

Universidad de San Carlos de Guatemala
 Facultad de Ingeniería
 Escuela de Ciencias y Sistemas
 Inteligencia Artificial 1
 Vacaciones de diciembre de 2020
 Catedrático: Ing. Jorge Gutiérrez
 Auxiliar: Nery Galvez



Práctica 1

Tabla de Contenido

OBJETIVOS	2
OBJETIVO GENERAL.....	2
OBJETIVOS ESPECÍFICOS	2
DESCRIPCIÓN DE LA SOLUCIÓN.....	2
CARGAR ARCHIVO.....	3
CRITERIO DE FINALIZACIÓN	3
CRITERIO DE SELECCIÓN DE PADRES.....	3
GENERAR MODELO	3
CALCULAR NOTA	3
MODELO	4
ALGORITMO GENÉTICO.....	4
INDIVIDUO.....	5
POBLACIÓN	5
CRITERIO DE FINALIZACIÓN	5
SELECCIÓN DE PADRES	6
EMPAREJAR	6
CRUZAR	6
MUTAR	7
CÁLCULO DEL VALOR FITNESS.....	7
BITÁCORA	9
CONSIDERACIONES	9
FORMA DE ENTREGA	9
ENTREGABLES	9
FECHA DE ENTREGA	9

Objetivos

Objetivo General

Implementar una aplicación web donde se apliquen los conceptos adquiridos en el laboratorio de Inteligencia Artificial 1.

Objetivos Específicos

- Aplicar un algoritmo genético para encontrar una solución óptima a un problema en específico.
- Que el estudiante comprenda el funcionamiento de cada una de las fases de un algoritmo genético.
- Que la aplicación sea capaz de predecir la nota final de un alumno tomando como entrada la nota obtenida en sus 4 proyectos de clase.

Descripción de la Solución

El estudiante debe de crear una aplicación web que permita encontrar un modelo matemático que se adecue a los datos definidos en un archivo con extensión CSV y que permita predecir un valor tomando ciertas entradas, este modelo será calculado utilizando un algoritmo genético. A continuación, se muestra una posible interfaz y se definen las funcionalidades de la aplicación.

The screenshot shows a web browser window titled "A Web Page" with a URL bar containing "https://practica1". The main content area is divided into two columns. The left column contains three sections: "Seleccione un Archivo" with a "Cargar" button and an input field; "Seleccione un Criterio de Finalización" with three checkboxes labeled "Criterio 1", "Criterio 2", and "Criterio 3"; and "Seleccione un Modo de Selección de Padres" with three checkboxes labeled "Modo 1", "Modo 2", and "Modo 3". A "Generar Modelo" button is located at the bottom of the left column. The right column is titled "Predecir Nota del Estudiante" and contains four input fields labeled "Proyecto 1", "Proyecto 2", "Proyecto 3", and "Proyecto 4". A "Calcular Nota" button is positioned below these fields. The browser's address bar and navigation icons are visible at the top.

Cargar Archivo

Esta función permitirá cargar un archivo con extensión CSV a la aplicación, este archivo es el que contendrá los datos que ayudarán a encontrar el mejor modelo.

Criterio de Finalización

Esta función permitirá definir el criterio de finalización que utilizará el algoritmo genético.

Criterio de Selección de Padres

Esta función permitirá definir cómo el algoritmo genético seleccionará a los mejores padres de una generación.

Generar Modelo

Esta función ejecutará el algoritmo genético utilizando los parámetros definidos en la interfaz y cargará a memoria los datos de la solución encontrada por el algoritmo

Calcular Nota

Esta función solicitará como entrada las notas obtenidas por un estudiante en sus 4 proyectos de clase y mostrará la nota estimada por el modelo encontrado con el algoritmo genético.

.

