

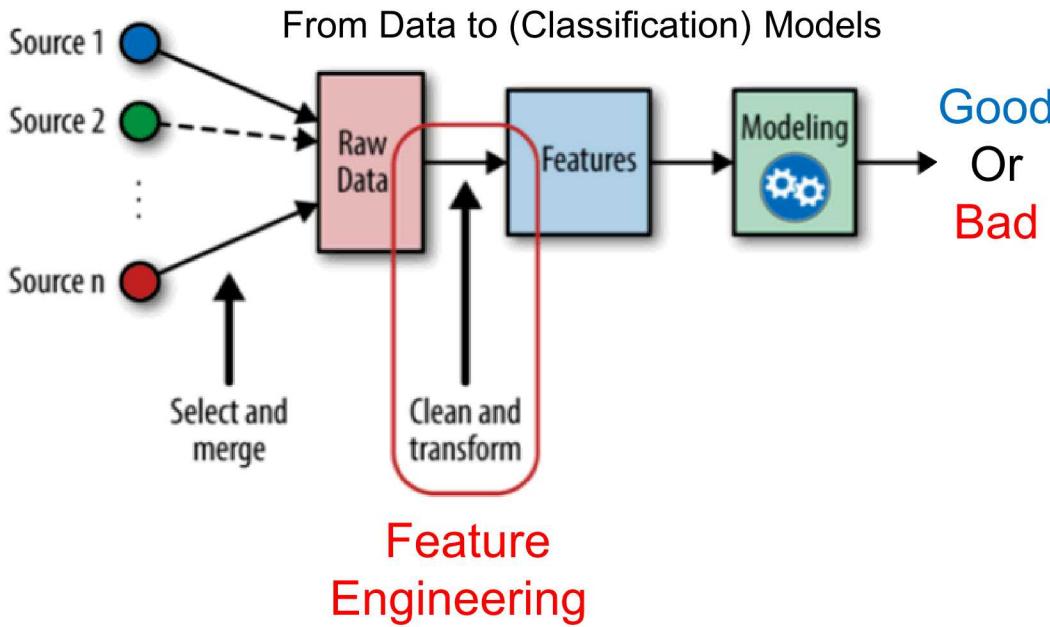
# Data Engineering for Deep Learning

Tim Draelos / 6362

SAND Number: SAND2019-?

# The Past

- Feature Engineering
  - Measure specific attributes in data
  - Still do when it's easy or necessary



## List of common wood features:

- Large knot (>30 mm)
- Medium knot (>20 mm)
- Small knot (>10 mm)
- Pin knot (< 30 mm)
- Hole
- Pith
- Bark inclusions
- Resin pocket
- Cracks
- Splits
- Wane
- Hit and miss
- Bow
- Spring
- Cupping
- Bluestain

# The Present

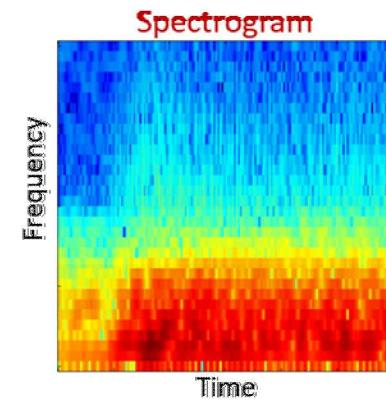
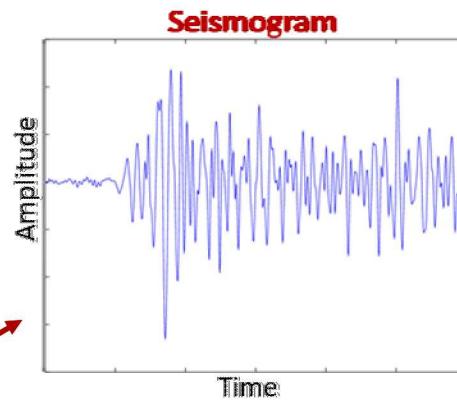
- Data Engineering
  - Deep Neural Networks will learn patterns in the data
    - Local Spatial Structure with Convolutional Neural Networks
    - Temporal patterns with Recurrent Neural Networks
    - Local Spatial-Temporal patterns with Recurrent Convolutional Networks
- Data Issues
  - Class Imbalances
  - Artifacts
  - Few Samples
    - Overfitting
  - Mixing Training/Validation/Test Data
    - Training Error – Capability feedback
    - Validation Error – Sign post of progress
    - Test Error – Estimate of ability to generalize



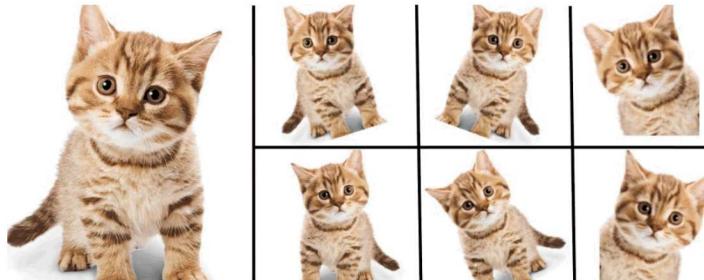
# First Things First

## ■ Do Good Data Engineering

- Data Collection
- Data Conditioning, Pre-processing
- Management of Training/Validation/Test Data
- Data Augmentation



## Variation



## Simulation

DELTAGEN



Honda R&D Americas, Inc. May 2014

**HONDA**  
The Power of Dreams

Simulation Postprocessor

**HONDA**  
The Power of Dreams

