

Predicción de Ingresos de Causas Penales utilizando Programación Genética Lineal

Alberto David Garcete Rodríguez¹ y Benjamín Barán²

¹ adgr_x@hotmail.com – ² bbaran@cba.com.py

Universidad Nacional del Este

Campus Km. 8 Acaray, Calle Universidad Nacional del Este y Rca. del Paraguay

Ciudad del Este – Paraguay

Resumen. Este trabajo propone una metodología de predicción de ingresos de causas penales utilizando Programación Genética Lineal (*Linear Genetic Programming* - LGP). El estudio se realizó en base a datos mensuales recogidos durante siete años (2007 a 2013), en los siete Juzgados Penales de Garantías de Ciudad del Este - Paraguay. La verificación del método propuesto se hizo mediante la comparación del método LGP implementado con modelos estadísticos conocidos como la regresión lineal, promedio móvil, suavizado exponencial y suavizado exponencial con tendencia, prediciendo valores sobre una serie de tiempo, de forma a comparar los errores promedios de cada metodología de predicción. Se utilizaron dos métricas de error: (1) el error cuadrático medio y (2) el error absoluto medio. Resultados experimentales demuestran la superioridad del LGP implementado sobre los demás métodos estadísticos para la predicción de ingresos de causas penales.

Palabras clave: Programación Genética - GP, Programación Genética Lineal - LGP, ingresos penales, predicción.

1 Introducción

El aumento de delitos comunes moviliza todo el aparato judicial afectando el número de juicios en los estrados judiciales que se va incrementando año a año [1] y [2].

Investigar el aumento de delitos resulta beneficioso para mejorar la eficiencia del sistema judicial y analizar los factores que pueden incidir en estos hechos delictivos. En consecuencia, este trabajo propone realizar por primera vez predicciones relativas al número de delitos utilizando un algoritmo evolutivo y compararlo con métodos estadísticos tradicionales como la regresión lineal, promedio móvil, suavizado exponencial y suavizado exponencial con tendencia, para permitir una adecuada predicción que facilite la planificación de los recursos necesarios para que el Poder Judicial pueda cumplir con sus funciones específicas, sean estos: recursos humanos, espacio físico o insumos, entre otras necesidades específicas. El trabajo fue elaborado en base a datos históricos que corresponden a un periodo de siete años transcurrido entre enero/2007 y diciembre/2013.

2 Trabajos previos

Azamathulla, Guven y Demir presentan en [3] la metodología LGP, como una herramienta alternativa en la predicción de profundidad para tuberías sumergidas. El modelo LGP propuesto se comparó con sistemas adaptativos de inferencia *neuro-fuzzy* (ANFIS), observando que las predicciones utilizando LGP logran resultados mejores que los obtenidos con ANFIS o con una tradicional ecuación de regresión.

En [4], Guven y Kisi estudiaron los datos climáticos diarios para estimar la evaporación utilizando la programación genética lineal. Las estimaciones del modelo LGP se compararon con los resultados de otras técnicas de predicción, concluyéndose que la técnica LGP es apropiada para el modelado del proceso de evaporación.

En [5], Shavandi y Ramyani utilizaron LGP para predecir la radiación solar global. Los resultados indicaron que los modelos LGP dan estimaciones precisas de la radiación solar global y superan significativamente al modelo tradicional.

A su vez, Alavi, Gandomi y Mollahasani presentan en [6] un modelo del LGP así como un algoritmo híbrido de búsqueda que combinaba el algoritmo LGP con *Simulated Annealing* (SA), denominado LGP/SA, logrando predecir características de rendimiento de un suelo estabilizado. En particular, los modelos basados en LGP encontraron valores más precisos que los modelos híbridos basados en LGP/SA.

3 Programación Genética (GP)

La computación evolutiva propone un conjunto de modelos computacionales, que desarrollan automáticamente programas de computadoras que evolucionan generación tras generación resolviendo determinados problemas [7]. En este contexto, la Programación Genética GP, es una técnica de aprendizaje automático que se utiliza para optimizar una *población* de programas utilizando una función aptitud (*fitness*) [11], responsable de evaluar la calidad de cada programa.

