1. Khi dùng AdaBoost và bị underfitting, ta nên điều chỉnh những hyperparameter? Nên điều chỉnh tăng hay giảm?

Khi dùng AdaBoost và bị underfitting, ta có thể thử tăng estimators, giảm regularization hyperparameters, hoặc tăng nhẹ learning rate.

2. Khi dùng Gradient Boosting và overfitting, ta nên điều learning\_rate và n\_estimators tăng hay giảm? Tại sao?

Khi dùng Gradient Boosting bị overfitting ta nên thử giảm learning rate và n\_estimators.

3. Nếu muốn huấn luyện một ensemble của nhiều models trên nhiều servers khác nhau, ta nên chọn phương pháp nào: bagging, boosting, hay stacking? Tại sao?

* Ta nên sử dụng bagging
* Vì từ tập dữ liệu gốc sẽ chia random ra các tập con và các tập con này có quyền được training song song trên nhiều models trên nhiều server
* Các models này chạy song song tìm ra các kết quả khác nhau
* Lấy các kết quả khác nhau hợp lại tìm ra kết quả cuối cùng

1. Extra-Trees là gì? Nó có ưu điểm nào so với Random Forests?

* Thay vì phải học dữ liệu tính ra threshold ở từng node thì extra-trees sẽ chọn 1 cách ngẫu nhiên
* Extra đảm bảo tính diverse tốt hơn random forest
* Khi chạy trên các tập có dữ liệu lớn và số lượng cây có số node khổng lồ thì extra tree cũng sẽ chạy nhanh hơn random forest

1. Cho 5 classification models khác nhau. Sau khi train trên cùng một tập dữ liệu, các model này đạt độ chính xác khoảng 90%. Có cách nào để kết hợp các models này cho độ chính xác cao hơn không? Nếu có, mô tả giải pháp.

* Chúng ta có thể sử dụng voting để kết hợp kết quả của các models
* Chúng ta sử cho models train trên cùng 1 tập dữ liệu sau đó bỏ phiếu dựa trên kết quả prediction của các models.