Modelo

El archivo de entrada con los datos tendrá la siguiente estructura:

PROYECTO 1	PROYECTO 2	PROYECTO 3	PROYECTO 4	NOTA FINAL
13	17	3	14	11
93	2	32	36	38.25
0	12	1	55	24.1
82	48	2	36	34.5
22	10	75	43	44.5
23	88	33	5	28.55
84	83	4	74	55.85
94	56	70	84	77.1
92	82	30	75	65.1
19	32	52	34	36.85

Como el objetivo es predecir la nota final de un alumno según las notas obtenidas en los proyectos de clase, se propone el siguiente modelo:

$$NC = w_1P_1 + w_2P_2 + w_3P_3 + w_4P_4$$

Donde:

- NC es la Nota Calculada según el modelo
- p1, p2, p3 y p4 son las notas de cada uno de los proyectos
- w1, w2, w3 y w4 son constantes que relacionan las notas de los proyectos con la nota final del alumno.

Algoritmo Genético

El algoritmo genético será utilizado para calcular el valor de las constantes w1, w2, w3 y w4 que hagan que la predicción de la nota final sea lo más acertada posible. A continuación, se definen cada uno de los pasos del algoritmo que deben de implementarse.

Individuo

Un individuo representará una solución para las constantes del modelo definido anteriormente, por lo tanto, cada individuo tendrá un tamaño de 4. Ejemplo de un individuo:

$$\text{Individuo1} = [-0.5, 0.15, 0.80, -0.03]$$

$$\text{Individuo2} = [0.02, 0.55, -0.08, 0.01]$$

Población

La población es un conjunto de soluciones para el problema que se está resolviendo. El estudiante debe definir según su criterio el tamaño de la población, la única restricción será que este tamaño debe de ser mínimo de 10.

Cada elemento de una solución será definido por un número aleatorio en el rango de $[-2, 2]$.

Criterio de Finalización

Para finalizar el algoritmo el estudiante deberá de implementar 3 criterios distintos, recordar que este criterio se seleccionará desde la aplicación web. Los criterios a utilizar quedan a discreción del estudiante. A continuación, se presentan ejemplos de criterios de finalización:

- Un valor Máximo o Mínimo alcanzado por una solución de la población
- Máximo número de generaciones
- Un porcentaje de la población que tenga un valor fitness igual
- Un valor fitness promedio dentro de la población

Selección de Padres

La selección de los mejores padres de una población y la cantidad que se selecciona quedará a criterio del estudiante. Se deben de implementar 3 criterios de selección distintos. A continuación, se presentan ejemplos de criterios de selección.

- Selección aleatoria
- Selección por torneo
- Selección de los padres con mejor valor fitness
- Selección de padres en posiciones impares o solo en posiciones pares

Emparejar

Quedará a discreción del estudiante cómo se va a emparejar a los padres para generar nuevos hijos. Recordar que la cantidad de hijos a generar depende de la cantidad de padres seleccionados y el tamaño de su población. Por ejemplo, si se tiene una población de 30 individuos y se seleccionan 12 padres, debe de crear 18 hijos.

Cruzar

Cada una de las posiciones del nuevo hijo tendrá una probabilidad de 60% de ser del primer padre y un 40% de ser del segundo padre. Una forma de manejar la probabilidad es obtener un número aleatoriamente entre $[0, 1]$, si el número generado es menor o igual que 0.6 el hijo toma el valor (gen) del padre 1, de lo contrario toma el valor del padre 2.

Por ejemplo: si se generan 4 números aleatorios que son 0.25, 0.61, 0.45, 0.1 el resultado sería el siguiente:

Padre 1

w1	w2	w3	w4
----	----	----	----

Padre 2

w1	w2	w3	w4
----	----	----	----

Hijo

w1	w2	w3	w4
----	----	----	----

Mutar

Para la función de mutación un individuo tendrá un 50% de probabilidad de mutar y un 50% de no hacerlo. Si el individuo tiene que mutar, cada uno de sus elementos tiene la probabilidad del 50% de mutar y 50% de no hacerlo, por lo que por cada elemento se debe verificar si se aplica la mutación o no. La mutación consiste en sustituir el elemento por un valor aleatorio en el rango de $[-2, 2]$.