La potencialidad de la programación genética reside en que permite desarrollar programas de forma automatizada, una vez definido el objeto de este programa y la métrica de comparación entre diferentes alternativas (o soluciones). Su base biológica es la misma que la de los algoritmos genéticos, es decir, el principio de evolución de las especies propuesto por Darwin [12]. Las primeras propuestas de GP utilizaban árboles sintácticos aplicados a un lenguaje de programación para representar a los individuos [10]. Posteriormente, aparecieron versiones más sencillas de implementar, como la Programación Genética Lineal (*Linear Genetic Programming - LGP*) [3], tratado en la siguiente sección.

4 Programación Genética Lineal (LGP)

En la metodología LGP los programas son representados como una secuencia de instrucciones de un lenguaje imperativo de programación; un ejemplo sería:

$$r[0] = r[2] + r[6]$$

En este ejemplo, $r[i]$ representa al registro i . El registro $r[0]$ es el registro destino y los registros $r[2]$ y $r[6]$ son los registros donde están los *operandos* de la operación realizada, en este caso una suma. Para este trabajo, las instrucciones están compuestas de la siguiente forma:

- Un registro destino donde se almacena el resultado.
- Uno o dos registros *operandos* donde se almacenan los valores con los que se operará.
- Una operación (de un conjunto de operaciones definidas a priori).

Entonces una representación típica de un programa obtenido por un algoritmo LGP podría ser de la siguiente forma:

$$\begin{aligned}r[6] &= r[3] + r[1] \\r[1] &= \text{seno } r[15] \\r[0] &= r[1] / r[2]\end{aligned}$$

El Algoritmo 4.1 que sigue fue diseñado para atender las necesidades propias del problema de predicción de causas penales estudiado en el presente trabajo. Para ello, se ha realizado una adaptación del algoritmo presentado por Guven y Kisi en [4].

Algoritmo 4.1: Programación Genética Lineal implementada

- 1: Inicializar aleatoriamente una población de programas.
- 2: Seleccionar los programas de la población y compararlos de acuerdo a su *fitness*, usando el *SelectorRuleta* [9].
- 3: Modificar los programas (individuos) utilizando uno o más de los siguientes operadores de variación:
 - **Reproducción:** copia un programa sin cambiarlo.
 - **Cruzamiento:** intercambia subestructuras (genes) entre 2 programas (o individuos).
 - **Mutación:** se utilizan dos tipos de mutaciones. La *micro-mutación* que modificar un único elemento (registro u operador) de la instrucción mutada, y la *macro-mutación* que inserta o borra una instrucción completa.
- 4: Se construye una nueva población con los programas variados en el paso 3 y se calcula el *fitness* de cada nuevo programa.
- 5: Si el criterio de parada no se cumple, volver al paso 2.
- 6: Parar: el programa con mejor *fitness* en la última población representa la solución encontrada por el LGP implementado.

El conjunto de registros junto con el conjunto de operaciones o funciones son los componentes a partir de los cuales la metodología LGP, trata de construir un programa de computador para solucionar un problema. El número de registros se define a priori [8] y para este trabajo solo se utilizaron 22 registros. El registro $r[0]$ es el que almacena la salida del algoritmo. Los registros $r[10]$ al $r[15]$ son de entrada, conteniendo información introducida desde la fuente de datos conocida, típicamente datos de la serie temporal estudiada. El conjunto de funciones escogidas para este trabajo incluye 4 funciones aritméticas, 3 exponenciales y 2 trigonométricas.

En este trabajo, la aptitud [9] de un programa se estima ponderando el error que se produce al predecir un valor futuro utilizando un programa generado por el algoritmo LGP (individuo cuyo *fitness* se desea estimar). Entonces, el mejor individuo (o programa) será aquel que tenga el menor error de predicción.

