Machine Learning & Data Science:

What does it mean and how it may impact utilities

Joshua West, Manager, Analytics, TEA

Dr. Jian Hu, Research Scientist, Strategic Innovations, TEA

Sarah Valovcin, Structure & Pricing Analyst, TEA
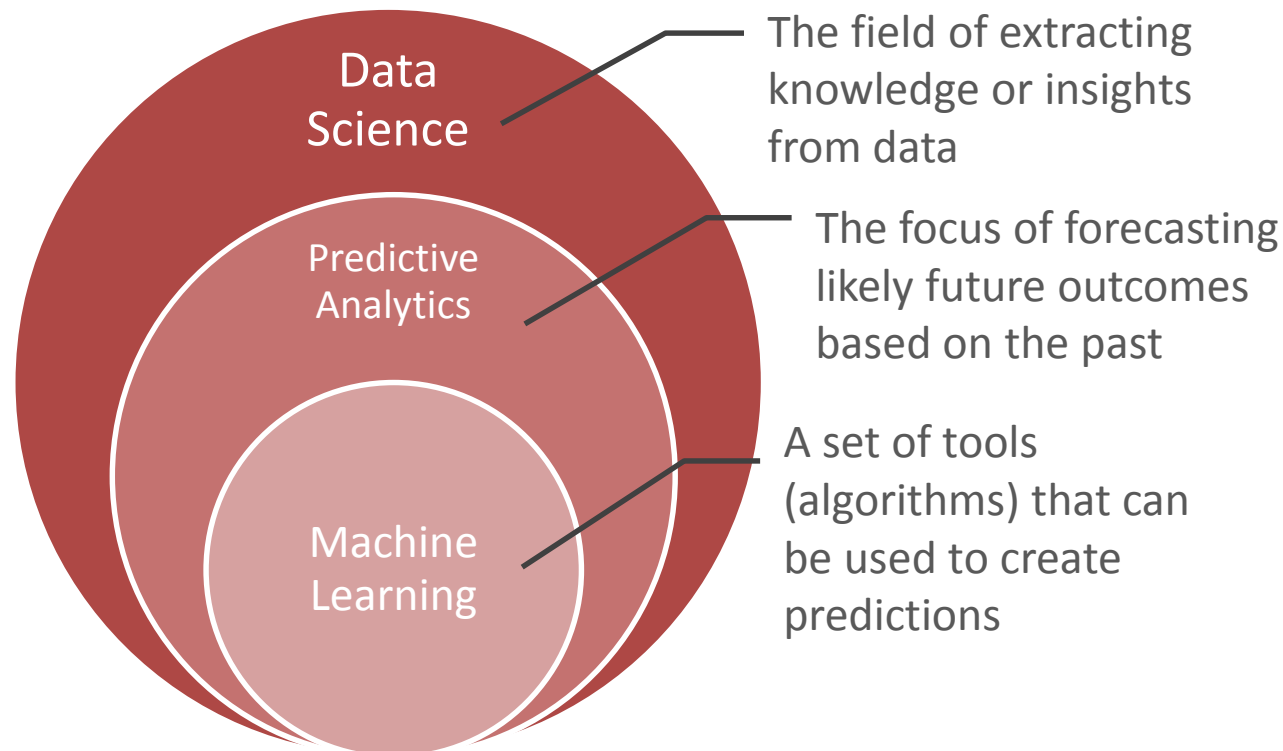
Agenda

- What is Machine Learning?

- Machine Learning at TEA

- The Future of Machine Learning & Data Science at Utilities

THE Energy Authority

Introduction

What are *Machine Learning, Predictive Analytics,* and *Data Science*?

Data Science — The field of extracting knowledge or insights from data

Predictive Analytics — The focus of forecasting likely future outcomes based on the past

Machine Learning — A set of tools (algorithms) that can be used to create predictions

EnergyAuthority

Predictive Analytics

- Can we predict the future based on history?
    - What will the market price be tomorrow?
    - What does the load profile look like in the next 48 hours?
    - How likely is it that a transformer will break in the coming year?

