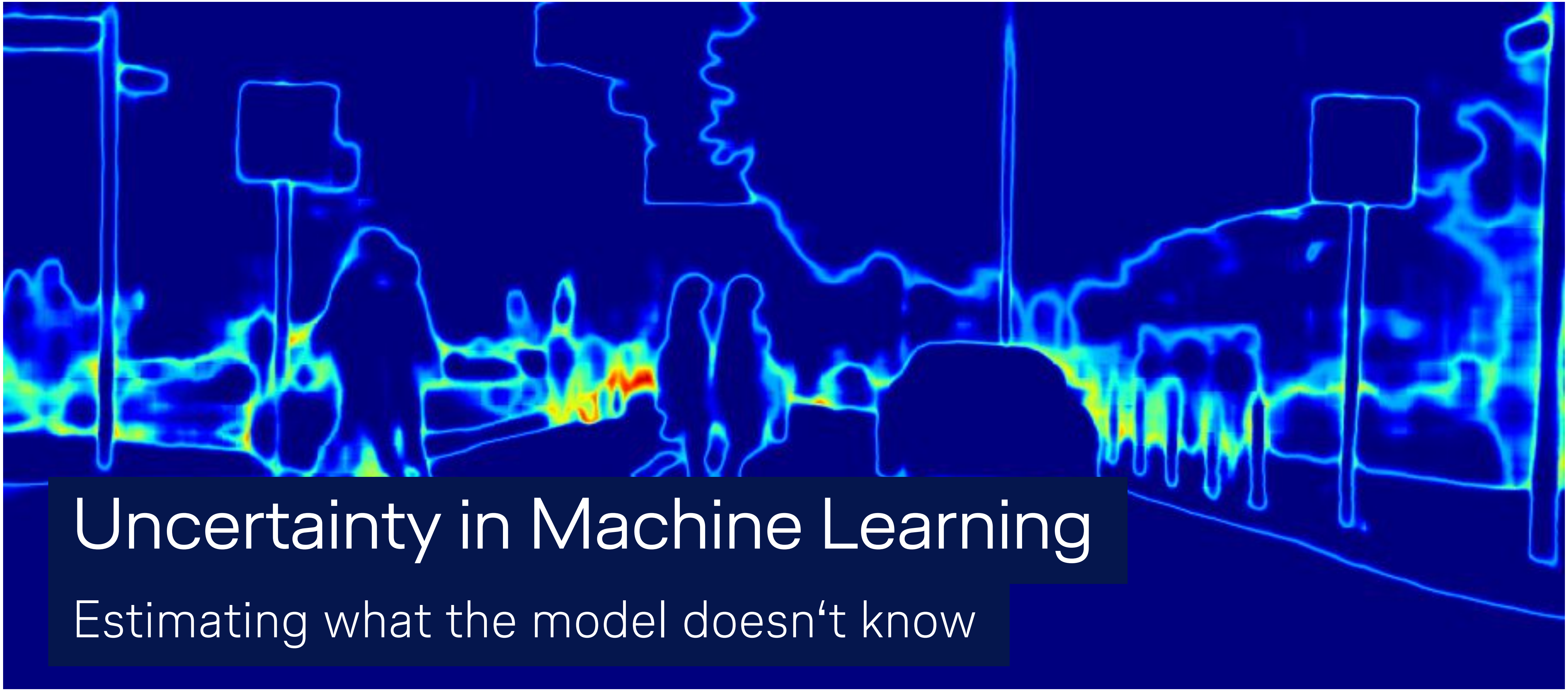


Uncertainty in Machine Learning

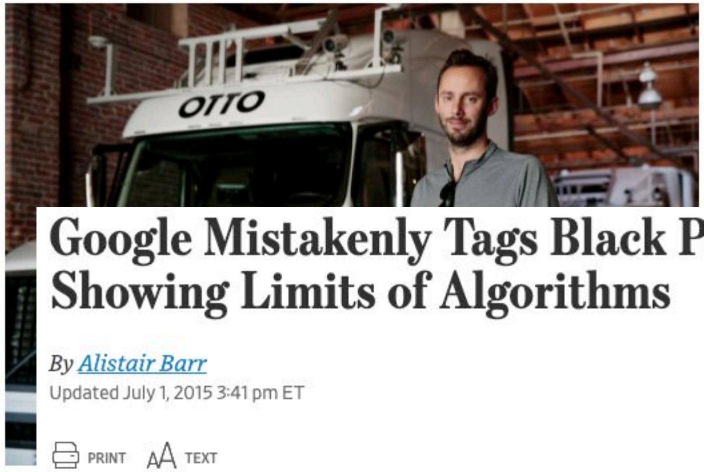
Estimating what the model doesn‘t know

Lufthansa
Industry Solutions

Why do need uncertainty estimates in machine learning?

We can live with models that aren't perfect but we want to know when they fail

A Lawsuit Against Uber Highlights the Rush to Conquer Driverless Cars



Last
Ube
driv

Google Mistakenly Tags Black People as 'Gorillas,' Showing Limits of Algorithms

By Alistair Barr
Updated July 1, 2015 3:41 pm ET

PRINT TEXT

Google is a leader in artificial intelligence and machine learning. But the company's computers still have a lot to learn, judging by a major blunder by its Photos app this week.

The app tagged two black people as "Gorillas," according to Jacky Alciné, a Web developer who spotted the error and tweeted a photo of it.

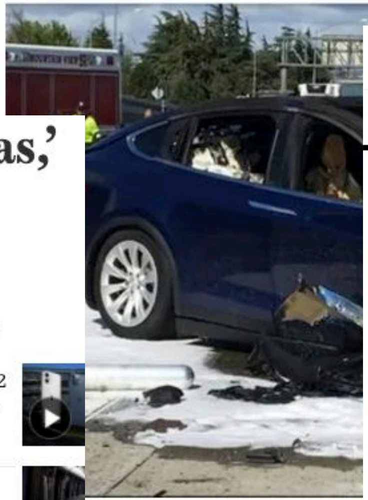
"Google Photos, y'all f**ked...

