

Sistemas de Inteligencia Artificial

Algoritmos Genéticos

Obtención de pesos de una red neuronal mediante algoritmos genéticos

José Ignacio Galindo (50211), Federico Homovc (50418), Nicolás Loreti (49479)

Instituto Tecnológico de Buenos Aires (ITBA)

10 de Junio - 2013

Palabras clave - algoritmos genéticos, redes neuronales, evolución, reemplazo, selección, mutación, individuo, gen, cromosoma, genoma, alelo, lotus.

1 - Introducción

Este trabajo describe el comportamiento de algoritmos genéticos para resolver el problema de obtención de pesos para redes neuronales multicapa. El objetivo será predecir el comportamiento de una serie de tiempo. Además se agregará la posibilidad de utilizar backpropagation combinando ambos métodos para obtener mejores resultados.

2 - Descripción del problema

El problema consiste en un conjunto de individuos que representan los pesos de una red neuronal dada una serie temporal que se debe predecir. El objetivo es predecir de la mejor manera posible los puntos de la serie mediante los valores anteriores utilizando algoritmos genéticos con distintas variantes y analizar los resultados obtenidos con cada uno.

3 - Resolución

3.1 - Modelado

3.1.1 Individuos

Una red neuronal se encuentra representada por varias matrices, cada una de ellas modelando una capa diferente. Cada individuo representa los pesos de una red neuronal.

Para poder facilitar el manejo de esta información a la hora de aplicar los algoritmos genéticos fue necesario construir un vector (*Cromosoma*) para almacenar los pesos. De esta forma cada posición del vector (*Locus*) representa un peso (*Alelo*) de la estructura original.

Luego a la hora de dar una respuesta será necesario aplicar la función inversa para transformar ese vector en la red neuronal que corresponda a ese individuo.

3.1.2 Algoritmo Genético

Dadas las numerosas opciones de configuración y pasaje de parámetros que puede recibir el programa, sumadas a las variaciones en los métodos de selección y reemplazo y la posibilidad de generar mutaciones y/o aplicar back-propagation, se pueden obtener resultados muy diferentes bajo la misma base teórica.

Otro factor importante en el desempeño del algoritmo es la Brecha generacional (G). Suponiendo una población de N individuos, para los métodos de reemplazo 2 y 3 se seleccionarán $k =$

$G * N$ individuos para aparear, mutar y realizar back propagation. Es decir, mientras mayor sea G más tenderá a cambiar la población y viceversa.

Luego, estos k individuos serán cruzados por alguno de los métodos de crossover, dependiendo de ciertas probabilidades podrán sufrir mutaciones o se le aplicará el algoritmo de backpropagation para finalmente formar parte de la generación siguiente.

De esta forma, la siguiente generación cuenta con k de los N individuos. Los $N-k$ restantes serán elegidos de entre los de la generación actual que bien no participaron de la selección anterior o son los padres de sus actuales hijos.

Por último, se permitirá el uso de métodos mixtos para la selección y el reemplazo de los individuos. El usuario puede elegir seleccionar una parte de los individuos a cruzar con el método de Ruleta o Boltzmann y el resto con Elite.

3.1.3 Función de Fitness

La función de fitness mide el grado de aceptación de un individuo dado un entorno particular. Para este problema se decidió tomar como función de fitness la inversa del error cuadrático medio (ECM). Es decir que el individuo más apto será el que menor ECM obtenga en la población analizada.

3.2 - Criterios de Selección y Reemplazo

3.2.1 Elitismo

El criterio de selección/reemplazo de Elitismo consiste en elegir los k individuos más aptos de determinada población, en este caso particular, a través de su *fitness*.

3.2.2 Ruleta

Este criterio de selección/reemplazo consiste en ponderar la probabilidad según el fitness del individuo en comparación con el fitness total de la población. Para esto se usan probabilidades acumuladas y la generación de un número random entre 0 y 1 para identificar qué individuo elegir.

3.2.3 Boltzmann

El criterio de Boltzmann se basa en dos parámetros: Una función no lineal de fitness y la tolerancia evolutiva (T). La tolerancia disminuye monótonamente con las generaciones. Al principio se busca diversidad, por lo que las probabilidades de elección de cada individuo son similares, pero luego de varias generaciones la tolerancia desciende haciendo que los individuos más aptos tengan más probabilidad de ser elegidos.

