ICML @ NYC

International Conference on Machine Learning

JUNE 19-24 2016 NEW YORK

CONFERENCE REPORT

MARCIN ZADROGA

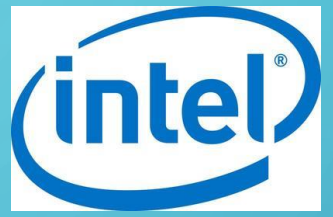
CONFERENCE

- 33rd edition
- 19–24 June 2016 (6 days, sic!)
- Marriot Marquis Hotel on Time Square
- 3000 participants, 1 day of tutorials (9), 3 days of conference (300+ accepted papers), 2 days of workshops (23)

BIG TOPICS ON THE CONFERENCE

- Neural Networks & Deep Learning
- (Deep) Reinforcement Learning
- Speech Recognition
  - (Google) “Towards End-To-End Speech Recognition with Recurrent Neural Networks”
  - (Baidu) “Deep Speech 2: End-to-End Speech Recognition in English and Mandarin”
- Alpha GO
- Machine Learning Systems and Old Ideas revision
- Other: Kernel Methods, Gaussian Processes, Time-Series Analysis, Large Scale Learning

(SOME) COMPANIES PRESENT AT ICML

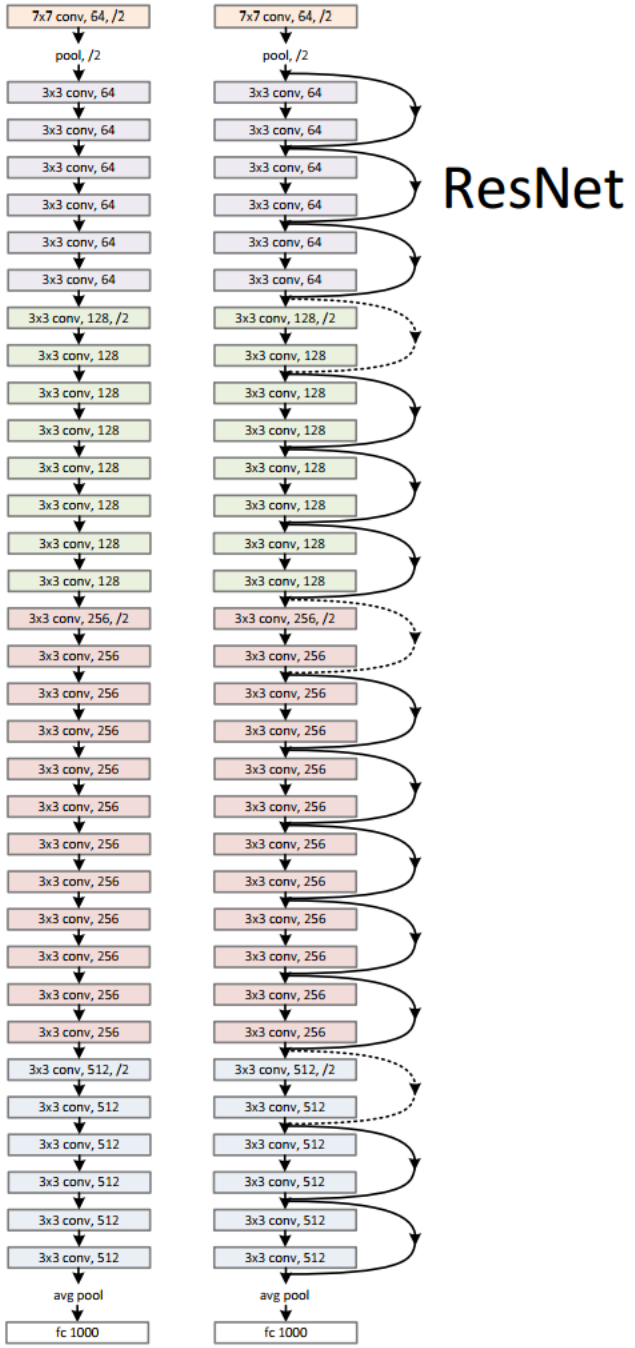


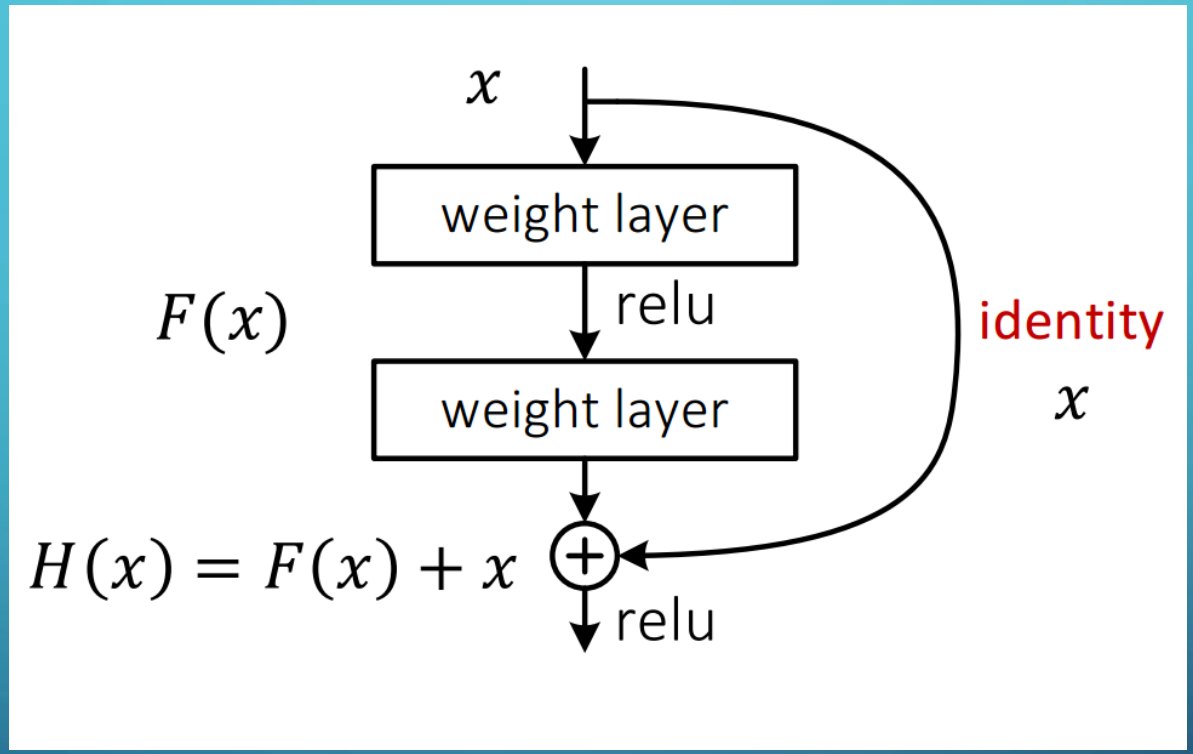
TUTORIALS – HIGHLIGHTS

- Deep Learning
  - https://github.com/KaimingHe/deep-residual-networks
  - http://icml.cc/2016/tutorials/icml2016_tutorial_deep_residual_networks_kaiminghe.pdf
- Reinforcement Learning
  - ALPHA GO
  - http://icml.cc/2016/tutorials/deep_rl_tutorial.pdf

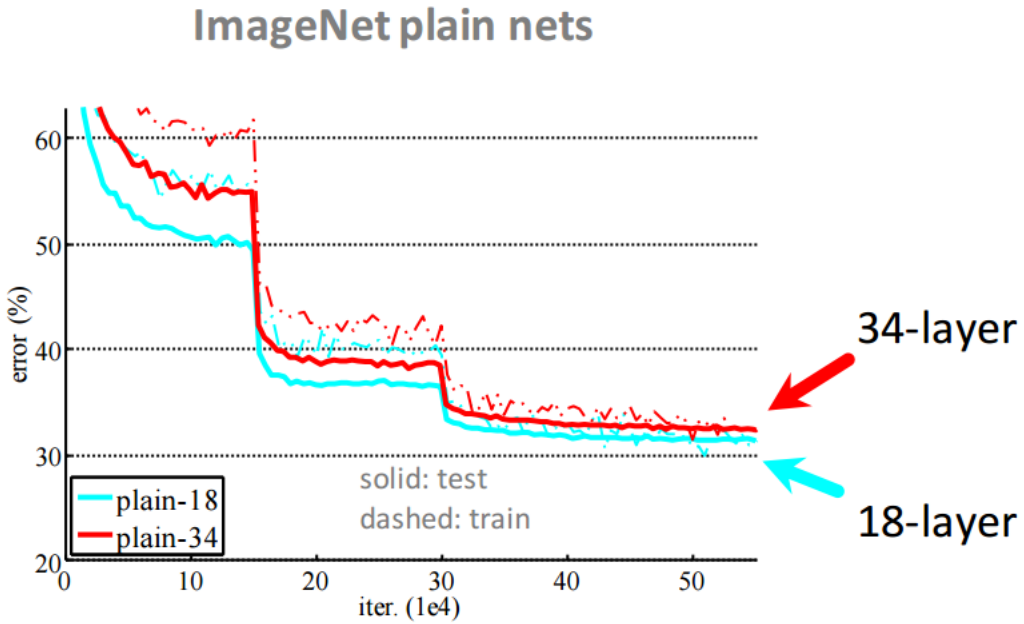
Network “Design”

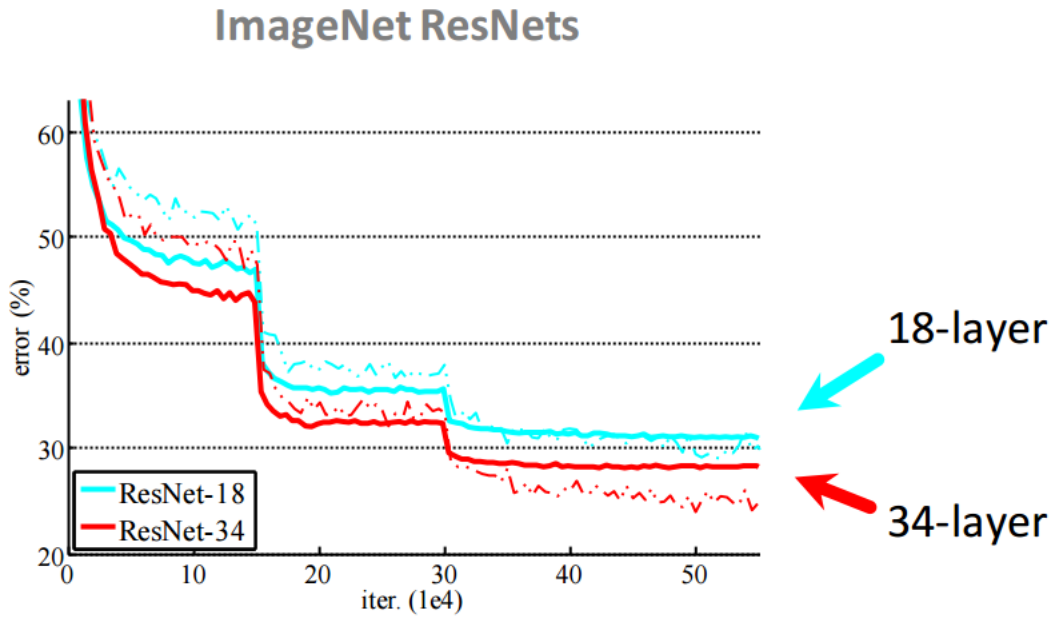
- Keep it simple
- Our basic design (VGG-style)
  - all 3x3 conv (almost)
  - spatial size /2 => # filters x2 (~same complexity per layer)
  - Simple design; just deep!
- Other remarks:
  - no hidden fc
  - no dropout



