Introduction to Machine Learning

Mahammad Valiyev

02.01.2022

Contents and timeline

1. Introduction to Machine Learning and use cases in O&G (Jan 2)
2. Overview of Machine Learning algorithms (Jan 8)
3. Machine Learning Life Cycle (Jan 15)
4. Overview of resources, skill sets, job types, general advice (Jan 22)

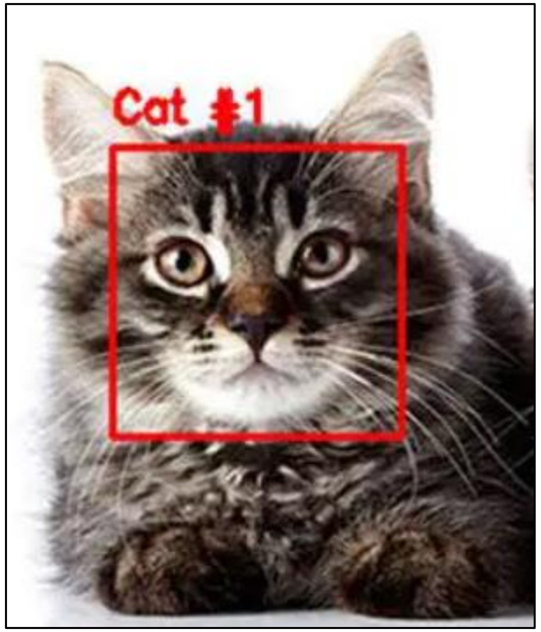
Part 1:

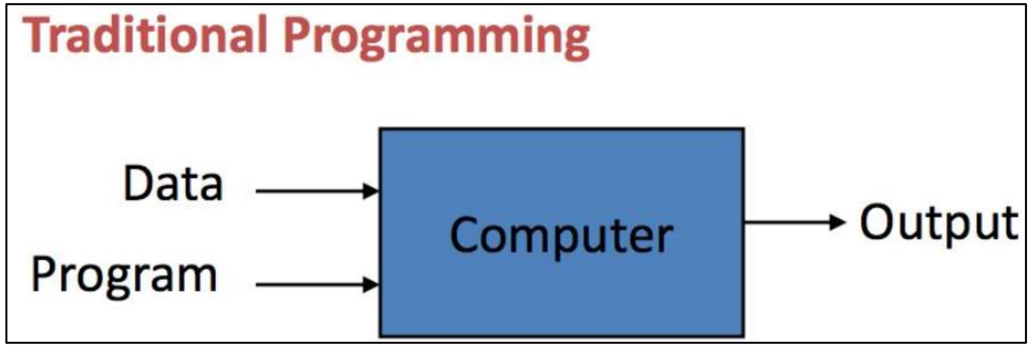
Introduction to Machine Learning and uses cases in O&G

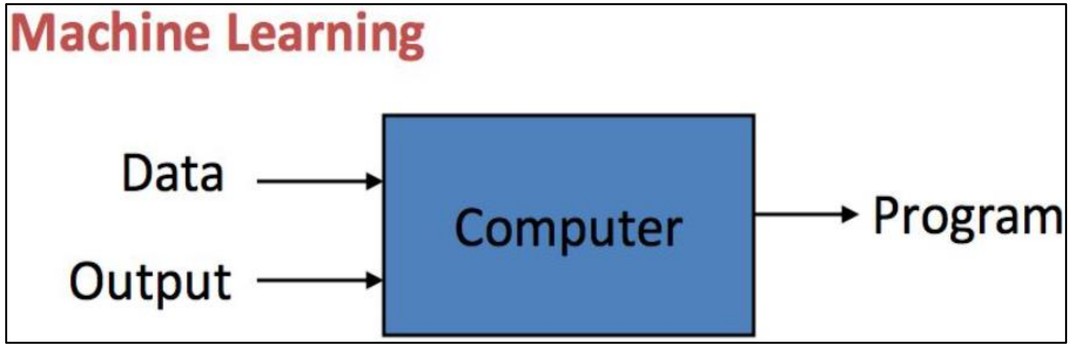
Machine Learning is new paradigm

- **Traditional programming**: develop algorithms as a logical series of steps and translate them into code

- **Machine Learning**: develop data-driven algorithms that learn on examples without explicitly programming all the rules/processes

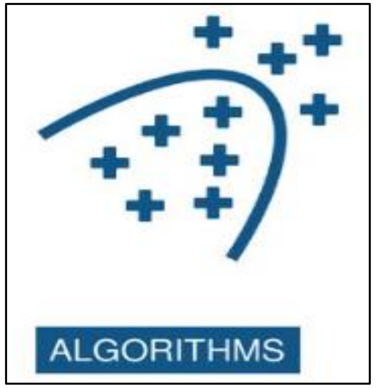






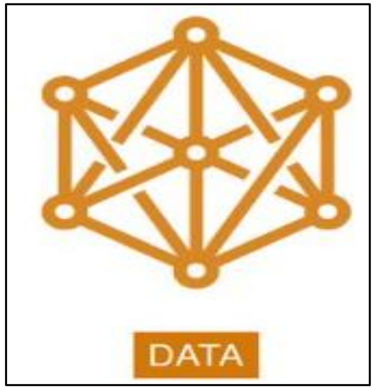
Why now?

Key enablers: algorithms, data, computing power



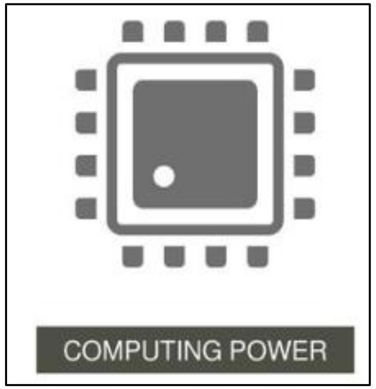
Algorithms:

- Initial developments, 1950-1960s: perceptron, nearest neighbors
- New hopes, 1980s: RNN, CNN, backpropagation, reinforcement learning
- Rise of Deep Learning, 2006: Deep Belief Network



Data:

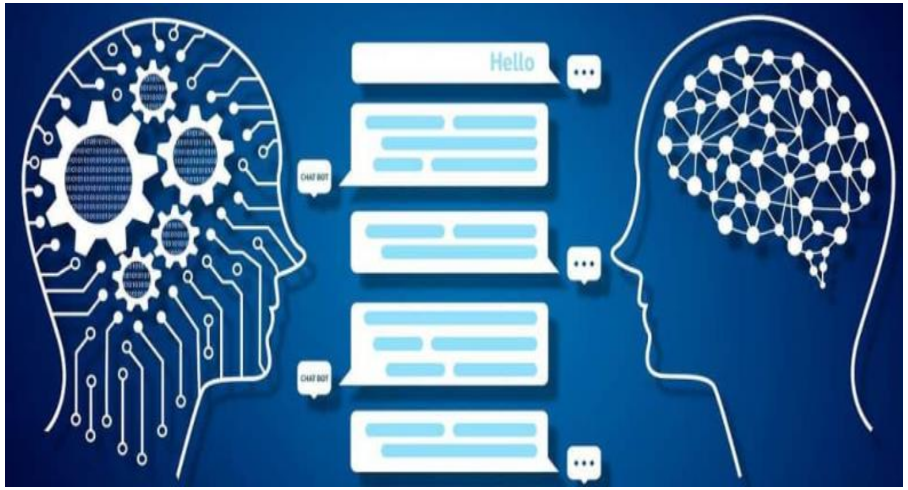
- World Wide Web: passive (1991), Web 2.0: interactive (2004)
- Facebook (2004), YouTube (2005), iPhone (2007)



