Machine Learning:

Going Beyond the Hype and Making it Work for Earth Science

Nina Marie Hernandez

March 2018
Manila, Philippines

www.irayaenergies.com

iraya
machine learning • geoscience

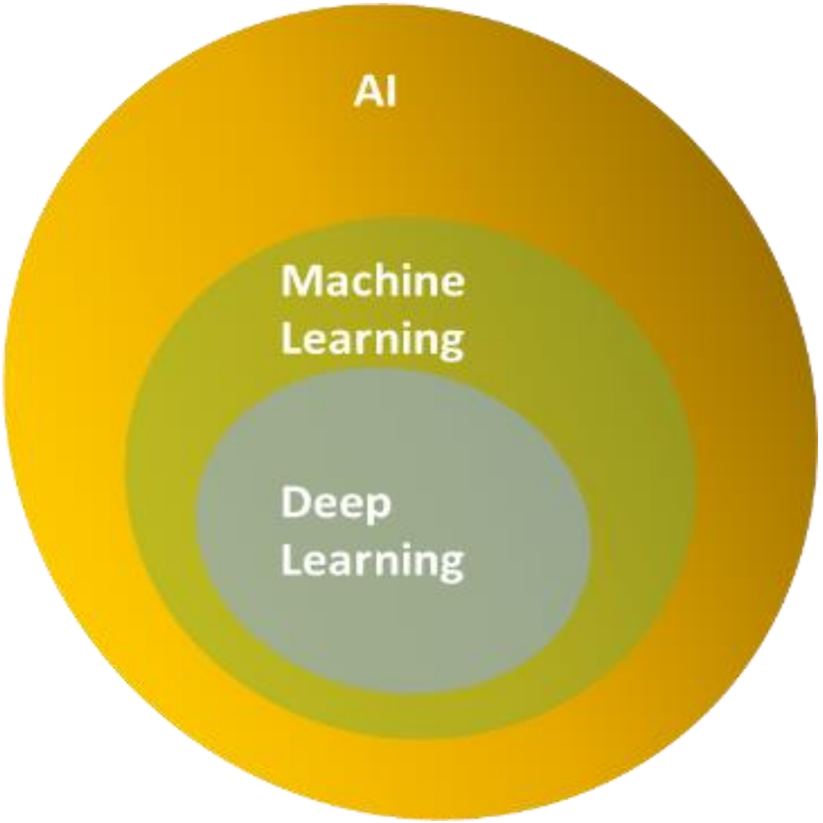
Outlines

1 AI – Is it Magic, or Math?

2 Making AI Work for Earth Science

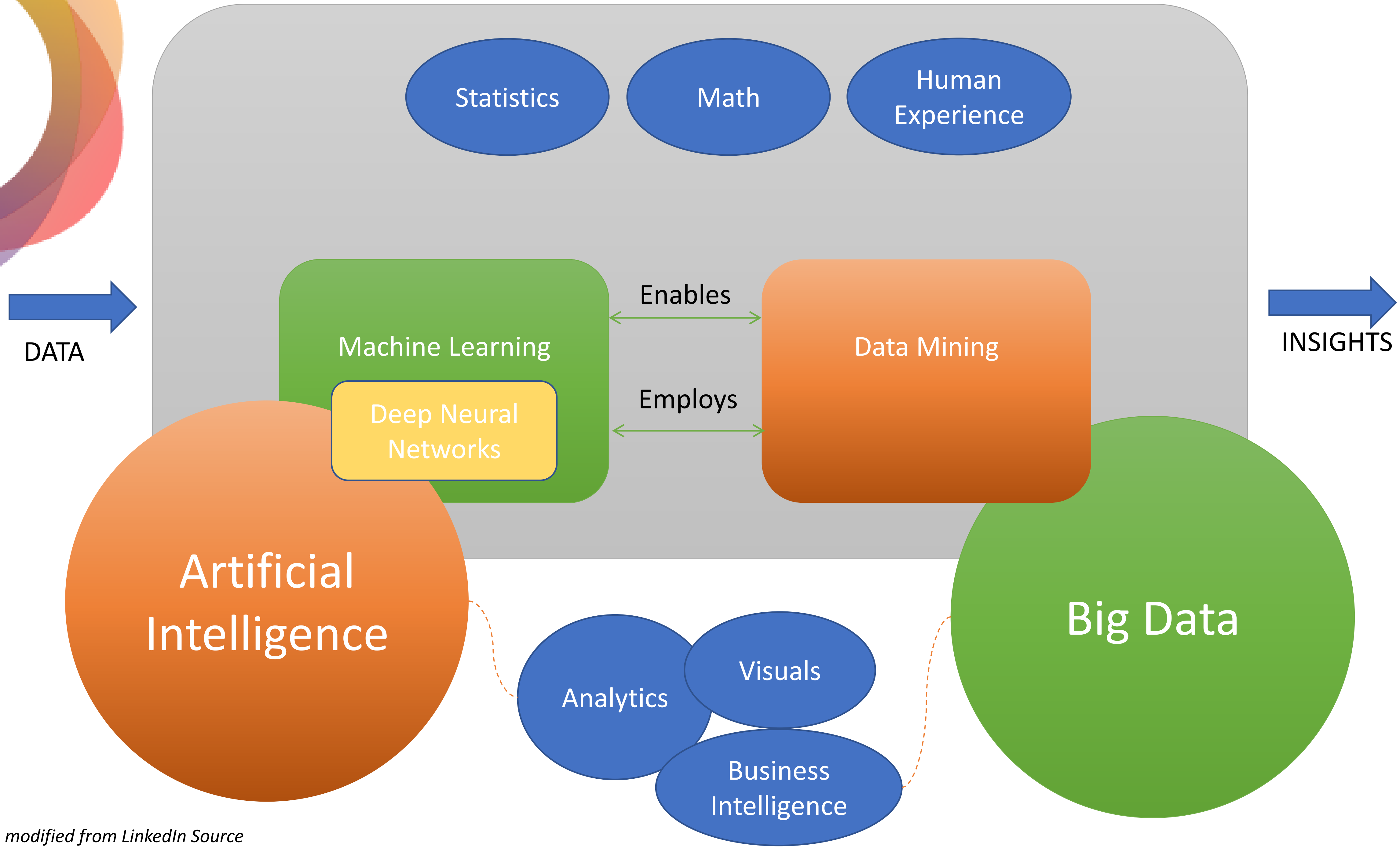
3 How to Build an AI Project

AI & ML - What is it?



Arthur Samuel (1959) on **Machine Learning**:

The field of study that gives computers the ability to learn without being explicitly programmed.



NMH modified from LinkedIn Source

History of Artificial Intelligence

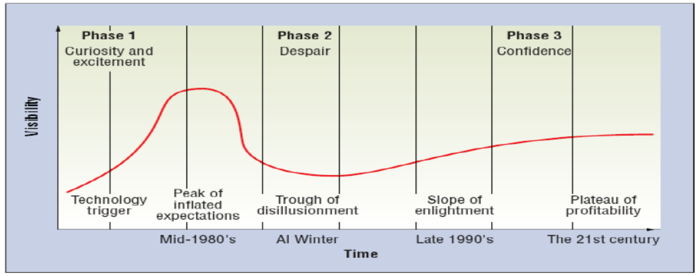


Figure 1. The Hype Cycle and AI winter [Menzies03]

1950

1950: TURING CAN MACHINE THINK?

1954: RUSSIAN TRANSLATION

1956 FIRST USE OF WORD AI

1970

1961 FIRST COMPUTER FOR SEISMIC DATA PROCESSING

1973- COMPUTER CHESS

1980- EXPERT SYSTEMS

AI WINTER 1980s-1990

2000

1997 – IBM DEEP BLUE BEATS KASPAROV

2000 – ANN PREDICTION

2018

MACHINE LEARNING IN GEOSCIENCE

AUTONOMOUS CARS

ALPHA GO

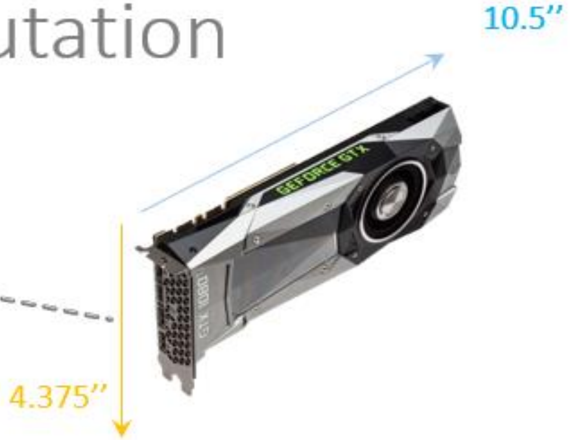
FACE RECOGNITION

EAGE PARIS 2017 ML

AI explosion in 2018

GPU computation

GPU computing



2018

Market Price 600$

Computing 11.3 TFLOPs

Power 230W

2001

Market Price 110M$

Power 3000 kW

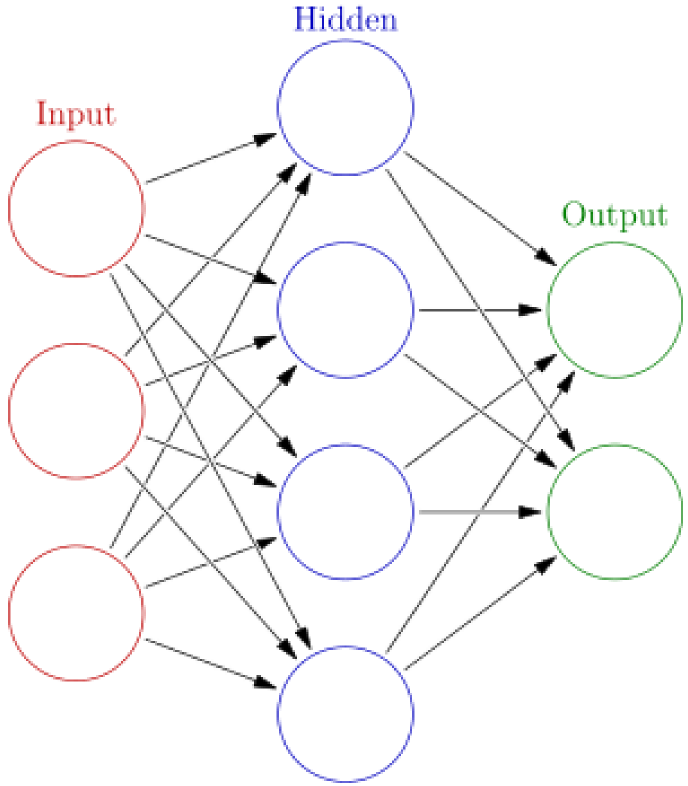
Source: top500.org & nvidia.com

Nature as an inspiration

Artificial neural networks (ANN) mimic neurons in a brain

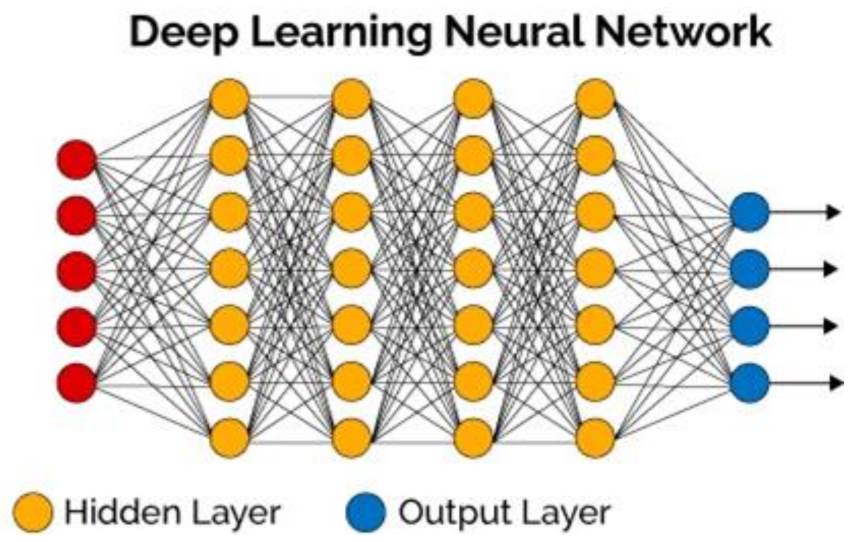
- Layers of nodes with weighted connections between layers

