

საქართველოს ტექნიკური უნივერსიტეტი

ელგუჯა ყუბანეიშვილი

ხელოვნური ნეირონული ქსელები მედიცინაში

ლექციების კურსი



თბილისი 2013

შესავალი

შეიძლება ითქვას, რომ ნეირონული ქსელების ისტორია დაიწყო მაშინ, როცა ადამიანი დაინტერესდა აზროვნების პრობლემით. არსებობს მრავალი ჰიპოთეზა ადამიანის აზროვნების ბუნების შესახებ. ამ საკითხების განხილვა, რომელიც მიმდინარეობდა ფილოსოფოსებსა, ფიზიოლოგებსა და ფსიქოლოგებს შორის მოიტანა მცირე შედეგი, რადგან თვით ეს პრობლემა, შესწავლის თვალსაზრისით, საკმაოდ რთულია.

ადამიანის ნერვული სისტემის ფუნქციებისა და სტრუქტურების გაძლიერებულმა კვლევებმა ნეირობიოლოგები და ნეიროანატომები მიიყვანა გარკვეულ პროგრესთან, კერძოდ თავის ტვინის ნეირონებში სიგნალების გავრცელების თვალსაზრისით. მაგრამ, თითქმის უცნობი რჩებოდა ნერვული სისტემის ფუნქციონირების საკითხები. დაგროვილი ცოდნის საფუძველზე დადგინდა, რომ ასეული მილიარდი ნეირონი, რომლებიც დაკავშირებული არიან ერთმანეთთან, ქმნიან სისტემას, რომელიც თავისი შესაძლებლობით დადასტურებულია თანამედროვე სუპერკომპიუტერებს.

ნეირონების ფუნქციონირების და მათი კავშირების უკეთ შესწავლამ მკვლევარებს საშუალება მისცა შეექმნათ ნეირონის მათემატიკური მოდელები თავიანთი თეორიების შესამოწმებლად. გაირკვა, რომ ეს მოდელები არა მარტო იმეორებენ თავის ტვინის ზოგიერთ ფუნქციებს, არამედ შეუძლიათ შეასრულონ გარკვეული ფუნქციები. აქედან გამომდინარე, ჩამოყალიბდა და დღემდე რჩება ნეირონის მოდელირების ორი მიზანი, რომლებიც ერთმანეთს ავსებენ. პირველი მათგანია ფიზიოლოგიის და ფსიქოლოგიის დონეზე ადამიანის ნერვული სისტემის ფუნქციონირების გაგება და მეორე – შეიქმნას გამოთვლითი სისტემები (ხელოვნური ნეირონული ქსელები), რომლებიც შეასრულებენ თავის ტვინის ფუნქციების მსგავს ქმედებებს.

ტერმინი „ნეირონული ქსელები“ წარმოიშვა მე-20 საუკუნის 40-ან წლებში. ამ სფეროში ძირითადი შედეგები დაკავშირებულია უ. მაკკალოხის, უ. პიტსის, დ. ჰების, ფ. როზენბლანტის, მ. მინსკის, დ. ჰოპფილდის და სხვა მეცნიერების კვლევებზე.

პირველი ხელოვნური ნეირონის მოდელი 1943 წ. უ. მაკკალოხმა და უ. პიტსმა შექმნეს და ჩამოაყალიბეს თავის ტვინის ფუნქციონირების თეორიის ძირითადი დებულებები.

1949წ. დ. ჰებმა გამოთქვა იდეა თავის ტვინის ნეირონების შეერთების შესახებ და ჩამოაყალიბა ნეირონის სწავლების კანონი, რომელიც შემდგომ გადაიქცა ხელოვნური ნეირონული ქსელების სწავლების სასტარტო წერტილად.

1957წ. ფ. როზენბლანტმა დაამუშავა პერსეპტრონის ორგანიზების და ფუნქციონირების პრინციპები და შემოგვთავაზა ტექნიკური რეალიზაცია მსოფლიოში პირველი ნეიროკომპიუტერის *MARK-* ის სახით.

1959წ. გამოქვეყნდა მ. მინსკის და ს. პეპერტის ნაშრომი „პერსეპტრონები“, სადაც ნაჩვენებია იყო პერსეპტრონების (ლაპარაკია ერთშრიან პერსეპტრონზე) გამოყენების შეზღუდვებზე, რამაც გამოიწვია ხელოვნური ნეირონული ქსელების მიმართ ინტერესის ორი ათეული წლით გაქრობა. მიუხედავად ამისა, ზოგიერთი მეცნიერი (კოჰონენი, გროსბერგი, ანდერსონი და სხვები) შეუპოვრად განაგრძობდნენ კვლევებს და უკვე 80-იანი წლების დასაწყისში მივიდნენ იმ დასკვნამდე, რომ მრავალშრიანი პერსეპტრონის საშუალებით შესაძლებელია მრავალი პრაქტიკული ამოცანის გადაწყვეტა და მათ შორის ისეთების, რასაც ერთშრიანი პერსეპტრონი ვერ ახერხებდა.

მეცნიერების აქტიობამ აფეთქებითი ხასიათი მიიღო 1987წ. როდესაც ჩატარდა ხელოვნურ ნეირონულ ქსელებში ოთხი უმსხვილესი თათბირი და 500-ზე მეტმა სამეცნიერო ნაშრომების გამოქვეყნებამ, რამაც გამოიწვია პრობლემის ფენომენალური ზრდის სიჩქარე.

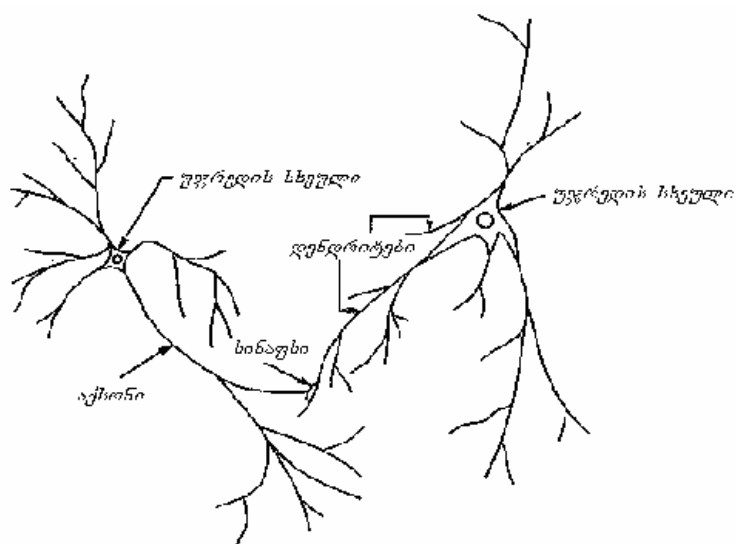
გაკვეთილი, რომელიც შეიძლება მივიღოთ ნეირონული ქსელების განვითარების ისტორიიდან გამოისახება კლარკის კანონით, რომელიც ჩამოაყალიბა მწერალმა და მეცნიერმა არტურ კლარკმა. ამ კანონით მტკიცდება, რომ თუ დიდი საპატიო მეცნიერი ლაპარაკობს, რომ რაიმე შეიძლება შესრულდეს, მაშინ ის თითქმის ყოველთვის მართალია. თუ მეცნიერი ამბობს, რომ ეს შეუძლებელია შესრულდეს, მაშინ ის თითქმის ყოველთვის ცდება. მეცნიერების განვითარების ისტორია წარმოადგენს შეცდომების და ნაწილობრივ სიმატლის მატანეს. ის რაც დღეს ეჭვს არ იწვევს, ხვალ შეიძლება იგნორირებული იყოს.

ფაქტების არაკორექტულმა აღქმამ შეიძლება გამოიწვიოს სამეცნიერო კვლევების პარალიზირება. მინსკის ბრწყინვალე სამეცნიერო ნაშრომმა დაამუხრუჭა ხელოვნური ნეირონული ქსელების განვითარება. მაგრამ, ეჭვს არ იწვევს ის, რომ ამით ნეირონული ქსელების განვითარება, დაუსაბუთებული ოპტიმიზმით და საჭირო თეორიული ბაზის უქონლობით, არ დაზარალებულა.

1. ბიოლოგიური ნეირონი და მისი მათემატიკური მოდელი

1.1. ბუნებრივი ნეირონი და მასში მიმდინარე პროცესები

ნეირონი წარმოადგენს ნერვული სისტემის უჯრედს და სხვა უჯრედებისაგან იმით განსხვავდება, რომ მას გააჩნია მთელი რიგი გამოთვლითი ფუნქციები. მიუხედავად ნეირონის მრავალფეროვნებისა ისინი ძირითადად შედგებიან ერთი და იგივე დანიშნულების ელემენტებისაგან: უჯრედის სხეული (სომა), დენდრიტები და აქსონები. სქემატურად ნეირონი წარმოდგენილია შემდეგ ნახაზზე:



ცნობილია, რომ მოზრდილი ადამიანის ტვინის ნეირონები არ აღდგებიან, ისინი კვდებიან. ეს იმას ნიშნავს, რომ ნეირონის ყველა კომპონენტი უწყვეტლევ უნდა იცვლებოდეს, ხოლო მასალა უნდა განახლდეს საჭიროების შემთხვევაში.

უჯრედის სხეული ანუ სომა, რომელიც მოქცეულია მემბრანულ გარსში, არეგულირებს უჯრედში მიმდინარე პროცესებს. იგი მართავს ნეირონის ენერგიის ხარჯვას და არეგულირებს უჯრედში მიმდინარე მრავალ სხვა პროცესებს. უჯრედის სხეულის გარე მემბრანას გააჩნია ნერვული იმპულსების (მოქმედების პოტენციალები) გენერირების უნიკალური უნარი, რომელიც წარმოადგენს ნერვული სისტემის სიცოცხლისათვის საჭირო ფუნქციას და მისი გამოთვლითი უნარის ცენტრს. სომა თავისი ფორმის მიხედვიდან გამოდინარე არჩევენ ათასობით ნეირონის ტიპს, რომელთაც გააჩნიათ ჩვეულებრივ 5-დან 100 მიკრომეტრის სიგრძის დიამეტრი.

სომას დანიშნულებაა შემოსული იმპულსების აკუმულირება (აჯამვა). თუ სომას პოლარიზაციის პროცესი უფრო მზარდია ვიდრე დეპოლარიზაციისა, მაშინ აჯამვის გარკვეული ხარისხის მიღწევისას იგი განიმუხტება და აქსონში ჩნდება შესაბამისი სიგნალი. განმუხტვის შემდეგ ნეირონს გარკვეული დროის განმავლობაში აღარ შეუძლია შემოსულ სიგნალზე რეაგირება. დროის ამ მონაკვეთს, რომელიც შეიძლება გაგრძელდეს წამის მეასედებიდან რამდენიმე წამამდე, ლატენცური პერიოდი ეწოდება. ლატენცური პერიოდის გავლის შემდეგ ნეირონი მოდის საწყის „მუშა“ მდგომარეობაში.

დენდრიტების დანიშნულებაა გარემოდან მიღებული სიგნალების გატარება და გარდაქმნა. გარდაქმნაში იგულისხმება სიგნალის დამუხრუჭება ანუ შემცირება, ან გაზრდა. დენდრიტების საშუალებით გარემოდან ან სხვა ნეირონებიდან გამომავალი სიგნალები მოხვდებიან სომაში. ამრიგად, დენდრიტები წარმოადგენენ განშტოებად სტრუქტურას, რომლებიც გამოდიან სომიდან და მათზე განლაგებულია სინაფსური შენაერთები, რომლებიც დებულობენ სხვა აქსონებიდან სიგნალებს. დენდრიტების მიერ სიგნალების გატარების პროცესი დამოკიდებულია სინაფსურ კავშირებზე და მისი დიამეტრის ცვლილებებზე. ასე მაგალითად, დიამეტრის შემცირება შეიძლება იწვევდეს სიგნალის გაზრდას, ხოლო დიამეტრის გადიდება – შემცირებას ანუ დამუხრუჭებას.

აქსონი წარმოადგენს განუტოტებელ ნერვულ ბოჭკოს, რომლის დანიშნულებაა ნეირონის მიერ გამოიმუშავებული სიგნალის გატარება. აქსონს გააჩნია განშტოებები, რომელთა საშუალებით იგი უკავშირდება სხვა აქსონებს, დენდრიტებს და მათი საშუალებით სხვა ნეირონებს. აქსონში სიგნალი გვაქვს იმ შემთხვევაში, როდესაც სომა განიმუხტება. იმპულსების ამპლიტუდა განმუხტვის ძალის მიუხედავად მუდმივია, მაგრამ იცვლება მათი სიხშირე. დენდრიტებში მისი დიამეტრის ცვლილებებით და სინაფსების საშუალებით ხდება აკრძალვის, არჩევის და თანაკვეთის ოპერაციები, ანუ ლოგიკური „და“, „ან“ და „არა“ მოქმედებები. აქსონი შეიძლება იყოს მოკლე (0,1 მმ) და გრძელი, რომლის სიგრძე აღემატება 1მ და რომელიც განთავსებულია ადამიანის სხეულის სხვა ნაწილებში.

აქსონის დაბოლოებას გააჩნია უამრავი განშტოება, რომელთა დაბოლოება წარმოადგენს სინაფსებს, საიდანაც სიგნალი დენდრიტების საშუალებით გადაეცემა სხვა ნეირონებს, ხოლო ზოგიერთ შემთხვევაში სომას.

სინაფსს გააჩნია სფეროსებრივი ფორმა, სადაც განთავსებულია ბუშტულები, რომლებიც შეიცავენ ნეიროტრანსმიტერულ მოლეკულების დიდ რაოდენობას. როდესაც იმპულსი გადის აქსონში ზოგიერთი მისი ბუშტულები გამოანთავისუფლებენ თავის შემადგენლობას სინაფსურ ხვრელში, რომლებსაც

იჭერენ დენდრიტის სპეციალური რეცეპტორები და ხდება მათი დანერგვა უჯრედის სხეულში. არსებობს 30-ზე მეტი სახის ნეიროტრანსმიტერები, რომელთაგან ზოგი წარმოადგენს ამგზნებს და იწვევენ უჯრედის აგზნებას და ზოგიერთი გამოიმუშავებს გამოსავალ იმპულსებს. სხვები წარმოადგენენ დამუხრუჭების ნეიროტრანსმიტერებს, რომლებიც ცდილობენ იმპულსის შემცირებას ან ჩაქრობას.

სომა შემოსულ სიგნალებს აჯამავს და თუ აჯამული სიგნალი ზღურბლურ სიგნალზე მეტია, გამოიმუშავებს იმპულსს, რომელიც აქსონის გავლით გადაეცემა სხვა ნეირონებს.

დასკენის სახით ჩამოვსავლით შემდეგი:

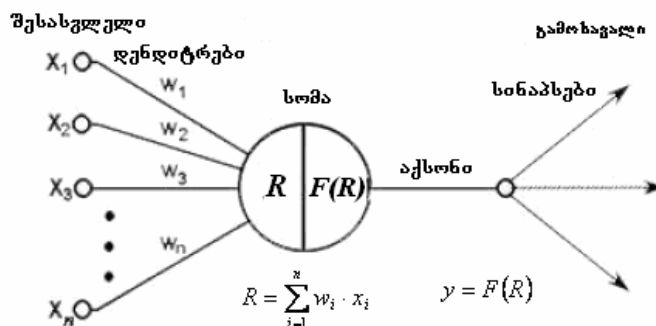
1. ფუნქციონალურად ნეირონი სიგნალებს ღებულობს შემავალი არხების ანუ დენდრიტების საშუალებით გარემოდან ან სხვა ნეირონებისაგან სინაფსების საშუალებით. აქედან სიგნალები მოხვედებიან სომაში.

2. სომაში მოხვედრილი სიგნალები აიჯამებიან სხვა ანალოგურ სიგნალებთან ერთად და როდესაც ჯამური სიგნალი გადააჭარბებს გარკვეულ ზღურბლურ მნიშვნელობას მოხდება სომას განმუხტვა და ნეირონის გამოსავალზე ანუ აქსონზე გაჩნდება იმპულსი.

3. ნეირონის მიერ გამოიმუშავებულ სიგნალს აქსონი გაატარებს და მის ბოლოებზე არსებული სინაფსების საშუალებით გადასცემს სხვა ნეირონებს, რომლებიც თავის მხრივ შეიძლება აღიგზნონ ან პირიქით დამუხრუჭდნ. სინაფსს გააჩნია გარკვეული ინტენსიობა ანუ წონა, რომელიც შეესაბამება ნეირონის სინაფსურ აქტიობას.

1.2. ხელოვნური ნეირონის მოდელი

ყოველი ნეირონი, თავის ტვინის ნეირონის ანალოგიურად, ხასიათდება თავისი მიმდინარე მდგომარეობით, რომელიც შეიძლება იყოს აგზნებული ან დამუხრუჭებული. მას გააჩნია სინაფსების ჯგუფი ანუ ერთმიმართულებიანი შემავალი კავშირები, რომლებიც შეერთებულნი არიან სხვა ნეირონების გამოსასვლელებთან. გარდა ამისა მას გააჩნია აქსონი – მოცემული ნეირონის გამოსასვლელი კავშირი, რომლის საშუალებითაც სიგნალი (აგზნებული ან დამუხრუჭებული) მიეწოდება შემდეგი ნეირონების სინაფსებს. ხელოვნური ნეირონის მოდელი ასე შეიძლება წარმოვადგინოთ:



პირველი მიახლოებით ხელოვნური ნეირონი, რომელსაც ზოჯერ მათემატიკურ ნეირონსაც უწოდებენ, წარმოადგენს ბუნებრივი ნეირონის თვისებების იმიტაციას. ხელოვნური ნეირონის შესავალზე შედის გარკვეული

რაოდენობის სიგნალების ერთობლიობა $x_1, x_2, \dots, x_n$, რომელთაგან თითოეული მათგანი წარმოადგენს სხვა ნეირონის გამოსავალს. ეს შემომაველი სიგნალები შეესაბამებიან ბიოლოგიური ნეირონის სინაფსებში გამავალ სიგნალებს. თითოეული სინაფსი ხასიათდება სინაფსური კავშირის სიდიდით ან მისი წონის w_j კოეფიციენტით.

თითოეული შემავალი სიგნალი x_j მრავლდება შესაბამის წონით w_j კოეფიციენტებზე და იჯამება აჯამვის $\sum$ ბლოკში (სუმატორი). თითოეული წონითი კოეფიციენტი შეესაბამება სინაფსური კავშირის ე.წ. „ძალას“ და მათი საშუალებით ხდება ნეირონების აგზნების რეგულირება. აჯამვის ბლოკი შეესაბამება სომას და თუ მისგან გამოსულ სიგნალს ავლნიშავთ R სიმბოლოთი, მაშინ გვექნება:

$$R = \sum_i w_i \cdot x_i = W'X. \quad R = \sum_{i=1}^n w_i \cdot x_i \quad y = F(R)$$

სუმატორიდან მიღებული R სიგნალი, რომელიც ახასიათებს ნეირონის რეაქციას, მიეწოდება აქტივაციურ ანუ მახასიათებელ (გადაცემის) F ფუნქციას. თუ ამ ფუნქციით გამოთვლილი მნიშვნელობა, რომელიც წარმოადგენს ნეირონის გამოსავალ სიგნალს, რაიმე T ზღვრულზე მეტია, მაშინ ნეირონის გამოსავალზე ვდებულობთ სიგნალს. თუ ნეირონის გამოსავალ სიგნალს ავლნიშავთ Y სიმბოლოთი, მაშინ გვექნება:

$$Y = F(W'X - T) = F(R - T).$$

დასკვნის სახით ჩამოვყალიბოთ შემდეგი ცნებები:

1. ნეირონის მიმდინარე მდგომარეობა განისაზღვრება ფორმულით:

$$Y_i = \sum_{j=1}^n w_{ij} x_j + T_i, \quad (2.1)$$

სადაც x_j , $j=1,2,\dots,n$ შემავალი სიგნალებია,

w_{ij} – წონითი კოეფიციენტებია,

T_i – ზღვრული მნიშვნელობაა.

2. დადებითი წონის კოეფიციენტები შეესაბამებიან სინაფსების აგზნებას, ხოლო უარყოფითები – დამუხრუჭებას. თუ $w_{ij} = 0$, მაშინ კავშირი i -ურ და j -ურ ნეირონებს შორის არ არსებობს.

3. აქტივაციური ფუნქციით მიღებულ სიგნალს ნეირონი გარდაქმნის გამოსავალ $Y_i = F(R_i)$ სიგნალად. ზოგადად, აქტივაციის ფუნქცია წარმოადგენს არაწრფივ ფუნქციას, რომლითაც ხდება აგზნების გატარების პროცესის მოდელირება.

1.3. აქტივაციური ფუნქციების სახეები

აქტივაციური ფუნქცია სხვადასხვა სახისაა და იგი შეიძლება ჩაითვალოს ხელოვნური ნეირონის არაწრფივ მაძლიერებელ მახასიათებლად. გაძლიერების კოეფიციენტი განისაზღვრება, როგორც ფარდობა Y სიგნალის ნაზრდისა R სიგნალის ნაზრდზე. არსებობს აქტივაციის ფუნქციების ნაირსახეობა. განვიხილოდ ზოგიერთი მათგანი, რომლებიც უფრო ხშირად გამოიყენებიან პრაქტიკაში.

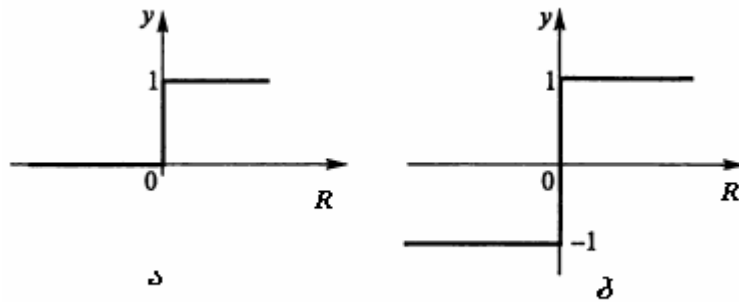
ზღურბლური აქტივაციური ფუნქციები. თუ ნეირონის ზღურბლურ მნიშვნელობას ანუ ნეირონის მგრძნობიარობას აღვნიშნავთ T -თი, მაშინ გვექნება:

$$R = \sum_i w_i \cdot x_i - T$$

ნეირონის გამოსავალ Y სიგნალს შეუძლია მიიღოს ერთი ან ნულის ტოლი მნიშვნელობა შემდეგნაირად:

$y = 1$, როცა $R \geq T$ და $y = 0$, როცა $R < 0$.

აქტივაციის ფუნქცია მიიღებს შემდეგ სახეს (ნახ. 2.2 ა):



ნახ. 2.2

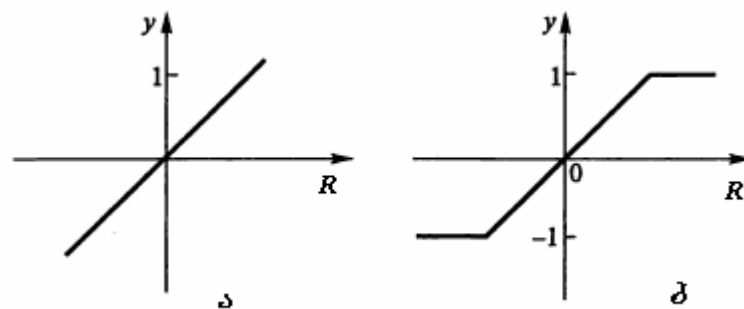
და მას ერთეულოვანი ნახტომის (ხვეისაიდის) ფუნქცია ეწოდება. გარდა ამისა, ნეიროქსელებში ხშირად გამოიყენება კოორდინატთა სათავეს მიმართ სიმეტრიული ზღურბლური მნიშვნელობა (ნახ.2.2 ბ):

$y = +1$, როცა $R > 0$ და $y = -1$, როცა $R \leq 0$

ასეთ აქტივაციურ ფუნქციას ნიშნის (კიბისეული) ფუნქცია ეწოდება.

წრფივი აქტივაციური ფუნქციები. ქვემოთ მოყვანილ ნახაზზე (ნახ.2.3 ა) წარმოდგეილია შეუზღუდავი ცვალებადობის $Y = R$ აქტივაციური ფუნქციის გრაფიკული გამოსახულება. წრფივი აქტივაციური ფუნქცია შეიძლება იყოს შეზღუდული ცვალებადობის (ნახ. 2.3 ბ):

$y = -1$, როცა $R < -1$, $y = R$, როცა $-1 \leq R \leq 1$, $y = 1$, როცა $R > 1$.

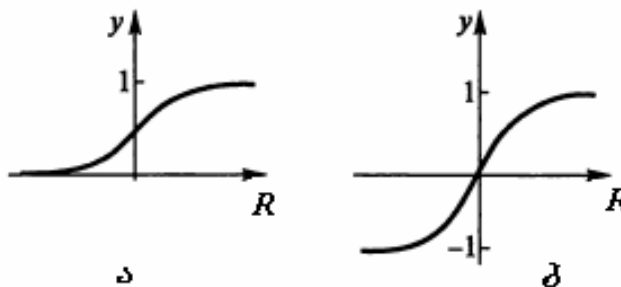


ნახ. 2.3

სიგმაიდალური ფუნქცია. პრაქტიკაში უფრო ხშირად გამოიყენება სიგმაიდალური (სიგმა) ფუნქცია. (ფუნქციას სიგმაიდალური უწოდეს იმის გამო, რომ გრაფიკულად იგი წააგავს „S“ ასოს). სიგმაიდალური აქტივაციური ფუნქცია მათემატიკურად ასე გამოისახება:

$$Y = \frac{1}{1 + \exp\{-aR\}}.$$

ხშირად ამ ფუნქციას სტანდარტულს უწოდებენ და როცა $(R - T) \geq 0$, მაშინ $Y = 0$. გრაფიკულად სიგმა ფუნქციას აქვს შემდეგი სახე:



ნახ. 2.4

თუ სიგმაიდალური ფუნქციის გამოსახულებაში a კოეფიციენტს შევამცირებთ, მაშინ ფუნქცია ხდება უფრო დამრეცი და როცა $a = 0$ იგი 0,5 ღონეზე გადაიქცევა პორიზონტალურ წრფედ. a -ს გაზრდისას ($R=0$) წერტილში სიგმაიდალური ფუნქცია გარკვეული ფორმით უახლოვდება ერთეულოვანი ნახტომის ფუნქციას. სიგმაიდალური ფუნქციის გამოსახულებიდან ჩანს, რომ ნეირონის გამოსავალი მნიშვნელობა იმყოფება $(0,1)$ დიაპაზონში.

სიგმაიდალური ფუნქციის ერთ-ერთ ღირებულებას წარმოადგენს მისი პირველი რიგის წარმოებულის მარტივი სახე:

$$y' = aF(R)[1 - F(R)]$$

უნდა აღინიშნოს, რომ სიგმაიდალური ფუნქცია დიფერენცირებადია მთელი აბსცისის ღერძის მიმართ. ამ თვისების გამო იგი ხშირად გამოიყენება ნეიროქსელების სწავლების პროცედურებში. გარდა ამისა, მას გააჩნია უნარი გააძლიეროს დაბალი სიგნალები უფრო ძლიერად, ვიდრე დიდი სიგნალები, რაც იწვევს დიდი სიგნალებით გაჯერების თავიდან აცილებას.

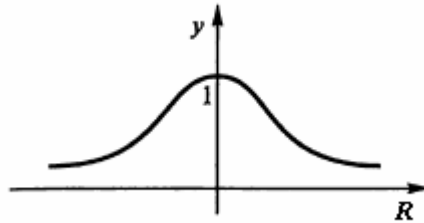
ჰიპერბოლური ტანგენსის ფუნქცია. გარდა სიგმაიდალური ფუნქციისა პრაქტიკაში ხშირად გამოიყენება ჰიპერბოლური ტანგენსის ფუნქცია, რომელიც სიგმა ფუნქციის მსგავს ფუნქციას წარმოადგენს იმ განსხვავებით, რომ იგი სიმეტრიულია კოორდინატთა სათავეს მიმართ და $R=0$ წერტილში ნულის ტოლია (ნახ. 2.4 ბ). ამიტომ მას ხშირად სიმეტრიულ სიგმაიდალურ ფუნქციასაც უწოდებენ.

$$Y = \frac{e^R - e^{-R}}{e^R + e^{-R}}$$

რადიალურ-ბაზისური აქტივაციური ფუნქციები. ბოლო წლებში გავრცელება ჰპოვა ნეირონულმა ქსელებმა, რომელთა ნეირონებს გააჩნიათ აქტივაციური ფუნქცია გაუსის ფუნქციის სახით (ნახ. 2.5):

$$y = \exp\left\{-\frac{R^2}{2s^2}\right\},$$

სადაც R ევკლიდეს მანძილია ნეირონის შემავალ X ვექტორსა და აქტივაციური ფუნქციის C ცენტრს შორის: $R = \|X - C\|$, σ - გაუსის ფუნქციის პარამეტრია, რომელიც გვიჩვენებს ფარჯრის სიგანეს.



ნახ. 2.5

დასკენისათვის უნდა აღვნიშნოთ, რომ ზემოდ მოყვანილი აქტივაციური ფუნქციები, გარდა ზღურბლურისა (ნახ. 2.2ა) და შეზღუდული ცვალებადობის წრფივი ფუნქციებისა (ნახ. 2.3ბ), წარმოადგენენ უწყვეტ დიფერენცირებად ფუნქციებს. წრფივი ფუნქცია (ნახ. 2.3ა) ასრულებს შემავალი R მნიშვნელობების სიმრავლის გარდაქმნას მისივე იდენტურ უსასრულო ცვალებადების Y სიმრავლეში. ზღურბლური აქტივაციური ფუნქციები R სიმრავლის ერთობლიობას გარდაქმნის ბინარულ სიმრავლეში ($y = 0$ და $y = 1$ ან $y = -1$ და $y = +1$). დანარჩენი აქტივაციური ფუნქციები R სიმრავლეს გარდაქმნიან შეზღუდული სიდიდის $(0,1)$ ნეირონის Y გამოსავალ სიმრავლეში.

2. ხელოვნური ნეირონული ქსელის დახასიათება

ნეირონული ქსელის აქტუალობა.

ადამიანის თავის ტვინში ინფორმაციის დამუშავება მნიშვნელოვნად განსხვავდება იმ მეთოდებისაგან, რომლებიც მიღებულია კომპიუტერში. თავის ტვინი წარმოადგენს უაღრესად რთულ, არაწრფივ, პარალელურ კომპიუტერს. მეცნიერებმა დაამტკიცეს, რომ თავის ტვინი შედგება უამრავი რაოდენობის სტრუქტურულ ელემენტებისაგან – ნერვის უჯრედებისაგან, რომლებსაც ნეირონები ეწოდება.

ადამიანის ტვინს გააჩნია ნეირონების ისეთი მუშაობის ორგანიზაცია, როდესაც ხდება კონკრეტული ამოცანების (სახეთა გარჩევა, მგრძნობიარეობის ორგანოების სიგნალების დამუშავება, მოტორული ფუნქციები და ბევრი სხვა მოქმედებები) მრავალჯერ უფრო სწრაფად დამუშავება ვიდრე ამას ასრულებს თანამედროვე ყველაზე სწრაფი კომპიუტერი. ასეთი ამოცანის მაგალითად შეგვიძლია მოვიყვანოთ ჩვეულებრივი მხედველობა. ტვინი თანმიმდევრულად ასრულებს სახეთა გარჩევის მთელ რიგ ამოცანებს (მაგალითად, ნაცნობი სახის გარჩევა უცნობ გარემოში). ამისათვის მას სჭირდება დაახლოებით 100 – 200 მილიწამი, მაშინ როცა ანალოგიური ამოცანის და თუნდაც უფრო მარტივი სირთულის ამოცანების ამოხსნისათვის კომპიუტერს სჭირდება გაცილებით დიდი დრო (ზოჯერ დღეები).

სხვა მაგალითად შეგვიძლია მოვიყვანოთ ღამურას ლოკატორი, რომელიც წარმოადგენს ექო-ლოკატორის აქტიურ სისტემას. ღამურას ლოკატორს, გარდა ობიექტამდე მანძილისა, გააჩნია ინფორმაცია ობიექტის სისწრაფეზე, მის ზომაზე და მისი ნაწილების ზომაზე, აგრეთვე მის აზიმუტზე და ობიექტის მოძრაობის სიმაღლეზე. მიღებულ სიგნალიდან ამ ინფორმაციის გამოყოფისათვის ღამურას მცირე ზომის ტვინი ასრულებს რთულ ნეირონულ გამოთვლებს. მფრინავი ღამურას ექო-ლოკაცია თავისი მახასიათებლების ხარისხით და სწრაფმოქმედებით მნიშვნელოვნად აჭარბებს თანამედროვე ყველაზე უფრო რთულ საინჟინრო სისტემებს.

ნეირონული ქსელების შექმნა გამოწვეული იყო ადამიანის თავის ტვინის მუშაობის პრინციპების შესასწავლად, რაც თავის მხრივ ეს გამოიწვევს ხელოვნური ნეირონული ქსელების განვითარებაზე ზემოქმედებას. დღეისათვის ნეირონული ქსელები, ადამიანის ტვინთან შედარებით, წარმოადგენენ მეტად უმნიშვნელო მოდელებს მიუხედავად მათი გარკვეული წარმატებული გამოყენებისა.

