Machine Learning in Engineering Applications and Trends

Sebastian Pokutta

David M. McKenney Family Early Career Professor
Associate Director for Research, Center for Machine Learning @ GT
Director, Laboratory for Interactive Optimization and Learning

Georgia Institute of Technology

NASA Workshop
Machine Learning Technologies and Their Applications to Scientific and Engineering Domains Workshop
August, 2016

Georgia Tech

Machine Learning

Why now?

Gartner predicts that by 2017, 20% of all market leaders will lose their number one position to a company founded after the year 2000 due to a lack of digital business advantage.

Overview

Part 0. **Machine Learning @ GT**

Part 1. Overview Data Science and Machine Learning

Part 2. Current Trends and Game Changers

Part 3. Success Stories

Machine Learning @ GT Center

An effort to focus ML resources on Campus

A joint effort of Computing, Engineering, and Sciences on GT Campus.

- Effort to unify and focus Machine Learning expertise on GT campus
- Brings together 50 - 80 faculty on campus involved in Machine Learning, Analytics, and Data
- Facilitate interaction of industry and other outside entities with ML @ GT
- Catalyst to define our leadership in Machine Learning
  - Strong focus on combining Computing, Engineering, and Sciences
  - Application focus areas: Aerospace, Manufacturing, Logistics/Supply Chains, Mechanical Eng, Industrial and Systems Engineering, …
- Strong focus on collaborations with industry and government to translate innovation
- Leadership.
  - Irfan Essa, College of Computing (Director)
  - Sebastian Pokutta, College of Engineering (Associate Director for Research)
  - Justin Romberg, College of Engineering (Co-Associate Director for Academics)
  - Karim Lounici, College of Sciences (Co-Associate Director for Academics)

Overview

Part 0. Machine Learning @ GT

Part 1. **Overview Data Science and Machine Learning**

Part 2. Current Trends and Game Changers

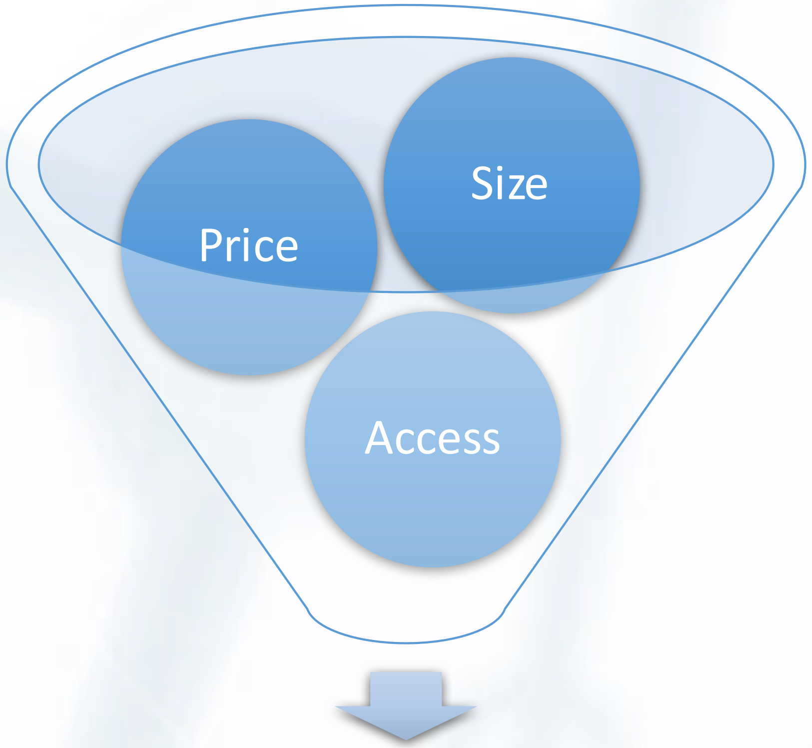
Part 3. Success Stories

Data Science, Machine Learning, and Analytics

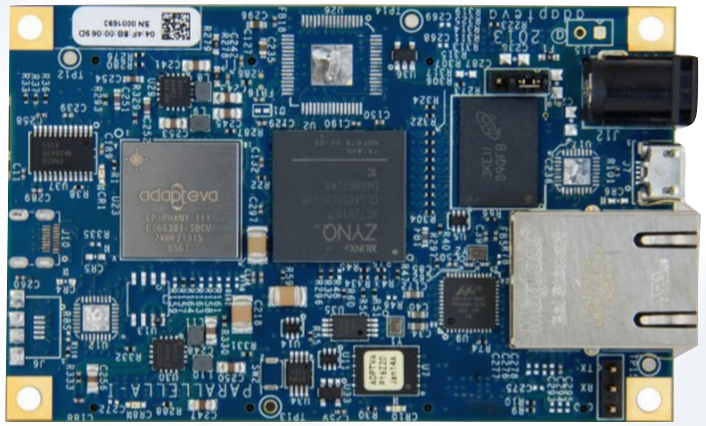
Convergence of three key enablers

Three major factors accelerated Machine Learning.

- Advances in Computing (Hardware)
  - Extreme performance via GPU computing
  - Very small and cheap
- Advances in Algorithms (Software)
  - New generation of Machine Learning algorithms
  - Deep Learning and Reinforcement Learning
- Advances in Sensor Technology (Data)
  - High-performance and cheap sensors
  - Large amounts of data



