Problems in Machine Learning – Week 12 – 25/05/2022

Inspecting steel sheets for defence

Given a baseline of 76.2%

Improving the Model resulted in an increase of 0%

Improving the quality of the data increased the accuracy by 16.9%

**Lifecycle of a ML project**

Scope project

Define and collect data

Train model – training, error analysis and iterative improvement

Deploy in production

**Iguana Detection example**

Labelling instructions: use bounding boxes to indicate the position of iguanas

If different labels use different labelling conventions then this creates inconsistencies and makes it harder for a model to learn.

Making data quality systematic: MLOps

Ask two independent labellers to label a sample of images

Measure consistency between labellers

Making it systematic: MLOps