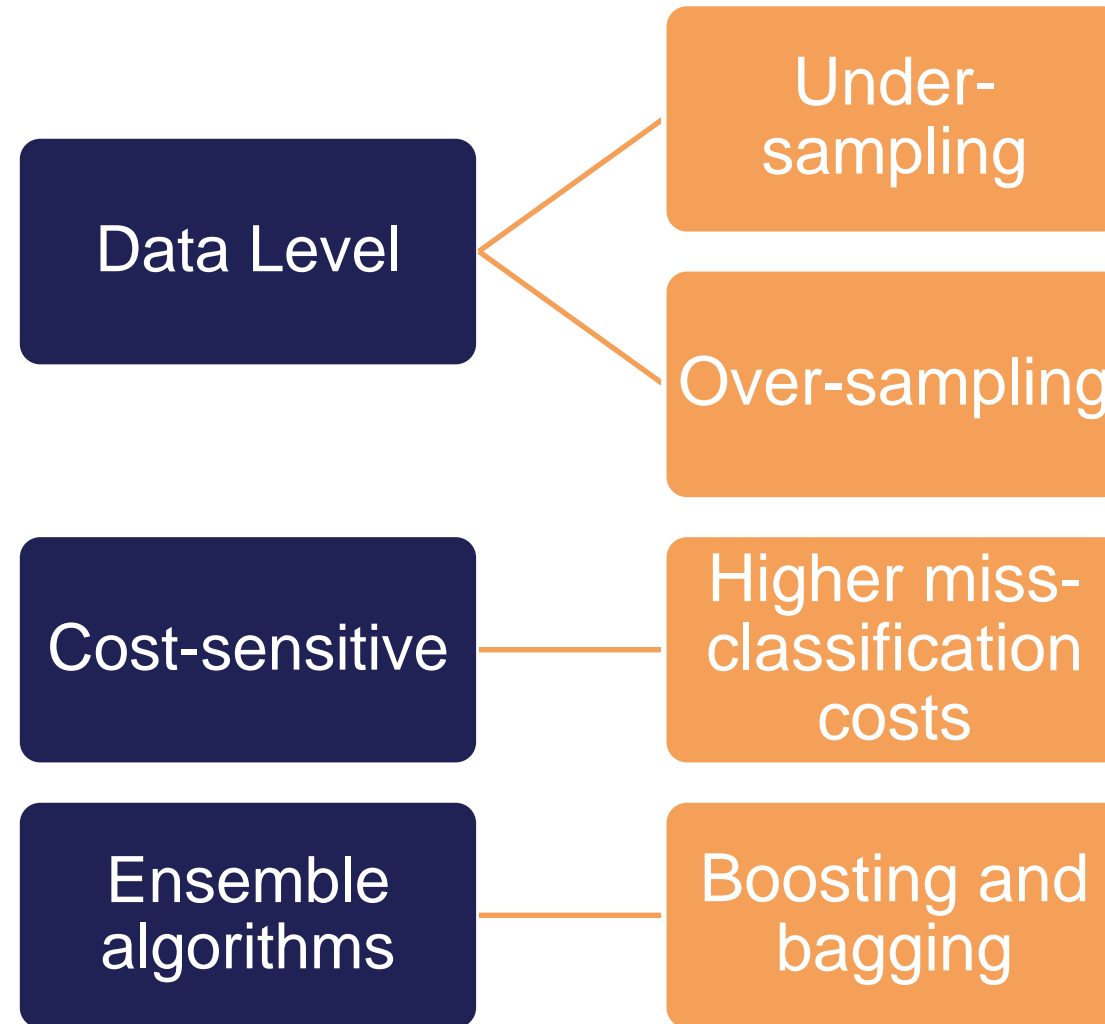# Ensemble methods for Imbalanced Data
