



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).

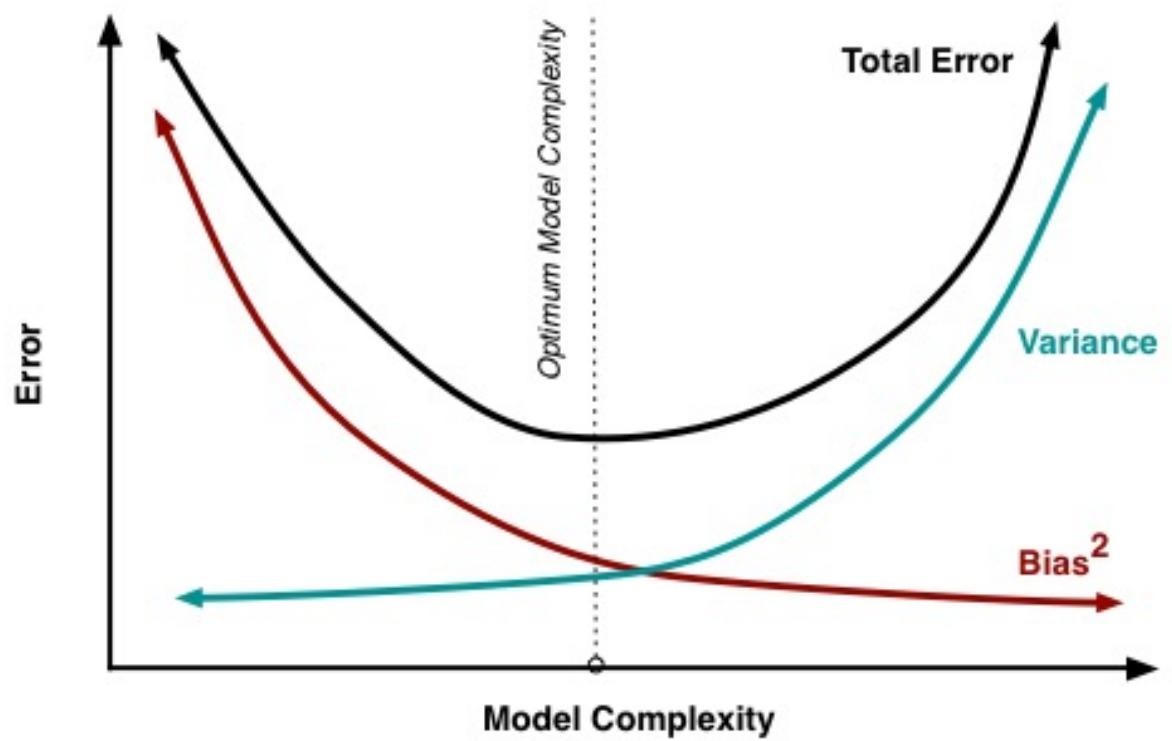
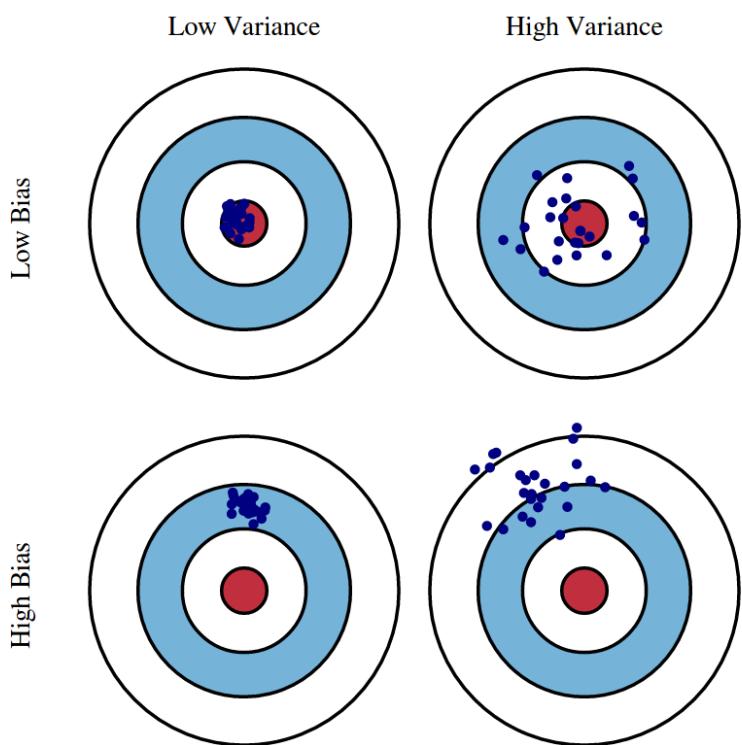
The final prediction is made by averaging (for regression) or voting (for classification) the predictions of all models.

