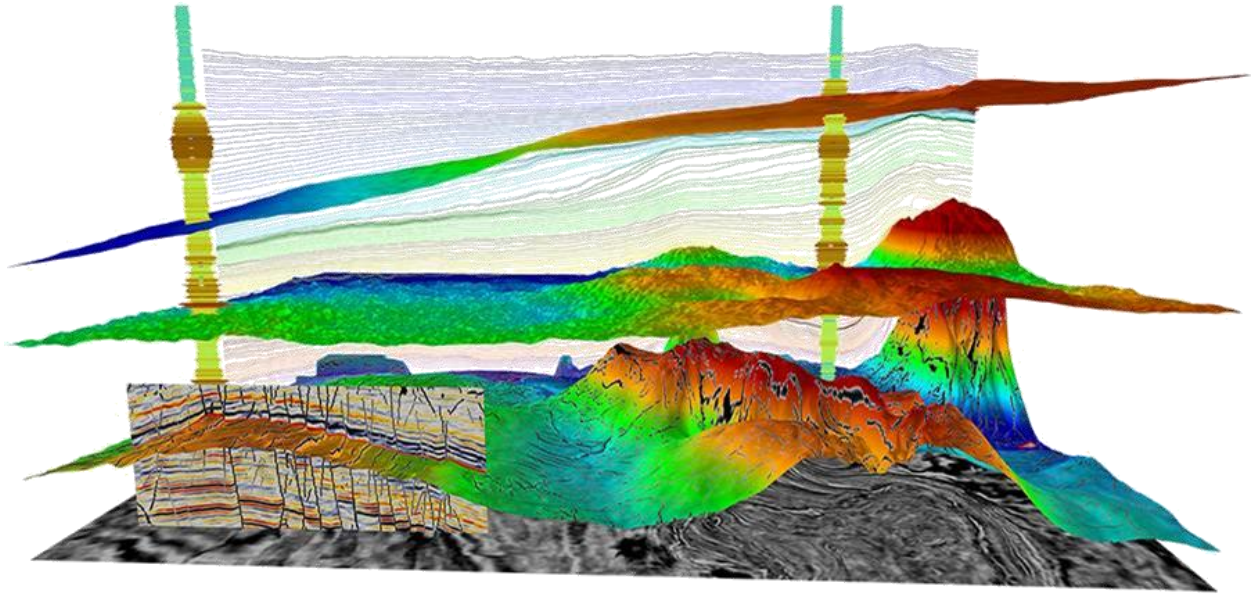
AUGMENTING SEISMIC INTERPRETATION

PREDICTING ROCK TYPES FROM 3D SEISMIC DATA TO IMPROVE WORKFLOW PIPELINES



PROJECT

Interpreting seismic data is hard!

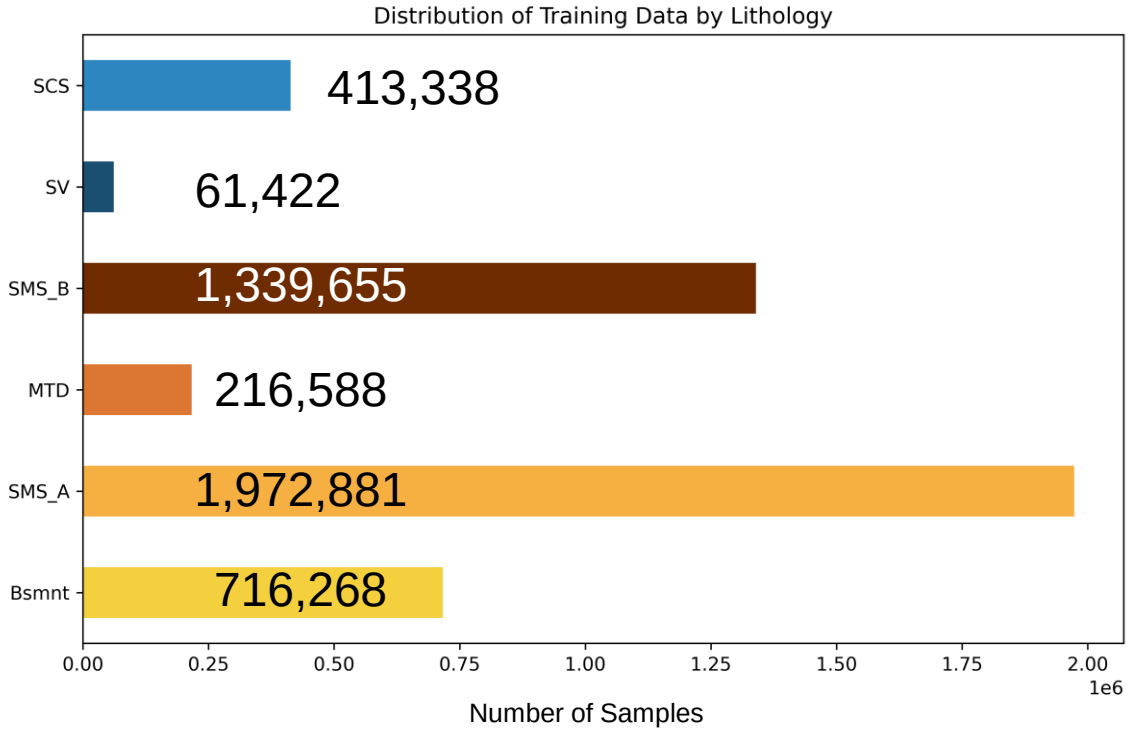
Prediction of seismic lithologies using feature engineering and tree-based classifiers

Goal 1: Robust prediction with atypical geophysical features, and data augmentation

Goal 2: Deeper understanding of impactful feature types for tree-based methods

DATA & TOOLS

- 464,148,280 samples (rows)
- 25 features (seismic volumes)