- What we need: **Machine Learning & Data Science!**

Machine Learning

- Teach machines to learn from data
- Many tasks are easy for humans but (traditionally) quite difficult for a computer
    - e.g., to recognize a "cat" is surprisingly HARD!
    - However, lots of data is available
- Think of how babies learn – just show them pictures of cats over and over!
- Machines can process thousands/millions of images to develop the concept (recognize the pattern)


'CAT'

Big Data Analytics

- *Big Data* is a collection of data so large/complex that traditional processing can't be used

- *Big Data Analytics* is the process of uncovering patterns, trends, and other useful information from big data
  - Machine Learning is a set of techniques that we can apply

Machine Learning Applications

- The most disruptive technology nowadays
- When fed the right data, computers become smarter and smarter every day...
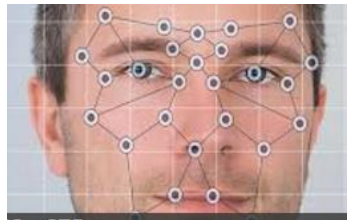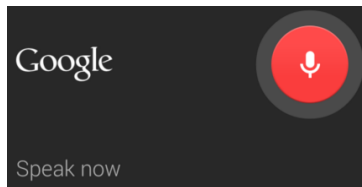
Machine Learning in the News

- Money is pouring into this area with over 1,700 startups…

The Race For AI: Google, Twitter, Intel, Apple In A Rush To Grab Artificial Intelligence Startups

Machine Learning Startups

$5.1M AVERAGE VALUATION

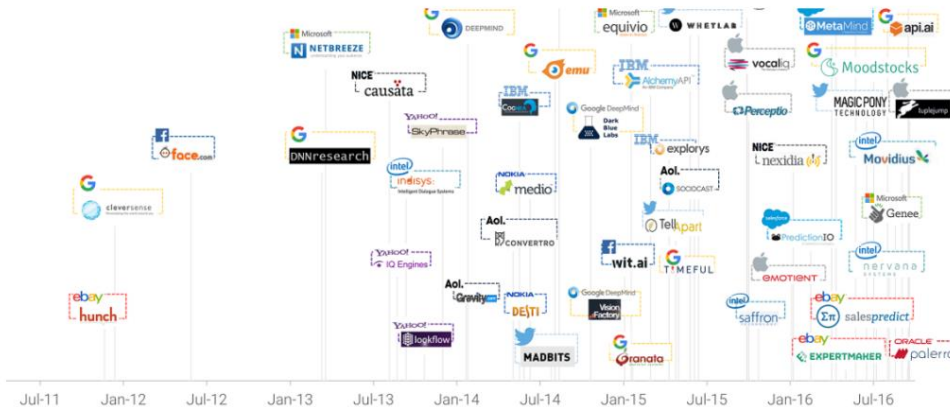1,769 COMPANIES   2,317 INVESTORS   16,391 FOLLOWERS   1,084 JOBS

2017 will be big year for AI thanks to tech giants
CIO - Dec 29, 2016
Machine learning and other variations of artificial intelligence
Facebook, Google and **Microsoft** all open-source or share their

How **Machine Learning** Is Transforming Bioscience Research
Forbes - Jan 12, 2017
The relationship between biology and **machine learning** is not new and has existed for decades, even before data science and machine …

Baidu hires ex-Microsoft exec Qi Lu as president, to lead new AI push for Chinese search giant

THE **Energy**Authority

Questions for Machine Learning

- How likely will a transformer fail prematurely?
  - Age, capacity (MVA), voltage, peak load, time of overloading, lightning, hit by car, exposure to heat…

- Can we tell if a meter is tampered/broken without a field check?
  - Daily usage pattern, signals from meter, customer info, house info…

- Not simple questions – let's learn from data!
  - Apply **machine learning** to identify patterns, statistical correlations, and distill models that can help us make smart decisions

modified wirings

THE **Energy**Authority
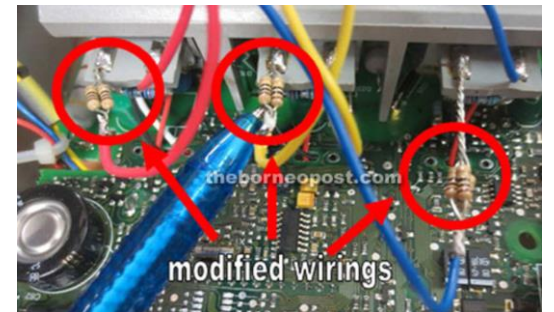
Should Utilities Care about Machine Learning and Data Science?

