



# Experimento de Montana y Davis

- 1 Experimento de Montana y Davis
- 2 La red neuronal
- 3 Algoritmos genéticos para redes neuronales
- 4 Experimentos

# Temas

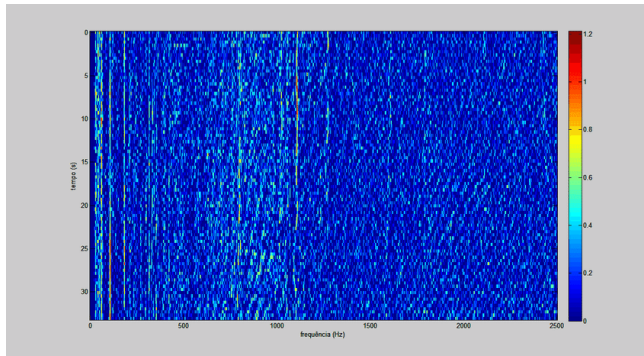
- 1 Experimento de Montana y Davis
  - Escenario

## Escenario (Experimento en 1989)

- Una etapa del procesamiento de datos de un sonar pasivo.
- Los datos son arreglos de receptores acústicos bajo el agua.
- Sistema experto desarrollado por BBN para analizar señales interesantes en el océano.
- Los datos que entran al sistema experto son *lofargramas* (*lofargrams*)<sup>[1]</sup>, que (similar a los espectrogramas) son imágenes basadas en la intensidad que ilustran la distribución de energía acústica como función de la frecuencia y el tiempo.

---

<sup>[1]</sup>LOFAR es *Low frequency analyser and recorder*.



**Figura:** Ejemplo de lofargrama. Cuando un barco navega, irradia un ruido particular que lo identifica con un categoría específica. La identificación de las clases de barcos es una tarea no trivial que se realiza utilizando sistemas de sonares pasivos.

# Escenario

- Bandas de energía aparecen como líneas que tienden a ser verticales predominantemente.
- Las características de cada línea proveen claves a los operadores del sonar acerca de los objetos que las están produciendo.
- Se busca obtener un algoritmo que capture todo lo que los operadores pueden ver a partir de las líneas reales.

# La red neuronal

- 1 Experimento de Montana y Davis
- 2 La red neuronal
- 3 Algoritmos genéticos para redes neuronales
- 4 Experimentos

# Temas

- 2 La red neuronal
  - Datos de entrada
  - Arquitectura de la red



# Datos

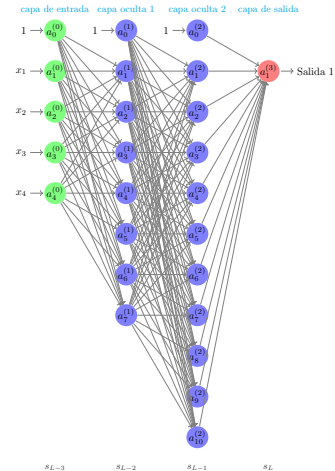
- 1200 lofargramas rectangulares de tamaño fijo.
- 236 fueron usados para el conjunto de entrenamiento.
- centrados en un tipo de línea determinado por un experto.
- 30 % líneas interesantes
- 70 % líneas de otros tipos
- Problema de clasificación: ¿las líneas son interesantes o no?

# Temas

- 2 La red neuronal
  - Datos de entrada
  - Arquitectura de la red

# Detalles de la red

- 4 entradas
- 1 salida
- capa oculta con 7 nodos
- capa oculta con 10 nodos
- dando un total de 126 pesos



# Algoritmos genéticos para redes neuronales

- 1 Experimento de Montana y Davis
- 2 La red neuronal
- 3 Algoritmos genéticos para redes neuronales
- 4 Experimentos

# Temas

- 3 Algoritmos genéticos para redes neuronales
  - Algoritmos genéticos
  - Operadores

# Algoritmos genéticos

Los algoritmos genéticos son una técnica de búsqueda que permite optimizar funciones. Para ello requieren:[2]

- ❶ Codificar soluciones como cromosomas.
- ❷ Una función de evaluación.
- ❸ Método para inicializar la población de cromosomas.
- ❹ Operadores para obtener a la siguiente generación. Ej:
  - mutación,
  - cruza y
  - operadores específicos para el dominio.
- ❺ Asignaciones a los varios parámetros del algoritmo.

---

[2]Montana y Davis 1989.

