

Sistemas de Inteligencia Artificial

Algoritmos genéticos

Informe TP4

Grupo 7:

Nicolás Buchhalter

Francisco Depascuali

Agustina Fainguersch

1. Introducción

Para este TP se implementó un motor de algoritmos genéticos con el objetivo de obtener los mejores resultados para la red neuronal implementada durante el TPE2 con la diferencia de que se utilizó una función diferente. Este motor es capaz de responder a variaciones de operadores genéticos, condiciones de corte, métodos de selección y reemplazo. Hacia el final del informe se compararon los resultados obtenidos con los del trabajo anterior, se los analizó y se obtuvieron las conclusiones correspondientes.

2. Objetivo

Estimar los valores de la función solicitada por la cátedra encontrando los pesos de una red neuronal mediante el uso de algoritmos genéticos.

3. El problema

$$y = \tanh(0,1x) + \sin(3x), \text{ con } x \in [-4, 4]$$

4. Modelado

Los algoritmos genéticos requieren de operadores genéticos de cruza, mutación, criterios de selección/reemplazo, métodos de reemplazo y condiciones de corte. A continuación se listarán los utilizados en este trabajo.

Operadores Genéticos:

Cruza:

- ☐ Clásico
- ☐ Dos puntos
- ☐ Uniforme
- ☐ Anular

Mutación:

- ☐ Clásica
- ☐ No uniforme

Backpropagation

Criterios de selección/reemplazo:

- ☐ Elite
- ☐ Ruleta
- ☐ Boltzmann
- ☐ Torneos probabilísticos.
- ☐ Mixto(Elite + Ruleta)
- ☐ Mixto(Elite + Universal)

Métodos de reemplazo

- ☐ Método de reemplazo 1
- ☐ Método de reemplazo 2
- ☐ Método de reemplazo 3

Condiciones de corte:

- ☐ Cantidad de generaciones
- ☐ Estructura
- ☐ Contenido

- Cota (error en torno al óptimo)

5. Representación del individuo

Cada individuo se representó con un objeto que tiene como variables de instancia todos los parámetros del perceptrón para poder correr el backpropagation, en conjunto con la matriz de pesos, representada como un vector que tiene todos los pesos de forma contigua y el error cuadrático medio que quedaría al evaluar esos pesos con los valores de entrada y compararlos con los valores de salida esperados. También se guarda la aptitud de cada individuo de la población, en un vector que los contiene a todos.

6. Función de aptitud

Para trabajar a lo largo de todo el TP se utilizó la siguiente función de aptitud:

$$f = \frac{1}{ECM^2}$$

Donde ECM es el error cuadrático medio calculado como $ECM = \frac{1}{2n} * \sum_{i=1}^n (S_i - O_i)^2$ siendo n la cantidad de patrones, S la salida esperada y O la salida obtenida.

7. Cruza

Se define $P_c = 0.75$ como la probabilidad de cruce. Las cruces se dan con probabilidad mayor a 0.5 para que ocurra alguna variación en la población en los métodos de reemplazo.

8. Mutación

Se define $P_m = 0.1$ como la probabilidad de mutación para asegurarse de que el algoritmo no se vuelva aleatorio y evitar estancarse en máximos locales.

9. Backpropagation

Luego de haber sido cruzados y mutados, los nuevos individuos son entrenados con el método de backpropagation implementado en el trabajo práctico anterior. En caso de no realizarse este entrenamiento, los algoritmos genéticos hallarían una solución, aunque el entrenamiento le proporciona un “empujón” para llegar más rápido a la solución.

La probabilidad de entrenamiento se fijó en 0.75 para que la mayoría de los hijos sean entrenados.

10. Resultados

Para hacer las pruebas tomamos un perceptrón con los siguientes valores:

eta: 0.02, beta: 0.5, capas: [1 15 5 1], momentum : 0.9, sin eta adaptativo, cota de aptitud máxima: 160000, función de activación: tanh, épocas a entrenar en backpropagation: 5, criterio de corte por contenido: 5 generaciones, criterio de corte por estructura: 25 generaciones siendo igual el 75% de la población, máxima cantidad de generaciones: 500.

