

UNIVERSIDAD DE SANTIAGO DE CHILE
FACULTAD DE INGENIERIA
DEPARTAMENTO DE INGENIERIA INFORMÁTICA

Propuesta de Tesis

Informe

Nombre: Felipe Alberto Reyes González

Programa: Magíster en Ingeniería Informática

Profesor patrocinante: Victor Parada

Cel.: 890 26 317

email: felipe.reyesg@usach.cl

20 de abril de 2017

Tabla de contenido

1. Introducción	3
1.1. Las redes neuronales	3
2. Algoritmo de retropropagación	5
2.1. Regla delta	5
2.2. Gradiente descendente	5
2.3. Gradiente descendente estocástico	5
3. Desvanecimiento del gradiente decendente	5
3.1. Soluciones	5
3.2. LEEA	5
3.3. Simulated Annealing	5
4. Experimentación	5
4.1. Diseño del experimento	5
4.2. Resultados de la experimentación	5

1 Introducción

El primer modelo de red neuronal artificiales (NN) fue propuesto en 1943 por McCulloch y Pitts en términos de un modelo computacional de actividad nerviosa. Las NN se han inspirado en las redes neuronales biológicas y las conexiones que construyen. Este modelo era un modelo binario, donde cada neurona posee un umbral, y sirvió de base para los modelos posteriores.

Las características principales de las NN son las siguientes:

1. Auto-organización y adaptabilidad: utilizan algoritmos de aprendizaje adaptativo y auto-organizativo, por lo que ofrecen mejores posibilidades de procesamiento robusto y adaptativo.
2. Procesado no lineal: aumenta la capacidad de la red para aproximar funciones, clasificar patrones y aumenta su inmunidad frente al ruido.
3. Procesado paralelo: normalmente se usa un gran número de nodos de procesamiento, con alto nivel de interconectividad.

1.1 Las redes neuronales

El elemento básico de las NN es el nodo, que recibe un vector de entrada para producir una salida como muestra en la figura 1. Cada entrada tiene asociado un vector de pesos w , que se va modificando durante el proceso de aprendizaje. Cada unidad aplica una función f sobre la suma de las entradas ponderada por el vector de pesos

$$y_i = \sum_j w_{ij} y_j$$

Donde el resultado puede servir como entrada de otras unidades.

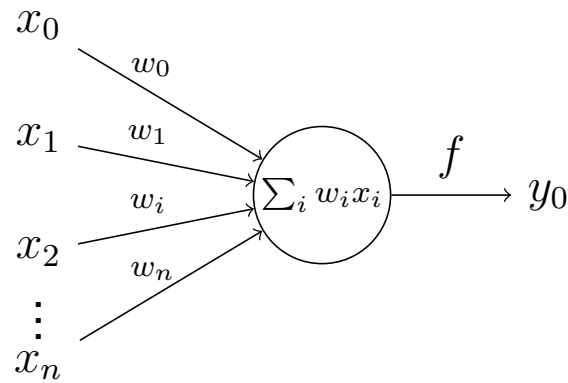


Figura 1: Neurona

Existen dos fases importante dentro del modelo

- Fase de entrenamiento: Se usa un conjunto de datos o patrones de entrenamiento para determinar los pesos que definen el modelo de la NN. Se calculan de manera iterativa, de acuerdo con los valores de entrenamiento, con el objeto de minimizar el error cometido entre la salida obtenida por la NN y la salida deseada.
- Fase de prueba: Durante el entrenamiento, el modelo se ajusta al conjunto de entrenamiento, perdiendo la habilidad de generalizar su aprendizaje a casos nuevos, a esta situación se le llama sobreajuste. Para evitar el sobreajuste, se utiliza un segundo grupo de datos diferentes, el conjunto de validación, que permitirá controlar el proceso de aprendizaje.

Los pesos óptimos se obtienen minimizando una función. Uno de los criterios utilizados es la minimización del error cuadrático medio entre el valor de salida y el valor real esperado.

2 Algoritmo de retropropagación

2.1 Regla delta

2.2 Gradiente descendente

2.3 Gradiente descendente estocástico

3 Desvanecimiento del gradiente decendente

3.1 Soluciones

3.2 LEEA

3.3 Simulated Annealing

4 Experimentación

4.1 Diseño del experimento

Decripción detallada de las redes, descripción de los datos.

4.2 Resultados de la experimentación