

Algoritmos Genéticos y el problema del desvanecimiento del gradiente

Informe III

Felipe Reyes González

Universidad de Santiago de Chile
Facultad de Ingeniería
Magíster en Ingeniería Informática

4 de mayo de 2017

1 El Gradiente Descendente

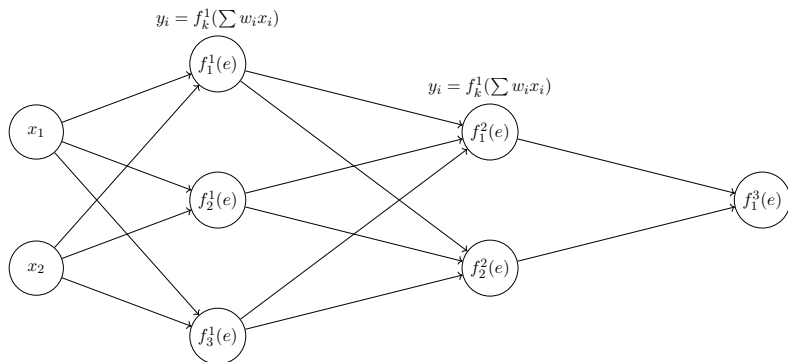
2 El Gradiente Descendente Estocástico

- El desvanecimiento del gradiente

- ▶ Se busca un algoritmo que permita encontrar pesos y sesgos para que la salida de la red aproxime los valores de $y(x)$ a los valores correspondientes con cada entrada x .
- ▶ Para cuantificar qué tan bien estamos logrando este objetivo definimos una función de costo

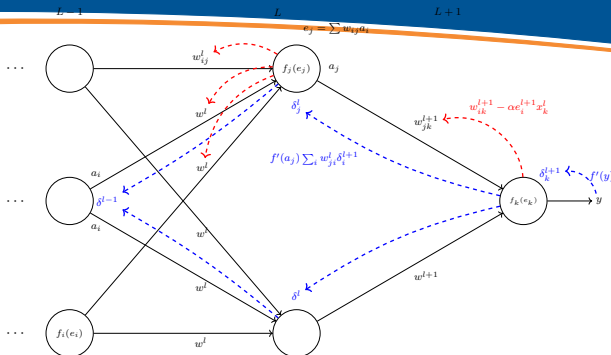
- El método del gradiente descendente estocástico (SGD) actualiza los parámetros en cada ejemplo x_i y etiqueta y_i de la siguiente manera

$$\theta = \theta - \eta \nabla_{\theta}$$



El Gradiente Descendente Estocástico

El desvanecimiento del gradiente



$$f'(a_j) \sum_i w_{ji}^l f'(y)$$

$$\underbrace{f'_{a_j}(a_j) \sum_m w_{jm}^l \left(\overbrace{f'(a_j) \sum_n w_{jn}^l f'(y)}^{\delta_k^{l+1}} \right)}_{\delta_j^l}$$

Algoritmos Genéticos y el problema del desvanecimiento del gradiente

Informe III

Felipe Reyes González

Universidad de Santiago de Chile
Facultad de Ingeniería
Magíster en Ingeniería Informática

4 de mayo de 2017