საქართველოს ტექნიკური უნივერსიტეტი

რევი ჩოგოვაძე, რამაზ ხუროძე

ხელოვნური ნეირონული ქსელები

დამტკიცებულია სტუ-ს
სასწავლო-მეთოდური
საბჭოს მიერ

თბილისი
2006

GEORGIAN TECHNICAL UNIVERSITY

Revi Chogovadze, Ramaz Khurodze

Artificial Neural Networks

Approved by the Educational Methodical Council of GTU

Tbilisi
2006

Dr. Prof. Revi (Revaz) Alex **Chogovadze** was born in Tbilisi, Georgia in 1944. He graduated from the Georgian Technical University, and later completed the post graduate studies of the Scientific Research Institute of Control Problems of the USSR Academy of Sciences in Moscow. Between 1973-2002 was: the Department Head, and later General Manager of the Republican Computing Center; Vice-president of the State Agricultural Technical Services Corporation; and State Counsellor of the Government of Georgia. Since 2003 is the Head of the Scientific Research Laboratory and the Associate Professor of the Artificial Intelligence Department at the Georgian Technical University. He is the author of numerous scientific publications and monographs in the field of Associative Parallel processors, Pattern recognition and Artificial Neural Networks theory.

Dr. Prof**.** Ramaz Adolf **Khurodze** was born in Tbilisi, Georgia in 1944.He graduated from the Georgian Technical University, and later completed the post graduate studies of the Scientific Research Institute of Control Problems of the USSR Academy of Science in Moscow. Between 1972-1993 was: the Department Head of the Institute of Control Systems; Deputy Ministry of Public Education; Prorector of Georgian Technical University, since 1994 he is a Rector of GTU. In 2002 became a corresponding member of the Georgian Academy of Sciences. He has published more than 100 articles and monographs, his scientific interests include: Reliability Theory, Recognition systems and Artificial Neural Networks theory.

Uuak 621.391

განხილულია ხელოვნური ინტელექტის დისციპლინის პერსპექტიული მიმართულება - ხელოვნური ნეირონული ქსელების თეორია, რომელიც ინფორმაციის დამუშავების ალტერნატიულ ტექნოლოგიას წარმოადგენს. ნაშრომში სისტემატიზებულია შესაბამისი სამეცნიერო და სასწავლო-მეთოდოლოგიური მასალები, ყურადღება ეთმობა ნეიროინფორმატიკის ქართული სამეცნიერო ტერმინოლოგიის დადგენასა და დახვეწას. წარმოდგენილი თეორიული მასალის დაუფლების შემდეგ აღწერილი ხელოვნური ნეირონული ქსელების უმეტესობის რეალიზება შესაძლებელია პერსონალურ კომპიუტერზე.

გათვალისწინებულია ბაკალავრიატის, მაგისტრატურისა და დოქტორანტურის სტუდენტებისათვის, გამოადგებათ აგრეთვე სათანადო მომზადების მქონე სხვადასხვა დარგის სპეციალისტებს, რომლებიც დაინტერესებულნი არიან თავიანთ სფეროში ნეიროკომპიუტინგის გამოყენებით.

რეცენზენტები: პროფ. ოთარ ვერულავა
პროფ. ზურაბ წვერაიძე

წინასიტყვაობა

მეცნიერების ის სფეროები, რომლებიც შეისწავლის ადამიანის ინტელექტუალური მოქმედების მოდელირების შესაძლებლობას, განსაკუთრებულ ადგილს იკავებს თანამედროვეობის მრავალი პრობლემის გადაწყვეტაში. ბუნებრივი ინტელექტუალური სისტემების შესწავლამ აჩვენა, რომ ეს მიმართულება უაღრესად პერსპექტიულია მეცნიერებისა და ტექნიკის პრაქტიკულად ყველა დარგისათვის, პირველ რიგში ძნელად ფორმალიზებადი და/ან რთული პრობლემების გადასაწყვეტად. ასეთი ტიპის ამოცანა მრავლად არის მედიცინაში, კრიმინალისტიკაში, ფსიქოლოგიაში, სამხედრო საქმეში, ფიზიოლოგიაში, ბიოლოგიაში, სოციოლოგიაში, ქიმიაში, ფიზიკისა და ტექნიკის იმ სფეროებში, რომლებიც დაკავშირებულია სახეთა ამოცნობასა და საუკეთესო გადაწყვეტილების არჩევასთან.

ხელოვნური ინტელექტის დისციპლინა კომპლექსური ხასიათისაა, იგი ინტენსიური ფორმირების სტადიაზეა. მისი შემადგენლობის შესახებ სხვადასხვა ავტორებს სხვადასხვა ხედვა აქვთ, თუმცა ხელოვნური ინტელექტის სფეროს ცალსახად მიაკუთვნებენ ისეთ მიმართულებებს, როგორიც არის სახეთა ამოცნობა, ნეიროინფორმატიკა (ხელოვნური ნეირონული ქსელები), რობოტიკა, მანქანური სწავლება, კოგნიტიური სწავლება, ექსპერტული სისტემები, გენეტიკური ალგორითმები. ხელოვნური ინტელექტის სფერო შთამბეჭდავია არა მარტო მეცნიერებისა და ტექნიკის, არამედ პრაქტიკული გამოყენების თვალსაზრისითაც. მსოფლიოს განვითარებულ ქვეყნებში ერთმნიშვნელოვნად არის აღიარებული ხელოვნური ინტელექტის სისტემების დანერგვის აუცილებლობა ბიზნესში და სახელმწიფო სტრუქტურებში. მაგალითად, ამერიკის შეერთებული შტატების წარმატებული კომპანიების უმეტესობას დანერგილი აქვს ხელოვნური ინტელექტის სისტემები

ისეთი პროცესების სამართავად, როგორიც არის სტრატეგიული დაგეგმვა, ფინანსების და მარკეტინგის მენეჯმენტი, რესურსების გადანაწილება, კლიენტების მომსახურება, წარმოების ეფექტურობისა და მაღალპროდუქტიულობის გაზრდა, პროდუქციაზე მოთხოვნის განსაზღვრა და ა.შ. სახელმწიფო სტრუქტურები ხელოვნური ინტელექტის სისტემებს იყენებენ მაკროეკონომიკური პროცესების, საგადასახადო და ფისკალური სფეროების, საბანკო სექტორის, სოფლის მეურნეობის, ენერგეტიკის, მრეწველობისა და სხვა დარგების პროგნოზირებისა და მართვისთვის. ხელოვნური ინტელექტის სისტემებს უმნიშვნელოვანესი ფუნქცია აკისრია პროგნოზირების სფეროშიც, ხოლო ზუსტი პროგნოზების გაკეთება XXI საუკუნის საზოგადოების უდიდესი მოთხოვნილება და ამასთანავე აუცილებლობაც არის. ასეთი პროგნოზების გაკეთების ეფექტური მეთოდების ძიებას კი ხელოვნური ინტელექტის სისტემებთან მივყავართ. ხელოვნური ინტელექტის მეთოდებს შეუძლია საკმარისად ზუსტი პროგნოზების გაკეთება ისეთი ძნელადპროგნოზირებადი მოვლენების, როგორიც არის ადამიანის ქცევა და ჯანმრთელობა, ამინდი, სპორტული თამაშები, ეკონომიკური მაჩვენებლები, საფონდო ბირჟის ტრანზაქციები და მრავალი სხვა.

მსოფლიოს წამყვანი სამეცნიერო ცენტრები ბოლო წლებში განსაკუთრებულ ყურადღებას უთმობენ ხელოვნური ინტელექტის დისციპლინის უაღრესად საინტერესო და პერსპექტიული მიმართულების - ხელოვნური ნეირონული ქსელების თეორიისა და პრაქტიკის განვითარებას. ხელოვნური ნეირონული ქსელების აგების პრობლემას, რომელიც მჭიდროდ უკავშირდება ადამიანის ცენტრალურ ნერვულ სისტემაში ინფორმაციის შენახვისა და გადამუშავების პროცესებს, გააჩნია გადაწყვეტის ორი მიმართულება: ერთი განიხილავს ბუნებრივი ნეირონისა და ნეირონული ქსელების ანალოგიური ხელოვნური ნეირონული

ქსელების აგებას, რასაც მნიშვნელოვნად ზღუდავს ბიოლოგიური ნეირონისა და ნეირონული ქსელების ფუნქციონირების ურთულესი მექანიზმების მხოლოდ ნაწილობრივი ცოდნა; მეორე მიმართულება გულისხმობს ბუნებრივი ინტელექტუალური სისტემებისაგან განსხვავებული ხელოვნური სისტემების აგებას, რაც პრაქტიკულად ხსნის შეზღუდვებს ინფორმაციის დამუშავების ახალი გზების ძიებისას. ინფორმაციის დამუშავების იმ ასპექტებში, სადაც ბუნებრივი ნეიროქსელების მკვლევართა აზრით დასკვნები უტყუარია, შესაძლებელია ბიოლოგიური ნეირონისა და ნეირონული ქსელების მოდელირება ანალოგიის დიდი ხარისხით, სხვა შემთხვევაში შეიძლება დავუშვათ ადეკვატურობის ნებისმიერი ხარისხი, ანუ ამ ორი მიმართულების სინთეზი უნდა მივიჩნიოთ ძნელად ფორმალიზებადი პრობლემების ეფექტური გადაწყვეტის ერთ-ერთ პერსპექტიულ გზად.

ჯონ ფონ ნეიმანის კომპიუტერული პარადიგმისა და გამოთვლითი ტექნიკის ელემენტური ბაზის მნიშვნელოვანი განვითარების მიუხედავად, რამაც მაღალი პარალელიზმის გამოთვლითი სისტემებისა და სუპერკომპიუტერების შექმნა განაპირობა, ბოლომდე ვერ წყდება რიგი პრობლემები, მათ შორის ადამიანისა და კომპიუტერის ეფექტური ურთიერთქმედების, ისეთი ამოცანების ხარისხიანი გადაწყვეტის, რომელთა ალგორითმების შექმნა გაძნელებულია ან ზოგჯერ შეუძლებელიც, ამასთან მრავალპროცესორიანი სისტემების მწარმოებლობა იზრდება მხოლოდ როგორც პროცესორების რაოდენობის ლოგარითმი (მიახლოებით), ანუ რაც უფრო მეტად იზრდება პროცესორების რაოდენობა, მით ნაკლებად იზრდება სისტემის მწარმოებლობა, მით მეტია გამოთვლების ღირებულება, ხოლო ინფორმაციისა და ბრძანებების გაცვლის ეფექტური ორგანიზაცია ასეთ სისტემებში თავად ხდება რთული მეცნიერული პრობლემა. ნათელი ხდება, რომ მაღალი პარალელიზმის მრავალპროცესორიანი

გამოთვლითი სისტემების განვითარების ამ გზას ნაკლებად აქვს პერსპექტივა ბევრი რთული გამოთვლითი ამოცანისათვის, სწორედ ამიტომ კვლევები წარმოებს ახალი კომპიუტერული პარადიგმების შესაქმნელად. ცხადი ხდება, რომ სახეთა ამოცნობისა და მსგავსი ძნელად ფორმალიზებადი, რთული ამოცანები უნდა გადაწყდეს ინფორმაციის დამუშავების ახალი, განსხვავებული მეთოდებით. ერთ-ერთ ასეთ ალტერნატიულ ტექნოლოგიას წარმოადგენს ნეიროკომპიუტინგი, ანუ ინფორმაციის დამუშავება ხელოვნური ნეირონული ქსელებით, რომლებშიც სიგნალების პარალელური დამუშავების პრინციპი ხორციელდება დიდი რაოდენობით პროცესორული ელემენტების - ფორმალური ნეირონების სხვადასხვა კონფიგურაციის მქონე ჯგუფებად ან შრეებად გაერთიანების გზით, ხოლო პროცესორული ელემენტების დაპროგრამების ნაცვლად ხორციელდება ფორმალური ნეირონების ”სწავლება”გადასაწყვეტი ამოცანის მაგალითების გამოყენებით. მიჩნეულია, რომ ამ მიმართულების განვითარებით მნიშვნელოვანი ნაბიჯი იქნება გადადგმული ხელოვნური ინტელექტის შექმნისაკენ. ხელოვნური ინტელექტუალური სისტემის აგება სწავლების პროცესების ეფექტური განხორციელების გარეშე შეუძლებელია, ამიტომ ხელოვნური ნეირონული ქსელების კვლევების უმნიშვნელოვანეს საგანს წარმოადგენს მათი სწავლების ეფექტური ალგორითმების შემუშავება. სხვა ალტერნატიული პარადიგმების ძიებასთან ერთად, აუცილებელია როგორც კონექციონისტური, ასევე ოსცილატორული ხელოვნური ნეირონული ქსელების სწავლების ახალი მეთოდებისა და პროცედურების შემუშავება და შესწავლა, რასაც ხელი უნდა შეუწყოს წინამდებარე ნაშრომმაც.

ხელოვნური ნეირონული ქსელების იდეოლოგიას, ჩვენი აზრით, გააჩნია სამი ძირითადი წყარო – ადამიანის ცენტრალური ნერვული სისტემის ფიზიოლოგიის,

მაღალი პარალელიზმის გამოთვლითი სისტემებისა და სახეთა ამოცნობის თეორიის დებულებები და მეთოდები. ეს წყაროები, აგრეთვე ბიოლოგიური ნეირონის მოდელის -ხელოვნური (ფორმალური) ნეირონის სტრუქტურა და თვისებები, ერთშრიანი და მრავალშრიანი ხელოვნური ნეირონული ქსელები დეტალურად არის განხილული [1] ნაშრომში.

წინამდებარე ნაშრომში ყურადღება გამახვილებულია დისკრეტული და უწყვეტი ნეირონული ქსელების სწავლების სხვადასხვა მეთოდებისა და ალგორითმების მიმართ, განხილულია ხელოვნური ნეირონული ქსელების სწავლების სტოქასტიკური მეთოდები, აგრეთვე ჰოპფილდის ხელოვნური ნეირონული ქსელები, ასოციაციური ქსელების სხვადასხვა სახეობები და შესაბამისი სწავლების მეთოდები, გადასაწყვეტი ამოცანების ადექვატური ნეირონული ქსელების სინთეზის პროცედურები სხვ.

ხელოვნური ნეირონული ქსელების თეორიაში გამოყენებული ცნებების შესაბამისი სამეცნიერო ტერმინოლოგია საბოლოოდ არ არის ჩამოყალიბებული, ამიტომ ყურადღება ეთმობა რიგი განსაზღვრებების დაზუსტებას. ავტორები ეყრდნობოდნენ როგორც საზღვარგარეთის მრავალფეროვან პუბლიკაციას ნეიროინფორმატიკაში, ასევე საკუთარ გამოკვლევებს ხელოვნური ნეირონული ქსელებისა და სახეთა ამოცნობის თეორიაში, წინამდებარე ნაშრომში სისტემატიზებულია შესაბამისი სამეცნიერო და სასწავლო-მეთოდოლოგიური მასალები.

ნაშრომის ზოგიერთი ასპექტის დახვეწას ხელი შეუწყო იმ გამოცდილებამ, რომელიც დაგროვდა საქართველოს ტექნიკური უნივერსიტეტის ხელოვნური ინტელექტის დეპარტამენტში ნეიროინფორმატიკის საფუძვლების ლექციების კურსის წაკითხვის შედეგად. ყველა საქმიან შენიშვნასა და წინადადებას ავტორები კმაყოფილებითი მიიღებენ და გაითვალისწინებენ მომდევნო პუბლიკაციებში.

თავი I. ხელოვნური ნეირონი და ნეირონული ქსელები

1.1. ნეიროკომპიუტინგის პრობლემის მოკლე მიმოხილვა

ბუნებრივი ინტელექტუალური სისტემების მოდელირება აქტუალურია პირველ რიგში რთული და ძნელად ფორმალიზებადი პრობლემების გადასაწყვეტად, რომლებიც მრავლად არის მეცნიერების, ტექნიკის, ეკონომიკისა და კულტურის ყველა სფეროში. ასეთი პრობლემების რიგს მიეკუთვნება სახეთა ამოცნობის, კლასტერირების, ფუნქციათა აპროქსიმაციის, ასოციაციური მეხსიერების, პროგნოზირების, ოპტიმიზაციის, ინფორმაციის ფილტრაციისა და შეკუმშვის, დინამიური სისტემების იდენტიფიკაციისა და სხვა მსგავსი რთული ამოცანები. მათი გადაწყვეტის ალტერნატიული გზების ძიებისას პერსპექტიული მიმართულებაა ხელოვნური ინტელექტუალური სისტემების გამოყენება. ინფორმაციის დამუშავების ერთ-ერთ ასეთ ალტერნატიულ ტექნოლოგიას წარმოადგენს ნეიროკომპიუტინგი, ანუ ინფორმაციის დამუშავება ხელოვნური ნეირონული ქსელებით.

ხელოვნური ნეირონის პირველი ლოგიკური მოდელი 1943წ. აღწერეს თავის პიონერულ ნაშრომში ვ. მაკ-კალოკმა და ვ. პიტსმა, მათ შემოიტანეს "ზღურბლური ლოგიკური ნეირონის ცნება". მაკ-კალოკისა და პიტსის ფორმალური ნეირონი შედარებით მარტივია, იგი ბაზისურია სხვა მოდელებისათვის და დღესაც ყველაზე გავრცელებულია.

ხელოვნური ნეირონული ქსელების თეორიის განვითარებისათვის განსაკუთრებით მნიშვნელოვანი იყო დ. ჰების 1949წ. გამოქვეყნებული ნაშრომი. მან წამოაყენა ბიოლოგიური ნეირონის სინაფსების "სწავლების პოსტულატი", რომელიც ეყრდნობოდა ჰიპოთეზას სწავლების პროცესების დროს ტვინში ნერონთაშორისი კავშირების განუწყვეტელი ცვლილებების შესახებ, რის

შედეგადაც იქმნება ნეირონული ანსამბლები. ეს იდეა აღმოჩნდა ძალზე ნაყოფიერი მრავალი მკვლევარისათვის. ნ. როჩესტერმა, ჯ. ჰოლანდმა და სხვ. 1956 წ. ჩაატარეს ჰების სწავლების პოსტულატის მოდელირება ელექტრონული გამოთვლითი მანქანის საშუალებით.
ა. უიტლიმ 1956 წ. აჩვენა, რომ მოდიფიცირებადი სინაფსების გამოყენებით ხელოვნურ ნეირონულ ქსელს შეუძლია "ისწავლოს" ბინარული რეალიზაციების კლასიფიკაცია. ხელოვნური ნეირონული ქსელების თეორიის განვითარებისათვის მნიშვნელოვანია
ფ. როზენბლატის, მ. მინსკის და ს. პეიპერტის ნაშრომები. საეტაპო იყო ფ. როზენბლატის პერსეპტრონისადმი მიძღვნილი ნაშრომი (1958 წ.), შემდგომში პერსეპტრონულ მიდგომას განიხილავდა მრავალი მკვლევარი. პერსეპტრონისაგან განსხვავებული სწავლების პროცედურები შეიმუშავეს ბ. ვიდროუმ და მ. ჰოფმა ნეიროქსელისათვის adaline "neurons" (1960 წ.). შემდგომში ბ. ვიდროუმ სტუდენტებთან ერთად შეიმუშავა ერთ-ერთი პირველი სწავლადი ნეირონული სტრუქტურა Madaline (1960 წ.). ახალი ნაბიჯი იყო ჯ. კოუენის მიერ ფორმალურ ნეირონში სიგმოიდური აქტივაციის ფუნქციის შეყვანა (1967 წ.).

მ. მინსკის და ს. პეიპერტის ცნობილმა ნაშრომმა "პერსეპტრონები" (1969 წ.), სადაც თეორიულად მაღალ დონეზე მკაცრად იყო დამტკიცებული ერთშრიანი პერსეპტრონების შესაძლებლობების შეზღუდულობა, დიდი ხნით შეაფერხა ხელოვნური ნეირონული ქსელების თეორიის განვითარება, ვინაიდან ბევრმა მკვლევარმა დაკარგა ნეირონული ქსელებისადმი ინტერესი.

ხელოვნური ნეირონული ქსელების თეორიის განვითარების მეორე ეტაპი იწყება ჯ. ჰოპფილდის მიერ სრულკავშირებიანი ნეიროქსელის შემუშავებით, აგრეთვე ტ. კოჰონენის 1982წ. გამოქვეყნებული ნაშრომით. განსაკუთრებით მნიშვნელოვანი იყო მრავალშრიანი ნეირონული ქსელების სწავლების უკუგავრცელების ალგორითმისა და მისი მოდიფიკაციების შემუშავება,

რომელიც სხვადასხვა დროს, ერთმანეთისაგან დამოუკიდებლად, შეიმუშავა რამდენიმე მკვლევარმა. დღეისათვის შექმნილია ხელოვნური ნეირონული ქსელების არქიტექტურის მრავალი სახეობა და მათი შესაბამისი სწავლების მეთოდი. ხელოვნური ნეირონული ქსელების ყველაზე მნიშვნელოვანი მოდელები განხილულია მომდევნო თავებში.

1.1.1. ნეიროკომპიუტინგით გადასაწყვეტი ამოცანების პროტოტიპები

წარმოვადგინოთ ხელოვნური ნეირონული ქსელების გამოყენებით გადასაწყვეტი ძირითადი პრობლემები, რომელთა წრე ძალზე ფართოა. ნეირონული ქსელი ზოგადად წყვეტს მრავალგანზომილებიანი ფუნქციების აპროქსიმაციის ამოცანას, ანუ მოცემული მაგალითების ნაკრების განმაზოგადებელი მრავალგანზომილებიანი ასახვის აგების ამოცანას.

ფუნქციათა აპროქსიმაცია. ვთქვათ, მოცემულია სასწავლო ნაკრები $\{(x_1,y_1), (x_2,y_2),\ldots,(x_n,y_n)\}$, ანუ ნეირონული ქსელის შემავალი და გამომავალი მონაცემების (ვექტორების) წყვილები, რომლებიც ხმაურით დამახიჯებული უცნობი F(x) ფუნქციის მიერ არის გენერირებული. აპროქსიმაციის მიზანია უცნობი F(x) ფუნქციის შეფასების პოვნა. ნეირონული ქსელის გამომავალი სიგნალების ტიპი განსაზღვრავს უცნობი ფუნქციის აპროქსიმაცია მიიღებს კლასიფიკაციის თუ რეგრესიის სახეს. იმ შემთხვევაში, თუ გამომავალი სიგნალები დისკრეტულია, აპროქსიმაცია მიიღებს კლასიფიკაციის სახეს, ხოლო თუ გამომავალი სიგნალები უწყვეტია, ფუნქციის აპროქსიმაცია მიიღებს რეგრესიის სახეს. ნაშრომში [8] დამტკიცებულია თეორემა, რომ მრავალი ცვლადის ნებისმიერი უწყვეტი ფუნქციის მიახლოება ნებისმიერი სიზუსტით შეიძლება

წრფივი ფუნქციების წრფივი კომბინაციების და სუპერპოზიციების და ერთი არგუმენტის ნებისმიერი ერთი არაწრფივი ფუნქციის მეშვეობით. ამ თეორემიდან გამომდინარეობს, რომ ნეირონული ქსელი აპროქსიმაციის უნივერსალურ საშუალებას წარმოადგენს და შეუძლია ნებისმიერი უწყვეტი ავტომატის ნებისმიერი სიზუსტით იმიტაცია. ფუნქციათა აპროქსიმაცია აუცილებელია მრავალი საინჟინრო და სამეცნიერო ამოცანის მოდელირების დროს.

სახეთა კლასიფიკაცია. ამოცანა ცნობილია აგრეთვე როგორც სახეთა ამოცნობა მასწავლებლით; წინასწარ განსაზღვრულია სახეთა რამდენიმე კლასი, მათგან ერთ-ერთს უნდა მიეკუთვნოს ამოსაცნობი სახე, რომელიც წარედგინება ნეირონულ ქსელს შემავალი სიგნალებით - სახის ნიშანთა ვექტორით (რეალიზაციით). ამოცანების ამ ჯგუფს მიეკუთვნება ანბანის ამოცნობა, სისხლის უჯრედების კლასიფიკაცია, ელექტროკარდიოგრამის სიგნალების კლასიფიკაცია, სამეტყველო სიგნალების ამოცნობა და მრავალი სხვა.

კლასტერირება (კატეგორირება). ამოცანა ცნობილია აგრეთვე როგორც სახეთა კლასიფიკაცია მასწავლებლის გარეშე. კლასტერი განსაზღვრებით არის ნებისმიერი სიმრავლის ისეთი ქვესიმრავლე, რომელშიც ობიექტები გაერთიანებულია რაიმე ნიშნის ან ნიშნების მიხედვით, ხოლო კლასტერირება წარმოადგენს ობიექტთა მოცემული სიმრავლის რაიმე ნიშნის ან ნიშნების მიხედვით არაგადამკვეთ ქვესიმრავლეებად დაყოფა-დაჯგუფების პროცესს. კლასტერირების ამოცანების განმასხვავებელი ნიშანია, რომ არ გვაქვს სასწავლო ნაკრები კლასების მითითებით, ანუ ვექტორების მასწავლებელი წყვილები. კლასტერირების პროცედურა ეფუძნება სახეთა მსგავსებას, იგი ათავსებს ერთ კლასტერში მსგავს სახეებს. ნეირონული ქსელებით კლასტერირება გამოიყენება მონაცემთა შეკუმშვისათვის, მონაცემთა თვისებების კვლევისათვის, ცოდნის მოპოვების ამოცანებში და სხვა.

პროგნოზირება. ზუსტი პროგნოზების გაკეთება საზოგადოების უდიდესი მოთხოვნილება და ამასთანავე აუცილებლობაც არის, ვინაიდან პროგნოზირება მნიშვნელოვან გავლენას ახდენს გადაწყვეტილებების მიღებაზე.

ვთქვათ, მოცემულია n დისკრეტული ანათვალი $\{y(t_1),\ldots\ldots,y(t_k)\}$ დროის t_1, $t_2,\ldots t_k$ მიმდევრობით მომენტში, საჭიროა დროის რომელიმე მომავალ t_{k+1} მომენტში $y(t_{k+1})$ მნიშვნელობის პროგნოზირება.. ნეიროკომპიუტინგის მეთოდებს შეუძლია ისეთი ძნელადპროგნოზირებადი მოვლენების საკმარისად ზუსტი პროგნოზების გაკეთება, როგორიც არის ადამიანის ჯანმრთელობა და ქცევა, ამინდი, ეკონომიკური მაჩვენებლები, სპორტული თამაშები, კორპორატიული გაყიდვები, საფონდო ბირჟის ტრანზაქციები და მრავალი სხვა.

ოპტიმიზაცია. ოპტიმიზაციის ალგორითმის მიზანია ისეთი გადაწყვეტილების მონახვა, რომელიც აკმაყოფილებს შეზღუდვების სისტემას და ამავე დროს უზრუნველყოფს მიზნობრივი ფუნქციის მაქსიმიზირებას ან მინიმიზირებას. ოპტიმიზაციის ამოცანები მრავლად არის ეკონომიკაში, მედიცინაში, მათემატიკაში, ფიზიკისა და ტექნიკაში.
ოპტიმიზაციის ამოცანის კლასიკური მაგალითია კომივოიაჟორის ცნობილი ამოცანა.

მართვა. ვთქვათ, მოცემულია დინამიური სისტემა $\{u(t), y(t)\}$, სადაც $u(t)$ არის შემავალი მმართველი ზემოქმედება, $y(t)$ არის სისტემის გამოსასვლელის მნიშვნელობა დროის t მომენტში. ეტალონური მოდელით მართვის სისტემებში მართვის მიზანი არის ისეთი $u(t)$ შემავალი მმართველი ზემოქმედების გაანგარიშება, რომელიც ეტალონური მოდელის შესაბამისად უზრუნველყოფს სისტემის სასურველ ტრაექტორიას. მაგალითად შეიძლება მოვიყვანოთ ძრავის ოპტიმალური მართვის ამოცანა.

ასოციაციური მეხსიერება. ადამიანის მეხსიერება ასოციაციურია, ანუ რაიმე ინფორმაციის გაგებამ შესაძლებელია მასთან დაკავშირებული სხვადასხვა მოვლენებისა და საგნების გახსენება გამოიწვიოს, მაშინ როდესაც ტრადიციული ფონ ნეიმანის არქიტექტურის კომპიუტერის მეხსიერებიდან სასურველი ინფორმაციის მოძიება მხოლოდ ამ ინფორმაციის ზუსტი მისამართის მითითებით არის შესაძლებელი, ხოლო აღნიშნულ ინფორმაციასთან დაკავშირებული სხვადასხვა მოვლენებისა და საგნების შესახებ ინფორმაცია მათი ზუსტი მისამართების მითითების გარეშე მიუწვდომელია. უკუკავშირებიანი ნეირონული ქსელის გარკვეული თვისებები განაპირობებს ასოციაციური მეხსიერების ფორმირების შესაძლებლობას, რომელიც ადამიანის ასოციაციური მეხსიერების მსგავსია. ასოციაციურ მეხსიერებას, რომელიც ფორმირებულია ნეირონული ქსელით, შეუძლია ადამიანის მეხსიერების მსგავსად, საჭირო ინფორმაციის მოცემული ნაწილის მიხედვით მოიძიოს მეხსიერებიდან სრული ინფორმაცია, მას გააჩნია განზოგადების უნარი, შეუძლია გამოიმუშაოს სწორი რეაქციები მაშინაც, როდესაც ნეიროქსელის შესასვლელს მიეწოდება ნაწილობრივ არასწორი (დამახინჯებული) ან არასრული სახე.

1.2. ხელოვნური ნეირონული ქსელების აქსიომატიკა

ხელოვნური ნეირონული ქსელების თეორია, როგორც აღნიშნული იყო ფორმირების ეტაპზეა, რაც განაპირობებს პრობლემების დასმისა და ძირითად განსაზღვრებათა მრავალფეროვნებას. არსებობს ნეიროქსელების მრავალი არქიტექტურა და სწავლების მრავალი ალგორითმი, ხოლო სხვადასხვა ავტორები ნეიროქსელებთან დაკავშირებულ ცნებებს იყენებენ ან მკაცრი განსაზღვრებების გარეშე, ან განსხვავებული ინტერპრეტაციით იმისდა მიხედვით, თუ მოდელირების რა პრინციპები და მეთოდებია გამოყენებული. საჭირო

ხდება ამ ცნებების აქსიომატიკურად განსაზღვრა ან დაზუსტება. ბუნებრივი ნეირონისა და ნეირონული ქსელების ბიომორფოლოგიური აგებულებიდან გამომდინარე აქსიომატიკური ცნებები, რომელთა მეშვეობით უნდა აიგოს ამ ობიექტების მოდელები, შეიძლება რამდენიმე ჯგუფად დაიყოს. ქვემოთ მოცემულია სხვადასხვა ნეიროობიექტების აქსიომატიკური განსაზღვრებები.

განსაზღვრება 1.2.1. ერთ ცალკე აღებულ ბიოლოგიური (ბუნებრივი) ნეირონის მოდელს ეწოდება ხელოვნური ანუ ფორმალური ნეირონი.

განსაზღვრება 1.2.2. ბუნებრივი ნეირონის სომაში შემავალი სიგნალების აკუმულირების პროცესის მოდელად მივიჩნევთ აჯამვის პროცედურას, ხოლო თვით სომას მოდელს ეწოდება ამჯამავი ელემენტი.

განსაზღვრება 1.2.3. ბიოლოგიური ნეირონის სინაფსური კავშირის მოდელი წარმოადგენს გამრავლების ელემენტს, რომელიც ახორციელებს სიგნალის მნიშვნელობის სკალარულ გამრავლებას რაიმე რიცხვზე, რომელიც ასახავს ან მაძლიერებელ კავშირს, ან მამუხრუჭებელ კავშირს.

განსაზღვრება 1.2.4. სკალარულ სიდიდეებს, რომელზეც მრავლდება სიგნალების მნიშვნელობები, სინაფსური კავშირების წონითი კოეფიციენტები, ანუ შემოკლებით სინაფსური წონები ეწოდება.

თუ წონითი კოეფიციენტის მნიშვნელობა ერთზე მეტია, მაშინ სინაფსური კავშირი არის მაძლიერებელი (ამგზნები), თუ წონითი კოეფიციენტის მნიშვნელობა ერთზე ნაკლებია, მაშინ სინაფსური კავშირი არის მამუხრუჭებელი.

ცნობილია, რომ ბიოლოგიური ნეირონის აქსონურ არხებს გააჩნია გადამრთველის თვისება გარკვეული

ზღურბლური მნიშვნელობით, რომელიც სწავლების პროცესში შეიძლება შეიცვალოს.

განსაზღვრება 1.2.5. ბიოლოგიური ნეირონის აქსონური კავშირის მოდელს, რომელშიც განხორციელებულია გადამრთველის თვისება, ეწოდება ფორმალური ნეირონის არაწრფივი გარდამქმნელი ელემენტი, ანუ აქტივაციის ელემენტი.

ფორმალური ნეირონის არაწრფივი გარდამქმნელი ელემენტი ახდენს ამჯამავი ელემენტიდან გამოსული სიგნალის მოდიფიცირებას, შედეგად ნეირონის აქსონზე მიიღება ნეირონულ ქსელში შემდგომი პროცედურების ჩასატარებლად უფრო მოხერხებული მოდიფიცირებული სიგნალი.

განსაზღვრება 1.2.6. ნეირონული შრე ეწოდება ფორმალური ნეირონების გარკვეული წესით დაკავშირებულ ერთობლიობას, რომელსაც გააჩნია საერთო რეცეპტორული ველი (სიგნალების ელემენტების ნაკრები), გარკვეული ფუნქციური დანიშნულება და სადაც შესაძლებელია სწავლების პროცესის განხორციელება.

განსაზღვრება 1.2.7. ნეირონულ შრეს, სადაც ხდება გარემოდან შემოსული სიგნალების აღქმა და ამ გარე ზემოქმედების შესაბამისი ვექტორის ელემენტების ნეირონული ქსელის მომდევნო შრის ფორმალურ ნეირონებზე რაიმე ცვლილების გარეშე განაწილება, ეწოდება მიმღები ანუ ნულოვანი შრე.

განსაზღვრება 1.2.8. ნეირონული შრეს, რომელიც წინამდებარე შრესა და გარემოს უკავშირდება, ეწოდება ნეირონული ქსელის გამომავალი შრე.

განსაზღვრება 1.2.9. ნეირონულ შრეებს, რომლებიც მიმღებ, ანუ ნულოვან შრესა და გამომავალ შრეებს შორის მდებარეობს ეწოდება საშუალედო ანუ ფარული შრეები.

განსაზღვრება 1.2.10. ხელოვნური ნეირონული ქსელი წარმოადგენს ფორმალური ნეირონების ან/და ნეირონული შრეების გარკვეული წესით დაკავშირებულ ერთობლიობას, რომლებიც ერთმანეთსა და გარემოს უკავშირდება ფორმალური ნეირონების პარამეტრებით განსაზღვრული კავშირებით და სადაც შესაძლებელია სწავლების პროცესის განხორციელება.