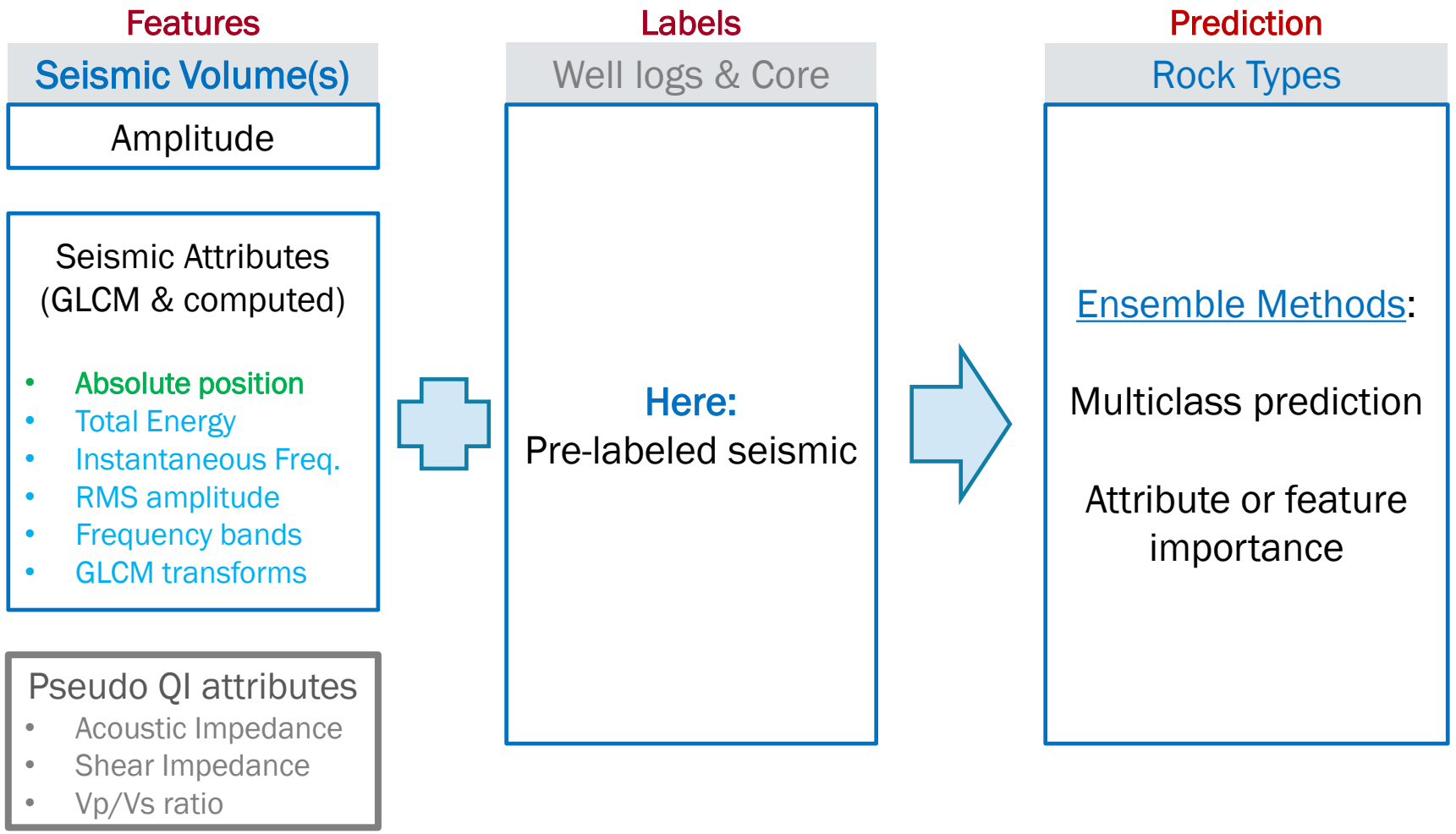
Alcrowd

pandas

$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$

scikit learn

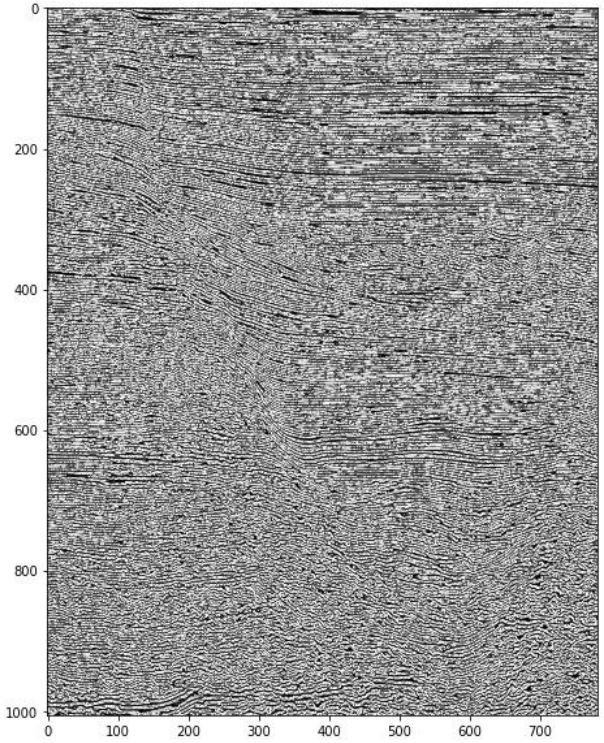
AUGMENTED WORKFLOW



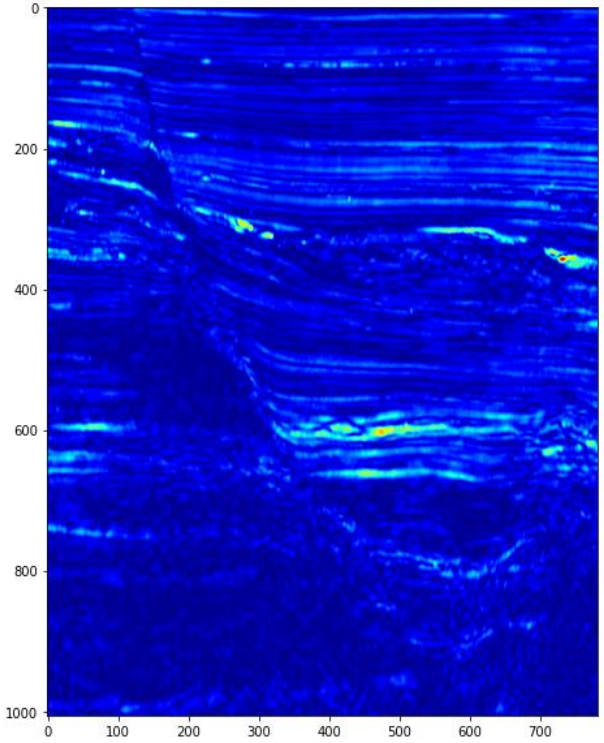
FEATURES

Seismic Amplitude	Max Peak Amplitude	Glcm_asm	Glcm_maxamp	Fband_04_08hz
Standardized Amp	RMS Amplitude	Glcm_contrast	Glcm_mean	Fband_08_16hz
Cos of Inst. Phase	X-position	Glcm_dissimilarity	Glcm_stdev	Fband_16_32hz
Envelope Amplitude	Y-position	Glcm_entropy	Glcm_total_energy	Fband_32_64hz
Hilbert Transform	Z-position	Glcm_homogeneity	Fband_02_04hz	Fband_64_128hz

