2016 FEDERAL FORUM

Intelligent Networking in the Modern Age: Brocade's Approach to Machine Learning for Networking

David Meyer
Chief Scientist and Brocade Fellow

Presented by BROCADE Produced by fedscoop

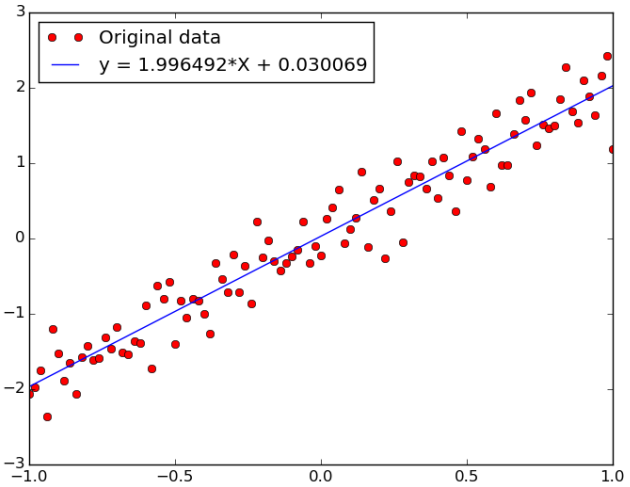
BROCADE

"You might be surprised but what is going to drive innovation in the enterprise and in the public cloud is machine learning."

Bill Coughran, Sequoia Capital at #ONUGSpring16

Agenda

- What Is Machine Learning?
- What Is All the Machine Learning Excitement About?
- Brocade's Approach to Machine Learning
- Overview of Our Demo: Open Network Insight
- Technical Explanations/Code
  - http://www.1-4-5.net/~dmm/ml



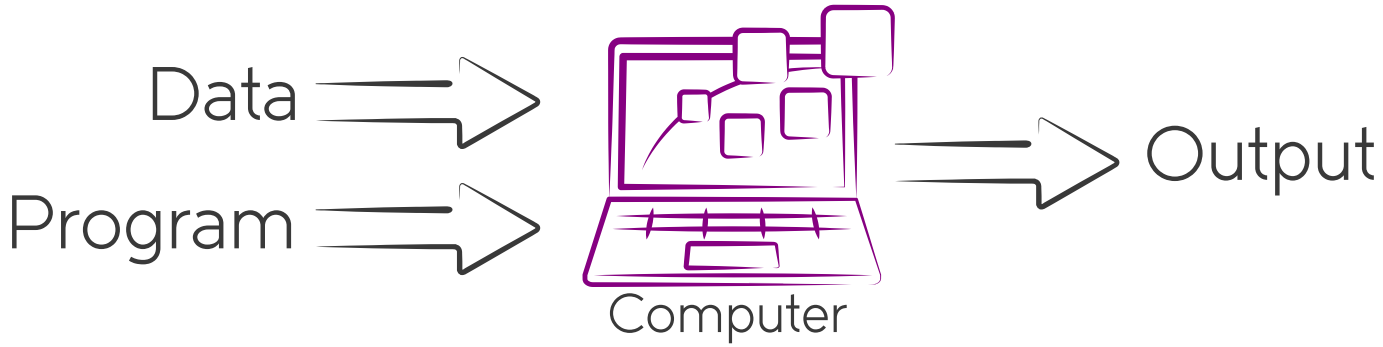
First, What Is Machine Learning?

> "The complexity in traditional computer programming is in the code (programs that people write). In machine learning, learning algorithms are in principle simple and the complexity (structure) is in the data. Is there a way that we can automatically learn that structure? That is what is at the heart of machine learning."
>
> **Andrew Ng**

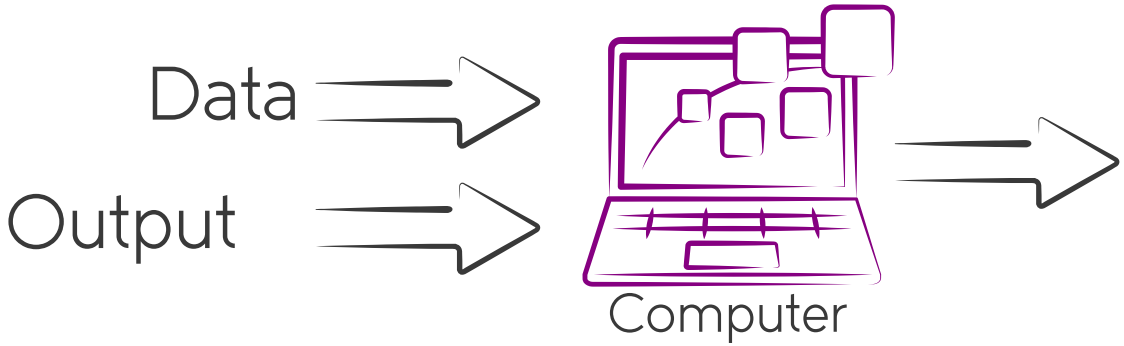
- Said another way, we want to discover the Data Generating Distribution that underlies the data we observe. This is the function that we want to learn.
- Moreover, we care about primarily about the generalization accuracy of our model (function)
  - *Accuracy on examples we have not yet seen -- How can this be?*
  - As opposed the accuracy on the training set (note: overfitting)

Same Thing Said in Cartoon Form

Traditional Programming

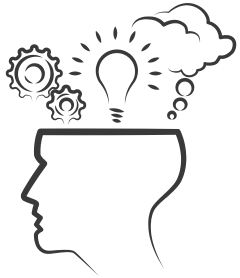


Machine Learning

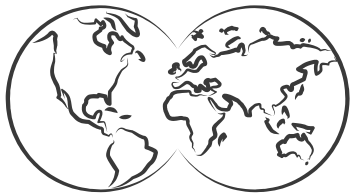


Ok, We Know That Machines Are Getting Smarter, But Where Does Knowledge Come From?

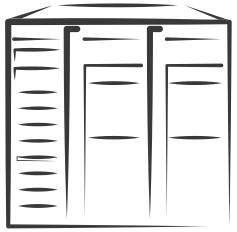
EVOLUTION



EXPERIENCE



CULTURE



MACHINES

Many orders of magnitude faster and larger

So how can machines discover new knowledge?

How Machines Can Discover New Knowledge?

Fill the gaps in existing knowledge.

- Symbolists
- Technology: Induction/Inverse Deduction

Emulate the brain.

- Connectionists
- Technology: Deep Neural Nets

Emulate evolution.

- Evolutionaries
- Technology: Genetic Algorithms

Systematically reduce uncertainty.

- Bayesians
- Technology: Bayesian Inference

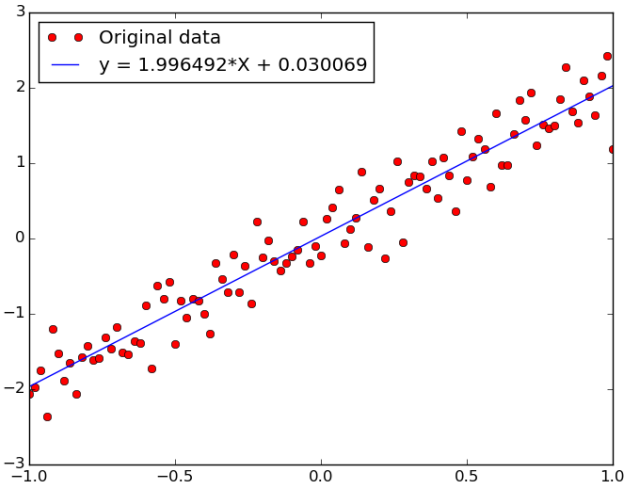
Notice similarities between old and new.

