



# **Universidad Autónoma del Estado de México**

Centro Universitario UAEM Valle de Chalco

DISEÑO DE UN ALGORITMO GENÉTICO PARA ESTABLECER  
PROFORMAS EN PROYECTOS DE INGENIERÍA CIVIL

## **T E S I S**

QUE PARA OBTENER EL GRADO DE

**MAESTRO EN CIENCIAS DE LA COMPUTACIÓN**

**P R E S E N T A**

Ing. José Manuel Pérez Esquivel

### **DIRECTORA**

Dra. Cristina Juárez Landín

CO-DIRECTOR

Dr. Marco Alberto Mendoza Pérez

TUTOR

Dr. Samuel Olmos Peña

**VALLE DE CHALCO SOLIDARIDAD, MÉXICO ENERO 2025**



**CUVCH**

## RESUMEN

La tesis se centra en el diseño de un algoritmo genético para la optimización en ingeniería civil, específicamente en la generación de proformas, documentos esenciales en la planificación y ejecución de proyectos de construcción, que permiten evaluar costos, tiempos y recursos necesarios para llevar a cabo una obra. A través de la generación eficiente de proformas, se busca facilitar la toma de decisiones y mejorar la gestión de proyectos en el ámbito de la ingeniería civil. El desarrollo del algoritmo incluye varias etapas clave, tales como la definición de la población inicial, la formulación de funciones de aptitud que evalúan la calidad de las soluciones generadas, y la implementación de operadores de selección, cruce y mutación para explorar el espacio de soluciones. Para evaluar el rendimiento del algoritmo, se utilizan gráficos de barras que permiten una visualización clara y comparativa de métricas relevantes como la desviación estándar, el cumplimiento de objetivos y la eficiencia de recursos, facilitando la identificación de tendencias y patrones en el rendimiento del algoritmo, lo cual es crucial para mejorar las futuras implementaciones. Los resultados obtenidos de la configuración óptima del algoritmo incluyen: un tamaño de población de 285, una probabilidad de cruce de 0.5901, una probabilidad de mutación de 0.0426, un número de generaciones de 516 y una probabilidad de elitismo de 0.0442 demostrando un rendimiento sobresaliente en las métricas evaluadas y proporcionan una base sólida para futuros análisis estadísticos y conclusiones del estudio.

## **ABSTRACT**

The thesis focuses on the design of a genetic algorithm for optimization in civil engineering, specifically in the generation of proformas, essential documents in the planning and execution of construction projects, which allow the evaluation of costs, times and resources necessary to carry out a work. Through the efficient generation of proformas, the aim is to facilitate decision making and improve project management in the field of civil engineering. The development of the algorithm includes several key stages, such as the definition of the initial population, the formulation of fitness functions that evaluate the quality of the generated solutions, and the implementation of selection, crossover and mutation operators to explore the solution space. To evaluate the performance of the algorithm, bar graphs are used that allow a clear and comparative visualization of relevant metrics such as standard deviation, compliance with objectives and resource efficiency, facilitating the identification of trends and patterns in the performance of the algorithm, which is crucial to improve future implementations. The results obtained from the optimal configuration of the algorithm include: a population size of 285, a crossover probability of 0.5901, a mutation probability of 0.0426, a generation number of 516 and an elitism probability of 0.0442 demonstrating outstanding performance on the evaluated metrics and providing a solid basis for future statistical analysis and conclusions of the study.

# ÍNDICE GENERAL

<b>Resumen</b>	<b>1</b>
<b>Abstract</b>	<b>2</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Antecedentes . . . . .	1
1.2. Planteamiento del Problema . . . . .	5
1.3. Objetivos . . . . .	13
1.3.1. General . . . . .	13
1.3.2. Específicos . . . . .	13
1.4. Hipótesis . . . . .	14
1.5. Justificación . . . . .	14
1.6. Delimitación del Problema . . . . .	15
1.7. Estructura del Trabajo . . . . .	16
<b>2. Marco Teórico y Estado del Arte</b>	<b>18</b>
2.1. Optimización Costo-Tiempo en Ingeniería Civil . . . . .	18
2.1.1. Técnicas de Optimización Costo-Tiempo . . . . .	19

2.2. Proformas en Proyectos de Ingeniería Civil . . . . .	23
2.3. Error Cuadrático Medio . . . . .	26
2.4. Conceptos Fundamentales de Algoritmos Genéticos . . . . .	27
2.4.1. Representación de individuos . . . . .	30
2.4.2. Inicialización de la población . . . . .	32
2.4.3. Evaluación de la población . . . . .	34
2.4.4. Selección de padres . . . . .	35
2.4.5. Recombinación o cruza . . . . .	36
2.4.6. Mutación . . . . .	38
2.4.7. Elitismo . . . . .	40
2.4.8. Iteración . . . . .	41
2.4.9. Gráfica de convergencia . . . . .	42
2.5. Estado del Arte . . . . .	44
<b>3. Metodología</b>	<b>48</b>
3.1. Diseño del Algoritmo Genético . . . . .	49
3.1.1. Representación de la Solución . . . . .	49
3.1.2. Función aptitud (fitness) . . . . .	50
3.1.3. Inicialización de la Población . . . . .	52
3.1.4. Operadores Genéticos . . . . .	53

3.2. Criterio de paro . . . . .	56
3.3. Algoritmo . . . . .	57
<b>4. Resultados Experimentales</b>	<b>59</b>
4.1. Recursos utilizados . . . . .	59
4.2. Representación del problema . . . . .	60
4.3. Ejecuciones de prueba . . . . .	61
4.4. Validación . . . . .	65
4.4.1. Configuración de la Simulación . . . . .	65
4.4.2. Rango de Valores para los Parámetros . . . . .	66
4.4.3. Grid Search con Validación Cruzada . . . . .	67
4.4.4. Selección de los Mejores Parámetros . . . . .	68
4.4.5. Visualización de Resultados . . . . .	69
<b>5. Conclusiones</b>	<b>73</b>
5.1. Conclusiones . . . . .	73
5.2. Trabajo Futuro . . . . .	74
5.3. Impacto y Aplicabilidad . . . . .	76
<b>Bibliografía</b>	<b>76</b>
<b>A. Apéndice</b>	<b>77</b>

## ÍNDICE DE TABLAS

4.1. Resultados de las ejecuciones del algoritmo genético. . . . .	63
4.2. Valores de cada parámetro para la validación cruzada . . . . .	67

## ÍNDICE DE FIGURAS

1.1. Ejemplo de una EDT de un proyecto de construcción de autopista (Elaboración propia) . . . . .	6
1.2. Ejemplo de una ECD de un proyecto de construcción de autopista (Elaboración propia) . . . . .	7
1.3. Ejemplo de un precio unitario (Elaboración propia) . . . . .	7
1.4. Análisis del costo directo de una tarea en el proyecto (Elaboración propia) . . . . .	8
1.5. Curva S en el inicio del proyecto, la línea en color cian representa el corte en el tiempo (Elaboración propia) . . . . .	9
1.6. Captura del avance físico. (Elaboración propia) . . . . .	10
1.7. Curva S con corte en febrero, en la gráfica de observan los presupuestos, original y trabajado, y el costo real de la obra (Elaboración propia) . . . . .	11
1.8. Desviación de los presupuestos con respecto al valor planeado (Elaboración propia) . . . . .	12
2.1. Actividad con remanente del 70% para terminarla en la que al establecer un reporte proforma se redistribuyen los recursos (Elaboración propia) . . . . .	25
2.2. Diagrama de flujo para un AG (Elaboración propia) . . . . .	30

2.3. Representación binaria del cromosoma de un individuo (Elaboración propia) . . . . .	32
2.4. Representación de individuos formando una población (Elaboración propia) . . . . .	33
2.5. Diagrama de flujo para un AG (Elaboración propia) . . . . .	35
2.6. Aplicando el cruce mediante la técnica de la cruce de punto (Elaboración propia) . . . . .	37
2.7. Mutación de dos alelos en uno de los hijos resultantes (Elaboración propia) . . . . .	40
2.8. Ejemplo de una gráfica de convergencia de una AG. (2020). Gráfica convergencia [Figura]. Ícaro Agostino. <a href="https://icaroagostino.github.io/post/sbo/">https://icaroagostino.github.io/post/sbo/</a> . . . . .	44
4.1. Gráfica de las mejores aptitudes en las ejecuciones (Elaboración propia) . . . . .	64
4.2. Gráficas con los resultados de la validación cruzada (Elaboración propia) . . . . .	70
4.3. Comparativa de campana de Gauss (Elaboración propia) . . . . .	72
4.4. Comparativa acumulado (Elaboración propia) . . . . .	72

# Capítulo 1

## INTRODUCCIÓN

### 1.1. Antecedentes

Existen varias organizaciones que se centran en promover las mejores prácticas proporcionando certificaciones y recursos además de servir como redes para profesionales de la gestión de proyectos en la industria de la construcción. Por mencionar algunas se encuentran; Construction Management Association of America (CMAA), es una asociación dedicada a la gestión de la construcción y la dirección de proyectos, ofrecen certificaciones como el Certified Construction Manager (CCM), brindan recursos y oportunidades de capacitación para profesionales de la gestión de proyectos en construcción ([www.cmaanet.org](http://www.cmaanet.org)); International Construction Project Management Association (ICPMA) una organización internacional que reúne a profesionales y académicos en el campo de la gestión de proyectos de construcción, organizan conferencias y colaboran en investigaciones relacionadas con la gestión de proyectos de construcción ([www.icpma.net](http://www.icpma.net)).

Project Management Institute (PMI) es una organización internacional sin fines de lucro que se dedica a la promoción y desarrollo de la gestión de proyectos, es ampliamente reconocido por su influencia en el campo de la dirección de proyectos así como su terminología y estándares descritos en su Project Management Body of Knowledge (PMBOK) ampliamente aceptado en la industria

para la gestión de proyectos ([www.Project Management Institute.org](http://www.Project Management Institute.org)).

La siguiente investigación se apoyará basándose en los lineamientos establecidos en el Project Management Institute, concretamente con lo decretado en el PMBOK versión 7 y en la Construction Extension to the PMBOK Guide, específicamente en el capítulo titulado: Project Cost Management. Aclarando todo esto por que para realizar la planeación, seguimiento y control de un proyecto de construcción en ingeniería civil la mayoría de los ingenieros y arquitectos deben de cumplir con una formación como Project Management Professional (PMP), esta certificación brinda las habilidades que se requieren así como la validación por el Project Management Institute.

Un proyecto es una iniciativa con un tiempo definido que tiene como objetivo crear algo nuevo y único, ya sea un producto, un servicio o un resultado determinados con un inicio y final claros. Los proyectos pueden ser independientes o formar parte de un programa o desarrollo más grande, contribuyendo a un objetivo general más amplio. (**pmbok21**). La gestión del proyecto engloba el proceso de iniciar, planificar, ejecutar, controlar y cerrar las actividades necesarias para llevar a cabo un proyecto con éxito. (**mattos2019metodos**), y en el entorno de la construcción es fundamental de manera que se pueda asegurar que los proyectos concluyan con éxito, más no se aplica tan solo de cumplir con las fechas pactadas, se trata, desde luego, de ejecutar la obra dentro del presupuesto cumpliendo con los estándares de calidad y seguridad. Con las exigencias de la industria moderna, el incremento de la competencia y la complejidad de los proyectos es imprescindible invertir en los procesos de gestión y control.

Una herramienta importante para el control y seguimiento de un proyecto se le da el nombre de proforma que se puede definir como un pronóstico de costos para un proyecto de construcción, ayudando a predecir los costos asociados con la construcción incluyendo una lista detallada de los costos asignados

al proyecto, como los materiales, mano de obra, equipo y los permisos. Las funciones principales de esta herramienta conllevan a revisar las cantidades con rendimientos en los insumos en la que se lleva a cabo la asignación de recursos a las actividades de trabajo del cronograma establecido para el proyecto realizando su distribución en el tiempo, establecido por el comienzo y fin de la actividad, a esta acción se le conoce como análisis de costos.

Cuando se tiene la aprobación por parte del cliente de un presupuesto de obra, después de varios eventos a esta como licitaciones, evaluaciones de ofertas, diseños de planos y estimaciones de costos preliminares, por mencionar algunos, se hace el establecimiento de líneas base, que se podrían ver como proformas para dar el seguimiento y control de la obra, al comienzo del proyecto se establece la Línea Base Original (LBo), un presupuesto que es inalterable en el tiempo de ejecución del proyecto y tiene la función de ser la guía en la ejecución, en el instante en que se establece LBo se realiza un “copia” del mismo análisis para poder establecer la Línea Base Actual (LBa), este análisis es el que se podrá modificar a lo largo de la ejecución del proyecto.

El establecimiento del proforma no es un evento único en el proyecto cuando se establece un proforma se nombra como Línea Base Proforma (PF), se elabora de manera continua tomando en cuenta que los cambios que se van presentando a lo largo de la ejecución, el objetivo es mantener o modificar la planeación de tal forma que se mantenga cercana a la LBo, en la ejecución del proyecto este método se realiza cada trimestre, aunque es posible en cualquier momento del año si hay un suceso extraordinario que cause una afectación a la obra, teniendo 4 proformas en el año denominados 00 + 12, 03 + 09, 06 + 06 y 09 + 03, los primeros dos dígitos se refieren a los meses con el valor del costo real del proyecto y los últimos dos a los meses restantes en la planeación.

En un proyecto de ingeniería civil los objetivos pueden ser múltiples y variados,

y a menudo están relacionados con diferentes aspectos del proyecto, como el alcance, el tiempo, el costo, la calidad y otros factores por lo que la gestión de proyectos implica equilibrar y optimizar estos diferentes objetivos para lograr un resultado exitoso, es por esta razón que la gestión del proyecto es un proceso multiobjetivo y los problemas que surgen durante la ejecución del proyecto son usualmente considerados de dos tipo; determinísticos y estocásticos, por lo que los Algoritmos Genéticos (AGs) son ideales para la optimización de estos problemas.

Los AGs son una subcategoría del campo de la informática nombrada Computo Evolutivo (CE), de aquí que esta técnica de optimización y búsqueda está inspirada en la evolución biológica, a este propósito, es simular el proceso de selección natural y evolución para encontrar soluciones óptimas o aproximadas a un problema, es necesario recalcar que los AGs son solo algunos de los enfoques más conocidos del CE. Esta herramienta puede ser de gran ayuda para la gestión de proyectos ya que la complejidad de un proyecto de ingeniería civil en donde se abordan problemas de optimización y toma de decisiones complejas tienden a desequilibrar la planeación de la obra quizá, entonces los AGs proporcionarían datos al personal de control de proyectos y así pueden tomar decisiones informadas, lo que conduce a una gestión más efectiva.

En esta investigación se diseña un algoritmo genético que realice una minimización en la curva referente al seguimiento en el proyecto con relación en el pronóstico creado para cada una de las actividades en un proyecto de ingeniería civil, con el objetivo de ofrecer una solución al ajuste en desviaciones causadas por los atrasos o adelantos de la obra lo cual ayudara al personal de control de proyectos a establecer proformas y, con ello, a reducir el tiempo en la toma de decisiones, sin mencionar que auxiliará al personal que este comenzando en la gestión de proyectos.

## 1.2. Planteamiento del Problema

Cuando se decide realizar una obra de construcción se llevan a cabo varios concursos para contratar a la empresa o empresas que van a ejecutar el proyecto, con la aprobación del proyecto por parte del cliente se inicia la gestión con una planificación y análisis de costo con la cual servirá de guía a lo largo de la ejecución de la obra civil, para llevarla a cabo, el personal de control de proyectos reúne la información necesaria de documentos generados en el concurso.

El cronograma de ejecución de la obra civil funciona como base para el control del proyecto, permite planificar y programar las actividades de la obra en secuencia temporal con tareas que se realizan en orden correcto y que no haya superposiciones ni retrasos innecesarios. La estructura de trabajo (EDT) es una herramienta que permite organizar y dividir el trabajo total de un proyecto en tareas más pequeñas y manejables, se caracteriza por su estructura sistematizada, donde cada nivel inferior representa un mayor nivel de detalle del trabajo que debe realizarse. De esta manera, la EDT permite visualizar claramente el alcance total del proyecto y los diferentes componentes que lo conforman (**pmbok21**). En la figura 1.1 se presenta un ejemplo de una EDT de un proyecto de ingeniería civil.



Figura 1.1: Ejemplo de una EDT de un proyecto de construcción de autopista (Elaboración propia)

En ciertas ocasiones se desarrolla una estructura de costos (ECD) como una herramienta de mapeo entre el proyecto EDT y el código de cuentas designado por la organización para ayudar a informar los costos, (**Project Management Institute 2016 project**), esta estructura, también conocida como cuenta de costos, se relaciona con cada una de las actividades para poder dar un seguimiento del costo real por actividad, básicamente son las facturas que se pagan por los recursos utilizados en la actividad. En la figura 1.2 es un ejemplo de una ECD.

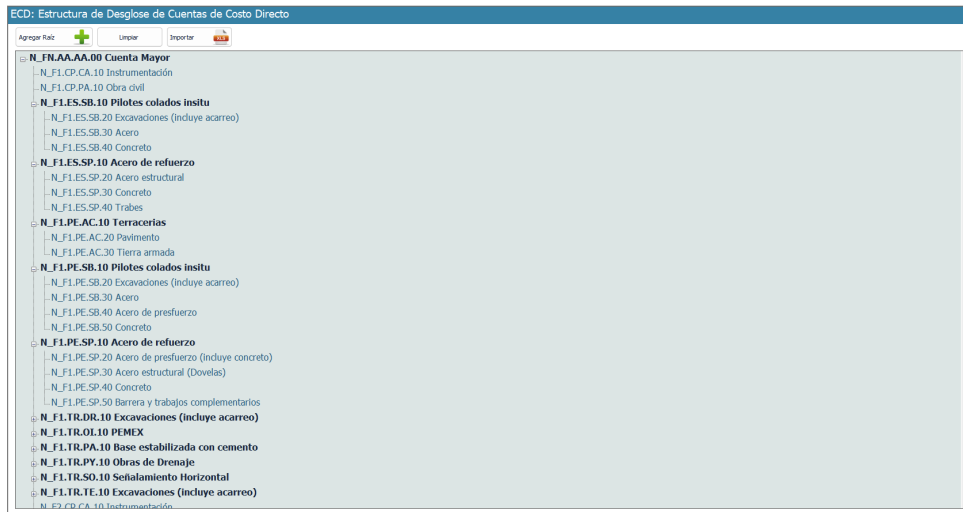


Figura 1.2: Ejemplo de una ECD de un proyecto de construcción de autopista (Elaboración propia)

Los precios unitarios (PU) son valores que representan el costo o precio de una unidad específica de un artículo, producto o servicio, se utilizan comúnmente en la estimación de costos, la presupuestación y en contabilidad de proyectos en la construcción, obras públicas, fabricación y otros campos donde es necesario calcular los costos de elementos individuales. En la figura 1.3 se muestra el ejemplo de un precio unitario.

PU100 - Fabricación y habilitado de placas de acero estructural A-36 en apoyos de neopreno como lo establece el proyecto, p.u.o.t.