La función de error más utilizada es el error cuadrático medio [8], dado en (1). Así, se definió el error “e” dado por (1) como el promedio de los errores cuadrados para “N” valores de la muestra, donde “ x_i ” representa el valor real de la muestra “i” mientras que “ p_i ” representa el pronóstico (o estimación) para el instante discreto “i”.

$$e = \frac{1}{N} \sum_{i=1}^N (x_i - p_i)^2 \quad (1)$$

Cabe recordar además que las operaciones genéticas son aplicadas sobre una población que evoluciona de generación en generación. En consecuencia, se deben definir los parámetros que controlarán la ejecución del algoritmo LGP, como ser: el tamaño de la población (150 individuos), la longitud máxima de un individuo (200 instrucciones), las probabilidades de aplicación de los operadores genéticos de cruce y mutación (20% y 30% respectivamente); entre otros parámetros [3], cuyos detalles no se presentan por falta de espacio.

5 Métodos Clásicos de Predicción

Este trabajo utiliza también métodos estadísticos de predicción de series de tiempo para compararlos con LGP implementado. Estos métodos son brevemente presentados a continuación.

5.1 Regresión Lineal

En una regresión lineal, se asume que existe una relación entre una variable x (independiente) y otra variable y (dependiente) que se representa como una recta [13].

$$y = mx + b \quad (2)$$

Los valores de “b” y “m” se eligen de manera a minimizar el error dado en (1).

5.2 Promedio Móvil

Es el promedio aritmético de las “N” observaciones más recientes [13], dado por:

$$S_t = \frac{1}{N} \sum_{i=t}^{t-N+1} D_i \quad (3)$$

donde “ D_i ” representa el dato disponible para el instante de tiempo “i”. El inconveniente de este modelo es la pérdida de datos iniciales de la serie original [13].

5.3 Suavizado Exponencial

Es uno de los métodos más utilizados [14]. Utiliza un mecanismo de corrección que ajusta los pronósticos, donde las ponderaciones disminuyen exponencialmente y está dado por:

$$F_t = \alpha D_{t-1} + (1 - \alpha)F_{t-1} \quad (4)$$

donde α representa la constante de suavizado que determina la ponderación relativa en la observación. Alternativamente, si existe una clara tendencia en los datos de la serie temporal, este método puede ser mejorado estimando esta tendencia, en lo que se da en llamar *Suavizado Exponencial con Tendencia*, que se presenta a continuación.

5.4 Suavizado Exponencial con Tendencia

Además de utilizar el parámetro α arriba presentado, este método requiere un segundo parámetro β que suavice la tendencia. En la mayoría de las aplicaciones [14], se utiliza:

$$\beta \leq \alpha \quad (5)$$

Las ecuaciones usualmente utilizadas con el método suavizado exponencial con tendencia son [14]:

$$S_t = \alpha D_{t-1} + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (6)$$

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1} \quad (7)$$

$$F_t = S_t + T_t \quad (8)$$

6 Pruebas Experimentales y Principales Resultados

Las pruebas experimentales fueron realizadas comparando el método LGP, con los métodos estadísticos seleccionados (regresión lineal, promedio móvil, suavizado exponencial y suavizado exponencial con tendencia) [13] y [14] utilizando los datos disponibles para el periodo comprendido entre enero del 2007 y diciembre del 2013, extraídos del sitio web de la Excma. Corte Suprema de Justicia [1] y [2]. Las pruebas se realizaron con la siguiente distribución de datos:

<i>Entrenamiento o prueba</i>	:	72 meses.
<i>Validación o Pronóstico</i>	:	12 meses.
<i>Total de datos disponibles</i>	:	84 meses (enero/2007 a diciembre/2013)

Las evaluaciones de las metodologías implementadas se realizaron considerando dos métricas de desempeño: (1) el error cuadrático medio y (2) el error absoluto medio dado por:

$$e = \frac{1}{N} \sum_{i=1}^N |x_i - p_i| \quad (9)$$

Procedimiento utilizado con el método LGP:

- La población de individuos es de estado de reposo, donde unos padres son reemplazados por los hijos a cada generación.
- El algoritmo realiza la selección de individuos utilizando el método de la ruleta [9]. A los individuos seleccionados se les aplican los operadores evolutivos (selección, cruzamiento y mutación), para generar los individuos de la siguiente población.
- El pronóstico se realizó con los últimos 12 meses, aplicando la regla obtenida con el mejor individuo generado por el algoritmo LGP utilizando los 72 meses de prueba como entrenamiento.

