

UN ALGORITMO INMUNOLÓGICO DE SELECCIÓN CLONAL APLICADO AL CALCULO DE MONTOS EN UN PORTAFOLIO DE INVERSIÓN

García-Mejía Juan Fernando, Barrera-Guzmán Diana, Martínez Garduño Yenit, Torres-Reyes Carlos Eduardo, Granda Gutiérrez Everardo Efrén, Pérez-Martínez José Arturo
Universidad Autónoma del Estado de México
Centro Universitario UAEM Atlacomulco
Carretera Libre Toluca-Atlacomulco Km 60 Atlacomulco Estado de México
Tel: 7121220446
fgarciam@uaemex.mx

RESUMEN.

Un problema de las ciencias económico-administrativas, en específico del área de finanzas es el cálculo de montos a invertir en un portafolio de inversión, que se define como una colección de activos financieros con la finalidad de obtener una ganancia máxima con un riesgo mínimo, esto por medio de una serie de ecuaciones matemáticas denominadas modelo de Markowitz, el cual suele ser solucionable por un experto financiero. Una alternativa de solución que garantiza la minimización de riesgo con una ganancia máxima es constituida por medio de algoritmos evolutivos, como el algoritmo genético canónico el cual implica un procedimiento con un número de iteraciones (convergencia) alto. En este trabajo se resuelve este inconveniente usando un algoritmo inmunológico de selección clonal, una técnica que se encuentra soportada por la teoría de inmunidad clonal de los mamíferos superiores

Palabras Clave: portafolio de inversión, selección clonal, convergencia, algoritmo genético

ABSTRACT.

One problem of economic-administrative sciences, specifically in the finance area is the calculation of amounts to invest in an investment portfolio, which is defined as a collection of financial assets in order to obtain a maximum profit with a minimal risk, this through a series of mathematical equations called the Markowitz model, which is usually solved by a financial expert. A solution alternative that ensures risk minimization with a maximum gain is made up of evolutionary algorithms, such as the canonical genetic algorithm which involves a procedure with a high number of iterations (convergence). This work solves this drawback using an immune algorithm of clonal selection, a technique that is supported by the theory of clonal immunity of the superior mammals

Keywords: investment portfolio, clonal selection, convergence, genetic algorithm

1. INTRODUCCIÓN

La determinación de los valores máximos y mínimos de una o varias funciones matemáticas representativas de un problema determinado de un dominio de conocimiento como la ingeniería química, la manufactura, el control automático o las finanzas puede conceptualizarse como un proceso de elección de variables denominado Optimización y puede o no estar sujeto a

determinadas condiciones denominadas restricciones, modelables por medio expresiones matemáticas.

La Optimización puede clasificarse numérica, cuando el dominio del problema es representado por medio de números reales, o combinatoria, cuando la representación de los elementos que conforman el dominio es determinada por medio de números enteros. Los conceptos antes definidos se puede expresar por medio de expresiones matemáticas 1,2,3 [1]

$$\min(\max)f(x), x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n \quad (1)$$

$$g_j(x) \leq 0 \forall (j = 1, 2, \dots, J) \quad (2)$$

$$h_j(x) = 0 \quad (3)$$

Donde, $f(x)$, $g_j(x)$ y $h_j(x)$ son funciones escalares del vector x . Los componentes de $x = [x_1, x_2, \dots, x_n]^T$ son denominados variables, $f(x)$ es la función objetivo, $g_j(x)$ y $h_j(x)$ son funciones que describen las condiciones de iniquidad e igualdad respectivamente. El vector óptimo x que resuelve la ecuación se denota por x^* con el correspondiente valor de optimización $f(x^*)$ [2].