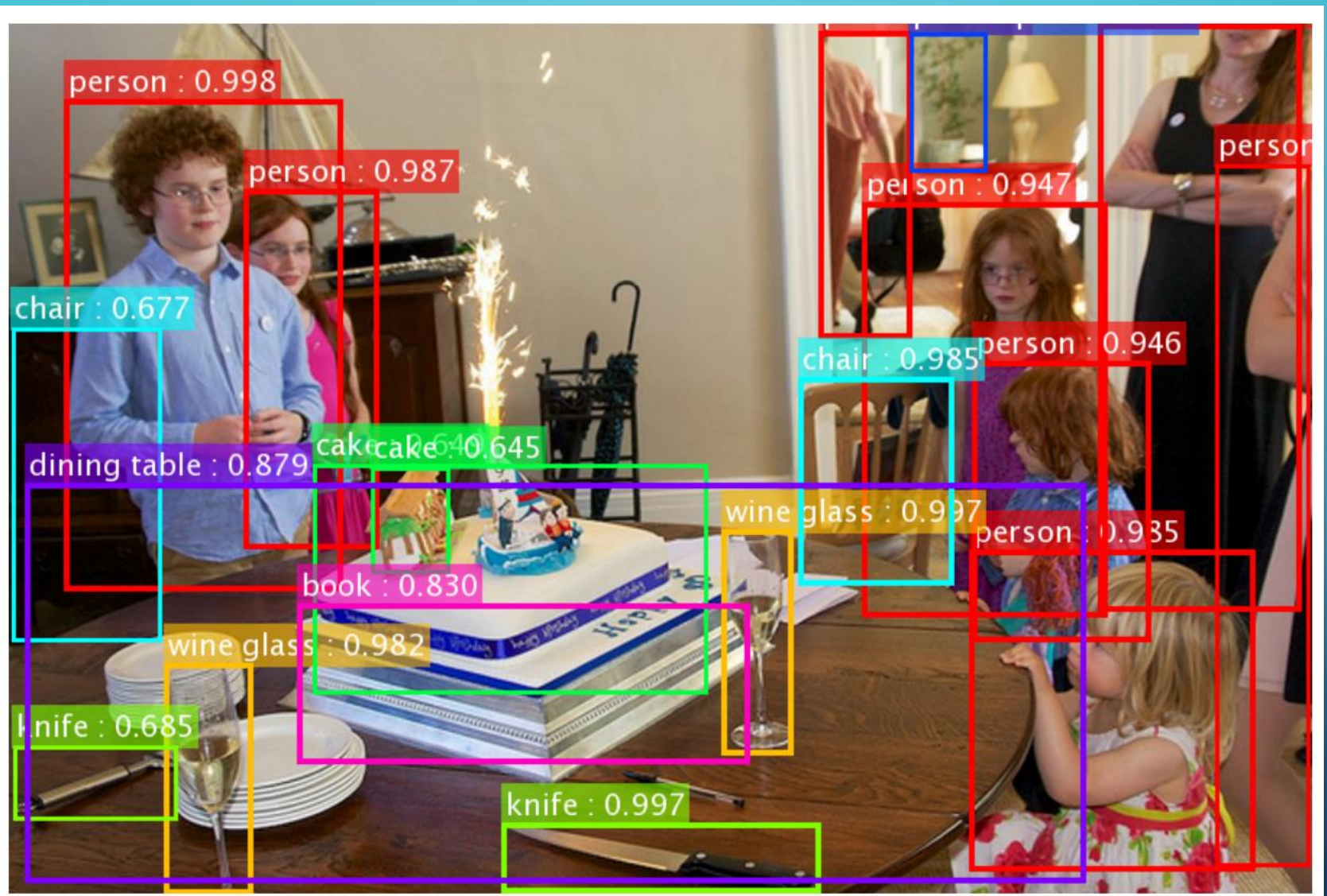

x
weight layer
$F(x)$
relu
identity
x
weight layer
$H(x) = F(x) + x$
relu

ImageNet experiments





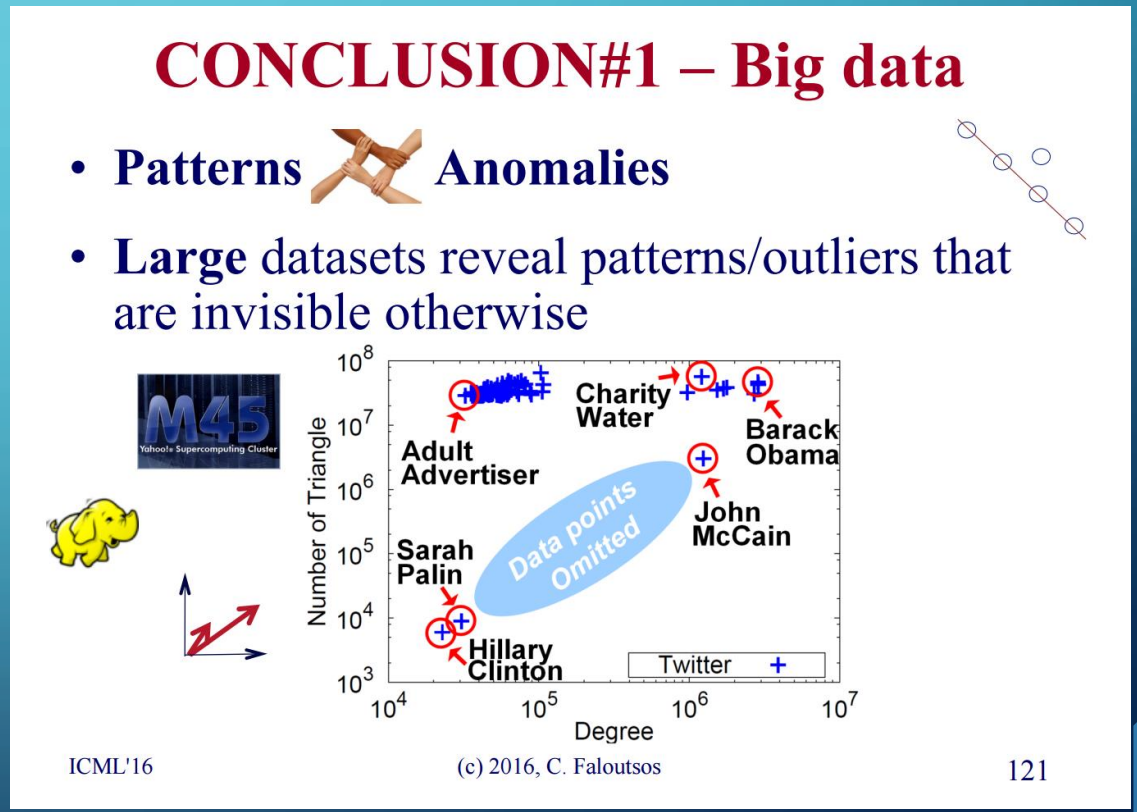
- Deep ResNets can be trained without difficulties
- Deeper ResNets have **lower training error**, and also lower test error