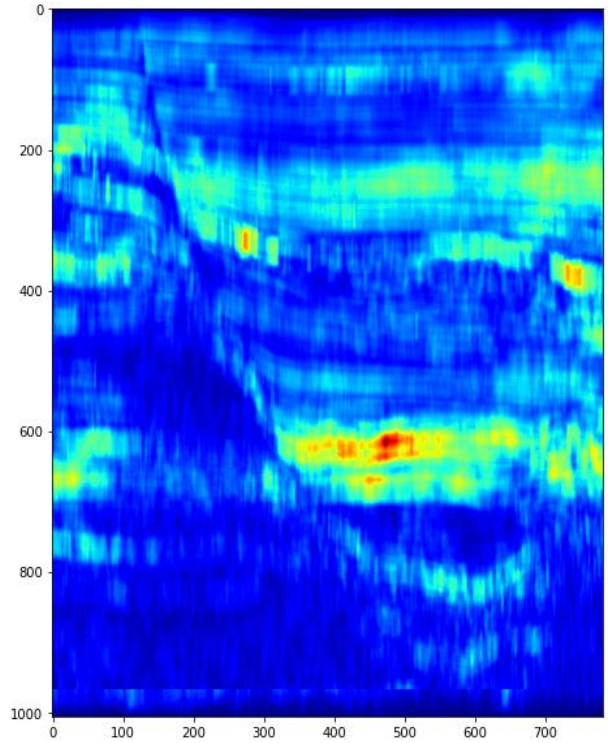
Instantaneous Phase



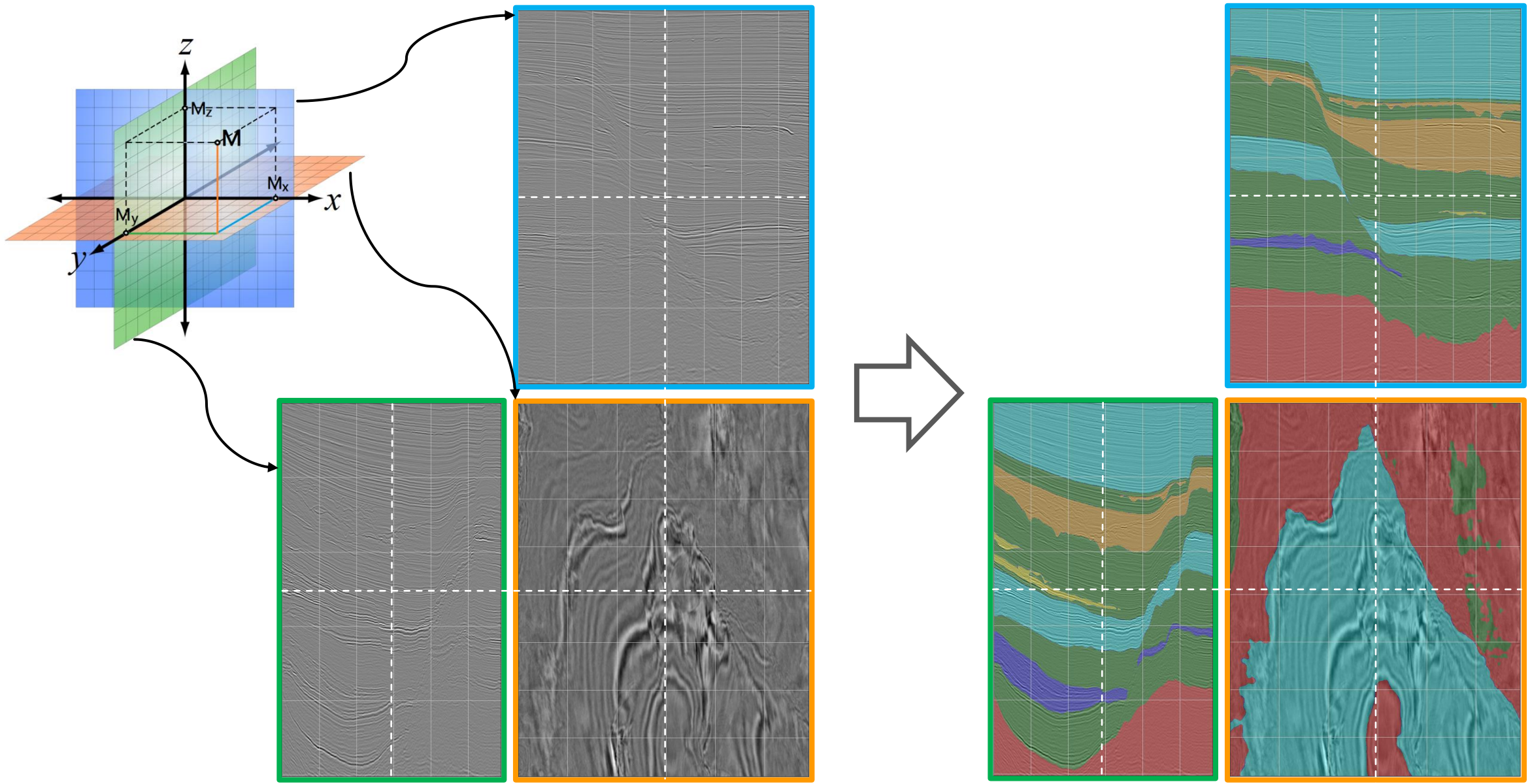
Envelop Amplitude



RMS Amplitude



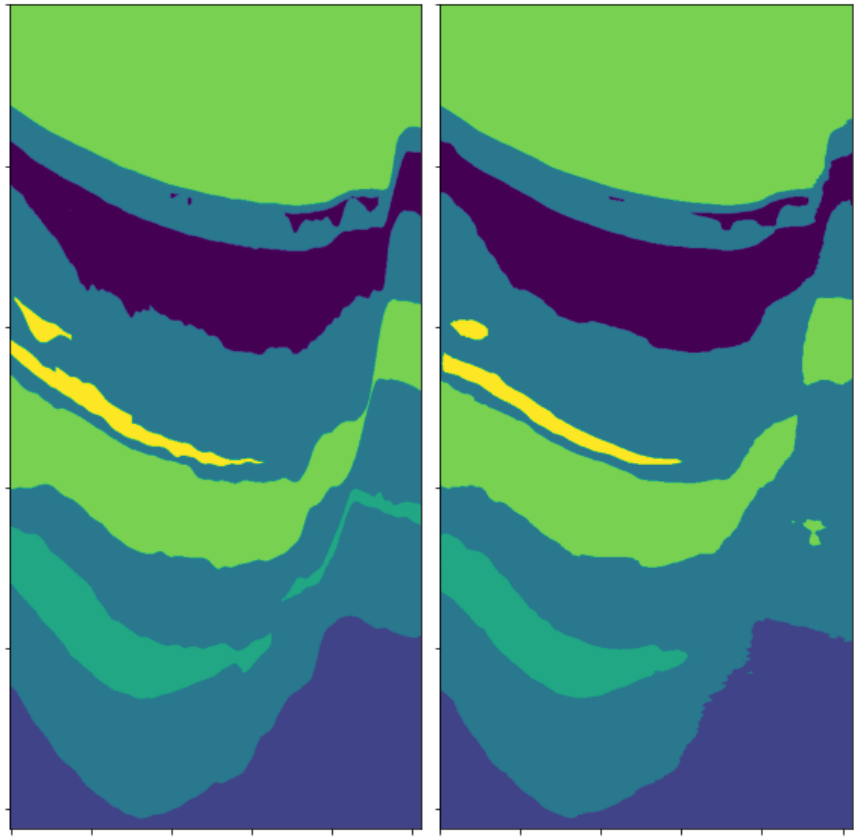
SUPERVISED LEARNING: AIDING SEISMIC INTERPRETATION



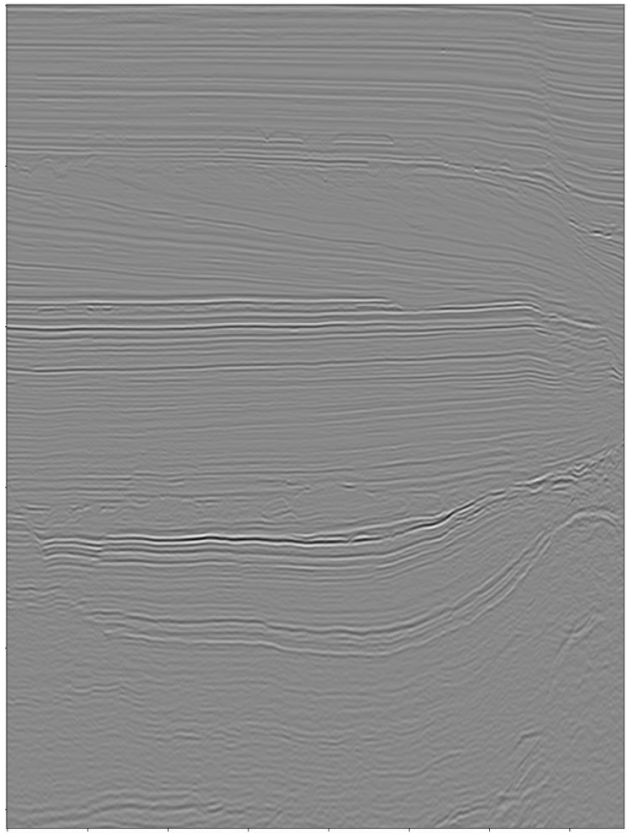
