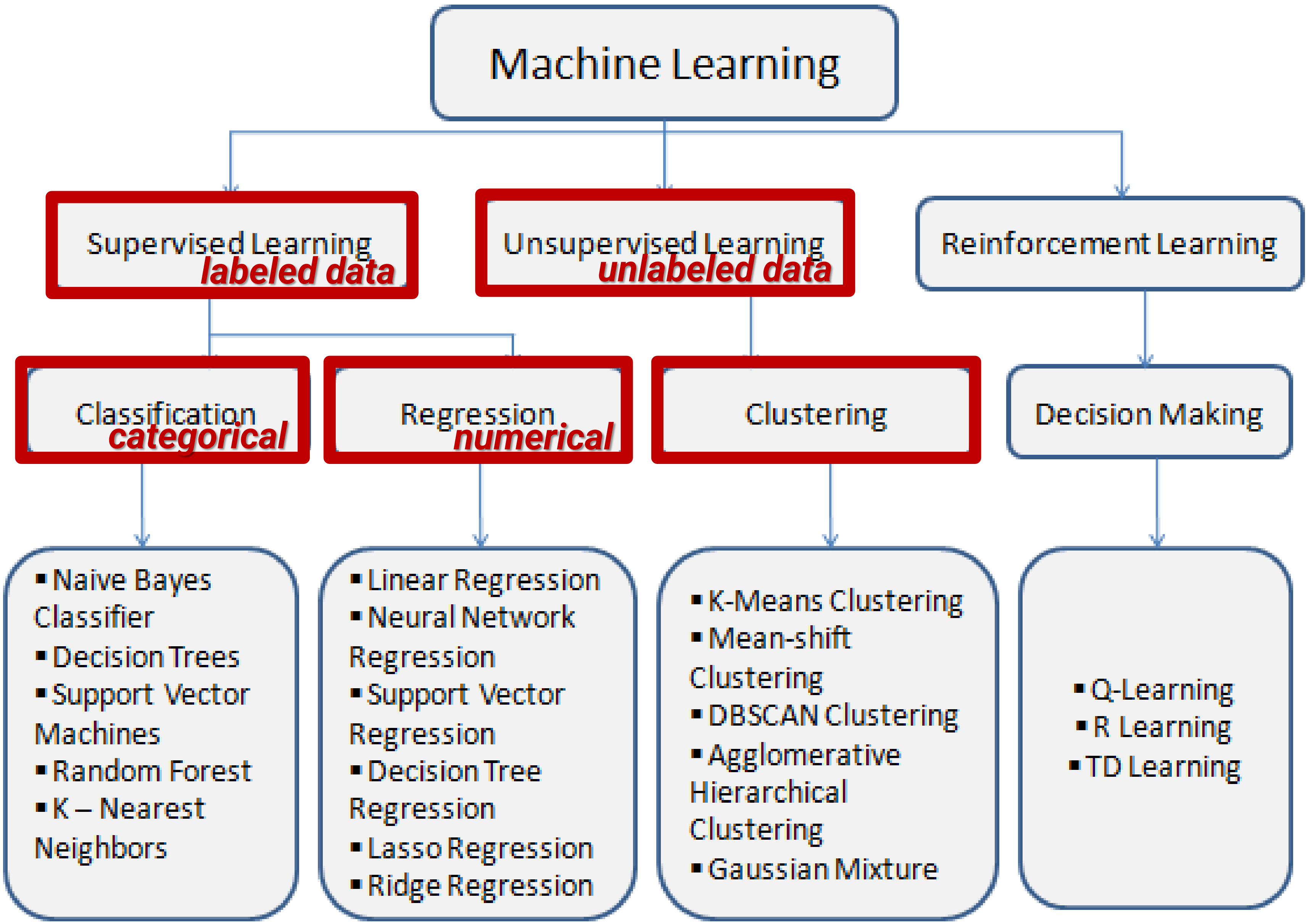
Introduction to AI within Watershed Management

