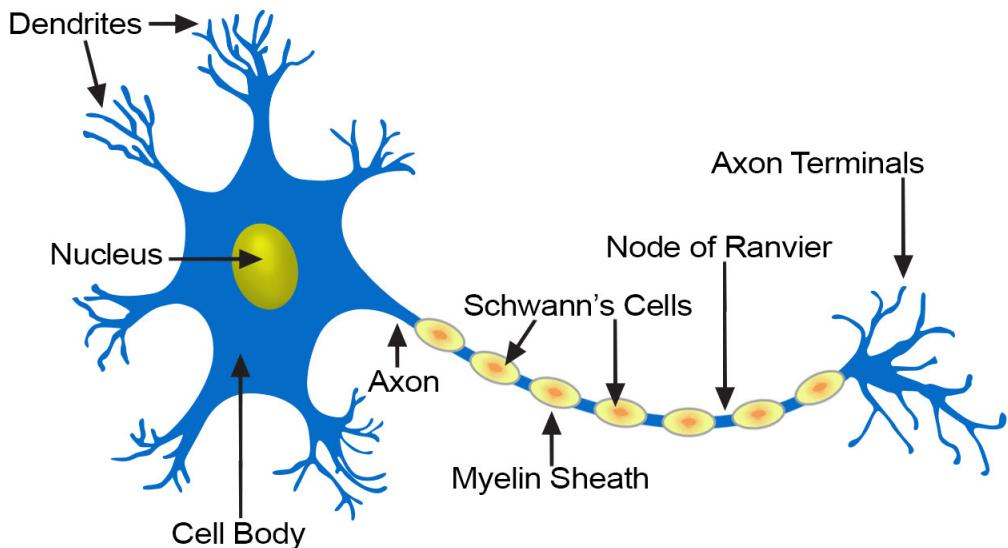


დენდრიტები და აქსონები ნეირონული უკრედების ნაწილებია, რომლებიც ნერვულ სისტემაში ინფორმაციის გადაცემას უზრუნველყოფს.

Structure of a Typical Neuron



დენდრიტები ნეირონების განშტოებული დაბოლოებების, სადაც სიგნალები სხვა ნეირონებიდან ან სენსორული რეცეპტორებიდან შედის.

მათი აგებულება ხის ტოტებს წაგავს, უკრედის სხეულიდან გაშლილი მრავალი პატარა გამონაზარდით.

დენდრიტები შეიცავს რეცეპტორებს მეზობელი ნეირონების მიერ გამოყოფილი ნეიროტრანსმიტერების, იგივე ქიმიური მესენჯერებისთვის.

როდესაც ნეიროტრანსმიტერი უერთდება დენდრიტებზე არსებულ რეცეპტორებს, მას შეუძლია აღავზნოს ან დათრგუნოს ნეირონი, რაც გავლენას ახდენს იმაზე, გამოიმუშავებს თუ არა ის ელექტრულ სიგნალს.

დენდრიტების მიერ მიღებული სიგნალები ინტეგრირდება ნეირონში. საკმარისი სიმძლავრის შემთხვევაში, წარმოიქმნება ელექტრული იმპულსი, რომელსაც მოქმედების პოტენციალი ეწოდება.

აქსონები ნეირონების გრძელი, თხელი ნაწილებია, რომლებიც ელექტრულ იმპულსებს უკრედის სხეულიდან სხვა ნეირონების, კუნთების ან ჰირკვლების მიმართულებით გადასცემს.

აქსონის ბოლოს არის ტერმინალური ტოტები, რომლებიც ქმნის კავშირებს, იგივე სინაფსებს, დენდრიტებთან ან სხვა ნეირონების უკრედის სხეულებთან.