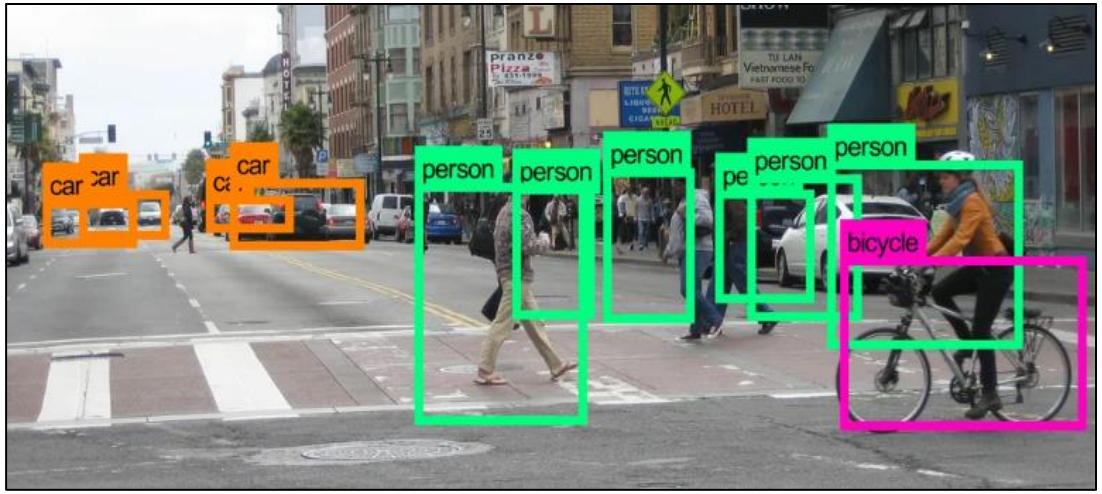
Computing Power:

- Moore's law (1965), Deep Blue (1997), 1st GPU (1999)
- AWS cloud (2006), CUDA for GPU (2007)

Top use cases



Natural language processing: language generation, translation

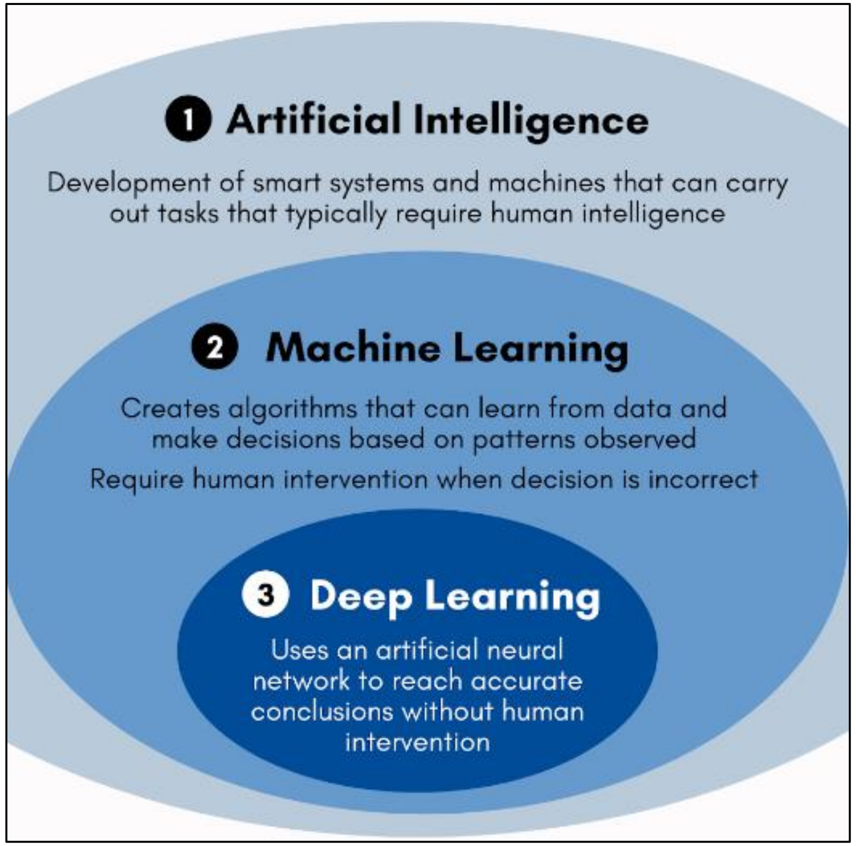


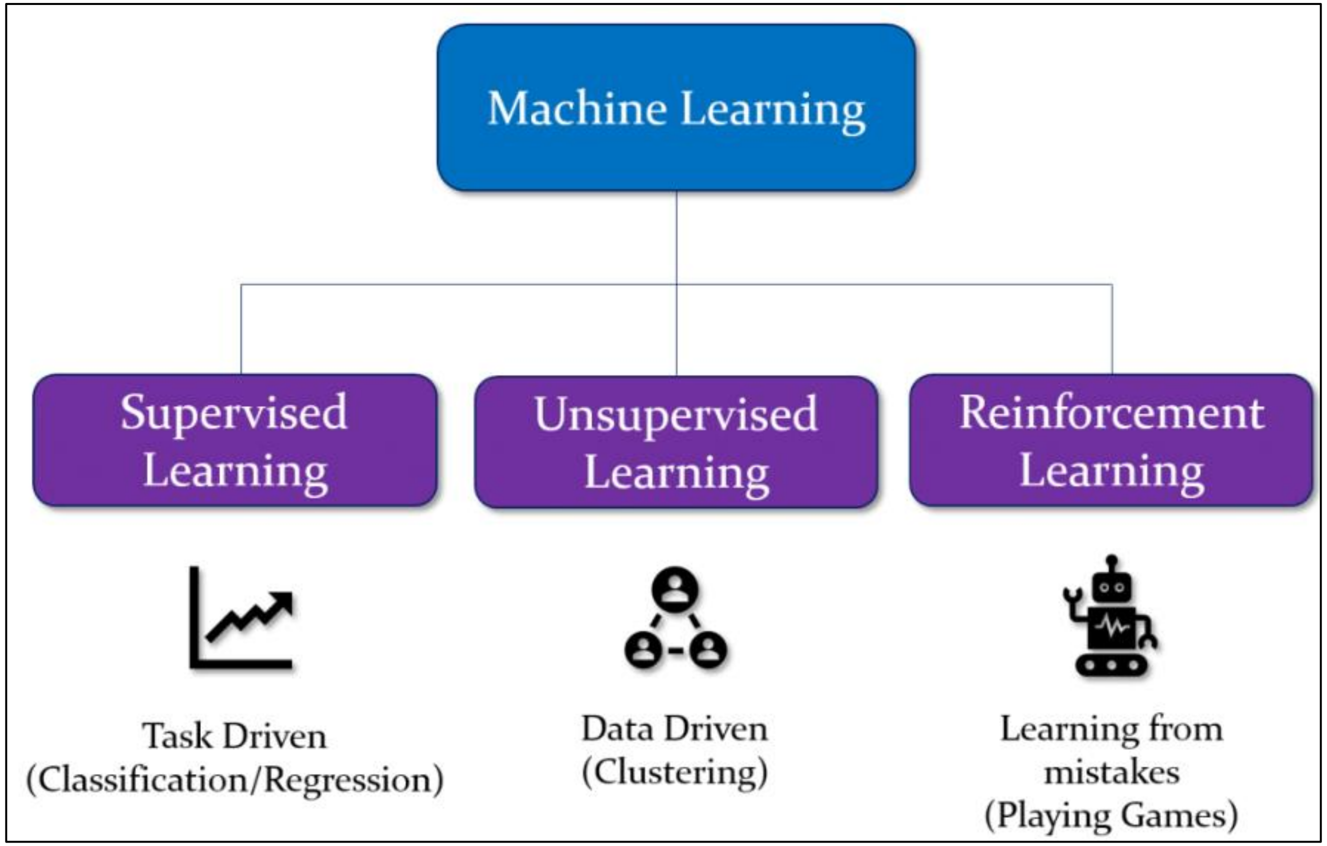
Computer vision: image recognition, object detection

