Quantitative Analyse neuronaler Netzwerke als universelle Funktionsapproximatoren

Stephan Trenn

Diplomarbeit - Zwischenverteidigung

Ilmenau, 4. Juli 2006

Gliederung

Gliederung

Grundstruktur des Multi-Layer-Perceptrons (MLP)

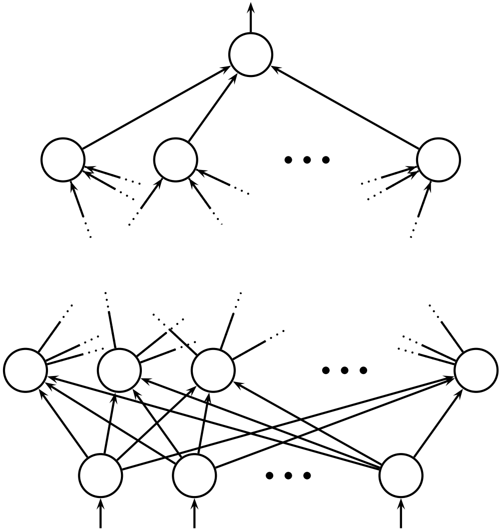
Ausgangsschicht

h-te Hidden-Schicht

⋮

1. Hidden-Schicht

Eingangsschicht



MLP als mathematisches Objekt

Definition MLP

Ein MLP ist ein Quadrupel $(h, \mathbf{n}, \sigma, \mathbf{P})$.

- h: Anzahl Hidden-Schichten
- $\mathbf{n} = (n_0, n_1, \ldots, n_h)$: Anzahl Neuronen pro Schicht
  - n_0: Anzahl Eingänge
  - $n_1, \ldots, n_h$: Anzahl Neuronen in der jeweiligen Hidden-Schicht
- $\sigma : \mathbb{R} \to \mathbb{R}$: Aktivierungsfunktion innerhalb der Hidden-Neuronen
- $\mathbf{P} = (\mathfrak{w}^1, \mathfrak{w}^2, \ldots, \mathfrak{w}^h, \mathbf{w}^y)$: variable Netzwerkparameter
  - Kantengewichte
  - Schwellenwerte

MLP als mathematisches Objekt

Definition MLP

Ein MLP ist ein Quadrupel $(h, \mathbf{n}, \sigma, \mathbf{P})$.

- h: Anzahl Hidden-Schichten
- $\mathbf{n} = (n_0, n_1, \ldots, n_h)$: Anzahl Neuronen pro Schicht
  - n_0: Anzahl Eingänge
  - $n_1, \ldots, n_h$: Anzahl Neuronen in der jeweiligen Hidden-Schicht
- $\sigma : \mathbb{R} \to \mathbb{R}$: Aktivierungsfunktion innerhalb der Hidden-Neuronen
- $\mathbf{P} = (\mathfrak{w}^1, \mathfrak{w}^2, \ldots, \mathfrak{w}^h, \mathbf{w}^y)$: variable Netzwerkparameter
  - Kantengewichte
  - Schwellenwerte

MLP als mathematisches Objekt

Definition MLP

Ein MLP ist ein Quadrupel $(h, \mathbf{n}, \sigma, \mathbf{P})$.

- h: Anzahl Hidden-Schichten
- $\mathbf{n} = (n_0, n_1, \ldots, n_h)$: Anzahl Neuronen pro Schicht
  - n_0: Anzahl Eingänge
  - $n_1, \ldots, n_h$: Anzahl Neuronen in der jeweiligen Hidden-Schicht
- $\sigma : \mathbb{R} \to \mathbb{R}$: Aktivierungsfunktion innerhalb der Hidden-Neuronen
- $\mathbf{P} = (\mathfrak{w}^1, \mathfrak{w}^2, \ldots, \mathfrak{w}^h, \mathbf{w}^y)$: variable Netzwerkparameter
  - Kantengewichte
  - Schwellenwerte

MLP als mathematisches Objekt

Definition MLP

Ein MLP ist ein Quadrupel $(h, \mathbf{n}, \sigma, \mathbf{P})$.