ResNet's object detection result on COCO

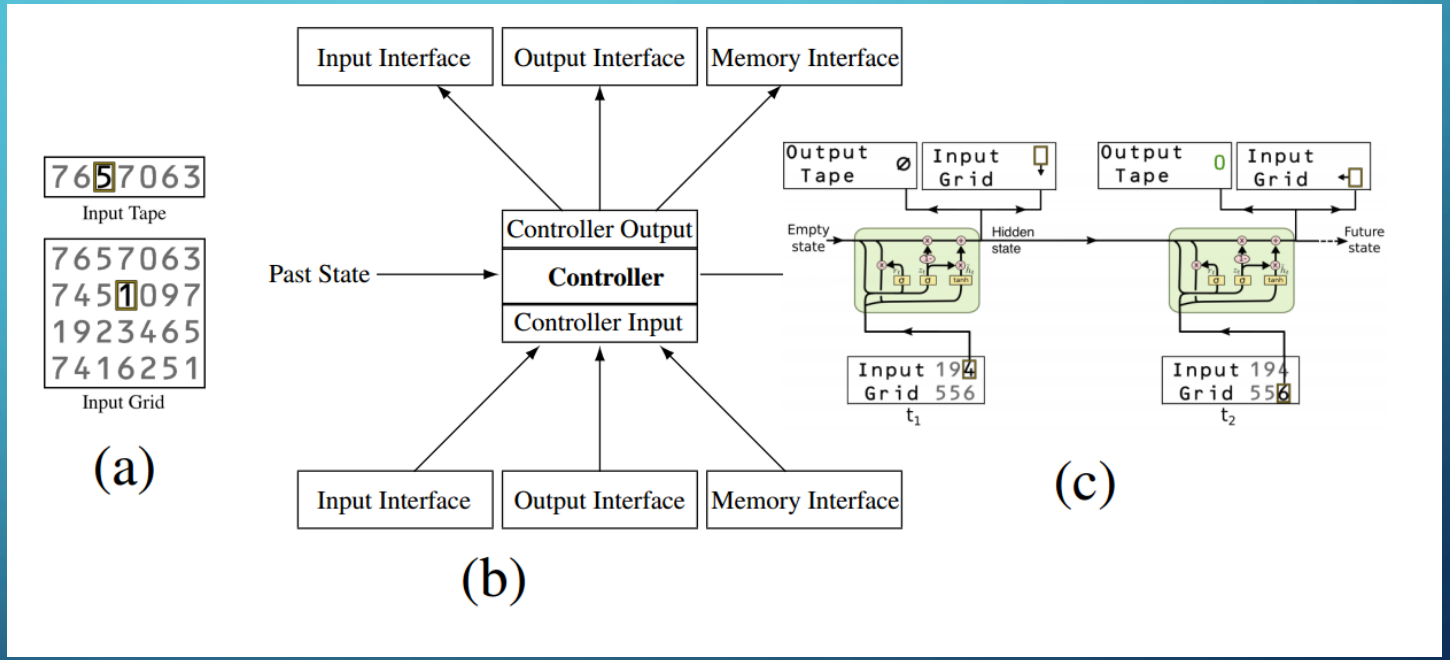
INVITED SPEAKERS

- "Mining Large Graphs: Patterns, Anomalies and Fraud Detection" – Ch. Faloutsos
  - http://icml.cc/2016/tutorials/faloutsos_ICML_2016.pdf



CONFERENCE – HIGHLIGHTS

- "Learning Simple Algorithms from Examples" – W.Zaremba, T.Mikolov, A.Joulin, R.Fergus, OpenAI



- https://www.youtube.com/watch?v=GVe6kfJnRAw

CONFERENCE – HIGHLIGHTS

- A lot of mathematics! (Convex optimization, Learning Theory, LargeScale Classification, Kernel Methods)

- A lot of engineering ☺ (Deep neural networks, Reinforcement Learning)