Speech recognition : speech to text

Recommender systems: retail, movies, ads

Major types

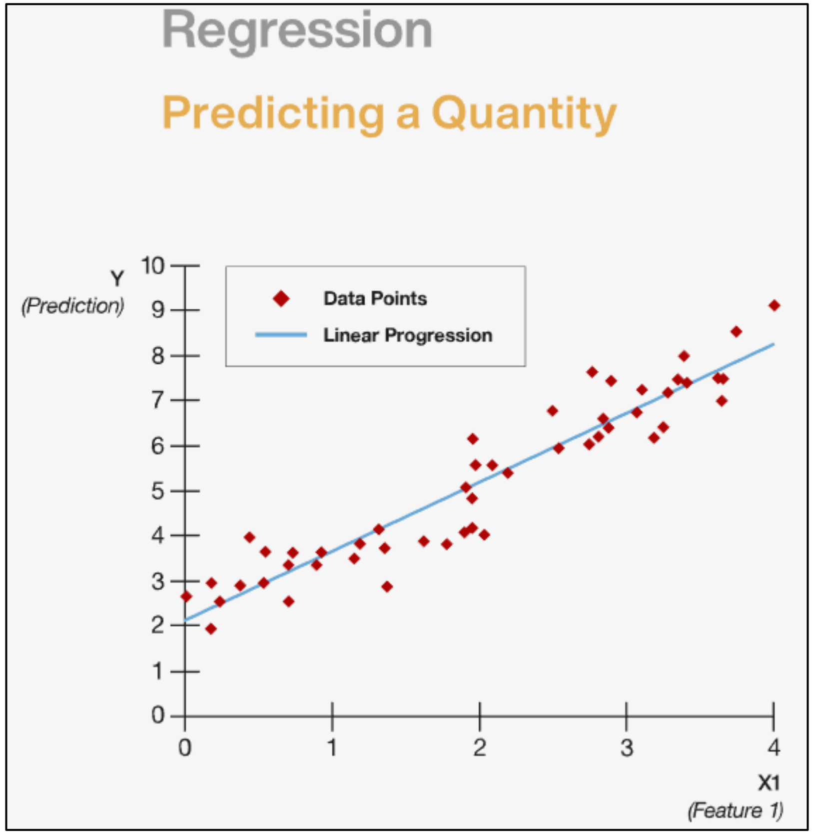

1 Artificial Intelligence
Development of smart systems and machines that can carry out tasks that typically require human intelligence
2 Machine Learning
Creates algorithms that can learn from data and make decisions based on patterns observed
Require human intervention when decision is incorrect
3 Deep Learning
Uses an artificial neural network to reach accurate conclusions without human intervention

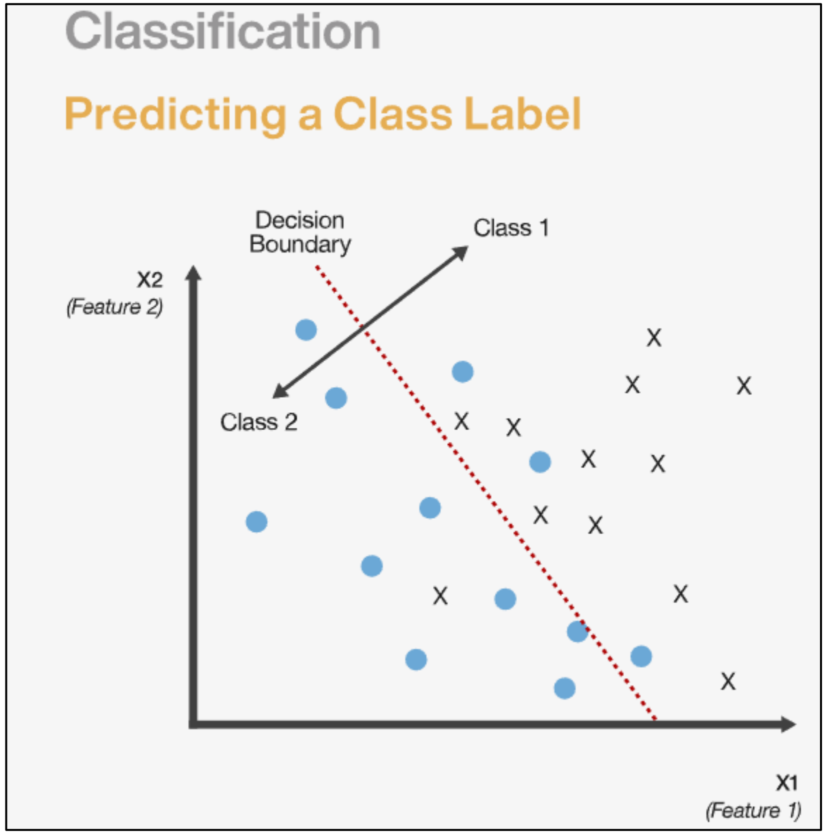

Machine Learning
Supervised Learning
Unsupervised Learning
Reinforcement Learning
Task Driven (Classification/Regression)
Data Driven (Clustering)
Learning from mistakes (Playing Games)

Supervised learning

Supervised learning: learning a functional mapping from input to outputs

- **Regression**: output is continuous variable
- **Classification**: output is categorical variable (2 categories: binary, or more: multiclass)

