

ხელოვნური ნეირონული ქსელის ზოგიერთი პარამეტრის კვლევა

მიხეილ კოტიშაძე

აკაკი წერეთლის სახელმწიფო უნივერსიტეტი, ქუთაისი, საქართველო
mmmkot56@mail.ru

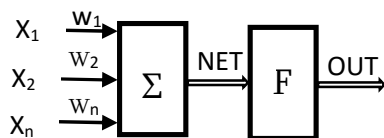
ანოტაცია - სტატიაში განხილულია ხელოვნური ნეირონული ქსელის ზოგიერთი პარამეტრის ექსპერიმენტული კვლევა, კერძოდ ქსელის შეცდომის დამოკიდებულება ქსელის კონფიგურაციაზე. კვლევა ჩატარებულია ქსელზე, რომელშიც გამოყენებულია შეცდომის უკუგავრცელების ალგორითმი. საკვანძო სიტყვები - ხელოვნური ინტელექტი, ნეირონული ქსელები.

I. შესავალი

კონკრეტული ამოცანის გადანყვეტისას ქსელის კონფიგურაციის შერჩევა ხელოვნური ნეირონული ქსელების ერთ-ერთ პრობლემას წარმოადგენს [1]. ამჟამად არ არსებობს უნივერსალური მეთოდი, რომლის საშუალებითაც შესაძლებელი იქნებოდა პრობლემის პარამეტრების მიხედვით ქსელის კონკრეტული კონფიგურაციის არჩევა. სრულკავშირიანი ქსელებისთვის კონფიგურაციის ძირითად პარამეტრებს წარმოადგენს შრეთა რაოდენობა და თითოეულ შრეში ნეირონების რაოდენობა. მიუხედავად იმისა, რომ შექმნილია ქსელების ავტომატური კონსტრუირების ალგორითმები [2] [3], პრაქტიკულად თითოეული ამოცანისათვის ქსელის პარამეტრები შეირჩევა ექსპერიმენტულად. აქედან გამომდინარე ხელოვნური ნეირონული ქსელის კონფიგურაციის პარამეტრების ექსპერიმენტული კვლევა აქტუალურ ამოცანას წარმოადგენს.

II. ხელოვნური ნეირონული ქსელი

ნეირონული ქსელი, რომელზედაც ტარდებოდა ექსპერიმენტები, წარმოადგენს მრავალშრიან პერსეპტონის კომპიუტერულ მოდელს შეცდომის უკუგავრცელების ალგორითმით [1]. ერთი ნეირონის სქემა, მოცემულია ნახ. 1-ზე.



ნახ. 1. ხელოვნური ნეირონის სქემა

სადაც F(NET) აქტივაციური ფუნქციაა და ყველაზე ხშირად აიღება კლასიკური სიგმოიდა. ეს ფუნქცია მოსახერხებელია, რადგანაც მას ძალიან მარტივი წარმოებულები აქვს, რაც გამოყენებულია უკუგავრცელების ალგორითმის რეალიზაციისას.

სიგმოიდი, რომელსაც ზოგჯერ ლოგისტიკურ ან შემკუმშავ ფუნქციას უწოდებენ, ისე კუმშავს (ავიწროვებს) NET-ის ცვლილების დიაპაზონს, რომ OUT-ის მნიშვნელობა ძვეს ნულსა და ერთს შორის. მრავალშრიან ნეირონულ ქსელებს მხოლოდ არანრფივობის შემთხვევაში გააჩნიათ უფრო დიდი წარმოდგენითი სიმძლავრე, ვიდრე ერთშრიანებს. შემკუმშავი ფუნქცია უზრუნველყოფს საჭირო არანრფივობას[1].

$$NET = \sum_{i=1}^n (x_i w_i), \text{ ხოლო } OUT = \frac{1}{1 + e^{-\alpha NET}}$$

