

## Mašinsko učenje veštačke neuronske mreže sa radijalnim aktivacionim funkcijama Gausovog tipa na bazi Kalmanovog filtra – rezultati primene

NAJDAN L. VUKOVIĆ, Univerzitet u Beogradu,

Inovacioni centar Mašinskog fakulteta, Beograd

ZORAN Đ. MILJKOVIĆ, Univerzitet u Beogradu,

Mašinski fakultet, Beograd

Originalni naučni rad

UDC: 004.85:004.032.26]:681.5.015.44

*U ovom radu su prikazani eksperimentalni rezultati primene tri nova sekvencijalna algoritma mašinskog učenja u cilju optimizacije parametara veštačke neuronske mreže sa radijalnim aktivacionim funkcijama Gausovog tipa na bazi Kalmanovog filtra. Uvedena su tri nova sekvencijalna algoritma mašinskog učenja: linearizovani Kalmanov filter, linearizovani informacioni filter, algoritam specifične aproksimacije momenata Gausove raspodele. Nakon prikaza odgovarajućih matematičkih modela datih u prvom delu ovog rada, u ovom delu razvijeni algoritmi su testirani u MATLAB® programskom okruženju razvojem odgovarajućeg softverskog koda i primenom test skupova podataka. Iako svi izabrani test skupovi podataka predstavljaju realne probleme, razvijeni algoritmi su testirani i na realnom inženjerskom problemu modeliranja izgleda segmenta obrađene površine. Sva tri algoritma su prilikom modeliranja ovih problema pokazala visok stepen tačnosti.*

**Ključne reči:** mašinsko učenje, veštačke neuronske mreže, Kalmanov filter

### 1. UVOD

U prvom delu ovog rada prikazana su tri sekvencijalna algoritma mašinskog učenja sopstvenog razvoja: linearizovani Kalmanov filter, linearizovani informacioni filter i algoritam specifične aproksimacije momenata. U ovom delu su dati eksperimentalni rezultati koji su dobijeni tokom testiranja karakteristika razvijenih algoritama za optimizaciju parametara veštačke neuronske mreže sa Gausovim aktivacionim funkcijama. Za sva tri algoritma mašinskog učenja razvijen je odgovarajući softverski kod u MATLAB® programskom okruženju [1]. Eksperimentalni rezultati su dobijeni na laptop računaru Pentium 4 2,0 GHz CPU sa 2 GB operativne memorije. Ocena performansi razvijenih algoritama mašinskog učenja izvršena je primenom veštačke neuronske mreže sa radijalnim aktivacionim funkcijama Gausovog tipa za modeliranje sledećih problema:

- Skupovi (realnih) podataka za proveru algoritama mašinskog učenja;

Adresa autora: Najdan Vuković, Univerzitet u Beogradu, Inovacioni centar Mašinskog fakulteta, Beograd, Kraljice Marije 16

Rad primljen: 07.05.2014.

Rad prihvaćen: 12.06.2014.

- Modeliranje topologije obrađene površine.

### 2. SKUPOVI REALNIH PODATAKA ZA PROVERU ALGORITAMA MAŠINSKOG UČENJA

Pomoću ovih skupova za proveru algoritama mašinskog učenja proveravaju se konvergencija, tačnost i preciznost novih algoritama [1], [2], [3]. Na ovaj način se novi algoritmi mogu direktno uporediti sa postojećim algoritmima koji su dobro poznati naučnoj zajednici. Od velikog broja postojećih skupova podataka navedenih u [4] izabrana su tri karakteristična skupa, čije osnovne karakteristike su date u Tabeli 1.

Od polaznog skupa podataka 50% je upotrebljeno za proces optimizacije parametara veštačke neuronske mreže, a preostalih 50% za verifikaciju rezultata i ocenu kvaliteta sprovedenog procesa mašinskog učenja (test skup podataka). Za sva tri razvijena algoritma testirane su arhitekture sa pet, deset, 15 i 20 neurona u skrivenom sloju (označavanje je usvojeno shodno preporuci u [5]):

- Test skup br. 1: 8-[5]<sub>1-1</sub>; 8-[10]<sub>1-1</sub>; 8-[15]<sub>1-1</sub>; 8-[20]<sub>1-1</sub>;
- Test skup br. 2: 10-[5]<sub>1-1</sub>; 10-[10]<sub>1-1</sub>; 10-[15]<sub>1-1</sub>; 10-[20]<sub>1-1</sub>;

- Test skup br. 3: 8-[5]<sub>1-1</sub>; 8-[10]<sub>1-1</sub>; 8-[15]<sub>1-1</sub>; 8-[20]<sub>1-1</sub>.
- Tabela 1. Karakteristike izabranih test skupova [4]

