

UDK: 004.85:621.61(045)=163.41

DOI: 10.5937/bakar2401021P

NAUČNI RAD

Oblast: Softversko inženjerstvo

Primljen: 17.04.2024.

Prerađen: 27.04.2024.

Prihvaćen: 30.04.2024.

**PRIMER MODELOVANJA VIBRACIJA DUVALJKE ZA SO<sub>2</sub>  
METODOM MAŠINSKOG UČENJA**

**EXAMPLE OF MODELING SO<sub>2</sub> BLOWER VIBRATIONS USING  
THE MACHINE LEARNING METHOD**

Marijana Pavlov-Kagadejev<sup>1a</sup>, Milan Radivojević<sup>1b</sup>, Predrag Stolić<sup>2a</sup>

<sup>1</sup>Institut za rudarstvo i metalurgiju Bor, Alberta Ajnštajna 1, 19210 Bor

<sup>2</sup>Univerzitet u Beogradu, Tehnički fakultet u Boru, Vojske Jugoslavije 12,  
19210 Bor

<sup>1a</sup> E-mail: marijana.pavlov@irmbor.co.rs, Orcid: <https://orcid.org/0000-0003-1090-6351>

<sup>1b</sup> E-mail: milan.radivojevic@irmbor.co.rs, Orcid: <https://orcid.org/0000-0003-2337-0306>

<sup>2a</sup> E-mail: pstolic@tfbor.bg.ac.rs, Orcid: <https://orcid.org/0000-0002-2574-4765>

**Izvod**

*U radu je opisan primer primene mašinskog učenja za predikciju vibracija duvaljke za SO<sub>2</sub>. Za formiranje modela korišćen je SVR (Support Vector Regression) algoritam. Prikazana je podela mašinskog učenja, kao i matematičke osnove SVM (Support Vector Machines) modela. Na osnovu velikog broja podataka, dobijenih pomoću monitoring sistema, formiran je matematički model. Nakon testiranja, kao mere tačnosti modela date su vrednosti za MSE i R<sup>2</sup>.*

**Ključne reči:** mašinsko učenje, SVM, algoritam, predikcija, vibracije

**Abstract**

*This paper describes an example of applying the machine learning to predict the SO<sub>2</sub> blower vibrations. The SVR (Support Vector Regression) algorithm was used to form the model. The paper includes an overview of machine learning and mathematical fundamentals of the SVM model. Based on a large amount of data, obtained from the monitoring system, a mathematical model was constructed. After testing, the measures of model accuracy such as the MSE (Mean Squared Error) and R<sup>2</sup> were provided.*

**Keywords:** machine learning, SVM, algorithm, prediction, vibration

**1. UVOD**

Do početka 2015. godine u RTB-u Bor radile su stara Topionica bakra i stara Fabrika sumporne kiseline, nakon čega su ove dve fabrike zamenjene novijom tehnologijom.

Tehnologija proizvodnje bakra u staroj Topionici je klasična pirometalurgija, kod koje se SO<sub>2</sub> gas dalje koristi u proizvodnji H<sub>2</sub>SO<sub>4</sub> sa

relativno malim stepenom iskorišćenja <60%, što je dovelo do zagađenja životne sredine uzrokovane većim koncentracijama SO<sub>2</sub>. Kontrola emisije SO<sub>2</sub> gasa iz Topionice vršila se u postrojenju Fabrike sumporne kiseline [1].

Da bi se poboljšao sistem kontrole u staroj Fabrici sumporne kiseline, instaliran je industrijski PLC. Razvijena je odgovarajuća softverska aplikacija za praćenje parametara procesa kako bi se smanjilo zagađenje vazduha. Glavni ciljevi kontrolnog sistema bili su obrada podataka u realnom vremenu, prezentacija podataka (u obliku dinamičkih sinoptičkih šema, grafova u realnom vremenu i tabela) i upravljanje bazama podataka. Podaci koji su rezultat merenja instaliranog PLC-a, a postoje u arhivi Odeljenja za Energetsku informatiku Instituta za rudarstvo i metalurgiju u Boru, iskorišćeni su u ovom radu.

