
BAGGING

BAGGING

- What is bagging?
- Benefit of bagging
- How bagging works
- Discussion about bagging
- Bagging with supervised learning
- Bagging with unsupervised learning
- Bagging with reinforcement learning
- Code

WHAT IS BAGGING?

Bagging, short for Bootstrap Aggregating, is a machine learning ensemble technique used to improve the stability and accuracy of algorithms. It works by training multiple versions of a model on different subsets of the data, which are created through bootstrapping (sampling with replacement).

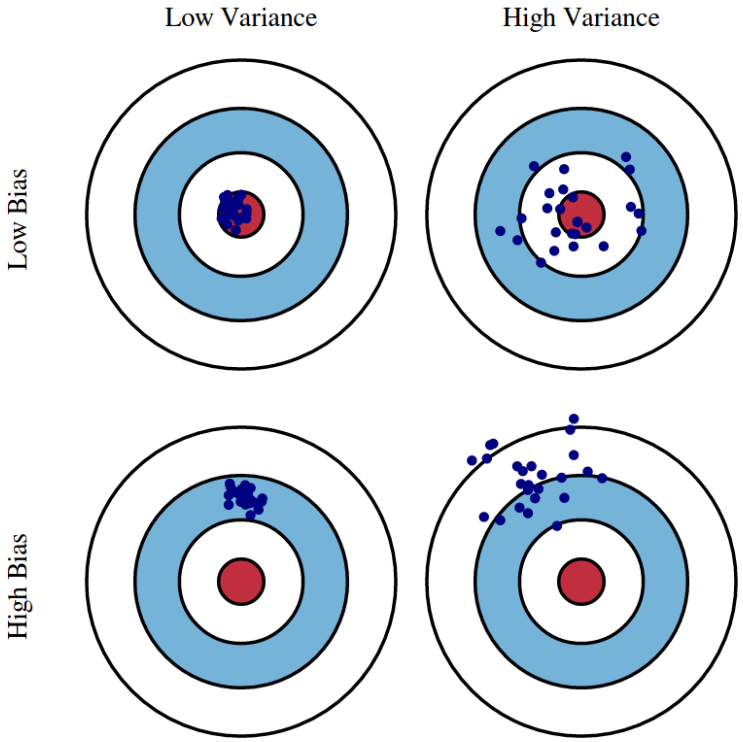
The final prediction is made by averaging (for regression) or voting (for classification) the predictions of all models.

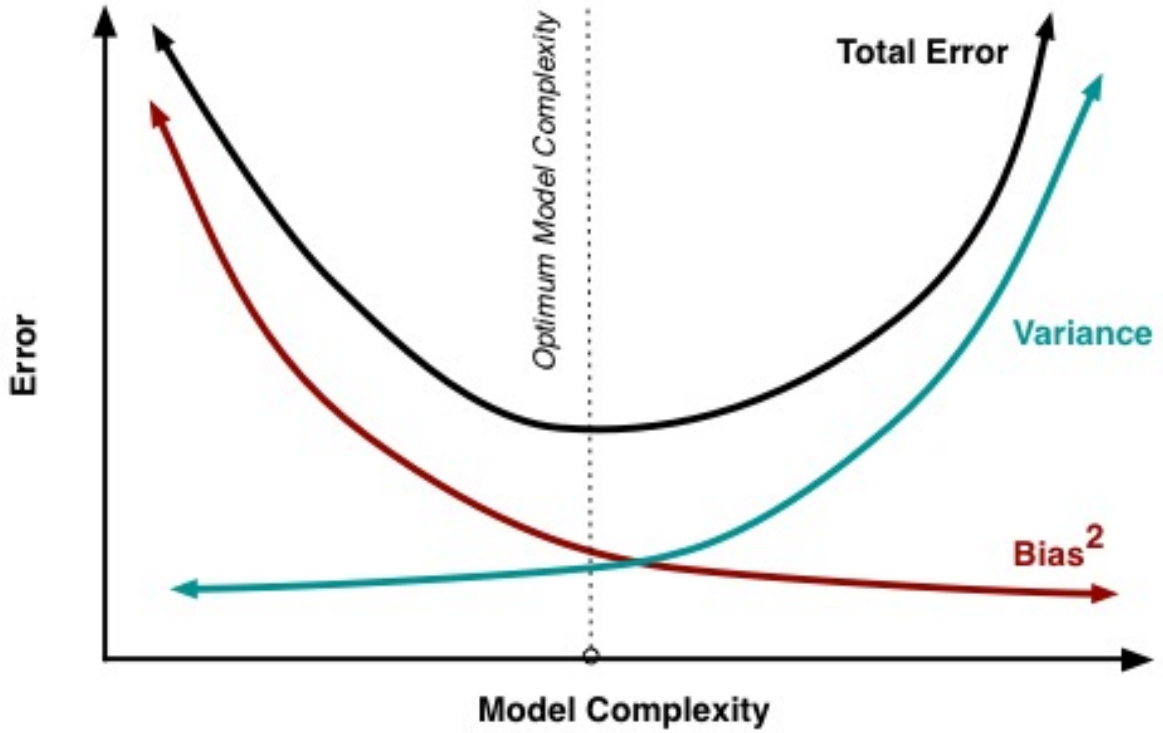
This method reduces variance and helps prevent overfitting.

