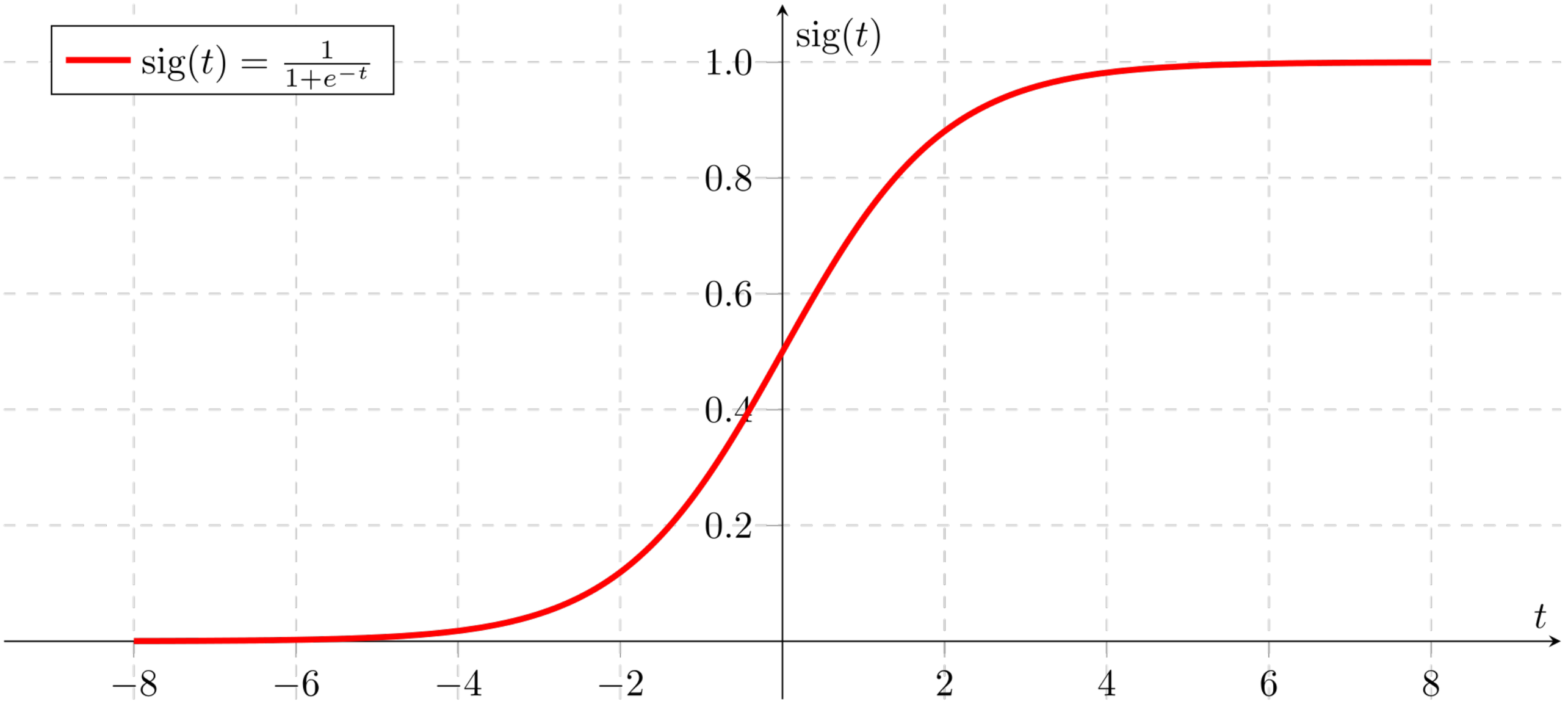
LOGISTIC REGRESSION

WHAT IS LOGISTIC REGRESSION?

- Logistic Regression is a standard classification machine learning algorithm.
- Like Linear Regression, Logistic Regression finds weights for each column of data.
- Unlike Linear Regression, Logistic Regression splits the data in half.
- Logistic Regression uses the sigmoid function to convert all outputs to 0 or 1.
- Logistic Regression is a classifier, not a regressor, even though it's called "Logistic Regression."