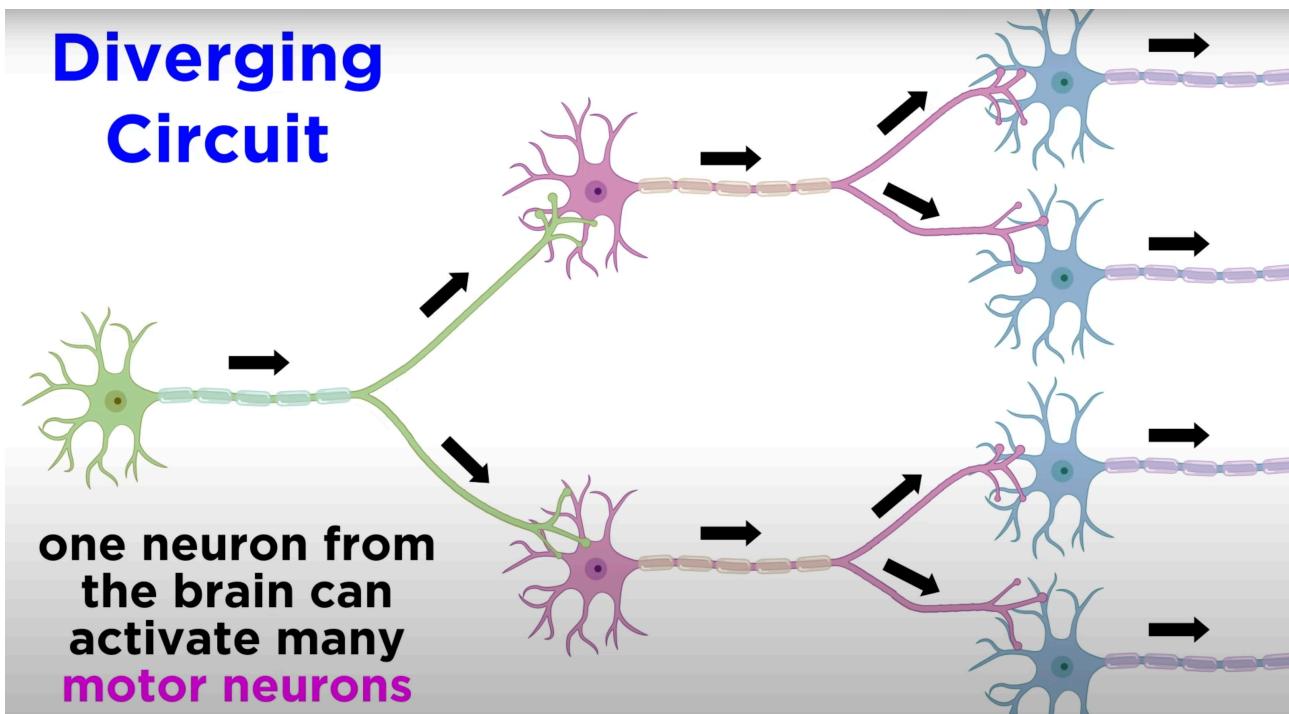
აქსონები დაფარულია მიელინის გარსით, ცხიმოვანი საიზოლაციო ფენით, რომელიც ხელს უწყობს ელექტრული იმპულსების გადაცემის დაჩქარებას.

ელექტრული იმპულსი, ანუ მოქმედების პოტენციალი, მოძრაობს აქსონის გასწვრივ, სინაფსში ნეიროტრანსმიტერების განთავისუფლებას იწვევს, რითაც ნეირონს ან სამიზნე უკრედის უკავშირდება.

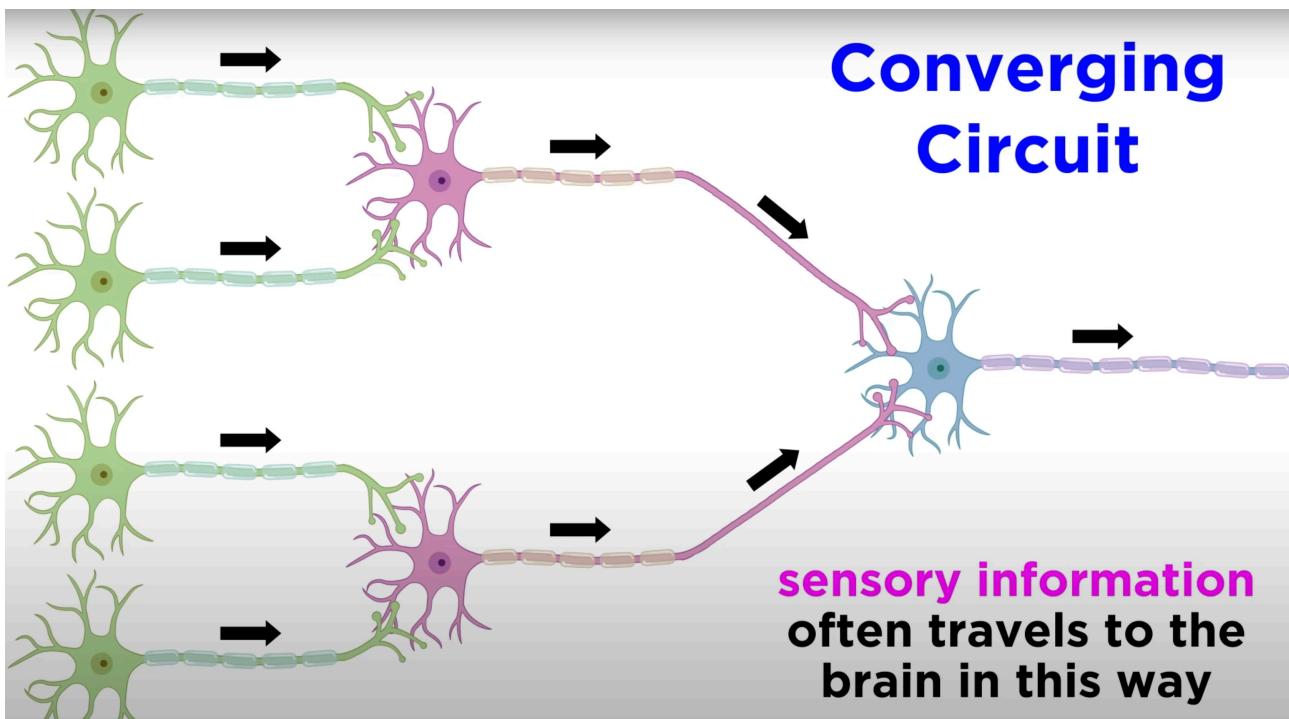
დენდრიტები იღებს სიგნალებს სხვა ნეირონებისგან, ხოლო აქსონები ამ სიგნალების სხვა უკრედებში გადაცემაზეა პასუხისმგებელი. ამ გზით ისინი ერთად ქმნიან კომუნიკაციის საფუძველს ნერვულ სისტემაში.