- **Criterios de reemplazo/selección:**

A partir de los resultados notamos que los mejores criterios de selección/reemplazo fueron boltzmann y ruleta. Interpretamos que es así dado que a medida que avanzan las generaciones, ambos se van adecuando al problema porque se diferencian mejor las aptitudes de los individuos,

entonces ambos criterios se quedan con las mejores aptitudes, sin embargo se mantiene la diversidad dado que son estocásticos. Más aún, a diferencia de ruleta, boltzmann separa mejor los individuos porque utiliza también la temperatura de boltzmann que se adecúa a medida que pasan las generaciones.

La selección mixto dió en general buenos resultados porque permitió elegir 2 subconjuntos utilizando dos criterios: elite, que posibilitó quedarse con los individuos más aptos y universal, que se queda con individuos aptos pero también proporciona diversidad.

- **Mutaciones:**

Nos permitió realizar una modificación mínima que puede resultar beneficiosa al acercarse a la solución, generando nueva información del problema.

Analizando los resultados no encontramos una diferencia razonable entre utilizar mutación no uniforme y clásica. En la mutación no uniforme la probabilidad de mutación se disminuyó a medida que avanzaban las generaciones. Intuimos que la no uniforme no resultó mejor que clásica porque la probabilidad de mutación se reduce muy rápido, por lo que hubiera sido bueno reducir menos esta probabilidad (para que continúe mutando) o empezar con una probabilidad de mutación más alta.

- **Cruza:**

Nos permitió combinar información del problema al cruzar dos padres. La cruce de dos puntos resultó la mejor elección, interpretamos que es así dado que la cruce de un punto es un caso particular de la cruce de dos puntos en el cual se cambia a partir de un punto hasta el final, mientras que en la cruce de dos puntos ese punto final no es el último.

Consideramos pertinente realizar una prueba sin utilizar cruce y solo mutación y vice versa. Mediante estos datos (que se encuentra en la tabla de la figura 1.4) podemos ver que:

1. con cruce alto y sin mutación: La aptitud se estanca porque no se introduce nueva información.
2. sin cruce y con mutación alta: Funciona al estilo “random walk”. Puede llegar a la solución, pero no combina la información de los padres.

1. Método de reemplazo 1

Como podemos ver en el gráfico de la figura 2.1, al combinar elite con este método se puede verificar efectivamente que, inicialmente, debido a que elegirá siempre a los mismos para la cruce, el algoritmo llegaría a un máximo local del que le costaría mucho salir. El algoritmo estuvo estancado durante casi 400 épocas en un fitness de alrededor de 50, y luego, por alguna mutación, llegó a generar una mejora para salir de ese máximo local y empezó a subir rápidamente la aptitud máxima de la población y terminó teniendo una buena aptitud en la época 500.

2. Método de reemplazo 2

Como éste método, ya que hay siempre individuos que pasan a la siguiente generación sin cambiar, si se utiliza ruleta para seleccionar puede llegar a pasar muchas veces que los elegidos sin cambiar sean los individuos que no tengan buena aptitud, y si los hijos generados tampoco resultaron ser buenos, todo esto llevará a una caída en la aptitud máxima que se pudo haber

perdido. Por ésta razón, como se puede ver en el gráfico de la figura 2.3, la aptitud máxima tenderá a oscilar mucho.

3. Método de reemplazo 3

En este método, se puede ver como el máximo y el promedio nunca se llegan a separar demasiado. Esto se debe a que en este método de reemplazo, como se selecciona siempre entre toda la comunidad en conjunto con los hijos generados, la tendencia va a ser siempre a elegir a los mejores de todos, por lo que la probabilidad de elegir a los hijos decrece, y de esta forma de a poco se va volviendo más homogénea la población. Esto se puede ver especialmente marcado en los métodos de selección en el reemplazo más deterministas como elite o mixto. Esto lleva a cortes por estructura o contenido de forma bastante prematura.