# Procedimiento

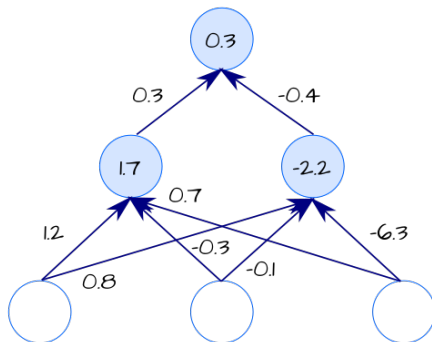
---

## Algoritmo 1 Búsqueda genética.

---

- 1: Inicializar población
  - 2: Ordenar población según función de evaluación.
  - 3: **while**  $\neg$ criterioDeParo **do** ▷ Reproducir población:
  - 4:     Seleccionar uno o más padres aleatoriamente, los mejores con mayor probabilidad.
  - 5:     Aplicar operadores para producir hijos.
  - 6:     Insertar hijos/reemplazar padres.
  - 7: **return** Mejor individuo.
-

# Cromosoma



**Codificación:**

(0.3, -0.4, 0.3, 1.2, 0.8, -0.3, -0.1, 0.7, -6.3, 1.7, -2.2)

**Figura:** Ejemplo. Pesos y sesgos como lista de números reales.



# AG

**Cromosomas** Vector de pesos y sesgos.

**Función de error** Suma de errores al cuadrado.

**Inicialización** ¿Gaussiana? (Pesos con valores absolutos pequeños y unos pocos con valores absolutos grandes)

**Parámetros** Sólo manejan uno:

- Población: 50.

# Temas

- 3 Algoritmos genéticos para redes neuronales
  - Algoritmos genéticos
  - Operadores

# Operadores

Montana y Davis 1989 probaron varios tipos de operadores:

- *Mutación*

(0.3	-0.4	1.2	0.8	-0.3	-0.1	0.7	-6.3)
<hr/>							
(0.3	-0.4	1.2	-2.2	-0.3	-0.1	1.2	-6.3)

- *Cruza* (Crossover)

(0.3	-0.4	1.2	0.8	-0.3	-0.1	0.7	-6.3)
<hr/>							
(-0.7	-0.9	1.3	0.4	1.8	2.1	-0.2	-1.2)
<hr/>							
(-0.7	-0.4	1.3	0.4	1.8	2.1	0.7	-6.3)

- *Gradiente*

(0.3	-0.4	1.2	0.8	-0.3	-0.1	0.7	-6.3)
<hr/>							
(0.4	-0.2	1.2	0.7	0	-0.2	0.5	-6.6)

# Mutación

## Mutación de pesos sin sesgo

- Para cada entrada en el cromosoma, remuestrear el valor con probabilidad  $p = 0.1$ .

## Mutación de pesos con sesgo

- Para cada entrada en el cromosoma, agregar ruido al valor con probabilidad  $p = 0.1$ .

## Mutación de nodos

- Para cada nodo en la red, agregar ruido a cada uno de los **pesos entrantes**.

## Mutación de nodos débiles

### Definición

Se define la *fortaleza* de un nodo oculto como la diferencia entre la evaluación de la red intacta y su evaluación con ese nodo lobotomizado (todos pesos de salida se fijan en cero).

- ❶ Para un ejemplar, probar la fortaleza de cada uno de sus nodos.
- ❷ Para los  $m$  nodos más débiles mutar cada uno de los pesos que llegan o salen de él:
  - Si la fortaleza del nodo es negativa, mutar sin sesgo.
  - Si es positiva, mutar con sesgo.

# Cruces

## Cruce de pesos

- Para cada entrada en el cromosoma, elegir aleatoriamente de qué padre se copia el valor.

## Cruce de nodos

- Para cada nodo en el hijo, elegir aleatoriamente qué nodo padre copiar. Lo que se copia son todos los pesos que entran.

## Cruce de características

- Alinea a los nodos en cada capa de los dos padres del tal manera que nodos con funciones parecidas (a mismas entradas a la red se activan de modos semejantes) queden en la misma posición. Luego realiza cruce de nodos.

# Ascenso de colinas

- ❶ Calcula el gradiente para cada elemento del conjunto de entrenamiento, los suma para obtener el total y lo **normaliza**.
- ❷ Realiza un paso modificando los pesos en la dirección del gradiente con:
  - $\alpha = 0.4\alpha$  si el hijo es peor que el padre.
  - $\alpha = 1.4\alpha$  si el hijo es mejor que el padre.
- ❸ Se distingue de propagación hacia atrás porque el ajuste no es proporcional a la magnitud del gradiente.

# Experimentos

- 1 Experimento de Montana y Davis
- 2 La red neuronal
- 3 Algoritmos genéticos para redes neuronales
- 4 Experimentos



# Temas

- 4 Experimentos
  - Experimentos
  - Resultados

# Experimentos

Se corrió cada experimento 10 veces con valores aleatorios iniciales distintos y se tomó el promedio del desempeño del mejor individuo.