აქ x_1, x_2, \dots, x_n - ნეირონის შემავალი სიდიდეებია, რომლებიც აღწერენ შემავალ სახეებს; w_1, w_2, \dots, w_n - კავშირების წონებია, რომლებსაც სწავლების საწყის ეტაპზე ენიჭებათ მცირე შემთხვევითი სიდიდეები (0-დან 1-მდე); α -სიგმოიდის პარამეტრია. ასეთი ნეირონებისაგან შექმნილია ქსელი რამოდენიმე შრით, თითოეულ შრეში ნეირონების განსაზღვრული რაოდენობით. შრის ყველა ნეირონის გამოშავალი სიდიდე წარმოადგენს შემდეგი შრის ყველა ნეირონის შემავალ სიდიდეს (სრულკავშირიანი ქსელი). ნეირონების პირველი (შესავლებთან შეერთებული) შრე მხოლოდ სიგნალების გამანაწილებელ ნერტილებს წარმოადგენს და შემავალი სიგნალი

უბრალოდ გადის მათში, მათსავე გამოსავალზე არსებულ წონებთან. ხოლო შემდეგი შრეების თითოეული ნეირონი იძლევა NET და OUT სიგნალებს, როგორც ნაჩვენებია იყო ზევით. ქსელის სწავლების მიზანს წარმოადგენს წონების ისეთი შეწყობა, რომ შემავალი სიგნალების რაიმე სიმრავლემ მოგვცეს გამომავალი სიგნალების საჭირო სიმრავლე. შემავალი და გამომავალი სიგნალების სიმრავლეებს ვუნოდოთ ვექტორები. სწავლებისას იგულისხმება, რომ თითოეული შემავალი ვექტორისთვის არსებობს მისი მენყვილე მიზნობრივი ვექტორი, რომლითაც მოცემულია მოთხოვნილი გამოსავალი. ორივეს ერთად ეწოდება შემსწავლელი წყვილი. როგორც წესი, ქსელს ასწავლიან მრავალ წყვილზე.

უკუგავრცელების ქსელის სწავლება მოითხოვს შემდეგი ოპერაციების შესრულებას:

1. შემსწავლელი სიმრავლიდან შევარჩიოთ მორიგი შემსწავლელი წყვილი; მივანოდოთ შემავალი ვექტორი ქსელის შესავალზე.
2. გამოვთვალოთ ქსელის გამოსავალი.
3. გამოვთვალოთ სხვაობა ქსელის გამოსავალსა და მოთხოვნილ გამოსავალს (შემსწავლელი წყვილის მიზნობრივი ვექტორი) შორის.
4. მოვახდინოთ წონების კორექტირება ისე, რომ მოხდეს შეცდომის მინიმიზირება.
5. ვიმეოროთ ბიჯები 1-დან 4-მდე შემსწავლელი სიმრავლის თითოეული ვექტორისთვის მანამ, სანამ შეცდომა მთელ სიმრავლეზე არ მიაღწევს მისაღებ მნიშვნელობას.

1 და 2 ბიჯებით შესრულებული ოპერაციები იმ ბიჯებს ჰგავს, რომლებიც სრულდება უკვე ნასწავლი ქსელის ფუნქციონირებისას, ანუ მიენოდება შემავალი ვექტორი და გამოითვლება მიღებული გამოსავალი. გამოთვლები სრულდება შრეობრივად. ჯერ გამოითვლება j შრის ნეირონების გამოსავლები, შემდეგ ისინი გამოიყენება როგორც k შრის ნეირონების შესავლები, ხოლო k შრის ნეირონების

გამოსავლები შეადგენს ქსელის გამომავალ ვექტორს.

ბიჯ3-ზე ქსელის გამოსავლებიდან თითოეული, რომლებიც აღნიშნულია OUT-ით, გამოაკლდება მიზნობრივი ვექტორის შესაბამის კომპონენტს, რომ მივიღოთ შეცდომა. ეს შეცდომა გამოიყენება ბიჯ4-ზე ქსელის წონების კორექციისთვის.

ამ ოთხი ბიჯის საკმაო რაოდენობით განმეორების შემდეგ ქსელის რეალურ გამოსავლებსა და მიზნობრივ ვექტორს შორის სხვაობა უნდა შემცირდეს მისაღებ სიდიდემდე. ამ დროს ამბობენ, რომ ქსელმა ისწავლა. შემდგომში ქსელი გამოიყენება სახეების გამოსაცნობად და წონები არ იცვლება.