ადამიანის თავის ტვინი და კომპიუტერი ასრულებენ სხვადასხვა ამოცანებს და გააჩნიათ სხვადასხვა თვისებები. ადამიანის ტიპურ ტვინს გააჩნია 1000- ჯერ მეტი ნეირონი, ვიდრე მაღალი კლასის კომპიუტერის ლოგიკური ელემენტებია. მურას კანონის თანახმად და იმის გათვალისწინებით, რომ ზოგიერთი გამოთვლებით ტვინში ნეირონების რაოდენობა უნდა გაომარგდეს დაახლოებით ყოველ 2 – 4 მილიონი წლის შემდეგ, შესაძლებელია იმის პროგნოზირება, რომ კომპიუტერის ლოგიკური ელემენტების რაოდენობა გაუტოლდება თავის ტვინის ნეირონების რაოდენობას დაახლოებით 2020 წლისათვის. რასაკვირველია ეს პროგნოზი არაფერს გვეუბნება, რადგან ელემენტებს შორის არსებული სხვაობა უმნიშვნელოა გადართვის სისწრაფესა და პარალელიზმის ხარისხთან შედარებით. დღეისათვის კომპიუტერის მიკროსქემებს შეუძლიათ თითოეული ბრძანების შესრულება ნანოწამზე ნაკლებ სიჩქარით, მაშინ როდესაც ნეირონები მუშაობენ მილიონჯერ უფო ნაკლებ სიჩქარით. ამ ნაკლს ტვინი ანაზღაურებს იმით, რომ ტვინში ყველა ნეირონი მოქმედებს ერთდროულად, მაშინ როცა თანამედროვე კომპიუტერებს გააჩნიათ მხოლოდ ერთი პროცესორი (მაგრამ რამოდენიმე ბირთვით) ან მცირე რაოდენობის პროცესორები. ამრიგად, მიუხედავად იმისა, რომ კომპიუტერს გააჩნია მილიონზე მეტი გადართვის სიჩქარე აღმოჩნდა, რომ თავის ტვინის გადართვის სიჩქარე დაახლოებით გამოდის 100000 – ჯერ უფრო სწრაფი.

ნეირონული ქსელების თვისებები.

ხელოვნური ნეირონული ქსელები ახდენენ თავის ტვინის ზოგიერთი მოვლენების დემონსტრირებას. მაგალითად, ცდების შედეგებით სწავლობენ, განაზოგადებენ წინა პრეცედენტებს და შემოსულ ინფორმაციიდან იღებენ მნიშვნელოვან თვისებებს. მიუხედავად ასეთი მსგავსებისა, ყველა რთული ნეირონული ქსელებისათვის რეალური „ინტელექტის“ დემონსტრირება იმყოფება ჭიაყელის ინტელექტზე დაბლა. მაგრამ, არ შეიძლება იგნორირება გაუკეთოდ ზოგიერთი ნეიროქსელების ფუნქციონირების მსგავსებას ადამიანის თავის ტვინის ქმედებასთან შედარებით. ეს შესაძლებლობები, დღეისათვის რაგინდ შეზღუდული უნდა იყოს, მიუყვართ იმ აზრამდე, რომ ადამიანის ინტელექტუალურ შესაძლებლობაში სიღრმისეული ჩაძიება, აგრეთვე მრავალი რევოლუციური შედეგები, არც ისე შორეული პერიოდის საქმეა. განვიხილოთ ნეიროქსელების ზოგიერთი თვისებები.

1. სწავლება. ხელოვნურ ნეირონულ ქსელები გარე სამყაროსთან მიმართებაში შეუძლიათ თავისი ქცევის შეცვლა. შემაჯავლი სიგნალების მიწოდების შემდეგ ნეიროქსელებს შეუძლიათ თვითრეგულირება, რათა მიღწეული იყოს საჭირო რეაქცია. დამუშავებული იყო ნეიროქსელების სწავლების მრავალი ალგორითმი თავიანთი მძლავრი და სუსტი მხარეებით. მიუხედავად ამისა, მაინც რჩება გარკვეული პრობლემები ქსელის სწავლებასთან დაკავშირებით.

2. განზოგადოება. სწავლების შემდეგ ქსელის გამოსავალი გარკვეულ ზღვრამდე შეიძლება იყოს არამეძნობიარე შემაჯავლი სიგნალების უმნიშვნელო ცვლილებების მიმართ. ქსელის შინაგანი თვისება ხმაურის ფონზე დაინახოს სახე, მეტად მნიშვნელოვანია სახეთა გარჩევის ამოცანების გადასაწყვეტად. ეს გვაძლევს საშუალებას გადავლახოთ მკაცრი სიზუსტის ის მოთხოვნები, რომლებიც სჭირდება კომპიუტერს და გზა გაუხსნას ისეთი სისტემების შექმნას, რომელთაც შეუძლია საქმე ჰქონდეს არასრულყოფილ გარემოსთან, სადაც რეალურად ჩვენ ვცხოვრობთ.

3. აბსტრაგირება. ზოგიერთ ხელოვნურ ნეიროქსელს გააჩნია შემომაჯავლი სინალებიდან არსის გამოყოფა. მაგალითად, ქსელს შეიძლება შევასწავლოთ დამახინჯებული „ა“ ასოს თანმიმდევრობაზე. გარკვეული სწავლების შემდეგ თუ ქსელს მივაწვდით ასეთ დამახინჯებულ ობიექტს, მაშინ ქსელი გვაძლევს ობიექტის სრულყოფილ სახეს (ამ შემთხვევაში „ა“ ასოს). ქედან გამომდინარე, გარკვეული აზრით ქსელი სწავლობს იმის წარმოქმნას, რაც ადრე მას არ უსწავლია.

4. გამოყენება. ხელოვნური ნეირონული ქსელი პანაცეას არ წარმოადგენს. ისინა ალბათ არ გამოდგებიან ისეთი ამოცანების ამოსახსნელად, როგორიცაა მაგალითად ხელფასის დარიცხვა, მაგრამ ისინი შეუცვლელნი არიან ისეთი ამოცანების ამოსახსნელად, რომელთაც ვერ უძლავდებიან გამოთვლითი სისტემები.

ნეირონული ქსელების უპირატესობები.

ცხადია, რომ ნეირონული ქსელების სიძლიერე დამოკიდებულია ჯერ ერთი ინფორმაციის პარალელურ დამუშავებაზე და მეორე თვითსწავლების ანუ განზოგადოების უნარზე. ტერმინ „განზოგადოების“ ქვეშ იგულისხმება საჭირო შედეგის მიღება ისეთ მონაცემებისაგან, რომლებიც არ გვხვებიან ნეიროქსელების სწავლების პროცესში. ეს თვისება საშუალებას იძლევა ნეიროქსელმა გადაწყვიტოს რთული (მაშტაბური) ამოცანები, რომლებიც დღეისათვის ითვლებიან ძნელად ამოსახსნელები.

მოვიყვანოთ ნეიროქსელების უპირატესობები ტრადიციულ გამოთვლით სისტემებთან შედარებით.

1. ამოცანების ამოხსნა უცნობი კანონზომიერების შემთხვევაში. მრავლობით მაგალითებზე სწავლების შემდეგ ნეიროქსელებს შესწევთ ისეთი ამოცანების ამოხსნის უნარი, სადაც უცნობია დამოკიდებულება ქსელში შემაჯავლ და გამომაჯავლ სიდიდეებს შორის. როგორც წესი, ტრადიციული მათემატიკური მეთოდები და ექსპერტული სისტემები ამ შემთხვევებში უძლურნი არიან.

ცდომილების უკუგავრცელების ალგორითმის შექმნამ ხელი შეუწყო ნეიროქსელური ტექნოლოგიების ფართო გამოყენებას სხვადასხვა ტიპის ამოცანების გადასაწყვეტად. მრავალშრიანი პერსპექტივის საშუალებით შესაძლებელი გახდა ისეთი მათემატიკური მოდელების შექმნა, რომლებიც

ასრულებენ რთულ მრავალგანზომილებიან ასახვას შემავალი პარამეტრების X ვექტორისა გამოსავალ Y ვექტორზე.

ტრადიციული მათემატიკური მოდელები აიგებოდა ბუნების ფუნდამენტალური კანონების გამოყენებით. ამ მოდელების განხილვის შედეგებიდან გამომდინარეობდა მათემატიკური ფორმულების ან დიფერენციალური განტოლებებისათვის სასაზღვრო ამოცანების ჩამოყალიბება. ნეიროქსელებმა შექმნეს მათემატიკური მოდელების აგების სხვა მიდგომები. შეიქმნა იმის საშუალება, რომ ავავთოთ მათემატიკური მოდელები ფიზიკის, ქიმიის, ბიოლოგიის, მედიცინის და სხვა კანონების გარეშე, მხოლოდ ემპირიული ცდის (ამონარჩევის სწავლება) საშუალებით. ასეთი მოდელები თვითონ იყენებენ არსებულ კანონებს და მათ ეფექტურად იყენებენ ფართო სპექტრის პრაქტიკული ამოცანების ამოხსნისათვის.

მონაცემებიდან ცოდნის ამოღების ასეთი ახალი ინსტრუმენტი საშუალებას იძლევა ხელახლა აღმოვაჩინოთ ბუნების განვითარების ფუნდამენტალური კანონები, გამოვავლინოთ ადრე უცნობი და შეუსწავლელი დამოკიდებულებები და კანონზომიერებები და აქტიურად გამოვიყენოთ ისინი კონკრეტული პრაქტიკული ამოცანების ამოხსნისათვის.

2. შემავალი მონაცემების ხმაურთან მდგრადობა. ნეირონულ ქსელებს გააჩნიათ მუშაობის უნარი იმ შემთხვევების დროს, როდესაც ქსელში შემომავალი ინფორმაცია შეიცავს მრავალ არაინფორმატიულ მონაცემებს და ხმაურს. ამ შემთხვევაში არ არის აუცილებელი საწყის მონაცემებიდან მათი წინასწარი გამორიცხვა, რადგან ნეიროქსელი თვითონ გამოავლენს ამოცანის ამოხსნისათვის ნაკლებად მისაღებ მაჩვენებლებს და უგულებელყოფს მათ.

3. გარემომცველი გარემოს შეცვლასთან ადაპტირება. ნეირონულ ქსელებს გააჩნიათ უნარი მოახდინონ ადაპტირება გარემოს შეცვლის დროს. კერძოდ, ნეიროქსელი, რომელიც ერთ გარემოში იყო ნასწავლი, ადვილად შეძლებს გადამზადებას ახალ გარემოში. უფრო მეტიც, არასტაციონარულ გარემოში, სადაც სტატისტიკური მაჩვენებლები დროში იცვლება, შესაძლებელია შეიქმნას ნეიროქსელქები, რომელთა გადამზადება ხდება რეალურ დროში. რაც უფრო დიდია ქსელის ადაპტირების შესაძლებლობები, მით უფრო მდგრადია ქსელის მუშაობა არასტაციონარულ გარემოში.

უნდა გვახსოვდეს, რომ ადაპტიურობას ყოველთვის არ მივეყვართ მდგრადობისაკენ, ზოგჯერ მას მოაქვს სრულად საწინააღმდეგო შედეგი. მაგალითად, დროში სწრაფად ცვალებადი პარამეტრების მქონე ადაპტურ სისტემას შეუძლია სწრაფი რეაგირება უცხო შემოვლოებებზე, რაც იწვევს მუშაობაში შეფერხებას. იმისათვის, რომ გამოვიყენოთ ადაპტურობის ყველა თვისება, საჭიროა სისტემის ძირითადი პარამეტრები იყვნენ საკმაოდ სტაბილური, რათა უგულებელყოთ გარემოს ხმაური. გარდა ამისა, სისტემა უნდა იყოს საკმაოდ მოქნილი, რათა უზრუნველყოს გარემოს ცვლილებებზე რეაქცია.

4. ზემოდადებული სიჩქარე. ნეირონულ ქსელებს გააჩნიათ ზემოდადებული სიჩქარე, რასაც უზრუნველყოფს ინფორმაციის მასიური პარალელური დამუშავება.

5. ნეიროქსელის აპარატული რეალიზაციის უარისთქმის მდგრადობა. ნეირონული ქსელი პოტენციურად უარისთქმის მიმართ მდგრადია. ეს იმას ნიშნავს, რომ არახელსაყრელი პირობებისას ქსელის წარმადობა მნიშვნელოვნად ეცემა. მაგალითად, თუ დაზიანებულია რომელიმე ნეირონი ან მისი კავშირები, მაშინ დამახსოვრებული ინფორმაციის ამოღება ძნელდება. მაგრამ, მიიღება რა მხედველობაში ქსელში ინფორმაციის შენახვის განაწილება, ამიტომ მხოლოდ ნეირონული ქსელის სტრუქტურის სერიოზული დაზიანება გამოიწვევს მის

მუშაობაზე ზეგავლენას. აქედან გამომდინარე, ნეიროქსელის მუშაობის ხარისხის შემცირება წარმოებს ნელა.

ნეირონულ ქსელში ცოდნის წარმოდგენა.

ნეირონულ ქსელში ცოდნის წარმოდგენა მდგომარეობს შემდეგში: როგორი ინფორმაციაა საჭირო და ეს ინფორმაცია ფიზიკურად როგორ წარმოვადგინოთ მისი შემდგომი გამოყენებისათვის. აქვე უნდა შევნიშნოთ, რომ ინტელექტუალ სისტემებში წარმატებული შედეგი დამოკიდებულია ცოდნის კარგ წარმოდგენაზე. ეს ეხება ნეირონულ ქსელებსაც, რომლებიც განეკუთვნებიან ხელოვნური ინტელექტის ცალკე კლასს. რადგან შემაჯავლი სიგნალების წარმოდგენის ფორმა შეიძლება იყოს მრავალფეროვანი, ამიტომ მისაღები ნეიროქსელის შექმნა წარმოადგენს შემოქმედებით პროცესს.

ნეირონული ქსელის ძირითად ამოცანას წარმოადგენს გარე სამყაროს მოდელის საუკეთესოდ შესწავლა დასმული ამოცანების ამოსახსნელად. გარე სამყაროს ცოდნაში იგულისხმება ორი ტიპის ინფორმაცია:

– გარე სამყაროს ცნობილი მდგომარეობა წარმოდგენილი სარწმუნო ფაქტებით. ასეთ ინფორმაციას ეწოდება აპრიორული;

– გარე სამყაროზე დაკვირვებები (გაზომვები) მიღებული სენსორების საშუალებით, რომლებიც ადაპტირებულნი ანიან კონკრეტულ პირობებთან, სადაც უნდა იფუნქციონროს ნეირონულმა ქსელმა. ჩვეულებრივ ასეთი გაზომვები საკმარის დოზით შეიცავენ ხმაურს, რომელიც პოტენციურად შეიძლება გახდეს ცდომილების წყარო.

სასწავლო მაგალითები შეიძლება იყოს მარკირებული ან მარკირების გარეშე. მარკირებულ შემთხვევაში სიგნალს შეესაბამება შესაბამისი გამოძახილი (რეაქცია). არამარკირებული მაგალითები შედგებიან ერთი შემაჯავლი სიგნალის რამდენიმე სხვადასხვა რეალიზაციებისაგან. ნებისმიერ შემთხვევაში მაგალითების ნაკრები, იქნებიან ისინი მარკირებული თუ არამარკირებული, წარმოადგენს შესასწავლი დარგობრივი სფეროს ცოდნას, რომლის საფუძველზე წარმოებს ნეირონული ქსელის სწავლება.

წყვილ სიგნალთა სიმრავლე, რომელიც შედგება ერთი შემაჯავლი სიგნალით და მეორე მისი შესაბამისი გამოსავალი სიგნალით, ეწოდება ნეიროქსელის სასწავლო მონაცემები ან ამონარჩევი. მონაცემები, რომლებიც გამოიყენებიან ნეიროქსელის სწავლებისათვის, უნდა შეიცავდნენ როგორც დადებით ასევე უარყოფით მაგალითებს.

მაგალითისათვის განვიხილოთ ციფრების გარჩევის ამოცანა. აქ შემაჯავლი სიგნალი (გამოსახულება) წარმოადგენს შავი და თეთრი წერტილებისაგან შემდგარ მატრიცას. თითოეული გამოსახულება წარმოადგენს თეთრ ფონზე ათიდან ერთ-ერთ ხელნაწერ ციფრს, რომელიც მიეწოდება ნეირონული ქსელის შესასვლელს. ზოგადად სწავლების ამონარჩევი შედგება დიდი რაოდენობის ხელნაწერი ციფრებისაგან, რომლებიც ასახავენ სიტუაციას, რომელიც რეალურად შეიძლება წარმოიშვას პრაქტიკაში. ასეთი ამონარჩევის შემთხვევაში ნეიროქსელი შეიქმნება შემდეგნაირად:

პირველ რიგში შეირჩევა ნეირონული ქსელის შესაბამისი არქიტექტურა, სადაც შემაჯავლი შრის განზომილება შეესაბამება გამოსახულების წერტილების რაოდენობას, ხოლო გამოსავალი შრე შედგება 10 ნეირონისაგან, რომლებიც შეესაბამებიან ციფრებს (0,1,2,3,4,5,6,7,8,9). ამის შემდეგ ხდება არსებული სწავლების ამონარჩევით ნეიროქსელის წონითი კოეფიციენტების რეგულირება. ქსელის მუშაობის ამ რეჟიმს სწავლება ეწოდება.

ნეირონული ქსელის სწავლების ეფექტიანობა (ტესტირება) მოწმდება მრავალ მაგალითზე, რომლებიც განსხვავდებიან სწავლების პროცესში გამოყენებული მაგალითებისაგან. ამ დროს ქსელის შესასვლელს მიეწოდება გამოსახულება, რომლისთვისაც ცნობილია მიზნობრივი გამოსავალი. ქსელის ეფექტური სწავლება მოწმდება გარჩევის შედეგის რეალურ ციფრთან შედარებით. ნეირონული ქსელის მუშაობის ამ ეტაპს ეწოდება განზოგადოება.

ნეირონულ ქსელში ცოდნის წარმოდგენა საკმაოდ რთული პრობლემაა. მიუხედავად ამისა, შესაძლებელია გამოვეყოთ ოთხი ზოგადი წესი.

წესი 1. მსგავსი შემავალი სიგნალები საჭიროა ერთიანი წარმოდგენით ჩამოყალიბება. კერძოდ, ისინი უნდა კლასიფიცირდებოდნენ როგორც ერთი კატეგორია. არსებობს მრავალი მიდგომა შემავალი სიგნალების მსგავსების დასადგენად. ზოგადად მსგავსება განისაზღვრება ევკლიდეს მანძილით.

წესი 2. ნეირონულ ქსელში სხვადასხვა კლასის ელემენტებს უნდა გააჩნდეთ რაც შეიძლება მეტი განსხვავებული წარმოდგენა.

წესი 3. თუ რომელიმე თვისება მნიშვნელოვანია, მაშინ ნეიროქსელში მისი წარმოდგენისათვის საჭიროა დიდი რაოდენობის ნეირონების გამოყენება. მაგალითისათვის მოვიყვანოთ რადარის მიერ ობიექტის (მაგალითად თვითმფრინავის) აღმოჩენის ამოცანა, როდესაც გარემოში არსებობს გარკვეული ხმაური, რომელიც გამოწვეულია, მაგალითად, სხვა ობიექტებიდან სიგნალის არეკვლით, როგორიცაა სახლი, ხეები, ღრუბელი და სხვა. ასეთი რადარის ეფექტიანობა იზომება ორი ალბათური სიდიდით:

- აღმოჩენის ალბათობა ანუ ალბათობა იმისა, რომ ობიექტის არსებობის შემთხვევაში რადარი მას აღმოაჩენს;

- ცრუ განგაშის ალბათობა ანუ ალბათობა იმისა, რომ სისტემა ობიექტს აღმოაჩენს მაშინ, როცა ის სინამდვილეში არ არსებობს.

ზოჯერ ამ ორ მნიშვნელობას პირველი და მეორე გვარის შეცდომებს უწოდებენ. ნეიმან-პირსონის კრიტერიუმის თანახმად ობიექტის აღმოჩენის ალბათობა უნდა იყოს მაქსიმალური, ხოლო ცრუ განგაშის ალბათობა მინიმალური. ასეთი სიტუაციის დროს მეტად მნიშვნელოვანია ის, რომ შემავალი სიგნალი უნდა შეიცავდეს მიზნის ინფორმაციას. მესამე წესი გვიჩვენებს, რომ გადაწყვეტილების მიღების პროცესში სიგნალში ობიექტის არსებობის შესახებ ნეირონულ ქსელს უნდა გააჩნდეს დიდი რაოდენობის ნეირონები. ნებისმიერ შემთხვევაში ნეირონების გაზრდა იწვევს მიღებული შედეგის სარწმუნოების გაზრდას და ხმაურთან მდგრადობას.

წესი 4. ნეიროქსელის სტრუქტურაში უნდა იყოს ჩაშენებული აპრიორული ინფორმაცია, რაც იწვევს ქსელის არქიტექტურის და სწავლების პროცესის გამარტივებას. ქსელის სწორი კონფიგურაცია უზრუნველყოფს მის სპეციალიზაციას, რომელიც მეტად მნიშვნელოვანია შემდეგი მოსაზრებიდან გამომდინარე:

- ბიოლოგიური ქსელი, რომელიც უზრუნველყოფს მხედველობითი და სმენითი ინფორმაციის დამუშავებას, მეტად სპეციალიზირებულია;

- სპეციალიზირებული ქსელის სწავლებისათვის საჭიროა მცირე მონაცემები. მ დროს სწავლების პროცესს სჭირდება მცირე დრო და ქსელს გააჩნია განზოგადოების კარგი თვისება;

- სპეციალიზირებულ ნეირონულ ქსელებს გააჩნიათ დიდი გამტარუნარიანობა;

- სპეციალიზირებული ნეირონული ქსელების ხარჯები მცირეა, რადგან მათი ზომები მნიშვნელოვნად მცირეა, ვიდრე სრულკავშირებიანი ქსელებისა.

3. ხელოვნური ნეირონული ქსელის არქიტექტურა

3.1 ნეირონული ქსელის სტრუქტურა

ხელოვნური ნეირონული ქსელი წარმოადგენს ნეირონების ერთობლიობას, რომლებიც ერთმანეთთან დაკავშირებულნი არიან. ყველა ნეირონის აქტივაციური (გადაცემის) ფუნქციები ქსელში ფიქსირებულნი არიან, ხოლო წონები წარმოადგენენ ქსელის პარამეტრებს და შესაძლოა მათი ცვლილება.

ქსელის შესასვლელზე რიცხვების მიწოდება იწვევს ქსელის გამოსავალზე გარკვეულ რიცხვთა ერთობლიობის წარმოქმნას. ამრიგად, ნეირონული ქსელის მუშაობა შედგება შემავალი X ვექტორის გარდაქმნაში გამოშვებულ Y ვექტორში, თანაც ეს გარდაქმნა ხორციელდება ქსელის წონითი კოეფიციენტების საშუალებით. პრაქტიკულად ნებისმიერი ამოცანა შეიძლება დავიყვანოთ ისეთ ამოცანამდე, რომლის ამოხსნა შესაძლებელია ნეირონული ქსელის საშუალებით.

ნეირონული ქსელის შესაქმნელად საჭიროა ორი ეტაპის გავლა:

- ნეიროქსელის სტრუქტურის (არქიტექტურის) შერჩევა;
- ქსელის წონების შერჩევა (სწავლება).

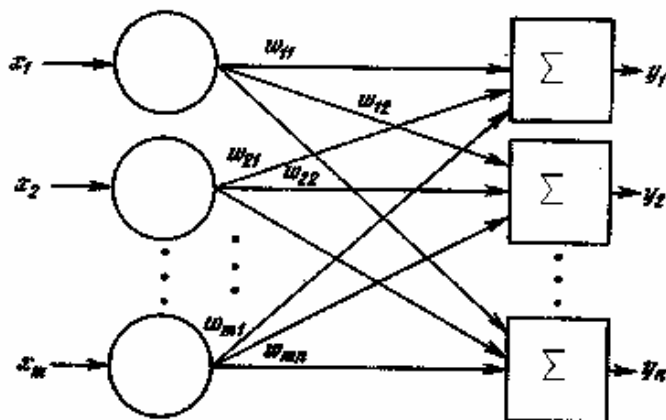
მიუხედავად იმისა, რომ ერთ ნეირონს შეუძლია შეასრულოს უმარტივესი გარჩევის ოპერაცია, უფრო რთული ამოცანების გადასაჭრელად საჭიროა ნეირონების გაერთიანება ქსელის სახით. ქსელში ნეირონების მიერ შესრულებული ფუნქციებიდან გამომდინარე, ნეირონები შეიძლება დავეყოთ სამ ტიპად:

– შემავალი ნეირონები. ეს ისეთი ნეირონებია, რომლებსაც მიეწოდებათ შემავალი ვექტორი. აქ, როგორც წესი, არ ხდება რაიმე გამოთვლითი ოპერაციები, ინფორმაცია გადაეცემა ქსელის გამოსავალ ნეირონს მისი აქტივაციური ფუნქციის საშუალებით/

– გამოსავალი ნეირონები. ეს ისეთი ნეირონებია, რომლებიც წარმოადგენენ ნეირონული ქსელის გამოსავალს.

– შუალედური ნეირონები. ეს რეირონები წარმოადგენენ ხელოვნური ნეირონული ქსელის საფუძველს, სადაც ხდება ინფორმაციის გარდაქმნა.

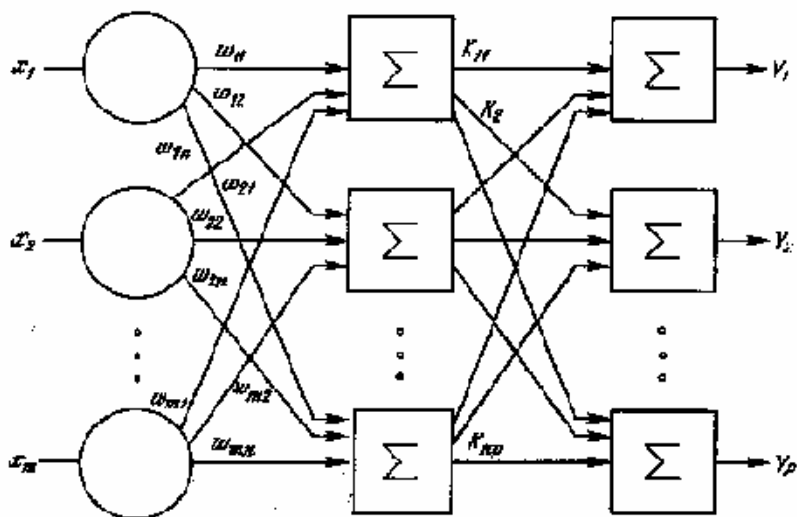
ნეირონების ერთობლიობა ქსელში ქმნიან შრეს. შრე შეიცავს ერთგვაროვან ნეირონებს საერთო შემავალი სიგნალებით. შერეების რაოდენობა ნეირონულ ქსელში შეიძლება იყოს ნებისმიერი. ქედან გამომდინარე, არჩევან ერთშრიან და მრავალშრიან ხელოვნურ ნეირონულ ქსელებს. ერთშრიანი ნეირონული ქსელის ბლოკ-სქემა წარმოდგენილია შემდეგ ნახაზზე:



სქემაზე აღნიშნული რგოლები გამოიყენება მხოლოდ შემავალი სიგნალის განაწილებისათვის. მათ არ გააჩნიათ რაიმე გამოთვლითი ფუნქცია და ამიტომ ისინი შრედ არ შეიძლება ჩაითვალოს.

წონითი კოეფიციენტების W მატრიცას გააჩნია m სტრიქონი და n სვეტი, სადაც n ქსელის შესასვლელის რაოდენობაა, m – ნეირონების რაოდენობაა. მაგალითად, w_{23} არის წონა, რომელიც აერთიანებს ქსელის მეორე შესასვლელს მესამე ნეირონთან. ამრიგად, გამოსავალი Y ვექტორის განსაზღვრა, რომლის კომპონენტებია y_i ნეირონების გამოსასვლელები, დაიყვანება შემდეგ მატრიცულ ნამრავლზე: $Y = XW$, სადაც Y და X – ვექტორი-სტრიქონებია.

ორშრიანი ხელოვნური ნეირონული ქსელი შეიძლება ასე წარმოვადგინოთ:



როგორც ვხედავთ, მრავალშრიანი ქსელი შეიძლება შევქმნათ შრეების კასკადებით, სადაც ერთი შრის გამოსასვლელი მეორე შრის შესასვლელია. შრის გამოსასვლელი სიგნალი განისაზღვრება შემავალი ვექტორის პირველი წონით მატრიცაზე გამრავლებით და შემდგომ მიღებული შედეგის გამრავლებით მეორე შრის წონით მატრიცაზე, ე.ი. $(XW_1)W_2$ ანუ $X(W_1W_2)$. მრავალშრიანი კავშირი ნეიროქსელის გამოსავალთან, ფარული შრე ეწოდება. ნეირონული ქსელის შესავალი შეიძლება განვიხილოთ როგორც ნულობანი შრე, რომლის

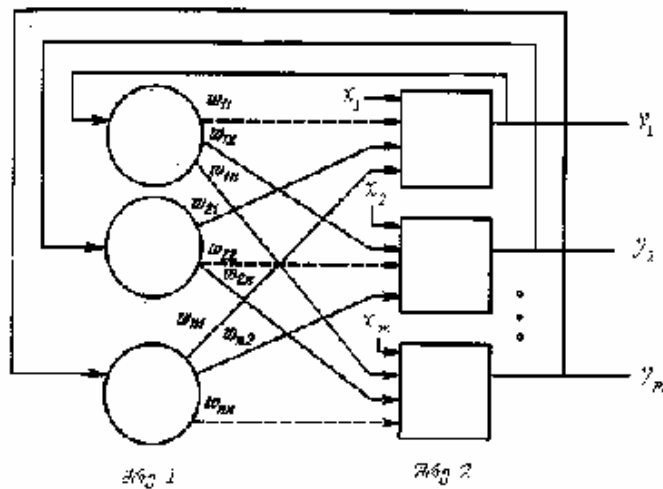
ნეირონები გამოიყენებიან მხოლოდ როგორც გამანაწილებელი, სადაც არ სრულდება სიგნალების აჯამების და გარდაქმნის ოპერაციები. გარდა შემავალი და გამომავალი შრეებისა, ნეირონული ქსელები შეიცავენ ერთ ან რამოდენიმე ფარულ შრეებს/

ერთშრიანი წრფივი ნეირონული ქსელი ექვივალენტურია ერთშრიანი ქსელისა, რომლის წონის მატრიცა მიიღება ორივე შრის წონითი მატრიცების გადამრავლებით. აქედან გამომდინარე, ნებისმიერი მრავალშრიანი წრფივი ნეირონული ქსელი შეგვიძლია შევცვალოთ მისი ექვივალენტური ერთშრიანი ქსელით.

ერთშრიანი ქსელი საკმაოდ შეზღუდულია თავისი გამოთვლითი შესაძლებლობებით. იმისათვის, რომ ქსელის შესაძლებლობა ერთშრიან ქსელთან შედარებით გაიზარდოს, საჭიროა გამოვიყენოთ არაწრფივი აქტივაციის ფუნქცია.

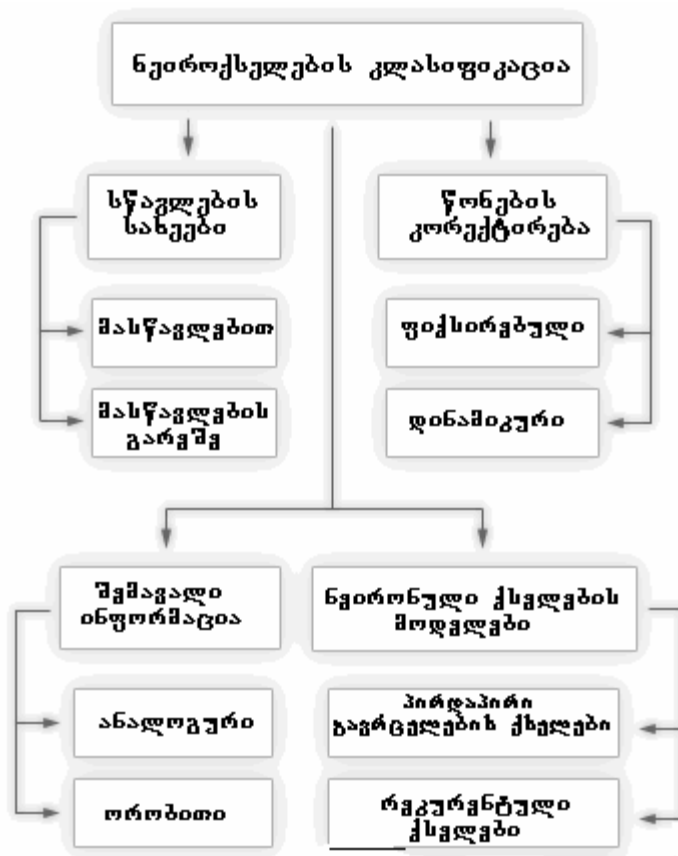
ზემოთ განხილულ ნეირონულ ქსელებს უკუკავშირი არ გააჩნიათ და ამიტომ მათ პირდაპირი გავრცელების ქსელებს უწოდებენ. ასეთ ქსელებს მესხიერება არ გააჩნიათ და მათი გამოსავალი მთლიანად განისაზღვრება მიმდინარე შემავალი სიგნალებით და წონებით.