- h: Anzahl Hidden-Schichten
- $\mathbf{n} = (n_0, n_1, \dots, n_h)$: Anzahl Neuronen pro Schicht
  - n_0: Anzahl Eingänge
  - $n_1, \dots, n_h$: Anzahl Neuronen in der jeweiligen Hidden-Schicht
- $\sigma : \mathbb{R} \to \mathbb{R}$: Aktivierungsfunktion innerhalb der Hidden-Neuronen
- $\mathbf{P} = (\mathfrak{w}^1, \mathfrak{w}^2, \dots, \mathfrak{w}^h, \mathbf{w}^y)$: variable Netzwerkparameter
  - Kantengewichte
  - Schwellenwerte

MLP als mathematisches Objekt

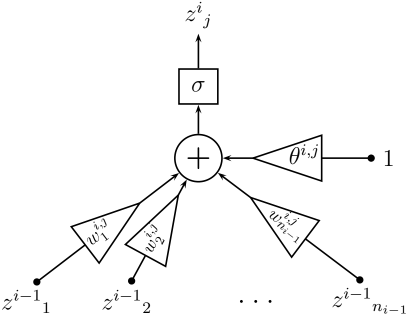
Definition MLP

Ein MLP ist ein Quadrupel $(h, \mathbf{n}, \sigma, \mathbf{P})$.

- h: Anzahl Hidden-Schichten
- $\mathbf{n} = (n_0, n_1, \ldots, n_h)$: Anzahl Neuronen pro Schicht
  - n_0: Anzahl Eingänge
  - $n_1, \ldots, n_h$: Anzahl Neuronen in der jeweiligen Hidden-Schicht
- $\sigma : \mathbb{R} \to \mathbb{R}$: Aktivierungsfunktion innerhalb der Hidden-Neuronen
- $\mathbf{P} = (\mathfrak{w}^1, \mathfrak{w}^2, \ldots, \mathfrak{w}^h, \mathbf{w}^y)$: variable Netzwerkparameter
  - Kantengewichte
  - Schwellenwerte

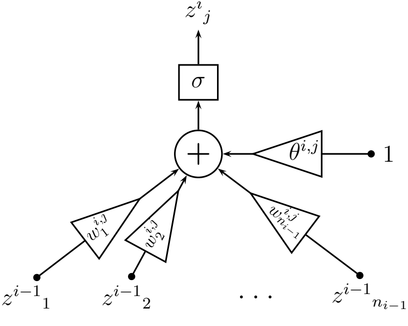
MLP-Funktion

Hidden-Schicht-Neuron:



MLP-Funktion

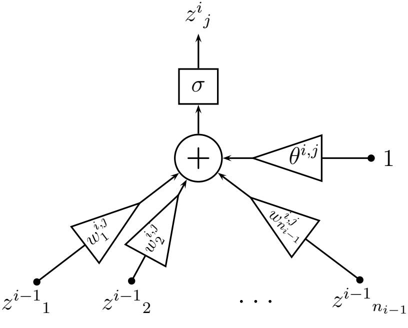
Hidden-Schicht-Neuron: Ausgangsneuron:

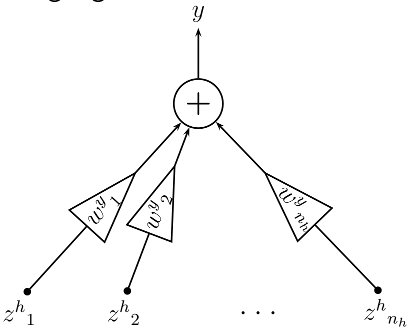


MLP-Funktion

Hidden-Schicht-Neuron:

Ausgangsneuron:





MLP-Funktion für fest gewählte Parameter

$$f_{\mathrm{MLP}} : \mathbb{R}^{n_0} \longrightarrow \mathbb{R}, \quad \mathbf{x} = (x_1, x_2, \ldots, x_{n_0}) \mapsto y$$

Allgemeine Fragestellung

Problemstellung

Gegeben: Zielfunktion $f : \mathbb{R}^{n_0} \to \mathbb{R}$.

Gesucht: Ein MLP mit $f_{\text{MLP}} \approx f$.

