PHY654

Machine learning (ML) in particle physics

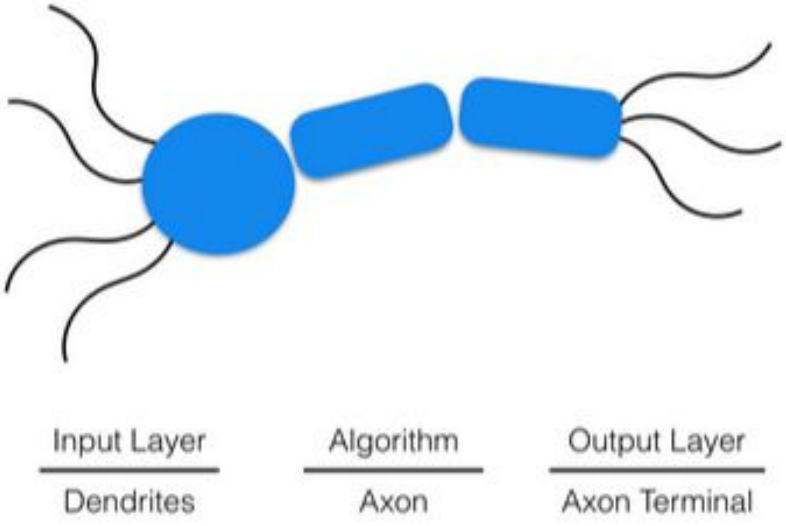
Swagata Mukherjee • IIT Kanpur

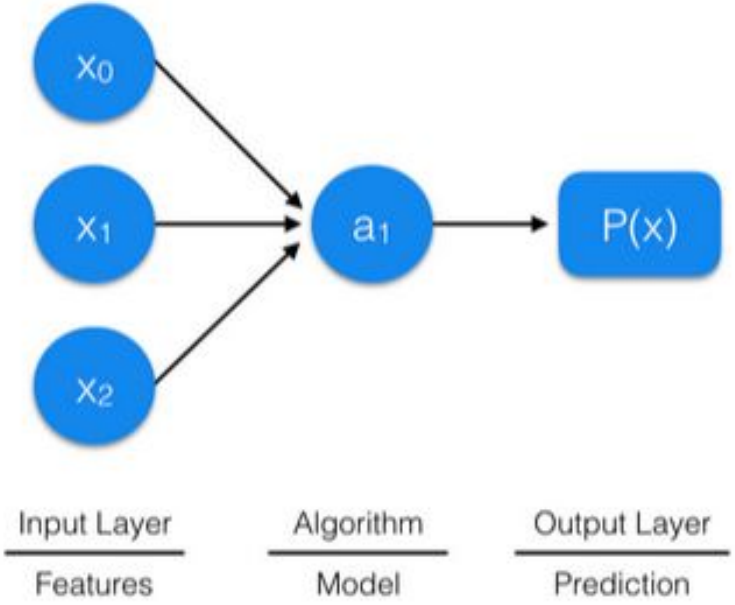
19th August 2024

Today’s hands-on exercises

1) https://github.com/swagata87/IITKanpurPhy654/blob/main/NN_python_output_complexity.ipynb
2) https://github.com/swagata87/IITKanpurPhy654/blob/main/Gradient_descent_simpleFit.ipynb

Human brain and neural network: is there a connection?

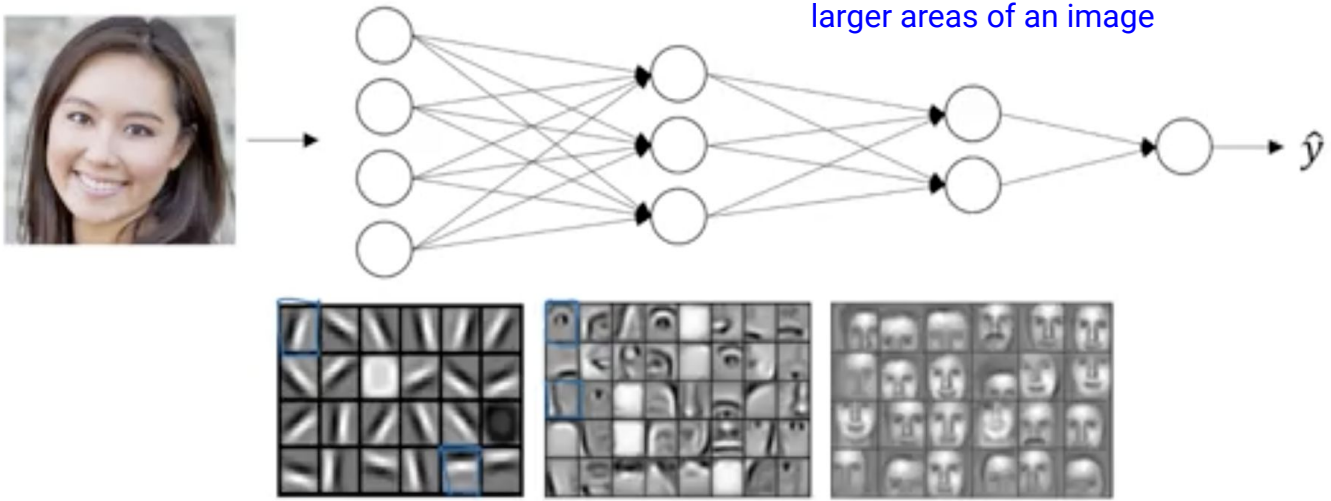




Why deep?

