

FAKULTETU TEHNIČKIH NAUKA U NOVOM SADU

OBRAZAC ZA PRIJAVU TEHNIČKOG REŠENJA

Obavezni podaci:

Autor/Autori rešenja:

Vuk Vranjković, Rastislav Struharik

Naziv tehničkog rešenja:

Rekonfigurabilno IP jezgro za akceleraciju homogenih i heterogenih ansambala prediktivnih modela

Kategorija tehničkog rešenja:

Softver, M85

Za koga je rešenje rađeno i u okviru kog projekta MNTR:

- Fakultet tehničkih nauka u Novom Sadu
- Projekat tehnološkog razvoja TR-32016

Tehničko rešenje je prihvatio i koristi:

Fakultet tehničkih nauka u Novom Sadu za potrebe daljih istraživanja

Godina kada je rešenje urađeno:

2016.

Kako su rezultati verifikovani i od strane kog tela:

- Razvojem HDL opisa IP jezgara i njihovom verifikacijom kao i proverom performansi na FPGA kolu.
- Rezultate je verifikovalo Naučno-nastavno veće Fakulteta tehničkih nauka u Novom Sadu.

Oblast na koju se tehničko rešenje odnosi:

Hardversko ubrzavanje algoritama u FPGA tehnologiji.

Problemi koji se tehničkim rešenjem otklanjaju ili minimizuju:

Jedinstvena arhitektura realizovana kao IP jezgro u predlogu tehničkog rešenja može da implementira heterogene i homogene ansamble koji se sastoje od tri najčešće korišćena tipa klasifikatora u oblasti mašinskog učenja: DT (Decision Trees, stabla odluka), ANN (Artificial Neural Networks, veštačke neuronske mreže) i SVM (Support Vector Machine). Arhitektura omogućuje i višestruku implementaciju pojedinačnih prediktivnih modela. Pri tome se dobija značajno povećana protočnost podataka, a kao posledica toga i ubrzanje procesa klasifikacije. Veoma je mali broj hardverskih arhitektura koje implementiraju ansamble. Pri tome niti jedna arhitektura ne podržava implementaciju heterogenih ansambala niti je u stanju da implementira homogene ansamble sastavljene od različitih vrsta prediktivnih modela. Po tome je predloženo jezgro jedinstveno.

Stanje rešenosti pitanja istog problema u svetu:

U oblasti hardverskih implementacija algoritama mašinskog učenja postoji veliki broj radova koji predlažu različite arhitekture. Što se tiče stabala odluka arhitekture za treniranje ovih prediktivnih modela su prikazane u [1] i [2]. Arhitekture samih prediktivnih modela su prikazane u [3], [4] i [5].

Od pojavljivanja SVM prediktivnih modela u [6], postoji veliki interes za hardversku implementaciju ovih modela, zbog velike računске složenosti. Hardverske implementacije za treniranje SVM modela su objavljene u [7], [8] i [9]. Implementacije samih klasifikatora su razmatrane u [10] i [11].

ANN su klasifikatori kojoj je istraživačka zajednica posvetila najveću pažnju. Ovi klasifikatori su i najduže u upotrebi pa postoje mnogobrojne hardverske implementacije od kojih su neke: [12], [13], [14] i [15].

Istraživačka zajednica predložila je značajno manje rešenja za hardversku implementaciju ansambala prediktivnih modela u odnosu na pojedinačne prediktivne modele. Jedna hardverska arhitektura koja implementira homogene ansamble stabala je prikazana u [16]. Niti jedna arhitektura ne implementira heterogene ansamble.

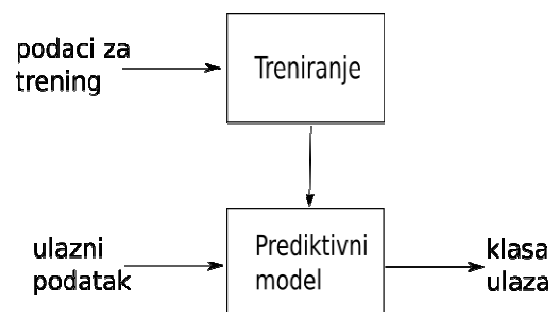
Reference:

- [1] Rastislav JR Struharik and Ladislav A Novak. Evolving decision trees in hardware. Journal of Circuits, Systems, and Computers, 18(06): 1033–1060, 2009.
- [2] Grigorios Chrysos, Panagiotis Dagritzikos, Ioannis Papaefstathiou, and Apostolos Dollas. Hc-cart: A parallel system implementation of data mining classification and regression tree (cart) algorithm on a multi- fpga system. ACM Trans. Archit. Code Optim., 9(4):47:1–47:25, January 2013.
- [3] Rastislav JR Struharik and Ladislav A Novak. Hardware implementation of decision tree ensembles. Journal of Circuits, Systems, and Computers, 22(05), 2013.
- [4] Rastislav JR Struharik and Ladislav A Novak. Intellectual property core implementation of decision trees. IET computers & digital techniques, 3(3):259–269, 2009.
- [5] JR Struharik. Implementing decision trees in hardware. In Intelligent Systems and Informatics (SISY), 2011 IEEE 9th International Symposium on, pages 41–46. IEEE, 2011.
- [6] Corinna Cortes and Vladimir Vapnik. Support-vector networks. Machine learning, 20(3):273–297,

- 1995.
- [7] Davide Anguita, Andrea Boni, and Sandro Ridella. A digital architecture for support vector machines: theory, algorithm, and fpga implementation. *Neural Networks, IEEE Transactions on*, 14(5):993–1009, 2003.
 - [8] Ta-Wen Kuan, Jhing-Fa Wang, Jia-Ching Wang, Po-Chuan Lin, and Gaung-Hui Gu. Vlsi design of an svm learning core on sequential minimal optimization algorithm. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 20(4):673–683, 2012.
 - [9] Chih-Hsiang Peng, Po-Chuan Lin, Shovan Barma, Jhing-Fa Wang, Hong-Yuan Peng, Karunanithi Bharanitharan, and Ta-Wen Kuan. Low-power enhanced system-on-chip design for sequential minimal optimisation learning core with tri-layer bus and butterfly-path accelerator. *IET Computers & Digital Techniques*, 2015.
 - [10] Christos Kyrkou and Theodoris Theodoridis. A parallel hardware architecture for real-time object detection with support vector machines. *Computers, IEEE Transactions on*, 61(6):831–842, 2012.
 - [11] V Vranjkovic and Rastislav Struharik. New architecture for svm classifier and its application to telecommunication problems. In *Telecommunications Forum (TELFOR)*, 2011 19th, pages 1543–1545. IEEE, 2011.
 - [12] Amos R Omondi and Jagath Chandana Rajapakse. *FPGA implementations of neural networks*, volume 365. Springer, 2006.
 - [13] Hirokazu Madokoro and Kazuhito Sato. Hardware implementation of back-propagation neural networks for real-time video image learning and processing. *Journal of Computers*, 8(3):559–566, 2013.
 - [14] Antony Savich, Medhat Moussa, and Shawki Areibi. A scalable pipelined architecture for real-time computation of mlp-bp neural networks. *Microprocessors and Microsystems*, 36(2):138–150, 2012.
 - [15] Dmitri Vainbrand and Ran Ginosar. Scalable network-on-chip architecture for configurable neural networks. *Microprocessors and Microsystems*, 35(2):152–166, 2011.
 - [16] Rastislav JR Struharik and Ladislav A Novak. Hardware Implementation of Decision Tree Ensembles. *Journal of Circuits, Systems, and Computers*, 2013.