Filtrar  
 Clave:  Descripción:  TR:

Insertar Insumo  
 TI:  Insumo:

Cifras en Peso mexicano (MXN)

RECURSO	DESCRIPCION	CANTIDAD	PRECIO	IMPORTE
%MO	H. MENOR (3% DE M.O.)	15,39907273	\$ 0,03	\$ 0,46
Maq18	Camioneta estacas 3Ton. Chev	0,00181818	\$ 189,27	\$ 0,34
Mat1	ACETILENO	0,04681818	\$ 126,00	\$ 5,90
Mat101	OXIGENO (CARGA)	0,05618182	\$ 27,28	\$ 1,53
Mat66	EQUIPO DE OXICORTE	0,00090909	\$ 3.600,00	\$ 3,27
MN002	AYTE. DE SOLDADOR	0,00909091	\$ 582,74	\$ 5,30
MN012	CHOFER DE CAMIONETA	0,00181818	\$ 699,64	\$ 1,27
MN028	OFICIAL SOLDADOR	0,00909091	\$ 971,23	\$ 8,83
MN091		0,00181818	\$ 0	\$ 0
PLACA-1	PLACA DE ACERO ESTRUCTURAL	1,03000000	\$ 16,00	\$ 16,48

Registros: 10 Importe: \$ 43,39

Figura 1.3: Ejemplo de un precio unitario (Elaboración propia)

Al obtener los datos de la EDT, ECD y PU, información descrita anteriormente, ya es posible comenzar con la gestión del proyecto, en específico con la sección de establecimiento del proforma, el siguiente paso es realizar el análisis de costo directo en el cual se van asignando los recursos que provienen de los PU en las actividades de nivel del trabajo, es decir, que se encuentran en el último nivel del EDT también conocidas como tareas, posteriormente se les pondera una cantidad lo cual genera un importe, con esto la actividad o tarea tendrá un presupuesto para su ejecución. En la 1.4 se muestra el análisis de costo directo con una actividad de trabajo.

PU100 - Fabricación y habilitado de placas de acero estructural A-36 en apoyos de neopreno como lo establece el proyecto, p.u.o.t.

Filtrar  
 Clave:  Descripción:  TR: AM | Amortizaciones

Insertar Insumo  
 TI: AM | Amortizaciones Insumo:

Cifras en Peso mexicano (MXN)

RECURSO	DESCRIPCION	CANTIDAD	PRECIO	IMPORTE
%MO	H. MENOR. (3% DE M.O.)	15,39907273	\$ 0,03	\$ 0,46
Maq18	Camioneta estacas 3Ton. Chev	0,00181818	\$ 189,27	\$ 0,34
Mat1	ACETILENO	0,04681818	\$ 126,00	\$ 5,90
Mat101	OXIGENO (CARGA)	0,05618182	\$ 27,28	\$ 1,53
Mat66	EQUIPO DE OXICORTE	0,00090909	\$ 3.600,00	\$ 3,27
MN002	AYTE. DE SOLDADOR	0,00909091	\$ 582,74	\$ 5,30
MN012	CHOFER DE CAMIONETA	0,00181818	\$ 699,64	\$ 1,27
MN028	OFICIAL SOLDADOR	0,00909091	\$ 971,23	\$ 8,83
MN091		0,00181818	\$ 0	\$ 0
PLACA-1	PLACA DE ACERO ESTRUCTURAL	1,03000000	\$ 16,00	\$ 16,48

Registros: 10 Importe: \$ 43,39

Figura 1.4: Análisis del costo directo de una tarea en el proyecto (Elaboración propia)

Finalizando la asignación de recursos en todas las actividades de trabajo de la EDT es como se culmina el análisis de costo directo, en este momento se procede a establecer la LBo, para esto se calcula una distribución de los presupuestos en el tiempo, en días laborables, que se obtiene de la duración de la tarea. Actualmente se contemplan dos distribuciones principales de recursos, la primera es de forma lineal ( $f(x) = \frac{1}{x}$ ) en la que la cantidad del recurso es dividida entre los días (x) de su ejecución y la segunda con una distribución normal a través de

la función de densidad  $(f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2})$ , al tener la distribución de los recursos ya se establece la LBo.

Para comenzar con la ejecución del proyecto se hace una “copia” de la LBo, con el objetivo de establecer la LBa, que en este instante del tiempo es equivalente a la LBo pero tendrá que ir cambiando acorde a la ejecución del proyecto, estos cambios se deben de realizar mediante proformas que se van estableciendo durante la ejecución de la obra civil, siendo que se establecen proformas cada trimestre, el personal de control de proyecto puede visualizar el avance del proyecto. En la figura 1.5 se muestra una gráfica en la cual se observan la LBo y LBa en una forma semejante.

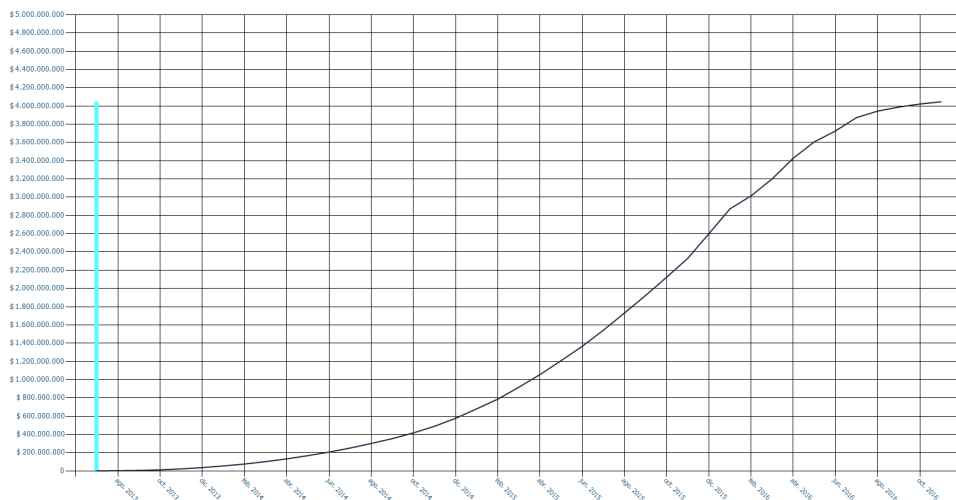


Figura 1.5: Curva S en el inicio del proyecto, la línea en color cian representa el corte en el tiempo (Elaboración propia)

Ahora que ya se tiene la LBo y la LBa comienza el seguimiento a la obra civil mediante la captura del avance físico por cada tarea que se está ejecutando, este avance físico se va almacenando de dos formas, en porcentaje o en volumen, es prudente advertir que a las tareas se les asigna una unidad de medida para poder realizar el registro de información, para ilustrar mejor; para remover  $3,000m^3$

de sustrato o tierra se requieren recursos con diversas unidades de medida (maquinaria (horas), mano de obra (jornada), materiales (kilogramos, metros y piezas)) de tal forma que la tarea queda como nombrada como EXCAVACIÓN a la cual se le da una asignación de unidad de medida de  $m^3$ , así se puede decir que se ha tenido un avance físico del 10% o  $300m^3$  de volumen. En la figura 1.6 se muestra la captura del avance físico.

134 - Concreto hidráulico de $f'c= 200 \text{ kg/cm}^2$ , p.u.o.t.	
Cantidad:	987,90 $m^3$
Avance:	50,7 %
Avance (Volumen)	500,87 $m^3$

Unidad: 50,7 %

Porcentaje  
 Volumen

Avance = 500,87  $m^3$

Inicio Real: 20/01/2014  
Fin Real: 03/10/2023

Guardar

Figura 1.6: Captura del avance físico. (Elaboración propia)

La captura del avance físico se va realizando de acuerdo con las necesidades del proyecto, esto quiere decir que se tendrán registros diarios, semanales o mensuales, dependiendo de la administración y naturaleza de la tarea. Con este avance es posible obtener el valor ganado que se abordará con más detalle en el subcapítulo 2.1.1 Técnicas de Optimización Costo-Tiempo, el cual se calcula como el producto del avance físico y el presupuesto aprobado de la tarea, así ya se pueden realizar las comparaciones del presupuesto original con el presupuesto trabajado. En la figura 1.7 se observa cómo va avanzando el proyecto en términos de presupuestos.

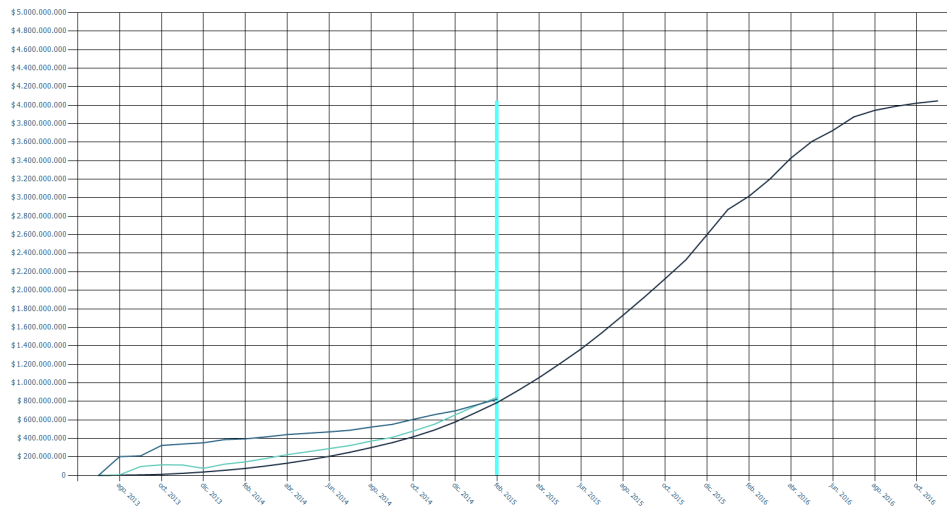


Figura 1.7: Curva S con corte en febrero, en la gráfica de observan los presupuestos, original y trabajado, y el costo real de la obra (Elaboración propia)

El objetivo de establecer proformas es para poder modificar la LBo del proyecto de tal manera que se ajuste a la LBo, durante la ejecución del proyecto donde van surgiendo inconvenientes que van afectando el costo y/o tiempo del proyecto generando una desviación en la curva de la LBo, la planeación original, el personal de control de proyectos se ven obligados, en ocasiones, a realizar malas prácticas referente a la administración del proyecto con el objetivo de generar los reportes “adecuados”, por ejemplo, el rehacer la planeación original para mostrar “resultados aceptables”. En la figura 1.8 se ve el momento en el tiempo donde se produce una desviación en los presupuestos propiciando la acción de establecer un pronóstico proforma.

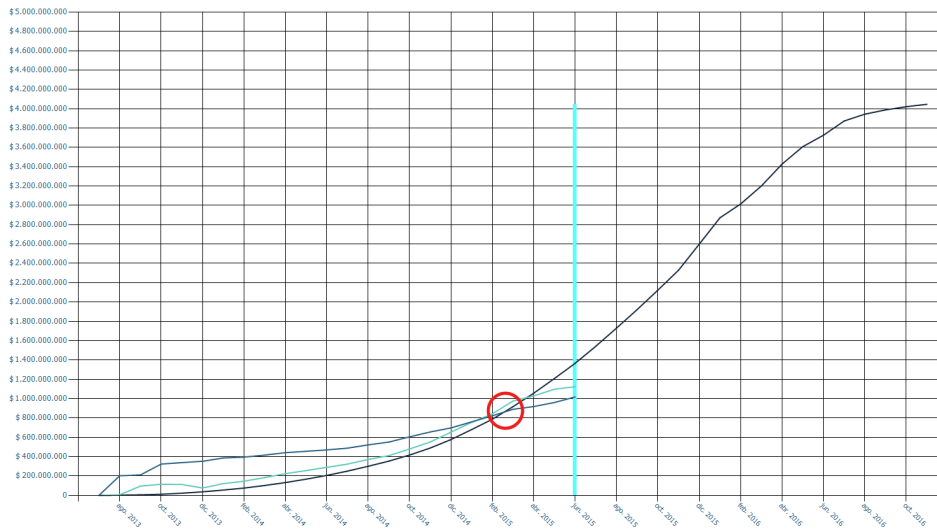


Figura 1.8: Desviación de los presupuestos con respecto al valor planeado (Elaboración propia)

Dada la evaluación del rendimiento del proyecto hasta ahora, se requiere que el gerente de proyectos determine la cantidad restante de presupuesto hasta la conclusión del mismo, esta estimación o proyección en el valor ganado se le conoce como estimación del costo hasta la conclusión, para esto existen varios enfoques lógicos, visiones que se retomará en el capítulo 2.1.1 Técnicas de Optimización Costo-Tiempo, para determinar el ETC, este proceso resulta ser en ocasiones muy extenso debido a que se debe de analizar cada una de las tareas, teniendo en ocasiones como consecuencia la omisión en el análisis de algunas estas, surgiendo como problema en el futuro y el dilema más común en la gestión de proyectos: Quieres avanzar una tarea de una duración planeada en semanas en días. El cómo establecer proformas se retomará en el capítulo 2.2 Proformas en Proyectos de Ingeniería Civil. Considerando la información presentada en este apartado, se plantean las siguientes preguntas de investigación:

¿Cómo puede un algoritmo genético distribuir de manera óptima el presupuesto asignado para ajustarse al valor planeado?

¿Cómo puede el algoritmo genético ajustar dinámicamente la distribución del presupuesto a medida que se obtienen datos en tiempo real sobre la ejecución del proyecto?

¿Cómo puede el algoritmo genético ser diseñado para minimizar el tiempo necesario para generar proformas, manteniendo al mismo tiempo la precisión en la distribución del presupuesto?

### **1.3. Objetivos**

#### **1.3.1. General**

Diseñar un algoritmo genético que distribuya el presupuesto de las tareas para de esta manera favorecer en el establecimiento de proformas en proyectos de obra civil para mejorar la eficiencia y eficacia en la gestión de proyectos brindando un apoyo en la predicción y toma de decisiones.

#### **1.3.2. Específicos**

1. Describir la representación más natural del problema por medio de una cadena de variables o genes que representen los parámetros de la función a optimizar.
2. Definir la función de aptitud (fitness), tomando en cuenta las variables identificadas para determinar la calidad de las soluciones generadas por el algoritmo genético, asignando ponderaciones de acuerdo con las necesidades del proyecto de obra civil, considerando restricciones y objetivos específicos.
3. Evaluar la función de aptitud, obteniendo y preparando un conjunto de

datos para realizar las pruebas, incluyendo proformas reales o simulados utilizando la función de aptitud para calcular los valores de calidad del conjunto de datos de prueba, identificando desviaciones y tendencias se podrá ajustar la función de aptitud.

4. de los AGs es la parte fundamental del desarrollo de un algoritmo genético, los operadores que se incluirán son **selección, cruce (recombinación), mutación y elitismo:**

#### **1.4. Hipótesis**

La implementación de un algoritmo genético que distribuya el presupuesto de las tareas puede tener un impacto significativo en la precisión y la eficiencia en la elaboración de proformas en los proyectos de obra civil, al realizar la optimización de variables clave, esto conduce a generar proformas de mayor precisión con la identificación de soluciones óptimas y la mejora general de la toma de decisiones en la gestión de proyectos de construcción.

#### **1.5. Justificación**

Para la mayoría de las empresas constructoras el tener la gestión de proyectos implica tener un personal dedicado a cada una de las actividades generando principalmente un retraso en los tiempos de entrega en los proyectos de construcción, la implementación de AGs puede mejorar significativamente la precisión y eficiencia de la elaboración de proformas en la ejecución de proyectos de construcción, lo que puede reducir los costos mejorando la rentabilidad de los proyectos. La gestión de proyectos de construcción puede permitir la integración de tecnologías de vanguardia, como la inteligencia artificial, lo que puede ayudar

a el establecimiento de proformas.

Lo que la investigación plantea es ayudar a reducir los tiempos en el establecimiento de proformas, así como evaluar el algoritmo genético en los presupuestos ajustados observando que tan cercanos son del valor planeado, permitiendo la predicción de tarea, administración de costos y tiempo con todo esto se pudiera creer que se mejora la toma de decisiones en la planificación dando una mejor calidad de servicio.

La investigación es viable ya que se cuenta con un sistema de optimización con programación tradicional como antecedente el cual ayuda en la curva de aprendizaje, además de tener los recursos que se obtendrán de Grupo Constructor REGMIX S.A de C.V.

## **1.6. Delimitación del Problema**

El algoritmo genético se diseña para establecer proformas considerando únicamente el presupuesto que se tienen establecido del proyecto y el tiempo de las actividades de trabajo o tareas, con la intención de ayudar al personal de control de proyectos para ajustar la LBa lo más cercano a la LBo. En cuanto a los costos se tomará en cuenta exclusivamente en el Costo Directo de la obra, sin entrar en consideraciones sobre Costos Indirectos, por consiguiente, se trata de una investigación de tipo cuantitativa dejando aparte, por un momento, al menos el enfoque cualitativo (el juicio de expertos) para futuras investigaciones convirtiendo la presente en una investigación mixta.

Para la obtención de estos datos se establece el contacto con el Grupo Constructor REGMIX S.A de C.V ([www.gcregmix.com](http://www.gcregmix.com)), estableciendo una colaboración con el Ingeniero Daniel Meza Urquiza. El ingeniero Meza, formalizando un acuerdo de confidencialidad por ambas partes, proporcionara

información de las obras que se hallasen en ejecución al momento, así como de proyectos concluidos, además desempeñará un papel crucial como experto en la gestión de proyectos de ingeniería civil, aportando sus conocimientos en la definición de ponderaciones y pesos en la función de aptitud (fitness).

Este trabajo de investigación se centra en el diseño de un algoritmo genético destinado a resolver problemas de optimización costo-tiempo en la gestión de proyectos de construcción, especialmente en la fase de ejecución y seguimiento. La implementación práctica del algoritmo se quedará pendiente para futuras investigaciones, por lo que las pruebas se enfocarán en evaluar la función de aptitud, observando el comportamiento del algoritmo y realizando los ajustes necesarios.

## **1.7. Estructura del Trabajo**

El presente documento se estructura en cuatro capítulos, siendo que cada uno de ellos está dedicado a la información específica de acuerdo con la siguiente estructura:

1. Marco teórico y estado del arte, se abordarán conceptos claves como: algoritmos genéticos, proforma, gestión de proyectos en obras de construcción además de una revisión del estado actual de la gestión de proyectos de construcción y aplicaciones previas de AGs en la construcción.
2. Metodología se describirá el diseño del algoritmo genético para la generación de proformas, recopilando datos de prueba y definiendo una función aptitud realizando su validación y evaluación para integrarla a los operadores del algoritmo genético.
3. Resultados experimentales, se presentarán los conjuntos de datos

utilizados mostrando la evaluación del algoritmo genético en términos de eficiencia y realizando una comparación de resultados con enfoques tradicionales en la elaboración de proformas. Con el análisis de los resultados obtenidos se hará una interpretación del impacto de los AGs en la elaboración de proformas.

4. Conclusiones, se realizará un resumen de los hallazgos dando respuesta a las preguntas de investigación, así como las implicaciones de los resultados obtenidos. Con esto se podrán detallar los aportes y contribuciones de esta investigación.

## Capítulo 2

### MARCO TEÓRICO Y ESTADO DEL ARTE

#### 2.1. Optimización Costo-Tiempo en Ingeniería Civil

La gestión de proyectos es una disciplina fundamental en el mundo empresarial, así como en una amplia variedad de industrias como la construcción con ella se obtiene la capacidad de planificar, ejecutar y controlar proyectos de manera eficiente es crucial para el éxito de las organizaciones a esto se adjuntan técnicas de optimización costo-tiempo se ha convertido en un enfoque esencial en la ingeniería civil debido a la creciente demanda de proyectos de construcción que requieren entregas más rápidas con costos más bajos estableciendo e implementando estrategias que buscan encontrar el equilibrio necesario para completar los proyectos.

Una gestión eficaz implica una comprensión sólida de las restricciones de tiempo y presupuesto en donde a menudo se enfrentan a presiones para ser entregados antes de lo previsto o con un presupuesto limitado, este es un desafío para los ingenieros en la tomar decisiones que deben de ser informadas sobre cómo maximizar la eficiencia y minimizar los riesgos, es interesante examinar el problema también desde los enfoques tradicionales de gestión de proyectos que se centran en la reducción de costos a menudo a sin considerar el tiempo o al contrario, en el cumplimiento de plazos sin tener en cuenta el costo es por eso que una optimización costo-tiempo busca una estrategia más equilibrada al

considerar ambos aspectos.