უკუკავშირიანი ქსელის შემთხვევაში გამოსავალი სიგნალი უბრუნდება შემოსასვლელს და აქედან გამომდინარე, გამოსავალი სიგნალი განისაზღვრება როგორც მიმდინარე შემომავალი სიგნალით, ასევე წინა გამოსავალი სიგნალებით. ამ მიზეზის გამო უკუკავშირებიან ქსელს შეიძლება გააჩნდეს ადამიანის მოკლევადიანი მესხიერების მსგავსი თვისებები. ერთშრიანი უკუკავშირებიანი ქსელი შეიძლება ასე წარმოვადგინოთ:



3.2. ნეირონული ქსელების კლასიფიკაცია

ნეირონული ქსელების კლასიფიკაცია წარმოდგენილია შემდეგ ნახაზზე:



სწავლების სახეებით ნეირონული ქსელები იყოფა:

- ქსელები, რომლებიც იყენებენ სწავლებას მასწავლებლით,
- ქსელები, რომლებიც იყენებენ სწავლებას მასწავლებლის გარეშე.

სწავლება მასწავლებლით გულისხმობს შემდეგს. ყოველ შემავალ ვექტორისათვის არსებობს მიზნობრივი ვექტორი (სწორი პასუხი), რომელიც წარმოადგენს სასურველ გამოსავალ ვექტორს. მათ ერთად ეწოდებათ სწავლების წყვილი. ჩვეულებრივ ნეიროქსელის სწავლება ხორციელდება გარკვეული რაოდენობის ასეთი სწავლების წყვილების საშუალებით. შესასწავლი ვექტორების სიმრავლე თანმიმდევრულად მიეწოდება ნეიროქსელის შესასვლელს. ცდომილების განსაზღვრისათვის ქსელის გამოსავალი დარდება შესაბამის მიზნობრივ ვექტორს და ხდება ყოველი ვექტორისათვის სინაფსური წონითი კოეფიციენტების კორექტირება. წონითი კოეფიციენტები ისე უნდა დარეგულიდეს, რომ გამოსავალი ვექტორი (პასუხი) რაც შეიძლება ახლოს იყოს მიზნობრივ ვექტორთან.

სწავლება მასწავლების გარეშე ანუ თვითსწავლების ქსელები წარმოადგენენ, ნეირონის ბიოლოგიური საფუძვლების თვალსაზრისით, სწავლების გაცილებით დამაჯერებელ მოდელს. ასეთ ქსელებს არ სჭირდებათ მიზნობრივი ვექტორები და შესაბამისად შედარების ოპერაციების ჩატარება. სწავლების სიმრავლე შდგება მხოლოდ შემავალი ვექტორებისაგან. სწავლების ალგორითმი წონითი კოეფიციენტების კორექტირებას ისე ახდენს, რომ

მიღებული იყოს შეთანხმებული გამოსავალი ვექტორი. ე.ი. ისე, რომ საკმაოდ ახლოს მდგომარი შემავალი ვექტორები იძლეოდნენ ერთნაირ გამოსავალ ვექტორს.

არსებობს აგრეთვე შერეული სწავლების ალგორითმი, რომლის დროსაც წონითი კოეფიციენტების ნაწილი განისაზღვრება მასწავლების, ხოლო მეორე ნაწილი – მასწავლების გარეშე.

წონების კორექტირება. აქ გვაქვს ფიქსირებული კავშირებიანი ქსელები, რომლისთვისაც წონითი კოეფიციენტები, ამოცანის პირობიდან გამომდინარე, შეირჩევა თავიდან და დინამიური კავშირების ქსელები, რომლისთვისაც სწავლების პროცესში ხდება სინაფსური წონების კორექტირება.

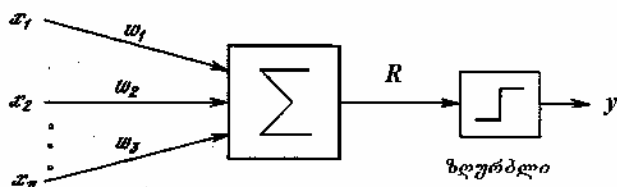
შემავალი ინფორმაცია ორი სახისაა: ანალოგური ანუ ნამდვილ რიცხვთა ფორმით წარმოდგენილი მონაცემები და ორობითი, სადაც შემავალი ინფორმაცია წარმოდგენილია ორობით (0,1) კოდის საშუალებით.

პირდაპირი გავრცელების ქსელების ყველა კავშირები მიმართულია მკაცრად შემავალი შრის ნეირონებიდან გამოსავალი შრის ნეირონებისაკენ. ასეთ ქსელებს მიეკუთვნებიან პერსეპტრონები, რადიალურ – ბაზისური ფუნქციების ქსელები და ალბათური ნეიროქსელები.

რეკურენტულ ანუ უკუკავშირებიან ქსელებში სიგნალი გამოსავალი შრის ნეირონებიდან ან ფარული შრის ნეირონებიდან გადაეცემა უკან შესასვლელი შრის ნეირონების შესასვლელებს. ასეთ ქსელებს მიეკუთვნებიან ჰოპფილდის, ჰემინგის, კოჰონენის, ადაპტური რეხონანსული თეორიის და სხვა ნეიროქსელები.

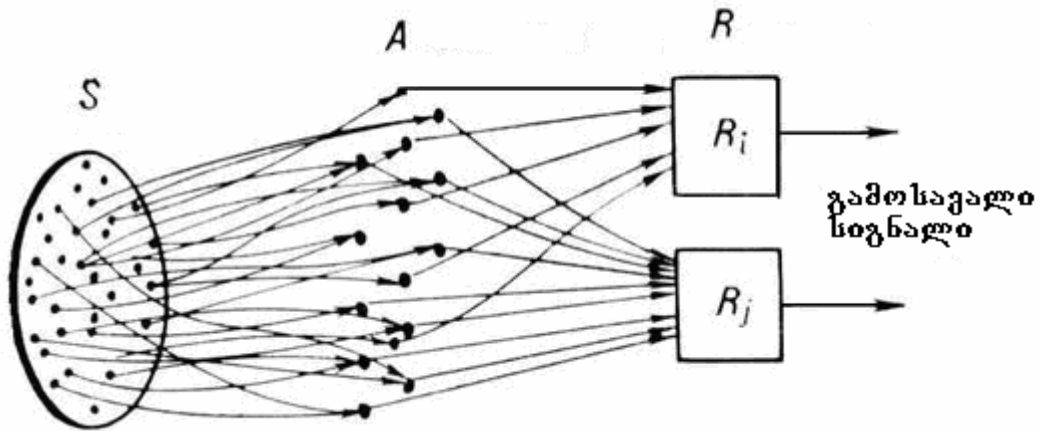
4. პერსეპტრონი

ერთ-ერთ პირველ ნეირონულ ქსელს წარმოადგენს პერსეპტონი, რომელიც 1957 წ. დაამუშავა ფ. როზენბლანტმა. როგორც წესი, პერსეპტონი ეწოდება ერთშირის ნეირონულ ქსელს, სადაც ყველა ნეირონი იყენებს ერთეულოვან ნახტომისებრივ აქტივაციურ ფუნქციას. ერთნეირონულ მარტივ პერსეპტრონს აქვს შემდეგი სახე:



თუ აჯამული სიგნალი $R = \sum_{i=1}^n w_i x_i$ მეტია ზღვრბლზე, მაშინ ნეირონის გამოსავალი სიგნალი $Y = 1$, წინააღმდეგ შემთხვევაში $Y = 0$.

განვიხილოთ ფ. როზენბლანტის მიერ დამუშავებული პერსეპტრონი, რომლის ბლოკ-სქემა წარმოდგენილია შემდეგ ნახაზზე:



პერსპექტონის მიმღებ მოწყობილობად აღებულია ბადურას ფოტოელექტრული მოდელი, კერძოდ რეცეპტორების S არე, რომელიც შედგება რამდენიმე ასეული ფოტოწინააღმდეგობისგან და რომელიც თვალის ბადურის ანალოგიურია და აღიქვამს გარედან შემოსულ სიგნალებს. რეცეპტორების არეს თითოეული ელემენტი შეიძლება იყოს ორ მდგომარეობაში – აღგზნებულ და არაღგზნებულში, იმისდა მიხედვით ეცემა თუ არა შესაბამის ფოტოწინააღმდეგობის არეში წარმოდგენილი გასარჩევი ფიგურის კონტური. თითოეული ელემენტის გამოსასვლელზე მიიღება სიგნალი, რომელიც ერთის ტოლია თუ ელემენტი აღგზნებულია და ნულის ტოლია, როცა ელემენტი არ არის აღგზნებული.

რეცეპტორების გამოსასვლელები მიერთებულია ასოციატური შრის ე.წ. A ელემენტებთან შემთხვევითობის პრინციპით და იგი უცვლელია ექსპერიმენტის ბოლომდე. A ელემენტი წარმოადგენს სუმატორს რამდენიმე შესასვლელით და ერთი გამოსასვლელით. A ელემენტი აწარმოებს მის შესასვლელზე მიწოდებული სიგნალების აღგებრულ აჯამებას და მიღებული მნიშვნელობა თუ აღემატება რაიმე T ზღურბლურ მნიშვნელობას, მაშინ A ელემენტი აღიგზნება და გამოსასვლელზე გვაძლევს სიგნალს, რომელიც ერთის ტოლია. წინააღმდეგ შემთხვევაში გამოსასვლელზე გვექნება ნული. ე.ი.

$$x_i = \begin{cases} 1, & \text{როცა } \left(\sum_i r_{ij} x_i - T \right) \geq 0, \\ 0, & \text{როცა } \left(\sum_i r_{ij} x_i - T \right) < 0, \end{cases}$$

სადაც $r_{ij}=1$, როცა i -ური რეცეპტორი მიერთებულია A_j ელემენტთან და $r_{ij}=-1$, როცა i -ური რეცეპტორი მიერთებულია A_j ელემენტთან უარყოფითი ნიშნით. როცა i -ური რეცეპტორი A_j ელემენტთან არ არის მიერთებული, მაშინ $r_{ij}=0$. A ელემენტის გამოსავალი სიგნალები მრავალდებიან შესაბამის w_i წონით კოეფიციენტებზე და შემდეგ აიჯამებიან

$$R = \sum_{j=1}^m w_j x_j$$

აჯამული სიგნალი მიეწოდება ე. წ. რეაგირების ელემენტს ანუ R -ელემენტს. თუ R დადებითია ან ნულის ტოლი, მაშინ R ელემენტი გამოსასვლელზე იძლევა ერთს, ხოლო როცა R უარყოფითია, მაშინ ნულს. ე. ი.

$$Y = \begin{cases} 1, & \text{როცა } R \geq 0, \\ 0, & \text{როცა } R < 0. \end{cases}$$

დაუშვათ, რომ რეცეპტორების არეში დაპროგრამებულია ფიგურები, რომლებიც მიეკუთვნებიან ორ სხვადასხვა კლასს. თუ შესაძლებელია პერსექტორი მივიყვანოთ ისეთ მდგომარეობაში, რომ საკმაოდ საიმედოდ გამოსასვლელზე მოგვცეს ერთი, როცა მის შესასვლელზე ერთი კლასის ფიგურაა და ნული, როცა მეორე კლასის ფიგურაა, მაშინ პერსექტორს გააჩნია ორი სახის გარჩევის უნარი.

როცა სახეთა რაოდენობა ორზე მეტია, მაშინ წონითი კოეფიციენტების სიმრავლე $\{w\}$ მოცემული უნდა იყოს ყველა სახისათვის ცალ-ცალკე. ასევე, რეაგირების R ელემენტების რაოდენობა ტოლი უნდა იყოს სახეთა რაოდენობისა. გასარჩევი ობიექტი მიეკუთვნება იმ სახეს, რომლის R ელემენტის გამოსავალი სიგნალი უდიდესია.

5 ხელოვნური ნეირონული ქსელების სწავლება

5.1 პრეშრიანი ნეირონული ქსელის სწავლება

როგორც ავღნიშეთ, ნეირონული ქსელები გამოიყენებიან სახეთა გარჩევის, პროგნოზირების და არაწრფივი დამოკიდებულების (არაწრფივი რეგრესია) ამოცანების გადასაწყვეტად. ამისათვის საჭიროა ნეირონული ქსელების სწავლება. მეტად მნიშვნელოვანია ის ფაქტი, რომ ნეირონული ქსელების სწავლება შესაძლებელია. გასული საუკუნის 60 წლებში როზენბლანტმა დაამტკიცა თეორემა ერთშრიანი პერსექტორის სწავლების შესახებ.

ნეირონის სწავლება ნიშნავს შევატყობინოთ მას თუ რა გვიდა მისგან. სწავლების პროცესი ძალიან წააგავს ბავშვის აღფავეტის სწავლებას. როცა ბავშვს ვუჩვენებთ „ა“ ასოს ჩვენ ვეკითხებით „რომელი ასოა ეს?“. თუ პასუხი არასწორია, ვეუბნებით სწორ პასუხს. ბავშვი იმახსოვრებს ამ პასუხს, ე.ი. მის მეხსიერებაში ხდება გარკვეული ცვლილებები საჭირო მიმართულებით. ვიმეორებთ ამ პროცესს სხვა ასოებისათვის მანამ, სანამ ბავშვი მტკიცედ არ დაიმახსოვრებს აღფავეტის ყველა ასოს. ზოგადად ასეთ პროცესს ეწოდება „სწავლება მასწავლებლით“.

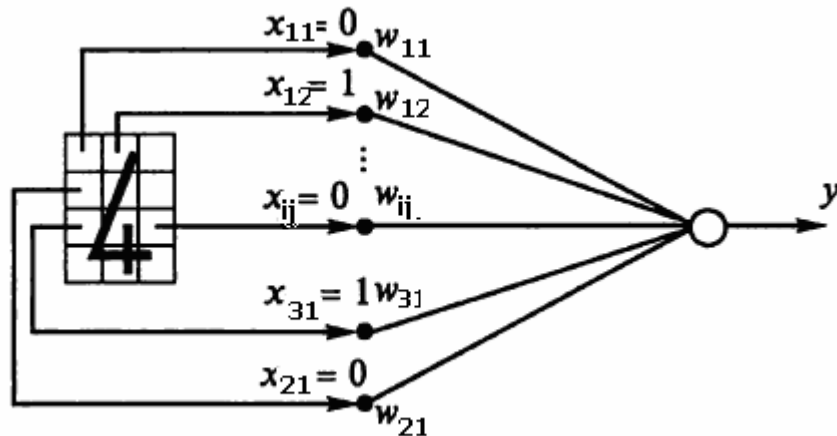
ნეირონული ქსელის სწავლებისას ვიქცევით ანალოგიურად. სადაც სწავლების ზოგადი იდეა შემდეგში მდგომარეობს: დასაწყისში ნეირონული ქსელის შესასვლელს მიეწოდება სასწავლო X ამონარჩევი ცნობილი შედეგებით და ვაკვირდებით ქსელის გამოსავალ სიგნალს $Y = F(X)w_{ij}$. წონით კოეფიციენტების და თითოეული ნეირონის აქტივაციის ზღურბლური მნიშვნელობის რეგულირებით ქსელს ვაძულებთ მის გამოსავალზე მივიღოთ სასურველი შედეგი. ამის შემდეგ, საგამოცდო ამონარჩევით ვამოწმებთ ნეირონული ქსელის მუშაობის სიზუსტეს. მაგალითად, კლასიფიკაციის ამოცანაში ჩვენ შეგვიძლია მოვითხოვოთ, რომ ქსელმა 90%-ით სწორად

ჩაატაროს სახეების კლასიფიკაცია. პროგნოზირების ამოცანებში ჩვენი მიზანი შეიძლება იყოს პროგნოზირების სიზუსტე, რომელიც არ უნდა აღემატებოდეს წინასწარ მოცემულ მნიშვნელობას. თუ ნეირონული ქსელი პასუხობს წაყენებულ მოთხოვნებს, მაშინ მისი გამოყენება პრაქტიკულ კვლევებში უკვე შესაძლებელია.

თეორიულად ნეირონული ქსელის შექმნის საფუძველს წარმოადგენს შემდეგი მტკიცებულება: ნებისმიერი განზომილების შემავალ-გამომავალ $\{X^k, Y^k\}$ $r=1,2,\dots,N$ ერთობლიობისათვის არსებობს ორშრიანი ერთგვაროვანი ნეირონული ქსელი ნებისმიერი რაოდენობის ნეირონებით და სიგმაიდალური აქტივაციის ფუნქციით, სადაც ყოველ შემომავალი X^k ვექტორი ქმნის გამოსავალ Y^k ვექტორს. ამრიგად, მრავალგანზომილებიანი ფურქციის წარმოდგენისათვის შეიძლება გამოყენებული იყოს ერთგვაროვანი ხელოვნური ნეირონული ქსელი, რომელსაც გააჩნია მხოლოდ ერთი ფარული შრე სიგმაიდალური გადაცემის ფუნქციის ნეირონებით.

სწავლების პების წესი

განვიხილოთ სწავლების პროცესის ალგორითმი მეტად მარტივ როზენბლატის ერთშრიან პერსეპტონის მაგალითზე. ქვემოთ მოყვანილ ნახაზე ნაჩვენებია პერსეპტონის უმარტივესი ვარიანტი, რომლის დანიშნულებაა გაარჩიოს რიცხვები ერთმანეთისაგან.



წარმოვიდგინოთ მატრიცა, რომელიც შეიცავს 12 ფოტოელემენტს, რომლებიც განლაგებული არიან 4 კორიზონტალურ მწკრივში. ყოველ მწკრივში მოთავსებულია 3 ფოტოელემენტი. ფოტოელემენტების მატრიცაზე ხდება ბარათის ზედღება, სადაც წარმოდგენილია ციფრი (ჩვენი მაგალითისათვის რიცხვი 4). თუ ფოტოელემენტზე მოხვდება ციფრის რომელიმე ფრაგმენტი, მაშინ ეს ფოტოელემენტი გამოიძუშავებს სიგნალს, რომელიც ერთის ტოლია, წინააღმდეგ შემთხვევაში – ნულის. როგორც ნახაზიდან ჩანს, პირველი ფოტოელემენტი იძლევა $x_{11}=0$ სიგნალს, მეორე ფოტოელემენტი – $x_{12}=1$ სიგნალს და ა.შ.

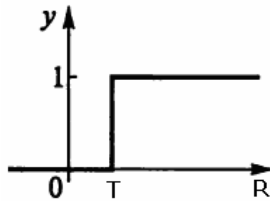
ამის შემდეგ სიგნალები გამრავლდება თავის შესაბამის წონით კოეფიციენტებზე და ხდება მათი აჯამება.

$$R = \sum_{j=1}^n w_j x_j$$

მიღებული ჯამური R სიგნალი დარდება ზღურბლურ მნიშვნელობას და თუ შედეგი დამაკმაყოფილებელია, მაშინ გამოსავალი სინალი $Y = 1$, წინააღმდეგ შემთხვევაში $Y = 0$. ე.ი.

$$Y = \begin{cases} 1, & \text{როცა } R \geq 0, \\ 0, & \text{როცა } R < 0. \end{cases}$$

როგორც Y ლოგიკურ გამოსახულებიდან ჩანს, ამ შემთხვევაში გამოიყენება შემდეგი სახის კიბისებური აქტივაციის ფუნქცია:



პერსეპტრონის მიზანია გამოსავალი სიგნალი იყოს $Y = 1$ ტოლი, როცა ბარათზე აღბეჭდილია ციფრი 4 და $Y = 0$ – სხვა რიცხვების დროს.

ეს მიზანი მიიღწეა პერსეპტრონის სწავლებით, რომელიც მდგომარეობს წონითი w_j კოეფიციენტების კორექტირებაში. თუ გამოსავალი სიგნალი $Y = 1$, მაშინ კოეფიციენტების კორექტირება არ არის საჭირო, რადგან პერსეპტრონის რეაქცია სწორია. მაგრამ, თუ გამოსავალი სიგნალი არასწორია, ე.ი. $Y = 0$, მაშინ საჭიროა იმ წონითი კოეფიციენტების გაზრდა, როლებიც ახდენენ ნეირონის ადგზნებას. ჩვენ შეთხვევაში საჭიროა w_{12} , w_{31} , w_{41} და ა.შ. კოეფიციენტების კორექტირება, ანუ ამ შეთხვევაში მათი მნიშვნელობების გაზრდა.

ამრიგად, შეგვიძლია ჩამოვყალიბოთ წონითი კოეფიციენტების კორექტირების შემდეგი იტერაციული პროცედურა:

ბიჯი 1. წარმოვადგინოთ შემავალი სახე და განვსაზღვროთ პერსეპტრონის გამოსავალი Y სიგნალი.

ბიჯი 2. ა) თუ გამოსავალი სწორია, მაშინ გადავიდეთ 1 - ბიჯზე.

ბიჯი 2. ბ) თუ გამოსავალი სიგნალი არასწორია და ნულის ტოლია, მაშინ იზრდება გააქტივებული შემომავალი სინალების კოეფიციენტები, მაგალითად, წონით კოეფიციენტებს დაუმატოთ ყველა შემომავალი სინალი:

$$w_j(t+1) = w_j(t) + x_j$$

ბიჯი 2. გ) თუ გამოსავალი სიგნალი არასწორია და ერთის ტოლია, მაშინ საჭიროა შემომავალი სიგნალის წონითი კოეფიციენტების შემცირება, მაგალითად ასე:

$$w_j(t+1) = w_j(t) - x_j$$

ბიჯი 3. გადავიდეთ 1 - ბიჯზე ან დავამთავროთ შესწავლის პროცესი.

მოყვანილ ალგორითმში ბიჯ 2. ბ)-ს უწოდებენ ჰების პირველ წესს, ხოლო ბიჯ 2. გ) – ჰების მეორე წესს. ეს ალგორითმი პირველად შემოვთავაზა ჰებმა 1949წ. უნდა შევნიშნოთ, რომ ჰების წესი მოგვაგონებს ცხოველების გაწვრთნის „მათრახი – ტკბილეული“ – ის მეთოდს ან ბავშვის აღზრდის პროცესს წახალისებისა და დასჯის პრინციპის გამოყენებას.

სწავლების დელტა-წესი.

ზემოთ განხილული სწავლების ალგორითმი შეიძლება წარმოვადინოთ ზოგადი ფორმის სახით. თუ d – თი აღვნიშნავთ საჭირო გამოსავალ სინალს, მაშინ ყოველ იტერაციაზე შესაძლებელია განისაზღვროს სხვაობა საჭირო d სიგნალსა და რეალურ Y სიგნალს შორის ე.ი. $e = (d - Y)$. როცა $e = 0$, მაშინ ის შეესაბამება ბიჯ 2. ა)-ს იმ შემთხვევაში, როცა გამოსავალი სიდიდე ჭეშმარიტია. როცა $e > 0$ შეესაბამება ბიჯ 2. ბ), ხოლო როცა $e < 0$ შეესაბამება ბიჯ 2. გ) –ს.

პერსეპტრონის წავლების ალგორითმი პეზის წესის საშუალებით მთავრდება, თუ იტერაციულ პროცესს წარვმართავთ შემდეგი ფორმულებით:

$$w_j(t+1) = w_j(t) + \Delta w_j$$

$$\Delta w_j = ex_j$$

სადაც $w_j(t)$ და $w_j(t+1)$ შესაბამისად პერსეპტრონის კოეფიციენტების ძველი და ახალი მნიშვნელობებია, j - შემავალი სიგნალის რიგითი ნომერია.

შესაძლებელია მივიღოთ ანალოგიური იტერაციული ფორმულა ნეირონის T ზღვრული მნიშვნელობის რეულირებისათვის, თუკი მას განვიხილავთ, როგორც ნეირონის დამატებითი შემომავალი x_0 სინალის წონით კოეფიციენტად. x_0 სინალის მნიშვნელობა -1 ტოლია.

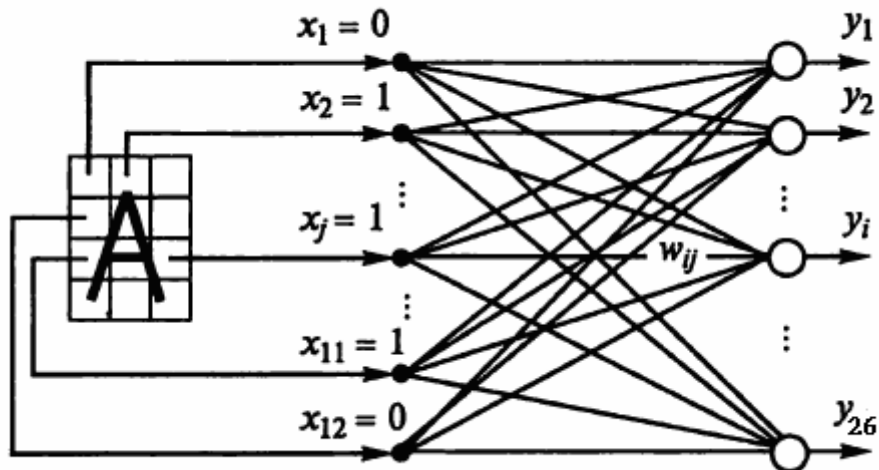
$$T(t+1) = T(t) + \Delta T, \quad \Delta T = -e.$$

იტერაციულ ფორმულებში შესაძლებელია სწავლების l კოეფიციენტის შემოტანა, რომლის საშუალებითაც შესაძლებელია წონითი კოეფიციენტების კორექტირება

$$\Delta w_j = lex_j, \quad \Delta T = -le.$$

პერსეპტრონის სწავლების ალგორითმს, რომელიც იყენებს ამ ფორმულებს, ეწოდება დელტა-წესი.

ქვემოთ მოყვანილ ნახაზზე წარმოდგენილია პერსეპტრონის სქემა, რომელიც განკუთვნილია ლათინური ალფავიტის ასოების გასარჩევად. ასეთ პერსეპტრონს გააჩნია 26 ნეირონი, ე.ი. ალფავიტის თითოეულ ასოს შეესაბამება ერთი ნეირონი. ითვლება, რომ პირველი ნეირონის გამოსავალი y_1 უნდა იყოს ერთის ტოლი როცა პერსეპტრონს გასარჩევად მიეწოდება ასო „A“ და ნულის ტოლი ყველა დანარჩენი ასოებისათვის. მეორე ნეირონის გამოსავალი y_2 უნდა იყოს ერთის ტოლი როცა ნეირონს „B“ ასო მიეწოდება და ნულის ტოლი ყველა დანარჩენი ასოებისათვის და ა.შ.



მოცემული პერსეპტრონის სწავლების ალგორითმი შედგება შემდეგი ბიჯებისაგან:

1. შემთხვევითი რიცხვების გადამწოდით ნეირონების ყველა წონით კოეფიციენტებს w_{ij} და ზღვრულ T_i სიდიდეებს მიენიჭებათ რაიმე მცირე მნიშვნელობები.

2. პერსეპტრონს მიეწოდება ალფავიტის რომელიმე ასო და ფორთოვლემენტების სისტემის საშუალებით გამოიმუშავებს გამოსავალ ვექტორს $x_j, j=1,2,...,12$.

3. თითოეული ნეირონი აწარმოებს შემომავალი სიგნალების აჯამებას:

$$R_i = \sum_{j=1}^{12} w_{ij} x_j$$

და გამოიმუშავებს გამომავალ სინალს $y_i=1$, როცა $R_i \geq T_i$ და $y_i=0$, როცა $R_i < T_i$.

4. ყოველი ნეირონისათვის განისაზღვრება ცდომილება:

$$e_i = (d_i - y_i),$$

სადაც d_i - პერსეპტრონის სწორი პასუხის ვექტორია (მაგალითად, „A“ ასოსათვის $d_1=1, d_2=0, ..., d_{26}=0$).

5. ხდება ნეირონის წონითი კოეფიციენტების და ზღვრული მნიშვნელობების კორექტირება:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}, \quad \Delta w_{ij} = |e_i x_j|,$$

$$T_i(t+1) = T_i(t) + \Delta T_i, \quad \Delta T_i = -|e_i|,$$

სადაც t – იტერაციის რიითი ნომერია.

6. გავიმეოროთ 2 – 5 ბიჯები საჭირო რაოდენობამდე.

ზემოთ მოყვანილი პერსეპტრონის სქემა, რომელიც განკუთვნილია ალფავიტის ასოების გასარჩევად, შეიძლება გამოვიყენოთ სხვა პრაქტიკული ამოცანების, მაგალითად, სამედიცინო დიაგნოსტიკური ამოცანების გადასაწყვეტად. ყველაფერი დამოკიდებულია იმაზე, თუ რა აზრს მივანიჭებთ შემომავალ x_i და გამომავალ y_i სიგნალებს.

გადასაწყვეტი ამოცანების სფერო მნიშვნელოვნად გაფართოვდება თუკი პერსეპტრონს ვასწავლით არა მარტო ბინარული მნიშვნელობების (0 და 1), არამედ უწყვეტი სიგნალების გამოტანასაც. ასეთი განზოგადოებული

პერსეპტრონის შექმნა შესაძლებელია, თუკი აქტივაციის ფუნქციად გამოვიყენებთ არა კიბისებურ, არამედ სიგმაიდალურ ფუნქციას. პრაქტიკულად სიგმაიდალური ფუნქცია უზრუნველყოფს კლასიკური ზღვრული ფუნქციის აპროქსიმაციას. სიგმაიდალურ ფუნქციისათვის მიღებულია $y = F_s(R)$ აღნიშვნა.

დელტა – წესის განზოგადოებული ალგორითმი

შემოვიტანოთ რამოდენიმე განსაზღვრება. პერსეპტრონს ეწოდება ადალანი, როცა მას გააჩნია ერთი გამოსავალი და სიგმაიდალური აქტივაციის ფუნქცია. თუ პერსეპტრონს გააჩნია მრავალი გამოსავალი, მაშინ მას ეწოდება მადალანი (ინგლისური სიტყვებიდან *ADaptive Linear Neuron* და *Many ADALINE*). ასეთი უწყვეტი აქტივაციური ფუნქციის პერსეპტრონის გამოჩენამ გამოიწვია მისი სწავლების მიმართ ახლებური მიდგომა. უიძროუმ და ჰოფმა შემოგვთავაზეს საჭირო d_i და რეალური y_i გამოსავალი სიგნალების სხვაობების საშუალო კვადრატული ცდომილების მინიმიზაცია:

$$e = \frac{1}{2} \sum_{i=1}^I (d_i - y_i)^2, \quad (5.1)$$

სადაც I – პერსეპტრონის გამოსავალი არხების რაოდენობაა.

განვიხილოთ პერსეპტრონის, რომელსაც გააჩნია J შესასვლელი და I გამოსავალი, წონითი კოეფიციენტების კორექტირების ალგორითმი. საშუალო კვადრატული ცდომილება e წარმოადგენს წონითი კოეფიციენტების მრავალგანზომილებიან ფუნქციას. ე.ი. $e = e(w_{ij})$ და w_{ij} კორდინანტთა სივრცეში გეომეტრიულად იგი წაავსებს მრავალგანზომილებიან ზედაპირს – ჰიპერზედაპირს.

პერსეპტრონის სწავლება შეიძლება წარმოადგინოთ ისეთი წონითი კოეფიციენტების w_{ij} მოძებნაში, რომლებიც შეესაბამებიან ჰიპერზედაპირის ყველაზე მინიმალურ წერტილს. ასეთი ტიპის ამოცანას ეწოდება ოპტიმიზაციის ამოცანა, რომლის დანიშნულებაა w_{ij} პარამეტრების სივრცეში $e = e(w_{ij})$ ფუნქციის მინიმიზაცია.

არსებობს ოპტიმიზაციის ამოცანის ამოხსნის მრავალი მეთოდი. ველაზე მარტივ მეთოდს წარმოადგენს წონითი კოეფიციენტების შემთხვევითი ცვლილება და შემდეგ ცდომილების გამოთვლა. მაგრამ, უფრო ეფექტური მეთოდია გრადიენტის დაშვების მეთოდი, რომლის თანახმად ყოველი წონითი კოეფიციენტის ცვლილება Δw_{ij} წარმოებს ჰიპერზედაპირის გრადიენტის საწინააღმდეგო მიმართულებით. ე.ი.

$$\Delta w_{ij} = -h \frac{\partial e}{\partial w_{ij}}, \quad (5.2)$$

სადაც h – სწავლების სიჩქარის კოეფიციენტი.

საშუალო კვადრატული ცდომილება e წარმოადგენს რთულ ფუნქციას, რომელიც პირველ რიგში დამოკიდებულია პერსეპტრონის გამოშვებულ y_i სიგნალებზე. მიტომ:

$$\frac{\partial e}{\partial \Delta w_{ij}} = \frac{\partial e}{\partial y_i} \frac{\partial y_i}{\partial \Delta w_{ij}} \quad (5.3)$$

აქ $y = f_s(R)$, სადაც $R_i = \sum_{j=1}^J w_{ij} x_j$. მაშასადამე:

$$\frac{\partial y_i}{\partial w_{ij}} = \frac{\partial f_s(R)}{\partial R_i} \frac{\partial R_i}{\partial w_{ij}} = f_s^T(R) x_j. \quad (5.4)$$

თუ (4.1) გამოსახულებას გავაწარმოებთ y_n -ით, სადაც $n \in [1, J]$, მაშინ მივიღებთ:

$$\frac{\partial e}{\partial y_n} = -(d_n - y_n), \text{ ე.ი. } \frac{\partial e}{\partial y_i} = -(d_i - y_i). \quad (5.5)$$

თუ ჩავსვამთ (5.4) და (5.5) (5.3) ფორმულაში და მიღებულ შედეგს (5.2) ფორმულაში, საბოლოოდ მივიღებთ:

$$\Delta w_{ij} = -h[-(d_i - y_i) f_s(R) x_j] = h(d_i - y_i) f_s^T(R) x_j. \quad (5.6)$$

ეს გამოსახულება მიღებულია ნეირონებისათვის, რომელთაც გააჩნიათ ნებისმიერი ტიპის აქტივაციური ფუნქციები. აგრამ, თუ $f_s(R)$ სიგმაიდალური ფუნქციაა, მაშინ

$$f_s^T(R) = \left(\frac{1}{1 + \exp(-R)} \right) = f(R)(1 - f(R)). \quad (5.7)$$

თუ ამ გამოსახულებას ჩავსვამთ (5.6) ფორმულაში, მაშინ მივიღებთ:

$$\Delta w_{ij} = h(d_i - y_i) f_s(R)(1 - f(R)) x_j = h(d_i - y_i) y_i (1 - y_i) x_j \quad (5.8)$$

ამრიგად მივიღეთ ერთშრიანი პერსეპტრონის სწავლების შემდეგ იტერაციული ფორმულა:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}, \quad (5.9)$$

სადაც $\Delta w_{ij} = h d_i x_j$, $d_i = (d_i - y_i) y_i (1 - y_i)$. (5.10)

ამ ალგორითმს ეწოდება დელტა - წესის განზოგადოებული ალგორითმი, რომლის უპირატესობა მდგომარეობს იტერაციის უფრო სწრაფ კრებალობაში და შემაჯავლი და გამომაჯავლი უწყვეტი სიგნალების უფრო ზუსტ დამუშავებაში. ყოველივე ეს იწვევს პერსეპტრონის მიერ ამოცანების ამოხსნის სფეროს გაფართოებას.

