2024 QCAM Machine Learning Workshop

Binder Session

- Go to: https://github.com/ucgmsim/qc_ml_workshop_2024

- Click Binder Icon -> Wait for session to start

Workshop Plan

- Introduction (~5 minutes)
- Jupyter Notebook (~1.5 hours)
- Break (~20-30 minutes)
- Hands-on (~1 hour)

- Problem Setup
- Data Preparation
- Model Fitting
- Model Evaluation

Why Machine Learning?

- Ability to develop complex models from data
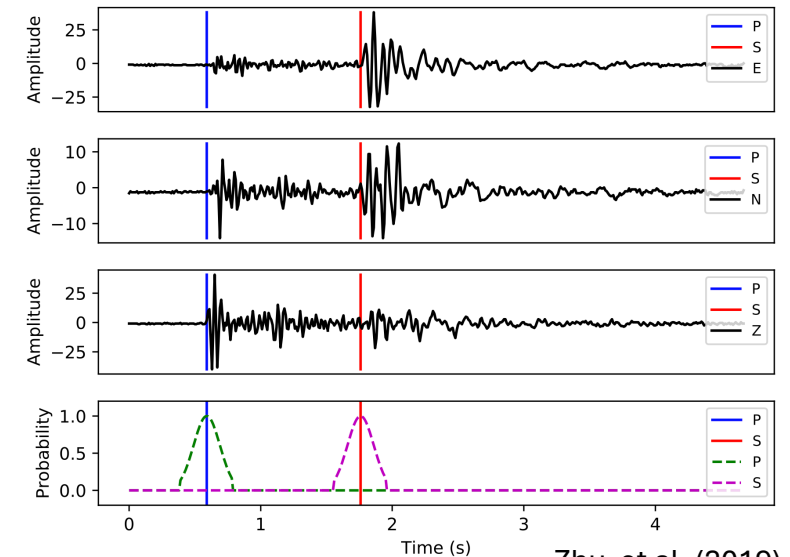 - Potential to improve on traditional methods
  - Allows for automation of tasks
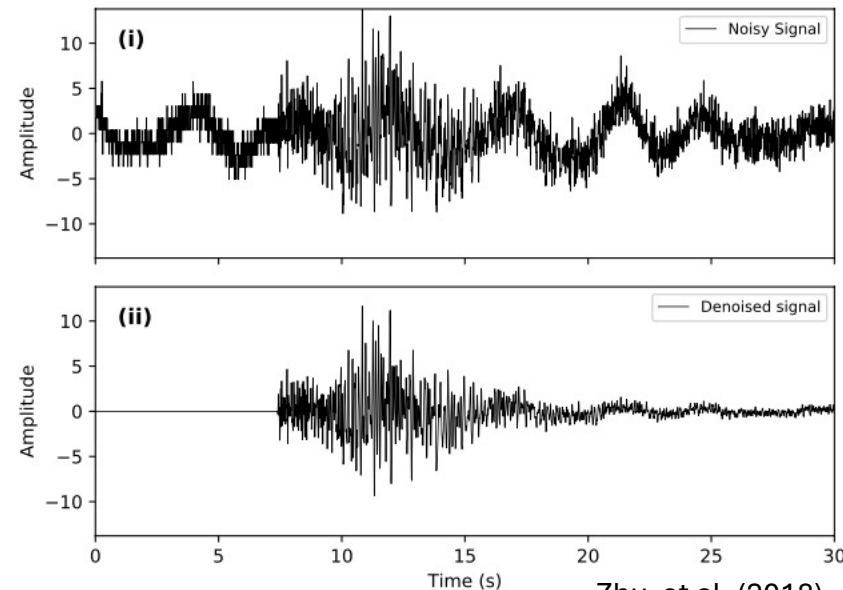 - Applicable to new complex problems

