

Ensemble methods

V, SCOTT

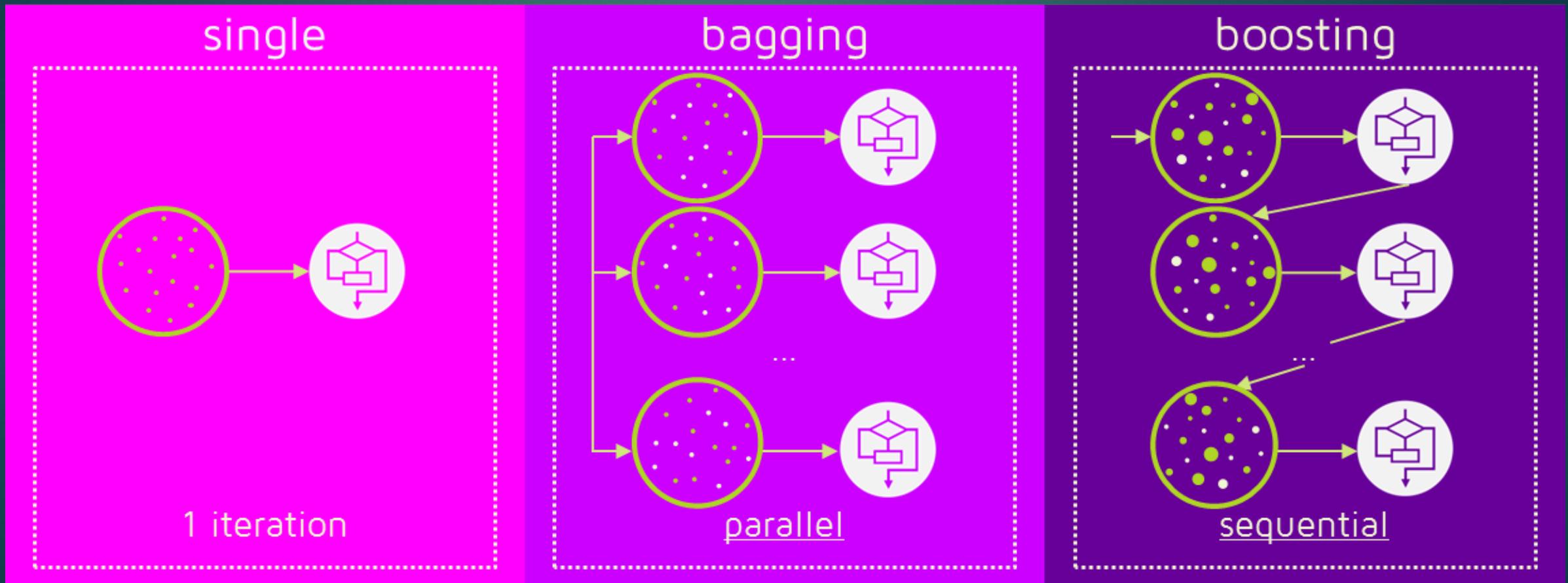
What is it?

- ▶ Machine learning technique
- ▶ Combines several model to make one super model
- ▶ Better results occur when there is diversity in the models
- ▶ Supervised learning

Common types

- ▶ Bootstrap aggregating, known as ‘bagging’.
- ▶ Bayes optimal classifier – can produce best average results, but cannot be practically implemented
- ▶ Boosting
- ▶ Bucket of models – Model selection algorithm is used to choose best model for problem