განსაზღვრება 1.2.11. ხელოვნური ნეირონული ქსელის გამოსასვლელი სიგნალების მიმდინარე მნიშვნელობების სიმრავლეს ეწოდება ნეირონული ქსელის მდგომარეობა.

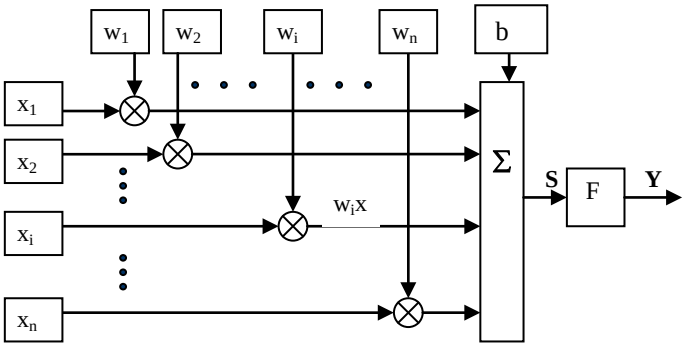
მივიღოთ, რომ ხელოვნურ ნეირონულ ქსელში, განსხვავებით ბუნებრივი ნეირონული ქსელისაგან, გვაქვს ე. წ. იდეალური არხები, რომლებიც დაუმახინჯებლად და მყისიერად გადასცემენ სიგნალებს მომდევნო ელემენტებს

1.3. ფორმალური ნეირონის სტრუქტურა და თვისებები

ხელოვნური ანუ ფორმალური ნეირონი წარმოადგენს ცალკე აღებული ბიოლოგიური ნეირონის ისეთ მოდელს, რომელშიც გათვალისწინებულია მისი მხოლოდ გამორჩეული ინტეგრალური თვისებების რეკონსტრუქცია, ანუ მოდელში უგულებელყოფილია ბიოლოგიური ნეირონის თვისებების გაცილებით უფრო რთული კომპლექსი. ბიოლოგიური ნერვული უჯრედების ანალოგიურად თითოეული ხელოვნური ანუ ფორმალური ნეირონი ხასიათდება თავისი მიმდინარე მდგომარეობით და შეიძლება იყოს აგზნებული ან დამუხრუჭებული. ფორმალურ ნეირონს აქვს მრავალი შესასვლელი (დენდრიტი) და სინაფსური კავშირების ჯგუფი, აგრეთვე ერთი – გამოსასვლელი (აქსონი), რომლითაც ამ ნეირონის გამომავალი სიგნალი (აგზნების ან დამუხრუჭების) გადაეცემა მომდევნო ნეირონს, ზოგჯერ კი საკუთარ თავსაც. ფორმალური ნეირონი

შედგება სამი ტიპის ელემენტისაგან: გამრავლების ელემენტები (სინაფსები, მათი რაოდენობა შემავალი ვექტორის განზომილების ტოლია); ამჯამავი ელემენტი (ზოგჯერ უწოდებენ ადაპტურ სუმატორს); არაწრფივი გარდამქმნელი ელემენტი (აქტივაციის ელემენტი).

სინაფსები აძლიერებს ან ასუსტებს ნეირონში შემავალ სიგნალს იმის მიხედვით, თუ როგორია შესაბამისი სინაფსური კავშირის ძალა ანუ წონითი კოეფიციენტი. თუ წონითი კოეფიციენტი ერთზე მეტია შემავალი სიგნალი გაძლიერდება, ხოლო თუ წონითი კოეფიციენტი ერთზე ნაკლებია, სიგნალი შესუსტდება. 1.1 სურათზე ნაჩვენებია ფორმალური ნეირონის სტრუქტურა.



სურ. 1.1. ფორმალური ნეირონის სტრუქტურა.

ფორმალური ნეირონის მიმდინარე მდგომარეობა განისაზღვრება როგორც მისი შესასვლელების შეწონილი ჯამი

$$S = \sum_{i=1}^{n} w_i x_i + b, \qquad (1.1)$$

ხოლო ფორმალური ნეირონის გამომავალი სიგნალი

არის მისი მიმდინარე მდგომარეობის ფუნქცია:

$$y = F(S) = F(\sum_{i=1}^{n} w_i x_i + b), \qquad (1.2)$$

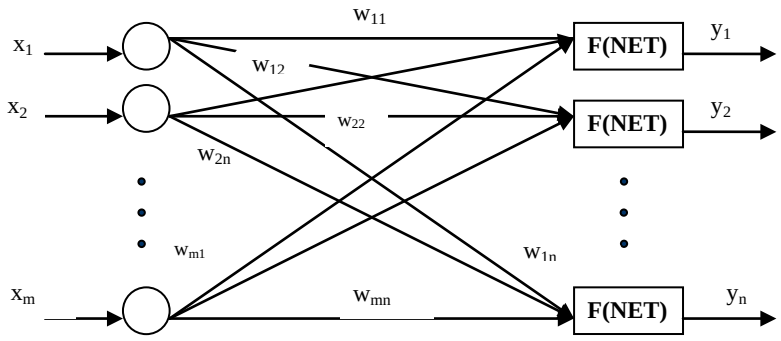
სადაც w_i არის სინაფსური წონა (weight), $i = 1, 2 \ldots n$;
x_i - შემავალი ვექტორის ელემენტი, $i = 1, 2 \ldots n$;
n - შემავალი ვექტორის განზომილება; b - წანაცვლების მნიშვნელობა (bias); S- შეჯამების შედეგი (sum) , ზოგჯერ იყენებენ NET აღნიშვნას; y - ნეირონის გამომავალი სიგნალი, იყენებენ OUT აღნიშვნასაც;
F - არაწრფივი გარდაქმნა (აქტივაციის ფუნქცია).
ფორმალური ნეირონის მათემატიკური მოდელი გვიჩვენებს, რომ არაწრფივი გარდამქმნელი ელემენტი ახორციელებს ერთი არგუმენტის – ამჯამავი ელემენტის გამოსასვლელის არაწრფივ ფუნქციას, ხოლო მთლიანად ფორმალური ნეირონი კი ვექტორული არგუმენტის სკალარულ ფუნქციას. არაწრფივი გარდამქმნელი ელემენტი ახდენს ამჯამავი ელემენტიდან გამოსული სიგნალის მოდიფიცირებას, რის შედეგადაც ნეირონის აქსონზე მიიღება ნეიროქსელში შემდგომი პროცედურების ჩასატარებლად უფრო მოხერხებული მოდიფიცირებული სიგნალი. ყველაზე ხშირად აქტივაციის ფუნქციებად იყენებენ ზღურბლურ და სიგმოიდურ ფუნქციებს. თუ აქტივაციის F ფუნქცია S შეწონილი ჯამის სიდიდის ცვლილების დიაპაზონს ავიწროებს ისე, რომ S-ის ნებისმიერი მნიშვნელობისას აქსონზე მიღებული სიგნალი y ეკუთვნის გარკვეულ სასრულ ინტერვალს, მაშინ F ფუნქციას უწოდებენ „შემკუმშავ" ფუნქციას. შემკუმშავ ფუნქციად იყენებენ სიგმოიდურ ფუნქციებს, კერძოდ ლოგ-სიგმოიდურ ფუნქციას

$$F(S) = \frac{1}{1 + e^{-\alpha s}}. \qquad (1.3)$$

ერთ ცალკე აღებულ ფორმალურ ნეირონს შეუძლია სახეთა ამოცნობის უმარტივესი პროცედურების შესრულება, ხოლო ნეირონული გამოთვლების ეფექტურობას განაპირობებს მრავალი ნეირონის გაერთიანება ხელოვნურ ნეირონულ ქსელებში.

1.4. ერთშრიანი და მრავალშრიანი ხელოვნური ნეირონული ქსელები

უმარტივესი ნეირონული ქსელი შედგება ნეირონების ჯგუფისაგან, რომლებიც გაერთიანებულია ერთ შრეში (სურ. 1.2), სადაც წრეებით აღნიშნულია ე.წ. ნულოვანი შრე, რომელიც არ ასრულებს გამოთვლით ოპერაციებს, არამედ ანაწილებს ქსელში შემავალ სიგნალებს გამომთვლელ ნეირონებზე, რომლებიც გამოსახულია მართკუთხედებით.

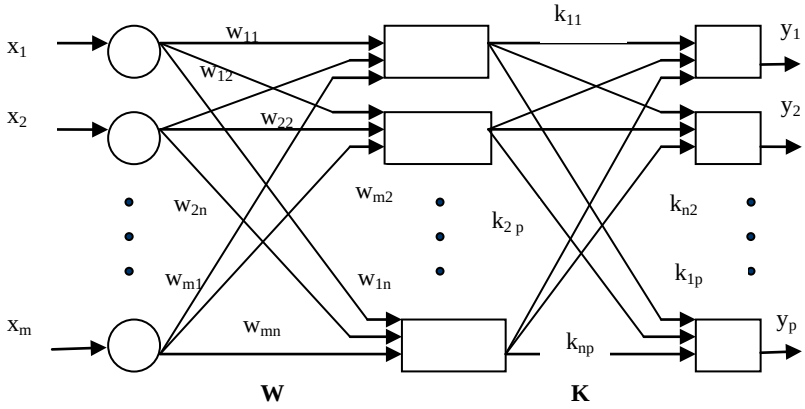


სურ. 1.2. ერთშრიანი ხელოვნური ნეირონული ქსელი.

X შემავალი ვექტორის ელემენტების (x_1, x_2,…. x_n) ნამრავლი შესაბამის წონით კოეფიციენტებზე (სინაფსურ წონებზე) მიეწოდება თითოეულ ნეირონს, ხოლო შესასვლელების შეწონილ ჯამს (y_1, y_2 , ….y_m) თითოეული

ნეირონი გადასცემს ნეირონული ქსელის მომდევნო კომპონენტს. საკმაოდ მოხერხებულია სინაფსური წონების წარმოდგენა **W** მატრიცის ელემენტებად. მატრიცას აქვს m სტრიქონი და n სვეტი, სადაც m არის შესასვლელების რიცხვი, ხოლო n – ნეირონების რიცხვი. მაგალითად, მატრიცის ელემენტი w_{12} არის წონა, რომელიც აკავშირებს პირველ შესასვლელს მეორე ნეირონთან, ხოლო w_{21} არის წონა, რომელიც აკავშირებს მეორე შესასვლელს პირველ ნეირონთან. ამრიგად, გამომავალი **N** ვექტორის გამოთვლა, რომლის კომპონენტებიც ნეირონების გამოსასვლელებია, დაიყვანება მატრიცულ გამრავლებაზე **N=XW**, სადაც **N** და **X** ვექტორ–სტრიქონებია. ხშირად ნეირონის გამოსასვლელს აღნიშნავენ OUT – ით, მაშინ OUT = **XW.**
მრავალშრიან ხელოვნურ ნეირონულ ქსელებს ერთშრიანი ნეირონულ ქსელებთან შედარებით გაცილებით მეტი გამოთვლითი შესაძლებლობები აქვს. დამუშავებულია მრავალშრიანი ხელოვნური ნეირონული ქსელების მრავალი სახეობა და შექმნილია მათი შესაბამისი სწავლების ალგორითმები, რომელთაც ახასიათებს თავისი ძლიერი და სუსტი მხარეები. მრავალშრიანი ნეირონული ქსელები შეიძლება შეიქმნას შრეების კასკადური შეერთებით, მაშინ ერთი შრის გამოსასვლელი იქნება მომდევნო შრის შესასვლელი. ითვლება, რომ სინაფსური წონები უშუალოდ დაკავშირებულია მათი მომდევნო შრის ნეირონებთან. ამრიგად, შრე შედგება სინაფსური წონებისა და მათი მომდევნო ნეირონებისაგან, რომლებიც შეწონილ ჯამებს გამოითვლიან. 1.3 სურათზე ნაჩვენებია ორშრიანი ხელოვნური ნეირონული ქსელი. მრავალშრიან ქსელებს ერთშრიან ქსელებთან შედარებით გაცილებით მეტი შესაძლებლობები აქვს მხოლოდ იმ შემთხვევაში, თუ ნეირონების აქტივაციის ფუნქცია იქნება არაწრფივი. წრფივი აქტივაციის ფუნქციის მქონე ორშრიანი ქსელის გამოსასვლელი გამოითვლება მასში შემავალი ვექტორის

გამრავლებით პირველ წონით მატრიცაზე, შემდეგ კი პირველი შრის მიერ გამომუშავებული შედეგის გამრავლებით მეორე წონით მატრიცაზე: $(XW_1)W_2$, იმის გამო, რომ მატრიცების გამრავლებას აქვს ასოციაციურობის თვისება, გვექნება $X(W_1W_2)$, ანუ ნათელია, რომ ორშრიანი ხელოვნური ნეირონული



სურ. 1.3. ორშრიანი ხელოვნური ნეირონული ქსელი.

ქსელი წრფივი აქტივაციის ფუნქციით ეკვივალენტურია ისეთი ერთშრიანი წრფივი ქსელისა, რომლის წონითი კოეფიციენტების მატრიცა არის პირველი და მეორე წონითი კოეფიციენტების მატრიცების ნამრავლი. აქედან გამომდინარეობს შედეგი: ნებისმიერი მრავალშრიანი ხელოვნური ნეირონული ქსელი, რომლის ნეირონებსაც აქვს წრფივი აქტივაციის ფუნქცია, შეიძლება შეიცვალოს მისი ეკვივალენტური ერთშრიანი ნეირონული ქსელით. ამრიგად ნეირონული ქსელის გამოთვლითი შესაძლებლობების გაზრდისათვის აუცილებელი პირობაა ქსელში არაწრფივი აქტივაციის ფუნქციის შეტანა.

1.5. ხელოვნური ნეირონული ქსელების კლასიფიკაცია

ხელოვნური ნეირონული ქსელების კლასიფიკაცია შესაძლებელია სხვადასხვა ნიშნის მიხედვით.

ნეირონულ ქსელში შესრულებული ფუნქციის მიხედვით თვით ნეირონები იყოფა შემავალ, საშუალედო (ფარულ) და გამომავალ ნეირონებად. შემავალ ნეირონებს მიეწოდება გარე ზემოქმედების შესაბამისი ვექტორი, რომლის ელემენტებსაც იგი გადასცემს სხვა ნეირონებს ცვლილებების გარეშე, ანუ შემავალი ნეირონები ზეგავლენას არ ახდენენ ქსელის გამოთვლით შესაძლებლობებზე, არამედ ემსახურება პირველი სიმრავლის განაწილებას პირველი შრის ნეირონებზე, ხოლო საშუალედო და გამომავალი ნეირონები (1.1) და (1.2) გამოსახულებების შესაბამისად, ასრულებს გამოთვლით პროცედურებს და შედეგებს გადასცემს სხვა ნეირონებს ან ქსელის გამოსასვლელს.

ფორმალური ნეირონის სტრუქტურის მიხედვით ნეირონული ქსელების დაყოფა შეიძლება ჰომოგენურ და ჰეტეროგენურ ქსელებად. ჰომოგენური ქსელები (ერთგვაროვანი ქსელები) შედგება ერთი ტიპის ნეირონებისაგან, რომელთაც ერთნაირი აქტივაციის ფუნქცია აქვს, ხოლო ჰეტეროგენურ ქსელებში შემავალ ნეირონებს აქვს სხვადასხვა აქტივაციის ფუნქციები.

ფორმალური ნეირონის აქტივაციის ფუნქცია განსაზღვრავს არის ნეირონული ქსელი დისკრეტული (ბინარული) თუ უწყვეტი (ანალოგური). დისკრეტული ნეირონული ქსელი ოპერირებს ბინარული სიდიდეებით, ანუ თითოეული ნეირონის გამოსასვლელმა შეიძლება მიიღოს ლოგიკური ნულის (დამუხრუჭებული მდგომარეობა) ან ლოგიკური ერთის (აგზნებული მდგომარეობა) მნიშვნელობა. უწყვეტი ნეირონული ქსელი კი ოპერირებს ნამდვილი რიცხვებით, ასეთი ქსელების აქტივაციის ფუნქციები წარმოადგენს სხვადასხვა სახის სიგმოიდურ ფუნქციებს.

ტოპოლოგიის თვალსაზრისით ნეიროქსელები ოთხ სახეობად იყოფა: სრულკავშირებიანი ნეიროქსელები; მრავალშრიანი ნეიროქსელები; ლოკალურკავშირებიანი ნეიროქსელები; არასტრუქტურირებული ნეიროქსელები. სრულკავშირებიან ქსელებში თითოეული ნეირონი თავის გამომავალ სიგნალს გადასცემს ყველა ნეირონს, მათ შორის საკუთარ თავსაც, ყველა შემავალი სიგნალი მიეწოდება ყველა ნეირონს, ხოლო ქსელის გამომავალ სიგნალებად შეიძლება გამოყენებული იყოს ქსელის ფუნქციონირების რამდენიმე ტაქტის შემდეგ ყველა ან ზოგიერთი ნეირონის გამომავალი სიგნალი. მრავალშრიან ნეირონულ ქსელებში ნეირონები გაერთიანებულია შრეებად. შრეს აყალიბებს ნეირონები, რომელთაც აქვთ საერთო შემავალი სიგნალები, ანუ საერთო რეცეპტორული ველი.

ნეირონების რაოდენობა შრეში შეიძლება იყოს ნებისმიერი და არ არის დამოკიდებული მათ რაოდენობაზე სხვა შრეებში. ზოგადად, ნეიროქსელში შეიძლება იყოს რამდენიმე შრე, რომლებიც მარცხნიდან მარჯვნივ ინომრება. შემავალ შრეს ნულოვან ნომერს ანიჭებენ, ამით ხაზი ესმება იმას, რომ შემავალი შრე არ აწარმოებს გამოთვლით პროცედურებს. შემავალი და გამომავალი შრეების გარდა, შეიძლება არსებობდეს რამდენიმე საშუალედო, ანუ ფარული შრე (როგორც წესი, მათი რაოდენობა ხუთს არ აღემატება). რომელიმე k შრის ნეირონების გამოსასვლელებიდან (k+1) მომდევნო შრის ნეირონების შესასვლელებთან კავშირებს ეწოდება მიმდევრობითი კავშირები. თავის მხრივ, მრავალშრიანი ნეიროქსელები სამ ჯგუფად იყოფა: მონოტონური ნეირონული ქსელები; პირდაპირი გავრცელების ნეირონული ქსელები; უკუკავშირებიანი ნეირონული ქსელები. მონოტონურ ქსელებში თითოეული შრე, გარდა უკანასკნელისა, შედგება ორი ბლოკისაგან, ანუ ამგზნები და მამუხრუჭებელი ნეირონების ჯგუფებისაგან. ბლოკებს შორის კავშირებიც ასევე იყოფა ამგზნებ და მამუხრუჭებელ კავშირებად. ერთი შრის შემადგენელ

ნეირონებს შორის კავშირებს, ეწოდება ლატერალური კავშირები. პირდაპირი გავრცელების ნეირონულ ქსელებში, როგორც წესი, თითეული k შრის ნეირონის გამომავალი სიგნალი მიეწოდება (k+1) შრის თითოეული ნეირონის შესასვლელს; თუმცა შესაძლებელია, რომ k შრე დაუკავშირდეს უშუალოდ რომელიმე (k+p) შრეს.

უკუკავშირებიან ნეირონულ ქსელებში სიგნალები გადაეცემა როგორც მომდევნო შრეს, ასევე რომელიმე წინამდებარე შრეს. ასეთ ქსელებში შეიძლება ორი ძირითადი სახეობის გამოყოფა:

ა) ციკლური ნეიროქსელები, სადაც უკანასკნელი შრე თავის გამომავალ სიგნალებს გადასცემს პირველ შრეს, ყველა შრე თანაბრად არის უფლებამოსილი მიიღოს ქსელში შემოსული სიგნალები და გასცეს ქსელიდან გამომავალი სიგნალები;

ბ) რეკურენტული სრულკავშირებიანი ნეირონული ქსელები, აქ თითოეული შრე სრულკავშირებიანია, სიგნალების გადაცემა ხდება როგორც შრიდან შრეზე, ასევე საკუთრივ შრის ნეირონებს შორის, ამასთან, სამუშაო ციკლი სამ ეტაპად იყოფა - სიგნალების მიღება წინამდებარე შრისაგან; სიგნალების გაცვლა შრის შიგნით; გამომავალი სიგნალების გამომუშავება და გადაცემა მომდევნო შრისათვის.

ლოკალურკავშირებიან ნეირონულ ქსელებში ნეირონები უკავშირდება თავის უახლოეს ოთხ, ექვს ან რვა ტოპოლოგიურ მეზობელს.

ისეთ ნეიროქსელებს, რომელთა მიკუთვნებაც ზემოთ აღწერილი სამი ჯგუფისათვის შეუძლებელია, უწოდებენ არასტრუქტურირებულ ნეირონულ ქსელებს.

არსებობს სინქრონული და ასინქრონული ხელოვნური ნეიროქსელები. სინქრონულ ქსელებში დროის ყოველ მომენტში თავის მდგომარეობას იცვლის მხოლოდ ერთი ნეირონი, ხოლო ასინქრონულ ქსელებში თავის მდგომარეობას ერთდროულად იცვლის ნეირონთა ჯგუფი ან მთლიანად შრე.

თავი II. ხელოვნური ნეირონული ქსელების სწავლება

2.1.ხელოვნური ნეიროქსელების სწავლების პრობლემა

ხელოვნური ნეირონული ქსელის ყველაზე მნიშვნელოვანი თვისება არის ის, რომ მას აქვს ”სწავლების” უნარი (ზოგჯერ იხმარება ტერმინი ”ტრენინგი”). სწავლების მიზანია ქსელის გამოსასვლელებზე შემავალი სიგნალების შესაბამისი სასურველი, ან მასთან მიახლოებული გამომავალი სიგნალების მნიშვნელობების მიღება. თითოეული შემავალი და გამომავალი სიგნალების სიმრავლე განიხილება, როგორც ვექტორი. სწავლება ხორციელდება ქსელისათვის შემავალი ვექტორების თანმიმდევრობითი წარდგენით და იმავდროულად, გარკვეული პროცედურების შესაბამისად, ქსელის პარამეტრების გადაწყობით ისე, რომ თითოეული შემავალი ვექტორი გამოიმუშავებს შესაბამის გამომავალ ვექტორს. ხელოვნურ ნეირონულ ქსელებშიც, ისევე როგორც ზოგადად ამომცნობ სისტემებში, განასხვავებენ სწავლების ალგორითმებს მასწავლებლით (supervised learning) და მასწავლებლის გარეშე (self-organized learning), სწავლების ამ პარადიგმების ჩარჩოებში შექმნილია ნეირონული ქსელების სწავლების მრავალი ალგორითმი და მათი მოდიფიკაცია.

ერთშრიანი ნეირონული ქსელების სწავლება შესაძლებელია მხოლოდ მასწავლებლით, ხოლო შესაბამისი ალგორითმი შედარებით მარტივია, ვინაიდან ერთადერთი შრის აქსონების (აქ და შემდგომში აქსონის ქვეშ იგულისხმება სიგნალის მნიშვნელობა ნეირონის გამოსასვლელზე) სასურველ მნიშვნელობებს განსაზღვრავს მასწავლებელი და სინაფსური წონების გადაწყობის პროცესი მიმდინარეობს აქსონებზე შეცდომის მინიმიზირებისს მიმართულებით. ცნობილია, რომ ერთშრიანი ნეირონული ქსელების გამოთვლითი და

შესაბამისად სწავლების შესაძლებლობები პრინციპულად არის შეზღუდული (მინსკი მ., პეიპერტი ს.), ამიტომ ასეთი ქსელები არ იძლევა რთული ამოცანების გადაწყვეტის შესაძლებლობას. მრავალშრიანი ნეირონული ქსელების (სადაც ეს შეზღუდვები მოხსნილია) სწავლების პროცესები კი დაკავშირებულია სერიოზულ პრობლემებთან, ვინაიდან უცნობია თითოეული შრის (უკანასკნელის გარდა) გამოსასვლელების - აქსონების ოპტიმალური მნიშვნელობები, ხოლო მთლიანად ქსელის სწავლების პროცესის განხორციელება მხოლოდ უკანასკნელი შრის გამოსასვლელებზე შეცდომების სიდიდეების განსაზღვრითა და გამოყენებით შეუძლებელია. ამ პრობლემის გადაჭრას ცდილობენ თითოეული შრისათვის შემავალი სიგნალების ნაკრებების შესაბამისი გამომავალი სიგნალების ნაკრებების შემუშავებით, რაც ხშირად არარეალურია, ვინაიდან საჭიროა ძალზე დიდი მოცულობის შრომატევადი გამოთვლითი ოპერაციებს შესრულება. მეორე გზაა სინაფსური წონების დინამიკური მოდიფიცირება, პროცედურას ახორციელებენ ბიჯურად, ამასთან ირჩევენ განსაკუთრებით სუსტ სინაფსურ კავშირებს, ცვლიან მათ მცირე სიდიდით ამა თუ იმ მიმართულებით და ინახავენ მხოლოდ იმ ცვლილებებს, რომლებმაც გამოიწვიეს შეცდომის შემცირება მთლიანად ქსელის გამოსასვლელებზე, მაგრამ ეს მეთოდიც მოითხოვს მეტად დიდი მოცულობის რუტინულ გამოთვლებს. მესამე გზაა სწავლების უკუგავრცელების ალგორითმის გამოყენება, თუმცა ამ ალგორითმით რთული ამოცანის გადაწყვეტისას ნეიროქსელის სწავლების პროცესი შეიძლება მიუღებლად დიდხანს გაგრძელდეს, ან საერთოდ ვერ დასრულდეს. ამას ჩვეულებრივ ორი მიზეზი აქვს, ნეირონული ქსელის ეგრეთ წოდებული ”დამბლა” ან მოხვედრა ლოკალურ მინიმუმში. სწავლების პროცესში კორექციების შედეგად სინაფსურმა წონებმა შეიძლება ძალზე დიდი მნიშვნელობები მიიღოს, ამის გამო ნეირონების

უმეტესობა სიგნალის დიდი მნიშვნელობების პირობებში ფუნქციონირებს იმ არეში, სადაც აქტივაციის ფუნქციის წარმოებული ძალზე მცირე სიდიდეა. იმის გამო, რომ სწავლების პროცესში უკუმიმართულებით გაგზავნილი შეცდომის სიგნალი წარმოებულის პროპორციულია, სწავლების პროცესი შეიძლება გაჩერდეს. პრობლემის გადალახვას ცდილობენ სწავლების სიჩქარის კოეფიციენტის შემცირებით, რაც თავის მხრივ იწვევს პროცესის გახანგრძლივებას. ლოკალური მინიმუმის მახეები ემუქრება ყველა იმ ალგორითმს, რომელიც აგებულია მინიმუმის ძიების პრინციპზე. ამ პრობლემის დაძლევას ცდილობენ სწავლების სტოქასტიკური მეთოდებით - ახორციელებენ ქსელის სინაფსური წონების ფსევდოშემთხვევით ცვლილებებს, ამასთან ინარჩუნებენ ისეთ ცვლილებებს, რომლებიც ერთმანეთთან აახლოებს ფაქტობრივი გამომავალი სიგნალებისა და სასურველი გამომავალი სიგნალების მნიშვნელობებს. სტოქასტიკური მეთოდების გამოყენებისას შეიძლება მოინახოს ქსელის სინაფსური წონების კორექციის ისეთი სტრატეგია, რომელიც აიძულებს წონებს მიიღოს გლობალური მინიმუმის მნიშვნელობა, თუმცა ამ მეთოდების გამოყენებაც დაკავშირებულია დიდ სირთულეებთან. მაგალითად, ბოლცმანის სტოქასტიკური მეთოდით ხელოვნური ნეირონული ქსელების სწავლება ხშირად ითხოვს კომპიუტერული დროის მიუღებლად დიდ რესურსს, მასთან შედარებით კოშის მეთოდი ამცირებს ქსელის სწავლების დროს, თუმცა ეს მაჩვენებელი შეიძლება მაინც დიდი აღმოჩნდეს. სწავლების ხანგრძლივობის შემდგომი შემცირება შესაძლებელია ე. წ. ხელოვნური თბოტევადობის მეთოდით, რომელიც რთულია და პრაქტიკულად არ გამოიყენება. ზემოთ ჩამოთვლილი პროცედურების ეფექტურობა მნიშვნელოვნად არის დამოკიდებული ექსპერიმენტატორის გამოცდილებასა და სუბიექტურ მოსაზრებებზე. სხვადასხვა არქიტექტურის ნეირონული ქსელები და მათი სწავლების მეთოდები

განხილულია მრავალ ნაშრომში, მაგალითად [1],[8],[11], შექმნილია ნეირონული ქსელების სწავლების ალგორითმების ფართო სპექტრი, განსაზღვრულია მათი ზღვრული შესაძლებლობები, მაგრამ მიუხედავად ამისა ნეირონული ქსელების სწავლების პრობლემები თეორიულად ბოლომდე ჯერ არ არის შესწავლილი.

ხელოვნური ნეირონული ქსელი წარმოადგენს პარალელურ განაწილებულ გამოთვლით სისტემას, რომლის აგების პრინციპები დაკავშირებულია ნეიროფიზიოლოგიურ წარმოდგენებთან, ვინაიდან ხელოვნური ნეირონული ქსელის შემადგენელი ელემენტების - ფორმალური ნეირონების თვისებები მსგავსია ბიოლოგიური ნეირონების ზოგიერთი ელემენტარული ფუნქციისა, ხოლო ქსელის სტრუქტურული ორგანიზაცია ნაწილობრივ ემსგავსება ტვინის ანატომიურ აგებულებას. მიუხედავად ზედაპირული ანალოგიისა, ხელოვნურ ნეირონულ ქსელს აქვს ადამიანის ტვინის მსგავსი ზოგიერთი თვისება, კერძოდ, ის მაგალითებზე ”სწავლობს”, ახალი მაგალითების მიმართ ანზოგადებს წარსულ გამოცდილებას, შეუძლია შემოსული ჭარბი ინფორმაციიდან ამოკრიბოს არსებითი თვისებრივი ნიშნები, ამასთან ქსელს შეუძლია შეცვალოს თავისი რეაქცია გარემოს ცვლილებების ადეკვატურად. სწორედ ეს თვისებები აძლიერებს მისდამი ინტერესს. შემუშავებული და შესწავლილია ხელოვნური ნეირონებისა და ნეირონული ქსელების მრავალი განსხვავებული მოდელი, რომელიც მეტ-ნაკლებად ანალოგიურია მისი ბიოლოგიური პროტოტიპისა.

კოჰონენისა და სხვა მკვლევარების მიერ განვითარებულია სწავლების მეთოდები მასწავლებლის გარეშე. ეს მეთოდები არ საჭიროებს მიზნობრივ ვექტორებს, ანუ ქსელის გამომავალი ვექტორი არ უნდა შედარდეს იდეალურ პასუხს, სასწავლო ნაკრები შედგება მხოლოდ შემავალი ვექტორებისაგან-რეალიზაციებისაგან. სინაფსური წონების გადაწყობის

ალგორითმი მუშაობს ისე, რომ ქსელის გამოსასვლელზე მიღებული იქნეს შეთანხმებული გამომავალი ვექტორები, ანუ ქსელისათვის საკმარისად მსგავსი შემავალი ვექტორების წარდგენის შედეგად ქსელმა უნდა გამოიმუშაოს ერთნაირი გამომავალი სიგნალები. ამრიგად, სწავლება მასწავლებლის გარეშე გამოჰყოფს სასწავლო ნაკრების სტატისტიკურ თვისებებს და მსგავს ვექტორებს აჯგუფებს კლასებად.

ნეირონული ქსელების სწავლების თანამედროვე ალგორითმების უმეტესობა ეფუძნება დ. ჰების კონცეფციებს. დ. ჰებმა წამოაყენა მასწავლებლის გარეშე სწავლების ისეთი მოდელი, რომელშიც სინაფსური წონა იზრდება, თუ აგზნებულია ორივე ნეირონი – გადამცემი და მიმღები. ამრიგად, ქსელის ის სინაფსები და შესაბამისი გზები, რომლებიც ხშირად გამოიყენება, ძლიერდება. ამ მოდელით კარგად აიხსნება მიჩვევისა და სწავლების ფენომენი ბიოლოგიურ ორგანიზმებში.

ხელოვნურ ნეიროქსელში, სადაც გამოიყენება სწავლების ჰების მოდელი, სინაფსური წონების გაძლიერება განისაზღვრება გადამცემი და მიმღები ნეირონების აგზნების დონეების ნამრავლით:

$$w_{ij}(n+1)=w_{ij}(n)+\alpha\, OUT_i OUT_j, \qquad (2.1.)$$

სადაც $w_{ij}(n)$ არის გადაწყობამდე სინაფსური წონის სიდიდე i ნეირონიდან j ნეირონამდე; $w_{ij}(n+1)$ – გადაწყობის შემდეგ სინაფსური წონის სიდიდე i ნეირონიდან j ნეირონამდე; α - სწავლების სიჩქარის კოეფიციენტი; OUT_i - i ნეირონის გამოსასვლელი და j ნეირონის შესასვლელი; OUT_j - j ნეირონის გამოსასვლელი.

სწავლების ჰების მოდელი შედარებით მარტივია. ბოლო წლებში დამუშავებულია სწავლების უფრო ეფექტური ალგორითმები, რომლებიც ხასიათდება სწავლების უფრო მაღალი სიჩქარით და სასწავლო

ნაკრების მახასიათებლების შედარებით დიდი დიაპა-ზონით. ხელოვნური ნეირონული ქსელის კიდევ ერთი მნიშვნელოვანი თვისება არის ის, რომ მას აქვს განზოგადების (Generalization) უნარი, რაც ნიშნავს, რომ ნეირონული ქსელი იძლევა სწორ პასუხთან მიახლოებულ შედეგს მაშინაც, როდესაც წარედგინება ისეთი რეალიზაციები, რომლებიც არ შედიოდა მასწავლებელ სიმრავლეში. ნეიროქსელს რომ არ ჰქონდეს ასეთი თვისება, მაშინ ის იქნებოდა მხოლოდ დამახსოვრების მექანიზმი.