Un problema de la vida cotidiana, que puede modelarse por medio de ecuaciones, es la diversificación de portafolios de inversión, los cuales son un conjunto de activos financieros que un inversionista (denominado tenedor) posee. Ejemplos de activos financieros son acciones de diversas bolsas de valores, opciones, bonos, *commodities* y moneda extranjera por mencionar algunos. La finalidad de un portafolio de inversión es generar para su tenedor ganancias económicas máximas con un margen de riesgo de pérdida mínimo. Esto se puede conseguir con diferentes posturas teóricas, que pueden ser sintetizadas en modelos. Un ejemplo de estas posturas/modelos es el modelo de varianza mínima o de Markowitz [3] el cual se fundamenta en el supuesto teórico de que el diseño del portafolio de inversión puede estructurarse en función de la expectativa de rendimiento medio y la desviación estándar de los rendimientos del portafolio.

El soporte teórico del modelo de Markowitz es atribuido a Harry Markowitz, el cual consiste en asumir que el valor de los activos financieros es el reflejo de las expectativas sobre el valor futuro, por lo tanto, se busca el portafolio que permita una maximización del rendimiento simultánea con una minimización del riesgo. Las expresiones matemáticas del modelo se muestran en las ecuaciones 4 y 5 [3]

$$G = \sum_i a_i * w_i \quad (4)$$

$$R = \sum_i \sum_j \sigma_{i,j} * w_i * w_j \quad (5)$$

Donde G es la ganancia del portafolio, R es el riesgo i, j son índices, $\sigma_{i,j}$ se denomina matriz de covarianza de las a acciones y w representan los porcentajes a invertir en estas y están sujetos a la restricción $\sum w = 1$.

Existen varios métodos de solución de las ecuaciones 4 y 5, es decir la determinación de los valores de w , los cuales se dividen en analíticos y estocásticos. El problema de diversificación de un portafolio de inversión puede ser resuelto por medio de técnicas de softcomputing, en específico por medio de algoritmos evolutivos.

La literatura especializada referencia a algunos ejemplos del uso de softcomputing y algoritmos evolutivos en la diversificación de portafolios de inversión. En [5] se optimizó un portafolio de inversiones con instrumentos bursátiles del año 2013 pertenecientes a la bolsa de valores de Colombia haciendo uso del modelo de multiplicadores de Lagrange, donde se define el procedimiento para encontrar los máximos y mínimos de funciones de múltiples variables sujetas a restricciones que permiten transformar el problema restringido de n variables a uno sin restricciones como $n+k$ variables, donde k es igual al número de restricciones.

Otro acercamiento al diseño de portafolios de inversión se encuentra en [6] donde se usaron redes neuronales para predecir el comportamiento futuro de estos siguiendo las siguientes etapas: recolección de los datos, entrenamiento de la red, ejecución de la red neuronal, posteriormente se obtienen los resultados y por último se realiza la predicción. Esto consiste en predecir cómo será el comportamiento de las variables en un tiempo determinado.

En [7] se realiza la optimización de un portafolio de inversiones mediante el uso del modelo Media-Varianza de Markowitz tomando como referencia los activos financieros de la Bolsa Mexicana de Valores en el periodo del 1 de enero de 1999 al 25 de febrero del 2014, usando como técnica de inteligencia artificial un algoritmo evolutivo multiobjetivo con codificación binaria, lo cual crea cromosomas de extensión larga. En [8] usando un contexto de solución basado en Frentes

de Pareto se optimizó un portafolio realizado con acciones de la Bolsa Colombiana de Valores usando como función objetivo el modelo de Markowitz.