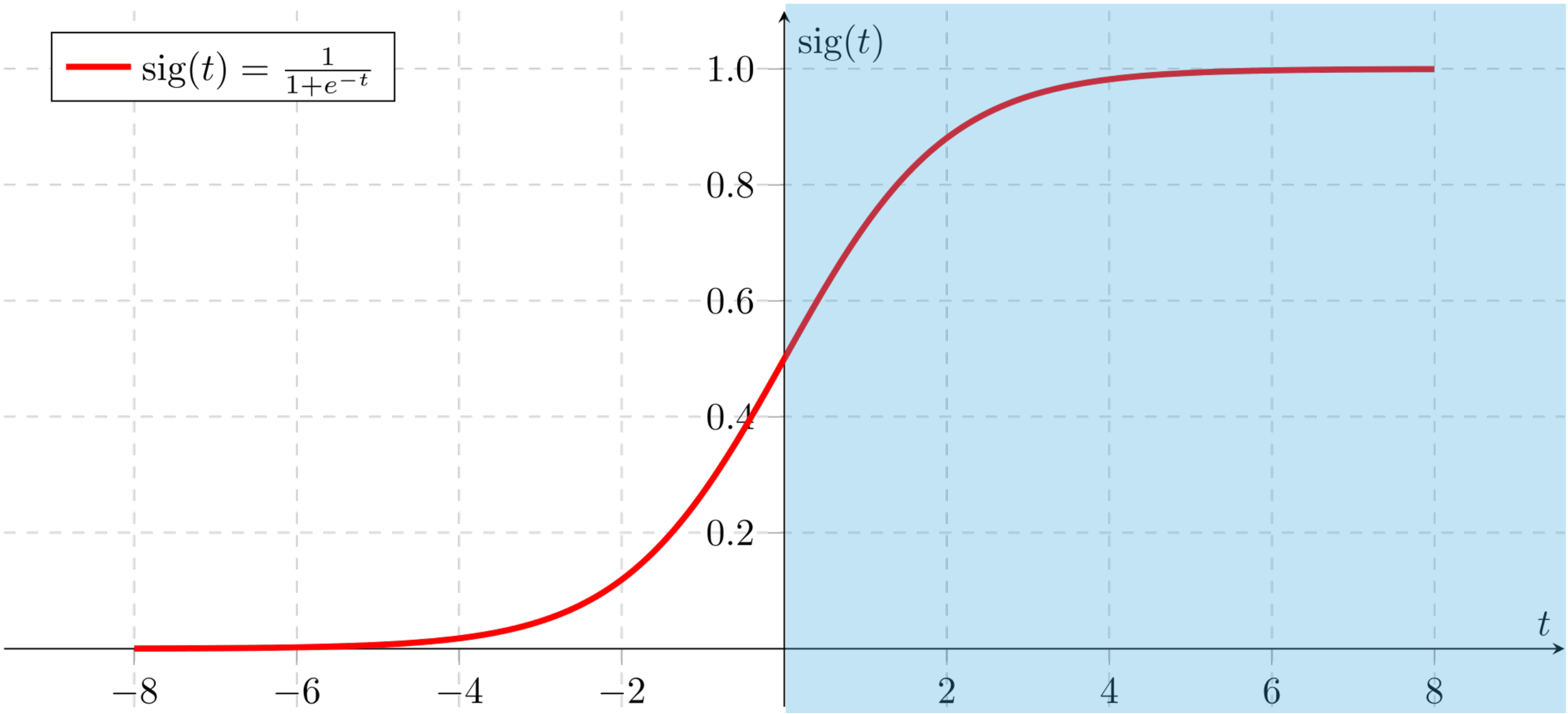
SIGMOID EQUATION

$$\text{sig}(t) = \frac{1}{1+e^{-t}}$$



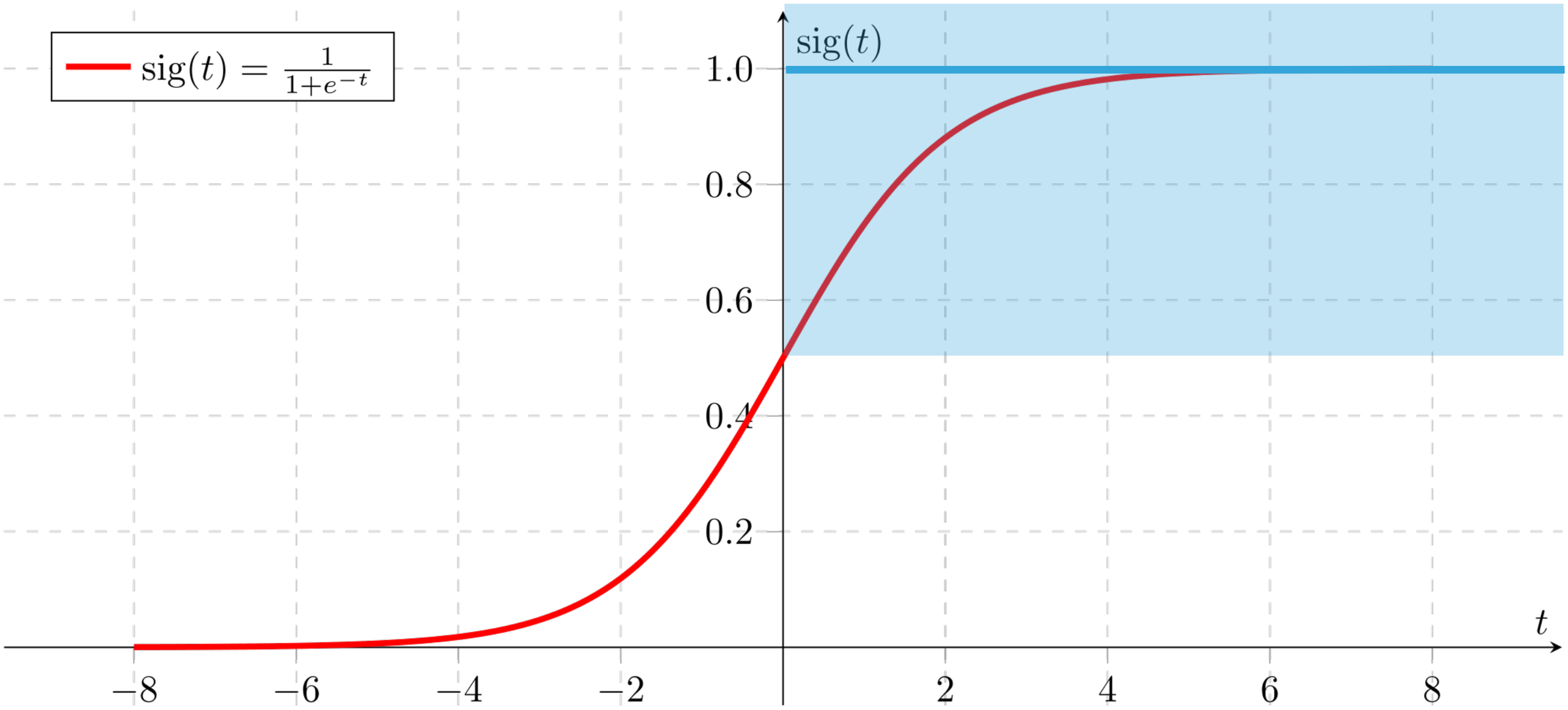
SIGMOID EQUATION

$$\text{sig}(t) = \frac{1}{1+e^{-t}}$$