2.2. ნეირონული ქსელების სწავლების ალგორითმები

სწავლების ალგორითმები, რომლებიც ხელოვნური ნეირონული ქსელების უმნიშვნელოვანესი ელემენტია, წარმოადგენს კვლევების ძირითად ობიექტს. სწავლების ალგორითმი განსაზღვრავს სისტემის ისეთ მახასიათებლებს, როგორიც არის სწავლების დრო, ამოცანის გადაწყვეტის სიზუსტე, გამოთვლითი ოპერაციების რაოდენობა სწავლების თითოეულ ბიჯზე და სხვ. პირველ ხელოვნურ ამომცნობ სისტემად და თანამედროვე ნეიროქსელების პროტოტიპად ითვლება ამერიკელი მეცნიერის ფ. როზენბლატის მიერ 1958 წელს შექმნილი პერსეპტრონი. პერსეპტრონული ალგორითმი, როგორც ტვინის მექანიზმების გამარტივებული მოდელი (როგორც მიიჩნევდა ფ. როზენბლატი) მაშინ განხორციელდა აპარატურულადაც. იმ პერიოდის ნაშრომებში ძირითად მათემატიკურ აპარატად გამოიყენებოდა გადაწყვეტილებათა მიღების სტატისტიკური მეთოდები, აგრეთვე ზღურბლური ლოგიკის მეთოდები. შემდგომ პერიოდში, კომპიუტერების ევოლუციამ შესაძლებელი გახადა გადასვლა მოდელირების პროგრამულ მეთოდებზე. ამომცნობი სისტემების თეორიის განვითარებასთან ერთად შეიქმნა ამოცნობის პროცესის მოდელირების ახალი

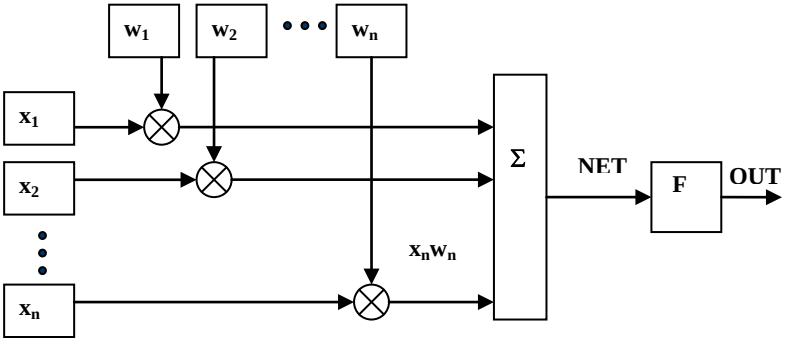
საშუალებები: სახეთა ანალიზისა და ამოცნობის სინტაქსური მეთოდი, არამკაფიო სიმრავლეების თეორია, ოპტიმალურობის თეორია, რანგული კავშირების თეორია, და ა.შ. სწავლების პროცესების ეფექტური განხორციელება მაღალსაიმედო ამომცნობი სისტემის აგების უმნიშვნელოვანესი ფაქტორია. ხელოვნური ინტელექტუალური სისტემების სწავლების პროცესები განხილულია მრავალ ნაშრომში, მათგან განსაკუთრებით მნიშვნელოვანია ი. ციპკინის და ვ. ვაპნიკის ნაშრომები, აგრეთვე მ. აიზერმანის, ე. ბრავერმანის, ლ. როზონოერის ნაშრომები, სადაც სწავლების ძირითადი მათემატიკური აპარატი სწავლების როგორც სტოქასტიკური, ასევე დეტერმინირებული პროცესებისათვის არის რეკურენტული პროცედურები. შედარებით ნაკლებად იყენებენ სწავლების არარეკურენტულ, როგორც წესი, ლოგიკურ მეთოდებს. ასეთ ნაშრომებს მიეკუთვნება მ. ბონგარდის, ო. ვერულავას და სხვ. ნაშრომები. ზოგადად ამომცნობ სისტემებში სწავლების პროცესები გამოიყენება სახეთა აღწერისა და გადაწყვეტილებათა მიმღები წესების ფორმირების დროს. სწავლების ისეთ ალგორითმს, სადაც გადამუშავებული ინფორმაციის ზრდასთან ერთად სწავლების მიზანი სულ უფრო უკეთ ხორციელდება და მიისწრაფის გარკვეული ცნობილი ზღვარისაკენ, ეწოდება კრებადი ალგორითმი (ან პროცესი). სწავლების დროს ხშირად გამოიყენება იტერაციული პროცედურები შესაბამისი მიზნობრივი ფუნქციებით. აგებული იტერაციული პროცედურების კრებადობის ხარისხის კვლევისათვის გამოიყენება სხვადასხვა მათემატიკური მეთოდები, კერძოდ ოპტიმიზაციის მეთოდების კარგად დამუშავებული აპარატი. იტერაციის ყოველ ბიჯზე ხდება მიზნობრივი ფუნქციის გამოთვლა და შედარება წინა ბიჯზე მიღებულ შედეგთან, შედეგის მიხედვით ხორციელდება შემდეგ ბიჯზე შესასრულებელი მოქმედებები. განასხვავებენ სწავლების დეტერმინირებულ (არაპარამეტრულ) და სტოქასტიკურ ანუ სტატისტიკურ (პარამეტრულ)

მეთოდებს. სწავლების დეტერმინირებულ პროცესებში გამოთვლებისას უშუალოდ გამოიყენება სასწავლო ნაკრების რეალიზაციები. სწავლების სტოქასტიკური პროცესებისათვის აპრიორულად უნდა იყოს ცნობილი რეალიზაციათა განაწილების კანონები. თუ სახეთა ნიშნები ერთმანეთზე არის დამოკიდებული, მაშინ საჭიროა ერთობლივი სტატისტიკური განაწილების კანონების დადგენა, რაც სასწავლო ნაკრების რეალიზაციათა ძალზე დიდ რაოდენობას მოითხოვს, ამიტომ მრავალგანზომილებიანი სახეებისათვის ამ მეთოდების გამოყენება შეზღუდულია. არასაკმარისი ინფორმაციის შემთხვევაში სწავლების პროცესებს განიხილავს ადაპტური სისტემების თეორია, სადაც შესწავლილია როგორც პარამეტრული, ასევე არაპარამეტრული სწავლების პროცესები. ადაპტურ სისტემებში ყოველი ახალი ინფორმაციის გამოჩენა წარმოადგენს იტერაციის ერთ ბიჯს, სადაც ხდება მიზნობრივი ფუნქციის ან მისი გრადიენტის მნიშვნელობის გამოთვლა და ამის საფუძველზე ამოცნობისათვის საჭირო ერთი ან რამდენიმე პარამეტრის ფორმირება. სწავლების პროცესი ჩაითვლება დამთავრებულად, თუ გამოთვლილ პარამეტრთა მნიშვნელობების ცვლილება არ სცილდება დასაშვებ ფარგლებს.

2.3. პერსეპტრონული ნეირონი და მისი სწავლება

ხელოვნური ნეირონული ქსელების სისტემატური შესწავლა პირველებმა დაიწყეს მაკ-კალოკმა და პიტსმა (1943წ). 2.1. სურათზე ნაჩვენებია მათ მიერ შემუშავებული ნეირონის მოდელი, სადაც ამჯამავი $\sum$ ელემენტი თითოეულ შემავალ x სიგნალს ამრავლებს w სინაფსურ წონაზე და აჯამებს შეწონილ ნამრავლებს. თუ ჯამი მეტია წინასწარ დადგენილი ზღურბლის მნიშვნელობაზე, მაშინ ნეირონის გამოსასვლელზე არის ლოგიკური

სიგნალი „ერთი”, თუ ნაკლებია ზღურბლის მნიშვნელობაზე, მაშინ ნეირონის გამოსასვლელზე იქნება სიგნალი „ნული”. ამ სისტემამ და მრავალმა მსგავსმა მიიღო პერსეპტრონის (perception - აღქმა) დასახელება. XX საუკუნის სამოციან წლებში პერსეპტრონმა დიდი ინტერესი და ოპტიმიზმი გამოიწვია, მაგრამ მალე



სურ. 2.1. მაკ-კალოკის და პიტსის ნეირონის მოდელი.

რომ პერსეპტრონს არ შეეძლო ზოგიერთი მარტივი ამოცანის გადაწყვეტა. მაგალითად, ერთშრიან პერსეპტრონს არ შეუძლია ისეთი მარტივი ლოგიკური ფუნქციის განხორციელება, როგორიც არის ლოგიკური არატოლმნიშვნელობა. მიუხედავად შეზღუდული შესაძლებლობისა, პერსეპტრონმა მნიშვნელოვანი როლი ითამაშა ნეიროკომპიუტინგის განვითარებაში.

იმისათვის, რომ ნეიროქსელს ჰქონდეს პრაქტიკული ღირებულება, უნდა ვიპოვოთ მისი სწავლების მეთოდის პოვნი, ანუ სინაფსური წონებისა და ზღურბლების საჭირო მნიშვნელობების გამოთვლისათვის აუცილებელია შესაბამისი ალგორითმი. ასეთი ალგორითმი შექმნა ფრენკ როზენბლატმა. პერსეპტრონის სწავლება შესაძლებელია მხოლოდ მასწავლებლით, რომელიც ნეიროქსელის მიმართ გარეშე ელემენტია. მასწავლებლის ფუნქციაა სისტემის ქცევის შეფასება და

მისი მოდიფიცირების მართვა, ანუ ნეიროქსელის პარამეტრების გადაწყობა.

2.2. სურათზე წარმოდგენილია პერსეპტრონის სწავლების ალგორითმი, სწავლება მიმდინარეობს შემდეგნაირად: სახეთა სიმრავლის რეალიზაციები რიგრიგობით მიეწოდება პერსეპტრონის შესასვლელს და სინაფსური წონების გადაწყობა გრძელდება სანამ პერსეპტრონის გამოსასვლელზე მიიღწევა სიგნალის სასურველი მნიშვნელობა ყველა რეალიზაციისათვის.

განვიხილოთ პერსეპტრონის სწავლების მაგალითი. ვთქვათ, მის შესასვლელს მიეწოდება სხვადასხვა ციფრების გამოსახულებები, ჩვენი მიზანია პერსეპტრონს ვასწავლოთ კენტი და ლუწი ციფრების გარჩევა, ანუ გამოსასვლელზე უნდა მივიღოთ "ერთი", თუ წარედგინება კენტი ციფრი და - „ნული", თუ წარედგინება ლუწი ციფრი. ვთქვათ, X შემავალი ვექტორი არის ციფრის რეალიზაცია, მაშინ ვექტორის თითოეული ელემენტი (x_1, x_2, …. x_n) უნდა გამრავლდეს W სინაფსური წონების ვექტორის შესაბამის ელემენტზე (w_1, w_2, …. w_n), ხოლო ნამრავლები აიჯამოს. ამის შემდეგ უნდა ჩატარდეს ზღურბლური ოპერაცია: თუ შეწონილი ჯამი აღემატება ნეირონის ზღურბლს, მაშინ გამოსასვლელზე იქნება სიგნალი "ერთი"(y = 1); თუ შეწონილი ჯამი ნაკლებია ნეირონის ზღურბლზე, მაშინ გამოსასვლელზე იქნება სიგნალი „ნული"(y = 0), ანუ პირველ შემთხვევაში პერსეპტრონი გვიჩვენებს, რომ ამოსაცნობად წარდგენილია კენტი ციფრი, ხოლო მეორე შემთხვევაში - წარდგენილია ლუწი ციფრი. პერსეპტრონის სწავლებისათვის ციფრის რეალიზაცია X წარედგინება შესასვლელს და გამოითვლება y გამოსასვლელი. თუ პასუხი სწორია სინაფსური წონები არ იცვლება; თუ პასუხი არასწორია, მაშინ იმ სინაფსური წონების, რომლებიც არასწორი შედეგის გამაძლიერებელ შესასვლელებს უკავშირდება, ისეთი მოდიფიცირება უნდა მოხდეს, რომ შეცდომა შემცირდეს.

2.4. უწყვეტი ნეირონული ქსელების დელტა-წესით სწავლების ალგორითმი

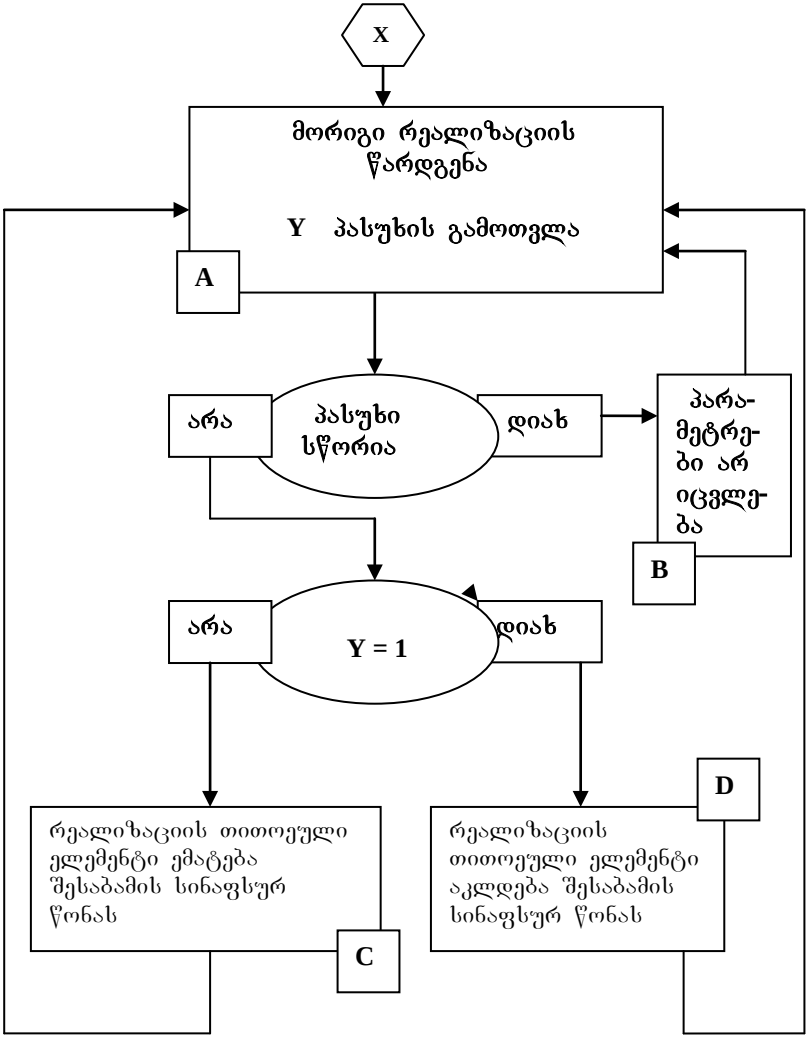
სწავლების აღწერილი ალგორითმი, რომელიც დისკრეტული ნეირონული ქსელებისათვის არის განკუთვნილი, შეიძლება განზოგადდეს უწყვეტი ნეირონული ქსელებისათვის. ამისათვის საჭიროა δ სიდიდის შემოტანა, რომელიც წარმოადგენს სხვაობას ქსელის გამოსასვლელზე სასურველი, ანუ მიზნობრივი (target) სიგნალისა და გამოსასვლელზე რეალურად მიღებულ Y სიგნალის სიდიდეებს შორის: δ = (T – Y). თუ δ=0, მაშინ T=Y, ანუ გამოსასვლელზე მიღებულია სწორი პასუხი. ეს შემთხვევა შეესაბამება 4.2. სურათზე წარმოდგენილი ალგორითმის B ბლოკს; თუ δ მეტია ნულზე, ეს შეესაბამება ალგორითმის C ბლოკს; ხოლო, თუ δ ნაკლებია ნულზე, ეს სიტუაცია შეესაბამება ალგორითმის D ბლოკს. ყველა ამ შემთხვევისათვის სწავლების აღწერილი ალგორითმი ძალაშია, თუ δ გამრავლდება შემავალი ვექტორის თითოეულ x_i ელემენტზე და ეს δx_i ნამრავლი აღემატება შესაბამის სინაფსურ წონას. განზოგადების მიზნით, უნდა შემოვიტანოთ სწავლების η სიჩქარის კოეფიციენტი, რომელიც უნდა გამრავლდეს δx_i სიდიდეზე, რაც მოგვცემს სინაფსური წონების საშუალო სიდიდის მართვის შესაძლებლობას. ალგებრულ ფორმაში გვექნება:

$$\Delta_i = \eta \delta x_i$$

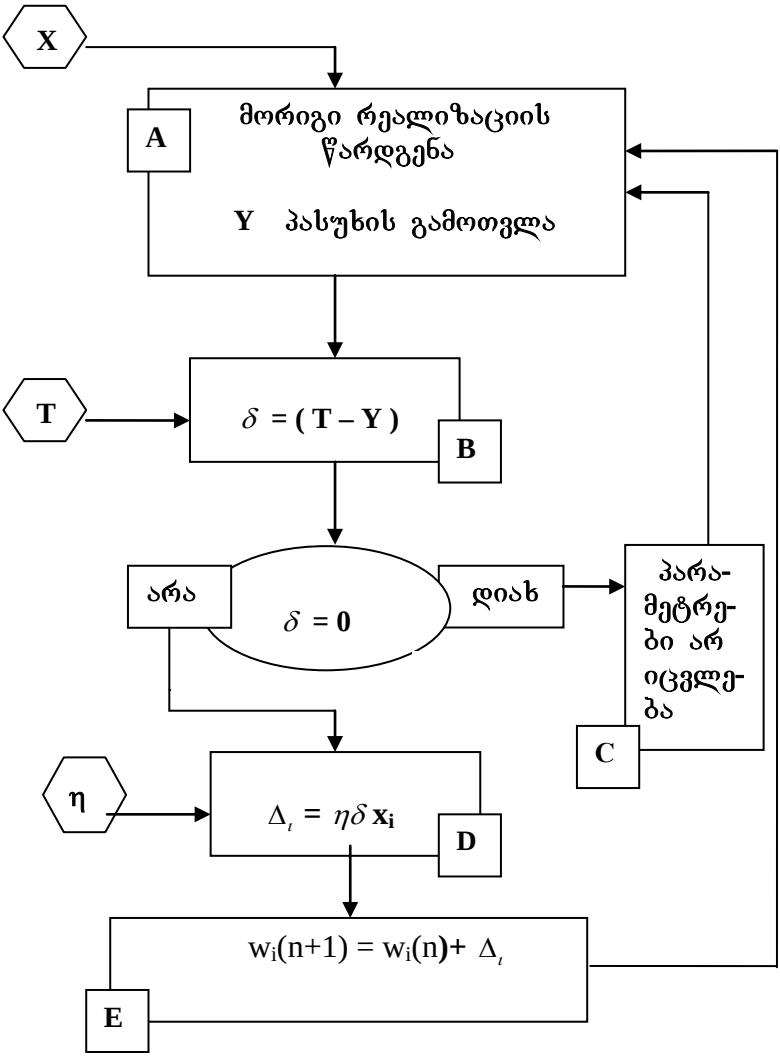
$$w_i(n+1) = w_i(n) + \Delta_i$$

სადაც Δ_i არის i-ური შესასვლელის კორექცია, ხოლო $w_i(n+1)$ – i-ური სინაფსური წონის მნიშვნელობა კორექ-ციის შემდეგ; $w_i(n)$ არის i-ური სინაფსური წონის მნიშვნელობა კორექციამდე. ნეირონული ქსელების დელტა-წესით სწავლების ალგორითმი წარმოდგენილია 2.3. სურათზე. გასათვალისწინებელია, რომ

პერსეპტრონის სწავლების ალგორითმის პრაქტიკული განხორციელების დროს შესაძლოა წავაწყდეთ შემდეგ სიძნელეებს: ზოგჯერ რთულია კონკრეტული სასწავლო ნაკრებისათვის იმის დადგენა სრულდება თუ არა წრფივად განმხოლოებადობის კრიტერიუმი; პრაქტიკულ ამოცანებში, როდესაც ქსელის შესასვლელზე სწრაფად იცვლება რეალიზაციები, შესაძლებელია კრიტერიუმი დროის ზოგ მომენტში შესრულდეს, ხოლო დროის სხვა მომენტში არა; შესაძლებელია დამტკიცდეს, რომ ქსელის სწავლება დამთავრდება სასრული რაოდენობის ბიჯების შემდეგ, თუმცა, ეს რაოდენობა შეიძლება აღმოჩნდეს ასტრონომიული სიდიდეების თანაზომადი და ცხადია, პრაქტიკისათვის მიუღებელი; არ არის დამტკიცებული, რომ ზემოთ განხილული სწავლების ალგორითმი უფრო სწრაფია, ვიდრე სინაფსური წონების გადაწყობა მათი ყველა შესაძლო მნიშვნელობის გადასინჯვის მეთოდით. ჩამოთვლილი პრობლემები აქტუალურია თანამედროვე ნეიროქსელური კვლევებისათვისაც.



სურ. 2.2. პერსეპტრონის სწავლების როზენბლატის ალგორითმი.



სურ. 2.3. ნეირონული ქსელების დელტა-წესით სწავლების ალგორითმი.

2.5. პირდაპირი გავრცელების მრავალშრიანი ნეირონული ქსელები

პირდაპირი გავრცელების მრავალშრიანი ნეირონული ქსელები, რომელსაც ზოგი ავტორი მრავალშრიან პერსეპტრონსაც უწოდებს, ხელოვნური ნეირონული ქსელების ნებისმიერ სხვა არქიტექტურასთან შედარებით ყველაზე უფრო ხშირად გამოიყენება და ნეიროკომპიუტინგის ძირითადი მუშა მოდელია. მრავალშრიან ნეირონულ ქსელს აქვს N შრე. პირველი შრის თითეული ნეირონი იღებს სიგნალებს ქსელის ყველა შესასვლელისაგან. მეორე შრის თითოეული ნეირონი იღებს სიგნალებს პირველი შრის ყველა აქსონისაგან, მესამე შრის თითოეული ნეირონი იღებს სიგნალებს მეორე შრის ყველა აქსონისაგან და ა.შ. გამომავალ შრემდე, რომლის აქსონები მიერთებულია ქსელის გამოსასვლელებთან. მრავალშრიანი ნეირონული ქსელის ფორმალური აღწერა მოხერხებულია მატრიცული ალგებრის ენაზე. ცხადია, რომ პირველი შრის ყველა ნეირონის წონითი კოეფიციენტი ($w_{11}, w_{12}, \ldots w_{1c}, \ldots w_{dc}$) შეიძლება წარმოვადგინოთ c x d განზომილების $\mathbf{W}^{(1)}$ მატრიცის ელემენტებად. მატრიცას აქვს c სტრიქონი და d სვეტი, სადაც c არის ქსელის შესასვლელების რიცხვი, ხოლო d – პირველი შრის ნეირონების რიცხვი. ანალოგიურად აიგება $\mathbf{W}^{(2)}$ მატრიცა, რომლის თითოეული ელემენტი $w_{ij}^{(2)}$ ახასიათებს პირველი შრის i -ური ნეირონის სინაფსური კავშირის წონას მეორე შრის j -ურ ნეირონთან. ამრიგად, $\mathbf{W}^{(n)}$ მატრიცის თითოეული ელემენტი $w_{ij}^{(n)}$ ახასიათებს (n-1) შრის i -ური ნეირონის სინაფსური კავშირის წონას n-ური შრის j -ურ ნეირონთან. მრავალშრიანი ნეირონული ქსელის მუშაობის პროცესი შეიძლება აღიწეროს ასე:

$$y^{(k)}=F(y^{(k-1)}W^{(k)}), \qquad k = 1\ldots N$$

სადაც F არის აქტივაციის არაწრფივი ფუნქცია (არაწრფივი გარდაქმნა); $y^{(0)}$ - ქსელის შემავალი ვექტორი; $y^{(N)}$ - ქსელის გამომავალი ვექტორი; $y^{(k)}$ - k - ური შრის გამომავალი სიგნალი; N – შრეების რაოდენობა. წრფივი აქტივაციის ფუნქციის მქონე ნეირონებისაგან შედგენილი მრავალშრიანი ქსელის გამოსასვლელი გამოითვლება ასე:

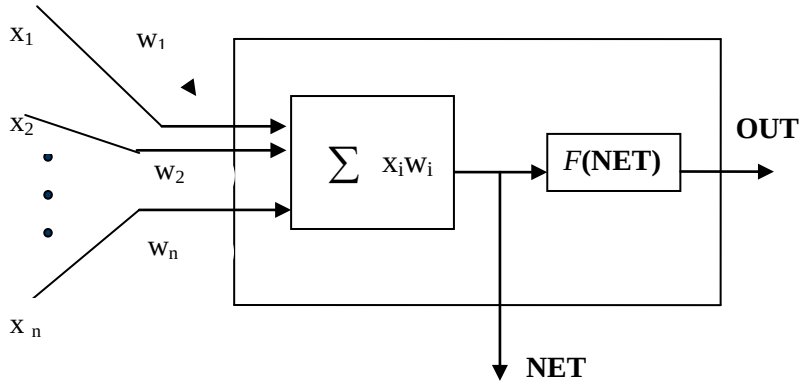
$y^{(N)}= y^{(0)}W^{(1)} W^{(2)}\ldots W^{(N)} = y^{(0)}W$ სადაც $W=W^{(1)} W^{(2)}\ldots W^{(N)}$.

აქედან გამომდინარეობს შედეგი დასკვნა: ნებისმიერი მრავალშრიანი ხელოვნური ნეირონული ქსელი, რომლის ნეირონებსაც აქვს წრფივი აქტივაციის ფუნქცია, შეიძლება შეიცვალოს მისი ეკვივალენტური ერთშრიანი ნეირონული ქსელით. ამრიგად ქსელის ფუნქციონირების ხარისხი მთლიანად არის დამოკიდებული ნეირონების აქტივაციის ფუნქციასა და სინაფსური კავშირების წონით კოეფიციენტებზე, ანუ პირდაპირი გავრცელების მრავალშრიანი ნეირონული ქსელების ფორმირება გულისხმობს, რომ ნეირონებს აქვს არაწრფივი აქტივაციის ფუნქცია, რომელიც განსაზღვრავს ასეთი ქსელების გაზრდილ შესაძლებლობებს. მრავალშრიანი ნეირონული ქსელის წარმოდგენა შეიძლება ”შავი ყუთის” სახით, რომელსაც აქვს N შესასვლელი და K გამოსასვლელი და რომელიც გარკვეული შიდა კანონებით გარდასახავს შემავალ ვექტორს გამომავალ ვექტორად. მრავალშრიანი ნეირონული ქსელი წარმოადგენს საფუძველს რიგი სხვა ქსელის აგებისას (მაგალითად, უკუგავრცელების ნეირონული ქსელი).

2.6. სწავლების უკუგავრცელების ალგორითმი

მრავალშრიან ხელოვნურ ნეირონულ ქსელებს, როგორც იყო აღნიშნული, გაცილებით მეტი გამოთვლითი შესაძლებლობები აქვს ერთშრიან

ნეირონულ ქსელებთან შედარებით, მაგრამ მრავალი წლის განმავლობაში არ არსებობდა მრავალშრიანი ნეირონული ქსელების თეორიულად დასაბუთებული სწავლების ალგორითმი. ამ გარემოებამ, დიდი ხნით შეაფერხა ნეიროინფორმატიკის განვითარება. უკუგავრცელების ალგორითმი წარმოადგენს ნეირონული ქსელების სწავლების მეთოდს, რომელსაც აქვს



სურ. 2.4. უკუგავრცელების ქსელებში გამოყენებული ფორმალური ნეირონი.

შესაბამისი მკაცრი მათემატიკური დასაბუთება. გარკვეული შეზღუდვების მიუხედავად, უკუგავრცელების პროცედურამ მნიშვნელოვნად გაზარდა ხელოვნური ნეირონული ქსელების გამოყენების არეალი.

2.4. სურათზე ნაჩვენებია ფორმალური ნეირონი, რომელიც გამოიყენება უკუგავრცელების ქსელების ასაგებად. რეცეპტორებისაგან ან წინამდებარე შრის ნეირონებისაგან ფორმალურ ნეირონს მიეწოდება სიგნალების სიმრავლე, რომლის თითოეული ელემენტი მრავლდება სინაფსურ წონაზე და შემდეგ ნამრავლები აიჯამება. ასეთი NET ჯამები უნდა იქნეს გამოთვლილი თითოეული ნეირონისათვის, რის შემდეგაც ხდება NET-

ის მოდიფიცირება F აქტივაციის ფუნქციით და საბოლოოდ მიიღება OUT სიგნალი:

$$NET = \sum x_i w_i \qquad OUT = F\,(NET). \qquad (2.2)$$

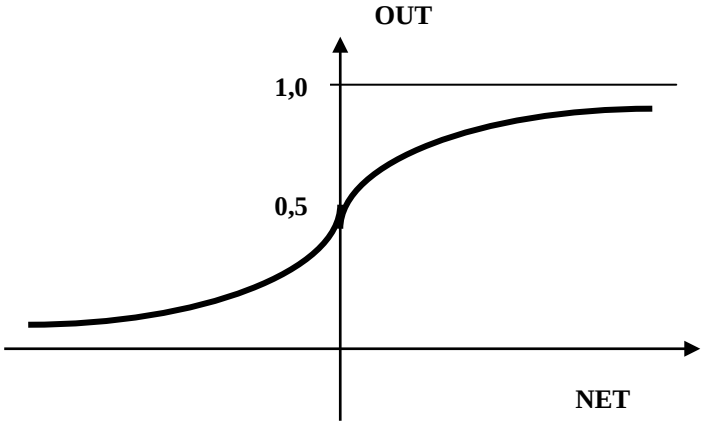
აქტივაციის ფუნქციად ყველაზე უფრო ხშირად უკუგავრცელების პროცედურაში გამოიყენება ლოგ-სიგმოიდური ფუნქცია

$$OUT = \frac{1}{1+e^{-NET}}. \qquad (2.3)$$

2.5. სურათზე ნაჩვენებია ეს შემკუმშავი ფუნქცია, რომელიც გამომავალი სიგნალის ცვლილებას ავიწროებს ისე, რომ OUT იცვლება ნულიდან ერთამდე. მრავალშრიან ნეირონულ ქსელებს ერთშრიან ქსელებთან შედარებით გაცილებით მეტი შესაძლებლობები აქვს მხოლოდ იმ შემთხვევაში, თუ ნეირონების აქტივაციის ფუნქცია იქნება არაწრფივი. ჩვენ შემთხვევაში, არაწრფივობას უზრუნველყოფს სიგმოიდი; იგი უზრუნველყოფს აგრეთვე გაძლიერების ავტომატურ კონტროლს – სიგნალის სიდიდის ზრდასთან ერთად მისი გაძლიერება მცირდება, ანუ ნეირონი ფუნქციონირებს სიგნალის ფართო დიაპაზონში. შედეგად ძლიერ სიგნალებს ქსელი აღიქვამს გაჯერების გარეშე, ხოლო სუსტი სიგნალები გაივლის ქსელს ზედმეტი შესუსტების გარეშე. სიგმოიდის გარდა, უკუგავრცე-ლების პროცედურაში შესაძლებელია სხვა არაწრფივი ფუნქციების გამოყენებაც, იმ პირობით, რომ ისინი, ისევე როგორც სიგმოიდი, დიფერენცირებადნი უნდა იყვნენ განსაზღვრის მთელს არეზე; თუმცა, სიგმოიდის გამოყენება განსაკუთრებით მოხერხებელია, ვინაიდან მისი წარმოებული, რომელიც აუცილებელია უკუგავრცელების პროცედურის შესასრულებლად, მარტივი სახისაა

$$F'(NET)=\frac{\partial OUT}{\partial NET}=OUT(1-OUT)\,. \qquad (2.4)$$

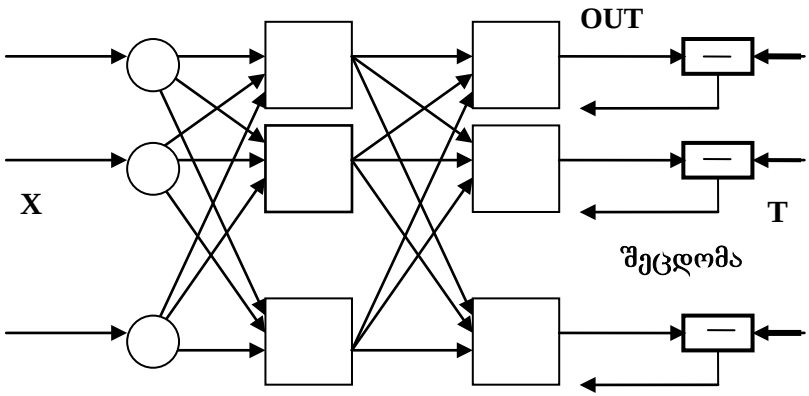
მრავალშრიანი ხელოვნური ნეირონული ქსელი, რომლის სწავლებაც შესაძლებელია უკუგავრცელების პროცედურით, ნაჩვენებია 2.6 სურათზე. ნეირონების



სურ. 2.5. სიგმოიდური აქტივაციის ფუნქცია.

პირველი შრე ასრულებს მხოლოდ შემავალი სიგნალების მიღებისა და განაწილების პროცედურას, ანუ შემავალი სიგნალები გაივლის პირველი შრის ნეირონებს როგორც გამანაწილებელ წერტილებს შესაბამისი სინაფსური წონებისაკენ რაიმე გარდაქმნის გარეშე, ხოლო მომდევნო შრის ნეირონები გამოითვლის NET და OUT სიგნალებს ისე, როგორც ეს აღწერილია ზემოთ. სწორედ ამიტომ სურათზე ნაჩვენები უკუგავრცელების ნეირონული ქსელი განიხილება როგორც ორშრიანი ქსელი, ხოლო გამანაწილებელ შრეს ეწოდება ნულოვანი. უკუგავრცელების პროცედურის გამოყენება შესაძლებელია ნებისმიერი რაოდენობის შრეების მქონე ხელოვნური ნეირონული ქსელებისათვის. ამასთან, ამ ალგორითმის გამოყენება შეიძლება როგორც

პირდაპირი გავრცელების, ასევე უკუკავშირებიანი ნეირონული ქსელების სწავლებისათვის. ნეირონული ქსელების სწავლების მიზანია მისი სინაფსური წონების ისეთი შეთანაწყობა, რომ შემავალი სიგნალების



სურ. 2.6. უკუგავრცელების ნეირონული ქსელი.

გარკვეულ სიმრავლეს (შემავალ **X** ვექტორს) შეესაბამებოდეს სასურველი გამომავალი სიგნალების სიმრავლე (მიზნობრივი **T** ვექტორი). სწავლების პროცესის განხორციელებისას ვთვლით, რომ ყოველი შემავალი ვექტორისათვის არსებობს მისი მეწყვილე მიზნობრივი ვექტორი, ორივე ვექტორს ერთად ეწოდება მასწავლებელი წყვილი. იმ შემთხვევაში, თუ ქსელმა უნდა ამოიცნოს სახეთა რომელიმე სიმრავლე, საჭირო გახდება იმდენივე მასწავლებელი წყვილის ფორმირება, რამდენი სახეც არის ამოსაცნობი. ასეთ ჯგუფს ეწოდება მასწავლებელი სიმრავლე. სწავლების დაწყების წინ თითოეულ სინაფსურ წონას უნდა მიენიჭოს რაიმე საწყისი მნიშვნელობა. პრაქტიკაში ცდილობენ, რომ

საწყის სინაფსურ წონებს მიენიჭოს მცირე შემთხვევითი რიცხვითი მნიშვნელობები, რათა სწავლების პროცესში ნეირონული ქსელი არ გადაიტვირთოს სინაფსური წონების დიდი მნიშვნელობებით და თავიდან იქნეს აცილებული ზოგიერთი პათოლოგიური შემთხვევა. უკუგავრცელების ალგორითმის გამოყენებით ნეირონული ქსელების სწავლებისათვის საჭიროა შემდეგი პროცედურების ჩატარება:

1. მორიგი მასწავლებელი წყვილის არჩევა და შემავალი ვექტორის წარდგენა ქსელისათვის;
2. ქსელის რეალური OUT გამოსასვლელის გამოთვლა;
3. ქსელის რეალურ OUT გამოსასვლელსა და მიზნობრივ (Target) გამოსასვლელს შორის სხვაობის გამოთვლა;
4. ქსელის სინაფსური წონების კორექტირება ისე, რომ შემცირდეს შეცდომა (სხვაობა);
5. ბიჯების (1, 2, 3, 4) განმეორება მასწავლებელი სიმრავლის თითოეული წყვილისათვის მანამ, სანამ შეცდომა არ გახდება მისაღები სიდიდის.

