadržaja: volatila, vlage, pepela i fiksnog ugljenika. Analiza je sprovedena na eksperimentalnoj bazi od 76 uzoraka uglja, od kojih je 55 korišćeno za obuku modela, dok je 21 uzorak izdvojen za nezavisnu evaluaciju performansi modela.

Radi pouzdane procene generalizacione sposobnosti modela, unutar trening skupa primenjena je unakrsna validacija pomoću 5 primeraka (*5-fold cross-validation*). Za Random Forest i Gradient Boosting modele izvršena je optimizacija hiperparametara primenom procedure rešetkastog pretraživanja (*grid search*), dok linearna regresija nije zahtevala dodatno podešavanje zbog svoje jednostavne strukture.

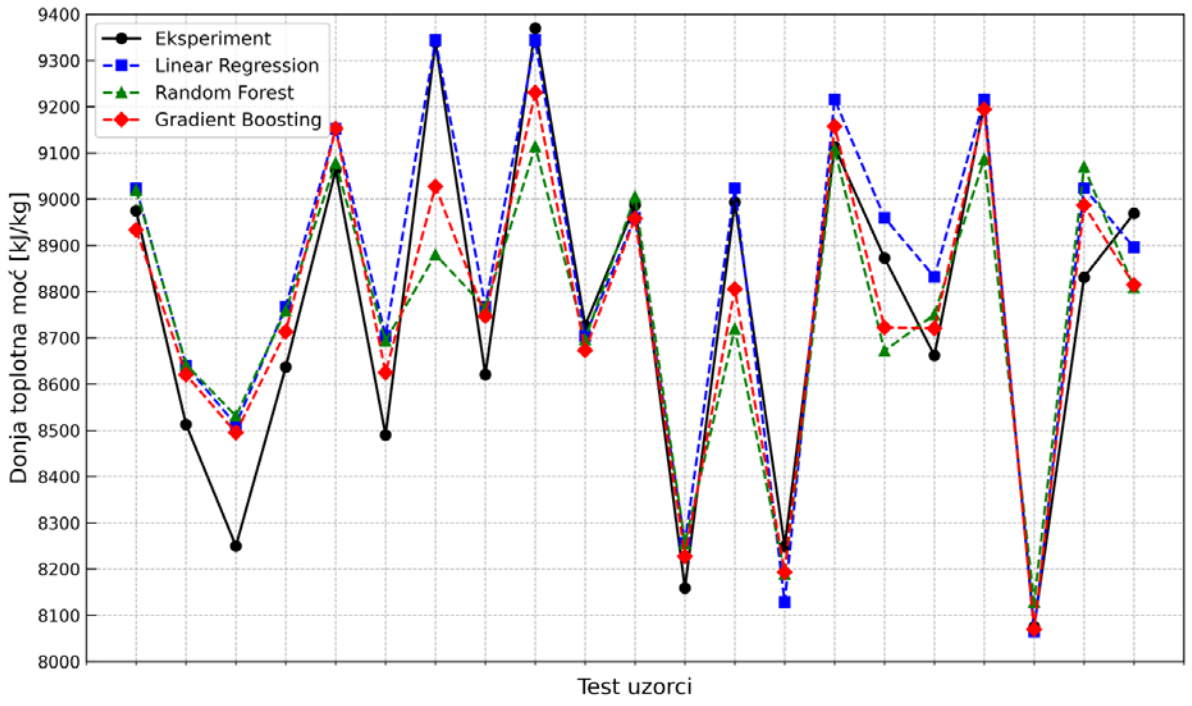
Linearni regresioni model pokazao je najvišu tačnost na test skupu podataka, sa vrednostima RMSE od 118,8 kJ/kg i MAPE od 1,1 %. Na trening skupu ostvarene su slične vrednosti greške (RMSE = 108,9 kJ/kg i MAPE = 1,0 %), što ukazuje na dobru saglasnost između trening i test performansi. Uporedivost grešaka na oba skupa potvrđuje stabilnost modela i odsustvo izraženog preprilagođavanja. Ovi rezultati sugerišu da se dominantna veza između parametara tehničke analize i donje toplotne moći može adekvatno opisati linearnim odnosom.