U radu su prikazani rezultati primene matematičkog modela vibracija duvaljke za SO<sub>2</sub>, dobijenog metodom mašinskog učenja. Cilj rada je da prikaže mogućnost primene mašinskog učenja u predikciji različitih vrsta merenih signala. Za modelovanje je korišćen SVR (eng. Support Vector Regression) algoritam. Korišćeni su mereni parametri (temperature ležajeva i vibracije duvaljke) iz stare Fabrike sumporne kiseline. Na osnovu analize velikog broja podataka iz postojeće baze, uočena je njihova međusobna zavisnost. Veliki broj podataka je omogućio izradu matematičkog modela, a dobijeni rezultati predikcije, pokazali su da ovakav način ima smisla primenjivati na različite vrste merenih signala u industriji.

SVM je relativno nova računarska metoda nadgledanog učenja zasnovana na teoriji statističkog učenja. Za razliku od većine metoda klasifikacije, SVM ne zahteva veliki broj uzoraka za obuku. Osim toga, SVM metoda može rešiti problem učenja čak i kada je dostupan samo mali broj uzoraka za obuku. Ovaj metod se primenjuje u različitim granama industrije. Jedna od primena je za automatsku dijagnozu neispravnosti valjakastih ležajeva [2].

Takođe, predložen je novi pristup izbora modela za prognoziranje vremenskih serija zasnovan na SVM klasifikaciji [3]. U radu [4], autori su predložili pristup mašinskog učenja za modeliranje i simulaciju izmenjivača toplote. Predložena tehnika koristi SVM metod za predikciju performansi izmenjivača toplote. Primer primene SVM modela u telekomunikacijama, dat je u radu [5], dok je u [6] opisana primena SVM modela u hemiji.

## **2. MAŠINSKO UČENJE**

### **2.1 Mašinsko učenje i njegova primena**

Oblast računarstva koja se najbrže razvija poslednjih decenija je veštačka inteligencija. Mašinsko učenje predstavlja oblast veštačke inteligencije koja se bavi izradom prilagodljivih računarskih modela, koji imaju sposobnost učenja iz

postojećih podataka. Ovi modeli mogu poboljšati svoje performanse upotrebom informacija iz iskustva. Oni su korisni tamo gde algoritamska rešenja nisu na raspolaganju, gde postoji nedostatak formalnih matematičkih modela, ili kada je ograničena stručnost u razumevanju složenih funkcija. Ovi modeli imaju mogućnost pronalazjenja novih zavisnosti između parametara, ispitujući odnose koji već postoje među njima. Primena mašinskog učenja postigla je dobre rezultate u mnogim oblastima, kao što je prepoznavanje govora, prepoznavanje rukom pisanog teksta, vožnja automobila, i sl. [7].

## 2.2 Tipovi mašinskog učenja

Postoje tri osnovna tipa mašinskog učenja:

- nadgledano učenje (učenje sa supervizijom),
- nenadgledano učenje (učenje bez supervizije), i
- polunadgledano učenje.

U prediktivnom, ili nadgledanom (nadziranom) pristupu, cilj je naučiti mapiranje od ulaza  $x$  do izlaza  $y$ , uzimajući u obzir označeni skup ulazno-izlaznih parova  $D = \{(x_i, y_i)\}_{i=1}^N$ . Oznaka  $D$  predstavlja *skup za obuku*, dok je  $N$  broj primera za obuku.

Slično tome, oblik izlazne ili odzivne promenljive u principu može biti bilo šta, ali većina metoda pretpostavlja da je  $y_i$  kategorička ili nominalna promenljiva, iz nekog konačnog skupa,  $y_i \in \{1, \dots, C\}$ , ili da je  $y_i$  realna skalarna vrednost. Kada je  $y_i$  kategorička, problem se naziva klasifikacija ili prepoznavanje obrazaca, a kada je  $y_i$  realna vrednost, problem se naziva regresija.

Drugi glavni tip mašinskog učenja je deskriptivni ili nenadgledani pristup učenju. Ovde su nam dati samo ulazi,  $D = \{x_i\}_{i=1}^N$ , a cilj je pronaći „interesantne obrasce“ u podacima. To se ponekad naziva otkrivanje znanja. Ovaj problem je znatno manje definisan, jer nije poznato kakve obrasce treba tražiti, i nema očiglednu meru greške koja se koristi (za razliku od nadgledanog učenja, gde možemo uporediti naše predviđanje  $y$  za dati  $x$  sa posmatranom vrednošću) [8].