თავის ტვინში განსხვავებული ნეირონული ქსელები დაკავშირებულია ნეირონული გზებით, სადაც, მაგალითად ერთი ნეირონი ასტრიმულირებს რამდენიმე ნეირონს მოცემულ ქსელში (როგორც ფოტოზე ნაჩვენები). ამ ტიპის ქსელები იძლევა სიგნალების გაძლიერებისა და გავრცელების საშუალებას, რის შედეგადაც ხდება კომპლექსური ინფორმაციის დამუშავება და სხვადასხვა ფუნქციების კოორდინაცია.

Diverging Circuit



Converging Circuit



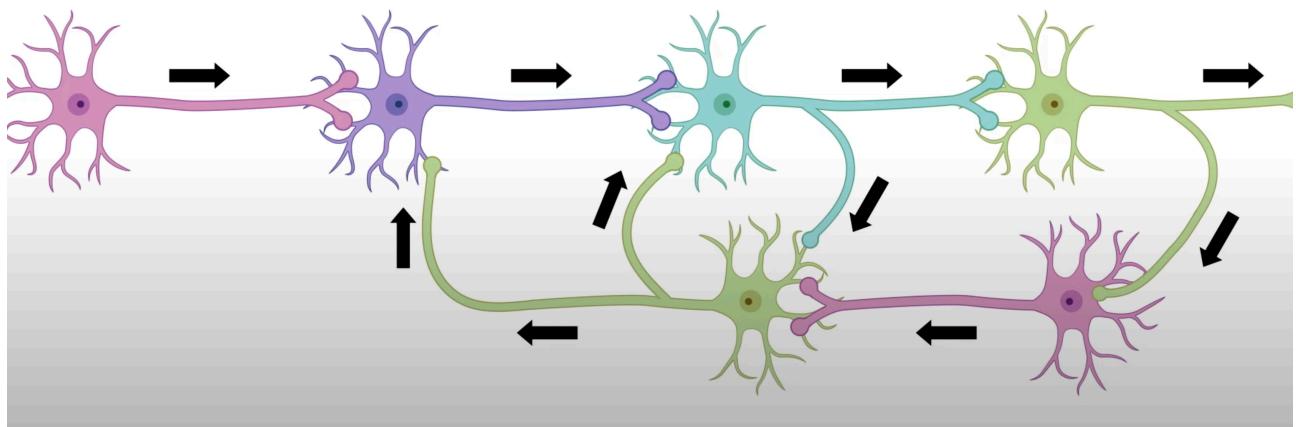
რიტმული აქტივობების განმახორციელებელი წრედი შემდეგნაირია:

ასეთი ქსელები პასუხისმგებელია მაგალითად სუნთქვაზე, სიარულზე...

უფრო კომპლექსური მოქმედებებისას აქტიურდება გაცილებით რთული ქსელები.

Reverberating Circuit

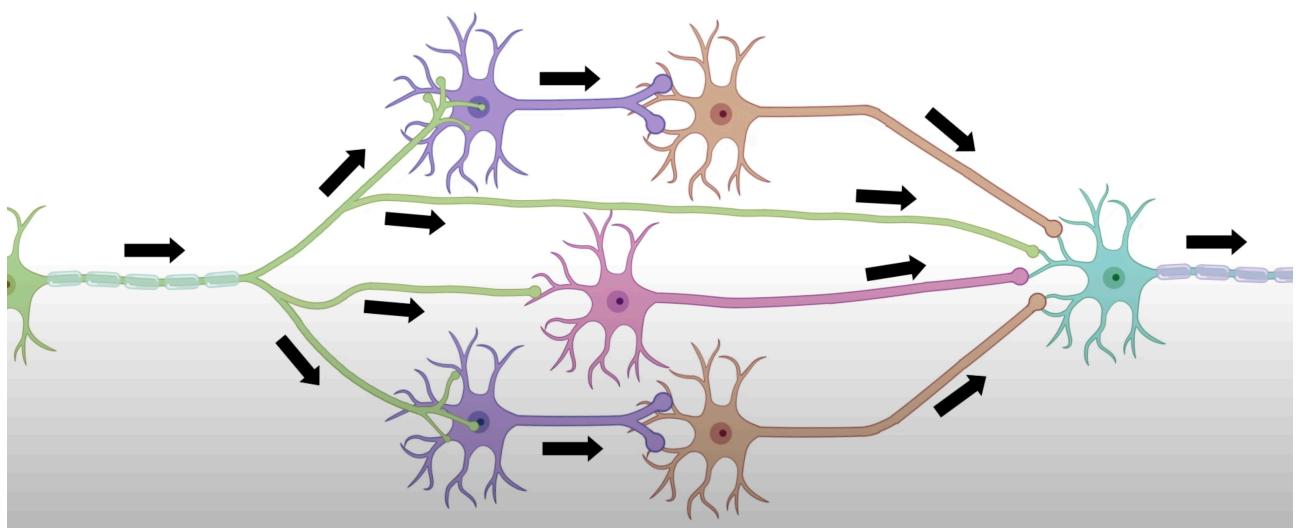
controls **rhythmic activity** (like breathing)



ხელოვნური ნეირონი არის მათემატიკური ფუნქცია, რომელიც შექმნილია ბიოლოგიური ნეირონების მოდელის მიხედვით, სადაც თითოეული ნეირონი იღებს **შენატანებს**, აწონის მათ ცალკე, აჯამებს და გადასცემს ამ ჟამს არანრთვივი ფუნქციის მეშვეობით **გამომავალი წარმოების**თვის.