Allgemeine Fragestellung

Problemstellung

Gegeben: Zielfunktion $f : \mathbb{R}^{n_0} \to \mathbb{R}$.

Gesucht: Ein MLP mit $f_{\text{MLP}} \approx f$.

In diesem Zusammenhang muss geklärt werden:

- Was bedeutet $f_{\text{MLP}} \approx f$ formal?
- Welche Art von Funktionen soll approximiert werden?
- Kann man mit fester Anzahl von Neuronen beliebig gut approximieren?
- Wie viele Neuronen braucht man mindestens, um mit einer vorgegeben Güte zu approximieren?

Allgemeine Fragestellung

Problemstellung

Gegeben: Zielfunktion $f : \mathbb{R}^{n_0} \to \mathbb{R}$.

Gesucht: Ein MLP mit $f_{\text{MLP}} \approx f$.

In diesem Zusammenhang muss geklärt werden:

- Was bedeutet $f_{\text{MLP}} \approx f$ formal?
- Welche Art von Funktionen soll approximiert werden?
- Kann man mit fester Anzahl von Neuronen beliebig gut approximieren?
- Wie viele Neuronen braucht man mindestens, um mit einer vorgegeben Güte zu approximieren?

Allgemeine Fragestellung

Problemstellung

Gegeben: Zielfunktion $f : \mathbb{R}^{n_0} \to \mathbb{R}$.

Gesucht: Ein MLP mit $f_{\text{MLP}} \approx f$.

In diesem Zusammenhang muss geklärt werden:

- Was bedeutet $f_{\text{MLP}} \approx f$ formal?
- Welche Art von Funktionen soll approximiert werden?
- Kann man mit fester Anzahl von Neuronen beliebig gut approximieren?
- Wie viele Neuronen braucht man mindestens, um mit einer vorgegeben Güte zu approximieren?

Allgemeine Fragestellung

Problemstellung

Gegeben: Zielfunktion $f : \mathbb{R}^{n_0} \to \mathbb{R}$.

Gesucht: Ein MLP mit $f_{\mathrm{MLP}} \approx f$.

In diesem Zusammenhang muss geklärt werden:

- Was bedeutet $f_{\mathrm{MLP}} \approx f$ formal?
- Welche Art von Funktionen soll approximiert werden?
- Kann man mit fester Anzahl von Neuronen beliebig gut approximieren?
- Wie viele Neuronen braucht man mindestens, um mit einer vorgegeben Güte zu approximieren?

Allgemeine Fragestellung

Problemstellung

Gegeben: Zielfunktion $f : \mathbb{R}^{n_0} \to \mathbb{R}$.

Gesucht: Ein MLP mit $f_{\text{MLP}} \approx f$.

In diesem Zusammenhang muss geklärt werden:

- Was bedeutet $f_{\text{MLP}} \approx f$ formal?
- Welche Art von Funktionen soll approximiert werden?
- Kann man mit fester Anzahl von Neuronen beliebig gut approximieren?
- Wie viele Neuronen braucht man mindestens, um mit einer vorgegeben Güte zu approximieren?

Gliederung

Relevanter Funktionenraum

Einschränkungen

- Zielfunktion: mindestens stetig
- Definitionsbereich: Einschränkung auf kompakte Menge $K \subseteq \mathbb{R}^{n_0}$, z.B. $K = [-1, 1]^{n_0}$, d.h. $x_1, x_2, \ldots, x_{n_0} \in [-1, 1]$

Dieser Funktionenraum sei mit $C(K; \mathbb{R})$ bezeichnet.

Relevanter Funktionenraum

Einschränkungen

- Zielfunktion: mindestens stetig
- Definitionsbereich: Einschränkung auf kompakte Menge $K \subseteq \mathbb{R}^{n_0}$, z.B. $K = [-1,1]^{n_0}$, d.h. $x_1, x_2, \ldots, x_{n_0} \in [-1,1]$

Dieser Funktionenraum sei mit $C(K;\mathbb{R})$ bezeichnet.