Surface Crack Detection

P- and S-wave picking
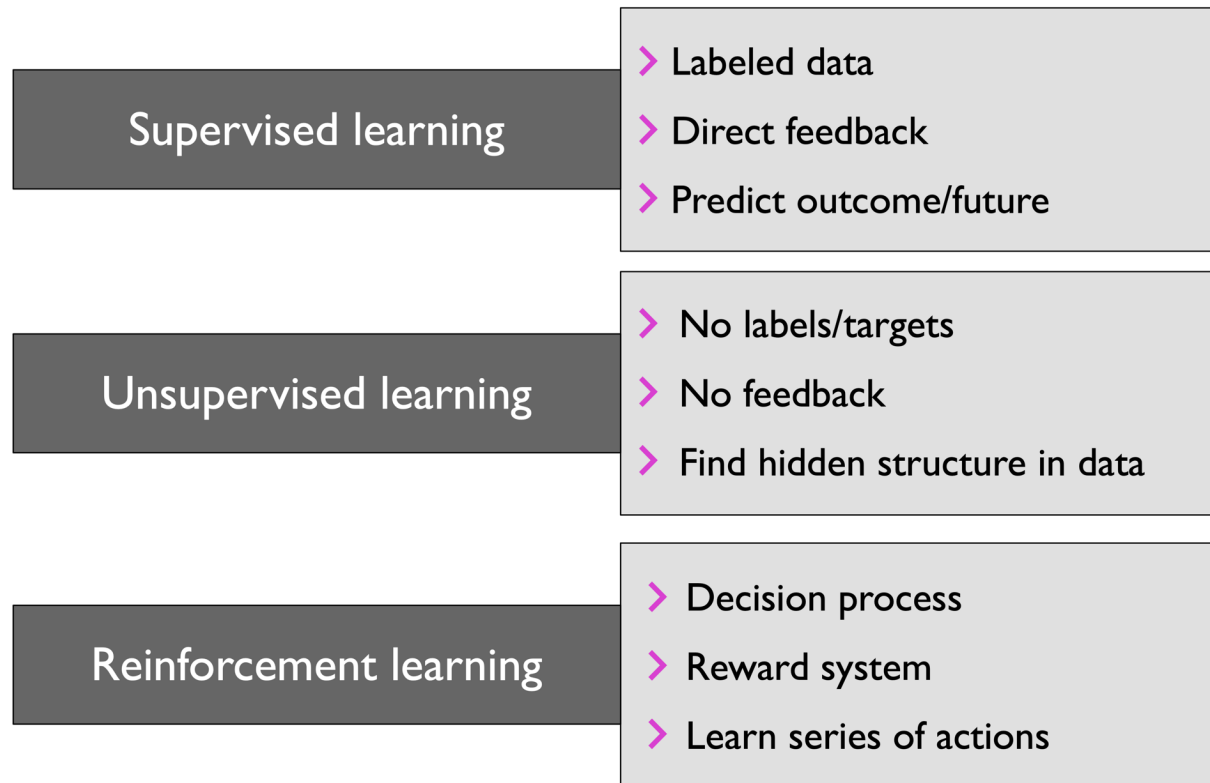

Zhu et al. (2019)

Denoising

Zhu et al. (2018)

Limitations of ML

- Require large amounts of high-quality data (garbage-in, garbage-out)
- Often difficult to interpret
- Models learn biases present in the training data
- Generally, models only interpolate

Overview – Types of ML

Supervised learning
- Labeled data
- Direct feedback
- Predict outcome/future

Unsupervised learning
- No labels/targets
- No feedback
- Find hidden structure in data

Reinforcement learning
- Decision process
- Reward system
- Learn series of actions

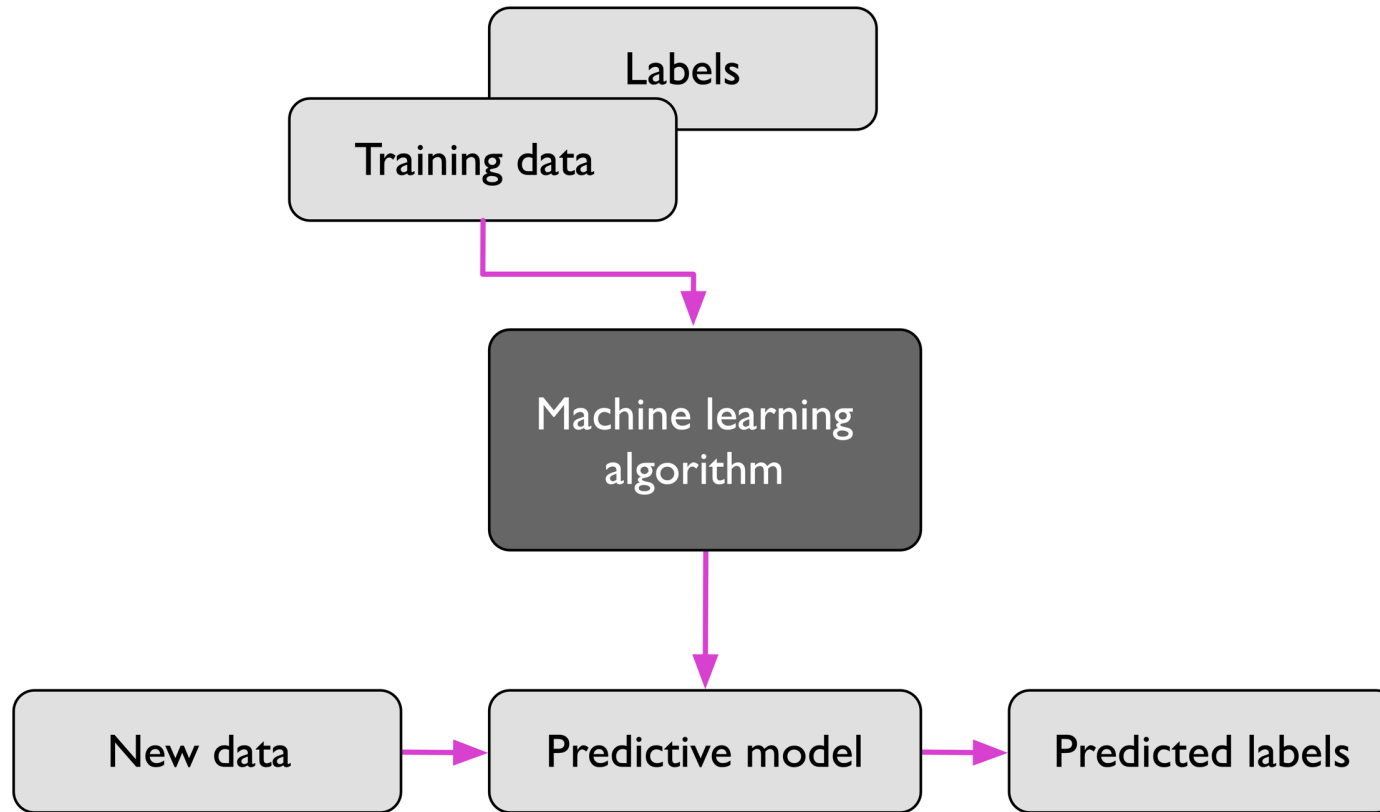Figure: Raschka (2022)

