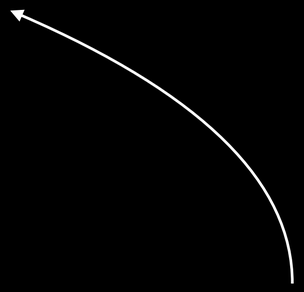
Machine Learning Roadmap 2020

(a machine-learning-flavoured-visual-iterative-living-mind-map/compass)

“What is machine learning?”

The curious internet-dwelling user (maybe you)

Machine learning is turning things (data) into numbers and finding patterns in those numbers.



The computer does this part.

How?

Math.

(we'll cover a little on this later)

Traditional programming (software 1.0)

(does not require 2.0)

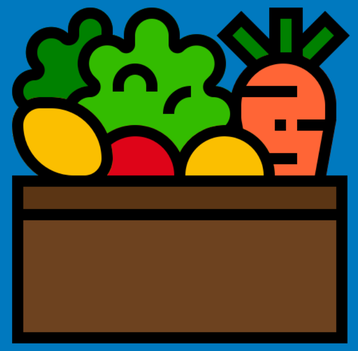
vs.

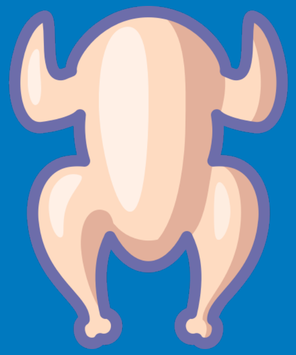
(requires 1.0)

Machine learning (software 2.0)

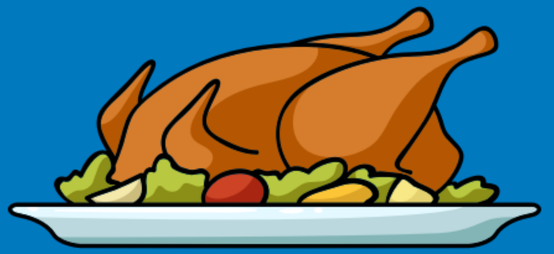
Andrei Karpathy post: Software 2.0

Traditional programming





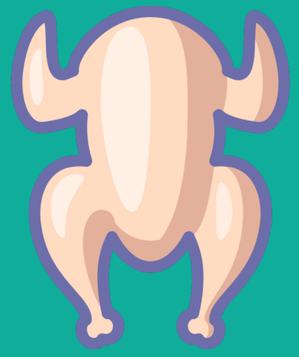
1. Cut vegetables
2. Season chicken
3. Preheat oven
4. Cook chicken for 30-minutes
5. Add vegetables



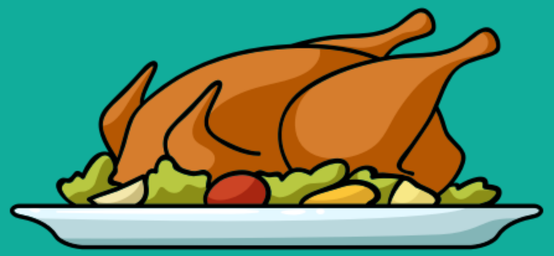
Starts with → **Makes**

Machine learning algorithm

Inputs



Output



1. Cut vegetables
2. Season chicken
3. Preheat oven
4. Cook chicken for 30-minutes
5. Add vegetables

Starts with → **Figures out**

“Why use machine learning?”

Another curious (perhaps even more curious than before) internet dweller

Good reason: ~~Why not?~~

Better reason: Can you think of all the rules?

(probably not)

(maybe not very simple...)

