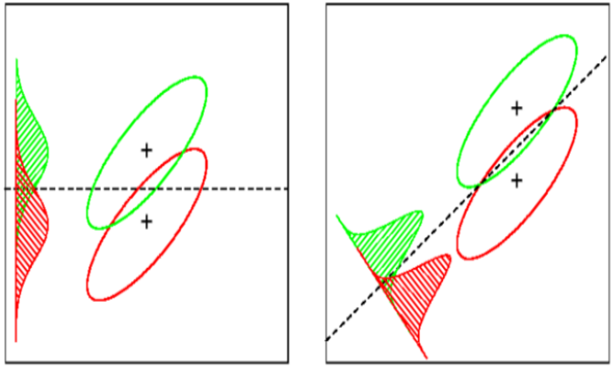
Вебинар 1.
Повторение основ машинного обучения.
Алгоритм Линейной Регрессии.
Выполнение оптимизации методом градиентного спуска.

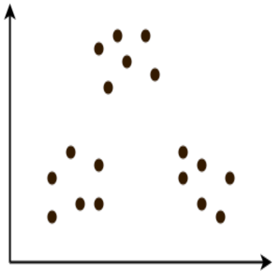
План урока

- Повторение основ машинного обучения
- Задачи оптимизация
- Линейная регрессия

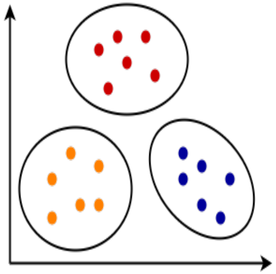
Каким бывает машинное обучение

- Supervised — поиск зависимости между распределением данных и таргетной переменной
- Unsupervised — поиск статистических паттернов данных, которые могут быть связаны со скрытыми свойствами, не заданными в явном виде
- Reinforcement- та или иная реакция машины на внешние данные подкрепляется неким стимулом.





Before K-Means



After K-Means

Каким бывает Supervised машинное обучение

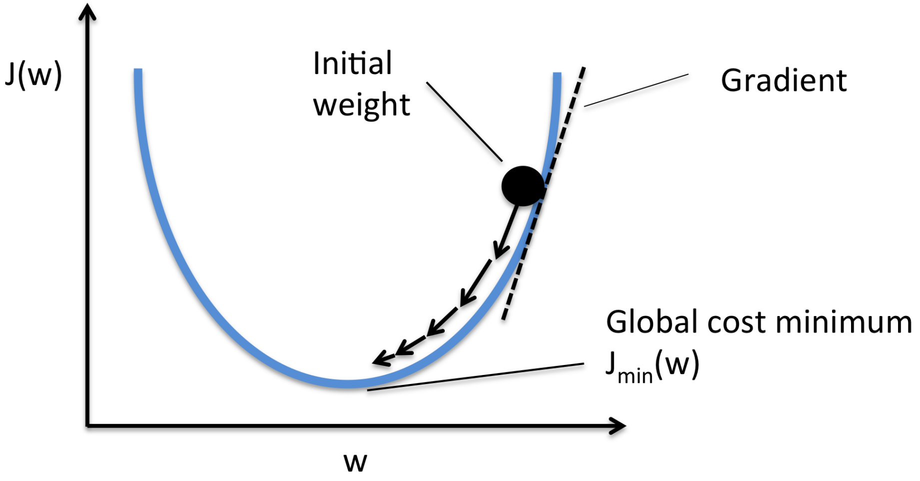
- Регрессия
- Классификация
  - Бинарная
  - Мультикласс
    - (Могут ли классы пересекаться?)
      - Обычная
      - Мультилейбл
        (Есть ли иерархия включения классов друг в друга?)
        - Иерархическая
        - Обычная
    - (Классы равнозначны или упорядочены?)
      - Обычная
      - Ординальная

Supervised машинное обучение

- Данные:
  - Матрица данных X
  - Таргетная переменная Y
- Модель
  - Выражает прямо или косвенно(например в виде вероятности) значение таргетной переменой в зависимости от входных данных
  - Может быть представлена соответствующей функцией или алгоритмом, или их комбинацией их наборов
  - Всегда включает в себя ряд параметров, которые позволяют подстроить модель под имеющиеся данные
- Функция потерь
  - Количественно оценивает предсказательное качество модели, принимая на вход наблюдаемые и предсказываемые моделью значения.
- Обучение модели:
  - Способ при заданной модели и заданной функции потерь настроить параметры модели так, чтобы функция потерь приняла максимально/минимально возможное значение.

Оптимизация

- Есть функция нескольких переменных, которая выражает оптимальность/неоптимальность. Например, цена/качество, качество/цена.
- У функции есть заданные переменные, и переменные подлежащие оптимизации.
- Для переменных, подлежащих оптимизации необходимо найти значения при которых функция будет минимальна/максимальна.



Линейная регрессия (без интерцепта)

- Выражение для моделируемой переменной: $\tilde{Y} = X \times w^T$
- Функция потерь $Penalty = \frac{1}{N}(\tilde{Y} - Y) \times (\tilde{Y} - Y)^T$, где Y реальные значения , $\tilde{Y}$ моделируемые значения.
- Нахождение производной для градиентного спуска:
  - Поскольку $Penalty = \frac{1}{N}\sum_i (X_i \times w^T - Y_i)^2$,

    $$\frac{\partial Penalty}{\partial w_j} = \frac{1}{N}\sum_i 2(X_i \times w^T - Y_i) X_j$$
  - В случае коэффициентов при факторах имеем:

    $$\nabla Penalty = 2(X \times w^T - Y)^T \times X$$
  - Для реализации использования нулевого коэффициента(интерцепта) можно в матрицу данных добавить столбец, значения которого равны 1. Тогда производная по этому коэффициенту будет равна

    $$\frac{\partial Penalty}{\partial w_0} = \sum_i 2(w_0 + X_i \times w^T - Y_i)^T = \sum 2(w_0 + X \times w^T - Y)$$

Линейная регрессия (с интерцептом)

- Выражение для моделируемой переменной: $\tilde{Y} = w_0 + X \times w^T$
- Функция потерь $Penalty = \frac{1}{N}(\tilde{Y} - Y) \times (\tilde{Y} - Y)^T$, где Y реальные значения , $\tilde{Y}$ моделируемые значения.
- Нахождение производной для градиентного спуска:
  - Поскольку $Penalty = \frac{1}{N}\sum_i (w_0 + X_i \times w^T - Y_i)^2$,

    $$\frac{\partial Penalty}{\partial w_j} = \frac{1}{N}\sum_i 2(w_0 + X_i \times w^T - Y_i) X_j$$
  - В случае коэффициентов при факторах имеем:

    $$\nabla Penalty = 2(w_0 + X \times w^T - Y)^T \times X$$
  - Для нулевого коэффициента(интерцепта)

    $$\frac{\partial Penalty}{\partial w_0} = \sum_i 2(w_0 + X_i \times w^T - Y_i)^T = \sum 2(w_0 + X \times w^T - Y)$$