WORKSHOP – HIGHLIGHTS

Neural Nets Back to the Future @ ICML 16

June 23rd 2016 at Crowne Plaza in NYC

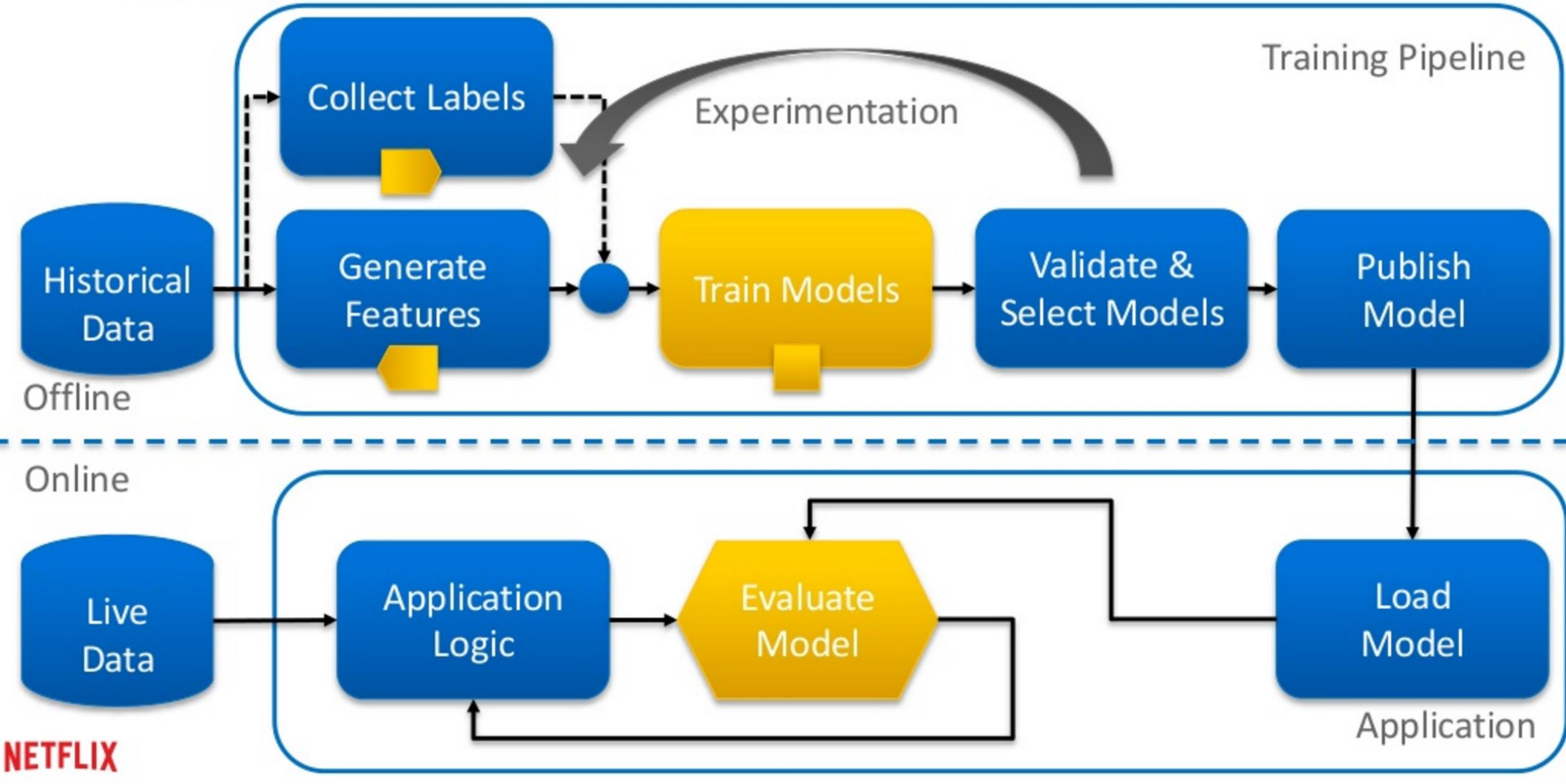
A workshop **linking** the past, present and future research on neural networks

- "Let your network grow" – John Platt
- "Recurrent Neural Networks" – C. Lee Giles
- "Learning Long Term Dependencies with Gradient Descent is Difficult" – Y. Bengio