This method reduces variance and helps prevent overfitting.

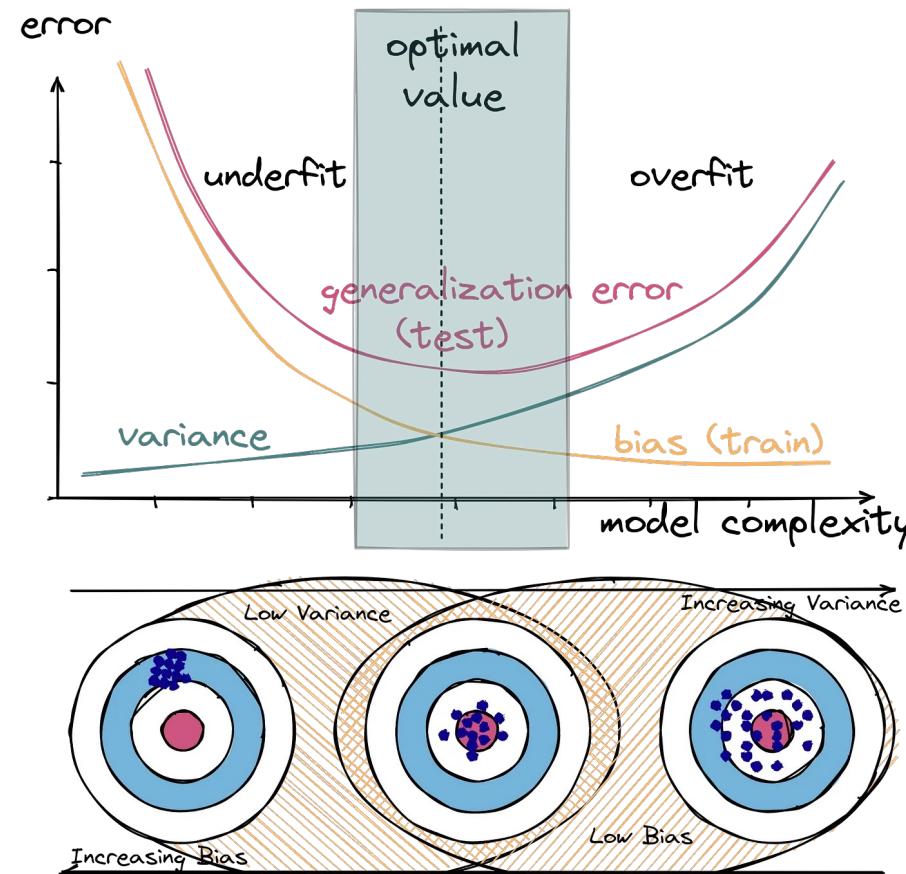
BENEFIT OF BAGGING

	Benefit of bagging
Supervised learning	<ul style="list-style-type: none">- Reduce variance- Improve generalization- Robustness to overfitting
Unsupervised learning	<ul style="list-style-type: none">- Enhance clustering stability- Noise reduction- Create diverse SL model
Reinforcement learning	<ul style="list-style-type: none">- Policy stability

