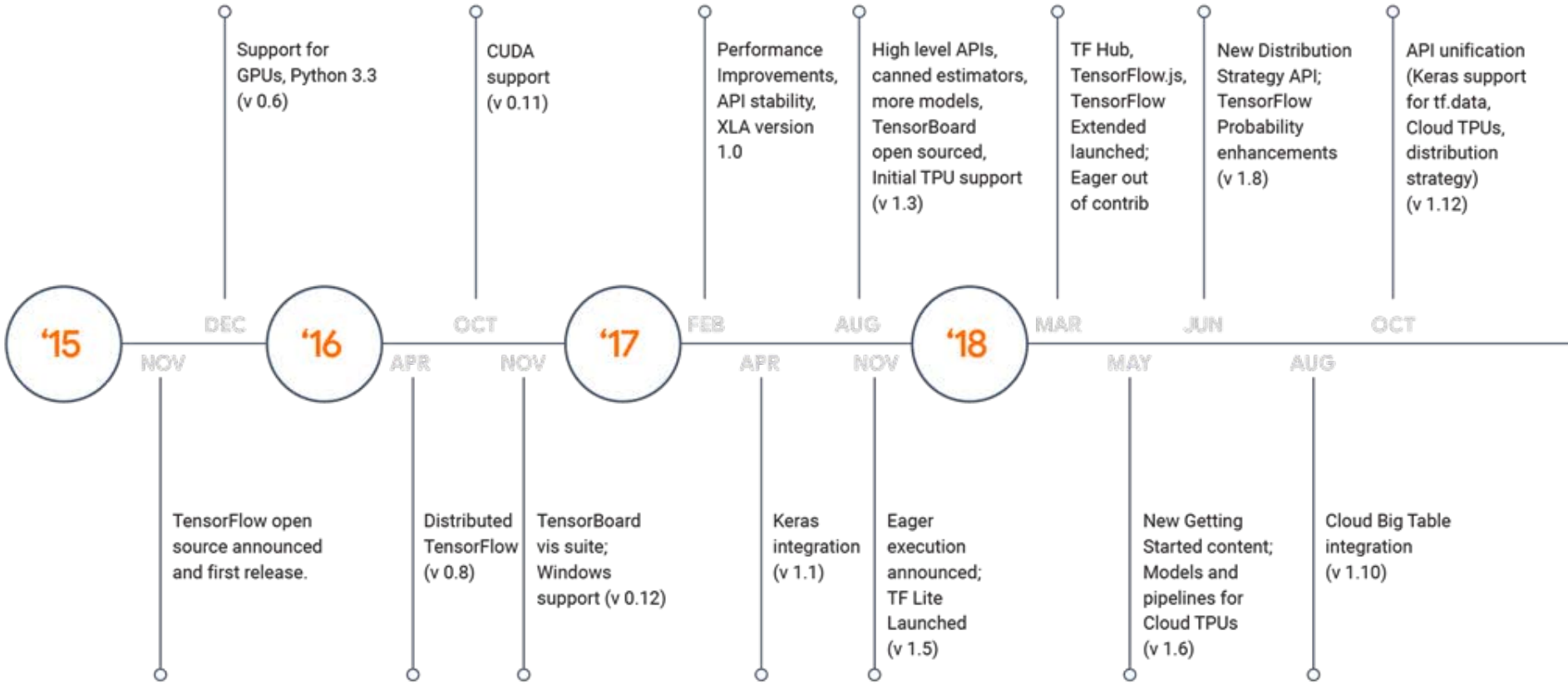
TF 2019 Summit Recap

for Tianjin GDG

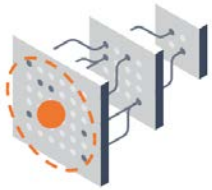
linsong


TensorFlow

TensorFlow


'15
NOV
TensorFlow open source announced and first release.
DEC
Support for GPUs, Python 3.3 (v 0.6)
'16
APR
Distributed TensorFlow (v 0.8)
OCT
CUDA support (v 0.11)
NOV
TensorBoard vis suite; Windows support (v 0.12)
'17
FEB
Performance Improvements, API stability, XLA version 1.0
APR
Keras integration (v 1.1)
AUG
High level APIs, canned estimators, more models, TensorBoard open sourced, Initial TPU support (v 1.3)
NOV
Eager execution announced; TF Lite Launched (v 1.5)
'18
MAR
TF Hub, TensorFlow.js, TensorFlow Extended launched; Eager out of contrib
MAY
New Getting Started content; Models and pipelines for Cloud TPUs (v 1.6)
JUN
New Distribution Strategy API; TensorFlow Probability enhancements (v 1.8)
AUG
Cloud Big Table integration (v 1.10)
OCT
API unification (Keras support for tf.data, Cloud TPUs, distribution strategy) (v 1.12)

TensorFlow 2.0



Easy

Simplified APIs.
Focused on Keras and eager execution

Powerful

Flexibility and performance.
Power to do cutting edge research and scale to > 1 exaflops

Scalable

Tested at Google-scale.
Deploy everywhere

TF Probability

TF Agents

Tensor2Tensor

TF Ranking

TF Text

TF Federated

TF Privacy

...