- Analogizers
- Technology: Kernel Machines/Support Vector Machines

These correspond to the 5 major schools of thought in machine learning

Agenda

- What Is Machine Learning?
- What Is All the Machine Learning Excitement About?
- Brocade's Approach to Machine Learning
- Overview of Our Demo: Open Network Insight
- Technical Explanations/Code
  - http://www.1-4-5.net/~dmm/ml

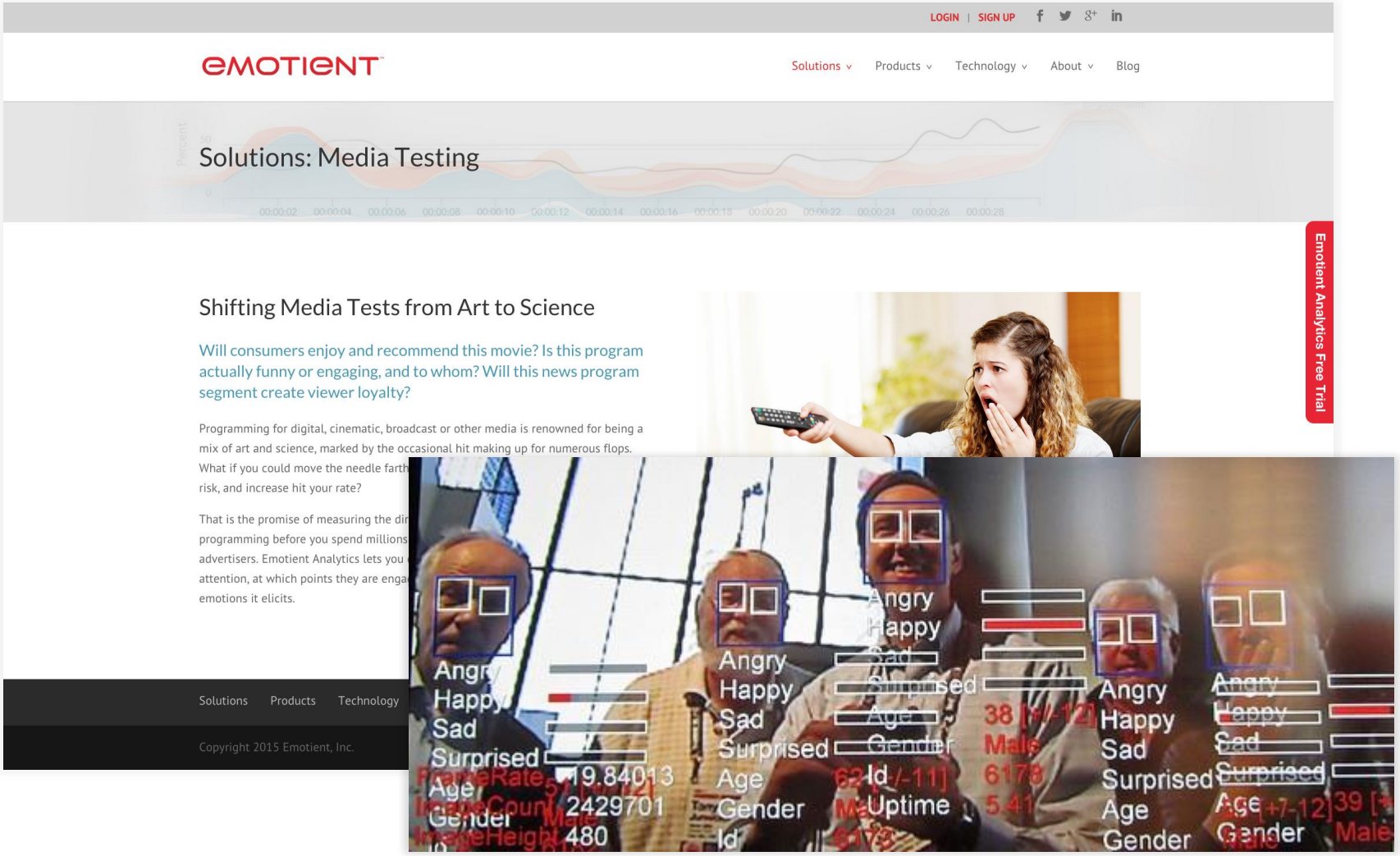


What Is All the ML Excitement About?

ML applications you interact with everyday

Why this is relevant: Compute, Storage, Networking, Security, and Energy (CSNSE) use cases will all use this technology

BTW, Think Object Recognition Is Impressive?



Lip Reading?

LIPREADING WITH LONG SHORT-TERM MEMORY

Michael Wand, Jan Koutník, Jürgen Schmidhuber

The Swiss AI Lab IDSIA, USI & SUPSI

ABSTRACT

Lipreading, i.e. speech recognition from visual-only recordings of a speaker's face, can be achieved with a processing pipeline based solely on neural networks, yielding significantly better accuracy than conventional methods. Feed-forward and recurrent neural network layers (namely Long Short-Term Memory; LSTM) are stacked to form a single structure which is trained by back-propagating error gradients through all the layers. The performance of such a stacked network was experimentally evaluated and compared to a standard Support Vector Machine classifier using conventional computer vision features (Eigenlips and Histograms of Oriented Gradients). The evaluation was performed on data from 19 speakers of the publicly available GRID corpus. With 51 different words to classify, we report a best word accuracy on held-out evaluation speakers of 79.6% using the end-to-end neural network-based solution (11.6% improvement over the best feature-based solution evaluated).

Index Terms— Lipreading, Long Short-Term Memory, Recurrent Neural Networks, Image Recognition

1. INTRODUCTION

It is well-known that humans understand speech not only by listening, but also by taking visual cues into account [1].

Local Binary Patterns [7]. Classification is frequently done with Support Vector Machines (SVMs), e.g. [7], or Hidden Markov Models (HMMs), e.g. [4, 5, 8, 9].

Our aim is to replace the complete visual speech recognition pipeline with a compact neural network architecture. Neural networks (NNs) have become increasingly popular in conventional speech recognition, first as feature extractors in an HMM-based architecture [10–12], more recently replacing the entire processing chain [13]. For the latter, the *Long Short Term Memory* (LSTM; [14]) architecture is typically used. Consequently, our approach to the lipreading problem uses a NN that chains feed-forward layers and LSTM layers, described in detail in subsection 4.2. Manual feature extraction is no longer required. The NN inputs are now the raw mouth images, as is common in modern computer vision tasks, but stands in stark contrast e.g. to [5, 7].