Tehnički detalji predloženog rešenja:

Sistemi mašinskog učenja se sastoje od dela za treniranje prediktivnog modela i samog prediktivnog modela, kao što je prikazano na slici 1.

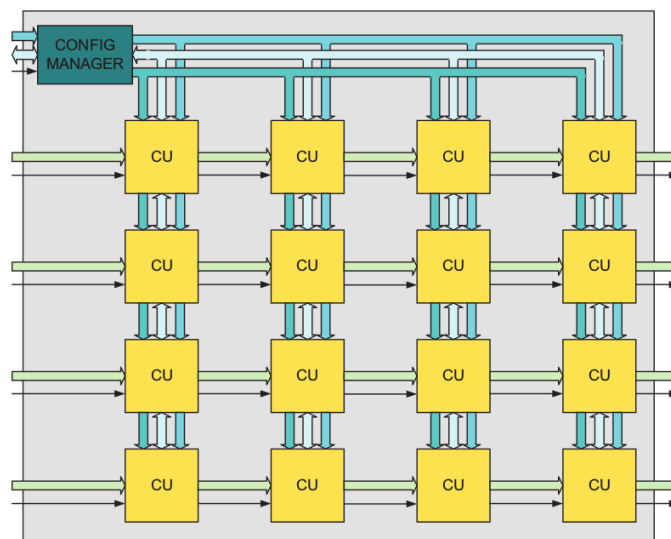


Slika 1: Mašinsko učenje

Ulazni podaci u prediktivne modele se mogu prikazati kao vektor realnih brojeva. Najčešće korišćeni

prediktivni modeli, ANN, SVM i DT sadrže veliki broj istovetnih operacija: računanje proizvoljne nelinearne funkcije jedne realne promenljive, oduzimanje dva vektora, skalarni proizvod dva vektora, dodavanje skalara. REC (Reconfigurable Ensemble Classifier) IP jezgro predstavlja univerzalnu digitalnu rekonfigurabilnu arhitekturu koja može da radi kao ansambl navedenih prediktivnih modela ili kao veći broj ovakvih klasifikatora u paraleli. Za proračun različitih prediktivnih modela bi se koristile samo neke od navedenih operacija u odgovarajućem redosledu.

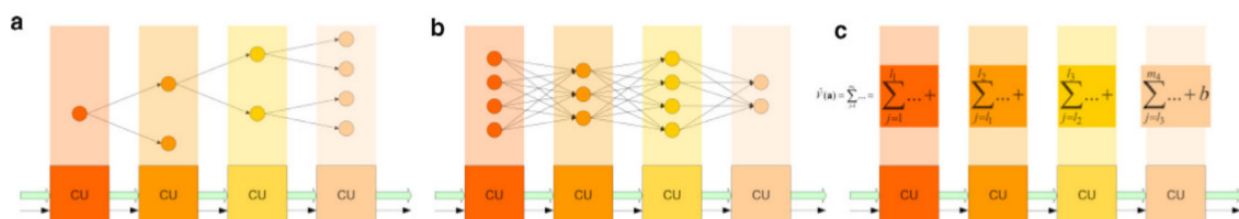
Arhitektura je organizovana kao matrica istovetnih blokova, kao što je prikazano na slici 2. Blokovi se nazivaju CU (Computation Unit). Modul Config Manager omogućava, tokom rekonfiguracije, da se odabere samo jedan CU blok. U toku implementacije REC jezgra, pre sinteze, moguće je korišćenjem konfiguracionih parametrima odrediti broj CU jedinica po vrstama i kolonama, pa se arhitektura može prilagoditi zadacima na kojima će biti primenjena. Na ovaj način obezbeđena je odlična skalabilnost arhitekture, pošto se ona može prilagoditi i veličini raspoloživih resursa.



Slika 2: Struktura REC jezgra

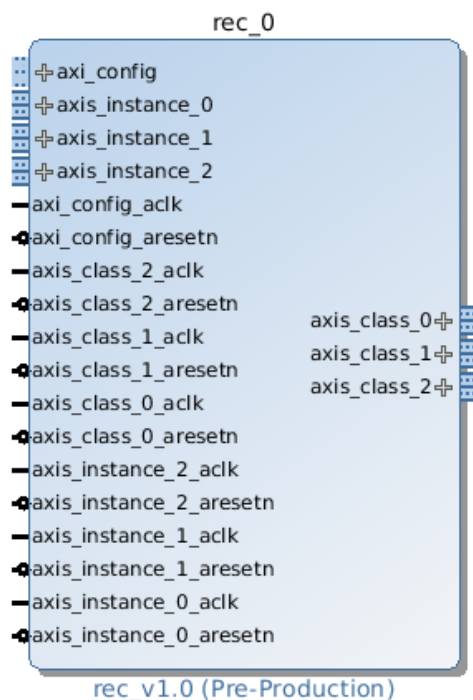
Svaka vrsta REC arhitekture može da implementira pojedinačni prediktivni model. Sve vrste koje implementiraju neki od modela mogu da se iskombinuju u jedan ansambl. Interpretacija računanja CU modela zavisi od konfiguracije REC arhitekture, kao što je prikazano na slici 3. Ukoliko vrsta REC arhitekture treba da implementira ANN tada CU blok predstavlja sloj neuronske mreže. Kada vrsta implementira stablo odluke tada CU računa neki od čvorova u jednom nivou stabla prediktivnog modela. I na kraju, ukoliko vrsta REC jezgra implementira SVM, tada CU modul računa deo sume neophodne za rezultat predikcije SVM modela. Arhitektura je organizovana kao matrica. Pošto svaka vrsta predstavlja pojedinačni klasifikator, to znači da svi pojedinačni prediktivni modeli mogu da rade zajedno u paraleli. Svaki pojedinačni klasifikator je podeljen u više CU jedinica. Dok jedna CU jedinica obrađuje jednu instancu, ostale CU jedinice mogu da računaju vrednosti za neke druge instance. Na ovaj način je ostvarena velika protočnost arhitekture. Svaka CU jedinica dodaje paralelizam u arhitekturu koji može

biti efektivno iskorišćen. Ovaj paralelizam dovodi do velikih ubrzanja koje arhitektura ostvaruje u odnosu na postojeća softverska rešenja.