Parallel After-Discharge Circuit

involved in **complex brain activity** (like coincidence detection)



ბიოლოგიური ნეირონი	ხელოვნური ნეირონი
სომა	ნოუდი

დენტრიტი	შესავალი
სინაფსი	წონები
აქსონი	გამოსავალი

ტერმინი Perceptron შემოიღო ფრენკ რობერტ ბლატმა 1957 წელს. მან შემოგვთავაზა პერცეპტრონის დასწავლის წესი, რომელიც დაფუძნებულია ორიგინალური **MCP** ნეირონზე. Perceptron ბინარული კლასიფიკატორების ზედამხედველობითი დასწავლის ალგორითმია, რომელიც საშუალებას იძლევა ერთდროულად მოხდეს **ელემენტების** დასწავლა და დამუშავება ნეირონების მიერ.

Perceptron არის ხელოვნური ნერვული უკრედიტის ტიპი, რომელიც მლ-ში ფუნდამენტური კონცეფციაა. პერცეპტრონის ძირითადი კომპონენტებია:

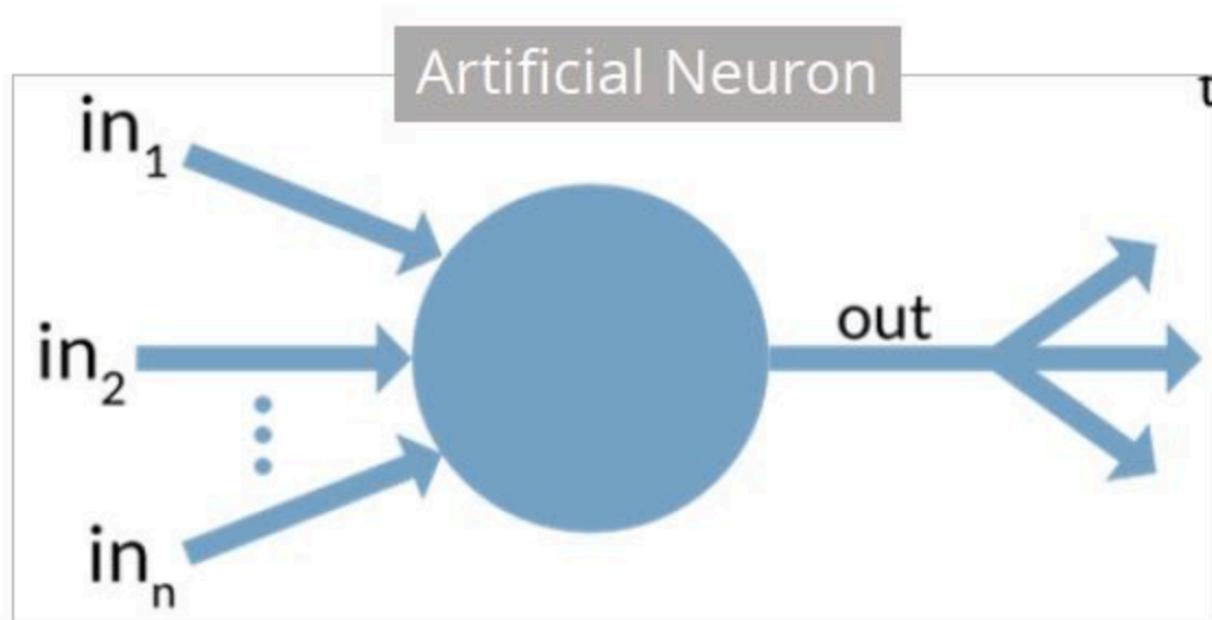
შესავალი: ინფუზ ფენა შედგება ერთი ან მეტი ინფუზ ნეირონისგან, რომლებიც სიგნალებს გარე სამყაროდან ან ნერვული ქსელის სხვა შრეებიდან იღებენ.

წონა: თითოეული ინფუზ ნეირონი ასოცირდება წონასთან, რომელიც წარმოადგენს ინფუზ ნეირონსა და აუთფუზ ნეირონს შორის კავშირის სიძლიერეს.

Bias: bias (მიკერძოების) ტერმინი ემატება ინფუზ ფენას, რათა მონაცემებში რთული შაბლონების მოდელირებისას უზრუნველყოს პერცეპტრონის დამატებითი მოქნილობა.

აქტივაციის ფუნქცია: აქტივაციის ფუნქცია განსაზღვრავს პერცეპტრონის **აუთფუზ ინფუზის** შეწონილი ჰამისა და bias ტერმინის საფუძველზე. პერცეპტრონებში გამოყენებული აქტივაციის საერთო ფუნქციები მოიცავს ნაბიჭის, სიგმოიდურ და ReLU ფუნქციას.