BENEFIT OF BAGGING

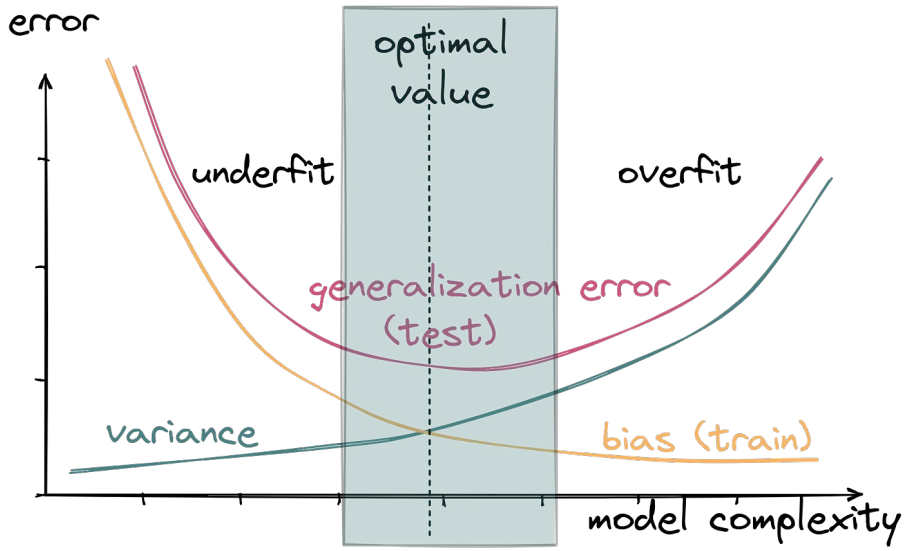
	Benefit of bagging
Supervised learning	- Reduce variance - Improve generalization - Robustness to overfitting
Unsupervised learning	- Enhance clustering stability - Noise reduction - Create diverse SL model
Reiinforcement learning	- Policy stability