Slika 3: Način korišćenja CU blokova za implementaciju različitih tipova prediktivnih modela

Za razvoj REC IP jezgra korišćen je Xilinx Vivado. Arhitektura je upakovana u jezgro uz pomoć Xilinx IP integratora. REC IP je modelovan na RT nivou abstrakcije. Izgled interfejsa REC IP jezgra unutar Xilinx alata je prikazan na slici 4.



Slika 4: REC IP

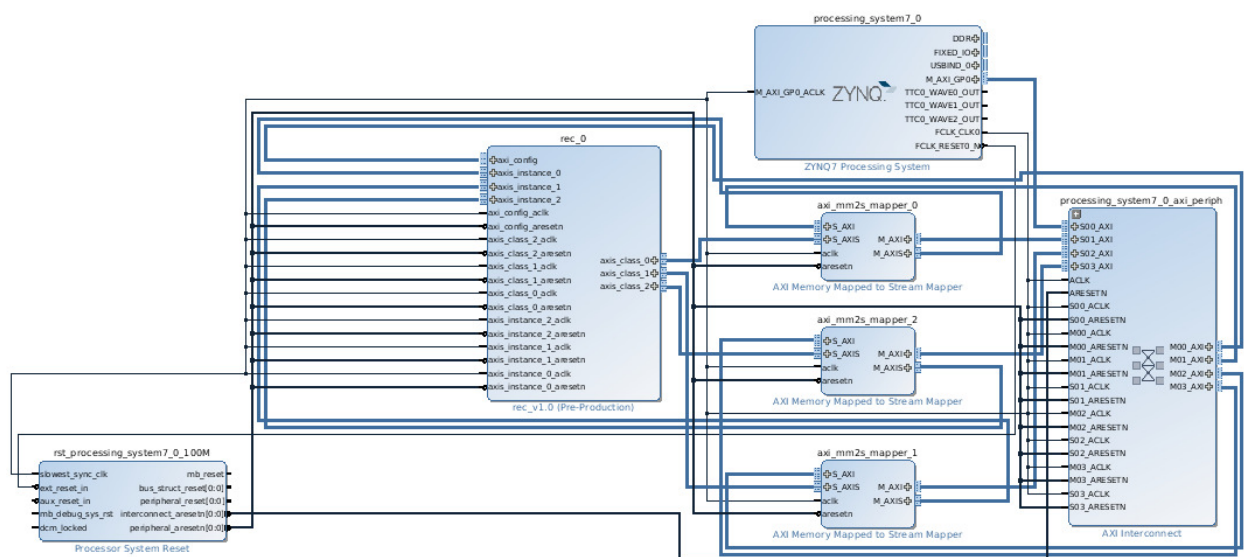
Xilinx IP integrator koristi skup standardnih magistrala koje poštuju AXI (AMBA eXtensible Interface) protokol, koji je deo AMBA (Advanced Microcontroller Bus Architecture) specifikacije. Prilikom projektovanja IP jezgara, unutar IP integratora, korisniku su na raspolaganju tri vrste protokola: AXI Full, AXI Lite i AXI Stream. Za konfigurisanje REC jezgra odabran je AXI Full protokol. REC jezgro sadrži veliki broj registara i memorija. AXI full je predviđen da se koristi za prenos podataka od mastera ka slejvu kada slejv sadrži i registre i memoriju. Iako je ovaj protokol najkompleksniji za implementaciju, on omogućuje brz transfer podataka, što je neophodno za efikasnu konfiguraciju memorija unutar REC IP-a. Ovaj

interfejs je slejv tipa.

REC IP može da ima različit broj interfejsa koji primaju instance i koji šalju rezultate klasifikacije. Broj interfejsa zavisi od broja predviđenih pojedinačnih prediktivnih modela za implementaciju. Za svaki pojedinačni prediktivni model, potreban je po jedan slejv AXI stream interfejs i po jedan master AXI stream interfejs. Pomoću slejv AXI stream interfejsa IP jezgro prima ulaznu instancu u odgovarajući pojedinačni klasifikator. Rezultat klasifikacije prediktivnog modela se šalje ostatku sistema pomoću AXI master interfejsa. Na slici 4, prikazano je REC IP jezgro koje radi sa tri pojedinačna prediktivna modela. AXI stream protokol je odabran za ove interfejse jer je najjednostavniji za implementaciju i daje odličnu protočnost.

Firma Xilinx u ponudi ima FPGA čipove koji sadrže integrisan konfigurabilni procesor i rekonfigurabilnu logiku i čipove koji imaju samo rekonfigurabilnu logiku. U prvoj grupi se nalazi i Zynq-7000 familija čipova u kojoj su integrisana dva hard IP ARM jezgra. Druga vrsta čipova ima samo rekonfigurabilnu logiku. Ovoj vrsti FPGA pripadaju Artix, Kintex i Virtex familije čipova. Prilikom projektovanja SoC dizajna, za procesorsku namenu na raspolaganju stoji Microblaze soft IP jezgro. Hard IP jezgra pružaju bolje performanse i manju konfigurabilnost u odnosu na soft IP jezgra. U narednim paragrafima biće opisano kako se REC IP jezgro može ugraditi u FPGA SoC, u zavisnosti od toga da li je na raspolaganju hard IP procesor, Zynq ARM, ili soft IP procesor, Microblaze.

Postoji više načina kako se REC IP jezgro može povezati sa Xilinx Zynq procesorom. Jedan od načina je prikazan na slici 5.