### **2.1.1. Técnicas de Optimización Costo-Tiempo**

Uno de los enfoques más efectivos en la gestión de proyectos es el método de la ruta crítica el cual fue desarrollado en 1957 en los Estados Unidos de América, una metodología que ha demostrado su valor no solo en proyectos de construcción sino también en desarrollo de software, investigación y muchas otras áreas, el método de la ruta crítica identifica las actividades críticas que deben completarse en el menor tiempo posible para evitar retrasos en el proyecto. Esto permite una gestión más efectiva del tiempo.

Una de las ventajas más significativas del método de la ruta crítica es su capacidad para identificar las tareas críticas y destacar las que pueden retrasarse sin afectar el plazo de finalización del proyecto como consecuencia proporciona información necesaria para tomar decisiones estratégicas sobre la asignación de recursos, la priorización de tareas y la gestión de riesgos de esta manera se puede decir que la método de la ruta crítica es la secuencia de actividades con la mayor duración, y cualquier retraso en cualquiera de estas actividades retrasará la finalización del proyecto en su conjunto.

Para poder hacer una implementación de la método de la ruta crítica se tiene que Identificar las actividades del proyecto que deben completarse para concluir el proyecto normalmente mediante la descomposición del proyecto en tareas más pequeñas y manejables posteriormente es necesario estimar la duración esperada basándose en el conocimiento y la experiencia del equipo del proyecto tomando en cuenta estas duraciones se establecen las relaciones de las actividades del proyecto, es decir, una actividad no puede comenzar hasta que otra haya terminado con esto se determinar la secuencia de las actividades. Una

vez que se han identificado las actividades, la duración de las actividades y las relaciones de precedencia, se puede calcular la ruta crítica mediante un software de gestión de proyectos o manualmente hay que tomar en cuenta que puede ser complejo y laborioso de implementar.

El método del valor ganado es una metodología de gestión de proyectos que se basa en el concepto de "trabajo en progreso" para predecir el desempeño futuro del proyecto en términos de costos y cronograma, utilizando el costo como unidad de medida común para evaluar el avance del proyecto y detectar desviaciones potenciales (**bhosekar2012cost**) y se utiliza para proporcionar una imagen precisa del estado del proyecto en un momento dado comparando el trabajo real realizado con el trabajo planificado y el costo real con el costo presupuestado para determinar si un proyecto está siguiendo su curso previsto.

En el método del valor ganado se manejan varios componentes, todos expresados en términos de valor monetario ya que mirándolo así es posible tener una visión clara de la ejecución del proyecto, de los cuales se expresa el valor planeado (VP), la cantidad de trabajo que se esperaba que estuviera completada en un punto específico del proyecto, según el plan original, es decir el presupuesto original que es aprobado, por otro lado, el valor ganado (VG) es la cantidad de trabajo que realmente se ha completado en el proyecto hasta la fecha por consiguiente se calcula a partir del progreso real del proyecto mediante un avance físico de ejecución de la obra y para finalizar el costo real (CR) es el costo real en el que se ha incurrido para realizar el trabajo hasta la fecha, se deduce que son las facturas registradas de los gastos realizados en el proyecto.

Con estos datos, se pueden calcular varias métricas como el índice de rendimiento de costos (IRC) el cual indica si el proyecto está dentro o por debajo del presupuesto. Se calcula como  $IRC = VG / CR$ , si el IRC es mayor que 1 ( $IRC > 1$ ), indica que el proyecto está utilizando los recursos de manera más eficiente de

lo planeado originalmente, en otras palabras, está obteniendo más valor por cada peso gastado pero si el IRC es igual a 1 ( $IRC = 1$ ), indica que el proyecto está utilizando los recursos exactamente según lo planeado por otro lado si el IRC es menor que 1 ( $IRC < 1$ ), significa que el proyecto está gastando más de lo que se había planeado originalmente en relación con el trabajo completado indicado una ineficiencia en el uso de los recursos financieros.

Otra de las métricas es el índice de rendimiento planeado (IRP) este índice muestra si el proyecto está avanzando según lo planeado, se calcula como  $IRP = VG / VP$ , si el IRP es mayor que 1 ( $IRP > 1$ ), indica que el proyecto está adelantado en cuanto a la planeación, es decir, el proyecto ha ganado más valor (completado más trabajo) de lo originalmente planificado en ese momento, en cierta situación si el IRP es igual a 1 ( $IRP = 1$ ), significa que el proyecto está progresando exactamente según la planeación original, sin embargo, si el IRP es menor que 1 ( $IRP < 1$ ), sugiere que el proyecto está retrasado en relación con la planeación ganado menos valor (completado menos trabajo) de lo planeado en ese punto.

La variación de costo (VCo) evalúa la diferencia entre el CR y el VP, se obtiene como  $VCo = VG - CR$ , si el VCo es positivo ( $VCo > 0$ ), significa que el proyecto está por debajo del presupuesto planificado en ese punto del tiempo, generalmente se considera un buen resultado, Si el VCo es negativo ( $VCo < 0$ ), significa que el proyecto está por encima del presupuesto planificado en ese punto del tiempo, indicado un problema de sobre costo en el proyecto y si el VCo es igual a cero ( $VCo = 0$ ), indica que el proyecto está gastando exactamente según lo planeado en ese punto del tiempo.

La variación del cronograma (VCr) mide la diferencia entre el VG y el VP, se calcula como  $VCr = VG - VP$ , si el VCr es positivo ( $VCr > 0$ ), significa que el proyecto está adelantado en términos de programación en ese punto del tiempo

indicando que el trabajo completado es más de lo que se había planeado, si el VCr es negativo ( $VCr < 0$ ) significa que el proyecto está retrasado en términos de programación en ese punto del tiempo revelando que el trabajo completado es menos de lo que se había planeado por otro lado si el VCr es igual a cero ( $VCr = 0$ ), indica que el proyecto está en línea con la programación planeada en ese punto del tiempo.

Existen cuatro enfoques lógicos que proporcionan diferentes perspectivas sobre cómo estimar el costo final del proyecto, en primer lugar, un enfoque basado en el presupuesto original (optimista), se admite que el trabajo futuro se realizará con el costo presupuestado inicialmente, es decir, con el valor planeado, otro enfoque está basado en el desempeño del costo (realista), lo que quiere decir que se ajusta la proyección considerando el Índice de Desempeño del Costo (IRC) hasta la fecha, por otro lado, el enfoque basado en el desempeño del costo y el cronograma (pesimista) incorpora tanto el IRC como el Índice de Desempeño del Cronograma (IRP) hasta la fecha para ajustar los costos futuros y las estimaciones de tiempo, y por último, realizar una nueva estimación donde se presupuesta nuevamente el trabajo restante sin mantener ninguna vinculación con estimaciones anteriores (**mattos2019metodos**).

El método del valor ganado es una herramienta valiosa para la gestión de proyectos, ya que proporciona una visión integral del desempeño del proyecto y permite a los gerentes y partes interesadas tomar decisiones informadas para corregir desviaciones, realizar ajustes en el plan o asignar recursos adicionales cuando sea necesario para garantizar el éxito del proyecto.

Una técnica que no ha sido utilizada muy frecuente en ingeniería civil es la simulación de Monte Carlo es una técnica numérica utilizada para resolver problemas matemáticos o de toma de decisiones mediante la generación de números aleatorios, se basa en la idea de que es posible estimar resultados

probabilísticos o cuantitativos utilizando muestras aleatorias y estadísticas, se llama así en honor al famoso casino de Monte Carlo en Mónaco.

Algunos de los pasos típicos en una simulación de Monte Carlo incluyen, por supuesto, identificar el problema que se quiere resolver o el sistema que se desea modelar definiendo las variables de entrada, las restricciones y los resultados deseados, posteriormente se generan números aleatorios siguiendo una distribución de probabilidad adecuada para representar incertidumbre en las variables de entrada, con estos datos se ejecutan numerosas simulaciones utilizando los valores de entrada representando una instancia del problema por simulación obteniendo sus correspondiente resultados que se analizan estadísticamente incluyendo la estimación de promedios, desviaciones estándar y percentiles, con estos análisis se utilizan para tomar decisiones informadas o para obtener una comprensión de la probabilidad de diversos resultados en el problema.

La simulación de Monte Carlo es especialmente útil cuando se trata de problemas complejos o no lineales, donde no es posible obtener una solución analítica ya que, al simular múltiples escenarios, se puede evaluar la sensibilidad del problema a diferentes condiciones iniciales y parámetros, lo que ayuda en la toma de decisiones o en la comprensión de sistemas complejos. Se incluye esta técnica ya que tiene cierta similitud a los AGs.

## **2.2. Proformas en Proyectos de Ingeniería Civil**

Un proforma es una herramienta importante para el desarrollo de un proyecto y ayuda a predecir los costos y gastos asociados con la construcción incluye una lista detallada de los costos asociados con el proyecto, como los materiales, la mano de obra, el equipo y los permisos además de incluir los

plazos para la finalización del proyecto, el calendario de pagos y los términos de financiamiento todo esto para asegurarse de que el proyecto se mantenga dentro del presupuesto es utilizado para obtener financiamiento o para establecer un plan de negocio para el proyecto con el objetivo de mantener o modificar la planeación de tal forma que se cumpla la planeación original al inicio del proyecto.

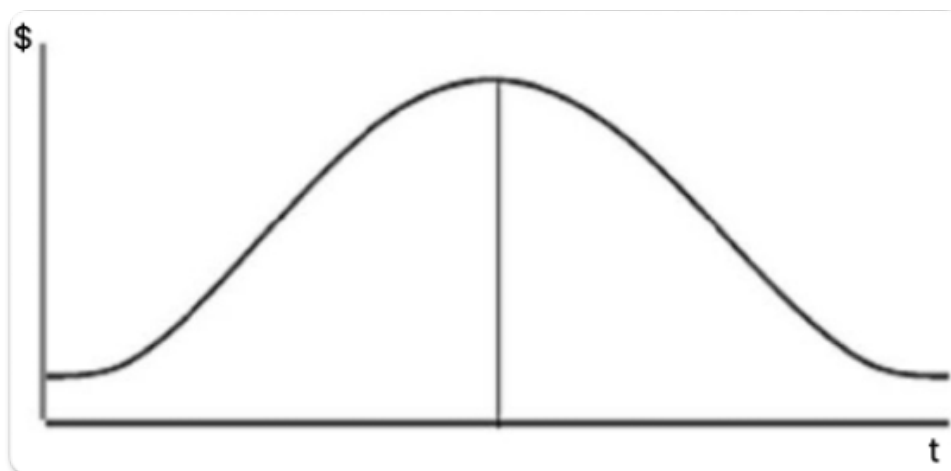
Para establecer proformas se puede conseguir mediante una serie de pasos, este método se ejecuta como mínimo cada trimestre teniendo cuatro proformas en el año denominados 00 + 12, 03 + 06, 06 + 06 y 09 + 03, los primeros dos dígitos se refieren a los meses con el valor del costo real del proyecto y los últimos dos a los meses de planeación. La elaboración de proformas es una parte importante de la gestión de proyectos, a continuación, se describen los pasos para establecer proformas. Antes de realizar cualquier modificación en el cronograma del proyecto, es una práctica recomendada hacer una copia de seguridad del cronograma original permitiendo comparar los cambios realizados y mantener un registro histórico del plan original.

El primer paso es revisar el estado de todas las actividades del proyecto y marcar como completas las que ya se hayan finalizado para proceder a eliminar estas actividades del plan original, ya que no es necesario incluirlas en la proforma para posteriormente calcular la ruta crítica con las actividades que determinan la duración total del proyecto ya que estas actividades no pueden retrasarse sin afectar la fecha de finalización del proyecto, si surgen nuevas tareas o cambios en el proyecto que no estaban contemplados en la planificación inicial se agregan al plan para incluirlas en la proforma con las actividades por completar se hace una revisión del progreso y se ajusta las cantidades de trabajo restante en función de los avances físicos actualizando las estimaciones de tiempo y recursos necesarios para completar cada tarea y se calcula los costos directos restantes para cada actividad, teniendo en cuenta el avance obtenido y las cantidades ajustadas dando una idea más precisa de los costos que quedan por incurrir en

el proyecto por último se redistribuyen los recursos en el tiempo, la redistribución de recursos en el contexto de la gestión de proyectos se refiere a ajustar la asignación de recursos a lo largo del tiempo de manera más uniforme o siguiendo una distribución de recursos que puede ser lineal o en forma de campana de Gauss (distribución normal) por lo que se obtiene una gestión más eficiente de los recursos a lo largo del proyecto y evitar sobrecargas o subutilización de los mismos en momentos específicos. En la figura 2.1 se muestra una actividad que no ha sido completada y es considerada que este en un reporte proforma.

### Identificación de la actividad en tiempo y costo

Actividad	Tiempo	Avance 30 %	Remanente 70 %
Tiempo	3 Meses	0.9 Meses	2.1 Meses
Costo	\$ 1 000 000.00	\$ 300 000.00	\$ 700 000.00



Redistribuir remanente de \$ 700, 000.00 en 2.1 meses

Figura 2.1: Actividad con remanente del 70% para terminarla en la que al establecer un reporte proforma se redistribuyen los recursos (Elaboración propia)

Una vez completado estos pasos se tendrá un reporte proforma actualizado que refleja la situación real del proyecto y ayudará a tomar decisiones informadas para su gestión continua, es prudente advertir que se debe seguir un monitoreo

constante del proyecto que avanza para garantizar que se cumplan los objetivos de tiempo y costo.

### 2.3. Error Cuadrático Medio

El Error Cuadrático Medio (ECM) es una medida comúnmente utilizada para evaluar la precisión de un modelo de regresión o de predicción. Es una medida de cuánto difieren los valores predichos por el modelo de los valores reales. Se calcula como la media de los cuadrados de las diferencias entre los valores predichos y los valores reales. La fórmula general para calcular el ECM es:

$$ECM = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Donde:

1.  $n$ : Es el número de observaciones en el conjunto de datos.
2.  $y_i$ : Son los valores reales de la variable dependiente.
3.  $\hat{y}_i$ : Son los valores predichos por el modelo para la variable dependiente.

El ECM se utiliza para comparar diferentes modelos o para evaluar la mejora de un modelo después de realizar ajustes. Un valor bajo de ECM indica que el modelo es capaz de predecir con precisión los valores reales, mientras que un valor alto de ECM indica que el modelo tiene un desempeño deficiente en términos de precisión de predicción. Es importante tener en cuenta que el ECM se mide en las mismas unidades que la variable dependiente al cuadrado, por lo que no es fácilmente interpretable en términos absolutos, sino que se utiliza más para comparar diferentes modelos entre sí.

## 2.4. Conceptos Fundamentales de Algoritmos Genéticos

El origen de las especies, escrito por Charles Darwin en 1859, es un libro fundamental en la historia de la ciencia que revolucionó la comprensión de la vida en la Tierra, en resumen, Darwin expuso su teoría de la evolución por selección natural que sostiene que los organismos con rasgos los cuales les permiten adaptarse mejor a su entorno tienen más probabilidades de sobrevivir y reproducirse, proceso que ocurre a lo largo de muchas generaciones y esto conduce la aparición de nuevas especies.

Las ideas de Darwin sobre la evolución natural son la base en la computación evolutiva de esta manera muchas de las propiedades de la evolución de los seres vivos, como la edad, la tendencia a la mutación en relación con el estado de la evolución, están siendo objeto de investigación para su incorporación a las técnicas de computación evolutiva, si bien es una subdisciplina de las ciencias de la computación, también se considera una rama de la inteligencia artificial que se inspira en los mecanismos de la evolución biológica para resolver problemas de optimización haciendo uso de métodos de naturaleza estocástica introduciendo un elemento de aleatoriedad en el proceso de búsqueda de soluciones, lo que puede llevar a resultados diversos (**suarez2021 fundamentos**), cabe señalar que los métodos de búsqueda estocástica son aquellos que no garantizan encontrar la solución óptima, pero que tienen una alta probabilidad de encontrar una buena solución.

Una investigación se desarrolló de forma independiente a la teoría de la evolución de Darwin en la misma época, el trabajo de Gregor Mendel en 1866 en cual realizó experimentos con guisantes y estableció las leyes fundamentales de la herencia genética convirtiéndose en la base de la genética moderna en la que se estudia la herencia y la variación de los rasgos biológicos en los seres vivos, esta

involucra una amplia gama de temas y procesos relacionados con la transmisión de información genética de una generación a la siguiente.

La genética investiga cómo los rasgos y características son transmitidos de padres a hijos a través de la información genética mediante la transmisión de genes, unidades de información que codifican las instrucciones determinando las características de un organismo, y cromosomas, estructuras en el núcleo de las células que albergan los genes. En la reproducción, los organismos heredan una combinación de genes de ambos padres en ciertas ocasiones surgen las mutaciones, cambios que pueden afectar a uno o varios genes dando lugar a nuevas variantes genéticas. Fue en el siglo XX, que las ideas de ambos científicos se unieron en lo que se conoce como la síntesis moderna de la evolución.

Los conceptos fundamentales de la genética son descritos, sin entrar a detalle, a continuación: Los genes son las unidades fundamentales de la herencia formando una secuencia particular que contiene información para una función específica en un organismo también se tiene que mencionar el genotipo que determina las características hereditarias de un organismo, como su color de ojos, grupo sanguíneo, predisposición a enfermedades y muchas otras características biológicas por otro lado el fenotipo es la expresión observable de los rasgos genéticos en un organismo estos incluyen características físicas, comportamientos y cualquier efecto de los genes en la apariencia y el funcionamiento del individuo.

Ya en el mecanismo y funcionamiento se encuentran la recombinación genética, este es un proceso de mezcla y de intercambio en el cual los genes experimentan una redistribución y reordenación contribuyendo a la diversidad genética, en ciertas ocasiones suceden mutaciones que son cambios en los genes los cuales puede ser causados por diversas razones produciendo nuevas variantes genéticas. Algunas mutaciones pueden ser beneficiosas, dañinas o

neutras en términos de la adaptación de un organismo.

Al computo evolutivo se agregó una técnica en la década de 1970 con el trabajo de John Henry Holland que dio lugar a una de las líneas más prometedoras de la inteligencia artificial, debido a la complejidad del problema, encontrar la solución óptima puede ser un proceso extenso y posiblemente interminable. Por lo tanto, es crucial aprovechar la mejor opción disponible en cada paso de la búsqueda (**holland1992adaptation**): Los AGs, se trata, desde luego, de algoritmos de búsqueda basados en la mecánica de la selección y la genética naturales. Combinan la supervivencia del más apto entre estructuras de cadenas con un intercambio de información estructurado pero aleatorio para formar un algoritmo de búsqueda con algo del estilo innovador de la búsqueda humana con una optimización con las mejoras en el rendimiento con el objetivo de alcanzar uno o varios puntos óptimos centrados en dos elementos clave: 1) la búsqueda de mejoras para acercarnos a 2) un punto óptimo (**goldberggenetic**). David E. Goldberg, cuyo director de tesis doctoral fue John H. Holland, popularizó el uso de AGs en problemas de optimización gracias a algunos de sus libros y artículos. En la figura 2.2 se muestra un diagrama de flujo genérico para implementar un AG.