- First, utilities have lots of data
  - Smart meters
  - Sensors on transmissions, transformers, generators…
- Second, what can we do with the data other than billing and reporting?
- Finally, this is great opportunity to advance your organization and unlock great value
  - Make smart decisions vs. shoot in the dark
  - Data -> Insight -> Action
  - Utility 2.0 – stay ahead of peers

THE EnergyAuthority
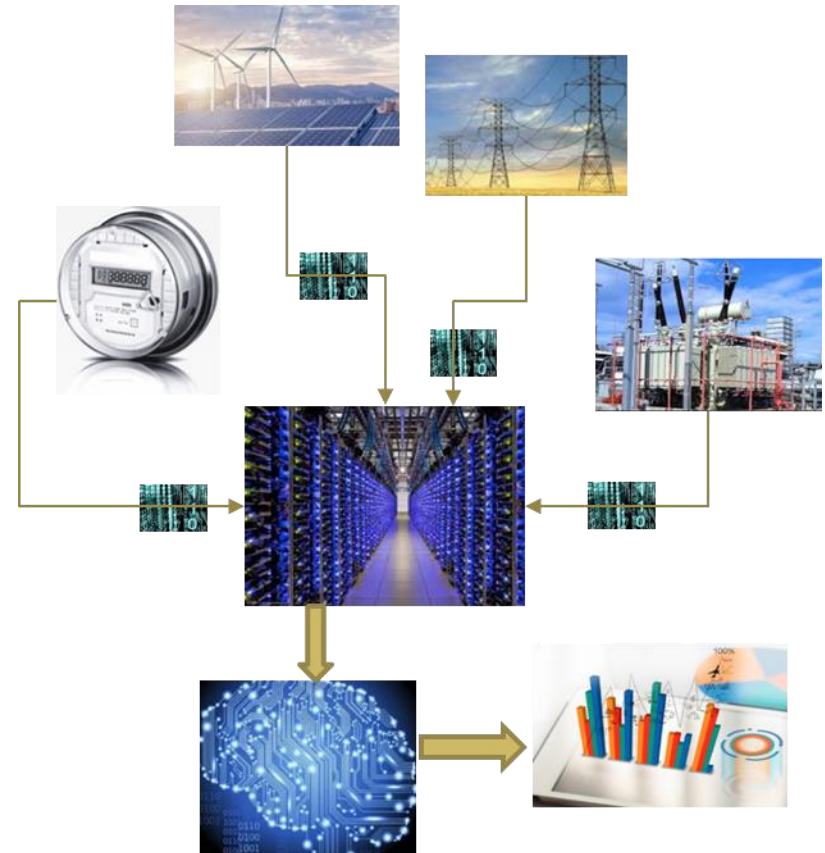
Big Machine Learning Players in Energy

How are Big Players Using Machine Learning?

- Google Nest
  - Smart thermostat learns your schedule and starts automatically optimizing your comfort and consumption
- Mercury Energy
  - GEM program uses smart meter data to predict high bills up to weeks in advance and sends customers alerts
  - Call volumes fell by 9% and customer churn reduced by 10%
- CenterPoint Energy
  - Partnered with IBM to use predictive analytics to help operators detect and respond to changes in infrastructure before customers are affected
  - Improved grid reliability

THE EnergyAuthority

MACHINE LEARNING AT TEA

**THE EnergyAuthority**

Machine Learning Research at TEA

- TEA Analytics as a Service
  - Retail Revenue forecasting
  - LMP forecasting
  - Load forecasting
  - Meter health management
  - Transformer asset management
  - Fraud detection
 - …

THE EnergyAuthority

Day-Ahead LMP Forecasting

- 19,788 models learn every day from previous 91 days of historical data to create forecasts for 69 nodes in MISO, PJM, and SPP
  - Inputs include load, natural gas prices, wind, outages...