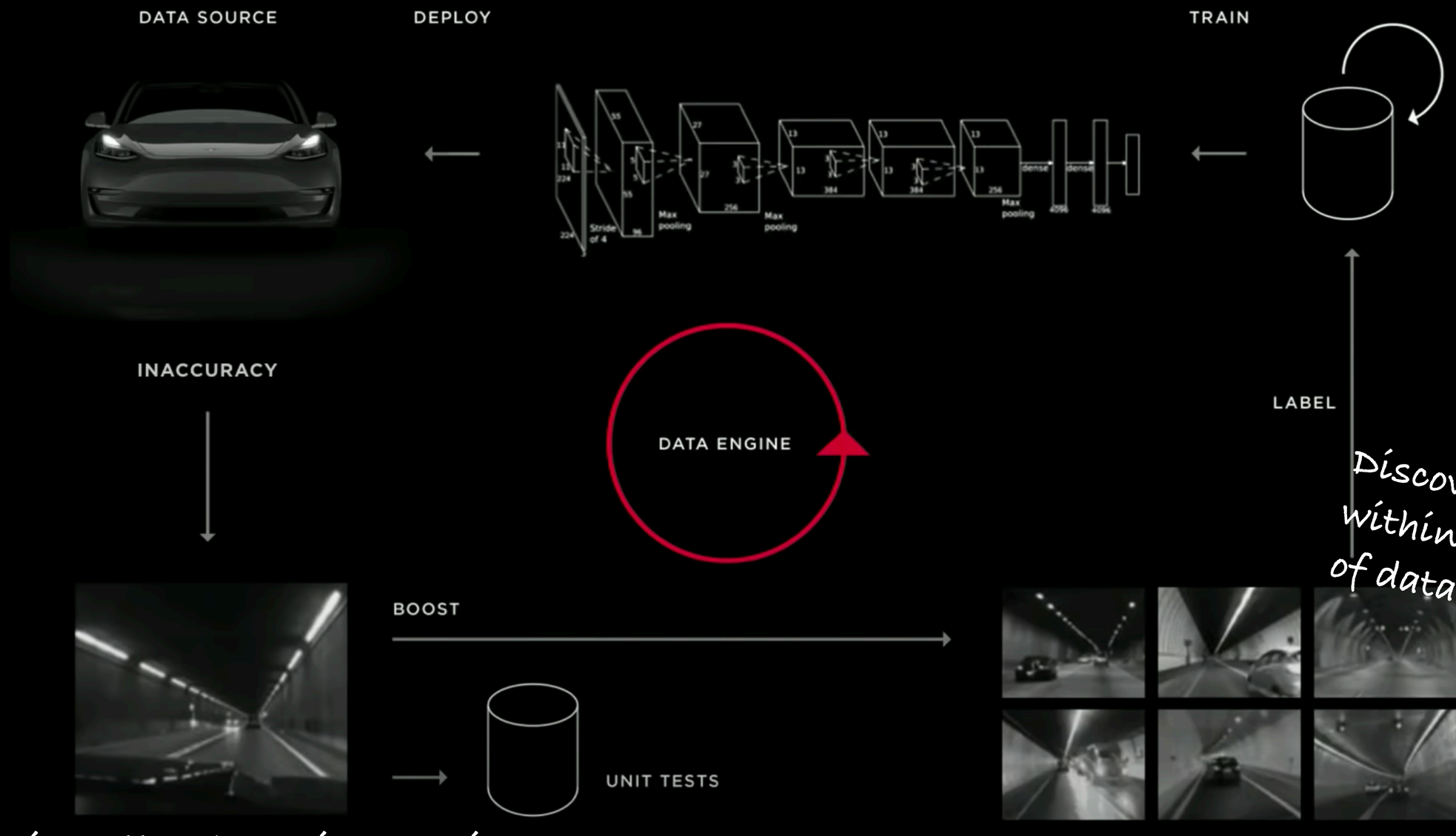
"If you can build a **simple rule-based** system that doesn't require machine learning, do that."

A wise software engineer… (actually rule 1 of Google's Machine Learning Handbook)

What machine learning is good for 

- **Problems with long lists of rules**—when the traditional approach fails, machine learning may help.
- **Continually changing environments**—machine learning can adapt (‘learn’) to new scenarios.
- **Discovering insights within large collections of data**—can you imagine trying to go through every transaction your (large) company has ever had by hand?

Problems with long lists of rules



Discovering insights within large collections of data

Continually changing environments

TESLA LIVE

Source: Tesla Autonomy Day video

What we're going to cover (broadly)

1. 🤔 **Machine Learning Problems**—what does a machine learning problem look like?
2. ♻️ **Machine Learning Process**—once you've found a problem, what steps might you take to solve it?
3. 🛠️ **Machine Learning Tools**—what should you use to build your solution?
4. 🧮 **Machine Learning Mathematics**—what exactly is happening under the hood?
5. 📚 **Machines Learning Resources**—okay, this is cool, how can I learn all of this?

