

Angewandte Mathematik

Kleines Projekt: Neuronale Netze - Adaline

Prof. Dr. Sarah Brockhaus

In diesem Projekt betrachten wir das *Adaline* (ADAPtive LInear NEuron), ein künstliches Neuron. Das Adaline besitzt die Parameter $w_0, w_1, \dots, w_n \in \mathbb{R}$ und stellt eine mathematische Funktion

$$g : \mathbb{R}^n \rightarrow \{-1, 1\}, \quad g(\mathbf{x}) = \alpha(f(\mathbf{x}))$$

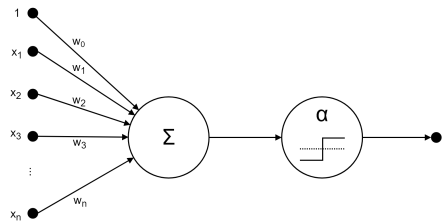
mit

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(\mathbf{x}) = w_0 \cdot 1 + \sum_{j=1}^n w_j \cdot x_j$$

und

$$\alpha : \mathbb{R} \rightarrow \{-1, 1\}, \quad \alpha(x) = \begin{cases} -1, & \text{falls } x \leq 0, \\ 1, & \text{falls } x > 0 \end{cases}$$

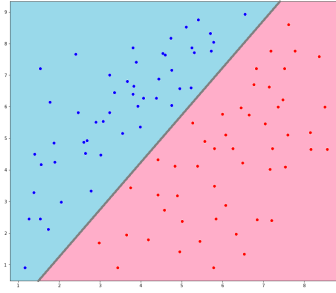
dar. Eine schematische Darstellung eines Adaline kann der folgenden Abbildung entnommen werden.



Das Adaline kann für die *Klassifikation* verwendet werden, d.h. es ordnet einen Datenpunkt $\mathbf{x} \in \mathbb{R}^n$ eine der beiden Klassen positiv 1 oder negativ -1 zu. Durch Anpassung der Parameter w_i an einen gegebenen Datensatz mit k Datenpunkten

$$\mathcal{D} = \{(\mathbf{x}^{(i)}, y^{(i)}) | 1 \leq i \leq k\} \subseteq \mathbb{R} \times \{-1, 1\}$$

dem Training, kann das Adaline eine solche Klassifikationsaufgabe lernen. Die folgende Abbildung zeigt exemplarisch eine Klassifikation von Datenpunkten $\mathbf{x}^{(i)} \in \mathbb{R}^2$ nach positiv 1 (blau) und negativ -1 (rot).



Für das Training wählt man eine Fehlerfunktion, wie z.B. den quadrierten Fehler $E^{(i)}$ für den Datenpunkt $(\mathbf{x}^{(i)}, y^{(i)})$ gegeben durch

$$E^{(i)} = \frac{1}{2} \left(y^{(i)} - f(\mathbf{x}^{(i)}) \right)^2$$

und passt iterativ über alle Datenpunkte $(\mathbf{x}^{(i)}, y^{(i)})$ für $i \in \{1, \dots, k\}$ des Datensatzes (innere Schleife im Code) und mehrere Durchläufe hinweg (äußerste Schleife im Code) ausgehend von einer (oft zufälligen) initialen Wahl, die Parameter w_0, w_1, \dots, w_n mit Hilfe der Regel

$$w_j \leftarrow w_j - \eta \cdot \frac{d}{dw_j} E^{(i)}$$

an (innerste Schleife im Code).

1. Bestimmen Sie die allgemeine Ableitung des quadratischen Fehlers bzgl. des i -ten Datenpunkts nach allen Parametern w_j für $j \in \{0, \dots, n\}$.
Es genügt, wenn Sie die Fälle $j = 0$ und $j > 0$ explizit betrachten.
2. Implementieren Sie den Trainingsalgorithmus des Adalines.
3. Implementieren Sie einen Algorithmus, der bestimmt wie viel Prozent der Daten korrekt klassifiziert wurden.
Achtung: Zum Lernen wird Funktion f angepasst, jedoch wird für die tatsächliche Anwendung die volle Funktion g verwendet!
4. Plotten Sie den Datensatz `adaline.csv`. Was fällt Ihnen auf?
5. Testen und Evaluieren Sie das Adaline bzgl. des gegebenen Datensatzes `adaline.csv`.
6. Konstruieren und plotten Sie einen Datensatz, bei dem Ihnen als Mensch eine Klassifikation der Datenpunkte leicht fällt, wo das Adaline jedoch versagt. Warum versagt die Methode in diesem Fall? Beschreiben Sie einen möglichen Ansatz zur Lösung des Problems (z.B. Transformation der Daten, Anpassung der Methode, ...).
7. *Pflicht für 4-Gruppen; optional für 3-er Gruppen:* Plotten Sie den Datensatz `evasion.csv` und wenden Sie Ihre Implementierung des Adaline darauf an. Was fällt Ihnen auf?
8. *Bonusaufgabe:* Suchen Sie sich einen frei verfügbaren Datensatz mit einem Klassifikationsproblem und wenden Sie Ihre Implementierung des Adaline darauf an. Was fällt Ihnen auf?