En este trabajo se propone, a partir de lo revisado en la literatura especializada, la construcción de un portafolio de inversión constituido por activos de la Bolsa Mexicana de Valores, cuyo cálculo de los valores w es realizado por medio de un micro Algoritmo de Selección clonal un método heurístico soportado en la teoría del sistema inmune los mamíferos superiores caracterizado por su número pequeño de posibles soluciones

2. ALGORITMO INMUNOLÓGICO

Una descripción simplificada del sistema inmune es un sistema de órganos destinado a proteger al organismo huésped de las amenazas que le plantean los patógenos y las sustancias tóxicas. Los patógenos abarcan una variedad de microorganismos como bacterias, virus, parásitos y polen. La perspectiva tradicional con respecto al papel del sistema inmune se divide en dos tareas principales: la detección y eliminación del patógeno. Este comportamiento generalmente se conoce como la diferenciación de uno mismo (moléculas y células que pertenecen a los organismos hospedantes) del no ser potencialmente dañino.[9]

Los algoritmos inmunes pertenecen al campo de estudio de los sistemas inmunes artificiales relacionados con los métodos computacionales inspirados en el proceso y los mecanismos del sistema inmunológico biológico, surgieron a principios de la década de 1990, basado en la propuesta de fines de la década de 1980 de aplicar modelos teóricos de sistemas inmunológicos al aprendizaje automático y a la solución automatizada de problemas [9].

El algoritmo de selección clonal (CLONALG) está inspirado en la teoría de la selección clonal de la inmunidad adquirida propuesta por Frank Macfarlane Burnet un biólogo australiano que en 1978 demostró y explicó el comportamiento y las capacidades de los anticuerpos en el sistema inmune biológico, de los mamíferos superiores. Cabe destacar que la teoría de Burnet es paralela a la teoría de la evolución de Darwin[9]

La teoría de la selección clonal se soporta en los principios de la teoría de la evolución natural de la selección darwiniana, la teoría propone que los antígenos seleccionen los linfocitos (células B y T). Cuando se selecciona un linfocito y se une a un determinante antigénico, la célula prolifera haciendo miles de copias más de sí misma y se diferencia en diferentes tipos de células (células plasmáticas y de memoria). Las células plasmáticas tienen una vida útil corta y producen grandes cantidades de moléculas de anticuerpos, mientras que las células de memoria viven durante un período prolongado en el

huésped anticipando el reconocimiento futuro del mismo determinante.

La teoría de selección clonal propone que, a partir de un repertorio inicial de células inmunes generales, el sistema puede cambiarse a sí mismo (las composiciones y las densidades de las células y sus receptores) en respuesta a la experiencia con el medio ambiente. Mediante un proceso ciego de selección y variación acumulada a gran escala de miles de millones de células, el sistema inmunitario adquirido es capaz de adquirir la información necesaria para proteger al organismo huésped de los peligros patógenos específicos del medio ambiente. También sugiere que el sistema debe anticipar el patógeno al que estará expuesto, y requiere exposición al patógeno que puede dañar al huésped antes de que pueda adquirir la información necesaria para proporcionar una defensa [10].

Los principios de procesamiento de información de la teoría de la selección clonal describen una estrategia general de aprendizaje, por lo tanto, es aprovechable para el diseño de un algoritmo evolutivo. Esta estrategia involucra una población de unidades de información adaptativas (cada una representando una solución o componente del problema) sometidas a procesos competitivos para la selección, que junto con la duplicación y variación resultante finalmente mejora el ajuste adaptativo de las unidades de información a su entorno.

Los equivalentes computacionales a los principios biológicos descritos en esta sección se muestran en la tabla 1

Tabla 1 Equivalencias biológicas del principio inmune de selección clonal y el algoritmo CLONALG

Teoría del principio inmune de selección clonal	Algoritmo CLONALG
antígeno	El antígeno es el equivalente a una función objetivo de la teoría de Optimización tanto numérica como combinatoria
linfocito	Es el vector solución a la función objetivo
afinidad	El valor que se obtiene de la función objetivo al aplicar un antígeno. La idea

Cabe destacar el el algoritmo CLONALG, en su versión canónica fue desarrollada por en 2002 por Nunes de Castro and Fernando J. Von Zuben de la Universidad Campinas de Brazil a partir del los principios desarrollados por Burnet[9].