How:  

How to approach this roadmap


(actually a compass)

- Explore
- Comment (what's missing?)
- Give feedback, give advice
- Share
- Follow your curiosity (there's a lot, so don't expect to get it all)
- Explore

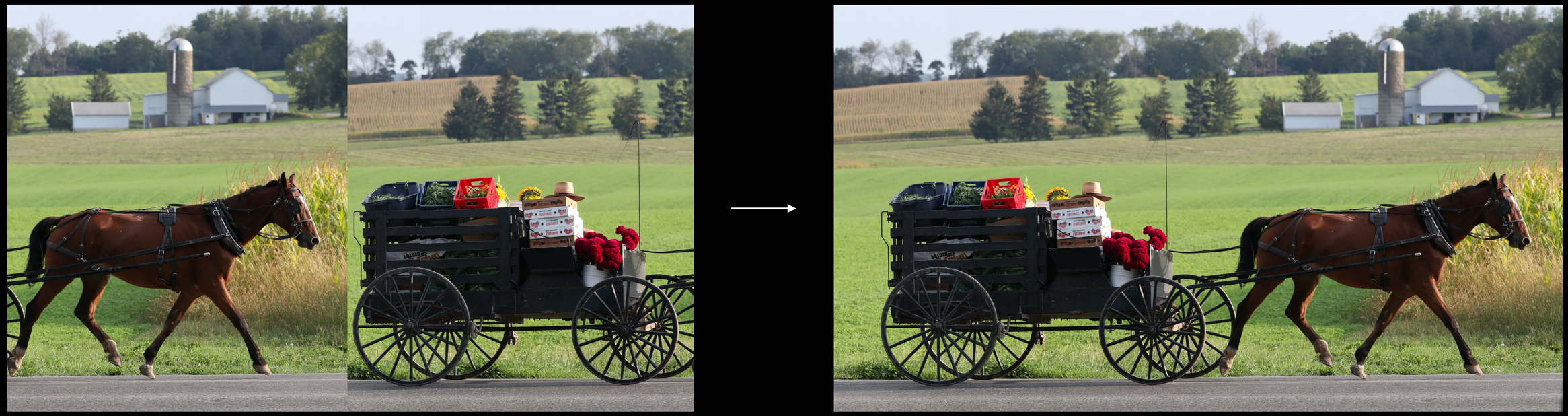
(twice on purpose)

Ready?

Okay, let’s go.

1. 🤔 Machine Learning Problems

1. 🤔 Machine Learning Problems

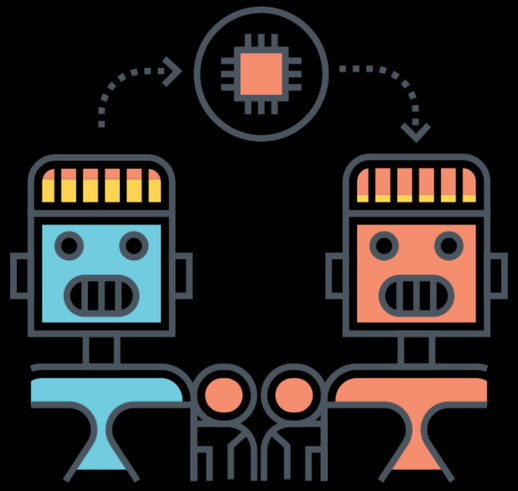


1. 🤔 Machine Learning Problems

(categories of learning)

Supervised Learning

Unsupervised Learning



Transfer Learning

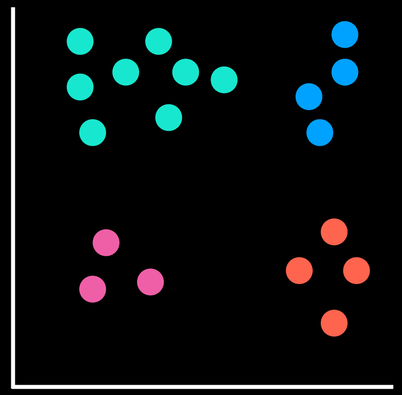
Reinforcement Learning

1. 🤔 Machine Learning Problems

(problem domains)

