

Übungsblatt 5

Alle Aufgaben unten sind für ihre Eigenimplementierungen – ggf. in Anlehnung an den Code im Buch gemeint – es geht nicht darum z.B. die `sklearn` zu verwenden.

1. Ein einfaches neuronales Netz mit dem Tangens Hyperbolicus



Auf den Seiten 190 – 195 wird erklärt wie das Netzwerk für die Regression mit einem Output in Python umgesetzt wird. In dieser Aufgaben geht es darum dies nachzuvollziehen und eine eigene Version zu implementieren die den Tangens Hyperbolicus als Aktivierungsfunktion nutzt. Gehen Sie dabei in folgenden Schritten vor.

1. Plotten Sie die Sigmoid-Funktion und den Tangens Hyperbolicus im Intervall von $[-5, 5]$. In der Datei `simpleMLP` finden Sie eine Anregung für die Umsetzung der Sigmoid-Funktion.
2. Vergleichen Sie den Steigung der beiden Funktionen und überlegen Sie, wie schnell die jeweilige Funktion wohl ggf. in die Sättigung geraten kann.
3. Nehmen Sie ein Blatt Papier und berechnen Sie für den Tangens Hyperbolicus die nötigen Ableitungen, analog zu den Seiten 186, 187

$$\begin{aligned}\frac{\partial y}{\partial w_{1,j}^{(3)}} &= \dots \\ \frac{\partial y}{\partial w_{j,k}^{(2)}} &= \dots \\ \frac{\partial y}{\partial w_{k,l}^{(1)}} &= \dots\end{aligned}$$

4. Laden Sie die Datei `simpleMLP.py` und speichern Sie diese unter einem neuen Namen.
5. Ändern Sie die Methoden `_calOut` und `train` so ab, dass dort nun der Tangens Hyperbolicus und die von Ihnen berechneten Ableitungen verwendet werden.
6. Kleinere Startgewichte sorgen, dafür das dieses Verfahren mit den Tangens Hyperbolicus weniger schnell in eine Sättigung geht. Skalieren Sie die Gewichte in `_initWeights` mit 0.5.
7. Da der Tangens Hyperbolicus empfindlicher ist, ändern Sie das η um sich mit kleineren Schritte vorwärts zu tasten. Tun Sie dies z.B. wie folgt:

```
def fit(self,X,Y,eta=0.25,maxIter=200,vareps=10**-3,scale=True):
```
8. Wenn Sie das fertig haben, festen Sie ihren Code auf dem Beispiel der letzten Folien bzw. Seite 193 und 194.
9. Welche Werte erhalten Sie für $\eta = 0.25$? Ab welchem η etwa beginnt das Verfahren für dieses Beispiel instabil zu werden?