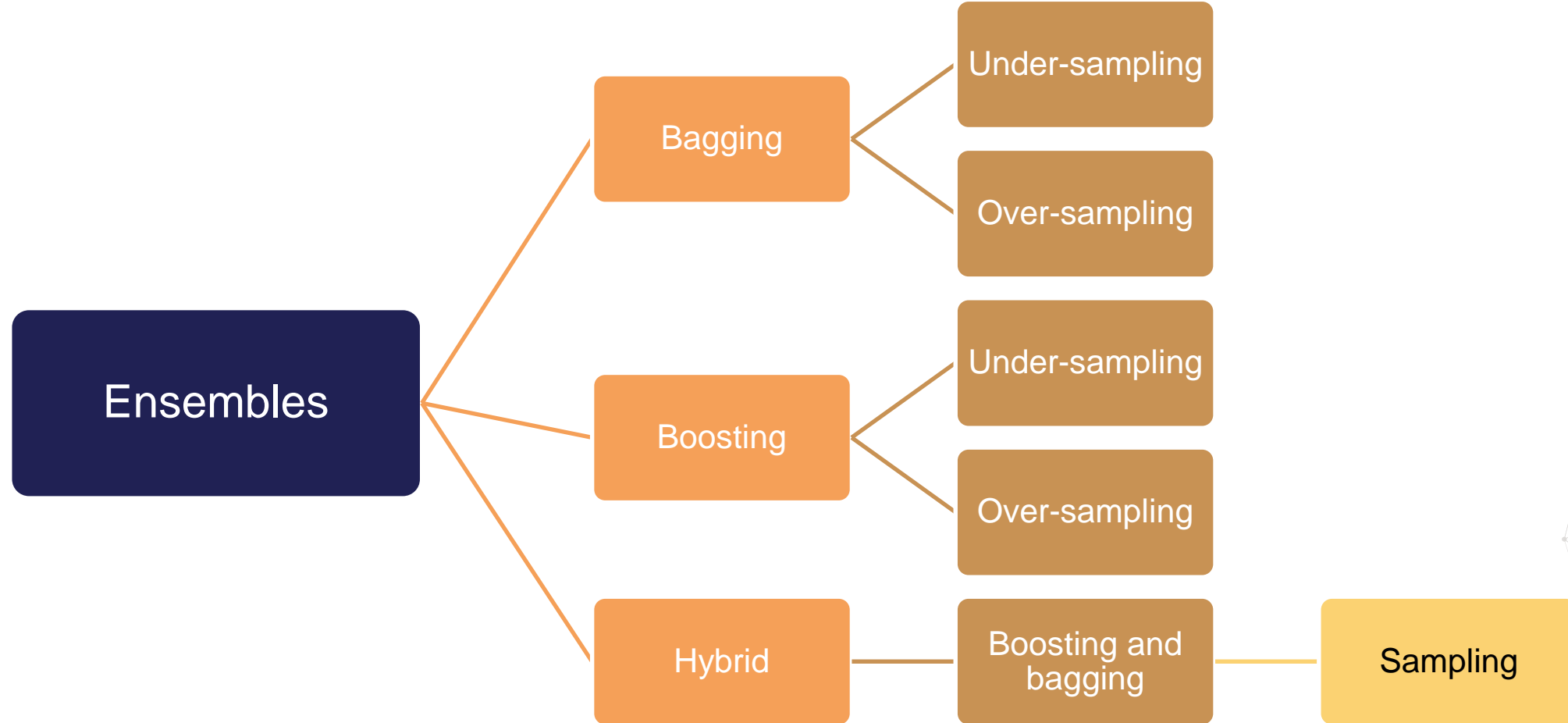
Solutions for Imbalanced Datasets