Im $C(K;\mathbb{R})$ kann man folgenden **Abstand** zwischen zwei Funktionen f_1 und f_2 definieren:

$$\|f_1 - f_2\| = \max_{\mathbf{x} \in K} \left| f_1(\mathbf{x}) - f_2(\mathbf{x}) \right|$$

Dichtheit

Definition

- $M \subseteq C(K;\mathbb{R})$ heißt **dicht** in $C(K;\mathbb{R})$
  $:\Leftrightarrow \quad \forall\, \varepsilon > 0\ \forall\, f \in C(K;\mathbb{R})\ \exists\, f_M \in M$:
  $$\|f - f_M\| < \varepsilon.$$
- $M \subseteq C(K;\mathbb{R})$ heißt **ε-dicht** in $C(K;\mathbb{R})$
  $:\Leftrightarrow \quad \forall\, f \in C(K;\mathbb{R})\ \exists\, f_M \in M$:
  $$\|f - f_M\| < \varepsilon.$$

Gliederung

Generelle Approximationsfähigkeit von MLPs

Positives Resultat:

Theorem

Falls σ stetig und kein Polynom ist, dann ist die Menge aller MLP-Funktionen von MLPs mit einer Hidden-Schicht dicht in $C(K;\mathbb{R})$.

Generelle Approximationsfähigkeit von MLPs

Positives Resultat:

Theorem

Falls σ stetig und kein Polynom ist, dann ist die Menge aller MLP-Funktionen von MLPs mit einer Hidden-Schicht dicht in $C(K;\mathbb{R})$.

Problem dieses Resultats:

Struktur des approximierenden MLPs hängt von der Zielfunktion ab, d.h. keine allgemeingültige Schranke für die Anzahl der benötigten Hidden-Neuronen!

Approximationsfähigkeit bei fester MLP-Struktur

Negatives Resultat:

Vermutung

Die MLP-Funktionen von MLPs mit *einer Hidden-Schicht* und *fester Neuronen-Anzahl* liegen **nicht** ε-dicht in $C(K;\mathbb{R})$.

Approximationsfähigkeit bei fester MLP-Struktur

Negatives Resultat:

Vermutung

Die MLP-Funktionen von MLPs mit *einer Hidden-Schicht* und *fester Neuronen-Anzahl* liegen **nicht** ε-dicht in $C(K;\mathbb{R})$.

Aber:

Theorem

Es gibt ein MLP mit **zwei** *Hidden-Schichten, einer speziell gewählten Aktivierungsfunktion und einer festen Anzahl von Hidden-Neuronen, welches jede stetige Funktion beliebig gut approximieren kann.*

Approximationsfähigkeit bei fester MLP-Struktur

Negatives Resultat:

Vermutung

Die MLP-Funktionen von MLPs mit *einer Hidden-Schicht* und *fester Neuronen-Anzahl* liegen **nicht** ε-dicht in $C(K;\mathbb{R})$.

Aber:

Theorem

Es gibt ein MLP mit **zwei** *Hidden-Schichten, einer speziell gewählten Aktivierungsfunktion und einer festen Anzahl von Hidden-Neuronen, welches jede stetige Funktion beliebig gut approximieren kann.*

Problem: Die Aktivierungsfunktion ist **nicht realisierbar**.

Zu hohe Zielstellung?

Problem

Stetige Funktionen als Zielfunktionen noch zu allgemein!

Zu hohe Zielstellung?

Problem

Stetige Funktionen als Zielfunktionen noch zu allgemein!

Mögliche Verschärfungen:

- Differenzierbarkeit der Zielfunktionen
- Beschränktheit der Zielfunktion und deren Ableitungen

Gliederung

Taylerpolynome

Theorem (Satz von Taylor)