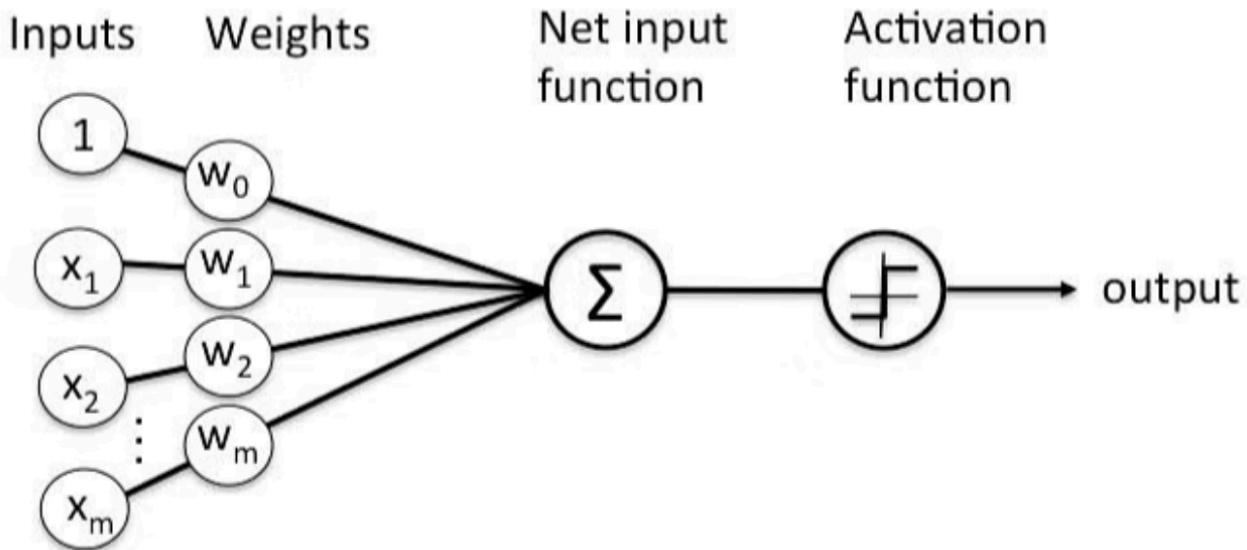
გამოსავალი: პერცეპტრონის აუთფუზთი არის ერთი ორობითი მნიშვნელობა, 0 ან 1, რომელიც მიუთითებს კლასს ან კატეგორიას, რომელსაც მიეკუთვნება ინფუზ მონაცემები.



წვრთნის ალგორითმი: პერცეპტრონი ჩვეულებრივ იწვრთნება ზედამხედველობის ქვეშ მყოფი დასწავლის ალგორითმის გამოყენებით, როგორიცაა **პერცეპტრონის დასწავლის ალგორითმი** ან უკან გავრცელება (ინგლისურადაც ჩაწერე აქ). წვრთნის დროს, პერცეპტრონის წონა და ბაიასი (მიკერძოება) რეგულირდება, რათა მინიმუმამდე

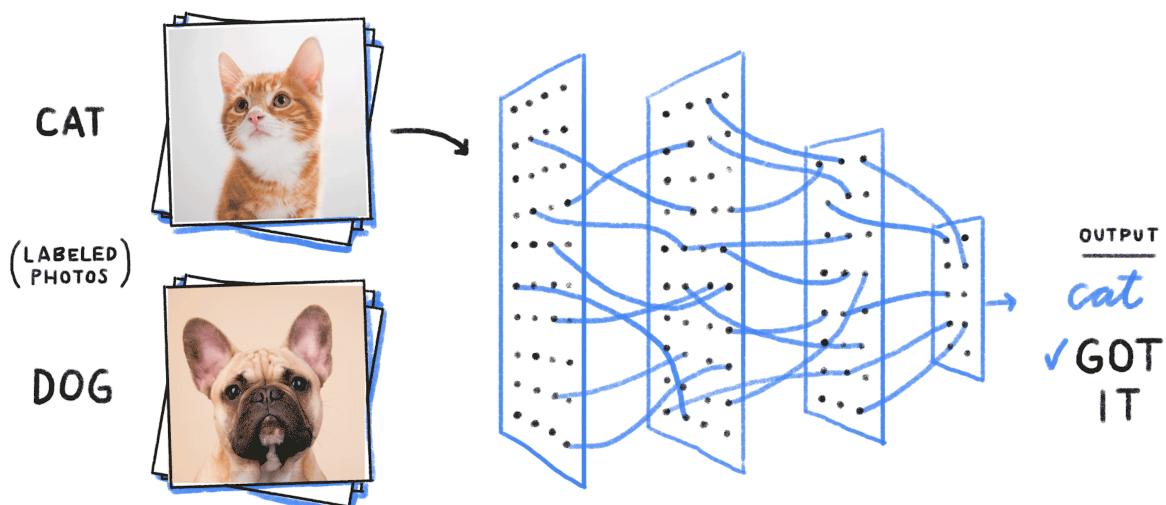
დაიყვანოს შეცდომა წინასწარმეტყველურ გამომუშავებასა და ნამდვილ გამომუშავებას შორის, საწვრთნელი მაგალითების მოცემული ნაკრებისთვის.