UNET BENCHMARK FOR TRAINED PREDICTIONS

Actual Lithology

Prediction



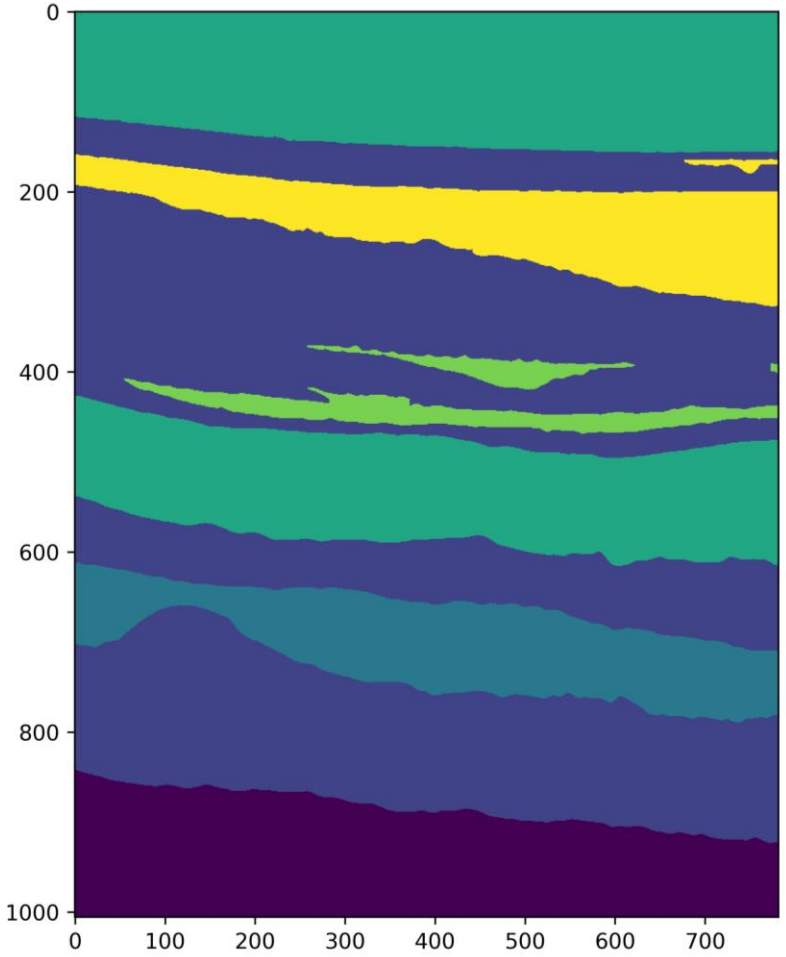
Holdout Seismic Data



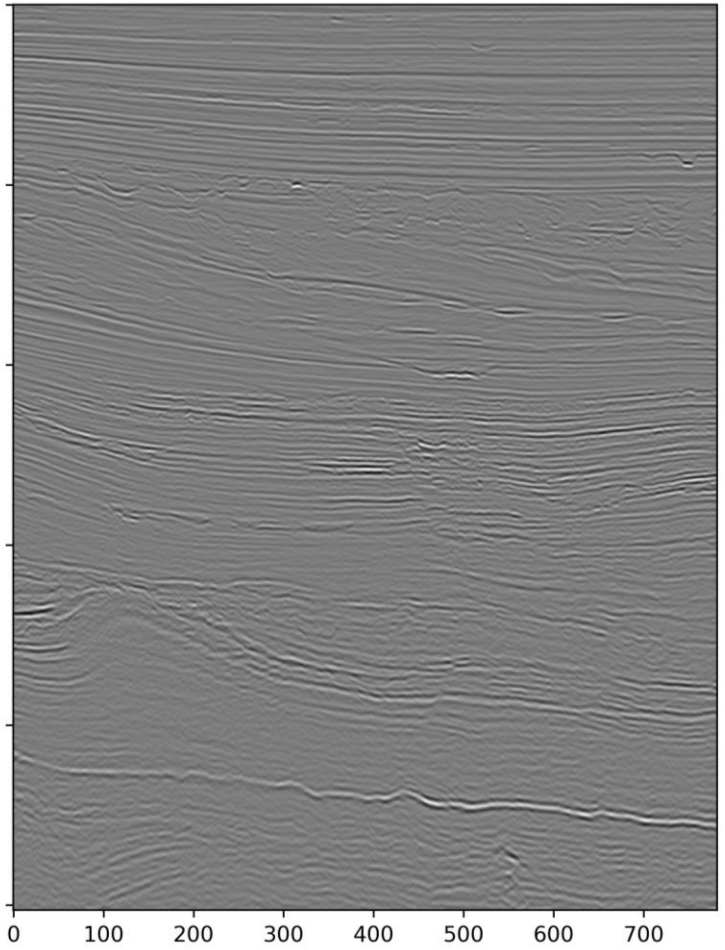
Lithology Prediction

DECISION-TREE PREDICTION: HOLDOUT DATA

Actual Label



Holdout Seismic Data