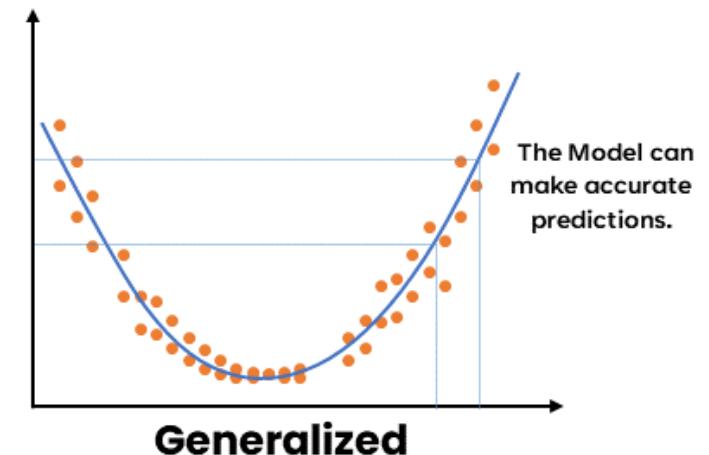
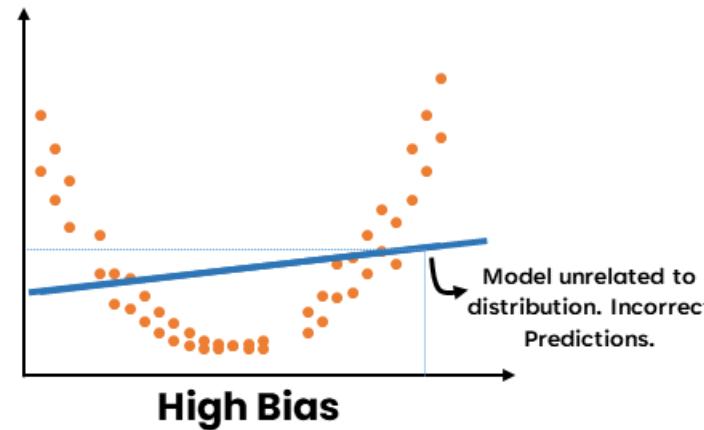
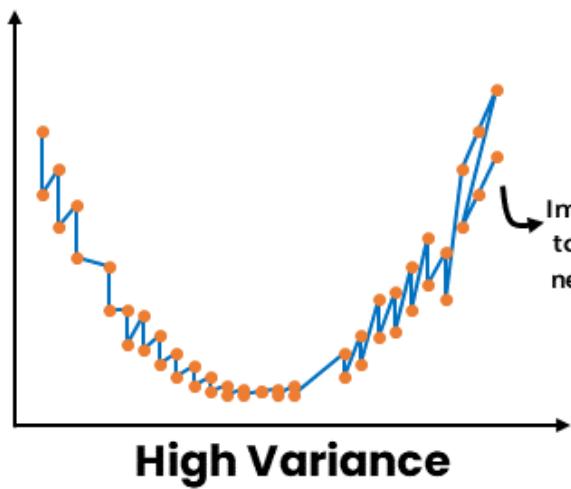
BIAS-VARIANCE TRADEOFF