- Information through network changes its structure – **it learns**



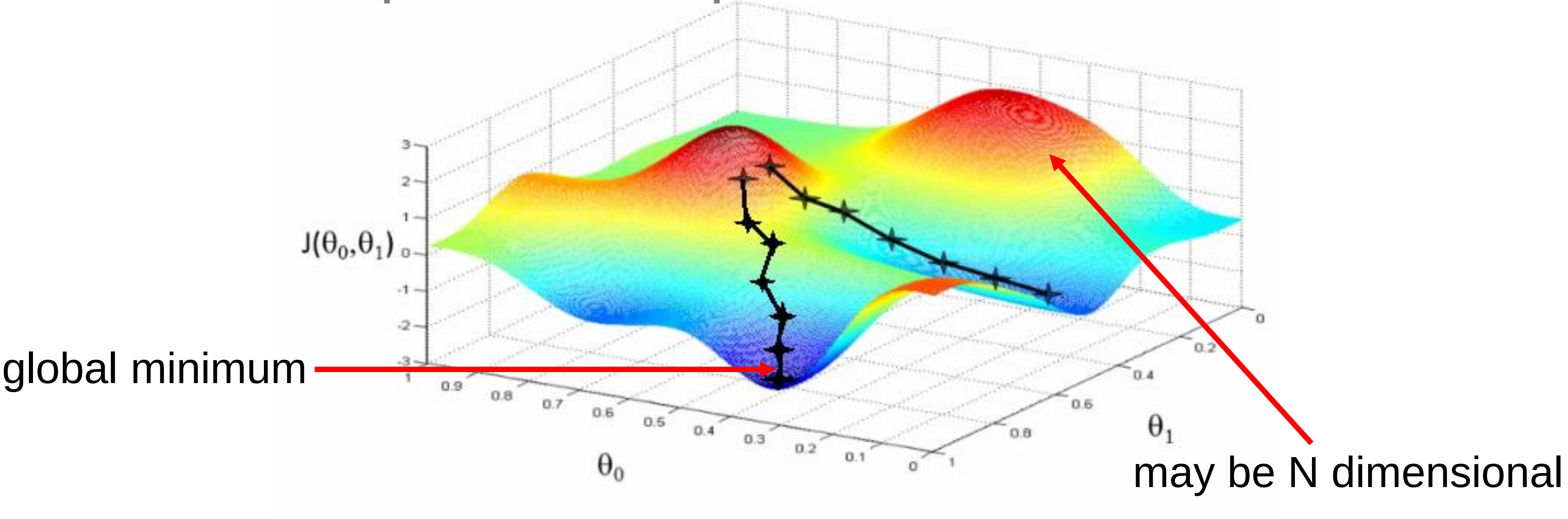
Architectural Principles of DNN

- **Features** : measurable property or characteristic of a phenomenon being observed
- **Layers**: information are passed from one layer to the other
- **Activation Function:** mathematical functions that act as filter
- **Loss Function:** define the difference between the predicted and actual features
- **Optimization Algorithms:** algorithms that minimize the loss functions or errors in the modeling
- **Hyperparameters:** adjustable parameters, such as number of layers, learning rate that can influence the effectivity a of neural network



Further Readings: Coursera Machine Learning, by Prof. Andrew Ng

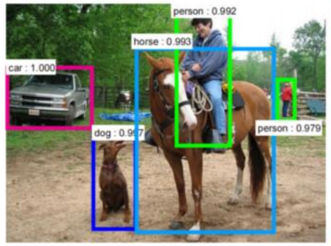
Mathematically, training a neural network is an optimization problem



- Show different input values and compute error
- Adjust weights in direction where error is minimized (along gradient)
- Eventually reach minimum value

Types of Neural Network

Convolutional Neural Nets



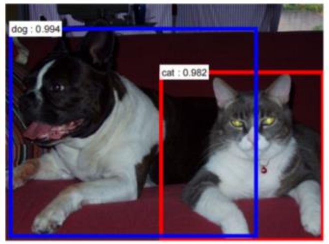
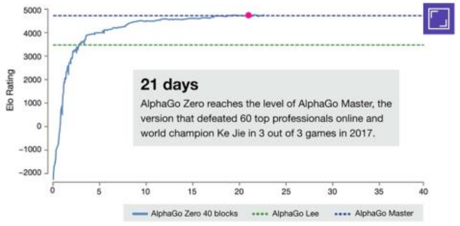


IMAGE CLASSIFICATION
and
FEATURE DETECTION

Recurrent / Recursive Neural Nets

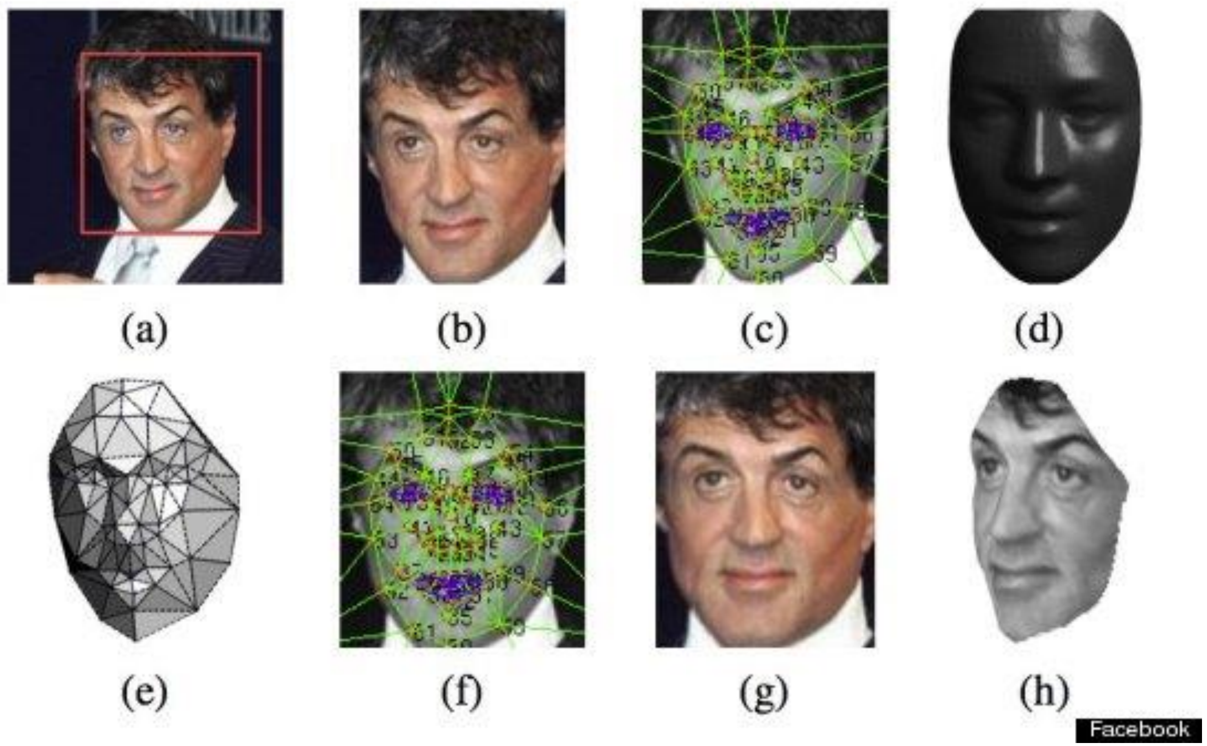
TEMPORAL BEHAVIOUR MODELING
AUDIO, TEXT
NATURAL LANGUAGE PROCESSING

Reinforced Neural Nets



PAVLOVIAN CONDITIONING

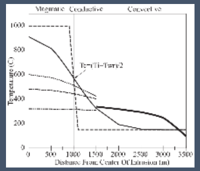
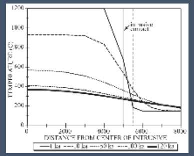
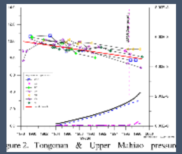
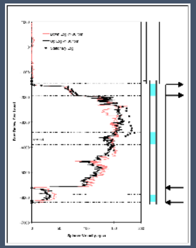
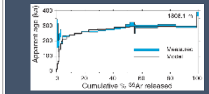
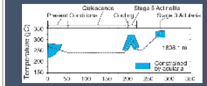
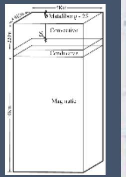
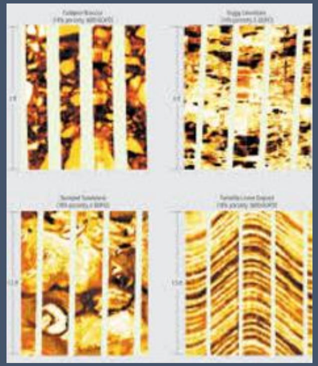
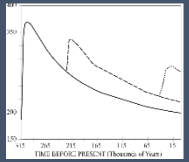
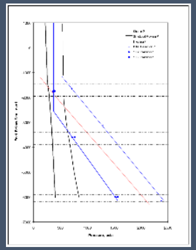
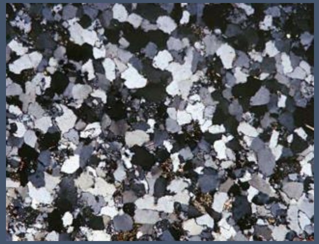
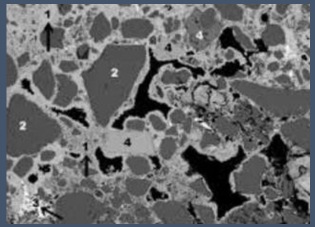
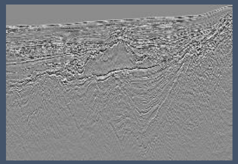
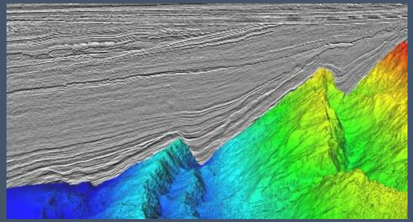
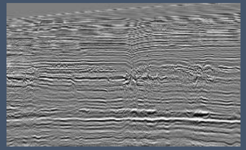
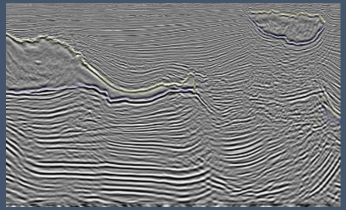
Deep learning has found many applications in image processing