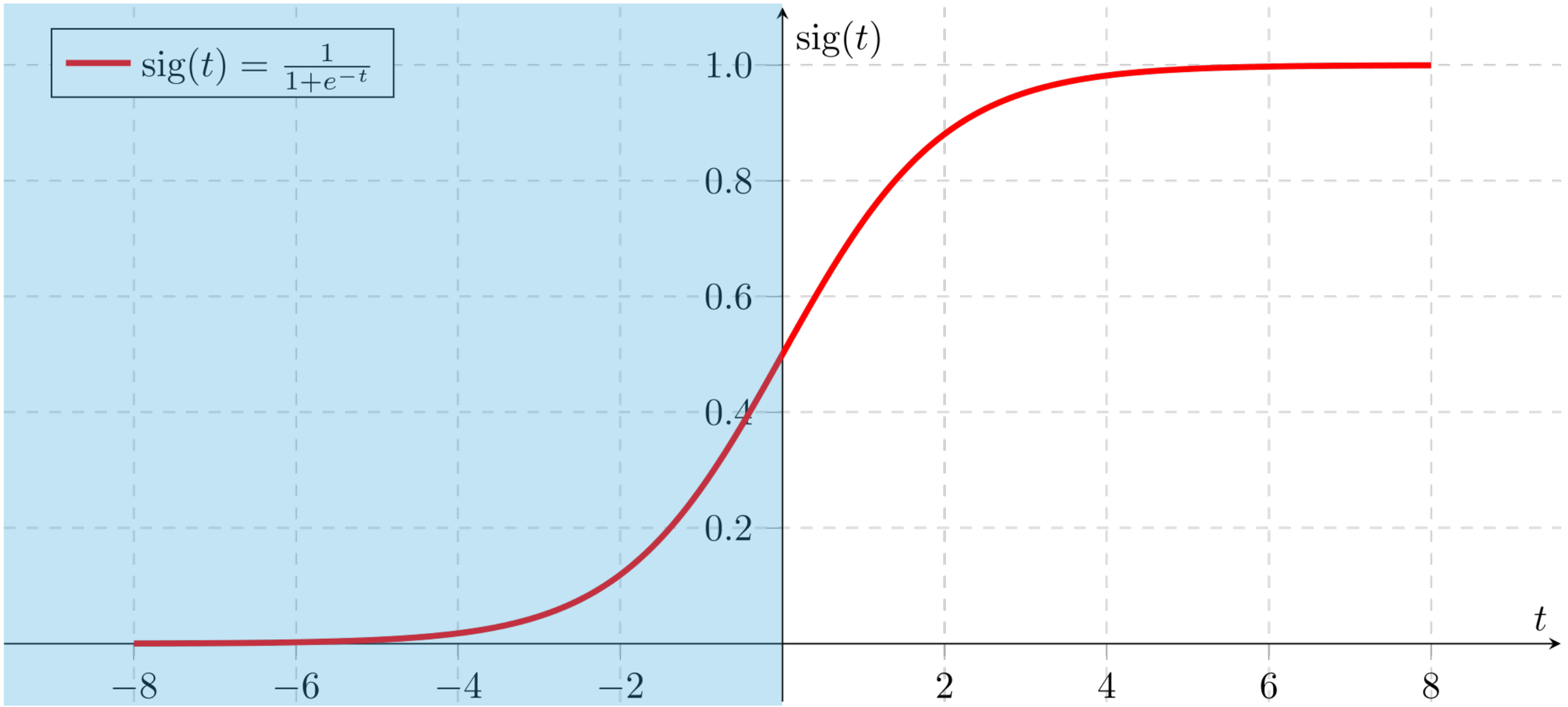
SIGMOID EQUATION

Positive X values are mapped to class 1



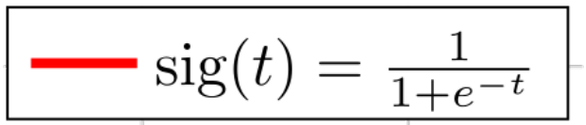
SIGMOID EQUATION

Negative X values