Reglas obtenidas con el método LGP

La aleatoriedad del método LGP para obtener soluciones distintas en cada corrida, ha generado buenos resultados en la mayoría de las ejecuciones realizadas. Como ejemplo, la Tabla 1 presenta tres reglas (o programas de predicción) obtenidas por el algoritmo LGP que fueron comparados con los métodos estadísticos arriba presentados.

Tabla 1. Tabla con tres individuos (o programas) propuestos por el algoritmo LGP implementado en este trabajo

<i>Individuo A</i>	<i>Individuo B</i>	<i>Individuo C</i>
r[9] = r[15] - r[21]	r[6] = r[17]	r[4] = r[14] + r[6]
r[0] = r [7] + r[0]	r[6] = r [18] ^ r[3]	r[9] = r[6] * r[4]
r[1] = r[15] / r[19]	r[3] = sen[12]	r[5] = r[7] + r[17]
r[7] = r[1] + r[1]	r[6] = r[16] * r[11]	r[3] = r[13] + r[10]
r[0] = r[9]	r[6] = r[17] - r[21]	r[3] = sen[15]
r[8] = r[21] + r[0]	r[2] = r[5] - r[7]	r[5] = r[7] + r[17]
r[8] = r[20] + r[0]	r[6] = cosen[9]	r[4] = r[10]
r[0] = r[9]	r[0] = r[9]	r[0] = r [13] ^ r[5]
r[0] = r [8] + r[1]	r[3] = sen[18]	r[8] = r[8] / r[7]
r[0] = r [7] + r[0]	r[3] = ln[14]	r[0] = r[7] * r[2]
r[8] = r[21] + r[0]	r[4] = ln[6]	r[0] = r[1] - r[0]
r[0] = r [7] + r[0]	r[9] = sen[13]	r[4] = cosen[5]
r[8] = r[21] + r[0]	r[9] = ln[16]	r[6] = r[13] - r[1]
r[0] = r [7] + r[0]	r[5] = r[10] * r[7]	r[7] = r [16] + r[16]
r[8] = r[21] + r[0]	r[5] = r[16]	r[3] = cosen[16]
r[0] = r [8] + r[1]	r[0] = cosen[11]	r[4] = r[4] * r[6]
r[7] = r[1] + r[1]	r[5] = cosen[19]	r[0] = r[7] * r[2]
r[0] = r [7] + r[0]	r[4] = r[15]	r[4] = r [14] + r[6]
r[8] = r[21] + r[0]	r[4] = r[19]	r[5] = r [18] + r[16]
r[0] = r[19] * r[8]	r[8] = r[14] * r[20]	r[7] = r [16] + r[16]
	r[2] = sen[21]	r[0] = r [19] + r[12]
	r[7] = r[3] / r[5]	
	r[4] = sen[14]	
	r[2] = r[7] / r[4]	
	r[7] = ln[3]	
	r[0] = r[2] - r[5]	
	r[3] = sen[0]	
	r[7] = r[8] + r[9]	
	r[0] = r[7] * r[20]	

La Tabla 2 presenta los errores de predicción para los 12 meses el año 2013 utilizando el error cuadrático medio. En las últimas filas se presentan: el error total, el error medio y el *Ranking* correspondiente, verificándose el excelente desempeño de los individuos LGP(A) y LGP(B) calculado por el LGP implementado, y destacando que el LGP(C) es el peor de todos, lo que demuestra que no siempre el LGP calcula una buena regla de estimación, debido a su aleatoriedad inherente.

