

Neuronale Netze (SS 2002)

8. Übungsblatt

Abgabe: Donnerstag, 13.6.02, 12⁰⁰ Uhr, Briefkasten ‚Neuronale Netze‘ im 4. Stock des AVZ

1. (5 Punkte) Beschreiben Sie umgangssprachlich, warum die Wahl der Schrittweite bei Gradientenverfahren sehr kritisch ist, und eine globale Schrittweite für alle Gewichte und alle Trainingsschritte ungeeignet sein kann.

Zeigen sie, daß der Effekt der Verstärkung der Schrittweite durch einen Momentum-Term in Hochplateaus beschränkt ist; gehen Sie dazu wie folgt vor: nehmen Sie an, die zu optimierende Funktion ist nur eindimensional $f : \mathbb{R} \rightarrow \mathbb{R}$ und im betrachteten Bereich (Hochplateau) angenähert linear, d.h. $f(x) = ax + b$. Die Schrittweite ohne Momentum-Term wäre also $\eta \cdot a$. Rechnen Sie die Schrittweite mit Momentum-Term nach der k -ten Iteration aus und zeigen Sie, daß diese durch ein Vielfaches von $\eta \cdot a$ beschränkt werden kann.

2. (5 Punkte) Sie möchten eine unbekannte Gesetzmäßigkeit anhand von Beispielen lernen (z.B. anhand des Gewichtes und der Art des Geräusches beim Schütteln eines Überraschungseis entscheiden, ob im Ei eine Figur steckt) und haben folgende Funktionenklassen, mit denen Sie lernen könnten, zur Verfügung:

- Perzeptrons,
- sigmoide neuronale feedforward-Netze mit einer verborgenen Schicht und beliebig vielen Neuronen,
- Polynome vom Grad maximal 20,
- alle stetigen Funktionen.

Welche Funktionenklasse nehmen Sie, warum, und wie gehen Sie dann vor?

3. (5 Punkte) Testen Sie verschiedene Trainingsverfahren zum Training für die Menge spirals_fies.pat und dokumentieren Sie Ihre Ergebnisse. Starten Sie mit relativ großen Netzen, und beachten Sie, daß die Ausgabe $-1/1$ -wertig ist (also tanh als Aktivierungsfunktion).
4. (5 Punkte) Trainieren Sie jeweils ein Netz auf die Zeitreihen daten1, daten2, daten3, daten4 (Patterndateien für ein Zeitfenster 4 stehen zur Verfügung; die Daten sind diesmal nicht trendbehaftet ;) und dokumentieren Sie das Ergebnis. Sie können (müssen aber nicht) auch Pruning verwenden. Wie schätzen Sie jeweils die Güte des sich ergebenden Netz ein, ohne dazu einen Testfehler zu berechnen? Ist die zugrundeliegende Gesetzmäßigkeit erkannt worden?