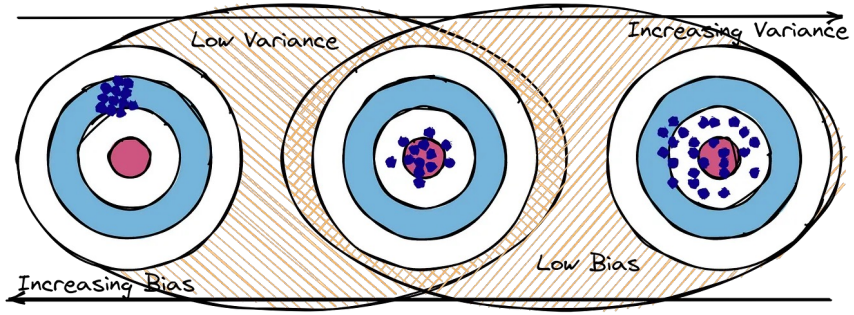
BIAS-VARIANCE TRADEOFF



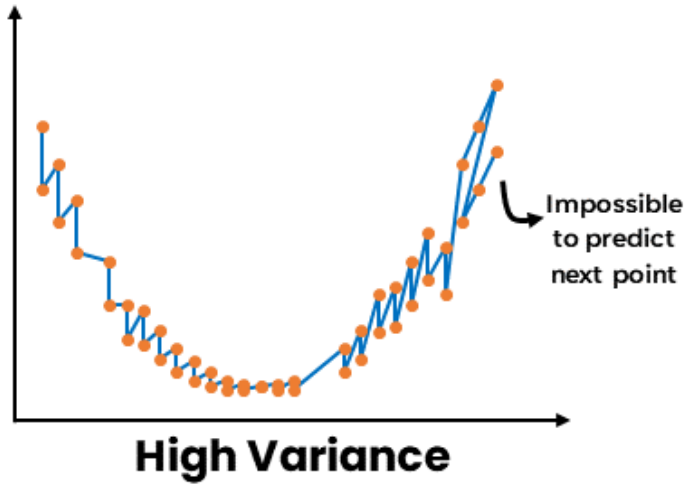


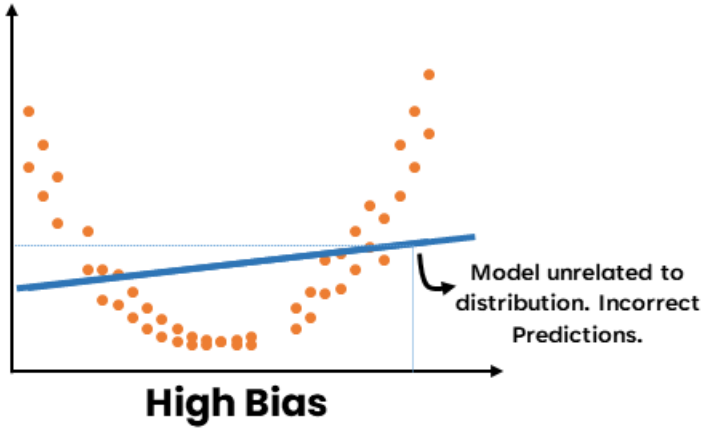
BIAS-VARIANCE TRADEOFF

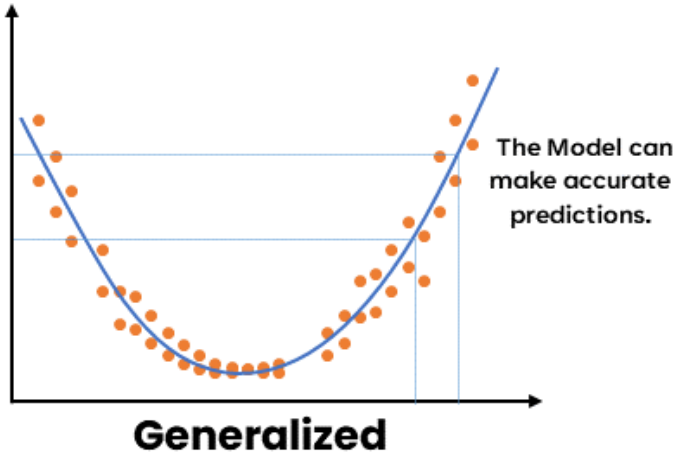




BIAS-VARIANCE TRADEOFF







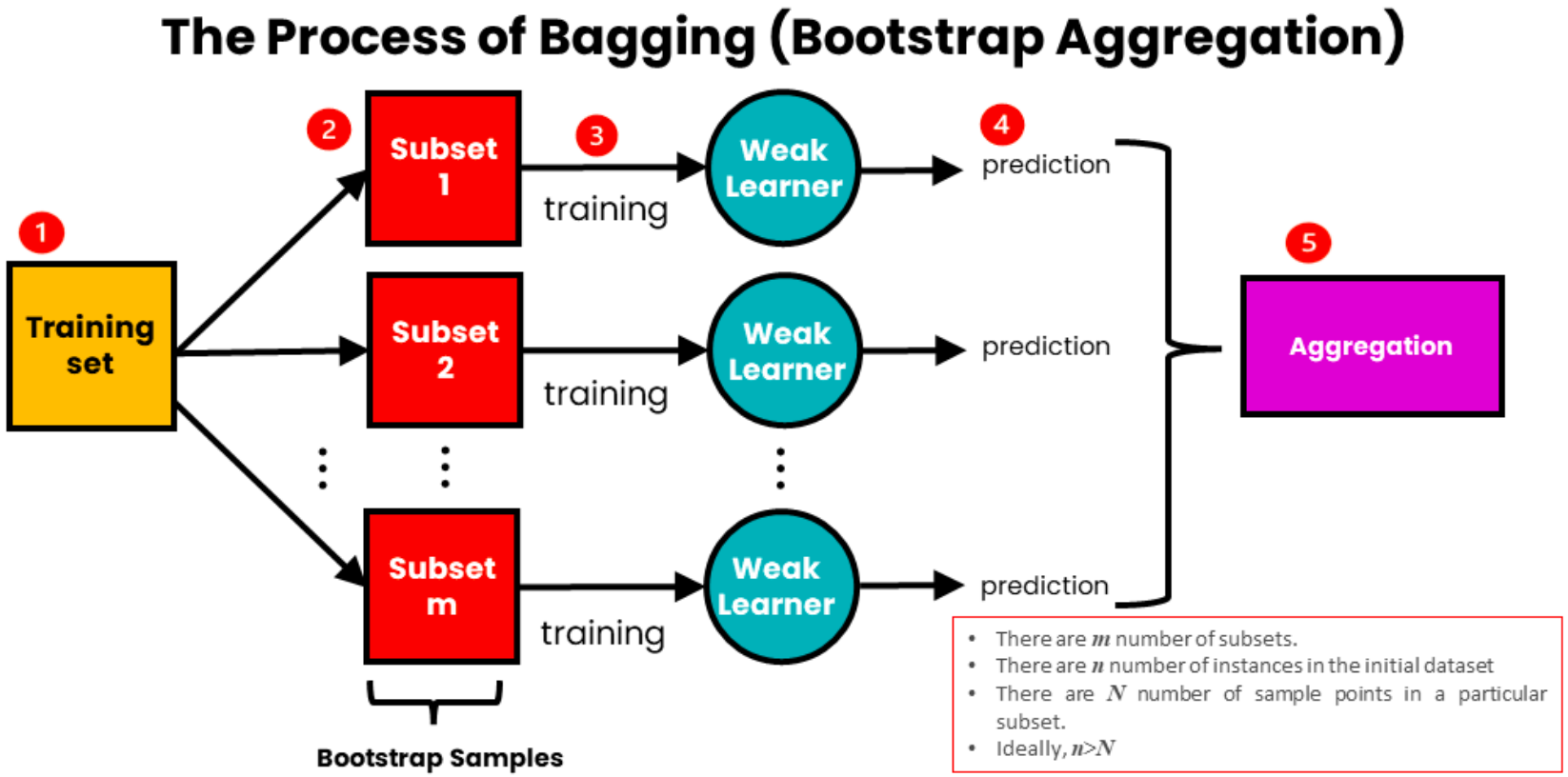
HOW BAGGING WORKS

Boostrapping

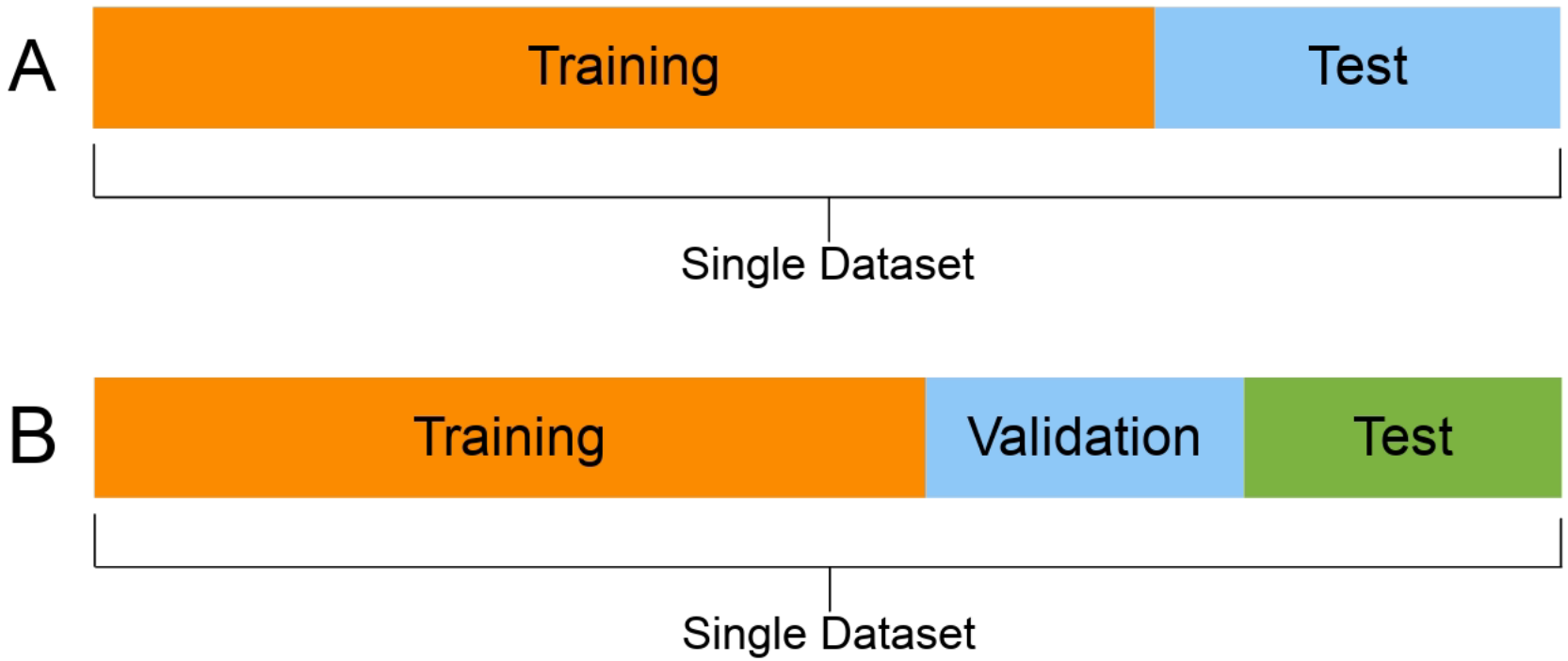
Training multiple models

Aggregating predictions

HOW BAGGING WORKS



HOW BAGGING WORKS



DISCUSSION ABOUT BAGGING

- What boostrapping?
- Why bagging reduce variance?
- Why bagging is robust to overfitting?
- Important condition for base models
- When to say that our base models have low correlation?

DISCUSSION ABOUT BAGGING

- Suitable n_models for bagging
- Suitable n_samples for creating base models
- Suitable n_features for creating base models