- Facebook's DeepFace for facial verification
- DNN with 9 layers
- Trained using millions of images uploaded by users
- Accuracy reaching 97.35%

What do geoscientists do on a daily basis?
We make (image) files

EARTH PROBLEMS



Making AI work for Earth Science

Use AI to (classify, predict, learn from)
archived, historical megadata

New Data Acquisition
is very costly

Optimize efficiency during exploration and early development phase

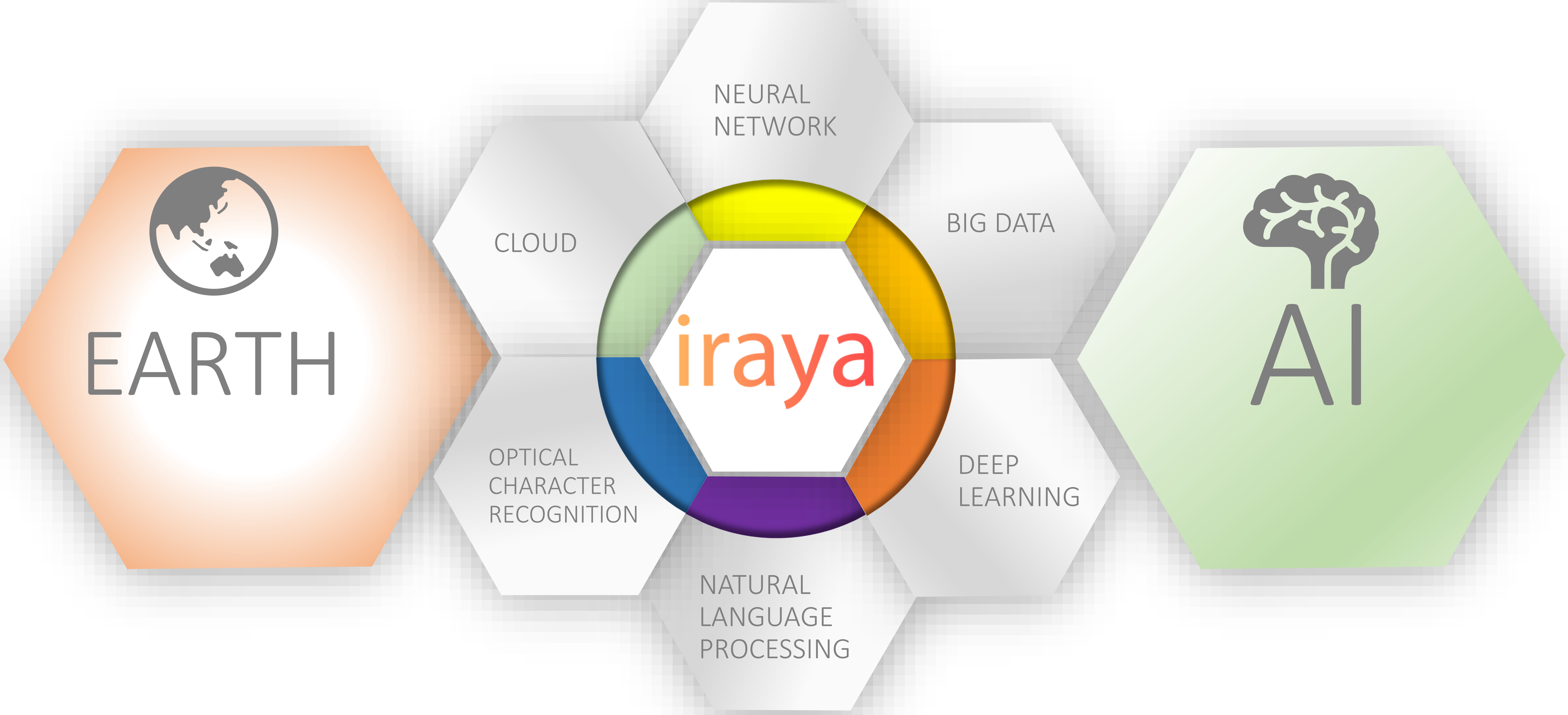
Learn Effectively

10 vs 1,000 wells

20 vs 2,000 seismic lines

Yet-to-Find becomes Easy-to-Find

Making AI work for Earth Science



Iraya Use Cases of AI

- Use Case # 1 : Data Mining
- Use Case # 2: Well Twinning
- Use Case # 3 : Clustering
- Use Case # 4: Deep Resolution

Use Case #1: Data Mining

- Problem Definition:

Extract information from a unstructured dataset

- Standard Solution:

Download data, manually read metadata and load in a spreadsheet

- Machine Learning Solution:

Apply mining robots, elastic search, natural language processing, optical character recognition to reduce timeframe by a factor of 100.

Data Mining Analogy

Harvesting

Transform

Sort

Enhance

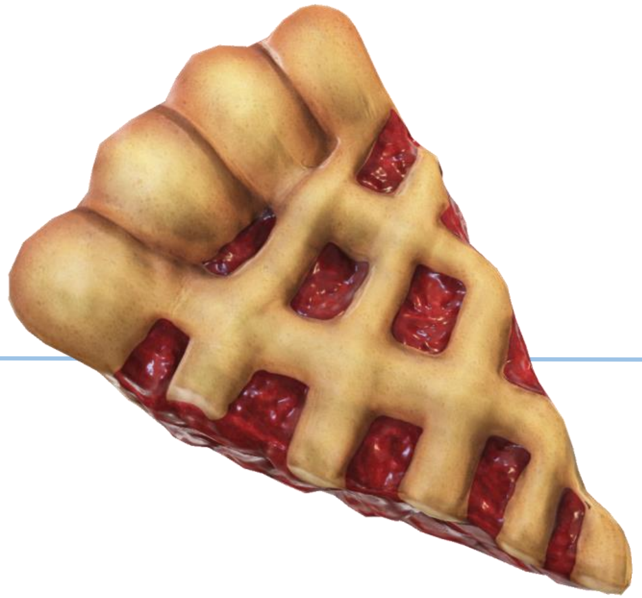
Data Harvesting Analogy

Fully automatic - AI driven

Harvesting

Transform

Sort



Enhance

MINE-BOT	CLASSIFY	VISUALIZE	BIG DATA ANALYTICS

MINE-BOT

Digital or Physical Storage
Web, USB, HDD, Vintage

↓

All types of file type
*las, segy, pdf,xls..

Catalogue
by file type or application

CLASSIFY

Wells

- ❑ Metadata
- ❑ Curve class
- ❑ Frequency count

Reports

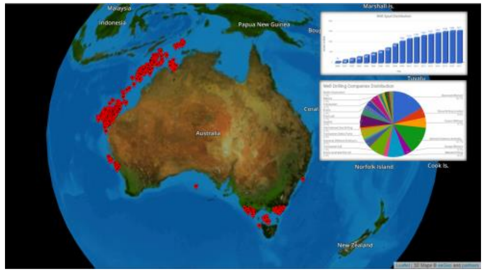
Seismic

- ❑ Metadata extraction
- ❑ OCR Digitization
- ❑ Text vs Image

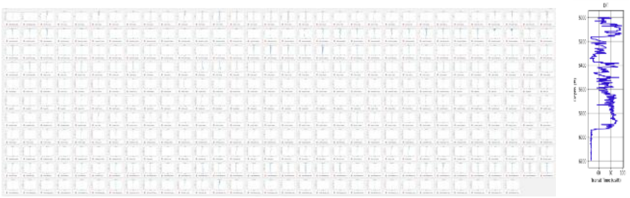
VISUALIZE

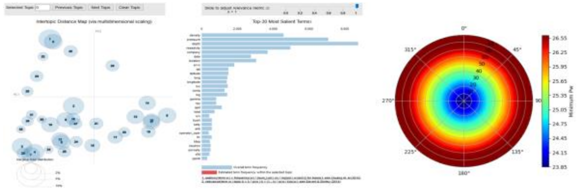
Cloud

WebGIS



Big Data Visuals





BIG DATA ANALYTICS

Dimension Reduction

T-SNE

PCA

Clustering

K-means

Classification

Supervised

Unsupervised

Convolutional Neural Network Applicications

Data Mining

LAS DATA IN DIFFERENT FORMAT

1,595 files

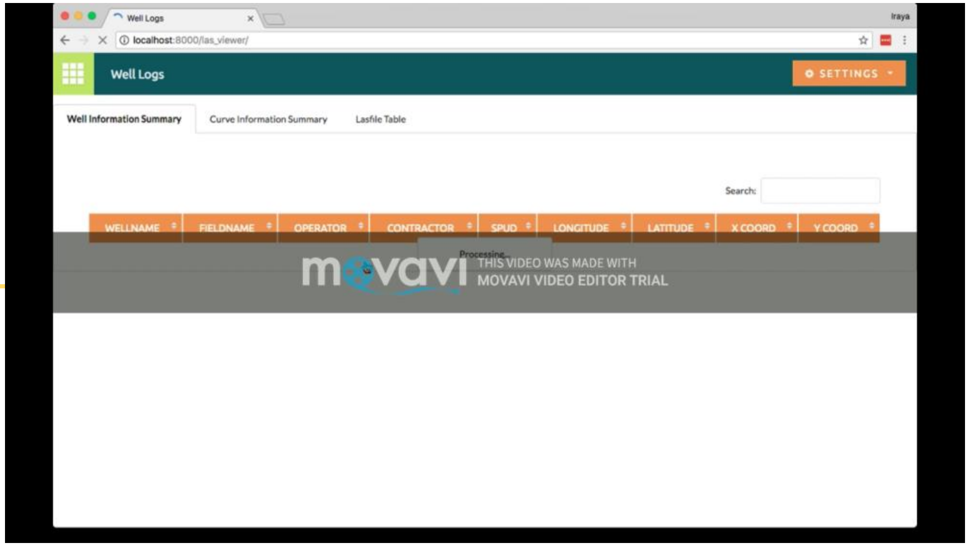
2 hrs 33.66 mins of Data Mining

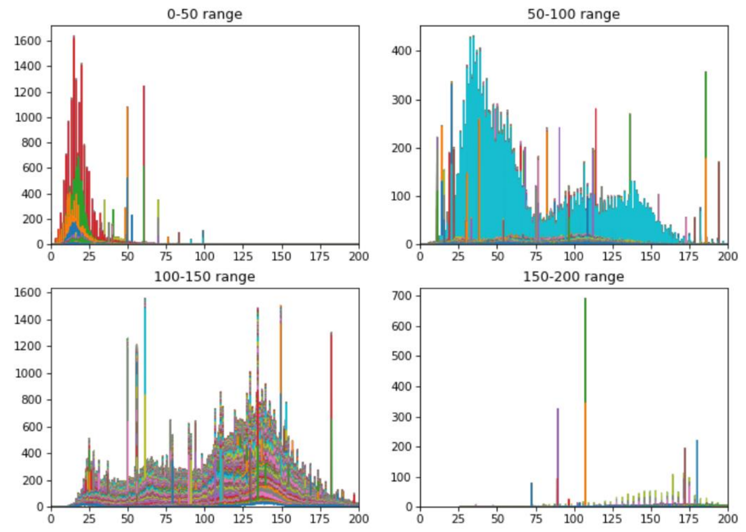
Identified:

66,515 curves

5,681 most used (10% of data)

90% of DATA REMAINS TO BE TAPPED





Use Case #2: Well Twinning

- **Problem Definition**:

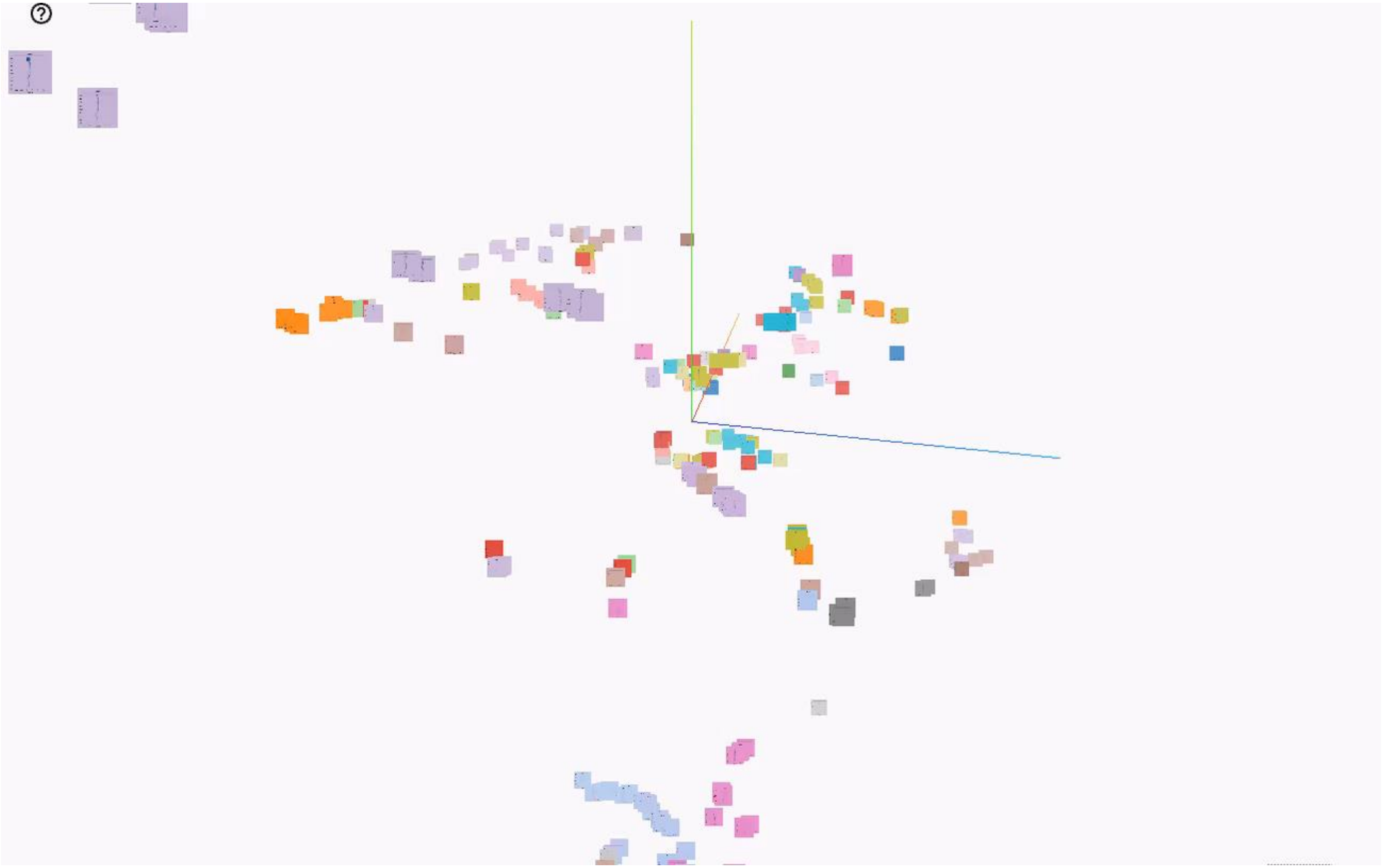
Find analog wells of a wildcat exploration area

- **Standard Solution:**

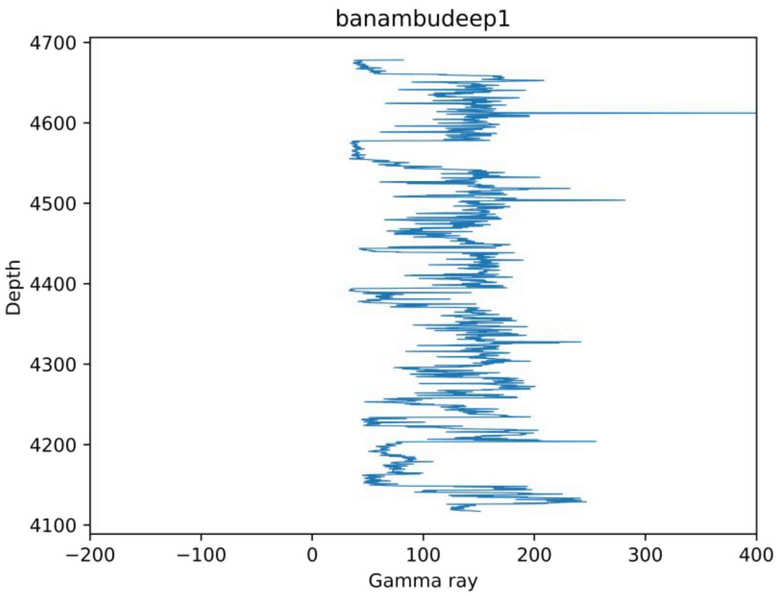
Find the nearest 1 or 2 wells in the nearest field (highly risky, does not capture all variabilities)

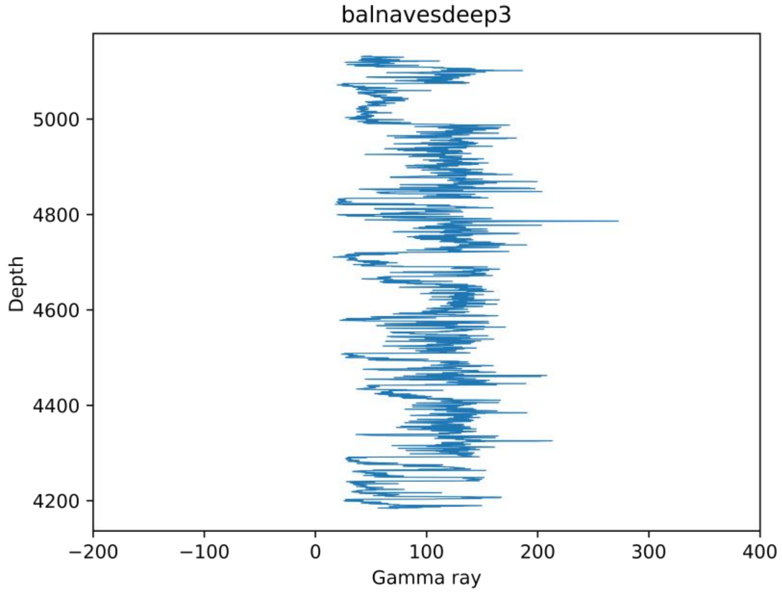
- **Machine Learning Solution:**

Leverage on big volume dataset to find geological analogs and de-risk potential prospect

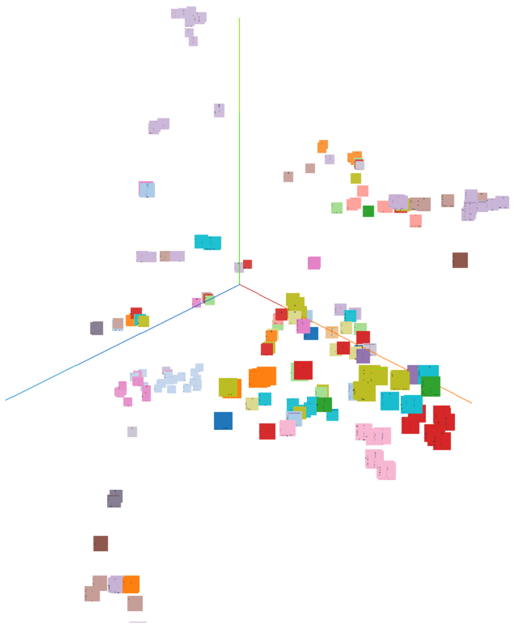


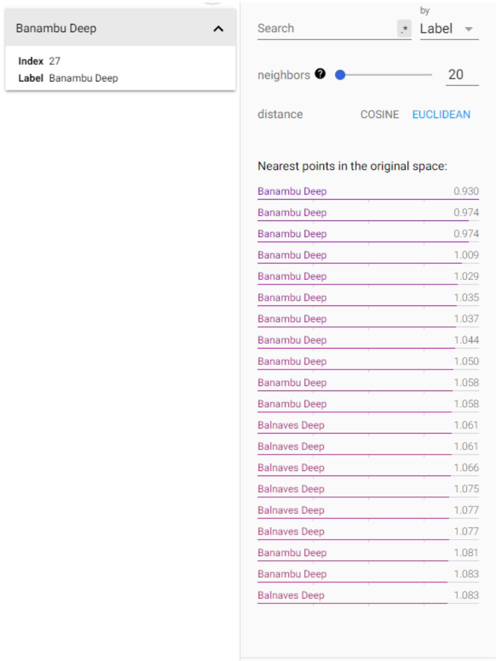
Automated Clustering of well data using t-SNE