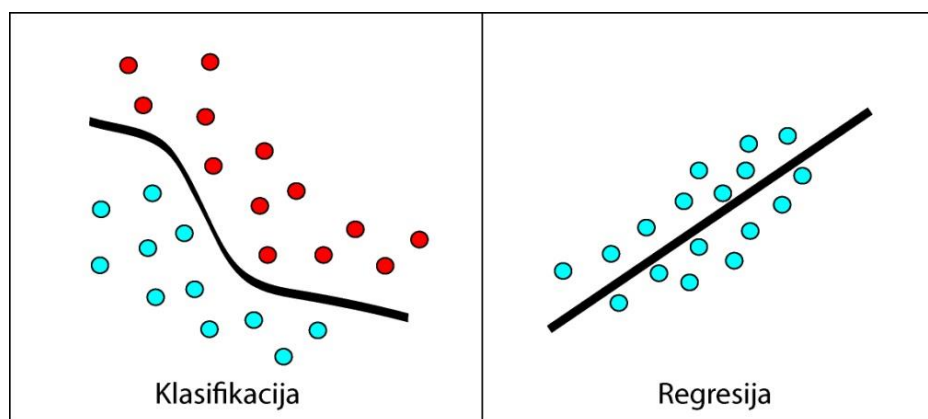
## 2.3 Primena nadgledanog mašinskog učenja

Postoje dva načina primene metoda nadgledanog mašinskog učenja: klasifikacija i regresija, slika 1.

*Klasifikacija* predstavlja problem razvrstavanja nepoznate instance u jednu od unapred ponuđenih kategorija – klasa i spada u jedan od najčešćih zadataka mašinskog učenja. Kako bismo bolje shvatili, ili organizovali stvari oko sebe, u

skladu sa našom prirodom je da ih klasifikujemo i kategorizujemo. Klasifikacija se uspešno primenjuje u: dijagnostifikovanju bolesti, klasifikaciji kreditnih zahteva klijenata, proceni da li će i koji korisnici kupiti određeni proizvod, izboru ciljne grupe klijenata za marketinške kampanje, analizi slike, analizi glasa za biometrijske potrebe, i sl. Kod klasifikacije funkcija cilja je diskretna. U opštem slučaju, oznakama klase se ne mogu dodeliti numeričke vrednosti niti uređenje. To znači da je atribut klase, čiju je vrednost potrebno odrediti, kategorički atribut [7].

Kod *regresije* izlaz je kontinualan i može se kvantitativno izraziti (objektima odgovaraju vrednosti iz skupa realnih brojeva). Regresija se najčešće koristi za *predviđanje*. Primer regresije je predviđanje kretanja cene akcija na berzi ili cena nekretnina.



Sl. 1. Primeri klasifikacije i regresije kod nadziranog mašinskog učenja

## 2.4 Metoda potpornih vektora

Metoda potpornih vektora (eng. Support Vector Machines) je algoritam nadgledanog mašinskog učenja, prvobitno kreiran za rešavanje problema binarne klasifikacije. Potrebno je podatke podeliti u dva linearno razdvojiva skupa, dve klase, tako da margina između elemenata tih skupova bude maksimalna, kao što je prikazano na slici 2 [9]. SVM je u osnovi binarni klasifikator, ali se može proširiti tako da vrši klasifikaciju u proizvoljan broj klasa.

Neka je dat skup od  $N$  podataka  $x_i, \{i = 1, 2, \dots, N\}$  za treniranje, koji pripada klasi  $y_i \in \{-1, 1\}$  (za binarnu klasifikaciju). Smatra se da je ovaj skup linearno razdvojiv, ukoliko postoji težinski faktor  $w$  i konstanta  $b$ , takav da važi [10,11]:

$$w \cdot x_i + b \geq 1 \text{ za } y_i = 1 \quad (1)$$

$$w \cdot x_i + b \leq -1 \text{ za } y_i = -1 \quad (2)$$

Prethodni izrazi se mogu napisati u sledećem obliku:

$$y_i (w \cdot x_i + b) \geq 1 \quad (3)$$

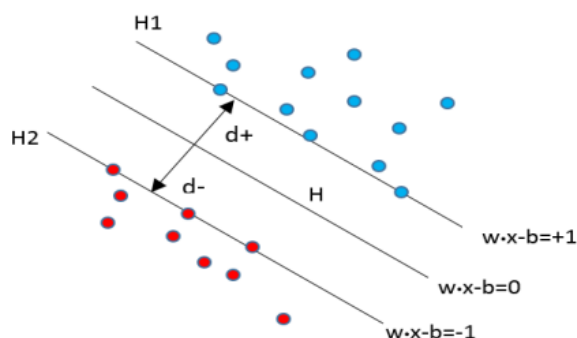
Ravni na kojima leže potporni vektori  $H_1$  i  $H_2$  mogu se prikazati sledećim izrazima:

$$w \cdot x_i + b = 1 \text{ za } H_1 \quad (4)$$

$$w \cdot x_i + b = -1 \text{ za } H_2 \quad (5)$$

Označimo rastojanja  $H_1$  i  $H_2$  od hiper-ravni, sa  $d_1$  i  $d_2$ , redom.  $H_1$  i  $H_2$  će biti podjednako udaljene od hiper-ravni, kada je  $d_1 = d_2 = \frac{1}{\|w\|}$ , pri čemu vrednosti  $d_1$  i  $d_2$  nazivamo marginama.

Vektori  $x_i$  za koje važi jednakost  $y_i \cdot (w \cdot x_i + b) = 1$ , nazivaju se potporni vektori [10]. Da bi se izabrala hiper-ravan maksimalno udaljena od potpornih vektora, potrebno je da se margina maksimizira [7].



Sl. 2. Princip metode potpornih vektora [9]

Ukoliko skup podataka nije moguće razdvojiti bez greške, uvodi se nenegativna vrednost  $\varepsilon_i$ , koja predstavlja toleranciju greške, odnosno:

$$y_i(w \cdot x_i + b) \geq 1 - \varepsilon_i \quad (6)$$

Ako skup podataka nije linearno razdvojiv u datom prostoru, moguće je izvršiti mapiranje ulaznog skupa podataka  $x_i$  u prostor više dimenzije, koristeći transformaciju  $\varphi(x_i)$ , gde je razdvajanje moguće. Uvodi se tzv. kernel funkcija:

$$k(x_i, x_j) = \varphi(x_i)^T \cdot \varphi(x_j) \quad (7)$$

Regresija primenom metode potpornih vektora (eng. Support Vector Regression - SVR) predstavlja algoritam koji je razvijen iz SVM metoda. Ovaj algoritam se koristi za regresiju. U slučaju regresije izlazna veličina nije

diskretna, kao kod klasifikacije, već je kontinualna i može se kvantitativno izraziti (kao realni broj, ili vektor realnih brojeva). Regresija se koristi često za predikciju različitih fizičkih veličina. Model se formira na osnovu poznatog skupa podataka za obučavanje (treniranje). Ovaj skup podataka sadrži ulazne veličine  $x$  i izlazne  $y$ . Nakon kreiranja modela, on će na osnovu poznatih ulaznih vrednosti  $x$ , biti u mogućnosti da predvidi nepoznate vrednosti izlaza  $y$ .

Pretpostavimo da nam je dat skup podataka za obuku  $\{x_i, y_i\}$ , gde su  $x_i$  ulazi, a  $y_i$  izlazi. Pomoću regresije potpornih vektora, cilj je pronaći funkciju  $f(x)$ , čije je najveće odstupanje od stvarne vrednosti  $\varepsilon$ . Za regresiju zasnovanu na potpornim vektorima linearni model se može opisati kao:

$$f(x_i) = w \cdot x + b; \quad w \in R^N, b \in R \quad (8)$$

gde: predstavlja  $w$  težinski faktor,  $b$  je konstanta, a  $w \cdot x$  označava skalarni proizvod. Veličine  $w$  i  $b$  se izračunavaju minimiziranjem na osnovu poznatog skupa podataka za treniranje.

$$\begin{aligned} & \min \frac{1}{2} \|w\|^2 \\ \text{uz uslove : } & \begin{cases} y_i - f(x_i) \leq \varepsilon \\ f(x_i) - y_i \leq \varepsilon \end{cases} \end{aligned} \quad (9)$$

