Forschungsmodul: Komplexe Systeme

Bericht zur Vorlesung vom 22. November 2007 von Jan-Philip Gehrcke

Künstliche neuronale Netze - Attraktor-/Mehrschichtnetzwerke

Attraktornetzwerk

Ein Attraktornetzwerk besteht aus untereinander voll vernetzten künstlichen Neuronen. Es ist jedes Neuron mit jedem anderen vernetzt. Diese Form des Netzwerkes nennt man auch Hopfield-Netz (John Hopfield machte dieses Modell des künstlichen neuronalen Netzes bekannt) oder Assoziativspeicher (Begriff wird später erläutert).

Der Gesamtzustand eines solchen Netzes lässt sich mathematisch einfach beschreiben. Jedes der N Neuronen befindet sich immer in einem Zustand $S \in \{+1,-1\}$ in Abhängigkeit seiner Eingänge. Die Neuronen sind im Allgemeinen untereinander unterschiedlich stark gekoppelt. Es gibt also N^2 Kopplungskonstanten $J_{ij} \in \mathbb{R}$.

Nimmt man bei Kenntnis der Kopplungen einen Anfangszustand $S_i(0) \in \{+1,-1\} \; mit\, i = 1,..., N$ zur Zeit oder zum Taktschritt $t = 0$ an, lässt sich taktweise (z.B. bei Betrachtung von synchronen Update-Schritten) eine Veränderung des Systems verfolgen. Ein Update-Schritt für das $S_i - te$ Neuron sieht dann so aus:

$$S_i(t+1) = sign(\sum_j J_{ij} S_j(t))$$

Gleichen sich für große t die linke und rechte Seite dieser Gleichung, ändert sich am Zustand mit fortschreitender Taktzahl nichts mehr. Dann spricht man von einem Fixpunkt des Systems. Ist das Netzwerk von einem Anfangszustand, der keinem Fixpunkt entspricht, durch die interne Dynamik zu einem solchen hinkonvergiert, nennt man diesen einen attraktiven Fixpunkt oder auch Attraktor. Besitzt ein Netzwerk Attraktoren, lassen sich somit Informationen speichern. Ein solches Netz kann aus einer bestimmten Vorgabe (kleinen Menge an Daten) ganze Datensätze rekonstruieren (z.B. Bilder aus Bildausschnitten). Bedingung hierfür ist, dass der Anfangszustand zum gewünschten Attraktor hinführt. Der Begriff des Assoziativnetzwerkes ist somit geprägt.

Die Fixpunkte des Netzes sind praktisch die “Bilder im Gedächtnis” (auch im Transfer zum menschlichen Gehirn). Ziel soll es nun sein, ein Netzwerk zu konstruieren, das eine vorgegebene Anzahl P von bestimmten Mustern $\vec{x}^{\nu}\,;\ \nu = 1,\dots,P$ als Fixpunkte besitzen soll. Die Vektoren $\vec{x}^{\nu}$ haben jeweils N Komponenten, die den Zustand S des Systems klassifizieren (siehe oben). Um das Netz mit den geforderten Fixpunkten aufzubauen, müssen die Kopplungen J_{ij} gefunden werden, für die gilt:

$$x_i^{\nu} = sign\Big(\sum_j J_{ij}\, x_j^{\nu}\Big)$$

Somit wäre $\vec{x}^{\nu}$ ein Fixpunkt. Ob dieser auch attraktiv ist, kann und soll an dieser Stelle nicht betrachtet werden.

Versuchsweise sollen Hebb'sche Kopplungen verwendet werden. Das mathematische Problem ähnelt dann stark dem eines Perzeptrons, welches bereits diskutiert wurde. Rechnet man hier analog mit den Summentermen und macht eine Wahrscheinlichkeitsbetrachtung der vorkommenden Größen, erhält man die obige Bedingung in statistischer Form: Mit $\alpha = \frac{P}{N}$ ergibt sich eine Gaußverteilung mit dem Mittelwert $1+\alpha$ und der Varianz $\sqrt{(\alpha)}$. Für alle in diesem Wahrscheinlichkeitsmodell vorkommende Werte größer als Null ist obige Bedingung an die Kopplungen erfüllt. Nur wenn man nun $\alpha \ll 1$ wählt, darf man von 100% Wahrscheinlichkeit im positiven Wertebereich ausgehen. Das erklärt, warum in einem Attraktornetzwerk deutlich weniger Muster (Fixpunkte) gespeichert werden können, als es Neuronen gibt.

Mehrschichtnetzwerk

Für komplexe Input-Output-Problemstellungen mit vielen Eingängen, die aber nur einer ja/nein-Entscheidung am Ausgang bedürfen, eignen sicht oft mehrschichtige neuronale Netze. Diese gehören, wie bereits angesprochen wurde, zu den feedforward-Netzen. Aufgrund ihrer Struktur und der praktisch beliebigen Anzahl innerer Schichten sind sie in der Lage, komplexe Logiken zu reproduzieren. Dies soll am Beispiel der XOR-Verknüpfung gezeigt werden.

XOR-Verknüpfung:

Gibt man zwei Eingänge x und y vor, die die Zustände -1 oder $+1$ annehmen können, dann ist der Ausgang $f(x,y)$, den die XOR-Verknüpfung zwischen diesen Eingängen vorgibt, folgendermaßen zu beschreiben:

$$f(x,y) = sign(sign(x+y+1) \; - \; sign(x+y-1))$$

Anhand dieser mathematischen Struktur erkennt man ein Perzeptron mit einer inneren Schicht, wobei die einzelnen teilhabenden Knoten unterschiedliche Schwellen θ besitzen (1 oder -1). Die Kopplungskonstanten sind ebenfalls entweder 1 oder -1 . Somit ist das XOR-Problem durch fünf Knoten lösbar: Zwei Eingänge, zwei Knoten in der inneren Schicht und ein Ausgang.

Mit jeder inneren Schicht können neue Klassifikationsebenen im Musterraum eingeführt werden. Im mathematischen Ausdruck können Trennebenen durch Signum-Funktionen geschaffen werden.

Der Backpropagation-Algorithmus

Der sogenannte Backpropagation-Algorithmus ist ein Lernalgorithmus für ein mehrschichtiges Netzwerk. Die einzelnen Kopplungen zwischen den Schichten werden während des Lernvorgangs fortlaufend angepasst. Genauso werden die

Schwellen der Knoten variiert. Der Algorithmus benötigt dabei den gewünschten Output F des Netzes bei vorgegebenem Input. Die quadratische Abweichung $E=\frac{1}{2}(F-f)^2$ zwischen F und dem aktuellen, fehlerbehafteten Output f wird von Lernschritt zu Lernschritt versucht zu minimieren. Zu diesem Zweck werden die partiellen Ableitungen der Größe E nach allen freien Parametern gebildet. Diese Parameter werden dann in Richtung des negativen Gradienten von E variiert. Aufgrund dieser Rückkopplung trägt der Algorithmus den Namen “Backpropagation”.