Tabela 1. Karakteristike izabranih test skupova [4]

RB	Naziv skupa	Ukupan broj primera	Broj primera za učenje	Broj primera za testiranje	Broj dimenzija ulaznog vektora	Broj dimenzija izlazne veličine
1.	„Abalone”	4177	2089	2088	8	1
2.	„Weather Ankara”	1609	805	804	10	1
3.	„Auto-mpg”	398	199	199	8	1

Tačnost veštačke neuronske mreže ocenjena je na osnovu određivanja statistike kvadratnog korena srednje kvadratne greške (KSKG):

$$s = \sqrt{\frac{\sum_{i=1}^N \|y_i - \mathbf{g}(\mathbf{w}, \boldsymbol{\mu}, \sigma, \mathbf{x}_i)\|_2^2}{N}} \quad (1)$$

Ocena tačnosti je izračunata za tri skupa podataka:

- $s_{test}$  - KSKG za test skup podataka pomoću kojih se ocenjuje kvalitet sprovedenog procesa mašinskog učenja;
- $s_{opt}$  - KSKG za primere koji su korišćeni tokom procesa optimizacije;
- $s_{svi}$  - KSKG za sve primere elemente skupa za optimizaciju.

Tabela 3. Eksperimentalni rezultati za test skup – „Weather Ankara”.

Alg.	$s$	Broj neurona u skrivenom sloju			
		5	10	15	20
LKF	$s_{test}$	0,0415±0,0027	0,0416±0,0042	0,0419±0,0034	0,0402±0,0032
	$s_{opt}$	0,0398±0,0010	0,0403±0,0006	0,0395±0,0009	0,0397±0,0009
	$s_{svi}$	0,0407±0,0015	0,0410±0,0022	0,0407±0,0020	0,0400±0,0017
LIF	$s_{test}$	0,0409±0,0029	0,0411±0,0030	0,0403±0,0026	0,0405±0,0032
	$s_{opt}$	0,0398±0,0008	0,0397±0,0012	0,0395±0,0011	0,0398±0,0008
	$s_{svi}$	0,0404±0,0016	0,0404±0,0017	0,0399±0,0017	0,0402±0,0019
ASAM	$s_{test}$	0,0405±0,0028	0,0393±0,0031	0,0404±0,0027	0,0395±0,0028
	$s_{opt}$	0,0398±0,0009	0,0387±0,0018	0,0393±0,0013	0,0388±0,0017
	$s_{svi}$	0,0402±0,0016	0,0390±0,0022	0,0399±0,0017	0,0392±0,0019

Ulazni skup je skaliran u domen [-1,1], a izlazni skup u domen [0,1]. Prilikom inicijalizacije početnih parametara, prototip vektori su slučajnim izborom birani iz odgovarajućeg skupa za optimizaciju. Prototip vektori izabrani na ovaj način prosleđuju se razvijenim algoritmima optimizacije.

Proces optimizacije je sproveden 20 puta za sve testirane arhitekture, a prilikom svakog ponavljanja procesa vršena je nova inicijalizacija početnih prototip vektora veštačke neuronske mreže. U tabelama 2, 3 i 4 su prikazani rezultati (srednja vrednost i standardna devijacija statistike  $s$  za 20 ponavljanja).

Prvi izabrani test skup ulazno-izlaznih vrednosti za optimizaciju parametara i proveru tačnosti algoritma je izrazito kompleksan i nelinearan. Na osnovu eksperimentalnih rezultata (tabela 2) se može zaključiti da je očekivana tačnost reda veličine oko  $s_{test} = 0,1$ .

Najmanju vrednost test statistike ima ASAM algoritam i iznosi 0,0993 za veštačku neuronsku mrežu sa 15 procesirajućih jedinica u skrivenom sloju.

Rezultati ukazuju i da je standardna devijacija statistike  $s$  mala, na osnovu čega se može zaključiti da je ponovljivost rezultata velika. Trebalo bi naglasiti da referentni izvori navode vrednost test statistike  $s_{test} \approx [0,08, 0,09]$ .

Vrednosti dobijene u ovom istraživanju su za nijansu (~0,01) veće od algoritama optimizacije parametara veštačke neuronske mreže koji su prikazani u [6].

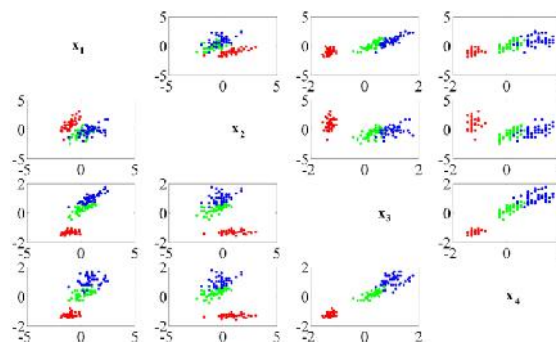
Tabela 4. Eksperimentalni rezultati za test skup – „Auto-mpg”.