Pretpostavljamo da postoji funkcija  $f(x_i)$  koja aproksimira sve parove  $(x_i, y_i)$  sa preciznošću  $\varepsilon$ . U ovom slučaju, pretpostavljamo da je problem izvodljiv. Ukoliko ne postoji model koji zadovoljava navedene uslove, uvode se tolerancije  $\xi_i$  i  $\xi_i^*$  i traži se rešenje problema optimizacije sa najmanjom greškom [12].

$$\begin{aligned} & f(x_i) - y_i \leq \varepsilon + \xi_i \\ & \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N |y_i - f(x_i)| \quad (10) \\ \text{uz uslove : } & \begin{cases} y_i - f(x_i) \leq \varepsilon + \xi_i \\ f(x_i) - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \end{aligned}$$

gde je  $C > 0$  konstanta koja predstavlja parametar koji određuje meru odstupanja većeg od  $\varepsilon$ , koje se može tolerisati.

### 3. REZULTATI DOBIJENI PRIMENOM SVR ALGORITMA

Za realizaciju regresije primenom metode potpornih vektora korišćena je LIBSVM biblioteka u MATLAB-u, sa radijalnom kernel funkcijom, odnosno:

$$k(x_i, x) = \exp(-\gamma \|x_i - x\|^2), \gamma > 0 \quad (11)$$

Svi podaci su podeljeni na podatke korišćene za treniranje, tj. učenje modela i podatke za testiranje, tj. proveru tačnosti modela, tako da je 80% podataka korišćeno za treniranje, a 20% podataka za testiranje. Kada je formiran odgovarajući model na osnovu podataka iz skupa za treniranje, vrši se predikcija vrednosti podataka iz skupa za testiranje, koji nisu korišćeni za učenje modela, a zatim poređenje dobijenih aproksimativnih vrednosti sa stvarnim vrednostima, kako bi se utvrdila tačnost modela.

Kao mera tačnosti modela korišćena je srednja kvadratna greška (*MSE* - *Mean Squared Error*), koja predstavlja srednje kvadratno odstupanje između stvarne i procenjene vrednosti izlaznog parametra:

$$MSE(y, f(x)) = \frac{1}{N} \sum_{i=1}^N (y_i - f(x_i))^2 \quad (12)$$

gde  $y_i$  predstavlja stvarne vrednosti izlaznog parametra, a  $f(x_i)$  procenjenju vrednost dobijenu primenjenim modelom. MSE uvek ima nenegativnu vrednost i vrednosti koje su bliže nuli označavaju tačniji model.

Pored toga, korišćen je i koeficijent determinacije  $R^2$  koji je definisan jednačinom:

$$R^2(y, f(x)) = 1 - \frac{\sum_{i=1}^N (y_i - f(x_i))^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (13)$$

gde  $\bar{y}$  predstavlja srednju vrednost stvarnih vrednosti izlaznog parametra. Što su vrednosti  $R^2$  bliže jedinici, veća je tačnost modela.

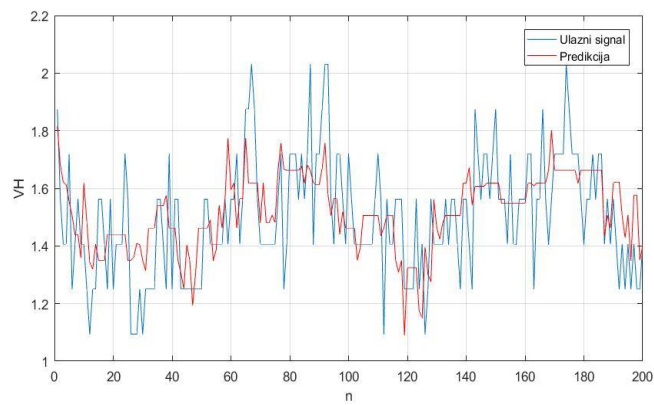
Za ispitivanje je korišćeno 1000, 2000, 5000 i 20000 podataka, pri čemu je 80% iz svake grupe podataka (tj. 800, 1600, 4000 i 16000), korišćeno za učenje (treniranje) modela, a 20% (tj. 200, 400, 1000 i 4000) podataka za testiranje modela, tabela 1. Kao ulazni parametri izabrane su temperature na ležajevima duvaljke (5 različitih temperatura), dok izlazni parametar koji je trebalo proceniti, predstavljaju vibracije duvaljke za SO<sub>2</sub>.