- 1 Mutaciones. El cruce es sólo sobre pesos.
- 2 Cruzas. La mutación es sobre nodos.
- 3 Mutación de nodos débiles. Se mutan nodos y se cruzan características.
- 4 Ascenso de colinas. Se compara con mutar y cruzar nodos nada más vs usar ascenso de colinas para las últimas 300 iteraciones.
- 5 Propagación hacia atrás vs mutar y cruzar nodos.

# Temas

- 4 Experimentos
  - Experimentos
  - Resultados

# Mutaciones y cruces

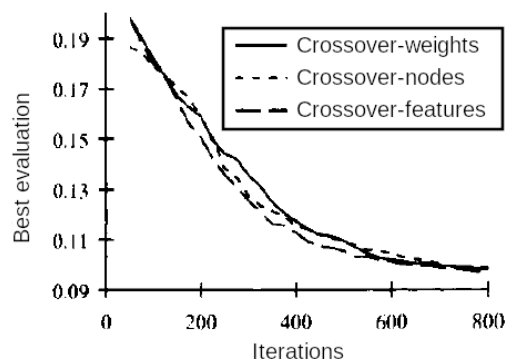
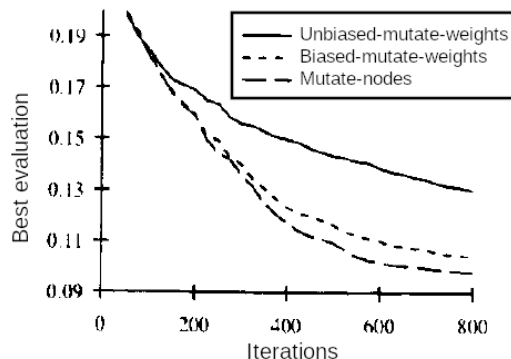
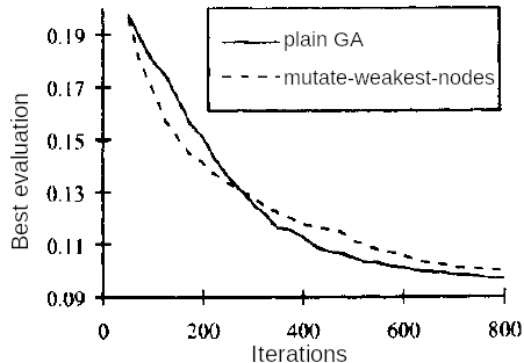


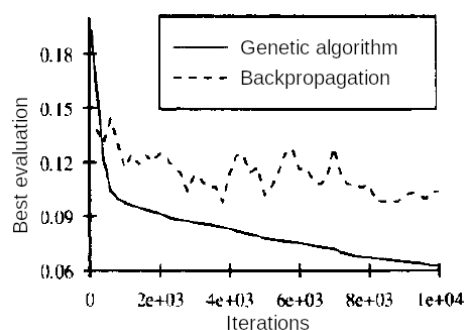
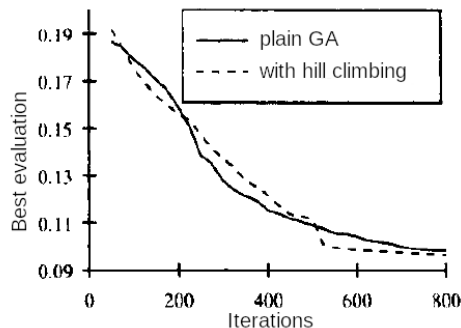
Figura: *Izquierda: Mutación* de a) pesos sin sesgo, b) pesos con sesgo, c) **nodos**. *Derecha: Cruce* de a) pesos, b) nodos, c) características.

# Mutación de nodos débiles



**Figura:** La función de *fortaleza del nodo* ayudó al inicio del entrenamiento, después la empeoró, por lo que sugieren modificarla.

# Ascenso de colinas y propagación hacia atrás



**Figura:** *Izquierda:* **Ascenso de colinas** es propenso a mínimos locales, se recomienda sólo usarlo cerca de un mínimo global. *Derecha:* Propagación hacia atrás simple vs el algoritmo genético con su población de 50 redes se ejecuta rápidamente y logra un buen desempeño.

# Referencias I



Montana, David J. y Lawrence Davis (1989). «Training Feedforward Neural Networks Using Genetic Algorithms». En: *Proceedings of the 11th International Joint Conference on Artificial Intelligence - Volume 1*. IJCAI'89. Detroit, Michigan: Morgan Kaufmann Publishers Inc., págs. 762-767. URL: <http://dl.acm.org/citation.cfm?id=1623755.1623876>.

# Licencia

Creative Commons  
Atribución-No Comercial-Compartir Igual