SIGMOID EQUATION

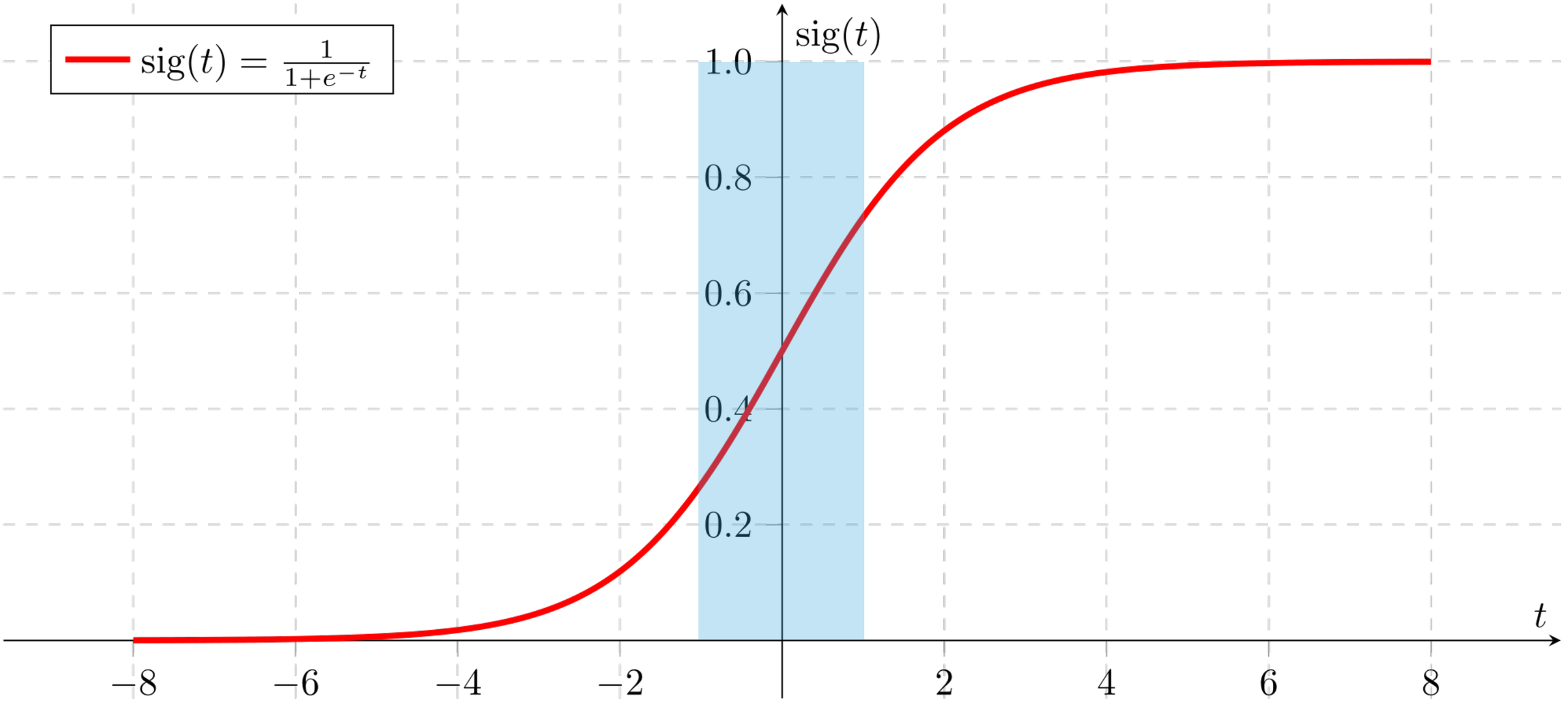
Negative X values are mapped to class 0



SIGMOID EQUATION

$$\text{sig}(t) = \frac{1}{1+e^{-t}}$$

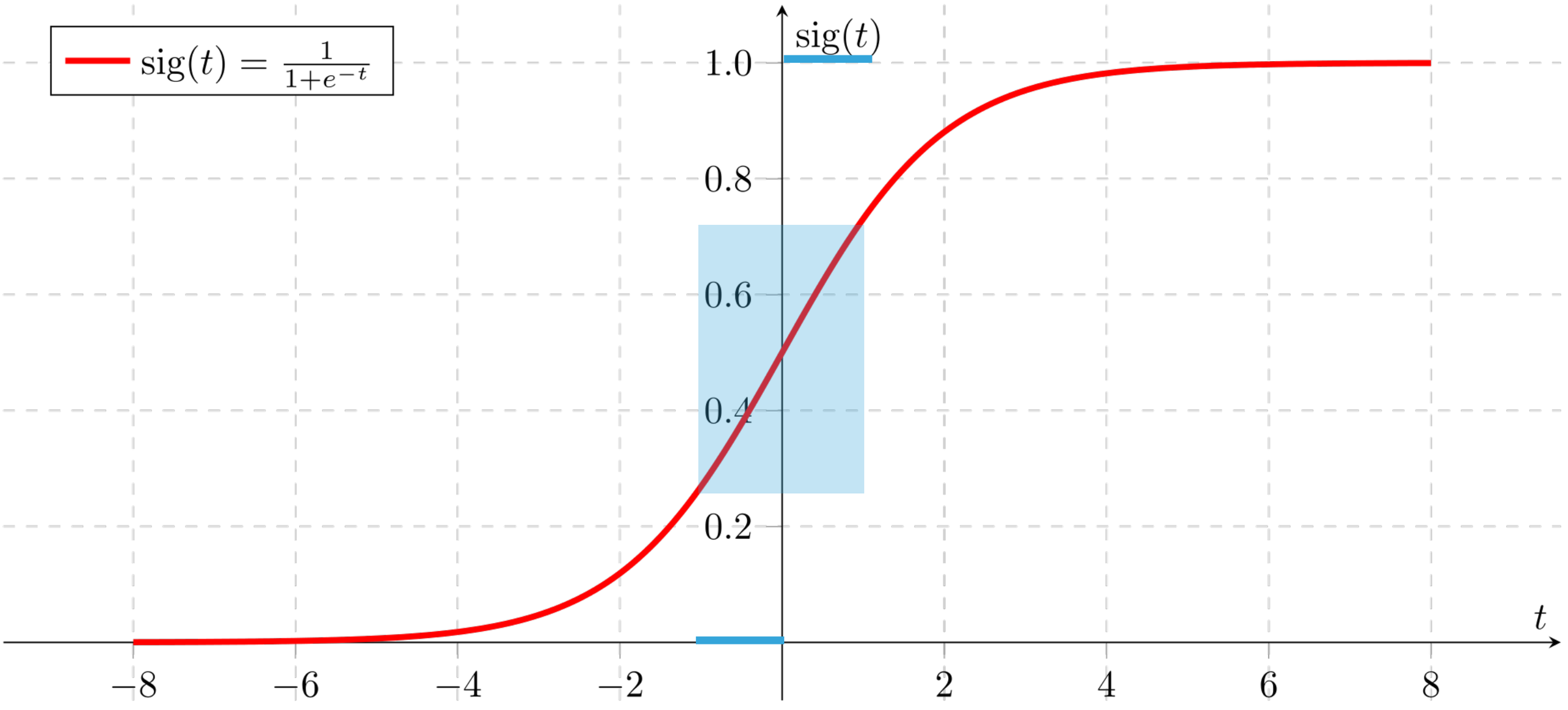