Na slici 3 dat je prikaz vrednosti vibracija  $V_H$  koje su dobijene merenjem, uz pomoć opisanog nadzorno-kontrolnog sistema (ulazni signal) i vrednosti koje su proračunate uz pomoć matematičkog modela (predikcija), za različit broj ulaznih podataka (1000, 2000, 5000 i 20000).

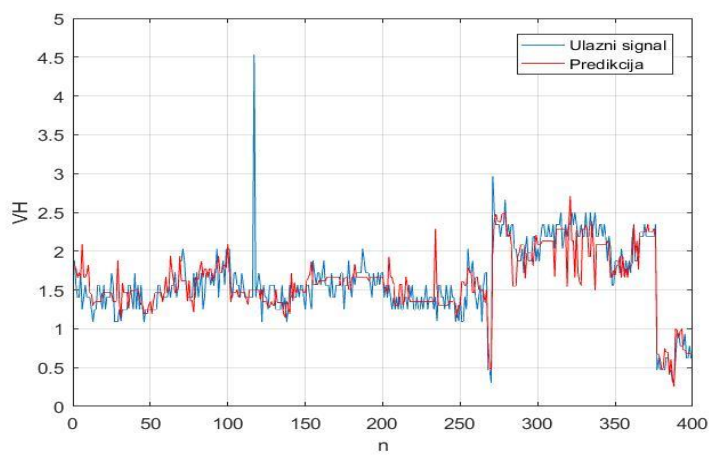
**Tabela 1.** Prikaz dobijenih vrednosti za MSE i  $R^2$

Br.podataka	1000	2000	5000	20000
MSE	0.03	0.08	0.11	0.52
$R^2$	0.33	0.66	0.94	0.96

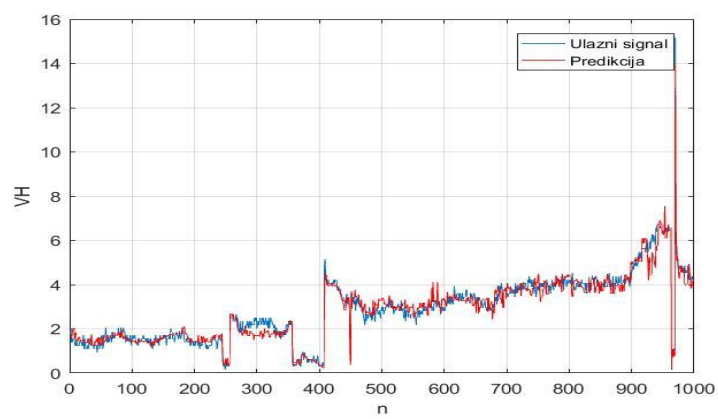
Kao rezultat primenjenog matematičkog modela na 20000 podataka, dobijena je vrednost srednje kvadratne greške  $MSE = 0.517944$ , a za koeficijent determinacije  $R^2 = 0.965034$ . Ovo se sa stanovišta tačnosti matematičkog modela, može smatrati zadovoljavajućim.



a)

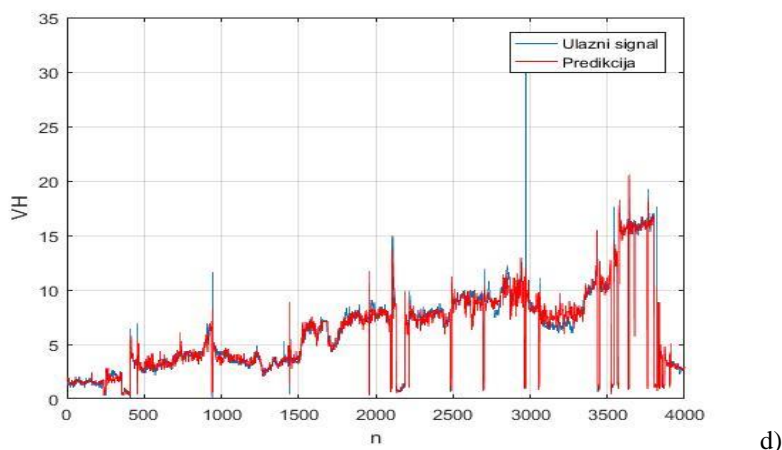


b)



c)