Deploy anywhere

Servers

TensorFlow
Extended

Edge devices

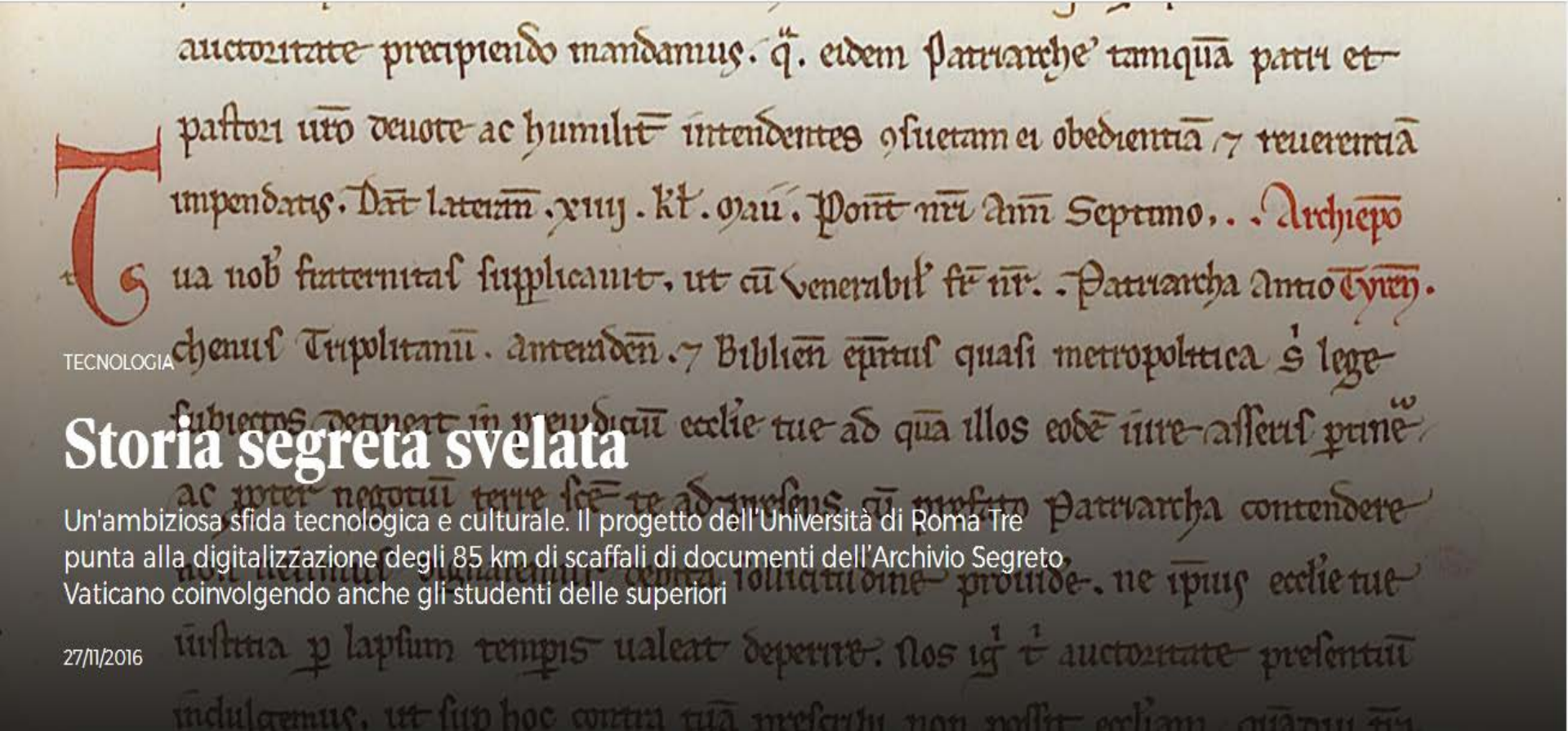
TensorFlow
Lite

JavaScript

TensorFlow
.JS

Content Today

- In Codice Ratio
- Reinforcement Learning
- TF Probability
- Airbnb
- 轻量化

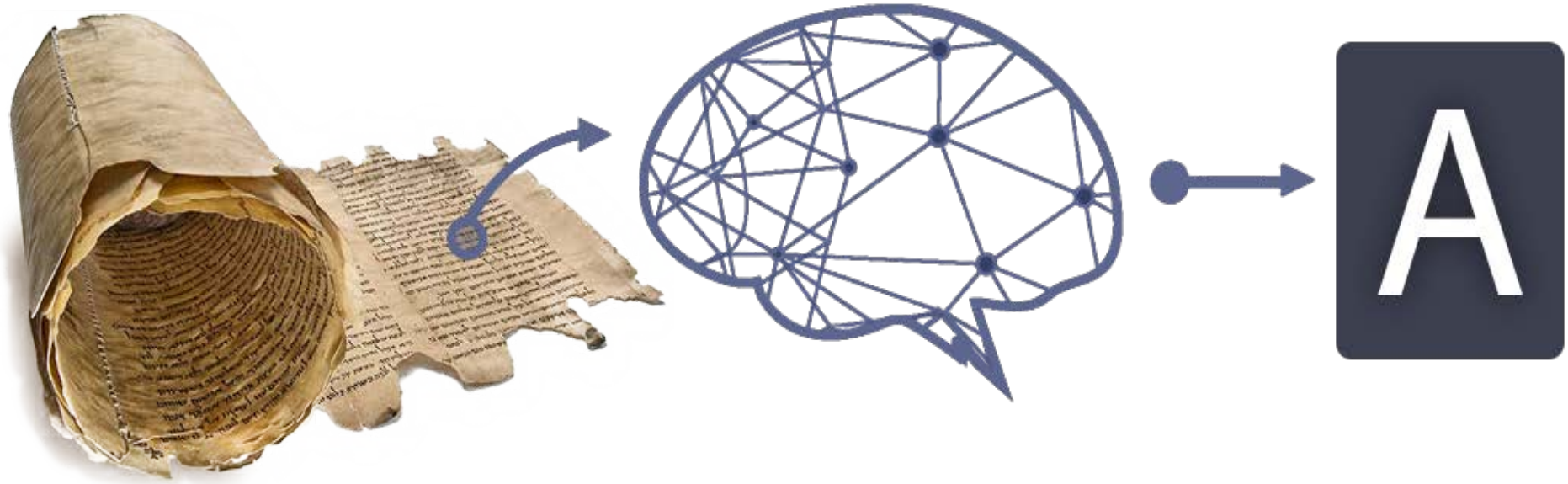


TECNOLOGIA

Storia segreta svelata

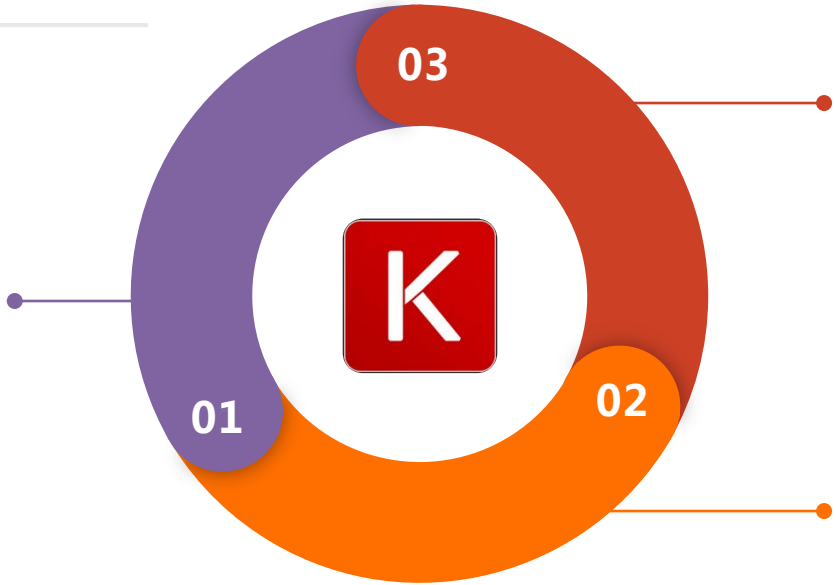
Un'ambiziosa sfida tecnologica e culturale. Il progetto dell'Università di Roma Tre punta alla digitalizzazione degli 85 km di scaffali di documenti dell'Archivio Segreto Vaticano coinvolgendo anche gli studenti delle superiori

27/11/2016


A

Prototyping with Keras

Implement incrementally



Train

Fast iterations using NVIDIA GPUs

Tune

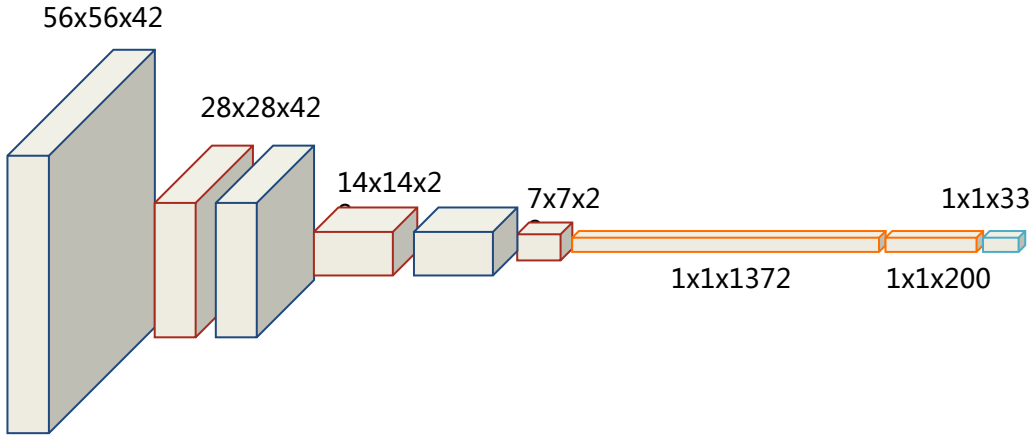
Search for best hyperparameters, increase model complexity, choose between optimizers, regularizers, initializers

Evaluate

K-Fold Cross-validation using 10% of dataset as test set

Final model

A simple, custom CNN



- Optimizer: Adam
- Initializer: Glorot normal
- Regularizers:
  - Dropout = 0.5
  - L2, λ = 0.001

94% avg accuracy

Convolution+ReLU

Max Pooling

Fully Connected+ReLU

Fully Connected+Softmax

Overview of the process

Input Image

High resolution picture of a manuscript page

Oversegmentation

Page is segmented into words
Words are segmented into strokes

Classification

CNN classifies combination of strokes (or discards them)

33 character classes
~1100 samples per class

Ranking

According to a statistical Latin Language Model

14M words from 716 publicly available Latin texts

Transcription(s)

