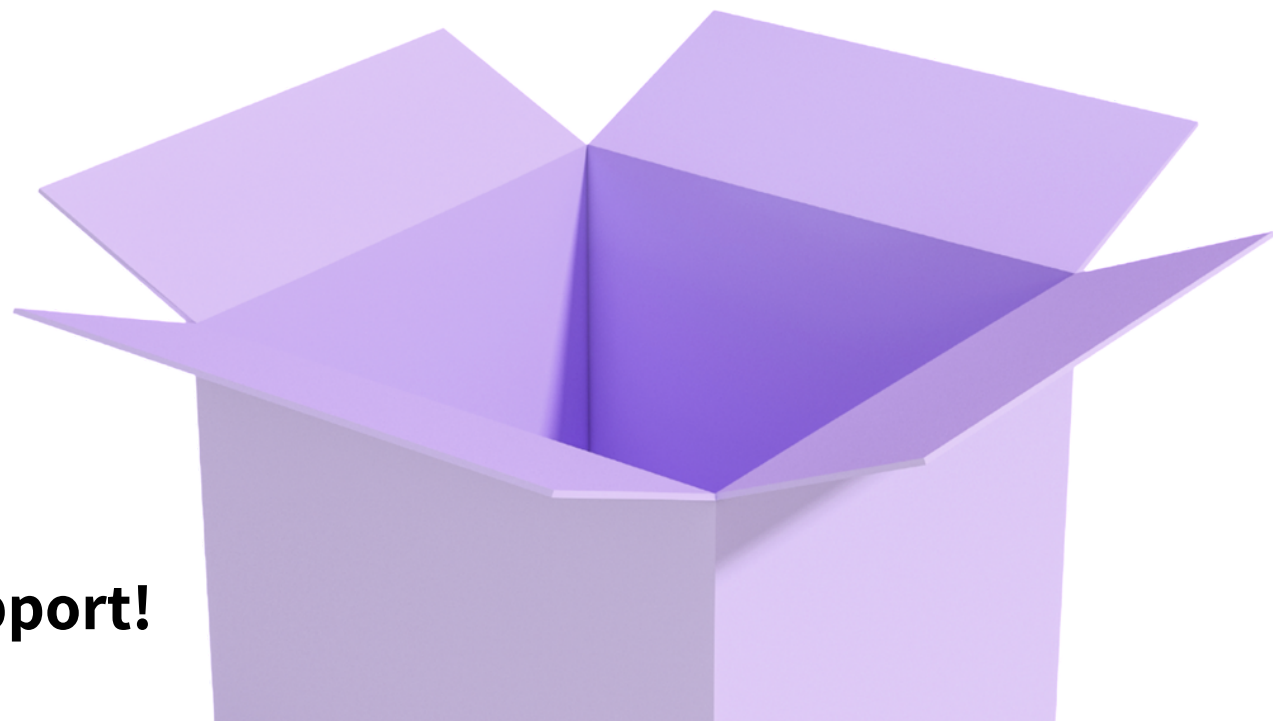
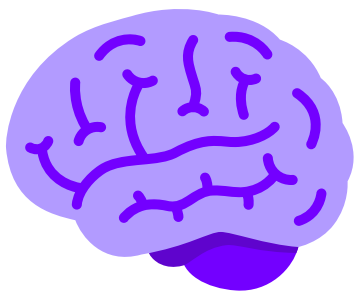
Data-Driven Science X MACHINE LEARNING INDIA

NON-LINEAR ACTIVATION FUNCTIONS

For Deep Learning!