BIAS-VARIANCE TRADEOFF



BIAS-VARIANCE TRADEOFF



HOW BAGGING WORKS

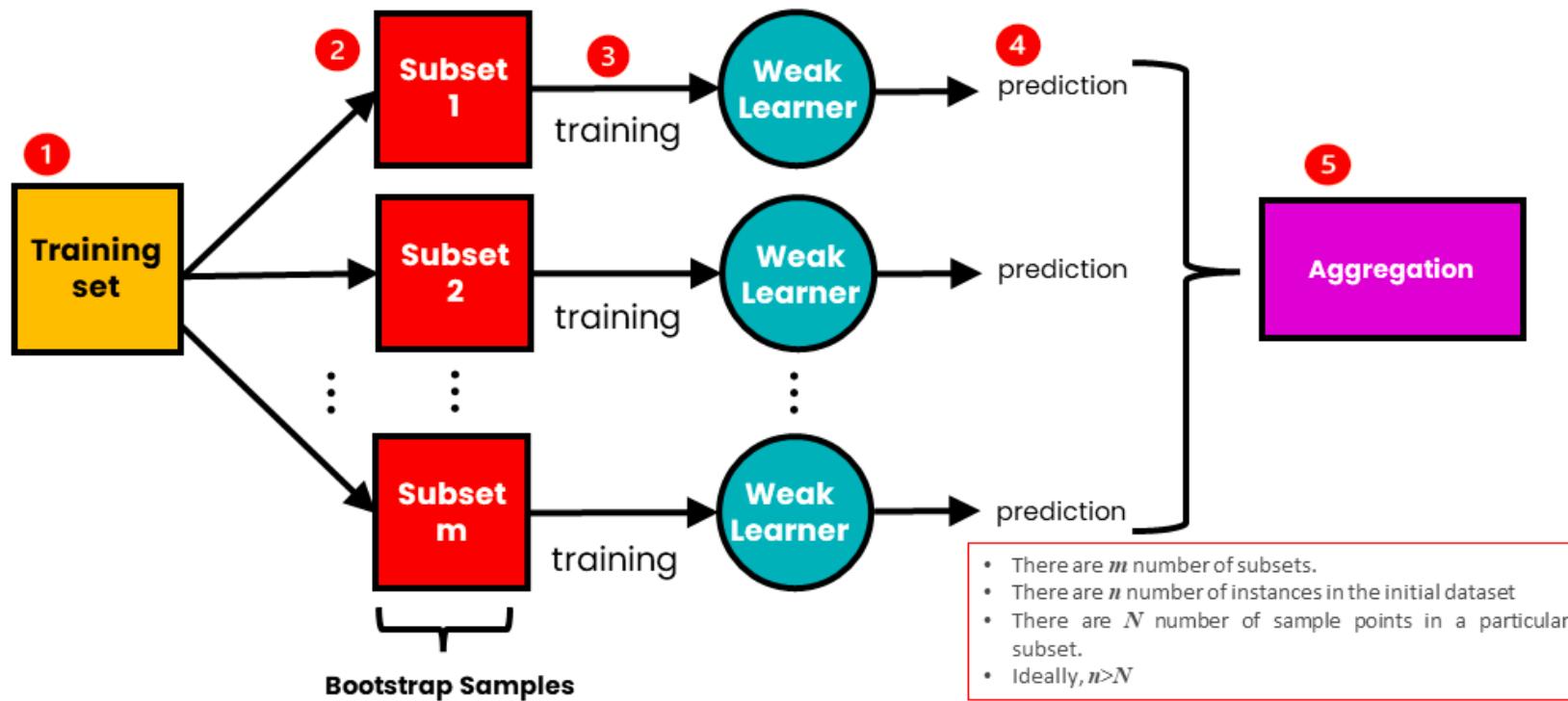
Bootstrapping

Training multiple
models

Aggregating
predictions

HOW BAGGING WORKS

The Process of Bagging (Bootstrap Aggregation)



HOW BAGGING WORKS



Single Dataset



Single Dataset



DISCUSSION ABOUT BAGGING

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



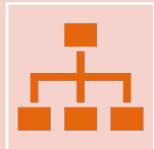
DISCUSSION ABOUT BAGGING

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

A stack of several books is shown, with a pair of dark-rimmed glasses resting on top of them. The books have various colored spines and covers, including blue, red, and white. The glasses are positioned centrally, with their lenses reflecting some of the surrounding environment.