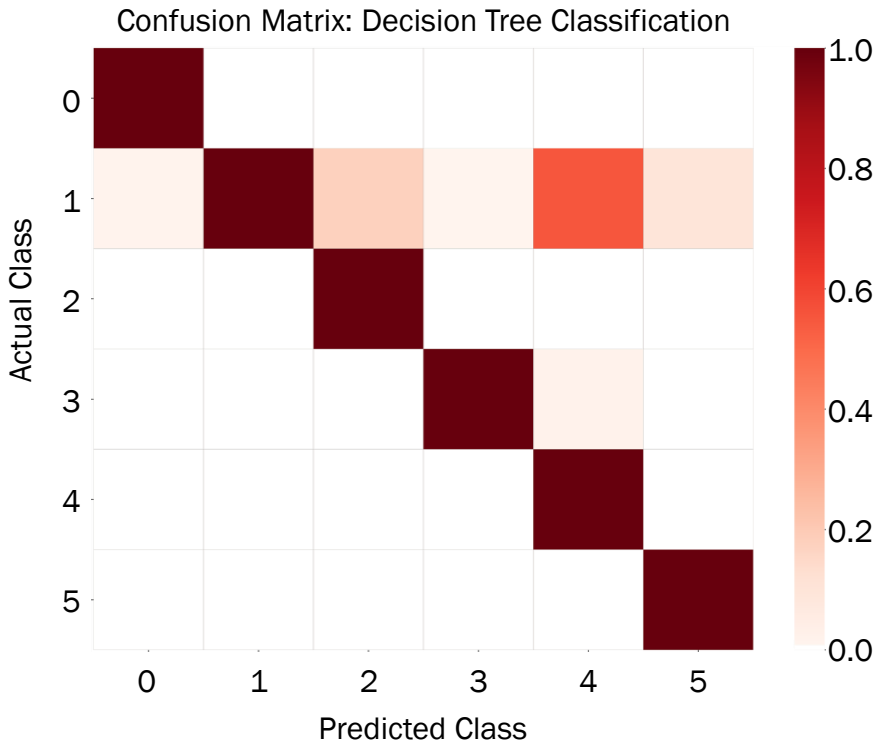


Prediction on Holdout Data

SCORING METRICS

- High Recall and F1 scores for a DecisionTree improve the overall quality of our seismic predictions, with plenty of room for improvement
