SAP: Large ERO system,

When you show up a client for intelligent automation:

Want to **ideate formulate and implement**  the clients AI solution/goal.

Many clients want/know about AI, but aren’t sure how they can use it to solve their problems. Good at optimising business processes.

Having an AI **solving** problems for you **frees up man power** for **more complex issues**.

**Challenges**

Gaining trust of client in the solution.

Conveying the solution and explaining metrics **in a way** that the client can **understand**.

**Businesses** can have a problem, but **don’t know** the solution. Vise versa can also occur.

**Log bark problem requirements:**

Needed continuous real time video feed.

Object detection, tracking and segmentation

Deployment on an edge device: At this stage you can optimise for what you need ie low powered device, maybe high visibility camera.

**YOLO model**

U-Net Model is another model mentioned (not gone into).

YOLO model is optimised for the edge.

They are doing segmentation and detection models in parallel. To both detect where the log begins and ends, and what is bark and not.

**For yolo**, you have the object, the (x,y) for centre and (width, height) – all must be as a **percentage.**

**Config** file tells you the info needed for **classification** of images. It is important to be able to read and change such a file, and know what values you need to change to adjust output.

For **training YOLO,** you also need to give it **negatives** (whats not a person – ie it says is X a person, and you say its not).

**Example of detection of cans**

Need a dataset of what a Monster and Redbull can looks like.

Research step of which model to use and why.

Need to train the model with dataset, then can use for detection.

**How to create a dataset**

You have a video, but annotations happen on frame by frame bases. The model will take a frame and do detection and such on that frame.

Having a **lower frame** rate is actually a **positive** thing, as having every frame **too**  **similar** when training the model is bad.

Now that the video is broken up into many image files (frames). You annotate each image. (VGG Image Annotator)

You can also use tools to do the annotation automatically (depending if you need accuracy vs quantity).

You must make sure that the dataset that the model trains on, represents the actual read world scenario data it will need to detect in the end.

Its important to have images which have background noise, which are blurred, skew ect, so that when that frame happens in reality, the model will be able to detect them.

**You** don’t train your model on **every** image. You test your model with non-trained images (that don’t have annotations).

You want to use greyscale if you don’t need any info from the colour of the image. Such a scenario would be character/licence plate recognition.

**What goes wrong with datasets**

When the dataset is can detect that the can is monster, but also detects that it is a red bull in that can. This can happen when features from one are mistaken – can happen if it picks up from when both were in image.

You can solve this issue by looking at % of how sure it is of that detection. So setting an eg 60% threshold.

**Real world**

Understanding how AI model works is needed is important for creating a solution and explaining things to the client. This is the case even when using a model such as YOLO. This helps in knowing what eg training data the model needs.

Theres no time to build a model yourself. More realistic in industry is to use something like YOLO, so knowing how to implement and use the tools is important.

Try out Imbuntu, Linux

**Resources**

Lambda labs – cheap GPU for model training

Colab – also good tool (was used in monster vs redbull demo)