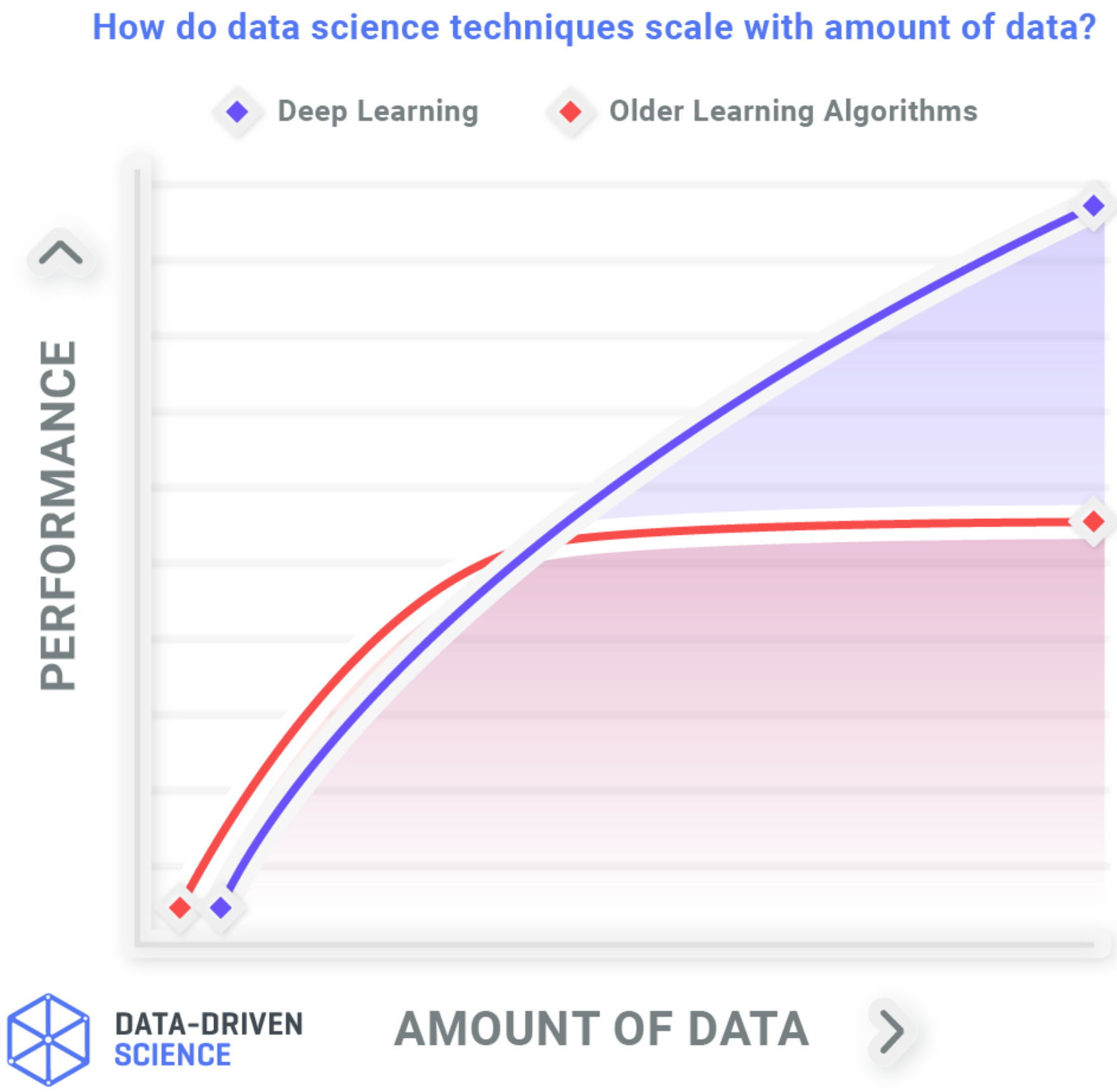




But, Why deep learning!?