6. Comparación con TPE2

A lo largo del TP2 se utilizaron diferentes arquitecturas para lograr el menor error de aproximación. En el TP3, si bien probamos varias arquitecturas, elegimos la que nos parecía óptima y alteramos los métodos de selección, reemplazo y mutación para lograr encontrar la mejor combinación de pesos posibles. Se decidió tomar un eta de 0.02, para obtener valores óptimos. Esto lleva a que las probabilidades de que queden individuos con una aptitud mayor al promedio sean mayores, y luego con cruza poder ir encontrando mejores individuos rápidamente.

7. Conclusiones

Se pueden explorar diferentes soluciones al mismo tiempo, en otras palabras paralelizar el problema, por lo cual si un individuo no resulta ser apto, se continúa con otro individuo.

Hay cierta información en común entre algunos individuos, similar a las clases de equivalencias. Por medio de esquemas podrían analizarse estas clases y detectar información en común de los individuos que pertenecen a esta clase (pudiendo agruparlos y analizarlos en conjuntos).

Los algoritmos genéticos son una buena herramienta para resolver problemas no lineales porque los individuos no son independientes entre sí y permite manipular múltiples parámetros al mismo tiempo.

Los algoritmos genéticos cuentan con métodos que en conjunto le permiten escapar de máximos locales:

1. Paralelismo: Analizar múltiples soluciones simultáneamente.
2. Selección: Guía a los individuos seleccionados hacia la solución óptima.
3. Mutación: Permite a un individuo explorar su “vecindario”, en otras palabras, introducir un cambio mínimo que puede llegar a mejorar la aptitud del individuo.
4. Cruza: Permite intercambiar información entre los diferentes individuos, por lo que el hijo hereda la información de los padres, beneficiándose de lo que ya habían aprendido estos.

8. Anexo

Crit. Selección	Población	Cruza	Mutación	Generaciones	Max F	Mean F	Min F	Corte
elite	10	dos puntos	clásico	312	40088	20346	661	cota
	20	dos puntos	clásico	500	23412	10849	531	generaciones
	10	clasico	clásico	174	40379	21897	1603	cota
	10	clasico	no uniforme	500	12901	5059	921	generaciones
ruleta	10	dos puntos	clásico	167	41540	17989	2585	cota
	20	dos puntos	clásico	261	43850	14047	104	cota
boltzmann	10	dos puntos	clásico	500	94359	47028	12006	generaciones
	20	dos puntos	clásico	438	164331	84221	12358	cota
torneos	20	dos puntos	clásico	500	77829	39934	454	generaciones

Figura 1.1: Tabla comparativa para el método de reemplazo 1

Crit. Selección	Población	Padres	Crit. Reemplazo	Cruza	Mutación	Generaciones	F Máx	F Prom	F Mín	Corte
ruleta	20	10	mixto(elite+rul)	dos puntos	clásica	148	10868	6426	421	contenido
	10	6	mixto(elite+rul)	dos puntos	clásica	194	25709	14942	800	contenido
	10	6	ruleta	dos puntos	clásica	500	64864	43439	9629	generaciones
	20	12	ruleta	dos puntos	clásica	262	36695	20607	254	contenido
	10	6	boltzmann	dos puntos	clásica	177	25780	18626	5650	contenido
torneos	20	10	mixto(elite+rul)	dos puntos	clásica	500	52	52	50	generaciones
	10	6	mixto(elite+rul)	dos puntos	clásica	261	9920	5679	122	contenido
	20	10	elite	dos puntos	clásica	198	10424	6593	162	contenido
	10	6	elite	dos puntos	clásica	177	6848	5230	1061	contenido
boltzmann	10	6	mixto(elite+univ)	dos puntos	clásica	320	14710	6658	80	contenido
	10	6	elite	dos puntos	clásica	219	20139	12642	248	contenido
elite	10	6	elite	clásico	clásica	174	11862	8672	7606	contenido
	10	6	elite	clásico	no uniforme	168	9727	7037	554	contenido
	10	6	elite	dos puntos	clásica	195	50	50	49	contenido
	10	6	elite	dos puntos	no uniforme	176	13877	9863	1561	contenido
	10	6	elite	uniforme	clásica	216	51	51	50	contenido
	10	6	elite	uniforme	no uniforme	229	17982	12112	2212	contenido
	10	6	elite	anular	clásica	176	51	49	47	contenido
	10	6	elite	anular	no uniforme	114	50	50	47	contenido
	10	6	ruleta	clásico	clásica	218	41155	15883	94	cota
	10	6	ruleta	clásico	no uniforme	318	41181	22377	4160	cota