Tabla 2. Error cuadrático medio para cada solución evaluada durante el periodo de validación, ordenados con los mejores resultados a la izquierda y los peores a la derecha (ver fila: *Ranking*)

Mes	LGP(A)	LGP(B)	S. Exp.	S. Exp. Tend.	Regresión	P. Móvil	LGP(C)
Enero	14361,47	35628,71	67667,37	52617,80	59413,70	81035,11	47552,66
Febrero	3764,75	6809,11	6573,60	316,19	28828,59	3927,11	15892,57
Marzo	19565,15	9093,21	24134,67	47484,53	1100,22	35344,00	76139,73
Abril	10631,37	408,67	3506,06	7544,58	3,50	16129,00	27867,05
Mayo	1002,25	7740,42	9307,00	6562,36	18201,12	7511,11	28583,22
Junio	4022,77	15009,00	7548,77	3351,48	27209,20	16469,44	26917,56
Julio	33926,80	65216,27	28944,29	17540,12	80651,84	32881,78	22519,72
Agosto	17413,24	34560,86	2216,35	98,19	52915,29	1111,11	4364,68
Setiembre	32860,20	15342,73	82801,86	116682,34	7212,46	98177,78	135375,64
Octubre	461,95	6413,29	747,92	658,03	13482,55	841,00	12754,15
Noviembre	4321,74	71,97	5415,65	6725,72	737,39	3844,00	12783,86
Diciembre	2118,48	8396,44	4204,92	4523,74	18822,61	13378,78	486,90
TOTAL Error:	144450,19	204690,68	243068,46	264105,08	308578,47	310650,22	411237,72
Error Medio:	12037,52	17057,56	20255,70	22008,76	25714,87	25887,52	34269,81
Ranking	1°	2°	3°	4°	5°	6°	7°

Al utilizar la métrica del error absoluto medio para comparar las mismas siete soluciones en el mismo año 2013, se obtiene la Tabla 3, donde nuevamente el LGP es el método que encuentra las mejores estimaciones con los individuos LGP(A) y LGP(B) que consistentemente resultan así los mejores en el Ranking, por lo que serían los recomendados por este trabajo para realizar predicciones del número de causas judiciales en los Juzgados de Ciudad del Este – Paraguay.

Finalmente, se hace notar que según cuál sea la métrica de comparación (error cuadrático medio o error absoluto medio), los métodos *Suavizado Exponencial* y *Suavizado Exponencial con Tendencia* pueden quedar terceros o cuartos en el Ranking. Algo similar ocurre con el *Promedio Móvil* y la *Regresión Lineal*, que pueden quedar quinto o sexto, pero sin ninguna variación en los primeros puestos del Ranking que consistentemente corresponden a los individuos LGP(A) y LGP(B) calculados con el LGP implementado. Así mismo, el último puesto en ambos Rankings corresponde al individuo LGP(C), confirmando que un LGP no siempre realiza buenas predicciones, por lo que resulta recomendable realizar varias corridas del método antes de escoger la solución que será efectivamente utilizada.

Tabla 3. Error absoluto medio para cada solución evaluada durante el periodo de validación, ordenados con los mejores resultados a la izquierda y los peores a la derecha (ver fila: *Ranking*)

Mes	LGP(A)	LGP(B)	S. Exp. Tend.	S. Exp.	P. Móvil	Regresión	LGP(C)
Enero	119,84	188,76	229,39	260,13	284,67	243,75	218,07
Febrero	61,36	82,52	17,78	81,08	62,67	169,79	126,07
Marzo	139,88	95,36	217,91	155,35	188,00	33,17	275,93
Abril	103,11	20,22	86,86	59,21	127,00	1,87	166,93
Mayo	31,66	87,98	81,01	96,47	86,67	134,91	169,07
Junio	63,43	122,51	57,89	86,88	128,33	164,95	164,07
Julio	184,19	255,37	132,44	170,13	181,33	283,99	150,07
Agosto	131,96	185,91	9,91	47,08	33,33	230,03	66,07
Setiembre	181,27	123,87	341,59	287,75	313,33	84,93	367,93
Octubre	21,49	80,08	25,65	27,35	29,00	116,11	112,93
Noviembre	65,74	8,48	82,01	73,59	62,00	27,15	113,07
Diciembre	46,03	91,63	67,26	64,85	115,67	137,20	22,07
TOTAL Error:	1149,95	1342,68	1349,69	1409,88	1612,00	1627,86	1952,26
Error Medio:	95,83	111,89	112,47	117,49	134,33	135,66	162,69
Ranking	1°	2°	3°	4°	5°	6°	7°