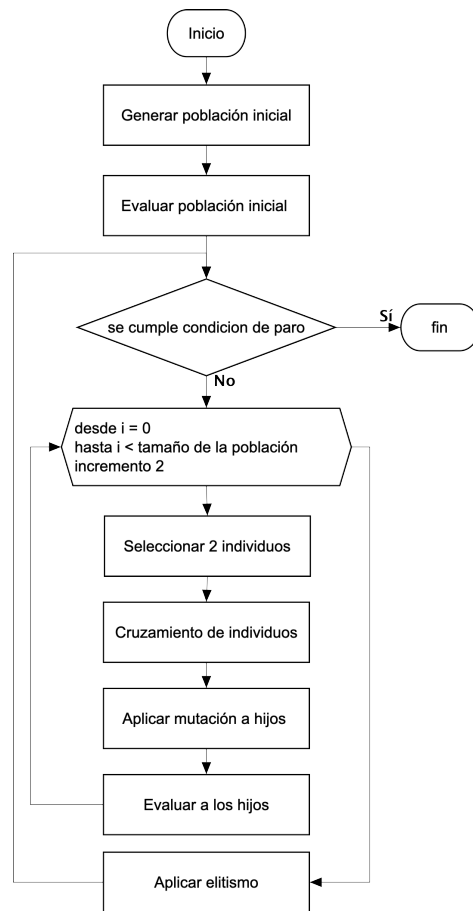


Figura 2.2: Diagrama de flujo para un AG (Elaboración propia)

A continuación, se enumeran los elementos del esquema general de los algoritmos evolutivos, con un enfoque en los AGs, en resumidas cuentas, es emular la evolución natural a través de la selección, recombinación y mutación de procedimientos para encontrar soluciones óptimas o cercanas a óptimas en problemas de optimización y búsqueda.

#### 2.4.1. Representación de individuos

La representación de individuos es crucial para que un algoritmo genético (AG) pueda abordar eficazmente un problema, naturalmente que con este primer

paso en el diseño de un AG es esencial definir la forma de representar los individuos para establecer una relación entre las soluciones candidatas en el dominio del problema original que se denomina fenotipos, es decir, la solución en el espacio original del problema y sus equivalentes en el dominio del AG, con la elección de la representación de los individuos en los algoritmos genéticos tiene un impacto significativo en la efectividad de la resolución de problemas ya que diversas representaciones conducen, por lo general, a resultados distintos. Por lo tanto, es crucial seleccionar la representación que mejor se adapte al problema específico para optimizar la búsqueda de soluciones. **(suarez2021fundamentos)**

Cada posición en la cadena debe tener una correspondencia significativa con el problema que se está abordando porque asegura que los genes que contribuyen a la calidad de un individuo continúen generando características de alta calidad en los nuevos individuos creados a través de operaciones genéticas. **(araujo2009algoritmos)** la representación de un individuo debe ser efectiva permitiendo que el AG encuentre soluciones de alta calidad, eficiente con el objetivo de ser fácil de codificar y decodificar a este propósito se refiere a cómo se transforma una solución del espacio fenotípico en el espacio genotípico, por último, flexible adaptándose a diferentes tipos de problemas. La elección de la estructura de datos utilizada en el espacio genotípico depende del tipo de AG que se esté implementando se pueden usar cadenas de símbolos, vectores de números reales, máquinas de estados finitos o estructuras arbóreas, entre otros, a este conjunto de información se denomina cromosoma. En la figura 2.3 se ejemplifica una representación binaria de un cromosoma.

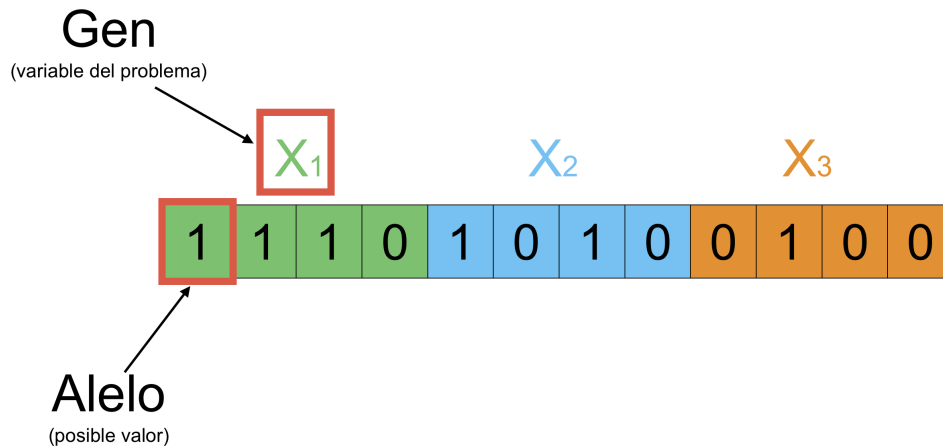


Figura 2.3: Representación binaria del cromosoma de un individuo (Elaboración propia)

#### 2.4.2. Inicialización de la población

La inicialización de la población es una de las fases fundamentales en la implementación de AGs, se busca encontrar soluciones óptimas o aproximadas a un problema mediante la evolución de una población de posibles soluciones a lo largo de generaciones, en este proceso de crear la primera generación de individuos o cromosomas que serán sujetos a la evolución. Hay un tipo de población denominada panmítica (**carson2017Genetic**) y es comúnmente utilizada en AGs y en la computación evolutiva, donde la probabilidad de reproducción entre individuos depende principalmente de su aptitud, lo que fomenta la diversidad genética en la población y la exploración del espacio de búsqueda.

Determinar el tamaño de la población inicial con un número que varía según el problema y considerando las restricciones computacionales, si se crea una población más grande puede aumentar la diversidad y la calidad de las soluciones, pero también puede requerir más tiempo de cómputo. Cada individuo en la población debe estar codificado de manera que represente una solución

potencial al problema esta puede ser binaria, real, entera, o de otro tipo, según las características del problema. Por ejemplo, en un problema de optimización, los individuos podrían representarse como cadenas binarias, y en un problema de diseño de circuitos electrónicos, los individuos podrían ser representados por árboles de expresión.

Para crear los individuos de la población inicial, puedes utilizar varios enfoques, como una inicialización aleatoria donde los individuos se generan aleatoriamente dentro del rango permitido para cada componente de la solución, una inicialización heurística utilizar conocimientos previos del problema para generar individuos iniciales que sean prometedores donde puede acelerar la convergencia del algoritmo y el muestreo estratificado en el cual si existen restricciones o características específicas en el problema, se pueden usar técnicas de muestreo estratificado para asegurar que la población inicial cubra adecuadamente el espacio de búsqueda. En la figura 2.4 se muestra un ejemplo de la población de individuos con cromosomas con valores binarios.

$C_1$	1	1	1	0	1	0	1	0	0	1	0	0
$C_2$	0	1	0	0	1	1	1	0	0	1	1	0
$C_3$	0	1	0	0	0	1	0	0	0	1	0	1
...	...	...	...	...	...	...	...	...	...	...	...	...
$C_{n-1}$	0	1	1	0	1	1	0	0	1	1	0	1
$C_n$	0	1	0	1	1	0	0	1	1	0	0	0

Figura 2.4: Representación de individuos formando una población (Elaboración propia)

La calidad de la inicialización de la población puede influir en la eficacia y

eficiencia del AG con una población inicial diversa y con individuos de alta calidad puede acelerar el proceso de evolución y mejorar las posibilidades de encontrar soluciones óptimas ya que a medida que el algoritmo evoluciona a lo largo de generaciones, los operadores genéticos, como la selección, el cruce y la mutación, se aplican a la población para crear generaciones subsiguientes y refinar las soluciones.

### **2.4.3. Evaluación de la población**

La evolución de la población en algoritmos genéticos está determinada por la calidad relativa de sus individuos, con mayor calidad tienen una mayor probabilidad de sobrevivir y reproducirse, lo que les permite aumentar su presencia en la siguiente generación. Esto se suele evaluar mediante una función de adaptación que refleja la adecuación del individuo como solución al problema. (**araujo2009algoritmos**) y es la que determina qué tan aptos son los individuos que la componen en relación con el problema que se está tratando de resolver, esto quiere decir que se realiza para asignar una puntuación o valor de aptitud a cada individuo, lo que permite a los AGs seleccionar a los individuos más adecuados para la reproducción y la creación de la próxima generación, para empezar, se debe definir una función de aptitud (fitness function) que cuantifique qué tan bueno es un individuo en términos del problema que se quiere resolver. Esta función de aptitud debe ser específica para el problema y reflejar los objetivos de optimización.

La función de aptitud es un componente en los AGs que cuantifica el desempeño de los individuos en una población en relación con los objetivos de optimización y las restricciones de tal manera que debe ser específica para el problema mediante la asignación de valores numéricos a cada individuo. Generalmente se diseña para que los valores más altos representen individuos más óptimos,

sin embargo, en algunos casos, se pueden utilizar funciones de aptitud inversas, donde los valores más bajos indican mejores soluciones con las cuales sirven para guiar la selección de individuos.

Para cada individuo en la población, se aplica la función de aptitud con la intención de asignar un valor numérico que representa la calidad del individuo en relación con el problema de esta manera cuanto mayor sea el valor de aptitud, mejor será el individuo en términos de la solución, por lo tanto, la función de aptitud generalmente está diseñada de manera que los valores más altos se asocien con individuos que cumplen mejor los objetivos de optimización. La elección de una función de aptitud adecuada es esencial para el éxito de un AG.

#### 2.4.4. Selección de padres

La selección de padres es una etapa crucial en un AG que determina cuáles individuos de una población serán elegidos para ser los padres de la próxima generación con el objetivo de dar preferencia a los individuos más aptos, de esta manera sus características se heredan y mejoran a lo largo de las generaciones. En la figura 2.5 se muestra la selección de los padres.

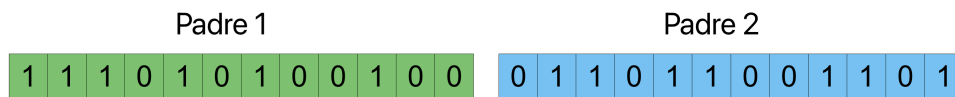


Figura 2.5: Diagrama de flujo para un AG (Elaboración propia)

La selección actúa como guía en la búsqueda genética, orientándola hacia zonas del espacio de búsqueda con mayor potencial de éxito. A lo largo de las últimas dos décadas, se han desarrollado, analizado y comparado numerosos métodos de selección. (**gen1999genetic**), la selección basada en la aptitud con la cual la probabilidad de selección de un individuo está directamente relacionada con su

aptitud, a mayor aptitud se tienen una mayor probabilidad de ser seleccionados como padre por otra parte la selección por torneo se seleccionan a los individuos directamente en función de su aptitud, se eligen al azar varios individuos (un "torneo") de la población y se selecciona al mejor de entre ellos para ser un padre tiene la ventaja de no calcular la aptitud de todos los individuos en una selección por rueda de ruleta ponderada con este enfoque en lugar de considerar la aptitud absoluta, esta estrategia clasifica a los individuos en función de su aptitud relativa en comparación con el resto de la población, posteriormente a esto, se utiliza una rueda de ruleta ponderada para seleccionar padres, otra que se puede encontrar es la selección de elite donde un cierto número de los individuos más aptos de la generación actual se seleccionan directamente para ser padres de la siguiente generación, sin necesidad de competir en el proceso de selección asegurando que las mejores soluciones perduren a través de las generaciones y finalmente la selección por sus propias preferencias es similar a la selección basada en la aptitud, pero en lugar de utilizar una rueda de ruleta, se selecciona un grupo de padres a intervalos regulares basados en su aptitud con la distancia entre los intervalos se determina por la aptitud acumulativa de la población.

La elección de la estrategia de selección de padres depende del problema que estás resolviendo y de los objetivos específicos de tu AG se pueden combinar varias estrategias de selección para lograr un equilibrio en la exploración y la explotación del espacio de búsqueda.

#### **2.4.5. Recombinación o cruza**

La cruza o recombinación, es una de las operaciones en los AGs en la que implica la combinación de información genética de dos individuos seleccionados de la población a los que se denominan padres con la intención de crear uno o más descendientes, es decir, los hijos con esto se pretende simular el proceso

de recombinación genética que ocurre en la reproducción sexual de organismos en la naturaleza. Existen varios tipos de cruza, de las más reconocidas esta la cruza de un punto con este método se elige un punto de referencia al azar en los cromosomas de los padres, los genes antes del punto de cruce se toman de un padre, y los genes después del punto de cruce se toman del otro padre mediante este proceso se crean dos descendientes, uno con la parte inicial de un padre y la parte final del otro, y viceversa. En la Figura 2.6 se puede apreciar el funcionamiento de este operador genético.

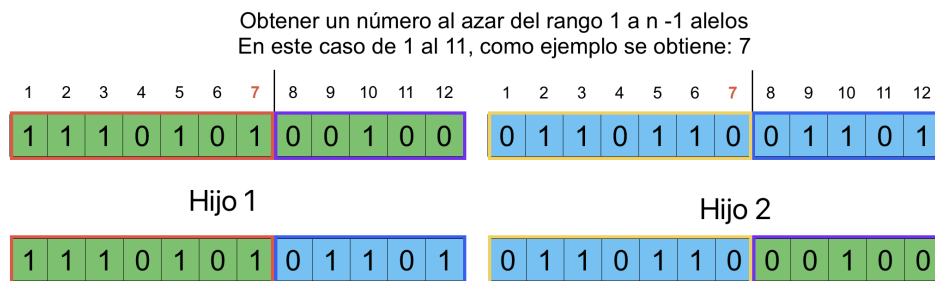


Figura 2.6: Aplicando el cruce mediante la técnica de la cruza de punto (Elaboración propia)

Aunque la principal forma de cruzar individuos es la cruza de un punto existen variaciones como la cruza de dos puntos, pero en este caso, se eligen dos puntos de cruce en donde los segmentos entre los dos puntos de cruce se intercambian entre los padres, otra técnica es realizar una cruza uniforme por la que cada gen en el cromosoma de los padres, se elige aleatoriamente de cuál de los dos padres se heredará ese gen, es decir se va tomando los genes de cada padre para crear a los hijos de tal forma que la repartición de genes quede distribuida en ellos de forma uniforme. Para realizar una optimización numérica existe la cruza de aritmética en esta, en lugar de intercambiar genes, se realiza una combinación lineal de los valores de los genes de los padres para generar los valores de los hijos y por ultimo la cruza basada en secuencias se utiliza en problemas donde los cromosomas representan secuencias ordenadas, como el problema del agente

viajero en esta opción los descendientes heredan secuencias de genes de ambos padres de manera que mantengan el orden relativo de los genes.

El operador de cruce se aplica a los padres seleccionados con una probabilidad generalmente alta. Cuando no se aplica el cruce, los padres se copian sin modificaciones, creando copias idénticas (clones) en la siguiente generación. (wirsansky2020hands) determinando la probabilidad de que se aplique el operador de cruce a un par de padres en todo caso un valor alto de en la tasa de cruce significa que es más probable que se aplique el cruce, lo que puede llevar a una mayor exploración del espacio de soluciones, pero también puede aumentar la diversidad genética de la población, en contraste, un valor bajo de probabilidad de cruce llevará a una exploración más limitada del espacio de soluciones.

La elección del tipo de cruce depende del problema que se esté resolviendo y la estructura de los cromosomas tiene la funcionalidad más importante en la creación de diversidad genética en cada generación ya que contribuye a la evolución y mejora de las soluciones a lo largo de las generaciones. Es necesario recalcar que existe la posibilidad de seleccionar varias opciones de cruce para la ejecución de un AG porque, al ser una búsqueda de optimización, la diversidad de los individuos ayudara a encontrar la solución al problema con mayor eficacia.

#### **2.4.6. Mutación**

La mutación es otra operación en AGs que introduce variabilidad en la población y permite la exploración de soluciones potencialmente nuevas y no representadas en la población actual que desempeña un papel crítico por varias razones por una parte introduce aleatoriedad en la población, lo que ayuda a prevenir que el algoritmo se estanque en óptimos locales y promueve la exploración de un espacio de búsqueda más amplio proporcionando soluciones

potencialmente nuevas que pueden ser diferentes de las de los padres que resulten beneficiosas para la resolución del problema.

En problemas con múltiples óptimos locales, la mutación puede ser la única manera de salir hacia una solución óptima global sin embargo existe una tasa de mutación o probabilidad de mutación que debe ser cuidadosamente ajustada, si es demasiado baja puede llevar a una convergencia prematura, mientras que una tasa demasiado alta puede reducir la convergencia y aumentar el tiempo de ejecución. Está claro que la mutación tiene un impacto directo en la diversidad genética de la población, pero la forma en que se implementa la mutación puede variar según el tipo de problema puesto que implica realizar cambios aleatorios en los cromosomas de los individuos seleccionados.

Para poner en marcha la mutación se selecciona un conjunto de individuos de la población para someterlos en los que son elegidos aleatoriamente o también por algún criterio específico de estos individuos se eligen uno o más genes dentro del cromosoma para realizar la mutación es prudente mencionar que variar según el diseño del algoritmo y el problema, dichos genes seleccionados se modifican de alguna manera que puede implicar cambiar el valor de un gen, intercambiar genes, eliminar genes o agregar nuevos genes, finalmente como se hizo referencia con anterioridad, definir una tasa de mutación (Araujo Cervigón, 2009) determinara la frecuencia en que se aplicará y resulta que, cuando una tasa de mutación es baja garantiza que la población actual se conserve en gran medida, mientras que una tasa alta aumenta la exploración, sin embargo puede afectar negativamente la explotación. La implementación cuidadosa y la configuración de la tasa de mutación son críticas para el éxito del algoritmo en la búsqueda de soluciones de alta calidad en un espacio de búsqueda en la optimización. En la figura 2.7 se ejemplifica el operador de la mutación a uno de los hijos obtenidos con anterioridad.



Figura 2.7: Mutación de dos alelos en uno de los hijos resultantes (Elaboración propia)

### 2.4.7. Elitismo

El algoritmo genético requiere de elitismo para poder converger al óptimo a partir de esto ha sido una estrategia comúnmente utilizada en AG teniendo su grado de importancia en la mejora del rendimiento ya que tiene como objetivo principal preservar a los individuos más aptos de una generación a la siguiente efectuando una conservación de soluciones óptimas con aquellos individuos con los valores de aptitud más altos en cada generación con esto se garantiza que las soluciones óptimas no se pierdan y permanezcan en la población teniendo como consecuencia una aceleración de la convergencia hacia soluciones de alta calidad, esto es especialmente útil en problemas donde se conocen soluciones cercanas al óptimo global.

La manera de realizar la implementación del elitismo en un AG es después de aplicar la función de aptitud a todos los individuos de la población, se seleccionan uno o más individuos élite, generalmente aquellos con los valores de aptitud más altos, para su inclusión en la próxima generación esta cantidad de individuos élite a preservar es un parámetro configurable y típicamente es un número pequeño en comparación con el tamaño total de la población, estos individuos reemplazan a un número equivalente de individuos en la población, generalmente aquellos de menor calidad de esta forma se asegura que los individuos élite se mantengan en la población y tengan la oportunidad de influir en las siguientes generaciones.

El elitismo es importante en los AG por varias razones se deduce que la principal es la preservación de soluciones óptimas que garantiza que las mejores

soluciones encontradas hasta el momento no se pierdan durante el proceso evolutivo acelerando la convergencia hacia soluciones de alta calidad, lo que puede ahorrar tiempo y recursos computacionales dando una estabilidad y consistencia del algoritmo, reduciendo la variabilidad en la calidad de las soluciones generadas y tiene la propiedad de poder ajustarse según la naturaleza del problema y las características de la población, lo que lo hace flexible y aplicable a una amplia gama de situaciones dando una implementación adecuada puede mejorar significativamente el rendimiento y la eficacia de los AG en la resolución de problemas de optimización y búsqueda.

#### **2.4.8. Iteración**

La iteración es el motor de mejora en los algoritmos genéticos esto es a través de múltiples generaciones que los AGs pueden refinar y evolucionar soluciones en la búsqueda de óptimos globales o soluciones de alta calidad para problemas de optimización y búsqueda. Cada generación es una oportunidad para mejorar las soluciones en comparación con la generación anterior porque tiende a mejorar su adaptación al problema, lo que puede conducir a soluciones de mayor calidad a lo largo del tiempo en el cual se permite un equilibrio entre la exploración (buscando soluciones nuevas) y la explotación (refinando soluciones existentes) a medida que se ajusta la población en cada generación.