MOST POPULAR VIDEOS

1. iPhone 12 vs. iPhone 12 Pro: 5G, Cameras and New Design Tested
2. Why East Asia Is Beating the West in Controlling Coronavirus
3. Justice Department to Sue Google for Alleged Anticompetitive Conduct
4. Wuhan, Former Pandemic Center,

Tesla car that crashed and killed driver was running on Autopilot, firm says

- **Company says driver took no action despite system's warnings**
- **Uber settles with family of woman killed by self-driving car**



nature View all Nature Research journals Search Login

Explore our content Journal information Subscribe

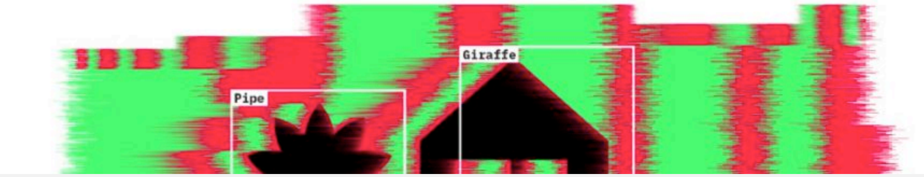
nature > news feature > article

NEWS FEATURE · 09 OCTOBER 2019

Why deep-learning AIs are so easy to fool

Artificial-intelligence researchers are trying to fix the flaws of neural networks.

Douglas Heaven

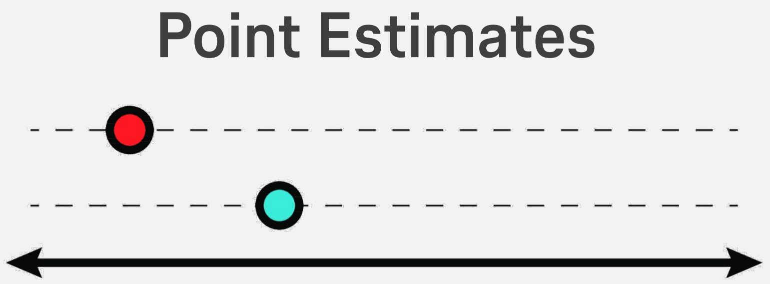


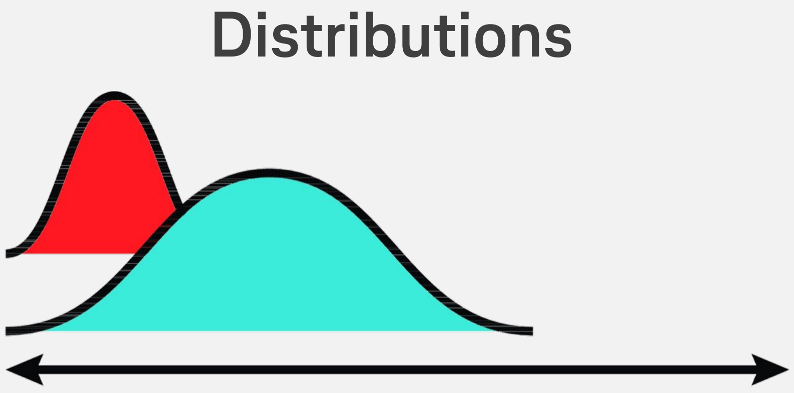
Lufthansa Industry Solutions

How can we interpret uncertainty on a mathematical level?

Using distributions we can judge how likely the most likely is

- Many ML models use point estimates to represent parameters or predictions
- If we can turn them into distributions we can tell a lot more about their implications
- Both examples give us the same expected value
- We are more certain about the expected value of the **red** distribution than of the **blue** because **red** has a smaller variance
- High variance → high risk → high uncertainty

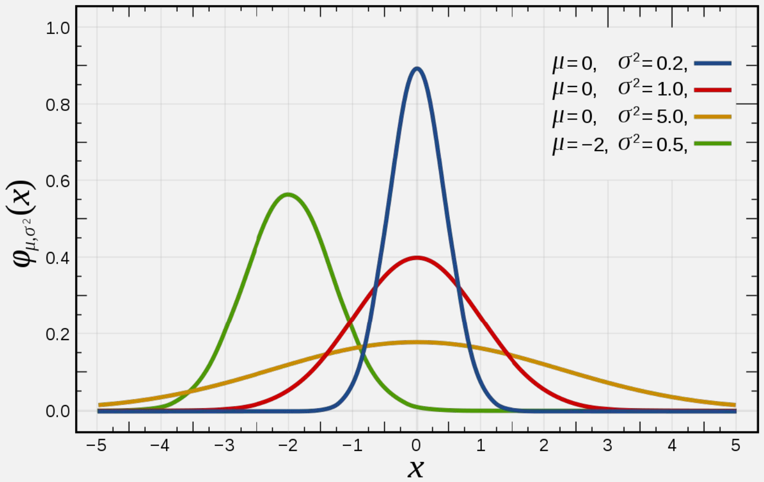




How can we model distributions in machine learning?