Disposable, in-situ sensing and computing

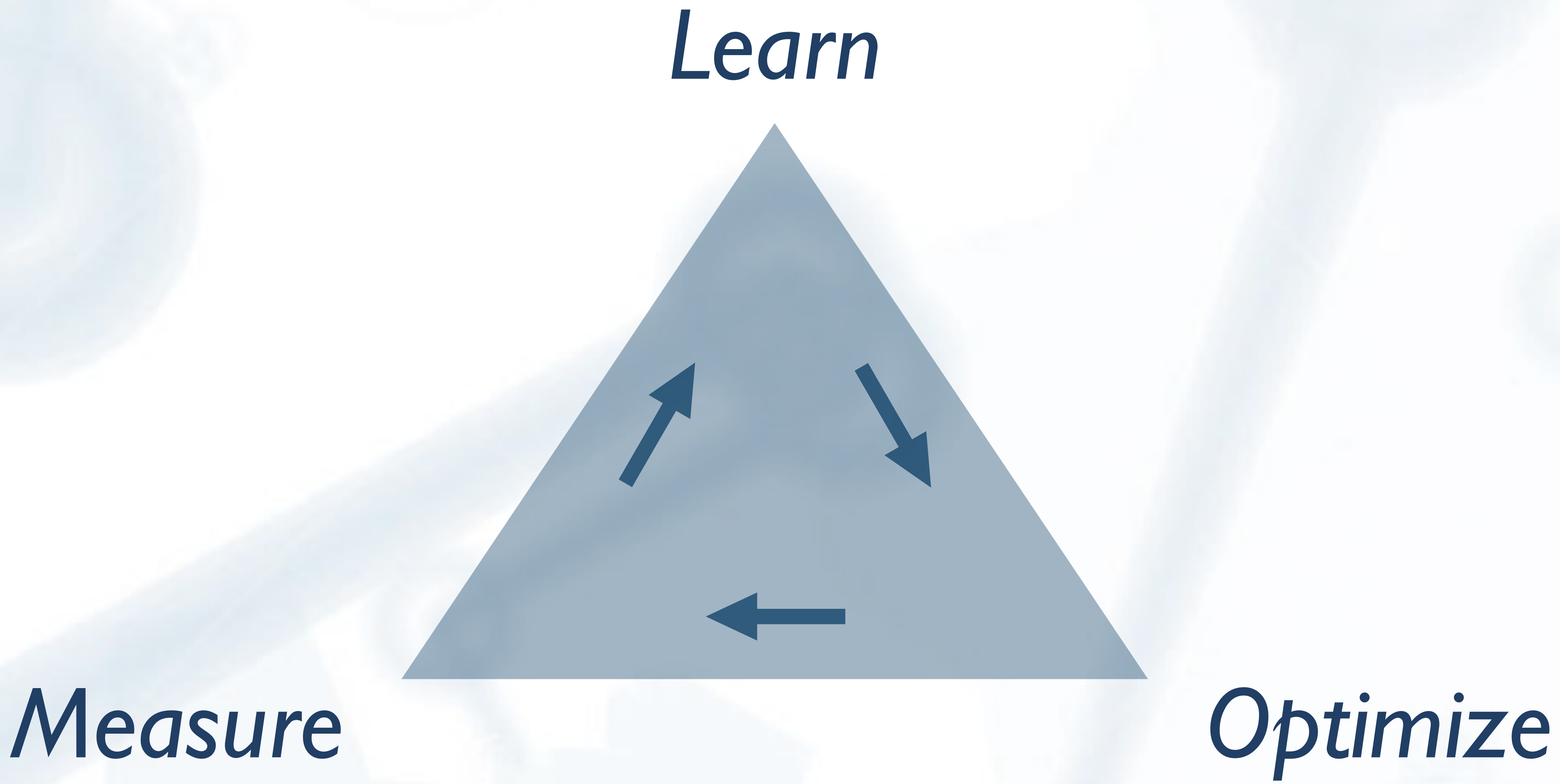


Parallela Board.
18 cores, 1 GB RAM
$149.00

Data Science, Machine Learning, and Analytics

Feedback Loop: Measure, Learn, Optimize

Machine Learning = Gaining insight from Data using Computers



Data Science, Machine Learning, and Analytics

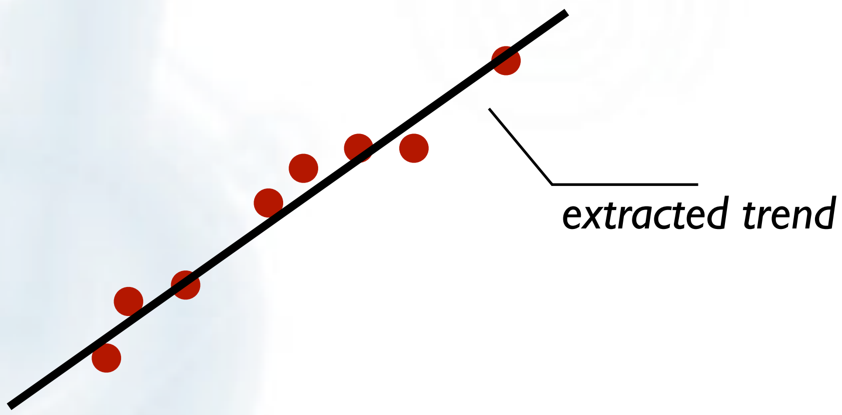
Feedback Loop: Measure, Learn, Optimize

Machine Learning = Gaining insight from Data using Computers

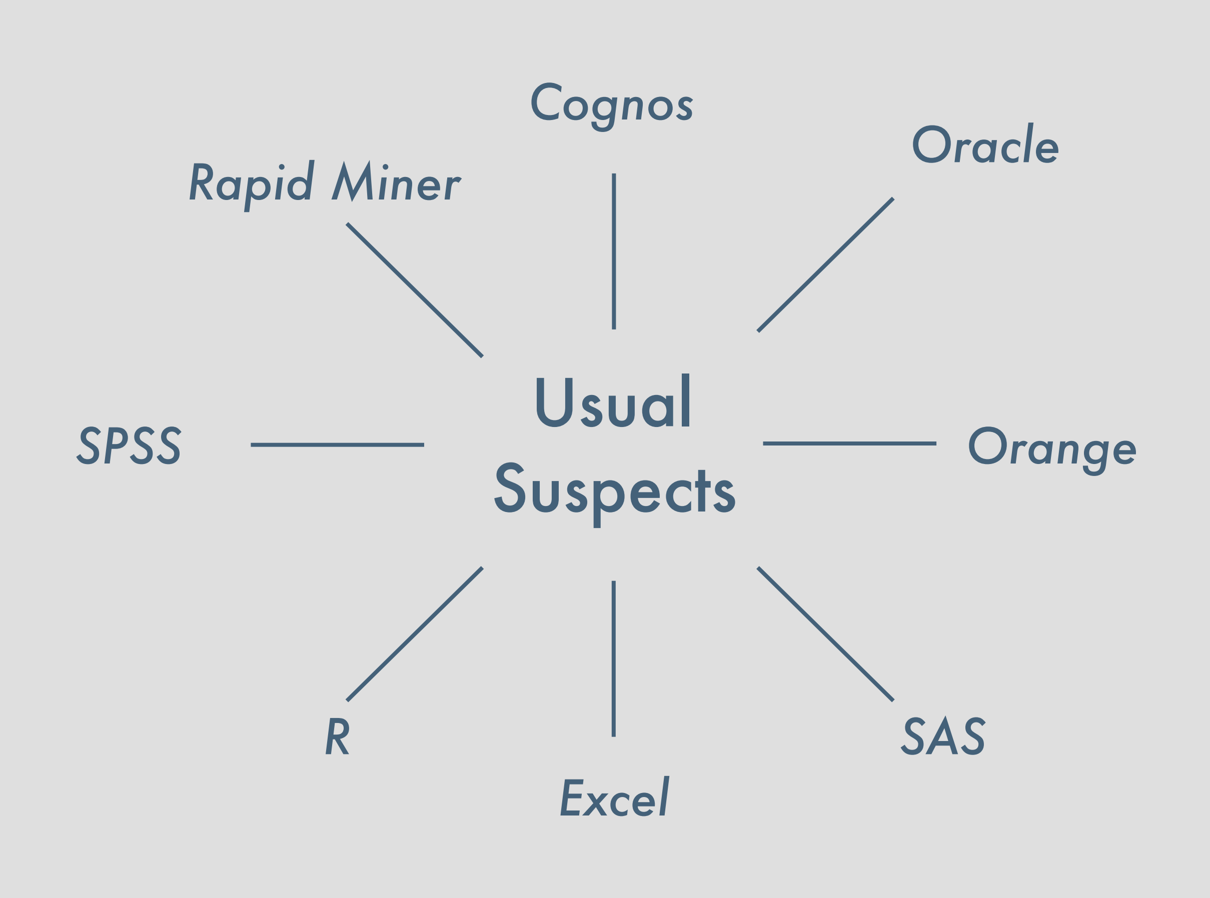
Data Analysis and Learning

Data-driven discovery

"If it is real it is in the data"