Slika 5: Povezivanje REC IP jezgra i Zynq procesora

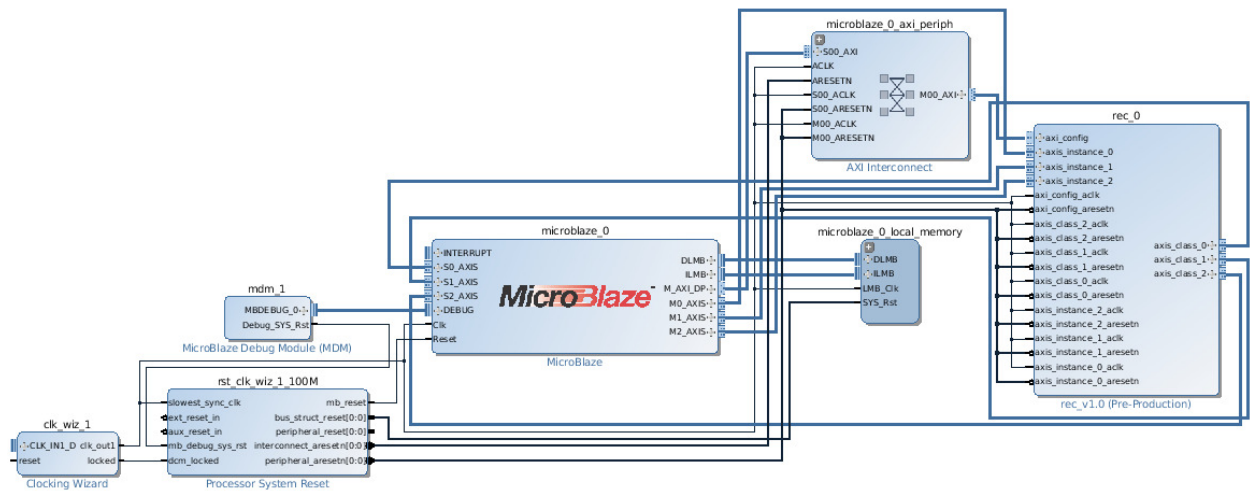
Naravno, prikazani način nije i jedini. Konkretno povezivanje zavisi od potreba samog dizajna. Ovo je samo ilustracija kako se REC jezgro može koristiti u SoC-u. Prva bitna komponenta u ovom primeru je AXI Interconnect jezgro. Ono je standardno jezgro koje se nalazi u standardnom repozitorijumu jezgara firme

Xilinx. Uobičajena primena ovog IP bloka je da se postavi između procesora i ostalih perifernih sistema. Ovaj IP obezbeđuje svim periferijama u sistemu jedinstven adresni prostor za softver koji se izvršava na Zynq ARM procesoru. AXI Interconnect slejv interfejs S00_AXI je povezan sa Zynq ARM AXI magistralom. Procesor preko ovog interfejsa šalje sve potrebne podatke ostalim delovima SoC-a. Na M00_AXI master interfejs AXI Interconnect jezgra je povezan konfiguracioni interfejs REC jezgra. Kada procesor pošalje podatak koji pripada adresnom prostoru ovog interfejsa AXI Interconnect jezgra, biće inicirana transakcija i odgovarajući registar ili memorija unutar REC jezgra će biti promenjeni, odnosno pročitani.

U prikazanom SoC-u se nalaze tri periferije koje konvertuju AXI Stream protokol u AXI light ili full. AXI stream protokol nema koncept adrese. Pošto je potrebno dodeliti AXI REC Stream interfejsima memorijske prostore, da bi im procesor mogao pristupiti, to se može uraditi preko Xilinx-ovih standardnih komponenti AXI Memory Mapped to Stream Mapper. Pošto u ovom primeru REC ima tri para AXI stream interfejsa za primanje instanci i slanje rezultata, potrebno je koristiti tri instance AXI Memory Mapped to Stream Mapper jezgra da bi se uradila konverzija protokola. Procesor ulaznu instancu, namenjenu prvom klasifikatoru, šalje preko svoje AXI magistrale. AXI Interconnect jezgro ga prosleđuje preko M01_AXI interfejsa do prvog konvertora i njegovog slejv interfejsa S_AXI. Protokol se konvertuje u AXI Stream i šalje se preko master interfejsa M_AXIS do ulaznog interfejsa za instance REC IP jezgra, axis_instance_0. Kada REC završi sa proračunom, rezultat se šalje konvertoru preko master AXI Stream interfejsa axis_class_0. Konvertor prebacuje informaciju u AXI Light/Full protokol i preko svog interfejsa M_AXI prosleđuje podatak AXI Interconnect jezgru, interfejsu S01_AXI. Rezultat se zatim prosleđuje procesoru. Na potpuno isti način su povezani i ostali interfejsi REC IP jezgra za ulazne instance i rezultate klasifikacije.

U slučaju da REC jezgro za ulazne instance i rezultate koristi AXI Lite ili Full protokol, konvertori bi se mogli ukloniti iz ovog dizajna. Sam procesor u Zynq familiji čipova može da radi na višestruko većoj frekvenciji od programabilne logike. Iz ovog razloga brzina slanja instanci i primanja rezultata nije posebno kritična. U FPGA familijama bez procesora ovo može biti veoma kritično za protočnost sistema. U tim familijama, kao što su Kintex i Virtex, odluka da se za ove interfejse koristi AXI Stream protokol pokazuje se opravdanom. Ovo će biti ilustrovano na primeru povezivanja REC IP jezgra sa Xilinx Microblaze procesorom.

Microblaze procesor je soft IP jezgro. To znači da se implementira u programabilnom delu FPGA čipa i da ima slične performanse kao i ostatak perifernih sistema. Frekvencije koje ovaj procesor može da postigne su višestruko slabije od Zynq ARM procesora. Microblaze ima mogućnost da mu se dodaju direktno AXI Stream magistrale, za razliku od Zynq ARM procesora. To znači da se REC IP jezgro može vezati direktno na Microblaze procesor. Brzina transfera je u ovom slučaju veoma bitna i zato se opravdava izbor AXI Stream protokola za interfejse koji rade sa instancama i rezultatima. Na slici 6, prikazan je Microblaze procesor kome su dodate tri AXI Stream slejv magistrale i tri AXI Stream master magistrale. Pomoću njih se Microblaze i REC IP jezgro mogu direktno povezati, kao što je prikazano na slici. I u ovom slučaju je veza sa axi_config interfejsom ostvarena preko AXI Interconnect jezgra, pošto će u SoC-u verovatno biti još dodatnih perifernih sistema. Ukoliko je REC IP jezgro jedina periferija u SoC-u, tada se može direktno vezati axi_config interfejs na M_AXI_DP magistralu procesora.



Slika 6: Povezivanje Microblaze procesora i REC jezgra

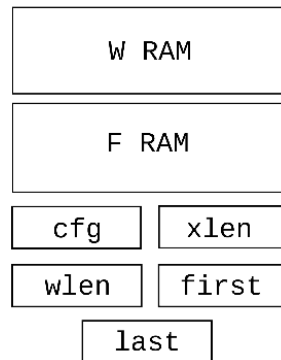
Adrese za REC jezgro se nameštaju pomoću alata Address Editor. Address Editor je deo Vivado IP integratora. Slika 7 prikazuje jednu moguću dodelu adresa za primer vezivanja REC jezgra sa Zynq ARM procesorom.