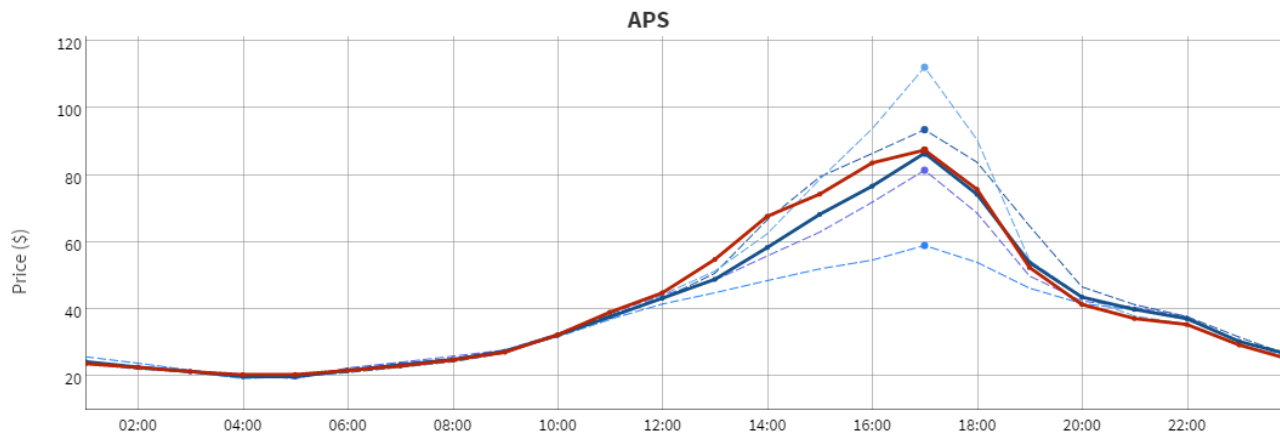| ISO	# LMPs	MAPE
MISO	17	9.2%
PJM	26	10.8%
SPP	26	15.4%

THE **Energy**Authority
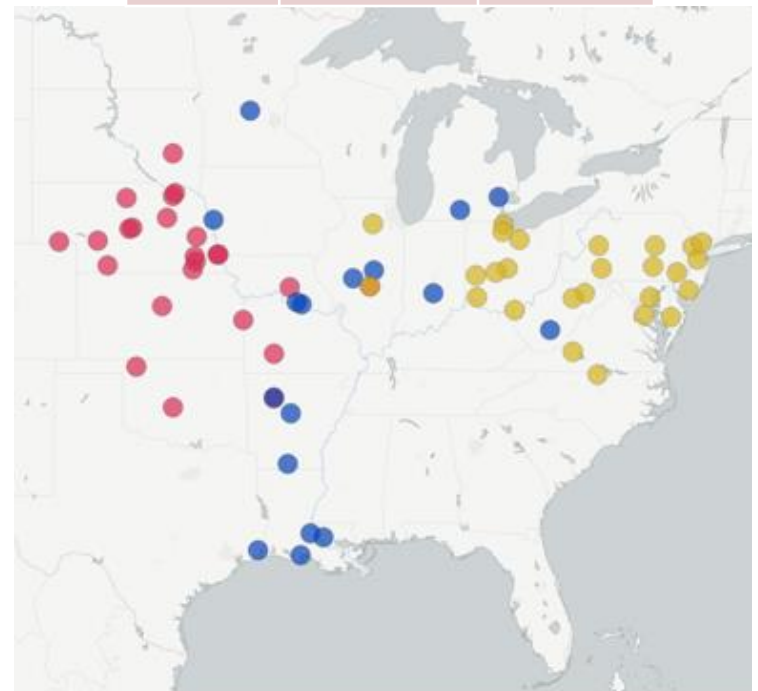
Day-Ahead LMP Forecasting (cont.)

- Ensemble forecast
  - Equally-weighted average of 4 different machine learning algorithms
 - Using an ensemble improves accuracy and stability
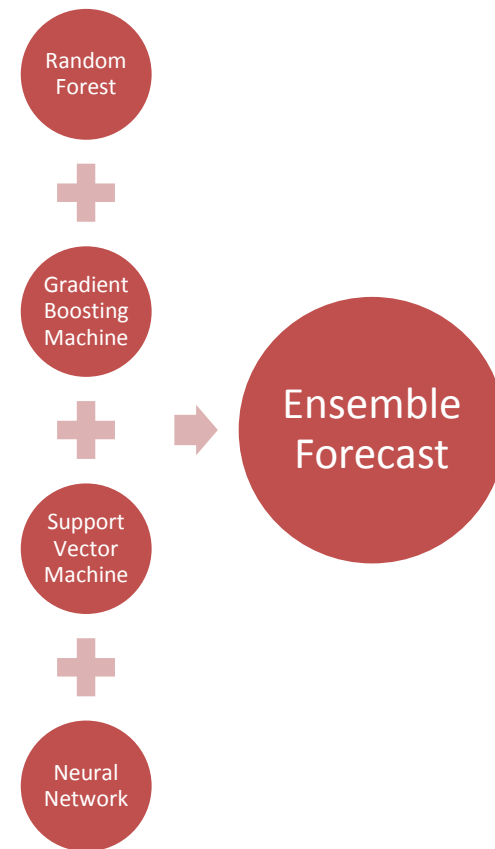
Friday, July 08, 2016 17:00:00 (EDT): **NN**: 81.44 **SVM**: 58.93 **RF**: 93.57 **GBM**: 112.24 **Ensemble**: 86.55 **Actual**: 87.47