Isabella Bertani, Qian Zhang, Mike Evans

Modeling Workgroup Quarterly Review
10/17/2023

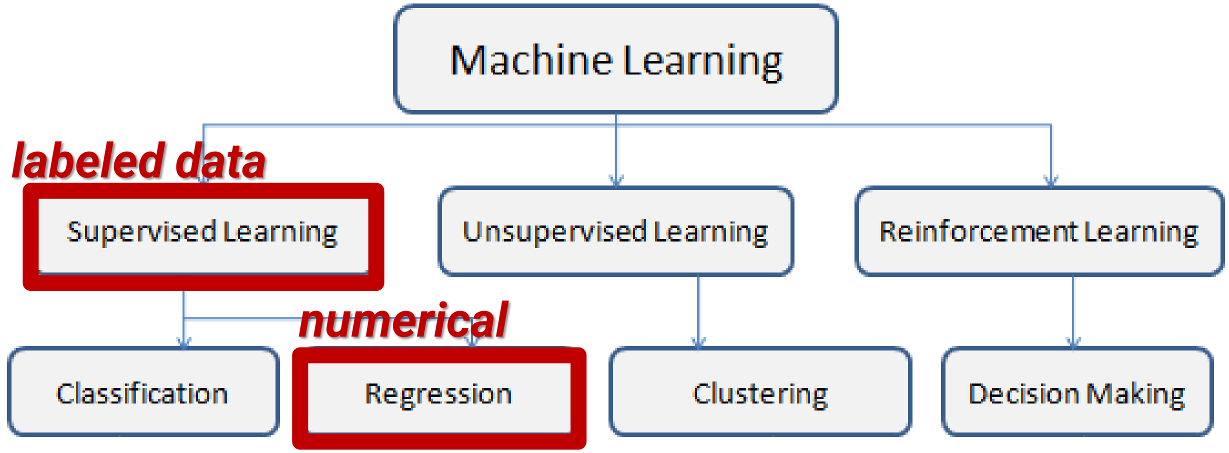
AI / ML / DL Examples for Chesapeake

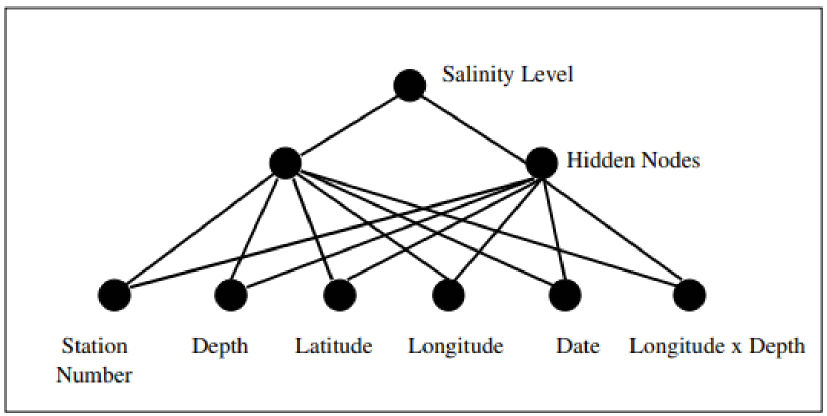
- I will provide several examples specific to Chesapeake Bay. This is not an exhaustive survey of the literature.
- I will show some ML and DL methods that frequently show up in the literature.
- I will use the classification diagram as a roadmap.
- I am not an AI/ML/DL expert by training but have been learning & working with these approaches for 10+ years.



Machine Learning
Supervised Learning
labeled data
Unsupervised Learning
unlabeled data
Reinforcement Learning
Classification
categorical
Regression
numerical
Clustering
Decision Making
Naive Bayes Classifier
Decision Trees
Support Vector Machines
Random Forest
K – Nearest Neighbors
Linear Regression
Neural Network Regression
Support Vector Regression
Decision Tree Regression
Lasso Regression
Ridge Regression
K-Means Clustering
Mean-shift Clustering
DBSCAN Clustering
Agglomerative Hierarchical Clustering
Gaussian Mixture
Q-Learning
R Learning
TD Learning

DeSilet et al., 1992, Predicting salinity in the Chesapeake bay using backpropagation. Computers & Operations Research.