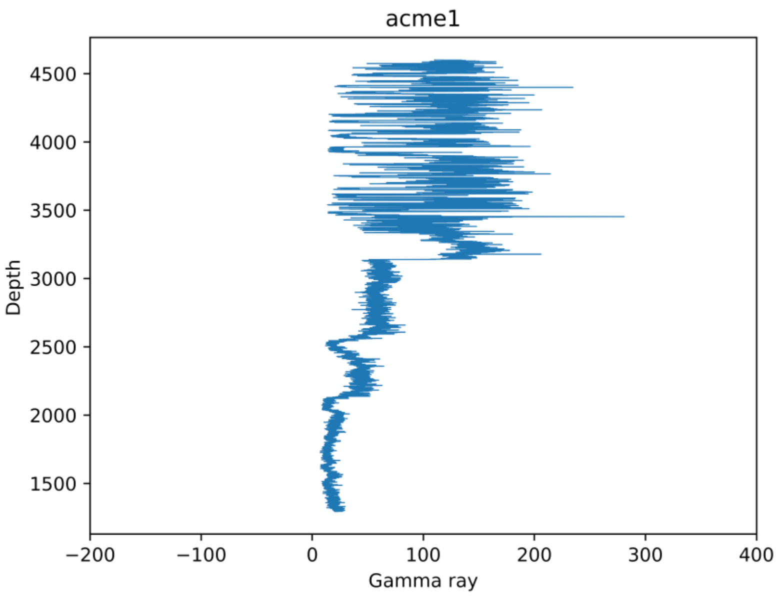


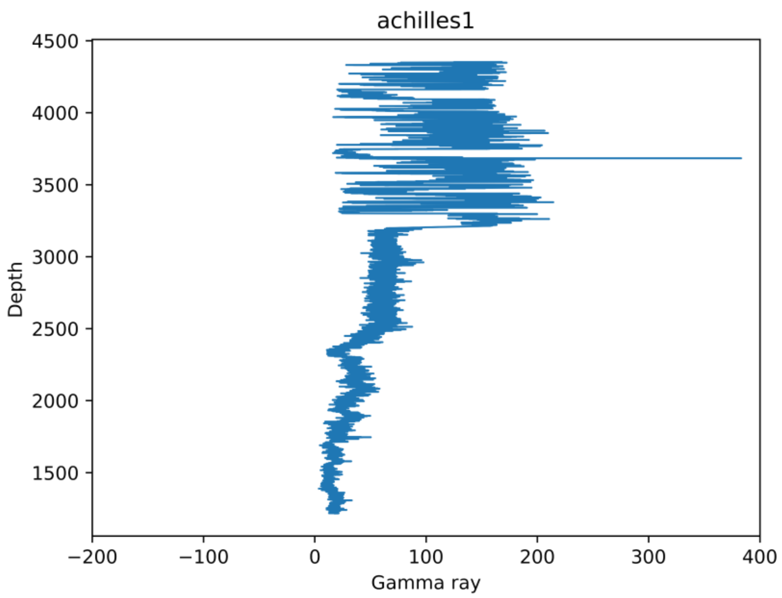


- Automated identification of the closest well "twin", without prior geological knowledge
- Applicable in ultra-wildcat area or cross-country analog search



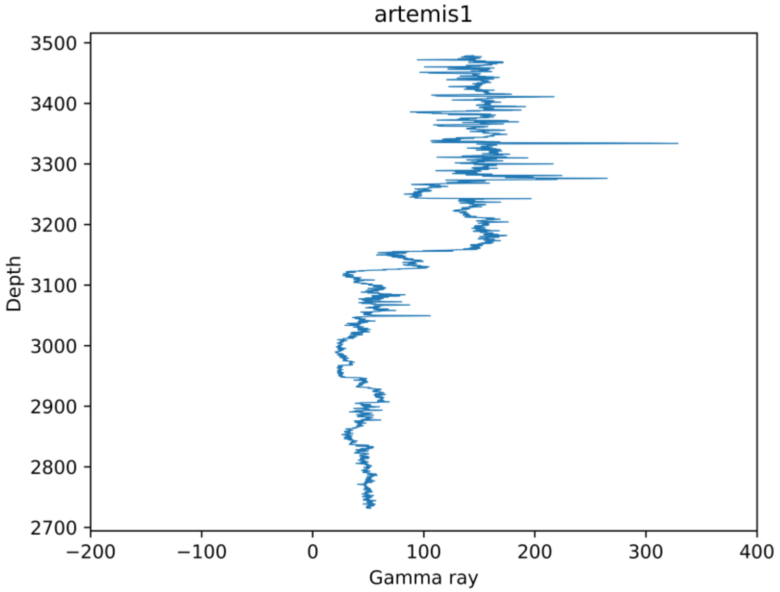


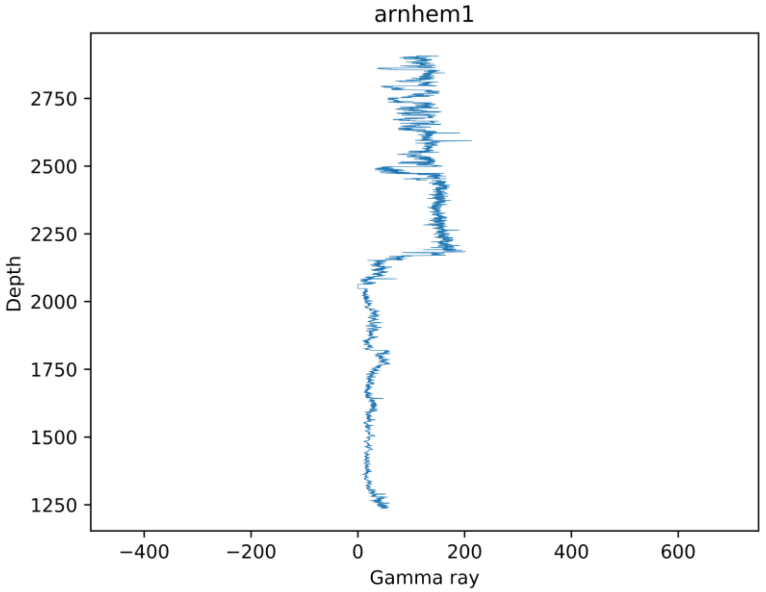




- Effective in automated identification of the closest genetic “twin” of the well

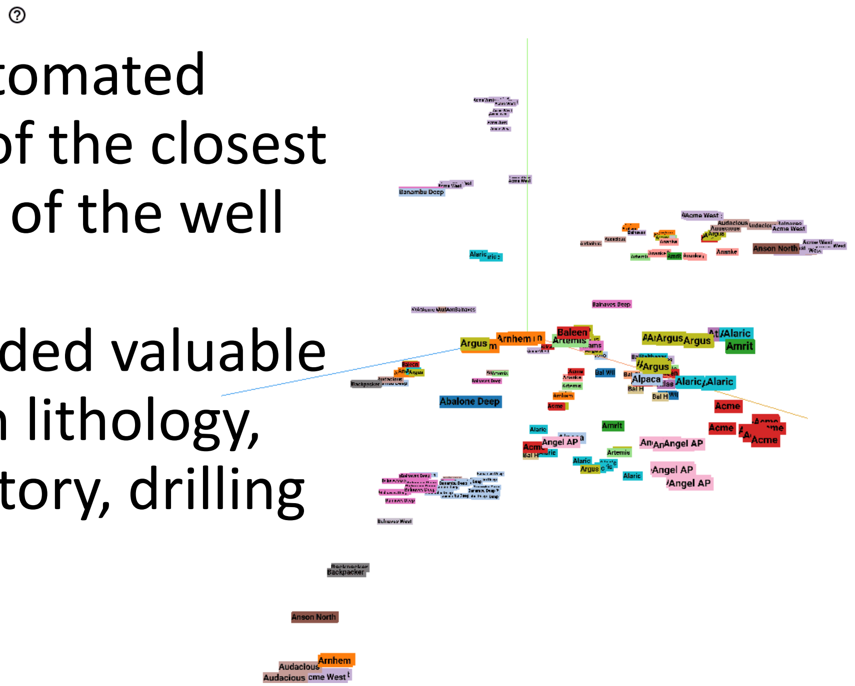
- Twin can provided valuable information on lithology, production history, drilling risks, etc.

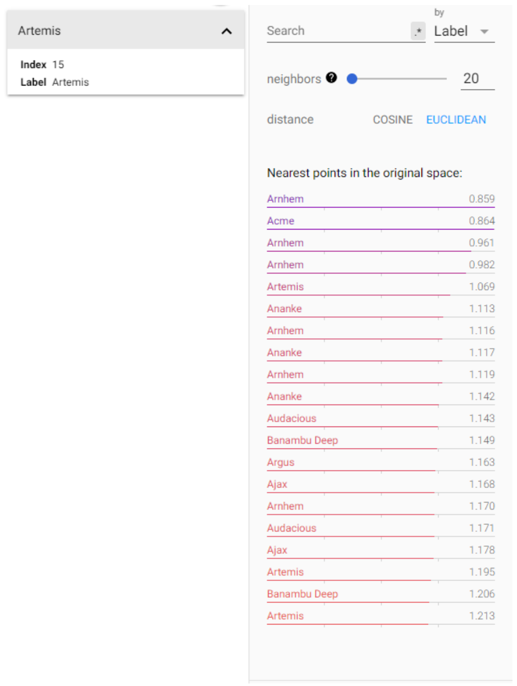




- Effective in automated identification of the closest genetic "twin" of the well

- Twin can provided valuable information on lithology, production history, drilling risks, etc.





Use Case #3: Clustering

- **Problem Definition**:

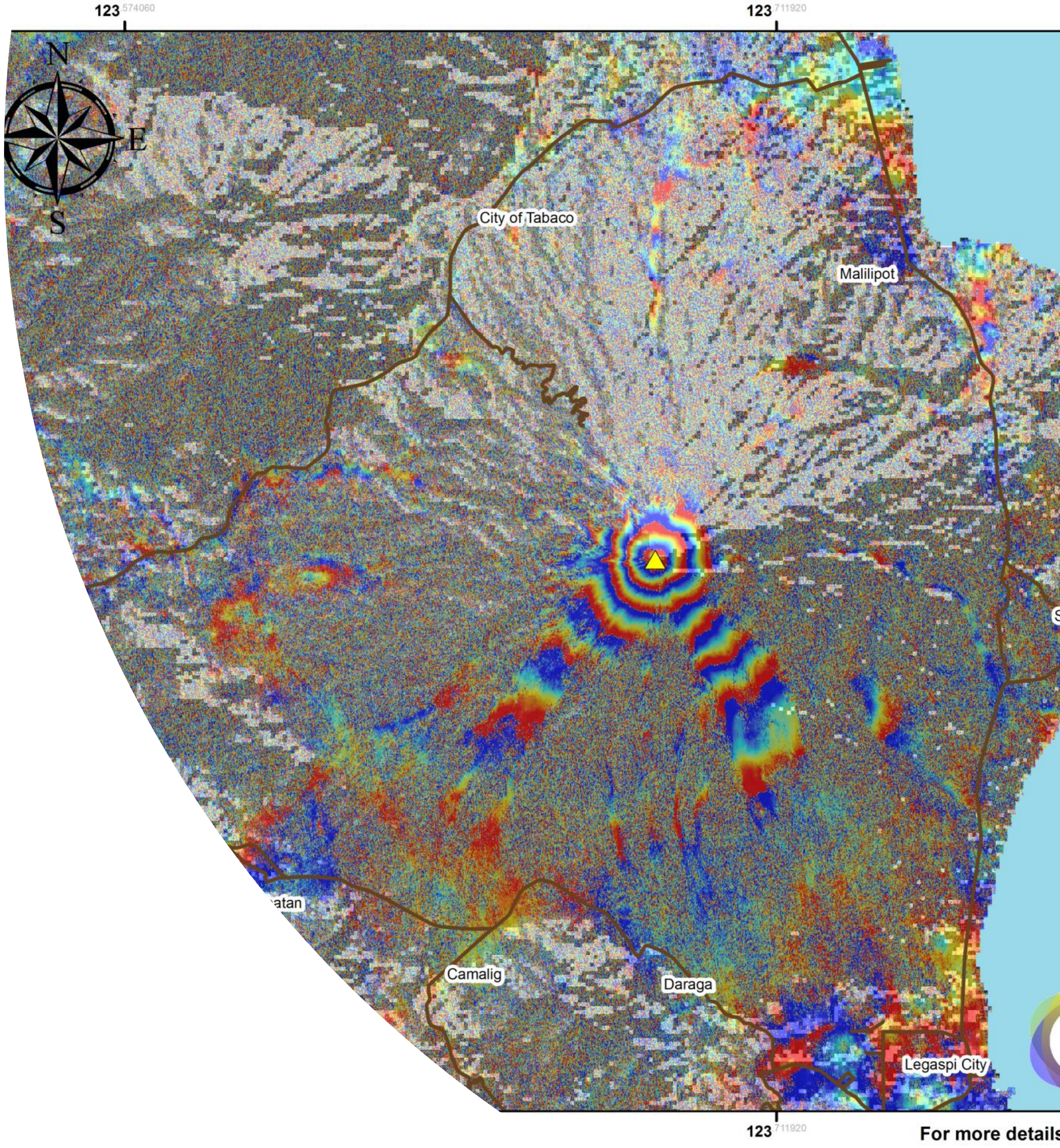
Identify surface features from satellite data