ერთშრიანი ნეირონული ქსელის შეზღუდვა

როგორც უკვე ავღნიშეთ, ფ. როზენბლანტმა შესძლო ერთშრიანი პერსეპტრონის საშუალებით ალფავიტის ასოების გარჩევა. ეს იმ დროისათვის იყო მნიშვნელოვანი მიღწევა და წინგადადგმული ნაბიჯი ადამიანის აზროვნების შემეცნების სფეროში. მაგრამ, პერსეპტრონის მიერ გადასაწყვეტი ამოცანების სფერო თანდათან ფართოვდებოდა. სამეცნიერო სფეროს გაშლასთან ერთად წარმოიშვა სიძნელეები, კერძოდ აღმოჩნდა, რომ ზოგიერთი ახალი ამოცანის გადაწყვეტა პერსეპტრონს არ შეუძლია, თუმცა ეს ამოცანები გარეგნულად მნიშვნელოვნად არ განსხვავდებოდნენ იმ ამოცანებისაგან, რომლებსაც პერსეპტრონი წარმატებულად წყვეტდა. საჭირო გახდა წარმოშობილი პარადოქსის ახსნა, სიღრმისეული ანალიზი და ნეირონული ქსელის თეორიული ბაზის შექმნა.

პერსეპტრონის განვითარების შემდეგი ეტაპი იყო მ.მინსკის და ს.პაიპერტის მიერ გამოქვეყნებული წიგნის „პერსეპტრონები“-ს გამოქვეყნება. ამ წიგნში მათემატიკური სიზუსტით დამტკიცებული იყო ის, რომ პრინციპში ერთშრიან პერსეპტრონს არ ძალუმს მრავალი პრაქტიკული ამოცანების გადაწყვეტა. მათ შორის იყო ლოგიკური ოპერაციის „გამომრიცხავი ან“-ის რეალიზაცია.

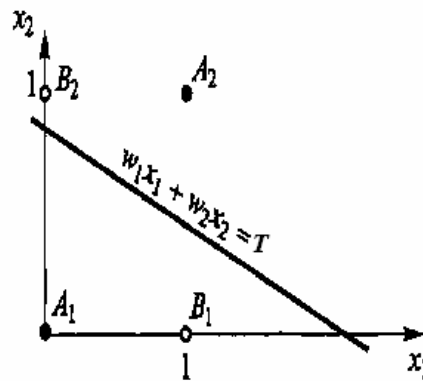
„გამომრიცხავი ან“ წარმოადგენს ბულის ორარგუმენტიან ფუნქციას, რომელმაც შეიძლება მიიღოს „ჭეშმარიტი“ ან „მცდარი“ მნიშვნელობა.

„ჭეშმარიტ“ მნიშვნელობას ფუნქცია ღებულობს მაშინ, როცა ფუნქციის ერთ-ერთი არგუმენტი (და არა ორივე) არის „ჭეშმარიტი“. სხვა შემთხვევებში ფუნქცია ღებულობს „მცდარ“ მნიშვნელობას. ფუნქციას აქვს შემდეგი სახე:

$$y = (x_1 \wedge x_2) \vee (x_2 \wedge x_1) \quad (1)$$

ამოცანა მდგომარეობს ერთშრიანი პერსპექტრით მოვახდინოთ (1) ფუნქციის რეალიზირება ორი x_1, x_2 შემაჯავლი და ერთი გამომაჯავლი y მნიშვნელობით.

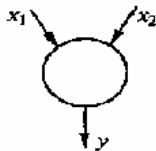
თუ ავღნიშნავთ „ჭეშმარიტ“ სიდიდეს ერთით, ხოლო „მცდარს“ ნულით, მაშინ შემაჯავლი სინალების ყველა შესაძლო კომბინაციები (x_1, x_2) სიბრტყეზე შეიძლება წარმოვადგინოთ ოთხი A_1, A_2, B_1, B_2 წერტილის სახით, როგორც ეს ნახვენებია შემდეგ ნახაზე:



მაგალითად, A_1 წერტილს შეესაბამება შემაჯავლი სიგნალების $x_1 = 0$ და $x_2 = 0$ მნიშვნელობები, ხოლო A_2 წერტილს კი $x_1 = 1$ და $x_2 = 1$. (1) ფორმულის თანახმად, პერსპექტრონის შემაჯავლი და გამომაჯავლი სიგნალების შესაბამისობები მოყვანილია შემდეგ ცხრილში:

წერტილები	x_1	x_2	y
A_1	0	0	0
A_2	1	1	0
B_1	1	0	1
B_2	0	1	1

ერთშრიანი პერსპექტრონი, რომელიც გამოსახულია ქვემოთ მოყვანილ ნახაზე.



აწარმოებს შემდეგ გარდაქმნებს: $R = w_1x_1 + w_2x_2$ (2). $y = 1$, როცა $R \geq T$ და $y = 0$, როცა $R < T$. თუ (2) განტოლებაში R -ს შევცვლით T -თი, მაშინ მივიღებთ: $w_1x_1 + w_2x_2 = T$ (3).

თუ მიღებულ (3) განტოლებაში x_1 და x_2 ცვლადებია, ხოლო T, w_1 და w_2 მუდმივები, მაშინ (x_1, x_2) სიბრტყეზე (3) განტოლება გამოისახება სწორი ხაზით, რომლის დახრილობა განისაზღვრება w_1 და w_2 წონითი კოეფიციენტებით და T ზღვრული მნიშვნელობით. იმ წერტილებებისათვის, რომლებიც ამ სწორ ხაზზე მდებარეობენ, სრულდება $R = T$ ტოლობა და პერსპექტრონის გამოსავალი ერთის ტოლია. იმ წერტილებებისათვის, რომლებიც განლაგებული არიან სწორი ხაზის ქვემოთ

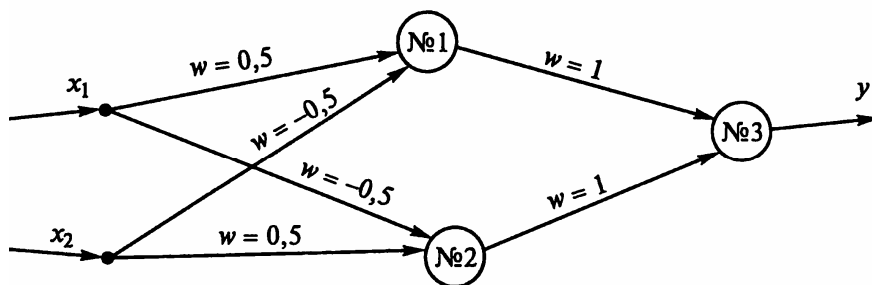
პერსექტრონის გამოსავალი სიდიდე ნულის ტოლია. იმ წერტილებებისათვის, რომლებიც განლაგებული არიან სწორი ხაზის ზემოთ პერსექტრონის გამოსავალი სიდიდე ერთის ტოლია. აქედან გამომდინარე, (3) განტოლებას უწოდებენ ზღვრულ სწორ ხაზს.

ზემოთ მოყვანილი ცხრილის თანახმად, A_1 და A_2 წერტილებში პერსექტრონის გამოსავალი უნდა იყოს ნულის ტოლი, ხოლო B_1 და B_2 წერტილებში – ერთის ტოლი. მაგრამ, ამისათვის ზღვრული სწორი ხაზი ისე უნდა იყოს წარმოდგენილი, რომ A_1 და A_2 წერტილები უნდა იმყოფებოდნენ ამ ხაზის ქვემოთ, ხოლო B_1 და B_2 წერტილები – ზემოთ, რაც ყოველად შეუძლებელია. ეს იმას ნიშნავს, რომ რა მნიშვნელობებიც არ უნდა მივანიჭოთ წონით კოეფიციენტებს და ზღვრულს, ერთშრიან ნეირონულ ქსელებს არ შეუძლია „გამომრიცხავი ან“ ფუნქციის წარმოდგენა.

გარდა „გამომრიცხავი ან“ ფუნქციის რეალიზების პრობლემისა, მინსკის წიგნში მოყვანილია სხვა ამოცანებიც, რომელთა გადაწყვეტა ერთშრიან ნეირონულ ქსელებს არ შეუძლიათ.

5.2 მრავალშრიანი ნეირონული ქსელის სწავლება

„პერსექტრონები“ წიგნის გამოსვლამ მეცნიერებს შორის შოკი გამოიწვია. საყოველთაო ოპტიმიზმი შეცვალა პესიმიზმმა, რამაც გამოიწვია ნეირონული ქსელის განვითარების გარკვეული შეფერხება. მიუხედავად ამისა, ცალკეული მეცნიერები მაინც განაგრძობდნენ მეცნიერულ კვლევებს. ბევრ მკვლევარს ესმოდა, რომ საჭირო იყო პერსექტრონის სტრუქტურის გართულება. აღმოჩნდა, რომ „გამომრიცხავი ან“ ფუნქციის პრობლემა შეიძლება გადაწყდეს ორშრიანი ნეირონული ქსელის საშუალებით, მაგალითად ისე, როგორც ეს წარმოდგენილია შემდეგ ნახაზზე:



წარმოდგენილი პერსექტრონის მუშაობა სრულდება შემდეგი ალგორითმით:

პერსექტრონი 1 :

$$R_1 = 0,5x_1 + (-0,5)x_2,$$

$$y_1 = 1, \text{ როცა } R_1 \geq T,$$

$$y_1 = 0, \text{ როცა } R_1 < T.$$

პერსექტრონი 2 :

$$R_2 = (-0,5)x_1 + 0,5x_2,$$

$$y_2 = 1, \text{ როცა } R_2 \geq T,$$

$$y_2 = 0, \text{ როცა } R_2 < T.$$

პერსექტრონი 3 :

$$R_3 = 1y_1 + 1y_2,$$

$$y_3 = 1, \text{ როცა } R_2 \geq T,$$

$$y_3 = 0, \text{ როცა } R_2 < T.$$

ამ ფორმულების საშუალებით, ადვილად შეიძლება წარმოვადგინოთ პერსეპტრონის შემავალი და გამოშვებული სიგნალების შესაბამისობები მოცემული $T=0,5$ ზღვრული მნიშვნელობისათვის შემდეგ ცხრილში:

x_1	x_2	R_1	R_2	y_1	y_2	R_3	y_3	y
0	0	0	0	0	0	0	0	0
0	1	-0,5	0,5	0	1	1	1	1
1	0	0,5	-0,5	1	0	1	1	1
1	1	0	0	0	0	0	0	0

მკვლევარებს ესმოდათ, რომ მრავალშრიანი ნეირონული ქსელები აფართოებენ ამოსახსნელ ამოცანათა კლასს, მაგრამ პრობლემა იქნებოდა ასეთი ნეირონული ქსელების სწავლება. მარტივი ჰობის წესი და მისი მოდიფიცირებული დელტა – წესი ვარგოდა მხოლოდ გამოსავალი შრის ნეირონების წონითი კოეფიციენტების კორექტირებისათვის, მაშინ როცა ნეირონული ქსელის შიგა(ფარული) შრეებისათვის იგი გამოუსადეგარი იყო.

ცდომილების უკუგავრცელების ალგორითმი.

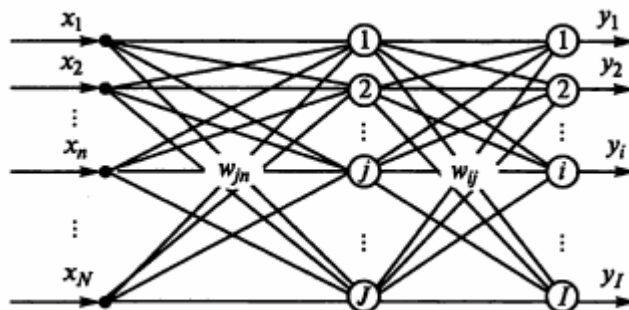
მრავალშრიანი ნეირონული ქსელის სწავლების ეფექტული ალგორითმი შეიქმნა 1986წ. რუმელჰარდის მიერ, რომელსაც ეწოდება ცდომილების უკუგავრცელების ალგორითმი. მეთოდმა თავისი დასახელება მიიღო იმიტომ, რომ მისი ფუნქციონირების დროს ქსელის გამოსავალი ცდომილება, რომელიც განისაზღვრება იტერაციის ყოველ ბიჯზე, ვრცელდება ნეირონულ ქსელში გამოსავალი შრიდან შემოსასვლელამდე (ე.ი. სიგნალის გავრცელების საწინააღმდეგოდ). სწორედ ამ დროს ხდება ფარული შრეების ნეირონების წონითი კოეფიციენტების განსაზღვრა.

განვიხილოთ ცდომილების უკუგავრცელების ალგორითმის იდეა ორშრიანი ნეოქსელის სწავლების დროს. ორშრიან ქსელს გააჩნია R შემავალი, I გამოშვებული და J ფარული შრის ნეირონები. ამოსავალი შრის სინაფსური წონების კოეფიციენტების კორექტირება დავტოვოთ იგივე, როგორც ერთშრიანი პერსეპტრონის დროს, ოღონდ x_j შევცვალოთ y_j სიმბოლოთი. ე.ი. გვექნება:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}$$

$$\Delta w_{ij} = h d_i y_i$$

$$d_i = (d_i - y_i) y_i (1 - y_i) \quad (5.11)$$



ფარული შრის სინაფსური ნეირონების კოეფიციენტები შევცვალოთ მათი კორექტირება იგივე ფორმულებით, სადაც i – ური ინდექსი შევცვალოთ j –თი, ხოლო $j - n$ – ით. ე.ი.

$$\Delta w_{ij} = \text{hd}_i y_n, \quad d_j = (d_j - y_j) y_j (1 - y_j).$$

ჩხადია, რომ ბოლო გამოსახულებაში y_n მაგივრად უნდა გამოვიყენოთ x_n . მაგრამ აქ გაურკვეველია $(d_j - y_j)$ ნეირონული ცდომილების განსაზღვრა, რომელიც ფარული შრისათვის უცნობია. ავტორის იდეა მდგომარეობდა იმაში, რომ ამ ცდომილების მაგივრად გამოვიყენოთ გამოსავალი შრის ჯამური ცდომილება გამრავლებული შესაბამის სინაფსური კავშირების ძალაზე. ე.ი.

$$(d_j - y_j) = \sum_{i=1}^I d_i w_{ij}.$$

ამრიგად ფარული შრისათვის საბოლოოდ გვექნება:

$$\Delta w_{jn} = \text{hd}_j x_n$$

$$d_j = y_j (1 - y_j) \sum_{i=1}^J d_i w_{ij} \quad (5.12)$$

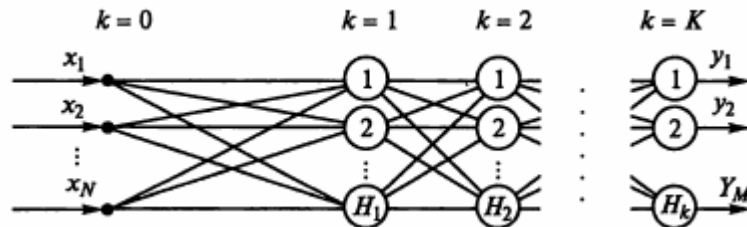
თუ ამ იდეას გამოვიყენებთ, ადვილია ნებისმიერი რაოდენობის ფარული შრიანი ნეიროქსელების სწავლებისათვის ცდომილების უკუგავრცელების ალგორითმის ჩამოყალიბება. ნეირონის მიერ გამოთვლილ ჯამს სასურველია დაუმატოთ რაიმე სიდიდე, რომელსაც წანაცვლება (ზღურბლი ადებული მინუს ნიშნით) ეწოდება

$$R = \sum_{j=1}^J x_j w_{ij} + w_{i0}. \quad (5.13)$$

w_{i0} წანაცვლება მოიცემა ისევე, როგორც სინაფსური წონები w_{ij} , ანუ შემთხვევითი რიცხვების გენერატორის საშუალებით. ნეირონის გარდაქმნის ფორმულაში წანაცვლების შეტანა შეიძლება ინტერპრეტირებული იყოს, როგორც კიდევ ერთი შემაჯავალი x_0 სიგნალის დამატება, რომელიც ყოველთვის ერთის ტოლია. ამიტომ (5.13) გამოსახულება შეგვიძლია უფრო კომპაქტურად წარმოვადგინოთ:

$$R = \sum_{j=1}^J x_j w_{ij}, \quad \text{თუ მივიღებთ } x_0 = 1.$$

ცდომილების უკუგავრცელების ალგორითმი წარმოვადგინოთ მრავალშრიან ნეიროქსელისათვის, რომელსაც გააჩნია შემაჯავალი შრე $k=0$, რამოდენიმე ფარული შრე $k=1, 2, \dots, K-1$ და გამოსავალი შრე $k=K$. ასეთი ნეირონული ქსელი წარმოდგენილია შემდეგ ნახაზე:



შემაჯავლი შრის ნეირონები მათემატიკურ გარდაქმნებს არ აწარმოებენ, ისინი მხოლოდ გადასცემენ შემომაჯავლ სიგნალებს პირველი შრის ნეირონებს. ჩავთვალოთ, რომ ყოველი k -ური შრე შეიცავს H_k ნეირონებს. ამრიგად, ქსელს გააჩნია $N=H_0$ შესავალი და $M=H_k$ გამოსავალი. ალგორითმში გამოყენებულია შემდეგი აღნიშვნები: $i - (k-1)$ -იური შრის ნეირონის რიგითი ნომერი, $l - (k+1)$ -იური შრის ნეირონის რიგითი ნომერი.

ბიჯი 1. წანაცვლების და სინაფსური წონების ინიციალიზაცია. ციკლში $k=1,2,\dots,K$, $i=1,2,\dots,H_k$, $j=1,2,\dots,H_{k-1}$ წანაცვლებას და სინაფსური w_{ij}^k წონებს შემთხვევითი რიცხვების გენერატორით ენიჭებათ მინიმალური მნიშვნელობები, მაგალითად $[-1, 1]$ ინტერვალიდან.

ბიჯი 2. სასწავლო ამონარჩევიდან ქსელს მიეწოდება შემაჯავლი $X_q = (x_1, x_2, \dots, x_N)_q$ ვექტორი და მისი შესაბამისი სასურველი გამოსავალი ვექტორი $D_q = (d_1, d_2, \dots, d_M)_q$ ვექტორი, სადაც q - სასწავლო ამონარჩევის მაგალითის რიგითი ნომერია.

ბიჯი 3. პირდაპირი გატარება. ციკლში $k=1,2,\dots,K$, $i=1,2,\dots,H_k$ გამოითვლება k -ური შრის i -ური ნეირონის გამოსავალი სიგნალები:

$$y_i^{(k)} = f_s \left(\sum_{j=0}^{H_{k-1}} w_{ij}^{(k)} y_j^{(k-1)} \right),$$

სადაც $y_j^{(0)} = x_j$, $x_0 = 1$, $y_0^{(k-1)}$ და ქსელის გამოსავალი სიგნალები $y_i = y_i^{(r)}$.

ბიჯი 4. უკუგატარება. ციკლში $k=K, K-1, \dots, 1$, $i=1,2,\dots,H_k$, $i=1,2,\dots,H_{k-1}$ განისაზღვრებიან სინაფსური წონები:

$$w_{ij}^{(k)}(t+1) = w_{ij}^{(k)}(t) + \Delta w_{ij}^{(k)}(t+1),$$

სადაც $\Delta w_{ij}^{(k)}(t+1) = \text{hd}_i^{(k)} y_j^{(k-1)}$,

თანაც გამოსავალი შრისათვის $k=K$ (5.11) თანახმად:

$$d_i^{(k)} = (d_i - y_i) y_i (1 - y_i),$$

ხოლო სხვა შემთხვევისათვის (5.12) თანახმად გვექნება:

$$d_i^{(k)} = y_i^{(k)} (1 - y_i^{(k)}) \sum_{l=1}^{H_{k+1}} d_l^{(k+1)} w_{li}^{(k+1)}.$$

ბიჯი 5. გავიმეოროთ 2 - 4 ბიჯები საჭირო რაოდენობით.

ალგორითმის მეორე ბიჯზე ხდება ახალი ვექტორის მიწოდება ქსელის შესასვლელზე, მეორდება 2 - 5 ბიჯები და ა.შ. მანამ სანამ არ მოხდება ყველა გამოსაცნობი სახის მიწორება ქსელის შესასვლელზე. მაგალითად, თუ გვინდა ქართული ალფავიტის ასოების გამოცნობა, მაშინ სწავლების ალგორითმი უნდა გავიმეოროთ 33 - ჯერ.

ალგორითმის მეხუთე ბიჯზე უკვე შესწავლილ მაგალითებზე განისაზღვრება გასაშუალოებული საშუალო კვადრატული ცდომილება:

$$e = \frac{1}{Q} \sum_{q=1}^Q \sum_{i=1}^M (d_i - y_i)^2, \quad (5.13)$$

სადაც Q - მაგალითების საერთო რაოდენობაა (ქართული შრიფტისათვის $Q = 33$).

გარდა საშუალო კვადრატული ცდომილებისა, შესაძლებელია განისაზღვროს აბსოლუტური ცდომილების მაქსიმალური მნიშვნელობა:

$$e = \max([d_i - y_i]_q), \quad i = 1, 2, \dots, M, \quad q = 1, 2, \dots, Q. \quad (5.14)$$

იტერაციული პროცესი მთავრდება, როცა ϵ განსაზღვრული (5.13) ან (5.14) ფორმულებით მიაღწევს წინასწარ დაშვებულ მნიშვნელობას.

ესლა დაუშვათ, რომ ქსელის შესასვლელ ფოტოელემენტების ტაბლოზე მიეწოდა რომელიმე ასო, რომელიც შესრულებულია სხვა შრიფტით. ფოტოელემენტები ჩამოაყალიბებენ X შემავალ სიგნალს, რომელიც არ დაემთხვევა არც ერთ სწავლების ასოების ვექტორს. თუ შრიფტი ძალზედ არ განსხვავდება შესასწავლი შრიფტებისაგან და ნეიროქსელი კარგად არის დაპროექტირებული და ნასწავლი, მაშინ ის განსაზღვრავს Y ვექტორს, სადაც ნეირონის გამოსავალი, რომელიც შეესაბამება შესასვლელ ასოს, ექნება მაქსიმალური მნიშვნელობა. ამრიგად, მიუხედავად შემომავალი სახის დამახინჯებისა და დაბრკოლებებისა ნეიროქსელი იძლევა სწორ დასკვნას სახის რომელიმე კლასთან მიკუთვნების შესახებ. ნეირონული ქსელის თვისებას, როდესაც სწავლაში არ მონაწილე შემავალი სახეების მიმართ ღებულობს სწორ გადაწყვეტილებას, ეწოდება განზოგადოების თვისება.

ცდომილების უკუგავრცელების ალგორითმის პრობლემები.

მიუხედავად იმისა, რომ ცდომილების უკუგავრცელების ალგორითმის მარტივი პროცედურისა და მისი გამოყენება ფართო კლასის ამოცანების ამოსხნისათვის, მას გააჩნია სერიოზული ნაკლოვანებები. უპირველეს ყოვლისა უნდა აღინიშნოს ნეირონული ქსელის სწავლების დიდი დრო, მაგალითად რთული ამოცანების ამოსახსნელად ზოგჯერ სჭირდება დღეები ან სულაც შეიძლება ვერ ისწავლოს. ეს ნაკლოვანებები შეიძლება გამოწვეული იყოს შემდეგი, ქვემოთ მოყვანილი ფაქტების გამო.

1. ქსელის პარალიზი (დამბლა). ნეიროქსელების სწავლების პროცესში წონითი კოეფიციენტებმა შეიძლება მიიღონ ძალზე დიდი მნიშვნელობები. ამან შეიძლება გამოიწვიოს ის, რომ ყველა ან ნეირონების უმეტესობა ქსელის გამოსავალზე მოგვცემს დიდ მნიშვნელობებს, სადაც აქტივაციის ფუნქციის წარმოებულის როლი იქნება საკმაოდ დაბალი. რადგან სწავლების პროცესში ცდომილება პროპორციულია ამ წარმოებულის, ამიტომ სწავლების პროცესი შეიძლება შეჩერდეს. თეორიულად ეს პრობლემა ნაკლებად არის შესწავლილი. პრაქტიკაში ამ პრობლემის გადაწყვეტის ერთ-ერთი მეთოდია სწავლების პროცესის ბიჯის შემცირება, მაგრამ ეს თავის მხრივ გამოიწვევს სწავლების დროის გაზრდას. სხვადასხვა ევრისტიკული მეთოდები ჯერ-ჯერობით განიხილებიან როგორც ექსპერიმენტალური მეთოდები.

2. ლოკალური მინიმუმები. როგორც აღვნიშნეთ, ცდომილების უკუგავრცელების ალგორითმი იყენებს გრადიენტულ მეთოდს ანუ ცდომილების ზედაპირზე წონების უწყვეტი კორექტირებით აწარმოებს ქვემოთ დაშვებას მინიმუმის მიმართულებით. რთული ამოცანის ცდომილების ზედაპირს შეიძლება გააჩნდეს რამოდენიმე მინიმუმი და ამის გამო ქსელი შეიძლება მოხდეს ისეთ ლოკალურ მინიმუმზე, რომლის გვერდით არსებობდეს უფრო დიდი სიდიდის მინიმუმი. ე.ი. საქმე გვაქვს გლობალური მინიმუმის განსაზღვრებასთან, რაც ამ მეთოდით არ მიიღწევა.

3. ბიჯის ზომა. ცდომილების უკუგავრცელების ალგორითმს გააჩნია კრებადობის თვისება. ეს ეფუძნება იმას, რომ წონების კორექტირება უნდა ხდებოდეს უსასრულოდ მცირე სიდიდით. ცხადია, რომ ამის პრაქტიკაში შესრულება შეუძლებელია. ბიჯის ზომა უნდა იყოს სასრული სიდიდის. თუ ბიჯის ზომა მცირეა, მაშინ კრებადობა საკმაოდ ნელია, ხოლო თუ ბიჯის ზომა

დიდია, მაშინ შეიძლება მოხდეს ნეიროქსელის პარალიზება ან მუდმივი არამდგრადობა.

4. დროებითი არამდგრადობა. თუ ქსელი სწავლობს, მაგალითად ალფავეიტის ასოებს, მაშინ არ აქვს აზრი ვასწავლოთ „B“, თუ ამ დროს ის ივიწყებს „A“-ს. სწავლის პროცესი ისეთი უნდა იყოს, რომ ქსელის სწავლება უნდა ხდებოდეს ყველა სასწავლო ობიექტით (ამ შემთხვევაში ასოებით) და მათ შორის უკვე ნასწავლი ობიექტებითაც. აქედან გამომდინარე, წონის კოეფიციენტების კორექტირება უნდა ხდებოდეს ყველა ობიექტისთვის, რაც მოითხოვს დამატებით მეხსიერებას. ეს მეთოდი შეიძლება აღმოჩნდეს უვარგისი როცა ნეიროქსელი იმყოფება მუდმივად ცვალებად გარემოში. ასეთ სიტუაციაში სწავლების პროცესი შესაძლებელია არასოდეს იყოს კრებადი.

დასკვნის სახით შეიძლება ითქვას, რომ მიუხედავად მრავალი ნაკლისა ცდომილების უკუგავრცელების ალგორითმი მაინც რჩება ძალზე პოპულარულ მეთოდად (განსაკუთრებით სახეთა გარჩევის ამოცანების გადასაწყვეტად), ორი პრაქტიკული მიზეზის გამო: ეს ალგორითმი რეალიზებისათვის მარტივია და ის უზრუნველყოფს რთული ამოცანების ამოხსნის ეფექტიანობას.

არსებობს მეთოდების ჯგუფი, რომლებიც არ იყენებენ მიზნობრივი ფუნქციის წარმოებულის გამოთვლას და ამიტომ მათ არაგრადიენტულ მეთოდებს უწოდებენ. მაგრამ ისინი მნიშვნელოვნად ჩამორჩებიან ეფექტიანობის თვალსაზრისით გრადიენტის მეთოდებს და ამიტომ მათი გამოყენება ნეიროქსელების სწავლებისათვის არარეკომენდირებულია. არსებობს აგრეთვე ნეიროქსელების სწავლების ევრისტიკული მეთოდები, რომლებსაც, როგორც წესი, არ გააჩნიათ მკაცრი თეორიული დასაბუთება.

5.3 ბენეზიკური ალგორითმი

ნეირონული ქსელების სწავლების ალგორითმებში პრობლემას წარმოადგენს სწავლების სიჩქარის h პარამეტრი, რომელიც განსაზღვრავს ოპტიმიზაციის ბიჯის სიგრძეს. მარტივ მეთოდს წარმოადგენს სწავლების მთელ პერიოდში h პარამეტრის მუდმივ სიდიდედ დაფიქსირება. ამ დროს მიზანშეწონილია ნეიროქსელის ყოველ შრისათვის h -ს მივანიჭოთ სიდიდე, მაგალითად შემდეგი ემპირიული ფორმულით:

$$h \leq \min \left(\frac{1}{n_i} \right),$$

სადაც n_i – შრის i -ური ნეირონის შესასვლელების რაოდენობაა.

სხვა მეთოდიკაში იგულისხმება სწავლების პროცესში h სიდიდის დინამიური ცვლილებები მიზნობრივი ფუნქციის ცდომილების გათვალისწინებით. პროცესის სწრაფი კრებადობისათვის ცდილობენ ცდომილების შემცირებისას h პარამეტრის გაზრდას, მაგრამ არ უნდა დაუშვათ მისი მნიშვნელობანი გაზრდა.

დღეისათვის ნეიროქსელების სწავლების ძირითად პრობლემას წარმოადგენს ცდომილების ფუნქციის მეტად რთული ფორმა მრავალი ლოკალური მინიმუმებით. ველა ზემოდ განხილულ მეთოდებს მივეყვართ ერთ-ერთ ლოკალურ მინიმუმამდე, რომელიც მდებარეობს სწავლების საწყის წერტილის მიდამოში. თუ ასეთი მინიმუმის პოვნის შემდეგ ნეიროქსელის

სწავლების ცდომილება არადამაკმაყოფილებელია, მაშინ წონით კოეფიციენტებს აძლევენ შემთხვევითი სიდიდის ნამატებს და სწავლება იწყება სხვა წერტილიდან. ხშირად ასეთი სწავლების პოცესი ღებულობს ხანგრძლივ ექსპერიმენტალურ ხასიათს, როდესაც იყენებენ სხვადასხვა ოპტიმიზაციის ალგორითმებს სხვადასხვა პარამეტრების გამოყენებით.

აქედან გამომდინარე, აქტუალურია გლობალური ოპტიმიზაციის მეთოდების განვითარება, ანუ ისეთი მეთოდების დამუშავება, რომლებიც საშუალებას იძლევიან მრავალექსტრემალური მიზნობრივი ფუნქციით გლობალური მინიმუმის მოძებნას. სხვა ასეთ მეთოდებს შორის ყველაზე მნიშვნელოვანია გენეტიკური ალგორითმის იდეა, რომელიც პირველად დ. ჰოლანდბამ 1970წ. შემოგვთავაზა. ალგორითმი შედგება ბუნებრივი ოპტიმიზაციური პროცესების იმიტაციაში, რომლებიც ვითარდებიან ცოცხალი ორგანიზმის ევოლუციის დროს.

როგორც ცნობილია, ევოლუციის თეორიის საფუძვლები ჩ. დარვინის მიერ იყო ფორმირებული 1859წ თავის ცნობილ ნაშრომში „სახეების წარმოშობა ბუნებრივი შერჩევის გზით“. ევოლუციის თეორიის თანახმად ბუნება ყველა ცოცხალის ოპტიმიზაციას აწარმოებს ორი ბიოლოგიური მექანიზმის – ბუნებრივი შერჩევის და გენეტიკური მემკვიდრეობის საშუალებით. ბუნებრივი შერჩევის არსი მდგომარეობს იმაში, რომ შეგუებული პირები უფრო გადარჩებიან და მოაქვთ დიდი შთამავლობა, ვიდრე ნაკლებად შეგუებულებს.