1 და 2 ბიჯებზე შეიძლება ითქვას, რომ ეს არის "წინსვლა", რადგანაც სიგნალი ვრცელდება შესავალიდან გამოსავალისკენ. ბიჯები 3 და 4 წარმოდგენენ "უკუსვლას", რადგანაც გამოთვლილი შეცდომის სიგნალი ვრცელდება ქსელში უკუმიმართულებით და გამოიყენება წონების კორექტირებისათვის. განვიხილოთ ორივე სვლა და წარმოვადგინოთ მათემატიკური ფორმით.

წინსვლა. ბიჯები 1 და 2 შეიძლება გამოვსახოთ ვექტორული ფორმით შემდეგნაირად: მიენოდება შემავალი ვექტორი $X (x_1, x_2, \dots, x_n)$ და გამოსავალზე მიიღება ვექტორი OUT ($OUT_1, OUT_2, OUT_3, \dots, OUT_n$). ვექტორული წყვილი შესავალი-მიზანი X და T აიღება შემსწავლელი სიმრავლიდან. გამოთვლები წარმოებს ვექტორ X -ზე, რომ მივიღოთ გამომავალი ვექტორი OUT.

როცა მიღებული იქნება გამოსავლების სიმრავლე, იგი წარმოადგენს შემავალ სიგნალებს შემდეგი შრის ნეირონებისათვის. პროცესი მეორდება ყველა შრეში მანამ, სანამ არ მიიღება ქსელის საბოლოო გამოსავლების სიმრავლე.

ერთი შრის გამომავალი ვექტორი წარმოადგენს შემავალ ვექტორს შემდეგი შრისთვის.

უკუსვლა. გამომავალი შრის წონების შეწყობა. რადგანაც გამომავალი შრის თითოეული ნეირონისთვის მოცემულია მიზნობრივი მნიშვნელობა, წონების შეწყობა ადვილად ხორციელდება მოდიფიცირებული დელტა-ნესის გამოყენებით. შიგა შრეებს "დაფარული შრეები" ეწოდებათ. მათი გამოსავლებისთვის არ არსებობს მიზნობრივი მნიშვნელობები შედარებისთვის, ამიტომ სწავლება რთულდება.

განვიხილოთ სწავლების პროცესი ერთი წონისათვის დაფარული j შრის p ნეირონიდან გამომავალი k შრის q ნეირონამდე. გამომავალი k შრის ნეირონის გამოსავალი, გამოაკლდება რა მიზნობრივ მნიშვნელობას (T), გვაძლევს შეცდომის სიგნალს. ის მრავლდება შემკუმშავი ფუნქციის წარმოებულზე $\alpha[OUT(1 - OUT)]$, რომელიც გამოთვლილია k შრის ამ ნეირონისთვის და გვაძლევს δ სიდიდეს.

$$\delta = \alpha OUT(1 - OUT)(T - OUT)$$

შემდეგ δ მრავლდება j შრის იმ ნეირონის OUT სიდიდებზე, რომლიდანაც გამოდის განსახილველი წონა. ეს ნამრავლი თავის მხრივ მრავლდება სწავლების სიჩქარის η კოეფიციენტზე (ჩვეულებრივ $0,01$ -დან $1,0$ -მდე) და შედეგი ემატება წონას. ასეთი პროცედურა სრულდება თითოეული წონისთვის დაფარული შრის ნეირონიდან, გამომავალი შრის ნეირონამდე.

საბოლოოდ, გამომავალი შრის ნეირონების წონების კორექციისთვის სრულდება შემდეგი გამოთვლები

$$\Delta w_{pq} = \eta \delta_q OUT_{pj}$$

$$w_{pq}(n+1) = w_{pq}(n) + \Delta w_{pq}$$

სადაც $w_{pq}(n)$ - დაფარული შრის p ნეირონის წონაა გამომავალი შრის q ნეირონამდე n ბიჯზე (კორექციამდე); $w_{pq}(n+1)$ - წონის სიდიდეა $n+1$ ბიჯზე (კორექციის შემდეგ);