Classification

Regression



Clustering

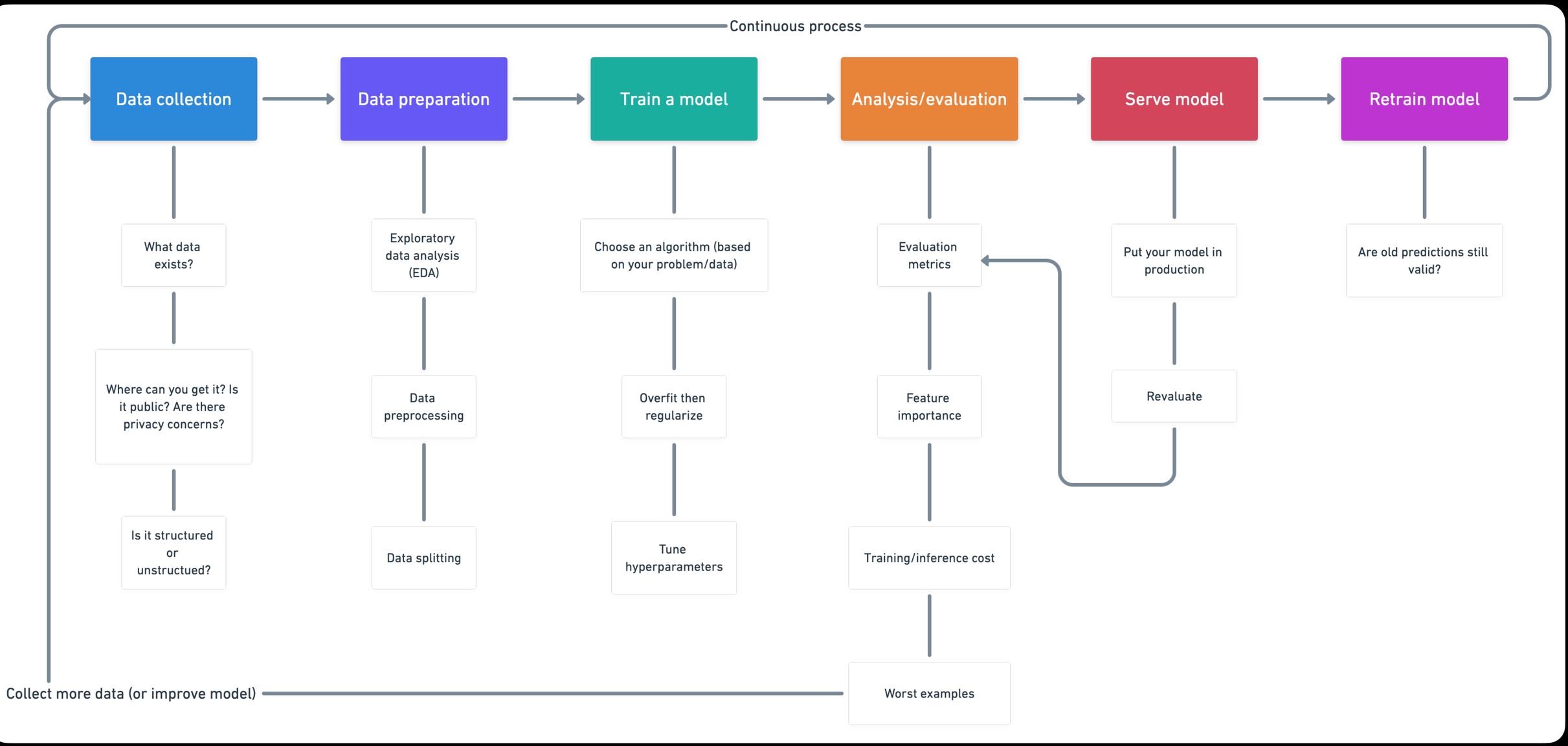
(dropped) (most important)