2. RELATED WORK

Lipreading has been used as a complementary modality for speech recognition from noisy audio data [2, 15], as well as for purely visual speech recognition [3, 16, 17]. The latter gives rise to a *Silent Speech interface*, which is defined as a system "enabling speech communication to take place when an audible acoustic signal is unavailable" [18]. Silent Speech technology has a large number of applications: it allows per-

Real-Time Language Translation

One More

WSJ | WSJ LIVE | REALTOR.COM | MANSION GLOBAL | BARRON'S | MEMBERSHIP | DJX | MORE — News, Quotes, Companies, Videos — SEARCH

Sign In

WSJ.D | TECH

Venture Capital Dispatch

An inside look from VentureWire at high-tech startups and their investors.

INVESTORS/RAISING CAPITAL
Norwest Venture Partners Closes Its Third $1.2 Billion Fund

COMPANY FUNDING
Internet-of-Things Security Company ForeScout Valued at $1 Billion

DATA | COMPANY FUNDING | VENTURE FUNDS | M&A | IPOS | PEOPLE

9:00 am ET
Jan 28, 2016 COMPANY FUNDING

Bay Labs Launches to Bring Artificial Intelligence to Ultrasounds

ARTICLE | COMMENTS

BAY LABS | ELEVEN TWO CAPITAL | KEVIN GOODWIN | KHOSLA VENTURES

Email | Print | 169 | Twitter

By CAT ZAKRZEWSKI CONNECT

PREVIOUS
Why Southeast Asia's GrabTaxi Is Removing 'Taxi' From Its Name

NEXT
Insurify Raises $2 Million for Virtual Insurance Agent

SEARCH VENTURE CAPITAL DISPATCH | GO

About Venture Capital Dispatch

Produced by the editors of **Dow Jones VentureWire**, Venture Capital Dispatch tracks the fast-moving developments at the intersection of high-tech innovation and venture capital finance. Featuring the VentureWire reporting team in the Silicon Valley, New York, Boston and Shanghai tech centers, Venture Capital Dispatch provides insight into the newest start-ups and latest trends in venture capital investing. Write us at **VCdispatch@dowjones.com**. For more information on Dow Jones products covering venture capital and other financial markets, go to **http://pevc.dowjones.com**.

Follow Venture Capital Dispatch on Twitter

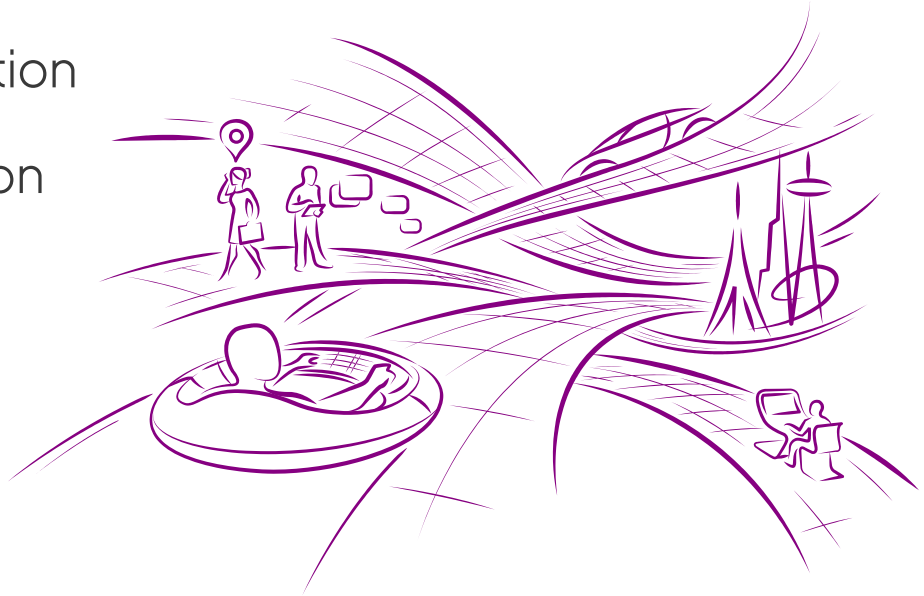
Like Venture Capital Dispatch on Facebook

Venture Capital Dispatch Bureau

Mike Billings | Scott Martin

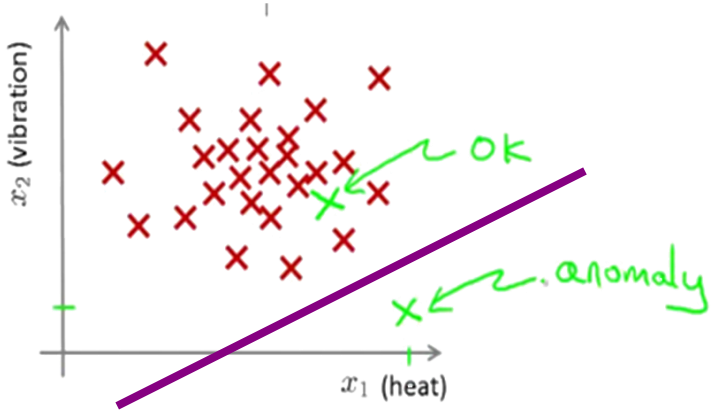
So What Kinds of Network Use Cases Are We Working On?

- Security/Anomaly Detection
- NFV orchestration and optimization
- New automation tools for DevOps
- Predicting and remediating problems in the mobile network
- Network control plane optimization
- Capture operator/analyst intuition
- …



Example:
Using Flow Data for Anomaly Detection

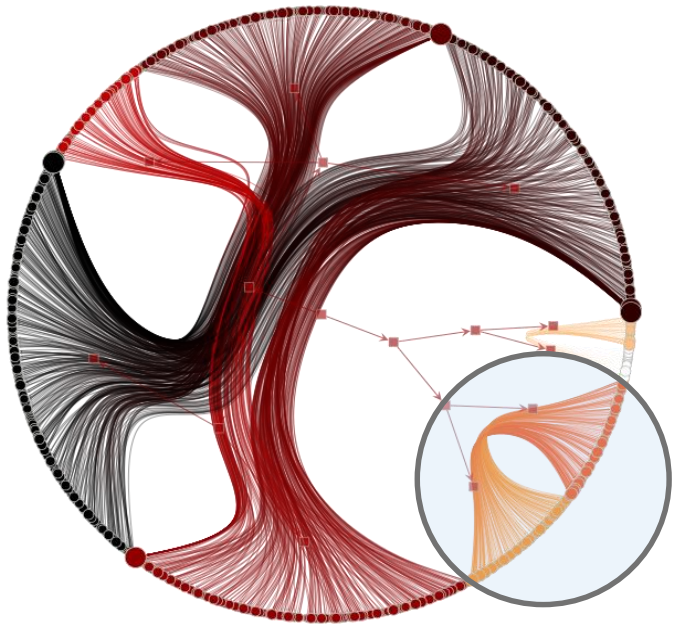
General Anomaly Detection Setting



Linear Decision Boundary

Generalization Graph

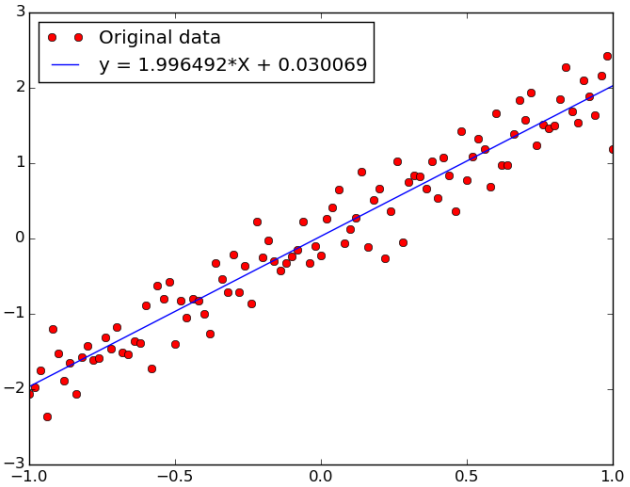
Radial Nested Block State Model with Edge Bundling