Random Forest model ostvario je niže greške na trening skupu (RMSE = 85,5 kJ/kg i MAPE = 0,6 %), ali značajno slabije performanse na test skupu (RMSE = 180,9 kJ/kg i MAPE = 1,6 %). Ovakvo performanse sugerišu preprilagođavanje, karakteristično za modele veće složenosti primenjene na ograničenim skupovima podataka.

Gradient Boosting model postiže vrlo visoku tačnost na trening skupu (RMSE = 57,6 kJ/kg i MAPE = 0,4 %), ali pokazuje smanjenu sposobnost generalizacije u poređenju sa linearnom regresijom (RMSE = 130,9 kJ/kg I MAPE = 1,2 %), iako je njegova generalizacija bolja nego kod Random Forest modela.

Ograničena veličina baze podataka, mali broj ulaznih promenljivih, kao i prethodno uočena homogenost parametara tehničke analize, upućuju na to da dominantni odnosi između ulaza i donje toplotne moći mogu biti adekvatno opisani linearnim modelom, dok složeniji nelinearni pristupi ne donose dodatno poboljšanje prediktivnih performansi.

Na Slici 3 prikazano je poređenje eksperimentalno određenih vrednosti donje toplotne moći i predikcija sva tri modela na test skupu podataka dobijenih na osnovu tehničke analize uglja. Vizuelni rezultati jasno ukazuju da linearna regresija daje najdoslednije predikcije duž celog opsega test uzoraka.



Slika 3. Poređenje eksperimentalno određenih i modelom predviđenih vrednosti donje toplotne moći uglja na osnovu tehničke analize

Na osnovu rezultata može se zaključiti da, u slučaju ograničenih skupova podataka i relativno jednostavnih ulaznih parametara, linearna regresija predstavlja optimalan balans između tačnosti, robusnosti i interpretabilnosti. Ovakav pristup omogućava pouzdanu procenu energetske vrednosti uglja isključivo na osnovu podataka tehničke analize, što ima značajnu praktičnu i inženjersku vrednost.

VI PREDIKCIJA DONJE TOPLOTNE MOĆI UGLJA NA OSNOVU PARAMETARA ELEMENTARNE ANALIZE

U ovom poglavlju analizirana je mogućnost predikcije donje toplotne moći uglja na osnovu parametara elementarne analize, koji obuhvataju sadržaj: ugljenika, vodonika, azota, sagorljivog sumpora i kiseonika. Za razliku od tehničke analize, ovi parametri neposrednije odražavaju hemijski sastav goriva, zbog čega se može očekivati izraženija i potencijalno nelinearna zavisnost između ulaznih promenljivih i energetske vrednosti uglja. Evaluacija performansi sprovedena uz istu metodološku postavku kao u prethodnom poglavlju, čime je obezbeđena uporedivost rezultata.