Cell	Slave Interface	Base Name	Offset Address	Range	High Address
processing_system7_0					
Data (32 address bits : 0x40000000 [1G])					
rec_0	axi_config	axi_config_reg	0x43C0_0000	64K	0x43C0_FFFF
axi_mm2s_mapper_0	S_AXI	Reg	0x43C1_0000	64K	0x43C1_FFFF
axi_mm2s_mapper_1	S_AXI	Reg	0x43C2_0000	64K	0x43C2_FFFF
axi_mm2s_mapper_2	S_AXI	Reg	0x43C3_0000	64K	0x43C3_FFFF
axi_mm2s_mapper_0					
axi_mm2s_mapper_1					
axi_mm2s_mapper_2					

Slika 7: REC IP dodela adrese

Adresa koja je dodeljena axi_config interfejsu je 0x43C00000. Prilikom razvoja softvera, sve adrese koje budu u opsegu od 0x43C00000 do 0x43C00000+64K biće namenjene za komunikaciju sa axi_config interfejsom. Na primer, kada se pristupi adresi 0x43C00001, informacije će teći od ili ka axi_config interfejsu i njegovom registru koji je na adresi 1 unutar samog REC IP jezgra.

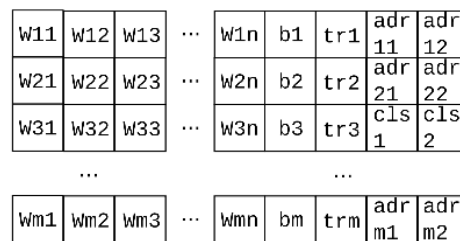
Programski model REC arhitekture je jednostavan. REC IP jezgro je organizovano kao matrica CU blokova. Svaki od CU blokova ima baznu adresu. Ukoliko je potrebno pristupiti nekom od registara ili memorija unutar odgovarajućeg CU bloka, potrebno je na baznu adresu dodati lokalnu adresu odgovarajućeg registra. Svaki od CU blokova ima svoje lokalne registre i memorije koji su prikazani na slici 8.



Slika 8: Programski model CU modula

Svaki CU modul ima 5 regisara i 2 programabilne memorije. Registri unutar CU modula su *cfg*, *xlen*, *wlen*, *first* i *last*. Svaki od ovih registara ima specifičnu namenu. Registra *cfg* određuje vrstu prediktivnog modela čije računanje CU modul izvršava. Registra *xlen* sadrži vrednost dužine ulazne instance. Interpretacija *wlen* registra zavisi od načina rada CU modula. Ovaj registar se i ne mora koristiti, u zavisnosti od režima rada cele REC arhitekture. Na primer, kada REC IP jezgro radi kao SVM prediktivni model ovaj registar određuje broj vektora podrške čija se suma računa u okviru trenutnog CU modula. Registri *first* i *last* su veličine 1 bit i određuju da li je trenutni CU modul prvi ili poslednji u vrsti koja implementira pojedinačni prediktivni model.

Svaki od CU modula sadrži i dve memorije: W RAM i F RAM. U okviru F RAM memorije nalaze se vrednosti nelinearne funkcije neophodne za proračun prediktivnog modela. Ukoliko model ne sadrži nelinearnu funkciju, ova memorija se ne koristi. W RAM memorija sadrži vektore neophodne za sračunavanje rezultata klasifikovanja. Vrsta vektora koji se nalaze u ovoj memoriji zavisi od trenutnog načina rada cele REC arhitekture. U nastavku će ukratko biti opisan sadržaj W RAM memorije zavisno od vrste klasifikatora. Struktura W RAM memorije, kada arhitektura radi kao stablo odluke, prikazana je na slici 9.



Slika 9: Organizacija W RAM memorije u slučaju stable odluke

Težinski vektori stabala odluka potrebni za proračun nekih vrsta stabala označeni su sa W_{xy} . Ukoliko se radi o ortogonalnim stablima, samo jedan od ovih koeficijenata će biti različit od 0. Dodatna vrednost, neophodna kod nekih vrsta stabala odluka, označena je sa b_x . Vrednost praga u odnosu na koji se određuje naredna adresa proračuna prikazana je sa tr_x . Kada je dobijena vrednost veća od ove granice u

narednom CU bloku računanje počinje od adrese adr_{x1} , inače počinje od adrese adr_{x2} . U slučaju da se dobije vrednost koja je kodovana kao klasa, cls_x , proračun se završava i rezultatom predikcije se proglašava vrednost cls_x .

Kada REC jezgro interpretira SVM prediktivni model tada je sadržaj u W RAM memoriji CU blokova prikazan na slici 10. Sa S_{xy} su označeni skalari vektora podrške, dok su sa α_x označeni Lagranžovi množitelji.

S11	S12	S13	...	S1n	α_1
S21	S22	S23	...	S2n	α_2
S31	S32	S33	...	S3n	α_3
...				...	
Sm1	Sm2	Sm3	...	Smn	α_m

Slika 10: W RAM za SVM

Kada REC IP jezgro radi kao neuronska mreža, tada je raspored vrednosti unutar W RAM memorije prikazan na slici 11. Težinski vektori ili centralni vektor u neuronima su označeni sa W_{xy} . U slučaju neuronskih mreža koje koriste prag, ta vrednost se nalazi nakon težinskih vektora i označena je sa b_x .

W11	W12	W13	...	W1n	b1
W21	W22	W23	...	W2n	b2
W31	W32	W33	...	W3n	b3
...				...	
Wm1	Wm2	Wm3	...	Wmn	b _m

Slika 11: W RAM za neuronske mreže

Performanse REC IP jezgra poređene su sa WEKA softverom. WEKA softverski paket ima odličnu podršku za rad sa ansamblima prediktivnih modela. Uz to je jedan od najkorišćenijih programa u oblasti mašinskog učenja. Za potrebe eksperimenata korišćen je AMD Phenom II 1090T procesor koji radi na učestanosti 3.2Ghz. REC IP jezgro je ograničeno da radi na 125 Mhz i implementirano je na Xilinx FPGA čipu. Testovi su sprovedeni na 17 standardnih testova iz UCI baze: Breast Cancer (bc), Wisconsin Breast Cancer (bcw), Credit Approval (ca), German Credit (cg), Horse Colic (col), Congressional Voting Records (cvr), Heart Disease (hd), Hepatitis (hep), Haberman's Survival (hs), Ionosphere (ion), Liver Disorders (ld), Labor Relations (lr), Mushrooms (mus), Pima Indians Diabetes (pid), King Rock vs King Pawn (rvp), Sonar, Mines vs Rocks (son), Tic-Tac-Toe (ttt).