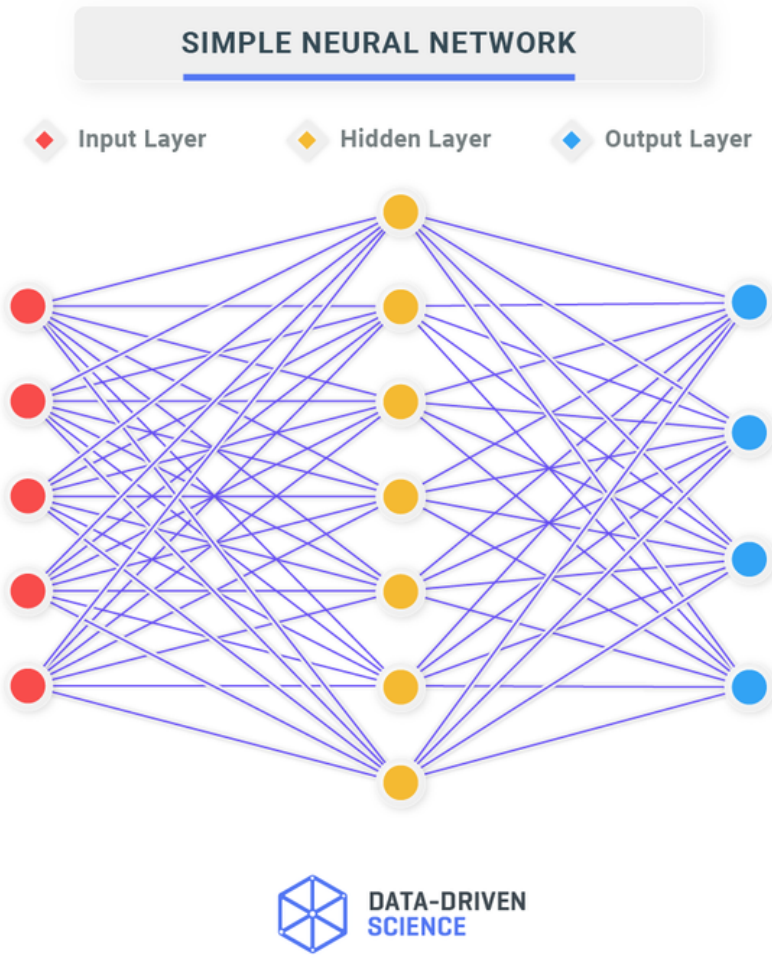
1. In traditional **Machine Learning** most features need to be identified by a domain expert, who may or may not be available.
2. On the contrary, in **Deep Learning**, algorithms learn features from the data itself in an incremental manner. This eliminates manual feature extraction to a great extent but requires **MUCH MORE** data.

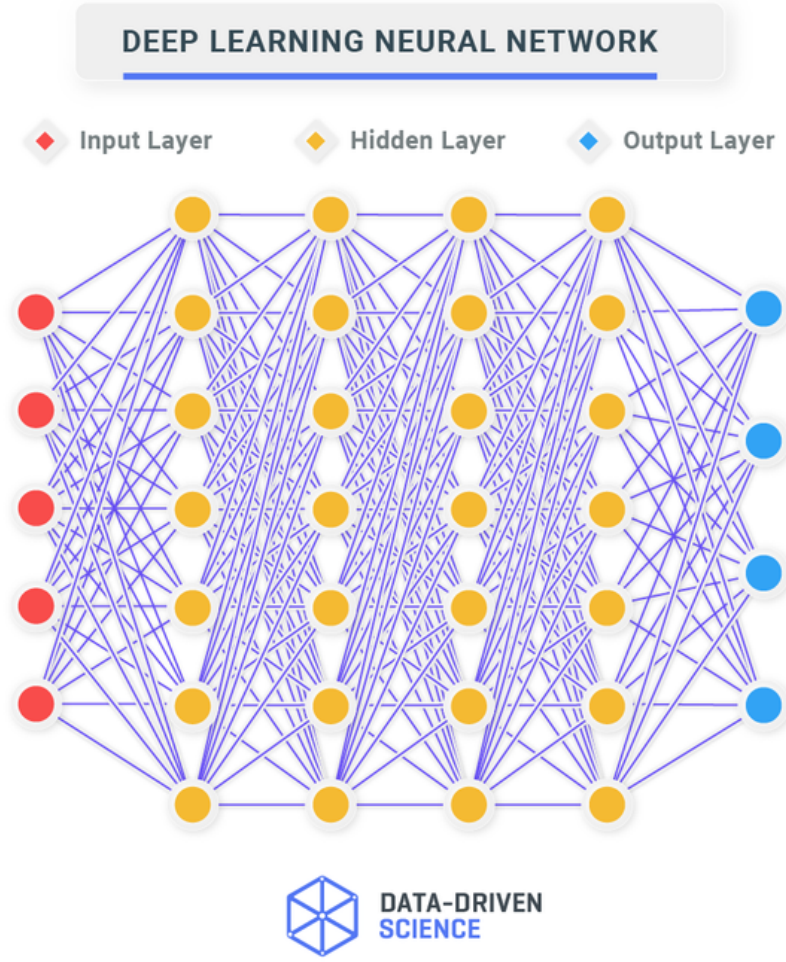
Another important factor:



Neural Networks: Simple vs Deep

Neural networks with some level of complexity, usually at least two layers, qualify as a **Deep Neural Network. Activation Functions** are a key component of **DNNs** because they determine the output of a deep learning model, its accuracy, and its computational efficiency.





How they operate:

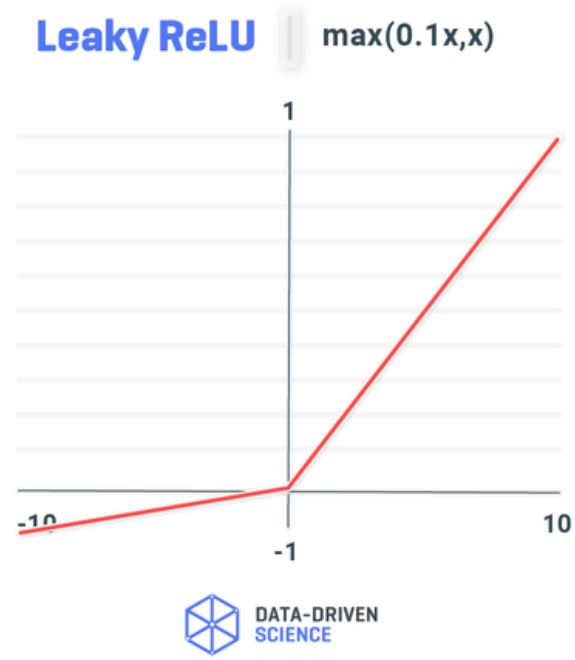
In a **Neural Network**, numeric data points, called inputs, are fed into the neurons in the input layer. Each neuron has a weight and multiplying the input number with the weight gives the output of the neuron, which is transferred to the next layer.

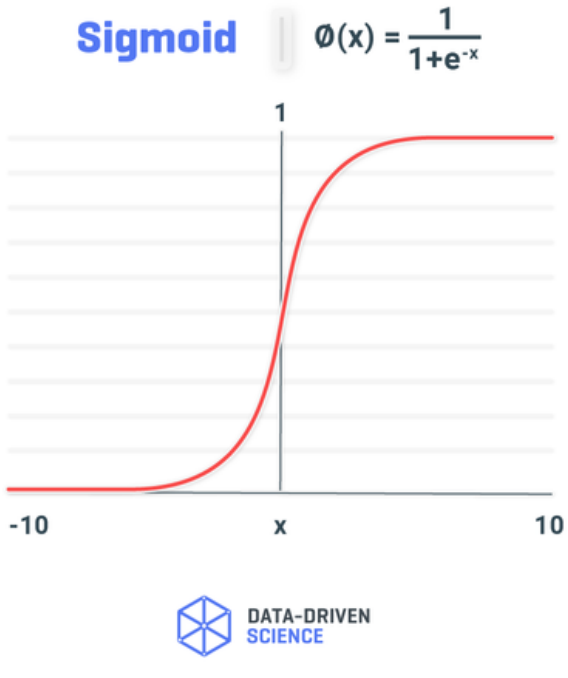
What are activation functions?