DNS Tunneling

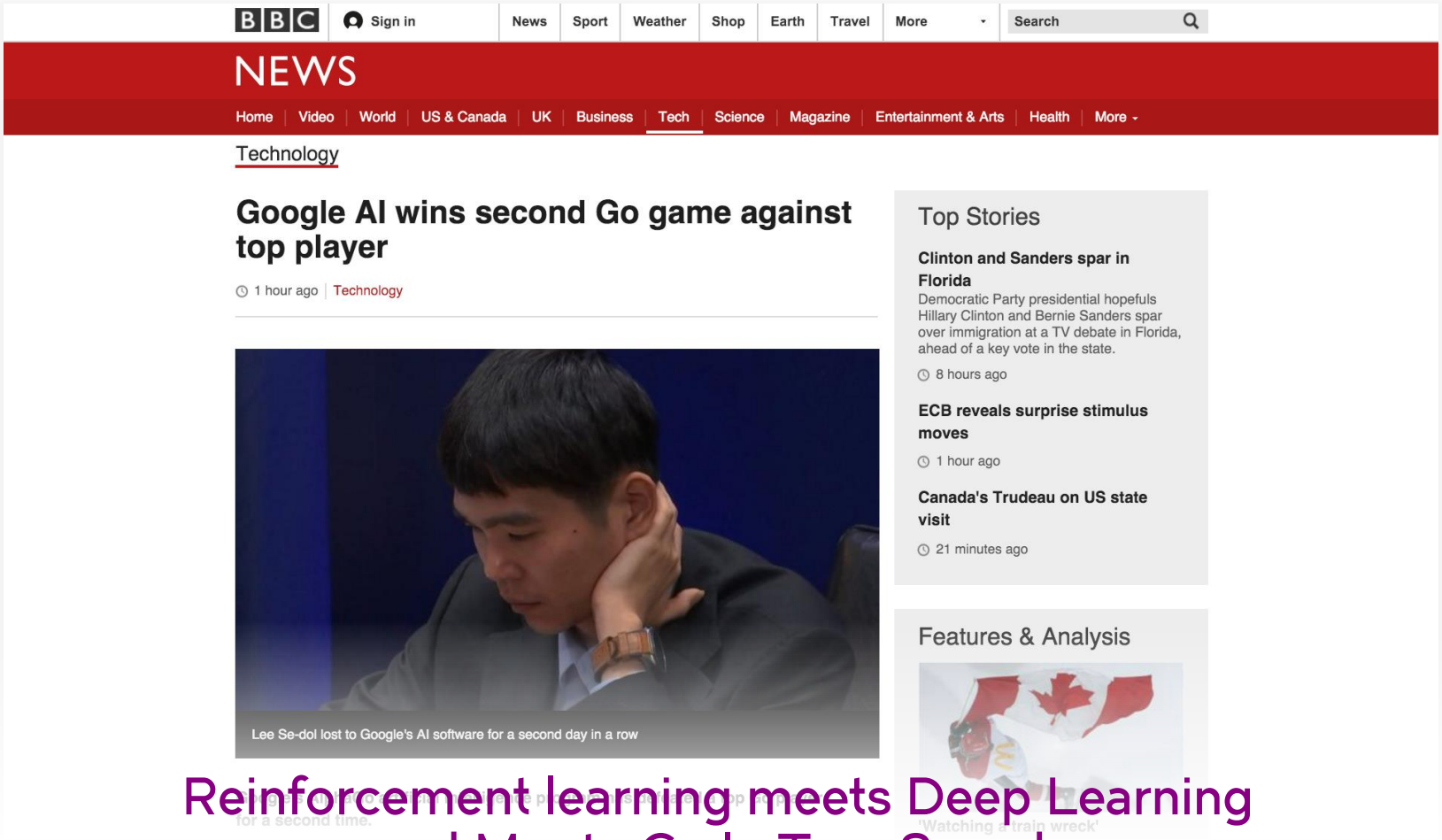
Agenda

- What Is Machine Learning?
- What Is All the Machine Learning Excitement About?
- Brocade's Approach to Machine Learning
- Overview of Our Demo: Open Network Insight
- Technical Explanations/Code
  - http://www.1-4-5.net/~dmm/ml



Pushing the Boundaries: Brocade's Approach

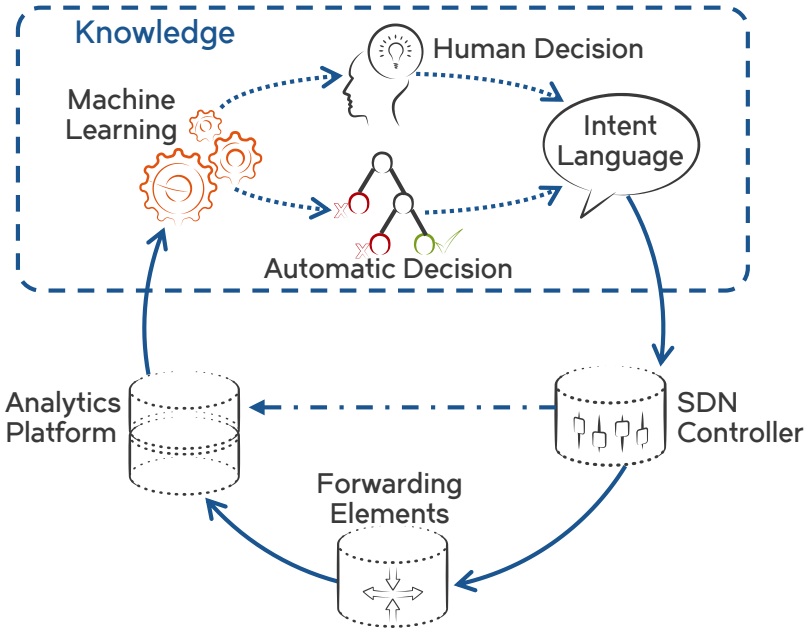
Bringing deep reinforcement learning to networking