X values in the middle



SIGMOID EQUATION

X values in the middle have the greatest uncertainty

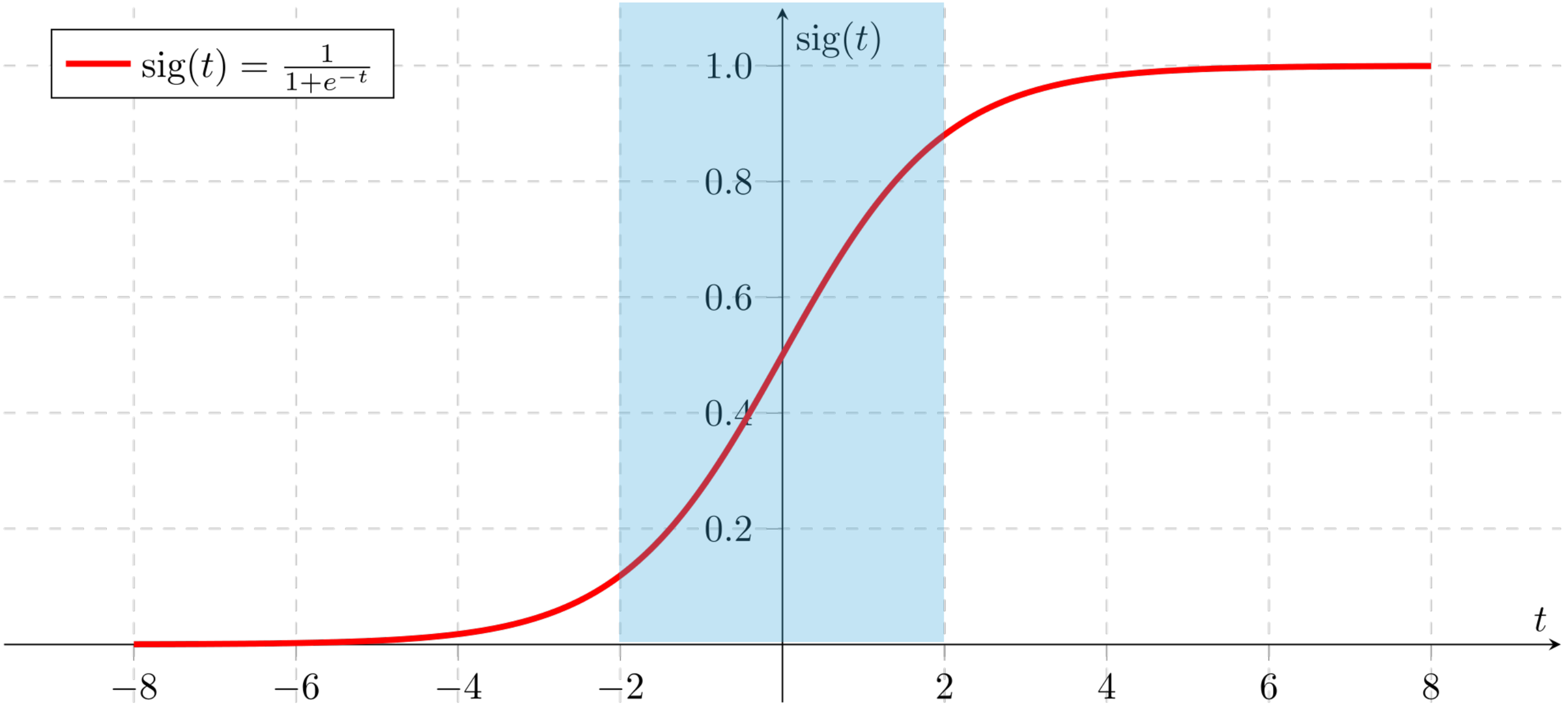
$$\text{sig}(t) = \frac{1}{1+e^{-t}}$$