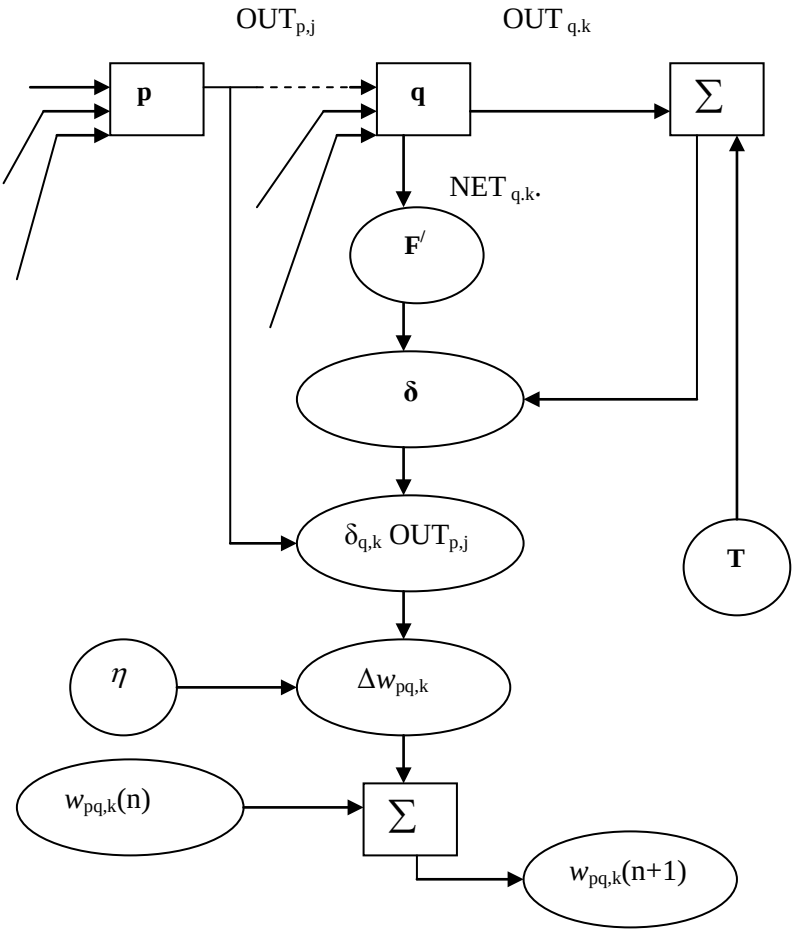
გამოთვლები თანმიმდევრობით ხორციელდება ნეირონული ქსელების შრეების მიხედვით, მაგალითად 2.6 სურათზე ნაჩვენები ქსელისათვის ჯერ უნდა გამოითვალოს j შრის გამოსასვლელები, რომლებიც k შრის შესასვლელებად გამოიყენება და მათი მეშვეობით გამოითვალოს k შრის გამოსასვლელები, რომლებიც ქმნის ქსელის გამოსასვლელ OUT ვექტორს. მესამე ბიჯზე მიზნობრივ (Target) გამოსასვლელს უნდა გამოაკლდეს ქსელის შესაბამისი რეალური OUT გამოსასვლელი, რაც მოგვცემს ქსელის შეცდომას (Error), რომელიც მეოთხე ბიჯზე გამოიყენება სინაფსური წონების კორექციისათვის. ამასთან, წონების ცვლილების ნიშანი და სიდიდე განისაზღვრება სწავლების ქვემოთ მოცემული ალგორითმის თანახმად. ამ ოთხი ბიჯის საკმარისი რაოდენობით განმეორების (იტერაციის)

შემდეგ, ქსელის რეალურ OUT გამოსასვლელსა და მიზნობრივ (Target) გამოსასვლელს შორის სხვაობა უნდა შემცირდეს მისაღებ სიდიდემდე, მაშინ ჩაითვლება, რომ ქსელის სწავლება დამთავრებულია. ამის შემდეგ სინაფსური წონების კორექცია აღარ ხდება და ქსელის გამოყენება შეიძლება პრაქტიკული ამოცანების გადასაწყვეტად. პირველი და მეორე ბიჯები განიხილება როგორც სიგნალის პირდაპირი გავრცელება, ხოლო მესამე და მეოთხე ბიჯები შეიძლება ჩავთვალოთ სიგნალის უკუგავრცელებად, ანუ ამ დროს გამოთვლილი შეცდომის სიგნალი ქსელში ვრცელდება ძირითადი სიგნალის გავრცელების საპირისპირო მიმართულებით და გამოიყენება სინაფსური წონების გადაწყობისათვის. სიგნალის პირდაპირი და უკუგავრცელება განვიხილოთ უფრო დაწვრილებით და ეს პროცედურები გამოვსახოთ მათემატიკური ფორმით.

პირდაპირი გავრცელების პროცედურა. ვექტორული ფორმით პირველი და მეორე ბიჯები შემდეგნაირად შეიძლება წარმოვადგინოთ: ქსელს წარედგინება **X** შემავალი ვექტორი და გამოსასვლელზე მიიღება **O** გამომავალი ვექტორი, ვექტორული წყვილი **X** და **T** უნდა ავიღოთ მასწავლებელი სიმრავლიდან. სინაფსური წონები განიხილება როგორც მატრიცა **W**, მაშინ ვექტორული ნოტაციით ქსელის შრისათვის გამოთვლითი პროცესი აღიწერება შემდეგი გამოსახულებით:

$$\mathbf{O} = F(\mathbf{XW}). \qquad (2.5)$$

მოცემული შრის გამომავალი ვექტორი იმავდროულად არის მომდევნო შრისათვის შემავალი ვექტორი, ამიტომ უკანასკნელი შრის გამოსასვლელი ვექტორის გამოსათვლელად აუცილებელია (2.5) პროცედურის განხორციელება თითოეული შრისათვის, დაწყებული ქსელის შესასვლელიდან.



სურ. 2.7. სინაფსური წონის გადაწყობა გამომავალ შრეში.

უკუგავრცელების პროცედურა. განვიხილოთ გამომავალი შრის სინაფსური წონების გადაწყობა. ვინაიდან გამომავალი შრის თითოეული ნეირონისათვის მოცემული გვაქვს მიზნობრივი მნიშვნელობა, ამიტომ სინაფსური წონების გადაწყობა ადვილად შეიძლება განვახორციელოთ მოდიფიცირებული დელტა-წესის მიხედვით, რაც განხილული იყო ზემოთ. შედარებით უფრო რთულია ფარული შრეების სწავლების პროცედურები, ვინაიდან შედარების ოპერაციის ჩასატარებლად მათთვის არა გვაქვს მიზნობრივი მნიშვნელობები. 2.7 სურათზე ნაჩვენებია ერთი სინაფსური წონის სწავლების პროცესი ***j*** ფარული შრის ***p*** ნეირონიდან ***k*** გამომავალი შრის ***q*** ნეირონამდე. შეცდომის სიდიდეს მივიღებთ, თუ მიზნობრივ მნიშვნელობას (Target) გამოვაკლებთ ***k*** გამომავალი შრის ნეირონის OUT სიგნალს. შემდეგ შეცდომის სიდიდე უნდა გამრავლდეს აქტივაციის ფუნქციის წარმოებულზე, რომელიც ***k*** გამომავალი შრის ამ ნეირონისათვის არის გამოთვლილი, რაც მოგვცემს **δ** სიდიდეს

$$\delta = \mathrm{OUT}(1 - \mathrm{OUT})(\mathrm{Target} - \mathrm{OUT}). \qquad (2.6)$$

ამის შემდეგ მრავლდება იმ **j** ნეირონის OUT სიდიდეზე, საიდანაც გამოდის ეს სინაფსური წონა. ნამრავლი თავის მხრივ უნდა გამრავლდეს **η** სწავლების სიჩქარის კოეფიციენტზე (ჩვეულებრივ, იგი 0,01-დან 1,0-მდე ფარგლებშია), ხოლო შედეგი უნდა დაემატოს სინაფსურ წონას. ასეთივე პროცედურები უნდა შესრულდეს თითოეული სინაფსური წონისათვის ფარული შრის ნეირონიდან გამომავალი შრის ნეირონამდე:

$$\Delta w_{pq,k} = \eta\, \delta_{q,k}\, \mathrm{OUT}_{p,j} \qquad (2.7)$$
$$w_{pq,k}(n+1) = w_{pq,k}(n) + \Delta w_{pq,k}, \qquad (2.8)$$

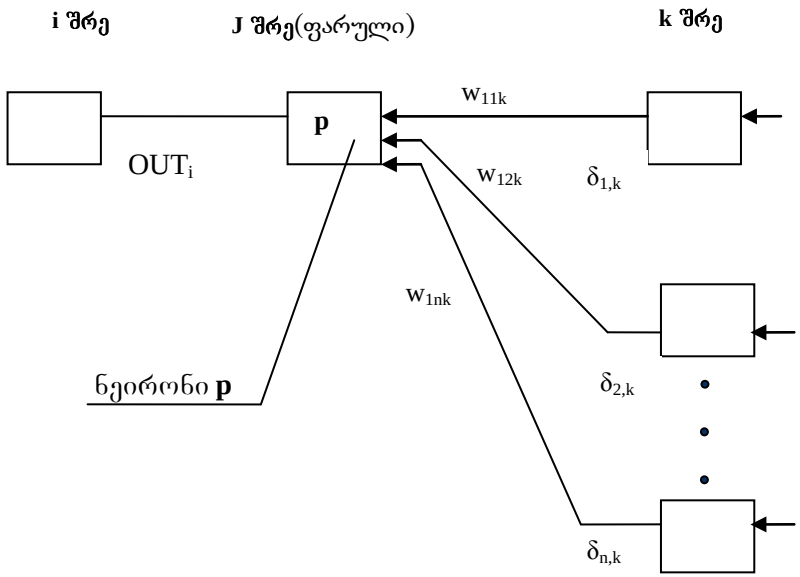
სადაც $w_{pq,k}(n)$ არის სინაფსური წონა ფარული შრის p ნეირონიდან გამომავალი შრის q ნეირონამდე n ბიჯზე

კორექციამდე. ამასთან, ინდექსი k ეკუთვნის იმ შრეს, სადაც მთავრდება მოცემული წონა; $w_{pq,k}(n+1)$ – სინაფსური წონა (n+1) ბიჯზე, ანუ კორექციის შემდეგ; $\delta_{q,k}$ არის δ სიდიდე გამომავალი k შრის q ნეირონისათვის; $OUT_{p,j}$ არის j ფარული შრის p ნეირონის OUT-ის მნიშვნელობა; η– სწავლების სიჩქარის კოეფიციენტი, რომელიც სინაფსური წონების ცვლილების საშუალო სიდიდის მართვის შესაძლებლობას იძლევა.

წონების გადაწყობა ფარულ შრეში. განვიხილოთ ერთი ნეირონი, რომელიც ეკუთვნის გამომავალი შრის წინამდებარე ფარულ შრეს. ეს ნეირონი პირდაპირი გავრცელების პროცესში თავის გამომავალ სიგნალს გადასცემს გამომავალი შრის ნეირონებს, მათი დამაკავშირებელი სინაფსური წონების საშუალებით. სწავლების პროცესში კი ეს სინაფსური წონები მუშაობს საპირისპირო მიმართულებით, ანუ სიდიდეს გაატარებს გამომავალი შრისაგან უკან, ფარული შრისაკენ, ამასთან, თითოეული ეს წონა უნდა გამრავლდეს გამომავალი შრის იმ ნეირონის სიდიდეზე, რომელსაც იგი უკავშირდება. ფარული შრის ნეირონის შეთანაწყობისათვის აუცილებელი სიდიდე მიიღება ყველა ასეთი ნამრავლის შეჯამებით და შედეგის გამრავლებით აქტივაციის ფუნქციის წარმოებულზე

$$\delta_{q,k} = OUT_{p,j}(1 - OUT_{p,j})\left[\sum_{q} \delta_{q,k}\, w_{pq,k}\right]. \qquad (2.9)$$

ფარულ შრეში სინაფსური წონის გადაწყობის პროცედურები ნაჩვენებია 2.8 სურათზე. მას შემდეგ, რაც δ სიდიდე მიღებულია, შეიძლება ჩატარდეს პირველი ფარული შრის კუთვნილი სინაფსური წონების გადაწყობის პროცედურები (2.7) და (2.8) განტოლებების მიხედვით, სადაც ინდექსები შრის შესაბამისად იქნება მოდიფიცირებული.



სურ. 2.8. სინაფსური წონის გადაწყობა ფარულ შრეში.

მოცემული ფარული შრის თითოეული ნეირონისათვის მიღებულ უნდა იქნეს δ სიდიდე და შესრულდეს ამ ფარულ შრესთან დაკავშირებული ყველა სინაფსური წონის გადაწყობა. ეს პროცესი მეორდება ყველა შრის მიმართ ნეიროქსელის შემოსასვლელის მიმართულებით, სანამ დამთავრდება ქსელის ყველა სინაფსური წონის გადაწყობა, ანუ ნეიროქსელის სწავლების მიზნით ხდება შეცდომის სიგნალის უკუგავრცელება, რაც აისახა ალგორითმის დასახელებაში.

2.7. უკუგავრცელების ალგორითმის დაჩქარების მეთოდები

ნეირონული ქსელების სწავლების უკუგავრცელების ალგორითმის მრავალი გაუმჯობესება და განზოგადება იქნა შემოთავაზებული მკვლევარების მიერ. განვიხილოთ

მათ შორის ყველაზე პერსპექტიული მიდგომები. უკუგავრცელების ალგორითმით სწავლების დაჩქარების ერთ-ერთ მეთოდს ეწოდება იმპულს-მეთოდი, სადაც სინაფსური წონის კორექციის გამოსახულებას ემატება წევრი, რომელიც სინაფსური წონის წინა ცვლილების პროპორციულია. როგორც კი კორექციის ოპერაცია ჩატარდება, მისი მნიშვნელობა უნდა იქნეს დამახსოვრებული და გამოყენებული ყველა შემდგომი კორექციის მოდიფიკაციისათვის. კორექციის განტოლებების მოდიფიკაცია ხდება შემდეგნაირად:

$$\Delta w_{pq,k}(n+1)= \eta\, \delta_{q,k}\, OUT_{p,j} + \alpha\, \Delta w_{pq,k}(n); \qquad (2.10)$$

$$w_{pq,k}(n+1) = w_{pq,k}(n) + \Delta w_{pq,k}(n+1), \qquad (2.11)$$

სადაც α არის იმპულსის კოეფიციენტი (ჩვეულებრივ, იღებენ 0,9 სიდიდეს). ლიტერატურაში აღნიშნულია, რომ იმპულს-მეთოდი კარგ შედეგს იძლევა მხოლოდ ზოგიერთი ამოცანის გადაწყვეტისას.

უკუგავრცელების ალგორითმის ერთი გაუმჯობესება ეფუძნება ექსპონენციალურ დაგლუვებას, რომელსაც ზოგიერთ შემთხვევაში შეიძლება ჰქონდეს უპირატესობა:

$$\Delta w_{pq,k}(n+1)= (1-\alpha)\, \delta_{q,k}\, OUT_{p,j} + \alpha\, \Delta w_{pq,k}(n). \qquad (2.12)$$

შემდეგ გამოითვლება სინაფსური წონის ცვლილება

$$w_{pq,k}(n+1) = w_{pq,k}(n) + \eta \Delta w_{pq,k}(n+1), \qquad (2.13)$$

სადაც α არის შესწორების კოეფიციენტი, რომლის ცვლილება ხდება 0,0-დან 1,0-მდე დიაპაზონში. თუ α უდრის 1,0, მაშინ ახალ კორექციას უკუვაგდებთ და მეორდება წინა კორექცია; ხოლო ნულსა და ერთს შორის არეში, სინაფსური წონის კორექცია მოხდება α-ს პროპორციული სიდიდით. სწავლების სიჩქარის კოეფიციენტი η სინაფსური წონების ცვლილების საშუალო სიდიდის მართვის შესაძლებლობას იძლევა.

უკუგავრცელების ალგორითმის კრებადობის დაჩქარების კიდევ ერთი მეთოდი არსებობს, რომელსაც ეწოდა მეორე რიგის უკუგავრცელება. იგი სინაფსური წონების საჭირო კორექციის უფრო ზუსტი შეფასებისათვის იყენებს მეორე რიგის წარმოებულებს. ლიტერატურაში ნაჩვენებია, რომ მეორე რიგის უკუგავრცელება ოპტიმალურია იმ გაგებით, რომ შეფასების გაუმჯობესება შეუძლებელია კიდევ უფრო მაღალი რიგის წარმოებულების გამოყენების გზით. ეს მეთოდი მოითხოვს დამატებით გამოთვლებს, ამიტომ ცალკე გამოკვლევაა საჭირო იმის დასამტკიცებლად, რომ გამართლებულია ასეთი მიდგომა.

ვ. სტორნეტას და ბ. ჰუბერმანს აღწერილი აქვთ უკუგავრცელების ქსელების სწავლების მახასიათებლების გაუმჯობესების საინტერესო მეთოდი, სადაც მითითებულია, რომ ფარული ნეირონების შესასვლელებისა და გამოსასვლელების ზოგადად მიღებული დინამიკური დიაპაზონი ნულიდან ერთამდე არ არის ოპტიმალური. ბინარული შემავალი ვექტორის შემთხვევაში საშუალოდ ნახევარი ამ ვექტორის ელემენტებისა იქნება ნული და შესაბამისად, ამ შესასვლელებთან დაკავშირებული სინაფსური წონების სწავლება არ მოხდება, ვინაიდან სინაფსური წონის კორექციის სიდიდე $\Delta w_{pq,k}$ პროპორციულია იმ ნეირონის გამომავალი სიგნალის დონისა, რომელიც $OUT_{p,j}$-ს გამოიმუშავებს. მაშინ ნულოვანი დონე იწვევს იმას, რომ სინაფსური წონა არ იცვლება. ნაპოვნია შემდეგი გამოსავალი: ბინარული მნიშვნელობები ქსელის შესასვლელებზე შეცვლილია +0,5 და –0,5 მნიშვნელობებით, ხოლო აქტივაციის ფუნქციას ემატება წანაცვლება –0,5; მაშინ აქტივაციის ფუნქცია იღებს შემდეგ სახეს:

$$OUT = -\frac{1}{2} + \frac{1}{1+e^{-NET}}. \qquad (2.14)$$

შედეგად უკუგავრცელების ალგორითმის კრებადობის დრო მცირდება საშუალოდ 30-დან 50 პროცენტამდე. ეს არის პრაქტიკული მოდიფიკაციის კარგი მაგალითი, როელმაც არსებითად გააუმჯობესა ალგორითმის მახასიათებლები. სწავლების უკუგავრცელების ალგორითმი შეიძლება სწავლების უკუგავრცელების ალგორითმი შეიძლება გამოყენებულ იქნეს აგრეთვე უკუკავშირებიანი ნეირონული ქსელებისათვის, რისთვისაც დამუშავებულია შესაბამისი მეთოდიკები.

სწავლების უკუგავრცელების ალგორითმი წარმატებით გამოიყენება მრავალფეროვანი პრაქტიკული ამოცანების გადაწყვეტის დროს. აქვე უნდა აღინიშნოს ისიც, რომ ზოგიერთი რთული ამოცანის გადაწყვეტისას სწავლების პროცესი შეიძლება ძალზე დიდხანს გაგრძელდეს, ან საერთოდ ვერ მოხდეს ნეიროქსელის სწავლება. ასეთ წარუმატებლობას ჩვეულებრივ ორი მიზეზი აქვს, ეგრეთ წოდებული ქსელის დამბლა ან მოხვედრა ლოკალურ მინიმუმში. სწავლების პროცესში კორექციების შედეგად სინაფსურმა წონებმა შეიძლება ძალზე დიდი მნიშვნელობები მიიღოს. ამის გამო, ნეირონების უმეტესობა OUT სიგნალის დიდი მნიშვნელობების პირობებში ფუნქციონირებს იმ არეში, სადაც აქტივაციის ფუნქციის წარმოებული ძალზე მცირე სიდიდეა. იმის გამო, რომ სწავლების პროცესში უკუმიმართულებით გაგზავნილი შეცდომის სიგნალი არის წარმოებულის პროპორციული, სწავლების პროცესი შეიძლება გაჩერდეს. ეს პრობლემა თეორიულად სუსტად არის შესწავლილი. მის გადალახვას ცდილობენ η ბიჯის შემცირებით, რაც, თავის მხრივ იწვევს სწავლების პროცესის გახანგრძლივებას, ამიტომ ამ პრობლემის შესწავლა გრძელდება.

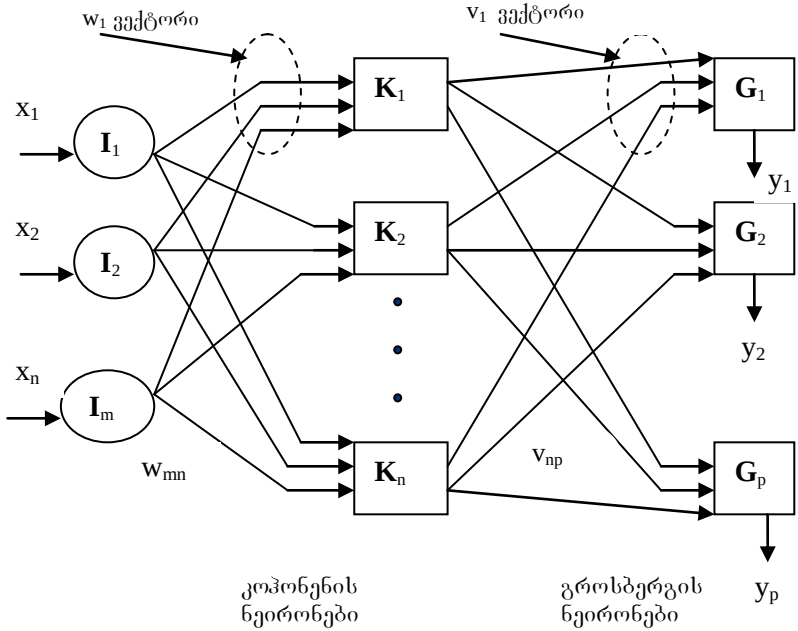
ქსელის ლოკალურ მინიმუმში მოხვედრის პრობლემა და მისი დაძლევის ხერხები ნაჩვენები იქნება სწავლების სტოქასტიკური მეთოდების განხილვის დროს.

ршага

2.8. შემხვედრი გავრცელების ქსელები

შემხვედრი გავრცელების ნეიროქსელების შესაძლებლობები გაცილებით აღემატება ერთშრიანი ნეიროქსელების შესაძლებლობებს, ხოლო უკუგავრცელების ნეიროქსელებთან შედარებით, სწავლებისათვის შეიძლება დასჭირდეს ასჯერ ნაკლები დრო, თუმცა შემხვედრი გავრცელების ნეიროქსელების გამოყენების არეალი შეზღუდულია, ვიდრე უკუგავრცელების ნეიროქსელების. შემხვედრი გავრცელების ნეიროქსელები კარგ შედეგს იძლევა ისეთი ამოცანების გადაწყვეტისას, სადაც შეუძლებელია სწავლების ხანგრძლივი პროცედურების ჩატარება. გარდა ამისა, ამ ქსელებს აქვს სხვა საინტერესო თვისებები, რასაც ქვემოთ განვიხილავთ. შემხვედრი გავრცელების ნეიროქსელებში გაერთიანებულია ორი ცნობილი ალგორითმი: კოჰონენის თვითორგანიზებადი რუკა და გროსბერგის ვარსკვლავი. მათი გაერთიანება გვაძლევს ისეთ თვისებებს, რომლებიც ცალ-ცალკე არც ერთს არა აქვს. შემხვედრი გავრცელების ნეიროქსელებს აქვს განზოგადების უნარი. სწავლების პროცესში შემავალი ვექტორები ასოცირდება შესაბამის გამომავალ ვექტორებთან. ვექტორები შეიძლება იყოს ბინარული ან უწყვეტი.

სწავლების პროცესის დამთავრების შემდეგ შემავალი ვექტორის წარდგენა გვაძლევს საჭირო გამომავალ ვექტორს. შემხვედრი გავრცელების ნეიროქსელების განზოგადების უნარი იძლევა იმის შესაძლებლობას, რომ მივიღებთ სწორ პასუხს მაშინაც კი, როდესაც ქსელს წარედგინება მცირედ დამახინჯებული ან არასრული შემავალი ვექტორი. ასეთი თვისებების გამო, შემხვედრი გავრცელების ნეიროქსელებს იყენებენ სიგნალების გასაძლიერებლად, სახეთა ამოცნობისათვის, აღდგენისათვის და სხვ.



სურ. 2.9. შემხვედრი გავრცელების ნეიროქსელის გამარტივებული სტრუქტურა.

ქსელის ფუნქციონირების განხილვის გასაადვილებლად 2.9 სურათზე ნაჩვენებია შემხვედრი გავრცელების პირდაპირი მოქმედების ნეიროქსელის გამარტივებული სტრუქტურა. სურათზე ნაჩვენები შემავალი შრის თითოეული ნეირონი უკავშირდება პირველი შრის (კოჰონენის შრის) თითოეულ ნეირონს ცალკეული w_{mn} სინაფსური წონით. ეს სინაფსური წონები განიხილება, როგორც **W** წონების მატრიცა. ანალოგიურად, კოჰონენის შრის თითოეული ნეირონი უკავშირდება მეორე შრის (გროსბერგის შრის) თითოეულ ნეირონს ცალკეული v_{np} სინაფსური წონით. ეს სინაფსური წონები შეადგენს **V** წონების მატრიცას. ყოველივე ეს

ზემოთ განხილული ნეიროქსელების მსგავსია, თუმცა განსხვავება მდგომარეობს იმ ოპერაციებში, რასაც ასრულებს კოჰონენის და გროსბერგის ნეირონები. შემხვედრი გავრცელების ნეიროქსელი, ისევე როგორც მრავალი სხვა ნეიროქსელი ფუნქციონირებს ორ რეჟიმში. პირველი რეჟიმი არის ნორმალური (მუშა) რეჟიმი, როდესაც ქსელს წარედგინება შემავალი ვექტორი **X** და მიიღება გამომავალი ვექტორი **Y;** მეორე არის სწავლების რეჟიმი, როდესაც ქსელს წარედგინება შემავალი ვექტორი **X** და ხდება სინაფსური წონების ისეთი კორექტირება, რომ მივიღოთ სასურველი **T** გამომავალი ვექტორი.

ნორმალური ფუნქციონირების რეჟიმი. შემხვედრი გავრცელების ნეიროქსელების პირველი შრე (კოჰონენის შრე) ფუნქციონირებს ისე, რომ მისი მხოლოდ და მხოლოდ ერთი ნეირონის გამოსასვლელზე შეიძლება იყოს ლოგიკური ერთი მოცემული შემავალი ვექტორისათვის, ყველა დანარჩენი ნეირონი იძლევა ლოგიკურ ნულს (ამბობენ, რომ შრეში არის ერთი გამარჯვებული ნეირონი). ქსელის თითოეულ შესასვლელს კოჰონენის ნეირონი უერთდება მასთან დაკავშირებული სინაფსური წონების საშუალებით. მაგალითად, 2.9 სურათზე ნაჩვენებია K_1 კოჰონენის ნეირონი, მისი სინაფსური წონები w_{11}, w_{21}, …, w_{m1} შეადგენს $\mathbf{W_1}$ წონების ვექტორს და ნულოვანი შრის მეშვეობით უკავშირდება **X** შემავალი ვექტორის შემადგენელ $x_1, x_2, \ldots, x_m$, სიგნალებს. კოჰონენის ნეირონის NET გამოსასვლელი გამოითვლება ჩვეულებრივად

$$\mathrm{NET_j} = w_{1j}x_1 + w_{2j}x_2 + \ldots + w_{mj}x_m, \qquad (2.15)$$

სადაც $\mathrm{NET_j}$ არის კოჰონენის j-ური ნეირონის NET გამოსასვლელი

$$\mathrm{NET_j} = \sum_i x_i w_{ij} \ . \qquad (2.16)$$

ვექტორული ფორმით იქნება

$$\mathbf{N} = \mathbf{XW}, \qquad (2.17)$$

სადაც **N** არის კოჰონენის შრის NET გამომავალი ვექტორი. კოჰონენის შრის ის ნეირონი არის გამარჯვებული, რომელსაც აქვს NET-ის მაქსიმალური მნიშვნელობა. გამარჯვებული ნეირონის გამოსასვლელზე იქნება ლოგიკური ერთიანი, ხოლო კოჰონენის შრის ყველა დანარჩენი ნეირონის გამოსასვლელზე იქნება ლოგიკური ნული.

გროსბერგის შრე ფუნქციონირებს მსგავსი წესით. მისი გამოსასვლელი NET არის კოჰონენის შრის $k_1, k_2, \ldots, k_n$ გამოსასვლელების, რომლებიც შეადგენს **K** ვექტორს, შეწონილი ჯამი. სინფსური წონები $v_{11}, v_{21}, \ldots, v_{np}$ შეადგენს **V** ვექტორს, მაშინ გროსბერგის ნეირონის NET გამოსასვლელი გამოითვლება შემდეგნაირად:

$$\mathrm{NET}_j = \sum_i k_i w_{ij}, \qquad (2.18)$$

სადაც NET_j არის გროსბერგის j-ური ნეირონის NET გამოსასვლელი. იგი ვექტორული ფორმით ასე ჩაიწერება:

$$\mathbf{Y} = \mathbf{KV}, \qquad (2.19)$$

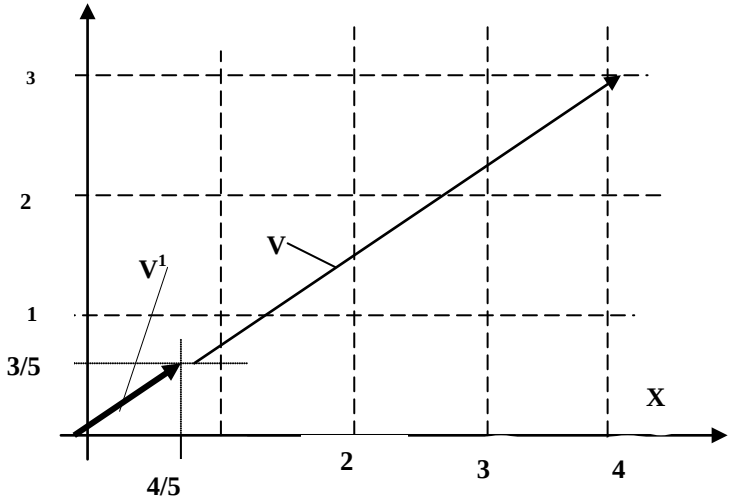
სადაც **Y** არის გროსბერგის შრის გამომავალი ვექტორი; **K** - კოჰონენის შრის გამომავალი ვექტორი; **V** – გროსბერგის შრის სინაფსური წონების მატრიცა. იმის გამო, რომ კოჰონენის შრე ფუნქციონირებს ისე, რომ მოცემული შემავალი ვექტორისათვის მხოლოდ ერთი ნეირონის გამოსასვლელზე შეიძლება იყოს ლოგიკური ერთიანი, ხოლო ყველა დანარჩენი ნეირონი იძლევა ლოგიკურ ნულს, ცხადია, **K** ვექტორის მხოლოდ ერთადერთი ელემენტია ერთის ტოლი და ამიტომ გამოთვლები ძალზე მარტივდება. ფაქტობრივად გროსბერგის შრის თითოეული ნეირონი გვაძლევს იმ სინაფსურ წონას, რომელიც აკავშირებს ამ ნეირონს კოჰონენის შრის არანულოვანი სიგნალის მომცემ ერთადერთ ნეირონთან.

კოჰონენის შრის სწავლება. კოჰონენის შრე შემხვედრი გავრცელების ნეიროქსელებში ახდენს შემავალი ვექტორების კლასიფიკაციას მსგავს ჯგუფებად, რისთვისაც სინაფსური წონების ისეთი გადაწყობა უნდა მოხდეს, რომ მსგავსმა შემავალმა ვექტორებმა კოჰონენის შრის ერთი და იმავე ნეირონის აქტივიზაცია გამოიწვიოს. ამის შემდეგ, გროსბერგის შრის ამოცანა არის სასურველი გამომავალი სიგნალების ფორმირება. კოჰონენის შრე არის თვითსწავლადი ანუ სწავლების მეთოდია მაწავლებლის გარეშე. ამიტომ წინასწარ არ არის ცნობილი, თუ რომელი ნეირონის აქტივიზაცია მოხდება, ამის აუცილებლობა არც არის. აუცილებელია მხოლოდ იმის უზრუნველყოფა, რომ სწავლების შედეგად ქსელმა შეძლოს მსგავსი ვექტორების ერთად დაჯგუფება, ანუ არამსგავსი ვექტორების განმხოლოება.

მიზანშეწონილია შემხვედრი გავრცელების ნეირო-ქსელებისათვის წარსადგენი ვექტორების წინასწარი ნორმალიზება. ამისათვის წარსადგენი ვექტორის თითოეული კომპონენტი უნდა გაიყოს ვექტორის სიგრძეზე. ვექტორის სიგრძე არის კვადრატული ფესვი ვექტორის კომპონენტების კვადრატების ჯამიდან, რომელიც ალგებრული ფორმით ასე გამოისახება:

$$x'_i = \frac{x_i}{\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}}. \qquad (2.20)$$

ამით წარსადგენი ვექტორი გარდაიქმნება იგივე მიმართულების ერთეულოვან ვექტორად, ანუ ერთეულოვანი სიგრძის ვექტორად n-გაზომილების სივრცეში.



სურ. 2.10. წარსადგენი ვექტორის ნორმალიზების მაგალითი.