We evaluate top-1 and top-3 word accuracy

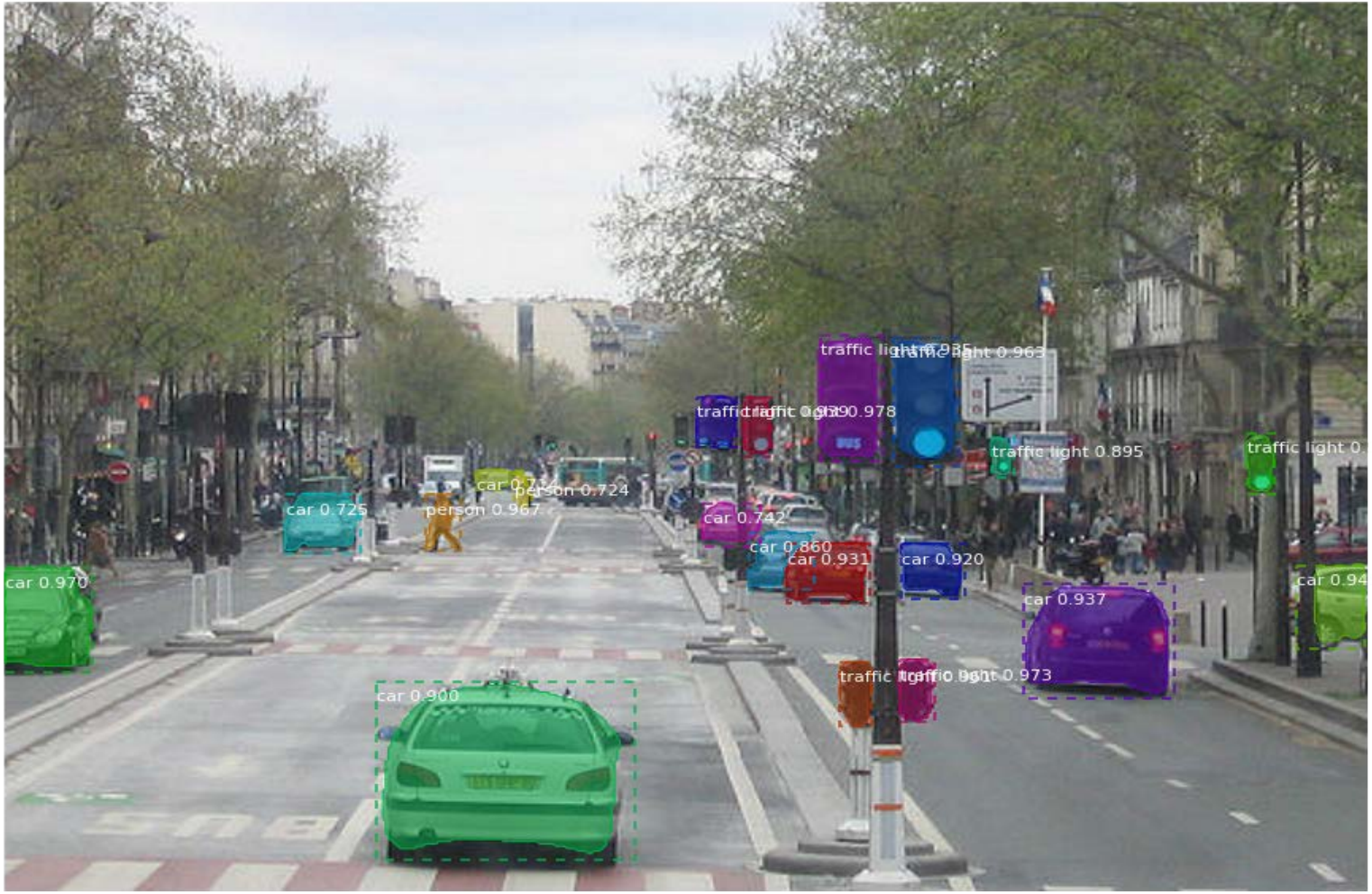
接下来做的

- Segmentation
- GAN
- Sequence

Semantic Segmentation

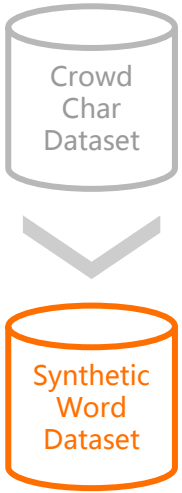
Input:

Prediction:


traffic light 0.985
traffic light 0.963
traffic light 0.978
traffic light 0.895
car 0.725
person 0.724
person 0.967
car 0.742
car 0.860
car 0.931
car 0.920
car 0.970
car 0.937
car 0.900
traffic light 0.973

GANs and Autoencoders

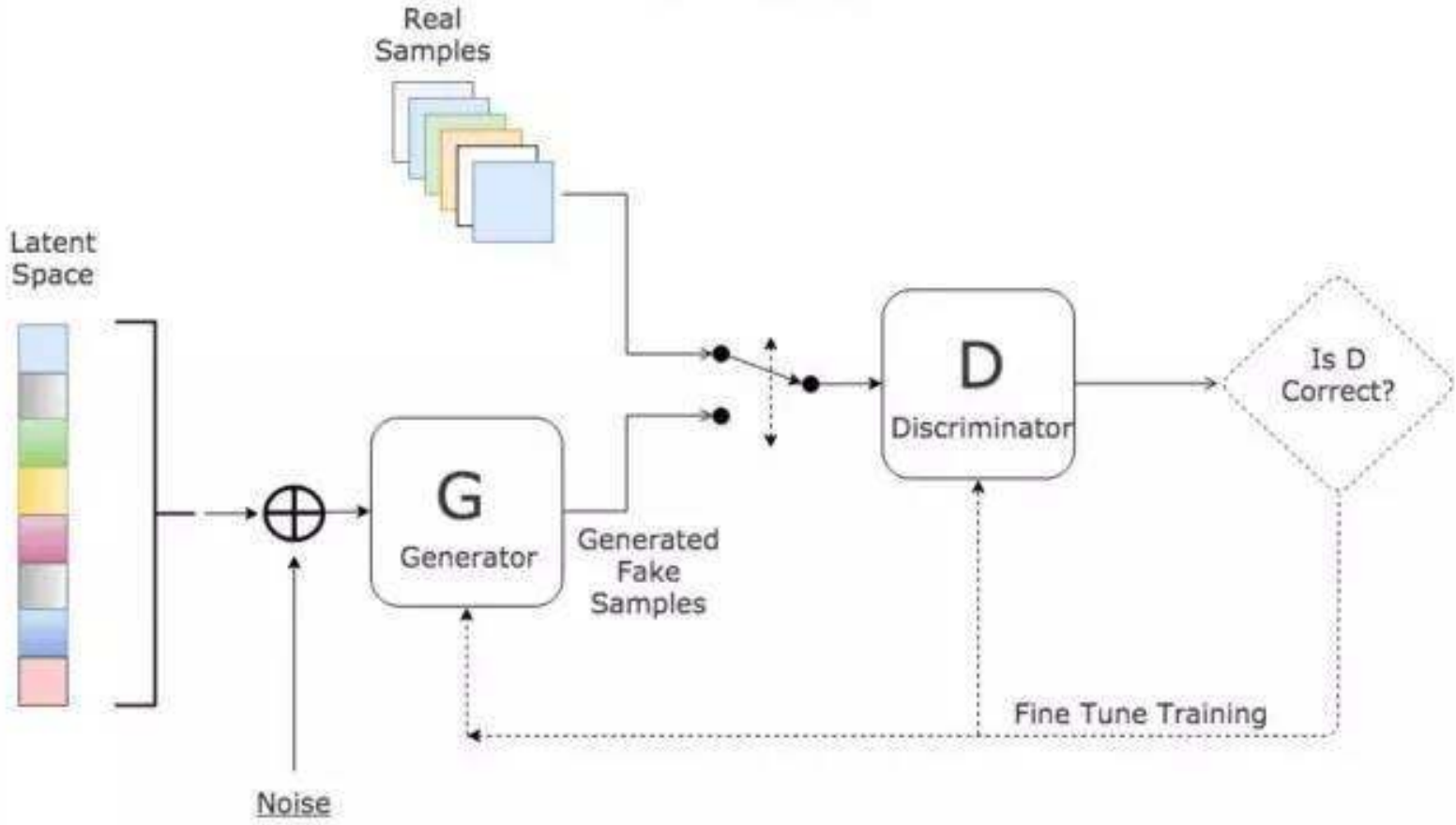
For synthetic dataset generation and embedding