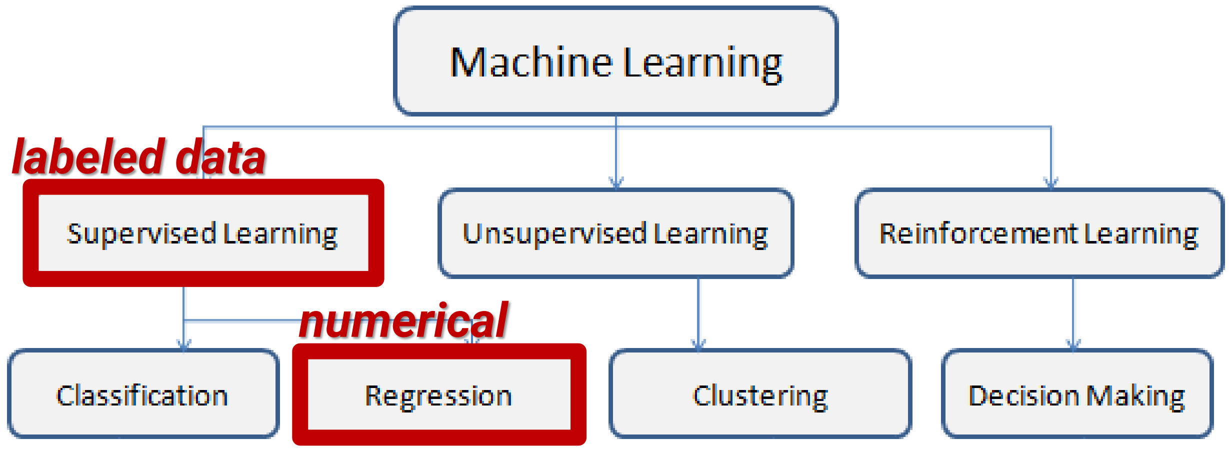


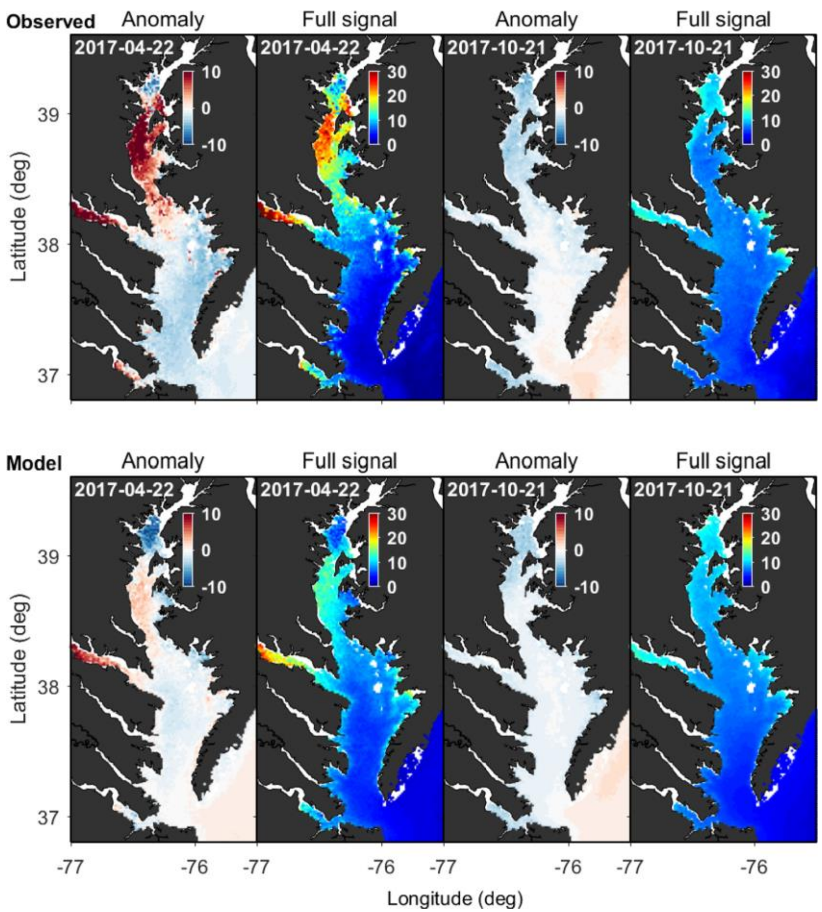


	Range of Average Percent Absolute Errors	
	10 old data sets	10 new data sets
Regression	9.60 – 16.46	9.19 – 20.15
Neural Network	9.54 – 16.18	7.70 – 19.37

- Managing an aquatic ecosystem requires frequent monitoring of salinity levels.
- Using nearly 40,000 observations from 34 stations in the Chesapeake Bay, the authors built and compared regression and neural network models.