მთლიანობაში, პერცეპტრონი არის მარტივი, მაგრამ ძლიერი ალგორითმი, რომლის გამოყენებაც ორობითი კლასიფიკაციის ამოცანების შესასრულებლად არის შესაძლებელი, რამაც, თავის მხრივ, გზა გაუხსნა მეტად რთული ნეირონული ქსელების წარმოებას, რომლებიც დღეს ღრმა დასწავლაში გამოიყენება.



ღრმა დასწავლაში, კლასიფიკაცია და რეგრესია არის **ზედამხედველობითი** სასწავლო ამოცანების ორი ფუნდამენტური ტიპი:

კლასიფიკაციის ამოცანები გულისხმობს კატეგორიის ან კლასის პროგნოზირებას. გამომავალი, როგორც წესი, არის დისკრეტული ლეიბლი ან კლასი. მაგალითად, ცხოველების სურათების კლასიფიკაცია კატეგორიებად, როგორიცაა "კატა", "ძაღლი" ან "ფრინველი" არის კლასიფიკაციის ამოცანა.



ღრმა დასწავლისას, კლასიფიკაციის პრობლემებს ხშირად წყვეტენ ისეთი ტექნიკის გამოყენებით, როგორიცაა კონვოლუციური ნერვული ქსელები (CNN), რომლებიც შესაფერისია ისეთი სტრუქტურირებული მონაცემების დასამუშავებლად, როგორიცაა სურათები, ან რეკურსიული ნერვული ქსელები (RNN) მაგალითად თანმიმდევრული მონაცემებისთვის, როგორიცაა ტექსტი.

რეგრესიის ამოცანები გულისხმობს უწყვეტი რიცხვითი მნიშვნელობის წინასწარმეტყველებას. გამომავალი არის რეალური რიცხვი ან რეალური რიცხვების ნაკრები. მაგალითად, სახლების ფასების პროგნოზირება ისეთი მახასიათებლების მიხედვით, როგორიცაა მდებარეობა, ზომა და საძინებლების რაოდენობა, არის რეგრესიის პრობლემა. ღრმა სწავლებისას, რეგრესიის პრობლემების გადაჭრა შესაძლებელია სხვადასხვა არქიტექტურის გამოყენებით, მათ შორის FNN-ით და RNN-ით რაც დამოკიდებულია უშუალოდ პრობლემის თავისებურებაზე.

როგორც კლასიფიკაციის, ასევე რეგრესიის პრობლემები ხშირად გვხვდება სხვადასხვა სფეროში, მათ შორის კომპიუტერული ხედვა, ბუნებრივი ენის დამუშავება, ჰანდაცვა, ფინანსები და სხვა.

აქტივაციის ფუნქცია ხელოვნური ნერვული ქსელების გადამწყვეტი კომპონენტია, რომელიც პასუხისმგებელია ქსელის გამომავალში არაწრფივობის იმპლემენტაციაზე. აქტივაციის ფუნქციები განსაზღვრავს, უნდა გააქტიურდეს თუ არა ნეირონი.

რატომ გვჭირდება არაწრფივობა?

არაწრფივობის გარეშე, ნერვულ ქსელებს შეეძლოთ მხოლოდ წრფივი ფუნქციების წარმოდგენა, რაც მნიშვნელოვნად ზღუდავს მათ უნარს, შექმნან რთული კავშირები მონაცემებში. არაწრფივი აქტივაციის ფუნქციები ნერვულ ქსელებს საშუალებას აძლევს ისწავლონ და წარმოადგინონ რთული შაბლონები, რის მეშვეობითაც იჭრება რეალური პრობლემების ფართო სპექტრი, გამოსახულების ამოცნობიდან ბუნებრივი ენის დამუშავებამდე.

გვაქვს აქტივაციის სხვადასხვა ფუნქციები:

Different Activation Functions:

$$1. \text{ Sigmoid: } \sigma(x) = \frac{1}{1+e^{-x}}$$

- Squashes the input values between 0 and 1.
- Historically popular but has some drawbacks like vanishing gradients.

$$2. \text{ Hyperbolic Tangent (tanh): } \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- Similar to sigmoid but squashes the input values between -1 and 1.
- Also susceptible to vanishing gradients.

$$3. \text{ Rectified Linear Unit (ReLU): } f(x) = \max(0, x)$$

- Outputs the input directly if positive, and zero otherwise.
- Simple and computationally efficient.
- Addresses the vanishing gradient problem and accelerates convergence in deep networks.

$$4. \text{ Leaky ReLU: } f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{otherwise} \end{cases}$$

- Similar to ReLU but allows a small, positive gradient for negative input values (α is a small constant).