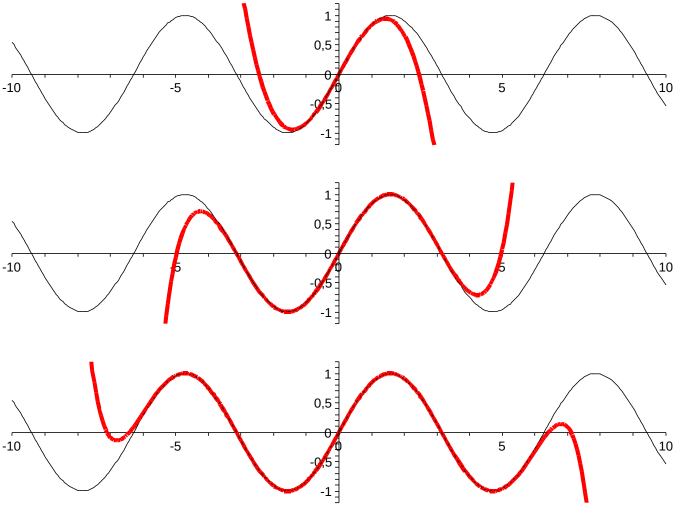
Für eine genügend oft differenzierbare Funktion $f : K \to \mathbb{R}$ gilt:

$$f(x) = \underbrace{\sum_{k=0}^{N} \frac{f^{(k)}(0)}{k!} x^k}_{\mathcal{T}_N\{f\}(x)} + \frac{f^{(N+1)}(\xi)}{(N+1)!}$$

für $N \in \mathbb{N}$ und ein $\xi \in K$.

Bemerkung: $n_0 > 1 \Rightarrow f^{(k)}(0)$ ist Tensor k-ter Ordnung.

Beispiel Taylorpolynome



Idee

Idee

1. $\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\text{MLP}}\} \quad \Rightarrow \quad f \approx f_{\text{MLP}}$
2. Ordnung des Taylorpolynoms $\leftrightarrow$ Maß für Approximationsgüte

Gliederung

Analytische Funktionen

Definition

$f : \mathbb{R}^{n_0} \to \mathbb{R}$ heißt *analytisch* in 0

$:\Leftrightarrow$

$\exists\ \delta > 0 : \quad \mathcal{T}_N\{f\}(x) \underset{N\to\infty}{\longrightarrow} f(x)$ für alle $\|x\| < \delta$.

Das größte δ heißt *Konvergenzradius*.

Analytische Funktionen

Definition

$f : \mathbb{R}^{n_0} \to \mathbb{R}$ heißt *analytisch* in 0

$:\Leftrightarrow$

$\exists\ \delta > 0 : \quad \mathcal{T}_N\{f\}(x) \underset{N\to\infty}{\longrightarrow} f(x)$ für alle $\|x\| < \delta$.

Das größte δ heißt *Konvergenzradius*.

Beispiele:

- Sinus ist analytisch in 0 mit $\delta = \infty$.
- Die Funktion $x \mapsto \frac{1}{x-1}$ ist analytisch in 0 mit $\delta = 1$.
- Die Funktion $x \mapsto \begin{cases} 0, & x \leq 0, \\ e^{-1/x}, & x > 0, \end{cases}$ ist **nicht** analytisch in 0.

Approximationsordnung

Theorem

Sei $f : [-1,1]^{n_0} \to \mathbb{R}$ *analytisch in* 0 *mit Konvergenzradius* $\delta > 1$. *Dann gilt:*

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{N} : \quad \left\| \mathcal{T}_N\{f\} - f \right\| < \varepsilon$$

Approximationsordnung

Theorem

Sei $f : [-1,1]^{n_0} \to \mathbb{R}$ *analytisch in* 0 *mit Konvergenzradius* $\delta > 1$. *Dann gilt:*

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{N} : \quad \left\| \mathcal{T}_N\{f\} - f \right\| < \varepsilon$$

Folgerung

Seien f und f_{MLP} beide analytisch in 0 mit Konvergenzradius $\delta > 1$, dann gibt es für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ so dass

$$\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\mathrm{MLP}}\} \quad \Rightarrow \quad \left\| f - f_{\mathrm{MLP}} \right\| < \varepsilon$$

Gliederung

Ansatz

Vorgabe

Ordnung N und Forderung dass $\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\mathsf{MLP}}\}$.

Ansatz

Vorgabe

Ordnung N und Forderung dass $\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\mathsf{MLP}}\}$.

Frage:

Welche Voraussetzung muss MLP erfüllen, damit Vorgabe erfüllt werden kann?

Ansatz

Vorgabe

Ordnung N und Forderung dass $\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\text{MLP}}\}$.