- In general, the neural network models predict salinity value better than the corresponding regression models.
- However, a major advantage of the regression models is that they are easily explained.

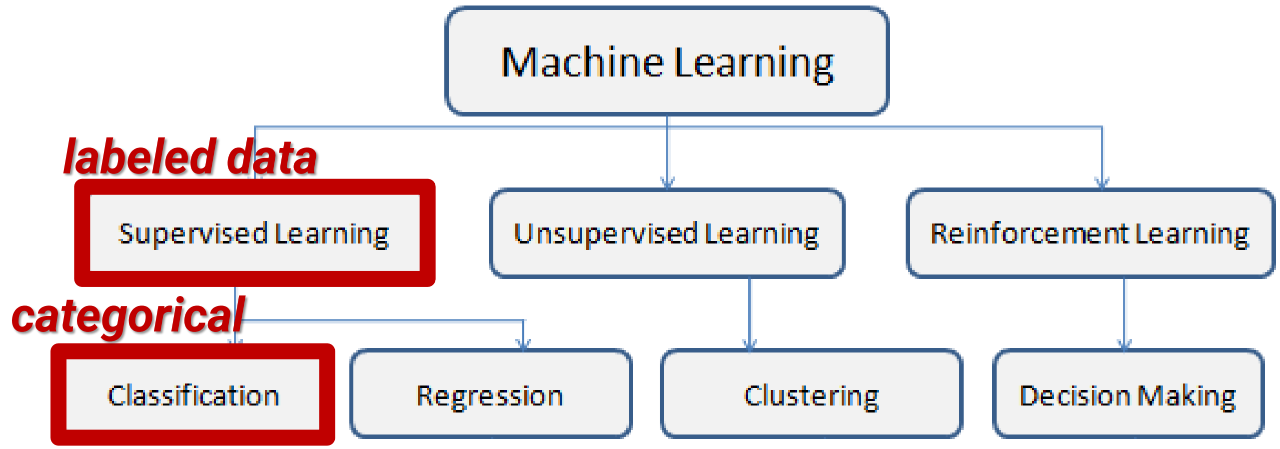
Yu et al., 2022, Chlorophyll-a in Chesapeake Bay based on VIIRS satellite data: Spatiotemporal variability and prediction with machine learning. Ocean Modelling.