WORKSHOPS – HIGHLIGHTS

- “Real-world Machine Learning design patterns” – J.Basilico, Netflix

“Typical” ML Pipeline: A tale of two worlds



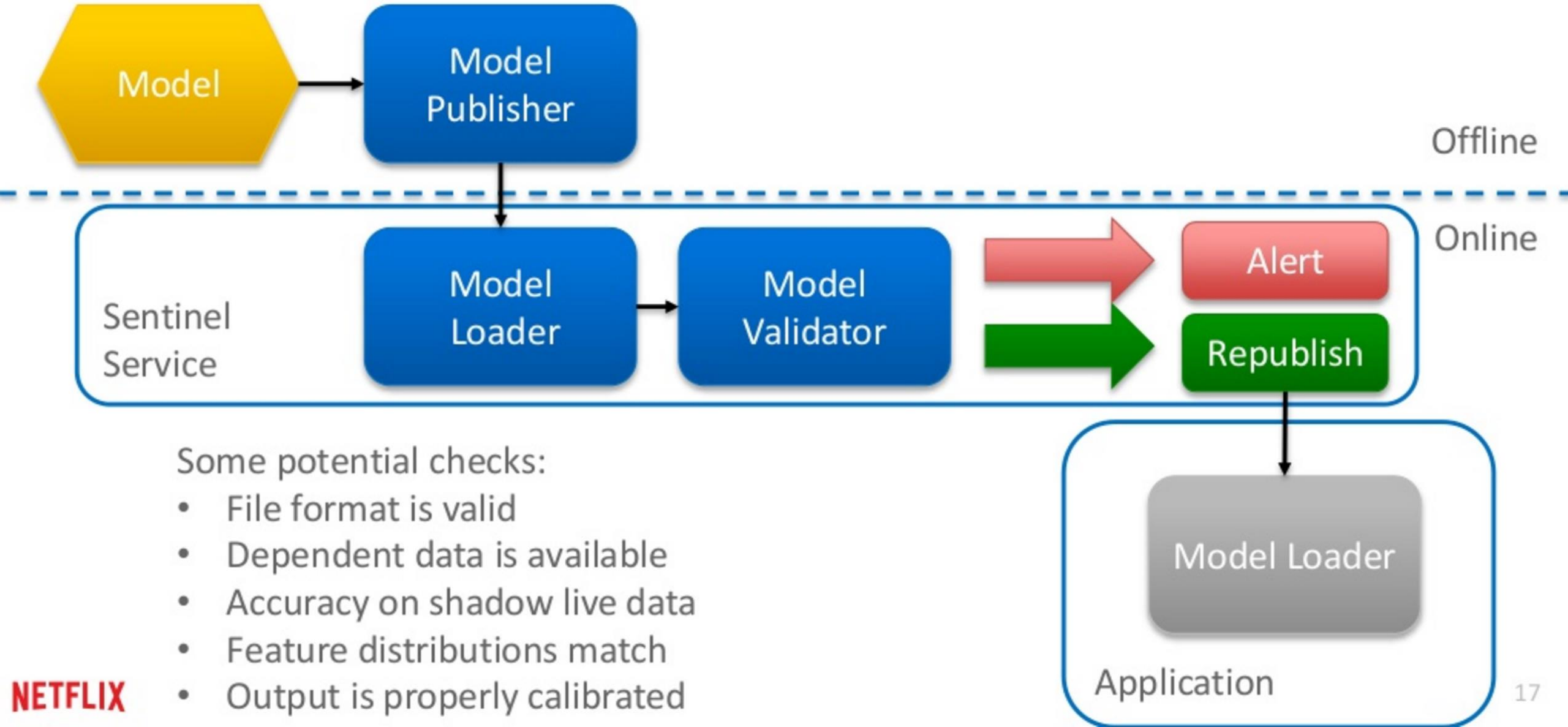
NETFLIX

The Sentinel

“You shall not pass!”

Validate model/data in online environment before letting it go live

Sentinel: Structure



Some potential checks:

- File format is valid
- Dependent data is available
- Accuracy on shadow live data
- Feature distributions match
- Output is properly calibrated

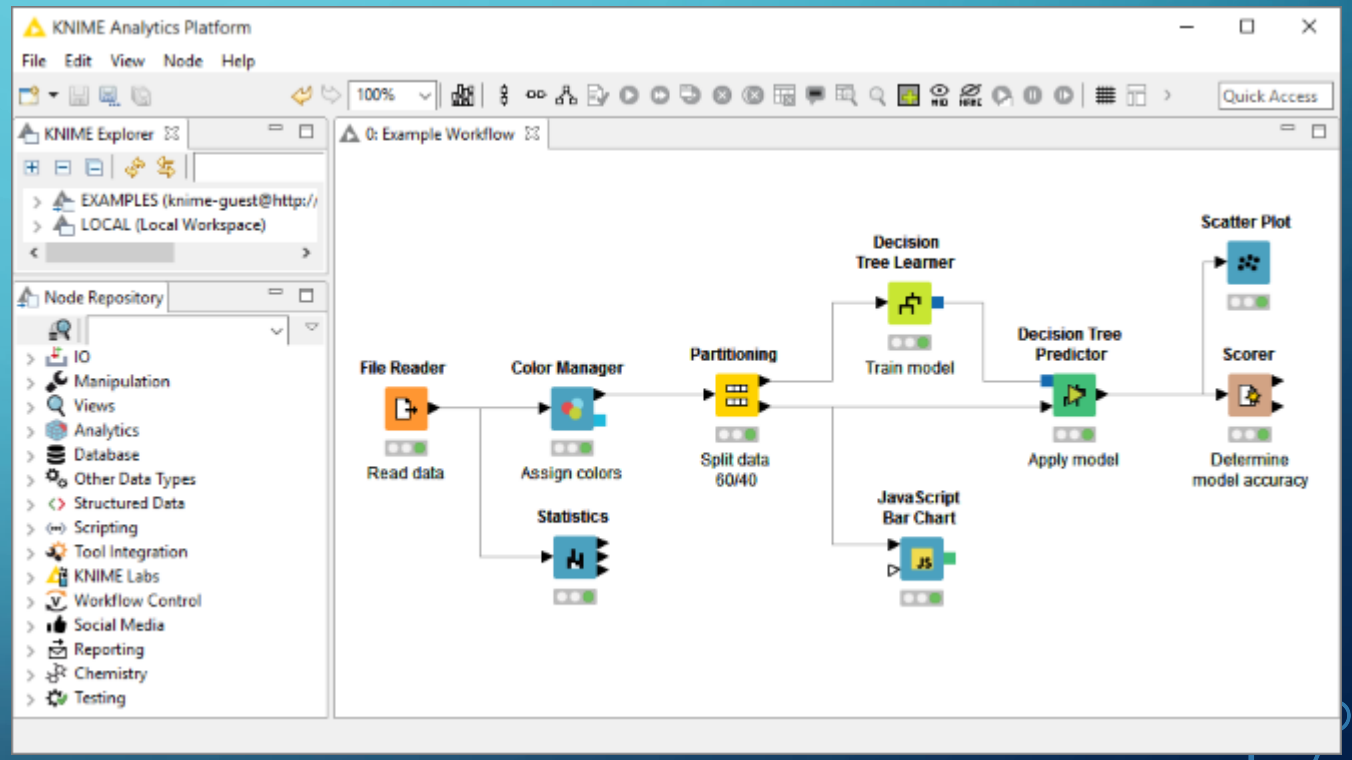
NETFLIX

Sentinel

- Example: Checking that new ranking model is valid and performs better than previous one
- Pros:
  - Using a model requires both code and data are available
  - Models may need to be versioned along-side code changes
  - Ensure that a new model is no worse than previous one
- Cons:
  - Sentinel needs to be in sync with application code
  - Difficult to choose failure thresholds for data-based checks