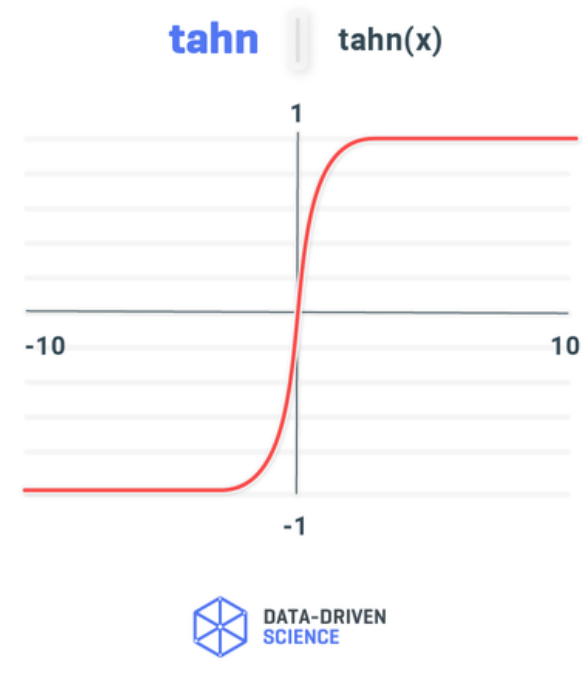
They are like a mathematical "**gate**" in between the input feeding the current neuron and its output going to the next layer.

It can be as simple as a **step function** that turns the neuron output on and off, depending on a rule or threshold. Or it can be a **transformation** that maps the input signals into output signals that are needed for the neural network to function.

Non-Linear Activation Functions can help the network learn complex data, compute and learn almost any function representing a question, and provide accurate predictions. Some are:



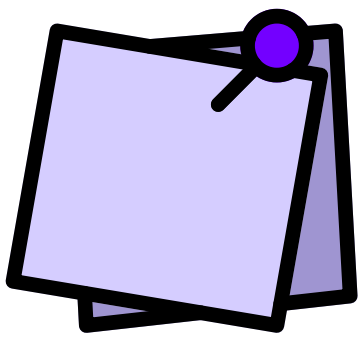




Some cool tips!

Which to use when?

- Sigmoid functions and their combinations generally work better in the case of classifiers.
- Sigmoids and tanh functions are sometimes avoided due to the vanishing gradient problem.
- ReLU function is a general activation function and is used in most cases these days.
- Encounter a case of dead neurons in your networks? Leaky ReLU is the best choice.
- ReLU function should only be used in the hidden layers.
- As a rule of thumb, you can begin with using ReLU function and then move over to other activation functions in case ReLU doesn't provide optimum results.



Notable references:

Fundamentals of DL – Activation Functions and When to Use Them? by **Analytics Vidhya**.

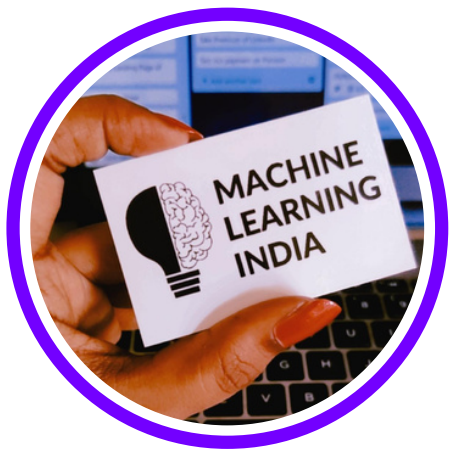
Important note:

The links to additional resources will be put up on our Telegram. Channel ID: **@machinelearning24x7**.

We have a newsletter!

Subscribe to our newsletter to receive latest updates on ML/AI research. **Link: in bio**.

Data-Driven Science X MACHINE LEARNING INDIA



@ml.india

@ml_india_

bit.ly/mli-linkedin

@machinelearning24x7

Shout out to Data Driven Science!

For making this post possible! Have you checked their awesome **Hands-On Data Science** course on **Udemy** out? Rated **4.2/5**, it's fantastic!

Check it out here: course.datadrivenscience.com/ml

Like. Comment. Share.