The machine learns

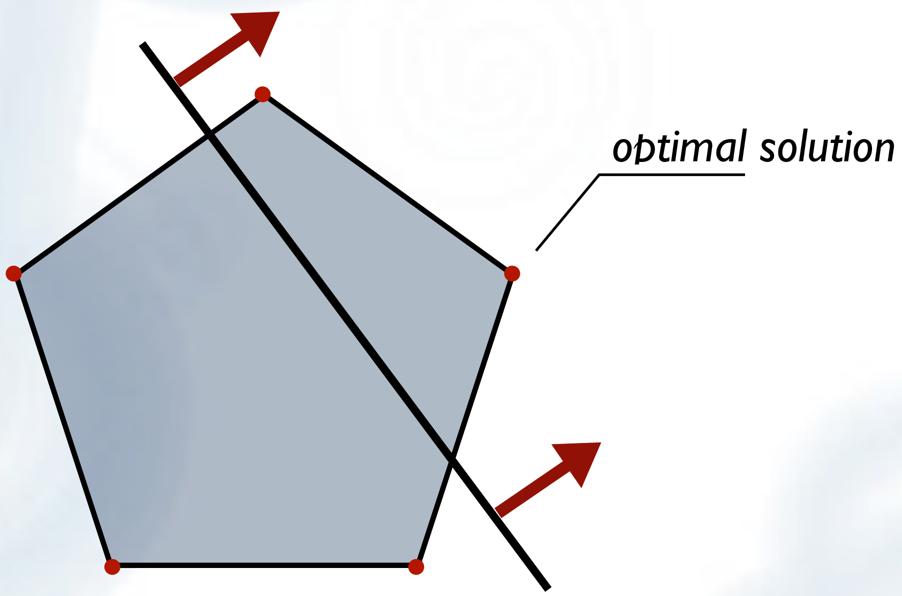


- Data analysis and curation is the basis for all other quantitative methods
  - Data consistency throughout company is key (master scales, data warehouses, etc.)
- Typically, weakest link: industry is not collecting the right data which inhibits use of analytics
- Recent trends from description to learning
  - machine learning at several large companies

Decision Making and Optimization

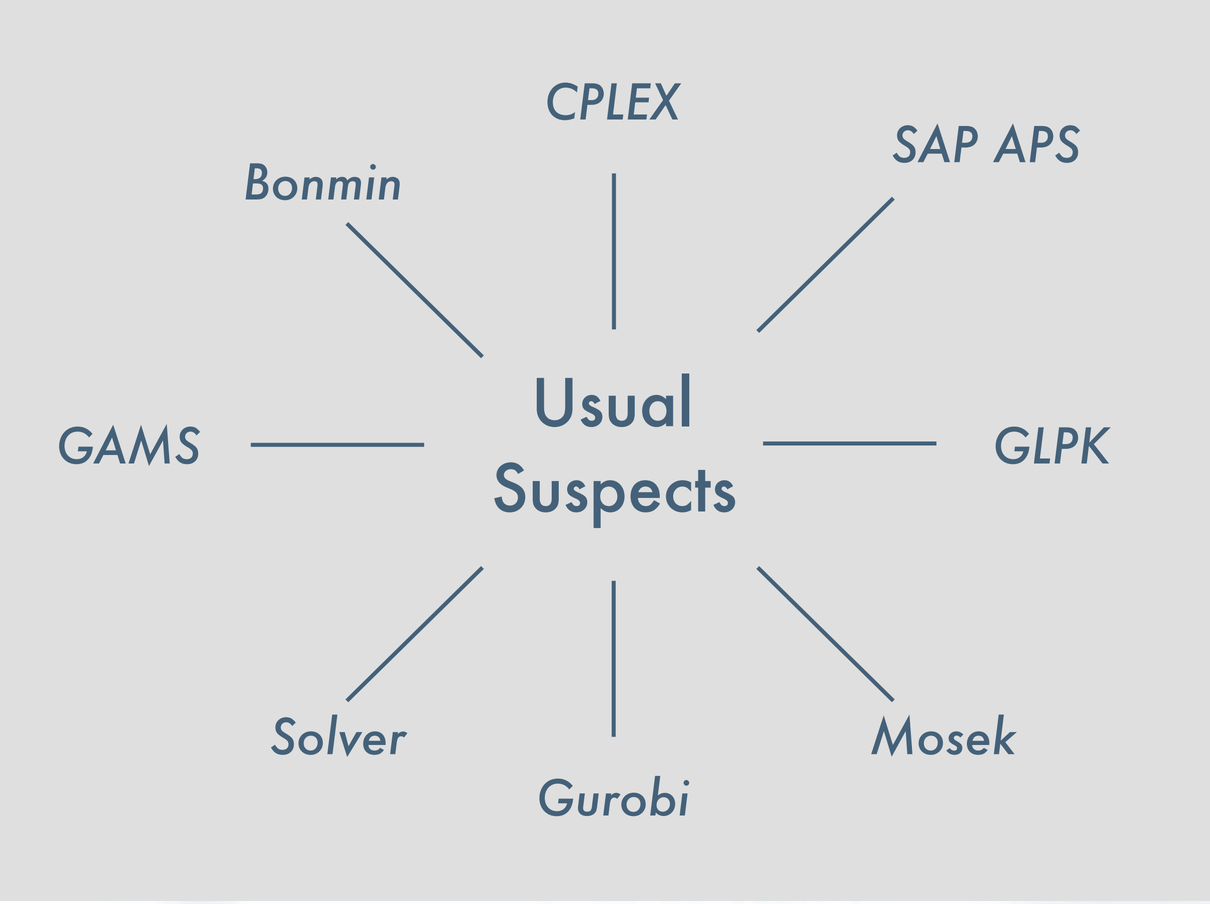
Optimal decisions

"Given current and future operating constraints what are the optimal decisions"



- A lot of production-ready methods available
  - black-box solvers that get a standardized problem file
- Very efficient for real-world problems (up to millions of decision variables)
- Dispatching/scheduling-heavy industries (e.g., airlines) rely on optimization