WORKSHOPS – HIGHLIGHTS

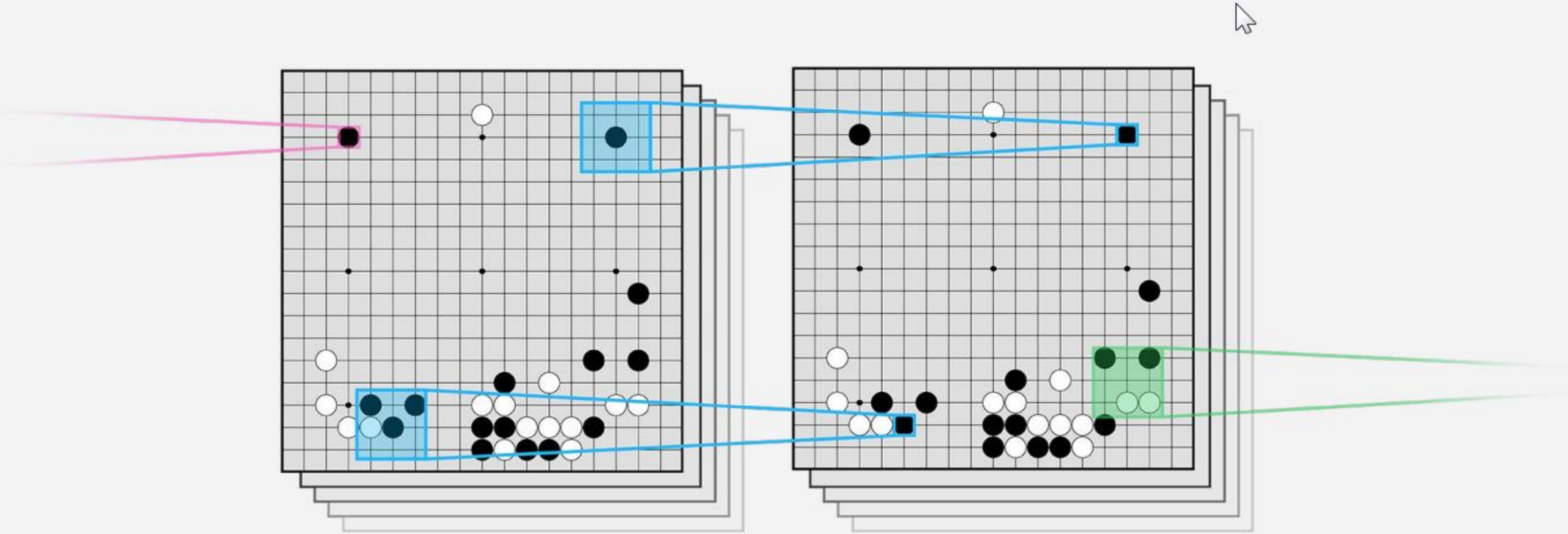
- “Modular Machine Learning with KNIME”
- “Probabilistic Demand Forecasting at Amazon Scale”
- “Intel® MKL-DNN – an Open-Source Library of Neural Network Primitives”



ALPHA GO

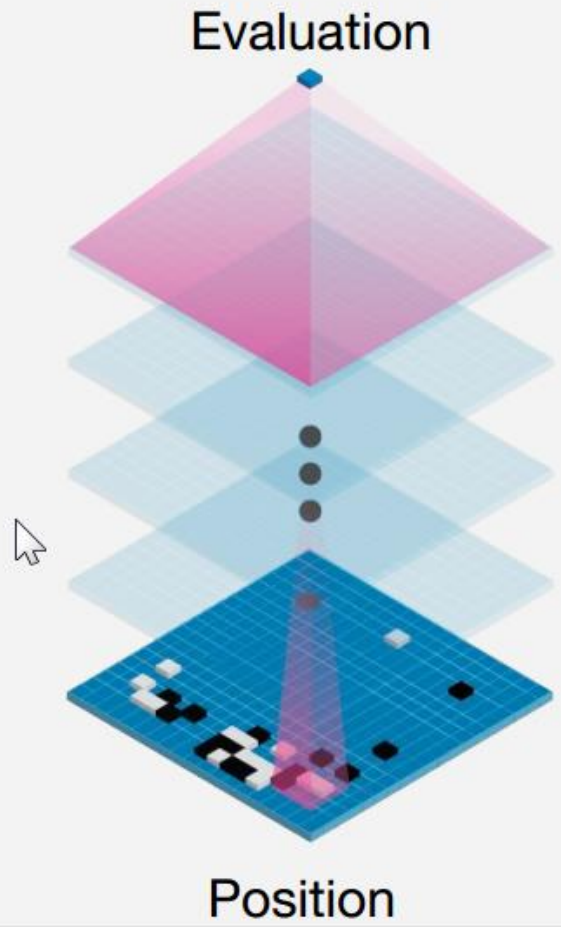

1/5 Match 1 - Google DeepMind Challenge Match: Lee Sedol vs Alp...
ALPHAGO
01:13:55
Google DeepMind
Challenge Match
LEE SEDOL
01:14:23
1:54:08 / 3:55:06
YouTube