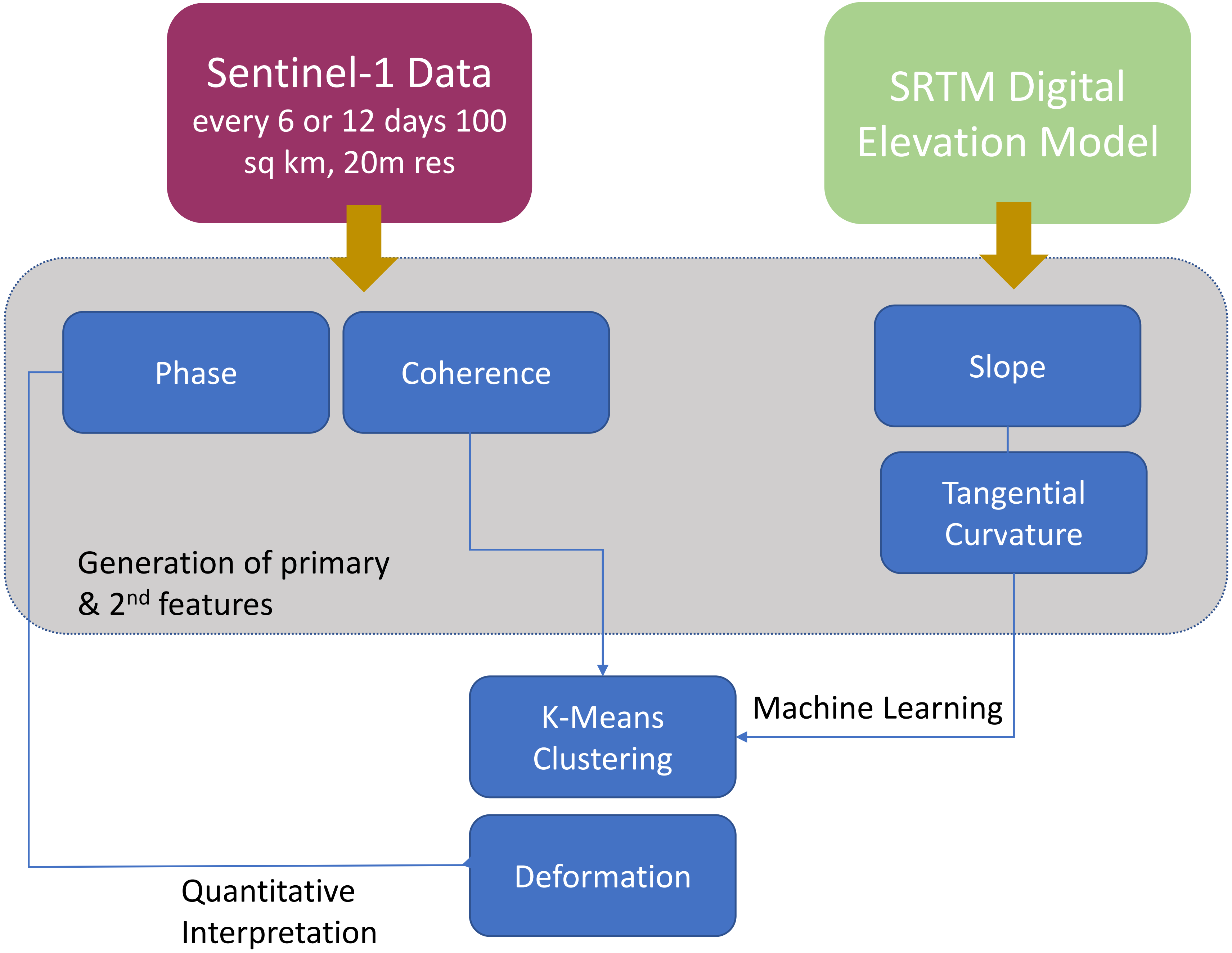
- **Standard Solution:**

Manual Interpretation

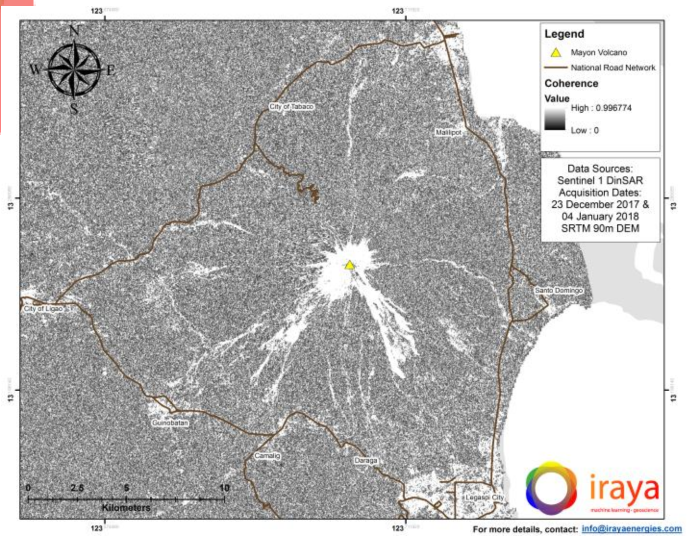
- **Machine Learning Solution:**

Unsupervised classification of multiple extracted features

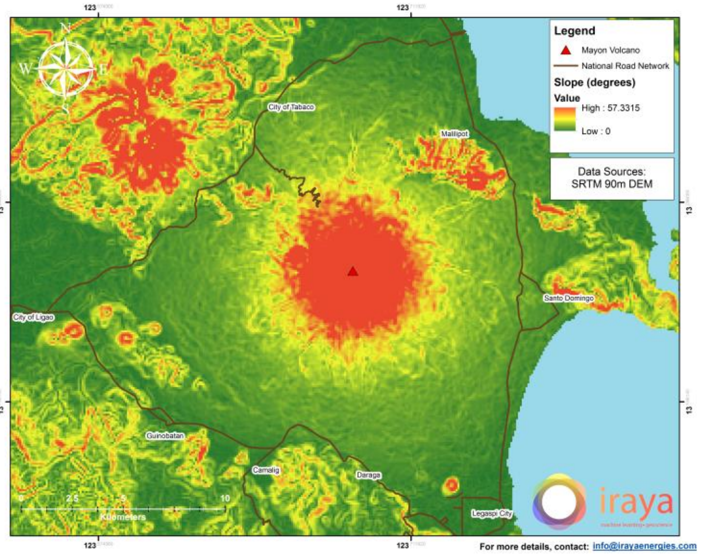



Sentinel-1 Data
every 6 or 12 days 100 sq km, 20m res
SRTM Digital Elevation Model
Phase
Coherence
Slope
Tangential Curvature
Generation of primary & 2nd features
K-Means Clustering
Machine Learning
Deformation
Quantitative Interpretation

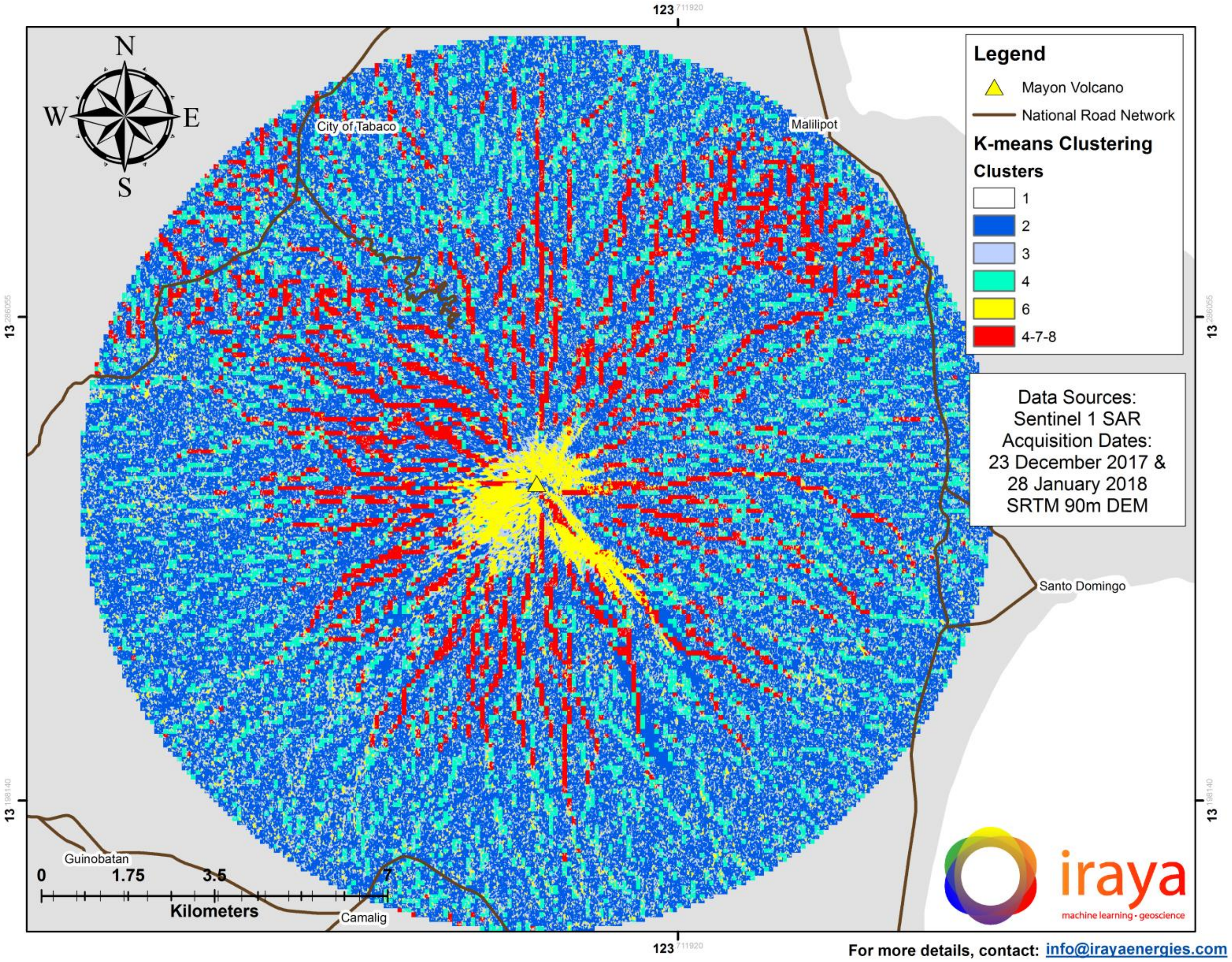
Coherence



Slope



K-means Cluster



Use Case #4: Resolution Enhancement

- **Problem Definition**:

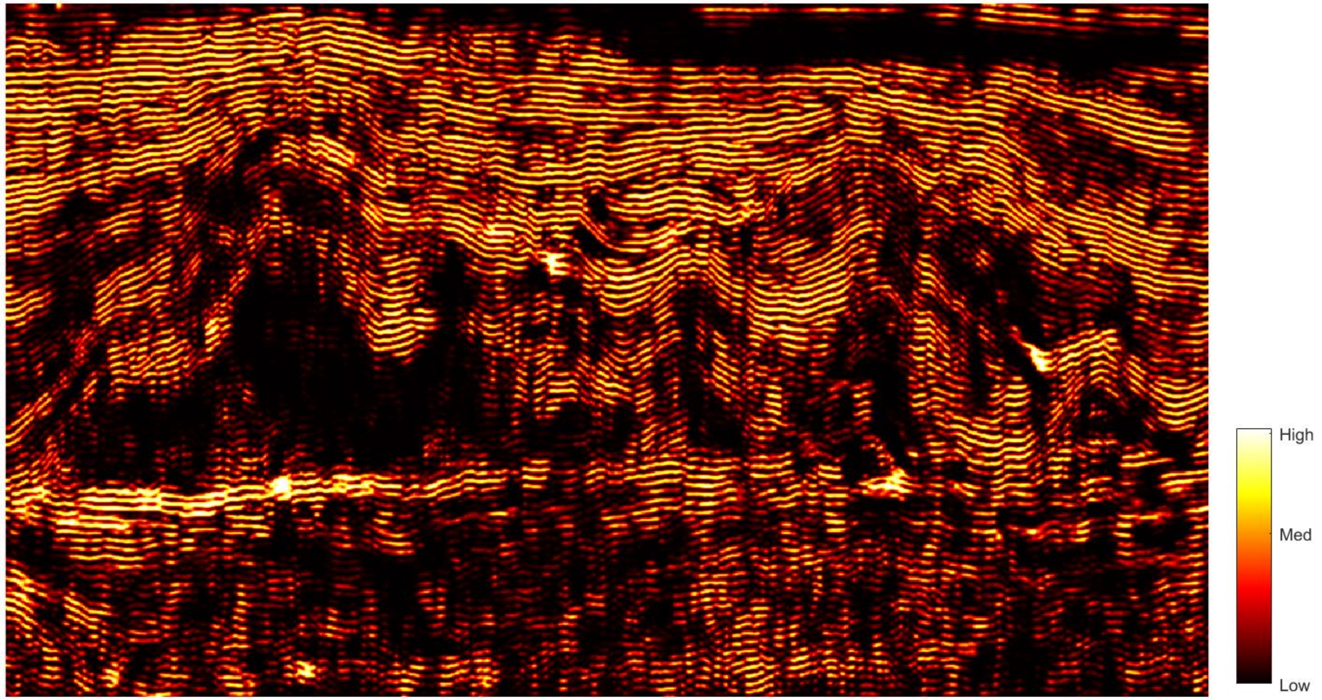
Increase seismic image Quality in Vintage Seismic acquisitions for better interpretation

- **Standard Solution:**

Traditional Seismic Processing + Stochastic Static Modeling

- **Machine Learning Solution:**

Model-based residual processing using deep convolutional neural network



Iraya Machine Learned (ML) broadband enhancement – Methodology

Architecture

Cascaded architecture using Adam optimizer, dropout 75% and transfer learning

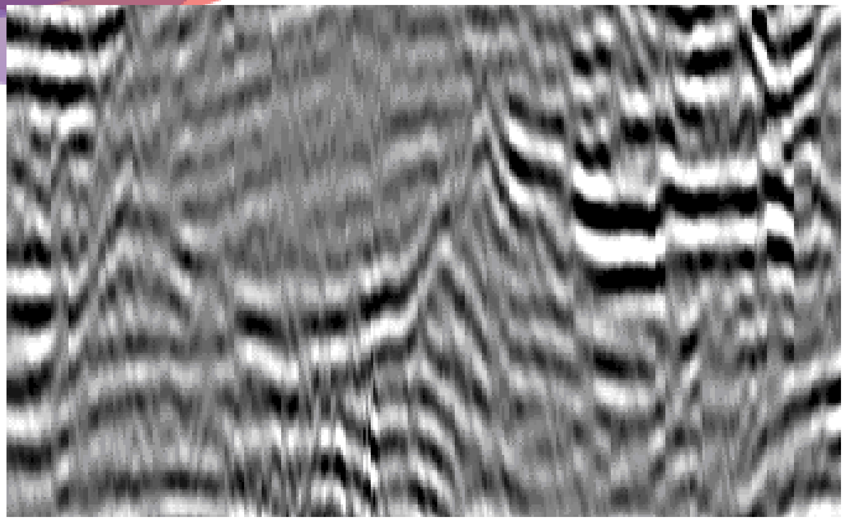
Training

Epochs of over 120,000+, global 2D seismic examples, 1 month per training

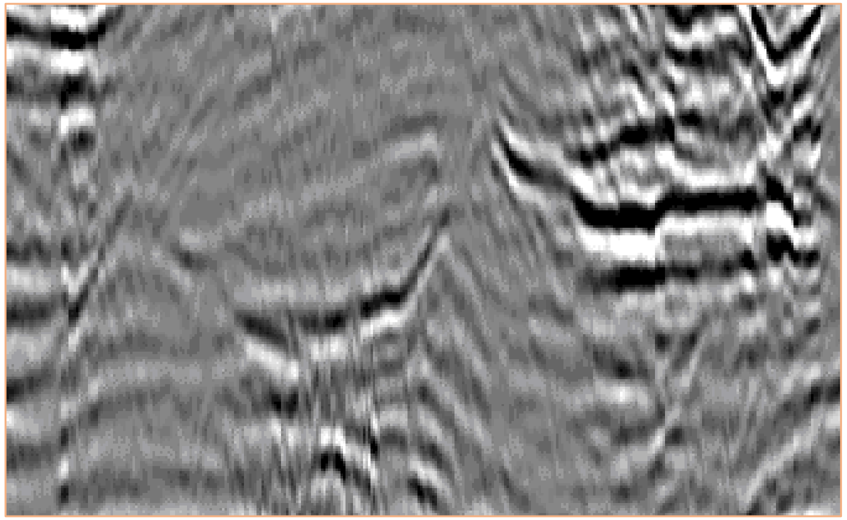
Inference

10minutes

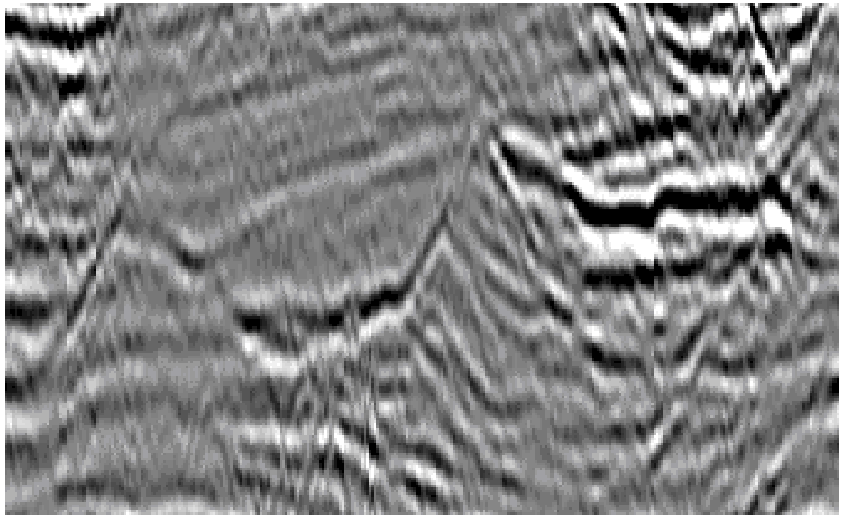
Iraya Machine Learned (ML) broadband enhancement – Training