Reinforcement learning meets Deep Learning and Monte Carlo Tree Search

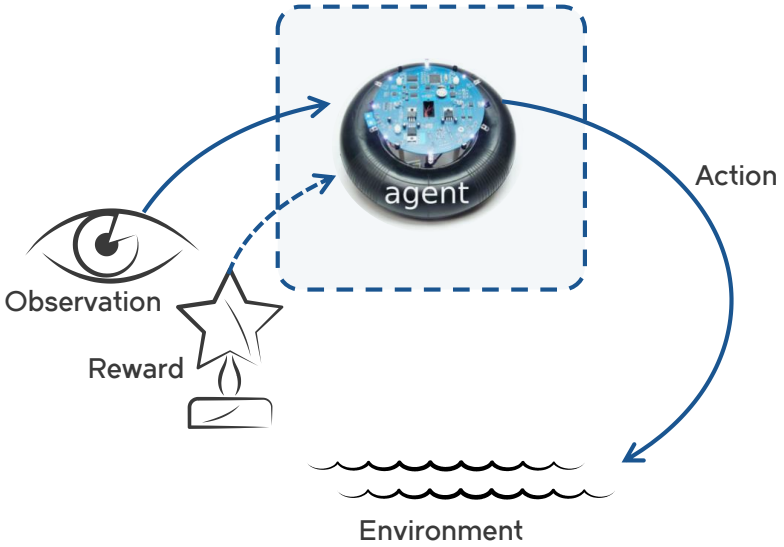
Static Architecture vs. Reinforcement Architecture

Static ML Architecture



- ML doesn't have "agency"
- Hard-coded or open loop control
- Doesn't learn control
- Assumes static underlying distribution

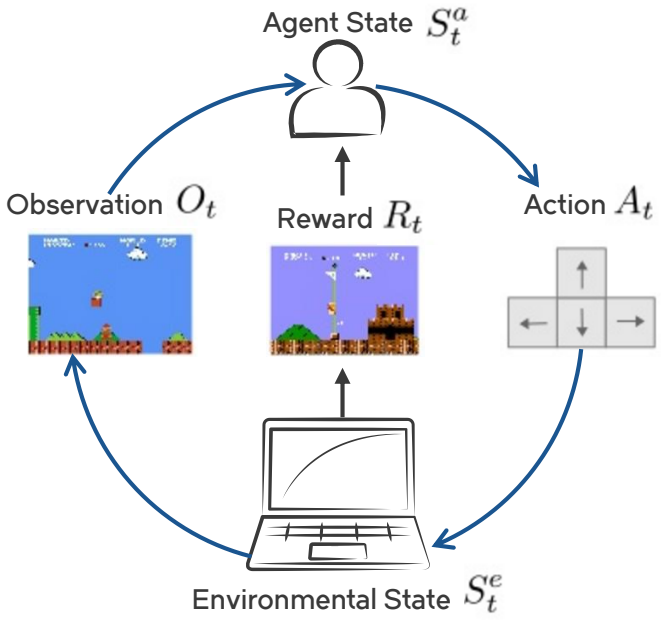
Reinforcement Learning Architecture



- Agent architecture
- Agent learns control/action selection
- Adapts to evolving environment
- Network gamification

Reinforcement Learning Example

Reinforcement Learning Setup

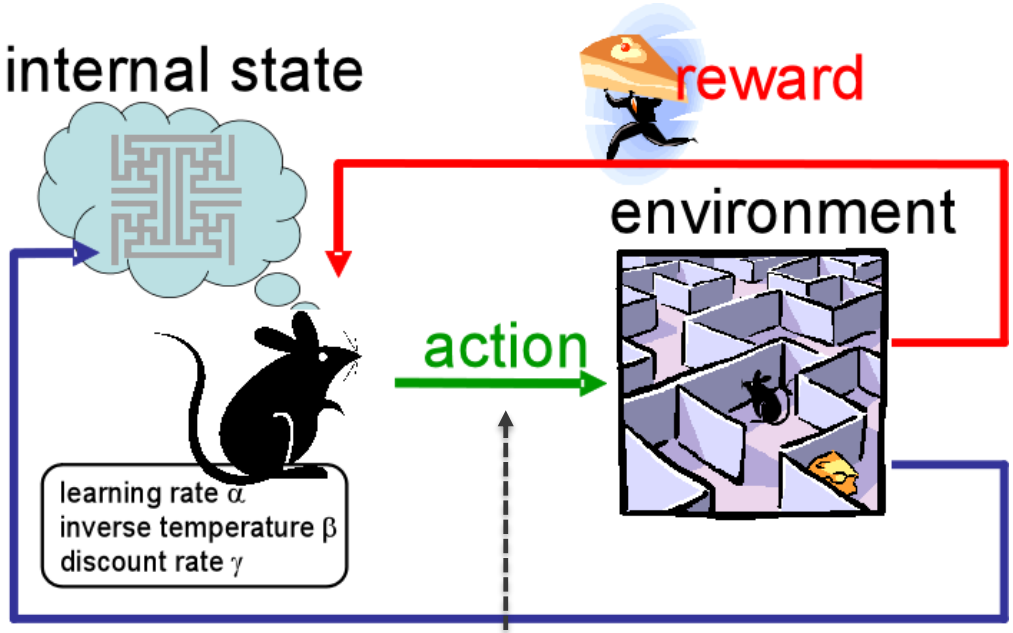


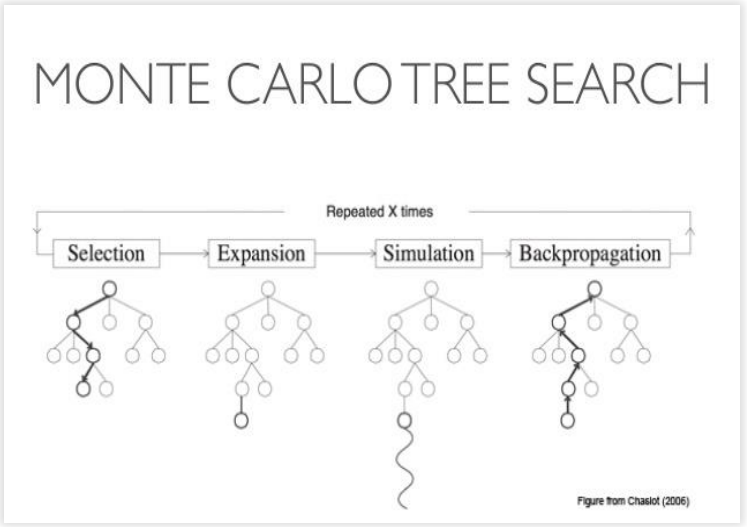
Reinforcement Learning Example

Environment:	You are in state 65. You have 4 possible actions.
Agent:	I'll take action 2.
Environment:	You received a reinforcement of 7 units. You are now in state 15. You have 2 possible actions.
Agent:	I'll take action 1.
Environment:	You received a reinforcement of -4 units. You are now in state 65. You have 4 possible actions.
Agent:	I'll take action 2.
Environment:	You received a reinforcement of 5 units. You are now in state 44. You have 5 possible actions.
⋮	⋮

- Rules of the game are unknown
- No supervisor, only a reward signal
- Feedback is delayed
- Agent's actions affect the subsequent data it receives

Drilling Down Just a Bit





$$Q_{t+1}(s_t, a_t) = \underbrace{Q_t(s_t, a_t)}_{\text{old value}} + \underbrace{\alpha_t(s_t, a_t)}_{\text{learning rate}} \times \left[\overbrace{\underbrace{R_{t+1}}_{\text{reward}} + \underbrace{\gamma}_{\text{discount factor}} \underbrace{\max_a Q_t(s_{t+1}, a)}_{\text{estimate of optimal future value}}}^{\text{learned value}} - \underbrace{Q_t(s_t, a_t)}_{\text{old value}} \right]$$