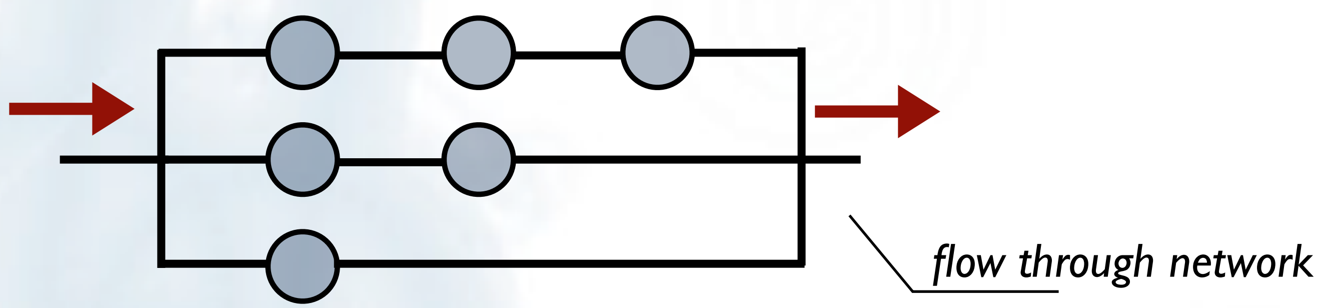
The machine decides



Scenario Analysis and Simulation

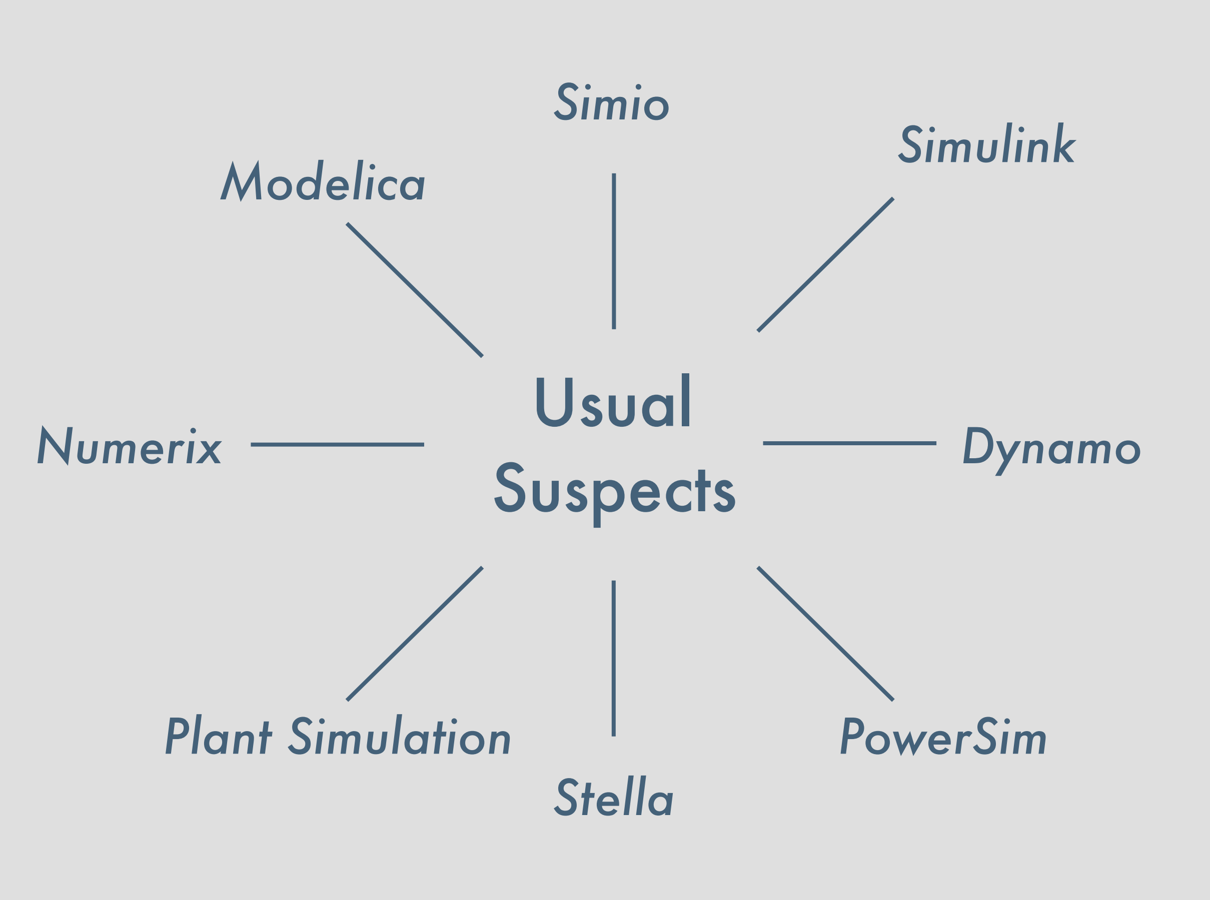
Exploring complex systems

"Complex systems typically do not admit closed-form solutions"



The machine explores

- Scenario analysis is a basic form simulation
- Simulation plays key role to model material flow through facilities
- Allows for exploring responses of dynamic systems to changing parameters
- Standard tool in Engineering (FEM), Banking (Pricing and Risk Management), and Supply Chain Management (Material Flow)



