

TEZĂ DE DOCTORAT

CONTRIBUȚII LA CREȘTEREA EFICIENȚEI NEURONULUI CU PONDERI NUMERE COMPLEXE ȘI FUNCȚIE PERIODICĂ DE ACTIVARE

Ing. Valentin Mircea LUPEA

Conducător științific: Prof.dr.ing. Ștefan HOLBAN

Timișoara, 2014

Rezumat

Tema abordată în cadrul tezei de doctorat vizează domeniul rețelelor neuronale (RN) axându-se în special pe neuronul cu ponderi numere complexe și funcție periodică de activare (MVN-P).

Punctul de start al acestui studiu îl constituie ideea enunțată de Igor Aizenberg, în lucrarea sa, "Periodic Activation Function and a Modified Learning Algorithm for the Multi-valued Neuron", (IEEE - IEEE Transaction on neural networks, Vol. 21, No. 12, December 2010) și anume de a integra MVN-P ca neuron pe stratul de ieșire al unei RN multi-strat. Astfel eficiența crescută a unui singur MVN-P ar putea fi îmbunătățită. În vederea îndeplinirii acestui obiectiv, un studiu al MVN-P a fost întreprins, analizându-se strategiile de învățare existente dorindu-se scăderea timpului de învățare însă cu păstrarea unei erori acceptabile. Aspectul reducerii timpului de învățare a unui singur neuron este important mai ales în cazul RN unde timpul total de învățare poate crește semnificativ datorită calculelor suplimentare introduse de straturile ascunse și neuronii aferenți. S-a propus astfel o variantă diferită a procesului de învățare, diferența constând în tehnica de căutare și selecție a sectorului (soluției) corect. Utilizând această tehnică de căutare s-au obținut rezultate îmbunătățite din punct de vedere al timpului de învățare (numărul de epoci/iterații). Următorul pas al studiului l-a constituit integrarea MVN-P într-o RN multi-strat (MLMVN-P). În vederea realizării acestui obiectiv s-a construit o RN cu arhitectură reglabilă, care înglobează pe stratul de ieșire un singur MVN-P. Pe stratul de intrare, respectiv pe straturile ascunse s-au utilizat MVN. Rezultatele obținute sunt încurajatoare, observându-se îmbunătățiri comparativ cu situația utilizării unui singur MVN-P. În paralel cu integrarea MVN-P într-o MLMVN-P, s-a dorit determinarea periodicității optime a posibilelor sectoare/soluții în momentul învățării MVN-P. S-a observat că un număr prea mare sau prea mic de sectoare influențează negativ procesul de învățare. S-a dorit astfel identificarea unei plaje comune de valori pentru parametrul "periodicitate", valori pentru care procesul de învățare converge mai repede.

Teza de doctorat „*Contribuții la creșterea eficienței neuronului cu ponderi numere complexe și funcție periodică de activare*” este structurată pe șase capitole având un număr de 97 de pagini și 50 de referințe bibliografice.

Capitolul 1 intitulat *INTRODUCERE* prezintă un scurt istoric al temei abordate, motivația alegerii temei, conținutul lucrării pe scurt și stabilește obiectivele științifice cât și metodele de cercetare folosite în vederea atingerii acestor obiective.

Capitolul 2 intitulat *REȚEAUA NEURONALĂ ȘI NEURONUL ARTIFICIAL* prezintă conceptele teoretice care stau la baza RN și a neuronului artificial clasic, respectiv, a neuronului cu ponderi numere complexe (MVN) și a neuronului cu ponderi numere complexe și funcție periodică de activare. (MVN-P).

Pentru RN sunt enumerate principalele tipuri de arhitecturi cum ar fi: perceptronul, perceptronul multi-strat, RN de tip BackPropagation, RN Hopfield și RN de tip Kohonen. Pentru fiecare arhitectură, sunt indicate caracteristicile principale cât și aria de aplicabilitate. De asemenea, sunt enumerate principalele tehnici de învățare folosite în cazul RN: învățarea pe baza propagării erorii dinspre stratul de ieșire spre cel de intrare (Backpropagation), învățarea hebbiană, învățarea Boltzman, învățarea competitivă și învățarea cu exemple negative. Aceste tipuri de învățări pot fi clasificate în două mari clase și anume: învățare supervizată, respectiv nesupervizată. Învățarea supervizată este acea învățare în care corecția erorii se face pe baza diferenței dintre ieșirea dorită și cea obținută, în vreme ce învățarea nesupervizată se bazează pe caracteristica legăturii dintre neuroni. Astfel, ca și învățare supervizată putem enumera: Backpropagation și învățarea cu exemple negative, iar ca și învățare nesupervizată putem enumera: învățarea hebbiană, învățarea Boltzman și învățarea competitivă.