El proceso de iteración en un AG consiste realizar los pasos anteriores hasta llegar a una solución, en cada generación, se seleccionan individuos de la población actual para ser padres ejecuta una cruce para generar una descendencia aplicando la mutación a los individuos de la descendencia con una probabilidad o tasa de mutación generalmente baja para garantizar la diversidad en los individuos que se evalúan mediante la función de aptitud para determinar su calidad en relación al problema eligiendo individuos para formar la

población de la siguiente generación aplicación de estrategias como el elitismo y finalmente existen varias condiciones para detener el GA, por ejemplo: Encontrar una solución aceptable, alcanzar un número determinado de generaciones o la disminución de las mejoras cualitativas en la población (**gridin2021learning**) por lo cual se verifica alguna situación como un número máximo de generaciones, convergencia de la población o mejora de la solución esto permite detener el AG en la solución que más se acerque a lo óptimo de acuerdo a las necesidades del problema.

#### **2.4.9. Gráfica de convergencia**

La gráfica de convergencia es una herramienta AG que proporciona información visual sobre cómo evoluciona la calidad de las soluciones (**padilla2023algoritmos**) a lo largo de las generaciones es esencial porque proporciona una representación visual de cómo cambia la calidad de las soluciones a medida que progresa el proceso de evolución a lo largo de las generaciones permitiendo a los investigadores y practicantes evalúen el rendimiento de un AG, identificado la rapidez con la que converge hacia soluciones óptimas y determinar si el algoritmo está estancado en óptimos locales se trata, desde luego de ayudar a ajustar los hiperparámetros del AG, como la tasa de mutación, la selección y la población, para lograr un equilibrio entre exploraciones.

Durante la ejecución del AG, se registra la calidad de la mejor solución (generalmente el valor de aptitud) en cada generación para cada población actual o para un conjunto de individuos élite, según la configuración del AG con ayuda de un software de visualización o una biblioteca de gráficos para representar los valores de aptitud registrados a lo largo de las generaciones en donde por correspondencia el eje x se le asigna el número de generación, y en el eje y

se representa la calidad de la mejor solución con ello se analizan los datos para evaluar la tendencia de mejora en la calidad de las soluciones en busca de identificar la convergencia hacia soluciones óptimas o cualquier patrón de estancamiento si muestra una disminución constante y pronunciada en el valor de aptitud a lo largo de las generaciones sugiere una convergencia rápida hacia una solución óptima o de alta calidad, una estabilización o un cambio mínimo en la calidad de las soluciones durante un período prolongado de generaciones indica un estancamiento en óptimos locales por lo que se procede a ajustar los hiperparámetros para promover una exploración más amplia, sin embargo si se muestran fluctuaciones excesivas en la calidad de las soluciones podría indicar que el AG existe una variabilidad excesiva y es necesario mejorar el equilibrio entre exploración. Resulta entonces que la gráfica de convergencia proporciona una forma de medir el éxito del algoritmo en la resolución del problema, cuanto más cerca esté el valor de aptitud de la solución de la convergencia al valor óptimo, mejor será el desempeño del AG conservando el rendimiento del algoritmo en la resolución de problemas de optimización y búsqueda. En la figura 2.8 muestra un ejemplo de una gráfica de convergencia de un AG.

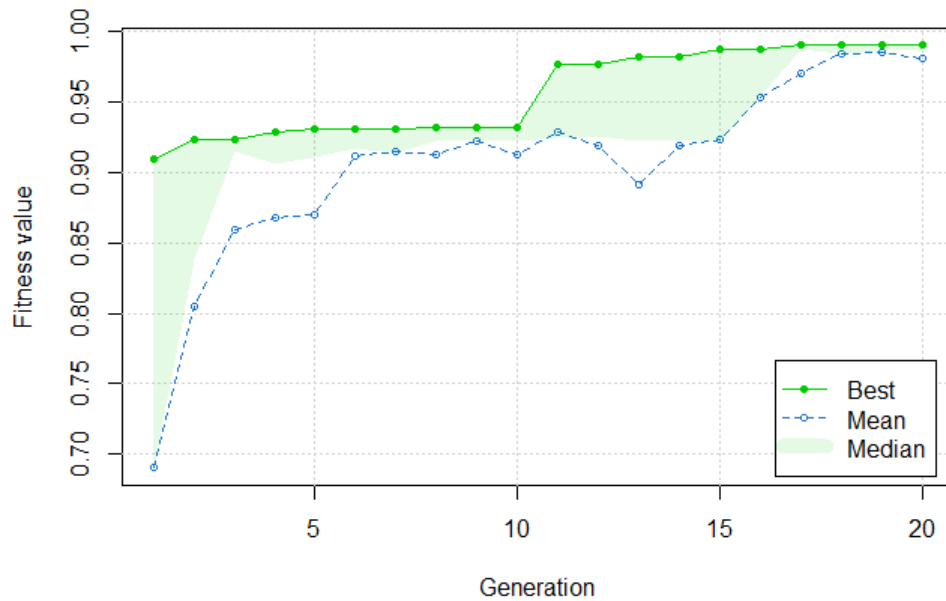


Figura 2.8: Ejemplo de una gráfica de convergencia de una AG. (2020). Gráfica convergencia [Figura]. Ícaro Agostino. <https://icaroagostino.github.io/post/sbo/>

## 2.5. Estado del Arte

Los algoritmos genéticos han demostrado ser robustos y eficaces en la optimización del Costo-Tiempo, superando a otras técnicas comúnmente utilizadas por lo que recomienda realizar investigaciones más profundas para validar estos resultados iniciales y promover mejoras continuas en este campo. (**zheng2002applying**), los AGs demuestran una ventaja significativa sobre las técnicas tradicionales, como el Método de la Ruta Crítica (método de la ruta crítica), los Métodos de Programación Lineal (LP) y por Enteros (IP) en el estudio se destaca que los AGs trabajan con una codificación del conjunto de parámetros permitiendo una mayor flexibilidad en la representación de las variables realizando búsquedas desde una población de soluciones garantizando una exploración global y evitando la convergencia prematura a mínimos locales con base en funciones objetivo en lugar de derivadas, lo que simplifica la

formulación de modelos matemáticos con una elección aleatoria al utilizarla como guía de la búsqueda por lo que se puede deducir que el uso de AG tiene el potencial de mejorar la eficiencia y la rentabilidad para estos proyectos.

Los AGs se presentan como una herramienta viable para abordar problemas de planificación de proyectos debido a su flexibilidad inherente con esta característica los hace particularmente útiles para problemas complejos donde la información disponible puede ser limitada o incompleta teniendo la capacidad de resolver problemas multiobjetivos de manera simultánea, considerando múltiples criterios de optimización al mismo tiempo. (**garcia2016experiencias**) donde se destaca la complejidad (NH-duro) de este problema y su aplicabilidad en organizaciones orientadas a proyectos, como una propuesta de solución se ha adoptado un enfoque basado en penalizaciones debido a la dificultad de obtener soluciones completamente correctas por la complejidad del problema de optimización, es importante señalar que, aunque las soluciones encontradas por los algoritmos a veces se consideran las mejores, en algunos casos no cumplen completamente con todas las restricciones del problema lo cual los autores recomiendan que el proceso de planificación sea asistido por algoritmos y herramientas computacionales, pero también sea verificado y ajustado por expertos humanos especializados en gestión de proyectos.

La aplicación del algoritmo NSGA-II al proyecto de construcción de la Etapa I de la Sede Universitaria Zarzal, partiendo de la programación inicial, ha generado un conjunto de soluciones cuasi óptimas (pares de tiempo-costo) que se representan en un frente de Pareto con estas soluciones permiten al equipo del proyecto evaluar distintas opciones para reducir o demorar la duración del proyecto a un costo mínimo. (**velasquez2019aplicacion**) En donde se abordó el Problema discreto de compensación tiempo-costo (DTCTP), que puede considerarse tanto de forma determinista como estocástica, y se evaluaron métodos exactos y no exactos, tanto de tipo mono objetivo como multi objetivo indicando que el

algoritmo NSGA-II generó soluciones cuasi óptimas en términos de tiempo y costo, lo que puede ser de gran relevancia en la toma de decisiones en la gestión de proyectos de construcción.

El algoritmo NSGA-II tradicional opera con enteros y números reales, pero en algunos escenarios como en el trabajo de Ralf Luis de Moura, Teresa Cristina Janes Carneiro, Teresa y Marcos Paulo Valadares de Oliveira fue necesario trabajar con variables binarias (**de2020unveiling**), para lograr esto, se realizaron los siguientes pasos: Inicialmente, se creó una población de 100 cromosomas (individuos) con 57 genes representados por "0" y "1".<sup>en</sup> donde "0" indicaba que el gen debía estar "apagado", es decir, excluido del cálculo, y "1" indicaba que el gen debía estar "encendido", es decir, incluido en el cálculo. El proceso se repitió hasta alcanzar la saturación, y se observó que el algoritmo convergía a una solución óptima después de 1000 generaciones esto quiere decir que la adaptación de NSGA-II para trabajar con variables binarias se ha convertido en una herramienta valiosa para abordar problemas de optimización en los que las variables son de naturaleza binaria permitiendo abordar una gama más amplia de problemas de gestión y optimización de proyectos.

Se ha creado un modelo de optimización que considera múltiples alternativas y propósitos, permitiendo un análisis tridimensional que incluye tiempo, costo y calidad, a diferencia del enfoque tradicional bidimensional de tiempo y costo, este modelo busca planes de utilización de recursos óptimos que minimicen tanto el tiempo como el costo de construcción, mientras maximizan la calidad final. (**isikyildiz2020multi**) al buscar planes para utilizar recursos óptimos que minimicen el tiempo y el costo de construcción, además de maximizar la calidad donde se investigó el efecto del tamaño de la población en el valor de Pareto en la aplicación del AG encontrando que el número de imágenes de Pareto aumentó con un aumento en el número de poblaciones resultado que la opción con el período de tiempo más corto y el costo mínimo existía en la aplicación con 100 y

1000 poblaciones, en términos de costo, se alcanzó el costo mínimo con el mismo valor en todas las poblaciones y, por lo tanto, se logró la duración mínima para el proyecto, sin embargo, en cuanto al valor de calidad, no se cumplió el objetivo para 50 poblaciones, y el valor de calidad siguió siendo inferior al de las otras poblaciones.

Un modelo de optimización tiempo-costo fue desarrollado para proyectos de construcción en condiciones de incertidumbre, utilizando el algoritmo genético y la teoría de conjuntos difusos con lo que permitió evaluar diferentes técnicas constructivas para un edificio en particular, considerando las incertidumbres presentes en el proyecto. (**acar2019time**) obteniendo diferentes pares de tiempo y costo óptimos utilizando un modelo de optimización basado en AG y la técnica de Pareto permitiendo a los gerentes de proyectos utilizar conjuntos de soluciones según el nivel de riesgo determinado, además, las componentes en el conjunto de soluciones se pueden modificar, y el modelo puede volver a analizarse. El trabajo demuestra que el modelo propuesto es capaz de encontrar soluciones óptimas que tengan en cuenta la incertidumbre y los niveles de riesgo siendo flexible según las necesidades combinando la lógica difusa y los AG.

## Capítulo 3

### METODOLOGÍA

En el presente trabajo se emplea un enfoque basado en algoritmos genéticos como metodología principal para abordar el problema de la optimización en la generación de proformas en proyectos de ingeniería civil, con este enfoque que es particularmente adecuado para problemas complejos y es necesario encontrar soluciones eficientes para la asignación de recursos, tiempos y costos.

Los algoritmos genéticos permiten explorar un amplio espacio de soluciones posibles de manera eficiente, lo que los hace adecuados para problemas de optimización como el que se presenta en este trabajo con la simulación de procesos evolutivos, este tipo de algoritmos busca mejorar iterativamente las soluciones candidatas, seleccionando las mejores con base en una función de aptitud que evalúa qué tan bien se ajustan a los objetivos planteados. En este capítulo se detallará la metodología utilizada para implementar el algoritmo genético y a lo largo de este proceso, se han seguido varias etapas fundamentales para garantizar la efectividad del algoritmos genéticos aplicado al contexto de la generación de proformas.

Estas etapas incluyen la definición de la población inicial, el diseño de la función de aptitud, la aplicación de operadores genéticos (selección, cruce, mutación y elitismo), y el criterio de convergencia para determinar cuándo detener la búsqueda de soluciones. A continuación, se explica cada una de estas etapas en detalle, describiendo el rol que desempeñan en la generación de soluciones

optimizadas para el problema planteado.

### **3.1. Diseño del Algoritmo Genético**

Diseñar un algoritmo genético implica definir la representación de las soluciones, la función de aptitud, y los operadores genéticos (elitismo, selección, cruzamiento y mutación) en las siguientes secciones se explicara cada uno de los pasos para comenzar con el diseño del algoritmo genético.

#### **3.1.1. Representación de la Solución**

Como se menciona en la sección 1.2 existen curvas de distribución que se utilizan para realizar el reparto en el tiempo del presupuesto de una tarea, sin embargo, uno de los objetivos de la presente investigación es comprobar la eficiencia de la distribución del presupuesto de las tareas en la gestión de proyectos mediante el uso de Algoritmos Genéticos. El primer paso es definir una representación de la solución, como resultado de un análisis de la información en la creación de proformas, se observa que cada una de las curvas de distribución se pueden representar como una serie de números reales en el rango de 0 a 1, los cuales de igual forma representan el avance de una tarea a lo largo del tiempo.

Dado que se está tratando con el avance de la tarea a lo largo del tiempo, se puede representar cada solución como un cromosoma donde cada gen representa el avance de un día específico, por ejemplo, si se tienen  $N$  días restantes de la tarea, el cromosoma se establece en una secuencia de  $N$  genes definiendo cada gen del cromosoma como un número real en el intervalo  $0 > n < 0.5$ . La representación del cromosoma se define de la siguiente manera:

$$Cromosoma = [a_1, a_2, a_3, \dots, a_n] \quad (3.1)$$

Donde  $a_n$  es el avance remanente de una tarea para el día  $n$ , y cada  $a_n$  debe de estar en el rango  $0 > n < 0.5$  ya que se considera que  $n < 50\%$  como el limite extremo del avance diario de una tarea, cabe señalar que la suma de los genes debe ser igual al avance total remanente. Por ejemplo, un cromosoma válido para una tarea con el 20% de avance podría ser  $[0.1, 0.2, 0.2, 0.2, 0.1]$ , donde la suma de los genes es igual al avance remanente ( $0.1 + 0.2 + 0.2 + 0.2 + 0.1 = 0.8$ ).

### 3.1.2. Función aptitud (fitness)

Se elige el Error Cuadrático Medio (ECM) como función de aptitud debido a su alta precisión en la evaluación, dado que es ampliamente utilizado en problemas de regresión y optimización midiendo la diferencia entre los valores predichos y los valores reales. Haciendo uso de la biblioteca NumPy en Python, la cual proporciona los métodos implementados para calcular la media aritmética de los elementos de un arreglo, la función llamada **mean** calcula tomando como argumento el array o la secuencia de valores y opcionalmente el eje a lo largo del cual se calculará la media.

El objetivo es medir la diferencia entre una distribución de avance planeada con las distribuciones ajustadas resultantes como soluciones candidatas, para esto se establece la función de aptitud como una inversa (ya que se busca minimizar) del ECM obtenida por la función mean de NumPy . Quedando de la siguiente manera:

$$aptitud = \frac{1}{1 + (\sum_{i=1}^n (y_i - \hat{y})^2)} \quad (3.2)$$

Este sería el resultado final, donde la función de aptitud se define como el inverso del ECM más uno, lo cual es común en algoritmos evolutivos para asegurar que la aptitud sea una función que minimice los errores y maximice la exactitud del modelo.

Donde:

1.  $y_i$ : Representa los valores observados o reales del conjunto de datos, en este caso el avance real de la actividad.
2.  $\hat{y}_i$ : Representa los valores predichos o calculados por el modelo para cada  $i$ -ésimo dato en el conjunto, en este caso el valor planeado de la actividad.
3.  $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ : Esta parte de la fórmula es la suma de los cuadrados de las diferencias entre los valores reales y los valores predichos, indicando qué tan lejos están los valores predichos de los valores reales, por lo tanto, si esta suma es pequeña, significa que los valores predichos están muy cerca de los valores reales, lo que indica un buen desempeño del modelo.
4.  $1 + (\sum_{i=1}^n (y_i - \hat{y}_i)^2)$ : Al añadir 1 a la suma de los errores, se asegura que el denominador nunca sea cero. Esto evita problemas de división por cero y garantiza que la aptitud siempre sea un valor positivo.
5.  $\frac{1}{1 + (\sum_{i=1}^n (y_i - \hat{y}_i)^2)}$ : Finalmente, la aptitud se define como el inverso del valor calculado lo que convierte la medida de error en una medida de aptitud, donde un error grande resulta en una aptitud pequeña y viceversa, cuanto menor sea la suma de los errores al cuadrado, mayor será la aptitud de la solución.

### 3.1.3. Inicialización de la Población

La inicialización de la población es un paso crucial en un algoritmo genético esto es porque debido a la calidad de la población inicial puede afectar significativamente el rendimiento del algoritmo, como se menciona en la sección 2.4.2 existen varias estrategias para la inicialización de la población, en el diseño algoritmo genético se pretende usar una inicialización utilizando la distribución normal.

En el primer paso, se genera una población inicial compuesta por múltiples individuos, cada uno representado como una serie de números reales de tamaño fijo determinado por los días por finalizar la tarea, dichos números aleatorios son generados dentro de un rango predefinido (entre 0 y 0.5) lo cual asegura que los valores iniciales sean suficientemente diversos, permitiendo explorar un espacio de soluciones amplio desde el inicio del algoritmo evitando la convergencia prematura del algoritmo.

Tras la generación de la población inicial, es necesario ajustar los valores de los genes de cada individuo para garantizar que la suma de estos valores sea igual a un remanente específico, esto se logra al realizar un normalizando de cada arreglo con respecto a la suma de sus componentes, y luego escalando dicha suma al valor del remanente, Asegurando que cada individuo en la población no solo respete una condición de suma predefinida, sino que también se mantenga en una escala comparable, lo cual es esencial para la evaluación de la aptitud en contextos donde los recursos o probabilidades totales deben ser conservados.

Finalmente, la población ajustada se convierte en un arreglo de NumPy y se retorna para su uso en las etapas posteriores del algoritmo genético, con este formato estructurado facilita las operaciones vectorizadas y manipulaciones eficientes durante la evolución de la población

### 3.1.4. Operadores Genéticos

La técnica de **elitismo** en un algoritmo genético, asegura que los mejores individuos (con la mayor aptitud) se preserven sin cambios en la nueva generación para evitar que las mejores soluciones se pierdan durante el proceso evolutivo y para garantizar que la calidad de la población mejore o, al menos, se mantenga a lo largo de las generaciones. Como parte de la selección natural, es una estrategia eficiente para equilibrar la explotación de las mejores soluciones conocidas con la exploración de nuevas soluciones potenciales.

En primer lugar se determina el número de individuos que se considerarán "élite" dentro de la población actual y representan las soluciones más aptas (con los valores de aptitud más altos) y, por tanto, se preservarán en la siguiente generación sin sufrir modificaciones y se obtiene multiplicando el tamaño total de la población por la tasa de elitismo y convirtiendo el resultado en un número entero. La implementación del elitismo en algoritmos genéticos es esencial para garantizar la retención de las mejores soluciones encontradas.

Una vez determinado cuántos individuos serán seleccionados como élite, se identifican estos individuos utilizando la función de Numpy `argsort` para ordena los índices de los individuos en la población según sus valores de aptitud en orden ascendente, para posteriormente invertir el orden con `[: :-1]` y así obtener los índices de los individuos con la aptitud más alta, así la nueva generación comienza con los individuos élite identificados y se copian a la nueva generación sin modificación alguna por lo que se facilita la transmisión de soluciones de alta calidad, lo que puede mejorar significativamente la eficiencia del algoritmo genético en la búsqueda de la solución óptima.

