Webinars

Rock Type Classification Using Machine Learning

Sridharan Chandrasekaran, PetroAnalytix
Taufan Rusady , PetroAnalytix

SPE International
Drilling Uncertainty Prediction Technical Section

Webinars

SPE Webinars are Sponsored by:

SPE
SPE Foundation

SPE International
Drilling Uncertainty Prediction Technical Section

Webinars

Sridharan Chandrasekaran

Taufan Rusady

SPE International
Drilling Uncertainty Prediction Technical Section

Agenda

- Drilling uncertainty
- Machine learning
- Society of Exploration Geophysicists (SEG) machine learning contest
  - Workflow
- Real-time lithology classification

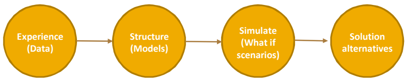
Poll Question

What is your role within the organization ?

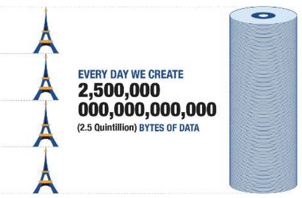
Agenda

- Drilling uncertainty
- Machine learning
- Society of Exploration Geophysicists (SEG) machine learning contest
  - Workflow
- Real-time lithology classification

Introduction



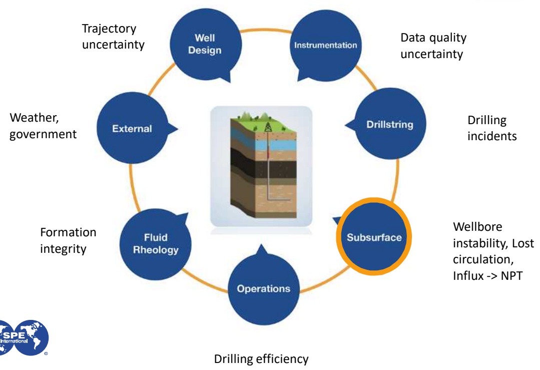
Data



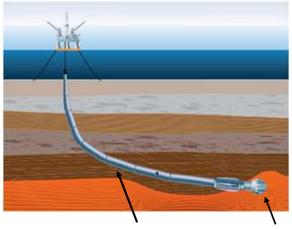
- The world is driven by data (people, sensors and devices)
- 2.5 Exabyte of data per day.
  (1 Exabyte = 10^18 bytes)
- Less than 5% of all data is ever analyzed
  (*Source : Report [1]*)
- Opportunity, Challenge

[1]

Drilling Uncertainty Factors



Drilling Uncertainty



Determination of lithology type at the bit

- Drill bit as sensor - vibrations
- Data driven models
  - Historical database
  - Computational resources
- Challenge : Data in silos

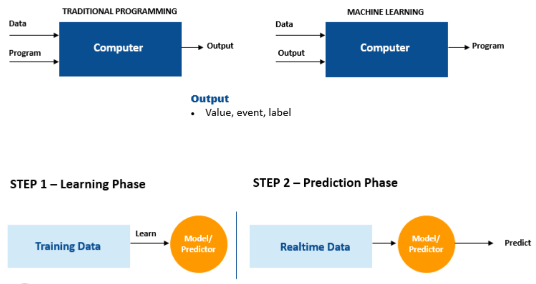
Agenda

- Drilling uncertainty
- Machine learning
- Society of Exploration Geophysicists (SEG) machine learning contest
  - Workflow
- Real-time lithology classification

Poll Question

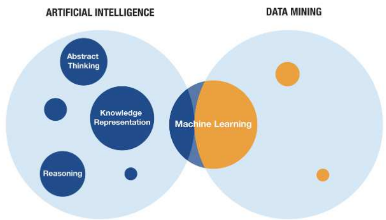
What type of analytics do you or your team/division deal with predominantly?

Machine learning



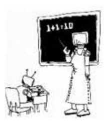
Machine Learning

- Where does it fit ?