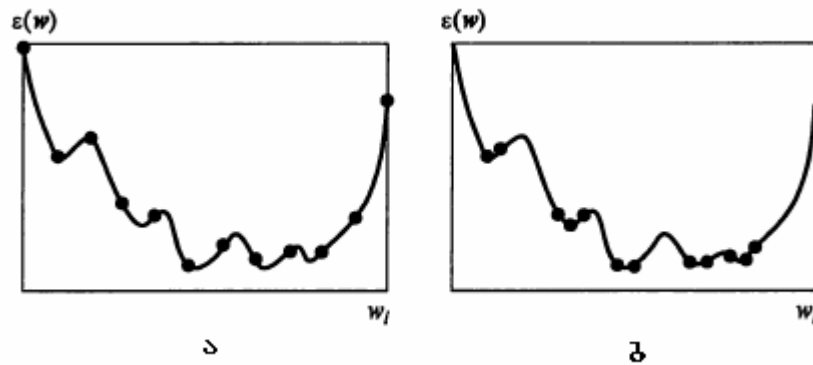
გენეტიკური მემკვიდრეობის მექანიზმი მდგომარეობს შემდეგში. ნებისმიერი ცოცხალი ორგანიზმის თითქმის ყველა უჯრედს გააჩნია ქრომოსომების ნაკრები, რომლებიც ატარებენ ინფორმაციას მთელი ორგანიზმის შესახებ. ქრომოსომის ძირითად ნაწილს წარმოადგენს დეზოქსირიბონუკლეინის მჟავის მოლეკულის (ღნმ) ძაფი, რომელიც შედგება ოთხი შენაერთისაგან – გარკვეული თანმიმდევრობით მიმდინარე ნუკლეოტიდებისაგან. ნუკლეოტიდები აღინიშნებიან A, T, C, G ასოებით და მათი მსვლელობის რიგითობით ხდება ორგანიზმის ყველა გენეტიკური თვისებების კოდირება.

გენი არის ღნმ-ს ჯაჭვის მონაკვეთი, რომელიც პასუხისმგებელია პირის გარკვეულ თვისებებზე, მაგალითად თვალის ფერზე, თმის ტიპზე, კანის ფერზე და ა.შ. დადგენილია, რომ ადამიანის გენეტიკური ნიშნების ერთობლიობა დაკოდირებულია დაახლოებით 60 ათასი გენით, რომელთა ჯამური სიგრძე შეადგენს 90 მილიონ ნუკლეოტიდებზე მეტი.

ცოცხალი ორგანიზმის გამრავლება წარმოებს მშობლების ორი სასქესო უჯრედის შერწყმით. კერძოდ, მშობლების ქრომოსომები მჭიდროდ უახლოვდებიან ერთმანეთს, შემდეგ მათი ღნმ-ის ძაფი წყდება რამოდენიმე შემთხვევით ადგილებში და ქრომოსომები ცვლიან თავიანთ ნაწილებს. ამრიგად შთამომავლის ღნმ-ის მოლეკულები შემთხვევით ღებულობენ როგორც მამის, ასევე დედის გენებს.

მემკვიდრეობის მიღებისას შესაძლებელია მუტაცია ანუ ერთ-ერთი მშობლის სასქესო უჯრედებში გენების შეცვლა. შეცვლილი გენები გადაეცემა მემკვიდრეს და ანიჭებს მას ახალ თვისებებს, რომლითაც ის განსხვავდება მშობლებისაგან. თუ ეს ახალი თვისებები აღმოჩნდებიან სასურველნი, ე.ი. შთამომავლი იქნება უფრო შეგუებული გარემოსთან და ბუნებრივი შერჩევის პროცესში ის გადარჩება, მაშინ ის შექმნის კიდევ უფრო სრულყოფილ შთამომავლებს. ამრიგად, ბუნებრივი შერჩევა, ცვალებადობა და მემკვიდრეობითი მექანიზმები წარმოადგენენ ბიოლოგიური არსებების სრულყოფილების საშუალებებს, ცოცხალი ორგანიზმის თვისებების ოპტიმიზაციის მეთოდებს, რომლებსაც ბუნება თვითონ ქმნის.

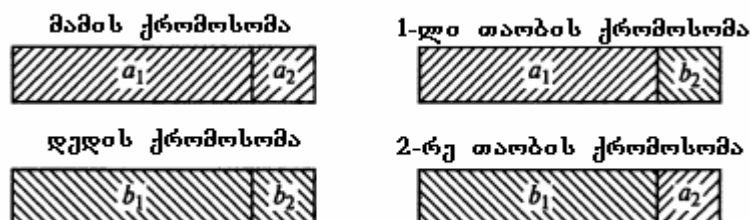
გენეტიკური ალგორითმი გამოიყენება ნეირონული ქსელის წონითი კოეფიციენტების ოპტიმიზაციისათვის. ალგორითმი მუშაობს შემდეგნაირად: ჯერ იქმნება გარკვეული რაოდენობის პირების საწყისი პოპულაცია, სადაც თითოეულ მათგანს გააჩნია თავისი საკუთარი ქრომოსომი ანუ ნეიროქსელის წონითი კოეფიციენტების ვექტორი $W = (w_1, w_2, \dots, w_n)$. თითოეული პირისათვის განისაზღვრება მიზნობრივი ფუნქცია $e(W)$, რომელიც წარმოადგენს პირის გარემოსთან შეგუების ხარისხს. პირველადი პოპულაცია თანაბრად განაწილდება ოპტიმიზირებული პარამეტრების სივრცეში. ამრიგად, წერტილები, რომლებიც შეესაბამებიან პირებს, მეტ-ნაკლებად თანაბრად ნაწილდებიან მიზნობრივი ფუნქციის ზედაპირზე, როგორც ეს ნაჩვენებია შემდეგ ნახაზზე:



აქ (ა) ნახაზზე ნაჩვენებია საწყისი პოპულაციის პირის ქრომოსომების წერტილების განლაგება მიზნობრივი ფუნქციის ზედაპირზე და ქრომოსომების განლაგება n შთამომავლის შემდეგ (ბ).

პირების შერჩევა, მათი შეჯვარებისათვის, საჭიროა ახალი თაობის შესაქმნელად და იგი შეიძლება ეფუძნებოდეს სხვადასხვა პრინციპებს. ერთ-ერთ ყველაზე გავრცელებულ პრინციპს წარმოადგენს ელიტარობის პრინციპი, რომლის თანახმად შეჯვარებისათვის დაიშვებიან ყველაზე უფრო შეგუებული პირები, ხოლო ნაკლებად შეგუებული პირები პროცესიდან გამოირიცხებიან და ხდება მათი ჩანაცვლება ახლად შექმნილი შთამომავლებით. ამ პროცესს სელექცია ეწოდება.

ზოგადად შეჯვარების პროცესი მდგომარეობს იმაში, რომ მშობლების ქრომოსომები შემთხვევითობის პრინციპით გაიყოფა ორ არათანაბარ ნაწილად, როს შედეგადაც ისინი შეერთდებიან ისე, რომ შთამომავლის ქრომოსომი შეიცავს როგორც მამის, ასევე დედის ქრომოსომის ნაწილებს, ისე როგორც ეს ნაჩვენებია ქვემოთ მოყვანილ ნახაზზე:



ამ ნახაზზე წარმოდგენილი სიტუაცია შეჯვარების შემდეგ მამის ქრომოსომი (a_1 და a_2 ფრაგმენტები) დედის ქრომოსომთან ერთად (ფრაგმენტები b_1 და b_2) ქმნიან

ახალ ქრომოსომის წყვილს, რომელთაგან პირველს გააჩნია a_1 და b_2 ფრაგმენტები, ხოლო მეორეს – b_1 და a_2 ფრაგმენტები.

შემდეგ გენეტიკურ ოპერაციას ეწოდება მუტაცია, რომელიც მდგომარეობს შემთხვევით შერჩეული ვექტორის (პირის) შემთხვევით შერჩეული ელემენტის (გენის) შეცვლაში შემთხვევით მოცემული დასაშვები მნიშვნელობით. ზოგადად, მუტაციას განიცდის მთელი პოპულაციის ქრომოსომების ბიტების არა უმეტეს 1 – 5%.

ყველა ამ გენეტიკური ოპერაციებით ფორმირდება ახალი თაობა, რომელთა პირების რაოდენობა ტოლია წინა თაობის პირების რაოდენობისა. როგორც პრაქტიკა აჩვენებს, ახალი თაობას, რომელიც შექმნილია სელექციით, შეჯვარებით და მუტაციით, საშუალოდ გააჩნია მიზნობრივი ფუნქციის მცირე მნიშვნელობები (როგორც ეს ნაჩვენებია ნახ.ბ), ე.ი. ახალი თაობა წარმოადგენს უფრო სრულყოფილს.

თაობების შეცვლის პროცესი მთავრდება როცა იტერაციის ბიჯი გახდება წინასწარ დასაშვებ რაოდენობის ტოლი ან მას შემდეგ, როცა ერთ-ერთი რომელიმე პირი იძენს მიზნობრივი ფუნქციის მინიმალურ მნიშვნელობას. ეს პირი წარმოადგენს გამარჯვებულს და მისი ქრომოსომა მიიღება გენეტიკური ალგორითმის საბოლოო შედეგად.

ნეირონული ქსელების სწავლებისას ჩვეულებრივ იყენებენ გენეტიკური ალგორითმის სხვადასხვა ვარიანტებს, რომლებიც დაკავშირებულნი არიან პარამეტრების შერჩევაზე და სელექციის, შეჯვარების და მუტაციის მეთოდებზე. შესაძლებელია გენეტიკური ალგორითმის შეთავსება სხვა ოპტიმიზაციის ალგორითმებთან. ასე მაგალითად, სწავლების საწყის სტადიაზე იყენებენ გენეტიკურ ალგორითმს, ხოლო შემდეგ სხვა ოპტიმიზაციის ალგორითმს, მაგალითად გრადიენტის მეთოდებს.

დასკვნისათვის შევნიშნოთ, რომ გენეტიკური ალგორითმები, ისე როგორც ნეირონული ქსელები, წარმოადგენენ ხელოვნური ინტელექტის ახალ მიმართულებას. ნეირონულ ქსელთან შედარებით გენეტიკურ ალგორითმებს გააჩნიათ დამოუკიდებლობა და გამოყენების ფართო სპექტრი. კერძოდ, ისინი გამოიყენებიან სხვადასხვა ტიპის ოპტიმიზაციის ამოცანების გადასაწყვეტად, რომლებიც წარმოიშებიან ბიზნესში, ეკონომიკაში, მედიცინაში და ადამიანის მოღვაწეობის სხვა სფეროებში.

6. შემხვედრი ბავრცელების ნეირონული ქსელები

6.1 ქსელის სტრუქტურა

რეალურ სიტუაციებში ნებისმიერი ცოცხალი ორგანიზმი, რომელიც მუდმივად ურთიერთქმედებს გარემოსთან, მუდმივად განიცდის გარემოს ქმედებას, რაც იწვევს მისი ქმედების კორექტირებას. მაგალითად, თუ კატა ერთხელ გადახტა ანთებულ ღუმელზე შემდგომში ის ამ ქმედებას არასოდეს არ გაიმეორებს. ასეთი მაგალითების მიყვანა შესაძლებელია მრავალჯერ, ამიტომ მასწავლებლი სწავლების მექანიზმი დამახასიათებელია ნებისმიერი ცოცხალი ორგანიზმის თავის ტვინისათვის.

მაგრამ ადამიანს თავის ტვინს გააჩნია სხვა სწავლების მექანიზმიც. ჩვენ შეგვიძლია გადავწყვიტოთ მრავალი ინტელექტუალური ამოცანა წინასწარი სწავლების გარეშე. მაგალითად, ჩვენ, დაფიქსირების გარეშე, შეგვიძლია

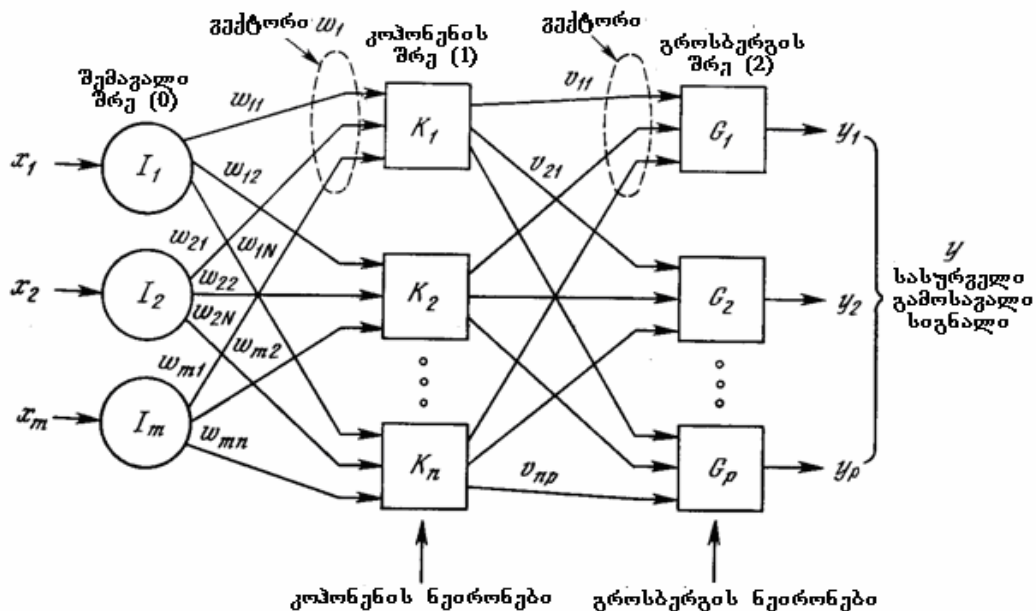
შევასრულოთ ობიექტების კლასიფიკაცია კერძოდ, ერთმანეთის მსგავსი ობიექტების გაერთიანება დამოუკიდებელ კლასებში, რომლებსაც კლასტერები ეწოდებათ.

შემხვედრი გავრცელების ნეინოქსელი (შბნძ) მიეკუთვნება ჰიბრიდულ ქსელებს, სადაც სწავლება ხორციელდება როგორც მასწავლებლით, ასევე მასწავლების გარეშე. შბნძ, რომელიც შედგება ორი ე.წ. კოჰონენის და გროსბერგის შრეებისაგან, თავისი მახასიათებლებით მნიშვნელოვნად სჭარბობს ერთი ფარული შრის მქონე ნეინოქსელებს. მაგალითად, ასეთი ქსელის სწავლების დრო, სახეთა გარჩევის და კლასიფიკაციის ამოცანების გადასაწყვეტად, ასჯერ უფრო მცირეა ვიდრე უკუკავშირის ქსელებისა. შბნძ ნაკლოვანებას წარმოადგენს ის, რომ ქსელი არ იძლევა ზუსტი აპროქსიმაციის საშუალებას. ამ შემთხვევაში ის მნიშვნელოვნად ჩამორჩება უკუკავშირის ქსელებს.

შბნძ მნიშვნელოვან თვისებას წარმოადგენს განზოგადების თვისება, რომელიც საშუალებას იძლევა მივიღოთ სწორი გამოსავალი მაშინაც კი, როცა ქსელის შესასვლელზე გვაქვს არასრული ან ხმაურის მქონე ვექტორი. სწავლების პროცესში შემაჯავლი ვექტორები ასოცირდებიან შესაბამის გამოსავალ ვექტორებად. ეს ვექტორები შეიძლება იყოს ორობითი ან უწყვეტი. სწავლების შედეგად ფორმირდება გამოსავალი ვექტორები, რომლებიც შეესაბამებიან შემაჯავალ სიგნალებს.

შბნძ გააჩნია თანმიმდევრულად შეერთებული ორი კოჰონენის და გროსბერგის შრე. ნახ. 6.1. ნაჩვენებია შბნძ გამარტივებული ვერსია. შემაჯავლი სიგნალის ყოველი ელემენტი მიეწოდება კოჰონენის შრის ყველა ნეირონს. თავის მხრივ, კოჰონენის შრის თითოეული ნეირონი მიერთებულია გროსბერგის შრის ყველა ნეირონთან.

ფუნქციონერების რეჟიმში ქსელს მიეწოდება X სიგნალი და ფორმირდება გამოსავალი სიგნალი Y . სწავლების რეჟიმში ქსელის შესასვლელს მიეწოდება შემაჯავალი სიგნალი, ხდება წონების კორექტირება რათა ქსელის გამოსავალზე მივიღოთ საჭირო სიგნალი.



ნახ. 6.1

ნულოვანი შემაჯავალი შრე არ აწარმოებს არავითარ არითმეტიკულ ოპერაციებს. მისი დანიშნულებაა მოახდინოს შემაჯავალი სიგნალის განშტოება.

ნულოვანი შრის ყოველი ნეირონი მიერთებულია პირველი შრის ყოველ ნეირონთან w_{mn} წონითი კოეფიციენტებით. ეს წონები განისაზღვრებიან როგორც W მატრიცის ელემენტები. ანალოგიურად, კოპონენის შრის (1 შრე) ყოველი ნეირონი მიერთებულია გროსბერგის შრის (2 შრე) ყოველ ნეირონთან v_{np} წონითი კოეფიციენტებით. ეს წონითი კოეფიციენტები ქმნიან V წონით მატრიცას.

როგორც სხვა ნეიროქსელები, შზნძ ფუნქციონირებს ორ რეჟიმში: ნორმალური რეჟიმი, როცა შემაგალი X ვექტორი გამოსავალზე იძლევა Y ვექტორს და სწავლების რეჟიმი, როცა ქსელს მიეწოდება შემაგალი ვექტორი და ხდება წონების კორექტირება, რათა ქსელის გამოსასვლელზე მივიღოთ საჭირო სახის გამოსავალი ვექტორი.

6.2 ქსელის ნორმალური ფუნქციონირება.

კოპონენის შრე. კოპონენის შრე ფუნქციონირებს პრინციპით „გამარჯვებული ღებულობს ყველაფერს“. მოცემული შემაგალი ვექტორისათვის კოპონენის შრის მხოლოდ ერთი ნეირონი იძლევა ლოგიკურ ერთიანს, ხოლო დანარჩენები – ნულს. კოპონენის ნეირონის გამოსავალი წარმოადგენს შემაგალი X სიგნალის ელემენტების შეწონილ ჯაჭვს

$$S_j = w_{1j}x_1 + w_{2j}x_2 + \dots + w_{nj}x_n = \sum_{i=1}^n x_i w_{ij}, \quad (6.1)$$

სადაც S_j – კოპონენის j -ური ნეირონის გამოსავალია, $W = (w_{1j}, w_{2j}, \dots, w_{nj})$ – კოპონენის j -ური ნეირონის წონების კოეფიციენტებია, $X = (x_1, x_2, \dots, x_n)$ – შემაგალი სიგნალის ვექტორი. (6.1) ფორმულის მატრიცული სახეა: $S = XW$, სადაც S – კოპონენის შრის გამოსავალი ვექტორია.

კოპონენის შრის ნეირონს, რომელსაც გააჩნია S_j მაქსიმალური მნიშვნელობა, ეწოდება „გამარჯვებული“, რომლის გამოსავალი ერთის ტოლია, ხოლო ყველა დანარჩენი ნეირონის – ნული.

გროსბერგის შრე. გროსბერგის შრეც ფუნქციონირებს იგივე მეთოდით, როგორც კოპონენის შრე. მისი გამოსავალი ვექტორი წარმოადგენს კოპონენის შრის გამოსასვლელების $k_1, k_2, \dots, k_n$ შეწონილ ვექტორს:

$$S_j = \sum_{i=1}^n x_i w_{ij},$$

სადაც S_j – გროსბერგის j -ური ნეირონის გამოსავალია. მატრიცულ სახით $Y = KV$, სადაც Y – გროსბერგის შრის გამოსავალი ვექტორია, K – კოპონენის შრის გამოსავალი ვექტორია, V – გროსბერგის შრის წონების მატრიცა.

თუ კოპონენის შრე ისე ფუნქციონირებს, რომ მხოლოდ ერთ ნეირონს გააჩნია ერთის ტოლი გამოსავალი, მაშინ K მატრიცის მხოლოდ ერთ ელემენტი განსხვავდება ნულისაგან და გამოთვლები ძალზე მარტივია. ფაქტიურად გროსბერგის შრის ყოველი ნეირონის წონა აკავშირებს ამ ნეირონს კოპონენის არანულოვან ნეირონთან.

6.3 ქსელის სწავლება

ქსელის სწავლებისას სასურველია შემავალი ვექტორების წინასწარი დამუშავება, კერძოდ მათი ნორმალიზაცია. ასეთი ნორმალიზაცია სრულდება თითოეული შემავალი ვექტორის კომპონენტების გაყოფით ვექტორის სიგრძეზე.

$$x_i^j = \frac{x_i}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}.$$

ნორმალიზება იწვევს შემავალი ვექტორის გარდაქმნას იგივე მიმართულების ერთეულოვან ვექტორად ანუ n – განზომილებიან სივრცეში ერთეულოვანი სიგრძის ვექტორად.

კოჰონენის შრის სწავლება. კოჰონენის შრის სწავლება ხდება მასწავლების გარეშე. სწავლების შედეგად შრე საშუალებას იძლევა გამოვეყნოთ რამოდენიმე არამსგავსი შემავალი ვექტორები. თუ რომელი ნეირონის აქტივაცია მოხდება წინასწარ ცნობილი არ არის.

სწავლებისას კოჰონენის შრეს მიეწოდება შემავალი ვექტორი და განისაზღვრება მისი სკალარული ნამრავლი ყველა თავის ნეირონების ვექტორებთან. სკალარული ნამრავლი წარმოადგენს შემავალი ვექტორების მსგავსების ზომას. მაქსიმალური მნიშვნელობის სკალარული ნამრავლის ვექტორი გამოცხადდება „გამარჯვებულად“ და ხდება მისი წონითი კოეფიციენტების რეგულირება ანუ წონითი ვექტორი უახლოვდება შემავალ ვექტორს.

სკალარული ნამრავლის გარდა შეიძლება გამოყენებული იყოს ევკლიდეს მანძილი. ამ შემთხვევაში შეირჩევა ის ნეირონი, რომელსაც გააჩნია უმცირესი მანძილი შემავალ ვექტორთან და ეს ვექტორი გამოცხადდება „გამარჯვებულად“.

სწავლების პროცესი აღიწერება შემდეგი განტოლებით:

$$w_N = w_c + h(x - w_c), \quad (6.2)$$

სადაც w_N – წონის ახალი მნიშვნელობაა, რომელიც აერთებს შემავალ x კომპონენტს გამარჯვებულ ნეირონთან, w_c – წონის წინა მნიშვნელობაა, h – სწავლების სიჩქარის კოეფიციენტი.

ყოველი წონა, რომელიც დაკავშირებულია კოჰონენის გამარჯვებულ ნეირონთან, იცვლება მისი მნიშვნელობის შემავალი მნიშვნელობასთან სხვაობის პროპორციულად. სწავლების სიჩქარის კოეფიციენტი ცვლადი სიდიდეა, რომელსაც დასაწყისში იღებენ 0,7 და შემდეგ სწავლების პროცესში შესაძლებელია მისი მნიშვნელობის თანდათანობით შემცირება. ეს საშუალებას გვაძლევს საწყისი ბიჯების დიდი მნიშვნელობის აღებისათვის, რათა სწრაფად მოვახდინოთ უხეში სწავლება და შემდეგ მცირე ბიჯების საშუალებით მივიღოთ სასურველი შედეგი.

თუ კოჰონენის შრის თითოეული ნეირონი ასოცირებულია ერთი შემავალი ვექტორით, მაშინ კოჰონენის შრის სწავლება შესაძლებელია ერთი ბიჯით. გამარჯვებული ნეირონის წონები უტოლდება შესასვლელი ვექტორის კომპონენტებს ($h=1$). როგორც წესი, სასწავლო ამონარჩევი შეიცავს მრავალ ერთმანეთის მსგავს ვექტორებს და ქსელის სწავლება ისე უნდა იყოს ორგანიზებული, რომ თითოეულმა შემავალმა ვექტორმა უნდა გააქტიუროს მხოლოდ ერთი და იგივე კოჰონენის ნეირონი. ამ ნეირონის წონები მიიღება შემავალი ვექტორების გასაშუალებით, რომლებმაც უნდა გააქტიურონ ის.

საწყისი წონითი ვექტორების შერჩევა. სწავლის დაწყებამდე ქსელის ყველა წონებს ანიჭებენ საწყის მნიშვნელობებს. ნეირონულ ქსელებთან მუშაობის საყოველთაო პრაქტიკას წარმოადგენს წონებისათვის მინიმალური მნიშვნელობების მინიჭება. კოჰონენის შრის სწავლისათვის ხდება შემთხვევით შერჩეული წონის ვექტორის ნორმალიზაცია. სწავლების შემდეგ წონითი ვექტორის საბოლოო მნიშვნელობა ემთხვევა ნორმალიზირებულ შემომაველ ვექტორს

კოჰონენის შრის წონების რანდომიზაციამ სწავლების დროს შეიძლება წარმოშვას სერიოზული პრობლემები. კოჰონენის იტერაციული ფორმულის (6.2) თანახმად ყოველი იტერაციის შემდეგ გამარჯვებული ნეირონი თავის წონის w_N ვექტორით სულ უფრო და უფრო უახლოვდება შემომაველ X ვექტორს. მაგრამ, შეიძლება აღმოჩნდეს ისეთი ნეირონები, რომლებიც თავისი საწყისი მდებარეობით ისე არიან დაშორებული შემავალ X ვექტორთან, რომ შეიძლება იგი არცერთხელ არ აღმოჩნდეს გამარჯვებულები. ამიტომ ასეთ ნეირონებს მკვდარ ნეირონებს უწოდებენ. მკვდარი ნეირონების არსებობა არასასურველია, რადგან მცირდება გამოთვლითი ალგორითმის ეფექტიანობა.

მკვდარი ნეირონების პრობლემის გადაწყვეტა ხდება ნეიროფიზიოლოგიურ დაკვირვებებში ცნობილი გადაღლის ეფექტის მექანიზმის მოდელირებით. ეს ეფექტი მდგომარეობს იმაში, რომ ბიოლოგიური ნეირონები გამარჯვებისთანავე გარკვეული დროით კარგავენ თავიანთ აქტიობას და არ მონაწილეობენ კონკრეტულ ბრძოლაში.

არსებობს ნეირონების გადაღლის ეფექტის მექანიზმის მოდელირების რამოდენიმე მეთოდი, მაგალითად ყველა ნეირონში აქტიური პოტენციალის შეყვანით. ყოველი შემავალი ვექტორის დროს ხდება აქტიობის პოტენციალის მოდიფიცირება

$$p_i(t+1) = \begin{cases} p_1(t) - p_{\min}, & i = w \\ p_1(t) + \frac{1}{N}, & i \neq w \end{cases}$$

ამ ფორმულაში $p_{\min}$ – პოტენციალის მინიმალური მნიშვნელობაა, რომელიც ნეირონს საშუალებას აძლევს მიიღოს კონკურენტულ ბრძოლაში მონაწილეობა. თუ რომელიმე ნეირონის პოტენციალი ეცემა $p_{\min}$ დაბლა, მაშინ ეს ნეირონი დროებით ისვენებს და თავის ადგილს უთმობს სხვა ნეირონს კონკურენტულ ბრძოლაში მონაწილეობისათვის.

ასეთი სწავლების შემდეგ ყველა ნეირონის წონითი კოეფიციენტების ვექტორები გადანაწილდებიან ისე, რომ ისინი ხდებიან სხვადასხვა კლასტერის ცენტრებად, რომლებიც შექმნილნი არიან შემავალი ვექტორებით, თანაც კლასტერების რაოდენობა ტოლი იქნება ქსელის ნეირონების რაოდენობისა.

ამრიგად, კოჰონენის შრე მასწავლების გარეშე ავტომატურად არეგულირებს თავის წონით კოეფიციენტებს, რათა გადაწყვიტოს შემომაველი ვექტორების კლასიფიკაციის ამოცანა.

გროსბერგის შრის სწავლება. გროსბერგის შრის სწავლება შედარებით ადვილია. მისი შემომაველი ვექტორი, რომელიც წარმოადგენს კოჰონენის შრის გამოსავალს, მიეწოდება გროსბერგის შრის ნეირონებს და ხდება გროსბერგის შრის გამოსავალი ვექტორის განსაზღვრა, როგორც ეს ხდება ქსელის ნორმალურ რეჟიმში მუშაობისას, როცა შემავალ X ვექტორი გამოსავალზე იძლევა Y ვექტორს. წონის კორექტირება ხდება მხოლოდ იმ შემთხვევაში, როცა ის მიერთებულია კოჰონენის შრის იმ ნეირონთან, რომელიც განსხვავდება

ნულისაგან. წონების კორექტირების სიდიდე პროპორციულია წონისა და მასთან მიერთებული გროსბერგის შრის გამოსავალი ნეირონის წონებს შორი სხვაობისა. სიმბოლურად ეს ასე ჩაიწერება:

$$v_{ijN} = v_{ijC} + h(y_i - v_{ijC})k_i,$$

სადაც k_i – ნულისაგან განსხვავებული კოჰოპეჰის შრის j -ური ნეირონის გამოსავალი, y_i – სასურველი ვექტორის j -ური კომპონენტი.

დასაწყისში იღებენ $h = 0,1$ ტოლად და შემდეგ სწავლების პროცესში თანდათანობით ამცირებენ. აქედან ჩანს, რომ გროსბერგის შრის წონითი კოეფიციენტები კრებადია სასურველი ვექტორების საშუალო მნიშვნელობასთან, მაშინ როცა კოჰონენის შრის წონების კორექტირება ხდება შემავალი ვექტორების საშუალო მნიშვნელობით.

გროსბერგის შრის სწავლება წარმოადგენს სწავლებას მასწავლებით, რომელიც იყენებს მოცემულ სასურველ გამოსავალ ვექტორებს.

7. რეკურენტული ნეირონული ქსელები

7.1 რეკურენტული ქსელის მდგრადობა

როგორც ნეიროფიზიოლოგიურმა კვლევებმა აჩვენა, ადამიანის თავის ტვინს გააჩნია გაცილებით რთული სტრუქტურა და ნეირონებშორისო ურთიერთობის მექანიზმები, კერძოდ, ბუნებრივ ნეირონებს შორის გამოვლინდა დიდი რაოდენობის არა მარტო პირდაპირი, არამედ უკუკავშირებიც.

პირდაპირი გავრცელების ნეიროქსელებს არ გააჩნიათ უკუკავშირი ანუ კავშირი გამოსავალ სიგნალებსა და შემავალ სიგნალებს შორის. უკუკავშირის არარსებობა საშუალებას იძლევა ნეირონული ქსელი მუშაობაში იყოს მდგრადი, კერძოდ, მათ არ შეუძლიათ ისეთი რეჟიმში შესვლა, როცა ქსელის გამოსავალი განუწყვეტლივ იცვლება მდგომარეობიდან მდგომარეობამდე. ქსელის არამდგრადობის მიუხედავად უკუკავშირიან ნეირონულ ქსელებს გააჩნიათ უფრო ფართო შესაძლებლობები, ვიდრე შემხვედრი გავრცელების ნეიროქსელებს.

რეკურენტული ანუ უკუკავშირის ქსელები, რომლებიც გამოსავალ სიგნალს გადაცემენ ქსელის შესასვლელს, წარმოადგენენ დინამიურ ქსელებს, რადგან ახალი შემომავალი სიგნალით განისაზღვრება გამოსავალი სიგნალი, რომელიც მიეწოდება ქსელის შესასვლელს და იწვევს შესასვლელის მოდიფიცირებას. შემდეგ გამოსავალი სიგნალი კვლავ განისაზღვრება და პროცესი მეორდება. მდგრადი ქსელებისათვის ასეთ იტერაციულ პროცესებს მიჟვავართ ისეთ სიტუაციამდე, როცა გამოსავალი სიგნალი ხდება მუდმივი.

მრავალი ნეიროქსელისათვის ასეთი იტერაციული პროცესი არასოდეს არ მთავრდება. მაშინ ქსელი არამდგრადია. არამდგრად ქსელებს გააჩნიათ საინტერესო თვისება და იგი შეიძლება განვიხილოთ როგორც ქაოტიური სისტემის მაგალითი.

ქსელების არამდგრადობამ გამოიწვია გარკვეული შეფერხებები ნეიროქსელების ტექნოლოგიის განვითარებაში. საბედნიეროდ 1983 წ. კონმა და გროსბერგმა დაამტკიცეს თეორემა, სადაც ჩამოყალიბებულია ნეირონული ქსელის მდგრადობის საკმარისი პირობები. კერძოდ, რეკურენტული ნეიროქსელი

მდგრადია, როცა წონითი მატრიცა $W = (w_{ij})$ სიმეტრიულია და მისი მთავარი დიაგონალის მნიშვნელობები ნულის ტოლია, ე.ი.

$$w_{ij} = w_{ji} \text{ ყველა } i \neq j$$

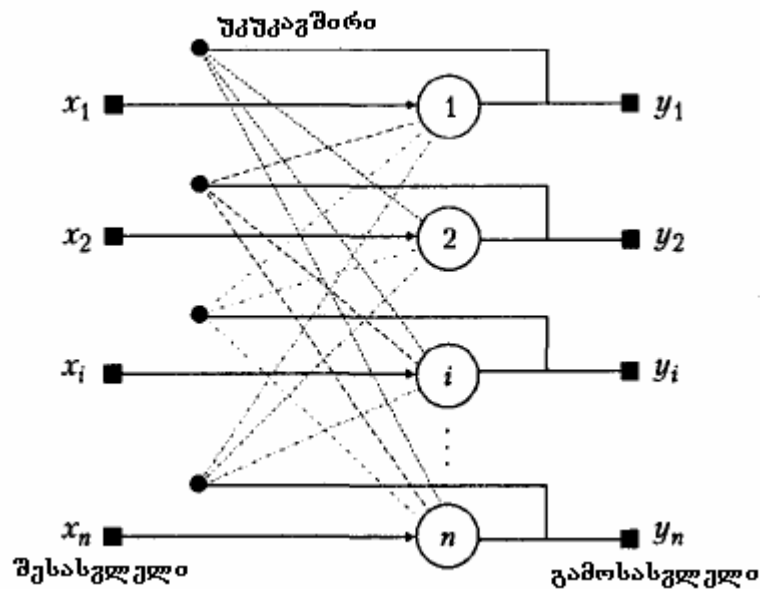
$$w_{ii} = 0, \text{ ყველა } i\text{-სთვის.}$$

უნდა აღინიშნოს, რომ თეორემის ეს პირობები საკმარისია, მაგრამ აუცილებელი არ არის, ე.ი. რეკურენტული ქსელებისათვის შესაძლებელია მათი მდგრადობა, როცა არ სრულდება ეს პირობები.