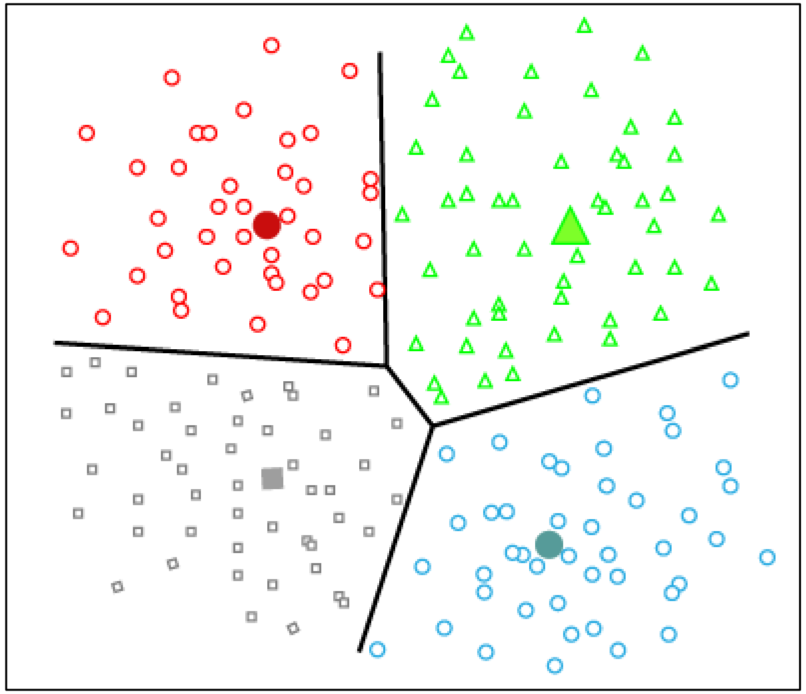




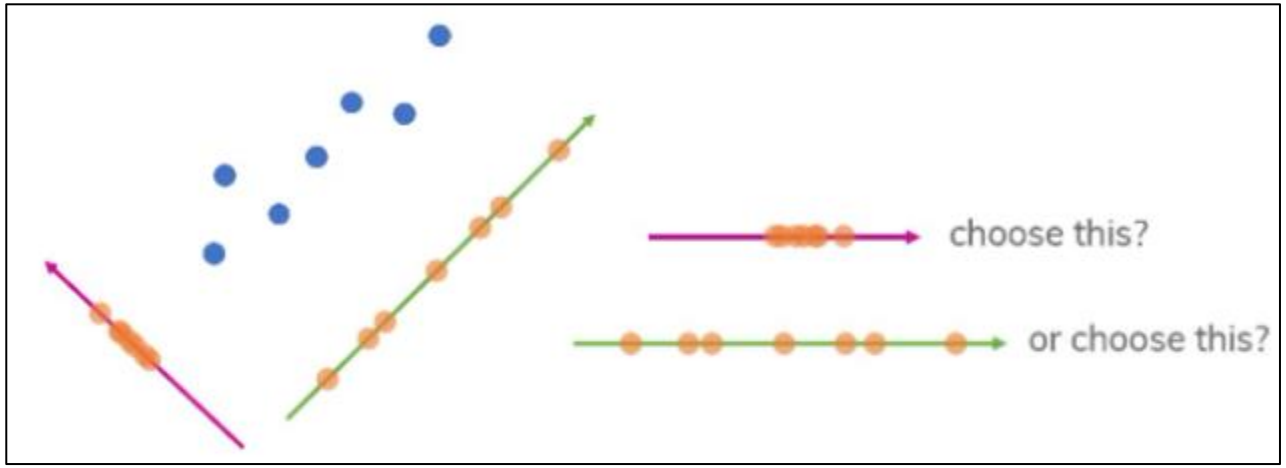
Unsupervised learning

Unsupervised learning: learning patterns from unlabeled data

- **Clustering**: dividing data into a number of groups
- **Dimensionality reduction**: reducing number of input variables



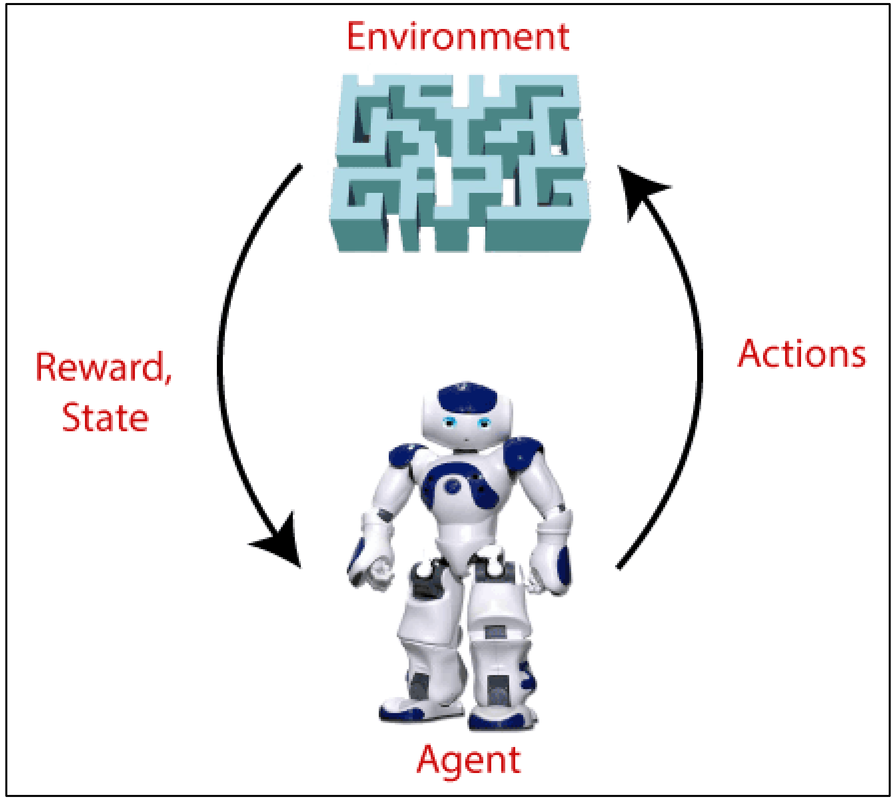
Clustering



Dimensionality reduction

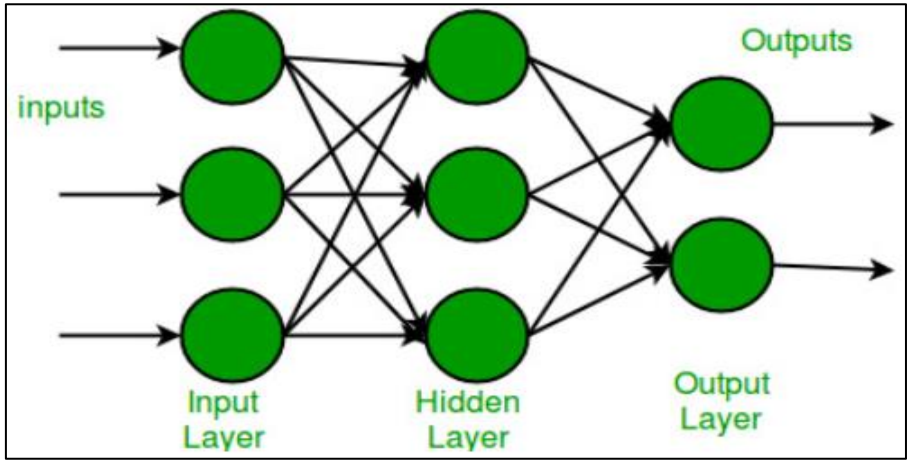
Reinforcement learning

- **Reinforcement learning:** learning by trial and error through interaction with environment
  - Environment, possible actions, rewards and penalties need to be defined
  - Once those defined, model tries to select actions that maximize the reward