Learning paradigms

- Supervised Learning
  - Knowledge of the output
    - Data is labelled
    - Goal : Predict the class (classification) or value (Regression)



GR	Resistivity	Porosity	Type
77	0.66	11	Siltstone
43.23	0.69	15	Dolomite

Label

- Unsupervised Learning
  - No knowledge of the output
    - Data is unlabelled
    - Goal : Determine data patterns

GR	Resistivity	Porosity
77	0.66	11
43.23	0.69	15

SEG Contest → Supervised Classification Problem

Agenda

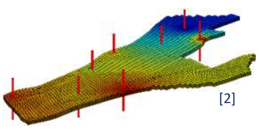
- Drilling uncertainty
- Machine learning
- Society of Exploration Geophysicists (SEG) machine learning contest
  - Workflow
- Real-time lithology classification

Poll Question

What is the degree of Machine Learning adoption in your organization/team?

SEG Machine learning contest

- Society of Exploration Geophysicists (SEG) Open Contest
- Accurate differentiation of facies (rock-type) by indirect methods
- Dataset – 10 Wells
- 40+ teams

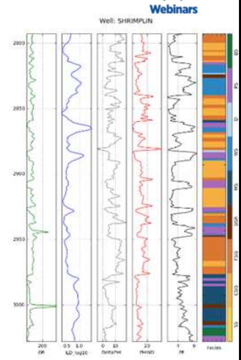


[2]

[3]

SEG Machine learning contest

- Objective : Predict the rock type for two wells
- Inputs : Wireline logs of 10 wells and their interpreted facies
- Predictor variables
  - Five Wireline logs
  - Two Geologic constraining variables
- No access to blind/test data



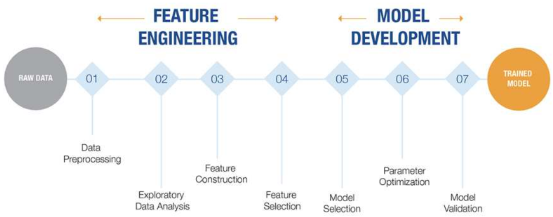
https://github.com/seg/2016-ml-contest [4]

Challenges

- Missing data
- Small sample size
- Facies (rock-type) are not discrete

Tag	Facies Label	Rock Facies Type	Adjacent Facies Label
SS	1	Nonmarine sandstone	2
CsiS	2	Nonmarine coarse siltstone	1,3
FSiS	3	Nonmarine fine siltstone	2
SiSh	4	Marine siltstone and shale	5
MS	5	Mudstone (limestone)	4,6
WS	6	Wackestone (limestone)	5,7
D	7	Dolomite	6,8
PS	8	Grainstone (limestone)	6,7,9
BS	9	Bafflestone (limestone)	7,8

Machine Learning Process

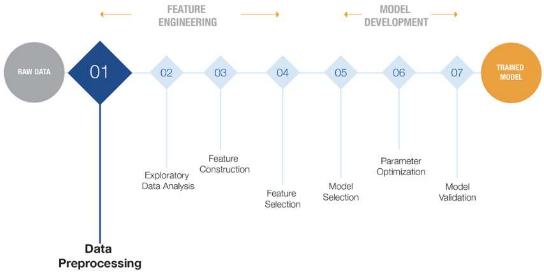


- Extract inherent information
- Offers flexibility
- Domain intensive

- Right choice of algorithms
- Computationally intensive
- Automation (Jupyter, Python)

Learning Process

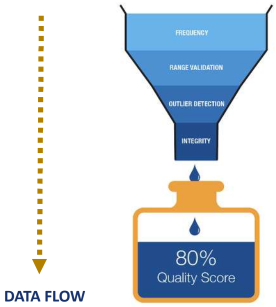
Webinars



SPE International

Data Quality Analysis

Webinars



Gaps in the data

0 < Block Height < 150

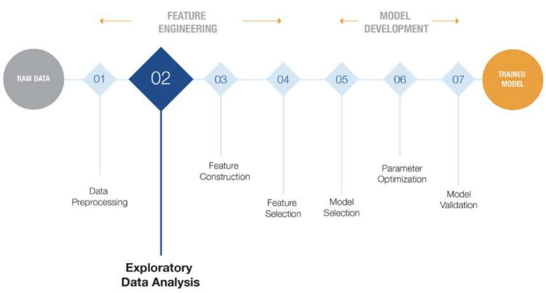
Statistical (Window based methods, Z-Score, Dixon test)

Is bit depth > hole depth ?