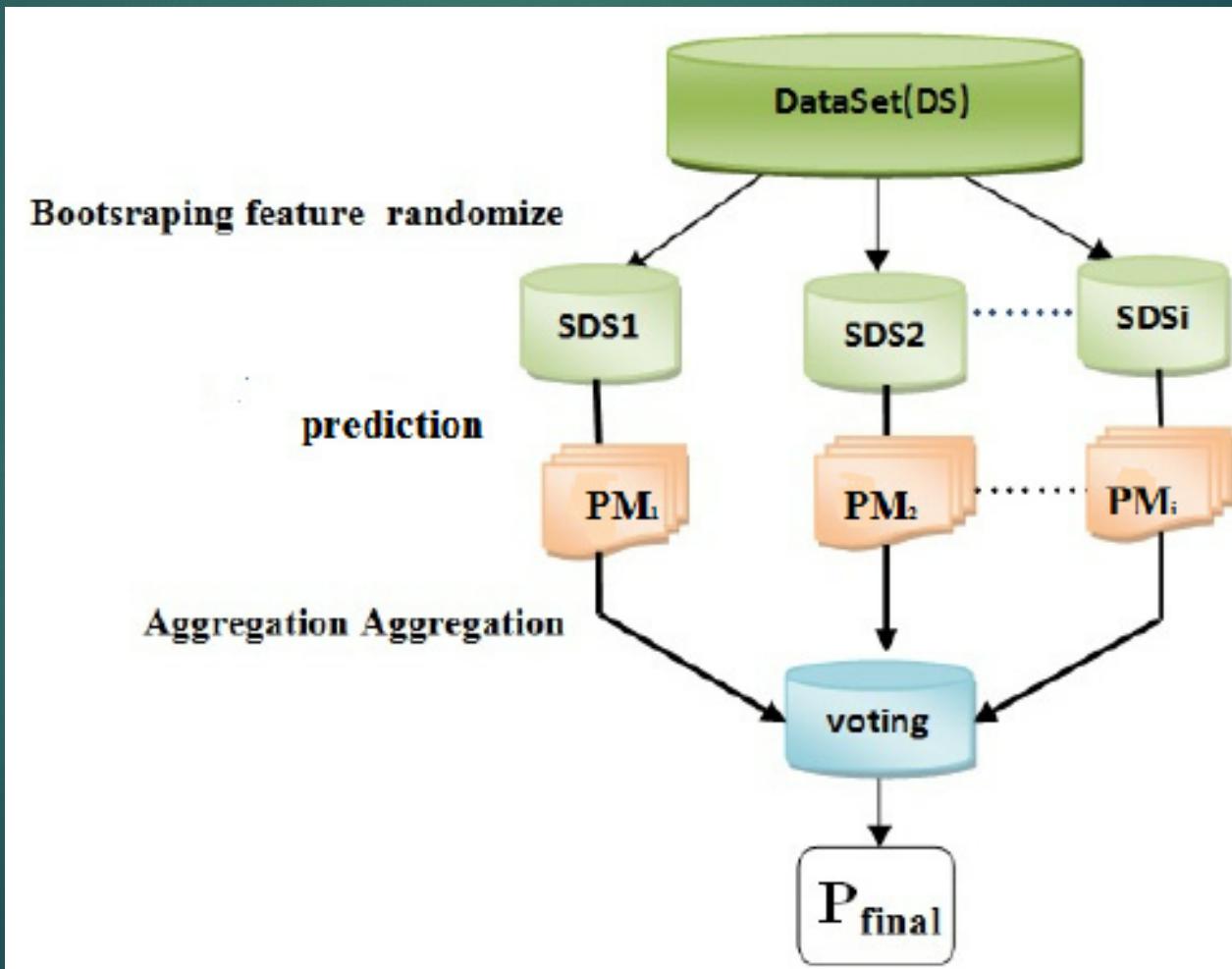
Bagging vs Boosting



Bagging

- ▶ Bagging creates smaller training sets from the overall training set
- ▶ Generally uses random sampling with replacement
- ▶ Resulting models from smaller sets are fitted and combined by averages
- ▶ Main purpose is to decrease variance to tune prediction
- ▶ Works in parallel, each model is built independently
- ▶ Example: multiple decision trees)

Bagging General Idea



Bagging vs Random Forest

- ▶ Not the same!
- ▶ Bagging has 1 parameter: The number of tree/algorithms
- ▶ Random Forests tend to have 2 parameters: Number of trees, as well as the number of features to search over

Pros of Bagging

- ▶ Reduces variance
- ▶ Can handle categorical features
- ▶ Better at handling overfitting
- ▶ Works best with unstable classifiers (classifiers sensitive to variations)

Cons of Bagging

- ▶ Has trouble reducing variance with correlated features
- ▶ Is actually worse with stable classifiers

