

Домашнее задание №2 по курсу «Машинное обучение»: основы машинного обучения

Скавыш Максим

Задание 1

Предложите алгоритм для генерации случайной линейно разделимой выборки в двумерном пространстве.

Пусть алгоритм принимает на вход следующие величины:

N, P – размеры отрицательной и положительной выборок соответственно

$\min_x \leq \max_x, \quad \min_y \leq \max_y$ – допустимы границы координат точек x и y соответственно.

Алгоритм:

Впишем в прямоугольник круг с радиусом:

$$R = \min\left(\frac{(\max_x - \min_x)}{2}, \frac{(\max_y - \min_y)}{2}\right)$$

И возьмем ϕ – произвольный угол

Сгенерируем 2 случайных выборки размера N .

Элементы первой выборки должны удовлетворять условию: $0 < r < R$ где r – элемент

Элементы второй выборки должны удовлетворять условию: $\phi < \alpha < \phi + \pi$ где α – элемент первой выборки.

Получили выборку размерности $N: \{(r_i, \alpha_i) \mid i \in 1, \dots, N; 0 < r_i < R, \phi < \alpha_i < \phi + \pi\}$ и, перейдя к полярным координатам, получим пары $(x_i, y_i) \mid i \in 1, \dots, N$ объектов отрицательного класса

Аналогично для выборки $P: \{(r_i, \alpha_i) \mid i \in 1, \dots, P; 0 < r_i < R, \phi + \pi < \alpha_i < \phi + 2\pi\}$ и перейдя к полярным координатам, получим пары $(x_i, y_i) \mid i \in 1, \dots, P$ объектов положительного класса.

Получили линейно разделимую выборку с центром в начале координат. Сложив каждый элемент выборки с координатами центра вписанного в прямоугольник

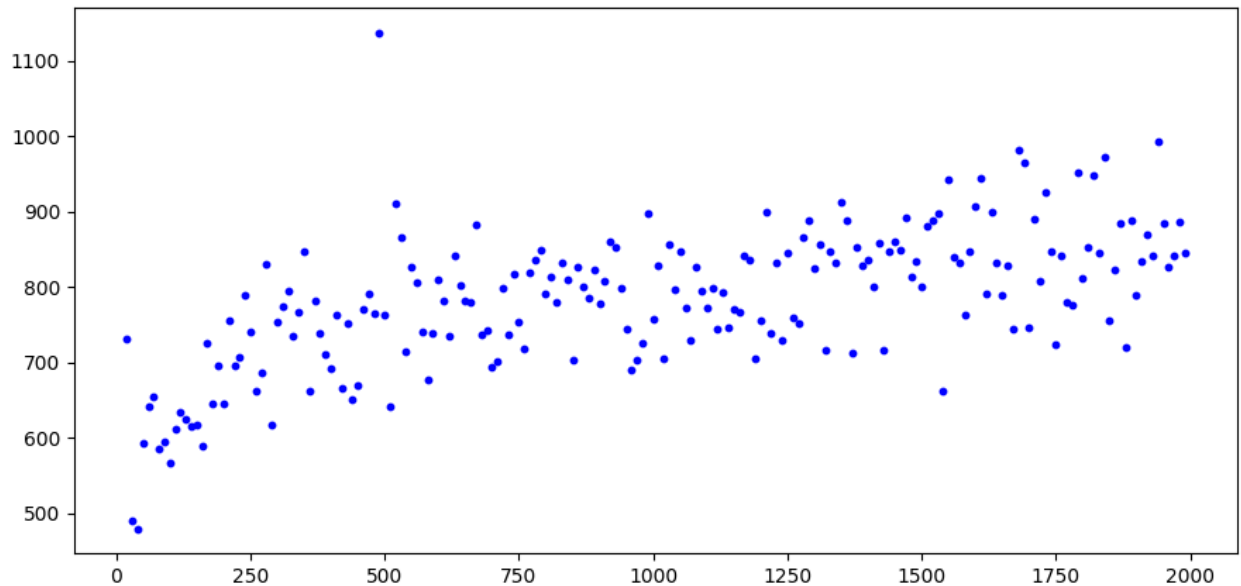
$\min_x \leq \max_x, \quad \min_y \leq \max_y$ получим необходимую нам линейно разделимую выборку:



Вероятность того что два объекта выборки имеют равные координаты равна нулю с вероятностью 1 так как точка имеет меру равную 0.

Задание 2

На данном графике представлена зависимость количества шагов в алгоритме batch perceptron (по вертикали) от размера выборки (горизонталь)



IPython: <https://github.com/MaksimSkavysh/Math/blob/master/ML-2017/hw2/batch-Copy1.ipynb>