Za potrebe testiranja nad svim navedenim setovima iz UCI baze sprovedeni su eksperimenti nad raznim

vrstama homogenih i heterogenih ansambala. Svi ansambli u eksperimentima su se sastojali od 30 pojedinačnih prediktivnih modela. Prvi deo testova je sproveden nad 6 vrsta homogenih ansambala: homogeni ansambli funkcionalnih stabala (F-DT), homogeni ansambli ortogonalnih stabala (AP-DT), homogeni ansambli SVM prediktivnih modela sa polinomijalnim kernelom (SVM-P) i radijalnim kernelom (SVM-R), neuronske mreže MLP tipa (MLP-ANN) i neuronske mreže koje koriste radijalne funkcije (RB-ANN). U tabeli 1 prikazana su ubrzanja u procesu klasifikacije instanci pomoću homogenih ansambala koja se mogu ostvariti korišćenjem REC IP jezgra, u odnosu na WEKA implementaciju koja se izvršava na standardnom PC računaru.

Tabela 1: Ubrzanja procesa klasifikacije u slučaju homogenih ansambala

Skup	F-DT	AP-DT	SVM-P	SVM-R	MLP-NN	RB-NN
bc	226867.7	4474.71	360.72	331.75	6768	10403.7
bcw	38942.77	2559.62	260.2	188.36	4022.26	4875.46
ca	195000	2444	281	296	4468	7052
cg	499509	2300	329	299	4584	7613
col	306018.13	4099.7	355.63	359.95	6470.92	10712.49
cvr	53352.94	3823.53	353.17	167.13	5220.59	6495.1
hd	61144	5679	272	184	6921	7454
hep	88194	8439	542	298	10238	11376
hs	79538	6022	236	169	6691	7527
ion	150047.62	3650.79	322.01	214.25	5339.29	7392.86
ld	61803.23	5861.29	150.11	108.4	6491.94	5873.66
lr	130784.31	13980.39	1992.01	820.29	18602.94	23112.75
mus	404272.74	843.11	458.06	462.73	4734.65	10255.78
pid	55420	2330	191	166	3703	4216
rvp	152587	960	183	195	2754	4179
son	239840	5781	427	306	6918	11275
ttt	122978	1628	185	180	3135	4558
Prosek	181309.25	4218.73	395.44	277.6	6110.37	8308.49

Eksperimenti su sprovedeni za sve skupove i za slučaj korišćenja arhitekture kao heterogenih ansambala. Korišćeni su sledeći heterogeni ansambli: heterogeni ansambli koji se sastoje od stabala odluka i SVM modela (DTSVM), heterogeni ansambli koji se sastoje od stabala odluka i neuronskih mreža (DTNN), heterogeni ansambli koji se sastoje od SVM modela i neuronskih mreža (SVMNN) i na kraju heterogeni ansambli koji se sastoje od sve tri vrste prediktivnih modela (DTSVMNN). U tabeli 2 prikazana su ubrzanja u procesu klasifikacije instanci pomoću heterogenih ansambala koja se mogu ostvariti korišćenjem REC IP jezgra, u odnosu na WEKA implementaciju koja se izvršava na standardnom PC računaru.

Tabela 2: Ubrzanja procesa klasifikacije u slučaju heterogenih ansambala

Skup	DTSVM	DTNN	SVMNN	DTSVMNN
bc	949.93	38035.02	382.68	1071.41
bcw	594.47	8836.78	378.56	837.25
ca	597	31451	273	660
cg	615	63953	318	670
col	899.05	45513.6	381.7	1034.82
cvr	641.22	13284.31	357.77	775.61
hd	592	14631	357	799
hep	1288	22200	756	1770
hs	378	15926	241	484
ion	1042.86	24429.89	412.74	1312.33
ld	263.86	14529.57	178.03	351.08
lr	4573.98	37957.52	2563.65	6519.69
mus	450.23	43868.88	396.21	436.05
pid	191	10485	147	224
rvp	274	20020	195	256
son	1449	37718	521	1684
ttt	257	18166	162	280
Prosek	852.31	28173.64	456.8	1080.85

Iz tabela 1 i 2 se može izvući zaključak da REC IP jezgro pruža značajno bolje performanse od WEKA softverskog alata za potrebe klasifikacije instanci korišćenjem homogenih ili heterogenih ansambala prediktivnih modela. Treba primetiti da REC IP jezgro radi na više od 25 puta manjoj učestanosti od WEKA softvera. Pri tome ostvaruje ubrzanja procesa klasifikacije i do 5 redova veličine veća. Ukoliko bi se arhitektura poredila sa embedded procesorom, umesto sa PC procesorom, razlika bi bila još veća pri čemu bi potrošnja ostala značajno manja.

Realizacija rešenja i mogućnost primene:

Prikazana arhitektura je implementirana kao IP jezgro na RT nivou abstrakcije pri čemu je korišćen jezik za opis digitalnog hardvera (Verilog). Zbog tehnološki nezavisnog opisa REC IP-a, njega je moguće upotrebiti kao hardverski akcelerator ansambala prediktivnih modela mašinskog učenja u ASIC ili FPGA tehnologijama. Dodatne modifikacije nisu neophodne.

Podnosilac prijave



УНИВЕРЗИТЕТ
У НОВОМ САДУ



ФАКУЛТЕТ
ТЕХНИЧКИХ НАУКА

Трг Доситеја Обрадовића 6, 21000 Нови Сад, Република Србија
Деканат: 021 6350-413; 021 450-810; Централа: 021 485 2000
Рачуноводство: 021 458-220; Студентска служба: 021 6350-763
Телефакс: 021 458-133; e-mail: ftndean@uns.ac.rs

ИНТЕГРИСАНИ
СИСТЕМ
МЕНАЏМЕНТА
СЕРТИФИКОВАН ОД:



Наш број: 01-сл

Ваш број:

Датум: 2016-11-03

ИЗВОД ИЗ ЗАПИСНИКА

Наставно-научно веће Факултета техничких наука у Новом Саду, на 25. редовној седници одржаној дана 26.10.2016. године, донело је следећу одлуку:

-непотребно изостављено-

ТАЧКА 10. Питања научноистраживачког рада и међународне сарадње

Тачка 10.2.4: У циљу верификације новог техничког решења усвајају се рецензенти:

- Проф. др Теуфик Токић, Електронски факултет у Нишу
- Доц. др Татјана Николић, Електронски факултет у Нишу

Назив техничког решења:

"РЕКОНФИГУРАБИЛНО IP ЈЕЗГРО ЗА АКЦЕЛЕРАЦИЈУ ХОМОГЕНИХ И ХЕТЕРОГЕНИХ АНСАМБАЛА ПРЕДИКТИВНИХ МОДЕЛА"

Аутори техничког решења: Вук Врањковић, Ратислав Струхарик.