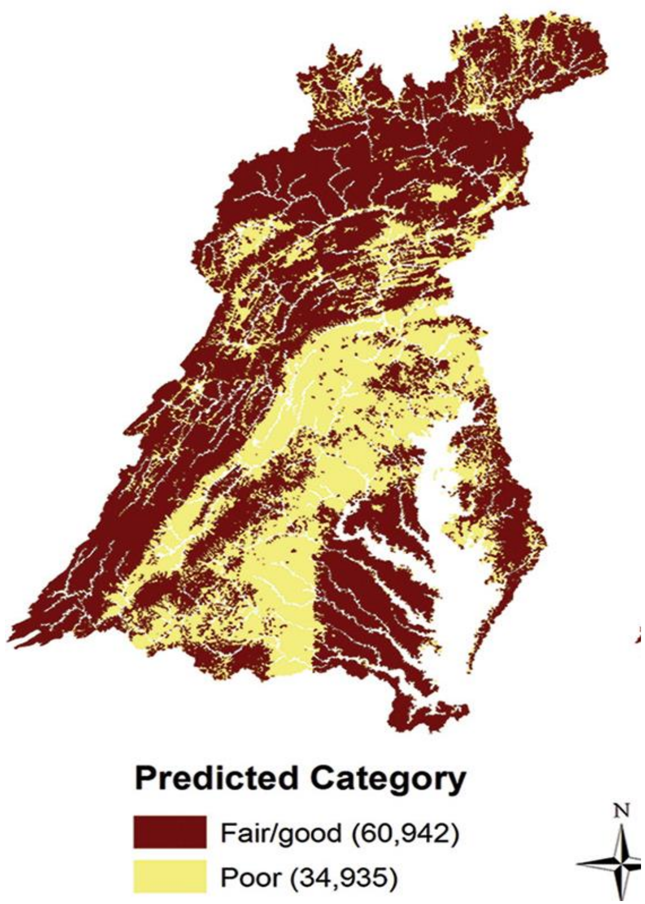




- Chl-a is a key WQ parameter. Satellite data can provide a better spatial & temporal coverage than conventional sampling.
- The authors trained a machine-learning, data-driven model (involving artificial neural network) to simulate high-resolution Chl-a variations in the Bay.
- External forcing included Q, nutrient loads, solar radiation, wind, and air temperature.
- The model shows an overall satisfactory performance in predicting Chl-a (bay-wide RMSE = 1.85 ug/l), highlighting the potential of using data-driven models for high-resolution water quality simulations.

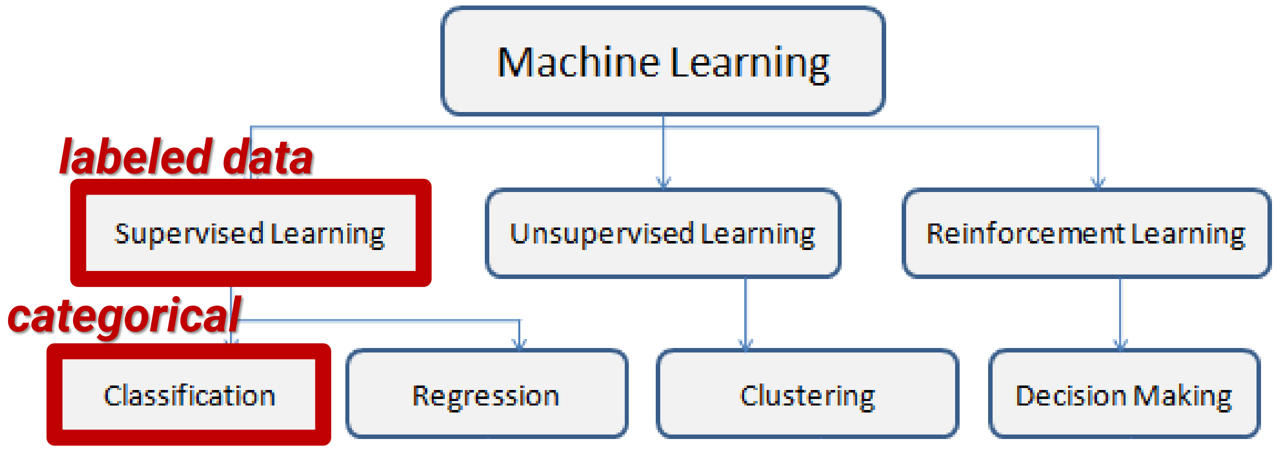
Maloney et al., 2018, Predicting biological conditions for small headwater streams in the Chesapeake Bay watershed, Freshwater Science.

