

Blatt 07 - RBF-Netze und Klassifikation

Abgabe: Bis Mittwoch, 17. Dezember 2014, 10:00 Uhr

An: Tobias Rothenberger, <rothenb@informatik.uni-frankfurt.de>

Aufgaben: Aufgabe 7.1 Aufgabe 7.2 Aufgabe 7.3

Aufgabe 7.1 Funktionsapproximation

Die Datei `aprx_noised.dt` enthält verrauschte Funktionswerte einer unbekannten Funktion. Trainieren Sie nun ein RBF-Netz darauf, diese Funktion zu lernen und versuchen Sie herauszufinden,

1. welchen Einfluss die **Anzahl** der RBF-Neuronen
2. und welchen Einfluss die **Breite** σ der RBF-Neuronen

auf den Fehler des RBF-Netzes (Differenz der Ausgabe zur Lehrervorgabe) hat.

nach oben

Aufgabe 7.2

Die Dateien `TrainingSet.dt`, `ValidationSet.dt` und `TestSet.dt` enthalten jeweils eine Trainings-, Validierungs- und Testmenge bestehend aus 2-dimensionalen Datenpunkten, die in 16 Klassen (1-aus-16 Kodierung) eingeteilt sind. Jede Zeile in diesen Dateien entspricht einem Muster, wobei die ersten 2 Werte einer Zeile die Eingaben und die letzten 16 Komponenten einer Zeile die

Klassenzugehörigkeit zu den Eingaben angeben.

Trainieren Sie nun ein RBF-Netz darauf, eine Klassifikation der in den Dateien `TrainingSet.dt`, `ValidationSet.dt` und `TestSet.dt` vorhandenen Muster zu lernen. Die Dateien sollen dabei entsprechend Ihrer Bezeichnung als Trainings-, Validierungs- und Testmenge während des Trainings verwendet werden.

Versuchen Sie dann herauszufinden,

1. welchen Einfluß die Anzahl der RBF-Neuronen,
2. welchen Einfluß die Breite σ der RBF-Neuronen
3. und welchen Einfluß die in den Folien erwähnte Normierung der Ausgaben der RBF-Schicht auf den Fehler des RBF-Netzes (Differenz der Ausgabe zur Lehrervorgabe) und die per Winner-Take-All (s. Folien) durchgeführte Klassifizierung hat.

Erstellen Sie zur Beantwortung dieser Fragen eine repräsentative Auswahl an Plots der Entscheidungsgebiete, auf dem die durch das Netz vorgenommene Klassenaufteilung des Eingaberaums und die Datenpunkte jeder Klasse zu erkennen sind.

•

nach oben

Aufgabe 7.3

- a. Trainieren Sie ein RBF-Netz darauf, eine Klassifikation des Card-Datensatzes (`card1.dt`) zu lernen. Versuchen Sie die Anzahl der Neuronen und die Breite σ der RBF-Neuronen so anzupassen, daß Sie ein möglichst

gutes Klassifikationsergebnis erzielen.

Geben Sie anschließend den Generalisierungsfehler als Prozentsatz fehlerhaft klassifizierter Muster Ihres Netzes an.

Weisen Sie durch Ausgabe der Trainingsstatistik (Gradientenlänge und Fehlerwerte auf der Trainings-/Validierungs-/Testmenge) während des Trainings nach, daß das Netz ausreichend (d.h. bis zum Erreichen eines lokalen Minimums) trainiert wurde und kein (oder nur ein sehr geringer) Overfitting-Effekt vorliegt.

- b. Nehmen Sie an, Sie hätten ein System, das Ihnen als Klassifikation für ein Muster x des Card-Datensatzes immer die selbe Klassenzugehörigkeit ausgeben würde, ungeachtet der Tatsache, wie x beschaffen ist. Welchen Prozentsatz fehlerhaft klassifizierter Muster hätte dieses System im besten Fall auf der Testmenge?
Diskutieren Sie vor diesem Hintergrund die Fehlerrate Ihres eigenen Systems.

•

nach oben