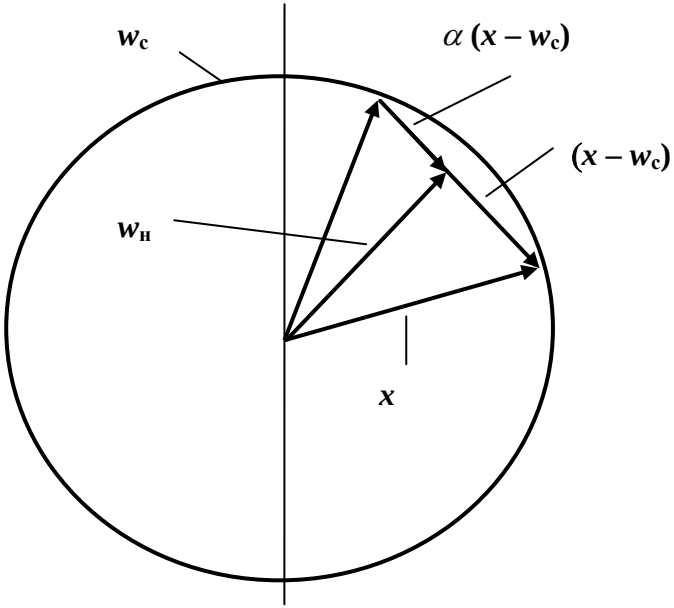
(2.20) განტოლება განაზოგადებს ორი განზომილების შემთხვევას, სადაც ვექტორის სიგრძე გამოითვლება პითაგორას თეორემით. 2.10 სურათზე ნაჩვენებია ორგანზომილებიანი **V** ვექტორი *x* - *y* კოორდინატებით, სადაც x = 4 და y = 5. ნორმალიზების შედეგად, (2.20) განტოლების თანახმად მივიღებთ ახალ **V**/ ვექტორს კოორდინატებით x = 4/5 და y = 3/5, რომელსაც ექნება იგივე მიმართულება, მაგრამ ერთეულოვანი სიგრძე. სამგანზომილებიანი ვექტორის შემთხვევაში, ის დამთავრდებოდა ერთეულოვანი სფეროს ზედაპირზე. ეს შეიძლება განვაზოგადოთ ნებისმიერი განზომილების ვექტორებზე, მაშინ ისინი შემოსაზღვრავდნენ ერთეულოვან ჰიპერსფეროს.

კოჰონენის შრის სწავლების დროს შესასვლელს წარედგინება შემავალი ვექტორი და გამოითვლება მისი სკალარული ნამრავლები იმ სინაფსურ წონებთან, რომლებიც უკავშირდება კოჰონენის შრის ყველა ნეირონს. სკალარული ნამრავლის მაქსიმალური

მნიშვნელობის მქონე ერთადერთი ნეირონი იქნება შრეში გამარჯვებული და მოხდება მისი წონების გადაწყობა. იმის გამო, რომ NET სიდიდეების გამოსათვლელი სკალარული ნამრავლი წარმოადგენს მსგავსების ზომას შემავალ ვექტორსა და წონების ვექტორს შორის, ამიტომ სწავლების პროცესი მდგომარეობს კოჰონენის შრის ისეთი ნეირონის შერჩევაში, რომლის წონების ვექტორი ყველაზე უფრო ახლოს არის შემავალ ვექტორთან და ამ წონების ვექტორის შემდგომ მიახლოებაში შემავალ ვექტორთან. მიმდინარეობს სწავლება მასწავლებლის გარეშე, ანუ ქსელის თვითორგანიზების ისეთი პროცესი, რომ კოჰონენის შრის მოცემული ნეირონის გამოსასვლელი იღებს მოცემული შემავალი ვექტორისათვის მაქსიმალურ მნიშვნელობას. სწავლების პროცესი აღიწერება შემდეგი განტოლებით:

$$w_{\text{н}} = w_{\text{c}} + \alpha\ (x - w_{\text{c}}), \qquad (2.21)$$

სადაც $w_{\text{н}}$ არის იმ წონის ახალი მნიშვნელობა, რომელიც აკავშირებს x შემავალ ელემენტს გამარჯვებულ ნეირონთან; w_{c} – აღნიშნული წონის მნიშვნელობა წინა ბიჯზე; α – სწავლების სიჩქარის კოეფიციენტი, რომლის ცვლა შესაძლებელია სწავლების პროცესში. თითოეული წონა, რომელიც უკავშირდება კოჰონენის შრის გამარჯვებულ ნეირონს, იცვლება მისი მნიშვნელობისა და მასთან დაკავშირებული შესასვლელის მნიშვნელობას შორის სხვაობის $(x - w_{\text{c}})$ პროპორციულად. ცვლილების მიმართულება ისეთია, რომ მცირდება ეს სხვაობა. 2.11 სურათზე მოცემულია ამ პროცესის ილუსტრაცია ორგანზომილებიანი ვექტორებისათვის. სურათზე $\mathbf{W}_{\text{н}}$ არის ახალი სინაფსური წონის ვექტორი, ხოლო $\mathbf{W}_{\text{c}}$ – ძველი სინაფსური წონის ვექტორი. ჯერ ვპოულობთ $\mathbf{X} - \mathbf{W}_{\text{c}}$ ვექტორს, რისთვისაც გავავლებთ მონაკვეთს ძველი სინაფსური წონის ვექტორის ბოლოდან $\mathbf{X}$ ვექტორის ბოლოსკენ; მიღებულ ვექტორს დავამოკლებთ ერთზე



სურ. 2.11. წონის ვექტორის მოძრაობა სწავლების პროცესში.

ნაკლები სიდიდის α სწავლების სიჩქარის კოეფიციენტზე გამრავლებით – შედეგად ვიღებთ ცვლილების δ ვექტორს. მაშინ ახალი სინაფსური წონის ვექტორი $\mathbf{W}_н$ იქნება მონაკვეთი მიმართული კოორდინატთა სათავიდან δ ვექტორის ბოლოსაკენ. ამ ილუსტრაციიდან ნათლად ჩანს, რომ სწავლების პროცესის ეფექტი შეიძლება წარმოვიდგინოთ როგორც სინაფსური წონის ვექტორის მოძრაობა მისი სიგრძის უმნიშვნელო ცვლილებით შემავალი ვექტორის მიმართულებით. α სწავლების სიჩქარის კოეფიციენტს სწავლების დასაწყისში ჩვეულებრივ იღებენ 0,7 სიდიდეს, რომელიც სწავლების პროცესში შეიძლება თანდათან შემცირდეს, ანუ პროცესის დასაწყისში კეთდება მოზრდილი ნაბიჯები, რომლებიც საბოლოო სიდიდესთან მიახლოებისას თანდათან მცირდება.

სინაფსური წონების საწყისი მნიშვნელობების შერჩევა. შემხვედრი გავრცელების ქსელის ყველა სინაფსურ წონას სწავლების პროცესის დაწყების წინ უნდა მიენიჭოს გარკვეული საწყისი მნიშვნელობა. ნეირონული ქსელების სწავლების პრაქტიკაში მიღებულია სინაფსური წონების საწყის მნიშვნელობებად მცირე შემთხვევითი სიდიდეების გამოყენება. კოჰონენის შრის სწავლების დაწყებამდე საჭიროა შერჩეული წონითი ვექტორების ნორმალიზება. სწავლების პროცესის დამთავრების შემდეგ წონითი ვექტორების საბოლოო მნიშვნელობები თანხვედრილი იქნება შემავალ ნორმალიზებულ ვექტორებთან. ამიტომ ნორმალიზება სწავლების დაწყებამდე ახდენს წონითი ვექტორების მიახლოებას მათ საბოლოო მნიშვნელობებეთან და შედეგად ამოკლებს სწავლების პროცესს. კოჰონენის შრის სწავლებისათვის სინაფსური წონების საწყის მნიშვნელობებად შემთხვევითი სიდიდეების გამოყენებამ (რანდომიზაციამ) შეიძლება სწავლებისას წარმოიქმნას პრობლემები, რადგანაც წონითი ვექტორები ჰიპერსფეროს ზედაპირზე ლაგდება თანაბრად, ხოლო შემავალი ვექტორები, როგორც წესი, განაწილებულია არათანაბრად და ჯგუფდება ჰიპერსფეროს ზედაპირის შედარებით პატარა უბნებზე. შედეგად, წონითი ვექტორების დიდი ნაწილი იმდენად იქნება დაშორებული ნებისმიერი შემავალი ვექტორიდან, რომ ისინი ვერ მოგვცემენ საუკეთესო შესაბამისობას. კოჰონენის შრის ეს ნეირონები უსარგებლონი აღმოჩნდებიან, ვინაიდან მათი გამოსასვლელები ყოველთვის ნულის ტოლი იქნება. შეიძლება წარმოიქმნას უარესი პრობლემაც: დარჩენილი წონითი ვექტორების რაოდენობა არასაკმარისი იყოს იმისათვის, რომ კლასებად იქნეს დაყოფილი ის შემავალი ვექტორები, რომლებიც ერთანეთთან ახლოს არის განლაგებული ჰიპერსფეროს ზედაპირზე.

დამუშავებულია ამ პრობლემებთან ბრძოლის რამდენიმე მეტ-ნაკლებად ეფექტური მეთოდი. ერთ-ერთი მათგანია ამოზნექილი კომბინაციის მეთოდი (convex

combination method), რომელიც წონებს ანიჭებს ერთნაირ საწყის მნიშვნელობებს

$$w_i = \frac{1}{\sqrt{n}}, \qquad (2.22)$$

სადაც n არის თითოეული წონითი ვექტორის ელემენტების რაოდენობა. ამის გამო ყველა წონითი ვექტორი თანხვედრილია და აქვს ერთეულოვანი სიგრძე. შემავალი ვექტორის თითოეულ ელემენტს მიენიჭება შემდეგი მნიშვნელობა:

$$x_i = \alpha x_i + \frac{1-\alpha}{\sqrt{n}}, \qquad (2.23)$$

სადაც n არის შესასვლელების რაოდენობა. თავდაპირველად α ძალზე მცირე სიდიდეა, რის გამოც შემავალ ვექტორებს აქვს $1/\sqrt{n}$-თან მიახლოებული სიგრძე და თითქმის ემთხვევა წონით ვექტორებს. ქსელის სწავლების პროცესში α თანდათანობით იზრდება და უახლოვდება ერთს. ეს იძლევა შემავალი ვექტორების განმხოლოების შესაძლებლობას და საბოლოოდ მათ ენიჭება ჭეშმარიტი მნიშვნელობები. ამოზნექილი კომბინაციის მეთოდი კარგია, მაგრამ ანელებს სწავლების პროცესს.

შემხვედრი გავრცელების ქსელებში სინაფსური წონებისათვის სწავლების პროცესის დაწყების წინ საწყისი მნიშვნელობების მინიჭების პრობლემა დამატებით შესწავლას მოითხოვს.

კოჰონენის სწავლების მეთოდს აქვს შემავალი მონაცემების სიმრავლიდან სტატისტიკური თვისებების გამოყოფის სასარგებლო უნარი.

გროსბერგის შრის სწავლება. გროსბერგის შრის სწავლება შედარებით მარტივია. სწავლების მეთოდია მასწავლებლით. გროსბერგის შრისათვის შემავალი ვექტორი არის კოჰონენის შრის გამოსასვლელი,

რომელიც წარედგინება გროსბერგის შრის ნეირონებს და მათი გამოსასვლელები გამოითვლება ისევე, როგორც ნორმალური ფუნქციონირების რეჟიმში. ამის შემდეგ, თითოეული წონის კორექტირება ხდება მხოლოდ იმ შემთხვევაში, თუ ის უკავშირდება არანულოვანი გამოსასვლელის მქონე კოჰონენის შრის ნეირონს. წონის კორექციის სიდიდე პროპორციულია წონასა და გროსბერგის შრის იმ ნეირონის გამოსასვლელს შორის სხვაობისა, რომელსაც იგი უკავშირდება:

$$v_{ijн} = v_{ijc} + \beta\ (y_j - v_{ijc})k_i, \qquad (2.24)$$

სადაც k_i არის i–ური კოჰონენის ნეირონის გამოსასვლელი (ნულისაგან განსხვავდება მხოლოდ ერთი კოჰონენის ნეირონისათვის); y_j არის სასურველი გამოსასვლელი ვექტორის j–ური ელემენტი. თავდაპირველად β-ს მნიშვნელობას იღებენ ~0,1-ის ტოლს და სწავლების პროცესში იგი თანდათან მცირდება.

შემხვედრი გავრცელების ნეირონული ქსელების შემქმნელი რ. ჰეხტ-ნილსენი თვითონ აღნიშნავს, რომ ამ ქსელებს აქვს შეზღუდული შესაძლებლობები უკუგავრცელების ქსელებთან შედარებით, თუმცა გარკვეული ამოცანების გადაწყვეტისას კარგ სტატისტიკურ მოდელებს იძლევა. შემხვედრი გავრცელების ქსელის სწავლება შედარებით სწრაფად ხდება და სწორად გამოყენებისას იგი ზოგავს მნიშვნელოვან მანქანურ დროს. ეს ქსელები სასარგებლოა განსაკუთრებით მაშინ, როდესაც საჭიროა სისტემების სწრაფი მოდელირება და ნაკლებ მნიშვნელოვანია დიდი სიზუსტე.

თავი III. ხელოვნური ნეირონული ქსელების სწავლების სტოქასტიკური მეთოდები

სტოქასტიკური მეთოდების გამოყენება კარგ შედეგებს იძლევა როგორც ხელოვნური ნეირონული ქსელების სწავლების, ასევე "ნასწავლი" ქსელებისაგან სასურველი გამოსასვლელი სიგნალების მიღების, ანუ ნორმალური ფუნქციონირების რეჟიმის დროს. სტოქასტიკური მეთოდები შესაძლებლობას იძლევა სწავლების პროცესში დავძლიოთ ლოკალური მინიმუმების პრობლემა, თუმცა ამ მეთოდების გამოყენება დაკავშირებულია ზოგიერთ სირთულეებთან.

3.1. სწავლების ლოკალური მინიმუმების პრობლემა

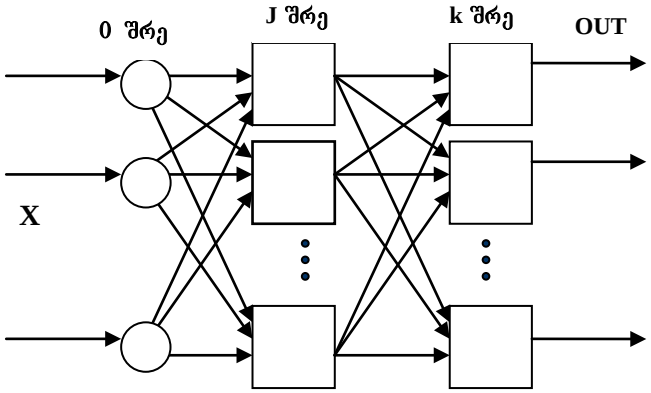
ხელოვნური ნეირონული ქსელების სწავლება წარმოადგენს მისი პარამეტრების (როგორც წესი, სინაფსური წონების) მოდიფიცირების გარკვეულ პროცესს. იმ შემთხვევაში, თუ სწავლების პროცესი წარმატებულია, მაშინ ქსელისათვის შემავალი სიგნალების სიმრავლის წარდგენის შედეგად ქსელის გამოსასვლელზე უნდა გამომუშავდეს სასურველი სიგნალების სიმრავლე. ცნობილია, რომ არსებობს სწავლების მეთოდების ორი ჯგუფი: დეტერმინირებული და სტოქასტიკური. სწავლების დეტერმინირებული მეთოდები სინაფსური წონების მოდიფიცირების პროცედურას ახორციელებს ბიჯურად, ამასთან იყენებს წონების მიმდინარე მნიშვნელობებს, აგრეთვე შემავალი სიგნალების, სასურველი გამომავალი სიგნალების და ფაქტობრივი გამომავალი სიგნალების მნიშვნელობებს. ასეთი დეტერმინირებული მიდგომის მაგალითი არის პერსეპტრონის სწავლების პროცესი. სწავლების სტოქასტიკურ მეთოდებში ახდენენ სინაფსური წონების ფსევდოშემთხვევით ცვლილებებს; ამასთან, ინარჩუნებენ ისეთ ცვლილებებს, რომლებიც ერთმანეთთან აახლოებს ფაქტობრივი გამომავალი სიგნალებისა და სასურველი

გამომავალი სიგნალების მნიშვნელობებს. იმისათვის, რომ ვნახოთ, თუ როგორ ხორციელდება ეს პროცედურები, განვიხილოთ ტიპური ორშრიანი ხელოვნური ნეირონული ქსელი, რომელიც 3.1 სურათზე არის ნაჩვენები. იგულისხმება, რომ ქსელის ნეირონების აქტივაციის ფუნქცია არის არაწრფივი. ნეირონული ქსელის სწავლებისათვის შეიძლება გამოყენებულ იქნეს შემდეგი პროცედურა:

1. შევარჩიოთ სინაფსური წონის შემთხვევითი სიდიდე და მოვახდინოთ მისი კორექცია მცირე შემთხვევითი სიდიდით. წარვუდგინოთ ქსელს შემავალი სიგნალების სიმრავლე და გამოვითვალოთ გამომავალი სიგნალების ფაქტობრივი მნიშვნელობები.
2. შევადაროთ გამომავალი სიგნალების ფაქტობრივი მნიშვნელობები სასურველი სიგნალების მნიშვნელობებს და გამოვთვალოთ მათ შორის სხვაობა. ეს პროცედურა უნდა შესრულდეს მასწავლებელი წყვილის თითოეული ელემენტისათვის, შემდეგ ელემენტების სხვაობები აიყვანება კვადრატში და გამოითვლება მათი ჯამი. სწავლების მიზანი არის ამ სხვაობის (მიზნობრივი ფუნქციის) მინიმიზაცია.
3. შევარჩიოთ სინაფსური წონის შემთხვევითი სიდიდე და მოვახდინოთ მისი კორექცია მცირე შემთხვევითი სიდიდით. თუ კორექცია სასარგებლოა, ანუ ამცირებს მიზნობრივ ფუნქციას, იგი უნდა იქნეს დამახსოვრებული, წინააღმდეგ შემთხვევაში უნდა დავუბრუნდეთ სინაფსური წონის ძველ მნიშვნელობას.
4. გავიმეოროთ 1, 2, 3 ბიჯები სანამ მიიღწევა მიზნობრივი ფუნქციის დასაშვები დაბალი დონე, ანუ ჩაითვლება, რომ ნეირონული ქსელის სწავლების პროცესი დასრულებულია.

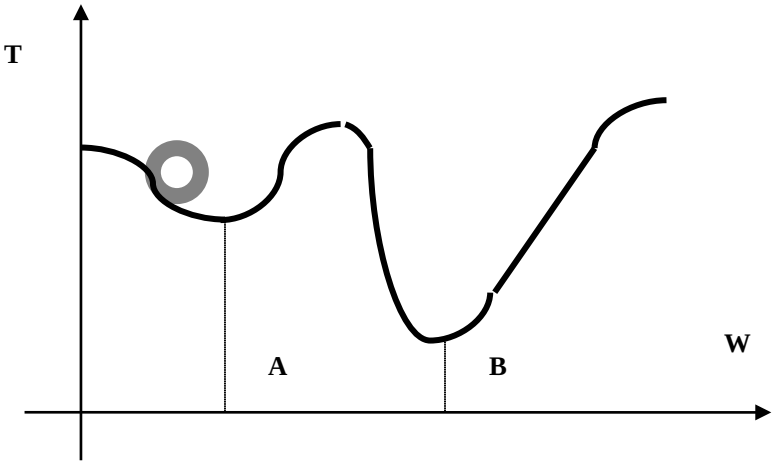
აღწერილი პროცესი მიისწრაფვის მიზნობრივი ფუნქციის მინიმიზაციისაკენ, მაგრამ შეიძლება მოხვდეს ლოკალურ მინიმუმში, როგორც მახეში, რადგანაც ქსელი ჩათვლის, რომ გადაწყვეტილება

მიღებულია, თუმცა შედეგი ფაქტობრივად წარუმატებელია.



სურ. 3.1. ტიპური ორშრიანი ხელოვნური ნეირონული ქსელი.

ლოკალური მინიმუმების პრობლემის გრფიკული ილუსტრაცია ერთი სინაფსური წონის კორექციის შემთხვევისათვის ნაჩვენებია 3.2 სურათზე.



სურ. 3.2. ლოკალური მინიმუმების პრობლემა.

ვთქვათ, **A** წერტილი შეესაბამება სინაფსური წონის საწყის მნიშვნელობას, თუ წონის ცვლილებები მცირე სიდიდის იქნება, ნებისმიერი გადახრა **A** წერტილიდან გაზრდის **T** მიზნობრივი ფუნქციის მნიშვნელობას და ამიტომ უარყოფილი იქნება, ხოლო სინაფსური წონის საუკეთესო მნიშვნელობა, რომელსაც **B** წერტილი შეესაბამება, ვერასოდეს იქნება მიღწეული. სისტემა **A** წერტილის ლოკალური მინიმუმის მახეში დარჩება. იმ შემთხვევაში, თუ წონის შემთხვევითი ცვლილებები დიდი სიდიდის იქნება, წონა განიცდის იმდენად მკვეთრ ცვლილებებს, რომ სასურველ (გლობალურ) მინიმუმში მისი მოხვედრა პრაქტიკულად შეუძლებელია. ამგვარი პრობლემების თავიდან აცილების სასარგებლო სტრატეგია მდგომარეობს ბიჯის ზომის თანდათანობით შემცირებაში დიდი საწყისი ბიჯიდან საშუალო შემთხვევით ბიჯის ზომამდე. ეს იძლევა ლოკალური მინიმუმის მახიდან თავის დაღწევისა და ქსელის საბოლოო სტაბილიზების შესაძლებლობას. ლოკალური მინიმუმის მახეები ემუქრება ყველა იმ ალგორითმს, რომლებიც აგებულია მინიმუმის ძიების პრინციპზე. სწავლების სტოქასტიკური მეთოდები კი ამ პრობლემის დაძლევის შესაძლებლობას იძლევა, ვინაიდან წონების კორექციის ისეთი სტრატეგია შეიძლება მოინახოს, რომელიც აიძულებს წონებს მიიღონ გლობალური მინიმუმის (B წერტილი) მნიშვნელობა. მიღებულია, რომ გლობალური მინიმუმის ძიება წააგავს ფოლადის წრთობის პროცესს, ამიტომ ხშირად იყენებენ ტერმინს – წრთობის იმიტაცია. გამდნარ ლითონში ატომებს ახასიათებს ძლიერი უწესრიგო მოძრაობა, ისინი ისწრაფვიან ენერგიის მინიმუმის მდგომარეობისაკენ (ამ შემთხვევაში ერთიანი კრისტალისაკენ), მაგრამ ამას მაღალ ტემპერატურაზე ეწინააღმდეგება ატომების მოძრაობის ენერგია. ლითონის

თანდათანობითი გაცივების პროცესში წარმოიშობა სულ უფრო დაბალენერგეტიკული მდგომარეობები, ვიდრე არ იქნება მიღწეული უმცირესი ენერგეტიკული მდგომარეობა, ანუ გლობალური მინიმუმის მნიშვნელობა. ენერგეტიკული მდგომარეობების P(e) განაწილებას აღწერს შემდეგი გამოსახულება:

$$P(e) = \exp(-e/kT), \qquad (3.1)$$

სადაც P(e) არის იმის ალბათობა, რომ სისტემა იმყოფება e ენერგიის მდგომარეობაში; k - ბოლცმანის მუდმივა; T – ტემპერატურა კელვინის სკალით. P(e) მაღალ ტემპერატურაზე უახლოვდება ერთის სიდიდეს ყველა ენერგეტიკული მდგომარეობისათვის. ტემპერატურის შემცირებასთან ერთად მაღალენერგეტიკული მდგომარეობების ალბათობა მცირდება, ხოლო ნულთან მიახლოებისას იმის ალბათობა, რომ სისტემა იმყოფება მაღალენერგეტიკულ მდგომარეობაში, ძალზე უმნიშვნელო სიდიდე ხდება.

3.2. სწავლების ბოლცმანის სტოქასტიკური მეთოდი

ხელოვნური ნეირონული ქსელების ბოლცმანის სტოქასტიკური მეთოდით სწავლების მიზნით უნდა ჩატარდეს შემდეგი პროცედურები:

1. განისაზღვროს T ცვლადის მნიშვნელობა (ე.წ. ხელოვნური ”ტემპერატურა”), თავდაპირველად მას ანიჭებენ დიდ რიცხვით მნიშვნელობას;

2. წარედგინოს ქსელს შესასვლელი სიმრავლე და გამოითვალოს გამოსასვლელები და მიზნობრივი ფუნქცია;

3. შემთხვევითი სიდიდით შეიცვალოს წონითი კოეფიციენტი და განისაზღვროს შესაბამისი ქსელის გამოსასვლელისა და მიზნობრივი ფუნქციის ცვლილება,

რომელიც გამოწვეულია წონითი კოეფიციენტის ცვლილებით;

4. მიზნობრივი ფუნქციის შემცირების შემთხვევაში წონითი კოეფიციენტის ცვლილებას ვიმახსოვრებთ, თუ წონითი კოეფიციენტის ცვლილება იწვევს მიზნობრივი ფუნქციის გაზრდას, მაშინ ამ ცვლილების შენახვის ალბათობა გამოითვლება ბოლცმანის განაწილების მიხედვით:

$$P(c) = \exp(-c/kT)\ , \qquad (3.2)$$

სადაც P(c) არის მიზნობრივ ფუნქციაში c ცვლილების ალბათობა; k – ბოლცმანის მუდმივას ანალოგიური მუდმივა, რომელიც ამოცანის შესაბამისად შეირჩევა; T - ხელოვნური ტემპერატურა.

შეირჩევა რაიმე r შემთხვევითი რიცხვი ნულიდან ერთამდე, თუ P(c) არის r -ზე მეტი, მაშინ ცვლილება შეინახება, წინააღმდეგ შემთხვევაში წონით კოეფიციენტს უნდა დაუბრუნდეს წინა მნიშვნელობა. ეს საშუალებას აძლევს სისტემას გააკეთოს შემთხვევითი ბიჯი მიზნობრივი ფუნქციის გაუარესებისაკენ, რითაც სისტემა ამოდის ლოკალური მინიმუმიდან, სადაც ნებისმიერი მცირე ბიჯი ზრდის მიზნობრივ ფუნქციას.

ბოლცმანის სტოქასტიკური მეთოდით ხელოვნური ნეირონული ქსელის სწავლების დასასრულებლად საჭიროა 3 და 4 ბიჯის გამეორება ქსელის თითოეული წონისათვის. ამასთან თანდათანობით მცირდება ”ტემპერატურა”, ვიდრე მიიღწევა მიზნობრივი ფუნქციის დასაშვები დაბალი დონე. ამ მომენტში ქსელს წარედგინება მომდევნო შემავალი ვექტორი და სწავლების პროცესი მეორდება. ეს პროცედურები მეორდება სასწავლო სიმრავლის ყველა ვექტორისათვის, ვიდრე მიზნობრივი ფუნქცია ყველასათვის არ გახდება დასაშვები სიდიდის. მესამე ბიჯზე წონითი კოეფიციენტის ცვლილების შემთხვევითი სიდიდე

შეიძლება განისაზღვროს სხვადასხვა მეთოდით, მაგალითად, გაუსის განაწილების შესაბამისად.

დადგენილია, რომ ტემპერატურის შემცირების სიჩქარე უნდა იყოს დროის ლოგარითმის უკუპროპორციული, რათა მიღწეულ იქნეს კრებადობა გლობალურ მინიმუმთან. გაცივების სიჩქარე იქნება

$$T(t) = \frac{T_0}{\log(1+t)}, \qquad (3.3)$$

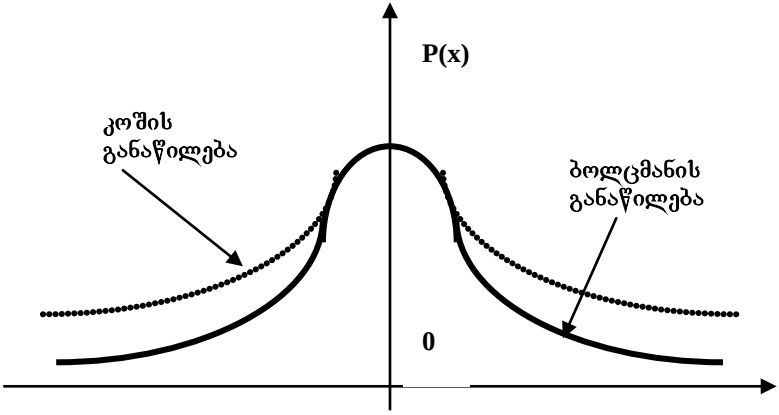
სადაც $T(t)$ არის ხელოვნური ტემპერატურა როგორც დროის ფუნქცია; T_0 - საწყისი ხელოვნური ტემპერატურა; t - ხელოვნური დრო.

(3.3) განტოლებიდან ჩანს, რომ მოსალოდნელია გაცივების დაბალი სიჩქარე, რაც დადასტურდა ექსპერიმენტულად. ბოლცმანის სტოქასტიკური მეთოდით ხელოვნური ნეირონული ქსელების სწავლება ხშირად ითხოვს მანქანური დროის ძალზე დიდ რესურსს.

3.3. სწავლების კოშის მეთოდი

ბოლცმანის სტოქასტიკური მეთოდის განვითარებას წარმოადგენს სწავლების კოშის მეთოდი. ამ მეთოდში ბოლცმანის განაწილება შეცვლილი არის კოშის განაწილებით, რომელსაც აქვს უსასრულო დისპერსია, ეს კი ზრდის დიდი ბიჯების ალბათობას. აღნიშნული მარტივი ცვლილების შედეგად ტემპერატურის შემცირების მაქსიმალური სიჩქარე იქნება დროის წრფივი სიდიდის უკუპროპორციული და არა დროის ლოგარითმის უკუპროპორციული, როგორც ეს იყო ბოლცმანის სტოქასტიკური მეთოდით სწავლებისას

$$T(t) = \frac{T_0}{1+t} \ . \qquad (3.4)$$



სურ. 3.3. ბოლცმანისა და კოშის განაწილება.

ეს გარემოება მკვეთრად ამცირებს ქსელის სწავლების დროს. კოშის განაწილებას აქვს შემდეგი სახე:

$$P(x) = \frac{T(t)}{T(t)^2 + x^2}, \qquad (3.5)$$

სადაც $P(x)$ არის x სიდიდის ბიჯის ალბათობა. ბოლცმანის განაწილება და კოშის განაწილება ნაჩვენებია 3.3 სურათზე. მიუხედავად იმისა, რომ ბოლცმანის მეთოდთან შედარებით კოშის მეთოდი მკვეთრად ამცირებს ქსელის სწავლების დროს, ეს მაჩვენებელი შეიძლება მაინც დიდი აღმოჩნდეს. ქსელის სწავლების ხანგრძლივობის შემდგომი შემცირება შესაძლებელია ე.წ. ხელოვნური თბოტევადობის მეთოდით, რომელიც რთულია და ნაკლებად გამოიყენება.

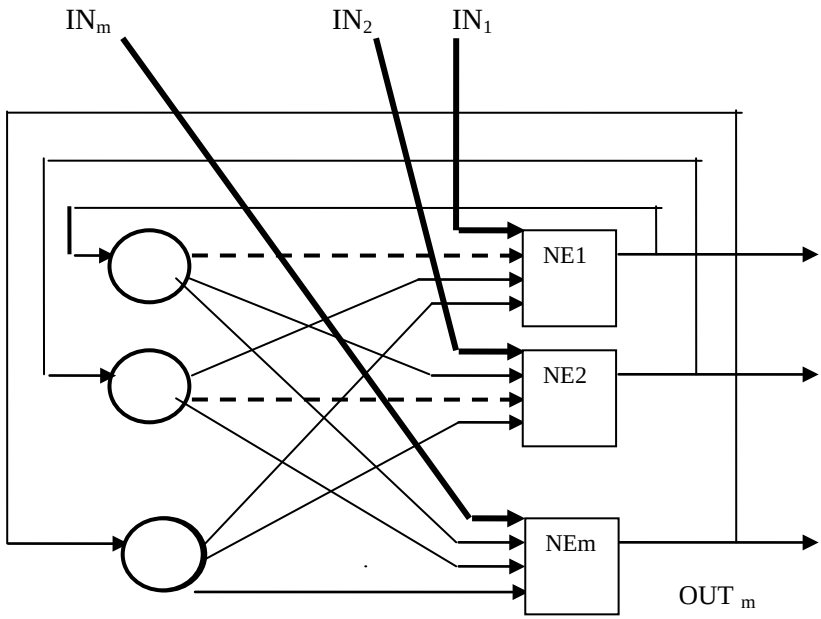
3.4. ჰოპფილდის ნეირონული ქსელები

მრავალშრიანი ხელოვნური ნეირონული ქსელების ის კლასი, რომელსაც არ გააჩნია ქსელის გამოსასვლელების დამაკავშირებელი არხები წინამდებარე შრეებთან, ანუ უკუკავშირები, განხილული იყო ზემოთ. ასეთი ნეირონული ქსელები ხასიათდება უპირობო მდგრადობით, რასაც განსაზღვრავს სწორედ ის გარემოება, რომ მათ არ გააჩნია უკუკავშირები. მართალია მდგრადობა სისტემების მეტად სასურველი მახასიათებელია, მაგრამ ასეთი ნეირონული ქსელების შესაძლებლობები შეზღუდულია უკუკავშირებიან ქსელებთან შედარებით. თავის მხრივ უკუკავშირებიანი ქსელებიც ზოგჯერ ავლენს არამდგრადობას, ასეთ შემთხვევაში კი ნეირონული ქსელი პრაქტიკული მიზნებისათვის ხდება უსარგებლო. უკუკავშირებიანი ქსელი დინამიურად რეაგირებს შესასვლელის ცვლილებაზე, ანუ ახალი რეალიზაციის მიღების შემდეგ ქსელი გამოითვლის სიგნალის მნიშვნელობებს გამოსასვლელზე, რომელიც უკუკავშირებით გადაეცემა ქსელს და გამოიწვევს შესასვლელის მოდიფიცირებას. ამის შემდეგ კვლავ გამოითვლება გამოსასვლელი და ეს პროცესი მეორდება მრავალჯერ. იმ შემთხვევაში თუ უკუკავშირებიანი ქსელი მდგრადია, იტერაციების შედეგად ნეიროქსელის გამოსასვლელის მნიშვნელობა შეიცვლება სულ უფრო მცირე სიდიდით, სანამ გამოსასვლელი გახდება მუდმივი სიდიდე (მოხდება გამოსასვლელის სტაბილიზება). ზოგიერთი კონფიგურაციის ნეირონულ ქსელში ეს პროცესი არასოდეს არ მთავრდება, ასეთ უკუკავშირებიან ქსელს უწოდებენ არამდგრად ნეირონულ ქსელს. ჩვენთვის საინტერესოა მდგრადი ნეირონული სისტემები, ანუ ისეთი ქსელები, რომლებიც იტერაციების შედეგად ნეირონული ქსელის გამოსასვლელზე საბოლოოდ გამოიმუშავებს სტაბილურ სიგნალებს.

ნეირონული ქსელის მდგრადობის გარკვევა ნეიროინფორმატიკის განვითარების პირველ ეტაპზე სერიოზულ პრობლემას წარმოადგენდა. 1983 წელს ს. გროსბერგმა და მ. კოჰენმა აღწერეს უკუკავშირებიანი ნეირონული ქსელების ისეთი ქვეკლასი, რომლის გამოსასვლელები იტერაციების შედეგად საბოლოოდ აღწევენ მდგრად მდგომარეობას და დაამტკიცეს შესაბამისი თეორემა. ამ მიღწევამ გზა გაუხსნა ამგვარი რთული ქცევის მქონე ნეირონული ქსელების შესაძლებლობების კვლევას. უკუკავშირებიანი ნეირონული ქსელების თეორიასა და პრაქტიკაში მნიშვნელოვანი წვლილი შეიტანა ჯ. ჰოპფილდმა, ამიტომ ასეთი სისტემების ზოგიერთი კონფიგურაციები ცნობილია როგორც ჰოპფილდის ქსელები.