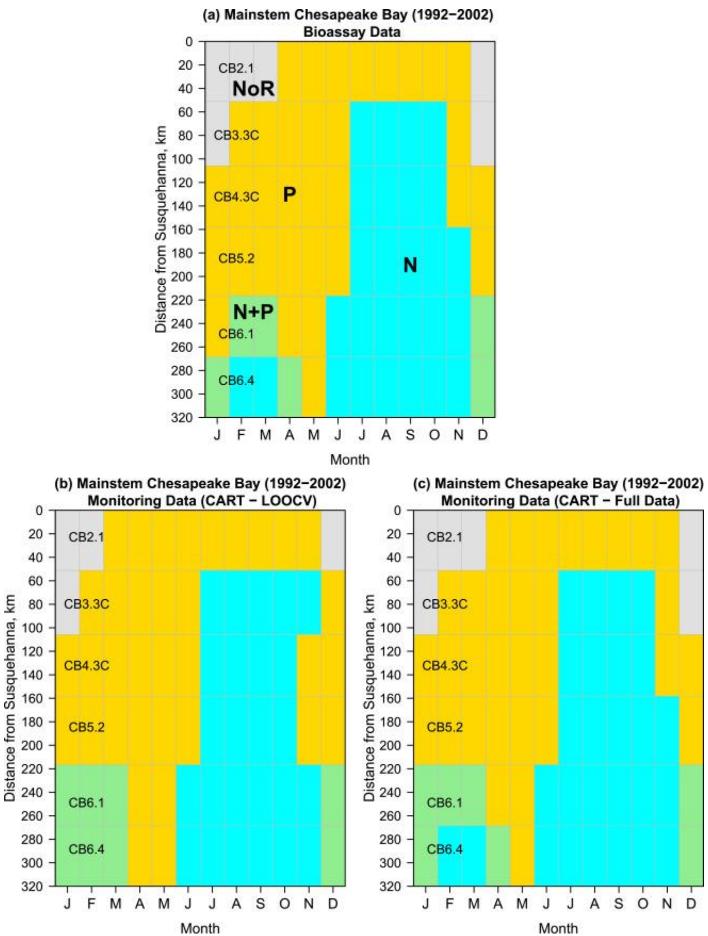




- A primary goal for the Bay watershed restoration is to improve stream health and function in 10% of stream miles by 2025.
- Biological condition was measured with the Chesapeake Bay Basin-wide Index of Biotic Integrity (Chessie BIBI), which was classified as either “poor” or “fair/good”.
- The authors developed random forest model to predict the index for small streams (<200 km2), using 12 geospatial variables (spatial location, bioregion, land cover, soil, precipitation, number of dams).
- Model predictions showed fair/good for 64% of the 95,877 small stream reaches.

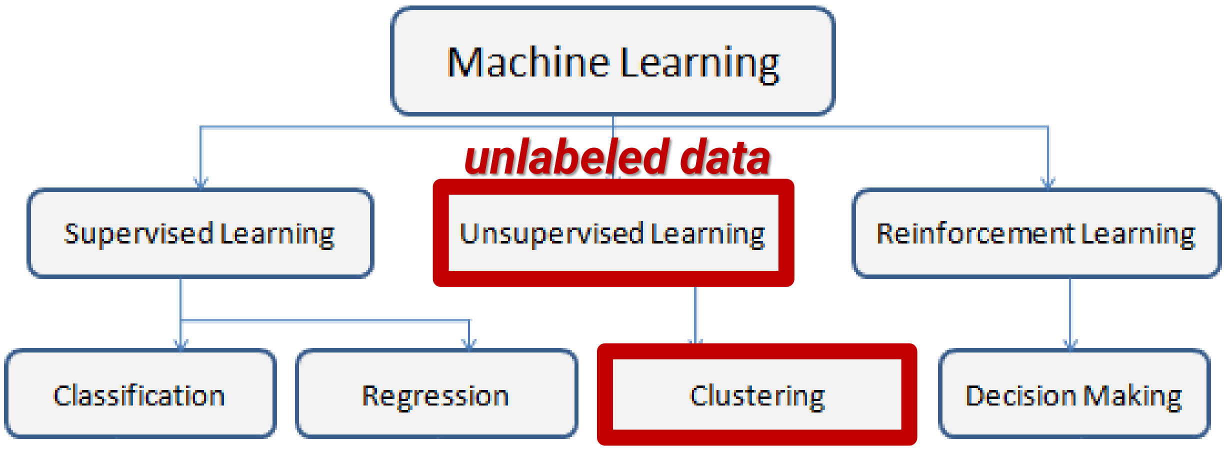
Zhang et al., 2021,
Nutrient limitation of phytoplankton in Chesapeake Bay: Development of an empirical approach for water-quality management, Water Research.