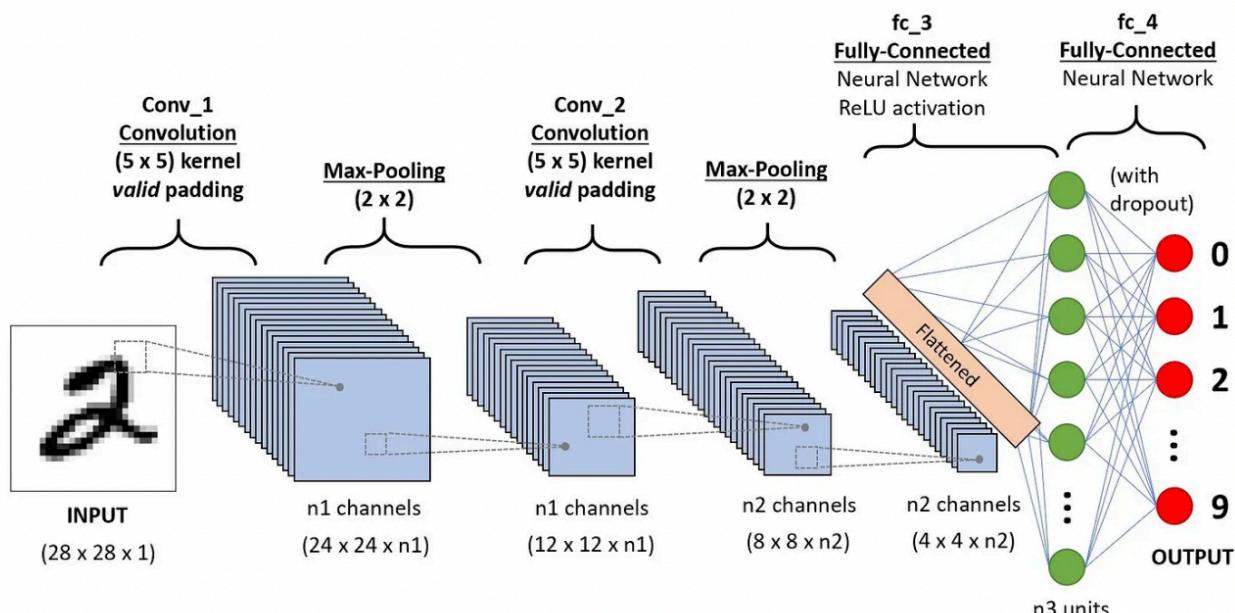
$$5. \text{ Exponential Linear Unit (ELU): } f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha(e^x - 1) & \text{otherwise} \end{cases}$$

- Smooth function that alleviates the dying ReLU problem and can help with learning robust representations.

ამ ფუნქციებს შორის ყველაზე სწრაფად შეიძლება ჩაითვალოს RELU, რადგან ის მოიცავს მარტივ ზღურბლოვან ოპერაციებს (მაქსიმუმ 0-დან შეყვანას შორის), რომელთა გამოთვლა ბევრად უფრო სწრაფია რთული ფუნქციებთან შედარებით. გარდა ამისა, RELU ხელს უწყობს გრადიენტის გაქრობის პრობლემის შერბილებას, რაც უზრუნველყოფს ტრენინგის დროს უფრო სწრაფ კონვერგენციას.

CNN

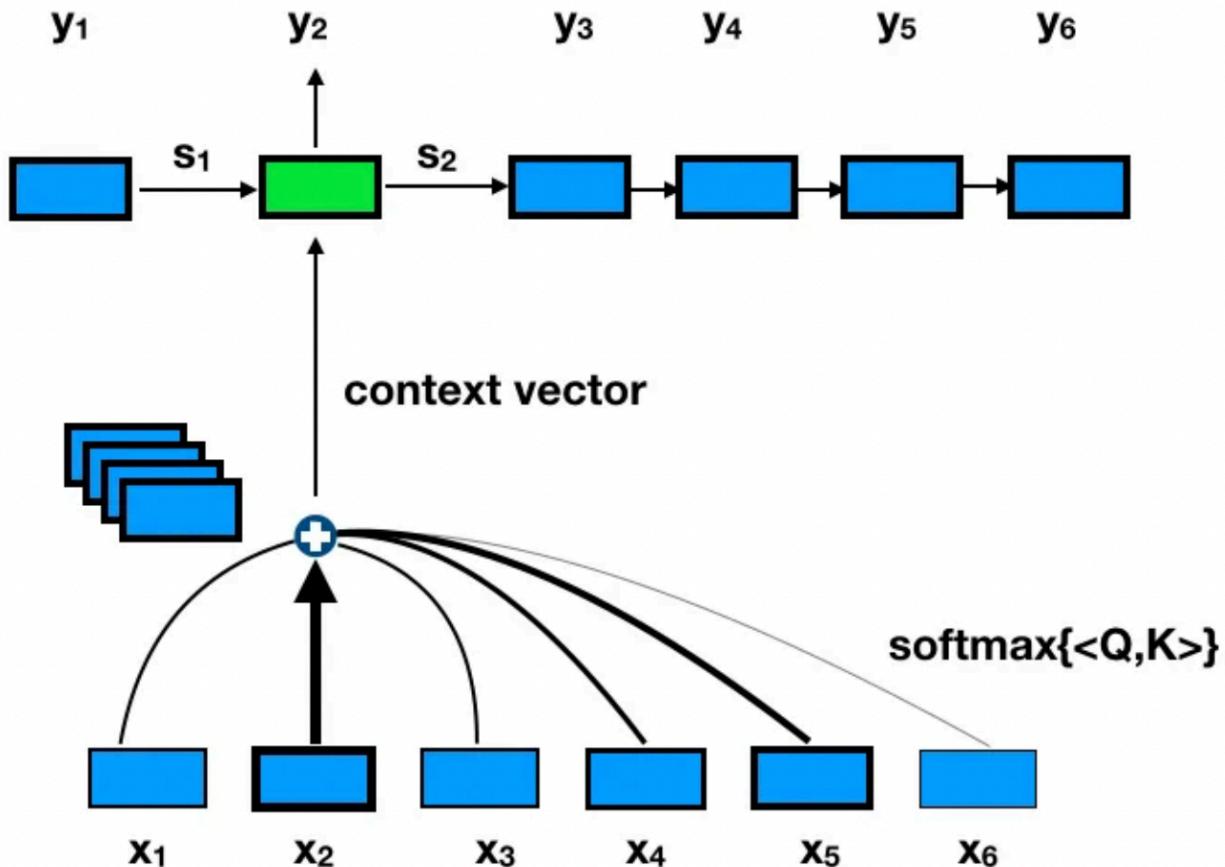
კონვოლუციური ნერვული ქსელი (CNN) არის ღრმა ნეირონული ქსელის ტიპი, რომელიც სპეციალურად შექმნილია სტრუქტურირებული ქსელის მსგავსი მონაცემების დასამუშავებლად, როგორიცაა სურათები ან აუდიო. CNN-ები ფართოდ გამოიყენება



