



USING MACHINE LEARNING TO EXTRACT DATA

HELLO!

I am Manoj

I am here to share a case study
on Machine learning

You can find me at
manoj@rbtsb.com



“

Nobody phrases it this way, but I think that artificial intelligence is almost a humanities discipline. It's really an attempt to understand human intelligence and human cognition — Sebastian Thrun

1. Background























Background - Problem statement

- ◆ Customer is an engineering consulting company in oil and gas industry
- ◆ As part of regular activities, engineering drawing schematics come in scanned copy image form
- ◆ A voluminous, repetitive and manual job is to extract various information, metadata from the schematics
- ◆ Manual information extraction was expensive, slow and error prone

2. Scope of Work

Scope of Work

Automation of the extraction of metadata from scanned engineering drawings (P&ID – piping and instrumentation diagrams) using AI techniques

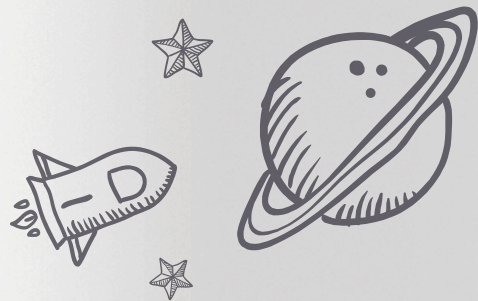
Line Symbols		Blinds	
	Piping		Line Reducer
	Instrument Air		Ejector
	Hydraulic		Line Strainer
	Electric		Steam Trap
	Capillary		Flexible Coupling
	Bursting (Rupture) Disk		
Valves			
	Gate Valve, Hand-operated		Control Valve
	Globe Valve, Hand-operated		Solenoid Valve
	Plug or Cock Valve, Hand-operated		Motor-operated
	Check Valve		Piston-operated
	Butterfly Valve		Safety Valve or Relief Valve
	Angle Valve, Hand-operated		

3. Challenges

Challenges

Automation of metadata of a P&ID has challenges:

- ◆ Large and complicated diagrams involving many components, pipelines etc.
- ◆ Labelling are sometimes handwritten
- ◆ Low resolution



Solution Approach

- ❖ Image processing and deep neural network models
- ❖ Deep learning platform

Solution

Objective

Automation refers to extracting relevant properties of each engineering component (like type of component, its size) and their connectivity in a particular engineering drawing (schematic)

Output

The output of the exercise is to generate a spreadsheet automatically from scan copy of an engineering drawing which captures all the above properties for each component.

Solution - Computer vision with deep learning

Design Decision

It is a computer vision problem solved with technologies like image processing and deep neural network models

Solution Approach

For example, detection of a particular component is a multi-classification problem

Tool

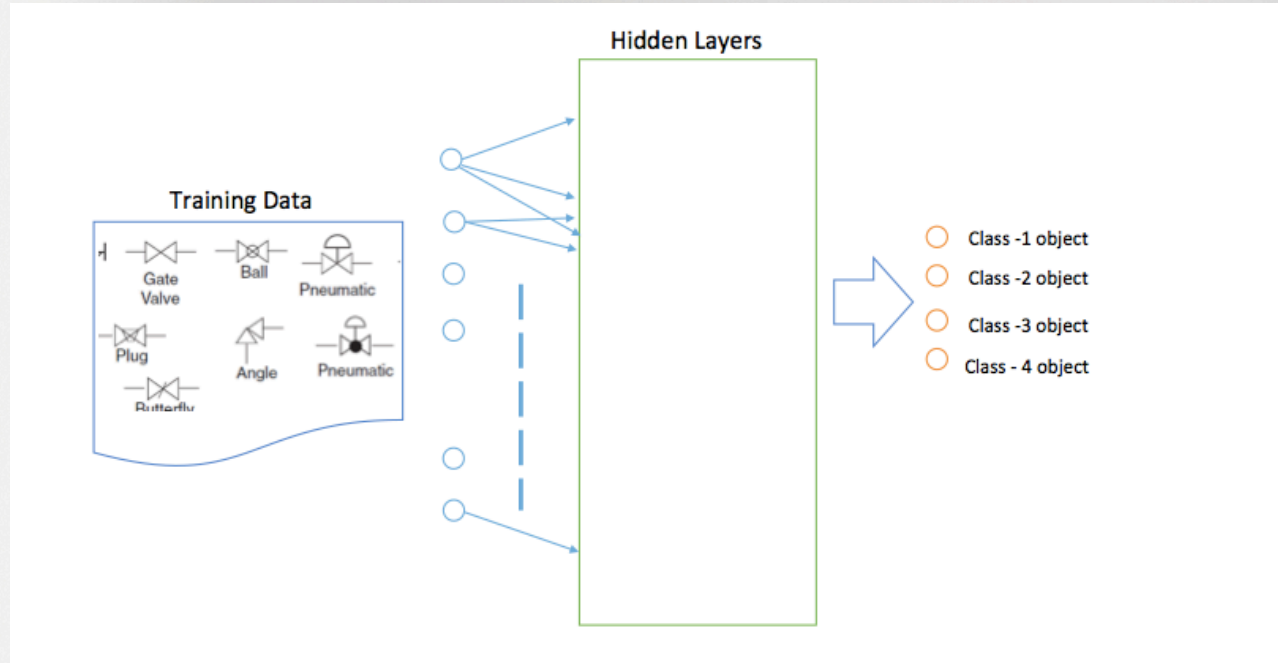
Model is trained with various labelled engineering components and deep learning platform

A Well trained model

System can automatically detect even if the components are scaled or rotated in the actual drawing



Deep Neural network architecture for object recognition



**With 200 dpi
resolution result is
100% correct**

“

*Aim for simplicity in Data Science.
Real creativity won't make things
more complex. Instead, it will
simplify them — Damian Duffy
Mingle*

Benefits of Machine learning in this project

First

Less amount of training data for AI model required, as shapes are well defined

Second

Modern deep learning platform like *TensorFlow* automatically generates other geometric transformations like scaling, rotation from the basic shape

Tangible benefits

ROI for the client comes in the form of below benefits:

- ◆ Much faster processing (from hours to few seconds)
- ◆ Elimination of human errors, omissions
- ◆ Direct manpower reduction (about 70%), as only final evaluations and cases where tool fails to perform job, go to staff

Next Step

Identify how valves are connected through pipes

- ◆ Identify connectivity and
- ◆ To read the labels (text) in the drawing

“

*Neither Man, nor machine can
replace its creator*

*— Tapan Ghosh, Faceless The Only
Way Out*

THANKS!

You can find me at
manoj@rbtsb.com