წონების შეწყობა დაფარულ შრეში. განვიხილოთ გამომავალი შრის წინ მდებარე დაფარული შრის ერთი ნეირონი. წინსვლის დროს ეს ნეირონი თავის გამომავალ სიგნალს,

შემაერთებელი წონის გავლით, გადასცემს გამომავალი შრის ნეირონებს. სწავლებისას ეს წონები ფუნქციონირებენ უკუ მიმართულებით, ანუ სიდიდე δ -ს გაატარებენ გამომავალი შრიდან დაფარული შრისკენ. ამ წონებიდან თითოეული მრავლდება გამომავალი შრის იმ ნეირონის δ -ზე, რომელთანაც ეს წონაა შეერთებული. დაფარული შრის ნეირონისთვის აუცილებელი δ სიდიდე მიიღება ყველა ასეთი ნამრავლის ჯამის გამრავლებით შემკუმშავი ფუნქციის წარმოებულზე:

$$\delta_q = \alpha OUT_{pj}(1 - OUT_{pj})[\sum \delta_k w_{pq}]$$

როცა δ -ს მნიშვნელობა მიღებულია, შეიძლება კორექტირება გავუკეთოთ პირველი დაფარული შრის მკვებავ წონებს ზემოთ მოცემული ფორმულების საშუალებით, სადაც ინდექსები იცვლება შრეების შესაბამისად.

III. ქსელის ექსპერიმენტული კვლევა

კვლევის ობიექტს წარმოადგენს ხელოვნური ნეირონული ქსელი შეცდომის უკუგავრცელების ალგორითმით [1]. ნეირონულ ქსელს წარდგენილი სასწავლო ნაკრებით უნდა ესწავლა ფესვის ამოღება ნოლიდან ასამდე რიცხვებიდან. სასწავლო ნაკრები შედგებოდა 21 და 31 შემსწავლელი წყვილისაგან. შემავალ შრეში 10 ნეირონია, ხოლო გამომავალ შრეში ერთი. დაფარული შრეები და მათში ნეირონების რაოდენობა იცვლებოდა. რაც შეეხება ქსელის სხვა პარამეტრებს, ისინი წარმოდგენილია შემდეგი სახით:

- გამოყენებული სასწავლო წყვილები < 80%
- სწავლების სიჩქარის კოეფიციენტი - 0,1
- წყვილი გამოცნობილია თუ შეცდომა < 0,05
- ეპოქების (სასწავლო ნაკრების 1 სრული გავლა ქსელში) რაოდენობა - 50000
- სიგმოიდის პარამეტრი - 0,5

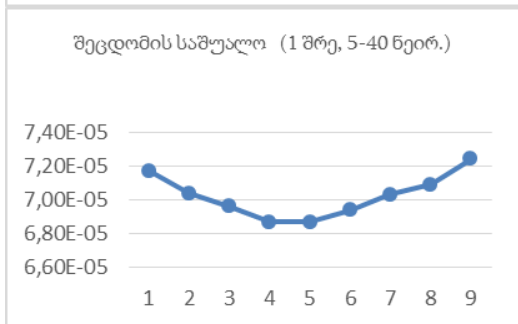
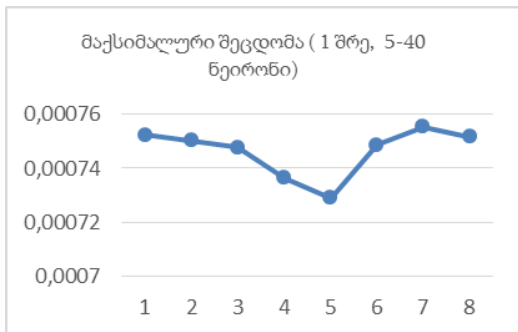
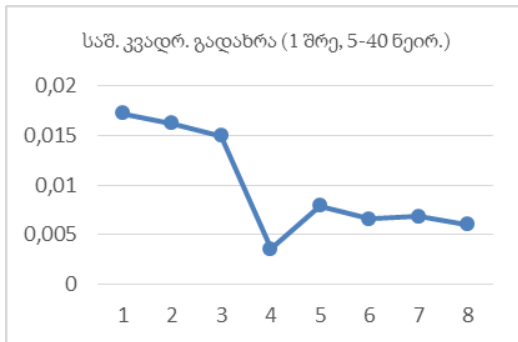
შეფასების კრიტერიუმად აღებული იყო :

- ქსელის მაქსიმალური შეცდომა სწავლების პროცესში
- სწავლების პროცესში ქსელის შეცდომის საშუალო
- საშუალო კვადრატული გადახრა უკვე ნასწავლი ქსელისთვის კონკრეტული რიცხვებიდან ფესვის ამოღებისას, კერძოდ რიცხვებიდან: 2, 3, 4, 9, 16, 25, 36, 49, 64, 81.