La figura 1 muestra el diagrama de flujo de un algoritmo CLONALG[10]

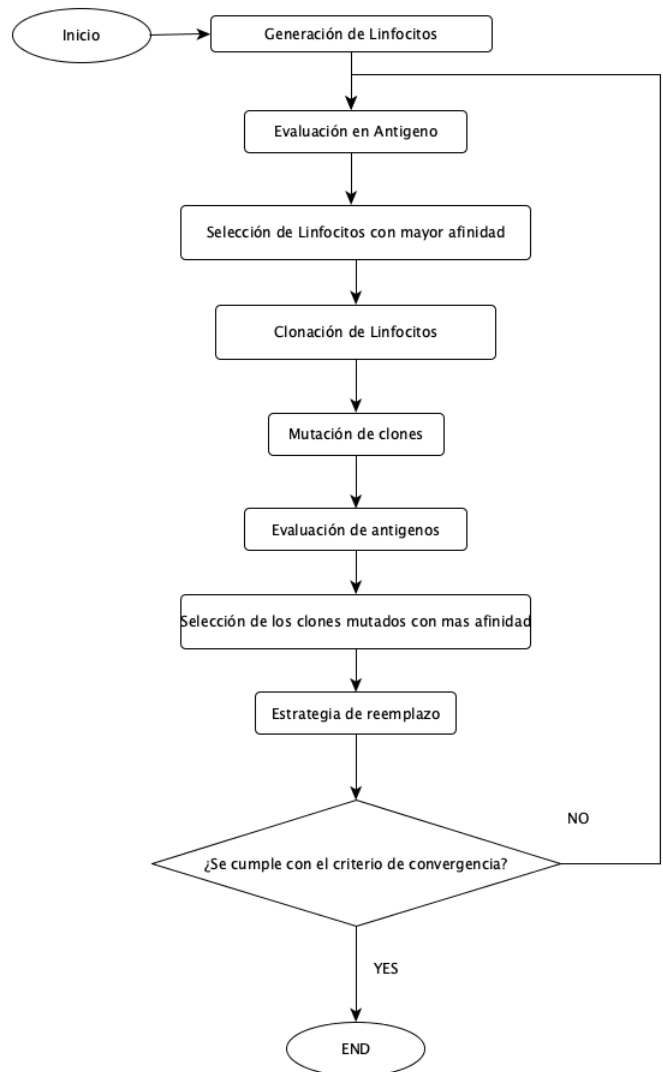


Figura 1 Diagrama de flujo del algoritmo CLONALG

3. METODOLOGÍA

Como se demostró en las secciones anteriores de este documento el algoritmo CLONALG necesita una función objetivo denominado antígeno, la cual es contruida por la relación entre la ganancia y el riesgo escritas en las ecuaciones (4) y (5) obteniendo (6)

$$\text{maximizar } \frac{\sum_{i=1}^N X_i * \bar{r}_i}{\sum_{i=1}^N \sum_{j=1}^N X_i * X_j * \sigma_{ij}} \quad (6)$$

Los linfocitos propuestos en este trabajo estan compuestos por un vector, cuyos componentes representan los montos a invertir

en cada uno de los activos del portafolio y pueden ser representados por medio de la expresión (7)

$$\vec{X} = (X_1, X_2, \dots, X_N)^T \quad (7)$$

Donde X_1 es el monto a invertir en el activo 1, X_2 es el monto a invertir en el activo 2 y X_N es el monto a invertir en el N-esimo activo, que para esta propuesta tiene un valor de 10

Cabe destacar que los activos del portafolio de inversión usado en este trabajo es construido a partir de activos de la Bolsa Mexicana de Valores en un ventana de tiempo de 1 año para poder construir la la matriz de riesgo covarianza σ , los cuales fueron sugeridos por un financiero. La tabla 2 muestran los valores de los activos seleccionados para esta propuesta

Tabla 2 Valores de los activos seleccionados para el portafolio de inversión

Stock	Stock value
Aeromex	40.02
Bimbo	51.27
Cemex	9.57
Elektra	351.7
Herdez	42.99
Ienova	73.92
Kimber	40.75
Oma	85.73
Soriana	40.51
Walmex	42.26

