1. Персептронный алгоритм получения линейных решающих правил

Персептронный алгоритм получения линейных решающих правил

Простейший методы получения линейных решающих функций на основе персептронных алгоритма обучения основывается на рекуррентном построении решающего правила путем коррекции ошибок.

Требуется найти W^T, Wn+1 для построения решающего правила $d(X)=W^T X+$ Wn+1 на основе использования конечных обучающих выборок .

Введем понятие расширенных векторов . Перейдем от размерности n к n+1 следующим образом:

$$\tilde{W} = \begin{pmatrix} W_1 \\ .. \\ W_n \\ W_{n+1} \end{pmatrix}$$

$$\tilde{X} = \begin{pmatrix} X_1 \\ .. \\ X_n \\ 1 \end{pmatrix}$$

Тогда наша система неравенств сводится к более простой задаче:

(*) $d(X) = W^T X > 0$ (или <) $x \in \mathbf{X}1$ ($x \in \mathbf{X}2$)

Персептронный алгоритм основан на последовательном просмотре обучающей выборки:

$$\underbrace{X_1, \dots X_{N1}}_{\mathbf{X}_1}\ \underbrace{\dots\dots\dots X_N}_{\mathbf{X}_2}$$

$$N = N_1 + N_2$$

Процесс обучения заключается в том, что мы циклически просматриваем выборку и подставляем получаемое значение в W в (*), и на каждом шаге просмотра производим или не производим коррекцию весового вектора.

$$1.\ \widetilde{W}_{n+1} = \widetilde{W}_{n},\ \text{если} \quad \begin{cases} x \in X_1, & \widetilde{W}_n^T \tilde{X} > 0 \\ x \in X_2, & \widetilde{W}_n^T \tilde{X} < 0 \end{cases}$$

В этом случае получен правильный ответ при классификации текущего вектора X.

$$2.\ \widetilde{W}_{n+1} = \widetilde{W}_{n} + C\overline{X},\ \text{если}\ \widetilde{W}_n^T \tilde{X} < 0,\ x \in X_1$$

$$\widetilde{W}_{n+1} = \widetilde{W}_{n} - C\overline{X},\ \text{если}\ \widetilde{W}_n^T \tilde{X} > 0,\ x \in X_2$$

Этот случай соответствует ошибочной классификации и соответственно производится коррекция весового вектора (должно быть C>0)

Эта процедура и является процедурой обучения персептронного типа.

Пусть мы имеем величину весового вектора после коррекции:

$$\tilde{W}_{n+1} = \tilde{W}_n + C\overline{X},$$

Подставим новый весовой вектор в выражение для решающей функции:

$$\tilde{W}_{n+1}^T \tilde{X} = (\tilde{W}_n + C\overline{X})^T \tilde{X} = \tilde{W}_n^T \tilde{X} + C\tilde{X}^T \tilde{X} = \tilde{W}_n^T \tilde{X} + C\|\tilde{X}\|^2 ,$$

Видно, что значение весовой функции увеличилось на положительную величину $C\|\tilde{X}\|^2$ то есть мы продвинулись к правильному решению.

Показано, что если решение существует, то алгоритм сходится за конечное число шагов.Различные варианты выбора коэффициента C позволяют улучшить данный алгоритм:

1. C – константа . Скорость сходимости может быть мала.
2. C = Cn = var(n)

Попробуем менять C на каждом шагу так .чтобы сразу получить на текущем векторе правильное решение. Здесь можно использовать такой выбор C

$\tilde{W}_n^T \tilde{X}$ + Cn$\|\tilde{X}\|^2$ > 0 , отсюда следует :

$$C_n > \frac{|\tilde{W}_n^T \tilde{X}|}{\|\tilde{X}\|^2}$$

Рассмотренный алгоритм появился на основе интуитивных соображений при разработке моделей работы головного мозга человека при решении задач обучения.

Дальше мы рассмотрим более формальный подход .