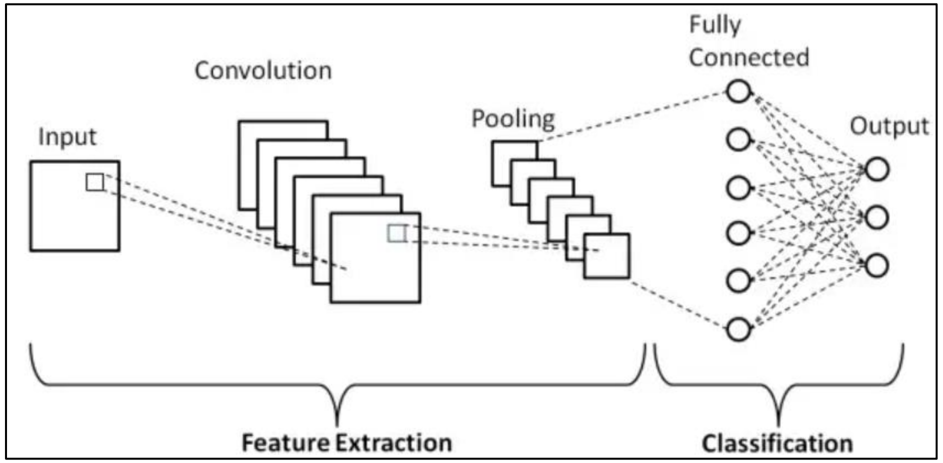


Deep learning

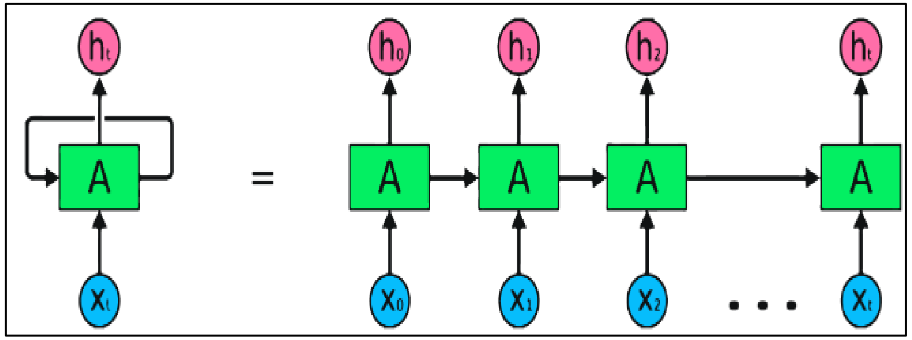
- **Deep learning**: subset of machine learning that involve use of neural networks
  - Applicable to all 3 ML subcategories (supervised, unsupervised, reinforcement learning)
  - More powerful models, applied to solve more complex problems
  - Requires more data and computing resources, but less data preprocessing



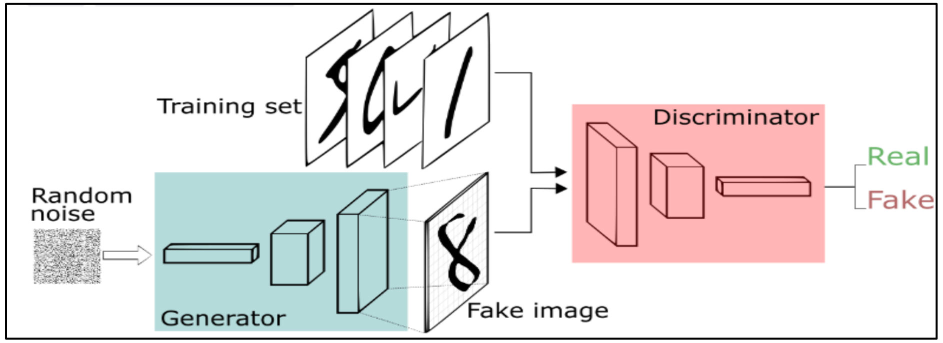
Multilayer perceptron: tabular data



CNN: image data



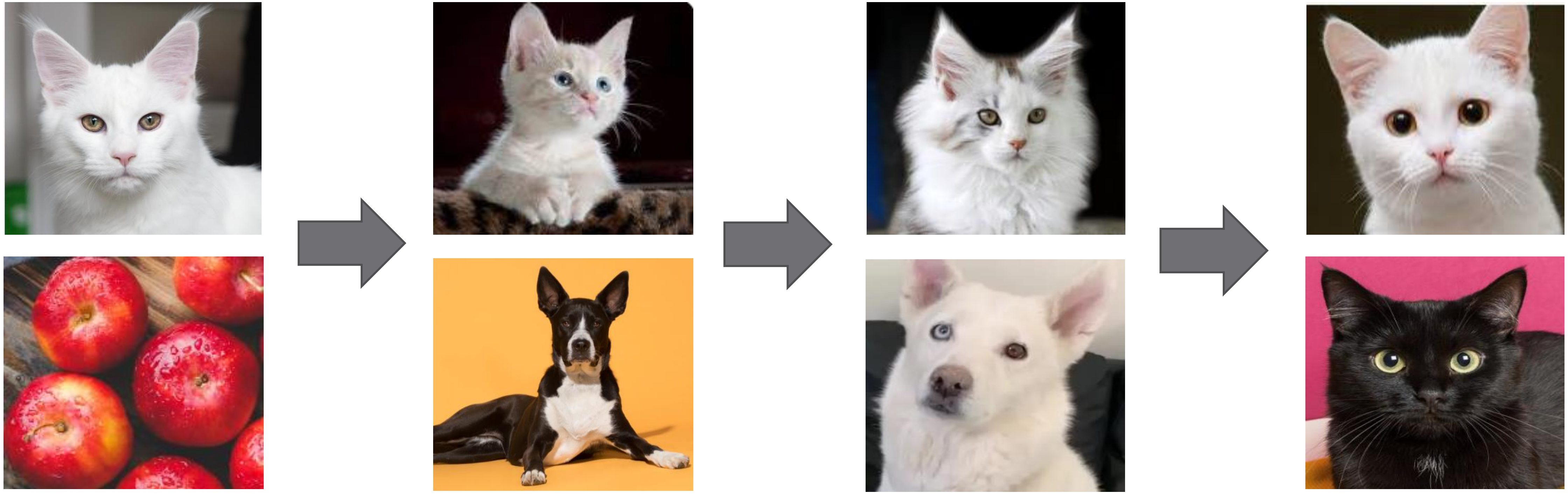
RNN: sequential (timeseries, text) data



GAN: (mostly) image and audio data

How does Machine Learning really learn?

- Building a cat recognition model..