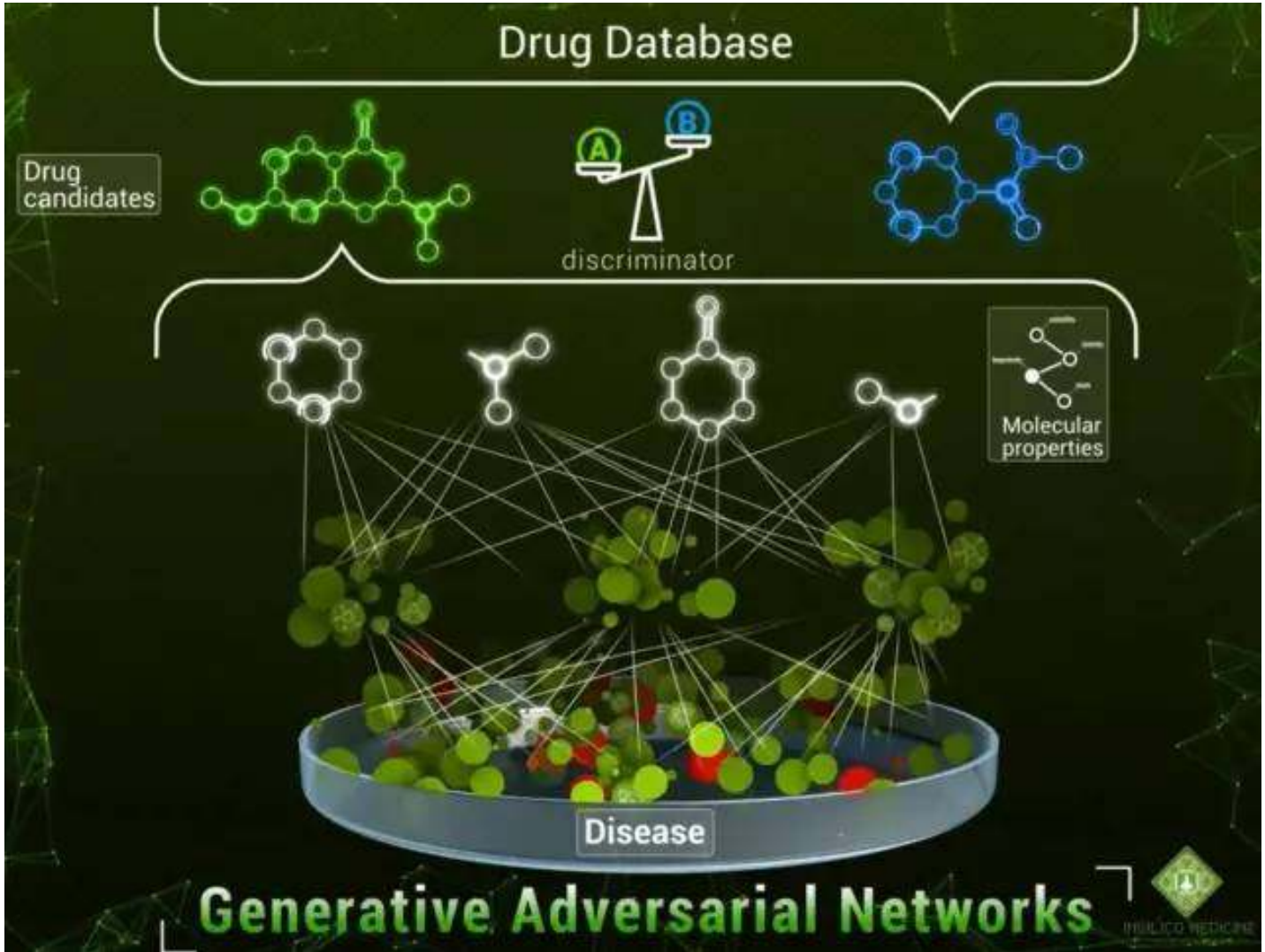


Input:

Generated:

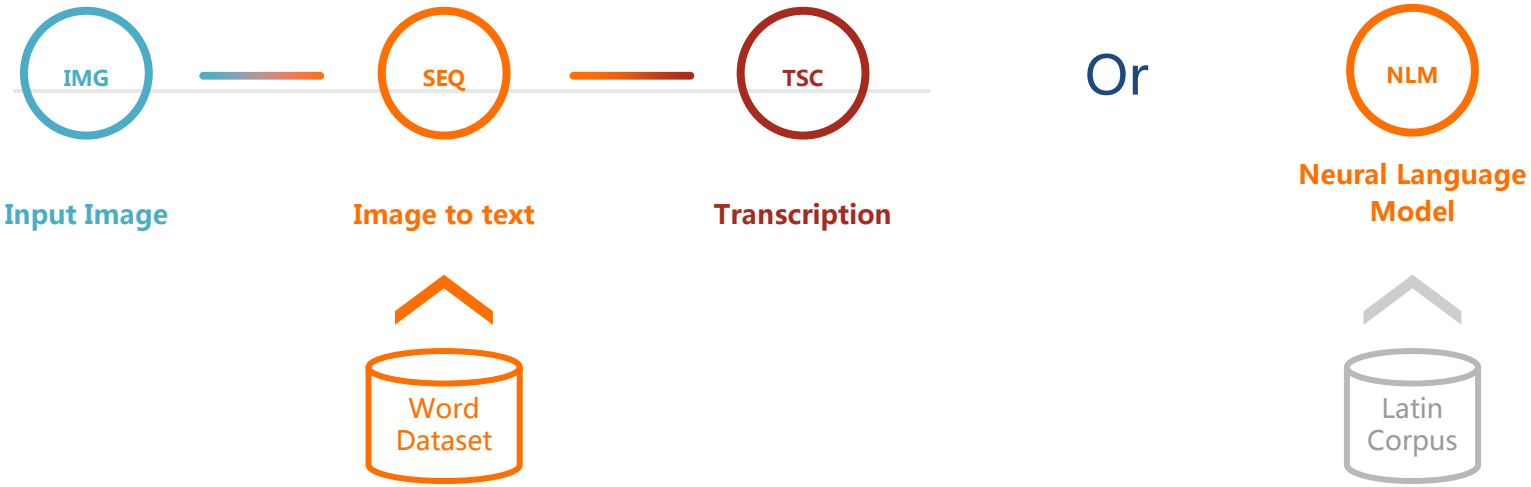
Generative Adversarial Network




Drug Database
Drug candidates
A
B
discriminator
Molecular properties
Disease
Generative Adversarial Networks

Sequence models

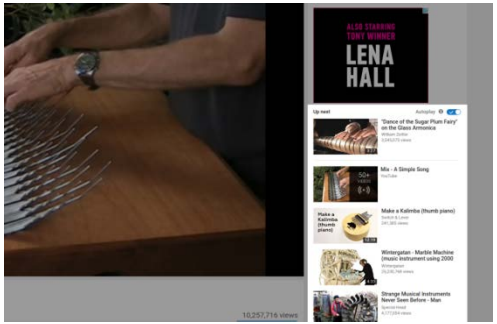
Neural LMs and segmentation free transcription