Alg.	$s$	Broj neurona u skrivenom sloju			
		5	10	15	20
LKF	$s_{test}$	0,0923±0,0074	0,0945±0,0105	0,0977±0,0095	0,0856±0,0095
	$s_{opt}$	0,0872±0,0012	0,0871±0,0022	0,0872±0,0021	0,0756±0,0010
	$s_{svi}$	0,0898±0,0043	0,0910±0,0057	0,0927±0,0053	0,0809±0,0049
LIF	$s_{test}$	0,0932±0,0074	0,0946±0,0096	0,0933±0,0092	0,0860±0,0099
	$s_{opt}$	0,0869±0,0032	0,0872±0,0022	0,0870±0,0024	0,0752±0,0010
	$s_{svi}$	0,0902±0,0041	0,0911±0,0053	0,0904±0,0048	0,0809±0,0055
ASAM	$s_{test}$	0,0887±0,0080	0,0968±0,0073	0,0917±0,0060	0,0846±0,0065
	$s_{opt}$	0,0871±0,0015	0,0871±0,0019	0,0868±0,0024	0,0759±0,0011
	$s_{svi}$	0,0921±0,0039	0,0921±0,0039	0,0893±0,0031	0,0804±0,0033

U drugom primeru se može videti da test statistika ima vrednost oko  $s_{test} = 0,04$ , kao i da standardna devijacija statistike sračunata za 20 ponavljanja ima malu vrednost. Najmanja vrednost test statistike  $s_{test}$  je  $s_{test} = 0,0393$  koju generiše veštačka neuronska mreža sa 20 procesirajućih jedinica u skrivenom sloju čiji parametri su optimizirani ASAM algoritmom. Referentni izvori [6] navode vrednosti oko  $s_{test} = 0,06$ , što je značajno veća vrednost od vrednosti test statistike dobijene u ovom istraživanju. Poslednji izabrani test skup ima vrednost test statistike oko  $s_{test} = 0,09$ . S obzirom na to da drugi istraživači prijavljuju vrednosti  $s_{test} = 0,1165$  [6], zaključuje se da predloženi algoritmi imaju višu tačnost i preciznost. Eksperimentalni rezultati (tabele 2, 3 i 4) test statistike izračunate za sve skupove (svi primeri iz početnog skupa za optimizaciju  $s_{svi}$ , test primera  $s_{test}$  i primera koji su korišćeni tokom optimizacije parametara  $s_{opt}$ ) pokazuju da test statistika  $s$  ima isti red veličine za sve pomenute skupove, što ukazuje da su svi algoritmi konvergirali ka optimalni vrednostima.

### 3. IRIS TEST SKUP

Iris test skup je predložen još 30-tih godina prošlog veka za ocenu tačnosti predikcije klasifikacije estimatora [7]. Iris test skup sadrži 150 primera, koji su klasifikovani u jednu u od tri klase, a svaka klasa je predstavljena sa 50 primera [4]. Primere čine četiri karakteristike odgovarajuće klase koje su predstavljene ulaznim vektorom  $\mathbf{x} \in \mathbb{R}^4$ . Iris test skup je prikazan na slici 1.



Slika 1 - Ulazni skup podataka za Iris test skup. Duž glavne dijagonale su navedeni elementi vektora  $\mathbf{x} \in \mathbb{R}^4$ , koji predstavljaju osnovne karakteristike tri klase. Svaki dijagram pokazuje zavisnost jednog elementa ulaznog vektora  $\mathbf{x}_i$  u odnosu na drugi  $\mathbf{x}_j$

Ulazni skup podataka je normalizovan koristeći očekivanu vrednost  $\mu_x$  i standardnu devijaciju  $\sigma_x$  podataka u ulaznom skupu  $\bar{\mathbf{x}}_i = (\mathbf{x}_i - \mu_x) / \sigma_x$ . Izlazni skup podataka (tj. odgovarajuće klase) transformisan je u vektorsku formu  $\mathbf{y} \in \mathbb{R}^3$ , tako da svaki element ovog vektora predstavlja binarnu promenljivu koja određuje pripadnost  $k$ -tog ulaznog vektora  $j$ -toj klasi

$$\mathbf{y}_i^k = \begin{cases} 1 & \forall i = j \\ 0 & \forall i \neq j \end{cases} \quad (2)$$

Na ovaj način je izlazni skup već normalizovan. U problemima klasifikacije nije uobičajeno da se navodi srednja kvadratna greška ili njen kvadratni koren, ali se ipak mogu naći i takvi primeri. Problem

klasifikacije omogućava direktnu ocenu tačnosti predikcije estimatora izraženu u procentima. 70 % primera iz polaznog skupa je upotrebljeno za optimizaciju parametara veštačke neuronske mreže, a preosta-

lih 30% za testiranje. U tabeli 5 su dati eksperimentalni rezultati za 20 ponavljanja (greška klasifikacije i procenat uspešnosti klasifikacije).