La elección del operador de **selección** en un algoritmo genético depende en gran medida de la naturaleza del problema y de los objetivos específicos en este

sentido la elección es la *selección por torneo*, una técnica comúnmente utilizada para elegir individuos que pasarán a la siguiente generación o que se utilizarán para la reproducción. Obteniendo un subconjunto de individuos de la población que se enfrentan en un "torneo", y el ganador (el individuo con la mayor aptitud) avanza a la siguiente ronda o se selecciona para su uso en el siguiente paso del algoritmo.

Para iniciar, se seleccionan aleatoriamente 16 individuos de la población con esto se asegura que cada individuo tenga una oportunidad justa de ser seleccionado para participar en el torneo, lo que contribuye a la diversidad genética y evita sesgos en la selección. El torneo se organiza en varias rondas hasta que solo queden dos individuos. que van compitiendo en pares, comparando sus valores de aptitud. El proceso se implementa en un ciclo que continúa hasta que el número de individuos seleccionados se reduce a dos, garantizando que solo los individuos más aptos, que han ganado repetidamente en sus enfrentamientos, avancen en el torneo.

El enfrentamiento se organiza de manera que el primer individuo compite contra el último, el segundo contra el penúltimo, y así sucesivamente, el individuo con la mayor aptitud se guarda en una lista que representará los ganadores de la ronda actual y se convertirá en los participantes de la siguiente ronda. Este proceso se repite hasta que solo queden dos individuos, que serán los ganadores finales del torneo, con una actualización constante se asegura que solo los individuos más fuertes (con la aptitud más alta) se mantengan en la competencia aumenta la probabilidad de que los mejores genes se transmitan a la próxima generación.

El **cruzamiento** es esencial para combinar información genética de dos padres y generar descendencia influyendo en la exploración del espacio de búsqueda y en la convergencia del algoritmo genético. Se ha elegido el **cruce de un punto** esta es un técnica específica en la que se selecciona un punto en los genomas de los

padres y se intercambia la información genética a partir de ese punto.

Para iniciar se decide si se llevará a cabo el cruce entre dos individuos basándose en una tasa de cruce, que define la probabilidad de que el cruce ocurra, esto se lleva a cabo si un número aleatorio generado es menor que la tasa de cruce entonces se procede con el; de lo contrario, los padres se devuelven sin cambios como los hijos. Esta condición probabilística permite controlar la frecuencia del cruce en la población, equilibrando la exploración de nuevas combinaciones genéticas con la conservación de individuos ya adaptados.

Si se decide realizar el cruce, se selecciona aleatoriamente un punto de cruce dentro de los genomas de los padres, el cual ocurre entre el segundo y el penúltimo gen para asegurar que ambos padres contribuyan significativamente a la genética de los hijos, afectando directamente la diversidad genética y la capacidad del algoritmo para explorar el espacio de soluciones. Una vez seleccionado el punto de cruce, se generan dos nuevos individuos (hijos) combinando la genética de los padres. El primer hijo se crea concatenando los genes de primer padre desde el inicio hasta el punto de cruce y los genes del padre 2 desde el punto de cruce hasta el final. De manera similar, el segundo hijo se genera invirtiendo los roles de los padres. Así los hijos heredan características de ambos padres.

En los algoritmos genéticos (AGs) que trabajan con números reales, la **mutación** juega un papel crucial para mantener la diversidad genética y evitar el estancamiento en soluciones locales, la mutación permite que el algoritmo explore nuevas soluciones y escape de óptimos locales, contribuyendo a la diversidad genética de la población.

El primer paso de la mutación consiste en generar una máscara booleana que indica qué genes del individuo serán mutados, se crea generando un arreglo de números aleatorios para cada gen y comparando cada valor con la tasa de

mutación, si el número aleatorio es menor que la tasa de mutación, se tomara esa posición de la máscara, lo que indica que el gen correspondiente será mutado.

Este enfoque probabilístico permite controlar la frecuencia de mutación, asegurando que no todos los genes se muten simultáneamente, lo cual podría desestabilizar al individuo, con los genes seleccionados para mutación se reemplazan por nuevos valores aleatorios generados entre 0 y 0.5, asignando nuevos valores aleatorios a las posiciones del individuo donde la máscara es elegida, lo que puede permitirle adaptarse mejor a su entorno si las mutaciones resultan beneficiosas.

### **3.2. Criterio de paro**

Se utilizan dos criterios de paro para este problema de optimización, en primer lugar el criterio de mejor individuo se realiza el paro del algoritmo cuando la aptitud de la mejor solución en la población iguale o supere el umbral objetivo, garantizando la eficiencia en la obtención de una solución de alta calidad. En segundo lugar se elige el criterio de número máximo de generaciones donde el algoritmo se detendrá tras alcanzar un número predefinido de generaciones, garantizando así que se obtendrá una solución en un tiempo limitado, lo cual es crítico para aplicaciones con restricciones de tiempo o recursos.

El criterio de paro basado en la mejor aptitud es eficaz en problemas donde es posible definir con precisión lo que constituye una solución óptima o suficientemente buena. Al detener el algoritmo una vez que se alcanza esta aptitud, se asegura que el proceso evolutivo sea eficiente, evitando iteraciones innecesarias y garantizando que se cumplan los objetivos del problema, esto es particularmente útil en situaciones donde el tiempo o los recursos son limitados, y se requiere una solución de alta calidad en un marco de tiempo específico.

### 3.3. Algoritmo

Para obtener la distribución con mayor acercamiento del avance remanente con el avance planeado y así ayudar a establecer proformas de manera eficiente se utiliza el siguiente procedimiento:

---

**Algorithm 1:** Algoritmo genético para redistribuir el presupuesto

---

**procedure** ALGORITMOGENETICO(*distribución planeada, avance, tamaño de población, número de generaciones, probabilidad de reproducción, probabilidad de mutación*)

    InicializarPoblación(Tamaño de población, Tamaño de cromosoma)

    EvaluarPoblación()

**while** *No se cumplan los criterios de paro* **do**

        AplicarElitismo()

        SeleccionarPadres()

        CruzarPadres(Probabilidad de reproducción)

        MutarDescendencia(Probabilidad de mutación)

        EvaluarIndividuos()

**end while**

**return** *Mejor individuo de la población final*

**end procedure**

---

Se inicia con la función iniciar población, realiza la creación de los individuos con la definición de los cromosomas que se explica en la sección 3.1.3, el número de individuos que se generan depende de el número de población que se indique al finalizar retornara una matriz.

Ya que se tiene la matriz de la población, se procede a evaluar la capacidad de cada individuo con utilizando la función de aptitud descrita en la sección 3.1.2 así se puede cuantificar el rendimiento o la idoneidad de cada solución potencial en

relación con los objetivos del problema.

Con la información que se tiene hasta el momento se inicia la iteración utilizando los operadores genéticos elegidos en la sección 3.1.4, comenzado con la selección de dos individuos, que a menudo se denominan padres, y con ellos se realiza la operación de cruce con una tasa de probabilidad mediana-alta para generar dos nuevos individuos con los cuales a su vez se leas aplica el operador de mutación con una tasa de probabilidad baja-media, al finalizar esta operación se realiza la evaluación de los dos nuevos individuos para finalmente realizar el reemplazo en la población de las peores soluciones con los nuevos individuos resultantes de las operaciones genéticas, este proceso continua hasta que se cumplan uno de los dos criterios de paro que se determinaron en la sección 3.2.

## Capítulo 4

# RESULTADOS EXPERIMENTALES

### 4.1. Recursos utilizados

Para realizar el desarrollo del algoritmo genético basado en el diseño que se describe en el capítulo 3 y por consiguiente poder ayudar a establecer proformas mediante la distribución del presupuesto en el seguimiento de un proyecto en el momento que exista un retraso o adelanto a nivel de tareas, se utiliza el lenguaje de programación **Python 3.9.6** ya que cuenta con una gran cantidad de bibliotecas para el análisis de datos entre las que destacan: **NumPy 1.26.4**, que proporciona una estructura de datos extremadamente eficiente para trabajar con matrices multidimensionales y matrices numéricas, y **Matplotlib 3.8.2** que ayudara a mostrar gráficamente los datos.

Como entorno de desarrollo se utiliza en primer plano el editor de código fuente **Visual Studio Code**, con el se realiza la codificación del algoritmo genético y para las ejecuciones, salidas y resultados se hace uso de un entorno con desarrollo interactivo como **Jupyter Notebook**, para poder hacer uso de esta herramienta se crea un entorno de trabajo en donde se instalan las librerías necesarias y adicionalmente en Visual Studio Code se realiza la instalación de una extensión denominada **Jupyter 2024.2.0**. Toda codificación y ejecución se realiza de manera local en una laptop.

Se utiliza un a MacBook Pro con chip M2 Pro con una memoria RAM LPDDR5 de

16 GB, como almacenamiento un SSD (Solid State Drive), disco de estado solido con capacidad de 1 TB, con un sistema operativo macOS Sonoma 14.4.

## 4.2. Representación del problema

En la gestión de proyectos, es común emplear distribuciones probabilísticas para modelar y prever el progreso esperado de una tarea a lo largo del tiempo, la más utilizada es la **distribución gaussiana** o **campana de Gauss**, que es simétrica y se centra en un valor medio, representando situaciones donde la mayor parte de los eventos se acumulan alrededor de la media con menos eventos en los extremos.

En este caso, se representa el progreso planificado de una tarea a lo largo de un período de 44 días utilizando una distribución gaussiana con un escenario donde el progreso real de la tarea refleja un atraso del 20 % durante los primeros 10 días de este modo es cómo se representa este problema y cómo se simula el impacto del atraso en el progreso de la tarea.

El primer paso consiste en representar el progreso planificado de la tarea utilizando una distribución gaussiana calculada con una función que se encarga de crear la distribución de probabilidad ajustada de tal forma que la suma de las probabilidades sea igual a 1, el 100 % de la tarea. Este ajuste se distribuye uniformemente entre todos los días, corrigiendo pequeñas desviaciones que puedan haber surgido durante el cálculo inicial, esto es importante para asegurar que el progreso planificado sea una representación normalizada, adecuada para compararse con otros valores o distribuciones.

La media se establece en la mitad del número de días de la tarea, lo que significa que el pico de la distribución, donde se espera que el progreso sea más rápido, se encuentra en el punto medio del período y un valor fijo de 8.8 para la desviación

estándar que determina la amplitud de la campana con lo cual refleja cuánto se espera que varíe el progreso a lo largo del período, con un mayor progreso concentrado en el centro.

En el siguiente paso, se introduce un atraso del 20% en el progreso real de la tarea durante los primeros 10 días, ajustando los valores de la distribución planificada para esos días específicos, para los días posteriores (del 11 al 44), el progreso aún no se ha registrado, esto representa un corte en la información en el seguimiento del proyecto en el cual los 10 días son el avance real de la tarea y los posteriores serán parte del reporte proforma que es lo que el algoritmo genético buscara como solución óptima.

### **4.3. Ejecuciones de prueba**

Para las ejecuciones de prueba se realiza una secuencia que ejecuta el algoritmo genético varias veces, cada vez con un conjunto diferente de parámetros seleccionados aleatoriamente en cada iteración del bucle se representa una ejecución independiente del algoritmo genético con un conjunto distinto de parámetro. Este proceso permite evaluar el rendimiento del algoritmo bajo diversas configuraciones, ayudando a identificar cuáles parámetros son más efectivos para resolver un problema específico.

Se generan valores aleatorios para los parámetros principales del algoritmo genético: el número de generaciones, el tamaño de la población, la tasa de cruce, la tasa de mutación y la tasa de elitismo.

1. Número de Generaciones: Se genera un número entero aleatorio entre 20 y 1,000. Este valor representa cuántas generaciones se ejecutarán en el algoritmo genético. Un número más alto permite que el algoritmo explore

más soluciones, aunque también aumenta el tiempo de cómputo.

2. **Tamaño de la Población:** Se genera un número entero aleatorio entre 16 y 1,000, que determina el tamaño de la población en cada generación del algoritmo genético. Un tamaño mayor de población puede mejorar la diversidad de las soluciones, pero también incrementa el costo computacional.
3. **Tasa de Cruce:** Se genera un número decimal aleatorio entre 0.6 y 1, redondeado a cuatro decimales. Este valor representa la probabilidad de que ocurra un cruce entre dos individuos seleccionados de la población. Una tasa alta promueve la recombinación de genes, lo cual es esencial para la exploración de nuevas soluciones.
4. **Tasa de Mutación :** Se genera un número decimal aleatorio entre 0.001 y 0.1, redondeado a cuatro decimales. Este valor representa la probabilidad de que ocurra una mutación en los genes de un individuo. La mutación introduce variación genética en la población, evitando que el algoritmo se estanque en óptimos locales.
5. **Tasa de Elitismo:** Se genera un número decimal aleatorio entre 0.001 y 0.1, redondeado a cuatro decimales. Este valor determina la proporción de los mejores individuos que se transfieren directamente a la siguiente generación sin modificación. El elitismo asegura que las mejores soluciones encontradas no se pierdan, mejorando la eficiencia del algoritmo.

Este enfoque permite identificar las configuraciones más efectivas para ir balanceando la exploración de nuevas soluciones con variabilidad en los parámetros otorgando una flexibilidad que ayuda a evitar que el algoritmo se estanque en óptimos locales, aumentando las probabilidades de encontrar una solución óptima global. Sin embargo, también se debe tener en cuenta que

Best Test	Fitness	Gen.	Pop.	CR.	MR.	ER.	Time (s)
217	0.999999992	979	973	0.8439	0.0168	0.0545	14.6566
063	0.999999991	968	962	0.9335	0.0309	0.0656	14.2351
186	0.999999990	740	867	0.9483	0.0370	0.0114	10.2239
24	0.999998793	598	235	0.9702	0.0538	0.0173	02.1189
182	0.999998748	938	362	0.8948	0.0624	0.0134	05.2734
52	0.999998741	593	118	0.6499	0.0538	0.0744	00.9776
181	0.999998715	479	628	0.8338	0.0791	0.0936	04.4603
250	0.999998660	537	125	0.7215	0.0061	0.0109	00.9816
185	0.999998658	564	566	0.9706	0.0753	0.0351	05.0266
226	0.997719703	501	085	0.7766	0.0671	0.0069	00.6437
241	0.995965040	176	304	0.7326	0.0732	0.0032	00.8231

Tabla 4.1: Resultados de las ejecuciones del algoritmo genético.

un aumento en la variabilidad y en los valores altos de parámetros puede incrementar significativamente el costo computacional, por lo que es importante encontrar un equilibrio adecuado entre exploración y eficiencia computacional.

Se realizan 251 ejecuciones del algoritmo genético, almacenando los resultados de cada iteración para posteriormente observar el desempeño del algoritmo genético con el objetivo de ir analizar máximos y mínimos de los parámetros aleatorios en cada ejecución, en la tabla 4.1, se muestran los valores obtenidos de los parámetros de forma ordenada de mayor a menor aptitud, mostrando los tres mejores, cinco en el rango medio y tres peores individuos.

En la gráfica 4.1 se aprecia el declive en la aptitud de los individuos a partir de la quinta ejecución como consecuencia de la disminución de los máximos de la población y las generaciones que, aunque beneficia en el tiempo de ejecución, castiga severamente en la aptitud.

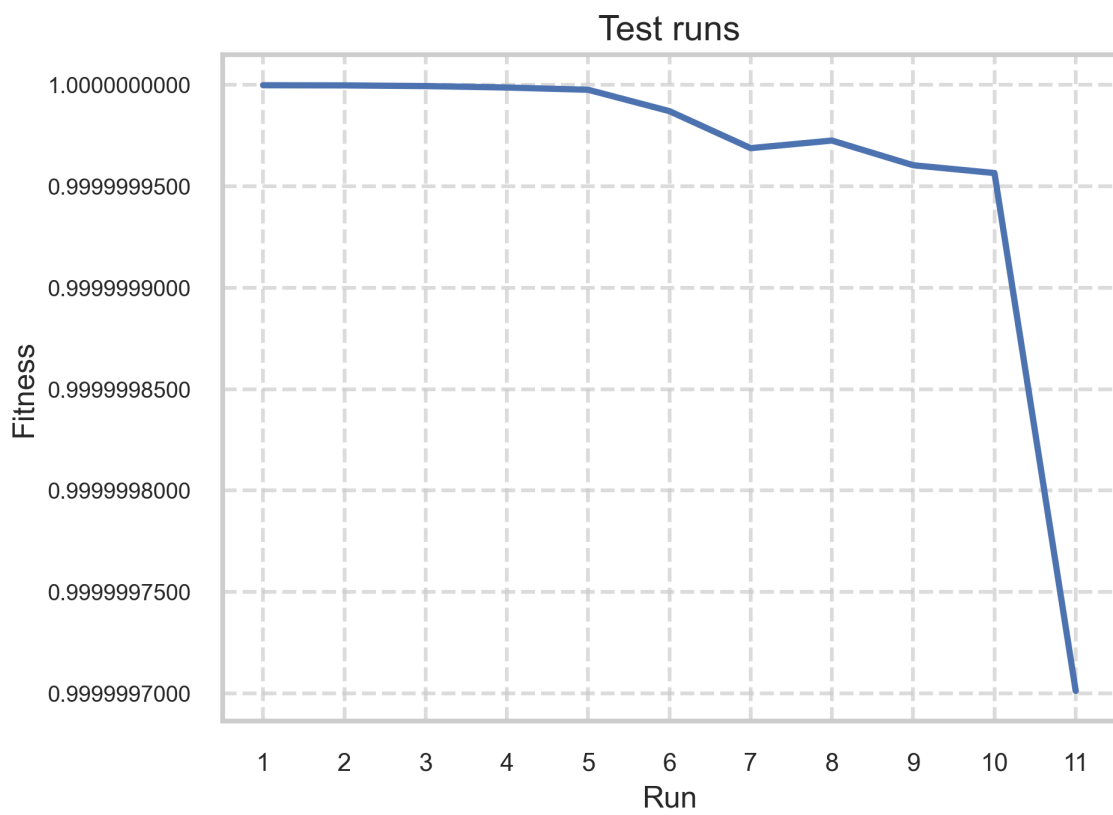


Figura 4.1: Gráfica de las mejores aptitudes en las ejecuciones (Elaboración propia)

## **4.4. Validación**

El uso de un grid search con validación cruzada para optimizar los parámetros de un algoritmo genético es una estrategia sólida para obtener resultados robustos, a continuación se explica tanto la metodología como las decisiones tomadas durante el proceso de optimización de los parámetros del algoritmo genético.

### **4.4.1. Configuración de la Simulación**

El primer paso se define la cantidad de proyectos de prueba generados para la simulación, en cada proyecto se puede representar un conjunto de avance en tareas o actividades dentro de un entorno de la ingeniería civil, donde el algoritmo genético será aplicado para optimizar alguna variable específica, en este caso la asignación de recursos, al generar múltiples proyectos de prueba se puede evaluar la robustez y la generalización del algoritmo en diferentes escenarios.

A cada proyecto de prueba se establece una duración en días, esto es relevante porque puede afectar la complejidad del problema a resolver. Por ejemplo, proyectos con una duración mayor pueden tener más avance en las actividades, lo que incrementa la dificultad de optimizar el proforma y finalmente se asigna un retraso con el objetivo de introducir una variabilidad o incertidumbre en los proyectos de prueba. Este valor representa un porcentaje de días de retraso que podría ocurrir en las tareas de los proyectos de construcción, lo que hace que la simulación sea más realista y relevante.

Para la generación de proyectos, se utiliza una distribución gaussiana para generar los recursos planeados a lo largo de los días del proyecto para

posteriormente aplicar un factor de reducción a los recursos planeados en los primeros diez días del proyecto, simulando así un escenario en el cual estos recursos experimentan un retraso del 10% debido a algún contratiempo imprevisto. Finalmente, los recursos planeados, ya ajustados por el retraso, se almacenan en una lista que representa el conjunto de proyectos generados para evaluar de manera efectiva cómo el algoritmo genético maneja y optimiza recursos bajo diferentes condiciones de operación y restricciones temporales.