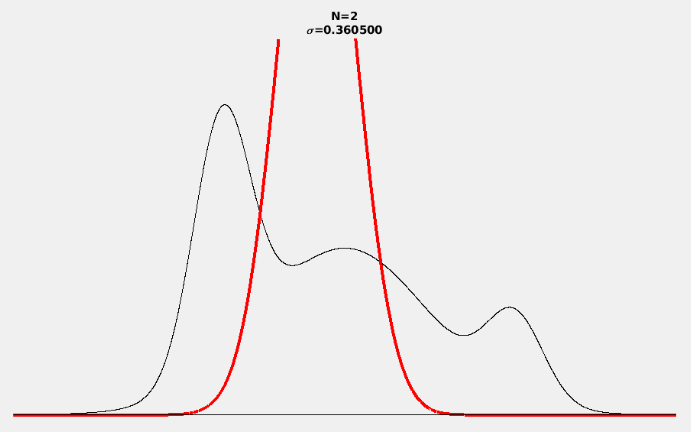
There are two common methods for using distributions in our models

Learning distribution parameters



- Instead of just learning the most likely value we can directly learn the parameters of a distribution
- E.g.: normal distribution → learn μ and σ
- Assumes we know the target distribution

Sampling from the distribution

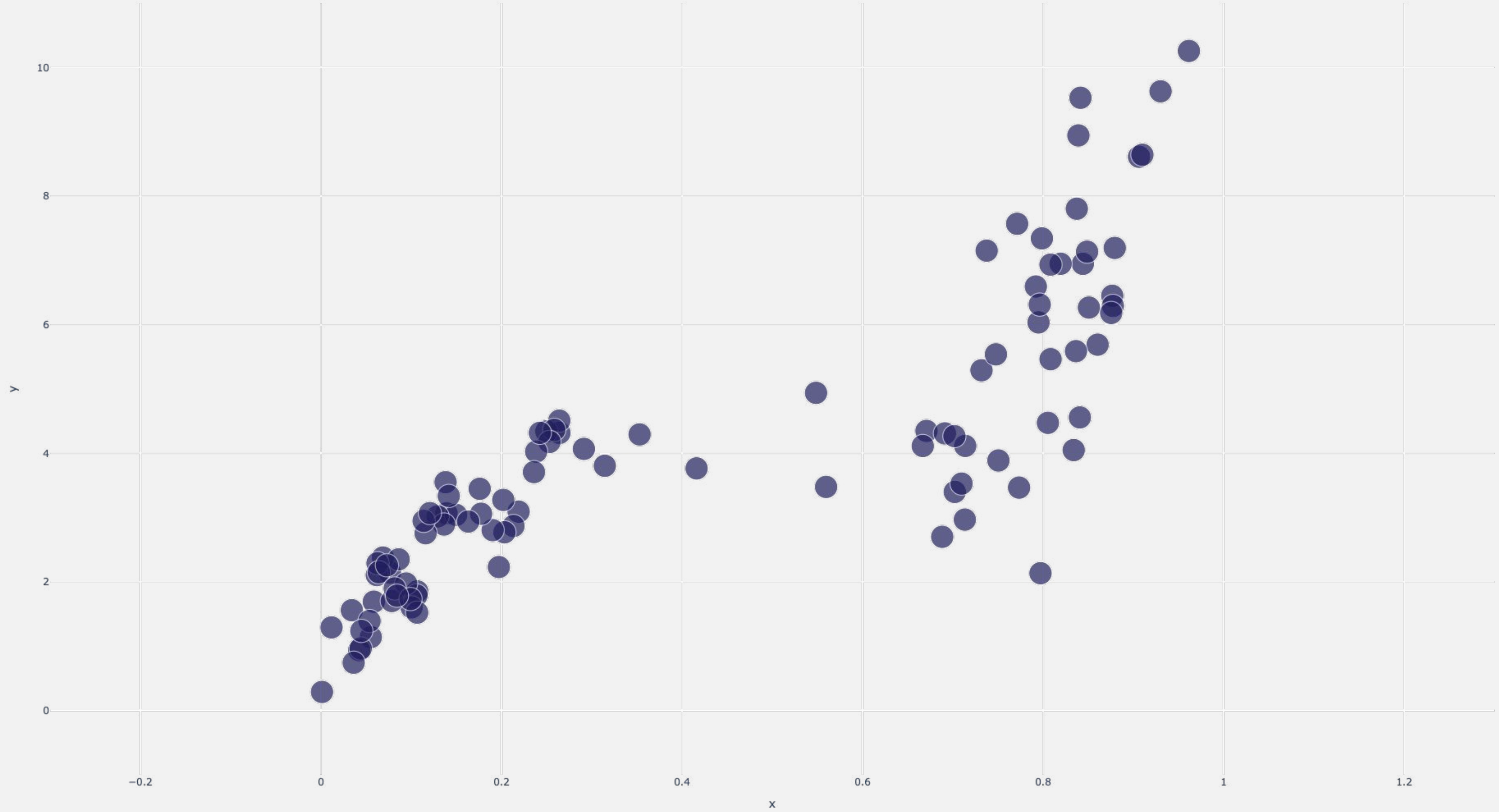


- Make the model stochastic and run multiple forward passes to get multiple predictions
- Each prediction is a sample in a distribution
- Can be very slow when we need many samples