Overview

Part 0. Machine Learning @ GT

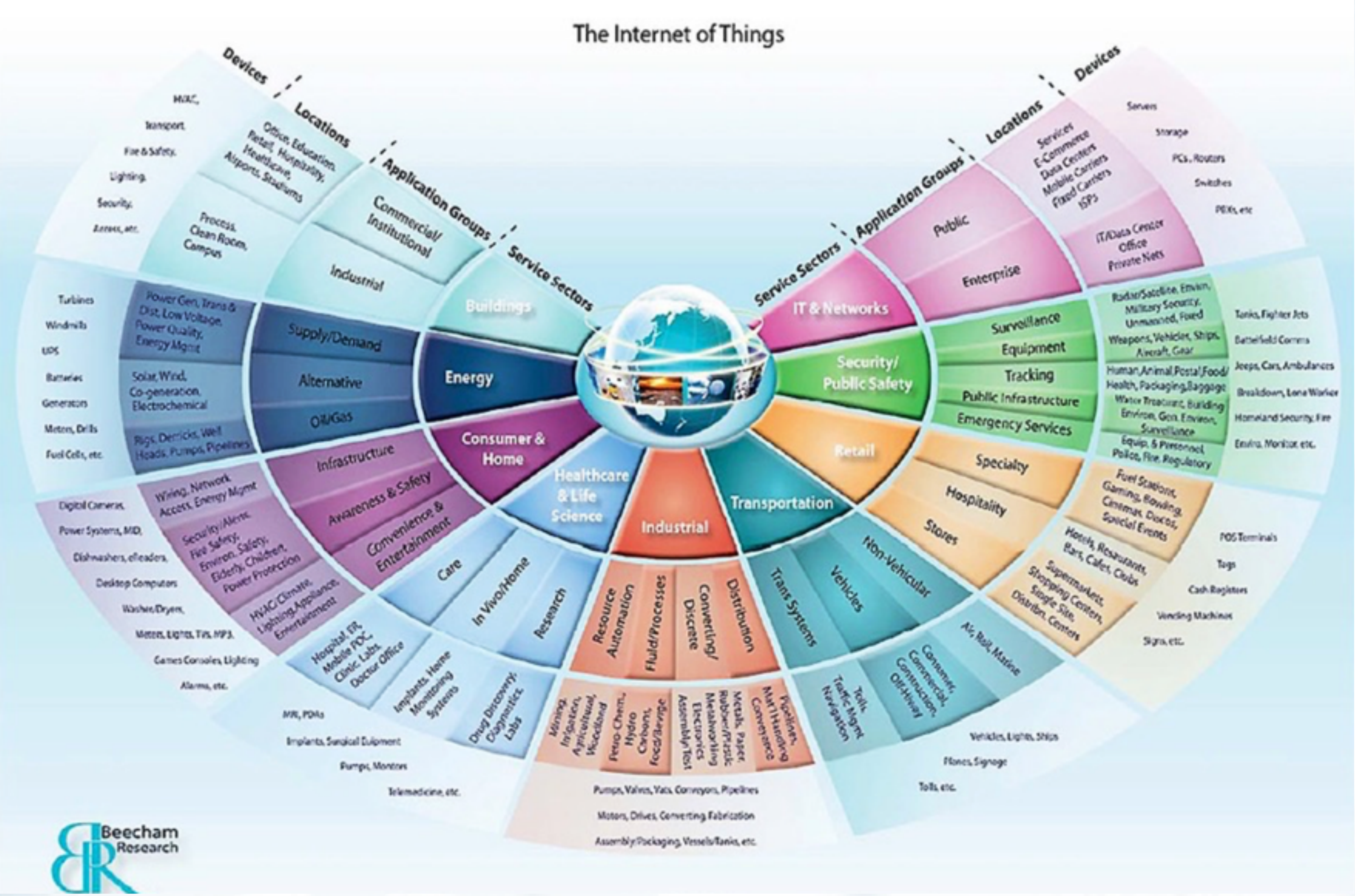
Part 1. Overview Data Science and Machine Learning

Part 2. **Current Trends and Game Changers**

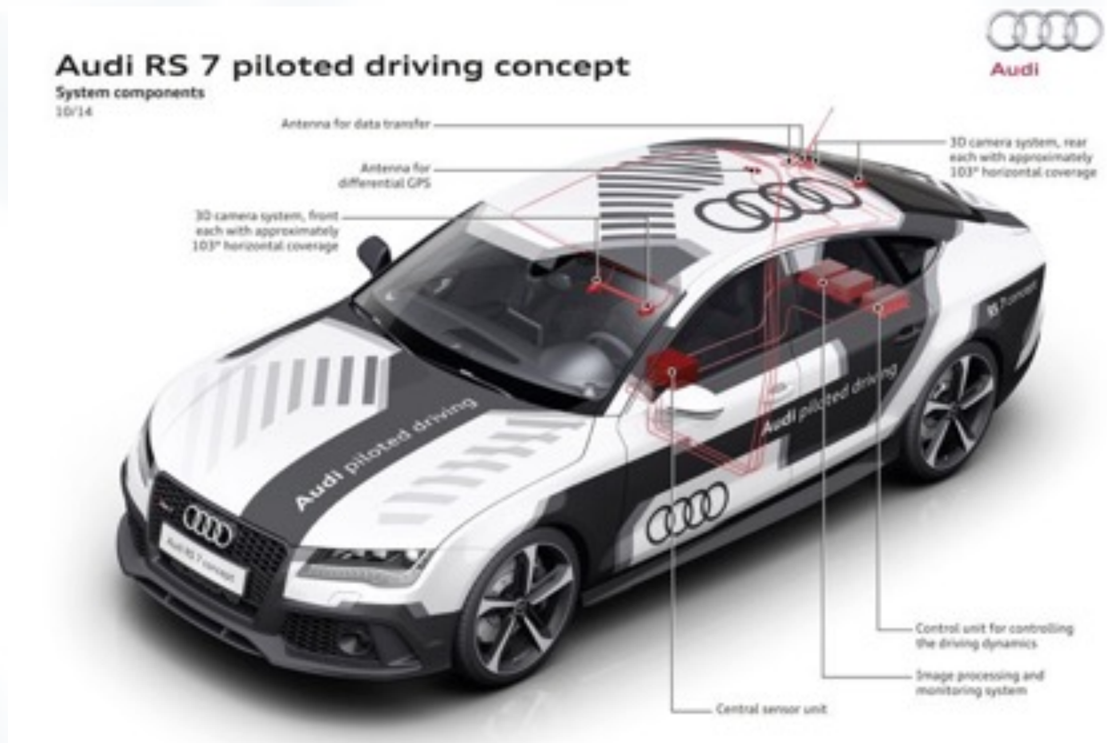
Part 3. Success Stories

Cyber-Physical Systems

Industry 4.0, Industrial Internet, and Internet of Things



Cyber-Physical Systems
Autonomous Vehicles



Cyber-Physical Systems = Machine + Sensors + Computing

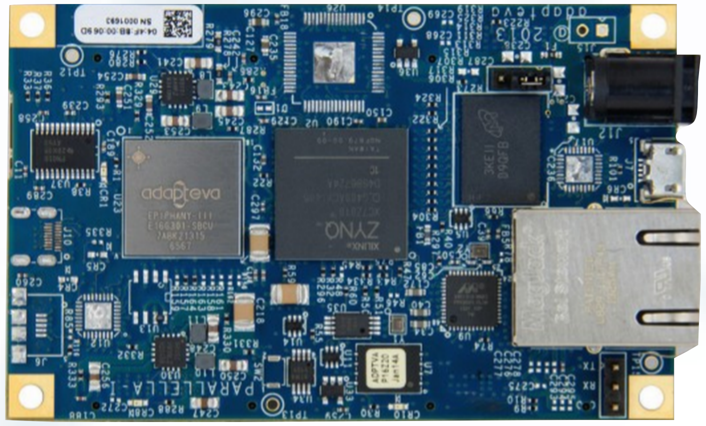
- Robotics and Intelligent machines (self-driving cars, drones, material handling, …)
  - Motivation: create *truly* intelligent machines
- Autonomous vehicle are a prime example of the fusion of physical and digital
  - Most technical challenges considered to be solved
- Many companies work on a car.

Cyber-Physical Systems
In-Situ Machine Learning

Ultra-smart embedded systems.