Figura 1.2: Tabla comparativa para el método de reemplazo 2

Crit. Selección	Población	Padres	Crit. Reemplazo	Cruza	Mutación	Generaciones	F Máx	F Prom	F Mín	Corte
ruleta	20	10	mixto(elite+rul)	dos puntos	clásica	148	10868	6426	421	contenido
	10	6	mixto(elite+rul)	dos puntos	clásica	194	25709	14942	800	contenido
	10	6	ruleta	dos puntos	clásica	500	64864	43439	9629	generaciones
	20	12	ruleta	dos puntos	clásica	262	36695	20607	254	contenido
	10	6	boltzmann	dos puntos	clásica	177	25780	18626	5650	contenido
torneos	20	10	mixto(elite+rul)	dos puntos	clásica	500	52	52	50	generaciones
	10	6	mixto(elite+rul)	dos puntos	clásica	261	9920	5679	122	contenido
	20	10	elite	dos puntos	clásica	198	10424	6593	162	contenido
	10	6	elite	dos puntos	clásica	177	6848	5230	1061	contenido
boltzmann	10	6	mixto(elite+univ)	dos puntos	clásica	320	14710	6658	80	contenido
	10	6	elite	dos puntos	clásica	219	20139	12642	248	contenido
elite	10	6	elite	clásico	clásica	174	11862	8672	7606	contenido
	10	6	elite	clásico	no uniforme	168	9727	7037	554	contenido
	10	6	elite	dos puntos	clásica	195	50	50	49	contenido
	10	6	elite	dos puntos	no uniforme	176	13877	9863	1561	contenido
	10	6	elite	uniforme	clásica	216	51	51	50	contenido
	10	6	elite	uniforme	no uniforme	229	17982	12112	2212	contenido
	10	6	elite	anular	clásica	176	51	49	47	contenido
	10	6	elite	anular	no uniforme	114	50	50	47	contenido
	10	6	ruleta	clásico	clásica	218	41155	15883	94	cota
	10	6	ruleta	clásico	no uniforme	318	41181	22377	4160	cota

Figura 1.3: Tabla comparativa para el método de reemplazo 3

Crit. Selección	Población	Padres	Pc	Pm	Crit. Reemplazo	Cruza	Mutación	Generaciones	F Máx	F Prom	F Mín	Corte
Torneos	10	6	0	0.75	mixed(elite + univ)	dos puntos	clásica	290	15272	7463	750	contenido
Torneos	10	6	0.75	0	mixed(elite + univ)	dos puntos	clásica	137	49	48	45	contenido