Lufthansa Industry Solutions

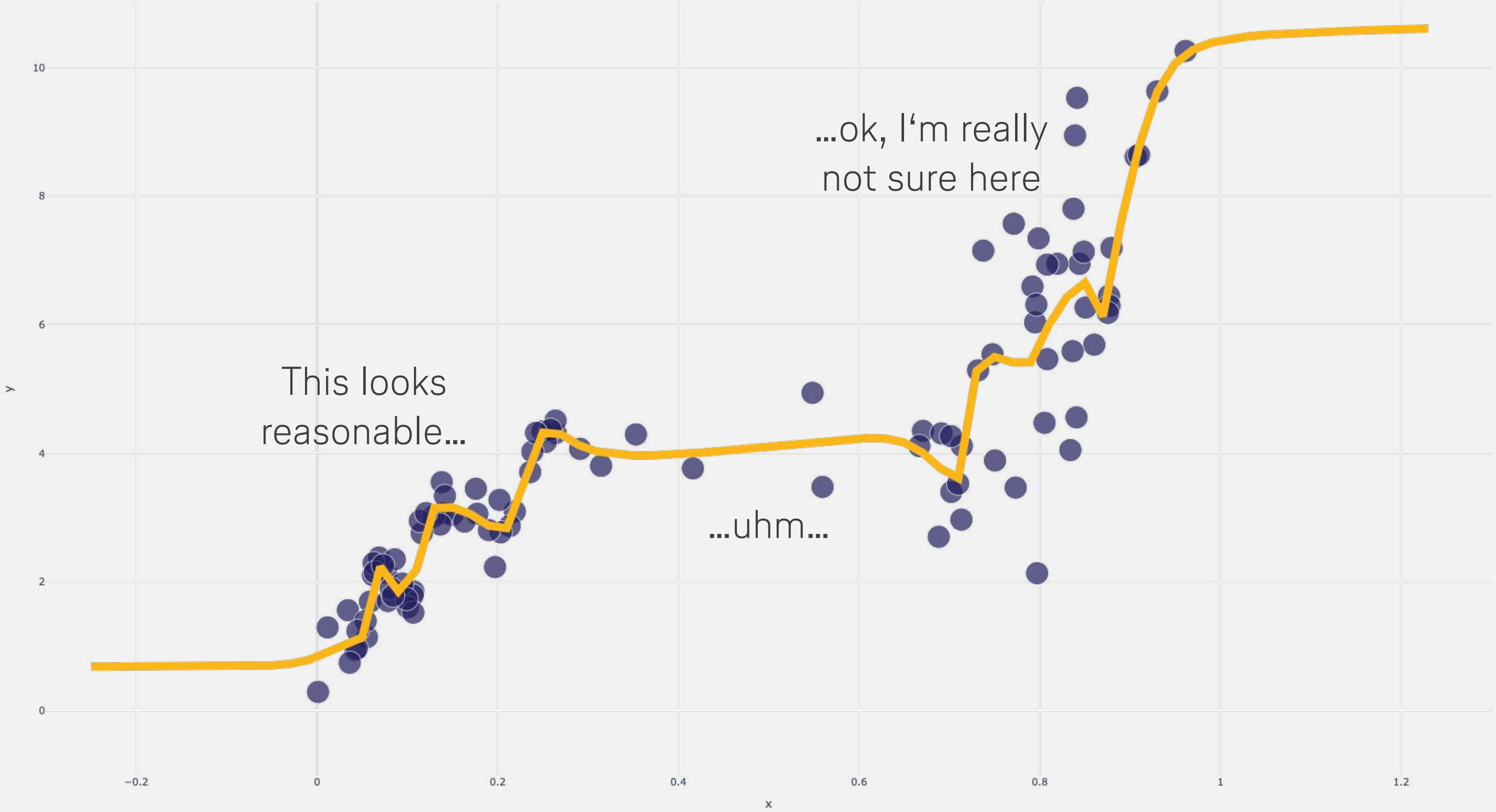
Let‘s take a look at an example

Fitting a 3-layer neural network to a toy dataset using the mean squared error (MSE)



Common neural networks only produce point estimates

In some regions we would like to express uncertainty about the prediction



We train neural networks with the maximum likelihood approach

Where the error is assumed to be normally distributed and the variance is constant

Normal Distribution

$$f(Y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i{}^2}} \exp\left(-\frac{(y_i - \mu_i)^2}{2\sigma_i{}^2}\right)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu_i)^2$$

