Generative Adversarial Networks (GANs) - 01

Main notions about GANs

DEFINITION

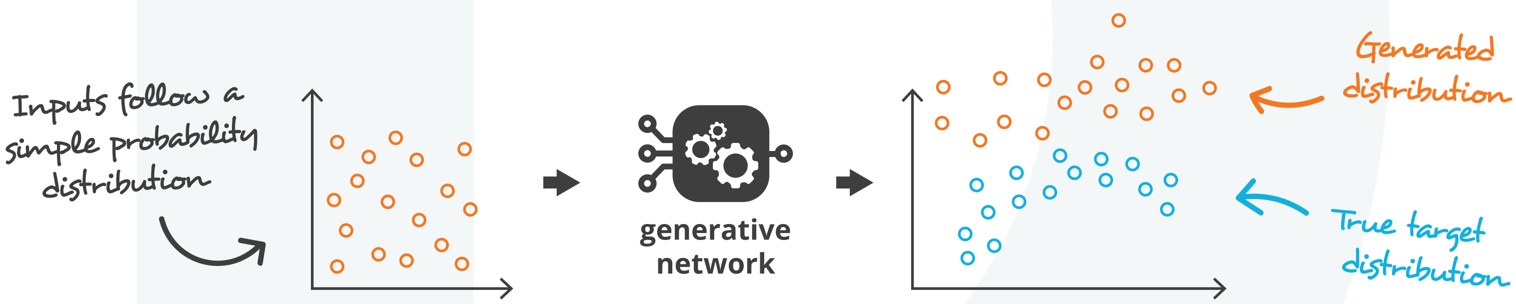
Generative Adversarial Networks (GANs) are deep generative models that can produce new pieces of content. This kind of neural network **can generate, for example, images, texts, or music**. It was introduced in 2014 by Ian J. Goodfellow and his co-authors in the article "Generative Adversarial Nets".

FIRST KEY POINT: THE TARGET PROBABILITY DISTRIBUTION

The first key point lies in the idea that **there exists a probability distribution describing the kind of data we try to generate**. We want to be able to sample new data from that distribution.

For example, the problem of generating a new human face is equivalent to the problem of generating a new data point following the "human faces probability distribution."

To generate data from our target distribution, we use the inverse transform method. Once trained, **the generative network in a GAN takes points following a simple distribution as input and turns them into points following the target distribution.**



SECOND KEY POINT: THE ADVERSARIAL TRAINING

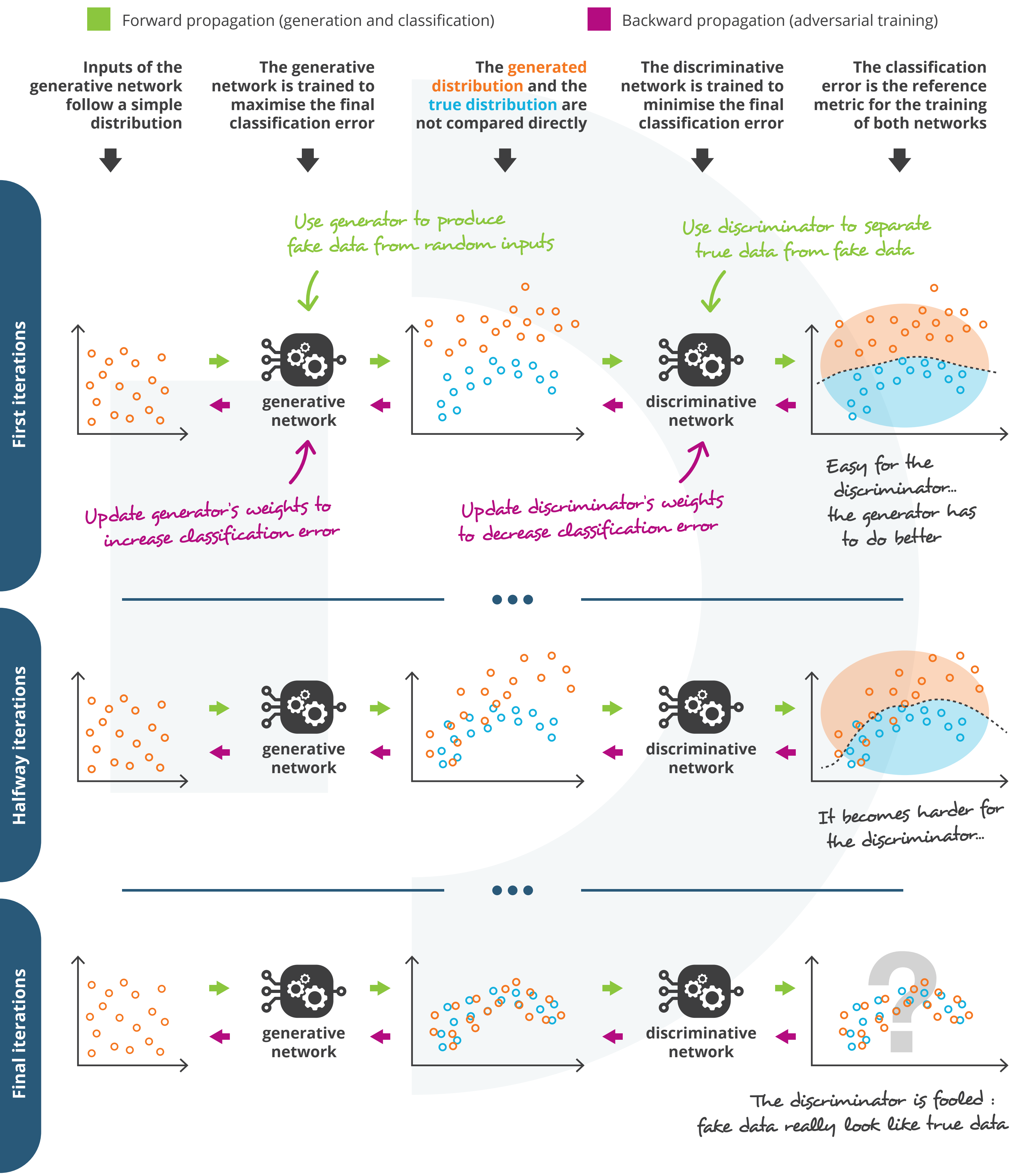
The second key point is the notion of **adversarial training** that defines how the generator learns the function that transforms a simple distribution into the correct target distribution.

When training the generative network, the target and the generated distributions are not directly compared. Instead, **a discriminative network is trained to take true data and generated data and to classify them.**

During training, **the two networks (generator and discriminator) have opposing goals**. The discriminator always wants to classify the data as accurately as it can. The generator always tries to produce fake data that looks like the true data to fool the discriminator.

Generative Adversarial Networks (GANs) - 02

Understanding the concept of adversarial training



More on this subject at:
https://towardsdatascience.com/understanding-gans-cd6e4651a29

Towards Data Science