# Тема 5. Композиции: бэггинг, случайный лес

Рассмотрим простые методы композиции: бэггинг и случайный лес. Вы узнаете, как можно получить распределение среднего по генеральной совокупности, если у нас есть информация только о небольшой ее части; посмотрим, как с помощью композиции алгоритмов уменьшить дисперсию и таким образом улучшить точность модели; разберём, что такое случайный лес, какие его параметры нужно «подкручивать» и как найти самый важный признак. Сконцентрируемся на практике, добавив немного математики.

План

Бэггинг

Ансамбли

Бутстрэп

Бэггинг

Out-of-bag ошибка

Случайный лес

Алгоритм

Сравнение с деревом решений и бэггингом

Параметры

Вариация и декорреляционный эффект

Смещение

Сверхслучайные деревья

Схожесть с алгоритмом k-ближайших соседей

Преобразование признаков в многомерное пространство

Оценка важности признаков

Плюсы и минусы случайного леса

Домашнее задание №5