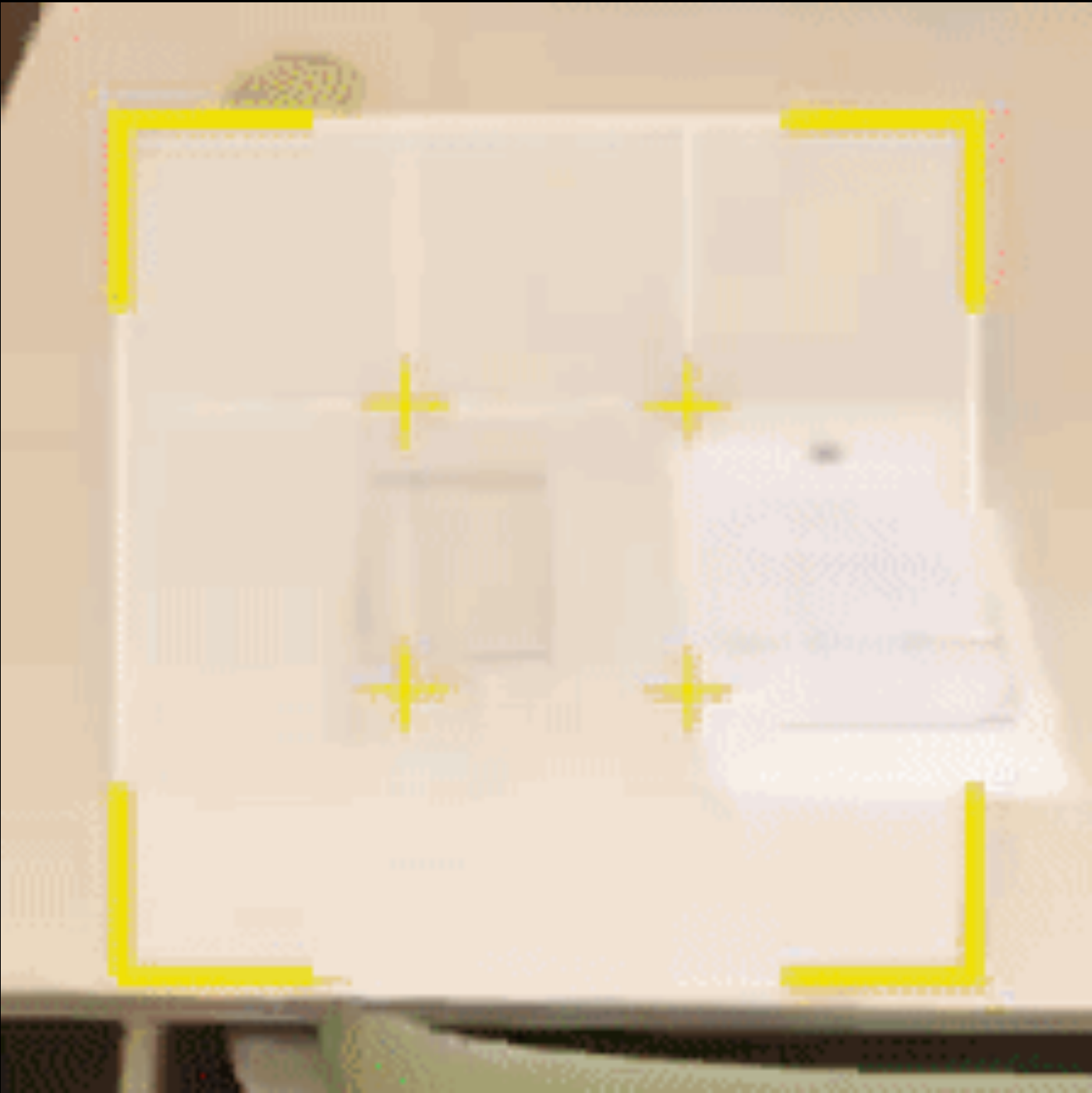
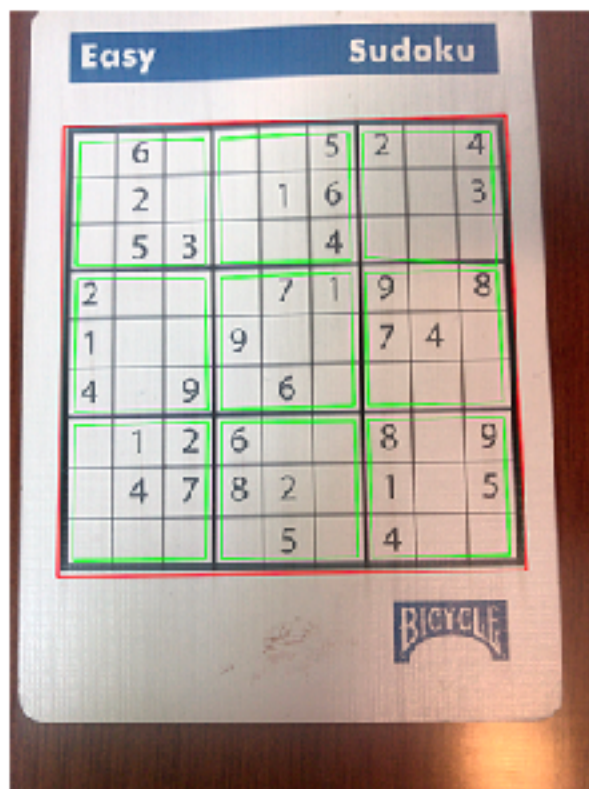