Earlier layers learning simple features, later layers learn more complex features.

Earlier layers focus on small areas of an image. Later layers focus on larger areas of an image



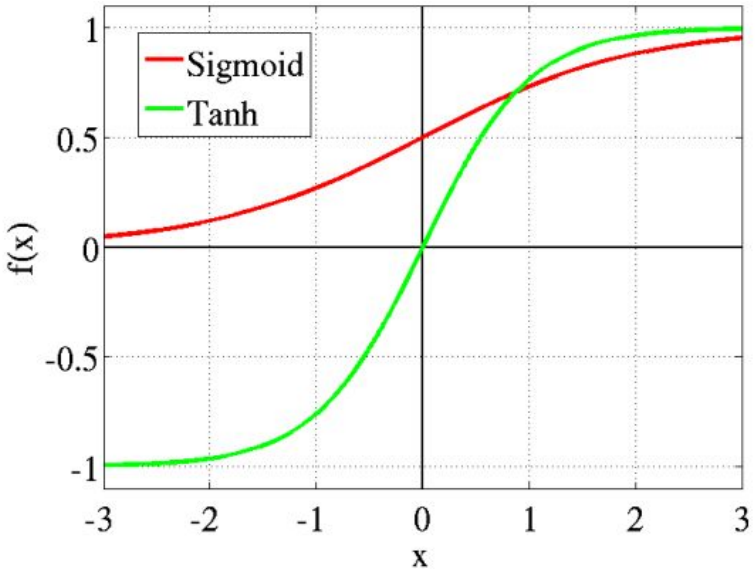
1st layer may be detecting edges

2nd layer may be detecting parts of faces (eye, nose etc)

3rd layer may be detecting different types of faces

Activation function

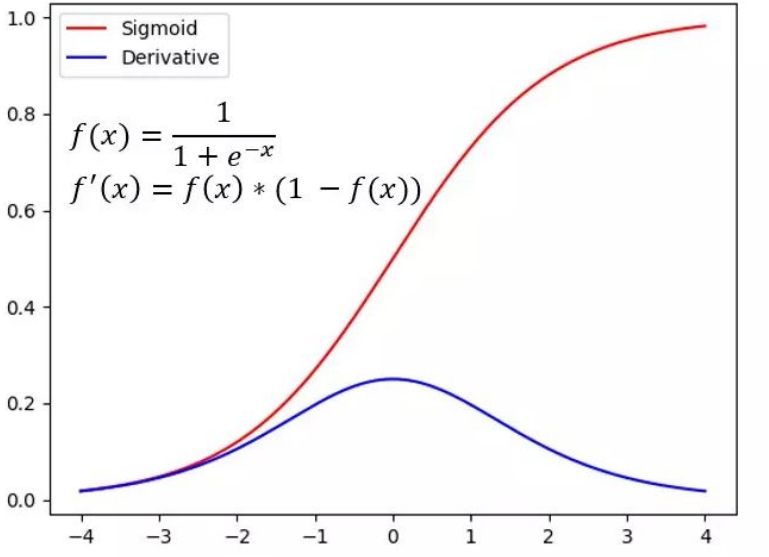
We have been using **sigmoid** so far. But there are other options too. Example **tanh** (sometimes works better than sigmoid in hidden layers).



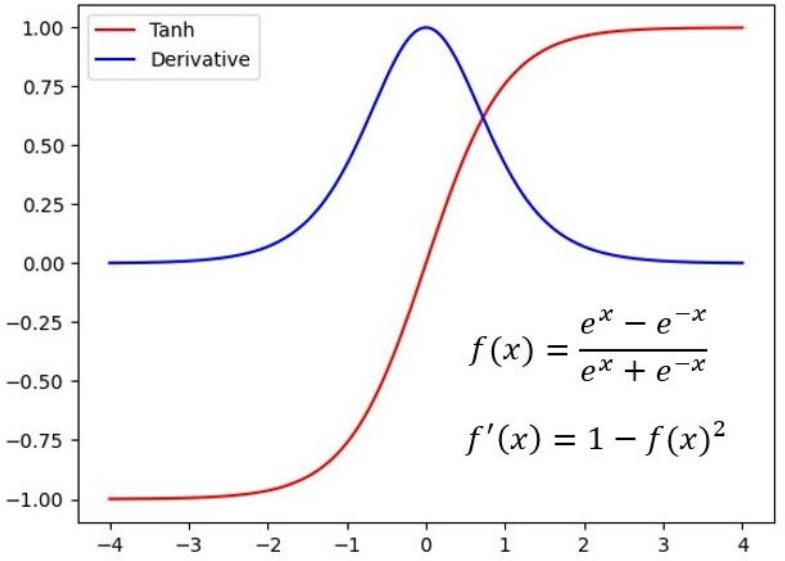
Sigmoid and tanh have issues for values too large or too small, because slope becomes close to zero and gradient descent is very slow.

Vanishing gradient problem.

Activation function



Sigmoid function (red) and first derivative(blue)



Tangent hyperbolic function(red) and first derivative (blue)

Activation function

ReLU or **Leaky ReLU** can be used to avoid vanishing gradient issue