- The large number of samples from Class 1 are classifying elsewhere and impact the overall interpretation

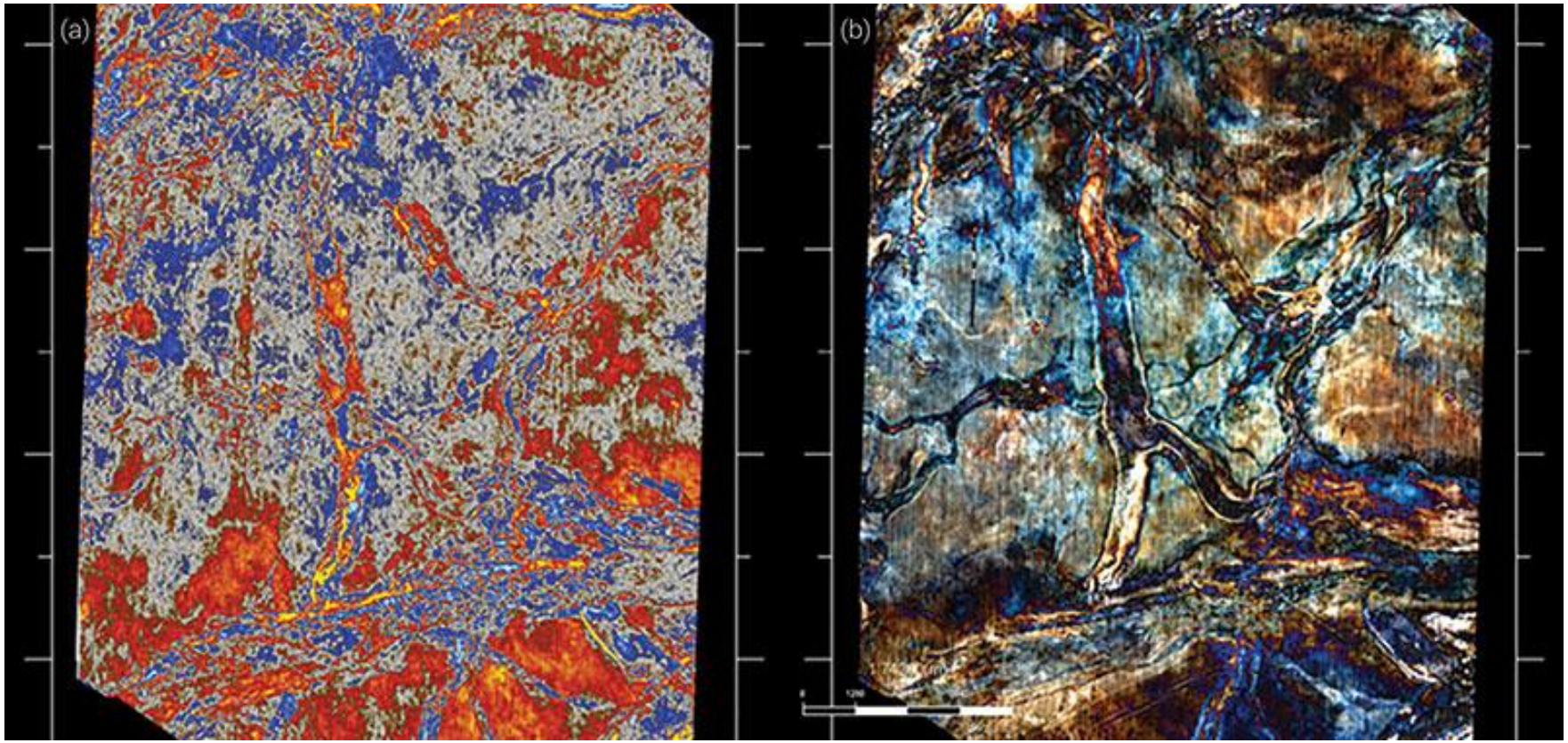
	0	1	2	3	4	5	Total
Precision	0.99	0.69	0.85	0.99	0.63	0.91	**0.97**
Recall	1.00	0.72	0.99	0.99	0.99	0.99	**0.97**
F-1	0.99	0.96	0.91	0.99	0.77	0.95	**0.97**