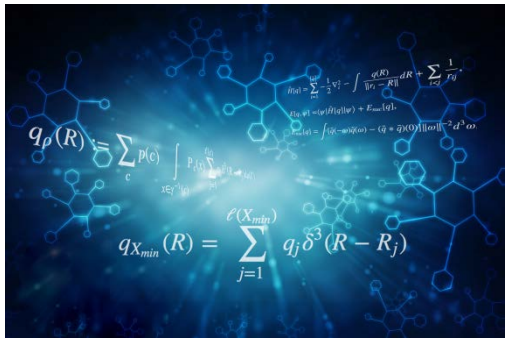


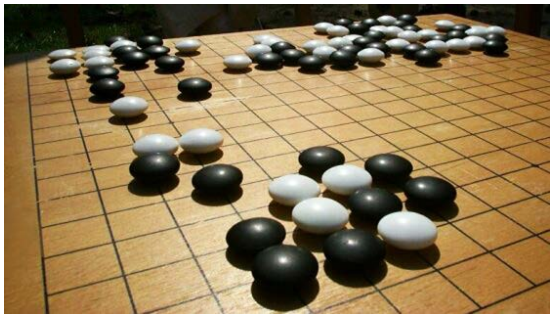
外面在下雨你别忘了带__。

校长：校服上除了校徽别别别的，让你别别别的别别别的你非得别别的！

Reinforcement Learning







Reinforcement Learning

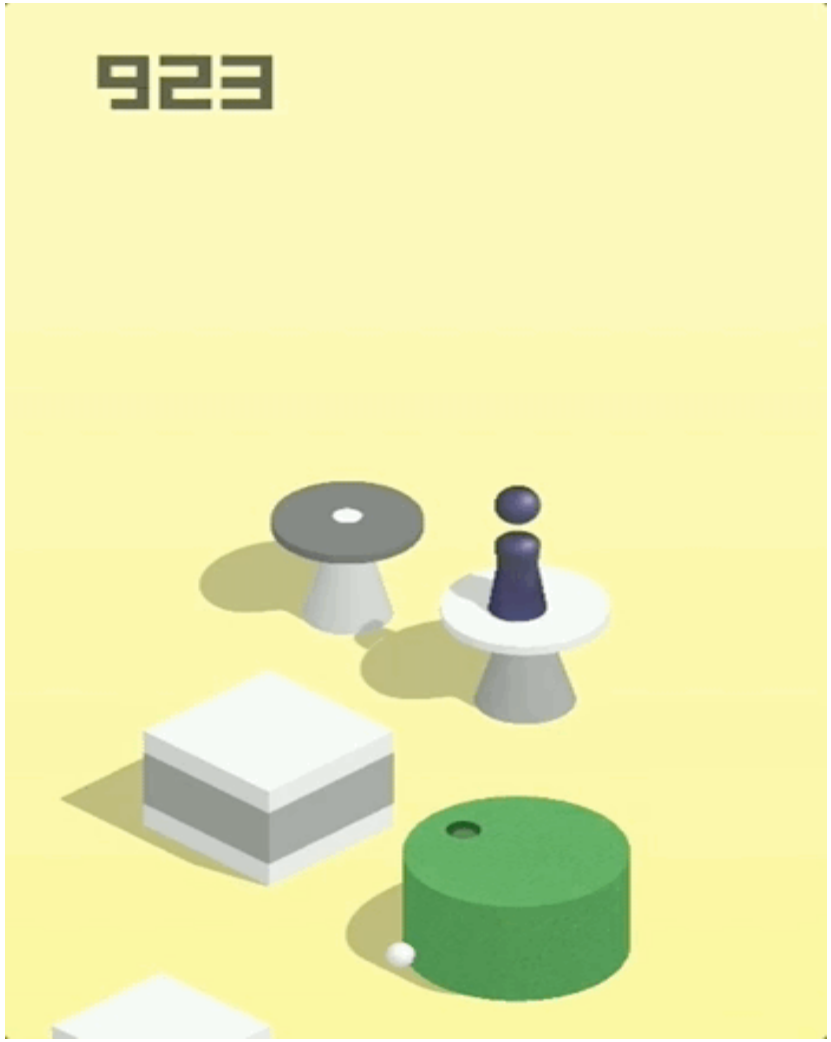
Environment

Agent

(1) Observation

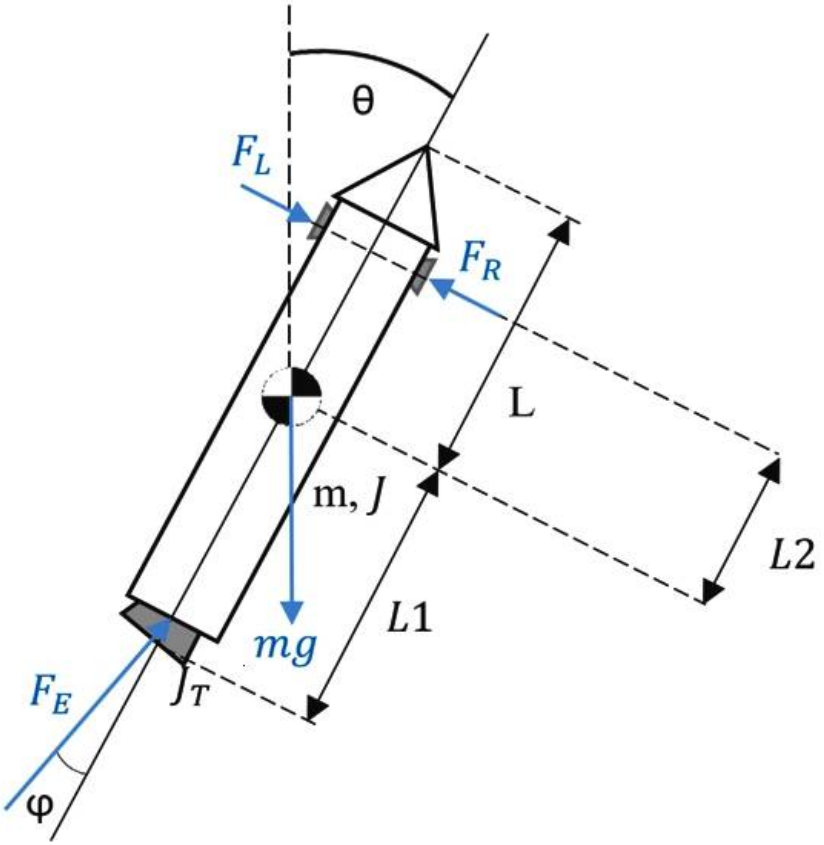
(2) Action

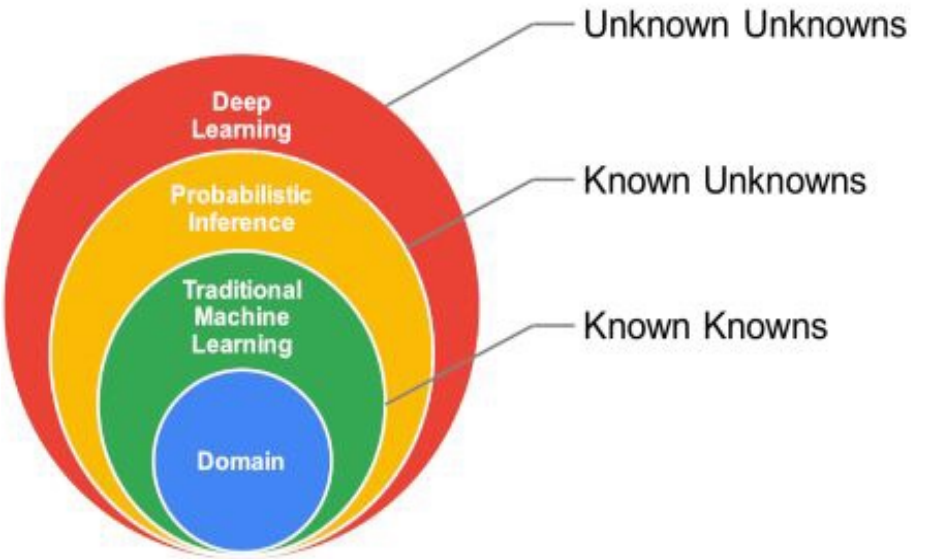
(3) Reward

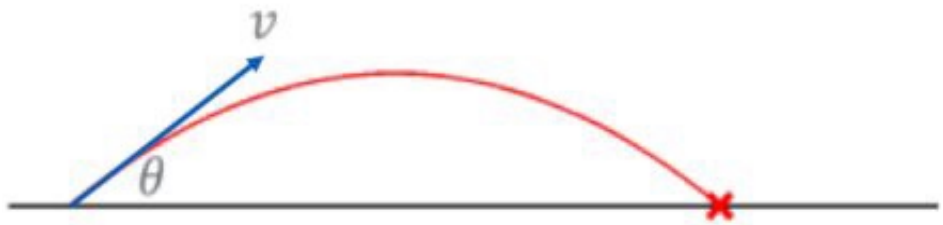

923

TF-Agents is easy to use

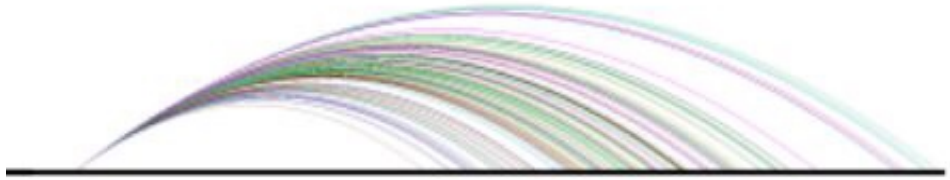
- https://github.com/tensorflow/agents
- Build for TF 2.0:
  - Develop and debug quickly with **TF-Eager.**
  - Use **tf.keras** to define your Networks.
  - Use **tf.function** to speed everything up.
  - **Modular** and **extensible**.
- Compatible with TF 1.14:
  - For those not ready to make the change.


θ
F_L
F_R
L
m, J
$L2$
$L1$
mg
F_E
J_T
φ





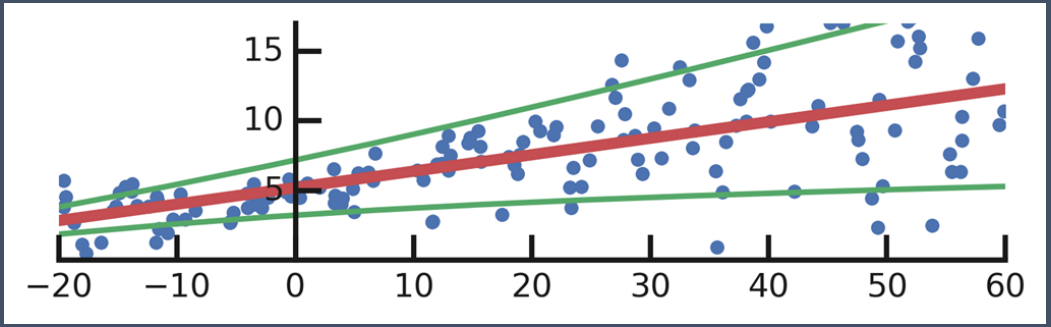
Wind Speed Uncertainty



Learn known unknowns.

```
model = tf.keras.Sequential([
  tf.keras.layers.Dense(hidden_units, ...),
  tf.keras.layers.Dense(1+1),
  tfp.layers.DistributionLambda(lambda t:
    tfd.Normal(loc=t[..., 0],
               scale=tf.softplus(t[..., 1]))),
])
```

} Linear Regression (Heteroskedastic)