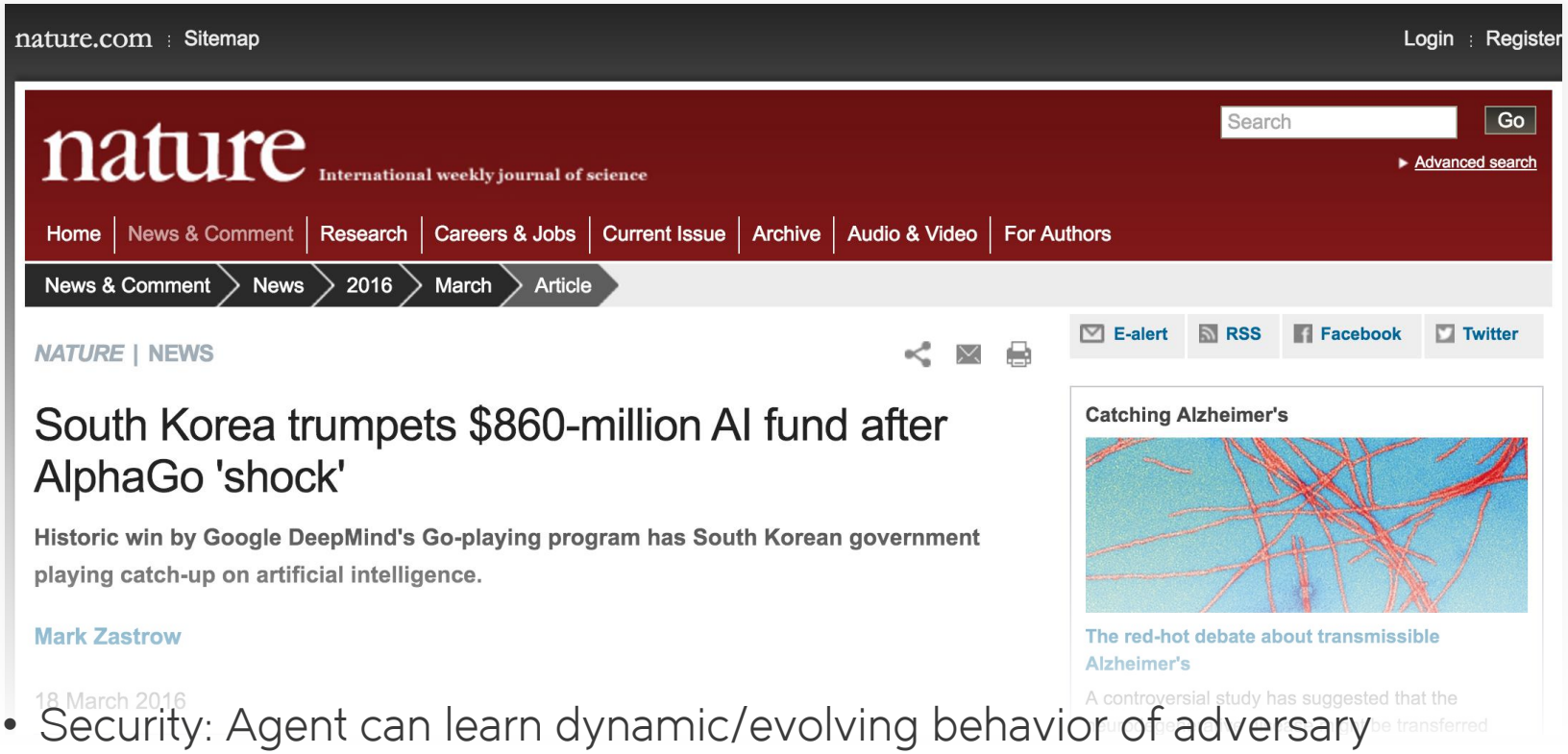
high values = pleasure

high values = pleasant anticipation

low values = pain

low values = fear

Why Is AlphaGo Important/Relevant?



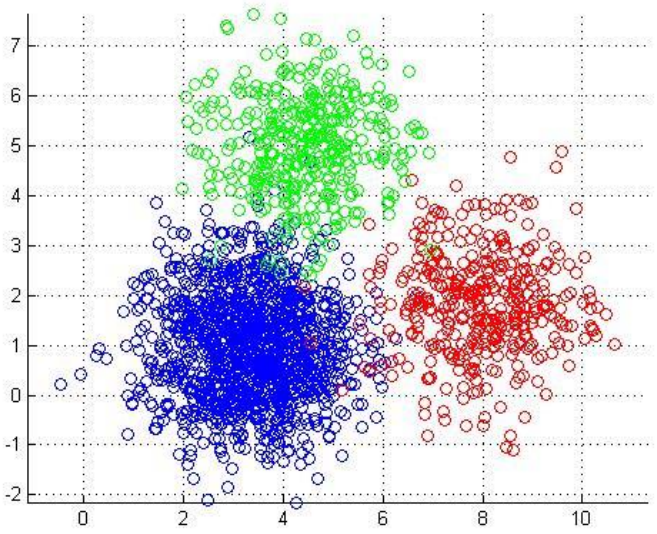
- Security: Agent can learn dynamic/evolving behavior of adversary
- DevOPs: Agent can learn workflow automation (e.g., Openstack Mistral/StackStorm)
- NFV: Agent can learn dynamic behavior of VNFs
- General: Deep learning can capture human intuition

Brocade's Approach to Machine Learning

Bringing rigor to ML for Networking

Clustering

- Categorical and continuous (e.g., LDA, K-means, ..
- Anomaly detection



Brocade's Approach to Machine Learning

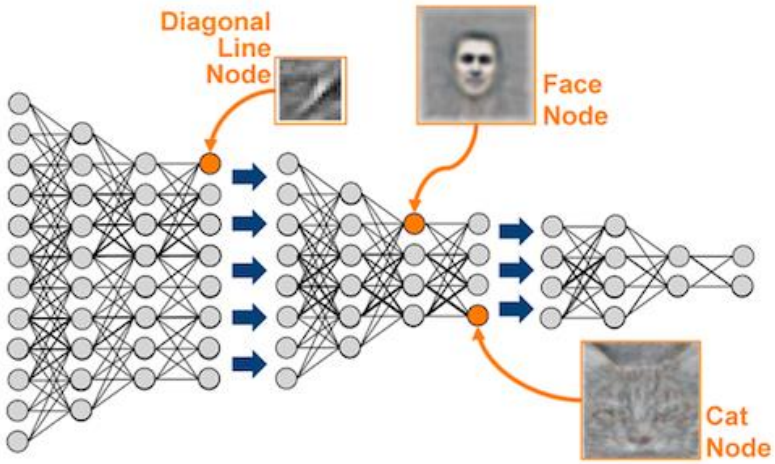
Bringing rigor to ML for Networking

Clustering

- Categorical and continuous (e.g., LDA, K-means, …)
- Anomaly detection

Deep Neural Networks

- Understand sequences such as network flows
- Capture expert intuition
- Recurrent/memory nets (LSTMs, NTMs, …)
- Time series/long range dependencies