BAGGING WITH SUPERVISED LEARNING

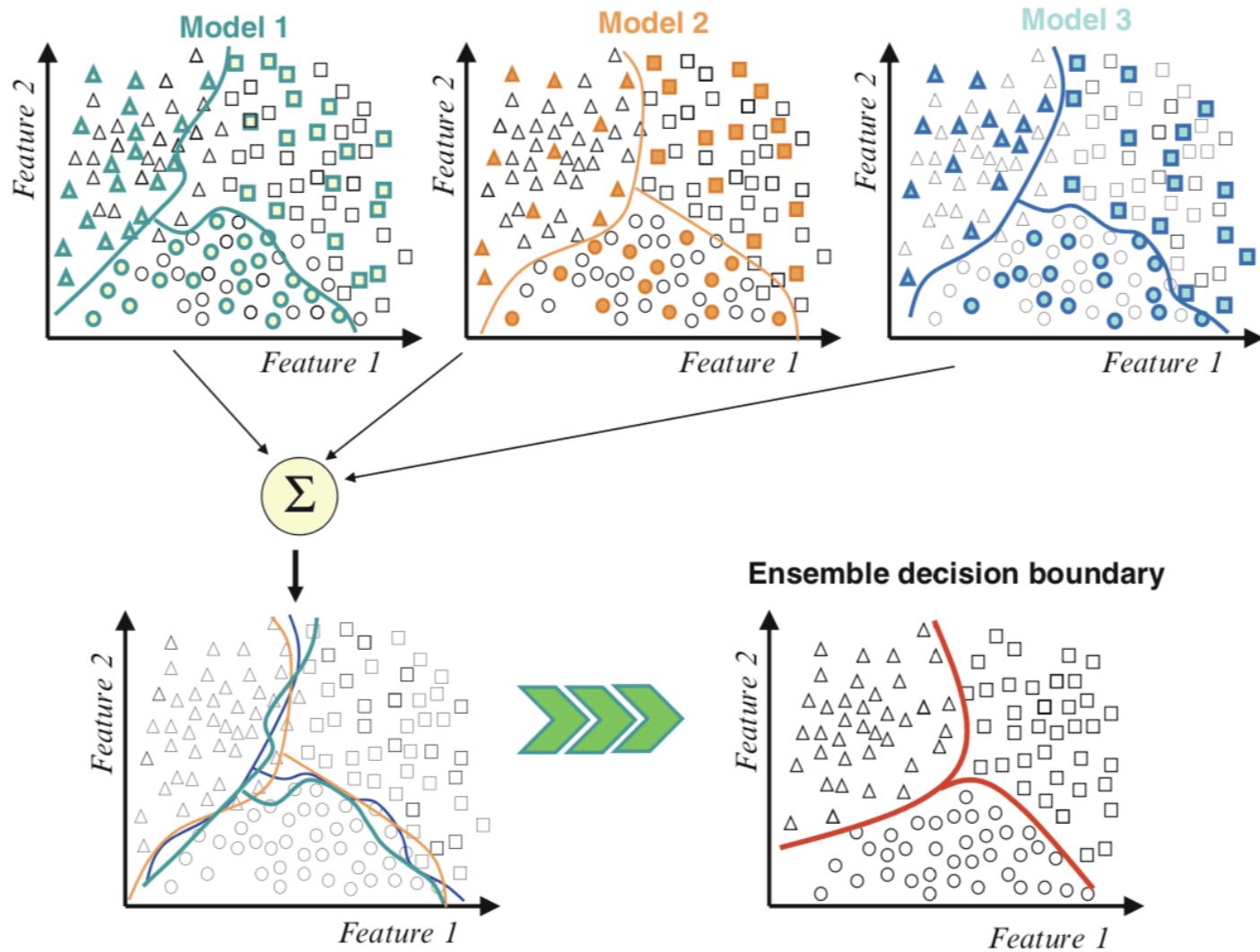
BAGGING WITH SUPERVISED LEARNING



Classification



Regression

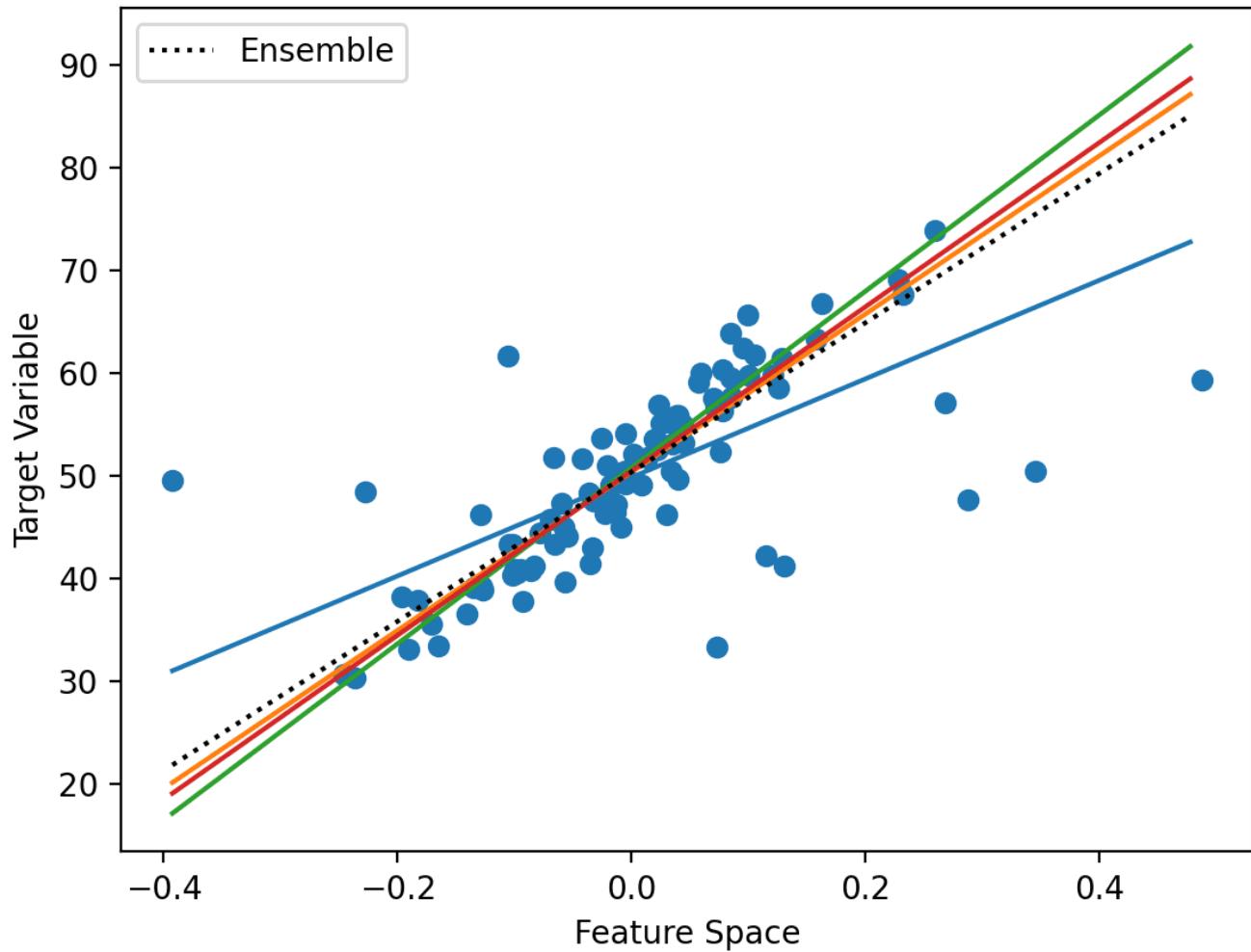