#### **4.4.2. Rango de Valores para los Parámetros**

El número de generaciones es un parámetro crítico en los algoritmos genéticos, que determina cuántas veces se renovará la población de soluciones posibles a lo largo del proceso evolutivo. Valores más altos permiten una exploración más exhaustiva del espacio de soluciones, pero incrementan el tiempo de cómputo. En esta tesis, se han seleccionado valores específicos para explorar cómo diferentes números de generaciones impactan la eficacia del algoritmo.

El tamaño de la población determina cuántas soluciones candidatas se mantienen en cada generación del algoritmo genético. Un tamaño de población mayor generalmente conduce a una mejor cobertura del espacio de búsqueda, aunque a costa de un mayor tiempo de procesamiento. Los valores seleccionados permiten analizar el equilibrio entre la calidad de las soluciones obtenidas y el tiempo de cómputo necesario.

La probabilidad de cruce es la probabilidad con la que dos individuos (soluciones) dentro de la población intercambian información genética para crear nuevas soluciones (hijos). Este parámetro influye en la capacidad del algoritmo para combinar características deseables de diferentes soluciones. La exploración de

Parameter	Value 1	Value 2	Value 3
Generations	864	575	285
Population	820	668	516
Crossover rate	0.7636	0.67685	0.5901
Mutation rate	0.0146	0.02860	0.0426
Elitism rate	0.0984	0.07130	0.0442

Tabla 4.2: Valores de cada parámetro para la validación cruzada

diferentes probabilidades de cruce permite entender su efecto en la diversidad genética de las soluciones generadas.

La probabilidad de mutación indica la probabilidad de que una solución sufra cambios aleatorios en su estructura genética. Este parámetro es crucial para evitar la convergencia prematura a soluciones subóptimas, asegurando que el algoritmo siga explorando nuevas áreas del espacio de búsqueda. Variar este parámetro ayuda a determinar su impacto en la diversidad de las soluciones y en la capacidad del algoritmo para escapar de óptimos locales.

El elitismo en los algoritmos genéticos asegura que las mejores soluciones de una generación se mantengan intactas en la siguiente generación. La probabilidad de elitismo controla qué proporción de las mejores soluciones se preserva. Explorar diferentes valores de elitismo permite entender cómo este mecanismo impacta en la calidad final de las soluciones y en la velocidad de convergencia del algoritmo.

#### 4.4.3. Grid Search con Validación Cruzada

El grid search es una técnica exhaustiva de optimización que consiste en probar sistemáticamente todas las combinaciones posibles de un conjunto predefinido de parámetros para identificar la configuración óptima. En este

contexto, el grid search se utiliza para encontrar la mejor combinación de parámetros (popsize, numgen, crossprob, mutprob, y elitism) que maximice la eficacia del algoritmo genético en la simulación.

La validación cruzada es una técnica de evaluación utilizada para estimar la habilidad de un modelo de predecir datos nuevos. En la k-fold cross-validation, los datos se dividen en k partes, y el modelo se entrena utilizando k-1 partes, mientras que la parte restante se utiliza para la validación. Este proceso se repite k veces, cambiando la parte utilizada para la validación en cada iteración. Finalmente, se promedian los resultados de las k iteraciones. Esta técnica asegura una evaluación robusta del rendimiento del modelo y evita el sobreajuste.

#### **4.4.4. Selección de los Mejores Parámetros**

El criterio de evaluación en este caso se basa en una métrica denominada eficiencia de recursos, que es un indicador clave del rendimiento del algoritmo genético. Aunque se pueden emplear múltiples métricas, la elección de un criterio específico como el puntaje global permite simplificar la selección de los mejores parámetros.

Después de realizar el grid search, los mejores parámetros se determinan basándose en la puntuación más alta obtenida en la métrica seleccionada. Esta configuración óptima se considera la que proporciona el mejor equilibrio entre calidad de soluciones y eficiencia computacional dentro del contexto de la simulación.

#### 4.4.5. Visualización de Resultados

Los gráficos de barras permiten una visualización clara y comparativa de las métricas evaluadas (desviación estándar, cumplimiento de objetivos, y eficiencia de recursos). Cada barra representa una configuración diferente de parámetros, y las alturas de las barras muestran la magnitud promedio de las métricas asociadas. Esta visualización facilita la identificación de tendencias y patrones, y permite una evaluación más intuitiva de cómo los diferentes parámetros afectan el rendimiento del algoritmo.

El cálculo de los promedios de las métricas para cada configuración de parámetros permite cuantificar y comparar el rendimiento de las distintas configuraciones de manera concisa. Esto es útil para realizar análisis estadísticos adicionales y proporciona una base sólida para las conclusiones del estudio, por lo que en las gráficas de la figura 4.2 son resultado de la validación cruzada realizada para obtener los parámetros del algoritmo genético.

Los mejores parámetros encontrados mediante este proceso fueron los siguientes:

- **Tamaño de la Población:** 285
- **Probabilidad de Cruce:** 0.5901
- **Probabilidad de Mutación:** 0.0426
- **Número de Generaciones:** 516
- **Probabilidad de Elitismo:** 0.0442

Los resultados obtenidos a partir de esta configuración fueron evaluados en términos de tres métricas clave: desviación estándar promedio, cumplimiento de objetivos promedio y eficiencia de recursos promedio.

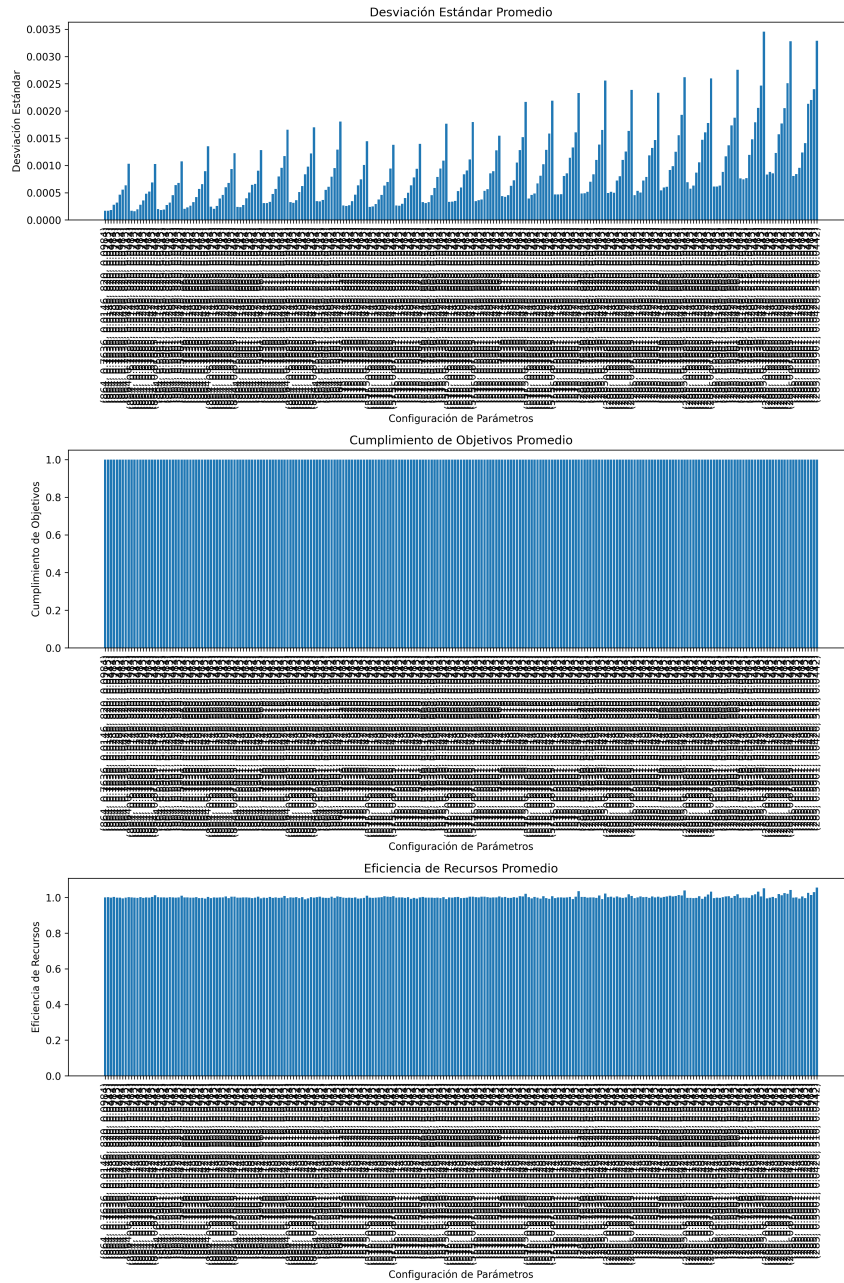


Figura 4.2: Gráficas con los resultados de la validación cruzada (Elaboración propia)

- **Desviación Estándar Promedio:** 0.00089658. Este valor sugiere una baja variabilidad entre las distintas iteraciones del proceso de validación cruzada, indicando que el rendimiento del algoritmo es consistente independientemente de la partición de los datos utilizada para el entrenamiento y la validación lo que implica que los resultados son estables y que el modelo tiene una alta capacidad de generalización.
- **Cumplimiento de Objetivos Promedio:** 1.0. El algoritmo logró cumplir con todos los objetivos planteados en cada una de las iteraciones del proceso de validación cruzada, lo cual refleja un excelente desempeño en términos de alcanzar las metas definidas en el problema de optimización.
- **Eficiencia de Recursos Promedio:** 1.0028. El algoritmo está utilizando los recursos de manera eficiente, logrando un rendimiento ligeramente superior al esperado bajo las condiciones establecidas con una eficiencia cercana a 1 sugiere un uso adecuado y balanceado de los recursos disponibles, lo cual es fundamental en la optimización de procesos complejos.

El uso de la técnica de validación cruzada k-fold permitió obtener una evaluación más precisa y confiable del rendimiento del algoritmo genético, asegurando que los resultados no estuvieran sesgados por la selección particular de los datos de entrenamiento y validación, estos valores obtenidos para las métricas evaluadas indican que el modelo no solo es efectivo, sino también robusto y eficiente, características fundamentales para su aplicación en problemas reales de ingeniería civil. En las figuras 4.3 y 4.4 se puede observar el resultado obtenido con el algoritmo genético, por una parte la representación de la actividad en la distribución normal y en seguida con la curva S.

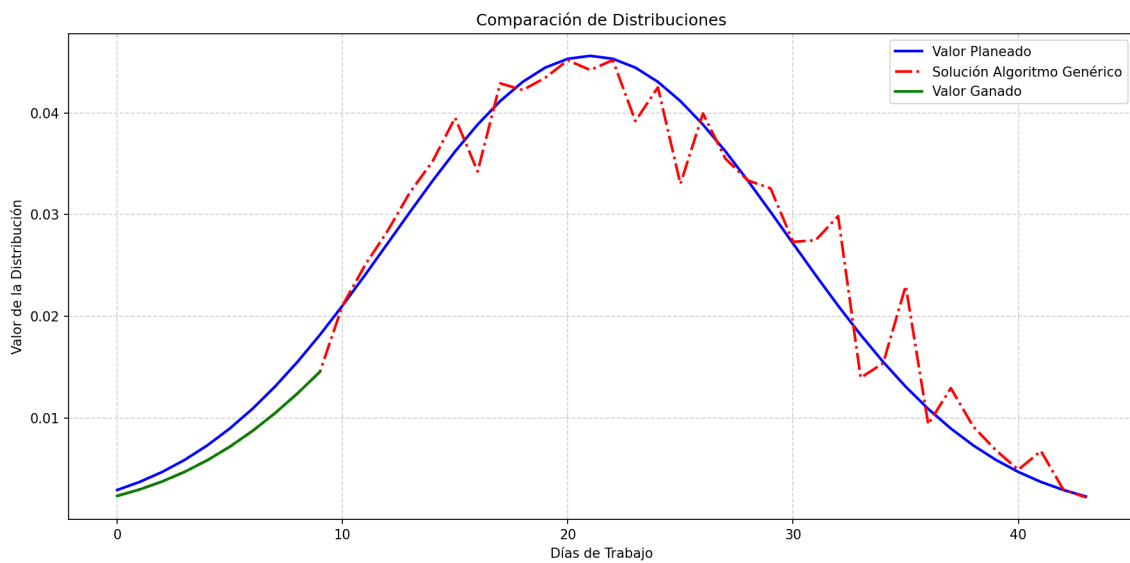


Figura 4.3: Comparativa de campana de Gauss (Elaboración propia)

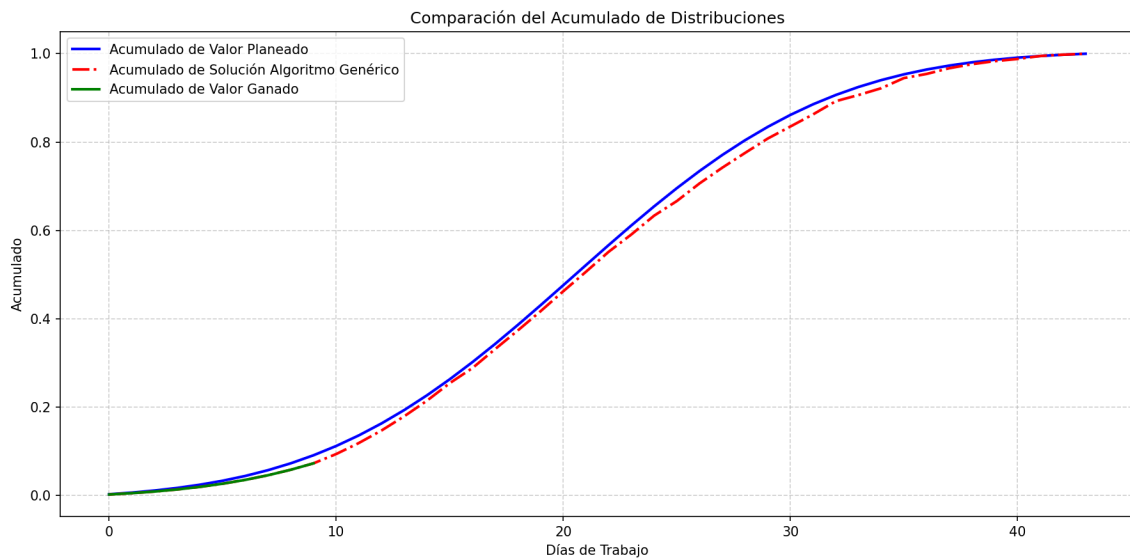


Figura 4.4: Comparativa acumulado (Elaboración propia)

## Capítulo 5

# CONCLUSIONES

### 5.1. Conclusiones

Tras un exhaustivo análisis de los resultados obtenidos durante la experimentación, se pueden establecer las siguientes conclusiones clave:

1. La elección de los parámetros obtenidos por ambas validaciones en el capítulo 4.4.5 se justifica plenamente, ya que se encuentra dentro de los rangos esperados de la distribución planeada y porque se realizó considerando el tipo de operadores del algoritmo genético, los parámetros, número de generaciones, el tamaño de la población, la probabilidad de cruce y la probabilidad de mutación fueron ajustados de manera meticulosa, asegurando una exploración exhaustiva del espacio de búsqueda y manteniendo una diversidad genética adecuada, lo que resultó en una solución robusta y confiable, tal como se documenta en la tabla 4.2.
2. La experimentación ha demostrado que, al configurar correctamente los parámetros del algoritmo genético, es posible obtener soluciones que se aproximan significativamente al óptimo, con este hallazgo se valida la eficacia del enfoque propuesto para ajustar el presupuesto en los proformas, destacando la capacidad del algoritmo para adaptarse a las necesidades específicas del proyecto.

3. Aunque la complejidad del ajuste de una distribución de recursos se hizo evidente durante las ejecuciones, el uso del algoritmo genético ha demostrado ser una técnica poderosa y adecuada para abordar este desafío por su capacidad para explorar eficazmente el espacio de soluciones, adaptarse a restricciones y cumplir con los objetivos específicos, con estas virtudes lo convierte en una herramienta valiosa para la optimización del presupuesto en proyectos de construcción.
4. Si bien este estudio se ha centrado en la gestión de proyectos de construcción, es importante subrayar que la distribución de recursos es un tema crucial que trasciende en este ámbito y se demuestra en los resultados y la metodología aplicada tienen un amplio potencial de aplicación en diversas disciplinas, como la economía, finanzas y logística, lo que subraya la relevancia y utilidad de los hallazgos obtenidos, más allá del contexto específico de la construcción.

Por lo tanto, los resultados de esta investigación no solo proporcionan una solución efectiva para el ajuste del presupuesto en proformas de proyectos de construcción, sino que también demuestran el potencial del enfoque propuesto para abordar problemas similares en diferentes contextos interdisciplinarios.

## **5.2. Trabajo Futuro**

Este trabajo sienta las bases para el estudio de problemas relacionados con la distribución de recursos en la gestión de proyectos, y abre la puerta a futuras investigaciones en varias direcciones:

1. Exploración de Diferentes Curvas de Distribución Una línea futura de investigación podría enfocarse en la experimentación con diversas curvas

de distribución y la evaluación de su impacto en los resultados, así como en la variación de los parámetros del algoritmo genético, como la población, el número de generaciones, y las probabilidades de cruce y mutación.

2. Incorporación del Juicio de Expertos Introducir el juicio cualitativo de expertos en la determinación de recursos en las tareas podría enriquecer el algoritmo genético, permitiendo una selección más precisa de las curvas de distribución ideales para tareas específicas, basadas en la experiencia y el conocimiento práctico.
3. Desarrollo de un Sistema Inteligente Considerando las dificultades prácticas para capturar el avance diario en proyectos de construcción, futuras investigaciones podrían enfocarse en el desarrollo de un sistema inteligente que integre técnicas avanzadas de inteligencia artificial. Este sistema podría automatizar la captura de datos en tiempo real, proporcionando al algoritmo genético información actualizada y precisa para optimizar el presupuesto en función del progreso real del proyecto.
4. Integración con BIM Una dirección relevante para el futuro es la integración del algoritmo genético dentro de la metodología Building Information Modeling (BIM). Esto permitiría no solo una visualización precisa del avance del proyecto, sino también la optimización dinámica de recursos en tiempo real, haciendo del algoritmo una herramienta aún más poderosa en la gestión digital de proyectos de construcción.
5. Optimización de Recursos Humanos y Maquinaria La incorporación de tipos de recursos específicos, como mano de obra y maquinaria, en el algoritmo genético podría mejorar la precisión de las estimaciones y optimizar la distribución de estos recursos críticos, impactando directamente en la eficiencia y el éxito de los proyectos de construcción.

### **5.3. Impacto y Aplicabilidad**

La metodología aplicada en este estudio, aunque específicamente aplicada a la optimización de presupuestos en proyectos de construcción, posee un alcance y aplicabilidad mucho más amplios, con la capacidad del algoritmo genético para adaptarse a diferentes restricciones y optimizar soluciones dentro de un amplio espacio de búsqueda lo convierte en una herramienta versátil con potencial para revolucionar la gestión de recursos en diversos campos.

Por ejemplo, en la economía y las finanzas, este enfoque podría ser utilizado para la distribución óptima de inversiones o la asignación de recursos en procesos logísticos complejos, por otro lado, en sectores como la salud o la educación, donde la gestión eficiente de recursos es crucial, las técnicas aquí desarrolladas podrían ser adaptadas para mejorar la toma de decisiones y la planificación estratégica con los hallazgos de esta investigación no solo aportan al campo de la gestión de proyectos en construcción, sino que también tienen el potencial de contribuir al avance de otras disciplinas, ampliando el impacto y la aplicabilidad de los métodos propuestos.