SPE International

Learning Process

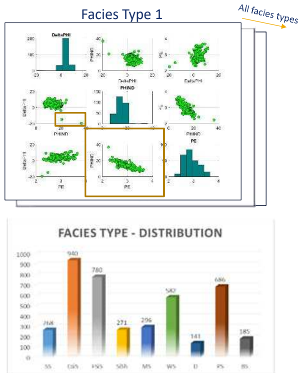
Webinars



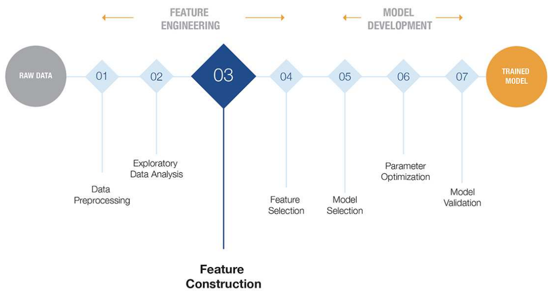
SPE International

Exploratory Data Analysis (EDA)

- Summarize data characteristics by visualization
- Objective
  - Discover patterns
  - Spot anomalies
  - Check assumptions

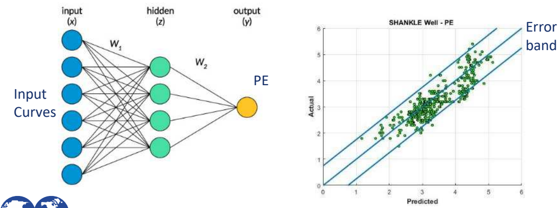


Learning Process



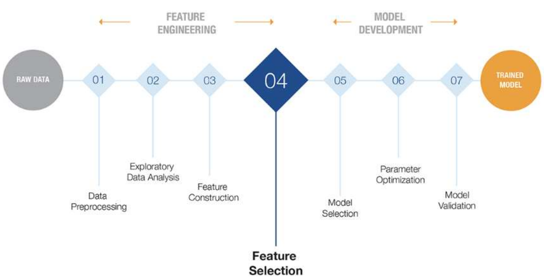
Missing data

- Missing data introduces bias, in-efficiency and handling issues
- Missing PE observations
- Model based imputation
- Neural Network regression to estimate missing data



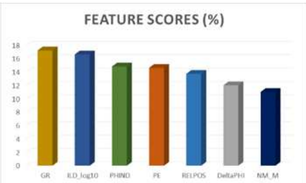
Increase in the data availability by 20 %

Learning Process



Feature Selection

- Select features that contribute to the accuracy
- Benefits
  - Alleviate the "curse of dimensionality"
  - Enhance accuracy
  - Speed up learning process



Feature contribution in training data

Feature Augmentation

Cappuccino ?

Cat

Rotation,
Translation,
Scaling

- Generate new feature from existing features
- Neighbouring layers are strongly related
- More time for learning

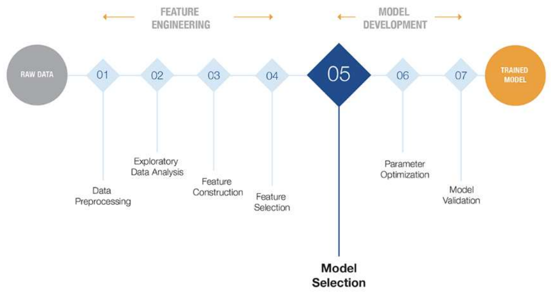
Dataset

Depth	GR	Feature 1	Feature 2	Feature 3	Feature 4
999	20	-	-	-	-
999.5	22	-	-	-	-
1000	**45**	**20**	22	48	52
1000.5	48	-	-	-	-
1001	52	-	-	-	-

[5]

