

Redes neuronales multicapa

Castiglione, Karpovsky, Sturla

Sistemas de Inteligencia Artificial

3 de Mayo de 2012

1 Introducción

- El problema

2 Modelado del problema

- Representación de la red neuronal
- Funciones de activación
- Simetría
- Cálculo del error
- Conjuntos de entrenamiento y testeo

3 Mejoras al algoritmo backpropagation

- Eta dinámico
- Ruido y momentum

4 Resultados

5 Conclusiones

Tabla de contenidos

- 1 **Introducción**
 - El problema
- 2 **Modelado del problema**
 - Representación de la red neuronal
 - Funciones de activación
 - Simetría
 - Cálculo del error
 - Conjuntos de entrenamiento y testeo
- 3 **Mejoras al algoritmo backpropagation**
 - Eta dinámico
 - Ruido y momentum
- 4 **Resultados**
- 5 **Conclusiones**

El problema

El problema planteado consiste en la estimación de funciones escalares a partir de un conjunto de puntos que las representan.

En nuestro caso particular hemos trabajado con el archivo **`samples7.txt`**

Gráfico de la función

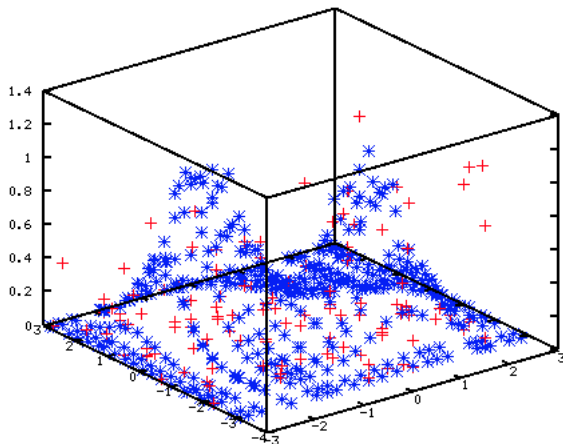


Tabla de contenidos

- 1 Introducción
 - El problema
- 2 **Modelado del problema**
 - Representación de la red neuronal
 - Funciones de activación
 - Simetría
 - Cálculo del error
 - Conjuntos de entrenamiento y testeo
- 3 Mejoras al algoritmo backpropagation
 - Eta dinámico
 - Ruido y momentum
- 4 Resultados
- 5 Conclusiones

Representación de la red neuronal

Se representó la red neuronal como una matriz de pesos.

- Cada neurona es una columna de pesos.
- Cada capa de neuronas es una matriz de pesos.
- La red neuronal, por consiguiente, es un vector de matrices.

Se utilizaron dos funciones de activación distintas:

Sigmoidea exponencial

$$g(h) = \frac{1}{1 + e^{-2\beta h}}$$

Derivada:

$$2\beta g(1 - g)$$

Tangente hiperbólica

$$g(x) = \tanh(x)$$

Derivada:

$$\beta g(1 - g^2)$$

Ruptura de la simetría

asfas

Ruptura de la simetría

asfsas

Conjunto de entrenamiento y testeo

Se decidió seguir el consejo de la cátedra y al realizar las pruebas se utilizó un subconjunto de los datos seleccionados al azar para la fase de aprendizaje y el subconjunto restante para testeo.

- Elección de puntos al azar? Puntos representativos?

Tabla de contenidos

- 1 Introducción
 - El problema
- 2 Modelado del problema
 - Representación de la red neuronal
 - Funciones de activación
 - Simetría
 - Cálculo del error
 - Conjuntos de entrenamiento y testeo
- 3 Mejoras al algoritmo backpropagation
 - Eta dinámico
 - Ruido y momentum
- 4 Resultados
- 5 Conclusiones

Eta dinámico (η adaptativo)

Sefasfas

Ruido y momentum

Seasfasfasfa

Tabla de contenidos

- 1 Introducción
 - El problema
- 2 Modelado del problema
 - Representación de la red neuronal
 - Funciones de activación
 - Simetría
 - Cálculo del error
 - Conjuntos de entrenamiento y testeo
- 3 Mejoras al algoritmo backpropagation
 - Eta dinámico
 - Ruido y momentum
- 4 Resultados
- 5 Conclusiones

Resultados

Se

Tabla de contenidos

- 1 Introducción
 - El problema
- 2 Modelado del problema
 - Representación de la red neuronal
 - Funciones de activación
 - Simetría
 - Cálculo del error
 - Conjuntos de entrenamiento y testeo
- 3 Mejoras al algoritmo backpropagation
 - Eta dinámico
 - Ruido y momentum
- 4 Resultados
- 5 Conclusiones

A sample slide

Theorem (The Poincaré inequality)

Suppose $\Omega \in \mathbf{R}^n$ is a bounded domain with smooth boundary. Then there exists a $\lambda > 0$, depending only on Ω , such that for any function f in the Sobolev space $H_0^1(\Omega)$ we have:

$$\int_{\Omega} |\nabla u|^2 dx \geq \lambda \int_{\Omega} |u|^2 dx.$$

Here is what *itemized* and *enumerated* lists look like:

- itemized item 1
- itemized item 2
- itemized item 3

- ① enumerated item 1
- ② enumerated item 2
- ③ enumerated item 3