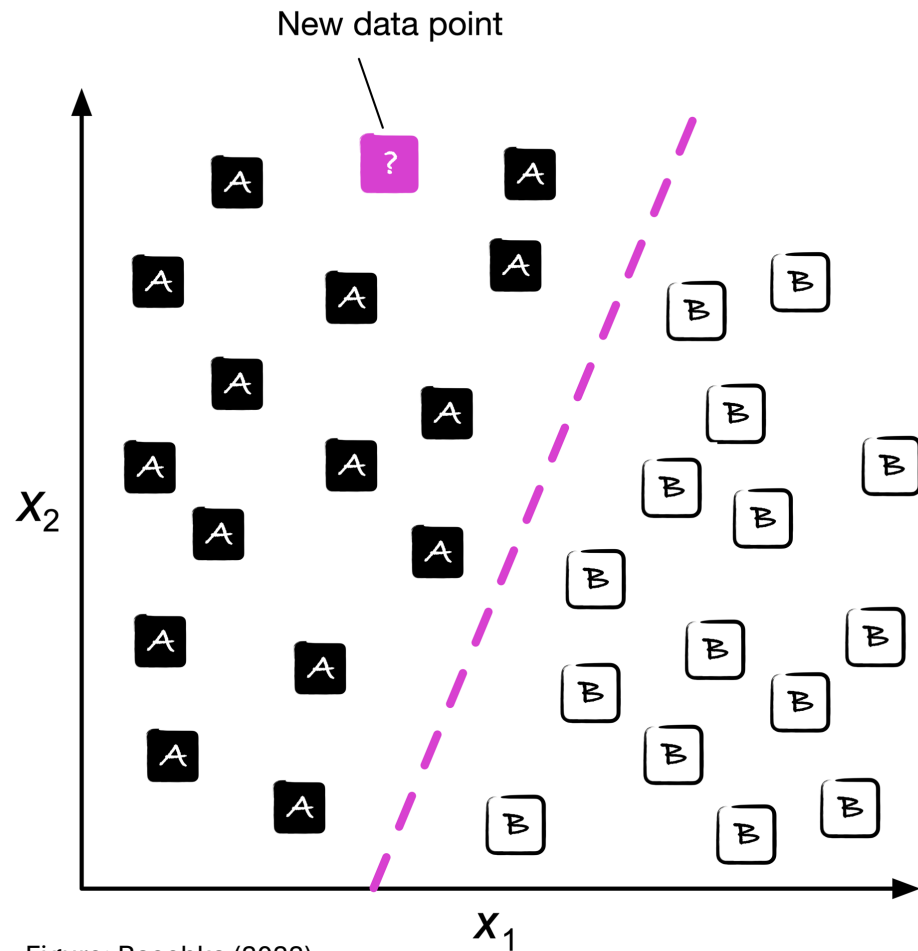# Supervised ML Process

Figure: Raschka (2022)

Classification vs Regression

Figure: Raschka (2022)

Figure: Raschka (2022)