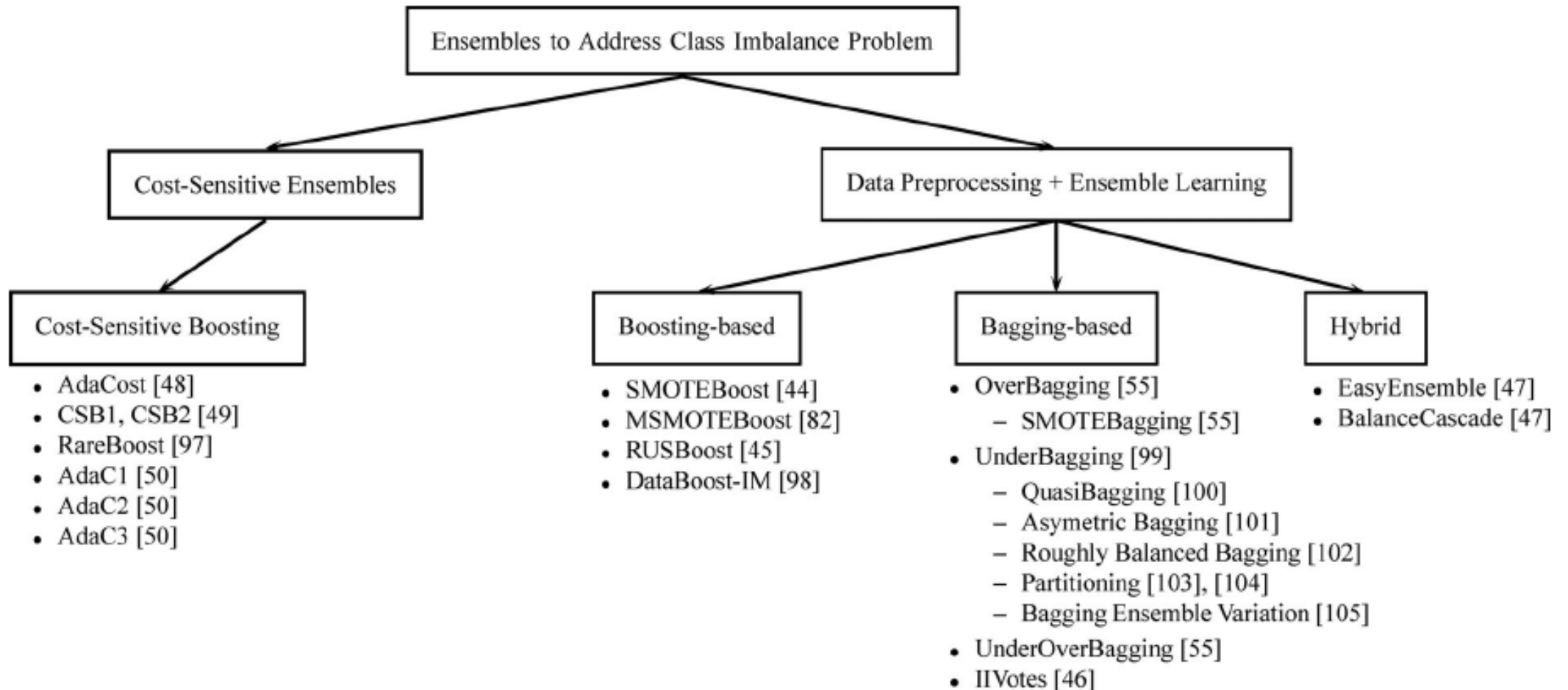
Data Level
- Under-sampling
- Over-sampling