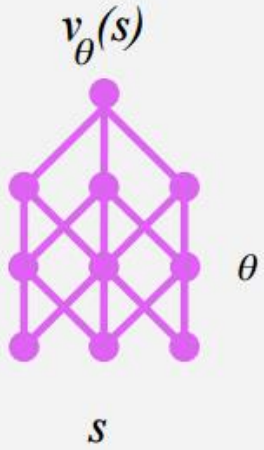
Convolutional neural network



Google DeepMind

Value network

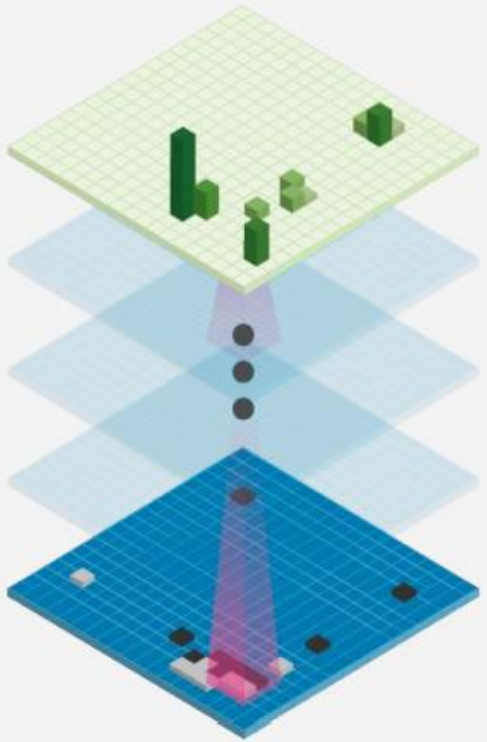




Google DeepMind

Policy network

Move probabilities



Position

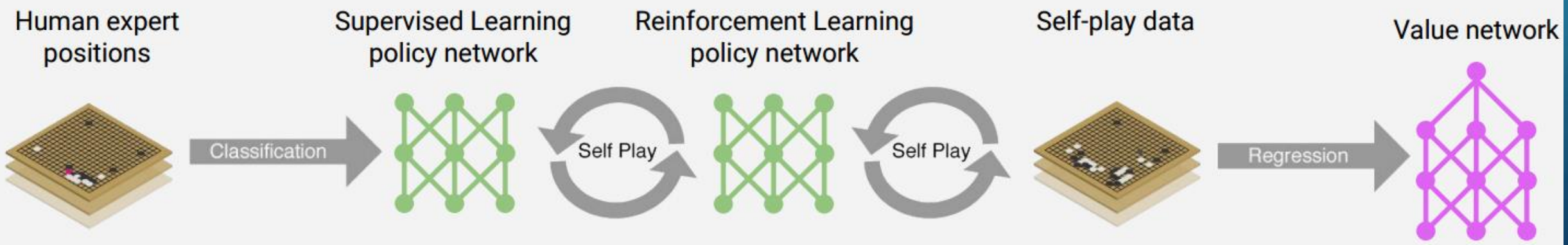
$p_\sigma(a|s)$

σ

s

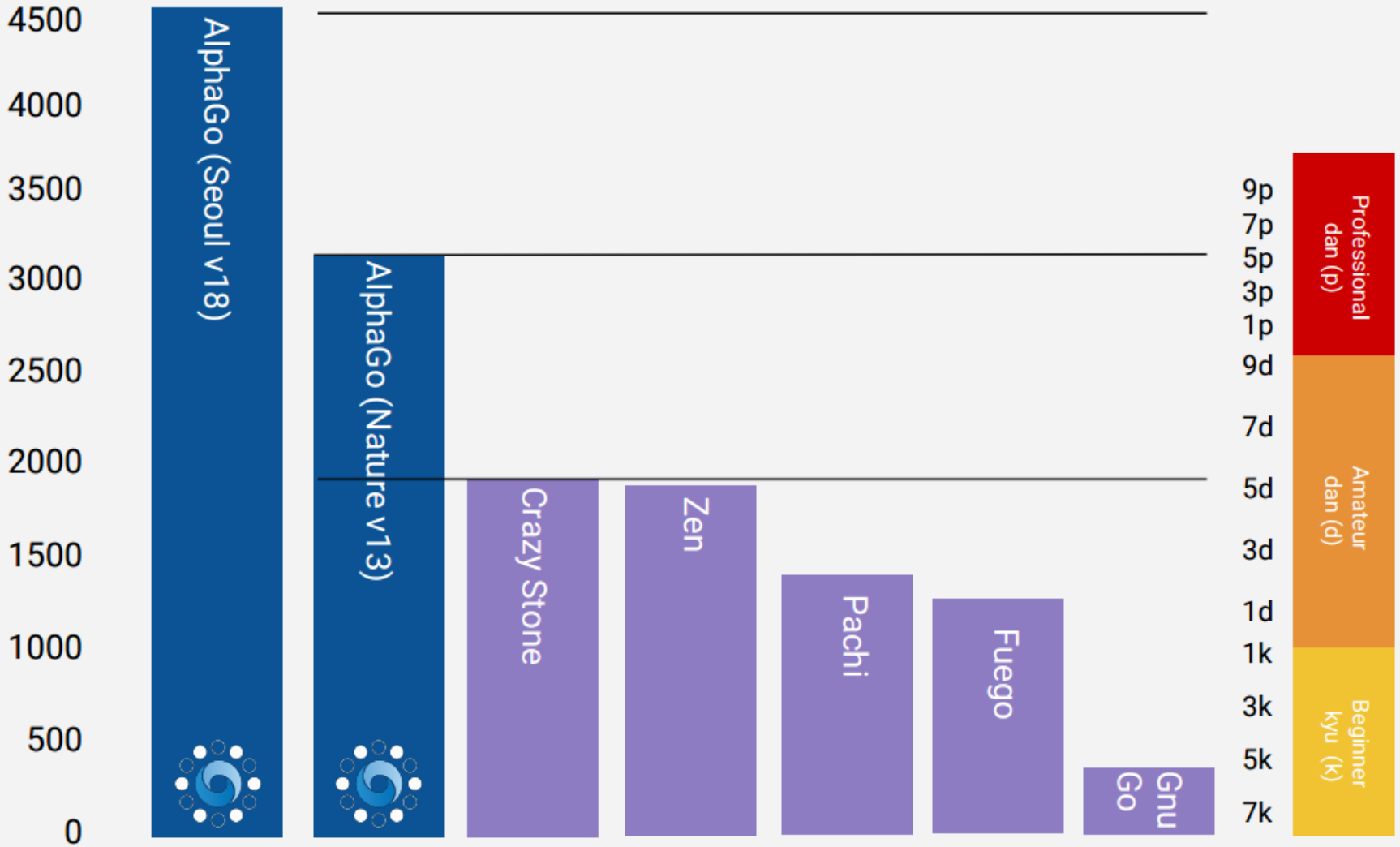
Google DeepMind

Neural network training pipeline



Google DeepMind

Evaluating AlphaGo against computers



LINKS

- https://icml.cc
- http://icml.cc/2016/wp-content/uploads/ICML-Book-Web-Version.pdf
- http://jmlr.org/proceedings/papers/v48/
- https://openai.com/blog/
- https://gogameguru.com/i/2016/03/deepmind-mastering-go.pdf
- http://www.nature.com/nature/journal/v529/n7587/full/nature16961.html
- http://jmlr.org/proceedings/papers/v48/zaremba16.pdf
- https://arxiv.org/pdf/1512.02595v1.pdf
- http://www.slideshare.net/justinbasilico/is-that-a-time-machine-some-design-patterns-for-real-world-machine-learning-systems


The End