BAGGING WITH SL

- CLASSIFICATION

BAGGING WITH SL

- REGRESSION





BAGGING WITH UNSUPERVISED LEARNING

BAGGING WITH UNSUPERVISED LEARNING

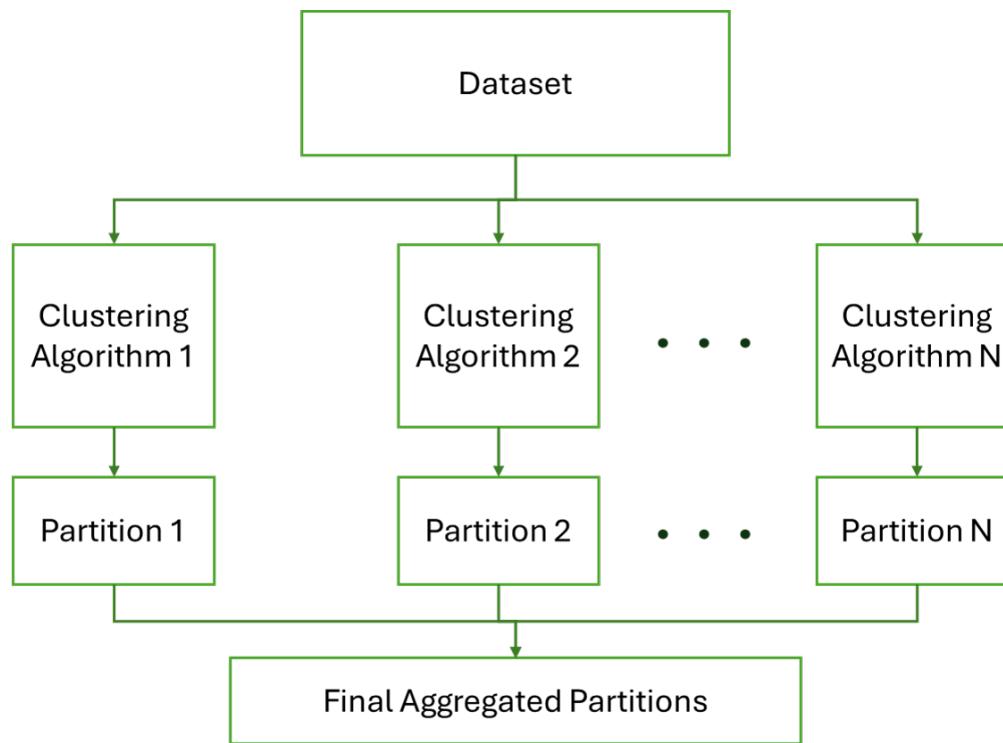