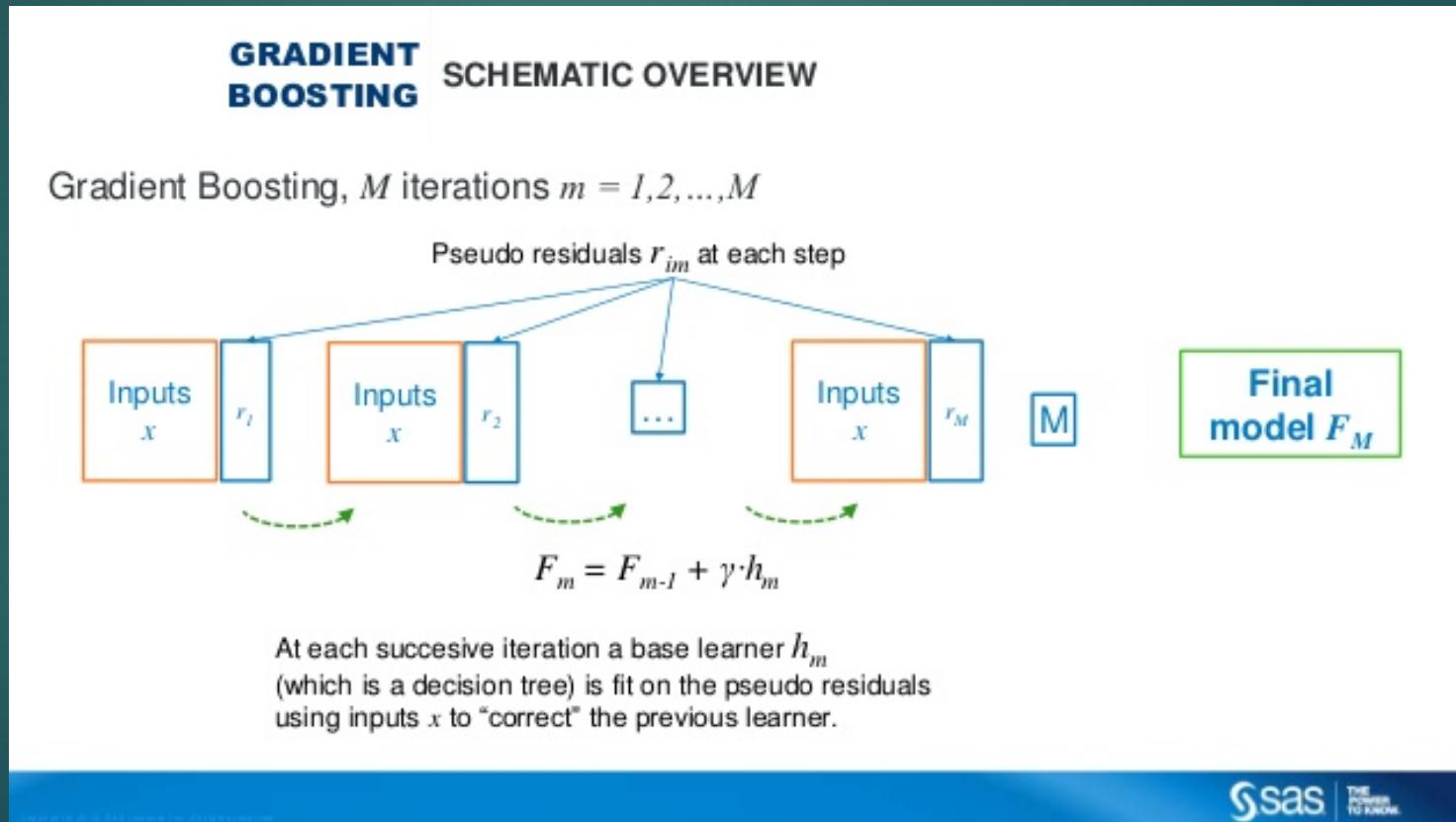
Bagging example

- ▶ <https://machinelearningmastery.com/bagging-and-random-forest-ensemble-algorithms-for-machine-learning/>

Boosting

- ▶ Similar to bagging
- ▶ Major difference is that in boosting, data has weights attached to its probability.
- ▶ Weights determine elements that were most likely misclassified by the previous models
- ▶ Works sequentially, working off of each previous model
- ▶ Aims to decrease bias
- ▶ Example: gradient boosting

Boosting General Idea



Pros of Boosting

- ▶ Better at classifying ‘difficult’ samples
- ▶ More often generates a model with lower errors
- ▶ Attempts to reduce bias

Cons of Boosting

- ▶ Prone to over-fitting
- ▶ Noise sensitive

Conclusion

- ▶ Yes