Augmented Features

Learning Process

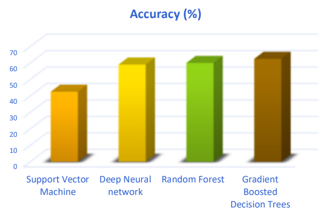


Model Selection

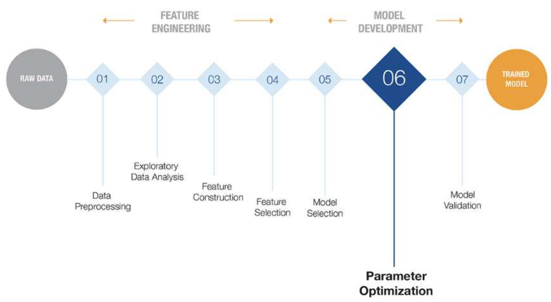
- Model/Algorithm
  - Boosted Tree
  - Neural Network
  - Random Forest

dmlc XGBoost

theano



Learning Process



Parameter Optimization

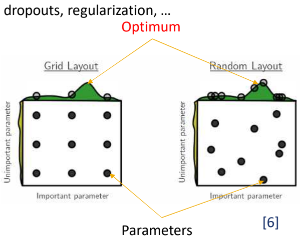
Purpose: find optimum combination of parameters which maximize accuracy
Example of hyper-parameters optimization:

- o Gradient Boosted Decision Tree
  - #trees, Max. tree depth, …
- o Neural Network
  - #hidden layer, #mini batch, dropouts, regularization, …

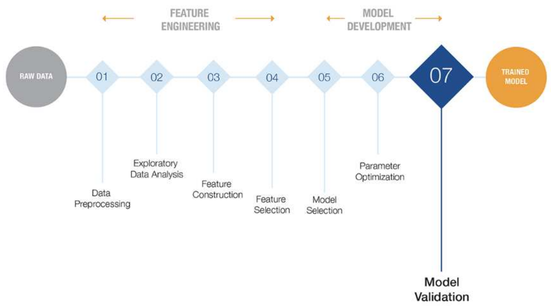
Problems:

- Curse of dimensionality
- Time and computing power

Methods: grid and random search

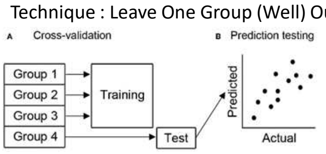


Learning Process



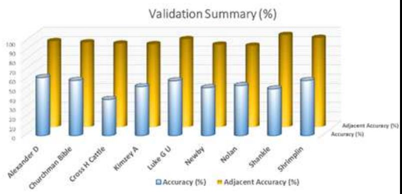
Model Validation

Technique : Leave One Group (Well) Out

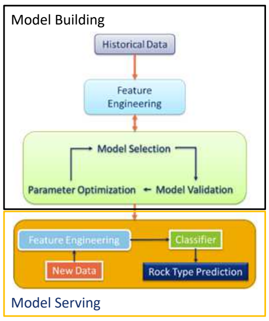


Facies	Label	Adjacent Facies
1	SS	2
2	CSiS	1,3
3	FSiS	2
4	SiSh	5
5	MS	4,6
6	WS	5,7
7	D	6,8
8	PS	6,7,9
9	BS	7,8

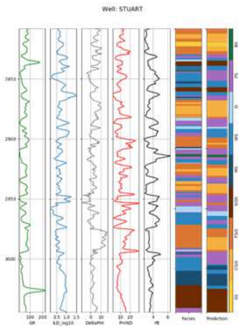
Facies		Metric	
Prediction	Actual	Accuracy	Adjacent
1	1	TRUE	TRUE
2	1	FALSE	TRUE
6	1	FALSE	FALSE



Workflow Summary



Blind Well Results



Actual vs Predicted comparison

Depth	*Facies*	*Prediction*
2878.5	2	2
2879	2	2
2879.5	3	2
2880	3	2
2880.5	2	2
2881	2	2
2881.5	3	3
2882	3	3
2882.5	7	3
2883	7	8
2883.5	7	8
2884	7	8
2884.5	7	8
2885	8	8

Facies	Label	Adjacent Facies
1	SS	2
2	CSiS	1,3
3	FSiS	2
4	SiSh	5
5	MS	4,6
6	WS	5,7
7	D	6,8
8	PS	6,7,9
9	BS	7,8