Cost-sensitive
- Higher miss-classification costs

Ensemble algorithms
- Boosting and bagging

Ensembles for Imbalanced Datasets
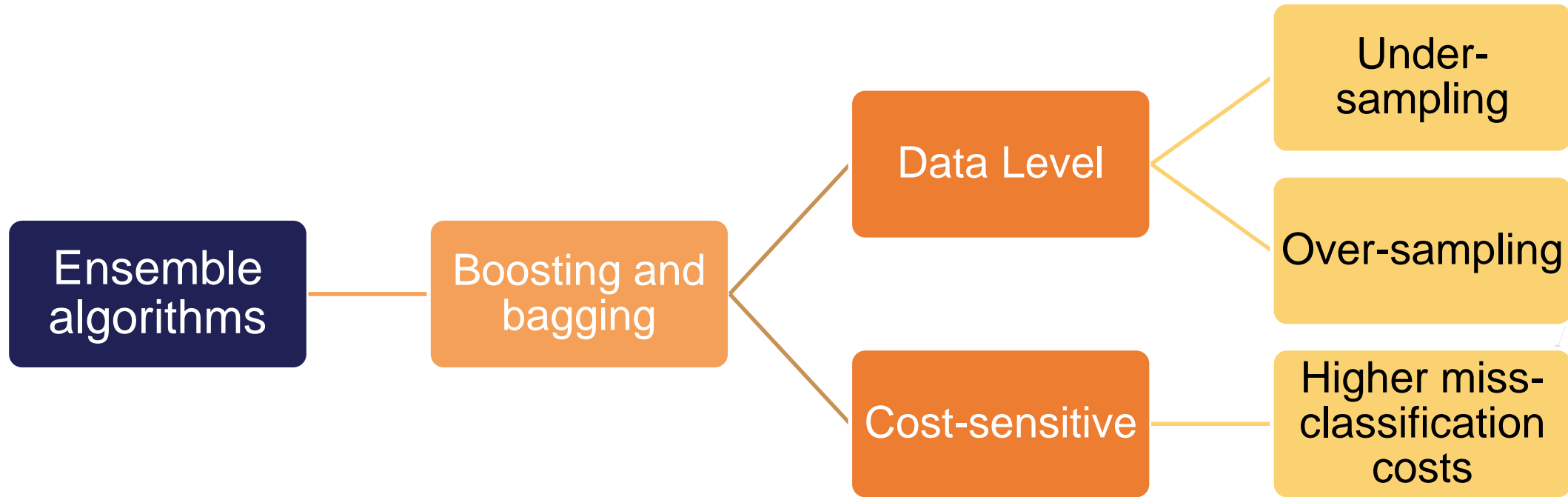
Ensembles for Imbalanced Datasets

Ensemble approaches

Ensembles to Address Class Imbalance Problem

Cost-Sensitive Ensembles

Data Preprocessing + Ensemble Learning

Cost-Sensitive Boosting
- AdaCost [48]
- CSB1, CSB2 [49]
- RareBoost [97]
- AdaC1 [50]
- AdaC2 [50]
- AdaC3 [50]

Boosting-based
- SMOTEBoost [44]
- MSMOTEBoost [82]
- RUSBoost [45]
- DataBoost-IM [98]

Bagging-based
- OverBagging [55]
  - SMOTEBagging [55]
- UnderBagging [99]
  - QuasiBagging [100]
  - Asymetric Bagging [101]
  - Roughly Balanced Bagging [102]
  - Partitioning [103], [104]
  - Bagging Ensemble Variation [105]
- UnderOverBagging [55]
- IIVotes [46]

Hybrid
- EasyEnsemble [47]
- BalanceCascade [47]

Ensemble approaches

Ensembles to Address Class Imbalance Problem

Cost-Sensitive Ensembles

Data Preprocessing + Ensemble Learning

Cost-Sensitive Boosting

- AdaCost [48]
- CSB1, CSB2 [49]
- RareBoost [97]
- AdaC1 [50]
- AdaC2 [50]
- AdaC3 [50]

Boosting-based

- SMOTEBoost [44] ✓
- MSMOTEBoost [82]
- RUSBoost [45] ✓
- DataBoost-IM [98]

Bagging-based

- OverBagging [55]
  - SMOTEBagging [55]
- UnderBagging [99] ✓
  - QuasiBagging [100]
  - Asymetric Bagging [101]
  - Roughly Balanced Bagging [102]
  - Partitioning [103], [104]
  - Bagging Ensemble Variation [105]
- UnderOverBagging [55]
- IIVotes [46]

Hybrid

- EasyEnsemble [47] ✓
- BalanceCascade [47] ✓

Ensemble approaches