-непотребно изостављено-

Записник водила:

Јасмина Димић, дипл. правник

Тачност података оверава
Секретар

Иван Нешковић, дипл. правник



Декан

Проф. др Раде Дорословачки

Softver:

Rekonfigurabilno IP jezgro za akceleraciju homogenih i heterogenih ansambala prediktivnih modela

Rukovodilac projekta: prof. dr Ljiljana Živanov

Odgovorno lice: dr Vuk Vranjković

Autori: Vuk Vranjković, Rastislav Struharik
Fakultet tehničkih nauka (FTN), Novi Sad

Razvijeno: u okviru projekta tehnološkog razvoja TR-32016

Godina: 2016.

Primena: novembar 2016.

Kratak opis

Ukoliko je prilikom projektovanja sistema neophodno veoma kratko vreme za izračunavanje rezultata homogenih i heterogenih ansambala prediktivnih modela mašinskog učenja tada hardverski akceleratori mogu biti jedino prihvatljivo rešenje za implementaciju sistema. Dodatno, ukoliko postoje ograničenja u potrošnji koju sistem ima na raspolaganju, hardverski akcelerator može biti od pomoći. Ova ograničenja postoje kod savremenih embedded sistema i SoC (System on Chip) rešenja. Razvijeno IP jezgro za akceleraciju homogenih i heterogenih ansambala prediktivnih modela mašinskog učenja olakšava izgradnju embedded sistema i SoC rešenja u kojima je neophodno mašinsko klasifikovanje visoke preciznosti.

Tehničke karakteristike:

Rešenje je realizovano kao IP jezgro u jeziku za opis digitalnog hardvera (Verilog) korišćenjem RTL metodologije. Stoga, ovo jezgro se može upotrebiti kao hardverski akcelerator u SoC rešenjima raznih tehnologija (ASIC, FPGA) bez dodatnih modifikacija.

Tehničke mogućnosti:

Jedinstvena arhitektura realizovana kao IP jezgro u predlogu tehničkog rešenja može da implementira heterogene i homogene ansamble koji se sastoje od tri najčešće korišćena tipa klasifikatora u oblasti mašinskog učenja: DT (Decision Trees, stabla odluka), ANN (Artificial Neural Networks, veštačke neuronske mreže) i SVM (Support Vector Machine). Predložena arhitektura je rekonfigurabilna i skalabilna i moguće je jednostavno menjanje pojedinačnih modela mašinskog učenja u okviru ansambla.

Realizator:

Fakultet tehničkih nauka – FTN

Korisnik:

Fakultet tehničkih nauka – FTN, Novi Sad

Podtip rešenja:

Softver – M85

Mišljenje

Tehničko rešenje "Rekonfigurabilno IP jezgro za akceleraciju homogenih i heterogenih ansambala prediktivnih modela" autora doc. dr Vuk Vranjkovića, i van. prof. dr Rastislava Struharika, realizovano 2016. godine, prikazano na 12 stranica A4 formata, grupisano je u ukupno četiri poglavlja:

1. Problemi koji se tehničkim rešenjem otklanjaju ili minimizuju
2. Stanje rešenosti pitanja istog problema u svetu
3. Tehnički detalji predloženog rešenja
4. Realizacija rešenja i mogućnost primene

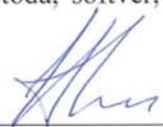
Tehničko rešenje pripada polju tehničko-tehnoloških nauka i oblasti elektrotehničkog inženjerstva. Naručilac tehničkog rešenja je Fakultet tehničkih nauka u Novom Sadu, Republika Srbija, koji je i korisnik tehničkog rešenja.

Tehničko rešenje je realizovano u okviru projekta "Nove generacije ugrađenih elektronskih komponenti i sistema u neorganskim i organskim tehnologijama za uređaje široke potrošnje" (Broj projekta TR 32016, Program istraživanja u oblasti tehnološkog razvoja za period 2011-2014., Tehnološka oblast - Elektronika, telekomunikacije i informacione tehnologije, Rukovodilac projekta: dr Ljiljana Živanov, redovni profesor).

Na osnovu analize tehničkog rešenja "Rekonfigurabilno IP jezgro za akceleraciju homogenih i heterogenih ansambala prediktivnih modela" autora doc. dr Vuk Vranjkovića, i van. prof. dr Rastislava Struharika, mogu se izvesti sledeći zaključci:

1. Dokumentacija tehničkog rešenja jasno prikazuje kompletnu strukturu tehničkog rešenja – opis problema, daje detaljniji osvrt na stanje u svetu, sadrži odgovarajući prikaz teorijskih osnova na kojima je zasnovano tehničko rešenje i posebno detaljno prikazuje strukturu i primenu realizovanog tehničkog rešenja.
2. Predloženo tehničko rešenje, "Rekonfigurabilno IP jezgro za akceleraciju homogenih i heterogenih ansambala prediktivnih modela", predstavlja efikasan alat za rešavanje problema u oblasti hardverske akceleracije homogenih i heterogenih ansambala prediktivnih modela mašinskog učenja.
3. Tehničko rešenje predstavlja originalan naučni doprinos sa praktičnom dimenzijom. Predložena arhitektura je rekonfigurabilna pa je njena primena fleksibilna i univerzalna.

Na osnovu prethodnog, predlažem da se "Rekonfigurabilno IP jezgro za akceleraciju homogenih i heterogenih ansambala prediktivnih modela", autora doc. dr Vuk Vranjkovića, i van. prof. dr Rastislava Struharika, prihvati kao novo tehničko rešenje i u skladu sa Pravilnikom o postupku i načinu vrednovanja, i kvantitativnom iskazivanju naučnoistraživačkih rezultata istraživača ("Službeni glasnik RS", broj 38/2008) klasifikuje kao rezultat "M85 Prototip, nova metoda, softver, standardizovan ili atestiran instrument, nova genska proba, mikroorganizmi".



Prof. Dr Teufik Tokić,
Univerzitet u Nišu
Elektronski fakultet

Softver:

Rekonfigurabilno IP jezgro za akceleraciju homogenih i heterogenih ansambala prediktivnih modela

Rukovodilac projekta: prof. dr Ljiljana Živanov

Odgovorno lice: dr Vuk Vranjković

Autori: Vuk Vranjković, Rastislav Struharik

Fakultet tehničkih nauka (FTN), Novi Sad

Razvijeno: u okviru projekta tehnološkog razvoja TR-32016

Godina: 2016.

Primena: novembar 2016.