El número de clones de cada linfocitos seran proporcionales a la afinidad con respecto al antígeno y es calculado por medio de la expresión (8) mientras que el porcentaje de mutación de cada clon es determinada por medio de la expresion 9 [11]

$$Cl_n = \text{round} \left(\frac{Nl^* \beta}{n} + 0.5 \right) \quad (8)$$

Donde

Cl_n número de clones para el n -esimo linfocito

Nl número total de linfocitos

β factor de clonación (constante) = 0.1

n número de linfocitos a clonar

$$\alpha = \exp \left(-\rho f^*(C_i) \right) \quad (9)$$

Donde

$f^*(X_i) \in [0,1]$ es la aptitude normalizada para casa (C_i)

ρ es la constante de decaimiento que determina la tasa de mutación

De tal forma que el operador de hipermutación es deeterminado por medio de la expresión 10 [12]

$$\tilde{C}_i = \begin{cases} C_i + \lambda(C_{r1} - C_{r2}), & \forall \alpha \leq p_m \\ C_i, & otherwise \end{cases} \quad (10)$$

\tilde{C}_i , i -esimo clon mutado

C_i clon a mutar

λ factor de mutación $\text{rand}[0,1]$

p_m probabilidad de mutación

Con la finalidad de realizar un contraste se presenta una comparación entre CLONALG y un algoritmo Genético Canónico (GA) como el presentado en [13], con un operador de cruzamiento BLX-alpha A partir de esto se propone un conjunto experimental descrito en la tabla 3, propuesto en terminos generales, despues de una revisión de la literatura especializada

Tabla 3 Conjunto experimental

Test 1 (CLONALG)			
linfocitos	linfocitos a clonar	Tasa de mutación	Probabilidad de mutación
100	10	Random [0,1]	Random [0,1]
Test 2 (GA)			
Cromosomas	Porcentaje de cruza	Porcentaje a mutar	--
100	0.8	0.1	--

4. METODOLOGÍA

Despues de aplicar 40 ejecuciones sobre el set experimental descrito en la la tabla 3 se aplicó el test de Kolmogorov-Smirnov con la finalidad de determinar las distribuciones estadísticas de las soluciones reportadas por cada uno de los algoritmos. El resultado se muestra en la tabla 4

Tabla 4 Resultados del Test de Kolmogorov

Test	p -valor test de Kolmogorov
1	1.727e-08
2	0.5418

Los valores p del test de Kolmogorov que se obtienen en el test 2 es mayor a 0.05 (valor de significancia), por lo tanto es posible asumir que los datos presentan una distribución que se ajustan a una distribución normal, por lo contrario el test, correspondiente a CLONALG que tiene un p valor menor a 0.05, lo que implica una distribución diferente a una campana

de Gauss implicando que se trata de dos algoritmo con efectos diferentes. Por lo tanto no es necesario aplicar el test de Anova, dado que las diferencias entre GA y CLONALG fueron determinadas a partir del hecho de tener distribuciones de datos diferentes.

El siguiente criterio de desempeño de los tests propuestos es la convergencia la cual se muestran en la figura 2 para el CLONALG y la figura 3 para GA