Load Forecasting

- Currently redesigning the 10-day hourly load forecast services to use ensemble machine learning methodology
  - Capture complex non-linear relationships
  - Inputs include:
    - Weather
    - Load
   - Calendar variables

THE
EnergyAuthority
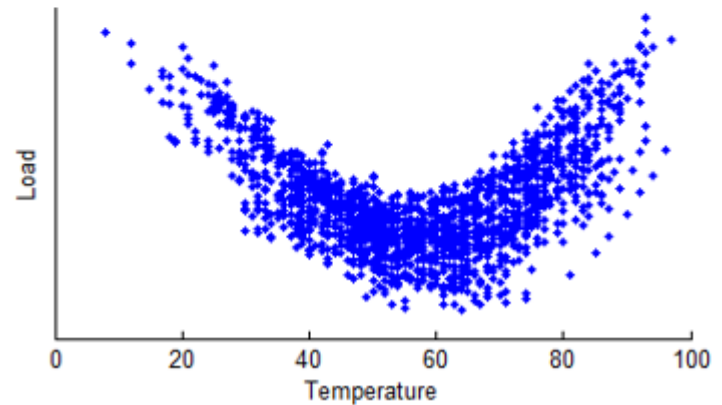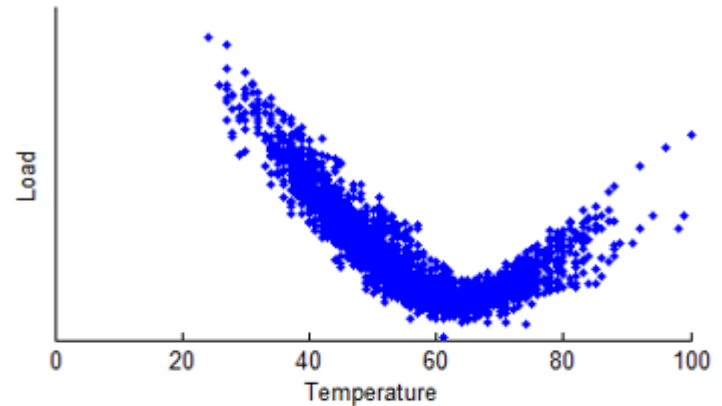
Load Forecasting (cont.)

- SCADA data remains a challenge
  - Errors in meter readings can have a large impact on forecast accuracy
- Use Machine Learning to remove and replace bad readings

THE EnergyAuthority
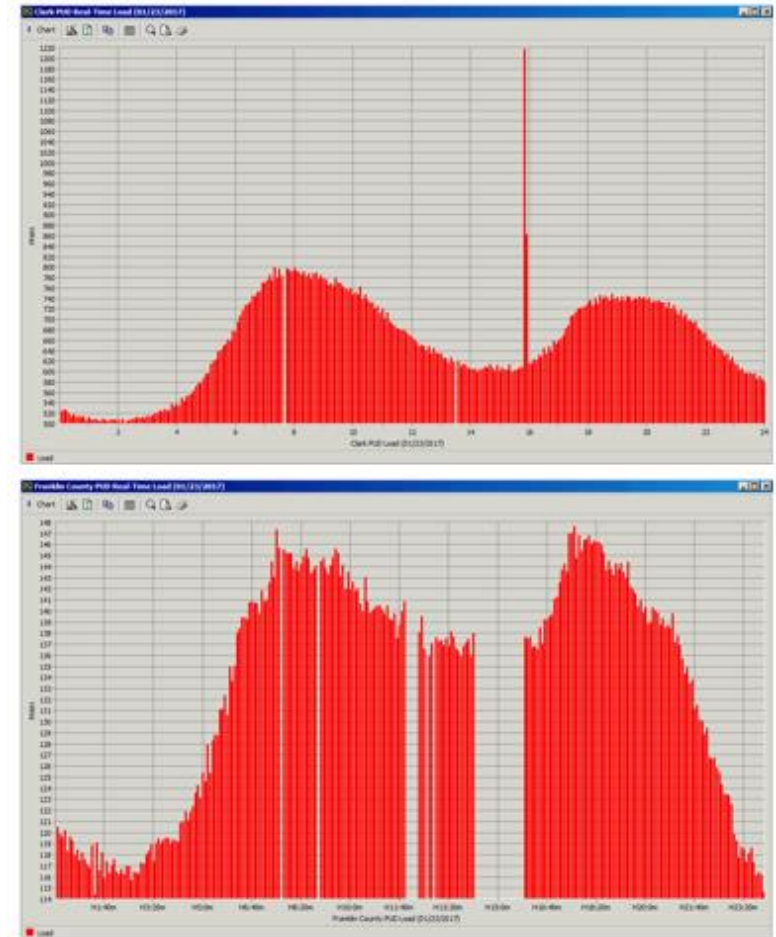
JEA Zero-Con Water Project