Kratak opis

Ukoliko je prilikom projektovanja sistema neophodno veoma kratko vreme za izračunavanje rezultata homogenih i heterogenih ansambala prediktivnih modela mašinskog učenja, tada hardverski akceleratori mogu biti jedino prihvatljivo rešenje za implementaciju sistema. Dodatno, ukoliko postoje ograničenja u potrošnji koju sistem ima na raspolaganju, hardverski akcelerator može biti od pomoći. Ova ograničenja postoje kod savremenih embedded sistema i SoC (System on Chip) rešenja. Razvijeno IP jezgro za akceleraciju homogenih i heterogenih ansambala prediktivnih modela mašinskog učenja olakšava izgradnju embedded sistema i SoC rešenja u kojima je neophodno mašinsko klasifikovanje visoke preciznosti.

Tehničke karakteristike:

Rešenje je realizovano kao IP jezgro u jeziku za opis digitalnog hardvera (Verilog) korišćenjem RTL metodologije. Stoga, ovo jezgro se može upotrebiti kao hardverski akcelerator u SoC rešenjima raznih tehnologija (ASIC, FPGA) bez dodatnih modifikacija.

Tehničke mogućnosti:

Jedinstvena arhitektura realizovana kao IP jezgro u predlogu tehničkog rešenja može da implementira heterogene i homogene ansamble koji se sastoje od tri najčešće korišćena tipa klasifikatora u oblasti mašinskog učenja: DT (Decision Trees, stabla odluka), ANN (Artificial Neural Networks, veštačke neuronske mreže) i SVM (Support Vector Machine). Predložena arhitektura je rekonfigurabilna i skalabilna i moguće je jednostavno menjanje pojedinačnih modela mašinskog učenja u okviru ansambla.

Realizator:

Fakultet tehničkih nauka – FTN

Korisnik:

Fakultet tehničkih nauka – FTN, Novi Sad

Podtip rešenja:

Softver – M85

Mišljenje

Fakultet tehničkih nauka je razvio IP jezgro koje može da implementira homogene i heterogene ansamble koji se sastoje od DT, SVM i ANN mašinskih prediktivnih modela. IP jezgra su opisana korišćenjem Verilog jezika za modelovanje hardvera. Arhitektura je modelovana na tehnološki nezavistan način. Pri tome, opisana arhitektura je skalabilna. Zbog dva navedena razloga, IP jezgro se može lako prilagoditi potrebama trenutne aplikacije, pa se stoga može koristiti u širokom spektru embedded i SoC rešenja.

U predloženom tehničkom rešenju razmatran je problem hardverske implementacije homogenih i heterogenih ansambala prediktivnih modela mašinskog učenja. Analizom postojećih rešenja utvrđeno je da do sada objavljena rešenja ne mogu da implementiraju homogene i heterogene ansamble koji se sastoje od tri najčešće korišćene vrste klasifikatora: DT, SVM i ANN.

Predložena arhitektura je nazvana REC (Reconfigurable Ensemble Classifier). Arhitektura se sastoji od dvodimenzionalnog niza istovetnih blokova za računanje, CU (Computation Unit). CU blokovi mogu da rade nezavisno, svaki na zasebnoj ulaznoj instanci. Arhitektura organizovana na prikazani način omogućava veliku protočnost. Kao posledica toga dobija se veliko ubrzanje procesa klasifikacije instanci.

Opisano je na koji način arhitektura može da se poveže u SoC sisteme. Kao primer korišćeni su Xilinx FPGA čipovi. Razmatrani su čipovi koji poseduju hard IP ARM procesor i oni koji poseduju samo rekonfigurabilnu logiku. U oba slučaja prikazan je po jedan slučaj SoC sistema u kojima se koristi REC IP jezgro. Objasnjeno je kako se jezgro može povezati sa hard IP ARM procesorom kao i soft IP Microblaze procesorom.

Potom je opisan programski model REC IP jezgra. Objasnjeno je koji registri i memorije se nalaze unutar CU blokova. Opisano je čemu registri služe i opisana je struktura memorija, u slučajevima kada se arhitektura koristi kao homogeni ili heterogeni ansambl DT, SVM i ANN prediktivnih modela.

Nakon opisa arhitekture, opisa načina povezivanja arhitekture u Xilinx FPGA čipove i opisa programskog modela arhitekture, prikazani su i eksperimentalni rezultati. Eksperimenti su sprovedeni nad 17 skupova iz UCI baze i prikazana su ubrzanja koja arhitektura pruža u odnosu na WEKA softverski alat. Eksperimenti su pokazali da arhitektura pruža značajna ubrzanja u odnosu na softverski paket WEKA.

U skladu sa gore iznetim činjenicama tehničko rešenje ispunjava uslove da bude priznato kao softver (odnosno M85 u skladu sa Pravilnikom o postupku i načinu vredovanja i kvantitativnom iskazivanju naučnoistraživačkih rezultata istraživača, Sl. gl. RS br. 38/08).

Dr Tatjana Nikolić

Vanredni profesor Elektronskog fakulteta, Niš

T. Nikolić



УНИВЕРЗИТЕТ
У НОВОМ САДУ



ФАКУЛТЕТ
ТЕХНИЧКИХ НАУКА

Трг Доситеја Обрадовића 6, 21000 Нови Сад, Република Србија
Деканат: 021 6350-413; 021 450-810; Централна: 021 485 2000
Рачуноводство: 021 458-220; Студентска служба: 021 6350-763
Телефакс: 021 458-133; e-mail: ftndean@uns.ac.rs

ИНТЕГРИСАНИ
СИСТЕМ
МЕНАџМЕНТА
СЕРТИФИКОВАН ОД:



Наш број: 01.сл

Ваш број:

Датум: 2016-12-07

ИЗВОД ИЗ ЗАПИСНИКА

Наставно-научно веће Факултета техничких наука у Новом Саду, на 26. редовној седници одржаној дана 30.11.2016. године, донело је следећу одлуку:

-непотребно изостављено-

ТАЧКА 11. Питања научноистраживачког рада и међународне сарадње

Тачка 11.4.: На основу позитивног извештаја рецензената верификује се техничко решење (М85) под називом:

"РЕКОНФИГУРАБИЛНО IP ЈЕЗГРО ЗА АКЦЕЛЕРАЦИЈУ ХОМОГЕНИХ И ХЕТЕРОГЕНИХ АНСАМБАЛА ПРЕДИКТИВНИХ МОДЕЛА"

Аутори техничког решења: Вук Врањковић, Растислав Струхарик.

-непотребно изостављено-

Записник водила:

Јасмина Димић, дипл. правник

Тачност података оверава:
Секретар

Иван Нешковић, дипл. правник

Декан



Проф. др Раде Дорословачки