Tabela 5 - Eksperimentalni rezultati za test skup – „Iris”

Alg.	s	Broj neurona u skrivenom sloju			
		5	10	15	20
LKF	Er	0,3635±0,0105	0,3611±0,0120	0,3619±0,0127	0,3645±0,0178
	%	94,8519±0,0105	95,3913±0,0120	94,5965±0,0127	95,1579±0,0178
LIF	Er	0,3662±0,0168	0,3683±0,0132	0,3635±0,0186	0,3603±0,0138
	%	94,6667±0,0168	95,1014±0,0132	94,7018±0,0186	95,6491±0,0138
ASAM	Er	0,3582±0,0138	0,3713±0,0192	0,3627±0,0175	0,3571±0,0132
	%	94,3333±0,0138	94,7246±0,0192	95,3333±0,0175	96,1053±0,0132

S obzirom na dužinu perioda rada istraživačke zajednice na Iris skupu podataka, on predstavlja jedan od najpoznatijih test skupova, zbog čega je i najduže izučavan. Od efikasnog estimatora se očekuje da dostigne tačnost od 95% u pogledu klasifikacije ulaznog vektora u jednu od tri klase. Na osnovu eksperimentalnih podataka datih u Tabeli 5, može se zaključiti da sva tri razvijena algoritma ostvaruju ciljani rezultat.

ASAM algoritam dostiže maksimum od 96,1053%, LKF 95,3913%, dok LIF 95,6491%. Zanimljivo je da tačnost klasifikacije veštačke neuronske mreže raste sa porastom broja neurona. U poređenju sa ostalim istraživačkim rezultatima [8] eksperimentalni rezultati dobijeni primenom tri nova algoritma mašinskog učenja veštačkih neuronskih mreža sa Gausovim aktivacionim funkcijama postizu istu ili za nijansu višu tačnost.

#### 4. PRIMENA VEŠTAČKIH NEURONSKIH MREŽA SA RADIJALNIM AKTIVACIONIM FUNKCIJAMA GAUSOVOG TIPNA NA MODELIRANJE TOPOLOGIJE OBRAĐENE POVRŠINE

U ovom delu su prikazani eksperimentalni rezultati modeliranja obrađene površine primenom veštačkih neuronskih mreža sa radijalnim aktivacionim funkcijama Gausovog tipa [3]. Osnovni cilj ovog eksperimentalnog procesa i prikaza eksperimentalnih rezultata nije ukazivanje na nedostatke postojećih metoda za modeliranje topologije obrađene površine, već želja da se pokažu mogućnosti primene razvijenih algoritama mašinskog učenja za modeliranje kompleksnih inženjerskih problema primenom analiziranih veštačkih neuronskih mreža. Važno je napomenuti da je u ovom eksperimentu topologija obrađene površine tretirana kao funkcionalna zavisnost, tj. kao preslikavanje ulaznog skupa u izlazni skup koje treba modelirati primenom veštačke neuronske mreže.

Eksperimentalni podaci hrapavosti obrađene površine dobijeni su na atomskom mikroskopu i dati su u vidu matrice tipa 512 x 512 (što znači da je ukupan broj primera 262,144). Zbog izrazito velikog broja primera, iz skupa ulazno-izlaznih parova slučajnim izborom su izabrani podskupovi tipa 30 x 30. Svi ulazni podaci su skalirani u domen [-1,1], a izlazni u domen [0,1]. 50% polaznog skupa je korišćeno za optimizaciju parametara veštačke neuronske mreže. Testirane su arhitekture sa 30 i 50 neurona u skrivenom sloju. Eksperimentalni rezultati izgleda izabranog segmenta obrađene površine nakon optimizacije parametara veštačke neuronske mreže dat je na slici 2.