3.4.1. უკუკავშირებიანი ნეირონული ქსელების კონფიგურაციები

ერთშრიანი უკუკავშირებიანი ნეირონული ქსელი მოცემულია 3.1. სურათზე, სადაც პუნქტირით ნაჩვენებია ნულოვანი წონითი კოეფიციენტები. შემავალი (ნულოვანი) შრის ნეირონები გამოთვლების ოპერაციებს არ ახორციელებს, არამედ წარმოადგენს განშტოების წერტილებს და გამოსასვლელებიდან მიღებულ სიგნალებს უკუკავშირით მიაწოდებს ქსელის შესასვლელებს. ნულოვანი შრის თითოეული ნეირონი უკავშირდება პირველი შრის თითოეულ ნეირონს. პირველი შრის ნეირონი გამოითვლის თავისი შესასვლელების შეწონილ ჯამს, ასეთი NET ჯამები უნდა იქნეს გამოთვლილი თითოეული ნეირონისათვის, რის შემდეგაც ხდება ჯამების მოდიფიცირება F არაწრფივი აქტივაციის ფუნქციით და საბოლოოდ



სურ. 3.1.

მიიღება OUT გამოსასვლელი სიგნალი:

$$NET = \sum x_i w_i \qquad OUT = F\,(NET). \qquad (3.1)$$

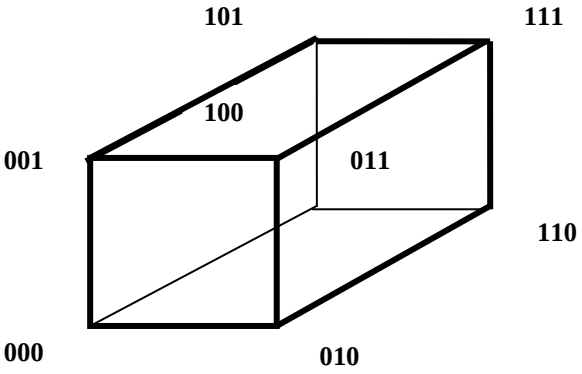
ჯ. ჰოპფილდის პირველ ნაშრომში (1982 წ.) არაწრფივ აქტივაციის F ფუნქციას წარმოადგენდა ზღურბლური ფუნქცია. ასეთი ნეირონის გამოსასვლელი ლოგიკური ერთის ტოლია, თუ სხვა ნეირონების გამოსასვლელების შეწონილი ჯამი T_j ზღურბლზე მეტია და ლოგიკური ნულის ტოლია, თუ შეწონილი ჯამი T_j ზღურბლზე ნაკლებია. იგი გამოითვლება შემდეგნაირად:

$$NET_j = \sum_{i \neq j} w_{ij} OUT_i + IN_j\,, \qquad (3.2)$$

OUT = 1, თუ $NET_j > T_j$, OUT = 0, თუ $NET_j < T_j$,

OUT სიდიდე არ იცვლება, თუ $NET_j = T_j$,

ჯ. ჰოპფილდის აღნიშნულ ნაშრომში აღწერილ ნეირონულ ქსელში თითოეული ნეირონის მდგომარეობა (ქსელის OUT გამოსასვლელი სიგნალების მიმდინარე მნიშვნელობების სიმრავლე) იცვლებოდა დროის შემთხვევით დისკრეტულ მომენტებში, ჰოპფილდის შემდგომ ნაშრომში აღწერილ ქსელში კი ნეირონებს შეუძლია შეიცვალოს მდგომარეობა ერთდროულად. იმის გამო, რომ ბინარული ნეირონის გამოსასვლელი შეიძლება იყოს მხოლოდ ლოგიკური ნული ან ლოგიკური ერთი, ამიტომ ასეთი ნეირონული ქსელის მიმდინარე მდგომარეობას წარმოადგენს ორობითი რიცხვი, რომლის თითოეული თანრიგი არის გარკვეული ნეირონის გამოსასვლელი OUT სიგნალი. ჰოპფილდის ხელოვნური ნეირონული ქსელების ფუნქციონირების მექანიზმი შეიძლება წარმოვიდგინოთ გეომეტრიულად. ნეირონული ქსელის ოთხი მდგომარეობა შეიძლება ჰქონდეს ისეთ ქსელს, რომელსაც გამოსასვლელ შრეში აქვს ორი ნეირონი (ორგანზომილებიანი სივრცე). ამ შემთხვევაში კვადრატის თითოეულ წვეროს შეესაბამება სისტემის ოთხი შესაძლებელი მდგომარეობიდან (00, 01, 10, 11) ერთი მდგომარეობა. ნეირონული ქსელის რვა მდგომარეობა შეიძლება ჰქონდეს ისეთ ქსელს, რომელსაც გამოსასვლელ შრეში აქვს სამი ნეირონი (სამგანზომილებიანი სივრცე). 3.2. სურათზე კუბის თითოეულ წვეროს შეესაბამება ერთი მდგომარეობა სისტემის რვა შესაძლებელი მდგომარეობიდან, რომლებიც აისახება სამთანრიგიანი ბინარული რიცხვებით ზოგად შემთხვევაში ნეიროქსელის ფუნქციონირება, რომელსაც გამოსასვლელ შრეში აქვს n ნეირონი და ნეირონული ქსელის 2^n რაოდენობის სხვადასხვა მდგომარეობა, შეიძლება წარმოვადგინოთ n -განზომილებიანი ჰიპერკუბით.



სურ. 3.2.

როდესაც ნეირონულ ქსელს მიეწოდება ახალი შემავალი ვექტორი (რეალიზაცია) ქსელის მდგომარეობა გადადის ერთი წვეროდან სხვა წვეროზე, სანამ მდგომარეობა სტაბილური გახდება. მდგრადი წვერო განისაზღვრება ქსელის წონითი კოეფიციენტების, ნეირონების ზღურბლის სიდიდისა და შესასვლელების მიმდინარე მნიშვნელობებით. იმ შემთხვევაში, როდესაც რეალიზაცია დამახინჯებულია ან არასრულია, მაშინ ნეირონული ქსელის სტაბილიზება მოხდება იმ წვეროში, რომელიც არის უახლოესი სასურველ მდგომარეობასთან.

ცნობილია, რომ ნეიროქსელის შრეებსშორისი წონითი კოეფიციენტები შეიძლება განვიხილოთ როგორც W მატრიცა. გროსბერგმა და კოჰენმა დაამტკიცეს, რომ უკუკავშირებიანი ნეირონული ქსელი მდგრადია, თუ მისი წონითი კოეფიციენტების W მატრიცა სიმეტრიულია და მთავარი დიაგონალი შედგება ნულებისაგან, ანუ თითოეული i-სათვის $w_{ij} = w_{ji}$ და $w_{ii} = 0$. წონითი კოეფიციენტების მატრიცის სიმეტრიულობა არის ნეირონული ქსელის მდგრადობის საკმარისი, მაგრამ არ არის აუცილებელი პირობა, ანუ

არსებობს მრავალი მდგრადი სისტემა (მაგალითად პირდაპირი მოქმედების ყველა ნეიროქსელი), რომელიც არ აკმაყოფილებს ზემოთ მოცემულ პირობას. არის ისეთი მაგალითებიც, როდესაც სიმეტრიისაგან უმნიშვნელო გადახრაც კი იწვევს ქსელის განუწყვეტელ ოსცილაციას, თუმცა როგორც წესი, მიახლოებითი სიმეტრია საკმარისია სისტემის მდგრადობისათვის.

3.5. ასოციაციური მეხსიერება

უკუკავშირებიანი ნეირონული ქსელის მდგრადობის თვისება განაპირობებს ასოციაციური მეხსიერების ფორმირების შესაძლებლობას, რომელიც ადამიანის ასოციაციური მეხსიერების მსგავსია. ადამიანის მეხსიერება ასოციაციურია, ანუ რაიმე ინფორმაციის მიღებამ შესაძლებელია გამოიწვიოს ადამიანის მიერ ამ ინფორმაციასთან დაკავშირებული სხვადასხვა მოვლენებისა და საგნების გახსენება, მაშინ როდესაც ტრადიციული ფონ ნეიმანის არქიტექტურის კომპიუტერის მეხსიერებიდან ინფორმაციის მოძიება მხოლოდ სასურველი ინფორმაციის ზუსტი მისამართის მითითებით არის შესაძლებელი, ხოლო ამ ინფორმაციასთან დაკავშირებული სხვადასხვა მოვლენებისა და საგნების შესახებ ინფორმაცია მათი ზუსტი მისამართების ცოდნის გარეშე მიუწვდომელია. ასოციაციურ მეხსიერებას, რომელიც ფორმირებულია უკუკავშირებიანი ნეირონული ქსელით, შეუძლია ადამიანის მეხსიერების მსგავსად, საჭირო ინფორმაციის მოცემული ნაწილის მიხედვით მოიძიოს მეხსიერებიდან სრული ინფორმაცია.

უკუკავშირებიანი ნეირონული ქსელით ასოციაციური მეხსიერების ორგანიზებისათვის ნეირონების პარამეტრები - წონითი კოეფიციენტები უნდა შეირჩეს ისე, რომ ერთეულოვანი ჰიპერკუბის სასურველ წვეროებში წარმოიქმნას მინიმალური შეცდომები (ენერგეტიკული მინიმუმები).

ჯ. ჰოპფილდმა შეიმუშავა ასოციაციური მეხსიერება, რომელიც ფორმირებულია უწყვეტი გამოსასვლელების მქონე უკუკავშირებიანი ნეირონული ქსელით, სადაც გამოსასვლელები იცვლება პლიუს ერთიდან მინუს ერთამდე, რაც შეესაბამება ლოგიკური ნულისა და ლოგიკური ერთის ბინარულ მნიშვნელობებს. დასამახსოვრებელი ინფორმაციის კოდირება ხდება ბინარული ვექტორებით, რომლებიც ინახება წონითი კოეფიციენტების სახით შემდეგი გამოსახულების შესაბამისად

$$w_{ij} = \sum_{d=1...m} (OUT_{i,d} OUT_{j,d}) \qquad (3.3)$$

სადაც m არის დასამახსოვრებელი ბინარული ვექტორების რაოდენობა; d - დასამახსოვრებელი ვექტორის ნომერი; $OUT_{i,j}$ - დასამახსოვრებელი გამომავალი ვექტორის i - ური კომპონენტი.

მას შემდეგ, რაც წონითი კოეფიციენტები მოცემულია, შესაძლებელია ქსელის გამოყენება დამახსოვრებული გამომავალი ვექტორის მისაღებად შემავალი ვექტორის მიხედვით, რომელიც შეიძლება იყოს ნაწილობრივ არასწორი (დამახინჯებული) ან არასრული. ამისათვის ქსელის გამოსასვლელებს პროცედურის დასაწყისში მიაწოდებენ შემავალი ვექტორის მნიშვნელობას, შემდეგ ქსელის გამოსასვლელებს მოაცილებენ შემავალ ვექტორს, შედეგად ქსელის მდგომარეობა გადადის ჰიპერკუბის ერთი წვეროდან სხვა წვეროზე, სანამ მდგომარეობა გასტაბილურდება უახლოეს ენერგეტიკულ მინიმუმში, თუმცა ეს შეიძლება იყოს არა გლობალური, არამედ ლოკალური მინიმუმი, ანუ ქსელის მდგომარეობა შეიძლება გასტაბილურდეს გლობალური მინიმუმის მოძიებამდე (საუკეთესო გადაწყვეტილების მიღებამდე). აღნიშნული არასასურველი ტენდენცია განაპირობებს ჰოპფილდის ქსელების ნაკლოვანი მხარეს.

ჰოპფილდმა განიხილა აგრეთვე მოდელები უწყვეტი აქტივაციის ფუნქციით, რომელიც უფრო ზუსტად ასახავს ბიოლოგიურ ნეირონის მოქმედებას, ზოგად შემთხვევაში ეს შეიძლება იყოს სიგმოიდური ფუნქცია:

$$F(x) = \frac{1}{1 + \exp(-\lambda \mathrm{NET})},$$

სადაც λ არის კოფიციენტი, რომელიც განსაზღვრავს სიგმოიდური ფუნქციის დახრილობას. თუ λ დიდია, მაშინ ადრე აღწერილ ზღურბლურ ფუნქციას უახლოვდება, ხოლო λ -ს მცირე მნიშვნელობები კი განაპირობებს ფუნქციის ნაკლებ დამრეცობას. ისევე როგორც ბინარული სისტემებისათვის, ასეთი ნეირონული ქსელიც მდგრადია, თუ მისი წონითი კოეფიციენტების **W** მატრიცა სიმეტრიულია და მთავარი დიაგონალი შედგება ნულებისაგან. თუ λ დიდია, მაშინ უწყვეტი სისტემები ფუნქციონირებს ბინარული სისტემების მსგავსად და საბოლოოდ სტაბილური ხდება ერთეულოვანი ჰიპერკუბის წვეროებში. λ -ს შემცირებით მდგრადი წერტილები სცილდება წვეროებს და თანდათანობით ქრება λ -ს ნულთან მიახლოვების შესაბამისად. ზემოთ აღინიშნა ჰოპფილდის ქსელების არასასურველი ტენდენცია რაც იმაში გამოიხატება, რომ ქსელის მდგომარეობა შეიძლება სტაბილური გახდეს გლობალური მინიმუმის მონახვამდე უახლოეს ლოკალურ ენერგეტიკულ მინიმუმში და არა ენერგიის ფუნქციის გლობალურ მინიმუმში. ამ სიძნელის გადალახვას ცდილობენ ერთი სახეობის ნეიროქსელებით, რომლებიც ცნობილია **ბოლცმანის მანქანების** სახელწოდებით, სადაც ნეირონების მდგომარეობის ცვლილებები განპირობებულია არა დეტერმინირებული, არამედ სტატისტიკური კანონზომიერებებით. არსებობს გარკვეული ანალოგია ამ მეთოდებსა და ლითონთა წრთობის პროცესებს შორის, რის გამოც ზოგი

მკვლევარი ასეთ მეთოდებს უწოდებს წრთობის იმიტაციას.

უკუკავშირებიანი ნეირონული ქსელების დინამიური ქცევა ნეიროსისტემების ახალი, საინტერესო თვისებების პრაქტიკული გამოყენების შესაძლებლობას იძლევა, თუმცა მეცნიერების წინაშე აყენებს სპეციფიურ პრობლემებს, ამდენად უკუკავშირებიანი ნეირონული ქსელები შემდგომი კვლევების პერსპექტიულ ობიექტს წარმოადგენს.

3.6. ორმხრივმიმართული ასოციაციური მეხსიერება

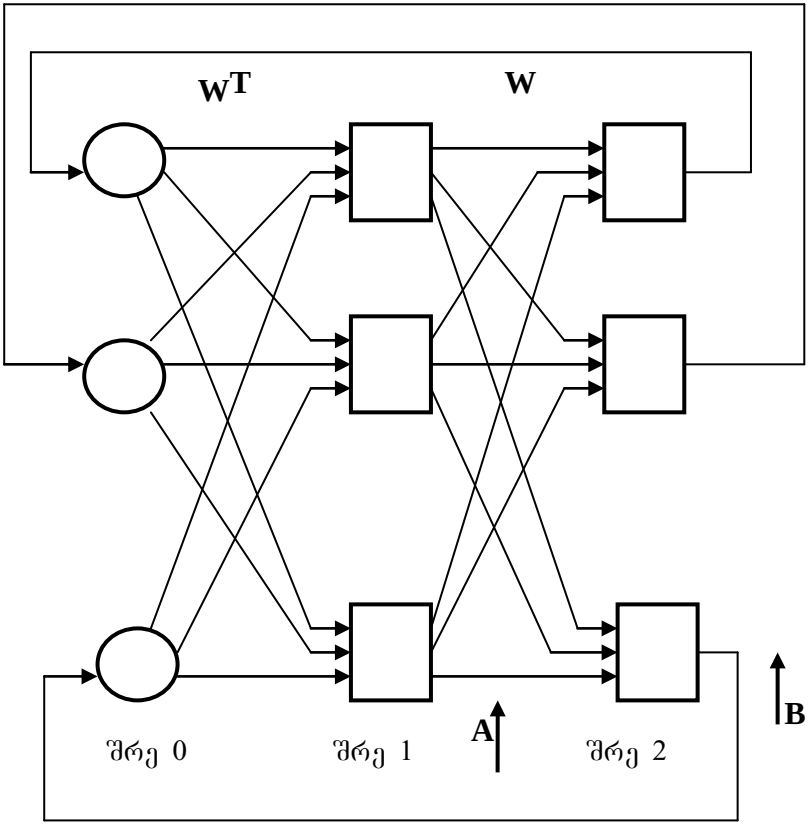
ადამიანის მეხსიერება ხშირად ავლენს ასოციაციურობის თვისებას, ანუ რაიმე საგანი ან მოვლენა გვახსენებს მეორე საგანს ან მოვლენას, ეს კი მესამე საგანს და ა. შ. ადამიანის აზრები შესაძლებელია გაყვეს ამგვარი ასოციაციების ჯაჭვს, სწორედ ამ თვისების გამოყენებით შეუძლია ადამიანს უკვე დავიწყებული სხვადასხვა მოვლენებისა და საგნების გახსენება.

ასოციაციური მეხსიერება, რომელიც ზემოთ იყო წარმოდგენილი, მკაცრად რომ განვიხილოთ არის **ავტოასოციაციური,** ეს ნიშნავს, რომ ნეირონულ ქსელში შემავალი ნაწილობრივ არასწორი (დამახინჯებული) ან არასრული სახე (რეალიზაცია) შესაძლებელია გასწორებული ან დასრულებული იქნას, მაგრამ შეუძლებელია მისი ასოცირება სხვა სახესთან. ეს ფაქტი გამოწვეულია იმით, რომ აღნიშნული ასოციაციური მეხსიერების სტრუქტურა არის ერთდონიანი, ანუ გამომავალი ვექტორი მიიღება იმავე ნეირონებზე, რომლებსაც მიეწოდება შემავალი ვექტორი. სრულფასოვან, ანუ **ჰეტეროასოციაციურ** მეხსიერებას წარმოადგენს ორმხრივმიმართული ასოციაციური მეხსიერება, რომლის სტრუქტურა ზოგად შემთხვევაში შეიძლება იყოს მრავალდონიანი. ორმხრივმიმართულ

ასოციაციურ მეხსიერებაში შემავალი ვექტორი წარედგინება ნეირონების ერთ ნაკრებს, ხოლო გამომავალ ვექტორს გამოიმუშავებს ნეირონების სხვა ნაკრები. ისევე როგორც ჰოპფილდის ქსელებს, ორმხრივმიმართულ ასოციაციურ მეხსიერებასაც გააჩნია განზოგადების უნარი, მას შეუძლია გამოიმუშაოს სწორი რეაქციები მაშინაც, როდესაც ნეიროქსელის შესასვლელს მიეწოდება ნაწილობრივ არასწორი (დამახინჯებული) ან არასრული სახე. აღსანიშნავია, რომ შესაძლებელია ორმხრივმიმართული ასოციაციური მეხსიერების ისეთი ადაპტური ვერსიის ფორმირება, რომელიც უზრუნველყოფს ეტალონური სახის გამოყოფას დამახინჯებული რეალიზაციებისაგან. ზემოთაღწერილი ქსელების შესაძლებლობები გვაგონებს ადამიანის ტვინში მიმდინარე პროცესებს, ამდენად ნეიროკომპიუტინგის ამ მიმართულბის განვითარება კიდევ ერთი გადადგმული ნაბიჯი იქნება ადამიანის აზროვნების პროცესების მოდელირების პრობლემის გადაწყვეტის გზაზე.

ორმხრივმიმართული ასოციაციური მეხსიერების სტრუქტურა. ორმხრივმიმართული ასოციაციური მეხსიერების საბაზო კონფიგურაცია ნაჩვენებია 3.4. სურათზე, სადაც ჩანს მისი მსგავსება ჯ. ჰოპფილდის ქსელებთან, ნათელია აგრეთვე შრეების რაოდენობის გაზრდის შესაძლებლობაც. **A** შემავალ ვექტორს დაამუშავებს ნეირონული ქსელის **W** მატრიცა, შედეგად ქსელი გამოიმუშავებს **B** გამომავალ ვექტორს. **B** გამომავალ ვექტორს დაამუშავებს ტრანსპონირებული $\mathbf{W}^{\mathrm{T}}$ მატრიცა, რომელიც გამოიმუშავებს ნეირონული ქსელის ახალ გამომავალ სიგნალებს, რომელიც წარმოადგენს ახალ **A** შემავალ ვექტორს. ეს პროცესი გრძელდება, სანამ ნეირონული ქსელი მიაღწევს სტაბილურ მდგომარეობას. სტაბილური მდგომარეობისას არც **A** ვექტორი და არც **B** ვექტორი არ იცვლება. ქსელის ფორმალური ნეირონების მიმდინარე

მდგომარეობა განისაზღვრება როგორც მისი შესასვლელების შეწონილი ჯამი, ხოლო ნეირონის გამომავალი სიგნალი როგორც მისი მიმდინარე



სურ. 3.4.

მდგომარეობის ფუნქცია:

$$b_i = F(\sum_j a_j w_{ij})$$

ან ვექტორული ფორმით:

$$\mathbf{B} = F(\mathbf{AW}),$$

სადაც **B** არის მეორე შრის ნეირონების გამომავალი სიგნალების ვექტორი, **A** – პირველი შრის ნეირონების გამომავალი სიგნალების ვექტორი, **W** – პირველ და მეორე შრეებს შორის სინაფსური კავშირების წონების მატრიცა, F – აქტივაციის ფუნქცია. ანალოგიურად

$$\mathbf{A} = F(\mathbf{BW}^{\mathrm{T}})$$

სადაც $\mathbf{W}^{\mathrm{T}}$ არის ტრანსპონირებული სინაფსური კავშირების წონების მატრიცა.

როგორც ზემოთ აღინიშნა (თავი I), გროსბერგმა აჩვენა სიგმოიდური აქტივაციის ფუნქციის გამოყენების უპირატესობა. ორმხრივმიმართული ასოციაციური მეხსიერების ძირითად ვერსიებში ირჩევენ სიგმოიდური აქტივაციის ფუნქციის λ კონსტანტის დიდ მნიშვნელობას, შედეგად აქტივაციის ფუნქციის სახე უახლოვდება ზღურბლურ ფუნქციას. შემდგომ განხილვაში მივიღოთ, რომ ვიყენებთ ზღურბლურ აქტივაციის ფუნქციას. მივიღოთ აგრეთვე, რომ პირველი და მეორე შრეების ნეირონებს გააჩნიათ მეხსიერება და ნეირონების გამომავალი სიგნალები ერთდროულად იცვლება სინქრონიზაციის თვითოეულ ტაქტზე, ხოლო ტაქტებს შორის სიგნალების მნიშვნელობა მუდმივია, მაშინ ნეირონების ქცევა აღიწერება შემდეგნაირად:

$$\begin{aligned}
&OUT_i(n+1) = 1, && \text{თუ } NET_i(n) > 0,\\
&OUT_i(n+l) = 0, && \text{თუ } NET_i(n) < 0,\\
&OUT_i(n+l) = OUT(n), && \text{თუ } NET_i(n) = 0,
\end{aligned}$$

სადაც $OUT_i(n)$ წარმოადგენს *i*-ური ნეირონის გამომავალი სიგნალის მნიშვნელობას დროის *n* მომენტში. ნულოვან შრეს არ გააჩნია მეხსიერება, იგი

ანაწილებს მეორე შრის გამომავალ სიგნალებს $\mathbf{W}^T$ მატრიცის ელემენტებზე.

3.7. დამახსოვრებული ასოციაციების აღდგენა

ორმხრივმიმართულ ასოციაციურ ნეიროქსელებში მეხსიერება, ანუ ასოციაციები რეალიზებულია **W** და $\mathbf{W}^T$ სინაფსური კავშირების წონების მატრიცებში. თითოეული სახე წარმოდგენილია ორი ვექტორით, ესენია პირველი შრის გამომავალი **A** ვექტორი და მასთან ასოცირებული - მეორე შრის გამომავალი **B** ვექტორი. ასოცირებული სახის აღსადგენად **A** ვექტორი ან მისი ნაწილი მოკლე ხნით მიეწოდება პირველი შრის გამოსასვლელებს, შემდეგ ეს სიგნალები (**A** ვექტორი) მოიხსნება და ქსელი გადადის სტაბილურ მდგომარეობაში, რაც იწვევს მეორე შრის გამოსასვლელებზე ასოცირებული **B** ვექტორის გამომუშავებას. ამის შემდეგ **B** ვექტორი ზემოქმედებს $\mathbf{W}^T$ ტრანსპონირებული სინაფსური კავშირების წონების მატრიცის მეშვეობით და აღადგენს საწყისი **A** ვექტორის ზემოქმედებას პირველი შრის გამოსასვლელებზე. ყოველი ასეთი ახალი ციკლი აზუსტებს პირველი და მეორე შრის გამომავალი ვექტორების მნიშვნელობებს, სანამ მიიღწევა ქსელის სტაბილიზების წერტილი. ამრიგად, ორმხრივმიმართული ქსელი ძირითადად ჰოპფილდის ავტოასოციაციური ქსელის მსგავსად ფუნქციონირებს. თითოეული ციკლი სისტემის მოდიფიცირებას ახდენს ენერგეტიკული მინიმუმის მიმართულებით, რომლის ადგილმდებარეობას განსაზღვრავს წონითი კოეფიციენტების მნიშვნელობები. სახეთა სიმრავლის ამოცნობის მიზნით ქსელის სწავლება მიმდინარეობს სასწავლო ნაკრების გამოყენებით, რომელიც **A** და **B** ვექტორების წყვილებისაგან შედგება. სინაფსური კავშირების წონების მატრიცა გამოითვლება

როგორც სასწავლო ნაკრების **A** და **B** ვექტორების წყვილების ნამრავლების ჯამი

$$\mathbf{W} = \sum_i \mathbf{A}_i^t \mathbf{B}_i$$

დამტკიცებულია, რომ ორმხრივმიმართული ასოციაციური ნეიროქსელების მწარმოებლობის ამაღლება შეიძლება ბიპოლარული ვექტორების გამოყენებით. ამ შემთხვევაში ვექტორების კომპონენტები, რომლებიც ნულზე ნაკლები ან ნულის ტოლია იცვლება მინუს ერთი სიდიდით (–1), ხოლო კომპონენტები, რომლებიც ნულზე მეტია, იცვლება პლუს ერთი სიდიდით (+1).

ორმხრივმიმართულ ასოციაციურ ნეიროქსელებს გააჩნია განზოგადების თვისება. მაგალითად, თუ **A** ვექტორის როლში ნეიროქსელს მიეწოდება ნაწილობრივ არასწორი (დამახინჯებული) ან არასრული სახე, მაშინ ქსელი ადრე დამახსოვრებული **B** ვექტორის გამომუშავების ტენდენციას ავლენს, ხოლო **B** ვექტორი თავის მხრივ ცდილობს გაასწოროს შეცდომები **A** ვექტორში. ამისათვის საჭირო იქნება რამდენიმე ციკლი, მაგრამ ქსელი საბოლოოდ მაინც ადადგენს უახლოეს დამახსოვრებულ სახეს.

ნაჩვენებია, რომ არსებობს ურთიერთკავშირი ორმხრივმიმართულ ასოციაციურ ნეიროქსელებსა და ჰოპფილდის ნეირონულ ქსელებს შორის. თუ **W** სინაფსური კავშირების წონების მატრიცა არის კვადრატული და სიმეტრიული, მაშინ $\mathbf{W} = \mathbf{W}^T$. ამ შემთხვევაში ორმხრივმიმართული ასოციაციური ნეიროქსელი გადაიქცევა ჰოპფილდის ავტოასოციაციურ ქსელად.

ისევე როგორც ჰოპფილდის ნეირონულ ქსელს, ორმხრივმიმართულ ასოციაციურ ნეიროქსელსაც გააჩნია იმ ასოციაციების რაოდენობის შეზღუდვა, რომლის ზუსტი აღდგენა შეუძლია. ამ ლიმიტის გადაჭარბების შემთხვევაში გამოიმუშავებს არასწორ

ასოციაციებს. ორმხრივმიმართული ასოციაციური ნეიროქსელის პრაქტიკაში ფართოდ გამოყენებას ხელს უშლის სხვა პრობლემებიც, თუმცა სიმარტივის გამო იგი რჩება ინტენსიური კვლევების ობიექტად.

3.8. ხელოვნური ნეირონული ქსელების სინთეზის პროცედურები

ნეიროქსელური ტექნოლოგიებით ძნელად ფორმალიზებადი ამოცანების გადაწყვეტა ნიშნავს ამ ამოცანების ადეკვატური ხელოვნური ნეირონული ქსელების სინთეზს. ნეირონული ქსელის სინთეზისას განსაკუთრებული მნიშვნელობა ენიჭება ნიშანთა სივრცის სათანადო შერჩევას. იმ შემთხვევაში, თუ პარამეტრების რაოდენობა მცირეა, შესაძლებელია ისეთი სიტუაციები, როდესაც რეალიზაციები (შემავალი ვექტორები) ერთნაირი ელემენტებისაგან შედგება, მაგრამ ეკუთვნის სხვადასხვა კლასტერებს, ანუ ასეთ სიტუაციის დროს ნეირონული ქსელის სწავლება შეუძლებელი იქნება. მეორე მხრივ, ნიშანთა სივრცის განზომილების გაზრდისას შეიძლება სასწავლო ნაკრების სიმძლავრე არასაკმარისი გახდეს ნეირონული ქსელის სწავლებისათვის, ასეთ დროს ქსელი მხოლოდ დაიმახსოვრებს რეალიზაციებს მათი განზოგადების ნაცვლად და ვერ მოახდენს საჭირო ამოცანის გადაწყვეტას (კლასტერირება, პროგნოზირება, სახეთა ამოცნობა და სხვა). ამოცანების დაყოფა შესაძლებელია სამ ჯგუფად: პირველ ჯგუფს შეადგენს ისეთი ამოცანები, სადაც ქვესიმრავლეებად დაჯგუფება შეიძლება ჰიპერსიბრტყეებით – ასეთი სიმრავლეები მიეკუთვნება წრფივად განმხოლოებად სიმრავლეებს; მეორე ჯგუფს შეადგენს არაწრფივად განმხოლოებადი სიმრავლეები; მესამე ჯგუფში შედის ისეთი ამოცანები, სადაც ქვესიმრავლეები იკვეთება. რეალური ამოცანების გადაწყვეტისას ზოგჯერ ვერ ხერხდება საწყისი მონაცემების ისეთი წინასწარი დამუშავება, რომ

უზრუნველყოფილი იყოს სახეთა წრფივად განმხოლოებადობა, რაც ართულებს არაგადამკვეთ ქვესიმრავლეებად დაჯგუფების პროცესს და შესაბამისად, საჭირო ნეირონული ქსელების სინთეზის პროცედურებს. ამრიგად, მნიშვნელოვანია იმის გარკვევა არის თუ არა მოცემული სახეთა ნაკრები წრფივად განმხოლოებადი. წრფივად განმხოლოებადობის პრობლემა განხილულია მეორე თავში და ცნობილია, რომ მისი დაძლევა შესაძლებელია ერთშრიანი ნეირონული ქსელისათვის დამატებითი შრეების კასკადური მიერთებით. ასეთ მრავალშრიან ნეირონულ ქსელებს შეუძლია რთული ამოცანების გადაწყვეტა.

ცნობილია, რომ ობიექტთა სიმრავლე შეიძლება წარმოდგენილი იყოს როგორც რაოდენობრივი, ასევე თვისებრივი სახით; ამასთან, რიცხვითი მონაცემები შეიძლება იცვლებოდეს ნებისმიერ დიაპაზონში, ნეირონული ქსელები კი ოპერირებენ ნამდვილი რიცხვებით ნულიდან ერთამდე ფარგლებში. ამის გამო აუცილებელია საწყისი მონაცემების წარმოდგენა რიცხვითი ფორმით და მათი ნორმირება. ნორმირების პროცესი გულისხმობს სახეთა ნიშნების რაოდენობრივი მნიშვნელობების მასშტაბირებას ისე, რომ შესაძლებელი იყოს მათი თანაბარმნიშვნელოვნად გამოყენება ნეიროკომპიუტინგის პროცესებში. არსებობს ნორმირების მრავალი მეთოდი, მარტივი წრფივი გარდაქმნებით დაწყებული და ექსპერტული შეფასებების ფორმირებით დამთავრებული, რომლებიც განხილულია შესაბამის ლიტერატურაში, მაგალითად, ნაშრომში [2].

გარკვეულ პრობლემას წარმოდგენს ნეირონული ქსელის მიერ გენერირებული გამომავალი სიდიდეების (შედეგების) კოდირება იმ შემთხვევებში, როდესაც კლასტერების რაოდენობა ორზე მეტია. კოდირების მრავალ მეთოდს შორის ყველაზე მარტივი და გავრცელებულია გამომავალი სიდიდეების წარმოდგენა ვექტორის სახით, რომლის ელემენტებიც შეესაბამება

კლასტერების ნომრებს, ანუ ვექტორის **j**-ური ელემენტი შეესაბამება **j**-ურ კლასტერს. ჩვეულებრივ მიღებულია, რომ კლასტერის ნომერს განსაზღვრავს ნეირონული ქსელის იმ გამოსასვლელის ნომერი, რომელზეც არის სიგნალის მაქსიმალური მნიშვნელობა. კოდირების ამ მეთოდის გამოყენებისას სასურველია სანდოობის ცნების შემოტანა, ანუ იმის რაოდენობრივად განსაზღვრა, თუ რამდენად სარწმუნოა ის, რომ რეალიზაცია ეკუთვნის ამორჩეულ კლასტერს. სანდოობის დადგენა მარტივად შეიძლება გამომავალი ვექტორის ელემენტების მაქსიმალურსა და მასთან უახლოეს სიდიდეებს შორის სხვაობის განსაზღვრით. იმის ალბათობა, რომ ნეირონულმა ქსელმა სწორი კლასტერირება მოახდინა მით უფრო მაღალია, რაც უფრო მაღალია სანდოობის მაჩვენებელი.

ნეირონული ქსელების სინთეზის ორი ძირითადი მიდგომა არის ცნობილი – კონსტრუქციული და დესტრუქციული. კონსტრუქციული მიდგომისას მინიმალური მოცულობის ნეირონული ქსელის მოცულობის თანდათანობითი გაზრდით მიიღწევა მოთხოვნილი სიზუსტე. აქ საჭირო ხდება ქსელის მოცულობის ყოველი გაზრდის შემდეგ სწავლების პროცესის ხელახალი ჩატარება. დესტრუქციული მიდგომისას დიდი მოცულობის ქსელიდან ხდება იმ ნეირონებისა და კავშირების მოშორება, რომლებიც ნეირონული ქსელის მიერ გენერირებულ შედეგებზე უმნიშვნელოდ მოქმედებენ. აუცილებელია იმ პირობის შესრულებაც, რომ სასწავლო ნაკრებში რეალიზაციების რაოდენობა უნდა აღემატებოდეს სინაფსური წონითი კოეფიციენტების რაოდენობას (ზოგიერთი შეფასებით ათჯერ), წინააღმდეგ შემთხვევაში, ქსელი მხოლოდ დაიმახსოვრებს მონაცემებს და ვერ მოახდენს იმ რეალიზაციების განზოგადებას, რომლებიც არ შევიდა სასწავლო სიმრავლეში. ნეირონული ქსელების არქიტექტურის შერჩევისას საჭირო სხვადასხვა

კონფიგურაციის სტრუქტურისა და ტოპოლოგიის მქონე ალტერნატივების შემოწმება წინასწარ განსაზღვრული კრიტერიუმების მიხედვით. ნეირონების აქტივაციის ფუნქცია არის ერთ-ერთი ძირითადი ფაქტორი, რომელიც განსაზღვრავს ნეიროქსელების სწავლების სიჩქარეს. მისი შერჩევა ამოცანის სპეციფიკაზეა დამოკიდებული.