Pentru neuronul artificial, ca și componentă fundamentală a RN, sunt identificate principalele caracteristici constructive cât și o serie de funcții de activare, cum ar fi: funcția prag, funcția semn, funcția sigmoid, funcția tangent hiperbolică, funcția lineară, funcția lineară cu saturație, funcția lineară cu saturație simetrică și funcția gaussiană. Pentru fiecare din funcțiile enumerate mai sus, caracteristica de activare este prezentată, atât în formă matematică cât și prin reprezentare grafică.

Neuronul cu ponderi numere complexe (MVN) este prezentat în detaliu, cu funcția specifică de activare, caracteristicile procesului de învățare cât și câteva reguli specifice de corecție a erorilor în timpul învățării. Acest tip de neuron reprezintă un concept relativ nou, noutatea constând în folosirea numerelor complexe pentru ponderi, respectiv valori de ieșire ale neuronului. Folosirea numerelor complexe a condus la plasarea claselor de ieșire pe cercul trigonometric. Astfel, cercul trigonometric a fost împărțit în sectoare egale, fiecare reprezentând o clasă de ieșire, procesul de învățare reducându-se la deplasarea dintr-un sector greșit într-unul corect. Această abordare a condus la posibilitatea rezolvării unor probleme clasice folosind un singur neuron în locul unei RN multi-strat, aceasta reprezentând principalul avantaj al folosirii MVN.

În continuare, este prezentat neuronul cu ponderi numere complexe și funcție periodică de activare (MVN-P). MVN-P poate fi considerat ca o evoluție a MVN. Această evoluție constă în introducerea unui factor numit ”periodicitate”, cu ajutorul căruia se multiplică, cu o valoare dată, numărul de sectoare în care se situează clasele de ieșire. Astfel, dacă pentru MVN avem doar k posibile sectoare (k reprezentând numărul claselor de ieșire) pentru MVN-P avem kl sectoare, l reprezentând factorul de ”periodicitate”. Pentru acest tip de neuron sunt prezentate funcția specifică (periodică) de activare cât și două strategii de învățare, definite în literatură. Pentru fiecare dintre aceste două strategii sunt identificate avantajele și dezavantajele, referitoare la durata totală de învățare, respectiv eroarea de învățare.

Plecând de la aceste avantaje și dezavantaje, în Capitolul 3, este prezentată o nouă strategie de învățare, care dorește să îmbine avantajele celor două strategii existente.

Capitolul 3 intitulat *CONTRIBUȚII PRIVIND ÎMBUNĂTĂȚIREA PROCESULUI DE ÎNVĂȚARE A MVN-P* prezintă o nouă strategie de căutare a sectorului corect în vederea îmbunătățirii eficienței de învățare a MVN-P. Ca și punct de start au fost considerate conceptele teoretice prezentate în Capitolul 2 și anume cele două strategii de învățare propuse de Aizenberg. O nouă strategie de căutare a sectorului corect este astfel propusă, dorindu-se pe de-o parte scăderea duratei totale de învățare, iar pe de altă parte, păstrarea unei erori de învățare scăzute (în jurul valorii de 5-10%). Noua strategie propune folosirea unui singur sens de parcurgere a cercului trigonometric (sensul anti-orar sau orar), iar dacă sectorul corect nu este returnat după un număr dinamic de încercări, sectorul 0 este returnat ca sector corect. Pașii compleți de implementare ai noii strategii, sunt prezentați în detaliu, indicându-se de asemenea parametrii configurabili ai algoritmului. În Capitolul 4, este prezentată o soluție de integrare a MVN-P, cu noua strategie propusă, într-o RN multi-start.

Capitolul 4 intitulat *INTEGRAREA MVN-P ÎNTR-O REȚEA NEURONALĂ* prezintă o soluție de integrare a MVN-P într-o RN multi-strat în vederea îmbunătățirii eficienței de învățare a MVN-P. Integrarea MVN-P într-o RN reprezintă o noutate în domeniu, încercându-se astfel creșterea suplimentară a eficienței MVN-P. Arhitectura folosită pentru integrare este detaliată și un mecanism de învățare, bazat pe corecția erorii (Backpropagation) este prezentat. În cadrul acestui mecanism sunt identificate funcțiile specifice de activare a neuronilor de pe fiecare strat în parte, cât și regulile de corecție a ponderilor acestora. În Capitolul 5, rezultatele obținute, pe baza conceptelor teoretice prezentate în Capitolul 3 și Capitolul 4, sunt ilustrate.