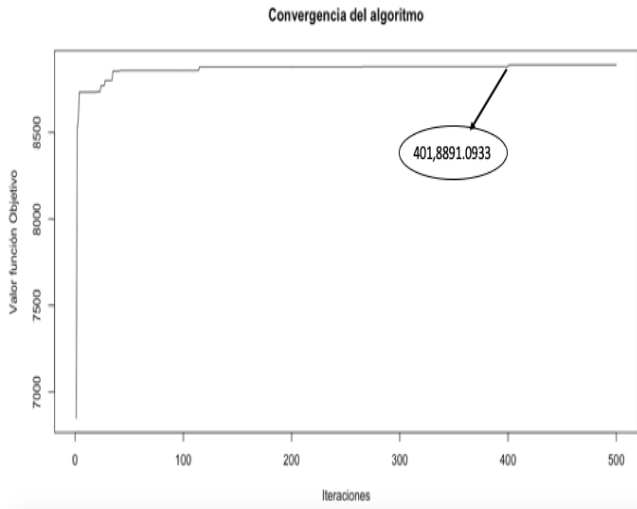


Figura 2 Convergencia para CLONALG

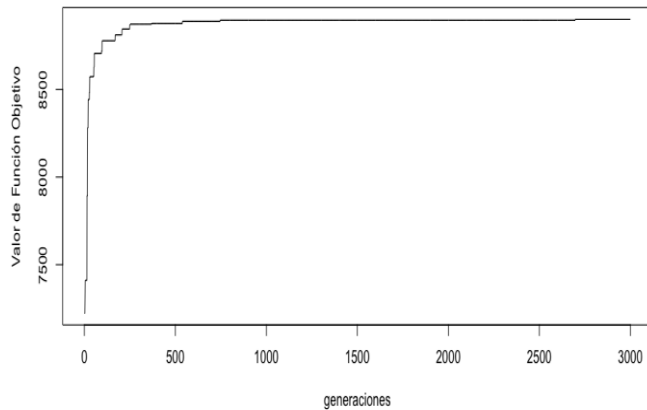


Figura 3 Convergencia para GA

Como lo muestran las figuras 2 y 3 la convergencia mas rápida es la del algoritmo CLONALG con respect al GA. Ahora a partir de las distribuciones de probabilidad es posible evaluar el coeficiente de variabilidad definido como la relación entre las desviación estandar y la media. Además se muestra la mediana y el rango intercuartilico para el CLONALG. Esto se observa en la tabla 5

Tabla 5 Medidas de Tendencia Central para el conjunto Experimental propuesto

Test 1		
Mediana	Distancia Intercuartilica	
8886.272	6.787923	
Test 2		
media	Desviación estandar	Cofiente de variación
8901.981701	3.587786303	0.0004030

Las implicaciones de estos resultados son discutidas en la sección de conclusiones. Las tablas 6 y 7 muestran los portafolios obtenidos a partir del conjunto experimental

Tabla 6 Portafolio obtenido con CLONALG

Activo	Monto a invertir
Aeromex	0.076631954
Bimbo	0.00027424159
Cemex	0.0063488394
Elektra	0.38579492
Herdez	0.1508086
Ienova	0.11485962
Kimber	9.1590199e-05
Oma	0.10589933
Soriana	4.8529663e-05
Walmex	0.15924238

Tabla 7 Portafolio obtenido con GA

Acción	Monto a invertir en la acción
Aerome x	0.06571839
Bimbo	0.007191815
Cemex	0.014509057
Elektra	0.375408833
Herdez	0.130308515
Ienova	0.10654552
Kimber	0.018581411
Oma	0.101224583
Soriana	0.063292582
Walmex	0.117219295

5. CONCLUSIONES

La asignación o calculo de los montos a invertir en un portafolio de inversión suele ser un problema de interes para las áreas económico administrativas, pero desde un punto de

problema de optimización en específico numérico, se vuelve un problema de softcomputing y en particular de los algoritmos evolutivos dado que la cantidad de combinaciones (que suele tender a infinito) implica una alta complejidad computacional, por lo tanto no es tratable por un especialista humano. Al contrastar la propuesta presentada (CLONALG) con un GA se tienen los siguientes hechos