Figura 1.4: Tabla comparativa con mutación sin cruza y cruza sin mutación

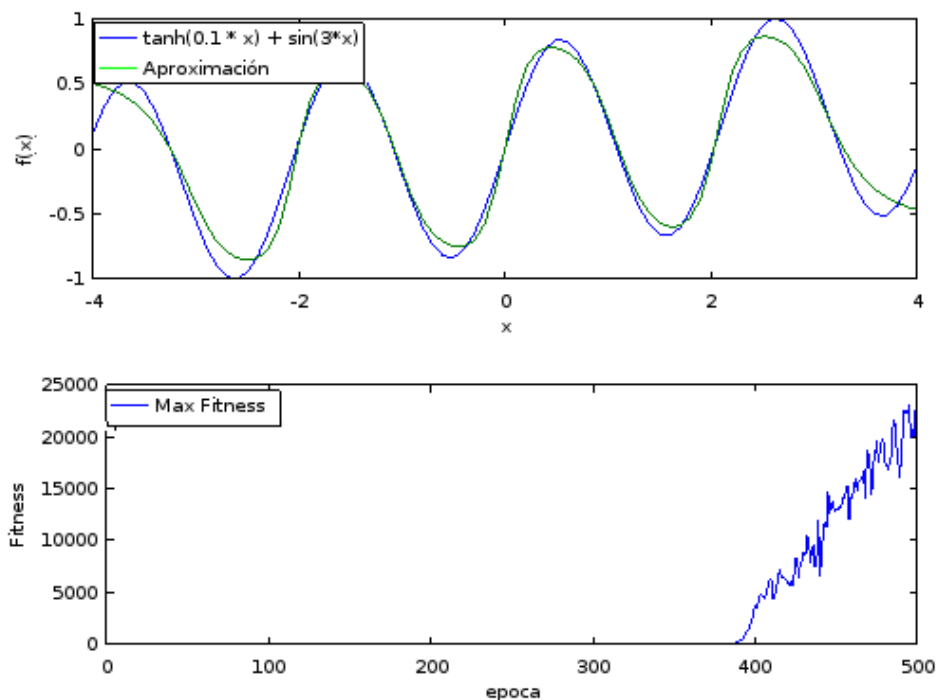


Figura 2.1: Gráfico de aptitud en método de reemplazo 1 con elite.