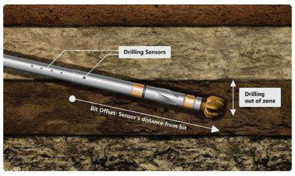
Blind Well	Accuracy (%)	
	True	Adjacent
Stuart	61.5	94
Crawford	63.5	88

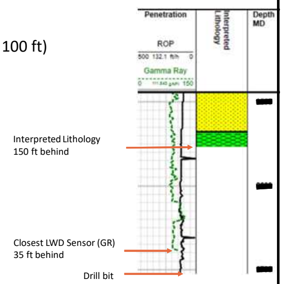
Agenda

- Drilling uncertainty
- Machine learning
- Society of Exploration Geophysicist (SEG) machine learning contest
  - Workflow
- Real-time lithology classification

Real-time Lithology Classification

- LWD data is proxy of lithology type
- Lithology vs Facies
- Each LWD sensors has an offset (35 ft – 100 ft)

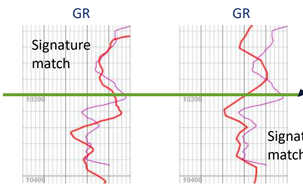




Challenge
1. Can we automate lithology interpretation and closing the gap?
2. Generate pseudo logs between sensors and bit?

Look Ahead

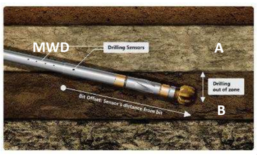
Correlate during drilling

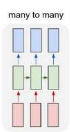


Assume LWD (GR) is still at this depth.
Which reference you trust to look ahead?

GR offset well (Reference)
GR current well

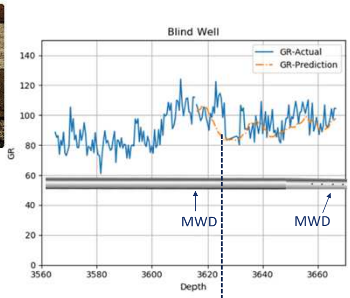
GR Projection Ahead of MWD





Recurrent Neural Network

[7]

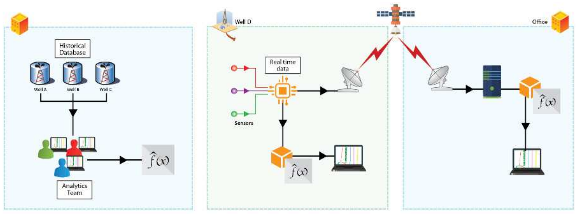


Real Time Lithology Classification

Topology of real-time classification

Model Building

Model Serving



Summary

Machine Learning has the potential to

- classify rock facies in wells with minimal supervision
- reduce uncertainty while looking ahead of the bit

Feature engineering improves model prediction (guided by domain experts)

Acknowledgements

- Matt Hall (Organizer of the contest) and participants
- SEG
- SPE Foundation

References

1. McKinsey Global Institute Analysis Report , June 2015, McKinsey & Company
2. Nilsen H.M., Lie K.A., Natvig J.R. (2010) Accurate Modelling of Faults by Multipoint, Mimetic and Mixed Methods, SPE 149690.
3. Martin K. Dubois, Geoffrey C. Bohling, Swapan Chakrabarti, Comparison of four approaches to a rock facies classification problem, In Computers & Geosciences, Volume 33, Issue 5, 2007, Pages 599-617, ISSN 0098-3004
4. https://github.com/seg/2016-ml-contest
5. Paolo Bestagini, Vincenzo Lipari, and Stefano Tubaro (2017) A machine learning approach to facies classification using well logs. SEG Technical Program Expanded Abstracts 2017: pp. 2137-2142.
6. Bergstra, Bengio, Random Search for Hyper-Parameter Optimization, Journal of Machine Learning Research 13 (2012) 281-305
7. The Unreasonable Effectiveness of Recurrent Neural Networks(http://karpathy.github.io/2015/05/21/rnn-effectiveness/)

Thank you

Questions?