Clustering

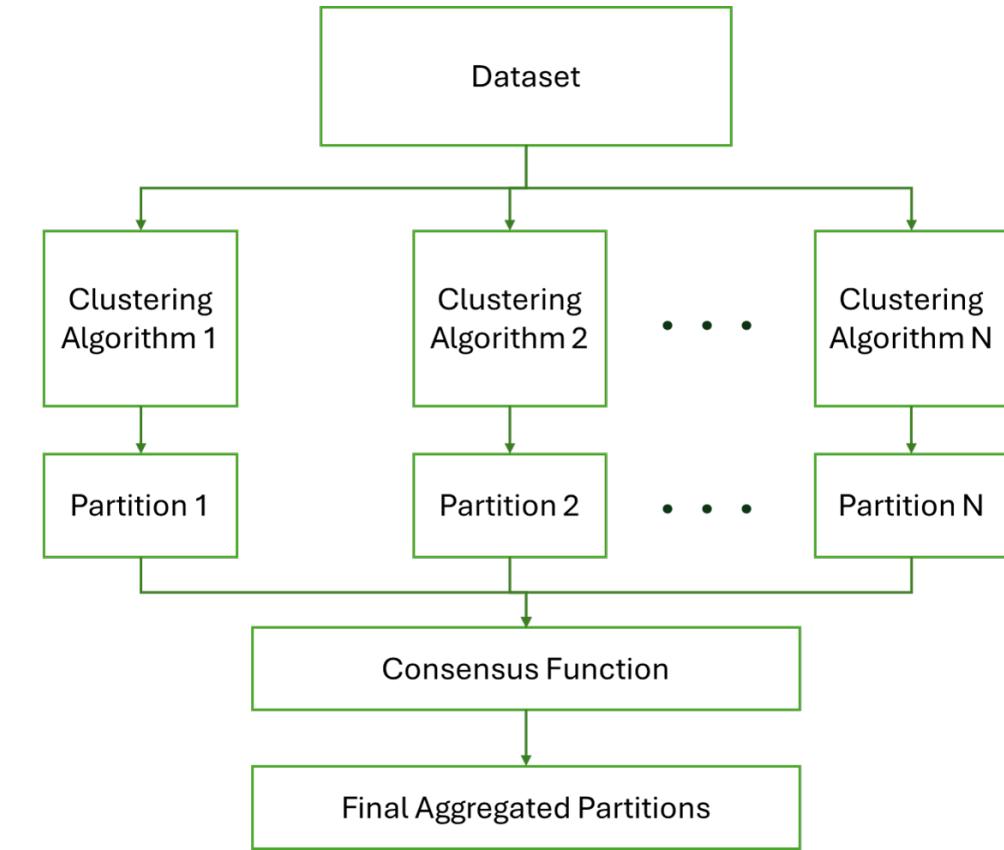


Dimensionality reduction

BAGGING WITH UL - CLUSTERING



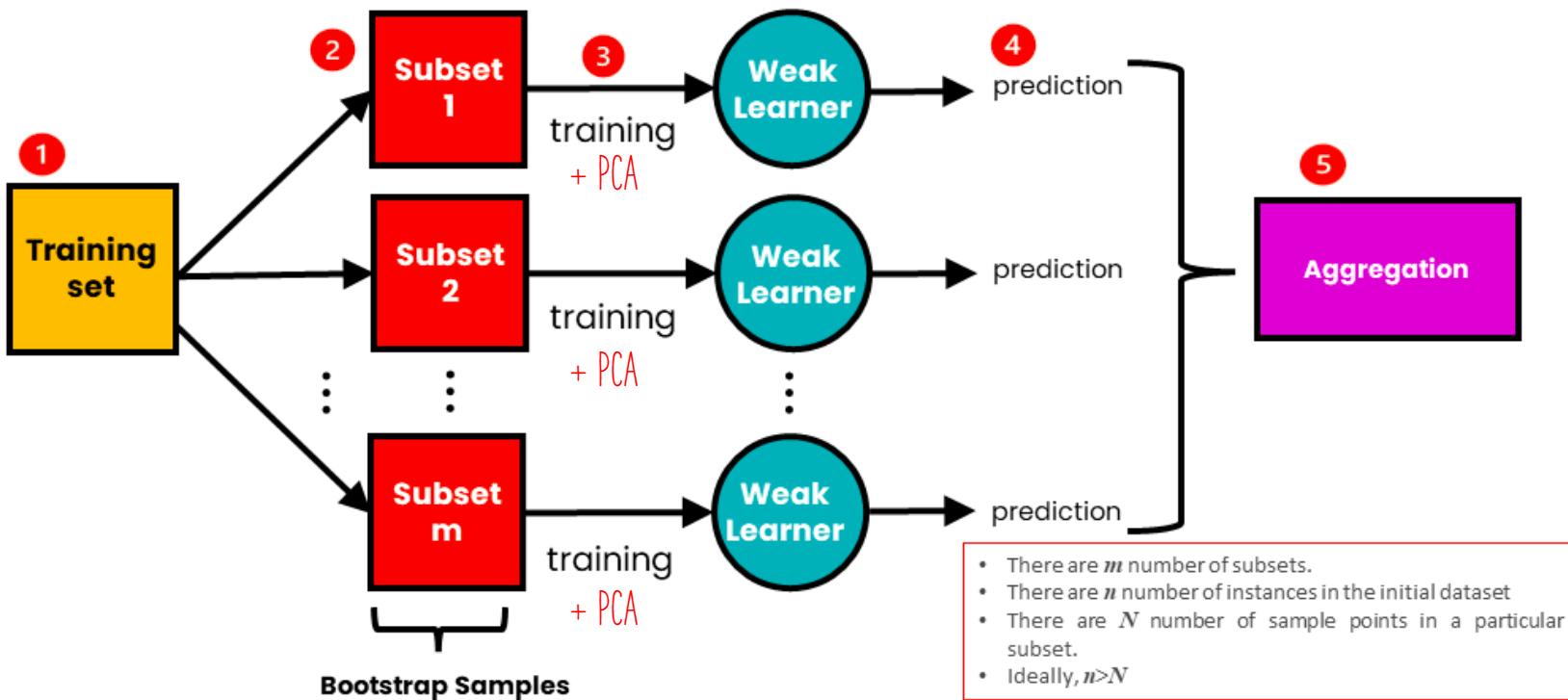
Same K



Different K