- Un GA permite la diversificación de los montos de inversión de un portafolio
- Un CLONALG también diversifica los montos de inversión de un portafolio
- Un CLONALG tiene una convergencia más rápida que el GA
- La distribución de los resultados de un GA presenta una distribución de datos normal, mientras que un CLONALG no la tiene, el tener dos distribuciones distintas significa que los resultados estadísticamente diferentes
- El rango intercuartil de CLONALG indica que el 50% de los resultados están en un rango de 8893.0599 a 8879.4841, prácticamente similar a GA
- No es conveniente usar la desviación estándar y la media sin determinar si la distribución es normal, si es diferente a esta se aconseja el rango intercuartil y la mediana
- A partir de los resultados, en virtud de la rápida convergencia que implica menos tiempo de cómputo se concluye que es mejor alternativa CLONALG con relación al GA

En la referencia [13] se usa una versión reducida microalgoritmo genético, como trabajo a futuro se propone la realización de una versión reducida del CLONAL y estudiar sus resultados por medio de estadística

Referencias.

- [1] Duarte Muñoz, A, Metaheurísticas, Madrid, España, Universidad Rey Juan Carlos Servicio de Publicaciones, 2007
- [2] Peñuela Meneses, C. A., y M., Granada Echeverri, "Optimización multiobjetivo usando un algoritmo genético y un operador elitista basado en un ordenamiento no-dominado (nsga-ii)". *Scientia et Technica*, 1(35), pp.175-180, 2007 doi:10.22517/23447214.5393
- [3] Blanco Ramos, F., Ferrando Bolado, M., y M. F., Martínez Lobato, Teoría de la Inversión. Madrid, España, Pirámide 2015.
- [4] Burger, C. Propeller performance analysis and multidisciplinary optimization using a genetic algorithm. Dissertation, Auburn University, Alabama 2007
- [5] Cruz Trejos, E. A., Medina Varela, P. D., y H. D., Salazar Arias (2016), Optimización de portafolios de acciones utilizando los multiplicadores de Lagrange. *Scientia et Technica*, 18(1), 114-119. doi:10.22517/23447214.8301
- [6] Robledo Escobar, J. P., y C., García Gómez Portafolios de Inversión a través de Redes Neuronales y Algoritmo Genético (Tesis de maestría). Pereira, Colombia: Universidad Tecnológica de Pereira 2008
- [7] Castro Enciso, S. F. Creación de Portafolios de Inversión utilizando Algoritmo Evolutivos Multiobjetivo (Tesis de Maestría). Ciudad de México, México: Instituto Politécnico Nacional 2005.
- [8] Henao Pérez, A. A. Creación de un Portafolio de Inversión constituido por valores primarios transados en la Bolsa de Valores de Colombia, empleando optimización multiobjetivo (Tesis de Maestría). Barranquilla, Colombia: Universidad del Norte 2007.
- [9] L. N. de Castro and F. J. Von Zuben, "Learning and optimization using the clonal selection principle," in *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 3, pp. 239-251, June 2002. doi: 10.1109/TEVC.2002.1011539
- [10] D. Ribatti, *Milestones in immunology: based on collected papers*. London, United Kingdom: Academic Press, an imprint of Elsevier, 2017.
- [11] A. I. Tauber and S. H. Podolsky, *The generation of diversity: clonal selection theory and the rise of molecular immunology*. Cambridge (Mass.): Harvard University Press, 2000.
- [12] V Cutello, G Nicosia, and MPavone. "Real coded clonal selection algorithm for unconstrained global optimization using a hybrid inversely proportional hypermutation operator." *Proceedings of the 2006 ACM symposium on Applied computing (SAC '06)*. Association for Computing Machinery, New York, NY, USA, pp 950-954. DOI:https://doi.org/10.1145/1141277.1141501
- [13] García Mejía Juan Fernando, Pérez Martínez José Arturo, Granda Gutiérrez Everardo Efren, Torres Reyes Carlos Eduardo, Martínez Garduño Yenit Micro algoritmo genético con codificación real para la optimización de un portafolio de inversión. Congr. Int. en Ing. Electrónica. Mem. ELECTRO, Vol. 39, pp. 253-258, Oct 2017, Chihuahua, Chih. México