- *Question:* *when a water meter continuously reports 0, is it truly broken?*

  – 80% of the 30,000 zero-con truck rolls were ineffective (~$50 per trip)

  – Many broken meters were not be identified in-time due to limited field crews - can we stop the bleeding earlier?

- *Project goal:* *can we reduce 80% of wasted trips to 20%?*
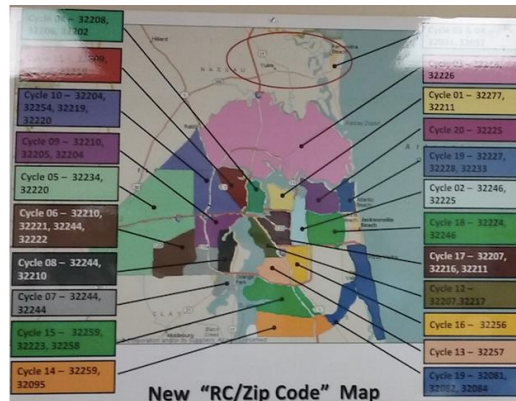
JEA Zero-Con Water Project (cont.)

- Consumes lots of smart meter data from JEA

- Three Supervised Machine learning models created

- Built on top of Microsoft Azure and eSmart Connected Grid, employed both open source and proprietary data science technologies

- Two field tests conducted

  – 89/102 new meters were identified correctly in 12/6 field trip —> 87% accuracy

  – Reduction of Unnecessary Truck rolls > 90%

- Annual economic benefit ~ 500k
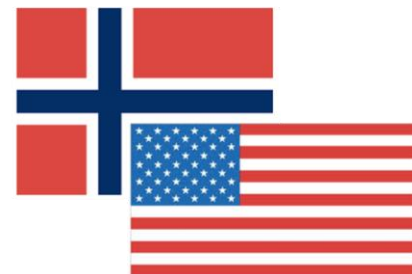
Two Ongoing Projects

- Fraud detection
  - Employ smart meter intervals data, meter signals, voltages etc. to automatically detect meter tampering

- Transformer Load Management
  - Predict the health of distribution transformers based on the transformer load, age, ambient temperatures etc.

EnergyAuthority

THE FUTURE OF MACHINE LEARNING AT UTILITIES

Machine Learning Applications for Utilities

- Load/Price/Revenue/Renewable Generation Forecasting

- Fault Detection/Predicative Maintenance/Asset Management

- Load Segregation

- Fraud Detection

- Improve grid reliability

- Demand Response

- EV integration

- …

THE **Energy**Authority
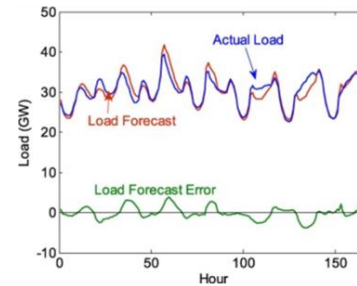
How Do Utilities Embrace
Machine Learning & Data Science?