ML + AR

Benzi Ahamed
Engineering
GO-JEK



Behind The Scenes



Vision
Detect Rectangles



Heuristics
Isolate Puzzle Squares

```
+---+---+---+
| 6 | 5|2 4|
| 2 | 16| 3|
| 53| 4|  |
+---+---+---+
|2  | 71|9 8|
|1  |9  |7 4|
|4  |9  |6  |
+---+---+---+
| 12|6  |8 9|
| 47|82 |1 5|
|  | 5 |4  |
+---+---+---+
```

CoreML
Classify Squares

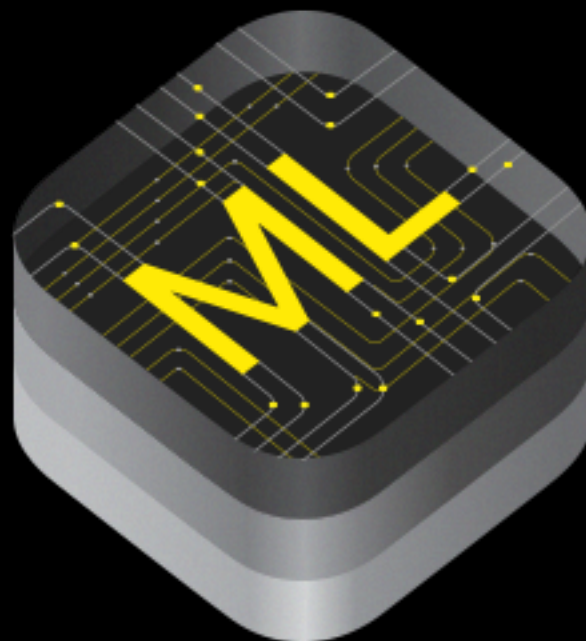
```
+---+---+---+
|761|385|294|
|924|716|583|
|853|294|617|
+---+---+---+
|236|471|958|
|185|932|746|
|479|568|321|
+---+---+---+
|512|643|879|
|347|829|165|
|698|157|432|
+---+---+---+
```

Algorithm
Solve Puzzle



ARKit
Place Into Real World

Machine Learning



Machine Learning is about making decisions based on trial and error.

A more application oriented version of statistics.

Making decisions based on certain amount of information you already have on hand.

Common ML Terms

- **Models**
- **Neural networks**
- **Supervised and unsupervised learning**
- **Common tools and libraries**

Model

Training Data



Algorithm

Neural Networks / Trees and
Jungles / Logistic Regression etc

Test Data



Classifier
= ?