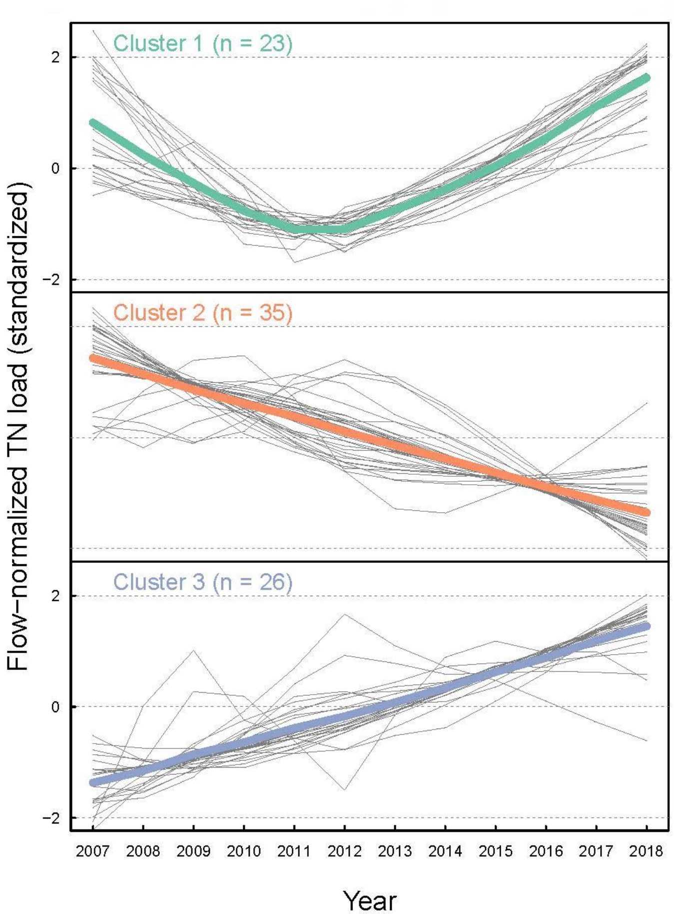




- Chesapeake Bay has well-documented seasonal and spatial variations in nutrient limitation, but it remains unknown whether these patterns have changed in response to nutrient management efforts.
- The authors analyzed historical data from nutrient bioassay experiments (1992–2002) and data from long-term, fixed-site WQ monitoring program (1990–2017).
- CART models satisfactorily reproduced bioassay-based nutrient limitation.
- CART predictions showed more space of N-limitation in 2007–2017 than 1992–2002, consistent with long-term N reduction.

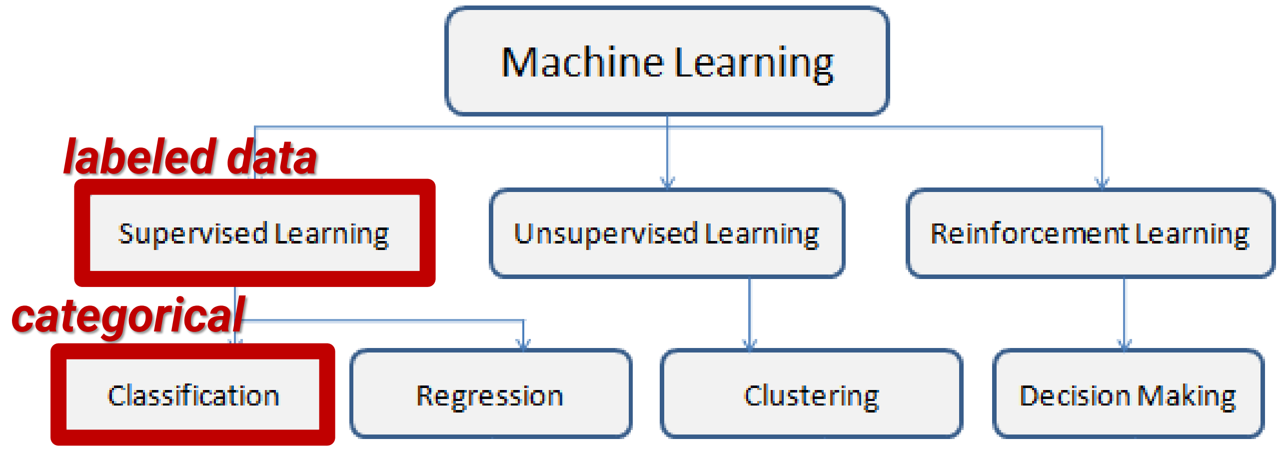
Zhang et al., 2022,
Regional patterns and drivers of total nitrogen trends in the Chesapeake Bay watershed: Insights from machine learning approaches and management implications, Water Research.