SIGMOID EQUATION

The Sigmoid ranges from 0 to 1

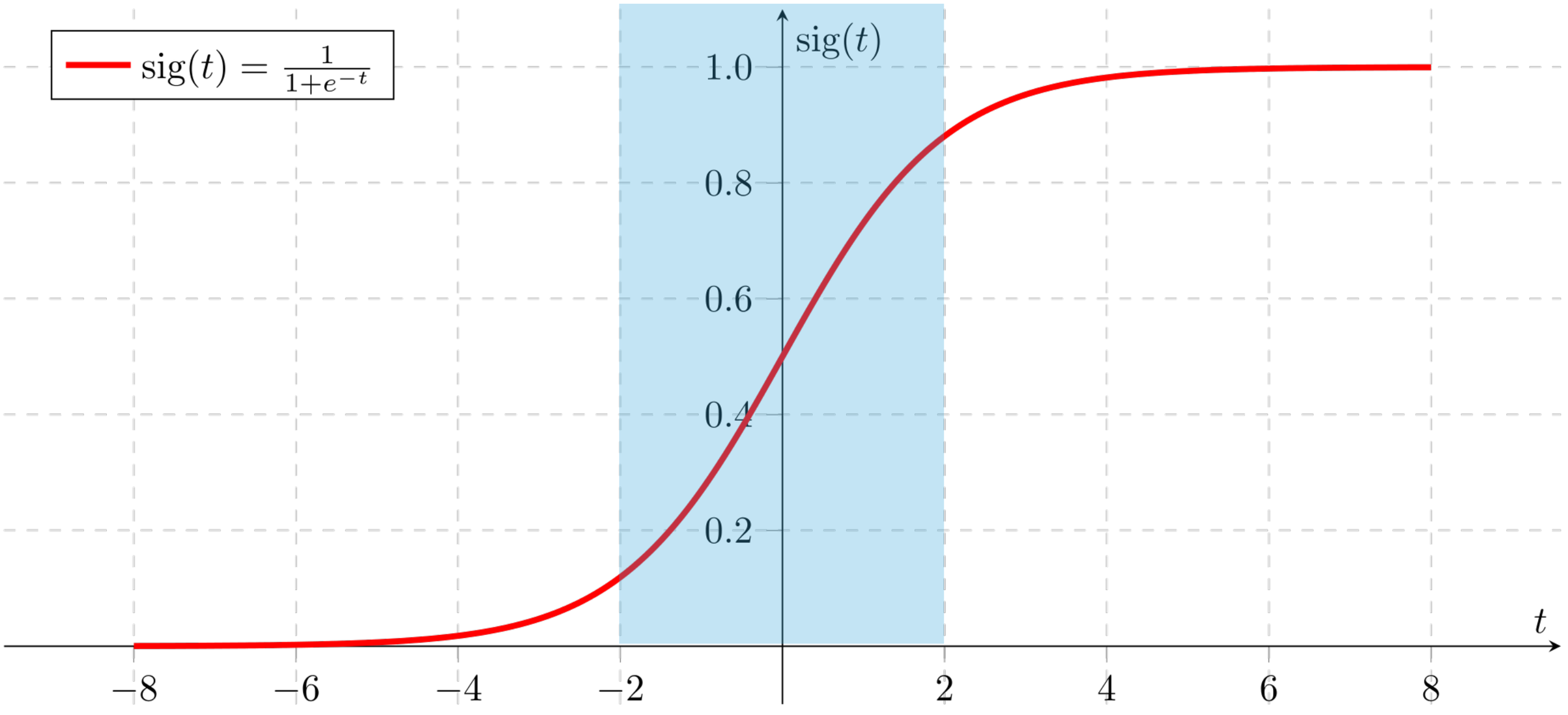
$$\text{sig}(t) = \frac{1}{1+e^{-t}}$$



SIGMOID EQUATION

These values may be interpreted as probabilities