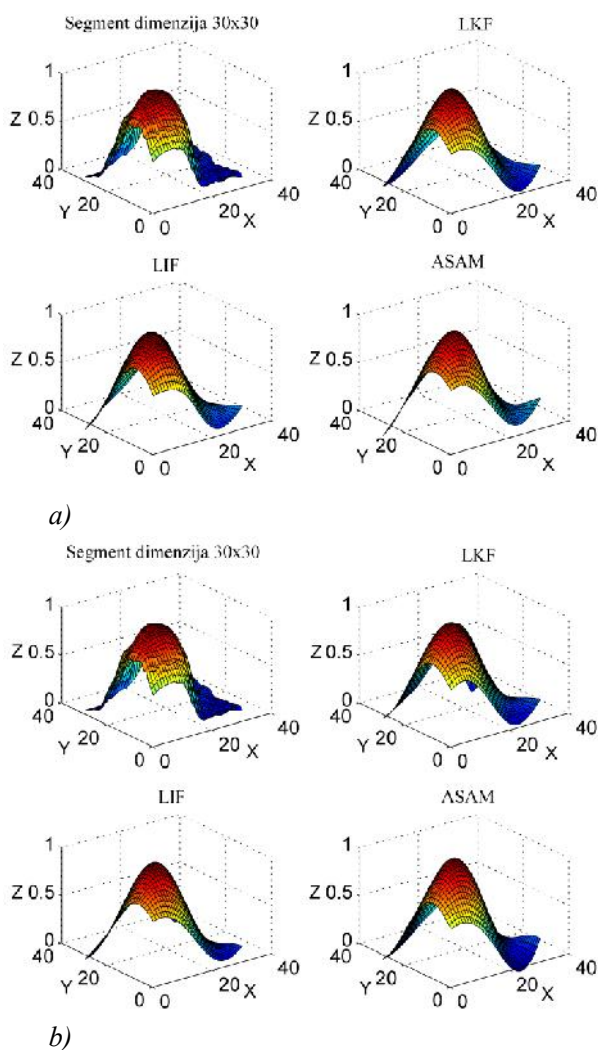
U tabeli 6 je prikazana KSKG (s statistika za sve primere) za testirane veštačke neuronske mreže sa 30 i 50 neurona u skrivenom sloju za 10 sprovedenih ponavljanja učenja i testiranja.

Tabela 6. Eksperimentalni rezultati modeliranja obrađene površine

Alg.	s	Broj neurona u skrivenom sloju	
		30	50
LKF	$S_{test}$	0,0560±0,0017	0,0594±0,0022
LIF	$S_{test}$	0,0556±0,0003	0,0579±0,0030
ASAM	$S_{test}$	0,0550±0,0004	0,0594±0,0061

Na osnovu grafičkog prikaza predikcije izgleda izabranog segmenta obrađene površine (slika 2 (a,b)), može se zaključiti da sva tri razvijena algoritma mašinskog učenja imaju problem sa neravnomernom promenom izlazne veličine, odnosno sa izrazitom nelinearnošću. Sva tri razvijena algoritma generišu glatku površ, što se može opravdati s obzirom na osnovnu formulaciju Kalmanovog filtra.

Zanimljivo je da veštačka neuronska mreža sa manjim brojem neurona u skrivenom sloju ima manju vrednost test statistike od veštačke neuronske mreže sa većim brojem neurona u skrivenom sloju. Međutim, ovaj zaključak ne može biti generalizovan, s obzirom na to da takva ocena zahteva teorijsku i eksperimentalnu analizu koja višestruko prevazilazi nameru ovog istraživanja. Može se uočiti i da je standardna devijacija test statistike za 10 ponavljanja veštačke neuronske mreže sa 30 neurona za red veličine manja kod LIF-a i ASAM-a u odnosu na LKF algoritam. Takođe, zanimljivo je da LKF i ASAM algoritmi mašinskog učenja imaju veću vrednost test statistike za arhitekturu sa 50 neurona od vrednosti test statistike LIF algoritma.



Slika 2 - Izgled segmenta obrađene površine: a) veštačka neuronska mreža sa 30 neurona u skrivenom sloju; b) veštačka neuronska mreža sa 50 neurona u skrivenom sloju

Razvoj teorijsko-eksperimentalnog modela topologije obrađene površine zahteva posebno istraživanje, a rezultati prikazani u ovom primeru predsta-

vljaju samo proveru karakteristika, tačnosti i performansi razvijenih algoritama mašinskog učenja veštačkih neuronskih mreža sa radijalnim aktivacionim funkcijama Gausovog tipa na ovom kompleksnom inženjerskom problemu.

## 5. DISKUSIJA

Mogućnosti inženjerske primene veštačkih neuronskih mreža sa radijalnim aktivacionim funkcijama Gausovog tipa protežu se od aproksimacije funkcionalnih zavisnosti, klasifikacije ulaznih podataka, procesiranja signala, identifikacije sistema do modeliranja procesa [1], [2], [3], [5]. U tom smislu, važno je još jednom naglasiti prednosti razvijenih algoritama mašinskog učenja, ali i navesti nedostatke.