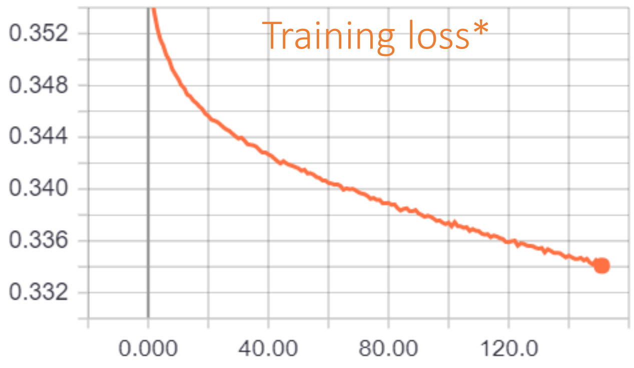
Input

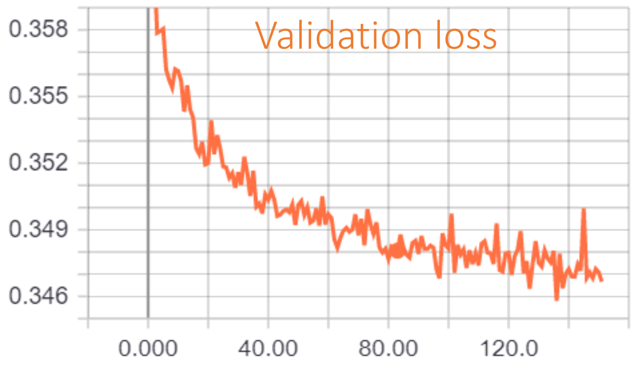


ML enhanced



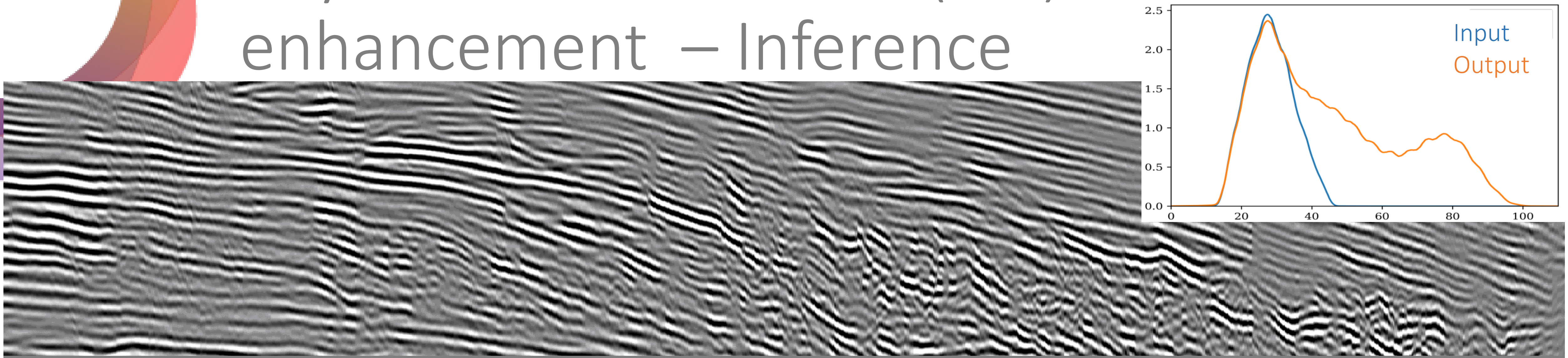
Original



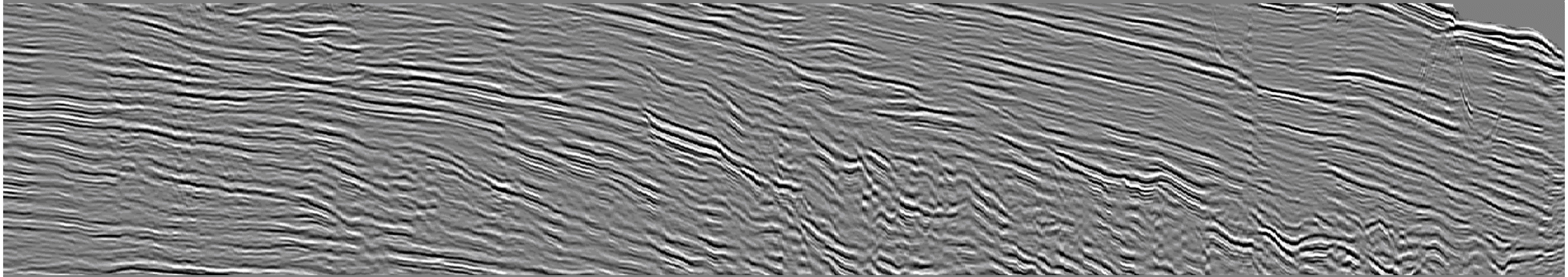


*Loss computed over 120,000+ 2D Seismic

Iraya Machine Learned (ML) broadband enhancement – Inference

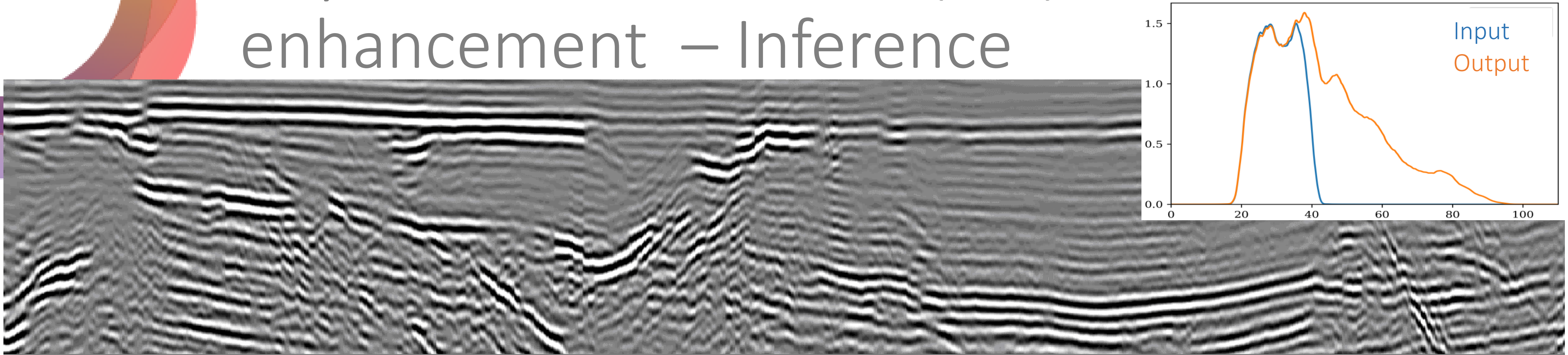


Original input

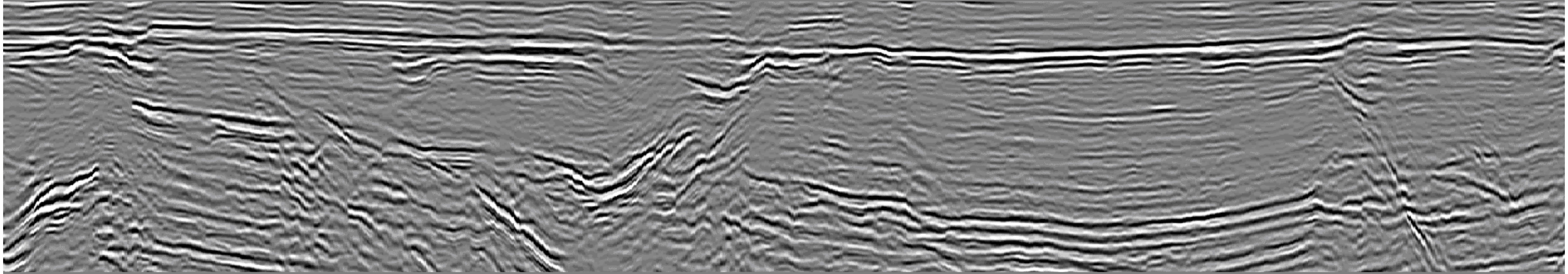


ML enhanced output

Iraya Machine Learned (ML) broadband enhancement – Inference



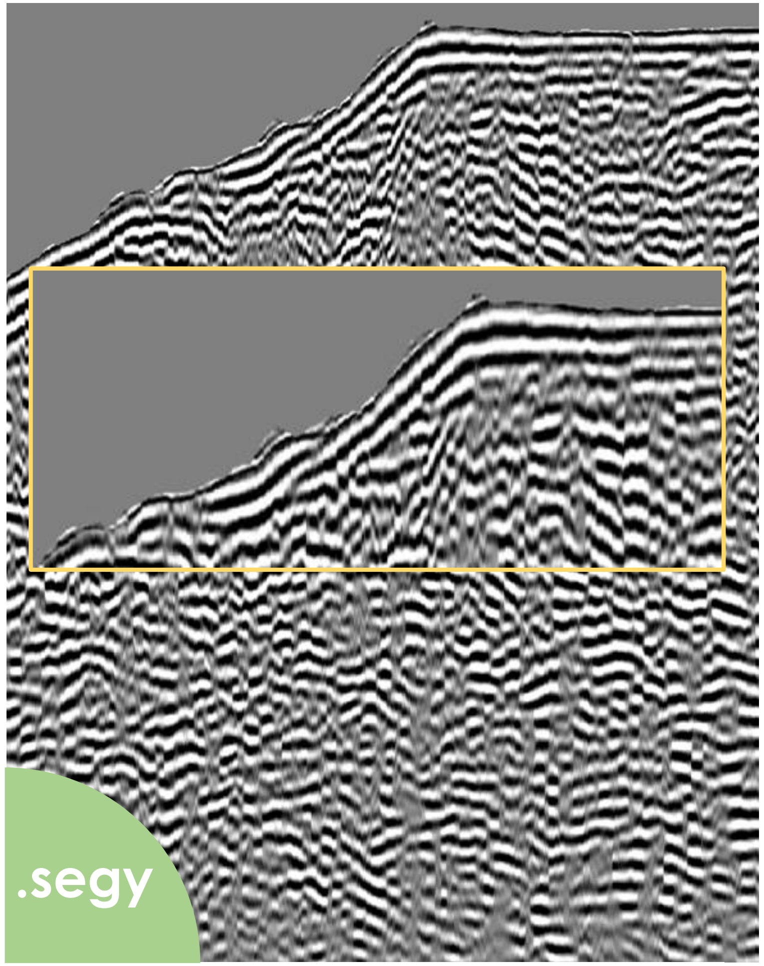
Original input



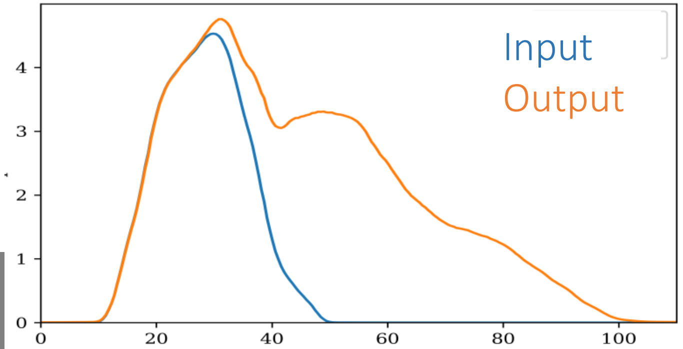
ML enhanced output

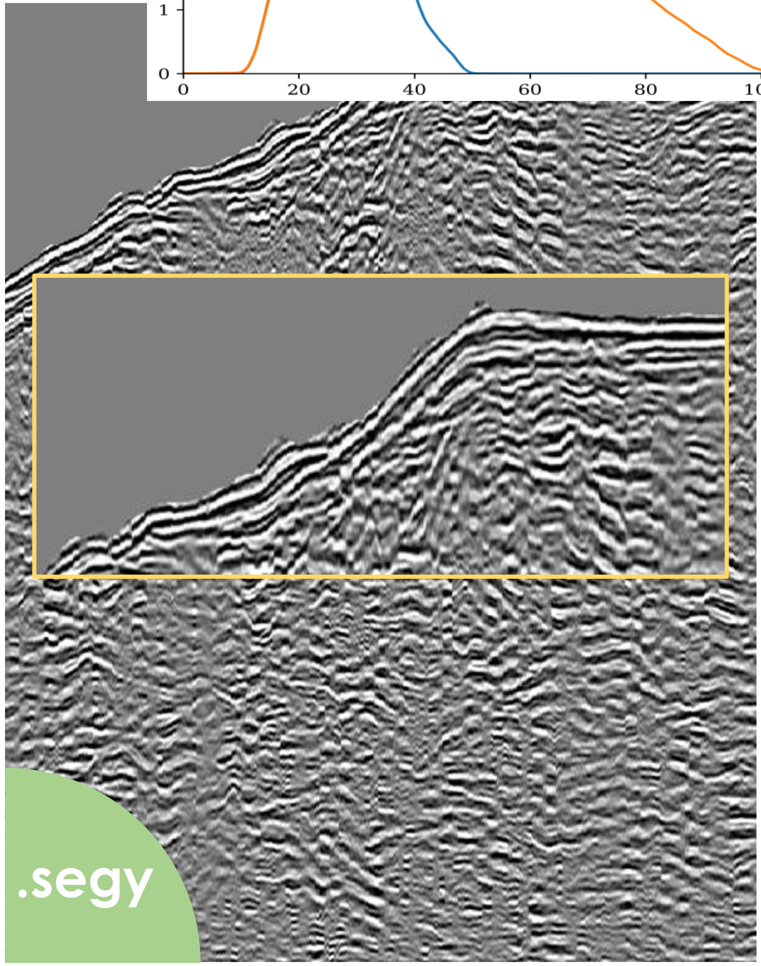
Let’s assume…

We have a powerpoint with somewhere a seismic image



We would like to find it, clean it and convert it into .segy





And why not enhance it at the same time

Fully automatic - AI driven

Conclusion:

How to build an AI project

Conclusion

D2V Approach

- ✓Define Problem
- ✓Define Data
- ✓Train
- ✓Test
- ✓Validate

Classification: Supervised vs Unsupervised
Prediction: Regression

Available : Check quality, quantity
Not Available: Find other sources? Public data, etc.

Thank you!

To build AI experiments or participate in beta test, you may contact:

info@irayaenergies.com or nmh@irayaenergies.com