ნეირონული ქსელების სინთეზის უმთავრესი კომპონენტი და პრობლემა არის სწავლების ალგორითმის შერჩევა. ქსელების სწავლება შეიძლება მრავალი ალგორითმის საშუალებით, ქსელების სხვადასხვა არქიტექტურისათვის შემუშავებულია სწავლების სხვადასხვა ალგორითმები, რომელთაც აქვს სპეციფიკური უპირატესობები და ამავე დროს ნაკლიც [1,8,11]. ყველაზე ხშირად ნეიროქსელების სწავლებისათვის გამოიყენება შეცდომის კორექციის ალგორითმები, აგრეთვე უკუგავრცელების ალგორითმი და მისი მოდიფიკაციები.

განხილული მოსაზრებების საფუძველზე ჩამოყალიბებული ნეირონული ქსელის სინთეზის პროცედურების ელემენტებია:

ბლოკი 1. მონაცემთა მომზადება.

1.1. მონაცემთა ბაზის შედგენა გადასაწყვეტი ამოცანის დამახასიათებელი მაგალითებისაგან;

1.2. მონაცემთა ბაზის დაყოფა სასწავლო და საკონტროლო სიმრავლეებად, აგრეთვე ვალიდაციური სიმრავლის ფორმირება;

ბლოკი 2. მონაცემთა წინასწარი დამუშავება.

2.1. ამოცანის შესაბამისი ნიშანთა სივრცის შერჩევა, მონაცემთა წარმოდგენა რიცხვითი ფორმით და მათი ნორმირება;

2.2 გამომავალი სიდიდეების (შედეგის) კოდირების სისტემის შერჩევა.

ბლოკი 3. ნეირონული ქსელის არქიტექტურის აგება, მისი სწავლება და ქსელის ხარისხის შეფასება.

3.1. ნეირონული ქსელის სტრუქტურისა და ტოპოლოგიის შერჩევა (შრეების რაოდენობა, ნეირონების რაოდენობა შრეში, კავშირები ნეირონებს შორის);

3.2. ნეირონების აქტივაციის ფუნქციის შერჩევა;

3.3. ნეირონული ქსელის სწავლების ალგორითმის შერჩევა;

3.4. ნეირონული ქსელის ფუნქციონირების კრიტერიუმების შერჩევა (სწავლების სიჩქარე, ამოცნობის დრო, გადაწყვეტილებათა სიზუსტე, სწავლების პროცესის გამჭვირვალობა და სხვ.);

3.5. ნეირონული ქსელის მუშაობის ხარისხის შეფასება შერჩეული კრიტერიუმებით.

ბლოკი 4. ნეირონული ქსელის არქიტექტურის რედაქტირება (სტრუქტურის, ტოპოლოგიის, ნიშანთა სივრცის კორექტირება).

4.1. ნეირონული ქსელის იმ ალტერნატივის საბოლოო ამორჩევა, რომელიც უზრუნველყოფს საუკეთესო შედეგებს;

4.2. ნეირონული ქსელის მუშაობის ხარისხის საბოლოო შეფასება.

ბლოკი 5. ნეირონული ქსელის პრაქტიკული გამოყენება და დიაგნოსტიკა.

5.1.გადაწყვეტილებათა მიღებაზე სხვადასხვა ფაქტორების გავლენის ევრისტიკული შეფასება;

5.2. ამოცანის გადაწყვეტის მოთხოვნილი სიზუსტის უზრუნველყოფა, რისთვისაც საჭიროების შემთხვევაში უნდა მოხდეს მეორე ეტაპზე დაბრუნება.

3.9. ნეირონული ქსელის სინთეზის მაგალითი

ხელოვნური ნეირონული ქსელის სინთეზის მეთოდიკის ილუსტრირების მიზნით, მაგალითად მოვიყვანოთ ქართული სტილიზებული ანბანის

ამოცნობის ავტომატიზაციის ამოცანის გადაწყვეტის პროცედურები.

ანბანის სიმბოლოების ამოცნობის ამოცანა დაისმება შემდეგნაირად: მოცემულია რასტრული შავ-თეთრი სტილიზებული ქართული ანბანის სიმბოლო, საჭიროა განისაზღვროს ანბანის 33 სიმბოლოდან რომელია ამოსაცნობად წარმოდგენილი. ჩამოვაყალიბოთ ამოცანის მოცემულობა ნეირონული ქსელების ტერმინებით, ფორმულირება იქნება:

მოცემულია ნეირონულ ქსელში შემავალი **N** განზომილების ბინარული ვექტორი, საჭიროა **N** შესასვლელის და 33 გამოსასვლელის მქონე ნეირონული ქსელის სინთეზი. ამასთან, გამომავალი ვექტორის ელემენტები შეესაბამება ანბანის სიმბოლოებს და მაქსიმალური მნიშვნელობის მქონე ვექტორის ელემენტის ნომერი განსაზღვრავს ამოსაცნობ სიმბოლოს.

სტილიზებული ქართული ანბანის სკანირებული სიმბოლო არის რასტრული შავ-თეთრი გამოსახულება, ანუ ბინარულელემენტებიანი მატრიცა:

$$X = \|x_{hv}\| \qquad \forall x_{hv} \in \{0,1\},$$

სადაც **X** არის N განზომილების მქონე რეალიზაციის ბინარული ვექტორი ანუ მატრიცა; $h = \overline{1,H}$ - სტრიქონის ნომერი; $v = \overline{1,V}$ - სვეტის ნომერი; H - ჰორიზონტალური განზომილების მნიშვნელობა; V- ვერტიკალური განზომილების მნიშვნელობა. ექსპერიმენტული კვლევების შედეგად შერჩეულ იქნა ნიშანთა სივრცის პარამეტრები: H = 22 პიქსელი და V = 27 პიქსელი, მაშინ რეალიზაციის ვექტორის განზომილება იქნება N =22*27=594. პარამეტრების ასეთი შერჩევა განპირობებულია შემდეგი მოსაზრებით: ერთი მხრივ, რეალიზაციის ვექტორის განზომილება უნდა იყოს იმდენად დიდი, რომ არ დაიკარგოს ქართული ანბანის

ასოების სპეციფიკური წანაზარდები - კბილები, აგრეთვე არ მოხდეს პრიმიტივების ზედმეტი მიახლოება, ხოლო მეორე მხრივ – იმდენად მცირე, რომ რასტრში არ იყოს არაინფორმაციული სტრუქტურები.

ამოცანის ნეიროქსელური ტერმინებით ფორმულირება განაპირობებს შედეგის (ქსელიდან გამომავალი სიგნალების) კოდირების სისტემის შერჩევას, ანუ გამომავალი შრის 33 ნეირონი ბუნებრივად განსაზღვრავს ანბანის თითოეული ასოსათვის უნიკალურ გამომავალ ვექტორს.

ნეიროქსელის არქიტექტურის აგებისათვის შეირჩევა დესტრუქციული მეთოდი, ხოლო ქსელი განისაზღვრება როგორც ორშრიანი პირდაპირი გავრცელების მიმდევრობითკავშირებიანი ნეიროქსელი, მაშინ 594 ნეირონი შეადგენს პირველ შრეს და 33 ნეირონი - მეორე შრეს. ფორმალური ნეირონი წარმოადგენს მაკ-კალოკისა და პიტსის ნეირონს საფეხურისებრი აქტივაციის ფუნქციით. სწავლება განხორციელდება შეცდომის კორექციის წესის ალგორითმის მიხედვით. სასურველია, რომ თითოეული სიმბოლოსათვის სასწავლო და საკონტროლო ნაკრებების რეალიზაციების რაოდენობა გაცილებით მეტი იყოს სინაფსური წონების რაოდენობაზე. მეთოდიკის ყველა რეკომენდაციის შესრულებით შესაძლებელია ამოცნობის მაღალი საიმედოობის უზრუნველყოფა.

თავი IV. სახეთა ამოცნობის პროცესების მოდელირება ხელოვნური ნეირონული ქსელებით

ნეირონული ქსელების სინთეზის უმთავრესი კომპონენტი არის სწავლების ალგორითმი. მრავალშრიანი ნეირონული ქსელების სწავლების პროცესები, როგორც ზემოთ იყო აღნიშნული, დაკავშირებულია გარკვეულ სიძნელეებთან, ვინაიდან გამოსასვლელი შრის გარდა უცნობია სხვა შრეების აქსონებზე სიგნალების მნიშვნელობები, ხოლო მთლიანად ნეიროქსელის სწავლების პროცესების განხორციელება მხოლოდ უკანასკნელი შრის გამოსასვლელებზე შეცდომების სიდიდეების განსაზღვრითა და მათ მიხედვით ქსელის პარამეტრების კორექტირებით უაღრესად რთულია. პრობლემის გადაჭრას ცდილობენ სხვადასხვა (მათ შორის ზემოთ აღწერილი) მეთოდებით, მაგრამ ყველა ცნობილი მეთოდი ძნელადფორმალიზებადი ამოცანის გადაწყვეტისას მოითხოვს მეტად დიდი მოცულობის რუტინულ გამოთვლებს, ამიტომ სწავლების პროცესი შეიძლება მიუღებლად დიდხანს გაგრძელდეს, ან საერთოდ ვერ დასრულდეს. ლოკალური მინიმუმის ”მახეები” ემუქრება ყველა ალგორითმს, რომელიც აგებულია მინიმუმის ძიების პრინციპზე. ამ პრობლემის დაძლევას ცდილობენ სწავლების სტოქასტიკური მეთოდებით, რომელთა პრაქტიკული გამოყენება აგრეთვე დაკავშირებულია დიდ სირთულეებთან.

ცხადია, რომ აქტუალურია ხელოვნური ნეირონული ქსელების სწავლების განსხვავებული მეთოდებისა და პროცედურების შემუშავება და შესწავლა. ქვემოთ განხილულია სახეთა ამოცნობის პროცესების მოდელირება ხელოვნური ნეირონული ქსელებით, სადაც გამოყენებულია ნეირონული ქსელის სწავლების განსხვავებული, ე. წ. ავტონომიური მეთოდი [4]. წარმოდგენილი მიდგომა არ ეფუძნება მინიმუმის ძიების პრინციპს და ამდენად გვერდს უვლის ზემოთ აღწერილ საფრთხეებს, ამავე დროს უზრუნველყოფს ამოცნობის

მაღალ საიმედოობას და სწავლების პროცესების განხორციელებას მისაღებ დროში პერსონალურ კომპიუტერზე.
ამ თავში გადმოცემული მასალის სრულფასოვანი ათვისება მკითხველისაგან მოითხოვს როგორც ნეირონული ქსელების თეორიის, ასევე სახეთა ამოცნობის თეორიის განსაზღვრებების, დებულებებისა და მეთოდების ცოდნას.

4.1. ნეირონული ქსელის სწავლების ავტონომიური მეთოდი

ხელოვნური ნეირონული ქსელი, რომელიც ფორმალური ნეირონების ერთობლიობას წარმოადგენს, ერთმანეთსა და გარემოს უკავშირდება ნეირონების პარამეტრებით განსაზღვრული კავშირებით. ფორმალური ნეირონის სკალარული გამოსასვლელი OUT და მისი **X** ვექტორული შესასვლელი დაკავშირებულია თანაფარდობით:

$$OUT = F\,(NET\,(x,w)),$$

$$NET\,(x,w) = \Sigma_n\, x_i w_i + ZNE_i, \quad i \in I, \quad n \in N, \qquad (4.1)$$

აქ და შემდგომში მიღებულია აღნიშვნები:
A არის ამოსაცნობ სახეთა სიმრავლე; A_i - i-ური სახე ამოსაცნობ სახეთა სიმრავლიდან, $i \in I$; **X** - ამოსაცნობი სახის რეალიზაციათა სიმრავლე ანუ ფორმალურ ნეირონში შემავალი ამოსაცნობი ვექტორების სიმრავლე; N - ფორმალურ ნეირონში შემავალი ვექტორის ანუ ნიშანთა სივრცის განზომილება; x_n - ამოსაცნობი სახის რეალიზაციის (შემავალი ვექტორის) ელემენტი, $n \in N$; NE_i - i - ური ფორმალური ნეირონი; ZNE_i - i - ური ფორმალური ნეირონის ზღურბლის მნიშვნელობა; NET - შემავალი ვექტორის (რეალიზაციის) ელემენტების შეწონილი ჯამი, ანუ ნეირონის ამჯამავი ელემენტის

გამომავალი სიგნალი; OUT - ფორმალური ნეირონის გამომავალი სიგნალი, ნეირონის სკალარული გამოსასვლელი; w_n - სინაფსური წონითი კოეფიციენტი, $n \in N$; F - არაწრფივი გარდაქმნა (აქტივაციის ფუნქცია); $\mathbf{X}_i[m_i]$ - i -ური სახის m -ური რეალიზაცია, $m_i \in M_i$ $M_i = \mathrm{Card}\ \{\mathbf{X}_i\}$; $ET\mathbf{X}_i[m_i]$ - i -ური სახის m -ური რეალიზაციის ნეირონული ეტალონური აღწერა; $SETW_i$ - i - ური ამოსაცნობი სახის საწყისი ნეირონული აღწერა; ETW_i - i - ური ამოსაცნობი სახის საბოლოო ნეირონული ეტალონური აღწერა;

h – ქართული ასომთავრული სიმბოლოთი აღინიშნება ფორმალური ნეირონისათვის შემავალი ვექტორის (რეალიზაციის) წარდგენის ოპერატორი.

მიღებული აღნიშვნებით ნეირონისათვის მიმდინარე რეალიზაციის წარდგენისა და რეალიზაციის ელემენტების შეწონილი ჯამის მიღების პროცედურა ჩაიწერება შემდეგი სახით:

$$X_i \ \mathbf{h}\ NE_i \Rightarrow NET_i, \qquad (4.2)$$

$$NET_i = \Sigma_n x_{ni} w_{ni}, \quad i \in I, \quad n \in N. \qquad (4.3)$$

ნეირონის ზღურბლის მნიშვნელობის გათვალისწინებით (ზღურბლური ოპერაციის შედეგად) ფორმალური ნეირონის სკალარული გამოსასვლელი OUT იქნება:

$$NET_i = \Sigma_n x_{ni} w_{ni} \geq ZNE_i \Rightarrow OUT_i = 1, \quad i \in I, \quad n \in N. \qquad (4.4)$$

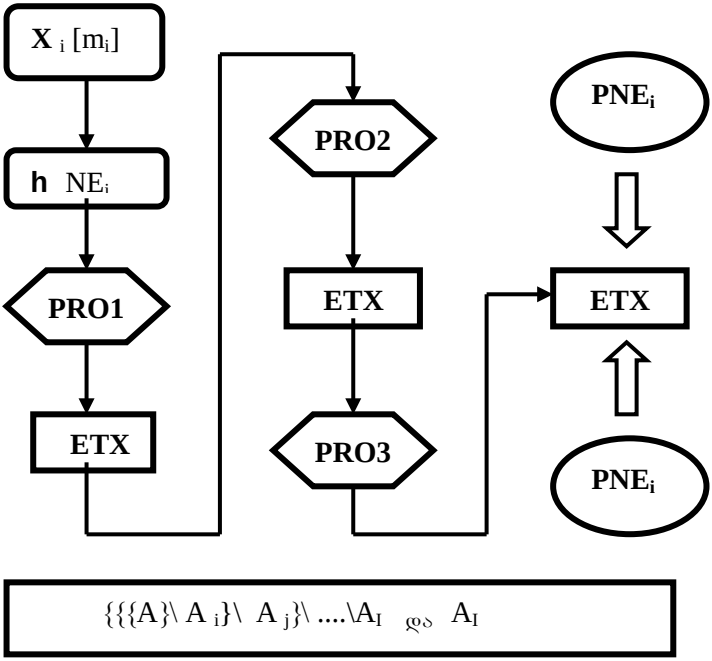
$$NET_i = \Sigma_n x_{ni} w_{ni} < ZNE_i \Rightarrow OUT_i = 0, \quad i \in I, \quad n \in N. \qquad (4.5)$$

ანუ ფორმალური ნეირონი შეიძლება იყოს აგზნებული - (4.4) გამოსახულება, ან დამუხრუჭებული - (4.5) გამოსახულება. (4.4) ფუნქციის ოპტიმუმის განსაზღვრა ზღურბლისა და N პარამეტრის მიხედვით ანუ ნიშანთა

სივრცის იმ განზომილების შემთხვევაში ($N \approx 10^3 - 10^8$), რაც სახეთა ამოცნობის ამოცანებში გვხვდება, თეორიაში ცნობილი მეთოდებით პრაქტიკულად შეუძლებელია სწავლების ალგორითმებისა და პარამეტრების მიმართ მკაცრი შეზღუდვების დაწესების გარეშე, რაც თავის მხრივ უარყოფითად მოქმედებს დასმული ამოცანის გადაწყვეტის ხარისხზე. აქედან გამომდინარე პრობლემის გადაწყვეტის მიზნით ნაშრომში [4] შემოთავაზებულია ფორმალური ნეირონების ავტონომიური სწავლების მეთოდი, რომელიც საშუალებას გვაძლევს სწავლების პროცესების სტრატეგია აიგოს ისე, რომ მისი განხორციელების შედეგად მივიღოთ გარკვეული კრიტერიუმების მიხედვით საუკეთესო სახეთა ნეირონული ეტალონური აღწერები (N პარამეტრით), რაც უზრუნველყოფს დასმული ამოცანის გადაწყვეტას და ამავე დროს სწავლებისა და ამოცნობის პროცესების განხორციელების მაღალ საიმედოობასა და ოპერატიულობას.

ნეირონების სწავლების ორგანიზაციის ზოგადი სტრატეგია. ფორმალური ნეირონების ავტონომიური სწავლების პროცესების ორგანიზაციის ზოგადი სტრატეგია ნაჩვენებია 4.1. სურათზე. ხელოვნური ნეირონული ქსელების სწავლების წარმოდგენილი სტრატეგია ეფუძნება ცალკეული ფორმალური ნეირონების ავტონომიური სწავლების პროცედურებს იმ მიზნით, რომ მათ გაარჩიონ ერთმანეთისაგან სახეთა ამოცნობის თეორიაში ცნობილი “საკუთარი” სახისა და “უცხო” სახეების რეალიზაციები. ამის მისაღწევად გამოყენებულია დიხოტომიის პრინციპი, რომელიც უზრუნველყოფს გადაწყვეტილებათა მიღებას შემდეგი თანმიმდევრობით: $\{A\}$ სახეთა სიმრავლიდან გამოიყოფა ამოსაცნობი A_i სახე და ყველა დანარჩენი – ანუ პირობითად მეორე სახე - $\{A\}\setminus A_i$, შემდეგ პირობითად მეორე სახე $\{A\}\setminus A_i$ ისევ იყოფა ორ სახედ – $\{\{A\}\setminus A_i\}\setminus A_j$

და A_j სახეებად და ა.შ. კვლევის მიზანია შევადგინოთ ამოსაცნობი ერთი სახის ნეირონული ეტალონური აღწერა გარკვეული კრიტერიუმების დაკმაყოფილებით, თუ ამ პროცედურას ჩავატარებთ დიხოტომიის პრინციპის გამოყენებით თითოეული ცალკე აღებული სახისათვის, მაშინ მივიღებთ თითოეული ამოსაცნობი სახის ნეირონულ ეტალონურ აღწერებს. თითოეული ამოსაცნობი სახისათვის ქვემოთ განსაზღვრული კრიტერიუმების დაკმაყოფილებითა და გარკვეული პროცედურების (**PRO1, PRO2, PRO3**) ეტაპობრივად შესრულების შედეგად ფორმირდება სამი ტიპის ნეირონული ეტალონური აღწერა. იმისათვის, რომ მივაღწიოთ A_i სახის სასწავლო ნაკრების ყველა რეალიზაციის უშეცდომო ამოცნობას შემოთავაზებულია



სურ. 4.1.

NE_i ფორმალური ნეირონისათვის ერთი და იგივე $[m_i]$ რეალიზაციის წარდგენის **h** პროცედურის მრავალჯერადი შესრულება (სანამ მიიღწევა მოცემული რეალიზაციის უშეცდომო ამოცნობა), ამის შემდეგ გადავდივართ იგივე A_i სახის მომდევნო რეალიზაციაზე. კერძოდ, პირველ ეტაპზე არჩეული NE_i ფორმალური ნეირონისათვის $X_i[m_i]$ რეალიზაციის მრავალჯერადი წარდგენით ხდება თითოეული რეალიზაციის შესაბამისი $ETX_i[m_i]$ - i -ური სახის m -ური რეალიზაციის ნეირონული ეტალონური აღწერის ფორმირება, ანუ ამ ეტაპზე ფორმირდება $M_i = Card\ \{X_i\}$ რაოდენობის ნეირონული ეტალონური აღწერა, მეორე ეტაპზე ამ { ETX } სიმრავლიდან ფორმირდება

მოცემული სახის ერთი $SETW_i$ საწყისი ნეირონული აღწერა. მესამე ეტაპზე საწყისი ნეირონული აღწერის, აგრეთვე მოცემული A_i სახის მინი ნეიროპორტრეტისა (PNE_i) და მაქსი ნეიროპორტრეტის (DNE_i) დახმარებით მიიღება ამ სახის ETW_i საბოლოო ნეირონული ეტალონური აღწერა, რომელიც უშეცდომოდ ამოიცნობს "საკუთარი" სახის რეალიზაციებს და უკუაგდებს ყველა "უცხო" სახის რეალიზაციებს.

4.1.1. სახეთა ნეირონული ეტალონური აღწერების კრიტერიუმების განსაზღვრა

ფორმალური ნეირონისათვის ერთი სახის ნეირონული ეტალონური აღწერა წარმოადგენს სინაფსური წონების მნიშვნელობათა სიმრავლეს

$$W = w_1, w_2, \ldots w_n, \ldots w_N. \qquad (4.6)$$

განსაზღვრება 4.1. ნებისმიერი A_i სახის საწყისი ნეირონული აღწერა $SETW_i$, $i \in I$ ეწოდება ისეთ ნეირონულ ეტალონს, რომელიც აერთიანებს მოცემული სახის იმ ნიშნებს, რომლებიც უზრუნველყოფს A_i სახის სასწავლო ნაკრების ყველა რეალიზაციის უშეცდომოდ ამოცნობას:

$$\forall X_i[m_i] \ \mathbf{h} \ NE_i \Rightarrow NET_i(X_i[m_i]) \Rightarrow OUT_i = 1, \qquad (4.7)$$

იმისათვის, რომ მივაღწიოთ A_i სახის სასწავლო ნაკრების ყველა რეალიზაციის ამოცნობას უშეცდომოდ შემოთავაზებულია NE_i ფორმალური ნეირონისათვის ერთი და იგივე $[m_i]$ რეალიზაციის წარდგენის **h** პროცედურის შესრულება რამდენჯერმე (სანამ მიიღწევა მოცემული რეალიზაციის ამოცნობა უშეცდომოდ), ამის შემდეგ გადავდივართ იგივე A_i სახის მომდევნო რეალიზაციაზე. ანუ ეს **h** პროცედურა

ანალოგიურია იმისა, როდესაც ადამიანს ასწავლიან რაიმე სიმბოლოების ამოცნობას თითოეული სიმბოლოს რამდენჯერმე წარდგენით.

განსაზღვრება 4.2. ნებისმიერი A_i სახის საბოლოო ნეირონული აღწერა (შემდგომში ეტალონური აღწერა) ეწოდება ETW_i ისეთ ნეირონულ ეტალონს, რომლის სინაფსური წონების სიმრავლე უზრუნველყოფს A_i სახის ნებისმიერი რეალიზაციისათვის (4.7) პირობის შესრულებას, და ყველა დანარჩენი $\{A\}\backslash A_i$ სახის სასწავლო ნაკრების ნებისმიერი რეალიზაციისათვის შემდეგი (4.8) პირობის შესრულებას:

$$X_j[m_j] \;\; \mathbf{h} \;\; NE_i \Rightarrow NET_i(X_j[m_j]) \Rightarrow OUT_i = 0,$$
$$\forall \mathbf{X}_j[m_j] \in \{X\}\backslash\{X_i\}, \quad j \in J, \; j \neq i, \quad m_j \in M_j \qquad (4.8)$$

გამოსახულებები (4.7) და (4.8) წარმოადგენს სახეთა საწყისი და საბოლოო ნეირონული ეტალონური აღწერების ფორმირების კრიტერიუმებს.

4.1.2. სახეთა ნეირონული ეტალონური აღწერების ფორმირება

სახეთა საწყისი და საბოლოო ნეირონული ეტალონური აღწერების ფორმირების კრიტერიუმების ჩამოყალიბების შემდეგ უნდა მოხდეს სახეთა ნეირონული ეტალონური აღწერების ფორმირება საკუთარი სასწავლო ნაკრების რეალიზაციების მიმართ. სახეთა ნეირონული ეტალონური აღწერების ფორმირების პროცესში ზღურბლური კოეფიციენტი არ იცვლება - $ZNE_i = const$. ზღურბლური კოეფიციენტის შერჩევის კრიტერიუმად შეგვიძლია გამოვიყენოთ ზოგადი დებულება დაპროგრამების მოხერხებულობის შესახებ, ანუ

ზღურბლური კოეფიციენტი უნდა იყოს ისეთი, რომ განხორციელებული იტერაციების რაოდენობა დიდი არ იყოს და ამავე დროს იძლეოდეს საშუალებას შევადგინოთ საკმარისად ზუსტი სახეთა ნეირონული ეტალონური აღწერები. საწყის ეტაპზე ზღურბლური კოეფიციენტი მივიღოთ $\forall\ ZNE_i = N/2$.

დავუშვათ მოცემულია ნეირონის სწავლების საწყისი, ანუ ნულოვანი ბიჯის პარამეტრები - სინაფსური წონები: $W_i[0] = w_{1i}, w_{2i}, \ldots w_{ni}, \ldots w_{Ni}$, სადაც

$$\forall\ w_n[0] = 0, \qquad n \in N. \qquad (4.9)$$

პროცედურა **PRO1** ხორციელდება NE_i ფორმალური ნეირონისათვის $X_i[m_i] \in \{X_i\}$ რეალიზაციის წარდგენით, რის შედეგად შეიძლება მივიღოთ ორი სიტუაცია:

1. $X_i[m_i]\mathbf{h}NE_i \Rightarrow NET_i\ \mathbf{h}\ ZNE_i \Rightarrow OUT_i = 1$ (4.10)
2. $X_i[m_i]\mathbf{h}NE_i \Rightarrow NET_i\ \mathbf{h}ZNE_i \Rightarrow OUT_i = 0$ (4.11)

პირველი სიტუაციის (4.10) განხორციელების შემთხვევაში ცხადია, რომ A_i სახის საწყისი ნეირონული ეტალონური აღწერა $SETW_i$ აკმაყოფილებს მისდამი წაყენებულ (4.4) კრიტერიუმს და შესაბამისად სინაფსური წონების შეცვლა საჭირო არ არის. მეორე სიტუაციის (4.11) განხორციელების შემთხვევაში ასევე ცხადია, რომ აუცილებელია სინაფსური წონების გადაწყობის პროცედურის ჩატარება. გადაწყობის პროცედურის ჩასატარებლად საჭიროა გავაძლიეროთ ის სინაფსური წონები, რომელთა შესაბამისი რეალიზაციის ელემენტები $x_{ni} = 1$. ასეთ შემთხვევაში A_i სახის n - ური სინაფსური წონა w_{ni} შეიცვლება შემდეგი გამოსახულების მიხედვით

$$w_{ni}\ [m_i + 1] = w_{ni}\ [m_i] + \Delta w_{ni} \qquad (4.12)$$

სადაც m წარმოადგენს იტერაციის ბიჯის ნომერს და ამავე დროს გვიჩვენებს, A_i სახის მერამდენე რეალიზაცია წარედგინება ნეიროქსელს; Δw_{ni}

არის სინაფსური წონითი კოეფიციენტის ცვლილების სიდიდე. ზოგადად იმ პირობიდან გამომდინარე, რომ სასწავლო ნაკრების თითოეული რეალიზაციის მიერ სინაფსური წონითი კოეფიციენტის ცვლილებაში შეტანილი წვლილი თანაბარი უნდა იყოს, შეგვიძლია მივიღოთ:

$$\Delta w_{ni} = MAK / M_i \qquad (4.13)$$

სადაც MAK არის მანორმირებელი კოეფიციენტი, $M_i = Card\ \{\mathbf{X}_i\}$; (როგორც ცნობილია, მანორმირებელი კოეფიციენტის სიდიდე შეიძლება შეირჩეს 1, 10, 100, 1000 და ა.შ.). მისი კონკრეტული მნიშვნელობა შეიძლება დავუქვემდებაროთ დაპროგრამების მოხერხებულობას. ნულოვანი ბიჯის პირობებიდან გამომდინარე პირველ ბიჯზე გვექნება

$$w_{ni}\ [1] = w_{ni}\ [0] + \Delta w_{ni} \qquad (4.14)$$

გამოსახულების შესრულების შედეგად მივიღებთ, რომ

$$NET_i = 0, \text{ რადგან } \forall\ w_{ni} = 0, \quad n \in N.$$

აქედან გამომდინარე

$$NET_i < ZNE_i \Rightarrow OUT_i = 0 \qquad (4.15)$$

(4.15) გამოსახულებიდან გამომდინარე და (4.14) გამოსახულების მხედველობაში მიღებით სინაფსური წონითი კოეფიციენტები გაიზრდება ყველგან, სადაც $x_{ni} = 1$ და გახდება Δw_{ni} –ის ტოლი. შემდეგ იტერაციაში ვიმეორებთ იგივე $\mathbf{X}_i$ რეალიზაციის NE_i ფორმალური ნეირონისათვის წარდგენას. ფაქტობრივად იგივე სინაფსურ წონით კოეფიციენტებს მიემატება Δw_{ni} -ის ტოლი მნიშვნელობები. შესაბამისად სწავლების მეორე ბიჯზე გვექნება: $w_{ni}\ [2] = w_{ni}\ [1] + \Delta w_{ni}$, რადგან $\Delta w_{ni} > 0$, ამიტომ $w_{ni}\ [2] > w_{ni}\ [1]$, აქედან გამომდინარე $NET_i\ [2] > NET_i\ [1]$, აქ შეიძლება გვქონდეს ორი შემთხვევა:

1. $NET_i\,[2] \geq ZNE_i \Rightarrow OUT_i = 1;$
2. $NET_i\,[2] < ZNE_i \Rightarrow OUT_i = 0.$

პირველ შემთხვევაში სწავლება დამთავრებულია, ხოლო მეორე შემთხვევაში სწავლების პროცესი გაგრძელდება იგივე $\mathbf{X}_i$ რეალიზაციის კვლავ წარდგენით. ეს პროცესი გრძელდება სანამ მივიღებთ (4.4) პირობის შესრულებას, მაშინ სწავლება დამთავრებული იქნება. A_i სახის სასწავლო ნაკრების ყოველი რეალიზაციისათვის განხორციელებული (4.12) პროცედურა გვაძლევს თითოეული რეალიზაციის ნეირონულ აღწერას, რომელსაც აღვნიშნავთ $ETX_i[m_i]$ - ით, მათი რაოდენობა M_i - ის ტოლია. ცხადია, რომ M_i რაოდენობის რეალიზაციების ნეირონული აღწერებიდან უნდა შევადგინოთ ერთი $SETW_i$ საწყისი ნეირონული ეტალონური აღწერა, რომელიც A_i სახის სასწავლო ნაკრების ყოველი რეალიზაციისათვის დააკმაყოფილებს (4.4) პირობას, რასაც უზრუნველყოფს **PRO2** პროცედურა. ამისათვის თითოეული რეალიზაციის ნეირონული ეტალონური აღწერის $ETX_i[m_i]$ რასტრის ყოველი პიქსელის შესაბამისი სინაფსებისაგან უნდა ავირჩიოთ მაქსიმალური მნიშვნელობის სინაფსური წონითი კოეფიციენტები და ჩავწეროთ $SETW_i$ საწყისი ნეირონული ეტალონური აღწერის შესაბამის პიქსელის ნაცვლად, რაც ნიშნავს, რომ გვაქვს:

$$w^*_{ni} = \max \{w_{ni}[m_i]\}, \quad m_i \in M_i, \; n \in N. \qquad (4.16)$$

თეორემა 4.1. (4.16) გამოსახულებით მოცემული პროცედურის მიხედვით ამოსაცნობი სახის საწყისი ნეირონული აღწერის $SETW_i$ ფორმირების შედეგი აკმაყოფილებს (4.4) პირობას.

დამტკიცება: დავუშვათ, რომ (4.16) გამოსახულებით მოცემული პირობა სრულდება და რომ

$$NET_i = \Sigma_n x_{ni}[m_i] w_{ni}[m_i] \geq ZNE_i \qquad (4.17)$$

თუ (4.17) გამოსახულება ჭეშმარიტია ყველა $m_i \in M_i$ რეალიზაციისათვის, მაშინ თითოეული n - ური პიქსელისათვის გვექნება ერთი ოპტიმალური w^*_{ni} სინაფსის მნიშვნელობა. (4.16) პირობის მიხედვით გვაქვს, რომ $w^*_{ni} \geq w_{ni}[m_i]$. აქედან გამომდინარე (4.17) ჯამის შესაბამისი წევრისათვის შესრულდება პირობა:

$$w^*[m_i]\, x_{ni} \geq w_{ni}[m_i]\, x_{ni} \qquad (4.18)$$

რაც ნიშნავს, რომ (4.17) ჯამის ყოველი წევრი ზრდის ან არ ამცირებს ჯამის მნიშვნელობას. შესაბამისად გვაქვს

$$\Sigma_n\, w^*_{ni}[m_i]\, x_{ni}[m_i] \geq \Sigma_n\, w_{ni}[m_i]\, x_{ni}[m_i] \qquad (4.19)$$

(4.17) გამოსახულების მხედველობაში მიღებით ცხადი ხდება, რომ თეორემა 4.1. დამტკიცებულია.