7.2 ჰოპფილდის ნეირონული ქსელი

ჰოპფილდმა არმოაჩინა, რომ იტერაციულ პროცესს შეუძლია რეკურენტული ქსელის გადაყვანა სტაციონარულ რეჟიმში. თანაც ქსელის ეს სტაციონარული მდგომარეობა დამოკიდებულია როგორც ქსელის საწყის მდგომარეობაზე, ასევე ქსელის წონის მატრიცის ელემენტების მნიშვნელობებზე.

ჰოპფილდის ნეირონული ქსელის სტრუქტურული სქემა მოცემულია ნახ. 7.1-ზე.



ნახ. 7.1

ჰოპფილდის ქსელი შედგება ნეირონების ერთი შრისაგან, რომელის ნეირონების რაოდენობა ერთდროულად წარმოადგენენ შემავალ და გამომავალ ნეირონებს რაოდენობის ტოლი. თითოეული ნეირონი სინაფსებით დაკავშირებულია სხვა ნეირონებთან. გარდა ამისა, ნეირონებს გააჩნიათ ერთი შემავალი სინაფსი, რომლის საშუალებითაც ხდება შემავალი სიგნალის შეყვანა, ხოლო გამოსავალი სიგნალი, როგორც წესი, წარმოიქმნებიან აქსონებზე.

ჰოპფილდის ქსელის მიერ ამოცანის ამოხსნა ფორმულირდება შემდეგნაირად: მოცემულია გარკვეული რაოდენობის ორობითი სიგნალი. ქსელმა უნდა შეძლოს მის შესასვლელზე მიწოდებული არაიდეალური სიგნალებიდან გამოყოს შესაბამისი სახ (თუ კი ასეთი არსებობს) ან გააკეთოს დასკვნა იმის შესახებ, რომ შემავალი სიგნალები არ შეესაბამებიან არც ერთ სახეს.

ზოგადად, ნებისმიერი სიგნალი შეიძლება აღიწეროს $X = (x_1, x_2, \dots, x_n)^T$ ვექტორით, სადაც n – ქსელის ნეირონების რაოდენობაა და აგრეთვე შემაჯავლი და გამომავალი ვექტორების განზომილებაა. X ვექტორის ყოველი კომპონენტი x_i ტოლია ან $+1$ ან -1 . აღნიშნოთ k -ური სახის ვექტორი X^k - თი, ხოლო მისი კომპონენტები შესაბამისად $x_i^k, i=1,2,\dots,m$, სადაც m – სახეების რაოდენობაა. როდესაც ქსელი ამოიცნობს რომელიმე სახეს, მაშინ ქსელის გამოსასვლელები შეიცავენ ამ გამოცნობილსახის ვექტორს, ე.ი. $Y = X^k$, სადაც Y – ქსელის გამოსასვლელების ვექტორია, ე.ი. $Y = (y_1, y_2, \dots, y_n)^T$. წინააღმდეგ შემთხვევაში, გამოსავალი ვექტორი არ დაემთხვევა არც ერთ გასარჩევ სახეს.

ქსელის ინიციალიზაციის სტადიაზე სინაფსური წონის კოეფიციენტები განისაზღვრებიან შემდეგნაირად:

$$w_{ij} = \begin{cases} \sum_{k=1}^N x_i^k x_j^k, & i \neq j \\ 0, & i = j \end{cases}$$

აქ i და j ინდექსები შეესაბამებიან წინასინაფსურ და შემდეგსინაფსურ ნეირონებს, x_i^k, x_j^k k -ური სახის ვექტორის i -ური და j -ური ელემენტებია.

ქსელის ფუნქციონირების ალგორითმი შემდეგია:

1. ქსელის შესასვლელს მიეწოდება უცნობი სიგნალი. ფაქტიურად მისი შეყვანა ხორციელდება აქსონების მნიშვნელობების უშუალოდ დაყენებით

$$y_i(0) = x_i, \quad i = 1, 2, \dots, n.$$

აქედან გამომდინარე, სქემაში აღნიშნული შემაჯავლი სინაფსები ცხადად ატარებენ პიობით ხასიათს, ხოლო y_i -ის ფრჩხილებში ნულის ყოფნა ნიშნავს ქსელის ნულობან იტერაციას.

2. განისაზღვრება ნეირონების ახალი მდგომარეობა

$$s_j(p+1) = \sum_{i=1}^n w_{ij} y_i(p), \quad j = 1, 2, \dots, n$$

და აქსონების ახალი მნიშვნელობები

$$y_i(p+1) = F[s_j(p+1)],$$

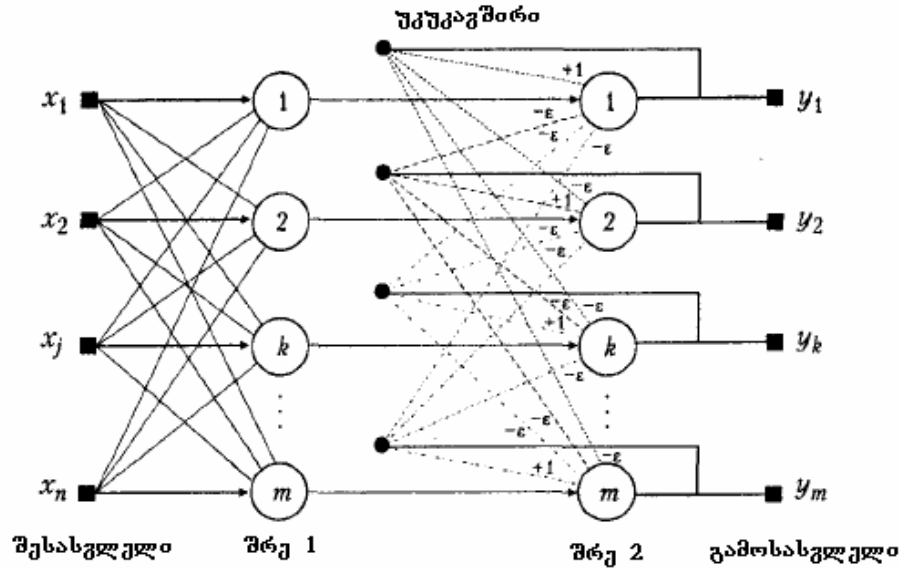
სადაც F – ერთეულოვანი ნახტომის აქტივაციის ფუნქციაა.

3. შემოწმება. შეიცვალა თუ არა ქსელის გამოსავალი აქსონების მნიშვნელობები. თუ შეიცვალა, მაშინ ხდება 2-ე ბიჯზე გადასვლა, თუ არა ალგორითმი დასრულდა. ამ დროს გამოსავალი ვექტორი წარმოადგენს სახეს, რომელიც არის შემაჯავლი სიგნალის საუკეთესო შეფასება.

ჰოპფილდის ქსელს გააჩნია გარკვეული შეზღუდვები. კერძოდ, ჰოპფილდის ქსელისათვის დასამახსოვრებელი სახეების რაოდენობა m არ უნდა აღემატებოდეს $0,15n$ სიდიდეს. გარდა ამისა, თუ ორი A და B სახე ძალზე წააგავენ ერთმანეთს ან დამახინჯებულია, მაშინ ქსელმა გამოსავალზე შეიძლება მოგვცეს მცდარი შედეგი: A -ს მაგივრად მივიღოთ B სახე ან პირიქით.

7.3 ჰემინგის ნაირონული ქსელი

ჰემინგის ქსელი, ჰოპფილდის ქსელთან შედარებით, ხასიათდება მცირე მეხსიერებით და გამოთვლების მცირე მოცულობით. ჰემინგის ქსელის სტრუქტურული სქემა მოცემულია ნახ. 7.2-ზე.



ნახ. 7.2

ქსელი შედგება ორი შრისაგან, რომელთაც გააჩნია m ნეირონი, სადაც m – სახეების რაოდენობაა. პირველი შრის ნეირონებს გააჩნიათ n რაოდენობის სინაფსები, რომლებიც მიერთებული არიან ქსელის შესასვლელებთან და ფაქტიურად ჰქმნიან ნულოვან შრეს. მეორე შრის ნეირონები შეერთებული არიან ერთმანეთთან უარყოფითი (დასაშუალებელი) სინაფსური უკუკავშირებით. ყოველი ნეირონის ერთადერთი დადებითი უკუკავშირის სინაფსი მიერთებულია თავისივე აქსონთან.

ჰემინგის ქსელის მუშაობის იდეა მდგომარეობს ტესტირებული სახის ყველა სხვა სახეებთან ჰემინგის მანძილის გამოთვლაში. ჰემინგის მანძილი ეწოდება ორ ერთნაირ განზომილებიან ბინარულ ვექტორში განსხვავებული ბიტების რაოდენობას და განისაზღვრება შემდეგი ფორმულით:

$$d(X_i X_j) = \sum_{k=1}^p |x_{ik} - x_{jk}|,$$

სადაც p – თანრიგების რაოდენობა. მაგალითად, $X_i = (1011101)$ და $X_j = (1001001)$, მაშინ ჰემინგის მანძილი ტოლია: $d(X_i X_j) = 2$.

ქსელმა უნდა შეარჩიოს ისეთი სახე, რომელიც უცნობ შემავალ ვექტორთან გააჩნია ჰემინგის მინიმალური მანძილი და ამის შედეგად ხდება ქსელის მხოლოდ ერთი გამოსავლის აქტივიზაცია.

ინიციალიზაციის სტადიაში პირველი შრის წონით კოეფიციენტებს და აქტივაციის ფუნქციის ზღურბლს ენიჭებათ შემდეგი მნიშვნელობები:

$$w_{ik} = \frac{x_i^k}{2}, \quad T_k = \frac{n}{2}, \quad i = 1, 2, \dots, n; \quad k = 1, 2, \dots, m,$$

სადაც x_i^k – k -ური სახის i -ური ელემენტი.

მეორე შრის დამამუხრუჭებელი სინაპსების წონით კოეფიციენტებს იღებენ $0 < e < 1/m$ გარკვეული მნიშვნელობების ტოლად. იმ ნეირონის სინაპსი, რომელიც მიერთებულია თავისივე აქსონთან გააჩნია +1 სიდიდის წონითი კოეფიციენტი.

ჰემინგის ქსელის ფუნქციონირების ალგორითმი შემდეგია:

1. ქსელის შესასვლელს მიეწოდება უცნობი $X = (x_1, x_2, \dots, x_n)^T$ ვექტორი, რომლის საშუალებითაც განისაზღვრება პირველი შრის ნეირონების მდგომარეობა:

$$y_j^{(1)} = s_j^{(1)} = \sum_{i=1}^n w_{ij} x_i + T_j, \quad j = 1, 2, \dots, m$$

აქ ზედა ინდექსის ფრჩხილებში ნაჩვენებია სიდიდე გვიჩვენებს ქსელის შრის ნომერს. ამის შემდეგ მიღებული მნიშვნელობებით ხდება მეორე შრის აქსონების ინიციალიზაცია

$$y_j^{(2)} = y_j^{(1)}, \quad j = 1, 2, \dots, m.$$

2. განისაზღვრება მეორე შრის ნეირონების ახალი მდგომარეობები

$$s_j^{(2)}(p+1) = y_j(p) - e \sum_{k=1}^m y_k^{(2)}(p), \quad k \neq j, \quad j = 1, 2, \dots, m$$

და მათი შესაბამისი აქსონების მნიშვნელობები

$$y_j^{(2)}(p+1) = F[s_j^{(2)}(p+1)], \quad j = 1, 2, \dots, m$$

აქ გამოიყენება წრფივი აქტივაციის ფუნქცია, თანაც F მნიშვნელობა უნდა იყოს მნიშვნელოვნად დიდი, რათა არგუმენტის ნებისმიერი შესაძლო მნიშვნელობამ პროცესი არ მიიყვანოს გაჯერებად.

3. შევამოწმოთ შეიცვალა თუ არა ქსელის მეორე შრის ნეირონების გამოსავალი სიდიდეები წინა იტერაციასთან შედარებით. თუ შეიცვალა, მაშინ გადავდივართ მეორე ბიჯზე, თუ არა – ალგორითმის მუშაობა დასრულებულია.

ამ ალგორითმიდან ჩანს, რომ ქსელის პირველი შრე საკმაოდ პირობითია, რადგან მას ვიყენებთ მხოლოდ ერთხელ, ალგორითმის პირველ ბიჯზე და შემდეგ აღარ გამოიყენება. ამიტომ პირველი შრე ქსელიდან შეიძლება გამოირიცხოს წონითი კოეფიციენტების მატრიცის ჩანაცვლებით.

ჰემინგის ქსელს, ჰოპფილდის ქსელისაგან განსხვავებით, გააჩნია რიგი უპირატესობები. მას შეუძლია მოძებნოს მინიმალური ცდომილება, როცა შემაჯავალი სიგნალის ცდომილება წარმოადგენს შემთხვევით და დამოუკიდებელს. ჰემინგის ქსელის ფუნქციონირებისათვის საჭიროა მცირე რაოდენობის ნეირონები, რადგან თითო სახეს სჭირდება ერთი ნეირონი. გარდა ამისა, ჰემინგის ქსელი ზოგადად უფრო სწრაფია და ზუსტი ვიდრე ჰოპფილდის ქსელი.

7.4 ორმიმართებული ასოციაციური მეხსიერება

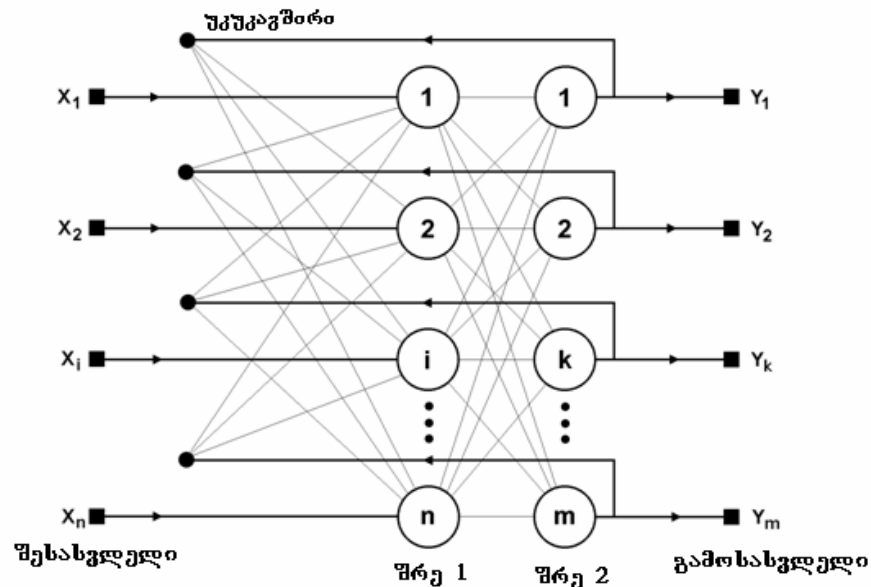
ასოციაციური მეხსიერება. ადამიანის მეხსიერება ასოციაციურია ანუ გრძელვადიანი მეხსიერებაა. ერთი საგანი მიგვანიშნებს მეორეს, მეორე მესამეს და ა.შ. ზოგიერთმა მოგონებამ შეიძლება წარმოშვას ამ მოგონებასთან დაკავშირებული სხვა მოგონებებიც. მაგალითად, რამოდენიმე მუსიკალურმა ტაქტმამ შეიძლება გამოიწვიოს მოგონებების მთელი გამა: პეიზაჟის, ბგერების და სურნელის ჩათვლით. გარდა ამისა, შესაძლებელია ასოციაციური აზროვნებით აღვადგინოთ დავიწყებული სახე და სხვა.

რეკურსიული ხელოვნური ნეირონული ქსელები ქმნიან ადამიანის მსგავს ასოციაციურ მეხსიერებას. ასოციაციურ მეხსიერება შეიძლება იყოს ავტოასოციაციური და გეტეროასოციაციური. ავტოასოციაციური მეხსიერება ეწოდება ისეთ მეხსიერებას, რომელსაც შეუძლია დაასრულოს ან შეცვალოს სახე (ობიექტი), მაგრამ არ შეუძლია მოცემული სახის სხვა სახეებთან ასოცირება. ეს ფაქტი წარმოადგენს ერთდონიან ასოციაციურ მეხსიერებას, სადაც გამოსავალი ვექტორი წარმოიქმნება იმ ნეირონებზე, რომლებსაც მიეწოდებათ შემავალი ვექტორი. ჰოპფილდის ნეირონული ქსელი რეალიზაციას ახდენს ავტოასოციაციური მეხსიერებას.

გეტეროასოციაციური მეხსიერება ეწოდება ისეთ მეხსიერებას, სადაც შემავალი ვექტორი მიეწოდება ნეირონების ერთ ნაკრებს, ხოლო გამოსავალ ვექტორებს ვღებულობთ ნეირონების სხვა ნაკრებზე.

ავტოასოციაციური მეხსიერების პირველი მოდელი დამუშავებული იყო ჰოპფილდის მიერ. შემდგომში ბ.კოსკომ განავითარა ჰოპფილდის იდეა და 1988წ დაამუშავა გეტეროასოციაციური მეხსიერება – ორმიმართებულიანი ასოციაციური მეხსიერების სახით.

ორმიმართებულიანი ასოციაციური მეხსიერება წარმოადგენს ჰოპფილდის ქსელის ლოგიკურ განვითარებას, სადაც საკმარისია დაემატის მეორე შრე. ორმიმართებულიანი ასოციაციური მეხსიერების სტრუქტურული აქემა წარმოდგენილია ნახ. 7.3-ზე



ნახ. 7.3

ქსელს შეუძლია დაიმახსოვროს ერთმანეთის მიმართ ასოცირებული ორი სახე. დაუშვათ, წყვილი სახე წარმოდგენილია $X^k = (x_i^k), i=1,2,\dots,n$ და $Y^k = (y_j^k), j=1,2,\dots,m, k=1,2,\dots,r$, სადაც r – წყვილთა რაოდენობაა.

ქსელის შესასვლელზე $P = (p_j^k), j=1,2,\dots,n$ ვექტორის მიწოდება იწვევს ქსელის მეორე შრის გამოსავალზე სხვა $Q = (q_j^k), j=1,2,\dots,m$ ვექტორის წარმოქმნას, რომელიც შემდგომ კვლავ მიეწოდება ქსელის პირველი შრის შესასვლელს. ყოველი ასეთი ციკლის დროს ორივე შრის გამოსავალი ვექტორები უახლოვდებიან საწყის ეტალონურ წყვილ ასოცირებულ ვექტორებს, რომელთაგან პირველი – X უფრო მსგავსია P ვექტორისა, რომელიც ქსელის შესასვლელზე პირველად მიაწოდეს, ხოლო Y მასთან არის ასოცირებული.

ვექტორებს შორის ასოციაცია დაკოდირებულია პირველი შრის $W^{(1)}$ წონითი კოეფიციენტების მატრიცის სახით. მეორე შრის წონითი კოეფიციენტების $W^{(2)}$ მატრიცა წარმოადგენს პირველი შრის $W^{(1)}$ მატრიცის ტრანსპონირებულ $(W^{(1)})^T$ მატრიცას. უნდა აღინიშნოს, რომ გრძელვადიანი (ან ასოციაცია) მექსიერება რეალიზირებულია W და W^T მატრიცების სახით.

ასეთი ჰოპფილდის ქსელის სწავლება მდგომარეობს W და W^T მატრიცების ელემენტების წინასწარ განსაზღვრაში შემდეგი ფორმულით:

$$w_{ij} = \sum_k x_i^k y_j^k, \quad i=1,2,\dots,n, \quad j=1,2,\dots,m.$$

ეს ფორმულა წარმოადგენს $W = \sum_k X^T Y$ მატრიცული განტოლების გაშლილ ჩანაწერს.

ორმიმართებულიანი ასოციაციური მექსიერების ნაკლს წარმოადგენს მისი მოცულობა, რომელიც შეზღუდულია. კერძოდ, თუ n არის შემავალი შრის ნეირონების რაოდენობა, მაშინ იმ ვექტორების რაოდენობა, რომელთა დამახსოვრება ქსელს არ აღემატება $L = 0,5n \log_2 n$ სიდიდეს. ასე მაგალითად, თუ $n=1024$, მაშინ ქსელს შეუძლია დაიმახსოვროს არაუმეტეს 25 სახე.

დადებით თვისებებს მიეკუთვნებიან:

- ჰოპფილდის ქსელისაგან განსხვავებით მას შეუძლია სხვადასხვა განზომილებიანი შემავალი ვექტორებისათვის ასოციაციის შექმნა, რაც იწვევს ქსელის გამოყენების სფეროს გაფართოებას;

- სინაფსური წონითი კოეფიციენტების ფორმირება საკმაოდ მარტივია და სწრაფი. ამასთან ქსელი მდრადობისაკენ სწრაფად კრებადია;

- ქსელში სიგნალები შეიძლება იყოს როგორც დისკრეტული ასევე უწყვეტი. ამ ორი შემთხვევისათვის დამტკიცებულია ქსელის სტაბილურობა.

დასკვნისათვის შეიძლება აღინიშნოს, რომ ჰოპფილდის, ჰემინგის და ორმიმართებულიანი ასოციაციური მექსიერების ნეიროქსელები საშუალებას იძლევიან მარტივად და ეფექტურად გადაწყვეტილი იყოს სახეთა გარჩევის ამოცანები სრული ან დამახინჯებული ინფორმაციისათვის.

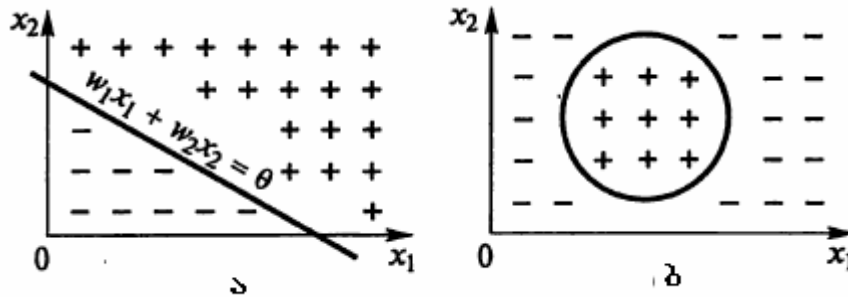
8. რადიალურ-ბაზისური ნეირონული ქსელები

რადიალურ-ბაზისური ნეირონული ქსელი წარმოადგენს ორშრიან უკუკავშირის გარეშე ქსელს, რომელსაც გააჩნია ფარული შრე რადიალურად სიმეტრიული ფარული ნეირონებით. რადიალურ-ბაზისური ფუნქცია (RBF) ეწოდება ფუნქციას, რომელიც რადიალურად იცვლება მოცემული რაიმე C ვექტორის ცენტრის მიმართ და ღებულობს არანულობან მნიშვნელობებს მხოლოდ ცენტრის გარშემო. RBF ფუნქციის არგუმენტად აღებულია მარძილი მიმდინარე X ვექტორსა და C ვექტორს შორის. ე.ი. $j = j(\|X - C\|)$.

გარკვეული აზრით ნეირონები, რომლებსაც გააჩნია რადიალურ-ბაზისური აქტივაციური ფუნქციები, წამოადგენენ ნეირონების ლოგიკურ დანამატებს კიბისებური და სიგმაიდალური აქტივაციური ფუნქციებით. მართლაც, მაკკალოკ-პიტსის ნეირონი აქტივიზირდება, როცა

$$\sum_i x_i w_i - q \geq 0,$$

ე.ი. მას გააჩნია ერთეულოვანი (დადებითი) გამოსავალი იმ წერტილებისათვის, რომლებიც განლაგებული არიან ჰიპერსიბრტყის $\sum_i x_i w_i - q = 0$ ერთ მხარეს და ნულობანი (უარყოფითი) – ჰიპერსიბრტყის მეორე მხარეს (ნახ. 8.1ა).

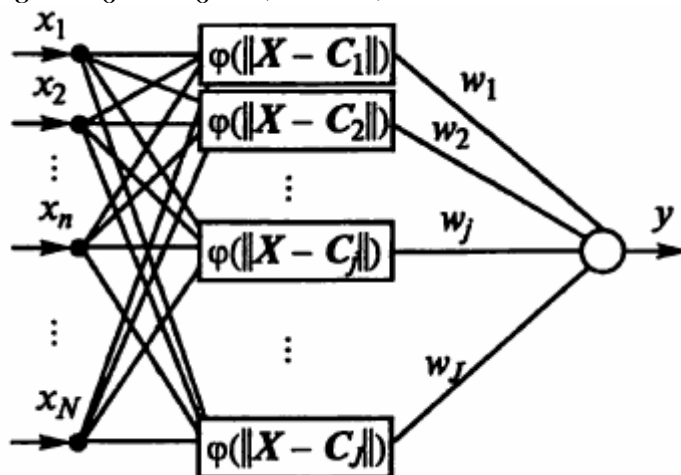


ნახ. 8.1

RBF ფუნქციის ნეირონი ყოფს აგრეთვე შემავალი პარამეტრების სივრცეს ორ ნაწილად, მაგრამ ამ შემთხვევაში გამოყოფ ზედაპირს წარმოადგენს ჰიპერსფერო (ნახ. 8.1ბ). სივრცის იმ წერტილებისათვის, რომლებიც განლაგებული არიან ჰიპერსფეროს შიგნით, ნეირონის გამოსავალი დადებითია, ხოლო იმ წერტილებისათვის, რომლებიც განლაგებული არიან ჰიპერსფეროს გარეთ – უარყოფითია ანუ ნულის ტოლია.

რადიალურ-ბაზისურ ნეირონებს გააჩნიათ ის უპირატესობა, რომ მათი საშუალებით ადვილია გამოყოფი ზედაპირის აგება, რომელიც უზრუნველყოფს შემავალი პარამეტრების დაყოფას კლასებად. აქედან გამომდინარე, რადიალურ-ბაზისურ ქსელებში არ არის აუცილებელი მრავალი ფარული შრეების გამოყენება. ტიპურ რადიალურ-ბაზისურ ქსელს გააჩნია მხოლოდ ერთი ფარული შრე, სადაც ნეირონების სინაფსური წონები ერთის ტოლია, ხოლო შემავალი და გამოსავალი შრეების ნეირონებს გააჩნიათ წრფივი აქტივაციური ფუნქციები. დამტკიცებულია, რომ თუ ასეთ ქსელს გააჩნია ფარული შრის საკმარის რაოდენობის ნეირონები, მაშინ გარანტირებულია სახეთა გარჩევის ნებისმიერი ამოცანის გადაწყვეტა.

განვიხილოთ რადიალურ-ბაზისური ნეირონული ქსელი, რომელსაც გააჩნია N შესასვლელი, ერთი გამოსასვლელი და ფარული შრის J რაოდენობის რადიალურ-ბაზისური ნეირონები (ნახ. 8.2).



ნახ. 8.2

დაუშვათ სწავლების მაგალითების ამონარჩევი შედგება Q რაოდენობის შემავალი N განზომილებიანი ვექტორებით X_q , $q = 1, 2, \dots, Q$ და მათი შესაბამისი

გამოსავალი პარამეტრებით d_q . თუ ფარული შრის ნეირონების რაოდენობა ტოლი იქნება სწავლების მაგალითების, ე.ი. $J=Q$, მაშინ ქსელის მუშაობა, რომელიც გარდაქმნის X_q შემომავალი ვექტორებს გამოსავალ d_q პარამეტრებად, შეიძლება წარმოვადგინო შემდეგნაირად:

$$\begin{bmatrix} \Phi(\|X_1 - C_1\|) & \Phi(\|X_1 - C_2\|) & \dots & \Phi(\|X_1 - C_j\|) & \dots & \Phi(\|X_1 - C_J\|) \\ \Phi(\|X_2 - C_1\|) & \Phi(\|X_2 - C_2\|) & \dots & \Phi(\|X_2 - C_j\|) & \dots & \Phi(\|X_2 - C_J\|) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \Phi(\|X_q - C_1\|) & \Phi(\|X_q - C_2\|) & \dots & \Phi(\|X_q - C_j\|) & \dots & \Phi(\|X_q - C_J\|) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \Phi(\|X_Q - C_1\|) & \Phi(\|X_Q - C_2\|) & \dots & \Phi(\|X_Q - C_j\|) & \dots & \Phi(\|X_Q - C_J\|) \end{bmatrix} \times$$

$$\times \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_j \\ \dots \\ w_J \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \dots \\ d_q \\ \dots \\ d_Q \end{bmatrix},$$

მატრიცული ფორმით გვექნება: $\Phi W = D$. საძიებელი წონითი კოეფიციენტების მატრიცა განისაზღვრება მარტივად: $W = \Phi^{-1} D$ (8.1).

ამრიგად, ქსელის სწავლებისთვის არ გვჭირდება იტერაციული პროცესი. მიღებულ შედეგს აქვს თეორიული ინტერესი, მაგრამ პრაქტიკული ფასეულობა მას არ გააჩნია. საქმე იმაშია, რომ სწავლების მაგალითების დიდი რაოდენობის დროს $J=Q$ პირობას მიუყვებათ გამოთვლით პრობლემებთან, რომლებიც დაკავშირებული არიან ფარული შრის ნეირონების რაოდენობის მნიშვნელოვან ზრდასთან. აქედან გამომდინარე, პრაქტიკულად ფარული შრის ნეირონების რაოდენობას იღებენ გაცილებით ნაკლებს, ვიდრე სასწავლო მაგალითების რაოდენობა ($J \ll Q$). ამ შემთხვევის დროს F მატრიცა არ წარმოადგენს კვადრატულს ($J \neq Q$) და წონითი კოეფიციენტები w_j შეიძლება განისაზღვროს ქსელის კვადრატული ცდომილების მინიმიზაციის პირობიდან:

$$e = \sum_{q=1}^Q \left[\sum_{j=1}^J w_j (\|X_q - C_j\|) - d_q \right]^2$$

მრავალშრიან პერსეპტრონისაგან განსხვავებით e ცდომილების ფუნქციას არ გააჩნია ლოკალური მინიმუმები. მისი მინიმიზაციის ამოცანა წრფივია და ამიტომ შესაძლებელია კარგად ნაცნობი წრფივი ოპტიმიზაციის ამოცანების გამოყენება, რომლებსაც სწავლების დროს გააჩნიათ კრებადობის დიდი სიჩქარე, ვიდრე მრავალშრიან ნეიროქსელებს.

რადიალურ-ბაზისურ ნეირონულ ქსელებში აქტივაციურ ფუნქციად ყველაზე უფრო ხშირად გამოიყენება გაუსის ფუნქცია:

$$\Phi(\|X - C_j\|) \exp \left\{ -\frac{\|X - C_j\|^2}{2s_j^2} \right\} \quad (8.2)$$

გაუსის ფუნქცია განისაზღვრება ორი პარამეტრით ცენტრის C_j ვექტორით და s_j სკალარული სიდიდით, რომლითაც მოცემულია ფუნქციის კლებადობა C_j

ცენტრსა და მიმდინარე X ვექტორს შორის გამოთვლილი ევკლიდეს მანძილის ზრდისას.

იმ შემთხვევაში, როცა ფარული შრის ნეირონების რაოდენობა ტოლია სწავლების მაგალითების რაოდენობისა ($J=Q$), მაშინ ლოგიკურია თუ აქტივაციის ფუნქციის ცენტრად ავიღებთ სწავლების ამონარჩევის ვექტორებს. ე.ი. $C_j = X_q$, $j=1,2,\dots,Q$, ხოლო s_j შეირჩევა იმ მოსაზრებით, რომ სივრცის ის ნაწილი, სადაც განსაზღვრულია X_q ვექტორები აქტივაციის ფუნქციის ზეგავლენის ქვეშ უნდა იმყოფებოდნენ. მაგალითად, s_j შეიძლება განისაზღვროს როგორც ევკლიდეს საშუალო კვადრატული მანძილი C_j ცენტრსა და უახლოეს P მეზობელს შორის:

$$s_j = \sqrt{\frac{1}{h} \sum_{p=1}^h \|C_j - C_p\|^2}, \quad (8.3)$$

სადაც P რეკომენდირებულია ავიღოთ [3–5] ინტერვალიდან. ამის შემდეგ წონითი w_j კოეფიციენტები შეიძლება განისაზღვროს (8.1) ფორმულით.

რეალურ სიტუაციაში $J \ll Q$, ამიტომ წინასწარ უნდა მოვახდინოთ X_q ვექტორების კლასიფიკაცია ანუ გავაერთიანოთ ახლოსმდგომი X_q ვექტორები კლასტერებად. შემდეგ განისაზღვრება კლასტერების ცენტრები, სადაც ათავსებენ აქტივაციური ფუნქციების ცენტრებს. ამის შემდეგ განისაზღვრება აქტივაციური ფუნქციის ფარჯრის სიგანე s , მაგალითად (8.3) ფორმულით. ბოლო ეტაპი შედგება ნეირონული ქსელის -სინაფსური წონების განსაზღვრა (8.2) ფორმულით.