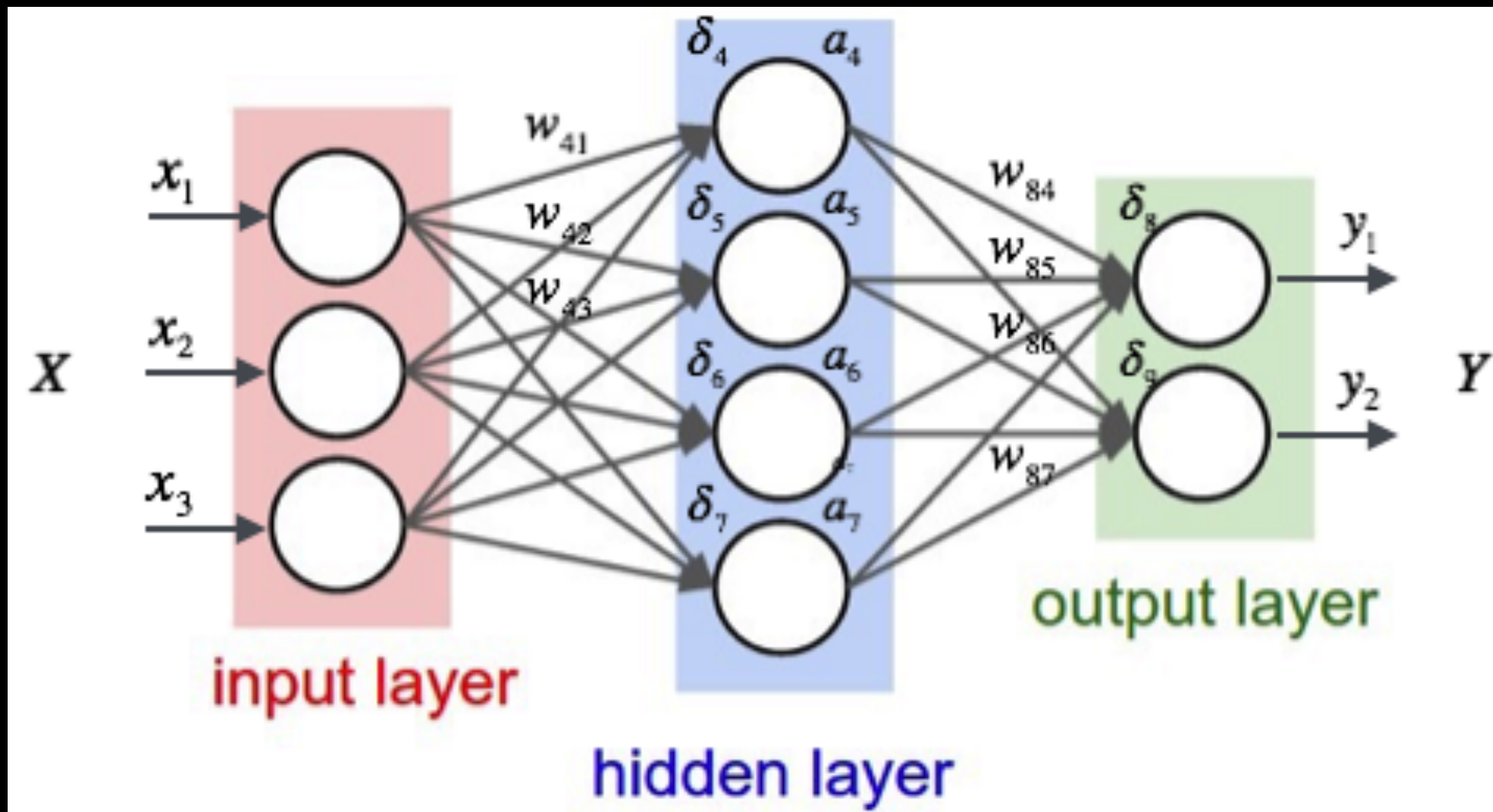
Prediction
= dog



Neural Networks

- Modelled loosely based on the structure of the human brain
- May contain thousands (maybe millions) of processing nodes that are interconnected