- Get all data in one place - a data analytic platform

- Work with data scientists who understand both the energy business and data science

- Start small and gradually build it up

    - Identify a clear-scoped pilot project

    - Learn along the way

EnergyAuthority

Typical Data Science Project Workflow
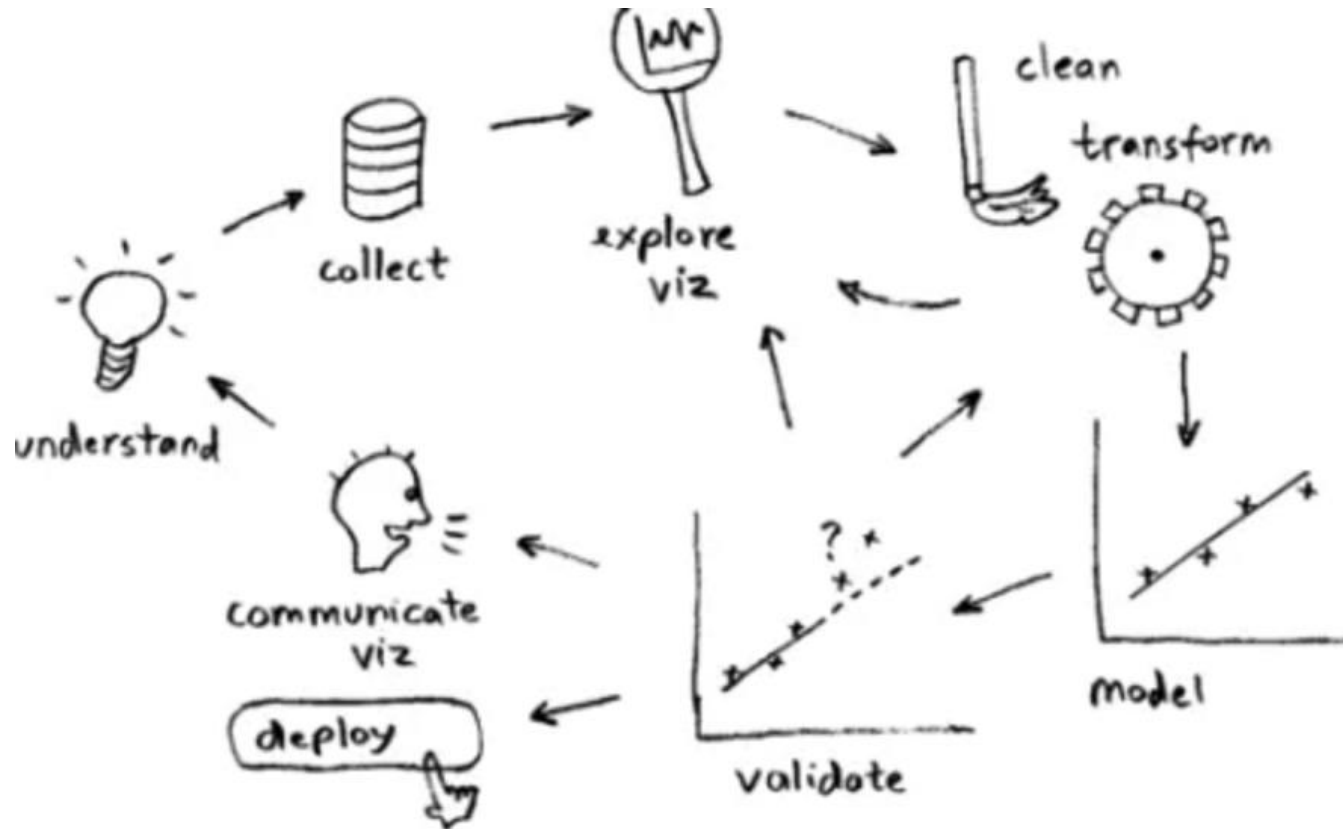
Concluding Remarks

- Machine Learning - teach computers to learn from data
  - Uncover intrinsic patterns, statistical correlations, and distill models that can help us make smart decisions
- Machine Learning & Data Science will revolutionize the energy industry just as it has in other industries
- There is massive opportunity to advance your organization and unlock great value
- **TEA is here to help!**

THE
EnergyAuthority

PARKING LOT SLIDES

The **Energy** Authority

Machine Learning, The Science

- "Sub field of AI that gives computers the ability to learn without being explicitly programmed" - Wikipedia

  - Inspired by the vastly complex structure of human brain, complete paradigm shift from the traditional way of programming

 - Rely on massive amount of data and computational
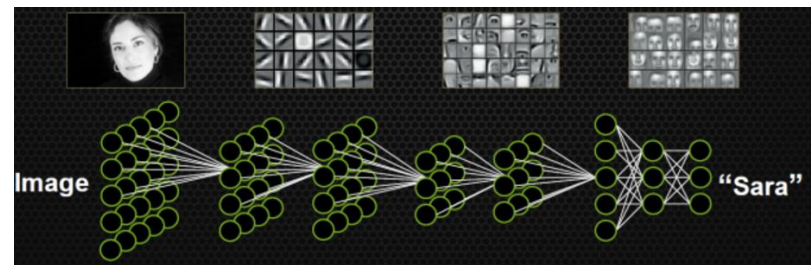
Data Analytics Platform