დასკერისათვის შეიძლება ითქვას, რომ ბოლო წლებში რადიალურ-ბაზისურ ნეირონულ ქსელებს წარმატებით იყენებენ ისეთი ამოცანების ამოსახსნელად, სადაც ტრადიციულად გამოიყენებოდა სიგმაიდალური აქტივაციური ფუნქციები. ამ ამოცანებს მიეკუთვნებიან სახეთა გარჩევის, ფუნქციის აპროქსიმაციის და პროგნოზირების ამოცანები.

ინტერესი რადიალურ-ბაზისურ ნეირონულ ქსელებს მიმართ აიხსნება შემდეგი უპირატესობებით: ქსელს გააჩნია მხოლოდ ერთი ფარული შრე, რის გამოც ქსელის კონსტრუქტორს მოეხსნა შრეების რაოდენობის დადგენა. რადიალურ-ბაზისური ნეირონული ქსელების სწავლება დაიყვანება წრფივი ოპტიმიზაციის ამოცანად, ამიტომ არ არსებობს ლოკალური მინიმუმში მოხვედრის საშიშროება, ხოლო თვით სწავლების პროცესს სჭირდება გაცილებით ნაკლები დრო ვიდრე მრავალშრიან ნეიროქსელებს.

მაგრამ რადიალურ-ბაზისური ნეირონულ ქსელების პროექტირებისას საჭირო ხდება ფარული შრის ნეირონების ოპტიმალური რაოდენობის დადგენა, შემაჯავლი სასწავლო ვექტორებისათვის კლასიფიკაციის ამოცანის გადაწყვეტა და აქტივაციური ფუნქციების ფარჯრის სიგანის დადგენა.

9 ადაპტური რეზონანსული თეორიის ნეირონული ქსელები

ადამიანის ტვინი ასრულებს გარე სამყაროდან მიღებულ უწყვეტ სენსორულ ინფორმაციის დამუშავების რთულ ამოცანებს. ტრივიალური ინფორმაციის ნაკადიდან მან უნდა შეძლოს სიცოცხლისათვის საჭირო ინფორმაციის გამოყოფა, დამუშავება და საჭიროების შემთხვევაში მისი გრძელვადიან მეხსიერებაში დამახსოვრება. ახალი სახეების დამახსოვრება ხდება ისეთი ფორმით, რომ ადრე დამახსოვრებული სახეები არ განიცდიან მოდიფიკაციას და არ დაივიწყებიან. ეს მოვლენა ქმნის დილემას: როგორ რჩება მეხსიერება პლასტიკური, რომელსაც შეუძლია ახალი სახეების აღქმა და იმავე დროს შეირჩაიროს სტაბილურობა, რომელიც იძლევა იმის საშუალებას, რომ სახეები ფუნქციონირების პროცესში არ განადგურდებიან.

ტრადიციული ხელოვნური ნეირონული ქსელები აღმოჩნდნენ უსუსური სტაბილურობის და პლასტიკურობის პრობლემების გადასწყვეტად. აქ საქმე გვაქვს ურთიერთგამომრიცხავ მოთხოვნასთან. ერთის მხრივ, მეტად მნიშვნელოვანია რომ ქსელმა აღმოაჩინოს ახალი კლასის სახე, რომელიც არ იყო წარმოდგენილი სასწავლო ამონაჩვენებში. ეს არის ქსელის პლასტიკურობის თვისება. მეორეს მხრივ, ნასწავლი სახეები უნდა დარჩნენ მეხსიერებაში. ამ თვისებას ქსელის მდგრადი მუშაობა ეწოდება. ეს ორი თვისება პლასტიკურობა და სტაბილურობა ურთიერთგამომრიცხავია.

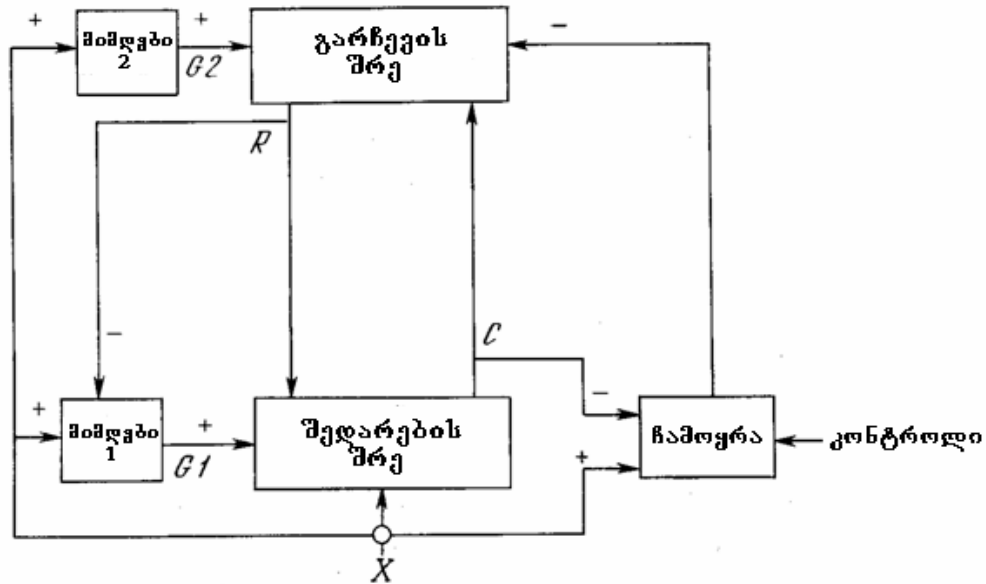
ნეიროქსელის პლასტიკურობის და სტაბილურობის პრობლემების გადასაწყვეტად გროსბერგმა დაამუშავა ადაპტური რეზონანსული თეორია (APT), რომლის საფუძველზე შექმნა APT ქსელის ორი ტიპი APT-1 და APT-2. APT-1 დამუშავებული იყო ორობითი ვექტორების დასამუშავებლად, ხოლო APT-2 შეუძლია როგორც ორობითი ასევე, უწყვეტი ვექტორების დასამუშავება. ჩვენ განვიხილავთ APT-1 ნეიროქსელს/

9.1 APT-1 ქსელის არქიტექტურა

APT ქსელის მუშაობის პრინციპი შედარებით მარტივია. ახალი ვექტორის შემოსვლისას ქსელი ცდილობს მიუსადაგოს მას უკვე არსებული კლასები. თუ ასეთი კლასი მოიძებნა, მაშინ ხდება კლასის პროტოტიპთან (ეტალონთან) შედარება და პროტოტიპის მცირეოდენი მოდიფიკაცია, რათა მან კარგად ასახოს ახალი სახე. ამ შემთხვევაში სახის კლასიფიკაცია ამით მთავრდება.

თუ ასეთი კლასის მოძებნა ვერ მოხერხდა, მაშინ იქმნება ახალი კლასი და შემოსული სახე რამდენადმე განიცდის მოდიფიკაციას და გამოიყენება როგორც ახალი კლასის პროტოტიპი. ამ დროს არსებული კლასები არ იცვლებიან.

ნახ. 9.1 წარმოდგენილია APT ქსელის გამარტივებული კონფიგურაცია.



ნახ. 9.1

APT ქსელი შედგება ხუთი ფუნქციონალური მოდულისაგან: ნეირონების ორი შრისაგან, ე.წ. „შედარების შრე“ და „გარჩევის შრე“ და სწავლების და კლასიფიკაციისათვის საჭირო სამი მართვის სპეციალიზირებულ ნეირონისაგან – ჩამოყრა, მიმღები 1 და მიმღები 2. განვიხილოთ თითოეული მათგანის ფუნქციები.

შედარების შრე. შედარების შრე ღებულობს ორობით შემომავალ X ვექტორს. პირველად იგი შეუცვლელად გაატარებს X ვექტორს, რათა მოხდეს გამოსავალი C ვექტორის ფორმირება.

შედარების შრის ყოველი ნეიონი ღებულობს სამ ორობით შესასვლელს შემავალი X ვექტორის x_i კომპონენტისაგან, უკუკავშირის R_i სიგნალიდან და მიმღები 1 გამოსავალიდან.

იმისათვის, რომ მივიღოთ ნეირონის გამოსავალზე ერთი ტოლი სიდიდე, საჭიროა, როგორც მინიმუმში სამი შესასვლელიდან ორი იყოს ერთი ტოლი. ამრიგად ხდება ორი მესამედის წესის რეალიზირება. დასაწყისში მიმღები 1-ის გამოსავალი $G1=1$, რათა უზრუნველყოს ნეირონის აქტივაცია. ასევე R ვექტორის კომპონენტები დაყენდება ნულის ტოლი. შესაბამისად ამ მომენტისთვის C ვექტორი X ვექტორის ტოლია ($C = X$).

გარჩევის შრე. შედარების შრის მიერ გამომუშავებული C სიგნალი მიეწოდება გარჩევის შრის შემავალ ნეირონებს, რომლებსაც გააჩნიათ წონის b_j ვექტორი, ამასთან აგიზნება შრის მხოლოდ ერთი ნეიონი, რომლის წონის ვექტორი ყველაზე უფრო ახლოს იქნება C სიგნალთან. ამ გამარჯვებული ნეირონის გამოსავალი დაყენდება ერთი ტოლად. დანარჩენი ნეირონები დამუხრუჭებულნი იქნებიან. გამატჯვებული ნეირონის სიგნალი უკუკავშირით კვლავ მიეწოდება შედარების შრეს სინაფსური წონის T ვექტორის საშუალებით. ფაქტობრივად T ვექტორი წარმოადგენს მსგავსობის კრიტერიუმის მატარებელს, რომელიც განსაზღვრულია გამარჯვებული ნეირონის ვექტორით.

მიმღები 1. გამოსავალი სიგნალი $G1=1$, როცა შემავალი ორობითი X ვექტორის თუნდაც ერთი კომპონენტი განსხვავდება ნულისაგან. მაგრამ, თუ R ვექტორის თუნდაც ერთი კომპონენტი განსხვავებულია ნულისაგან, მაშინ დაყენდება $G1=0$.

მიმღები 2. ისევე როგორც მიმღები 1, მიმღები 2-ის გამოსავალი სიგნალი $G2=1$, როცა შემავალი ორობითი X ვექტორის თუნდაც ერთი კომპონენტი

განსხვავდება ნულისაგან. ამ ნეირონის ამოცანას წარმოადგენს ჩაახშოს გარჩევის შრის აქტიობა, როცა ქსელ არ მიეწოდება არავითარი ინფორმაცია. გარდა ამისა, $G2$ წარმოადგენს X ვექტორის კომპონენტებისათვის ლოგიკურ ან ოპერაციას.

ჩამოყრა. ჩამოყრის მოდულის დანიშნულებაა დაადგინოს წინასწარ მოცემული მსგავსობის კრიტერიუმი საკმარისია თუ არა X ვექტორს მიენიჭოს გამარჯვებულის სტატუსი. ამისათვის ჩამოყრის მოდული ზომავს X და C ვექტორებს შორის მსგავსობის ზომას C ვექტორის ერთეულოვანი კომპონენტების რაოდენობის ფარდობით X ვექტორის ერთეულოვანი კომპონენტების რაოდენობასთან. ჩამოყრის სიგნალი გამოიშვება მხოლოდ მაშინ, როცა ეს ფარდობა ნაკლები იქნება r მსგავსობის პარამეტრზე, რაც იმას ნიშნავს, რომ X ვექტორის კლასიფიკაცია ვერ მოხერხდა. ჩამოყრის სიგნალი იწვევს გამარჯვებული ან წარუმატებელი ნეირონის დამუხრუჭებას, რომელიც შემდგომ ვერ მიიღებს ქსელის მუშაობაში მონაწილეობას.

9.2 APT ქსელის ფუნქციონირება

APT ქსელის ფუნქციონირების პროცესი შედგება ხუთი ფაზისაგან: ინიციალიზაცია, გარჩევა, შედარება, ძებნა და სწავლება. განვიხილოთ ისინი.

ინიციალიზაცია. ქსელის ფუნქციონირების დასაწყისში ნეირონების წონითი ვექტორები B_j და T_j , აგრეთვე მსგავსობის r პარამეტრი დებულობენ საწყის მნიშვნელობებს. B_j წონის ვექტორის მნიშვნელობები განისაზღვრებიან შემდეგი უტოლობით:

$$b_{ij} < \frac{L}{L-1+m} \quad \text{ყველა } i,j\text{-ისათვის,}$$

სადაც m – შემაგალი ვექტორის კომპონენტების რაოდენობა; L – მუდმივა მეტია ერთზე (ჩვეულებრივ $L=2$); i -შედარების შრის გამოსავალი ვექტორის i -ური კომპონენტა; j - გარჩევის შრის გამარჯვებული ნეირონის ნომერი.

T_j წონის ვექტორის მნიშვნელობები აიღება $t_{ij}=1$ ტოლი ყველა i,j -ისათვის.

მსგავსობის r პარამეტრი (ზომა) შეირჩევა ამოცანის მოთხოვნებიდან გამომდინარე $[0;1]$ ინტერვალიდან. თუ მსგავსობის r ზომა საკმაოდ დიდია, მაშინ სახეების უმეტესობის კლასიფიკაცია არ მოხდება და ქსელი იძულებული იქნება თითოეულ სახეს გამოუყოს ცალკე ნეირონი. ეს იწვევს ქსელის ცუდ განზოგადებას, რადგან ვექტორის თუნდაც მცირე განსხვავება კლასის პროტოტიპთან მაინც მოხდება ახალი კლასის შექმნა. პირიქით, თუ მსგავსობის ზომა საკმაოდ მცირეა, მაშინ ძალზე განსხვავებული სახეები შეიძლება მოხვდნენ ერთ კლასში.

სამწუხაროდ არ არსებობს მსგავსობის ზომის შერჩევის თეორიული დასაბუთება. აქედან გამომდინარე, ყოველ ცალკეულ შემთხვევაში საჭიროა შეირჩიოს მსგავსობის ზომის სიდიდე.

შემაგალი X ვექტორის ნულოვანი კომპონენტების დროს ხდება მიმღები 2-ის ნეირონების განულება, ამავე დროს განულებულია გარჩევის შრის გამოსავალი ნეირონიც. როცა X ვექტორის კომპონენტები განსხვავდებიან ნულისაგან, ორივე მიმღების სიგნალები $G1$ და $G2$ დაყენდება ერთი ტოლი. ამ

დროს შედარების შრის გამოსავალი ვექტორი C ორი მესამედის წესის თანახმად ზუსტად ტოლია X ვექტორისა.

გარჩევის ფაზა. ქსელის შესასვლელს მიეწოდება X ვექტორი, რომლის კლასიფიკაციაც უნდა მოხდეს. X ვექტორი უნდა შეიცავდეს ერთ ან მეტ ნულისაგან განსხვავებულ კომპონენტებს. გარჩევის შრის შესასვლელზე C ვექტორის მიწოდების შემდეგ, გარჩევის შრის ყოველი ნეირონისათვის განისაზღვრება ნაკეცი წონის B_j ვექტორსა და C ვექტორს შორის. ის ნეირონი, რომელსაც გააჩნია მაქსიმალური ნაკეცი გამოცხადდება გამარჯვებულად, რომლის გამოსავალი ერთის ტოლი იქნება, ხოლო დანარჩენი ნეირონების გამოსასვლელები იქნებიან ნულის ტოლი. ე.ი. ხდება ამ ნულოვანი ნეირონების დამუხრუჭება.

ამრიგად, R ვექტორის მხოლოდ ერთი r_j კომპონენტი ხდება ერთის ტოლი, ხოლო დანარჩენი ნულის ტოლია. რადგან R ვექტორი აღარ წარმოადგენს ნულოვანს, ამიტომ $G1$ სიგნალი ნულის ტოლი ხდება. გამარჯვებული ნეირონი უკუკავშირის საშუალებით გაგზავნის სიგნალს შედარების შრეში და დაიწყება შედარების ფაზა.

შედარების ფაზა. შედარების შრის გამოსავალი C ვექტორი შეიცავს ერთეულოვან კომპონენტებს მხოლოდ იმ პოზიციებში, სადაც ერთიანები გააჩნიათ X ვექტორსა და უკუკავშირის R ვექტორს. თუ X და C ვექტორების შედარების შედეგი არ აღმოჩნდება მნიშვნელოვნად განსხვავებული, მაშინ ჩამოყრის ნეირონი რჩება აქტიური. გარჩევის შრეში C ვექტორი კვლავ იწვევს იგივე გამარჯვებული ნეირონის აგზავნას, რაც იწვევს კლასიფიკაციის ამოცანის დამთავრებას. წინააღმდეგ შემთხვევაში, გამომუშავდება ჩამოყრის სიგნალი, რომელიც გარჩევის შრეში დამუხრუჭებს გამარჯვებულ ნეირონს და იწყება ძებნის ფაზა.

შედარების პროცედურა ტარდება შემდეგნაირად:

- გამოითვლება D სიდიდე, რომელიც ტოლია შემავალი X ვექტორის ერთეულოვანი კომპონენტების რაოდენობისა;
- გამოითვლება N სიდიდე, რომელიც ტოლია C ვექტორის ერთეულოვანი კომპონენტების რაოდენობისა;
- განისაზღვრება მსგავსობის S სიდიდე შემდეგი ფორმულით:

$$S = \frac{N}{D} .$$

მაგალითი. ვთქვათ მოცემულია

$$X = (1011101)^T, \quad D = 5.$$

$$C = (0011101)^T, \quad N = 4.$$

$$S = \frac{N}{D} = \frac{4}{5} = 0.8.$$

ძებნის ფაზა. ჩამოყრის დამამუხრუჭებელი სიგნალის მოქმედებით გარჩევის შრის ყველა ნეირონი გამოსასვლელზე ღებულობს ნულოვან მნიშვნელობას. შესაბამისად $G1=1$. კვლავ შედარების შრის გამოსავალი C სიგნალი დაყენდება X სიგნალის ტოლი, ისე, როგორც ქსელის მუშაობის დასაწყისში. მაგრამ, გარჩევის შრეში კონკურენტულ ბრძოლაში წინა ბიჯის გამარჯვებული ნეირონი მონაწილეობას არ ღებულობს და მოიძებნება ახალი კანდიდატი. ამის შენდეგ კვლავ მეორდება შედარების ფაზა

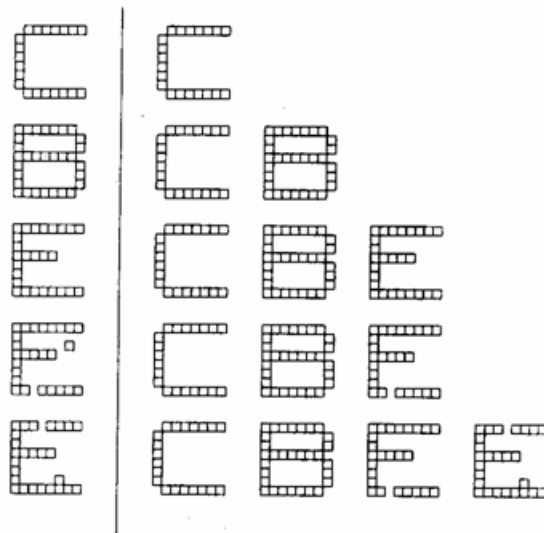
თუ X ვექტორმა ვერ მოგვცა მსგავსობა დამახსოვრებული კლასების ვერც ერთ პროტოტიპთან, მაშინ შემომავალი X ვექტორი გამოცხადდება

როგორც ახალი კლასი და მისთვის გარჩევის შრეში გამოიყოფა ახალი ნეირონი, რომლის წონითი კოეფიციენტები X ვექტორის წონითი კოეფიციენტების ტოლია.

როგორც ვხედავთ, APT ქსელის სწავლება და ფუნქციონირება ხდება ერთდროულად. უნდა აღინიშნოს, რომ სწავლების პროცესის გარკვეული სტაბილიზაციის შემდეგ კლასიფიკაციის სწავლება წარმოებს ძებნია ფაზის გარეშე.

ქსელის სწავლების მაგალითი.. ზოგადად ქსელის სწავლება ხდება წონების ისეთი შეცვლით, რომ შემომავალმა ვექტორმა მოახდინოს არჩევის შრეში ნეირონების გააქტიურება. გარდა ამისა, სწავლების პროცესში არ ნადგურდებიან ადრე დამახსოვრებული სახეები, რითაც აღმოიფხვრება ქსელის დროებითი არასტაბილურობა.

განვიხილოთ ქსელის სწავლების მარტივი მაგალითი. ნახ. 9.2-ზე წარმოდგენილია APT ქსელის სწავლების ასოების გარჩევის ტიპური სეანსი.



ნახ. 9.2

ასოები შედგებიან 8×8 ზომის პატარა კვადრატებისაგან. მარცხენა მხარეს წარმოდგენილი ყოველი კვადრატი შეესაბამება შემავალი X ვექტორის ერთი ტოლ კომპონენტას, ხოლო გამოტოვებული კვადრატები – X ვექტორის ნულოვან კომპონენტებს. ნახაზის მარჯვენა მხარეს ნაჩვენებია სახეები წარმოდგენენ დამახსოვრებულ სახეებს, რომლებიც წარმოდგენილი არიან T_j წონის ვექტორის კომპონენტებით.

დასაწყისში ქსელ მიეწოდება „C“ ასო. რადგან დამახსოვრებული სახეები არ არსებობენ, ამიტომ ხდება ამ სახის პირდაპირი დამახსოვრება, რომლის წონითი კოეფიციენტები X ვექტორის კოეფიციენტების ტოლია.

შემდეგ ქსელს მიეწოდება „B“ ასო, რომელიც იწვევს წარუმატებელ ძებნას და ამიტომ მეხსიერებაში ფიქსირდება ახალი სახე. ანალოგიურად ხდება „E“ ასოს მიწოდებისას. შემდეგ ქსელს მიეწოდება ოდნავ დამახინჯებული „E“ ასოს ვერსია, რომელიც გარკვეული სიზუსტით შეესაბამება დამახინჯებულ „E“ ასოს. „E“ ასოს ქვედა ფეხზე გამოტოვებული პიქსელი (ანუ კვადრატი) მის შესაბამის C ვექტორის კომპონენტს ანიჭებს ნულოვან მნიშვნელობას. დამატებითი იზოლირებული კვადრატი არ მოქმედებს დამახსოვრებულ ვექტორზე, რადგან ის არ შეესაბამება დამახსოვრებული სახის ერთიანს.

მეთოდე სიმბოლოს წარმოადგენს „E“ ასო, რომელსაც გააჩნია ორი დამახინჯება. რადგან ეს სახე არ შეესაბამება ადრე დამახსოვრებულ სახეებს ($S < r$), ამიტომ მის დასამახსოვრებლად გამოიყოფა ცალკე ნეირონი.

10. ნეიროქსელის პროექტირების საკითხები

10.1 არსებობის თეორემა

მათემატიკური მეთოდებით ამოხსნილი გამოყენებითი ამოცანების უმეტესობა დაიყვანება გარკვეული სახის რთული ფუნქციის მოძებნაში, რომელიც ასრულებს შემავალი მრავალგანზომილებიანი X პარამეტრის გარდაქმნას გამოსავალი Y პარამეტრზე. ასეთი ფუნქციის შექმნის უნივერსალურ ინსტრუმენტს წარმოადგენს ნეიროქსელური ტექნოლოგია.

ბუნებრივად ისმება კითხვა: ყოველთვის შესაძლებელია თუ არა ასაგოთ ნეირონული ქსელი, რომელიც ასრულებს მოცემული ნებისმიერი ამონარჩევის გარდაქმნას და რა მოთხოვნებს უნდა აკმაყოფილებდეს ნეირონული ქსელი? იმისათვის, რომ ვუპასუხოთ ამ შეკითხვას უნდა გავიხსენოთ, რომ ნეიროქსელის ყოველი ნეირონი ასრულებს სიგნალების სუპერპოზიციას (აჯამებას), რომლებიც მიიღებიან სხვა ნეირონებისაგან და რომლებმაც ამ ნეირონებში გაიარეს არაწრფივი (მაგალითად სიგმაიდალური) გარდაქმნა.

კითხვა იმის შესახებ შესაძლებელია თუ არა ნებისმიერი მრავალგანზომილებიანი ფუნქციის წარმოდგენა როგორც მცირე რაოდენობიანი პარამეტრების მქონე ფუნქციის სუპერპოზიცია, აინტერესებდათ მათემატიკოსებს ბოლო რამოდენიმე საუკუნის მანძილზე. კერძოდ, 1900 წ. პარიზის საერთაშორისო მათემატიკოსების კონგრესზე გამოჩენილმა გერმანელმა მათემატიკოსმა დავით გილბერტმა წარმოადგინა მის მიერ ჩამოყალიბებული 23 პრობლემა, რომელთა ამოხსნა შეთავაზა მათემატიკოსებს. ერთ-ერთ ამ პრობლემებიდან მეცამეტე რიგითი ნომრით დეკლარირებული იყო ასეთი გარდაქმნის შეუძლებლობა.

მაგრამ, შემდგომმა გამოკვლევებმა აჩვენა, რომ გილბერტის მე-13 პრობლემას გააჩნია სხვანაირი ამოხსნა. ა. კოლმოგოროვის და მისი მოსწავლის ვ. არნოლდის მიერ მიღებული იყო ფუნდამენტალური თეორიული დასაბუთება ასეთი გარდაქმნის შესაძლებლობის შესახებ. შემდგომში ჰესტ-ნილსენის მიერ ეს შედეგი დამუშავებული იყო ნეირონულ ქსელებში გამოსაყენებლად. კერძოდ, დამტკიცებული იყო, რომ ნებისმიერი წყვილი სიმრავლისათვის, რომლებიც განსხვავდებიან ერთმანეთისაგან შემავალი და გამომავალი ნებისმიერი განზომილების ვექტორებით (X_q, Y_q) , $q = 1, 2, \dots, Q$, არსებობს ორშიანი ნეიროქსელი (ერთი ფარული შრიით) სიგმაიდალური აქტივაციის ფუნქციით და სასრულო რაოდენობის ნეირონებით, რომლიც ყოველ შემავალ X ვექტორისათვის აყალიბებს შესაბამის გამოსავალ D_q ვექტორს. ამრიგად, დამტკიცებული იყო ნეირონული ქსელის შექმნის საშუალება, რომელიც ასრულებს მოცემული ნებისმიერი სასწავლო ამონარჩევის გარდაქმნას.

ფარული შრისათვის ნეირონების საჭირო რაოდენობის N_w განსაზღვრისათვის ჰეხტ-ნილსენის თეორემის შედეგად შემოთავაზებული იყო შემდეგი ფორმულა:

$$\frac{N_y Q}{1 + \log_2} \leq N_w \leq \left(\frac{Q}{N_y} + 1 \right) (N_x + N_y + 1) + N_y, \quad (10.1)$$

სადაც N_y – გამოშვებული სიგნალის განზომილება; Q – სასწავლო ამონარჩევის ელემენტების რაოდენობა; N_w – სინაფსური წონების რაოდენობა; N_x – შემავალი სიგნალის განზომილება.

თუ ამ ფორმულით განვსაზღვრავთ სინაფსური წონების საჭირო რაოდენობას, მაშინ შეგვიძლია გამოვთვალოთ ფარული შრის ნეირონების რაოდენობა:

$$N = \frac{N_w}{N_x + N_y}. \quad (6.2)$$

10.2 შემაჯავალი და გამომავალი პატამეტრების მომზადება

შესასწავლი მაგალითების წარმატებულ შერჩევაზე ბევრად არის დამოკიდებული დარგობრივი სფეროს მოდელირების ნეირონული ქსელების შექმნა. უნდა გვახსოვდეს, რომ დარგობრივი სფეროს ყველა პარამეტრი არ მოქმედებს ნეირონული ქსელის გამოსავალ პარამეტრებზე. მაგალითად, სხეულის ტემპერატურის ცოდნა ვერ იმოქმედებს პაციენტის დიაგნოზის დასმაზე, თუ კი ლაპარაკია მაღალითად, ფსიქიური დაავადების შესახებ,

პარამეტრები, რომლებიც არ მოქმედებენ ნეიროქსელის გამოსავალ Y ვექტორზე, ეწოდებათ არაინფორმატიული (უმნიშვნელო). ცხადია, რომ ასეთი უმნიშვნელო პარამეტრების ჩართვა შემოსავალ X ვექტორში არასასურველია.

პრაქტიკაში საკმაოდ რთულია და ზოგჯერ შეუძლებელიც დავადგინოთ დარგობრივი სფეროს რომელი პარამეტრია ინფორმატიული და რომელი არა. ამიტომ პირველ ეტაპზე X ვექტორში ჩართულია რაც შეიძლება დიდი რაოდენობის პარამეტრი რასაკვირველია იმ პარამეტრების გარდა, რომლებიც ცხადად არიან არაინფორმატიულები. პირველადი ნეიროქსელის შექმნის და მისი სწავლების შედეგად შესაძლებელია არაინფორმატიული პარამეტრების გამოვლენა შემდეგი ორი გზით:

1. შემაჯავალი ნეირონების წონითი კოეფიციენტების ანალიზით. კერძოდ, თუ აღმოჩნდება, რომ რომელიმე შემაჯავალი პარამეტრის სინაფსური წონა უმნიშვნელოა (ნაკლებია ნეირონის სხვა პარამეტრებთან შედარებით), მაშინ ეს შემაჯავალი ნეირონი შეესაბამება უმნიშვნელო პარამეტრს.

2. შემაჯავალი პარამეტრების შემფოთებით და ნეირონული ქსელის ამ შემფოთებაზე რეაქციით. თუ ქსელი არ რეაგირებს ან სუსტად რეაგირებს რომელიმე შემაჯავალი პარამეტრის შეცვლაზე, მაშინ ეს პარამეტრი არაინფორმატიულია.

არაინფორმატიული პარამეტრების ნეირონების გამოვლენის და მათი გამორიცხვის შემდეგ ნეიროქსელის ხარისხი უმჯობესდება, რადგან ხდება მისი განზომილების შემცირება. უნდა გვახსოვდეს, რომ საკმაოდ მცირე შემაჯავალი ნეირონების რაოდენობამ შეიძლება გამოიწვიოს ის, რომ ნეირონულ ქსელს

შეიძლება არ ეყოს შემავალი პარამეტრების რაოდენობა მის წინაშე დასმული ამოცანის გადასაწყვეტად.

დარგობრივი სფეროს პარამეტრები შეიძლება იყოს მრავალფეროვანი. ეს შეიძლება იყოს სხვადასხვა დიაპაზონში ცვალებადი სიდიდეები, ხარისხობრივი მაჩვენებლები, ისეთი როგორიცაა პაციენტის თმის და თვალების ფერი, თარიღი (რიცხვი, თვე, წელი), გრაფიკული ობიექტები და სხვა. რადგან ნეიროქსელს შეუძლია რიცხვითი მონაცემების დამუშავება, ამიტომ არარიცხვითი მონაცემები კოდირებული უნდა იყოს.

ნეირომული ქსელის შექმნისათვის რიცხვითი მონაცემები უმჯობესია წარმოვადგინოთ ნორმირებული სახით, ე.ი. მოვახდინოთ მათი მასშტაბირება, მაგალითად, $[0,1]$ ან $[-1,1]$ ინტერვალში. ეს შეიძლება განხორციელდეს შემდეგი მარტივი წრფივი გარდაქმნით:

$$\tilde{x}_n = \frac{x_n - x_{n,\min}}{x_{n,\max} - x_{n,\min}}(b - a) + a,$$

სადაც x_n და $\tilde{x}_n$ – i -ური პარამეტრის საწყისი და ნორმირებული მნიშვნელობებია; $[x_{n,\min}, x_{n,\max}]$ – n -ური პარამეტრის რეალური ცვალებადობის ინტერვალი; $[a, b]$ – შემავალი სიგნალის ცვალებადობის მისაღები დიაპაზონი. ნეიროქსელის სასურველი გამოსავალი სიგნალი უნდა იყოს კოდირებული და ნორმირებული მისაღებ $[a, b]$ დიაპაზონში ზემოდ მოყვანილი ფორმულის საშუალებით.

10.3 პროექტირების მეთოდები და პრობლემები

თეორიულად ნებისმიერი რთული ობიექტის ნეიროქსელური მოდელის ასაგებად საკმარისია გამოვიყენოთ ერთი ფარული შრის მქონე ნეიროქსელი ქსელი სიგნალიდან აქტივაციური ფუნქციით. მაგრამ, პრაქტიკაში მიზანშეწონილია მრავალშრიანი ნეირონული ქსელის გამოყენება. ასეთ ქსელებს შეიძლება ჰქონდეთ სინაფსური წონების მცირე განზომილებიანი მატრიცა, ვიდრე ორშიან ქსელს, რომელიც განკუთვნილია იგივე ამოცანების გადასაწყვეტად.

ფარული შრეების ოპტიმალური რაოდენობის განსაზღვრის მეთოდთა არ არსებობს. პრაქტიკაში უფრო ხშირად გამოიყენება ერთი ან ორი ფარული შრიანი ნეიროქსელები N_x – დან $3N_x$ – მდე.