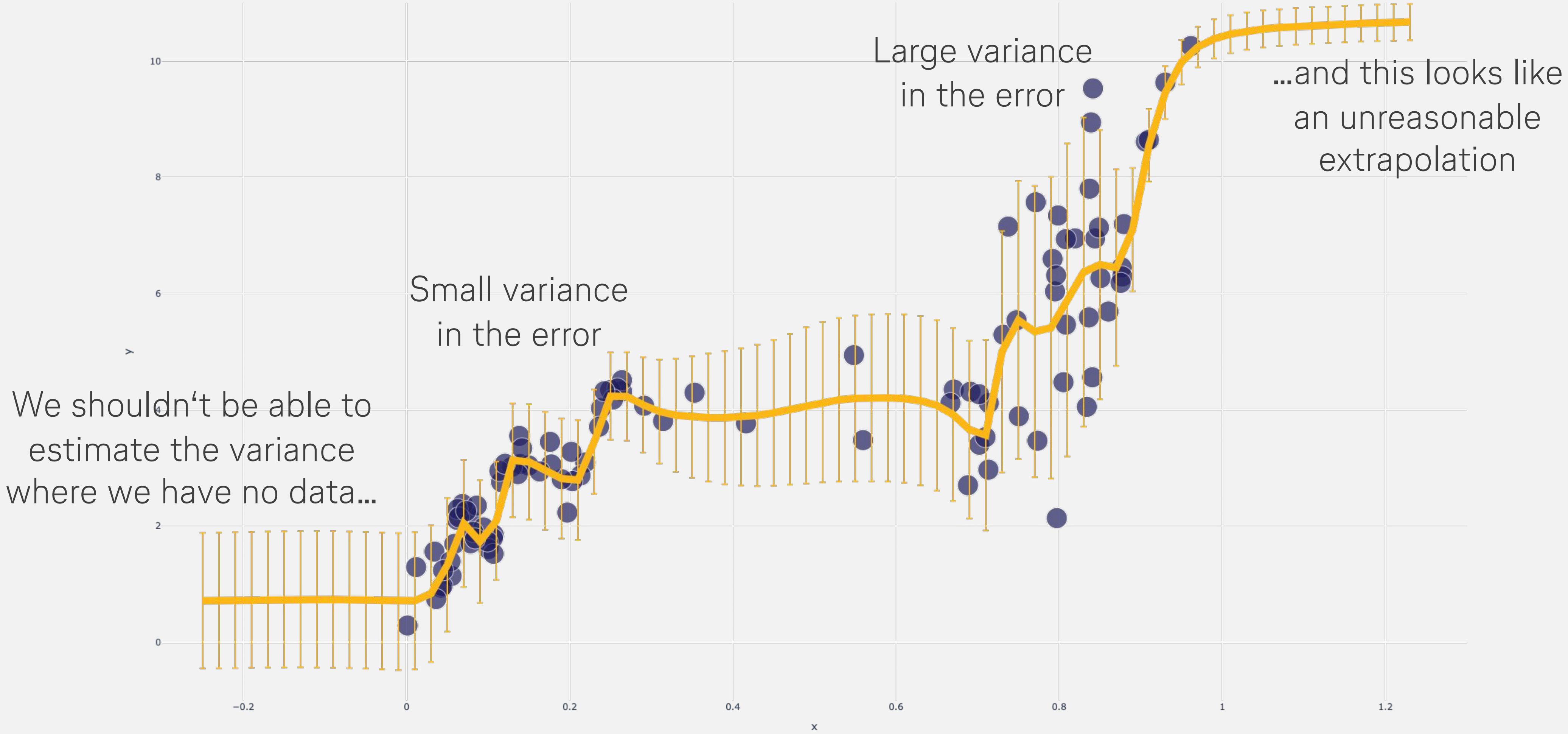
constant variance

$$MSE_\sigma = \frac{1}{n}\sum_{i=1}^{n} \log\sqrt{2\pi\sigma_i{}^2} + \frac{(y_i - \mu_i)^2}{2\sigma_i{}^2}$$