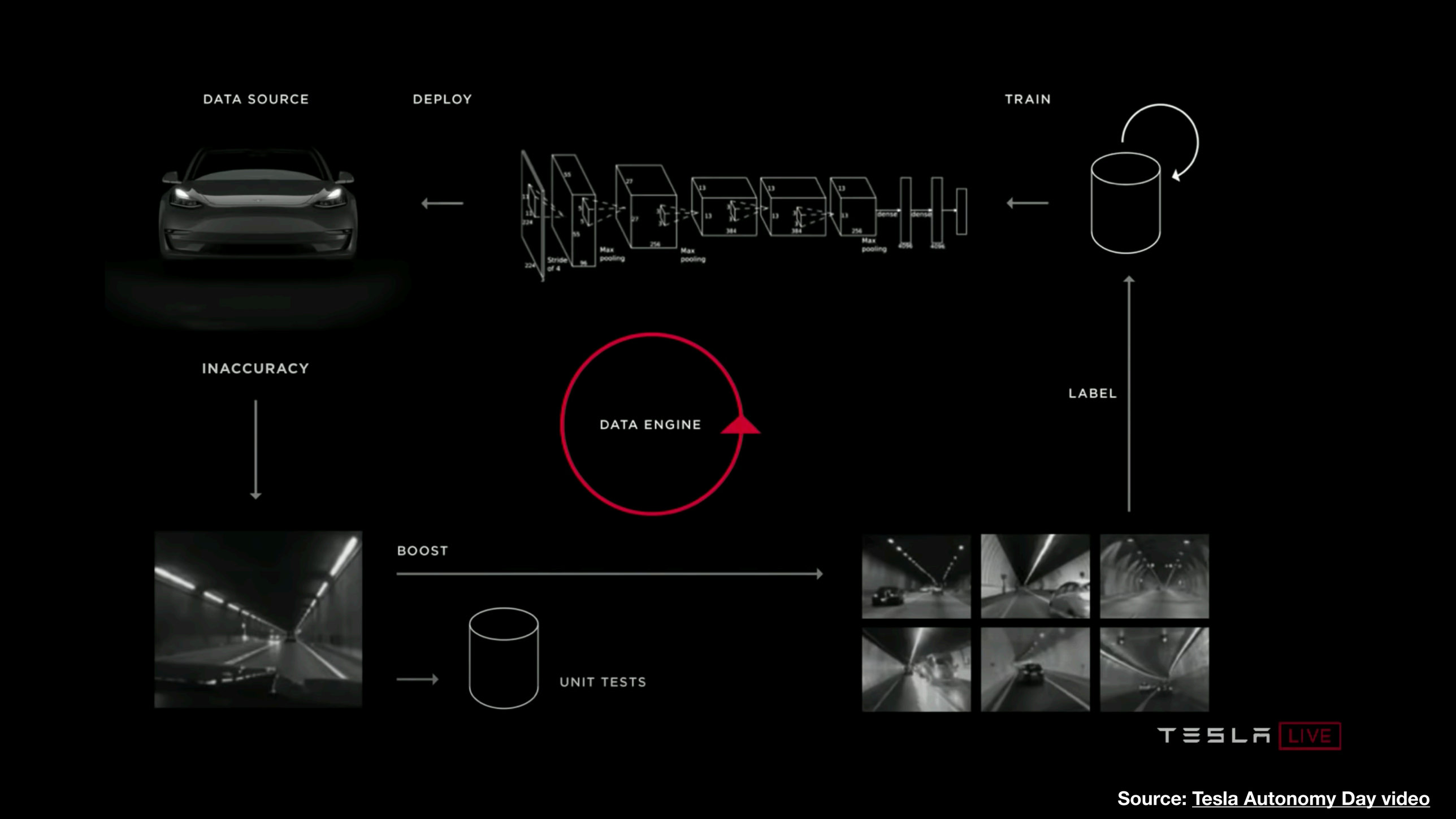
ID	Weight	Heartrate	Age	Heart Disease?
0	76	54	55	0
1	81	42	34	0
2	90	70	47	0
3	67	100	79	1

Dimensionality reduction

2. ♻️ Machine Learning Process

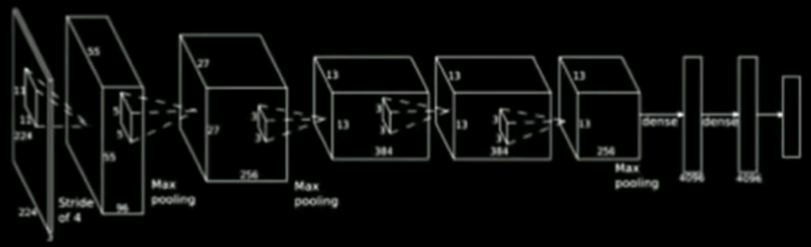
2. ♻️ Machine Learning Process





Source: Tesla Autonomy Day video

DATA SOURCE

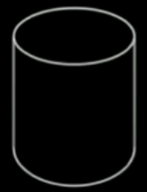


DATA ENGINE

Continuous process

Data collection → Data preparation → Train a model → Analysis/evaluation → Serve model → Retrain model