ამოცანებში, როგორიცაა გამოსახულების ამოცნობა, ობიექტების ამოცნობა და გამოსახულების სეგმენტაცია.

კონვოლუციური ფენები: CNN-ების ძირითადი სამშენებლო ბლოკები არის კონვოლუციური შრეები. ეს ფენები იყენებენ ფილტრების ერთობლიობას (ასევე ცნობილია როგორც ბირთვები) საწყის/შემავალ მონაცემებზე. თითოეული ფილტრი სურათიდან იღებს სხვადასხვა მახასიათებლ პარამეტრს, როგორიცაა კიდეები, ტექსტურები ან შაბლონები. ამ ინფორმაციების შეერთება და გაერთიანება ხდება კონვოლუციურ ფენებს შორის. ასეთი მიდგომა ამცირებს ზომაში შესავალს და აჩქარებს დასწავლის პროცესს.

არანრფივი აქტივაციის ფუნქციები, როგორიცაა RELU გამოიყენება კონვოლუციური და გაერთიანებული ფენების გამოსავალზე, რათა შემოიტანოს არანრფივობა ქსელში. რამდენიმე კონვოლუციური და გაერთიანებული ფენის შემდეგ, CNN, როგორც წესი, მთავრდება ერთი ან მეტი ffn-ით. ეს ფენები ასწორებენ მახასიათებლის **რუკებს** ერთგანზომილებიან ვექტორად და აკავშირებს ერთ ფენაში არსებულ ყველა ნეირონს მომდევნო ფენის ყველა ნეირონთან, რაც საშუალებას აძლევს მაღალი დონის მახასიათებლების წარმოდგენას და კლასიფიკაციას.



RNN

რეკურენტული ნერვული ქსელები (RNN) არის ხელოვნური ნერვული ქსელების კლასი, რომელიც შექმნილია თანმიმდევრული მონაცემების ეფექტურად დასამუშავებლად, როგორიცაა დროის სერიები, ტექსტი და აუდიო.

RNN მთავარი შინაარსი მდგომარეობს ინფორმაციის შენახვაში რადგან არის განმეორებადი კავშირები ბლოკებს შორის.

ყოველი დროის საფეხურზე, ქსელი იღებს შესავალს და აერთიანებს მას წინა დამალულ მდგომარეობასთან ახალი ფარული მდგომარეობის შესაქმნელად. ეს განმეორებადი სტრუქტურა საშუალებას აძლევს RNN-ებს აღბეჭდონ დამოკიდებულებები დროის ეტაპებზე.

ტრადიციული RNN-ებს აქვთ გრადიენტის პრობლემა. რაც აფერხებს გრძელვადიანი დამოკიდებულების შესწავლის პროცესს. ამ პრობლემის გადასაჭრელად შემუშავებულია ისეთი ვარიანტები, როგორიცაა გრძელვადიანი მოკლევადიანი მეხსიერება (LSTM) და Gated Recurrent Unit (GRU), რომლებიც აერთიანებს მექანიზმებს გრადიენტების ნაკადის უკეთ შესანარჩუნებლად და კონტროლისთვის.

RNN ფართოდ გამოიყენებოდა ბუნებრივი ენის დამუშავების ამოცანებში, როგორიცაა ენის მოდელირება, მანქანური თარგმანი, განწყობის ანალიზი და ტექსტის გენერირება. თუმცა დღესდღეობით ტრანსფორმერების არქიტექტურამ ჩაანაცვლა ეს სტრუქტურა.

Attention

ყურადღების მექანიზმები თანამედროვე ნეირონული ქსელის არქიტექტურის გადამწყვეტი კომპონენტია, განსაკუთრებით ბუნებრივი ენის დამუშავებისა და კომპიუტერული ხედვის ამოცანებში.

ყურადღების მექანიზმები საშუალებას აძლევს ნერვულ ქსელს კონცენტრირდეს კონკრეტულ ინფორმაციაზე და მათ შორის იპოვოს კანონზომიერება.