non-constant variance

Aleatoric uncertainty represents randomness in the data

Some error is irreducible and comes from the pure ambiguity of the data

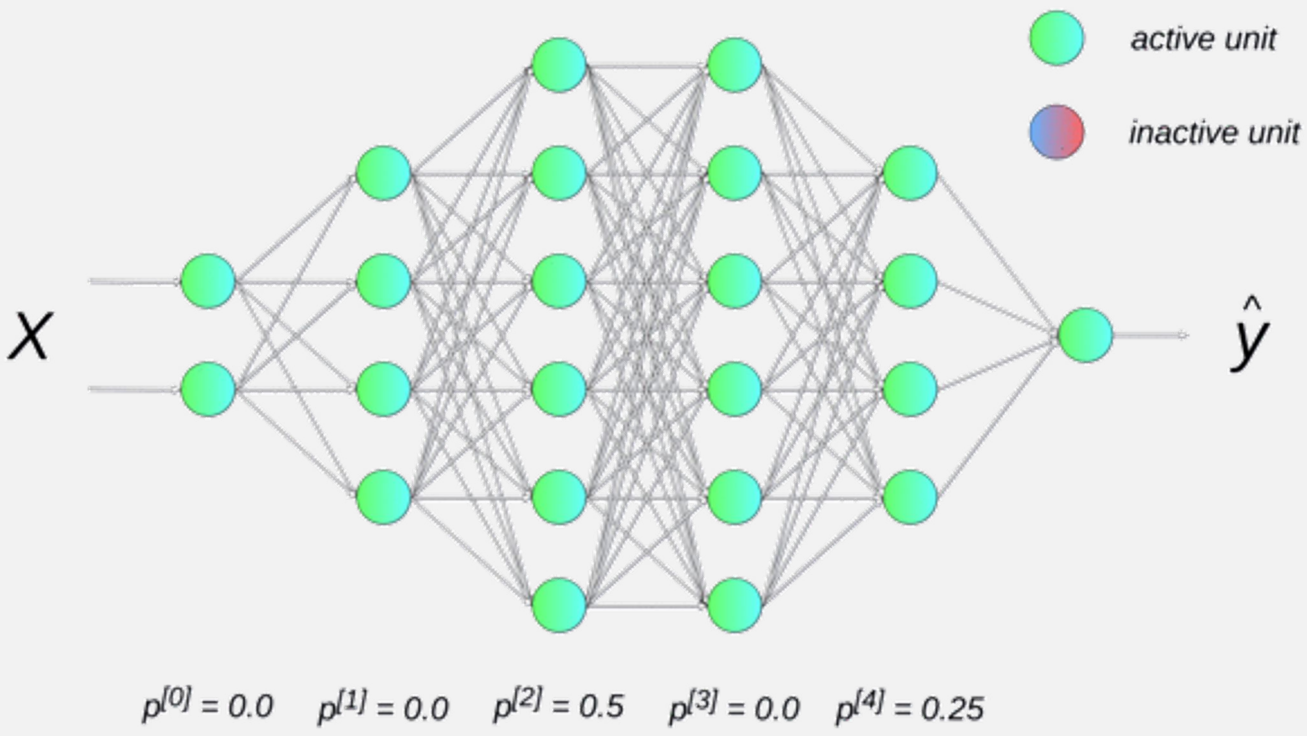


Lufthansa Industry Solutions

Dropout can be used to make a neural network stochastic

Using dropout can be interpreted as creating an ensemble of many subnetworks

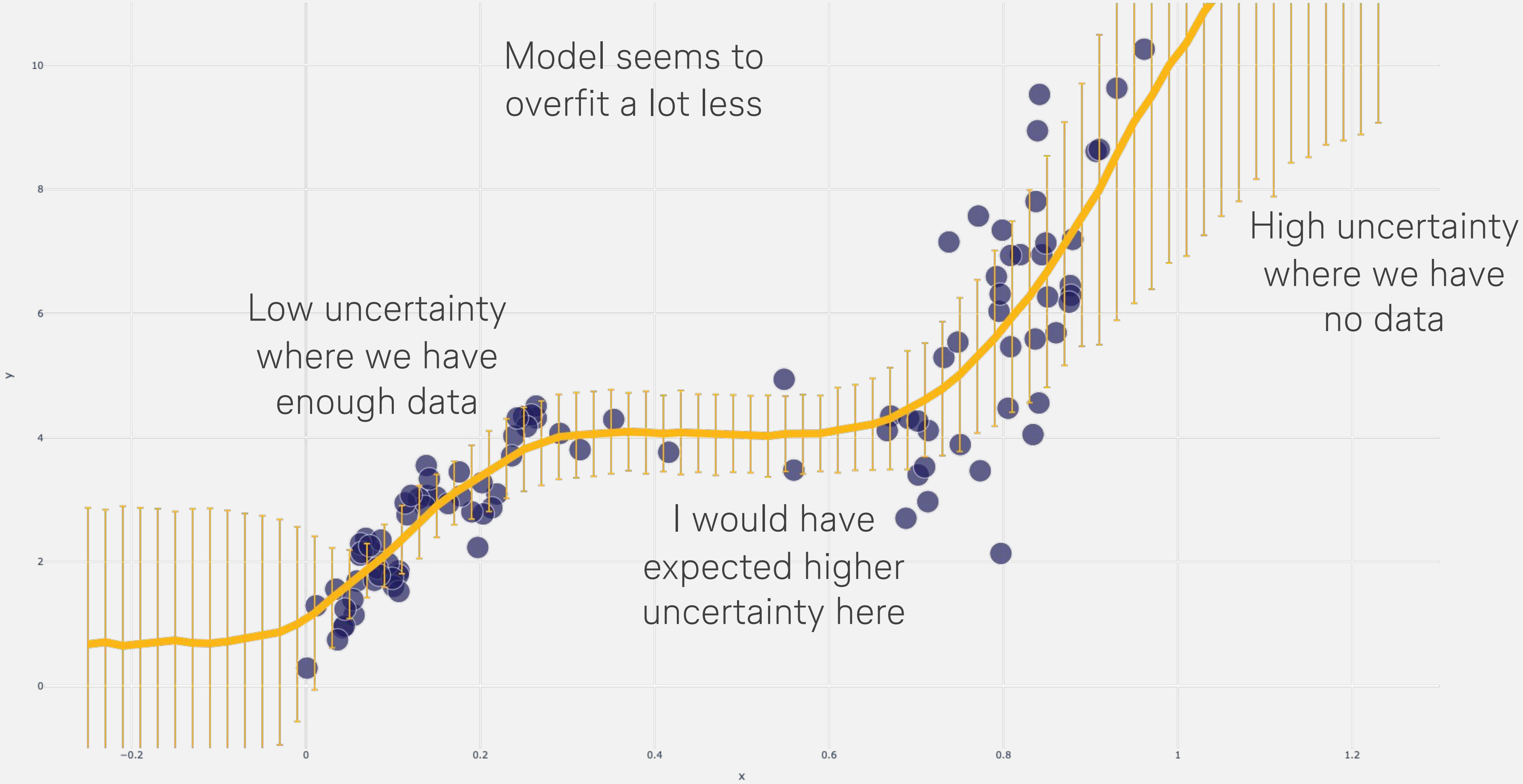
- Dropout sets a random fraction of parameters to zero during training
- Usually it does nothing during inference to access the full capacity
- However, by also using Dropout during inference we make the inference stochastic
- We get different predictions for the same input and can generate a distribution
- You can do this with models that are already trained
- Think of MC Dropout as having a very large ensemble of smaller networks we can sample from



Lufthansa Industry Solutions

Epistemic uncertainty represents what the model doesn‘t know

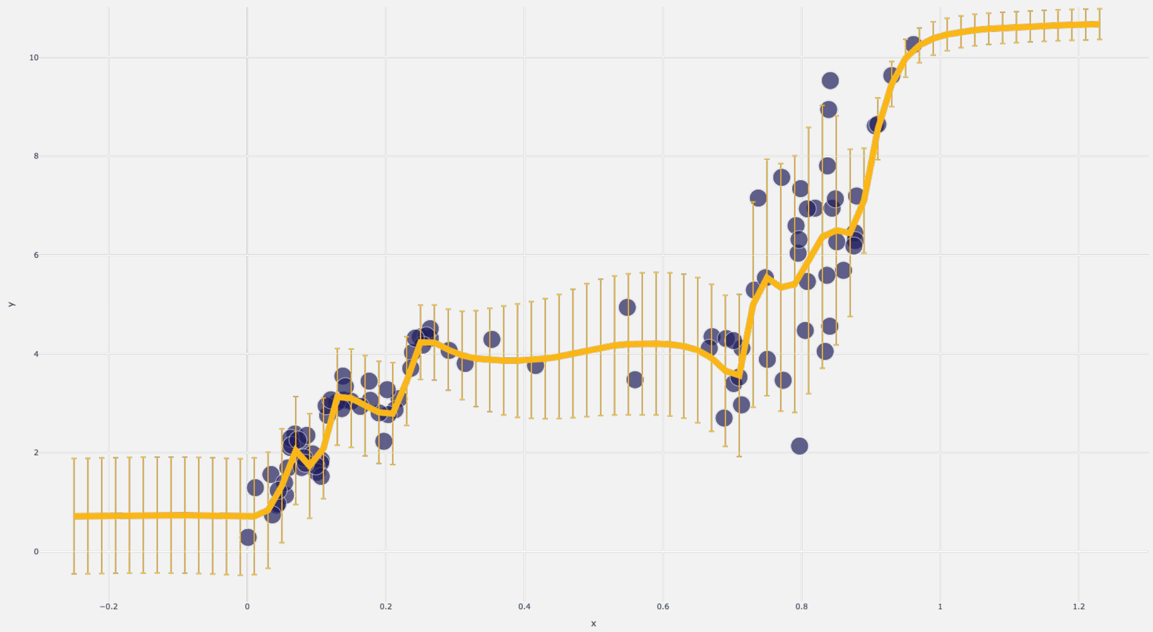
Where there is no data, different models will make different predictions