Neural Network



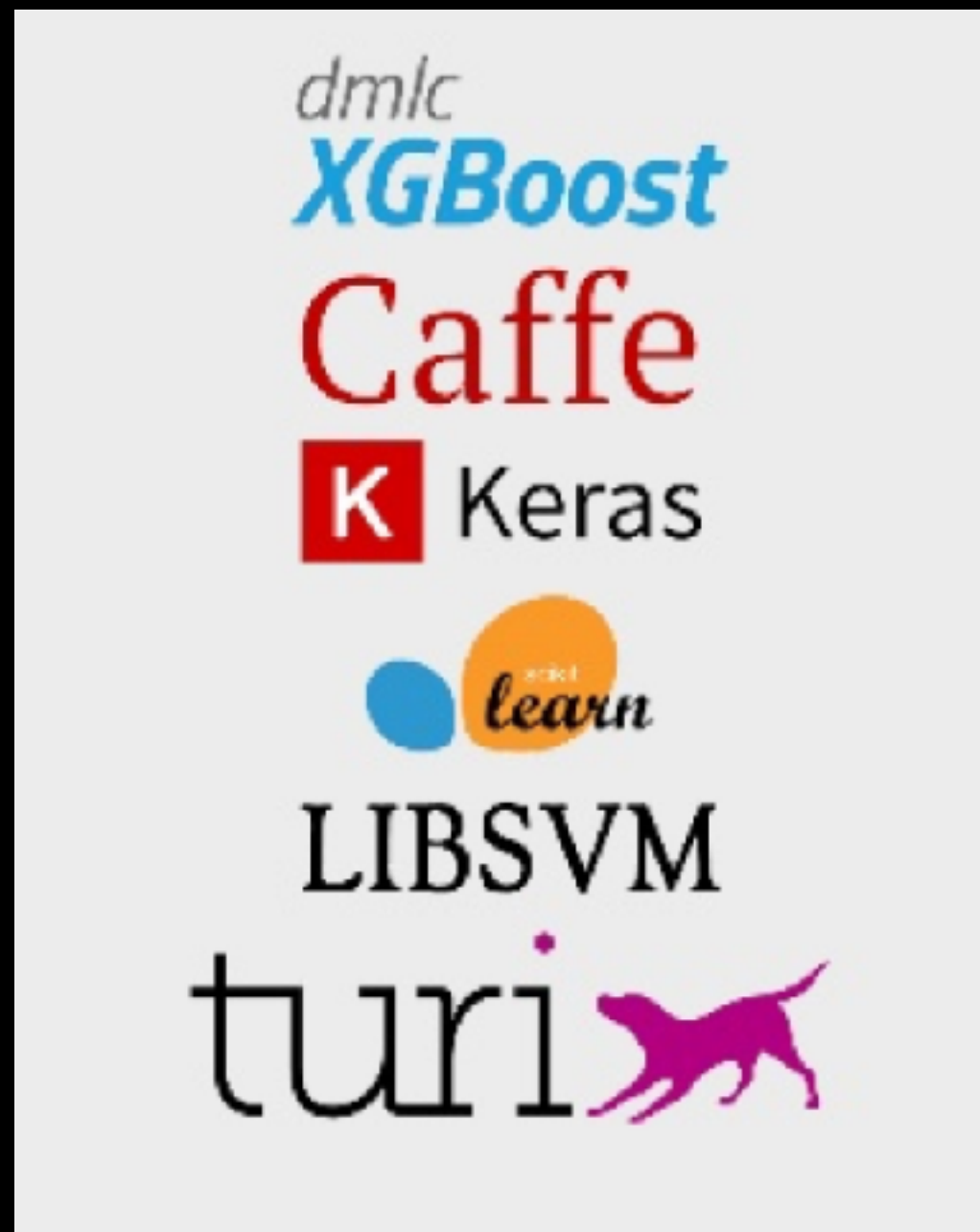
ML is not perfect



What is CoreML?