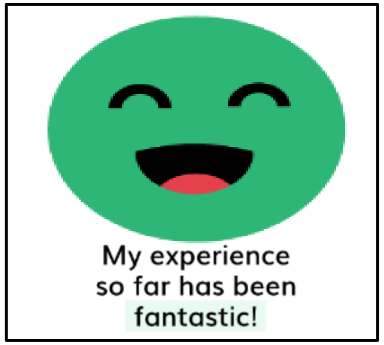


- **Machine learning learns on training examples in an iterative fashion**
- **Training data should cover wide range of possibilities**
- **Every data is represented in numbers: images, audio, videos, text**

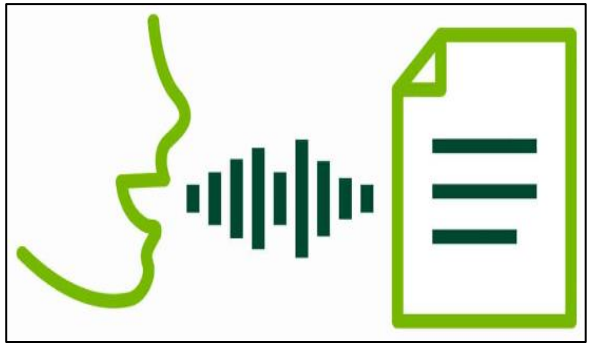
Power of Machine Learning

- Machine learning models can learn very complex nonlinear functions
- A lot of scientific problems & business decisions can be framed as functions

Image classification
Image to number (0,1)



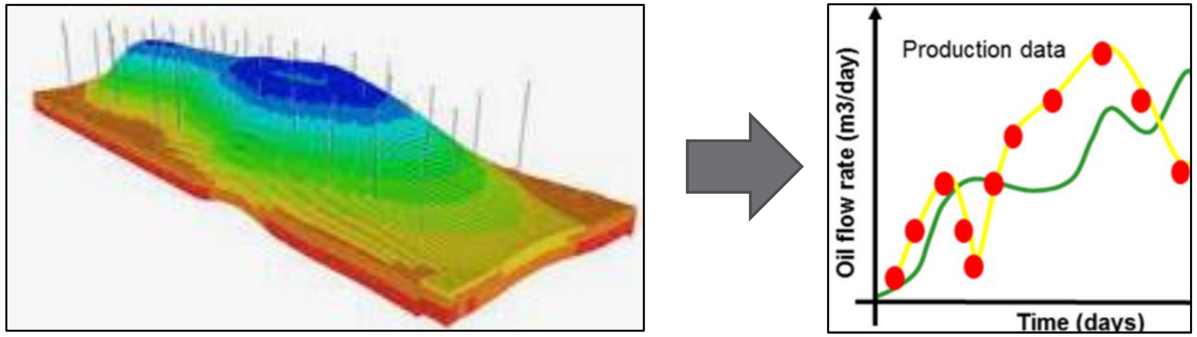
Sentiment analysis
Text to number (0, 1, 2)



Speech recognition
Audio to text

Language translation
Text to text

Fraud detection
Transaction data to number (0, 1)



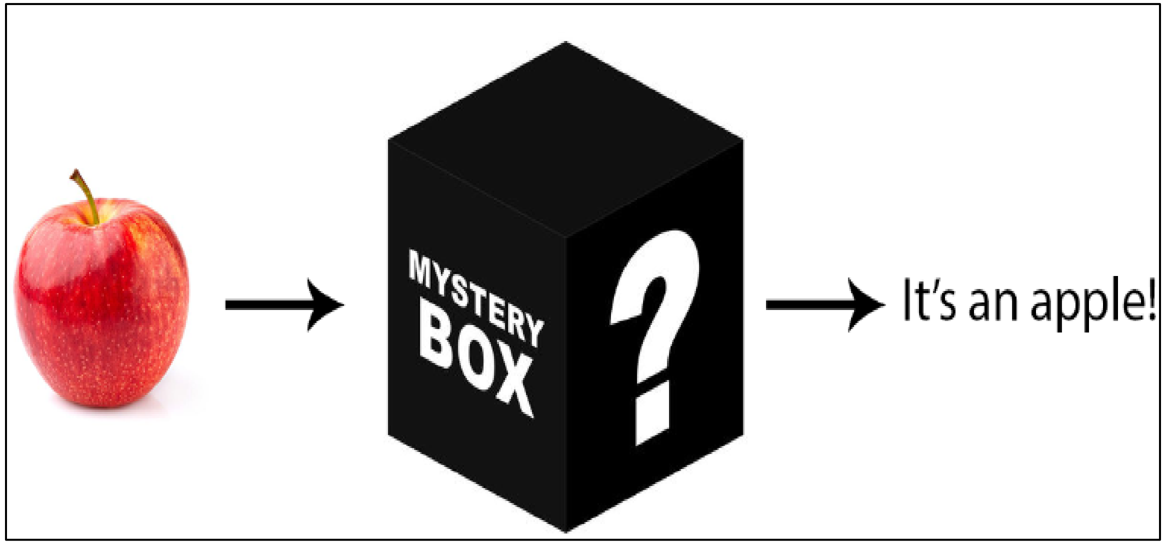
Oil rate prediction
Reservoir data to number / timeseries

Needs and limitations of Machine Learning

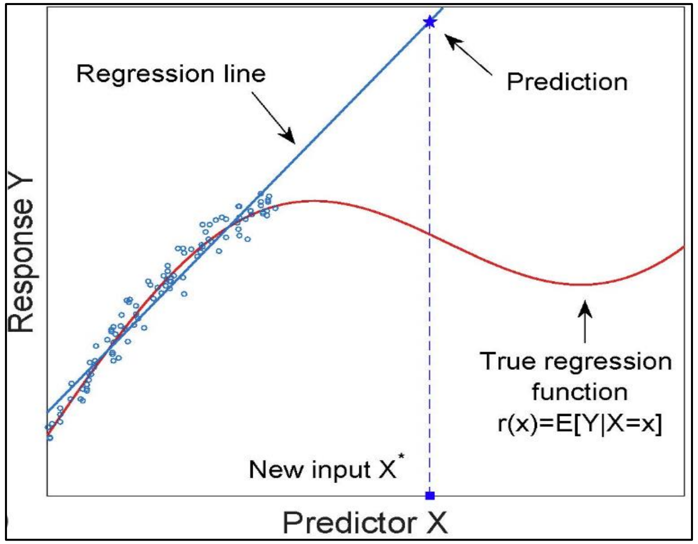
14,3	0,4	0,00
11,8	0,1	0,13
10,3	0,3	0,00
11,8	1,1	-0,06
13,2	1,9	-0,03
16,9	0,9	0,00