კვლევა დაიწყო ერთი დაფარული შრიტ და მასში 5 ნეირონით. ნეირონების მაქსიმალური რაოდენობა 40-მდე იზრდებოდა. შესაბამისი გრაფიკები ნაჩვენებია ნახ.1-ზე.

გრაფიკების ანალიზი გვიჩვენებს, რომ ოპტიმალური შედეგები ნაჩვენებია როცა ნეირონების რაოდენობა 20-25-ის ფარგლებშია.

რაც შეეხება შრეების რაოდენობის გავლენას ქსელის შეცდომაზე, კვლევის

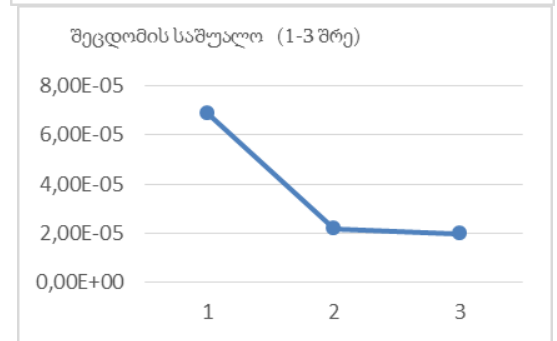
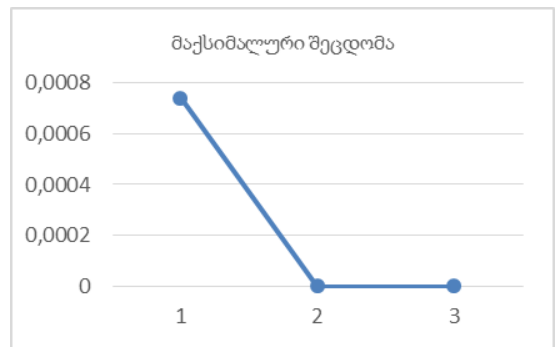


ნახ. 1. ექსპერიმენტის შედეგები ერთი დაფარული შრიტ

შედეგები მოცემულია ნახ.2-ზე, რომლებზეც ნათლად ჩანს, რომ შრეების რაოდენობის ზრდა ამკარად ამცირებს ქსელის შეცდომას.

დადასტურდა, რომ ქსელის შეცდომა არსებითადაა დამოკიდებული დაფარულ შრეში ნეირონების რაოდენობაზე და რაკი ქსელი სრულუკავშირებიანია, შესაბამისად წონებიანი კავშირების რაოდენობაზე.

დადასტურდა, რომ შრეების რაოდენობის ზრდასთან ერთად მკვეთრად მცირდება ქსელის შეცდომის მნიშვნელობა.



ნახ 2. ექსპერიმენტის შედეგები 1-3 დაფარული შრიტ

III. დასკვნა

ლიტერატურა

- [1] Ф. Уоссермен. Нейрокомпьютерная техника. М: Мир, 1992 г.
- [2] Царегородцев В.Г. Конструктивный алгоритм синтеза структуры многослойного персептрона. Вычислительные технологии, 2008. Т.13 – Вестник КазНУ, серия «математика, механика, информатика» , 2008. №4ю Часть 3. – с. 308-315.
- [3] მ.კოტიშაძე, ც. ბერიძე. ხელოვნური ნეირონული ქსელების განვითარების ზოგიერთი კონცეფციის შესახებ. VI საერთაშორისო სამეცნიერო-პრაქტიკული კონფერენცია „ინტერნეტი და საზოგადოება“ (INSO - 2013)ქუთაისი,2013.