- A framework that allows us to use Core ML models (an inference engine)
- Built on low level tech (Metal, Accelerate)
- Uses a new .MLModel file
- Ability to convert models from popular libraries

Library Integration



Tech Stack

Your Application

Vision

NLP

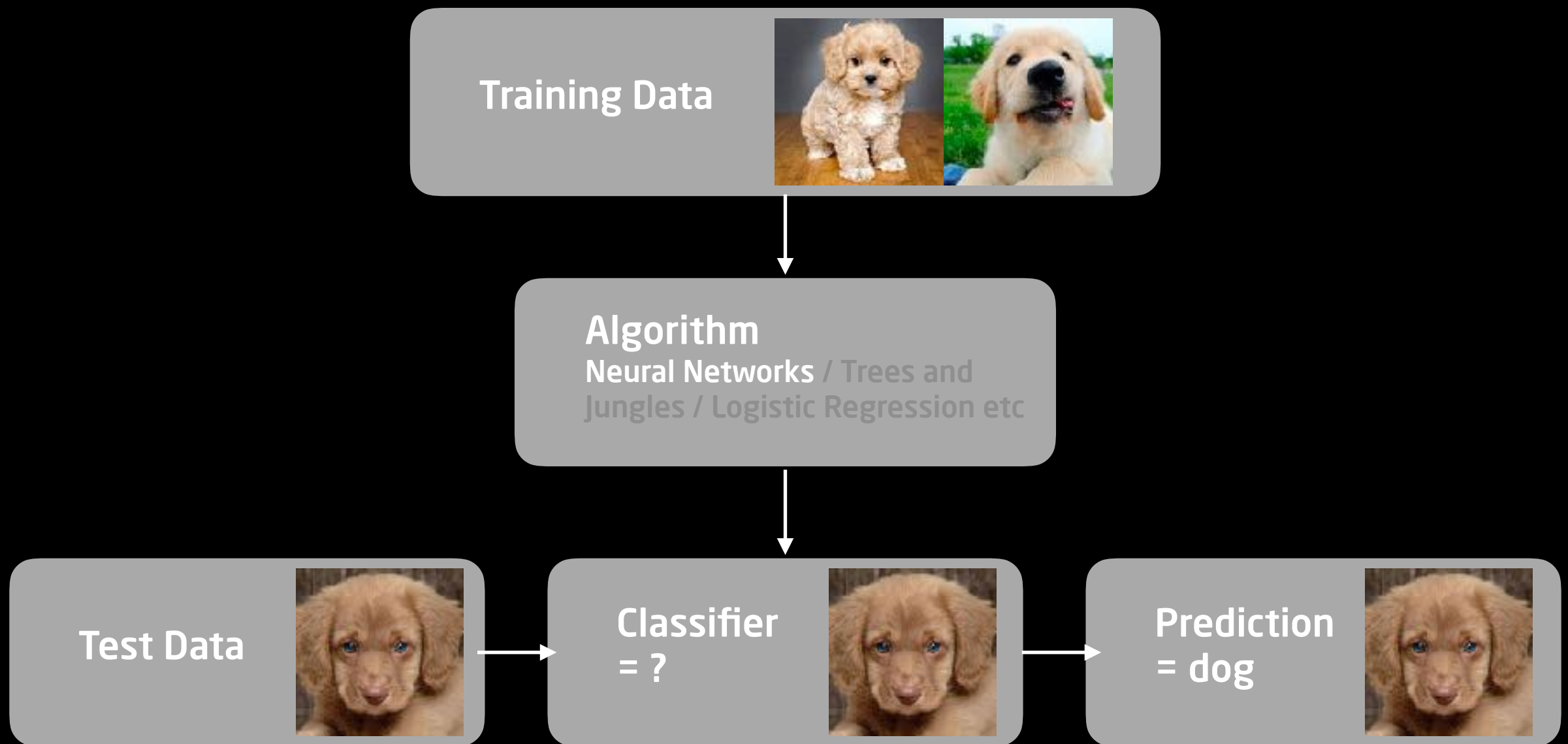
GameplayKit

CoreML

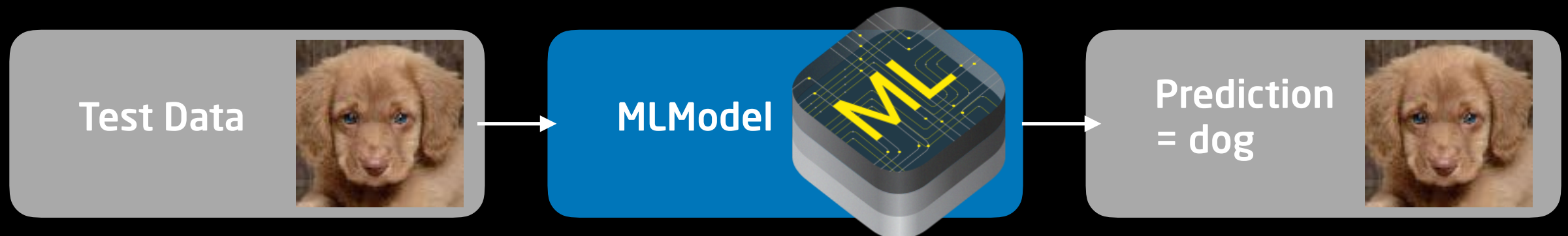
Accelerate, BNNS

Metal Performance Shaders

Models with .MLModel



Models with .MLModel



CoreML



CoreML



- Very little iOS code
- Easy to implement
- Powerful mobile experiences
- ML Tools and .MLModel are OSS

CoreML



- Requires a lot of ML training
- Only supports supervised training
- Starter resources to create new

CoreML models limited

- Every other library is in Python

CoreML



- No on-device model training
- .MLModel file is not encrypted
- Potential for huge file sizes
- No direct TensorFlow support (yet)

Don't Fret

- There are other options to do ML
- Use Accelerate, BNNS and Metal Performance Shaders for low-level / custom stuff
- For specialised cases, use Vision, NLP with Foundation, and GameplayKit for evaluating decision trees
- Other OS alternatives exists

Vision over CoreML

- Face detection, landmark identification (eyes, nose etc)
- Machine learning image analysis (custom feature extraction)
- Barcode detection
- Image alignment analysis (scanning documents)
- Text detection (OCR, card numbers, text)
- Horizon detection (auto correct photo orientation)
- Object detection (rectangle tracking, feature detection)