Data

Computing resources



Black box/non-interpretable models

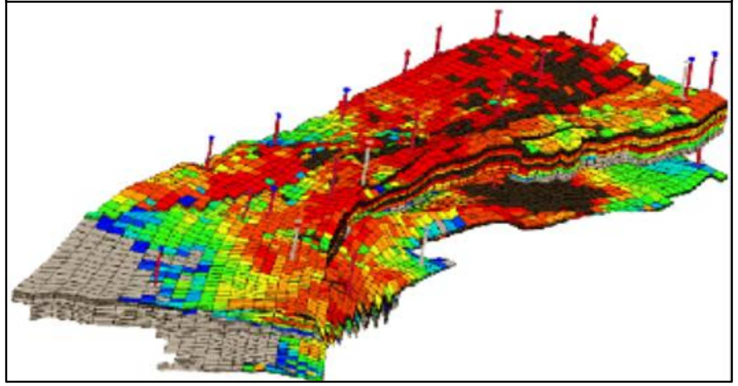


Extrapolation problem

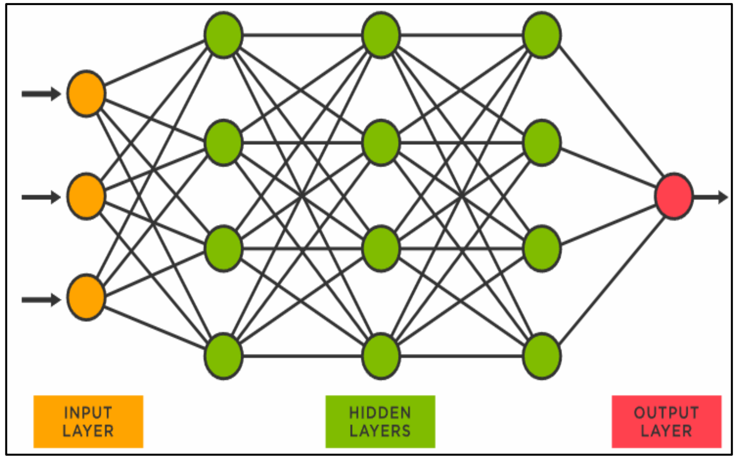
Use case 1: Proxy model development, Reservoir Engineering

$$\nabla \cdot \left(\frac{\lambda_w}{B_w} \mathbf{u} \left(\nabla P_w - \gamma_w \nabla Z \right) \right) = \frac{\partial}{\partial t} \left(\phi \frac{S_w}{B_w} \right) + q_w$$

$$\nabla \cdot \left(\frac{\lambda_n}{B_n} \mathbf{u} \left(\nabla P_n - \gamma_n \nabla Z \right) \right) = \frac{\partial}{\partial t} \left(\phi \frac{S_n}{B_n} \right) + q_n.$$



Reservoir simulation



Deep learning model

Inputs:

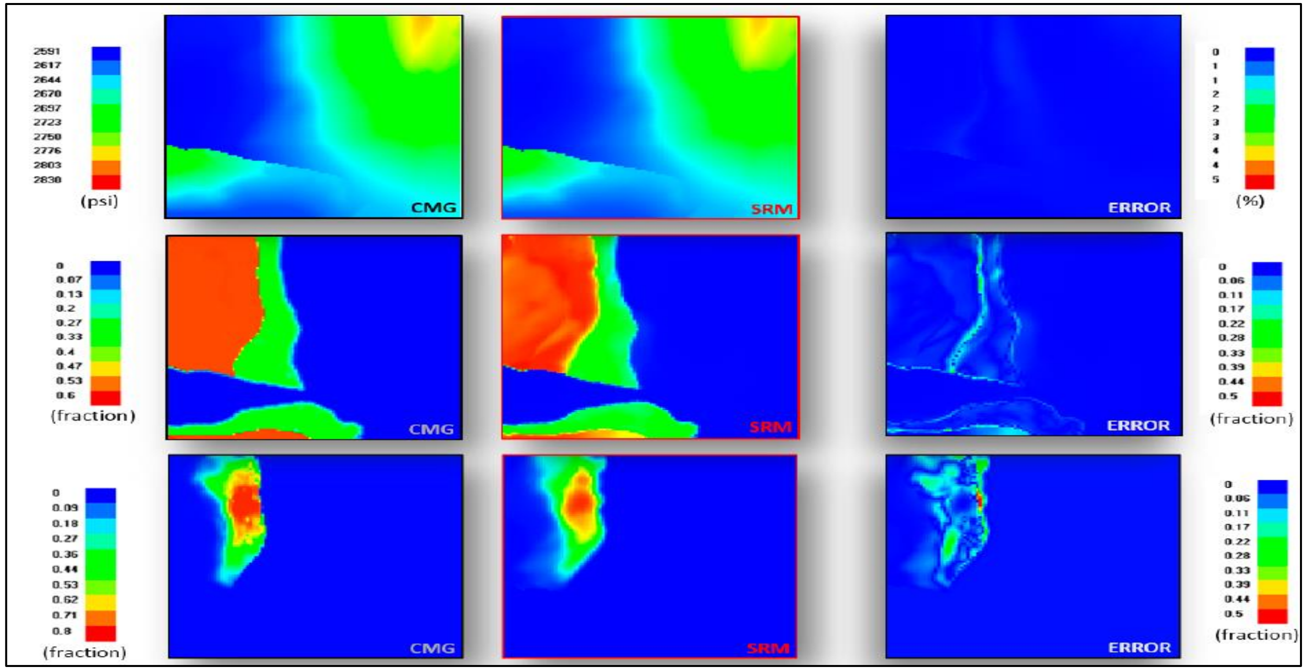
- reservoir info , dynamic data, well data

Outputs:

- pressure, gas saturation, CO2 mole %

Project value:

- reduced computation time (15 min vs 45 sec)



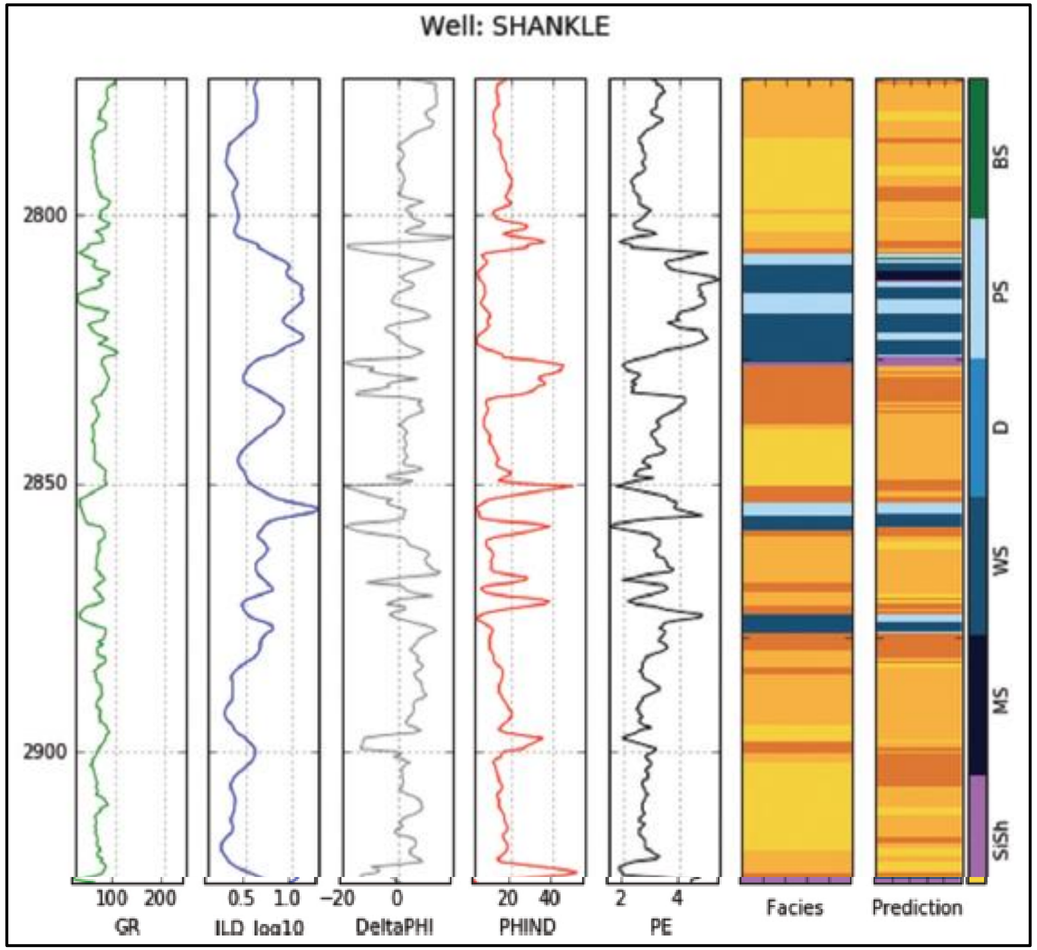
Maps of Pressure (top), gas saturation (middle), CO2 mole fraction (bottom) in the 1st layer of reservoir