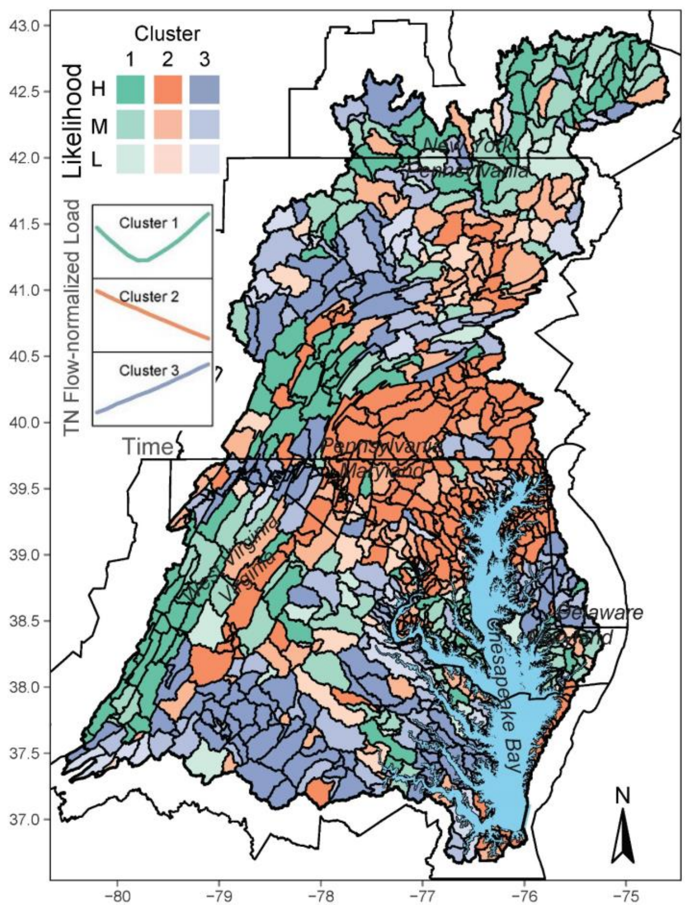




- The CBNTN stations (84) showed diverse trajectories of total nitrogen (TN) trends.
- Clustering methods are especially useful for categorizing regional WQ patterns.
- Hierarchical clustering identified 3 distinct clusters for short-term TN trends (2007-2018): V-shape (n = 23 stations), decline (n = 35), and increase (n = 26).

Zhang et al., 2022, Regional patterns and drivers of total nitrogen trends in the Chesapeake Bay watershed: Insights from machine learning approaches and management implications, Water Research.





- The CBNTN stations (84) showed diverse trajectories of total nitrogen (TN) trends.
- Clustering methods are especially useful for categorizing regional WQ patterns.
- Hierarchical clustering identified 3 distinct clusters for short-term TN trends (2007-2018): V-shape (n = 23 stations), decline (n = 35), and increase (n = 26).
- Random forest classification models were developed to predict TN trend clusters (i.e., 1, 2, or 3) for the entire Bay watershed, including unmonitored areas.