Brocade's Approach to Machine Learning

Bringing rigor to ML for Networking

Clustering

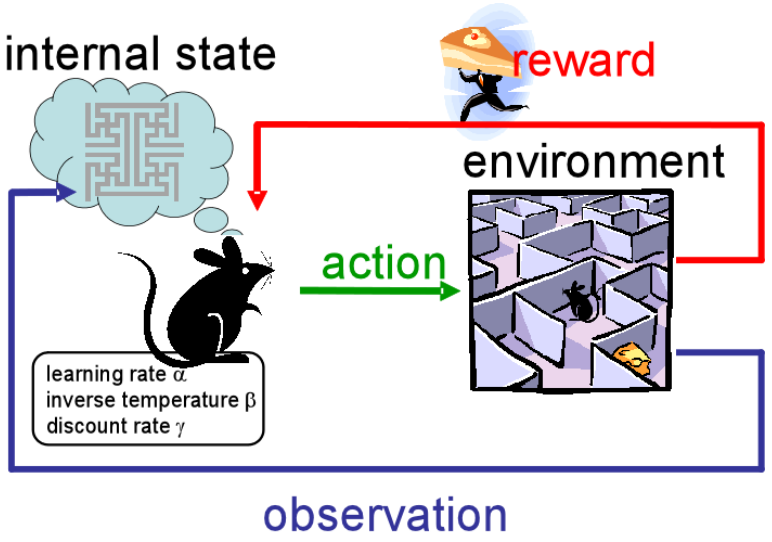
- Categorical and continuous (e.g., LDA, K-means, …)
- Anomaly detection

Deep Neural Networks

- Understand sequences such as network flows
- Capture expert intuition
- Recurrent/memory nets (LSTMs, NTMs, …)
- Time series/long range dependencies

Reinforcement Learning

- Give Machine Learning agency
  - Learn feedback control of actions
- Non-stationary distributions
- Deep neural networks (value/policy networks)
- Understand/react in adversarial environments



Brocade's Approach to Machine Learning

Bringing rigor to ML for Networking

Clustering

- Categorical and continuous (e.g., LDA, K-means, …)
- Anomaly detection

Deep Neural Networks

- Understand sequences such as network flows
- Capture expert intuition
- Recurrent/memory nets (LSTMs, NTMs, …)
- Time series/long range dependencies

Reinforcement Learning

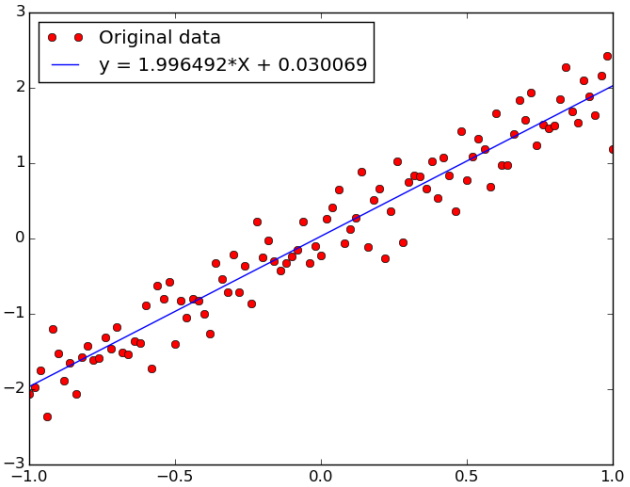
- Give Machine Learning agency
  - Learn feedback control of actions
- Non-stationary distributions
- Deep neural networks (value/policy networks)
- Understand/react in adversarial environments

Standardized (and public) data sets

- Required to evaluate techniques
- Move the field forward
  - e.g. MNIST , ImageNet, …

Agenda

- What Is Machine Learning?
- What Is All the Machine Learning Excitement About?
- Brocade's Approach to Machine Learning
- Overview of Our Demo: Open Network Insight
- Technical Explanations/Code
  - http://www.1-4-5.net/~dmm/ml



Brocade Machine Learning Partnerships

Intel Security/Cloudera

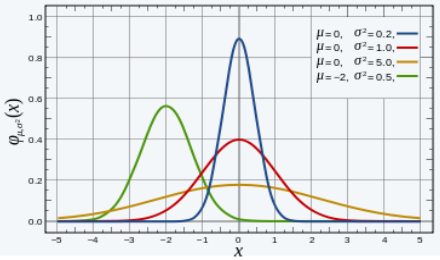
- Open Network Insight
  - http://open-network-insight.org/
  - https://github.com/Open-Network-Insight
- **Anomaly Detection** for Network Data
- Goal: Leverage flow and other packet data to help enterprises and service providers gain insight into events in their compute environments and identify potential security threats
  - In particular, DNS PCAP and netflow data
- Uses **Topic Modeling** is used to cluster categorical data
  - A low probability flow in a cluster (topic) is an anomaly

BTW, What Exactly Is Anomaly Detection?

- It is clear that one of the major challenges we face as a civilization is dealing with deluge of data that are being collected from our networks
  - While at the same time we are "knowledge starved"
  - Can't find the needles in an exponentially growing haystack
  - Anomaly Detection is one piece of the puzzle
  - Machine Learning is a fundamental part of the answer

- Key Assumption for Anomaly Detection
  - Anomalous events occur relatively infrequently (alternatively: most events normal)
  - Second order assumption: Common events follow a Gaussian distribution (likely to be wrong)

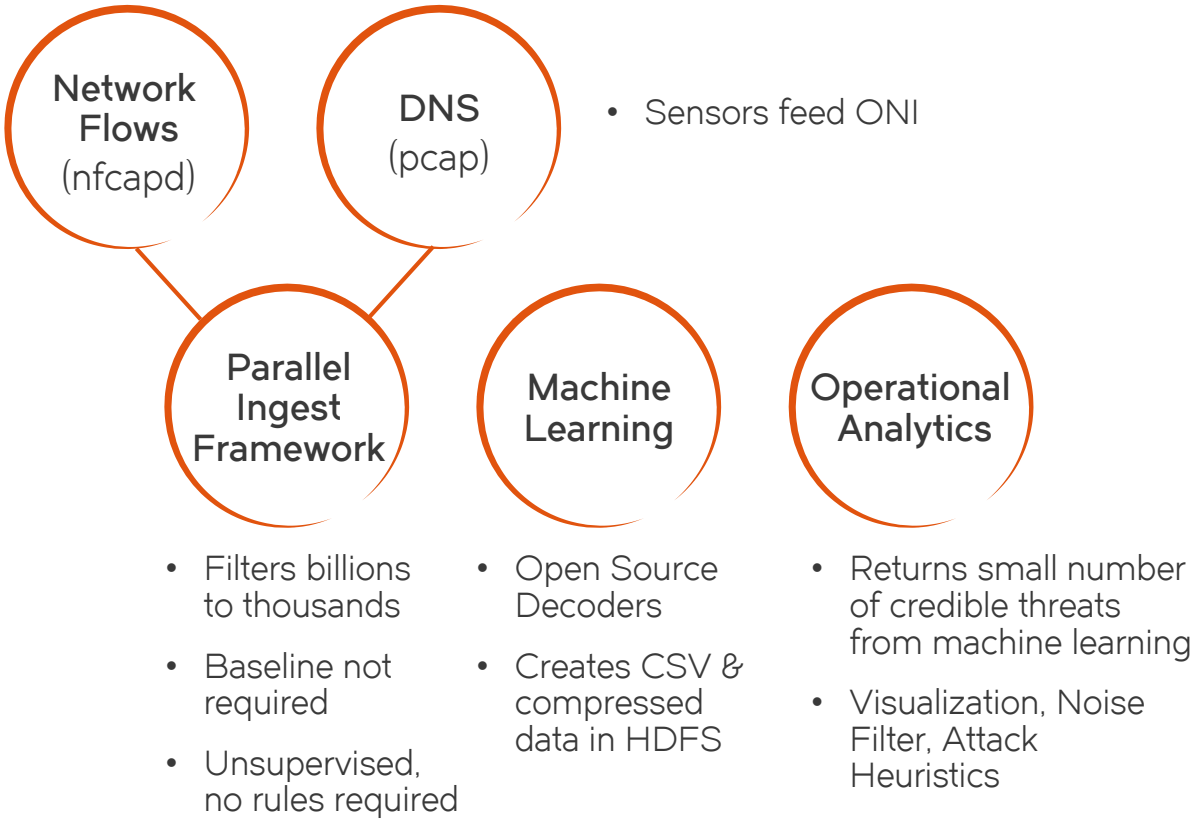
$$f(x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad Z = \frac{X - \mu}{\sigma}$$