- Process signals and data right where the sensors capture it
- Low energy consumption and price point
- Very high performance
- *Jetson TX1*
  - embedded GPU enabled for deep learning
  - 256 cores and 4 GB RAM
  - up to 1 TFLOP/s GPU performance @ 10 W energy cons.
- *Parallela board*
  - 18 cores and 1 GB RAM
  - up to 32 GFLOP/s @ 5W energy cons.
- *Fathom Neural Compute Stick*
  - VPU for Embedded Neural Networks
  - up to 150 GOPS/s @ < 1W energy cons.
  - USB plug-and-learn

NVIDIA Jetson TX1
256 cores, 4 GB RAM. $300.00 (est)



Parallela Board.
18 cores, 1 GB RAM. $149.00

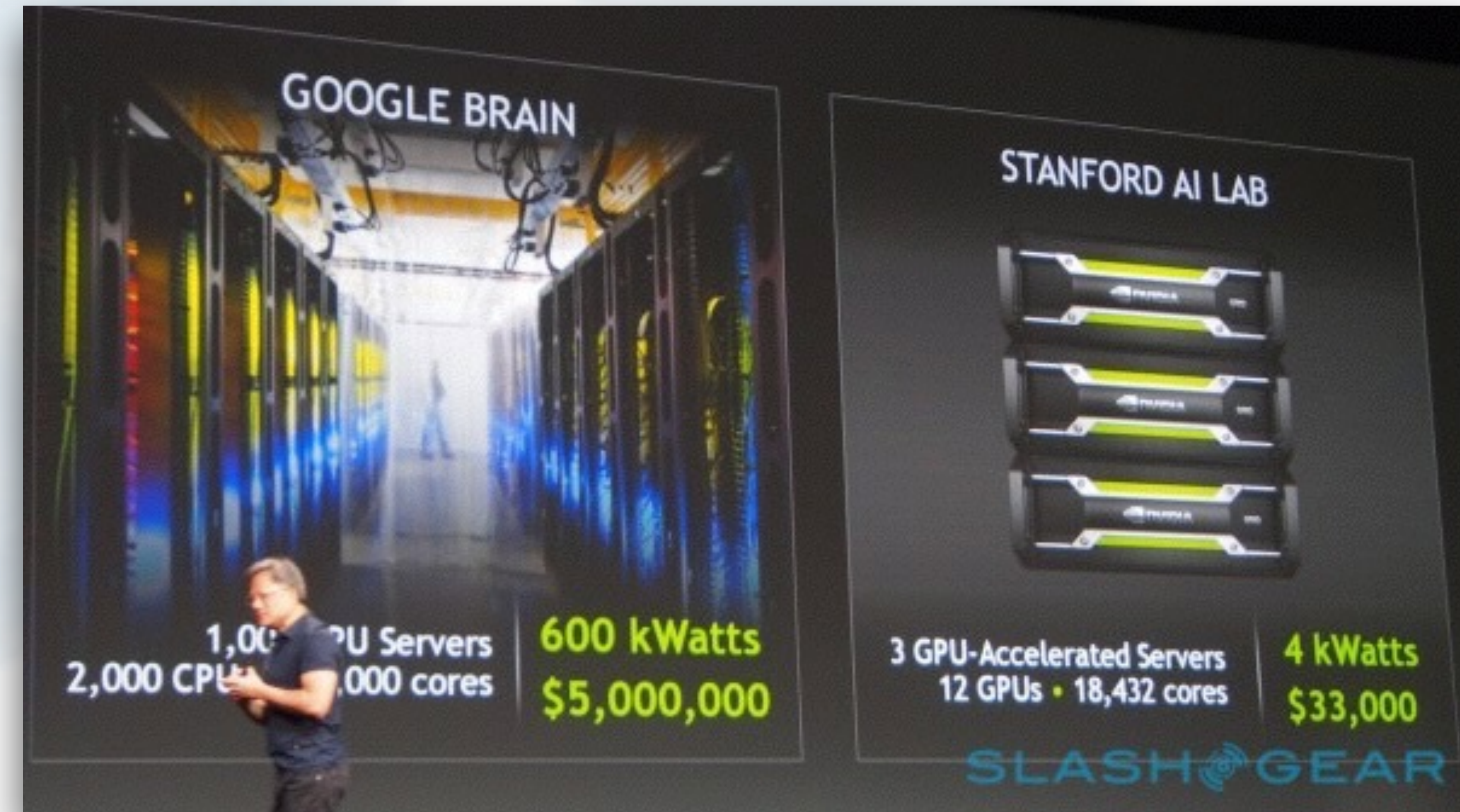
Fathom Neural Compute Stick
VPU, 512 MB RAM. $99.00 (est)

Deep Learning

A revolution in Machine Learning

2012 2014

"custom-made"