ნეირონული ქსელის პროექტირებისას საჭიროა გვახსოვდეს, რომ ქსელს უნდა შეეძლოს სასწავლო მაგალითზე არა მარტო სწორი რეაგირება, არამედ შეეძლოს შეძენილი ცოდნის განზოგადოება ანუ სწორად რეაგირებდეს ისეთი მაგალითების შემთხვევების დროს, რომლებიც სწავლების პროცესში არ აყვნენ წარმოდგენილნი. ამისათვის შემოაწვთ ტესტების გარკვეული რაოდენობა, რომლებიც იმავე დარგობრივი სფეროს განეკუთვნებიან და რომლებიც სწავლების პროცესში არ მონაწილეობდნენ. სწავლების შემდეგ განისაზღვრება საშუალო კვადრატული ცდომილება (5.1).

შესასწავლი მაგალითებით გამოწვეულ საშუალო კვადრატულ ცდომილებას სწავლების e ცდომილება ეწოდება, ხოლო ტესტური მაგალითებით გამოწვეულს e_T ცდომილებას განზოგადოების ცდომილება ეწოდება. ქსელის შიგა ფარული შრეების ნეირონების N რაოდენობის გაზრდა იწვევს სწავლების ცდომილების შემცირებას, მაშინ როდესაც განზოგადოების

ცდომილება ჯერ მცირდება, ხოლო შემდგომ დაწყებული გარკვეულ $N=N_0$ ოპტიმალური მნიშვნელობიდან იზრდება.

ამრიგად, ნეიროქსელის პროექტირების რდოს არსებობს ნეირონების საჭირო რაოდენობის განსაზღვრის საჭიროება. შემაჯავლი შრის ნეირონების რაოდენობა უნდა ემთხვეოდეს შემაჯავლი X პარამეტრების რაოდენობას, რომელიც განსაზღვრულია ამოსახსნელი ამოცანის პირობიდან. ქსელის გამოსავალი შრის ნეირონების რაოდენობა უნდა ემთხვეოდეს გამომავალი Y ვექტორის განზომილებას, რაც ასევე განსაზღვრულია ამოცანის პირობიდან. ნეიროქსელის ფარული შრეების რაოდენობა ჰესტ-ნილსონის თეორემის თანახმად უნდა იყოს არა ნაკლები ერთისა. ფარული შრეების ნეირონების რაოდენობა მიახლოებით განისაზღვრება (6.1) და (6.2) ფორმულებით.

არსებობს ფარული შრეების ნეირონების ოპტიმიზაციის ორი მეთოდი (ხერხი) დესტრუქციული და კონსტრუქციული. დესტრუქციული მეთოდი იმაში მდებარეობს, რომ დასაწყისში აიგება ქსელი თავიდანვე მაქსიმალური რაოდენობის ნეირონებით, შემდგომ სწავლების პროცესში თანდათანობით გამოირიცხება ზედმეტი სინაფსური ნავშირები და ნეირომები.

არსებობს ზედმეტი ნეირონების გამოირიცხვის უამრავი ალგორითმი. მათ შორის ყველაზე მარტივ მეთოდს წარმოადგენს იმ სინაფსური წონების განულება, რომლებიც ნეიროქსელის სწავლების პროცესში საშუალო მნიშვნელობასთან შედარებით მნიშვნელოვნად ნაკლებ სიდიდეს. მაგრამ, სინაფსური წონების მცირე მნიშვნელობები სულაც არ არის აუცილებელი, რომ იწვევდეს მცირე ზედავლენას ნეირონებზე. ამიტომ მრავალ თანამედროვე ნეირონულ ქსელებში დაპროგრამებულია ევრისტიკული ალგორითმები, რომლებიც გარდა ნეირონების თანმიმდევრული გამოირიცხვისა, ასევე იკვლევენ ასეთი გამოირიცხვის ზეგავლენის არა მარტო სწავლების e ცდომილებაზე, არამედ განზოგადოების e_T ცდომილებაზეც.

დესტრუქციული ალგორითმის საერთო ნაკლოვანებას წარმოადგენს მისი მუშაობის ხანგრძლივი დრო, რადგან დასაწყისში გამოითვლება წარმოებს ჭარბი რაოდენობის ნეირონების მქონე ქსელში. ამ ნაკლისაგან თავისუფალია ალტერნატიული კონსტრუქციული ალგორითმები, სადაც ხდება ნეირონების თანმიმდევრული დამატება. ნეირონების ყოველი დამატების დროს ხდება ქსელის განზოგადოების ცდომილების e_T განსაზღვრა. ყოველი ახალი ნეირონის დამატება იწვევს e_T ცდომილების გაზრდას, მაგრამ გარკვეული სწავლების პერიოდის შემდეგ ცდომილება ხდება შედარებით მცირე ვიდრე იყო ნეირონის დამატებამდე. თუმცა გარკვეული t_0 დროის მომენტიდან ახალი ნეირონის დამატება აღარ იწვევს e_T ცდომილების შემცირებას, არამედ ხდება მისი გაზრდა, რაც გულისხმობს ჰიპერგანზომილების ეფექტის დადგომას.

რჩება გაურკვეველი ერთი საკითხი, კერძოდ ნეირონული ქსელის კონკრეტულად რა ადგილას უნდა დაემატოს ახალი ნეირონი. ადგილის შერჩევა შეიძლება მოხდეს შემთხვევით ან გამოვიყენოთ, მაგალითად ნეირონების დანაწევრების ალგორითმი. ამ ალგორითმით საჭიროა თვალყური ვადევნოთ იმას, თუ როგორ შეიცვლება Δw_{ij} ვექტორი. თუ აღმოჩნდება ისეთი ნეირონი, როცა სწავლების ერთი მაგალითების ჯგუფი მიისწრაფვის ერთი მიმართულებით w_{ij} ვექტორის წონების შეცვლისაკენ, ხოლო მაგალითების მეორე კგუფი – სხვა მიმართულებით, მაშინ ლოგიკურია ამ ნეირონის გახლეჩა ორად. ამ გზით ქსელში ახალი ნეირონის დამატებას არ მივეყვართ ნახტომისებრივ ქსელის ცდომილების ფუნქციის გაზრდას. ამ შემთხვევაში, ქსელის ოპტიმალური სტრუქტურის და მისი სწავლების პროცესის რაოდენობა

მცირდება, მაგრამ ალგორითმის გართულება და დამატებითი გამოთვლები ნეირონის დაწვერიანობის დროს, ხშირად მიყვებათ ასეთი ალგორითმის გამოყენებაზე უარის თქმამდე.

11. ნეიროქსელები მედიცინაში

ბიოსამედიცინო ამოცანების გადაწყვეტა მათემატიკური მეთოდების გამოყენებითი მიმდინარეობს არა ერთი ასეული წელი. გასული საუკუნის 60-იან წლებში დამუშავებული იყო ანალიზის მეთოდები, რომლებიც ფართოდ გავრცელდა და გამოიწვია მკვლევარებს შორის დიდი ინტერესი. მიუხედავად იმისა, რომ ყველაზე ყვრო პოპულარული ალგორითმები დღემდე აქტიურად გამოიყენება სამედიცინო და ბიოლოგიურ კვლევებში, კლინიკურ პრაქტიკაში მათ ვერ პოვებს ფარო გამოყენება. ეს დაკავშირებულია პირველ რიგში იმაზე, რომ მეთოდები ორიენტირებულნი არიან ჯგუფური მონაცემების დამუშავებაზე და ნაკლებად ცალკეულ ობიექტისათვის. მეორეს მხრივ, ბიოსამედიცინო ინფორმაციის თავისებურებაზე. ასე მაგალითად, მეთოდი, როგორც წესი, კარგად მუშაობს იმ ჯგუფის ობიექტებისათვის, რომლებზედაც ხდებოდა გამოკვლევები. რასაკვირველია უნივერსალური ალგორითმის შექმნა შეუძლებელია და ამიტომ მეთოდის სხვა ჯგუფის მიმართ მოითხოვს პრაქტიკულად თითქმის მთლიანად შეცვლას.

სამედიცინო სფეროში ყველა ამოსახსნელი ამოცანა, ნეიროინფორმაციული ტექნოლოგიებიდან გამომდინარე, პირობითად შეიძლება დავეყთ ორ ჯგუფად.

1. ამოცანები, რომლებიც გააჩნიათ ცნობილი და გარკვეული პირობების ნაკრები, რომელთა საშუალებით საჭიროა მკვეთრი, ზუსტი, არაორაზროვანი პასუხები ცნობილი და განსაზღვრული ალგორითმის გამოყენებით;

2. ამოცანები, სადაც შესაძლებელი არ არის ყველა რეალური პირობების გათვალისწინება და შესაძლებელია მხოლოდ ყველაზე მნიშვნელოვანი პირობების გამოყოფა, რადგან პირობების გარკვეული ნაწილის გამოყენება არ ხერხდება. ამიტომ ამოცანის პასუხი ატარებს არაზუსტ, მიახლოებით ხასიათს და პასუხის ძეგნის ალგორითმი სრულად არ არის ჩამოყალიბებული.

პირველი ჯგუფის ამოცანების ამოსახსნელად წარმატებით შეიძლება გამოყენებული იყოს ტრადიციული კომპიუტერული პროგრამები. რაც არ უნდა რთული იყოს ალგორითმი, შეზღუდული პირობების (შემაგალი პარამეტრები) ნაკრები გვაძლევს ალგორითმის შექმნის და ამოცანის ამოსხნის კონკრეტული პროგრამის დაწერის საშუალებას. ასეთი ამოცანის ამოსახსნელად არავითარი აზრი არ არქვს ნეიროქსელური ტექნოლოგიების გამოყენებას. ერთ-ერთ გამონაკლის წარმოადგენს ისეთი შემთხვევა, როდესაც პასუხის განსაზღვრის ალგორითმი გვაძლევს საკმაოდ დიდ და დროში გაწელილ პასუხებს, რაც ეწინააღმდეგება ამოცანის პრაქტიკულ მოთხოვნებს.

მეორე ჯგუფის ამოცანების ამოსახსნელად ნეიროქსელების გამოყენება გამართლებულია ყველა პარამეტრებით, ოღონდ ორი პირობით: პირველ რიგში, ნეირონული ქსელი უნდა იყოს უნივერსალური არქიტექტურის და სწავლების უნივერსალური ალგორითმით, რათა თავიდან ავიცილოთ ყოველი კონკრეტული ამოცანისათვის ქსელის ხელახალი დამუშავება და მეორე – ნეიროქსელის სწავლებისათვის სასწავლო მაგალითების არსებობა. ორივე ამ პირობის შესრულება იძლევა ნეიროქსელის შექმნის სისწრაფის მკვეთრ გაზრდას, რაც იწვევს ხარჯების შემცირებას.

პრაქტიკულად თითქმის ყველა ბიოსამედიცინო მეცნიერება შედგება მეორე ტიპის ამოცანებისაგან და სწავლებისთვის საწყისი მაგალითების შექმნა არ წარმოადგენს სირთულეს. ასეთ ამოცანებს მიეკუთვნებიან დიაგნოსტიკის, დიფერენციალური დიაგნოსტიკის, პროგნოზირების, მკურნალობის სტრატეგიის და ტაქტიკის შერჩევისა და სხვა.

მრავალწლიანმა გამოკვლევებმა აჩვენა, რომ სამედიცინო ამოცანები, რომლებსაც გააჩნიათ არაცხადი ხასიათი და იხსნებიან ცხადი მეთოდებით გარკვეული სიზუსტით, სრულიად არადაამაკმაყოფილებელია ფართო პრაქტიკული გამოყენებისათვის, კონკრეტულად დიაგნოსტიკის, პროგნოზირების და გადაწყვეტილების მიღების ამოცანების გადასაწყვეტად. სწორედ ასეთი არაცხადი ამოცანები წარმოადგენენ იდეალურ სფეროს ნეიროქსელური ტექნოლოგიების გამოყენებისათვის და ამან გამოიწვია ნეირონული ქსელების პრაქტიკულ კვლევებში წარმატებული გამოყენება.

ამრიგად, ნეიროქსელების გამოყენების ერთ-ერთი წარმატებული სფეროა მედიცინა. განვიხილოთ როგორ იქმნება ნეირონული ქსელი და როგორ ხდება მისი სწავლება. ჯერ განვიხილოთ ექიმის მიერ დიაგნოზის დასმის პროცედურა. უპირველეს ყოვლისა ექიმი აფიქსირებს პაციენტის მონაცემებს (სახელი, გვარი, ასაკი, სამუშაო ადგილი და სხვა). შემდეგ, როგორც წესი, ზომავს არტერიულ წნევას, აწარმოებს გარეგნულ დათვალიერებას, მოისმენს პაციენტის ჩივილებს, ეცნობა დაავადების ისტორიას, ლაბორატორიული ანალიზის შედეგებს, შეისწავლის ელექტროკარდიოგრამას და სხვა. ყოველივე ამის შედეგად ექიმს უგროვდება 20-დან 100-მდე და შესაძლებელია უფრო მეტი პარამეტრები, რომლებიც ახასიათებენ პაციენტს და მის მდგომარეობას. სწორედ ესენი არიან საწყისი მაჩვენებლები (პარამეტრები), რომელთა დამუშავების შედეგად, მისი გამოცდილების და ცოდნის ბაზაზე ექიმი ღებულობს კონკრეტულ გადაწყვეტილებას პაციენტის დაავადების შესახებ ანუ სვამს დიაგნოზს.

თუ ჩვენ გვინდა ნეიროქსელის საშუალებით ექიმის მოდელის მოდელირება, მაშინ უპირველეს ყოვლისა საჭიროა განისაზღვროს შემავალი X ვექტორი და გამომავალი Y ვექტორი, მივცეთ მას განზომილება და დავაზუსტოთ მისი ყოველი კომპონენტი. X ვექტორში უნდა გავითვალისწინოთ ის პარამეტრები, რომლებსაც ექიმი ღებულობს ან განსაზღვრავს პაციენტისაგან. მაგალითად, x_1 კომპონენტად შეიძლება მივიღო დაბადების თარიღი, x_2 – სქესი (ნულისა და ერთის საშუალებით), x_3 – პაციენტის წონა, x_4 – არტერიული წნევა, x_5 – სხეულის ტემპერატურა და ა.შ. ზედმეტი არ იქნება თუ გავითვალისწინებთ პაციენტის თვალის ფერს, თმის ფერს, ზოლიაქოს ნიშანს და სხვა მონაცემებს, რომლებიც განსაზღვრავენ ორგანიზმის თავისებურებებს და შესაბამისად მოქმედებენ ამა თუ იმ დაავადების წარმოშობის ალბათობაზე. გამოსავალ Y ვექტორში საჭიროა მოცემული იყოს ყველა შესაძლო დაავადების კოდები, რომელთა აღმოჩენა შეუძლია ექიმს.

შემდეგ საჭიროა შესასწავლი ამონარჩევის შექმნა. ეს არის ნეიროქსელის შექმნის მეტად მნიშვნელოვანი ეტაპი, სადაც პროგრამისტი და ექიმი-ექსპერტი (ექსპერტები) პერსპექტონის სწავლებისათვის აყალიბებენ საჭირო რაოდენობის სასწავლო მაგალითებს (შემთხვევებს). თუ სამუშაოში ჩართული იქნება პათოლოგანატომი, მაშინ ცრუ დიაგნოსტიკის რაოდენობა საგრძნობლად შემცირდება და რასაკვირველია გაიზრდება ნეიროქსელის მუშაობის ეფექტიანობა.

ამრიგად, ექიმი-სპეციალისტების და პროგრამისტების ხანგრძლივი და შრომატევადი მუშაობის შედეგად გროვდება შესასწავლი ამონარჩევის ერთობლიობა, რომლებიც შედგებიან ორი X_q და D_q წყვილი ვექტორების

ერთობლიობისაგან ($q=1,2,\dots,Q$). შემდეგი ამოცანა მდგომარეობს პერსპექტონის პროექტირებაში და სწავლებაში. პერსპექტონის პროექტირება ნიშნავს ფარული შრეების და ნეირონების რაოდენობების განსაზღვრას და აქტივაციის ფუნქციის შერჩევას. სწავლებისათვის შეიძლება გამოყენებული იყოს ცდომილების უკუგავრცელების ალგორითმი.

ნეიროქსელმა უნდა შეძლოს სასწავლო ამონარჩევის ნებისმიერი X_q ვექტორის ასახვა Y გამოსავალ ვექტორზე, რომელიც ემთხვევა (ან თითქმის ემთხვევა) D_q ვექტორს. გარდა ამისა, ახალი პაციენტის შემთხვევაში, რომელიც ხასიათდება ახალი შემავალი X ვექტორით, ნეიროქსელმა უნდა განსაზღვროს მისთვის ახალი Y ვექტორი, რომელიც შეიცავს სწორ დიაგნოზს. სხვა სიტყვებით რომ ვთქვათ, პერსპექტონმა უნდა განსაზღვროს მასში ჩადებული ცოდნა, კერძოდ უნდა შეძლოს ახალი, აქამდე არ დაფიქსირებული დავადების დიაგნოზის დასმა.

განვიხილოთ მედიცინაში ნეიროქსელების გამოყენების ზოგიერთი ყველაზე უფრო საინტერესო მაგალითები. განსაკუთრებულ ინტერესს იწვევს ჯანდაცვის სისტემაში დიაგნოსტიკური და დიფერენციალური დიაგნოსტიკური ნეიროქსელების გამოყენება. ამ დროს შესაძლებელია გამოყენებული იყოს ნაირსახეობის მონაცემები: ანამნეზი, კლინიკური გამოკვლევის, ლაბორატორიული კვლევის შედეგები და სხვა პარამეტრები.

ნეიროქსელების გამოყენების ერთ-ერთ ყველაზე საინტერესო სფეროს წარმოადგენს კარდიოლოგია. იტალიაში დამუშავებული იყო არტერიული ჰიპერტონიის დიაგნოსტიკის და მკურნალობის ნეიროქსელი, რომელიც შედგებოდა სამი ნეირონული მოდულით. ერთი მოდულის პასუხი წარმოადგენს სხვა მოდულების შესავალ მონაცემებს. კვლევის დასაწყისში იზომება პაციენტის არტერიული წნევის (სისტოლური და დიასტოლური) მნიშვნელობები ყოველ ნახევარ საათში ერთი დღე-ღამის განმავლობაში. ყოველი საათის მაძვენებლები გასაშუალებიან. ამრიგად იქმნება მასივი, რომელიც შედგება 48 არტერიული წნევის მნიშვნელობებით (24 სისტოლური და 24 დიასტოლური).

ამის შემდეგ პირველი მოდული, რომელიც შედგება ორი სამშრიანი ნეირონული ქსელისაგან (თითოეული მათგანი შედგება 2 შემავალი, 4 ფარული და 24 გამომავალი ნეირონებისაგან). არსებული მონაცემებით და ავადმყოფის ასაკის მიხედვით განისაზღვრებიან შესაბამისი „საჭირო“ მნიშვნელობები და დარღვებიან რეალურს. პარალელურად მეორე მოდული, რომელიც წარმოადგენს ორშრიან ნეირონულ ქსელს 17 შემავალი და 4 გამომავალი ნეირონით, არსებული კლინიკური სიმტომებით (სიმპტომატიკა, ანამნეზი) განსაზღვრავს ჰიპერტონიის სამკურნალო საშუალებებს, რომლებიც შეიძლება გამოყენებული იყოს პაციენტის სამკურნალოდ. ორივე მოდულის გამოსავალი ინფორმაცია მიეწოდება მესამე მოდულის შესასვლელს.

მესამე მოდული წარმოადგენს ექვსშრიან ნეირონულ ქსელს, რომელიც ეფუძნება ჰიპერტენზიული პრეპარატების 4 ჯგუფს (დიურეტიკები, ბეტადრენობლოკატორები, ანგიოტენზიის ინჰიბიტორები, კალციუმის არხების ბლოკატორები). მესამე მოდულის მიზანია პაციენტს დაუნიშნოს დღეღამისეული (საათობრივი) წამლის მიღების გრაფიკი ამ ოთხივე (თუ საჭიროა) ჯგუფიდან. აქედან გამომდინარე, მოდულს გააჩნია 96 გამოსავალი ნეირონი (4 პრეპარატი გამრავლებული 24 საათზე). ყოველი გამოსავალი ნეირონიდან იხსნება დოზა, რომელიც შეესაბამება ერთ პრეპარატს, რომელიც დანიშნულია მიმდინარე ერთი საათისათვის. ცხადია, რომ რეალურ სიტუაციებში გამოსავალი მონაცემების უმეტესობა ნულის ტოლია.

ამრიგად, პაციენტისათვის იქმნება ჰიპერტონიის მეურნეობის ოპტიმალური სქემა. სისტემის განსხვავებულ თვისებას წარმოადგენს ექიმის გამოცდილების მიღების საშუალება. ამისათვის გათვალისწინებულია სპეციალური ბლოკი, რომელსაც კომპიუტერის ეკრანზე გამოყავს სადღეღამისო არტერიული წნევის მრუდები და ექიმს სთავაზობს შეიტანოს კომპიუტერში მისი აზრით ჰიპოტენზიის მეურნეობის საშუალებების სასურველი სადღეღამისო დოზა. შეტანილი მონაცემები მოთავსდება მონაცემთა ბაზაში და ნებისმიერ დროს შესაძლებელია მის გამოყენება ქსელის სწავლების დროს.

ანალოგურად წარმოებს პერიფერიული სისხლძარღვების დაავადების გიაგნოსტიკა. მაგალითად, აკუსტიკური სიგნალის ნეიროქსელური ანალიზი საშუალებას ძლევს გულის სარქველების ხმაურის დიაგნოსტიკის წარმოება. კარდიოლოგიის კლასიკურ პრობლემას წარმოადგენს მკბ სიგნალის ინტერპრეტაცია. გლახგოს უნივერსიტეტის თანამშრომლები აწარმოებდნენ გამოკვლევებს ნეიოქსელის გამოყენებისას მკბ დიაგნოსტიკისა მიოკარდიუმის ინფაქტის დროს. ამ შემთხვევაში ნეიროქსელის შემავალ მონაცემებს წარმოადგენდნენ 12 არხიანი მკბ და 12 არხიანი ვექტოროკარდიოგრამის პარამეტრების მაჩვენებლები. მკვლევარებმა ნეირონულ ქსელს შეასწავლეს 360 მკბ მონაცემები. შემდეგ ხდებოდა შესწავლილი ქსელის ტესტირება წინასწარ ცნობილი პასუხებით. ერთდროულად ტესტირების მაგალითებზე გამოყენებული იყო ლოგიკური მეთოდი (წინასწარ მოცემული ალგორითმით). შემდეგ ხდებოდა ნეიროქსელის პასუხების შედარება ლოგიკური ლოგარითმით მიღებულ პასუხებთან. შედარებამ აჩვენა, რომ ბევრ შემთხვევაში ნეიროქსელის მგრძნობელობა და სპეციფიურობა აღმოჩნდა მაღალი ვიდრე ლოგიკური მეთოდისა.

ნეიროქსელებში გამოყენებული არააღგორითმიზებული ან ძნელად ალგორითმიზებული ამოცანები შეიძლება დავყოთ ორ, პრინციპულად დანსხვავებული ტიპის – კლასიფიკაციის და პროგნოზირების ამოცანებად.

კლასიფიკაციის ამოცანები წარმოადგენენ სამედიცინო ამოცანების ფართო სპექტრს, სადაც პასუხი არის კლასი, ანუ წინასწარ მოცემულ კლასებიდან ერთ-ერთის შერჩევა. კლასიფიკაცია შეიძლება იყოს ბინარული, როცა მოცემულია ორი კლასი და n - განზომილებიანი, როცა მოცემულია ორზე მეტი კლასი. კლასიფიკაციის ამოცანის მნიშვნელოვან თვისებას წარმოადგენს ერთი და მხოლოდ ერთი კლასის შერჩევის საშუალება. ამიტომ დიაგნოზის დასმა არ შეიძლება ჩაითვალოს ერთ კლასიფიკაციურ ამოცანად, რადგან ერთ ადამიანს შეიძლება ერთდროულად სჭირდეს რამოდენიმე დავადება. ერთი პასუხის შერჩევის შეუძლებლობის შემთხვევაში (მრავლობითი შერჩევა) ამოცანა იყოფა ქვეამოცანებად, სადაც თითოეული მათგანი წარმოადგენს იდენტიფიკაციის დამოუკიდებელ ამოცანას.

ნეიროქსელების მეორე ამოცანას მიეკუთვნებიან პროგნოზირების ამოცანები. ისინი იყოფიან რიცხვის (ერთგანზომილებიანი პროგნოზირება) და ვექტორის (განზოგადოებული შემთხვევა) პროგნოზირების ამოცანებად. ვექტორული პროგნოზირება ითვალისწინებს იმას, რომ პასუხები შეიძლება იყოს ერთმანეთის მიმართ დამოუკიდებლად წარმოდგენილი, რომლებიც მრავალგანზომილებიან სივრცეში ქმნიან წერტილს, რომლის განზომილება პროგნოზირებადი რიცხვების ტოლია.

სამედიცინო ნეიროქსელური ექსპერტული სისტემები. როგორც ცნობილია, ტრადიციული ექსპერტული სისტემები ცოდნის ინჟინრების და ექსპერტების ერთობლივი მუშაობის შედეგია. ე.ი. ძირითადი სამუშაო ექსპერტებზე მოდის. ამ დროს წამოიჭრება ორი სერიოზული პრობლემა.

პირველი პრობლემა შედგება იმაში, რომ რთული რეალური ამოცანის ამოხსნისას ლოგიკური წესები მნიშვნელოვნად იზრდება. ზოგჯერ დგება ისეთი სიტუაცია, რომ შეუძლებელი ხდება მათი გააზრება. ამიტომ ქვეამოცანებად დაყოფა ყოველთვის არ შეეძლება, რადგან ჯერ ერთი, ამოცანის დაყოფა ქვეამოცანებად ყოველთვის არ არის ადვილი და მეორე – დაყოფის დროს ზოგჯერ იკარგება ზოგიერთი ურთიერთკავშირები.

მეორე, უფრო სერიოზული პრობლემა შედგება იმაში, რომ საკმაოდ ხშირად შეუძლებელი ხდება გამოთვლითი პროცედურის ლოგიკური წესებით გამოსახვა. ეს პრობლემა განსაკუთრებით ეხება მედიცინას, სადაც გადაწყვეტილების მიღების პროცედურა ძირითადად ეფუძნება ექიმის გამოცდილებას და მის ინტუაციას. ყოველ ასეთი შემთხვევის დროს აღნიშნავენ, რომ ამოცანა ალგორითმიზაციას არ ექვემდებარება. ყოველი ამ პრობლემისაგან თავისუფალია ნეირონული ქსელები.

ნეირონული ექსპერტული სისტემებს ჩვეულებრივ სისტემასთან შედარებით გააჩნია შემდეგი უპირატესობით:

1. ნეიროქსელები გადაწყვეტილებას დებულობენ გამოცდილების საფუძველზე, რომელსაც დებულობენ დამოუკიდებლად. „დამოკიდებულება“ ამ შემთხვევაში ნიშნავს იმას, რომ სისტემის შემქმნელს არ სჭირდება შემავალი მონაცემებსა და საჭირო გადაწყვეტილებებს შორის ურთიერთკავშირების დადგენა, გარკვეული სტატისტიკური დამუშავებაზე დროის დაკარგვა, მათემატიკური აპარატის შერჩევა, მათემატიკური მოდელების შექმნა და შემოწმება.

2. ნეიროქსელის მიღებული გადაწყვეტილება არ არის კატეგორიული. ქსელი იძლევა შედეგს გარკვეული დამაჯერებლობის ხარისხით, რაც მომხმარებელს უტოვებს შესაძლებლობას კრიტიკულად შეაფასოს პასუხი.

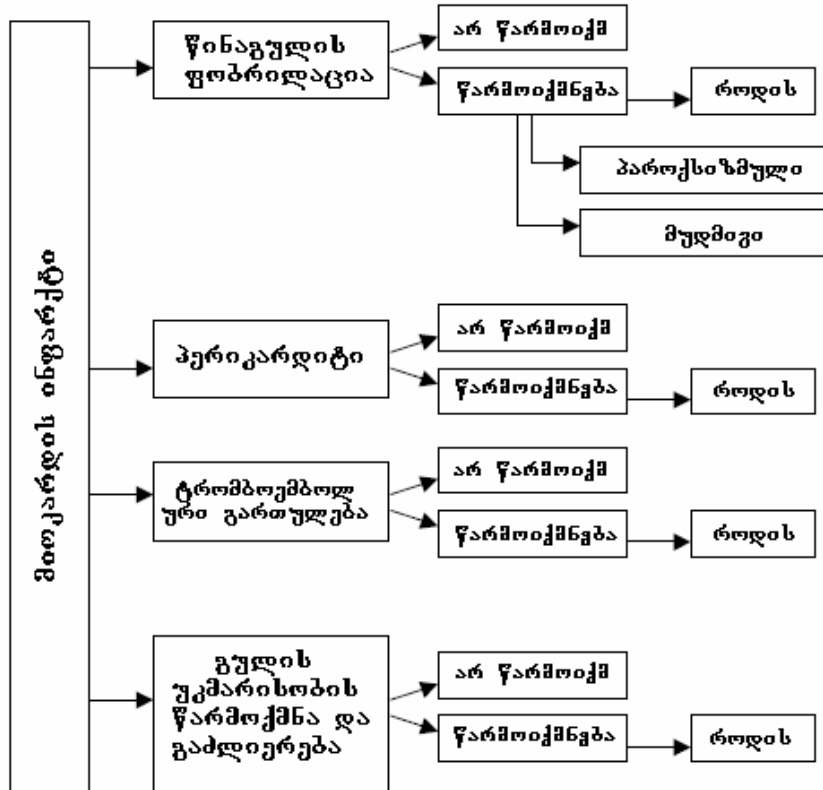
3. ნეიროქსელები იძლევიან გადაწყვეტილების მიღების სიტუაციის მოდელირების საშუალებას;

4. ნეიროქსელი პასუხს იძლევა ძალიან სწრაფად, რაც იძლევა მისი გამოყენების საშუალებას სხვადასხვა დინამიკურ სისტემებში, სადაც საჭიროა გადაწყვეტილების დაუყოვნებელი მიღება.

5. ნეიროქსელები ამარტივებენ ექსპერტული სისტემების შექმნის საშუალებებს.

მიოკარდის ინფარქტის გართულების პროგნოზირება. მიოკარდის ინფარქტის პროგნოზირება მეტად აქტუალურია და წარმოადგენს კარდიოლოგიის ერთ-ერთ ურთულეს ამოცანას. სტაციონარში ავადმყოფის მოთავსებისთანავე ჩატარებული გადაუდებელი პროცედურების შემდეგ საჭიროა მაშინვე პროგნოზირების განხორციელება. პროგნოზირება უნდა იყოს საკმაოდ სწრაფი და მრავალჯერადი. ინფარქტის გართულების პროგნოზის მიღება ექიმს ხელს უწყობს მიხანმძიმართულად წარმატოს პროფილაქტიკური ღონისძიებები. პროგნოზს შეუძლია განსაზღვროს ხანგრძლივი და ინტენსიური მეურნეობა ტრომბოემბოლური გართულების საშიშროებისას ანტიკოაგულანტებით, ანტიარითმიული პრეპარატებით – არითმიის პროფილაქტიკისათვის.

მიოკარდის ინფარქტის გართულების პროგნოზირების ნეირონული ქსელის შექმნისათვის შერჩეული იყო 4 გართულება: წინაგულის ფიბრილაცია. ტრომბოემბოლური გართულებები, პერიკარდიტი, და გულის უკმარისობის წარმოქმნა (ან გადრეკვა). მიოკარდის ინფარქტის გართულების პროგნოზირების ბლოკ-სქემა მოცემულია შემდეგ ნახაზზე:



შერჩეული იყო 32 სწავლების პარამეტრი, რომლებიც ახასიათებენ პაციენტს სტაციონარში მოყვანის მომენტში მიოკარდიის ინფარქტის კლინიკურ მდგომარეობას, ანამნეზის, ლაბორატორიული და ფუნქციონალური კვლევის მონაცემები; ასაკი, სქესი, ინფარქტის ლოკალიზაცია და სიღრმე (ელექტროკარდიოგრაფიული სიგნალიდან), ანამნეზში ინფარქტების რაოდენობა, სტენოკარდიის წინა პერიოდის მახასიათებელი, ჰიპერტონიური დაავადების არსებობა და სირთულე, ანამნეზში არსებული გულის რიტმის და გამტარებლობის დარღვევები, ენდოკრინული დაავადებები და სხვა.

შერჩეული პარამეტრების უმეტესობა მიეკუთვნება სუბიექტურ კატეგორიებს, რომლებიც შერჩეული იყო კლინიკაში დაგროვილი გამოცდილებით. ბევრი პარამეტრი ღებულობს მხოლოდ ორ მნიშვნელობას 0 (სიმპტომი არ არის) და 1 (სიმპტომი არის). შერჩეული პარამეტრებით ხდება ოთხივე გართულების პროგნოზირების ნეიროქსელის სწავლება. სწავლებაში გამოყენებული იყო როგორც სასწავლო, ასევე საკონტროლო ამონარჩევები. მიღებული შედეგები განისაზღვრება კონსოლიუმის საშუალებით.

გამოყენებული ლიტერატურა

1. Ф. Уоссермен. Нейрокомпьютерная техника: теория и практика. М., Мир, 1992 (эл. версия)
2. Заенцев И.В. Нейронные сети: Основные модели. Учебное пособие , Воронеж, 1999 (эл.версия).
3. Ясницкий Л.Н. Введение в искусственный интеллект.Учеб.пособие, М.,Академия 2008 (эл.версия)
4. Каширина И.Л. Нейросетевые технологии. Учебное пособие , Воронеж, 2008, (эл.версия)
5. В.Дьяконов, В.Круглов Математические пакеты расширения MATLAB. Питер, 2001 (эл.версия).