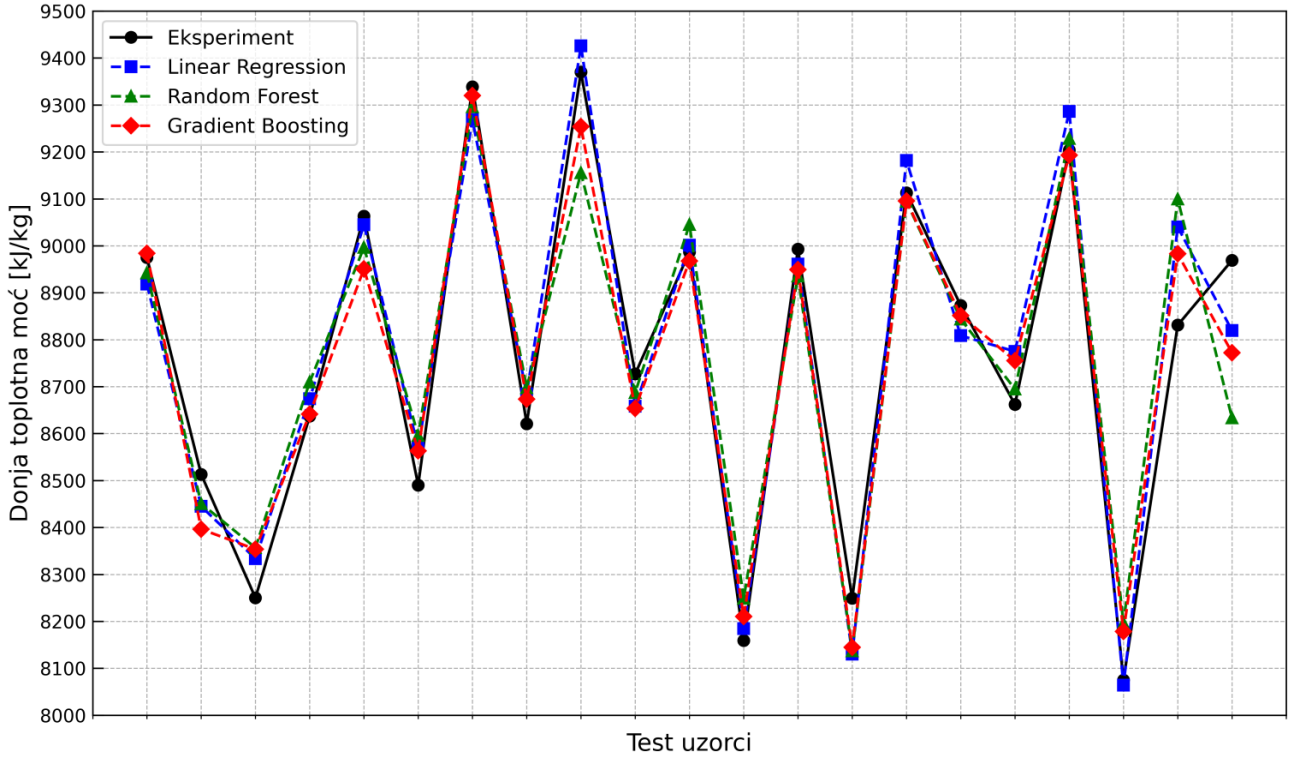
Linearni regresioni model pokazao je vrlo ujednačene performanse na trening i test skupu, sa vrednostima RMSE od 85,9 kJ/kg i 84,6 kJ/kg, kao i MAPE od 0,8 % i 0,8 % na oba skupa, što ukazuje na visoku stabilnost modela i odsustvo preprilagođavanja. Ovi rezultati potvrđuju da značajan deo veze između parametara elementarne analize i donje toplotne moći može biti opisan linearnim odnosom.

Random Forest model ostvario je male greške na trening skupu (RMSE = 64,6 kJ/kg i MAPE = 0,5 %), ali slabije tačnosti na test skupu (RMSE = 123,9 kJ/kg i MAPE = 1,0 %). Ovakav raskorak posledica je izraženog preprilagođavanje usled veće složenosti modela u odnosu na raspoloživi skup podataka.

Gradient Boosting model pokazao je najnižu grešku na trening skupu (RMSE = 54,6 kJ/kg i MAPE = 0,4 %), ali je za razliku od Random Forest pristupa zadržao dobru sposobnost generalizacije, sa test RMSE od 87,8 kJ/kg i MAPE od 0,8 %. Iako je njegova tačnost na test skupu blago slabija u poređenju sa linearnom regresijom po RMSE kriterijumu, razlike su male, dok model pokazuje veću fleksibilnost u opisivanju složenijih zavisnosti između hemijskog sastava i energetske vrednosti uglja.