Frage:

Welche Voraussetzung muss MLP erfüllen, damit Vorgabe erfüllt werden kann?

Beispiel für $n_0 = 1$ und $N = 3$:

$$\mathcal{T}_3\{f\}(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$$

und

$$\mathcal{T}_3\{f_{\text{MLP}}\}(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3.$$

Nichtlineares Gleichungssystem

Eigenschaft von Polynomen

$\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\mathsf{MLP}}\} \quad \Leftrightarrow \quad$ alle Koeffizienten stimmen überein.

Nichtlineares Gleichungssystem

Eigenschaft von Polynomen

$\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\mathrm{MLP}}\} \quad \Leftrightarrow \quad$ alle Koeffizienten stimmen überein.

Im Beispiel - Gleichungssystem mit 4 Gleichungen:

$$
\begin{aligned}
a_0 &= b_0 \\
a_1 &= b_1 \\
a_2 &= b_2 \\
a_3 &= b_3
\end{aligned}
$$

Nichtlineares Gleichungssystem

Eigenschaft von Polynomen

$\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\mathrm{MLP}}\} \quad \Leftrightarrow \quad$ alle Koeffizienten stimmen überein.

Im Beispiel - Gleichungssystem mit 4 Gleichungen:

$$
\begin{aligned}
a_0 &= b_0(\mathbf{P}) \\
a_1 &= b_1(\mathbf{P}) \\
a_2 &= b_2(\mathbf{P}) \\
a_3 &= b_3(\mathbf{P})
\end{aligned}
$$

Nichtlineares Gleichungssystem

Eigenschaft von Polynomen

$\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\text{MLP}}\} \quad \Leftrightarrow \quad$ alle Koeffizienten stimmen überein.

Im Beispiel - Gleichungssystem mit 4 Gleichungen:

$$\begin{aligned} a_0 &= b_0(\mathbf{P}) \\ a_1 &= b_1(\mathbf{P}) \\ a_2 &= b_2(\mathbf{P}) \\ a_3 &= b_3(\mathbf{P}) \end{aligned}$$

Erinnerung: $\mathbf{P} = (\mathfrak{w}^1, \mathfrak{w}^2, \ldots, \mathfrak{w}^h, \mathbf{w}^y)$
(Kantengewichte und Schwellenwerte)

Abstraktere Sichtweise

Nichtlineares Gleichungssystem

$$\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\mathsf{MLP}}\} \quad \Leftrightarrow \quad \mathbf{a} = \mathfrak{K}(\mathbf{P})$$

- $\mathfrak{K} : \mathbb{R}^{|\mathbf{P}|} \to \mathbb{R}^G$
- $|\mathbf{P}|$: Anzahl der variablen Netzwerkparameter, abhhängig von h und $\mathbf{n}$.
- G: Anzahl Gleichungen = Anzahl Koeffizienten im Taylorpolynom

Abstraktere Sichtweise

Nichtlineares Gleichungssystem

$$\mathcal{T}_N\{f\} = \mathcal{T}_N\{f_{\mathrm{MLP}}\} \quad \Leftrightarrow \quad \mathbf{a} = \mathfrak{K}(\mathbf{P})$$

- $\mathfrak{K} : \mathbb{R}^{|\mathbf{P}|} \to \mathbb{R}^G$
- $|\mathbf{P}|$: Anzahl der variablen Netzwerkparameter, abhhängig von h und $\mathbf{n}$.
- G: Anzahl Gleichungen $=$ Anzahl Koeffizienten im Taylorpolynom

Notwendige Forderung

$\mathfrak{K}$ muss surjektiv sein.

Gliederung

1. Qualitative Approximation
   - Das Multi-Layer-Perceptron (MLP)
   - Abstandsmessung zwischen Funktionen und Dichtheit
   - Dichtheitsaussagen für MLPs
2. Taylorpolynome
   - Taylerreihenentwicklung
   - Analytische Funktionen
3. Notwendige Anzahl von Hidden-Neuronen
   - Lösbarkeit von mehrdimensionalen Gleichungen
   - (Geplante) Ergebnisse der Diplomarbeit

Ergebnisse

Bisher erhaltene Ergebnisse:

- Notwendige Mindestzahl von Hidden-Neuronen kann berechnet werden!
- Auch die Anzahl der Hiddenschichten und die Verteilung der Neuronen darin kann bestimmt werden.
- Nebenergebnis: Mehr als zwei Hidden-Schichten sind nicht sinnvoll.
- Die klassische Aktivierungsfunktion ist nicht gut geeignet.