7 Aportes al Poder Judicial

Un punto relevante determinado por este estudio fue la tendencia creciente de los ingresos penales (ver Figura 1). El mismo algoritmo LGP puede de hecho ser utilizado para analizar esta preocupante situación.

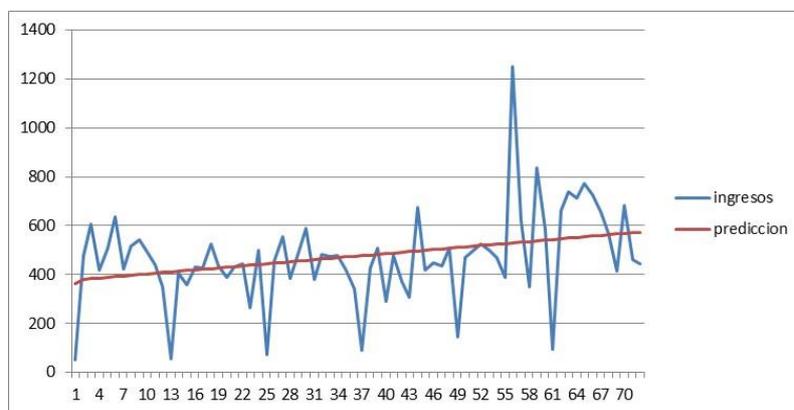


Fig. 1. Tendencia de crecimiento de los ingresos

La cantidad de causas estudiadas durante los 7 años son 38.305. El promedio anual de estas causas, arroja que cada despacho judicial estaría trabajando con un total aproximado de 5472 expedientes. Los datos del estudio demuestran que más de 6000 causas penales ingresaron en los últimos años y la tendencia es de un constante cre-

cimiento lo que significa que en un tiempo próximo, las autoridades deberán habilitar más juzgados penales de garantías en la localidad en cuestión, lo que conllevaría un gasto considerable, en base al costo que se genera en cada juzgado.

En consecuencia, este tipo de estudio realizado en diversas localidades aportaría a la Corte Suprema de Justicia con información relevante para ampliar en forma planificada los recursos que sean necesarios, minimizando así los contratiempos y demoras que hoy se vislumbra. Así, la herramienta propuesta permitirá advertir a las altas autoridades de las necesidades y tendencias antes de llegar a una situación extrema, proyecto con tiempo una solución adecuada, por ejemplo, invirtiendo más en prevención.

8 Conclusiones y trabajos futuros

El algoritmo LGP fue ejecutado varias veces y en la mayoría de las ejecuciones generaba resultados muy buenos al compararlos con otros métodos estadísticos más populares. Observando estos resultados y comparándolos con cada modelo matemático, se notó que las soluciones simulan y logran representar adecuadamente a las soluciones obtenidas por estos métodos matemáticos, con la diferencia relevante que el LGP logra generalmente mejores resultados.

El algoritmo LGP demuestra ser apto para ser aplicado como una metodología de predicción de los ingresos de causas penales de los Juzgados de Garantías de Ciudad del Este por la eficiencia obtenida en sus resultados, inclusive pudiendo ser extendido su aplicación para otras regiones del país. El éxito del LGP radica en que los resultados que genera se mantienen con un promedio de error bajo, quedando siempre entre los primeros lugares en cada predicción realizada. Cabe destacar que según sea la métrica de error considerada, el *Ranking* de comparación entre algoritmos presentados en las Tablas 2 y 3 puede diferir, por lo que no se logra establecer un único *Ranking*. Con todo, consistentemente, las mejores alternativas fueron siempre calculadas con el algoritmo LGP.

