# Random Forest

**Определение**: Случайный Лес — это ансамблевый метод машинного обучения, который используется для задач классификации, регрессии и других. Он строит множество деревьев решений на этапе обучения и затем выдает итоговый результат на основе голосования (в случае классификации) или среднего значения (в случае регрессии) предсказаний отдельных деревьев.

**Ключевая идея**: Метод использует технику **бутстрепа** (выборки с возвращением) для обучения каждого дерева на случайных подмножествах данных. Также на каждом шаге разбиения дерева случайным образом выбирается подмножество признаков, что помогает снизить переобучение и повысить точность модели.

# Bagging

**Определение**: Бэггинг — это общая техника увеличения точности и устойчивости моделей за счет обучения нескольких моделей на разных подмножествах данных (созданных с помощью **бутстрепа**) и объединения их предсказаний.

**Процесс**:

1. Создаются несколько случайных подмножеств данных (бутстреп-выборки).
2. На каждом подмножестве обучается отдельная модель.
3. Итоговые предсказания моделей объединяются (например, путем голосования или усреднения).

**Цель**: Бэггинг помогает снизить **дисперсию** модели (вариативность) и улучшить качество предсказаний, особенно для моделей с высокой вариативностью, таких как деревья решений.

# Бутстреп

* **Определение**: Бутстреп — это статистический метод повторной выборки, который используется для оценки распределения статистики путем случайной выборки с возвращением из исходных данных. Каждый набор данных, полученный таким образом, называется **бутстреп-выборкой**.
* **Применение в машинном обучении**: Бутстреп используется в методах, таких как Случайный Лес и Бэггинг. На каждой бутстреп-выборке обучается отдельная модель в ансамбле, что создает разнообразие моделей.
* **Преимущества**: Бутстреп позволяет строить более обобщенные модели, обучая их на слегка отличающихся версиях данных. Это особенно полезно для небольших выборок или при создании ансамблей моделей.