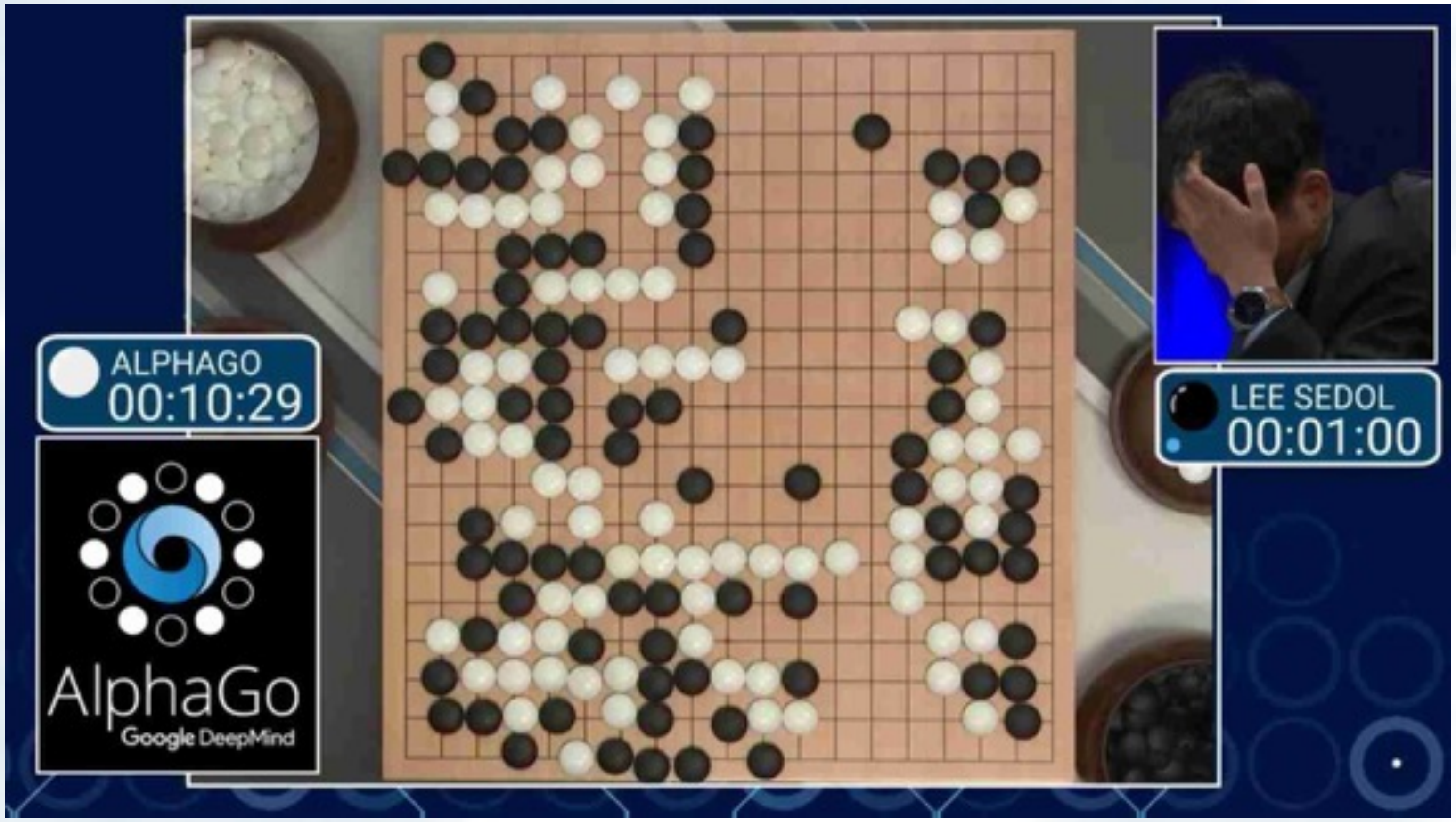


"off-the-shelf"

High-accuracy real-time image recognition: cats vs. dogs

- GPU based machine learning is a huge trend
- cheaper and extreme performance
- 1.2m training images
- 2 weeks training time = 25 exaflops to train system
- **Impossible 5 years back**

Deep Learning
A revolution in Machine Learning



"Go is a complex board game that requires intuition, creative and strategic thinking. […] Many in the field of artificial intelligence consider Go to require more elements that mimic human thought than chess."
Mathematician I. J. Good in 1965

AlphaGo's victory was a major milestone in artificial intelligence research.

- Go is extremely complex and cannot be solved via enumeration (unlike Chess)
- Compared to Deep Blue or Watson, AlphaGo's underlying algorithms are more general-purpose => potential evidence for progress toward artificial general intelligence
- **Go was believed to be outside of the realm of current technology by most experts**

Deep Learning

A revolution in Machine Learning

NVIDIA DGX-1
170 TFLOP/s, $130,000

Huge trend: Dedicated Machine Learning Hardware for Deep Learning applications

- Extreme performance: 170 TFLOP/s @ 3200W in 3U unit
  - 24 x faster than Titan X (state of the art GPU, 7 TFLOP/s)
  - 250 x faster than standard x86 server (two-socket Intel Xeon E5-2697 v3)

- All production capacity of NVIDIA has been absorbed by hyper-scalers up to end of 2017
  - Huge strategic advantage for these companies
  - Ability to solve problems that are inaccessible to other approaches

- **Machine Learning Arms Race has started**

Overview

Part 0. Machine Learning @ GT

Part 1. Overview Data Science and Machine Learning

Part 2. Current Trends and Game Changers

Part 3. **Success Stories**

Applications of Machine Learning

Overview

- Predictive maintenance or condition monitoring
- Warranty reserve estimation
- Propensity to buy
- Demand forecasting
- Process optimization
- Telematics

Manufacturing

- Predictive inventory planning
- Recommendation engines
- Upsell and cross-channel marketing
- Market segmentation and targeting
- Customer ROI and lifetime value

Retail

- Alerts and diagnostics from real-time patient data
- Disease identification and risk stratification
- Patient triage optimization
- Proactive health management
- Healthcare provider sentiment analysis

Healthcare and Life Sciences

- Aircraft scheduling
- Dynamic pricing
- Social media – consumer feedback and interaction analysis
- Customer complaint resolution
- Traffic patterns and congestion management

Travel and Hospitality

- Risk analytics and regulation
- Customer Segmentation
- Cross-selling and up-selling
- Sales and marketing campaign management
- Credit worthiness evaluation

Financial Services

- Power usage analytics
- Seismic data processing
- Carbon emissions and trading
- Customer-specific pricing
- Smart grid management
- Energy demand and supply optimization

Energy, Feedstock, and Utilities

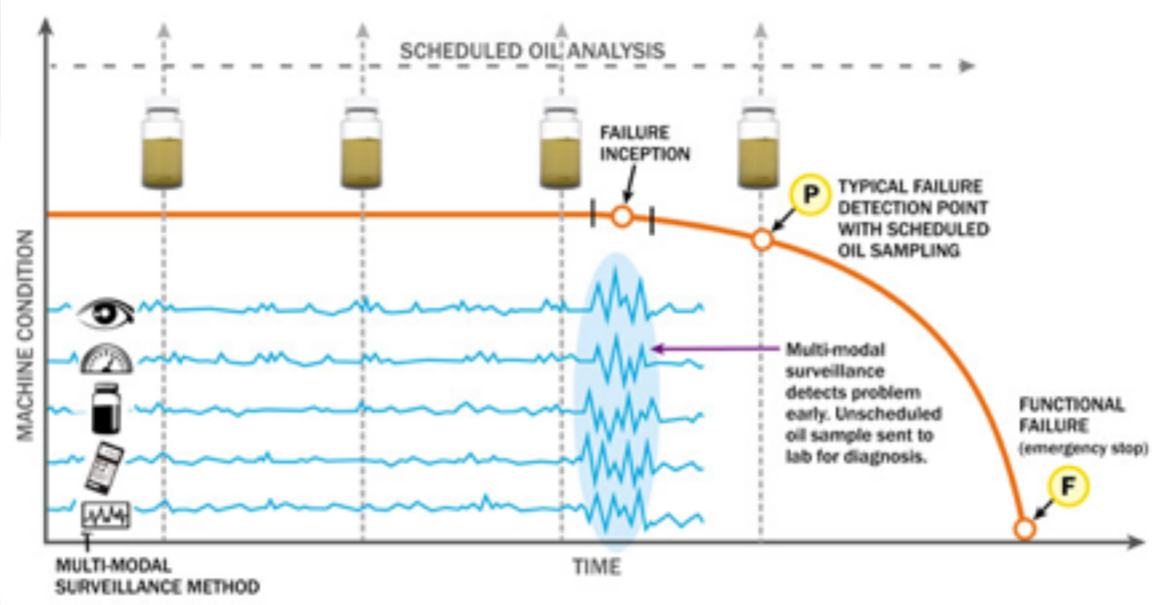
Source: TATA Consultancy Services

Predictive and Prescriptive Maintenance

From reactive to proactive

Three evolutionary stages of maintenance.