Eksperimentalni rezultati razvijenih algoritama mašinskog učenja pokazuju dovoljan nivo tačnosti neophodnih za primenu u cilju modeliranja inženjerskih problema. Da bi se ocenila upotrebnost vrednosti razvijenih algoritama mašinskog učenja, izabrane veštačke neuronske mreže su testirane na „standardnim” skupovima podataka koji predstavljaju realne probleme regresije/klasifikacije iz različitih domena. Poređenjem sa različitim algoritmima mašinskog učenja, konstatovano je da razvijeni algoritmi generišu isti red veličine statistike za test skupove podataka (u nekim slučajevima i znatno bolje rezultate) kao i konvencionalni pristupi. Ovaj rezultat je od izuzetne važnosti za ocenu kvaliteta predikcije izabranih veštačkih neuronskih mreža nakon optimizacije parametara primenom razvijenih algoritama.

Parametri LKF-a, LIF-a i ASAM-a (početni vektor stanja  $\lambda_{00}$ , početna matrica kovarijansi  $P_{00}$ , matrica šuma sistema  $Q$  i matrica šuma merenja  $R$ ) određuju početnu tačku procesa ocenjivanja, početno znanje/neznanje o sistemu, promenu sistema tokom vremena i poverenje u informacije koje sistem prima iz okruženja. U poređenju sa parametrima determinističkih algoritama mašinskog učenja, parametri LKF-a, LIF-a i ASAM-a imaju dublju fizičko-matematičku interpretaciju. Na ovaj način se omogućava primena izabranog tipa veštačkih neuronskih mreža za različite vrste problema.

Algoritmi su sekvencijalnog tipa čime je omogućeno i on line mašinsko učenje. Ova prednost je iskorišćena tokom rešavanja problema simultanog ocenjivanja položaja mobilnog robota i karakterističnih objekata u okruženju [1]. Uobičajeni pristup u toj oblasti baziran je na ocenjivanju primenom linearizovanog Kalmanovog filtra. Da bi kompenzovali nepoznate stohastičke uticaje izvršena je integracija LKF algoritma ocenjivanja vektora stanja i veštačke neuronske mreže. Eksperimentalni rezultati i zaključci izvedeni tokom razvoja novih algoritama mašinskog učenja

(LKF, LIF i ASAM) predstavljaju osnovu ideje o integraciji LKF algoritma simultanog ocenjivanja položaja mobilnog robota i karakterističnih objekata sa veštačkim neuronskim mrežama [1].

Međutim, pored prednosti u pogledu mogućnosti modeliranja kompleksnih dinamičkih sistema, trebalo bi naglasiti i nedostatke koji su zapaženi tokom sprovođenja eksperimentalnog procesa. Jedan od glavnih nedostataka odnosi se na numeričke nestabilnosti tokom procesa ocenjivanja. Često se dešavalo da usled nelinearnosti problema inverzija matrica postane nemoguća zbog numeričkih nepreciznosti. Razlog za pojavu nestabilnosti ove vrste je pre svega u „gruboj” primeni izvedenih jednakosti. Ovaj problem se jednostavno može otkloniti veštačkim povećavanjem ili smanjivanjem matrice kovarijansi  $\mathbf{P}_{k|k}$ . Primećen je i još jedan problem koji je direktno vezan za konvergenciju očekivane greške učenja. Ukoliko je problem nelinearan, a svi realni problemi to i jesu (test skupovi: Abalone, Wether Ankara, Auto-mpg, Iris i topologija obrađene površine), problem konvergencije greške postaje značajan. Uočeno je da algoritam konvergira ka jednoj vrednosti greške učenja, što rezultira ili završetkom procesa ili značajnim povećanjem broja iteracija. Konvergencija algoritma mašinskog učenja (ili algoritma optimizacije u opštem slučaju) ka lokalnom minimumu nije tako redak slučaj. U slučaju primene LKF-a, LIF-a i ASAM-a pomenuti problem se može rešiti direktnom promenom matrice kovarijansi u realnom vremenu tokom trajanja procesa učenja. Na taj način bi se izbeglo povremeno „zarobljavanje” algoritma u lokalnom minimumu.

Dalji razvoj LKF, LIF i ASAM algoritama mašinskog učenja će biti posvećen sposobnosti promene arhitekture veštačke neuronske mreže (tj. promene broja neurona) tokom trajanja učenja (u *on line* režimu) [2], [6], [9]. Ovakva sposobnost adaptacije bi značajno doprinela smanjenju broja parametara potrebnih za kodiranje ulazne informacije, a time i povećanju brzine učenja. U tom smislu, autori su već preduzeli odgovarajuće korak i u referenci [2] je uvedena i analizirana veštačka neuronska mreža za aproksimaciju funkcija čije aktivacione funkcije predstavljaju uopštenje Gausove funkcije.