Sl. 3. Uporedni dijagram ulaznog signala i signala dobijenog predikcijom pomoću SVR modela za a) 1000, b) 2000, c) 5000, d) 20000 podataka

#### 4. ZAKLJUČAK

Primenom metoda mašinskog učenja na različite vrste merenih signala u industriji, mogu se dobiti veoma dobri rezultati u različitim oblastima merenja, što može pomoći u predviđanju budućih događaja i ponašanja mernih sistema (kao što je na primer predikcija otkaza sistema, i sl.). Na ovaj način, moguće je vršiti modelovanje bilo kog signala, ukoliko imamo dovoljno veliku bazu poznatih vrednosti parametara od kojih taj signal zavisi. Rezultati koji su dobijeni u ovom radu, dobijeni su samo na osnovu međusobne zavisnosti temperatura ležajeva i vibracija duvaljke za SO<sub>2</sub>, iako su vibracije posledica uticaja velikog broja različitih faktora i parametara. Projektovani matematički model za testirane podatke, dao je MSE = 0.517944, a za koeficijent determinacije R<sup>2</sup>=0.965034, što se sa stanovišta tačnosti matematičkog modela može smatrati zadovoljavajućim.

#### ZAHVALNICA

*Ovaj rad je finansijski podržan od strane Ministarstva nauke, tehnološkog razvoja i inovacija Republike Srbije, Ugovor o realizaciji i finansiranju naučnoistraživačkog rada Instituta za rudarstvo i metalurgiju Bor u 2024. godini, br. 451-03-66/2024-03/ 200052 i Univerziteta u Beogradu, Tehničkog fakulteta u Boru ugovor br. 451-03-65/2024-03/ 200131.*

---

**LITERATURA**

- [1] V. Tasić, D. R. Milivojević, V. Despotović, D. Brodić, M. Pavlov, Improving Control System in the Sulfuric Acid Plant– ICEST2012, Proceedings, June 28-30 2012, Veliko Tarnovo, Bulgaria, pp. 573-576.
- [2] K.C. Gryllias, I.A. Antoniadis, A Support Vector Machine approach based on physical model training for rolling element bearing fault detection in industrial environments, *Engineering Applications of Artificial Intelligence*, 25(2) (2012) 326-344.
- [3] M. A. Villegas, D. J. Pedregal, J. R. Trapero, A support vector machine for model selection in demand forecasting applications, *Computers & Industrial Engineering*, 121 (2018) 1-7.
- [4] S. Muthukrishnan, H. Krishnaswamy, S. Thanikodi, D. Sundaresan, V. Venkatraman, Support vector machine for modelling and simulation of heat exchangers, *Thermal Science*, Vol. 24, 1B (2020) 499-503.
- [5] A. Rodan, H. Faris, J. Alsakran, O. Al-Kadi, A support vector machine approach for churn prediction in telecom industry, *International Journal on Information*, 17(8) (2014) 3961-3970.
- [6] H. Li, Y. Liang, Q. Xu, Support vector machines and its applications in chemistry, *Chemometrics and Intelligent Laboratory Systems*, 95(2) (2009) 188-198.
- [7] J. Đ. Novaković, Rešavanje klasifikacionih problema mašinskog učenja, Fakultet tehničkih nauka u Čačku, 2013.
- [8] K. P. Murphy, Probabilistic machine learning: an introduction, MIT press, Cambridge, Massachusetts, 2022.
- [9] D. Aju, R. Rajkumar, A multi-stage hybrid cad approach for MRI brain tumor recognition and classification, *IIOAB Journal*, 7.1 (2016) 7-20.
- [10] C. Cortes, Vladimir Vapnik, Support-vector networks, *Machine learning*, 20 (1995) 273-297.
- [11] H. Getachew Girma, A tutorial on support vector machine, Center of Experimental Mechanics, University of Ljubljana, 2009.
- [12] T. B. Trafalis, B. Santosa, M. B. Richman, Learning networks in rainfall estimation, *Computational Management Science*, 2 (2005) 229-251.