Augmented Reality



**How is mobile AR
different?**

Scale

AR on Mobile

- User base - 200M users (PSP has 0.9M)
- Entry cost - 40k INR

ARKit

- **Inside out world tracking**
- **Plane detection**
- **Face tracking**
TrueDepth camera iPhone X
- **Light estimation**
Colour correction, realtime shadows
- **Rendering**

World Tracking

- Visual Inertial Odometry (VIO)
- Accelerometer + Gyroscope 1000 Hz
- Camera 30 Hz
- 6dof (XYZ in 3D + Roll, Pitch and Yaw)
- Anchors

Plane Detection

- Horizontal plane detection
- Based on feature point extraction
- Allows you to place 3D content
- Vertical plane (not yet)

Light Estimation

- Determine ambient light intensity (lumens)
- Determine ambient colour temperature (degrees Kelvin)
- Can be in shading algorithms to account for real world conditions
- Automatically integrated with SceneKit

Rendering



unity



UNREAL
ENGINE

**ARKit is a real time
system**

- **Some experience in writing games**
- **Lots of experience with 3D math**
- **Understanding of the frame timing and the core update loop**
- **API provided by ARKit is pretty straightforward**
- **You need to know your rendering framework in and out**

Lifecycle

- You need to start an ARSession with an ARConfiguration (e.g. world tracking + plane detection)
- Choose a rendering tech ARSCNView / ARSKView
- Handling ARKit engine callbacks (plane detected, low quality of settings etc.)
- Hit rays, handling inputs, and placing content

Resources

Resources



Thanks

- Benzi Ahamed
- Product Engineer, GO-JEK
- benzi.a@go-jek.com