Drugim rečima, radijalna distanca  $\sigma_j$  (pogledati prvi deo ovog rada) više nije skalarna promenljiva  $\sigma \in \mathbb{R}^1$  već vektorska  $\sigma \in \mathbb{R}^{n_x}$ , čime je omogućeno skaliranje duž svih dimenzija ulaznog vektora  $\mathbf{x}_i \in \mathbb{R}^{n_x}$  tokom trajanja procesa učenja. Za detaljne informacije u vezi samog modela veštačke neuronske mreže i njene primene za aproksimaciju funkcija i identifikaciju nelinearnih sistema čitalac se upućuje na referencu [2].

Zanimljivo bi bilo videti da li bi integracija nenađgledanog mašinskog učenja [10], [11], [12] sa postojećim nadgledanim vidom mašinskog učenja rezultirala boljim rezultatima (npr. odgovarajući algoritam nenađgledanog mašinskog učenja generiše početne vrednosti vektora stanja  $\lambda_{0|0}$  i početne matrice kovarijansi  $\mathbf{P}_{0|0}$ ). Na ovaj način bi se iskoristio maksimum raspoloživih informacija potrebnih za inicijalizaciju algoritma nadgledanog mašinskog učenja. Sva navedena poboljšanja će biti u fokusu u budućim istraživanjima.

Analiza konvergencije razvijenih sekvencijalnih algoritama mašinskog učenja veštačkih neuronskih mreža sa radijalnim aktivacionim funkcijama Gausovog tipa je prikazana u [1]. Kao osnovni zaključak ove analize može se navesti da za iste početne vrednosti, ASAM algoritam brže konvergira od LKF-a i LIF-a, tj. potreban je manji broj iteracija za dostizanje željene ciljane vrednosti  $Er_{\min}$ . Drugi zaključak se odnosi na broj iteracija potrebnih za konvergenciju LKF-a i LIF-a. LKF-u je potreban veći broj iteracija od LIF-a za dostizanje željene vrednosti greške učenja  $Er_{\min}$ .

Pored ovog zaključka, trebalo bi ukazati i na uticaj parametara razvijenih algoritama mašinskog učenja. Može se uočiti i da smanjenje početne kovarijanse  $\mathbf{P}_{0|0}$  vektora stanja  $\lambda$  rezultira povećanjem broja iteracija potrebnih za uspešnu konvergenciju. Eksperimentalni podaci prikazani u okviru analize konvergencije LKF-a, LIF-a i ASAM-a dati u [1] ukazuju i na to da povećanje početne kovarijanse  $\mathbf{P}_{0|0}$  vektora stanja  $\lambda$  smanjuje broj potrebnih iteracija za dostizanje greške učenja  $Er_{\min}$ . Na osnovu sprovedene analize i rezultata prikazanih u [1] može se zaključiti da povećanje početne kovarijanse sistema  $\mathbf{P}_{0|0}$  za rezultat ima smanjenje broja potrebnih iteracija za red veličine. Ovaj rezultat je od velikog značaja, s obzirom na njen uticaj na konvergenciju razvijenih algoritama mašinskog učenja. Za detaljne informacije u vezi eksperimentalne analize konvergencije pogledati reference [1] i [2].

Povećanje matrice kovarijansi šuma sistema  $\mathbf{Q}$  utiče na smanjenje broja iteracija. Ovaj rezultat je logična posledica koraka predikcije Kalmanovog filtra, s obzirom na to da predikcija kovarijanse sistema predstavlja zbir trenutne najbolje ocene kovarijanse stanja sistema i matrice šuma sistema (prvi deo ovog rada, algoritmi 1, 2 i 3).

Veće vrednosti matrice kovarijansi šuma merenja  $\mathbf{R}$  takođe povećavaju broj iteracija. Kovarijansa vektora inovacije zavisi od matrice kovarijansi šuma

merjenja  $R$ , a njeno povećanje smanjuje pojačanje  $K$ , što direktno utiče na korak korekcije filtra.

## 6. ZAKLJUČAK

U ovom radu su prikazana tri nova sekvencijalna algoritma nadgledanog mašinskog učenja za veštačke neuronske mreže sa radijalnim aktivacionim funkcijama Gausovog tipa na bazi Kalmanovog filtra. U cilju primene algoritama prvi korak predstavlja formulisanje problema mašinskog učenja u vidu linearnog dinamičkog sistema u diskretnom obliku. Prvi algoritam primenjuje linearizovani Kalmanov filter, drugi je baziran na linearizovanom informacionom, dok treći na poseban način aproksimira prva dva momenta Gausove raspodele. U prvom delu ovog rada prikazani su matematički modeli za sva tri algoritma, dok su u drugom delu rada prikazani eksperimentalni rezultati ostvareni primenom na realnim problemima iz inženjerske prakse.