VALUE IN FREQUENCY (SPECTRAL) DECOMPOSITION

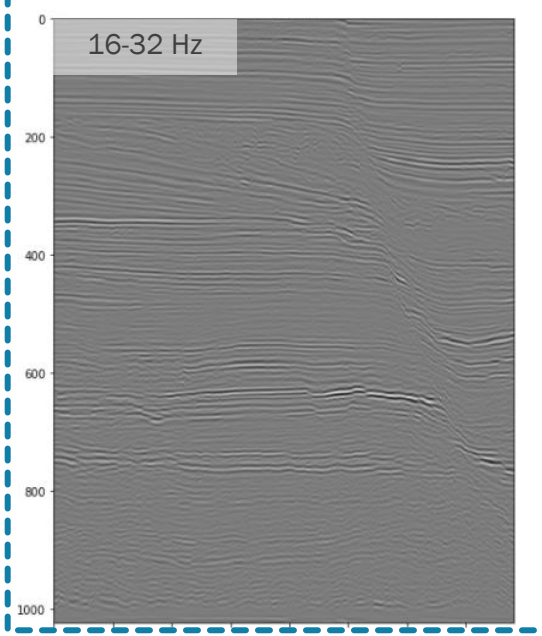
“Blended” Seismic Data

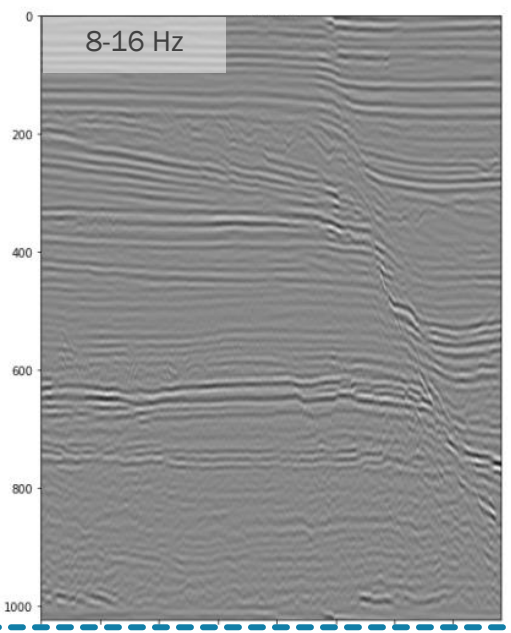
“Component” Frequencies

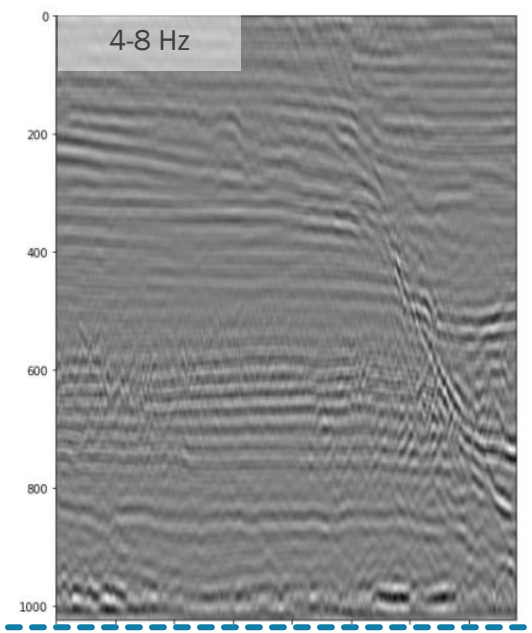


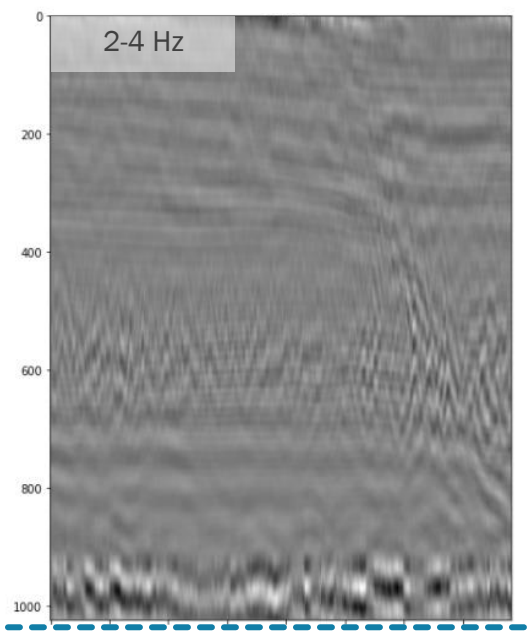
Zhang (2014): https://explorer.aapg.org/story/articleid/3184

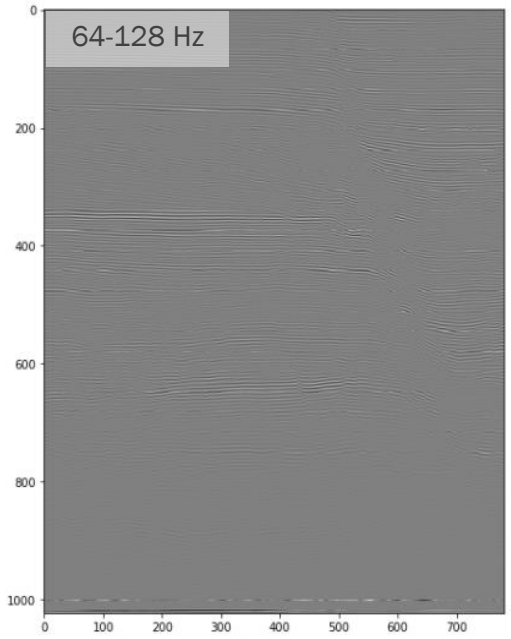
- **Spectral** (frequency-based) features can associate large-scale and small-scale features to the target classes