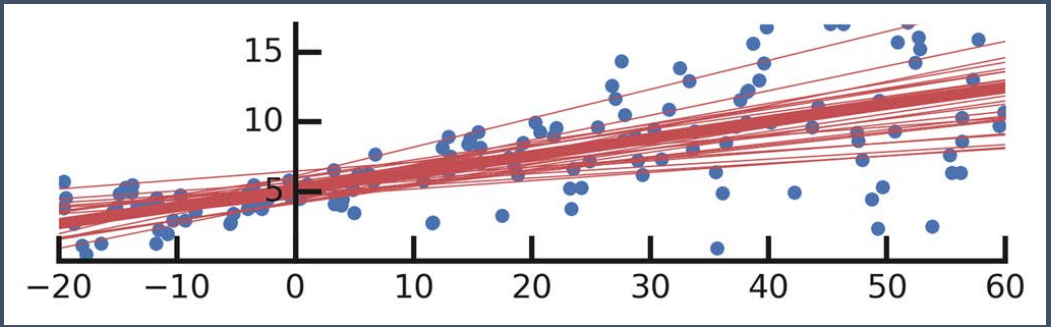


Learn unknown unknowns.

```
model = tf.keras.Sequential([
  tfp.layers.DenseVariational(hidden_units, ...),
  tfp.layers.DenseVariational(1),
  tfp.layers.DistributionLambda(lambda t:
    tfd.Normal(loc=t[..., 0],
               scale=1)),
])
```

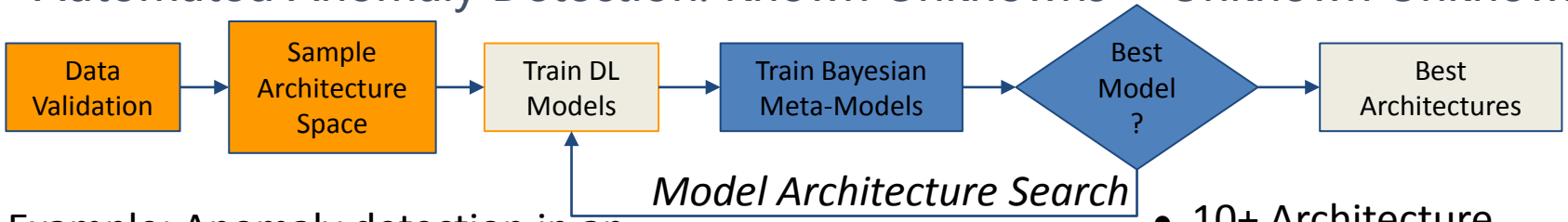
} "Bayesian Weights"

} Linear Regression

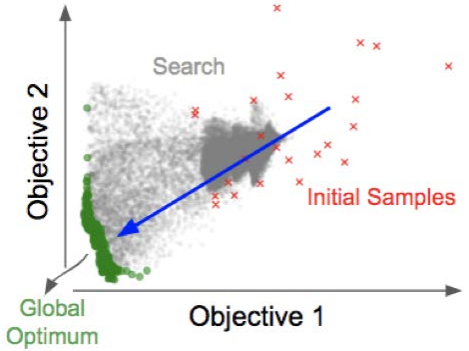


A Success Story: BHGE

Automated Anomaly Detection: Known Unknowns + Unknown Unknowns

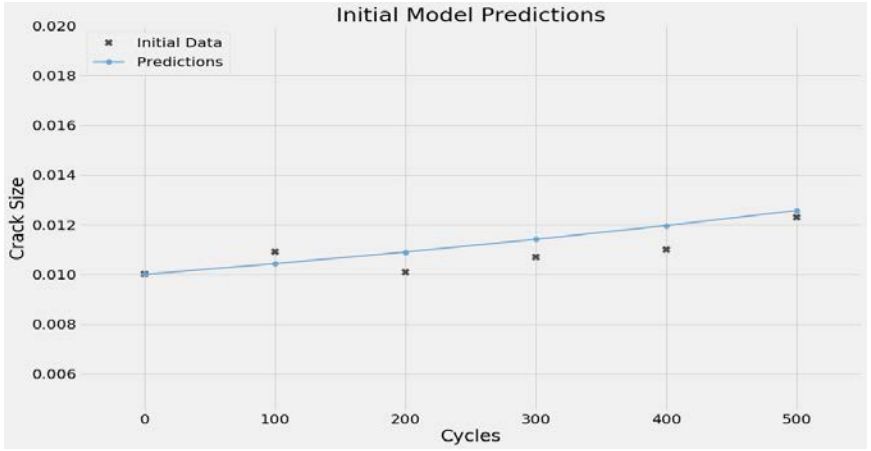


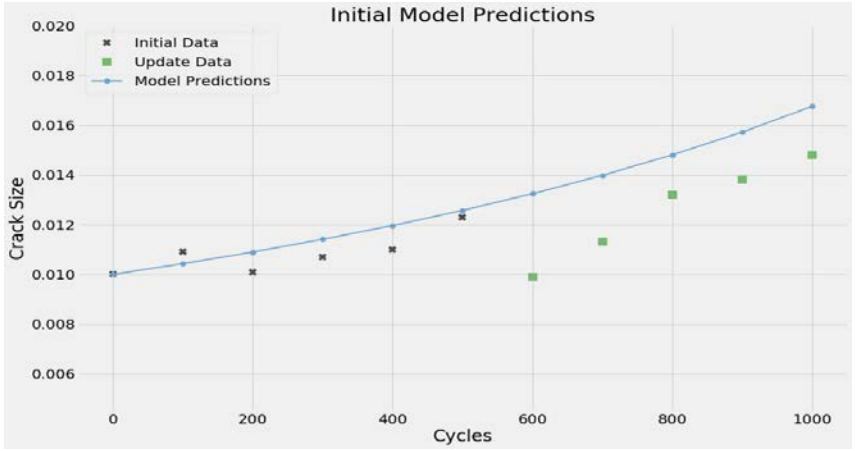
- Example: Anomaly detection in an industrial system:
  - Large system with 13 subsystems
  - 6TB - 2 year's operational data
  - **50% reduction in false positives, 300% reduction in false negatives**

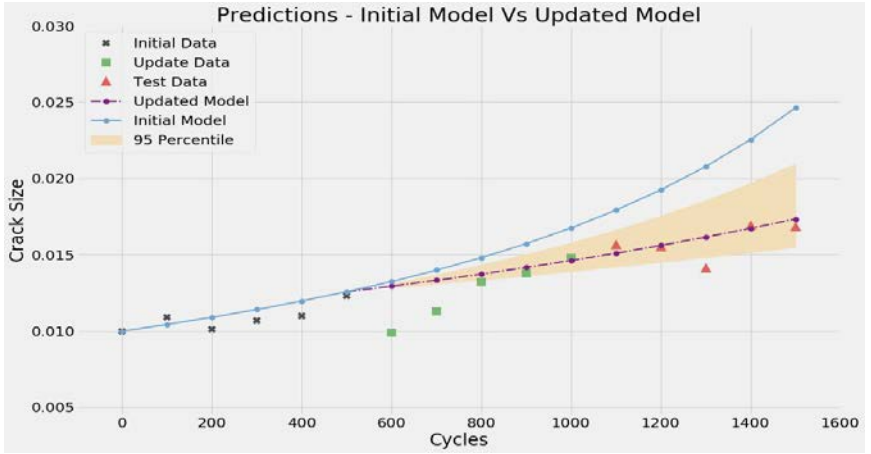


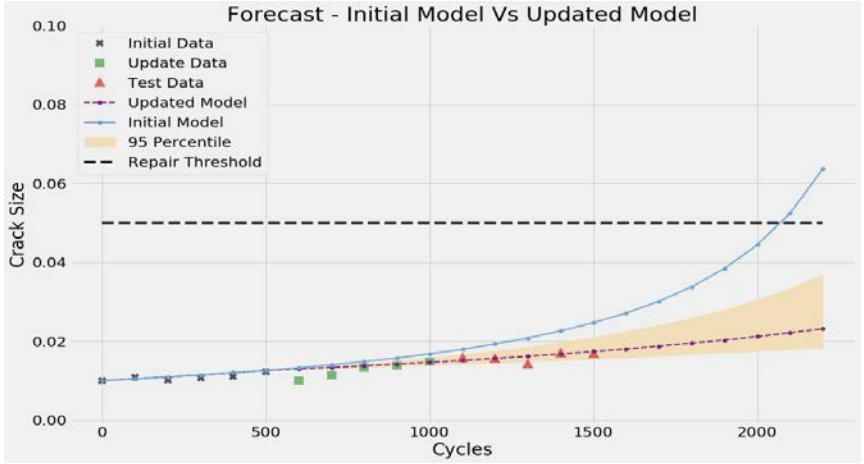
- 10+ Architecture parameters searched
- **0.25 Million** architectures searched for each model
- **10X improvement** in model/training performance over random search

- Probabilistic Deep Learning model for anomaly detection built and deployed in production with TFP
- Step change in performance realized on multiple applications under active deployment


Initial Model Predictions
Initial Data
Predictions
Crack Size
Cycles
0.020
0.018
0.016
0.014
0.012
0.010
0.008
0.006
0
100
200
300
400
500


Initial Model Predictions
Initial Data
Update Data
Model Predictions
Crack Size
Cycles
0.020
0.018
0.016
0.014
0.012
0.010
0.008
0.006
0
200
400
600
800
1000