Esencialmente es igual que ruleta sólo que los p_i se calculan

$$p_i = \frac{e^{\frac{f_i}{T}}}{\sum_j e^{\frac{f_j}{T}}} \text{ donde } \langle T \rangle \text{ es el valor temporal de la serie.}$$

3.2.4 Torneo

En el criterio de selección/reemplazo por torneo se realizan los siguientes pasos:

1. Elegir dos individuos al azar.
2. Generar un número aleatorio r entre 0 y 1.
3. Si $r < 0,75$ se selecciona al individuo más apto, en caso contrario se selecciona al menos apto de acuerdo a su fitness.

4. Ambos individuos se devuelven al conjunto población por lo que podrían ser elegidos nuevamente en la siguiente ronda.

3.2.5 Mixto

Al seleccionar este tipo de criterio el usuario podrá ingresar un parámetro k (menor al total de individuos a seleccionar/reemplazar) y se procederá a usar el método elite para los primeros k/a (con a un valor mayor que 1 que se pasa como parámetro) y luego boltzmann o ruleta para la cantidad que reste de acuerdo a lo indicado por otro parámetro.

3.3 - Mutación

3.3.1 Clásica

Para realizar las mutaciones se toma un individuo y para cada locus se calcula un número aleatorio. Si ese número aleatorio es menor que una determinada probabilidad de mutación, entonces se genera otro número aleatorio entre -0.5 y 0.5 y se lo coloca como el valor del alelo en ese locus.

3.3.2 No Uniforme

Para realizar las mutaciones con este método se hace el mismo proceso que en el método anterior, pero decrementando cada $N = 0.1 * G$ generaciones la probabilidad de mutación en un factor pasado como parámetro.

3.4 - Criterios de Corte

3.4.1 Cantidad de Generaciones

El usuario podrá determinar la cantidad máximas de generaciones a correr. Una vez alcanzada esa cantidad, el algoritmo termina.

3.4.2 Estructura

Se termina el algoritmo al detectar que una porción importante de la población no cambia de generación en generación. Si un alelo de un individuo cambia más que un determinado error, entonces se considera que ese individuo cambió. Si el 70% de la población no cambia durante 5 generaciones, entonces se termina la ejecución del algoritmo.

3.4.3 Contenido

Se corta la ejecución al identificar que el mejor individuo de cada generación es el mismo. Si el fitness del mejor individuo no cambia más de determinado error en 5 generaciones, entonces se termina la ejecución del algoritmo.

3.4.4 Entorno al óptimo

Se termina la ejecución si se alcanza un fitness mejor que la cota especificada.

3.5 - Método de Selección

1. **Método 1:** Se seleccionan parejas de entre los individuos de la generación actual y se los aparea para generar los de la generación siguiente.
2. **Método 2:** Se seleccionan K individuos y se los aparea para generar K hijos. Luego se seleccionan de la población actual $N-k$ para completar la generación siguiente.

3. **Método 3:** Se seleccionan K individuos y se los aparea para generar K hijos. Luego de entre los padres e hijos ($N+k$) se seleccionan N para la generación siguiente.

3.6 - Métodos de apareamiento (Crossing Over)

1. **Clásico (un solo punto):** Se elige un locus al azar y se intercambian los alelos a partir de ese locus hasta el final.
2. **Cruce de dos puntos:** Se eligen dos locus al azar y se intercambian los alelos entre los dos locus.
3. **Cruce uniforme:** Se produce el cruce de un alelo en cada locus con una probabilidad determinada (0.65).
4. **Anular:** Se elige un locus y luego un segmento hacia la derecha de longitud k, perteneciente al rango $[1, L/2]$ donde L es la cantidad de locus. Se intercambian k alelos a partir del locus seleccionado volviendo a empezar desde el principio en caso de llegar al final.

3.7 - Backpropagation

Una vez finalizada la etapa de mutación dependiendo de una probabilidad p , en nuestro caso fijada en 0.01, se puede ejecutar el algoritmo de backpropagation una cantidad finita de pasos (100 épocas). De esta forma se logra converger más rápido hacia la solución óptima.

4 - Análisis de resultados

Para que el análisis de resultados sea consistente, se optó por dejar fijada la población inicial para los distintos casos de prueba.

En este trabajo se optó por usar una configuración de red neuronal con 3 entradas, una capa oculta de 5 neuronas y una salida. Esta elección se fundó en que con esta red se habían obtenido muy buenos resultados (siendo 0.0144 el mejor con 500 épocas) además de ser una de las redes con menor número de neuronas estudiadas. El mínimo error cuadrático obtenido con esta red constituye nuestro punto de comparación para determinar el desempeño de los algoritmos genéticos para la solución del problema. Tal como se puede apreciar en la *Tabla 1*, los mejores resultados utilizando los métodos explicados en este trabajo mejoraron el error cuadrático medio de la red constituyendo un resultado positivo.