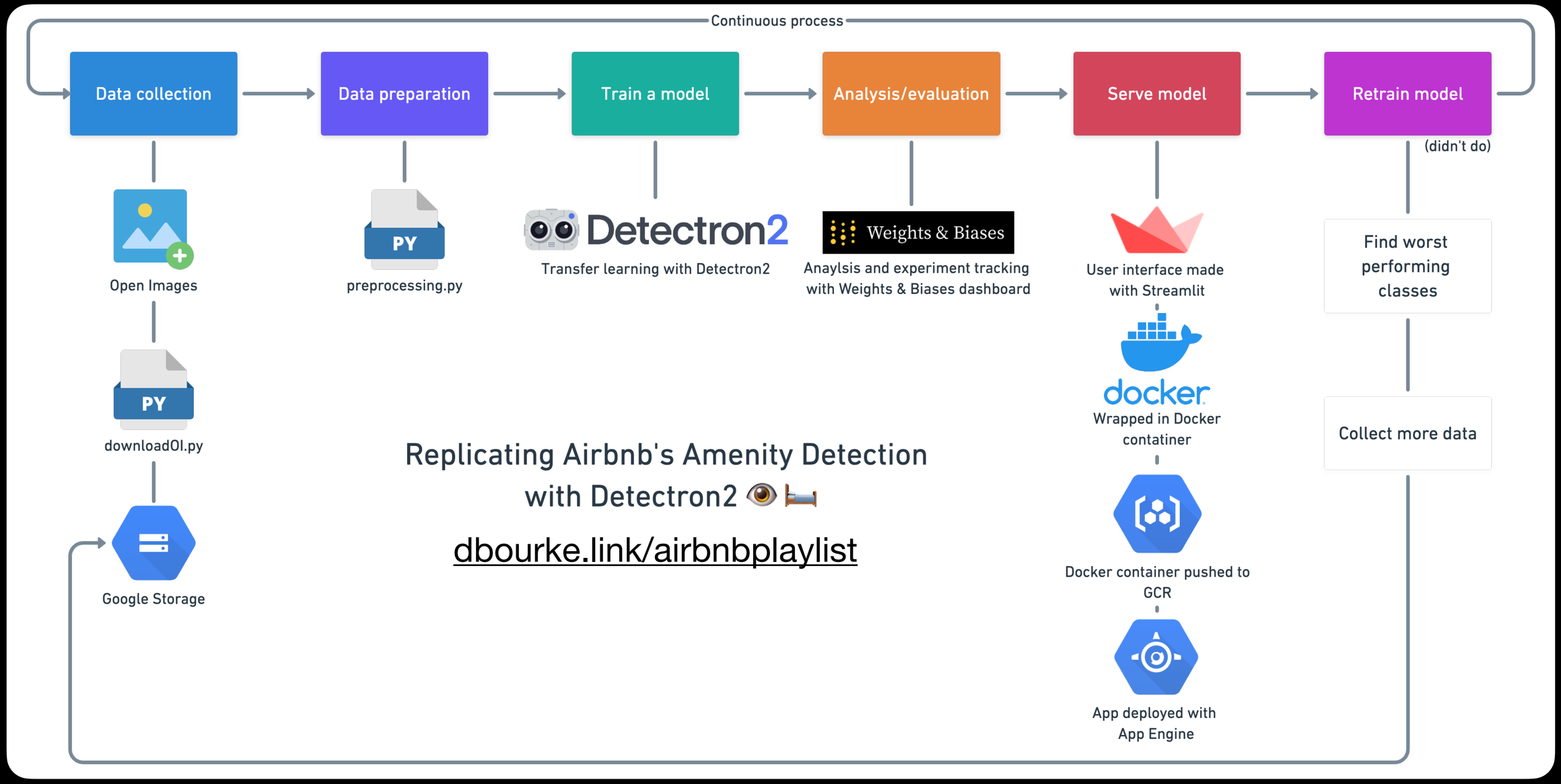
INACCURACY



UNIT TESTS

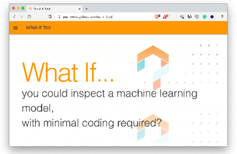
Source: Tesla Autonomy Day video

2. ♻️ Machine Learning Process



3. 🛠️ Machine Learning Tools

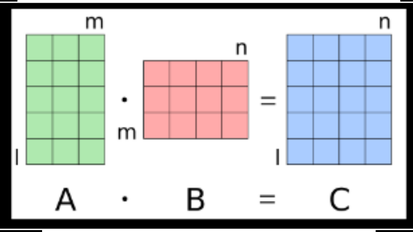
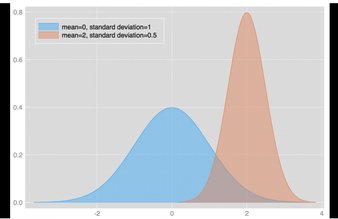
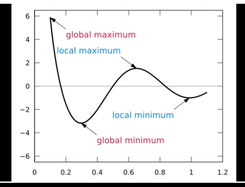
3. 🛠️ Machine Learning Tools