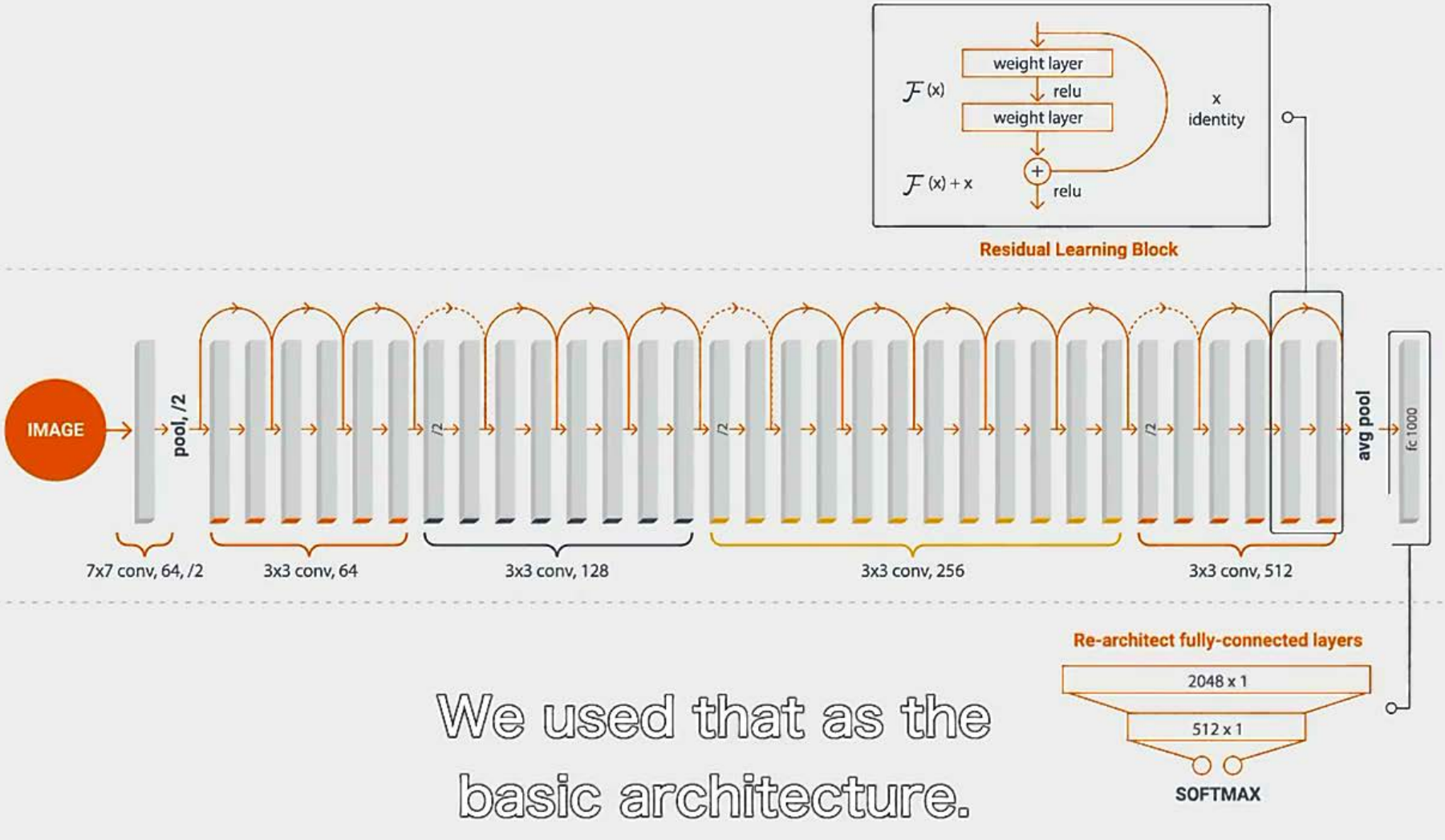

Predictions - Initial Model Vs Updated Model
Initial Data
Update Data
Test Data
Updated Model
Initial Model
95 Percentile
Crack Size
Cycles
0.030
0.025
0.020
0.015
0.010
0.005
0
200
400
600
800
1000
1200
1400
1600


Forecast - Initial Model Vs Updated Model
Initial Data
Update Data
Test Data
Updated Model
Initial Model
95 Percentile
Repair Threshold
Crack Size
Cycles
0.10
0.08
0.06
0.04
0.02
0.00
0
500
1000
1500
2000





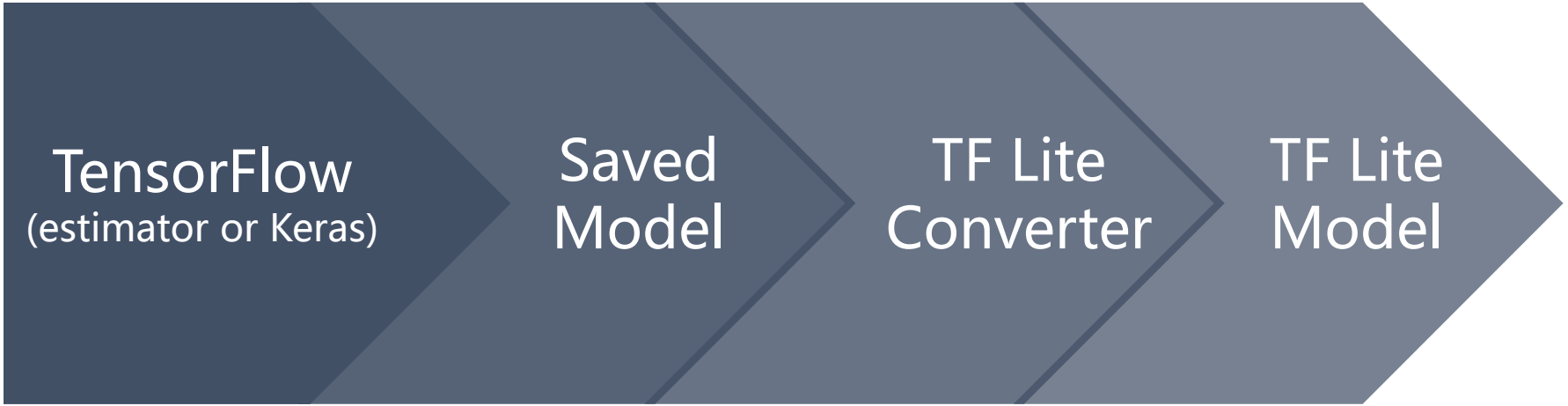

that they see aren't just
a picture of the garage



We used that as the basic architecture.