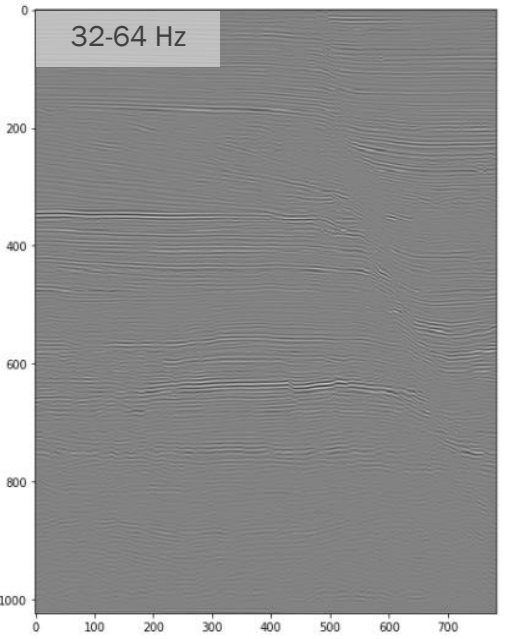




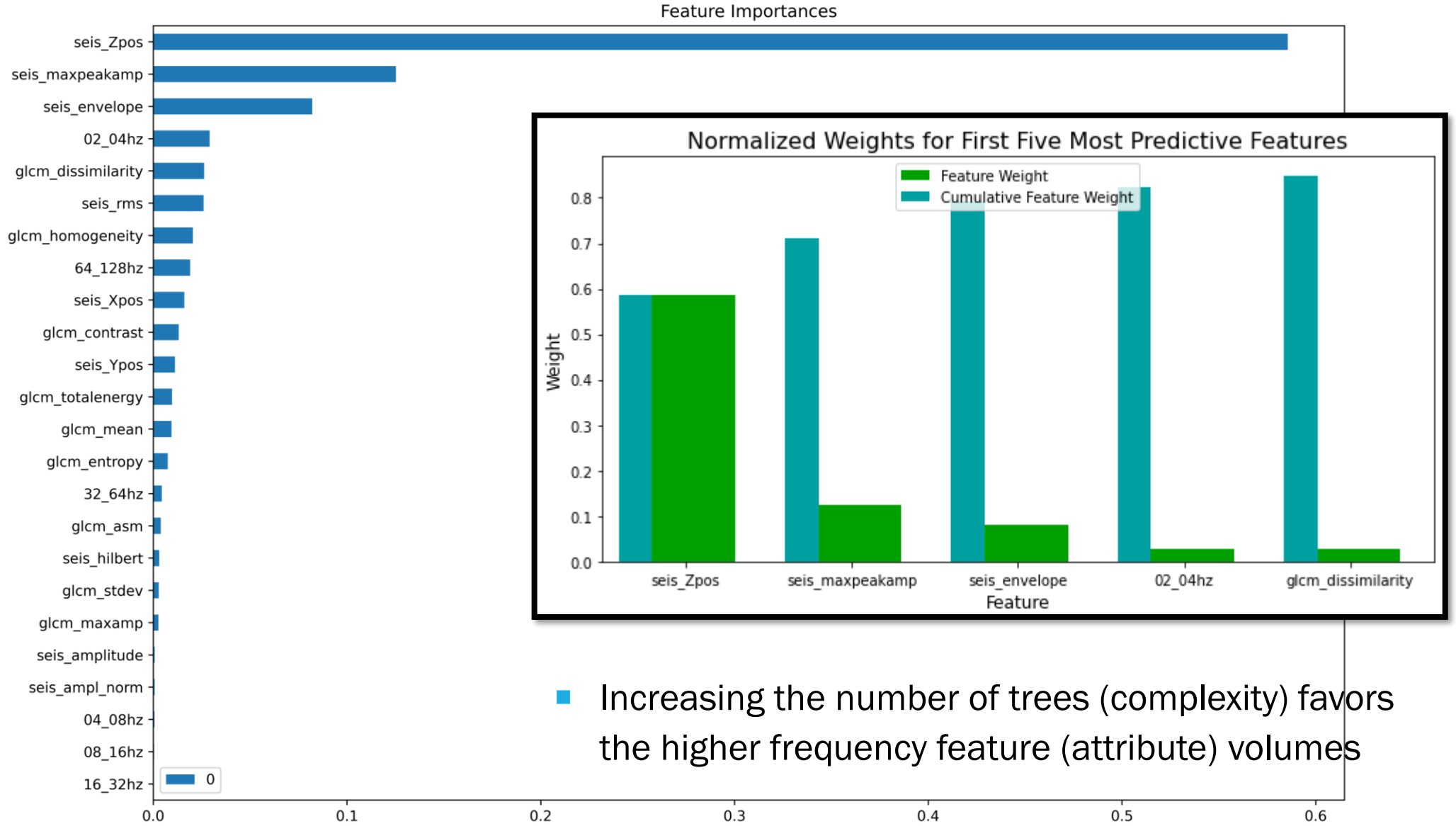








FEATURE IMPORTANCE: VALUE IN NEURAL NET AUGMENTATIONS



- Increasing the number of trees (complexity) favors the higher frequency feature (attribute) volumes

KEY INSIGHTS

- Locality (absolute position) is critical as lithologies tend to be “near” to one another
  - CNNs can naturally encode these positions in images, but we can do explicitly for non-NN models
  - Hypothesis confirmed in a subsequent literature search: https://arxiv.org/pdf/2001.08248.pdf

- Using a DecisionTree model with feature engineering provides an intermediary solution for real-time workflows, before considering more sophisticated neural nets

- Tree-based method feature importance may be biased by higher frequency features
  - i.e., feature importance has to be understood in the context of domain-knowledge

THANK YOU

Bao D. Nguyen

Email: bao.d.nguyen83@gmail.com

GitHub: https://github.com/bdnguyen-ds

LinkedIn: https://www.linkedin.com/in/bao-d-nguyen/