- What is obvious: When anomalous events do occur, their consequences can be quite serious and often have substantial negative impact on our businesses, security, …

Open Network Insight Components

Anomaly detection for network data

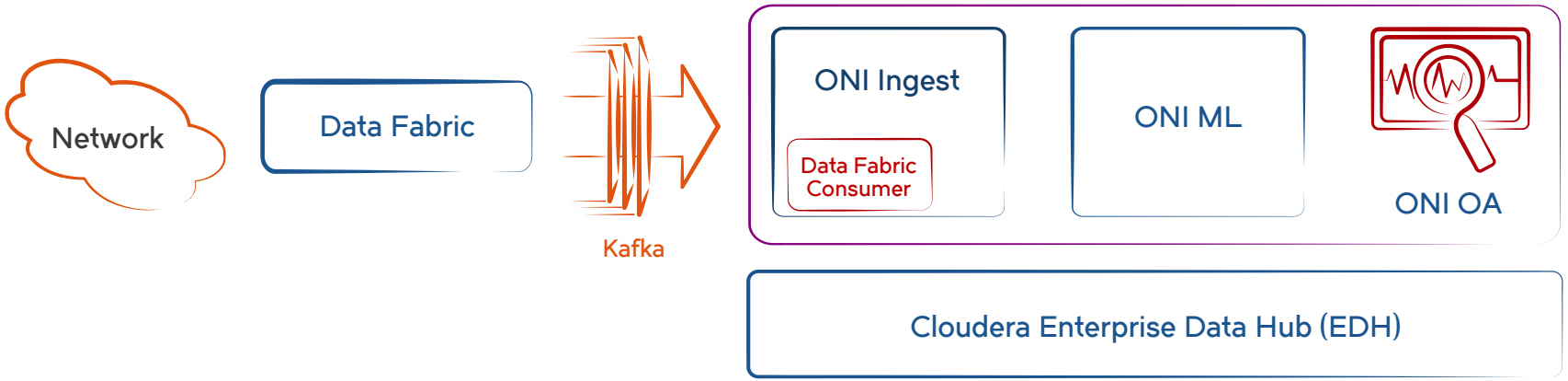


Open Network Insight uses machine learning as a filter to detect anomalies and to characterize the unique behavior of network traffic

http://open-network-insight.org/

Brocade Data Fabric and Open Network Insight

Demo deployment

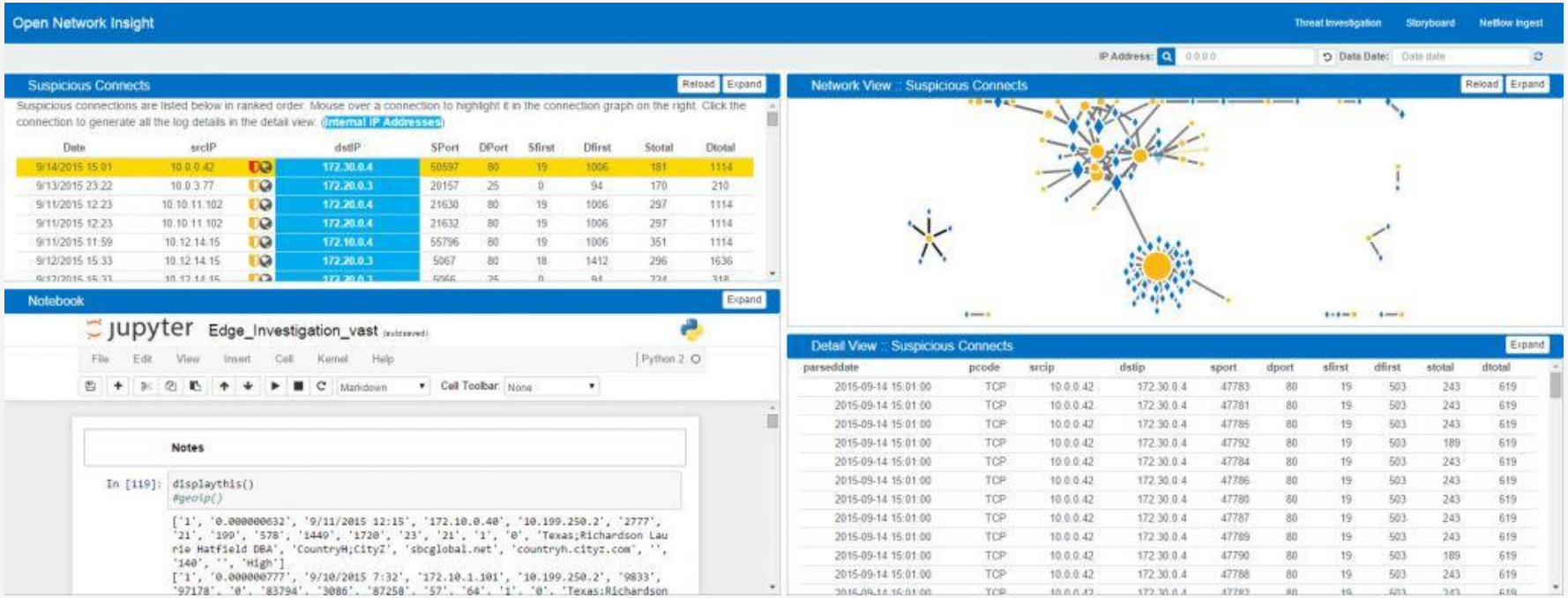


Brocade Data Fabric and Open Network Insight Deployment

GRAPHIC COURTESY VARMA BHUPATIRAJU

Don’t Miss Our Demo

Brocade Data Fabric Enabling Real Time Data Ingestion by Open Network Insight



BROCADE

Thank you