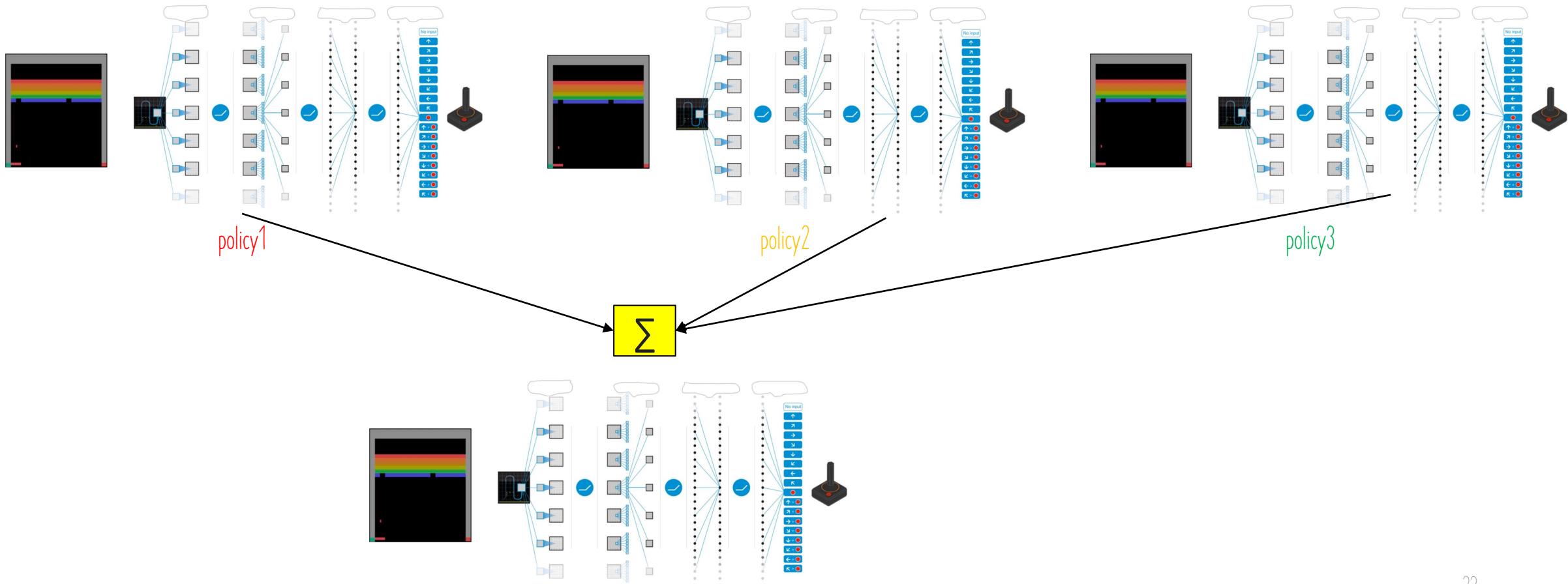
BAGGING WITH UL - DIMENESIONALITY REDUCTION

The Process of Bagging (Bootstrap Aggregation)



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