Libraries/code space	Experiment Tracking	Pre-trained models	Data and model Tracking	Cloud Compute Services	Hardware (building your own deep learning PC)	AutoML & hyperparmeter tuning	Explainability	ML Lifecycle	User Interface Design
jupyter	Dashboard by Weights & Biases	Detectron2	Artifacts by Weights & Biases	Google Cloud		Sweeps by Weights & Biases	SHAP	Kubeflow	Streamlit
TensorFlow.js TensorFlow Lite	 TensorBoard	TensorFlow Hub	DVC Data version control	aws		Google Cloud AutoML		SELDON	
PyTorch PyTorch Lightning	neptune.ai	PYTORCH HUB		Microsoft Azure		TPOT		mlflow	
ONNX		HuggingFace Transformers		CO		Microsoft Azure AutoML			
dmlc XGBoost catboost.ai									
scikit learn									

4. 🧮 Machine Learning Mathematics

4. 🧮 Machine Learning Mathematics

(some of the main ones)

Linear Algebra	$\begin{aligned} 2x + y - \ z &= \ \ 8 \\ -3x - y + 2z &= -11 \\ -2x + y + 2z &= \ -3 \end{aligned}$
Matrix Manipulation	
Multivariate Calculus	$f(x,y) = \frac{x^2 y}{x^4 + y^2}$
The Chain Rule	$\frac{d}{dx}\left[f\left(g(x)\right)\right] = f'\left(g(x)\right)g'(x)$
Probability + Distributions	
Optimization	

5. 📚 Machine Learning Resources

(where to start learning)

5. 📚 Machine Learning Resources

dbourke.link/mlcourse

Beginner (3-6 months+)

Machine Learning Concepts
(get your mind ready)

Learn within
Jupyter (or Colab)

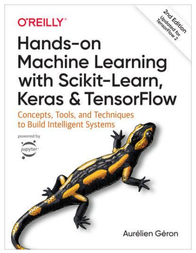
Python

NumPy

pandas

scikit learn

Part 1



Milestone project 1

Advanced (6-12 months+)

fast.ai Deep Learning
for Coders (part 1)

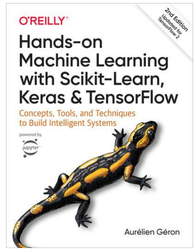
TensorFlow in Practice

fast.ai Deep Learning
from the foundations (part 2)

Deep Learning
Specialization

Full Stack Deep Learning

Milestone project 2 (deployed)



Part 2

Bonus (sprinkle in)