Cálculo del Valor Fitness

Para calcular el valor fitness de un individuo se utilizará la función del error cuadrático medio, la cual se define a continuación:

$$error = \frac{1}{N} \sum_{i=0}^N (NR_i - NC_i)^2$$

Donde:

- N es la cantidad de datos en el archivo de entrada.
- NR_i es la nota real del registro en la posición i del archivo de entrada.
- NC_i es la nota calculada con el modelo del registro en la posición i del archivo de entrada.

Ejemplo:

Se tiene la siguiente solución y los siguientes datos del archivo de entrada:

$$Individuo = [0.45, 0.2, 0.34, 0.15]$$

PROYECTO 1	PROYECTO 2	PROYECTO 3	PROYECTO 4	Nota Real
75	50	90	65	71.75
80	95	88	80	84.65
20	55	60	58	52.45
60	28	69	50	53.9

Para calcular el valor fitness del individuo se sustituyen sus valores (w_1 , w_2 , w_3 , w_4) en el modelo de predicción.

$$NC = 0.45P_1 + 0.2P_2 + 0.34P_3 + 0.15P_4$$

El siguiente paso es calcular la nota utilizando el modelo definido en el paso anterior (se obtiene la nota calculada).

PROYECTO 1	PROYECTO 2	PROYECTO 3	PROYECTO 4	Nota Real	Nota Calculada
75	50	90	65	71.75	84.1
80	95	88	80	84.65	96.92
20	55	60	58	52.45	49.1
60	28	69	50	53.9	63.56

El siguiente paso es aplicar la ecuación del error cuadrático medio.

$$error = \frac{1}{4} [(71.75 - 84.1)^2 + (84.65 - 96.92)^2 + (52.45 - 49.1)^2 + (53.9 - 63.56)^2]$$

$$error = \frac{1}{4} [152.52 + 150.55 + 11.22 + 93.32]$$

$$error = \frac{1}{4} [407.61]$$

$$error = 101.903375$$

El valor fitness asociado al individuo sería de 101.903375.

Como el valor fitness indica que tanto se aleja el valor calculado del valor real (error), para este proyecto las mejores soluciones son las que tendrán un valor fitness menor. Por ejemplo, de las siguientes soluciones la mejor solución es la solución 2.

Individuo1 = [w_1 , w_2 , w_3 , w_4], Fitness = 85

Individuo2 = [w_1 , w_2 , w_3 , w_4], Fitness = 50

Individuo3 = [w_1 , w_2 , w_3 , w_4], Fitness = 500

Individuo4 = [w_1 , w_2 , w_3 , w_4], Fitness = 240

Bitácora

La aplicación debe de llevar un registro de los resultados obtenidos de ejecutar el algoritmo (Generar Modelo). Los datos de cada registro en la bitácora son los siguientes:

- Fecha y hora de ejecución
- Nombre del documento CSV utilizado
- Criterio de finalización utilizado
- Criterio de selección de padres utilizado
- Número de generaciones generadas
- Mejor solución encontrada por el algoritmo

La forma de almacenar la información y cómo acceder a la bitácora queda a discreción del estudiante. Lo que se calificará es que exista y su contenido.

Consideraciones

- La aplicación deberá desarrollarse utilizando Python 3.
- El framework para la aplicación web queda a discreción del estudiante.
- El proyecto se realizará de manera individual, si se detecta algún tipo de copia el laboratorio quedará anulado.
- Se debe de crear un repositorio para llevar control de su proyecto.

Forma de Entrega

Entregables

- El entregable será el enlace al repositorio del proyecto, recordar que el auxiliar debe de tener siempre acceso a dicho repositorio, de lo contrario se anulará la nota obtenida en el proyecto.

Fecha de entrega

Domingo 13 de diciembre de 2020, hasta las 23:59