Razvijeni algoritmi mašinskog učenja veštačkih neuronskih mreža sa radijalnim aktivacionim funkcijama Gausovog tipa mogu se primeniti na obe osnovne klase zadataka: regresija i klasifikacija. Razvoj ovih algoritama mašinskog učenja je analiziran i sproveden uzimajući u obzir i opšti karakter primene veštačke neuronske mreže za širi spektar problema. Drugim rečima, proizvoljni problemi iz inženjerskog domena se mogu modelirati uz primenu veštačke neuronske mreže čiji su parametri modifikovani korišćenjem nekog od tri razvijena algoritma. Eksperimentalni rezultati pokazuju da veštačka neuronska mreža koja je obučena primenom ovih algoritama generiše bolje rezultate (bolje u smislu izabrane test statistike) od sličnih algoritama.

## ZAHVALNOST

Rad je nastao u okviru istraživanja na projektu „Inovativni pristup u primeni inteligentnih tehnoloških sistema za proizvodnju delova od lima zasnovan na ekološkim principima“ (TR-35004), koji je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Vlade Republike Srbije.

## LITERATURA

- [1] Vuković, N., Razvoj mašinskog učenja inteligentnog mobilnog robota baziran na sistemu veštačkih neuronskih mreža, doktorska disertacija, Univerzitet u Beogradu - Mašinski fakultet, 2012. (DOI: 10.2298/BG20120928VUKOVIC)
- [2] Vuković, N., Miljković, Z., A Growing and Pruning Sequential Learning Algorithm of Hyper Basis Function Neural Network for Function Approximation, *Neural Networks*, 46C, pp. 210-226, 2013. (DOI: 10.1016/j.neunet.2013.06.004)
- [3] Vuković, N., Miljković, Z., Babić, B., Bojović, B., Training of Radial Basis Function Networks with  $H_{\infty}$  Filter-Initial Simulation Results, *Proceedings of the 6th International Working Conference "Total Quality Management – Advanced and Intelligent Approaches"*, pp. 163-168, Belgrade, Serbia, 2011.
- [4] Bache, K. & Lichman, M. (2013). UCI Machine Learning Repository <http://archive.ics.uci.edu/ml>.
- [5] Irvine, CA: University of California. (poslednji datum pristupa: 21 maj 2014. godine)
- [6] Miljković, Z., Aleksendrić, D., Veštačke neuronske mreže - zbirka rešenih zadataka sa izvodima iz teorije, Univerzitet u Beogradu - Mašinski fakultet, Beograd 2009.
- [7] Bortman, M., Aladjem M., A Growing and Pruning Method for Radial Basis Function Networks, *IEEE Transactions on Neural Networks*, 20(6), pp. 1039-1045, 2009.
- [8] Fisher, R., The Use of Multiple Measurements in Taxonomic Problems, *Annals of Eugenics*, 7, pp. 179-188, 1936.
- [9] Simon, D, Training Radial Basis Function Neural Networks with the Extended Kalman Filter, *Neurocomputing*, 48, pp. 455-475, 2001.
- [10] Huang, G. B., Saratchandran, P., Sundararajan, N., A generalized growing and pruning RBF (GGAP-RBF) neural network for function approximation. *IEEE Transactions on Neural Networks*, 16(1), pp. 57-67, 2005.
- [11] Lazaro, M., Santamari I., Pantaleo C., A new EM-based training algorithm for RBF networks, *Neural Networks*, 16, pp. 69-77, 2003.
- [12] Schwenker, F., Kestler, H.A., Palm, G., Three learning phases for radial-basis-function networks. *Neural Networks*, 1(4-5), pp. 439-58, 2001.
- [13] Vuković, N., Miljković, Z., Babić, B., Čović, N., Towards implementation of intelligent mobile robots in a manufacturing environment, *Proceedings of the 4th International Conference on Manufacturing Engineering (ICMEN 2011)*, pp 367-376, Thessaloniki, Greece, 2011.

**SUMMARY****MACHINE LEARNING OF RADIAL BASIS FUNCTION NEURAL NETWORK BASED ON KALMAN FILTER - IMPLEMENTATION**

*In this paper we test three new sequential machine learning algorithms for radial basis function (RBF) neural network based on Kalman filter theory. Three new algorithms are derived: linearized Kalman filter, linearized information filter and unscented Kalman filter. Having introduced and derived mathematical model of each algorithm in the previous part of the paper, in this part we test and assess their performance using standard test sets from machine learning community. RBF neural network and three developed algorithms are implemented in MATLAB® programming environment. Experimental results obtained on real data sets as well as on real engineering problem show that developed algorithms result in more accurate models of the problem being investigated based on radial basis function neural network.*

**Key words:** *machine learning, artificial neural network, Kalman filter*