A su vez, en la *Tabla 1* se puede observar como los dos mejores resultados se obtuvieron con un criterio de selección mixto en combinación con otro criterio de reemplazo. Esto se debe a que cuando se usa un criterio mixto para selección con un 60% elegido mediante Elite y el resto con Ruleta, se logra mantener una buena cantidad de individuos aptos, pero permitiendo mantener diversidad en la población para hacer posible la evolución. De esta manera, los apareamientos generan individuos con características diferentes de las de sus progenitores pero manteniendo muchos de los mejores aspectos ya que en su mayoría fueron elegidos con Elite. Por razones similares también se obtuvieron buenos resultados usando Boltzmann.

Como se puede ver en la *Tabla 2*, para todos los casos de criterios de selección y reemplazo analizados, dejando de lado el caso de Elite, el método de reemplazo 3 fue el que produjo individuos con más alto grado de adaptación (lo que se refleja en menores errores cuadráticos mínimos).

También se puede ver en la *Tabla 2* los peores resultados se obtuvieron con el criterio de selección y reemplazo Torneo.

En general resultó conveniente usar un gap generacional G de aproximadamente 0.85 ya que de ser más alto se perdían muchos buenos individuos y si era más bajo no se generaban suficientes combinaciones nuevas para mejorar el error.

Backpropagation fué un operador clave para mejorar el resultado de manera mucho más rápida, principalmente cuando se ejecutaba dentro de las primeras 10 a 15 generaciones. Sin embargo, a partir de este punto las mejoras eran pocas, poco significativas o nulas y por lo general las mutaciones y apareamientos eran los que mejoraban a los individuos.

5 - Conclusión

A partir de lo realizado en este trabajo se pudo comprender el funcionamiento de los algoritmos genéticos y sus diferentes estrategias, como así también llegar a las siguientes conclusiones:

1. Los algoritmos genéticos implementados son una mejor solución al problema que el uso de redes neuronales multicapa con aprendizaje supervisado para la configuración estudiada.
2. El criterio de selección mixto en combinación con otro criterio de reemplazo dió en general los mejores resultados para el problema estudiado.
3. Con el método de reemplazo 3 se obtuvieron en general los mejores resultados para este problema.
4. Torneo es el que peor resultados tiene a la hora de usarse como criterio de selección y reemplazo ya que da mucho peso a individuos con fitness más bajo y puede hacer que se pierdan los mejores individuos, lo que lo aleja de la solución.
5. La introducción de backpropagation en el algoritmo es fundamental para obtener buenas aproximaciones a la solución.

6 - Anexo

Criterio Selección	Criterio Reemplazo	Método Reemplazo	Cant. Individuos	Nº Generaciones	Brecha Generacional	Mín. Generalización	Apareos	Back Prop.
Mixto	Ruleta	3	100	30	0.6	0.010881	29	6
Mixto	Boltzman	3	100	30	0.9	0.010437	21	11
Elite	Elite	2	40	200	0.8	0.0109	140	17
Ruleta	Elite	2	100	40	0.8	0.01153	27	19

Tabla 1: Mejores Resultados

Criterio Selección	Criterio Reemplazo	Método Reemplazo	Cant. Individuos	Nº Generaciones	Brecha Generacional	Mín. Generalización	Apareos	Back Prop.
Elite	Elite	1	100	50	0.8	0.0123	1886	14
Elite	Elite	2	40	200	0.8	0.0109	140	17
Elite	Elite	3	50	100	0.8	0.0113	81	22
Boltzman	Boltzman	1	100	30	0.6	0.0138	1136	15
Boltzman	Boltzman	2	100	40	0.8	0.0275	27	21
Boltzman	Boltzman	3	100	100	0.9	0.0122	7	56
Ruleta	Ruleta	1	10	50	0.8	0.438	179	4
Ruleta	Ruleta	2	20	50	0.8	0.0221	38	4
Ruleta	Ruleta	3	30	20	0.8	0.0142	18	5
Torneo	Torneo	1	30	30	0.8	0.4391	358	6
Torneo	Torneo	2	20	30	0.6	0.4892	24	1
Torneo	Torneo	3	30	30	0.5	0.3894	23	1
Mixto	Mixto	1	100	30	0.5	0.0734	1115	7
Mixto	Mixto	2	100	30	0.8	0.0168	19	9
Mixto	Mixto	3	100	30	0.9	0.0116	29	16

Tabla 2: Comparación entre métodos de reemplazo