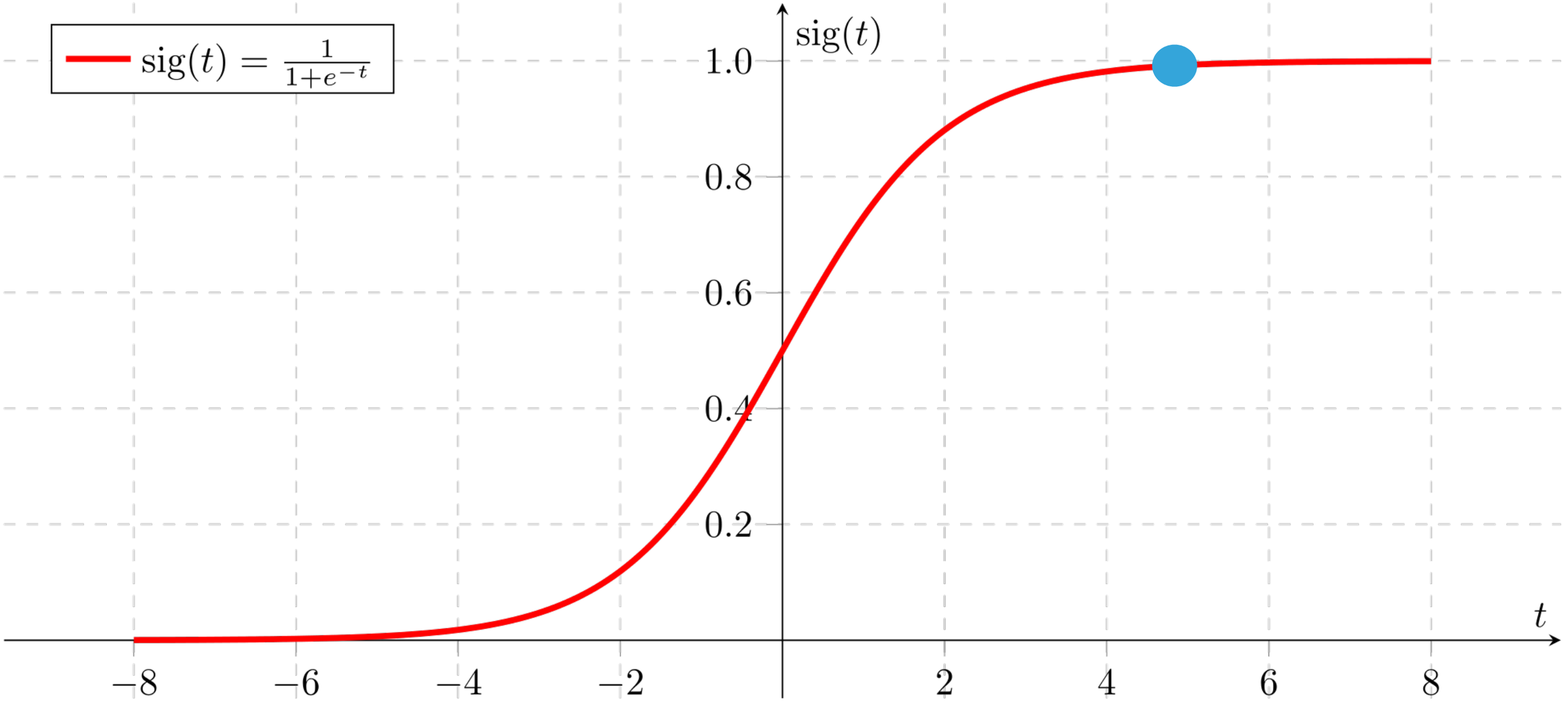
$$\text{sig}(t) = \frac{1}{1+e^{-t}}$$



SIGMOID EQUATION

This point has a y-value of 0.99 so there is a 99% probability it belongs in class 1.