ფორმალური ნეირონის ზღურბლის მნიშვნელობის შერჩევა. საწყისი ნეირონული ეტალონური აღწერისათვის საჭიროა ფორმალური ნეირონის ზღურბლის მნიშვნელობის შერჩევა. დავუშვათ (4.16) გამოსახულების მიხედვით ფორმირებული გვაქვს A_i სახის საწყისი ნეირონული ეტალონური აღწერა $SETW_i$. 4.1. თეორემით დამტკიცებულია, რომ (4.4) პირობა კმაყოფილდება შერჩეული ზღურბლისათვის, საჭიროა შევარჩიოთ ახალი ზღურბლის მაქსიმალური მნიშვნელობა იმ მიზნით, რომ სხვა A_j სახეების რეალიზაციებმა, რომელთათვისაც დაკმაყოფილდება (4.5) გამოსახულებით მოცემული პირობა, ვერ მოაგროვონ ZNE_i ზღურბლის გადასალახად საკმარისი NET_i ის მნიშვნელობები

$$\mathbf{X}_j\, \mathbf{h} \quad NE_i \Rightarrow NET_i[\mathbf{X}_j] < ZNE_i \qquad (4.20)$$

ამის მიზეზი შეიძლება იყოს სინაფსური წონების გაზრდის ფაქტი. ამ მიზნის განსახორციელებლად ფორმირებულ საწყისი ნეირონული ეტალონური აღწერას მივაწოდოთ A_i სახის სასწავლო ნაკრების რეალიზაციები და ყოველი m_i რეალიზაციისათვის

გამოვითვალით $NET_i[m_i]$. მივიღებთ შეწონილი ჯამების მნიშნელობათა სიმრავლეს $\{NET_i[m_i]\}$, რომლის $Card = M_i$. ამ სიმრავლიდან შევარჩიოთ მინიმალური ელემენტი,

$$NET^*_i = \min \{ NET_i[m_i] \} \qquad (4.21)$$

ცხადია, რომ თუ შესრულდება პირობა $ZNE_i = NET^*_i$, მაშინ A_i სახის სასწავლო ნაკრების ყველა რეალიზაცია გადალახავს ზღურბლს და გვექნება (4.10) პირობის OUT = 1 დაკმაყოფილება. ზღურბლისა და სინაფსური წონების შერჩევის პროცედურების განხორციელების შედეგად $\forall A_i$ სახისათვის მივიღებთ $SETW_i$ საწყის ნეირონული ეტალონურ აღწერებს.

4.1.3. ამოსაცნობი სახის ნეირონული მინი პორტრეტი და ნეირონული მაქსი პორტრეტი

ამოსაცნობი სახის ETW_i საბოლოო ნეირონული ეტალონური აღწერის მისაღებად "უცხო" A_j სახის რეალიზაცია წარედგინება A_i სახის $SETW_i$ საწყის ეტალონურ ნეირონულ აღწერას, ფორმირებულმა ETW_i საბოლოო ნეირონულმა ეტალონურმა აღწერამ უნდა გაარჩიოს ერთმანეთისაგან "საკუთარი" A_i და "უცხო" A_j სახის რეალიზაციები, ანუ ერთდროულად უნდა დაკმაყოფილდეს (4.7) და (4.8) პირობები, რასაც **PRO3** პროცედურა უზრუნველყოფს $SETW_i$ საწყისი ნეირონული ეტალონური აღწერის, აგრეთვე A_i სახის მინი ნეიროპორტრეტისა (PNE_i) და მაქსი ნეიროპორტრეტის (DNE_i) დახმარებით.

ამოსაცნობი სახის საბოლოო ნეირონული ეტალონური აღწერების ETW_i მისაღებად ანუ (4.8) პირობის დაკმაყოფილებისათვის საჭირო ხდება ნებისმიერი სახისათვის ისეთი ნიშნების და მათი შესაბამისი პიქსელების განსაზღვრა, რომლებიც სასწავლო ნაკრების ყველა რეალიზაციისათვის იღებენ ერთის ტოლ მნიშვნელობას.

მინი და მაქსი პორტრეტების თეორიაში [2] მოცემული განსაზღვრებიდან ცნობილია, რომ ასეთ ნიშანთა გაერთიანებას ეწოდება მინი პორტრეტი, რაც გულისხმობს ხდომილობათა სივრცის არსებობას, სადაც ნებისმიერი ნიშნისათვის (პიქსელისათვის) ითვლება, რომ მოვლენა მოხდა, თუ ეს ნიშანი იღებს ერთის ტოლ მნიშვნელობას და მოვლენა არ მოხდა, თუ ეს ნიშანი იღებს ნულის ტოლ მნიშვნელობას. მინი პორტრეტის $P_i = \{p_{ni}\}$ და მაქსი პორტრეტის $D_i = \{d_{ni}\}$ ფორმირება შესაბამის ნიშანთა სუპერპოზიციისა და რიგი ლოგიკური ოპერაციების განხორციელების შედეგად აღწერილია ნაშრომებში [12], [15].

(4.14) გამოსახულების განხორციელებით მიღებული აღწერები არ წარმოადგენს ხდომილობათა სივრცის ელემენტებს, ამიტომ თეორიაში ცნობილი მინი პორტრეტის განსაზღვრება შეუძლებელია უშუალოდ მივუსადაგოდ აქ ფორმირებულ ნეირონულ ეტალონურ აღწერებს, ამიტომ აუცილებელია აქსიომატიკურად განისაზღვროს A_i სახის ნეირონული მინი პორტრეტი და ნეირონული მაქსი პორტრეტი.

განსაზღვრება 4.1. A_i სახის ნეირონული მინი პორტრეტი PNE_i ეწოდება ისეთ N განზომილებიან ბინარულ ვექტორს (მატრიცას), რომლის ყველა ელემენტი აკმაყოფილებს შემდეგ პირობას:

$$p_{ni} = 1, \quad \text{თუ} \quad w_{ni} = w^*_{ni},$$
$$p_{ni} = 0, \quad \text{თუ} \quad w_{ni} \neq w^*_{ni}. \qquad (4.22)$$

სადაც PNE_i არის A_i სახის ნეირონული მინი პორტრეტი:

$$PNE_i = p_{1i}, p_{2i}, \ldots p_{ni}, \ldots p_{Ni}.$$

განსაზღვრება 4.2. A_i სახის ნეირონული მაქსი პორტრეტი DNE_i ეწოდება ისეთ N განზომილებიან ბინარულ ვექტორს (მატრიცას), რომლის ყველა ელემენტი აკმაყოფილებს შემდეგ პირობას:

$d_{ni} = 1$, თუ $w_{ni} > 0$,

$d_{ni} = 0$, თუ $w_{ni} = 0$. (4.23)

სადაც DNE_i არის A_i სახის ნეირონული მაქსი პორტრეტი:

$$DNE_i = d_{1i}, d_{2i}, \ldots d_{ni}, \ldots d_{Ni}.$$

მიღებული ბინარული PNE_i და DNE_i მინი და მაქსი ნეიროპორტრეტების გამოყენებით შესაძლებელი ხდება ფორმირებული $SETW_i$ საწყისი ნეირონული ეტალონური აღწერებიდან **PRO3** პროცედურით ETW_i საბოლოო ნეირონული ეტალონური აღწერების მიღება.

4.1.4. საბოლოო ნეირონული ეტალონური აღწერების ფორმირება

ETW_i საბოლოო ნეირონული ეტალონური აღწერის ფორმირება გულისხმობს $SETW_i$ საწყისი ნეირონული ეტალონური აღწერის არსებობას. დავუშვათ გვაქვს

$$SETW_i = s_{1i}, s_{2i}, \ldots s_{ni}, \ldots s_{Ni} \quad (4.24)$$

მიღებული აღნიშვნით (3.1.4) პირობას ექნება შემდეგი სახე:

$$NET_i = \Sigma_n x_{ni} s_{ni} \geq ZNE_i \Rightarrow OUT_i = 1, \quad i \in I, n \in N. \quad (4.25)$$

საბოლოო ნეირონული ეტალონური აღწერის ფორმირებისათვის, მაგალითად A_j სახის მიმართ, ზემოთ მოცემული აღნიშვნებით (4.8) გამოსახულების მაგივრად გვექნება

$$NET_i(\mathbf{X}_j[m_j]) = \Sigma_n x_{ni}[m_j] s_{ni}[m_i] < ZNE_i \Rightarrow OUT_i[m_j] = 0. \quad (4.26)$$

(4.26) გამოსახულებით მოცემული პირობის დაკმაყოფილება ნიშნავს, რომ A_i სახის საბოლოო ნეირონულმა ეტალონურმა აღწერამ უნდა განასხვაოს სხვა სახეების რეალიზაციები საკუთარი სახის რეალიზაციებისაგან. დავუშვათ, რომ A_j სახის რომელიმე m_j რეალიზაციისათვის (4.26) პირობა არ კმაყოფილდება

$$NET_i(\mathbf{X}_j[m_j]) = \Sigma_n x_{ni}[m_j]\, s_{ni}[m_i] > ZNE_i \Rightarrow OUT_i[m_j] = 1. \quad (4.27)$$

ცხადია, რომ A_i სახის ნეირონული ეტალონური აღწერის სინაფსური წონები უნდა გადავაწყოთ ისე, რომ ერთდროულად დაკმაყოფილდეს (4.25) და (4.26) პირობები. შემოვიტანოთ სინაფსური წონების გადაწყობის პროცესში ამ წონების ლიმიტირების პროცედურა, რომელიც გულისხმობს **სინაფსებისათვის ზღურბლის დაწესებას,** რომლის მიხედვით გაირკვევა, თუ რასტრის რომელი პიქსელის შესაბამისი სინაფსები განულდება და რომელი პიქსელების შესაბამისი სინაფსები გაძლიერდება განულებული სინაფსების ხარჯზე. აღვნიშნოთ ZS_i – ით სინაფსის ზღურბლი, მაშინ გვექნება

$$ZS_i = a \bullet w^*_i \qquad (4.28)$$

სადაც $a = const$, $(a \approx 0{,}3 - 0{,}9)$. განულებული სინაფსების სიმრავლე აღვნიშნოთ $\{W^0_i\}$ - ით, მაშინ $\{W^0_i\}$ É $\{W_i\}$, რომლის ელემენტები განისაზღვრება შემდეგი გამოსახულებით

$$w_{ni} \hat{I} \{W^0_i\}, \text{ თუ } 0 < w_{ni} < a \bullet w^*_i, \quad n \in N. \qquad (4.29)$$

ნეირონული აჯამვის შედეგი-შეწონილი ჯამი NET_i (4.29) გამოსახულების მხედველობაში მიღებით წარმოვადგინოთ შემდეგი გამოსახულებით

$$NET_i = NETO_i + NETH_i \qquad (4.30)$$

სადაც $NETO_i$ არის განულებული სინაფსური წონები, $NETH_i$ - დანარჩენი (არაგანულებული) სინაფსური წონები. თუ შევასრულებთ განულების ოპერაციას – $NETO_i = 0$, მაშინ (4.30) პირობის დასაკმაყოფილებლად საჭიროა $NETO_i$ -ს საწყისი მნიშვნელობა გადანაწილდეს (4.30) ჯამის მეორე წევრზე ($NETH_i$). ჯამის მეორე წევრი შედგება ცალკეული შესაკრებებისაგან, რომლებიც შეიძლება დავყოთ ორ ნაწილად:

$$NETH_i = NETH_i^{max} + NETH_i^{oth} \qquad (4.31)$$

სადაც $NETH_i^{max}$ შედგება A_i სახის PNE_i მინი ნეიროპორტრეტის იმ სინაფსური წონებისაგან, რომელთაც გააჩნიათ w_{ni}^*-ის მნიშვნელობები (ნამდვილი რიცხვები), $n \hat{I} N$, ხოლო $NETH_i^{oth}$ შედგება A_i სახის PNE_i მინი ნეიროპორტრეტის ყველა დანარჩენი სინაფსური წონების მნიშვნელობებისაგან.

თეორემა 4.2.

იმისათვის, რომ A_i და A_j სახეების სასწავლო ნაკრების რეალიზაციებისათვის ერთდროულად კმაყოფილდებოდეს (4.25) და (4.26) პირობები აუცილებელი და საკმარისია, რომ A_i სახის მინი ნეიროპორტრეტის დამატება A_j სახის მაქსი ნეიროპორტრეტით $PNE_i \setminus DNE_j$ არ იყოს ცარიელი სიმრავლე (სურ. 4.1).

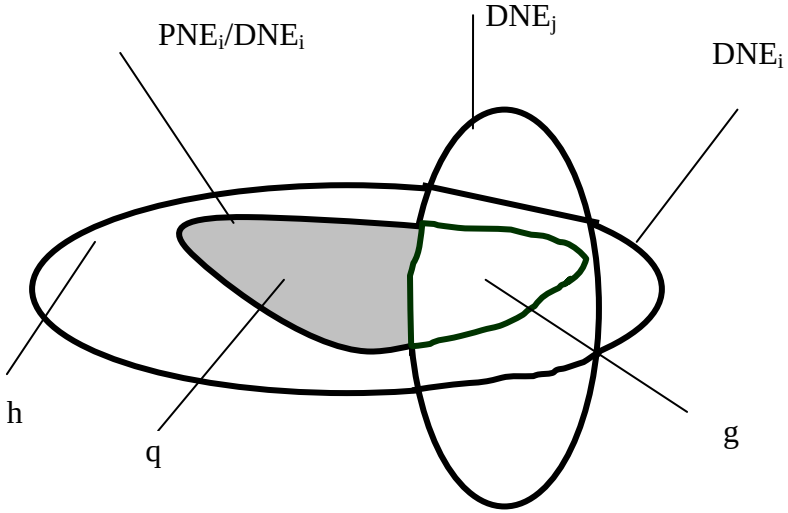
აღნიშნული პირობის ფორმალიზებული ასახვა იქნება:

$$PNE_i \setminus PNE_j \neq \emptyset \qquad (4.32)$$

დამტკიცება:

<u>აუცილებლობის პირობა:</u>

უტოლობა (4.25) ჭეშმარიტი დარჩება, თუ მისი მარცხენა წევრის ჯამის შესაკრებებს შევცვლით ტოლფასი შესაკრებებით. მაგალითად, რაიმე პირობის მიხედვით განულებული ჯამის წევრების მნიშვნელობებს მივუმატებთ ჯამის სხვა წევრების მნიშვნელობებს.



სურ. 4.1

განსაზღვრება 4.3. სინაფსური წონითი კოეფიციენტების ნულგადაწყობა ეწოდება (4.25) ან (4.26) უტოლობების ჯამის წევრების განულებას და მათი მნიშვნელობების მიმატებას ჯამის სხვა წევრებისათვის. ცხადია, რომ სინაფსური წონითი კოეფიციენტების ნებისმიერი ნულგადაწყობა 4.2 თეორემის თანახმად უნდა განხორციელდეს ისე, რომ ჭეშმარიტი იყოს (4.25) და (4.26) უტოლობები.

განვიხილოთ (4.25) უტოლობის მარცხენა ნაწილი, კერძოდ ჯამი წარმოვადგინოთ ორი შესაკრების სახით და დავუშვათ, რომ (4.26) უტოლობა არ არის ჭეშმარიტი

$$\Sigma_n\, x_{ni}\, s_{ni} = \Sigma_q\, x_{qi}\, s_{qi} + \Sigma_h\, x_{hi}\, s_{hi} \geq ZNE_i \qquad (4.33)$$

სადაც q ინდექსით მოცემულია მინი ნეიროპორტრეტის შემადგენელი პიქსელების (შესაბამისი სინაფსური

წონების) ნუმერაცია, ხოლო h ინდექსით მოცემულია რასტრის ყველა დანარჩენი პიქსელის ნუმერაცია. მინი ნეიროპორტრეტის პიქსელების შესაბამისი სინაფსური წონების გაზრდა სხვა პიქსელების მნიშვნელობების ხარჯზე ყოველთვის დააკმაყოფილებს (4.25) უტოლობას, რადგან მინი ნეიროპორტრეტის 4.1. განსაზღვრების თანახმად ეს პიქსელები, ანუ მათი შესაბამისი სინაფსები ყოველთვის რეაგირებენ A_i სახის საწავლო ნაკრების რეალიზაციებზე. განულებული პიქსელების სინაფსური წონების ჯამის ნულგადაწყობა ნიშნავს, რომ

$$s_{hi} \neq 0, \qquad \forall h \in H,$$

სადაც H არის განულებული პიქსელების რაოდენობა, ხოლო

$$s^{+}_{qi} = s_{qi} + \Delta s_h, \ \Delta s_h = \Sigma_h x_{hi} s_{hi} / Q, \ Q = \text{const}. \qquad (4.34)$$

სადაც s^{+}_{qi} არის სინაფსური წონების გაზრდის შედეგად მიღებული მნიშვნელობები, Q - მინი ნეიროპორტრეტის პიქსელების რაოდენობა. ცხადია, რომ

$$\Sigma_n x_{ni} s_{ni} = \Sigma_q x_{qi} s^{+}_{qi} \qquad (4.35)$$

რადგან $\Sigma_h x_{hi} s_{hj} = 0$, ნულგადაწყობის პირობის (განსაზღვრება 4.3) თანახმად. აქედან გამომდინარე (4.25) გამოსახულებით მოცემული პირობა გადაწყობის შემდეგაც სრულდება. ამრიგად აუცილებლობის პირობა დამტკიცებულია.

განვიხილოთ (4.25) უტოლობის მარჯვენა მხარე. ისევე როგორც წინა შემთხვევაში, ჯამი დავყოთ რამდენიმე ნაწილად

$$\Sigma_n x_{ni} s_{nj} = \Sigma_q x_{qj} s_{qi} + \Sigma_g x_{gj} s_{gi} + \Sigma_h x_{hj} s_{hi}, \qquad (4.36)$$

(4.36) გამოსახულებაში q და g ინდექსები არის მინი ნეიროპორტრეტის გარკვეული სინაფსური წონების

ნუმერაცია, ხოლო h ინდექსი არის დანარჩენი სინაფსური წონების ინდექსაცია. ამასთან q და g ინდექსების დასაზუსტებლად განვიხილოთ A_i სახის მინი ნეიროპორტრეტის დამატება A_j სახის მაქსი ნეიროპორტრეტით $PNE_i \setminus DNE_j$ (სურ. 4.1). დაშტრიხული ნაწილით სურათზე მოცემულია $PNE_i \setminus DNE_j$ დამატების შესაბამისი არე. ამ არის შესაბამისი ინდექსი (4.36) გამოსახულებაში მოცემულია q ინდექსით, ხოლო A_i სახის მინი ნეიროპორტრეტისა და A_j სახის მაქსი ნეიროპორტრეტის საერთო ნაწილის (თანაკვეთის) არე g ინდექსით. ცხადია, რომ ნულგადაწყობის ოპერაციას ექვემდებარება (4.36) გამოსახულების ჯამის მხოლოდ მესამე წევრი - $\Sigma_h\ x_{hj}\ s_{hi}$, კერძოდ შესაძლებელია h ინდექსიანი სინაფსური წონების განულება, ხოლო სინაფსური წონების გაზრდის (მიმატების) ოპერაცია უნდა განვახორციელოთ მხოლოდ q ინდექსიან სინაფსურ წონებზე, რადგან g ინდექსიანები საერთოა A_i და A_j სახეებისათვის, ამიტომ მათი გაზრდა არ მოგვცემს სასურველ ეფექტს. ზემოთმოყვანილი მსჯელობიდან გამომდინარე, h ინდექსიანი სინაფსური წონების განულების შედეგად მივიღებთ:

$$\Sigma_n\ x_{nj}\, s_{ni} = \Sigma_q\ x_{qj}\, s^{+}_{qi} + \Sigma_g\ x_{gj}\, s_{gi}, \qquad (4.37)$$

ნულგადაწყობის პირობის თანახმად $s_{hi} = 0, \quad \forall h \in H$ და შესაბამისად $\Sigma_h\ x_{hj}\ s_{hi} = 0$. ამ ოპერაციით საერთო ჯამი არ შეიცვლებოდა, რომ არა შემდეგი გარემოება: q ინდექსიანი სინაფსური წონები ვერ გაზრდის X_j რეალიზაციით მიღებულ (4.26) გამოსახულების ჯამს, რადგან ამ პიქსელების შესაბამისი სინაფსური წონები q ინდექსისათვის ნულის ტოლია. აქედან გამომდინარე ფაქტობრივად გვაქვს, რომ:

$$\Sigma_h\ x_{hj}\, s^{+}_{hi} = 0 \qquad (4.38)$$

და შესაბამისად ასეთ შემთხვევაში (3.36) გამოსახულების მიხედვით გვექნება:

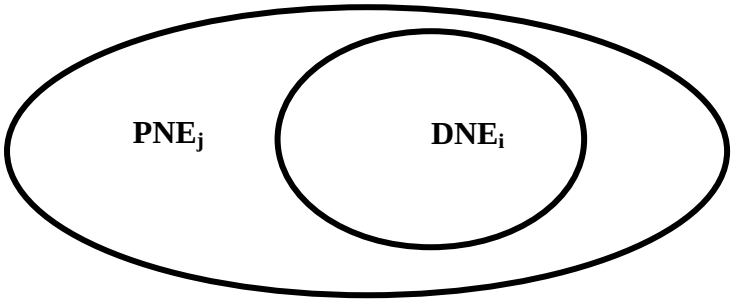
$$\Sigma_n\, x_{nj}\, s_{ni} = \Sigma_g\, x_{gj}\, s_{gi} < ZNE_i$$

რაც ნიშნავს, რომ (4.33) გამოსახულებით მოცემული პირობა შეიძლება დაირღვეს და ჭეშმარიტი გახდეს (4.26) პირობა, თუ (4.26) პირობა მაინც არ შესრულდა, მაშინ გადაწყობის ოპერაციას შევასრულებთ იმდენჯერ, სანამ ZNE_i ზღურბლის გაზრდა შეგვეძლება შემდეგი პირობის შესასრულებლად:

$$ZNE_i = \min_{mi} \{\Sigma_n \; s_{ni} x_{ni} [m_j]\} > \max_{mj} \{\Sigma_n \; s_{ni} x_{nj} [m_j]\} \qquad (4.39)$$

<u>საკმარისობის პირობა</u>

დავუშვათ, რომ (4.33) გამოსახულებით მოცემული პირობა ჭეშმარიტია, ხოლო $PNE_i \setminus DNE_j = \emptyset$, რაც გამოსახულია 4.2. სურათზე. ცხადია, რომ ასეთ შემთხვევაში (4.36) გამოსახულებაში q ინდექსიანი სინაფსური წონები ($PNE_i \setminus DNE_j$) არ იარსებებს, რის გამოც არ იარსებებს განულებული (h ინდექსიანი) სინაფსური წონების ნულგადაწყობის შესაძლებლობა, ანუ ვერ შევძლებთ (4.36) გამოსახულების შეცვლას, რაც გამოიწვევს (4.26) პირობის შესრულების შეუძლებლობას. ამრიგად საკმარისობის პირობაც დამტკიცებულია, შესაბამისად თეორემა 4.2. დამტკიცებულია.



სურ. 4.2

განვაზოგადოთ 4.2. თეორემა A სიმრავლის ყველა ელემენტის საბოლოო ნეირონული აღწერის ფორმირებისათვის. A_i სახის მინი ნეიროპორტრეტის დამატება A_j სახის მაქსი ნეიროპორტრეტით $PNE_i \setminus DNE_j$ აღვნიშნოთ PNE_i (J) – ით. მიღებული ვექტორის დამატება A_k სახის მინი ნეიროპორტრეტით აღვნიშნოთ PNE_i (J, k) – ით და ა. შ. თუ I არის სახეების რიცხვი, მაშინ A_i სახის მინი ნეიროპორტრეტის დამატება ყველა დანარჩენი სახეების მაქსი ნეიროპორტრეტის მიმართ წარმოვადგინოთ შემდეგი გამოსახულებით:

$$PNE_i (1,2, 3, \ldots, I - 1) = PNE_i (\bullet) \qquad (4.40)$$

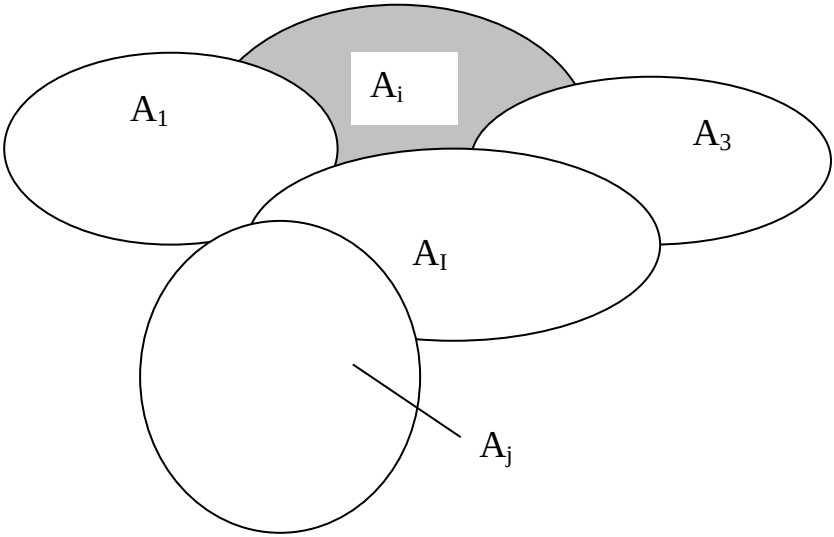
თეორემა 4.3. ნებისმიერი სახის საბოლოო ნეირონული ეტალონური აღწერა ფორმირებადია (4.2) განსაზღვრების მიხედვით, თუ (4.40) გამოსახულებით განსაზღვრული PNE_i (•) სიმრავლე არ არის ცარიელი სიმრავლე, ანუ $PNE_i (\bullet) \neq \emptyset$.

დამტკიცება: 4.1. თეორემით განსაზღვრულია A_i სახის მინი ნეიროპორტრეტის ის სინაფსური წონები, რომლებისთვისაც გადაწყობის ოპერაციის განხორციელება არ არღვევს (4.25) და (4.26) პირობებს. სანამ არსებობს A_i სახის მინი ნეიროპორტრეტის დამატება ნებისმიერი რაოდენობის სახეების მინი ნეიროპორტრეტებით, მანამდე გვაქვს q ინდექსით განსაზღვრული სინაფსური წონები, რომელთა გადაწყობით შეიცვლება A_i სახის ნეიროეტალონური აღწერა ისე, რომ არ დაირღვევა (4.25) და (4.26). უტოლობებით მოცემული პირობები, წინააღმდეგ შემთხვევაში გადაწყობის ოპერაცია დაარღვევს (4.25) ან (4.26) პირობას ერთად ან ცალ-ცალკე.

ამით 4.3. თეორემა დამტკიცებულია.

წარმოდგენილი ფორმალური ნეირონების ავტონომიური სწავლების მეთოდის თეორიული და ექსპერიმენტული გამოკვლევებით მიღებული შედეგები შესაძლებლობას იძლევა სინთეზირებული იქნეს გადასაწყვეტი ამოცანების ადექვატური დისკრეტული

ხელოვნური ნეირონული ქსელები და განხორციელდეს მათი ეფექტური სწავლება.



სურ. 4.3

სწავლების მეთოდის ავტონომიურობა გამოიხატება იმით, რომ ცალკეული ფორმალური ნეირონის სწავლება მიმდინარეობს სხვა ნეირონებისაგან დამოუკიდებლად, ხოლო სწავლების წარმატებას უზრუნველყოფს სინაფსური წონების ზრდაზე შეზღუდვის დაწესება. სწავლების ასეთი სქემა, ჩვენი აზრით, უფრო შეესაბამება ბიოლოგიურ ნეირონში მიმდინარე პროცესებს. მეთოდის მეორე განმასხვავებელი მომენტია სწავლების აქტიური პროცედურები, რაც რეალიზებულია ნეირონების სწავლების პროცესებში სახეთა ამოცნობის თეორიაში ცნობილი სხვადასხვა პროცედურების გამოყენებით (დიხოტომიის პრინციპი, მინი და მაქსი პორტრეტების მეთოდი და სხვა), რის შედეგადაც ამოსაცნობი რეალიზაციების სასწავლო ნაკრები გარდაიქმნება ახალ,

მოდიფიცირებულ სასწავლო ნაკრებად, რომელიც უზრუნველყოფს ნეირონული ქსელის უფრო ეფექტურ სწავლებას. აღსანიშნავია ფორმალური ნეირონების ავტონომიური სწავლების მეთოდის თვალსაჩინოება და გამჭვირვალობის მაღალი ხარისხი სწავლებისა და ამოცნობის პროცესების განხორციელების ყველა ეტაპზე. თეორიული კვლევის შედეგები აპრობირებული იქნა ქართული ნაბეჭდი სიმბოლოების ამოსაცნობად ნეირონული ეტალონური აღწერების მისაღებად ექსპერიმენტული გამოკვლევებისა და ნეირონული ქსელების სწავლების პროცესების ორგანიზებისათვის. ამოცნობის შედეგები თვალსაჩინოდ ადასტურებს ნეირონული ქსელების ავტონომიური სწავლების მეთოდისა და მიღებული ნეირონული ეტალონური აღწერების ეფექტურობასა და პერსპექტიულობას, როგორც საიმედოობის ასევე ოპერატიულობის თვალსაზრისით.

ნეირონული ქსელების ავტონომიური სწავლების მეთოდის გამოყენება შესაძლებელია ყველა იმ ობექტისათვის, რომლის რეალიზაციები წარმოდგენილია თანაბარგანზომილებიანი ბინარული ვექტორებით.

ლიტერატურა

1. ჩოგოვაძე რ., ხუროძე რ. „ნეიროინფორმატიკის საფუძვლები“ გამომცემლობა„ტექნიკური უნივერსიტეტი“, თბილისი, 2005.
2. ვერულავა ო., ხუროძე რ. “ამომცნობი სისტემების თეორიის საფუძვლები”. გამომცემლობა „ტექნიკური უნივერსიტეტი“, თბილისი, 2001.
3. ჩოგოვაძე რ. ნეირონული ქსელების სინთეზი გადაწყვეტილებათა მიღების პროცესებისათვის. საერთაშორისო სამეცნიერო კონფერენციის ”მართვისა და ენერგეტიკის პრობლემები PCPE-2004,“ მოხსენებათა კრებული № 8. თბილისი, 2004.
4. ჩოგოვაძე რ. ხელოვნური ნეირონული ქსელების სწავლების ერთი მეთოდის შესახებ. თბილისი, სტუ-ს შრომები №1(459), 2006.
5. ვერულავა ო., ჩოგოვაძე რ., ხუროძე რ. მსგავსების ზომის ფორმირება სიმრავლეებს შორის. საერთაშორისო სამეცნიერო კონფერენციის ”მართვისა და ენერგეტიკის პრობლემები PCPE-2004,“ მოხსენებათა კრებული № 8. თბილისი, 2004.
6. ჩოგოვაძე რ. ხელოვნური ნეირონული ქსელების სინთეზი სახეთა ამოცნობის პროცესებისათვის. თბილისი, სტუ-ს შრომები №№(454), 2004.
7. ნანეიშვილი თ. „ქცევის ფსიქოფიზიოლოგია″. გამომცემლობა „მეგობარი“, თბილისი, 1996.
8.Горбань А.Н. Россиев Д.А. Нейронные сети на персональным компьютере. Новосибирск, Наука,1996.
9.Горбань А. Н. Обучение нейронных сетей. Москва, СП ПараГраф, 1990.
10.Круглов В.В., Дли М.И. Голунов Р.Ю. Нечеткая логика и исскуственные нейронные сети. Москва. 2001.
11.Уоссермен Ф. Нейрокомпьютерная техника: теория и практика. Москва. Мир, 1992.

12.Чоговадзе Р. А., Верулава О. Г. О возможности использования ассоциативного параллельного процессора для распознавания образов в реальном масштабе времени. Труды НПО ЭЛВА, часть I, Тбилиси, 1973.
13.Чоговадзе Р. А. Трехмерный ассоциативный параллельный процессор // Сб. Вопросы кибернетики. Академия наук СССР. Научный совет по комплексной проблеме «Кибернетика». Москва, 1973.
14. Poznyak A., Sanchez E., Wen Yu. Dynamic Neural Networks for Nonlinear Control: Identification, State Estimation and Trajectory Tracking. World Scientific, Mexico, 2001.
15. Chogovadze R., Khurodze R., Verulava O., Decisions making modeling on basis of Mini and Maxi portraits by neural networks. Georgian technical university, Transactions, Vol.3 (453), Tbilisi, 2004.

შინაარსი

წინასიტყვაობა.....4

თავი I. ხელოვნური ნეირონი და ნეირონული ქსელები

1.1. ნეიროკომპიუტინგის პრობლემის მოკლე მიმოხილვა 9

1.1.1. ნეიროკომპიუტინგით გადასაწყვეტი ამოცანების პროტოტიპები 11

1.2. ხელოვნური ნეირონული ქსელების აქსიომატიკა 14

1.3. ფორმალური ნეირონის სტრუქტურა და თვისებები 17

1.4. ერთშრიანი და მრავალშრიანი ხელოვნური ნეირონული ქსელები 20

1.5. ხელოვნური ნეირონული ქსელების კლასიფიკაცია 23

თავი II. ხელოვნური ნეირონული ქსელების სწავლება

2.1. ხელოვნური ნეიროქსელების სწავლების პრობლემა 26

2.2. ნეირონული ქსელების სწავლების ალგორითმები 31

2.3. პერსეპტრონული ნეირონი და მისი სწავლება 33

2.4. უწყვეტი ნეირონული ქსელების დელტა-წესით სწავლების ალგორითმი 36

2.5. პირდაპირი გავრცელების მრავალშრიანი ნეირონული ქსელები 40

2.6. სწავლების უკუგავრცელების ალგორითმი 41

2.7. უკუგავრცელების ალგორითმის დაჩქარების მეთოდები 51

2.8. შემხვედრი გავრცელების ქსელები 55

თავი III. ხელოვნური ნეირონული ქსელების სწავლების სტოქასტიკური მეთოდები

3.1. ლოკალური მინიმუმების პრობლემა 66

3.2. სწავლების ბოლცმანის სტოქასტიკური მეთოდი 70
3.3. სწავლების კოშის მეთოდი 72
3.4. ჰოპფილდის ნეირონული ქსელები 74
3.4.1. უკუკავშირებიანი ნეირონული ქსელების კონფიგურაციები 75
3.5. ასოციაციური მეხსიერება 79
3.6. ორმხრივმიმართული ასოციაციური მეხსიერება 82
3.7. დამახსოვრებული ასოციაციების აღდგენა 86
3.4. ხელოვნური ნეირონული ქსელების სინთეზის პროცედურები 88
3.9. ნეირონული ქსელის სინთეზის მაგალითი 92

თავი IV. სახეთა ამოცნობის პროცესების მოდელირება ხელოვნური ნეირონული ქსელებით
4.1. ნეირონული ქსელის სწავლების ავტონომიური მეთოდი 95
4.1.1. სახეთა ნეირონული ეტალონური აღწერების კრიტერიუმების განსაზღვრა 100
4.1.2. სახეთა ნეირონული ეტალონური აღწერების ფორმირება 102
4.1.3. ამოსაცნობი სახის ნეირონული მინი პორტრეტი და ნეირონული მაქსი პორტრეტი 108
4.1.4. საბოლოო ნეირონული ეტალონური აღწერების ფორმირება 110
ლიტერატურა 120
შინაარსი 122