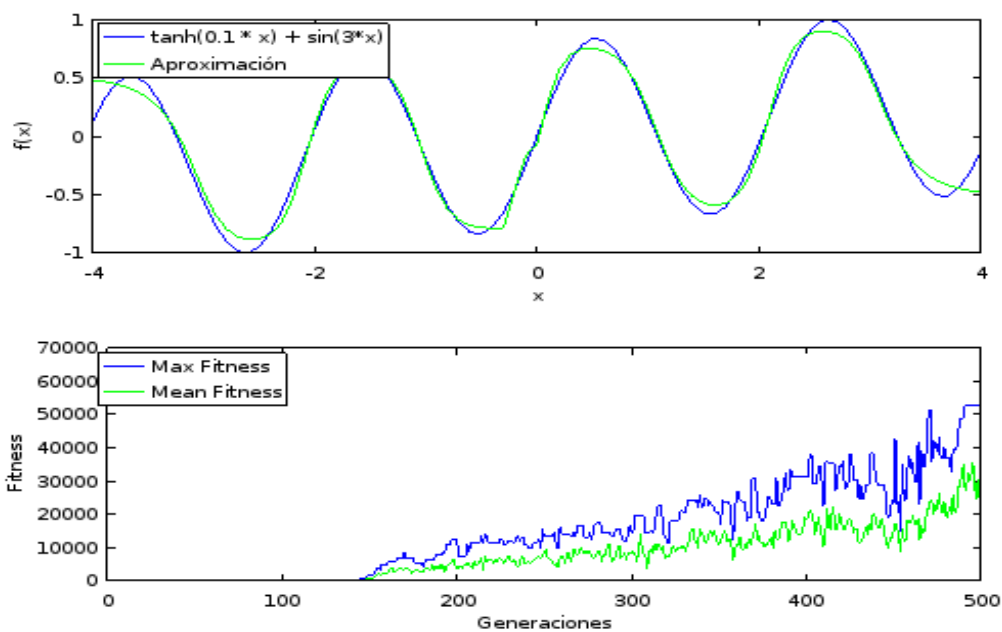


Figura 2.2: Gráfico de resultado óptimo para el método de reemplazo 1

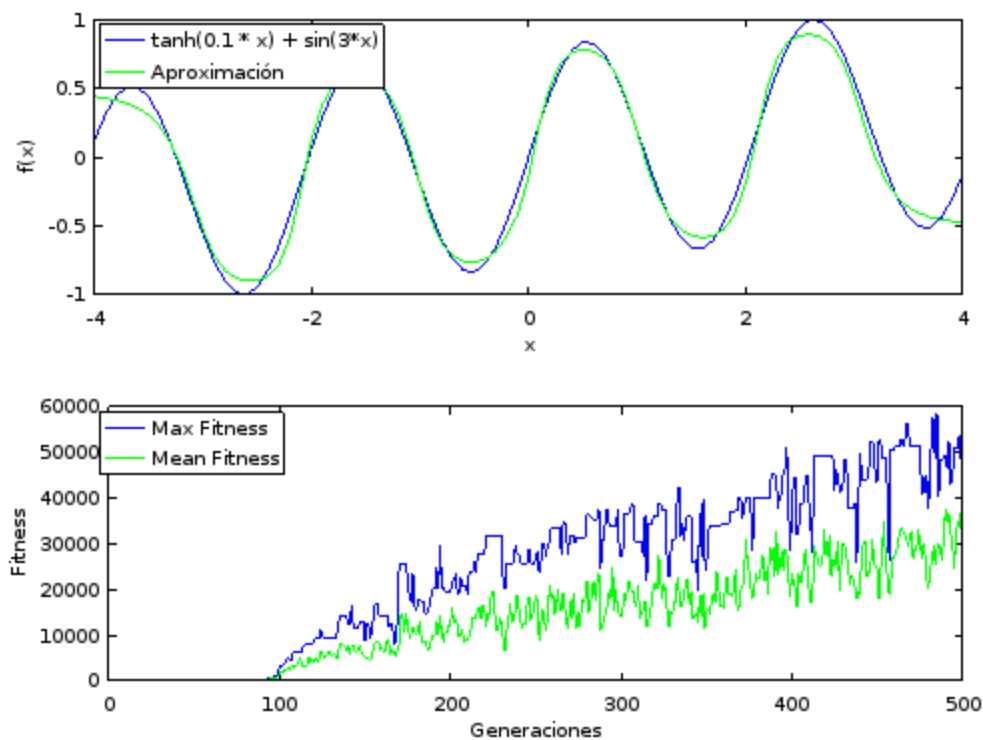


Figura 2.3: Gráfico de resultado óptimo para el método de reemplazo 2

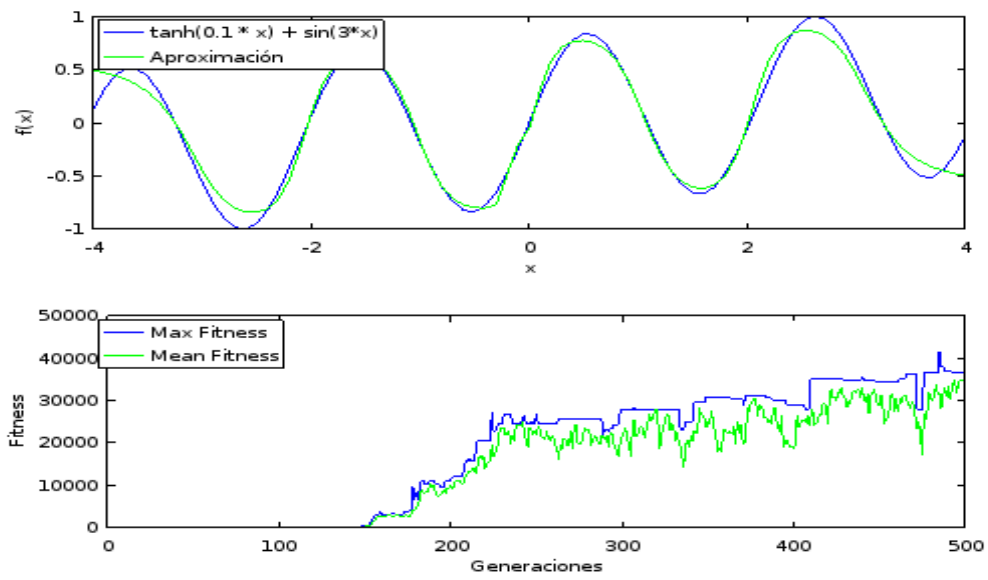


Figura 2.4: Gráfico de resultado óptimo para el método de reemplazo 3