Las variantes que se le pueden aplicar al algoritmo LGP para su ejecución permiten obtener resultados experimentales exitosos para modelos de predicciones tanto lineales como no lineales, otro de los motivos por los que un LGP genera soluciones eficientes en comparación a los métodos estadísticos clásicos que demuestran limitaciones para predicciones de series de tiempos no lineales [8].

Por último, observando el crecimiento de los delitos y el incremento en la estructura de recursos humanos de la Excma. Corte Suprema de Justicia [1] y [2], se puede mencionar que con este tipo de trabajo, se podría realizar una reingeniería de recursos, alocando infraestructura y recurso humano donde más se los necesite, con razonable previsión. En consecuencia, aplicando este tipo de técnicas se obtendrían mejores resultados en el ámbito de la justicia, ya que ajustando los procesos se podrá ofrecer un mejor servicio en el área en la que se encuentra especializada, LA JUSTICIA.

Como posibles trabajos futuros para continuar con lo hasta aquí realizado, se propone:

1. Análisis de los parámetros óptimos de inicialización y de los valores de entrada a ser utilizados por un algoritmo LGP.
2. Comparación del algoritmo LGP con otros modelos no lineales de predicción, como las redes neuronales.
3. Estudiar problemas con múltiples salidas.
4. Implementación del algoritmo LGP con enfoques multi-objetivo.
5. Considerar aplicaciones concretas a otros juzgados y diferentes áreas del sector público que necesitan realizar estimaciones para mejorar su planificación en base a predicciones cada vez mejores.

Referencias

1. <http://www.csj.gov.py/>, Sitio De La Excm. Corte Suprema De Justicia, 2014.
2. <http://www.pj.gov.py/>, Sitio del Poder Judicial, 2014.
3. Azamathulla, H. M., Guven, A., & Demir, Y. K. Linear genetic programming to scour below submerged pipeline. *Ocean Engineering*, 38(8), 995-1000, 2011.
4. Guven, A., & Kişi, Ö. Daily pan evaporation modeling using linear genetic programming technique. *Irrigation science*, 29(2), 135-145, 2011.
5. Shavandi, H., & Ramyani, S. S. A linear genetic programming approach for the prediction of solar global radiation. *Neural Computing and Applications*, 23(3-4), 1197-1204, 2013.
6. Alavi, A. H., Gandomi, A. H., & Mollahasani, A. A Genetic Programming-Based Approach for the Performance Characteristics Assessment of Stabilized Soil. In *Variants of Evolutionary Algorithms for Real-World Applications* (pp. 343-376). Springer Berlin Heidelberg, 2012.
7. Koza, J. R. Human-competitive results produced by genetic programming. *Genetic Programming and Evolvable Machines*, 11(3-4), 251-284, 2010.
8. R. Sánchez, J. Martínez y B. Barán. "Economic Time-Series Forecasting Using Linear Genetic Programming", *Computational Intelligence in Economics and Finance - CIEF. 11th Joint Conference on In-formation Sciences JCIS' Kylin Villa – Shenzhen, China, 2008.*
9. Koza, J. R. *Hierarchical Genetic Algorithms Operating on Populations of Computer Programs*. N. S. Sridharan, Editor, *Proceedings Of 11th International Joint Conference On Artificial Intelligence*, San Mateo, Morgan Kaufmann, California, 1989.
10. Koza, J. R. Introduction to genetic programming tutorial: from the basics to human-competitive results. *Proceedings of the 12th annual conference companion on Genetic and evolutionary computation*, 2137-2262, Julio 2010. ACM.
11. Poli, R., & Koza, J. *Genetic Programming*, 143-185. Springer US, 2014.
12. Darwin, Ch. *The Origin of Species By Means Of Natural Selection or the Preservation of Favored Races in the Struggle for Life*. Random House, New York, 1993.
13. Hernández, R. *Estadística, Segunda Edición*, Madrid España, 1997.
14. Triola, Mario F., *Estadística Elemental, Séptima Edición*, México 2000.