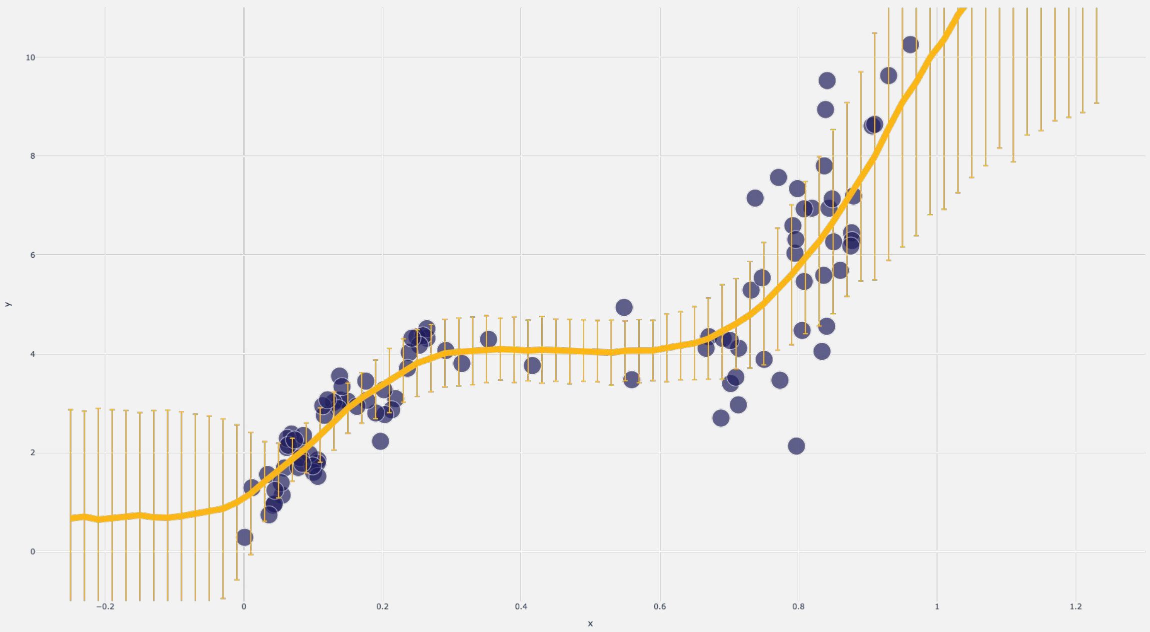
Uncertainties in regression

Two different types of uncertainties explain two different aspects of the experiment