- Keep balance between diversity and predictive performance
- What will happen to bagging if concept drift occurs?

DERIVE VARIANCE OF FINAL PREDICTION

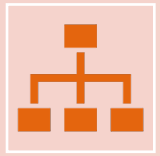
$$\begin{aligned}\mathrm{Var}(\bar{Y}) &= \mathrm{Var}\left(\frac{1}{m}\sum_{i=1}^{m} Y_i\right) \\ &= \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m}\mathrm{Cov}(Y_i, Y_j)\end{aligned}$$

Since $\mathrm{Var}(Y_i) = \sigma^2$ and $\mathrm{Cov}(Y_i, Y_j) = \rho\sigma^2$ for $i \neq j$

$$\begin{aligned}\mathrm{Var}(\bar{Y}) &= \frac{1}{m^2}(m\sigma^2 + m(m-1)\rho\sigma^2) \\ &= \frac{1}{m^2}(m\sigma^2 + m^2\rho\sigma^2 - m\rho\sigma^2) \\ &= \frac{\sigma^2}{m} + \rho\sigma^2\left(1 - \frac{1}{m}\right) \\ &= \frac{\sigma^2}{m} + \rho\sigma^2 - \frac{\rho\sigma^2}{m} \\ &= \rho\sigma^2 + \frac{(1-\rho)\sigma^2}{m}\end{aligned}$$