Na slici 4 prikazano je poređenje eksperimentalno određenih vrednosti donje toplotne moći uglja i predikcija razmatranih modela na test skupu podataka postignutih na osnovu elementarne analize. Za razliku od rezultata na osnovu parametara tehničke analize, gde linearna regresija nadmašuje nelinearne modele, u slučaju elementarne analize Gradient Boosting ostvaruje konkurentne i u pojedinim aspektima povoljnije performanse. Ovo sugeriše da parametri elementarne analize nose bogatiju informaciju o hemijskoj strukturi goriva, pri čemu umereno nelinearni modeli mogu efikasnije da iskoriste dostupne podatke bez izraženog gubitka sposobnosti generalizacije.

Na osnovu rezultata može se zaključiti da linearna regresija i Gradient Boosting predstavljaju dve komplementarne strategije za predikciju donje toplotne moći na osnovu elementarne analize. Linearni regresioni model obezbeđuje stabilnost i interpretabilnost, dok Gradient Boosting omogućava nešto veću fleksibilnost uz očuvanu sposobnost generalizacije. Izbor modela zavisi od konkretnih zahteva primene, pri čemu oba pristupa omogućavaju pouzdanu procenu energetske vrednosti uglja.



Slika 4. Poređenje eksperimentalnih i modelom predviđenih vrednosti donje toplotne moći uglja na osnovu elementarne analize

VII ZAKLJUČAK

U radu su razvijeni modeli mašinskog učenja za predikciju donje toplotne moći lignita korišćenog u blokovima B1 i B2 termoelektrane „Kostolac B", na osnovu rezultata laboratorijskih merenja tehničke i elementarne analize uglja. U tu svrhu sprovedena je deskriptivna statistička analiza hemijskog sastava i donje toplotne moći uzoraka, a zatim razvijeni modeli: linearna regresija, Random Forest i Gradient Boosting. Na osnovu dobijenih rezultata mogu se izdvojiti sledeća zapažanja:

- Analizom raspodele hemijskog sastava i donje toplotne moći utvrđeno je da su većina komponenti uglja relativno ujednačena (na primer, sadržaj volatila iznosi 23,77 % ± 0,88 %, vlage 39,98 % ± 1,12 %, a ugljenika 25,44 % ± 1,09 %). Najveću varijabilnost pokazuju sadržaj pepela 21,26 % ± 2,14 %, fiksnog ugljenika 14,99 % ± 0,97 % i sagorljivog sumpora 0,67 % ± 0,16 %. Donja toplotna moć je homogenog rasporeda, sa prosečnom vrednošću od 8759 kJ/kg i standardnom devijacijom od 403,9 kJ/kg.
- Pri predikciji na osnovu tehničke analize uglja, Linearni regresioni model je pokazao najstabilnije performanse (RMSE trening = 108,9 kJ/kg i RMSE test = 118,8 kJ/kg; MAPE trening = 1,0 % i MAPE test = 1,1 %), dok složeniji nelinearni modeli, Random Forest i Gradient Boosting, ispoljavaju preprilagođavanje i nešto slabije rezultate na test skupu. Rezultati ukazuju da linearan odnos između parametara tehničke analize i donje toplotne moći može biti adekvatno modelovan jednostavnim linearnim modelom.
- Pri predikciji na osnovu elementarne analize uglja, linearni regresioni model i Gradient Boosting ostvarili su sličnu tačnost na test skupu (linearna regresija: RMSE = 84,6 kJ/kg, MAPE = 0,8 %; Gradient Boosting: RMSE = 87,8 kJ/kg, MAPE = 0,8 %). Gradient Boosting je pokazao veću fleksibilnost u modelovanju složenijih hemijskih odnosa, dok linearni model obezbeđuje stabilnost i jednostavnu interpretaciju.
- Uporedna analiza otkriva da modeli zasnovani na parametrima elementarne analize ostvaruju niže greške u predikciji donje toplotne moći u odnosu na modele zasnovane na tehničkoj analizi, što je delom posledica većeg broja ulaznih promenljivih (pet naspram četiri) i njihove neposrednije veze sa hemijskim sastavom goriva.

