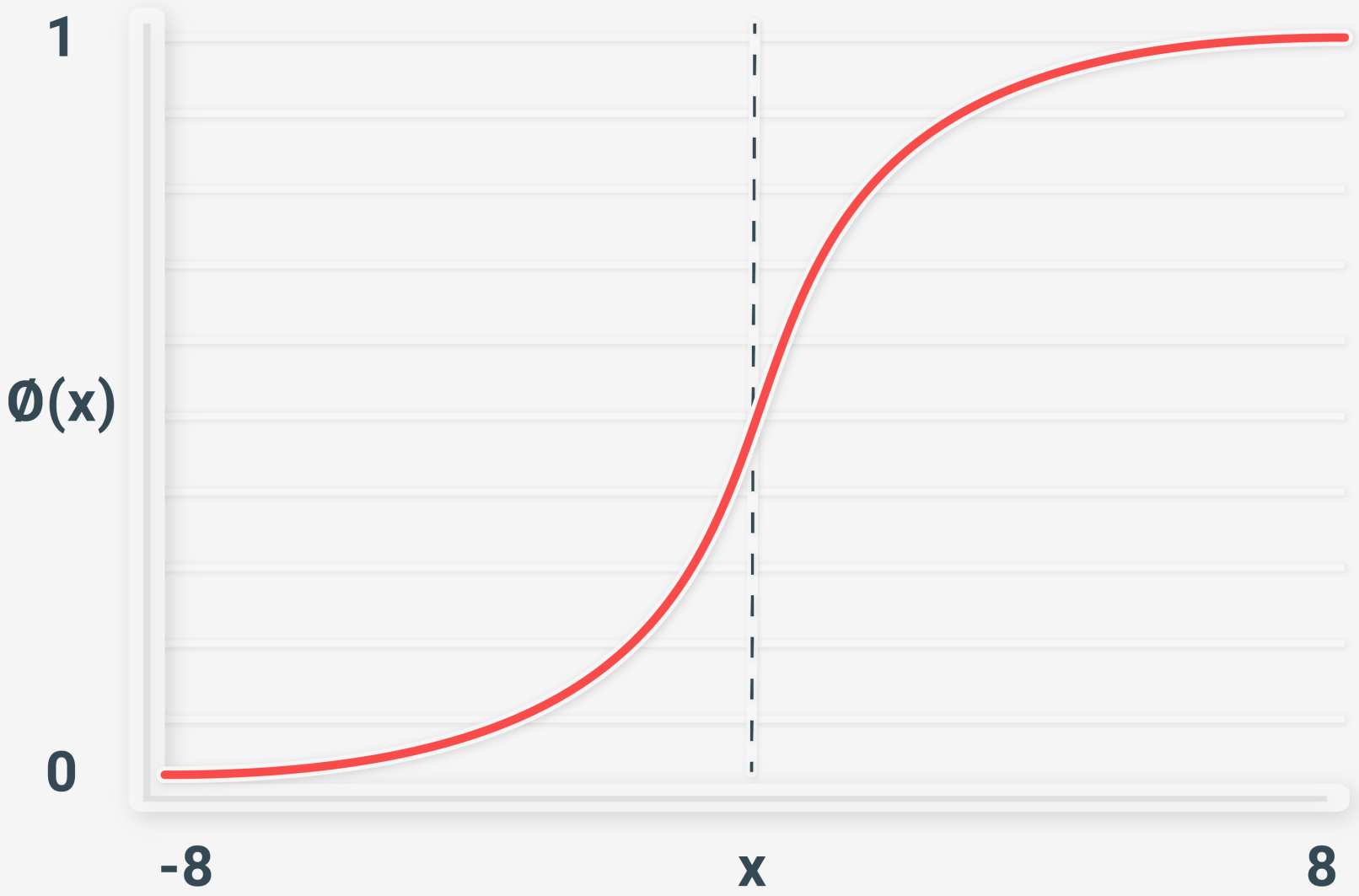
SIGMOID ACTIVATION FUNCTION

Often used in the output unit for binary classification

$$\emptyset(x) = \frac{1}{1+e^{-x}}$$

Returns a value between 0 and 1



DATA-DRIVEN
SCIENCE

Advantages

Smooth Curve

preventing “jumps” in output values

Output Values Bound b/w 0 and 1

normalizing the output of each neuron

Differentiable Function

Find the slope of the sigmoid curve at any two points

Function Monotonic, But

the function's derivative is not

DATA-DRIVEN
SCIENCE

Disadvantages

Neural Network Can Get Stuck

at the training time

Outputs Not Zero Centered

Vanishing Gradient

for very high or very low values of X, there is almost no change to the prediction

---

Computationally Expensive

DATA-DRIVEN
SCIENCE

Most-used

For models where we have to predict the probability as an output

Since probability of anything exists only between the range of 0 and 1, sigmoid is the right choice

Get daily updates on
Data Science concepts

FOLLOW NOW

DATA-DRIVEN
SCIENCE