`./missing-semester (+)`

The missing part of your
CS degree

Choose one

aws

Google Cloud

Microsoft Azure

freeCodeCamp

Web Development

Khan Academy

Math when needed

arXiv.org

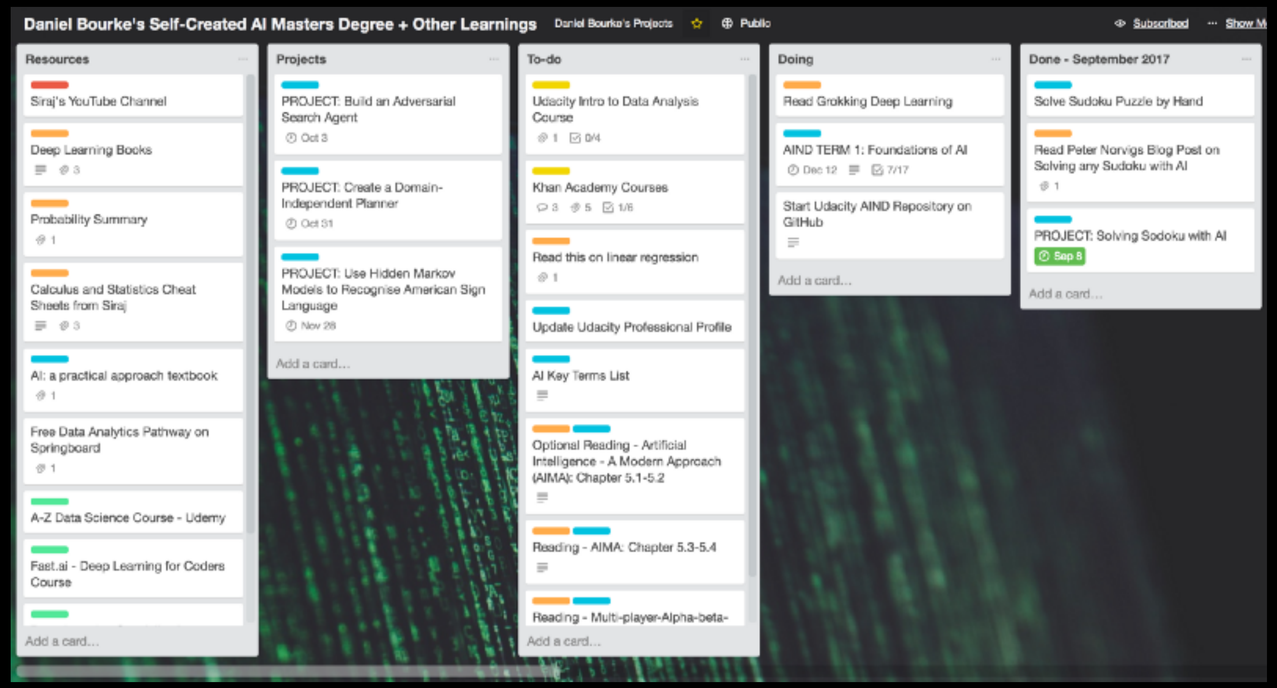
Implement a paper

git

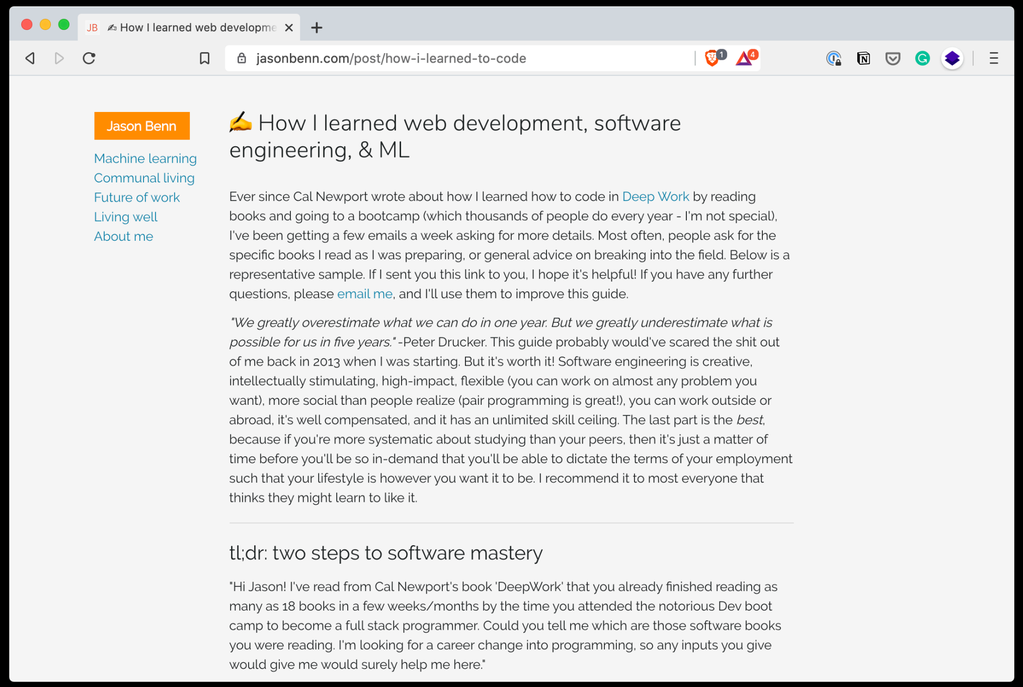
Version control

5. 📚 Machine Learning Resources

(example curriculums)



dbourke.link/aimastersdegree



https://jasonbenn.com/post/how-i-learned-to-code

5. 📚 Machine Learning Resources

(some useful places to visit)

sotabench

Papers with Code

Made with ML

End

Summary, plus next steps

What we~~'re going to~~ covered (broadly)

1. 🤔 **Machine Learning Problems**—what does a machine learning problem look like?
2. ♻️ **Machine Learning Process**—once you've found a problem, what steps might you take to solve it?
3. 🛠️ **Machine Learning Tools**—what should you use to build your solution?
4. 🧮 **Machine Learning Mathematics**—what exactly is happening under the hood?
5. 📚 **Machines Learning Resources**—okay, this is cool, how can I learn all of this?

How:  

What's missing?

- Instance vs. model-based learning
- Machine learning pipelines (still emerging on best practices)
- More specific topics (such as computer vision, NLP, RL…)
- Probably much more

(leave a comment)

Sources (thank you's)

- Daniel Formoso's machine learning mind maps
- Hands-on Machine Learning Book (2nd edition)
- 100-page Machine Learning Book
- Countless blog posts (check the mind map for links)

(machine)

Keep learning
Keep creating