Ergebnisse

Bisher erhaltene Ergebnisse:

- Notwendige Mindestzahl von Hidden-Neuronen kann berechnet werden!
- Auch die Anzahl der Hiddenschichten und die Verteilung der Neuronen darin kann bestimmt werden.
- Nebenergebnis: Mehr als zwei Hidden-Schichten sind nicht sinnvoll.
- Die klassische Aktivierungsfunktion ist nicht gut geeignet.

Ergebnisse

Bisher erhaltene Ergebnisse:

- Notwendige Mindestzahl von Hidden-Neuronen kann berechnet werden!
- Auch die Anzahl der Hiddenschichten und die Verteilung der Neuronen darin kann bestimmt werden.
- Nebenergebnis: Mehr als zwei Hidden-Schichten sind nicht sinnvoll.
- Die klassische Aktivierungsfunktion ist nicht gut geeignet.

Ergebnisse

Bisher erhaltene Ergebnisse:

- Notwendige Mindestzahl von Hidden-Neuronen kann berechnet werden!
- Auch die Anzahl der Hiddenschichten und die Verteilung der Neuronen darin kann bestimmt werden.
- Nebenergebnis: Mehr als zwei Hidden-Schichten sind nicht sinnvoll.
- Die klassische Aktivierungsfunktion ist nicht gut geeignet.

Ergebnisse

Bisher erhaltene Ergebnisse:

- Notwendige Mindestzahl von Hidden-Neuronen kann berechnet werden!
- Auch die Anzahl der Hiddenschichten und die Verteilung der Neuronen darin kann bestimmt werden.
- Nebenergebnis: Mehr als zwei Hidden-Schichten sind nicht sinnvoll.
- Die klassische Aktivierungsfunktion ist nicht gut geeignet.

Ergebnisse

Bisher erhaltene Ergebnisse:

- Notwendige Mindestzahl von Hidden-Neuronen kann berechnet werden!
- Auch die Anzahl der Hiddenschichten und die Verteilung der Neuronen darin kann bestimmt werden.
- Nebenergebnis: Mehr als zwei Hidden-Schichten sind nicht sinnvoll.
- Die klassische Aktivierungsfunktion ist nicht gut geeignet.

Weitere geplante Ergebnisse:

- Konkrete Beispiele zum Aufzeigen der Möglichkeiten und Grenzen des Ansatzes.
- Experimentelle Untersuchung anderer Aktivierungsfunktionen.

Ergebnisse

Bisher erhaltene Ergebnisse:

- Notwendige Mindestzahl von Hidden-Neuronen kann berechnet werden!
- Auch die Anzahl der Hiddenschichten und die Verteilung der Neuronen darin kann bestimmt werden.
- Nebenergebnis: Mehr als zwei Hidden-Schichten sind nicht sinnvoll.
- Die klassische Aktivierungsfunktion ist nicht gut geeignet.

Weitere geplante Ergebnisse:

- Konkrete Beispiele zum Aufzeigen der Möglichkeiten und Grenzen des Ansatzes.
- Experimentelle Untersuchung anderer Aktivierungsfunktionen.

Ergebnisse

Bisher erhaltene Ergebnisse:

- Notwendige Mindestzahl von Hidden-Neuronen kann berechnet werden!
- Auch die Anzahl der Hiddenschichten und die Verteilung der Neuronen darin kann bestimmt werden.
- Nebenergebnis: Mehr als zwei Hidden-Schichten sind nicht sinnvoll.
- Die klassische Aktivierungsfunktion ist nicht gut geeignet.

Weitere geplante Ergebnisse:

- Konkrete Beispiele zum Aufzeigen der Möglichkeiten und Grenzen des Ansatzes.
- Experimentelle Untersuchung anderer Aktivierungsfunktionen.