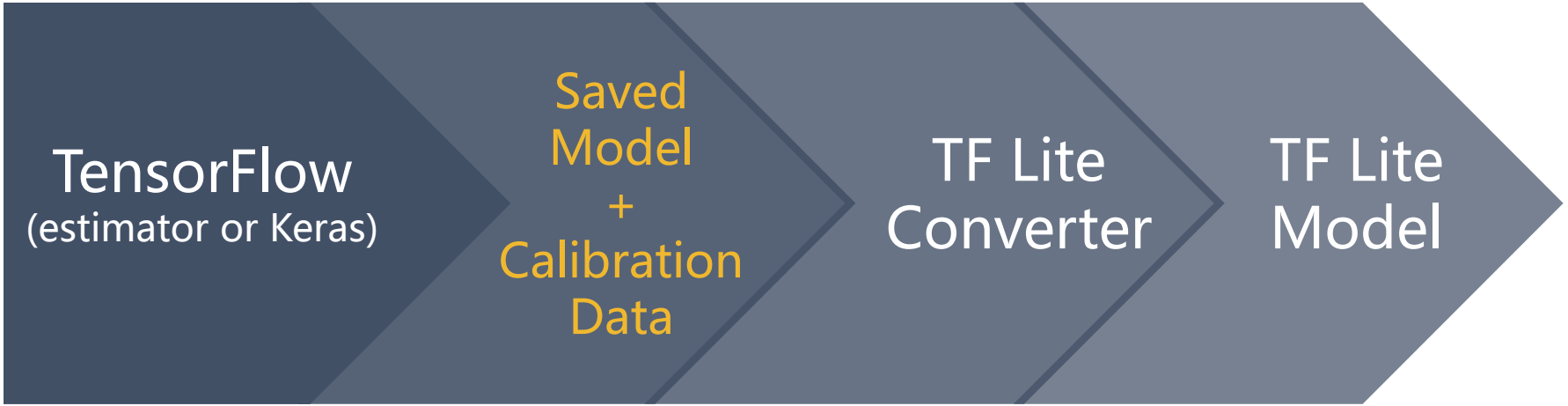
Model conversion

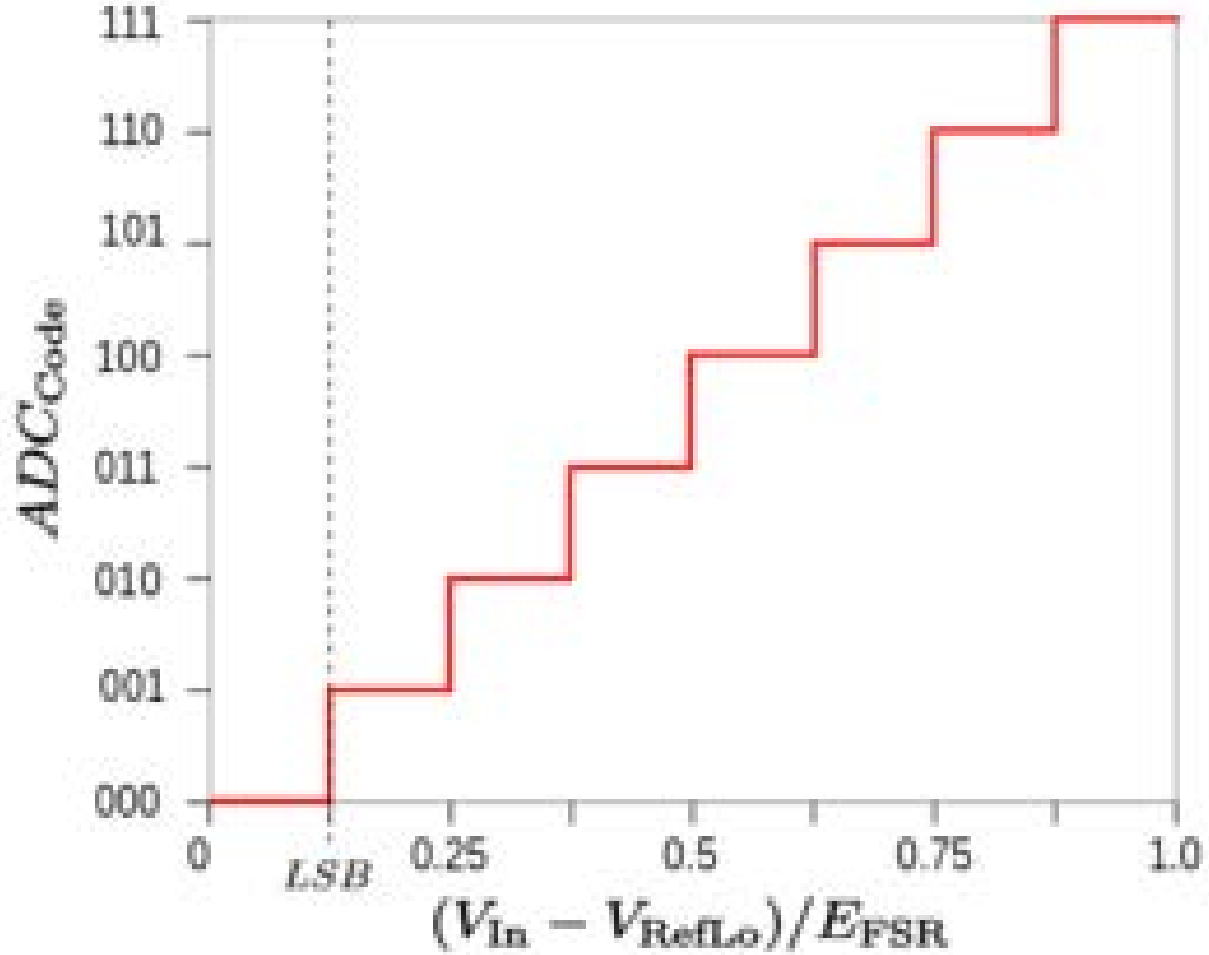
The conversion flow to TensorFlow Lite is simple ...

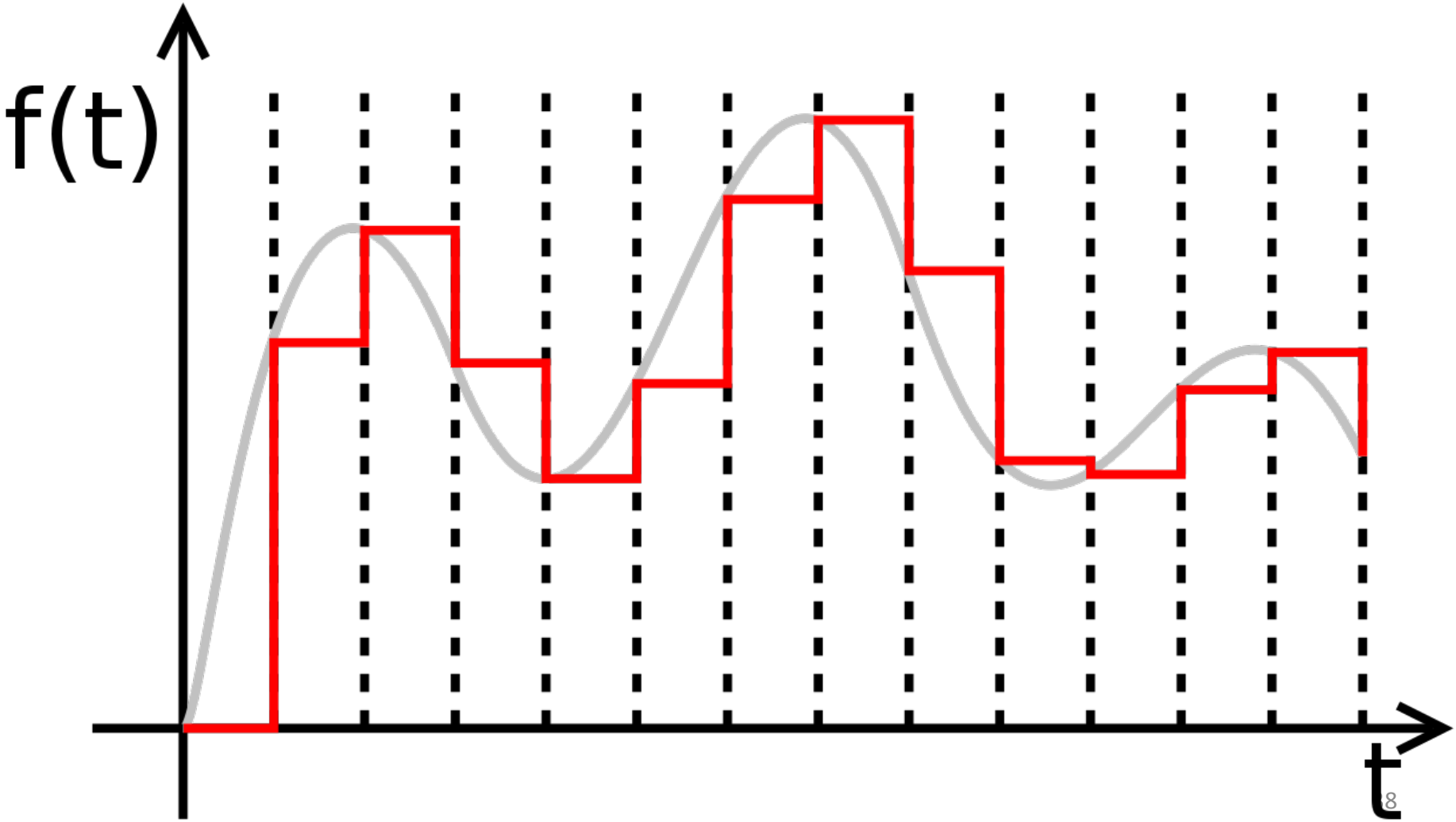


Optimization

Quantization (post-training)




111
110
101
100
011
010
001
000
ADC_{Code}
0
LSB
0.25
0.5
0.75
1.0
$(V_{In} - V_{RefLo})/E_{FSR}$


f(t)
t

Optimization

Quantization results: training vs post-training

Top 1 accuracy

Model	Float baseline	Quantization during training	Quantization after training
Mobilenet v1	70.95%	69.97%	69.54%
Resnet v2	76.8%	76.7%	76.6%
Inception v3	77.9%	77.5%	77.7%

New Courses

Introduction to TensorFlow
for AI, ML and DL

coursera.org/learn/introduction-tensorflow

Intro to TensorFlow
for Deep Learning

udacity.com/tensorflow

Why is Andrew Ng's Coursera Machine Learning course so ...

https://www.quora.com/Why-is-**Andrew**-**Ng**s-Coursera-Machine-Learning...

The math is tame, and **Ng** is very reassuring to those who **don't** follow the math. (He frequently says stuff like "**don't worry** if you **don't** understand this, there are good libraries to solve these problems".) It's available for free to anyone with an internet connection. Recommendations.

tf.thanks!