Analizirani rezultati pokazuju da razvijeni modeli omogućavaju pouzdanu i brzu procenu donje toplotne moći lignita, što može podržati operativno planiranje i optimizaciju sagorevanja u termoelektranama. Međutim, zbog relativno ograničenog obima baze podataka, dalja validacija na većim i raznovrsnijim skupovima uzoraka mogla bi dodatno poboljšati generalizaciju i pouzdanost modela.

ZAHVALNICA

Istraživanje je finansirano od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije (broj Ugovora: 451-03-33/2026-03/200017).

LITERATURA

[1] Milićević, A., Marković, Z., Erić, M., Belošević, S., Žarković, M., Marinković, A. Influence of fuel and operation mode on air pollutants emission from pulverized coal-fired power plant: Field experiments and ML predictions, Thermal Science and Engineering Progress, Vol. 57, 103151, 2025. https://doi.org/10.1016/j.tsep.2024.103151

[2] Büyükkanber, K., Haykiri-Acma, H., Yaman, S. Calorific value prediction of coal and its optimization by machine learning based on limited samples in a wide range, Energy, Vol. 277, 127666, 2023. https://doi.org/10.1016/j.energy.2023.127666

[3] Zhu, W., Xu, N., Hower J. C. Unveiling the predictive power of machine learning in coal gross calorific value estimation: An interpretability perspective, Energy, Vol. 318, 134781, 2025. https://doi.org/10.1016/j.energy.2025.134781

[4] Coal and coke — Determination of gross calorific value, ISO 1928:2025. https://www.iso.org/standard/86999.html [pristupljeno 06.02.2026]

[5] Kök, M.V., Matuschek, G., Varfolomeev, M.A., Nurgaliev, D.K. Combustion characteristics and efficiency of raw and clean lignite samples by thermal analysis techniques: TG–DTG–MID, DSC, and PDSC applications, Journal of Thermal Analysis and Calorimetry, Vol. 150, pp. 6013-6027, 2025. https://doi.org/10.1007/s10973-025-14204-w

[6] Guo, Y., Liu, X., Gao, Y., Wang, X., Ding, L., Pan, W., Hua, C., He, Y., Chen, X., Dai, Z., Yu, G., Wang, F. AutoML for calorific value prediction using a large database from the coal gasification practices in China, Journal of Thermal Analysis and Calorimetry, Vol. 12, No. 63, 2025. https://doi.org/10.1007/s40789-025-00763-8

[7] Goodfellow, I., Bengio, Y., Courville, A. *Deep Learning*, MIT Press, 2016.

[8] Raschka, S., Mirjalili V. *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2*, Birmingham, Packt Publishing, 2019. https://nuleren.be/ebooks/python-machine-learning-machine-learning-and-deep-learning-with-python-scikit-learn-and-tensorflow.pdf [pristupljeno 06.02.2026]

[9] Milićević, A., Belošević, S., Erić, M., Marković, Z., Tomanović, I., Crnomarković, N., Stojanović, A. Random forest model for determination of the lower heating value of TPP "Kolubara A" coal, in Proc. *Power Plants 2023*, Zlatibor, Serbia, pp. 745-752, 8-10 November 2023. https://vinar.vin.bg.ac.rs/handle/123456789/12766 [pristupljeno 06.02.2026]

[10] Montgomery, D. C., Peck, E. A., Vining, G. G. *Introduction to Linear Regression Analysis*, Wiley, 2021. https://www.kwcsangli.in/uploads/3--Introduction_to_Linear_Regression_Analysis__5th_ed._Douglas_C._Montgomery__Elizabeth_A._Peck__and_G._.pdf [pristupljeno 06.02.2026]

[11] Milićević, A., Belošević, S., Žarković, M., Tomanović, I., Crnomarković, N., Stojanović, A. Stupar, G., Deng, L., Che, D. Biomass and Bioenergy, Vol. 174, 106817, 2023. https://doi.org/10.1016/j.biombioe.2023.106817

[12] Breiman, L. Random Forests, Machine Learning, Vol. 45, pp. 5-32, 2001. https://doi.org/10.1023/A:1010933404324

[13] Friedman J. H. Greedy function approximation: A gradient boosting machine., Machine Learning, Vol. 29, No. 5, pp. 1189-1232, 2001. http://dx.doi.org/10.1214/aos/1013203451

[14] James, G., Witten, D., Hastie, T., Tibshirani, R. *An Introduction to Statistical Learning: with Applications in R*, Springer, 2013.