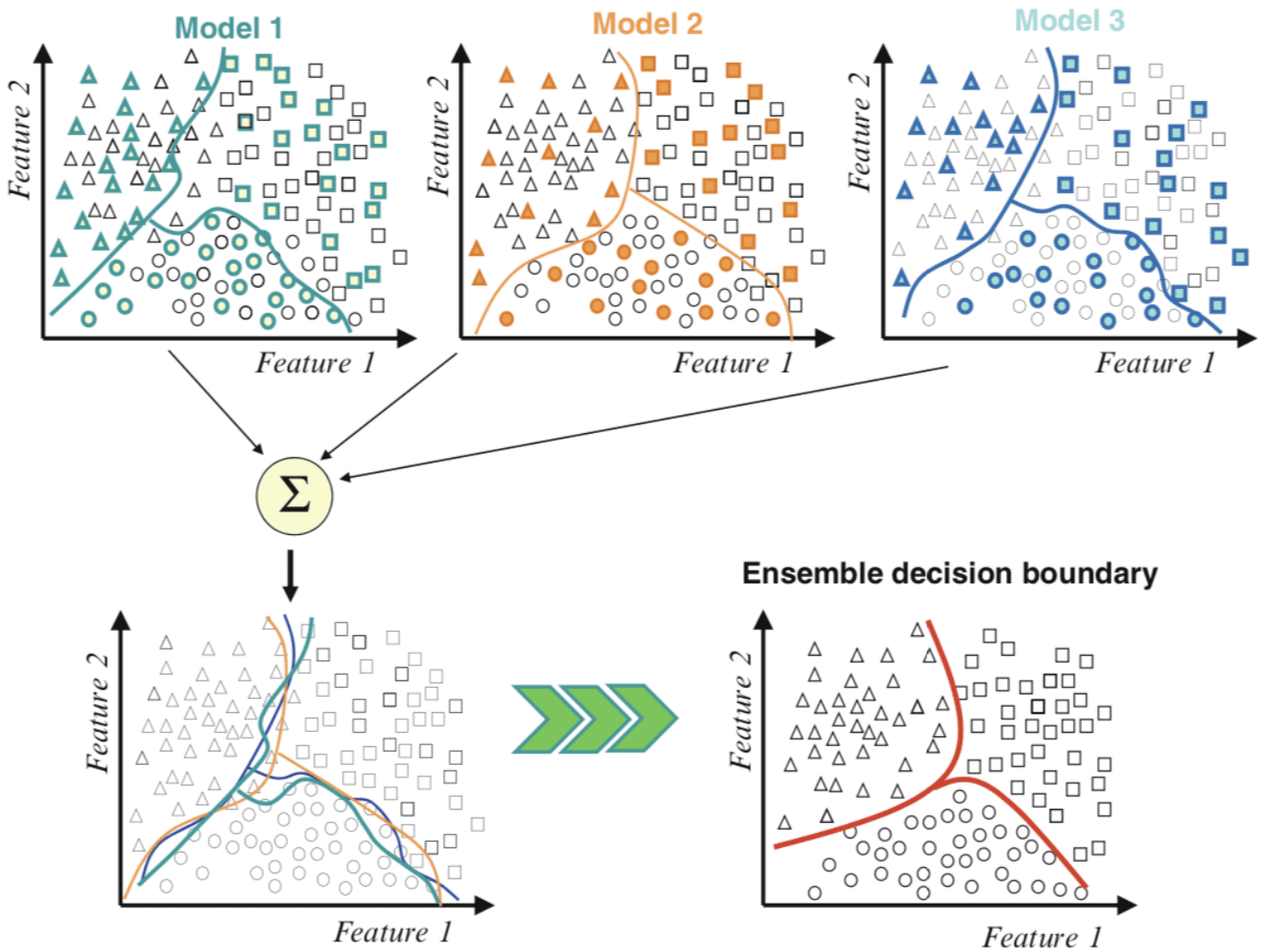
BAGGING WITH SUPERVISED LEARNING

BAGGING WITH SUPERVISED LEARNING



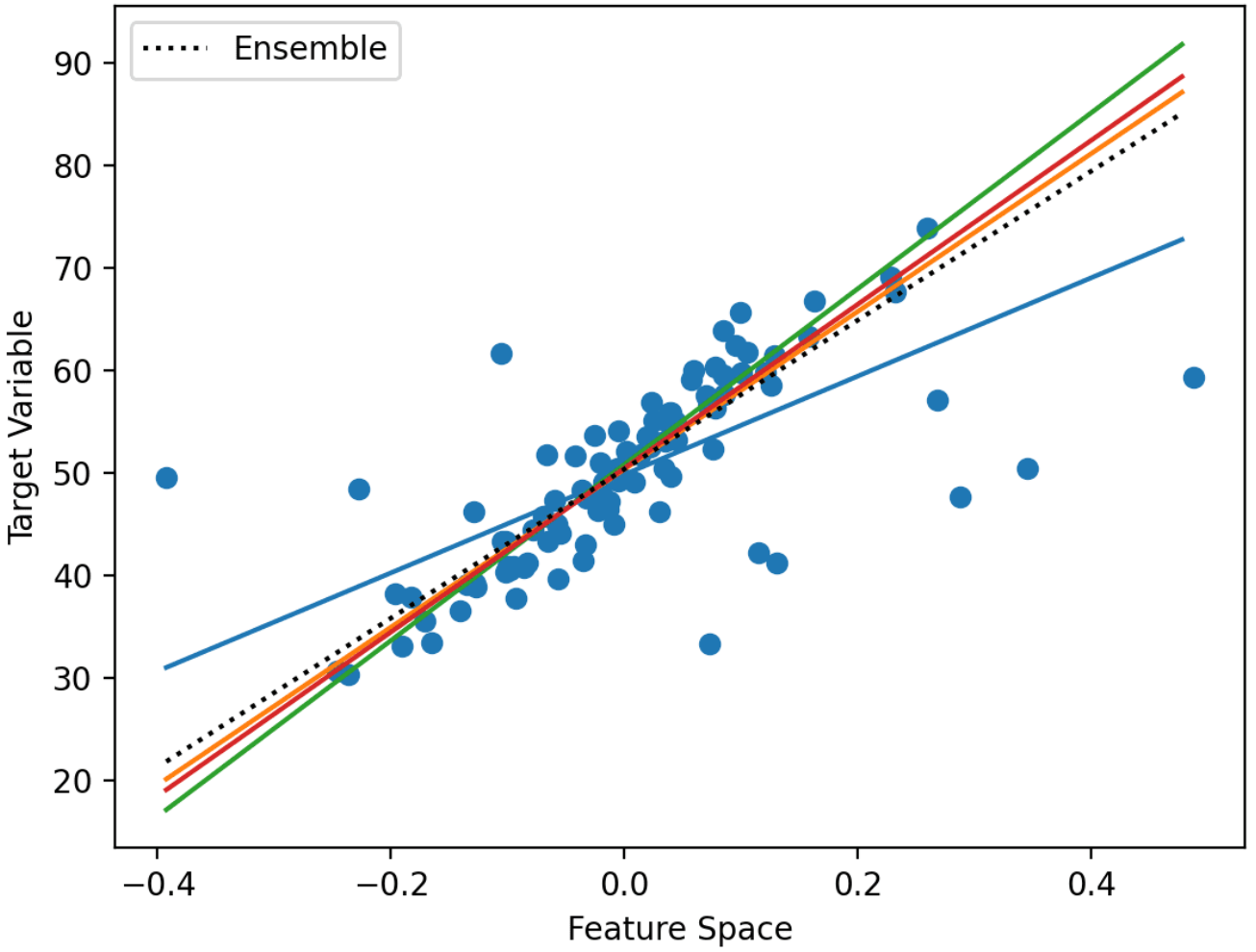
Classification

Regression



BAGGING WITH SL - CLASSIFICATION

BAGGING WITH SL - REGRESSION



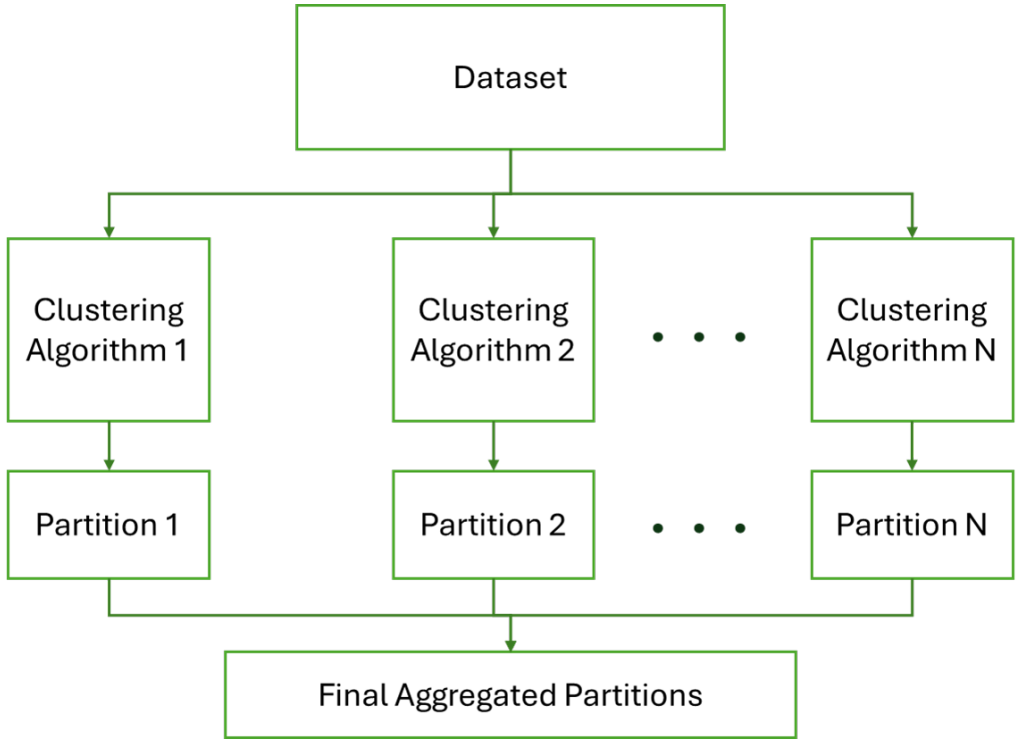
BAGGING WITH UNSUPERVISED LEARNING

BAGGING WITH UNSUPERVISED LEARNING