Capitolul 5 intitulat *REZULTATE EXPERIMENTALE* prezintă rezultatele obținute în urma aplicării conceptelor teoretice din Capitolul 3, respectiv Capitolul 4. În cazul noii strategii de căutare a sectorului corect, rezultatele obținute sunt prezentate comparativ cu cele din literatură pentru ambele strategii de căutare propuse de Aizenberg. În cazul MLMVN-P rezultatele obținute sunt prezentate comparativ cu cele obținute pentru un singur MVN-P, în ambele implementări (un singur MVN-P, respectiv MLMVN-P) fiind folosită noua strategie de căutare propusă.

Noua strategie de căutare a produs rezultate bune, în anumite cazuri semnificativ mai bune decât cele două strategii propuse de Aizenberg. În cazul seturilor de date “suma modulo k a n valori” cu noua strategie de căutare, numărul de iterații a scăzut, cu câteva excepții evidențiate. În cazul bazelor de date “IRIS” și “Breast Cancer Wisconsin (Diagnostic)” cu noua strategie, învățarea nu a produs o eficiență de 100%. Cu toate acestea, o eficiență de peste 95% a fost obținută pentru ambele seturi de date, această valoare încadrându-se în marja de eroare dorită (5-10%). De asemenea s-a analizat, comparativ, eficiența noii strategii de căutare pentru sensul orar, respectiv anti-orar, ajungându-se la concluzia că sensul anti-orar conduce la o eficiență mai mare a învățării.

Pentru MLMVN-P rezultatele indică o creștere a eficienței învățării (durata de învățare a scăzut, respectiv rata erorii a scăzut la 0) pentru toate seturile de date “suma modulo k a n valori”. Pentru bazele de date “IRIS” și “Breast Cancer Wisconsin (Diagnostic)” MLMVN-P nu a condus la o îmbunătățire semnificativă a procesului de învățare, rata erorii păstrându-se la aceeași valoare. Cu toate acestea s-a observat o variație a erorii mai mică în timpul învățării.

Ca și concluzie pentru MLMVN-P, integrarea MVN-P într-o rețea neuronală (cu folosirea unor funcții specifice de activare, respectiv corecție a ponderilor) a fost realizată cu

success, ducând la creșterea suplimentară, în majoritatea situațiilor, a eficienței MVN-P, obiectivul stabilit în Capitolul 4 fiind atins.

Capitolul 6 intitulat *CONCLUZII* prezintă contribuțiile proprii identificate cât și posibile direcții viitoare de dezvoltare.

Astfel, ca și principale contribuții personale la creșterea eficienței MVN-P se pot enumera:

1. Propunerea unei tehnici noi de căutare a sectorului/soluției corecte. Noua tehnică se bazează pe căutarea anti-orară, iar dacă sectorul corect nu este găsit după un număr de încercări (direct proporțional cu numărul de intrări) sectorul returnat ca și corect este sectorul 0 (echivalent clasei 0); tehnica de căutare propusă ajută la eficientizarea procesului de învățare (numărul de iterații este redus).

2. Identificarea plajei optime de valori a parametrului “periodicitate” pentru posibilele sectoare/soluții pentru o serie de seturi de date.

3. Integrarea MVN-P într-o RN cu structură configurabilă – acest lucru reprezintă o noutate în domeniu până la data redactării acestei teze de doctorat.

4. Analiza eficienței MLMVN-P comparativ cu cea a unui singur MVN-P.

5. Identificarea necesității folosirii tehnicii de învățare cu exemple negative pentru cazul MVN-P, respectiv MLMVN-P, în vederea creșterii suplimentare a eficienței învățării.

6. Identificarea necesității inserării, unde este nevoie, a unui sector suplimentar (ieșire fictivă), pentru creșterea eficienței învățării; un astfel de procedeu și rezultatele obținute sunt prezentate în Capitolul 5.1.3.

Ca și direcții noi de dezvoltare se pot identifica următoarele:

1. Aplicarea conceptului de neuron block în cazul MVN-P, respectiv MLMVN-P.

2. Aplicarea învățării nesupervizate, (e.x. învățarea hebbiană) pentru cazul MVN-P, respectiv MLMVN-P.

3. Aplicarea conceptului de cuaternion (numere complexe extinse) pentru cazul MVN-P.

Teza se încheie cu o bibliografie aferentă, care a făcut posibil acest studiu.

Cuvinte cheie: rețea neuronală, neuron artificial, neuron artificial cu ponderi complexe, funcție periodică de activare.