AUTORI

dr Aleksandar Milićević – viši naučni saradnik, Institut za nuklearne nauke „Vinča" - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu, amilicevic@vin.bg.ac.rs, ORCID 0000-0003-4615-8789

dr Srđan Belošević – naučni savetnik, Institut za nuklearne nauke „Vinča" - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu, v1belose@vin.bg.ac.rs, ORCID 0000-0001-9842-8408

dr Ivan Tomanović – viši naučni saradnik, Institut za nuklearne nauke „Vinča" - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu, ivan.tomanovic@vin.bg.ac.rs, ORCID 0000-0001-7573-7224

dr Nenad Crnomarković – naučni savetnik, Institut za nuklearne nauke „Vinča" - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu, ncrni@vin.bg.ac.rs, ORCID 0000-0001-5482-

082X
dr Milić Erić – viši naučni saradnik, Institut za nuklearne nauke „Vinča" - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu, milic@vin.bg.ac.rs, ORCID 0000-0001-6850-3112
dr Zoran Marković – viši naučni saradnik, Institut za nuklearne nauke „Vinča" - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu, zoda_mark@vin.bg.ac.rs, ORCID 0000-0003-0553-3263
dr Nikola Živković – viši naučni saradnik, Institut za nuklearne nauke „Vinča" - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu, nikolaz@vin.bg.ac.rs, ORCID 0000-0003-3399-1229
dr Ivan Lazović – viši naučni saradnik, Institut za nuklearne nauke „Vinča" - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu, ivan.lazovic@vin.bg.ac.rs, ORCID 0000-0002-3877-5157

Machine Learning Models for Predicting the Heating Value of Coal at the “Kostolac B” Thermal Power Plant

Abstract – The heating value of coal is one of the key parameters for managing the combustion process and the energy efficiency of thermal power plants. Its accurate determination is essential for controlling fuel consumption, optimizing boiler operation, and reducing negative environmental impacts. Although the heating value of coal is traditionally determined through laboratory measurements, this procedure is time-consuming for continuous monitoring of fuel quality under real operating conditions. With increasing demands for the digitalization of energy systems and the development of data-driven predictive methods, the aim of this study is to develop machine learning models for reliable prediction of the lower heating value of lignite used at the “Kostolac B” thermal power plant, based on laboratory measurements. A database containing the results of proximate and ultimate lignite analyses, as well as its lower heating value, was created from experimental measurements conducted in the accredited Department for Fuel Characterization within the Laboratory for Thermal Engineering and Energy at the “Vinča” Institute of Nuclear Sciences. Three machine learning models were developed using this database: Linear Regression as the baseline linear model, Random Forest as a robust ensemble model based on bagging, and Gradient Boosting as an advanced method based on sequential decision-tree enhancement. For predicting heating value based on the proximate analysis of coal, the linear regression model demonstrated stable and reliable results, while complex nonlinear approaches provided no additional benefit. When predicting heating value based on ultimate coal analysis, both linear regression and Gradient Boosting achieved similar accuracy on the test set, with nonlinear approaches enabling the modelling of more complex chemical relationships. The study results indicate that the choice of model depends on the complexity of the available input data and the specific application requirements. The application of these models enables faster assessment of coal quality and supports more efficient operational planning, thereby reducing the need for laboratory analyses.

Index Terms – Lignite, Heating value, Laboratory measurements, Machine learning, Thermal power plant

Rad primljen: 31.05.2026., prihvaćen: 22.08.2026.