(left: CMG, middle: proxy model, right: error)

Shohreh Amini, Shahab Mohaghegh, 2019

Use case 2: Facies classification, Geoscience

- **Input**:
  - well log data (GR, resistivity, photoelectric effect, neutron-density porosity difference, neutron-density porosity)
- **Output**:
  - facies class (1-9)
- **Project value**:
  - ability of predicting facies class of a sample from well log data that does not have a core description



Well logs and facies classification results from a single well
Brendon Hall, 2016

Use case 3: Prediction of ROP, Drilling Engineering

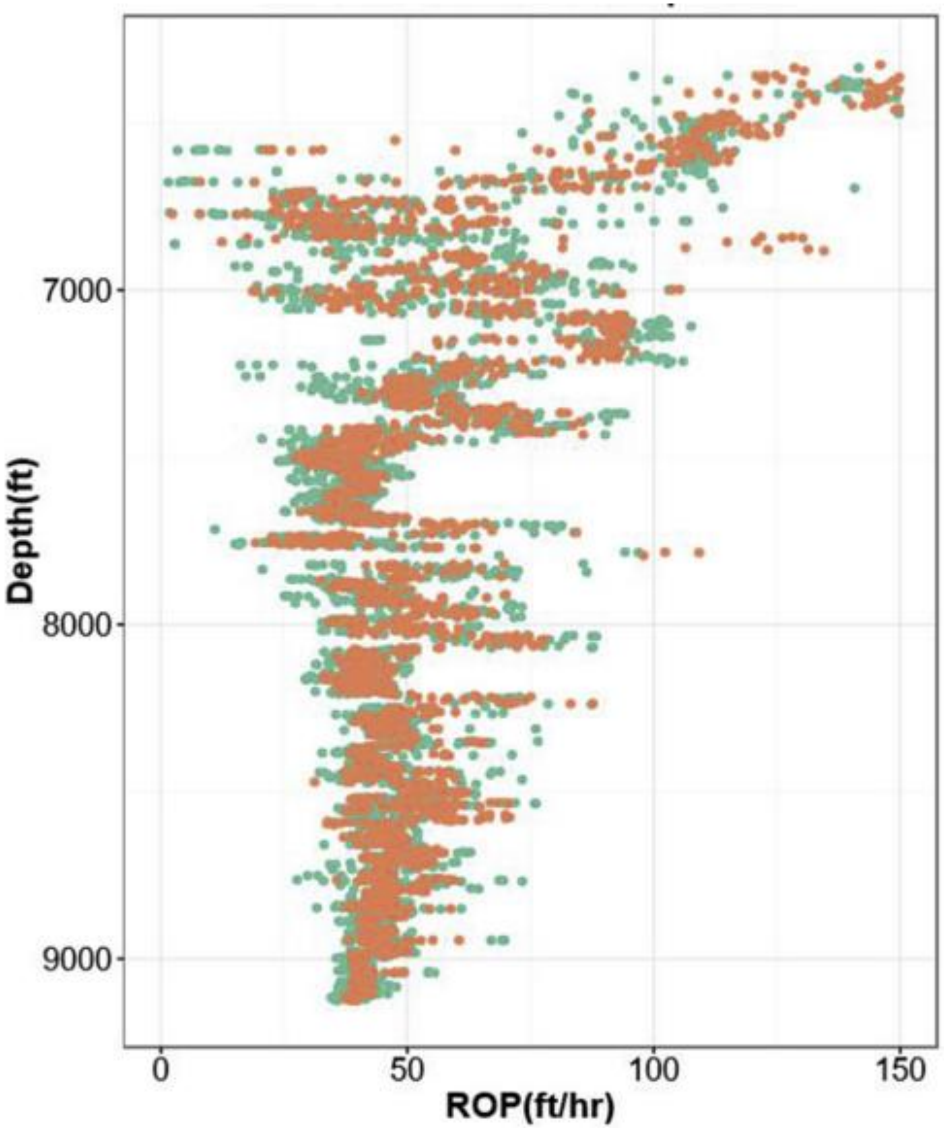
- **Inputs**:
  - WOB, RPM, depth, flow rate
  mud weight, bit diameter, rock strength etc
- **Output**:
  - Rate of penetration (ROP)
- **Project value**:
  - ROP prediction ahead
  of time, real-time ROP optimization

Other methods for ROP prediction:

- Experience
- Traditional / physics-based models

$$\hat{y} = f(x, \Theta)$$

$$ROP = W_f \frac{G.\, N^{\gamma}.\, W^{\alpha}}{d_b.\, S}$$



ROP prediction using ML model for the entire length of the well (green: actual ROP, orange: ML prediction

Chiranth Hegde, K.E. Gray, 2016

References and further resources

Websites:

- **An executive's guide to AI:**
  - https://www.mckinsey.com/capabilities/quantumblack/our-insights/an-executives-guide-to-ai
- **IBM: What is Machine Learning:**
  - https://www.ibm.com/topics/machine-learning
- **Machine Learning for Managers:**
  - https://c3.ai/introduction-what-is-machine-learning
- **Traditional programming vs Machine Learning:**
  - https://www.avenga.com/magazine/machine-learning-programming/
- **The Limitations of Machine Learning:**
  - https://towardsdatascience.com/the-limitations-of-machine-learning-a00e0c3040c6

Papers:

- **Proxy modeling:**
  - https://www.mdpi.com/2311-5521/4/3/126
- **Facies classification:**
  - https://library.seg.org/doi/full/10.1190/tle35100906.1
- **Prediction of ROP:**
  - https://www.sciencedirect.com/science/article/pii/S0920410519307533?casa_token=kcVdJE6hIJkAAAAA:ikWiqq-tdfzbVbsMiBYBbS_iUBRxD-QfghEg4hWmvpfiC_9NTzTdBSalcXrRfiJPmfeynFrT5ug
  - https://www.sciencedirect.com/science/article/pii/S0920410517307258?casa_token=HouxJBVkXLQAAAAA:Ihn7NjP9B7nVMGRUN2ex8-m-PfnSf2qFpEeujDe27T8y7mNVU6UiE8n18mMBtkK-9memZvZurK8

Recap

- ML vs traditional programming
- Key enablers of ML
- Major types of ML
- Intuition behind ML
- Power and limitations
- Use cases

Thank you