## **Apéndice A**

# **APÉNDICE**



# Certificado

OTORGADO A

**Lic. José Manuel Pérez Esquivel**

por su participación como Ponente en el:

**Simposio Internacional**

**Investigación interdisciplinaria y Modelos Educativos**

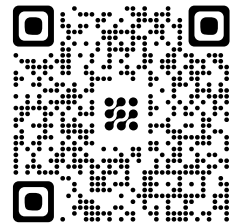
en calidad de Miembro de la CCH.



México, 1 de abril 2023

**Desiderio Ortegón**  
Presidente del Consejo Científico  
Revista Ciencia Latina

**Juan Carlos Manjarrez**  
Secretario Ejecutivo Internacional  
REDALAC



ARTÍCULO ACEPTADO

# Propuesta de un algoritmo genético para establecer proformas en proyectos de construcción

José Manuel Pérez Esquivel, Cristina Juárez Landín y Marco Alberto Mendoza Pérez

## Introducción

La Línea Base de una obra civil es un plan detallado que se utiliza como punto de referencia para medir el progreso del proyecto y como marco de información para la gestión de los cambios y la monitorización del avance trabajado. Se define como Línea Base de Alcance a la versión aprobada de un proyecto estableciendo su magnitud, que sólo se puede modificar a través de procedimientos formales de control de cambios y que se utiliza como base de comparación [2]. Esta línea base se establece en la fase de planificación del proyecto y es un elemento fundamental para la gestión de proyectos en la construcción. La elaboración de proformas en proyectos de construcción es un proceso crucial para determinar los costos y estimar los recursos necesarios para la ejecución de un proyecto, sin embargo, el establecimiento de nuevas proformas puede ser un proceso complejo y costoso en términos de tiempo y recursos humanos.

En la guía del Project Management Body of Knowledge (PMBOK) del Project Management Institute (PMI), se describen los procesos necesarios para la gestión de proyectos en la construcción, y se dedica una sección específica a la gestión del cronograma y la gestión de los costos. Estas secciones son esenciales para la elaboración de la línea base del proyecto porque permiten establecer los plazos, los costos y los recursos necesarios para la realización de la obra convirtiéndose en un marco de referencia para medir el desempeño del proyecto y comparar el progreso actual y real con el planificado. En caso de que se surjan desviaciones es posible identificar las causas y tomar las medidas necesarias para corregir el rumbo. Por lo tanto, la línea base es una herramienta esencial para la toma de decisiones informadas y la gestión eficaz del proyecto.

Otra técnica para el control de proyectos es la del Valor Ganado que determina el desempeño del proyecto durante su ejecución mediante el cual se controla la gestión integrada del alcance, el cronograma y los costos. Específicamente, esta técnica compara el desempeño de la línea base con respecto al desempeño real en términos del cronograma y costos [2], en otras palabras, la práctica del valor ganado mide la eficiencia y la eficacia

del proyecto, y como se están utilizando los recursos (tiempo y dinero) en comparación con lo que se planeó inicialmente.

Para llevar a cabo la gestión del valor ganado, se deben establecer tres medidas principales [10]: el Valor Planificado (*VP*), el Costo Real (*CR*) y el Valor Ganado (*VG*). El Valor Planificado representa el presupuesto asignado para el trabajo que se debe realizar durante un período de tiempo determinado. El *CR* representa el costo histórico de ese trabajo en ese período de tiempo, y el *VG* representa el valor del trabajo completado hasta el periodo que se esté revisando, de acuerdo con el plan inicial. Con estas tres medidas, se pueden calcular los indicadores claves de la gestión del valor ganado como la Variación de costo (*VC*) se calcula  $VC = VG - CR$ , el Índice de Desempeño de Costo (*IDC*) calculado con  $IDC = VG/CR$ , el Valor Ganado programado (*VGP*) se calcula  $VGP = VG - VP$  y el Índice de Desempeño del Cronograma (*IDC*) se calcula  $IDC = VG/VP$ .

En la Figura 1 se representa la gráfica Curva *S*, resultado de la técnica del *VG*, esta gráfica está en función del tiempo total del proyecto donde se puede apreciar la comparativa entre el *VP*, *VG* y *CR*.

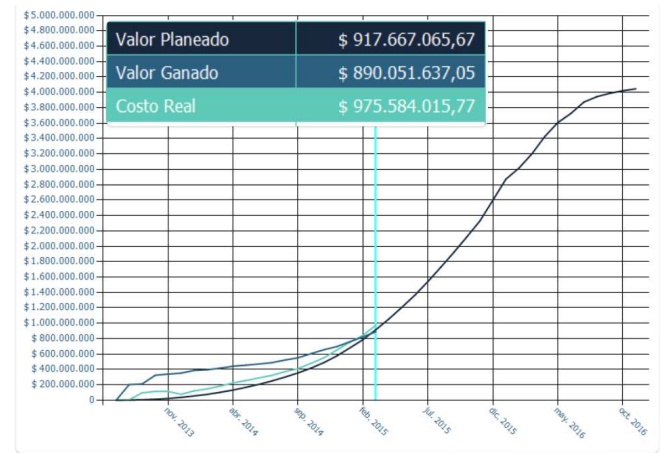


Figura 1. Curva S de un proyecto con fecha de corte en marzo 2015.

**En gestión de proyectos, la Línea Base de Alcance y Valor Ganado controlan los costos y el cronograma del proyecto asegurando el éxito.**

genéticos en la elaboración de proformas en proyectos de construcción ofrece una oportunidad para mejorar la eficiencia en la planificación y la toma de decisiones. Estos algoritmos pueden optimizar la asignación de recursos y el tiempo para cumplir con los objetivos del proyecto.

No existe una única solución a este problema por lo que la investigación se centrará en diseñar un algoritmo genético que pueda asistir al personal de control de proyectos en la toma de decisiones al momento de la elaboración de proformas y mejorar la planificación y estimación de costos en proyectos de construcción. Además, al reducir el tiempo y los recursos necesarios para generar proformas, se pueden mejorar la eficiencia y la rentabilidad de este proceso en proyectos de construcción.\*

## REFERENCIAS

- Urgilés, P., Claver, J., Miguel, S. (2018). Análisis de las técnicas del cronograma valorado y valor ganado para el seguimiento y control de proyectos de construcción complejos. *International Congress on Project Management and Engineering*. Madrid: Creative Commons.
- Institute, P. M. (2017). Construction Extension To The PM-BOK. Project Management Institute.
- Isikyildiz, S. y Akcay, C. (2020). Multi-objective optimization of time-cost-quality in construction projects using genetic algorithm. *Revista de la Construcción*, 19(3).
- Carmona, E.J. y Severino, F.G. (2021). Fundamentos de computación evolutiva. Marcombo.
- Gil-Velazquez, L., Sil-Acosta, M., Domiguez-Sanchez, Torres-Arreola, E. y LDP Medina Chavez, J. (2013). Diagnóstico y tratamiento de la diabetes mellitus tipo 2. *Rev Med Inst Mex Seguro Soc.*, 1-16.
- Banihashemi, S., Khalili, S., Sheikhhoshkar, M. y Fazeli, A. (2022). Machine learning-integrated 5D BIM informatics: building materials costs costs data classification and prototype development. *Innovative Infrastructure Solutions*, 7(215).
- Urgilés, P., Claver, J., Miguel Angel, S. (2018). Análisis de las técnicas del cronograma valorado y valor ganado para el seguimiento y control de proyectos de construcción complejos. *International Congress on Project Management and Engineering*. Madrid: Creative Commons.
- González, J., Solís, R. y Alcudia, C. (2010). Diagnóstico sobre la Planeación y Control de Proyectos en las PYMES de Construcción. *Revista de la Construcción*, 9(1), 17 - 25.
- Jiménez, F.C. y Castro Corrales, J.E. (2013). Implementación De Algoritmos Genéticos Para La Optimización De Un Presupuesto De Obra De Una Vivienda En Bogotá. Dc. Trabajo de Grado.
- Lledó, P., 2017. Director de proyectos: Cómo aprobar el examen PMP sin morir en el intento. 6ta ed, Estados Unidos.

## SOBRE LOS AUTORES



**José Manuel Pérez Esquivel** es Ingeniero en computación con 14 años de experiencia como Analista Programador. Sus intereses de investigación incluyen el desarrollo de sistemas inteligentes en la gestión de proyectos para la ingeniería civil y en el desarrollo de complementos en aplicaciones para la automatización de procesos en la metodología Building Information Modeling (BIM). Actualmente, está cursando una Maestría en Ciencias de la Computación en el Centro Universitario UAEM Valle de Chalco con una línea de investigación en Inteligencia Artificial (IA).



**Cristina Juárez Landín** es Profesora de tiempo completo en la Universidad Autónoma del Estado de México, adscrita al Centro Universitario Valle de Chalco y ha coordinado diversos programas educativos: Ingeniería en Computación, Informática Administrativa, Maestría y Doctorado en Ciencias de la Computación. Es miembro del Cuerpo Académico de Cómputo Inteligente y Sistemas y cuenta con perfil PRODEP, así también, pertenece al Sistema Nacional de Investigadores. Sus líneas de investigación son sistemas inteligentes con aplicación en los ámbitos social, educativo y médico, así como, procesamiento de señales e imágenes, temáticas en las cuales ha realizado publicaciones nacionales e internacionales.



**Marco Alberto Mendoza Pérez** es Profesor de Tiempo Completo, investigador, docente de Ingeniería en Computación, Maestría y Doctorado en Ciencias de la Computación, y responsable del Laboratorio de Electrónica en la Universidad Autónoma del Estado de México (UAEMex)/Centro Universitario Valle de Chalco. Miembro del Sistema Nacional de Investigadores (SNI) del Consejo Nacional de Humanidades Ciencias y Tecnologías (CONAHCYT). Sus áreas de interés son la robótica, la realidad aumentada y la inteligencia artificial en la educación.

## Capítulo 1

# Aprendizaje de las redes neuronales convolucionales mediante el análisis y modificación de la clasificación de imágenes para mejorar su precisión

José Manuel Pérez Esquivel, Marco Alberto Mendoza Pérez,  
José Roman Castro San Agustín, Cristina Juárez Landín

Universidad Autónoma del Estado de México  
Centro Universitario Valle de Chalco  
jpereze010@alumno.uaemex.mx, mamendozap@uaemex.mx,  
jcastros248@alumno.uaemex.mx, cjuarez1@uaemex.mx

**Resumen.** El presente trabajo explora cómo las Redes Neuronales Convolucionales (CNN) mejoran la precisión en la clasificación de imágenes, centrándose en la base de datos "fashion" de Zalando con 70, 000 ejemplos. Se introducen modificaciones en una CNN para clasificar prendas de vestir, utilizando Google Colab para el proceso. El artículo destaca la importancia de agregar capas ocultas para lograr una representación compleja y la extracción jerárquica de características. La elección de funciones de activación como "relu" y "sigmoid" también se considera esencial para mejorar la adaptación y la discriminación. Los resultados exhiben un progreso significativo en la precisión de la clasificación, con algunas categorías alcanzando niveles superiores al 97%. El enfoque en la elección de métricas, en particular la métrica exactitud, se destaca debido a su relevancia en la evaluación de la clasificación de imágenes. El artículo concluye que las adiciones de capas ocultas y las modificaciones en las funciones de activación impulsan la precisión de las CNN en la clasificación de imágenes. A pesar de los avances, persisten desafíos relacionados con la variabilidad en las condiciones de captura y la presencia de objetos similares. Las referencias a fuentes confiables como GitHub, Keras y TensorFlow validan el enfoque metodológico. En conjunto, el trabajo presenta una visión clara y concisa sobre cómo las modificaciones en redes neuronales pueden mejorar el desempeño en la clasificación de imágenes, resaltando la adaptabilidad y la elección de métricas adecuadas.

**Palabras Clave:** Clasificación de imagen, Inteligencia artificial, Métricas, Redes neuronales convolucionales.

en capas ocultas también mejoran el rendimiento al capturar patrones sutiles.

A pesar de los avances, hay desafíos. La variabilidad en las condiciones de captura y la presencia de objetos similares dificultan la clasificación precisa. Se requiere investigación continua para desarrollar técnicas de procesamiento de imágenes y aprendizaje más avanzadas, centradas en la adaptabilidad y robustez del modelo.

## Referencias

- Bardaji Serra, S. (2022). Learning contextual information via deep learning.
- Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.
- Caballero Hernández, J. A., Diaz Salazar, M., Moradillos Paz-Lago, M., & Pavoni Oliver, S. (2014). Implementación de la Función Sigmoidal Logarítmica en un FPGA. *Ingeniería Electrónica, Automática y Comunicaciones*, 35(2), 35-44.
- Chollet, F. (2017). *Deep Learning with Python*. Manning Publications.
- Fashion-MNIST (2024). <https://github.com/zalandoresearch/fashion-mnist>.
- Galvañ Sala, D. A. (2021). Comparativa de técnicas para la prevención del sobreajuste en redes neuronales.
- Google Colab (2024). <https://colab.research.google.com/>
- Keras. (s. f.). *Keras: The Python Deep Learning library*. Recuperado de <https://keras.io/>
- Kouretas, I., & Paliouras, V. (2019). Simplified hardware implementation of the softmax activation function. In 2019 8th international conference on modern circuits and systems technologies (MOCAST) (pp. 1-4). IEEE.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press.
- Nair, V., & Hinton, G. E. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines. *Proceedings of the 27th International Conference on International Conference on Machine Learning - Volume 27*, 807-814.
- Python (2024). <https://www.python.org>
- Revans, R. W. (1998). Sketches in Action Learning. *Performance Improvement Quarterly*, 11(1), 23..
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M.S., Berg, A.C., & Fei-Fei, L. (2014). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115, 211 - 252.
- Xiao, K. Rasul, and R. Vollgraf. Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms, (2017).
- Zalando Jobs. (s.f.). Software Engineering. (2024). <https://jobs.zalando.com/en/what-we-do/software-engineering>

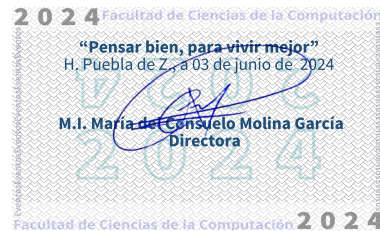


La Benemérita Universidad Autónoma de Puebla  
a través de la Facultad de Ciencias de la Computación  
otorga la presente

## CONSTANCIA

A: José Manuel Pérez Esquivel, Marco Alberto  
Mendoza Pérez, José Roman Castro San Agustín,  
Cristina Juárez Landín

Por presentar el artículo titulado "Aprendizaje de las redes neuronales  
convolucionales mediante el análisis y modificación de la clasificación  
de imágenes para mejorar su precisión", en el 8° Workshop de Objetos de  
Aprendizaje 2024.



## Optimización de la redistribución de recursos en proyectos con retrasos utilizando algoritmos genéticos

Ing. Comp. **José Manuel Pérez Esquivel**

Centro Universitario UAEM Valle de Chalco

ORCID: <https://orcid.org/0009-0009-2611-4763>

[jpereze010@alumno.uaemex.mx](mailto:jpereze010@alumno.uaemex.mx)

Dr. **Marco Alberto Mendoza Pérez**

Centro universitario UAEM Valle de Chalco

ORCID: <https://orcid.org/0000-0003-4911-4757>

[mamendozap@uaemex.mx](mailto:mamendozap@uaemex.mx)

Dra. **Cristina Juárez Landín**

Centro Universitario UAEM Valle de Chalco

ORCID: <https://orcid.org/0000-0002-0988-3060>

[cjuarezl@uaemex.mx](mailto:cjuarezl@uaemex.mx)

### RESUMEN

La gestión eficiente de recursos es crucial en la ejecución de proyectos, especialmente en la industria de la construcción, donde los retrasos son comunes y pueden impactar significativamente los costos y la calidad. Este estudio se enfoca en evaluar la eficacia de un algoritmo genético para la redistribución óptima de recursos en proyectos con retrasos en su inicio usando un diseño experimental, se manipularon variables independientes determinadas por los parámetros del algoritmo genético para observar sus efectos sobre la eficacia y eficiencia en la redistribución de recursos. La población de estudio consistió en actividades simuladas de planificación de recursos, representativas de entornos de la industria de la construcción con una selección muestra de 10 actividades con una duración de 44 días laborables. La distribución inicial se generó utilizando la distribución normal, con una media ( $\mu$ ) de 22 y una desviación estándar ( $\sigma$ ) de 8.8, posteriormente para simular retrasos, los valores de los primeros 10 días se ajustaron reflejando un retraso acumulado del 10%. Para evaluar la robustez y eficacia del algoritmo genético, se generaron múltiples instancias de actividades con diferentes distribuciones iniciales, simulando diversos escenarios para la implementación del algoritmo, se utilizó el lenguaje Python, usando técnicas como la selección por torneo, el cruzamiento de un punto y la mutación uniforme. La función de aptitud, basada en el Error Cuadrático Medio (ECM), permitió medir la calidad de la redistribución de recursos en términos de minimización del retraso acumulado y eficiencia en el uso de recursos.

**Palabras clave:** Algoritmos Genéticos, Gestión de Recursos, Optimización de Proyectos, Retrasos en el proyecto, Redistribución de Recursos.

Como trabajo a futuro está integrar datos reales de proyectos de diversas industrias para validar y mejorar la robustez del algoritmo genético, abriendo las posibilidades de explorar el uso de operadores genéticos más avanzados, como el cruzamiento multipunto, la mutación adaptativa y la selección elitista, para mejorar la exploración y explotación del espacio de soluciones, considerando desarrollar herramientas y plataformas de software que integren el algoritmo genético, facilitando su uso por parte de gestores de proyectos y profesionales en diferentes industrias para determinar cómo la colaboración y la comunicación entre los equipos de proyecto pueden influir en la eficacia del algoritmo genético.

## REFERENCIAS BIBLIOGRÁFICAS

- El-Shorbagy, M. A., & El-Refaey, A. M. (2020). *Hybridization of grasshopper optimization algorithm with genetic algorithm for solving system of non-linear equations*. *IEEE Access*, 8, 220944-220961.
- Galli, B. J. (2021). Statistical tools and their impact on project management—how they relate. *The Journal of Modern Project Management*, 9(2).
- Gridin, I. (2021). *Learning Genetic Algorithms with Python: Empower the performance of Machine Learning and AI models with the capabilities of a powerful search algorithm (English Edition)*. BPB publications.
- Holland, J.H. (1992). Genetic algorithms. *Scientific American*, 267(1), 66–73.
- James, G., Witten, D., Hastie, T. & Tibshirani, R. (2023). *An introduction to statistical learning with applications in R*, New York, Springer Science and Business Media, 2013, eISBN: 978-1-4614-7137-7.
- Mattos, A. D., & de Valderrama, F. G. F. (2019). *Métodos de planificación y control de obras*. Reverté.
- Nti, I. K., Nyarko-Boateng, O., & Aning, J. (2021). *Performance of machine learning algorithms with different K values in K-fold cross-validation*. *International Journal of Information Technology and Computer Science*, 13(6), 61-71.



La Universidad Autónoma del Estado de México por medio de la  
Red de Investigación Latinoamericana en Competitividad Organizacional y las  
Unidades Tecnológicas de Santander

Otorgan la presente

## CONSTANCIA

a:

Ingeniero José Manuel Pérez Esquivel

Doctor Marco Alberto Pérez Mendoza

Doctora Cristina Juárez Landín

Por su participación como PONENTE(S) con el trabajo de investigación  
**Optimización de la Redistribución de Recursos en Proyectos con Retrasos  
Utilizando Algoritmos Genéticos**

Impartida de manera Virtual Asíncrona.

En el marco del 14o Congreso Internacional en Competitividad  
Organizacional llevado a cabo del 30 de septiembre al 2 de octubre de  
2024 en la ciudad de Bucaramanga, Santander, Colombia.

Folio: 14074

**ATENTAMENTE**



Dr. en C. Ed. Julio Álvarez Botello

Presidente del Comité Científico del 14o Congreso Internacional en Competitividad Organizacional  
y Presidente de la Red de Investigación Latinoamericana en Competitividad Organizacional. (RILCO)