Aleatoric Uncertainty



Epistemic Uncertainty

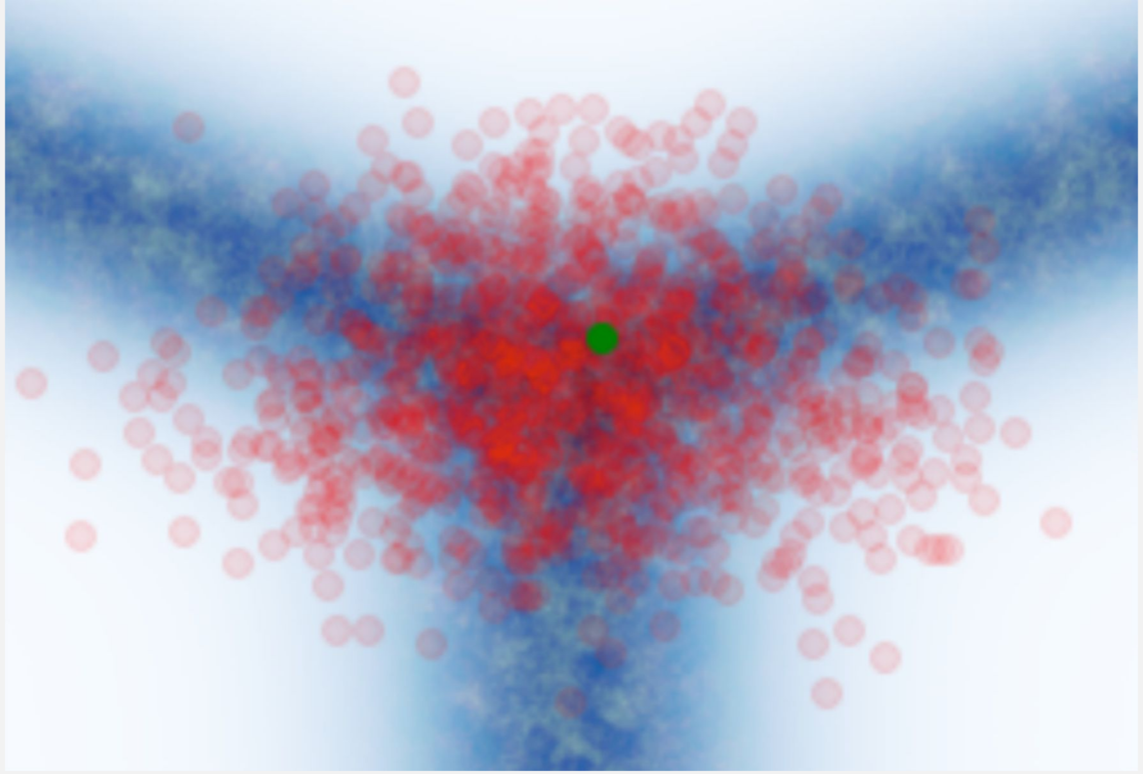


Lufthansa Industry Solutions

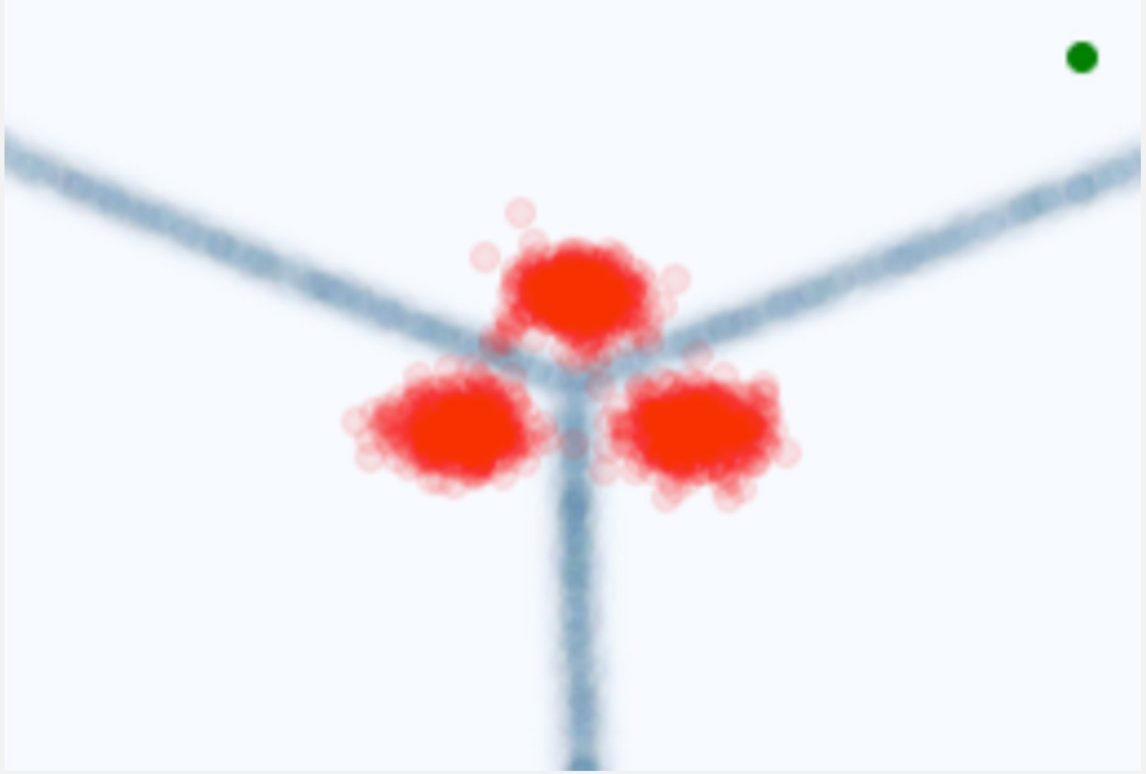
Uncertainties in classification

The same two uncertainties can be found in classification problems

Aleatoric Uncertainty

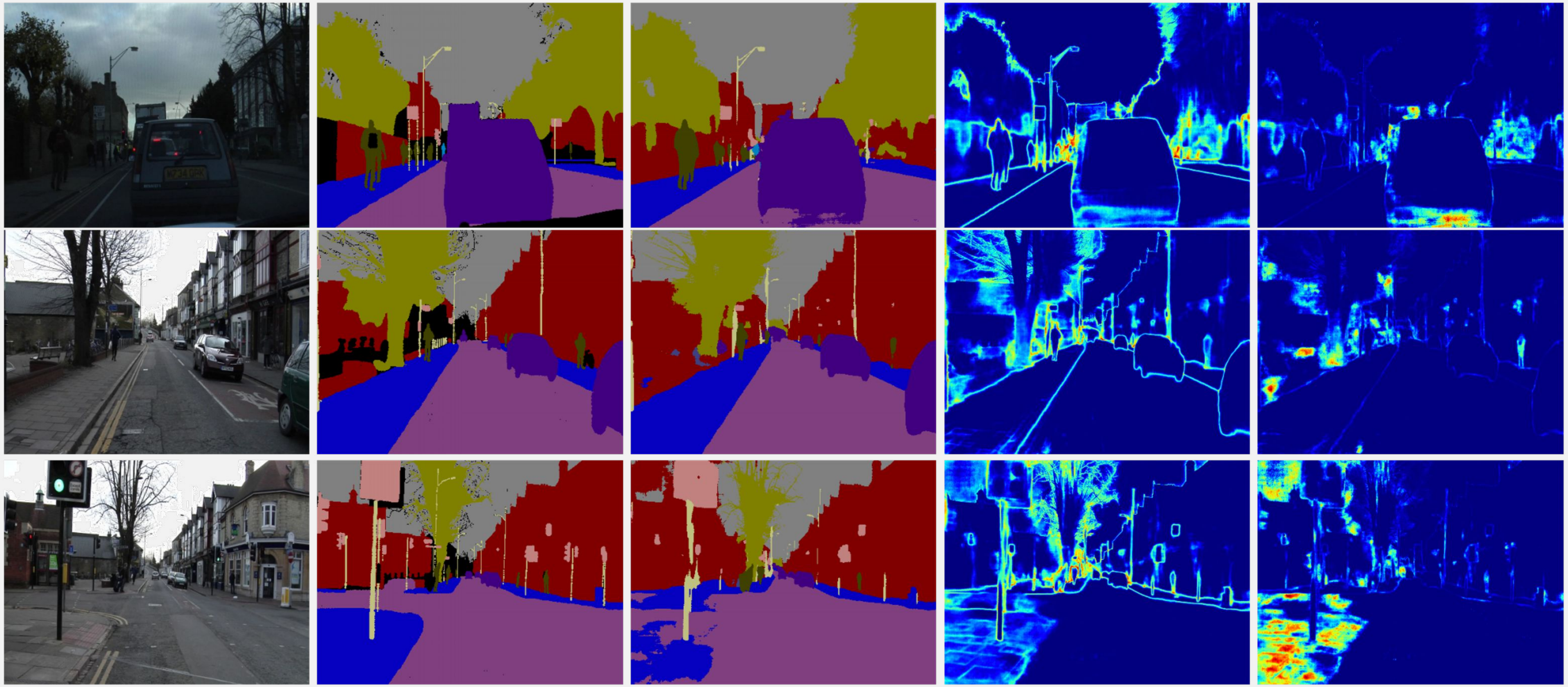


Epistemic Uncertainty



Uncertainties in semantic segmentation

Moving away from toy datasets, we can see the benefits of uncertainty estimates



(a) Input Image (b) Ground Truth (c) Semantic Segmentation (d) Aleatoric Uncertainty (e) Epistemic Uncertainty

Thank You

Paul-Louis Pröve

Data Scientist

+49 (0)1515 892 5503

paul-louis.proeve@lhind.dlh.de

www.Lufthansa-Industry-Solutions.com

Lufthansa Industry Solutions