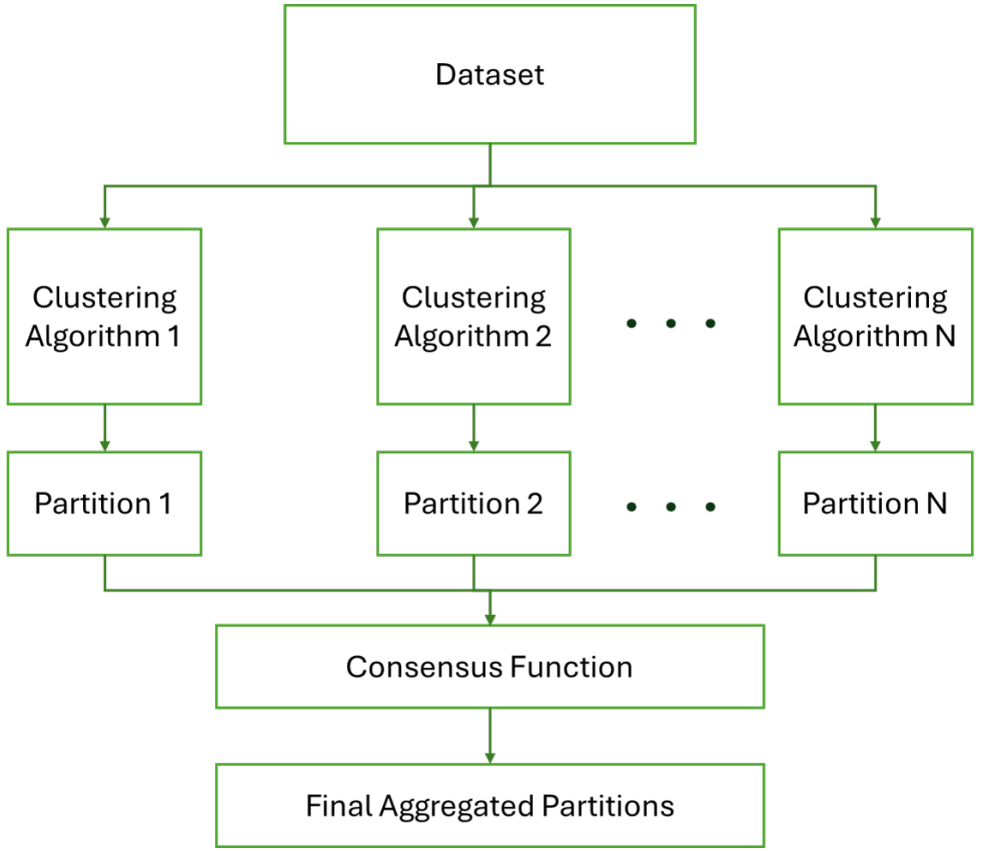
Clustering

Dimensionality reduction

BAGGING WITH UL - CLUSTERING

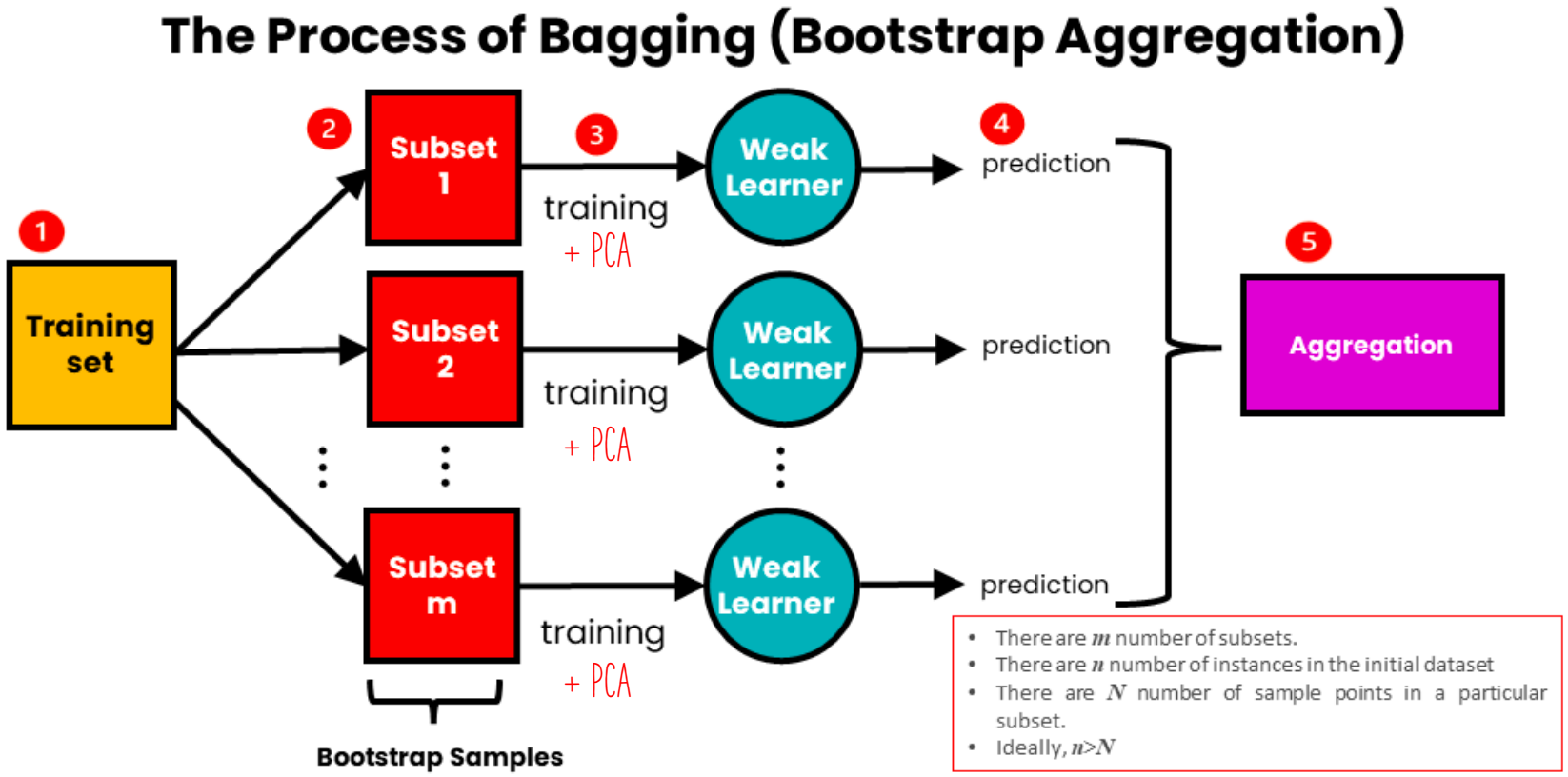


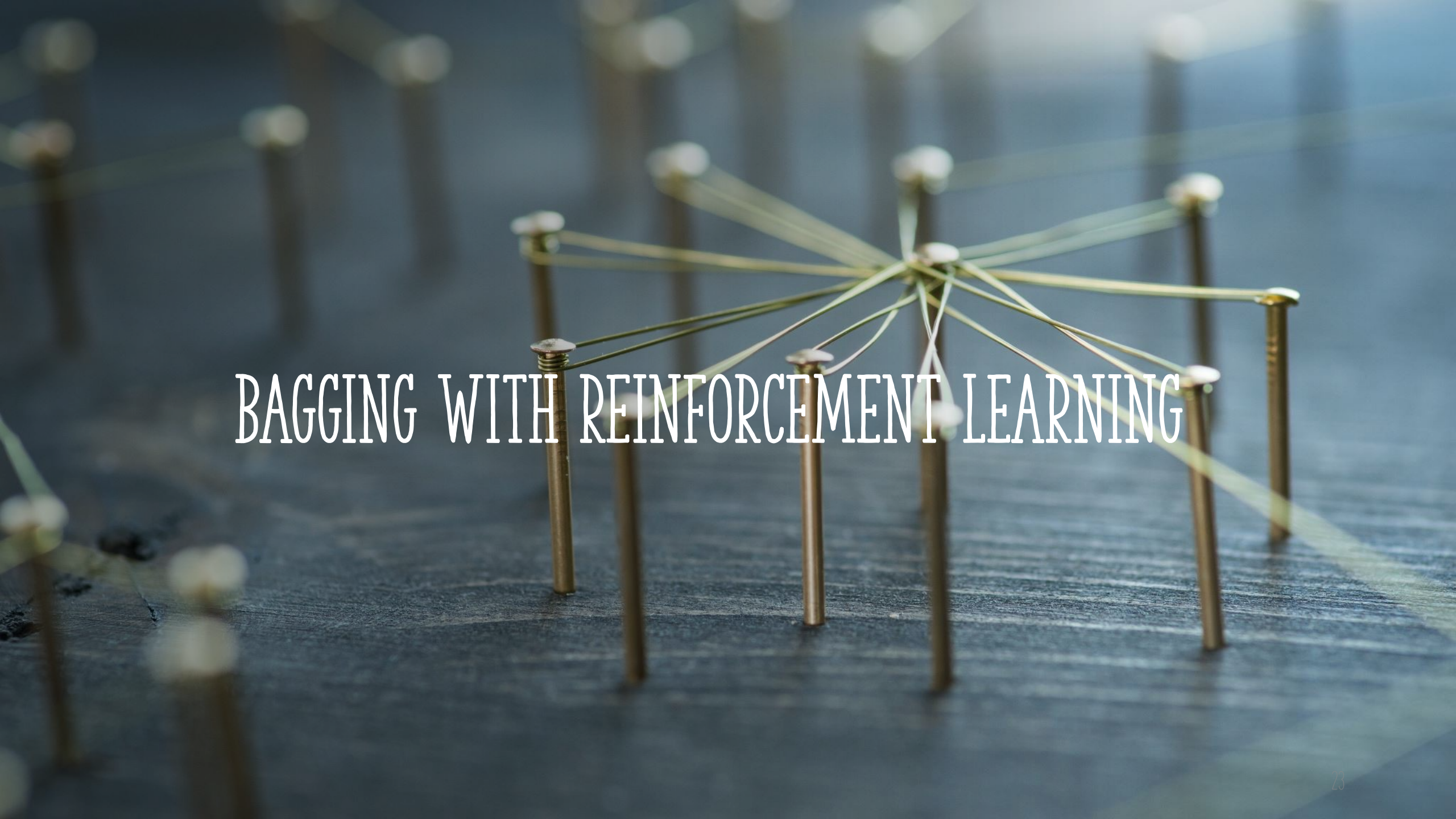
Same K



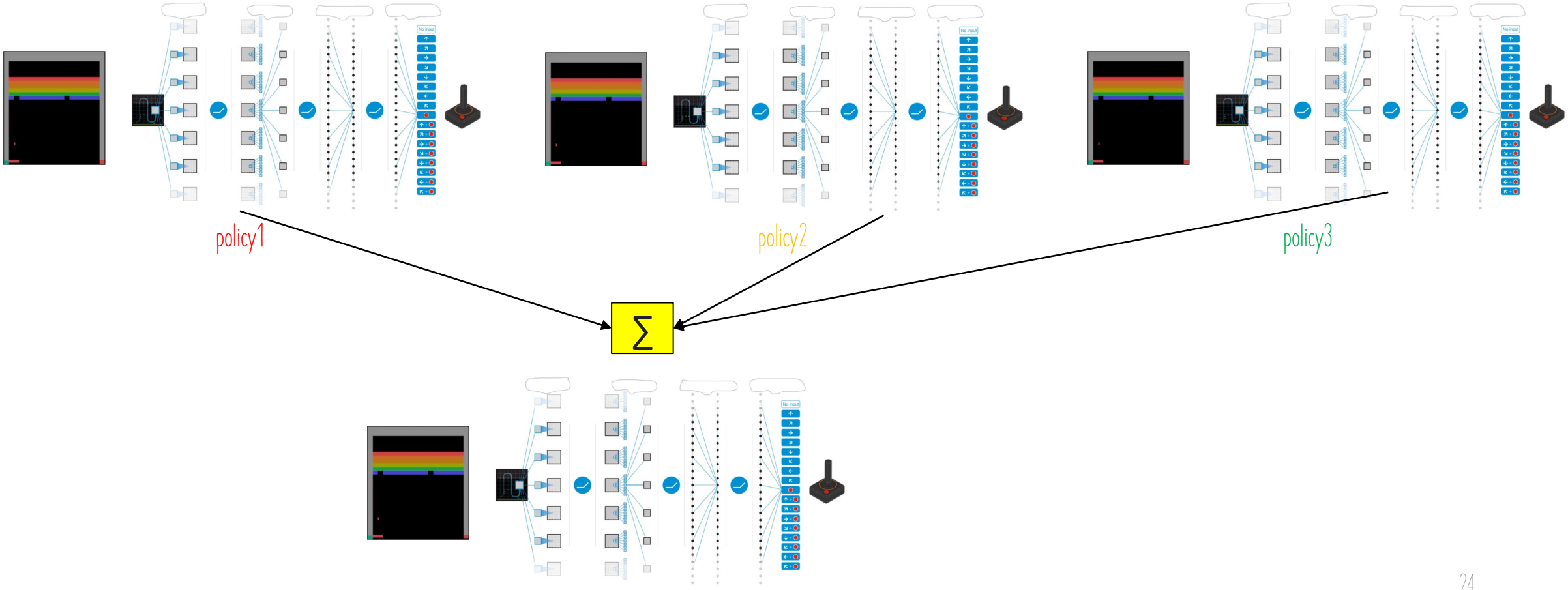
Different K

BAGGING WITH UL - DIMENESIONALITY REDUCTION




BAGGING WITH REINFORCEMENT LEARNING

BAGGING WITH REINFORCEMENT LEARNING



CODE

- Bagging - SL.ipynb
- Bagging - UL.ipynb
- Bagging - RL.ipynb

QUESTION & ANSWER

REFERENCE

- https://www.cs.cornell.edu/courses/cs4780/2018fa/lectures/lecturenote12.html
- https://medium.com/@ivanreznikov/stop-using-the-same-image-in-bias-variance-trade-off-explanation-691997a94a54
- https://www.analyticsvidhya.com/blog/2023/01/ensemble-learning-methods-bagging-boosting-and-stacking/
- https://medium.com/@ivanreznikov/stop-using-the-same-image-in-bias-variance-trade-off-explanation-691997a94a54
- https://machinelearningmastery.com/how-ensemble-learning-works/
- https://www.geeksforgeeks.org/consensus-clustering/
- https://commons.wikimedia.org/wiki/File:ML_dataset_training_validation_test_sets.png


THANK YOU