- Objectives:

  – Get all relevant data in one location and integrate with existing systems (head-end system, MDM, GIS)

  – Empower data analysis to get insight

  – Giga, Tera or even Peta bytes - performance at scale

- On-premise vs Cloud
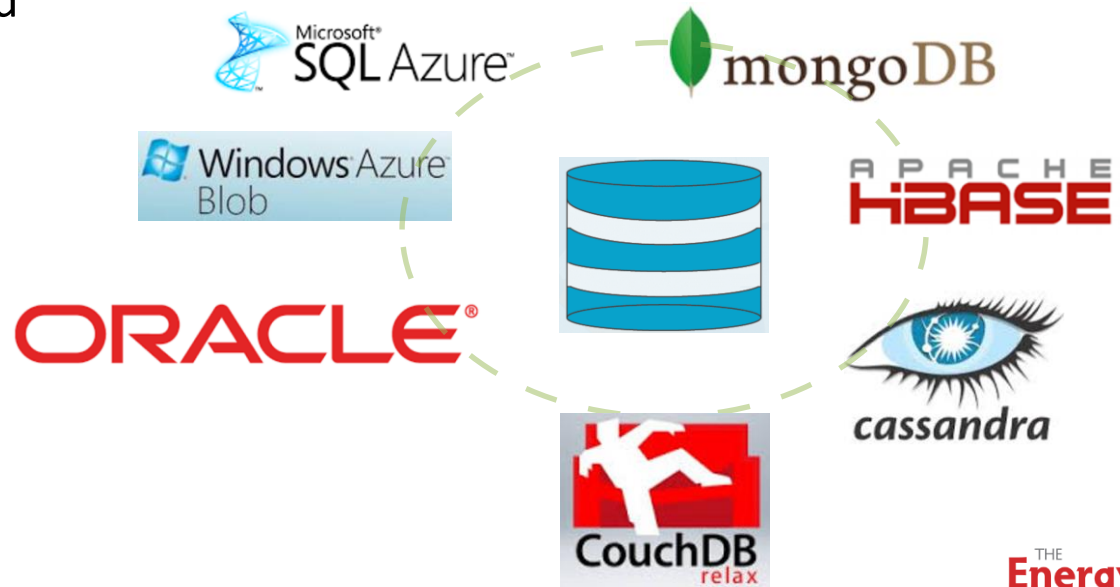
A Data Scientist

- Someone who turns data into actionable insight

- Someone who can navigate through the huge data jungle

- Some who can put all puzzle pieces together:
  - Identify statistical correlations in data

  - Design good features

  - build models that solve specific business problems
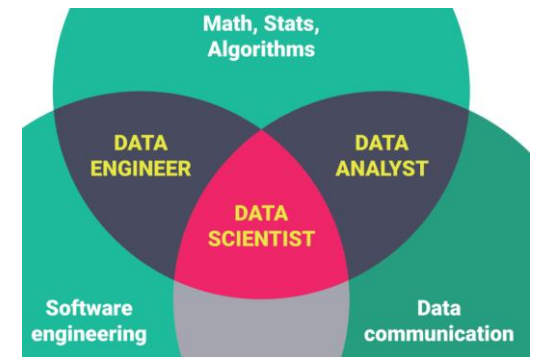
Collect Relevant Data

- Time series

  - Water/electricity daily consumptions from smart meters

  - Monthly billing data

- Other ancillary data

  - Meter brand/age/type/size…

  - Premise/Spid/account mapping

- Data saved in Azure Blob and Azure SQL, real-time data flow into system every hour

EnergyAuthority

Understand the Data

- Are there noises, or holes in data?

- What's the 'normal' daily consumption pattern for a typical unbroken meter?

- What exactly had happened around the truck rolls?

Machine Learning/Data Science Approach

- Data Collection
  - Time series (water/electricity consumption)
  - Other ancillary data
- Identify the features
- For every truck roll record, find what exactly happened in the history
- Train a ML model
- Cross-Validation
- Field Testing

A Data Scientist's Toolbox

Field testing - ready to go

Field testing - one worker & six supervisors