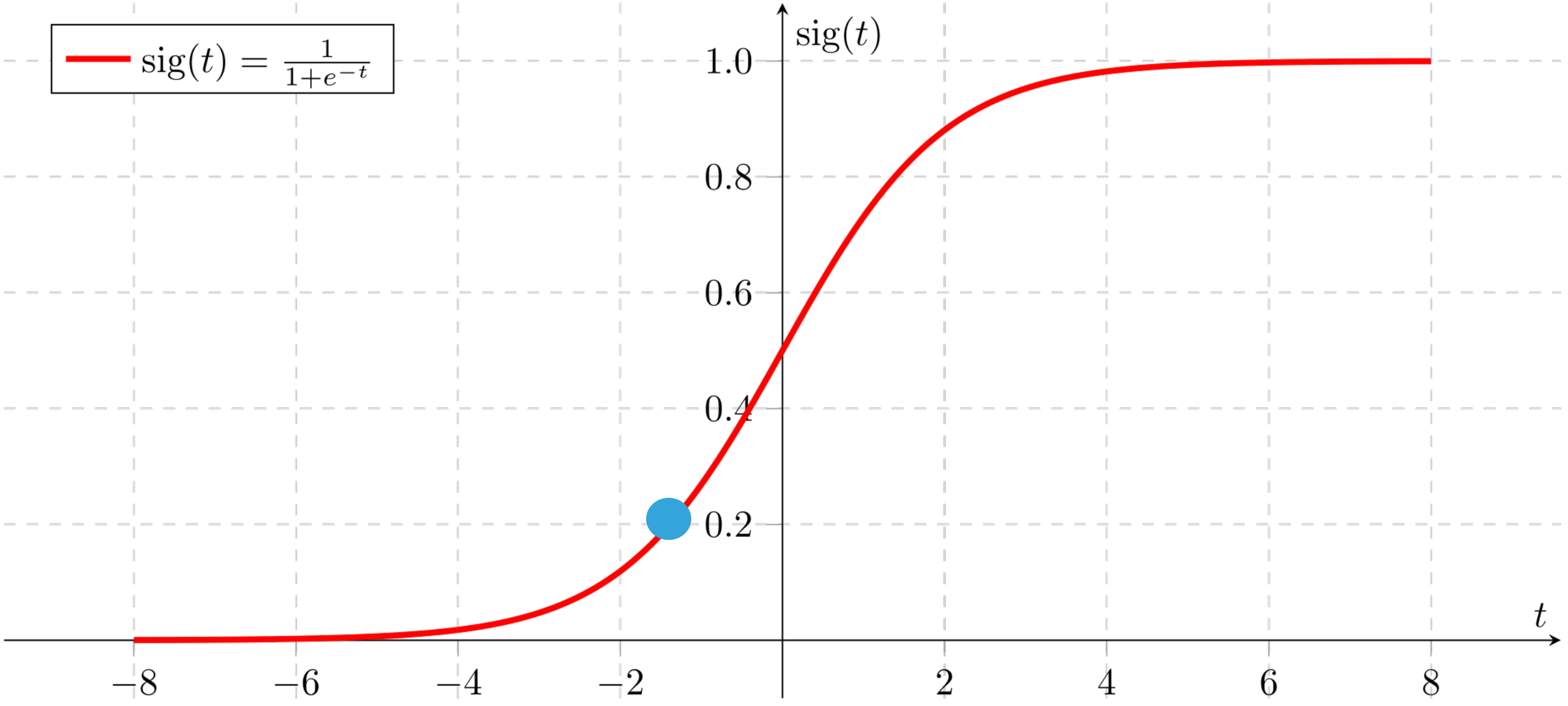
$$\text{sig}(t) = \frac{1}{1+e^{-t}}$$



SIGMOID EQUATION

This point has a y-value of 0.2 so there is a 20% probability it belongs in class 1, and 80% probability it belongs to class 0.

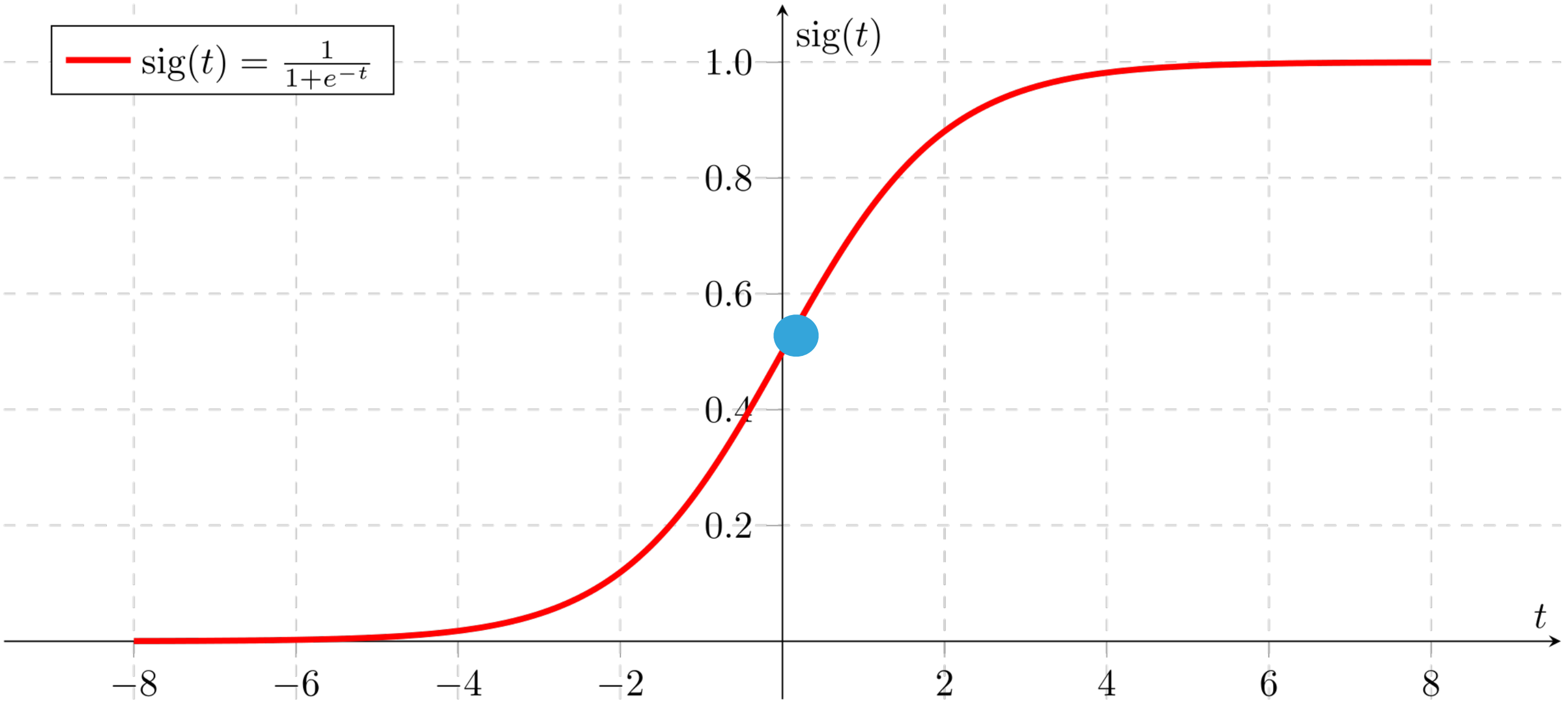
$$\text{sig}(t) = \frac{1}{1+e^{-t}}$$



SIGMOID EQUATION

This point is almost indeterminate, but it will be placed in class 1 since there is a probability of greater than 50% that it belongs in class 1.

$$\text{sig}(t) = \frac{1}{1+e^{-t}}$$



HOW LOGISTIC REGRESSION WORKS

- Weights are randomly chosen to multiply each column.
- The sum of the weighted columns becomes the t-value.
- The t-value is placed in the sigmoid equation and mapped to 0 or 1.
- The percentage of correct predictions is returned.
- Weights are adjusted depending on the error (using gradient descent).
- Another iteration continues.
- The iterations continue until the error is minimized.

HAPPY CODING!