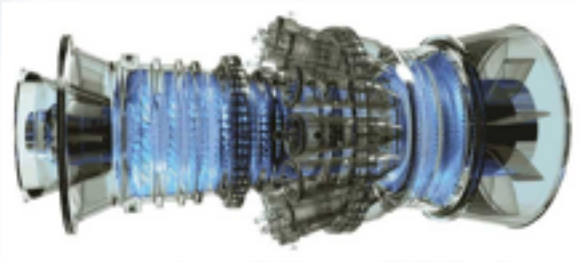
- Reactive maintenance
  - Mostly done today
- Predictive maintenance
  - Monitor system and predict imminent failure
  - Mostly predictive but no optimal decisions
- Prescriptive maintenance
  - Fully-integrated maintenance planning including spare parts logistics and workforce scheduling
  - Integrates machine learning and decision making
- Differentiator and strong value proposition



E-Commerce System

Printing presses

Gas turbines



Predictive and Prescriptive Maintenance

From reactive to proactive

- Goal: Minimize operational cost of assets and improve asset availability
- Preemptive Maintenance to reduce risk of unexpected failure

Typical setup.

- Starting point is a population-based statistical model of the failure time distribution
  - Model derived from historical data
- Sensors collect data about asset condition
  - Challenges arrive from fusing data from thousands of sensors
- Collected data is used to update the model
  - Traditionally, Bayesian approaches to update models
  - More recently, Recurrent Neural Networks (RNNs) to handle learning and updating

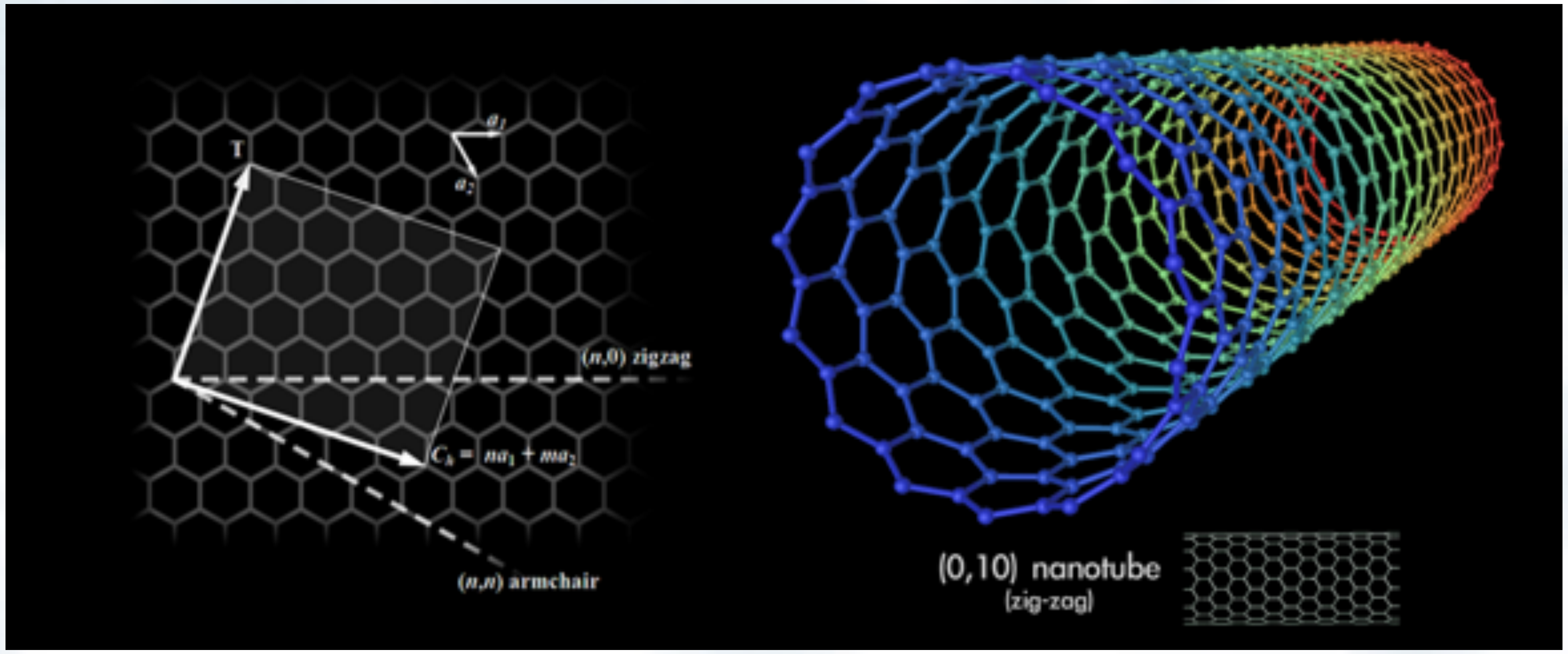
Predictive and Prescriptive Maintenance

From reactive to proactive

The next generation.

- Strong combination with online decision making
  - Dynamically adjust performance parameters and operational envelope as function of asset state
- In-situ learning and processing of data
  - Can handle higher data bandwidth in-situ
  - Sent-off preprocessed data for ex-situ analysis
- Derivation of high-dimensional failure mode features from neural networks
  - Provide compact representations for ex-situ processing
  - Can be fed into other statistical approaches as input

Real-time Manufacturing Optimization
Automatic Exploration and Optimization of Design Space



Example: Floating Catalyst Synthesis Process for Carbon Nanotubes (CNT)

- More than 20 design parameters (continuous and discrete) govern the synthesis process
- Parameters can be adjusted throughout the process
- Various surrogate models have to be learned throughout the experimentation process
- Physics-based models have to incorporated as priors of varying strength

Goal: Maximize yield given constraints on purity, alignment, etc. (scale-up manufacturing)

Real-time Manufacturing Optimization

Automatic Exploration and Optimization of Design Space

Two tasks that have to executed simultaneously

- Learn a model of the synthesis process
  - Predict effect of varying a parameter
  - Critical, as otherwise the whole design space has to be probed
- Determine optimal process parameters and parameter change
  - Optimize e.g., purity, alignment, yield
  - Given various synthesis constraints
- Two types of feedback provided
  - Actual outcome of synthesis process
  - In-line measurements, such as, Raman, x-ray, ccd, tension, furnace temperature

Real-time Manufacturing Optimization

Automatic Exploration and Optimization of Design Space

The next generation.

- Integration of Deep Learning techniques
  - Deep Reinforcement Learning for process control and integrated learning and optimization
  - Convolutional Neural Network (CNN) approaches for image analysis (CCD)
  - Temporal modeling via Recurrent Neural Networks (RNN)

- In-situ learning and processing of data
  - Deploy integrated system GPUs (e.g., Jetson TX-1) directly in the experimentation system
  - Shorter Feedback loops

Thank you for your attention!

Sebastian Pokutta

David M. McKenney Family Early Career Professor
Associate Director for Research, Center for Machine Learning @ GT
Director, Laboratory for Interactive Optimization and Learning

Georgia Institute of Technology