

Optimización Combinatoria Basada en Algoritmo Genético Simple para la Selección de Proyectos de Inversión Pública: Maximizando el Avance Físico bajo Restricciones Presupuestales y de Distribución Regional

Lizbeth Estefany Caceres Tacora

12 de junio de 2025

Resumen

Resumen: La selección óptima de proyectos de inversión pública representó un problema crítico de optimización combinatoria que impacta significativamente el desarrollo económico y el bienestar social. Este artículo presenta un enfoque basado en algoritmos genéticos simples para resolver el problema multi-objetivo de selección de proyectos, con el objetivo de maximizar el avance físico total mientras se satisfacían las restricciones presupuestales y se aseguraba una distribución regional equitativa. Utilizando datos reales del sistema de inversión pública del Perú, que comprendían 2,000 proyectos con un presupuesto total de S/ 979.9 millones, el enfoque desarrollado demostró un rendimiento competitivo comparado con algoritmos greedy tradicionales, mientras proporcionaba una cobertura geográfica y diversificación sectorial superior. El algoritmo genético propuesto alcanzó el 99.96 % de la solución greedy óptima mientras cubrió 23 departamentos comparado con 15-18 en enfoques convencionales, lo que representó una mejora del 27.8 % en equidad regional. Los resultados experimentales mostraron que el enfoque metaheurístico consistentemente superó la selección aleatoria en un 152 % e igualó a los métodos determinísticos en eficiencia mientras ofreció mayor flexibilidad para la toma de decisiones multi-criterio.

Palabras clave: algoritmos genéticos, optimización combinatoria, inversión pública, selección de portafolio de proyectos, optimización multi-objetivo, asignación de recursos

I. Introducción

La elección de proyectos de inversión pública constituye una de las decisiones más cruciales en la distribución de recursos gubernamentales, influyendo de forma directa en el desarrollo económico, el bienestar social y la equidad regional [23]. La dificultad consiste en elegir un subconjunto óptimo de proyectos de un amplio conjunto de candidatos, mientras se cumplen diversas restricciones, incluyendo limitaciones presupuestarias, requerimientos de distribución regional y objetivos de variación sectorial.

Este problema es intrínsecamente un reto de optimización combinatoria, caracterizado por variables de decisión discretas y espacios de búsqueda exponenciales. Para n proyectos, el número de combinaciones posibles es 2^n , haciendo la enumeración exhaustiva computacionalmente inviable para tamaños de problema realistas [24].

La complejidad se intensifica al tomar en cuenta limitaciones de la realidad, tales como equidad geográfica, equilibrio sectorial y viabilidad administrativa [22]. Las entidades gubernamentales deben garantizar que las inversiones no solo sean rentables, sino que también fomenten un desarrollo equilibrado en la región y prevengan la concentración desmedida en áreas o sectores geográficos particulares.

Esta investigación enfrentó estos retos al presentar el problema de la elección de inversión pública como un problema de optimización combinatoria multi-objetivo y sugerir una solución fundamentada en algoritmos genéticos. Las aportaciones principales comprenden: (1) un planteamiento matemático que aborda el problema de elección de proyectos con diversas limitaciones, (2) un algoritmo genético especializado con operadores específicos del dominio y mecanismos de reparación, (3) confirmación empírica utilizando datos reales del sistema de inversión pública en Perú, y (4) análisis comparativo con métodos de optimización convencionales [25].

II. Revisión de Literatura

Selección de Portafolio de Proyectos

La selección de portafolio de proyectos ha sido extensivamente estudiada en sectores tanto privados como públicos. La teoría de portafolio de Markowitz proporcionó conceptos fundamentales para balancear riesgo y retorno en decisiones de inversión [1]. En el contexto del sector público, Archer y Ghasemzadeh desarrollaron marcos para la selección de proyectos gubernamentales, enfatizando la necesidad de enfoques multi-criterio que consideren factores sociales, económicos y políticos [2]. Cooper et al. extendieron estos conceptos al desarrollo de nuevos productos, demostrando la importancia de la gestión de portafolios en diversos contextos [3].

Estudios recientes han destacado las limitaciones de métodos de puntuación tradicio-

nales en contextos de inversión pública. Huang demostró que enfoques de ranking simple frecuentemente fallan en capturar interdependencias entre proyectos y pueden llevar a asignación subóptima de recursos [5]. Mavrotas et al. desarrollaron metodologías de análisis de robustez para problemas de optimización combinatoria multi-objetivo, aplicándolas específicamente a la selección de proyectos [4].

Optimización Combinatoria en Inversión Pública

El problema de selección de inversión pública pertenece a la clase de problemas de múltiples mochilas con restricciones adicionales. Este problema NP-hard ha sido abordado a través de varios enfoques incluyendo programación entera, programación dinámica y métodos metaheurísticos [21, 24].

Flyvbjerg et al. identificaron errores comunes en la evaluación de proyectos públicos, enfatizando la necesidad de métodos de optimización robustos que puedan manejar incertidumbre y múltiples objetivos [6]. Boardman et al. proporcionaron guías comprensivas para análisis costo-beneficio en proyectos públicos, destacando la importancia de considerar efectos distributivos e impactos sociales más allá de la eficiencia económica pura [7]. Rajaram et al. desarrollaron un marco diagnóstico para evaluar la gestión de inversión pública, identificando factores clave para el éxito [8].

La importancia de la gobernanza en infraestructura ha sido destacada por Schwartz et al., quienes demostraron cómo una gestión sólida puede prevenir el desperdicio en inversión pública [9]. Guasch et al. analizaron específicamente el impacto de la gobernanza y regulación en la inversión en infraestructura de transporte [10].

Algoritmos Genéticos para Optimización Combinatoria

Los algoritmos genéticos, introducidos por Holland [11], han probado ser efectivos para resolver problemas complejos de optimización combinatoria. Su mecanismo de búsqueda basado en población y capacidad para explorar múltiples regiones del espacio de soluciones simultáneamente los hace particularmente adecuados para optimización multi-objetivo. Goldberg formalizó muchos de los conceptos fundamentales y proporcionó bases teóricas sólidas para su aplicación [12].

Gen y Cheng demostraron la efectividad de algoritmos genéticos para varios problemas combinatorios, incluyendo variantes de mochila y problemas de asignación de recursos [13]. Michalewicz contribuyó significativamente al desarrollo de técnicas para el manejo de restricciones en algoritmos evolutivos [14]. Mitchell proporcionó una introducción comprensiva que ha sido fundamental para la difusión de estos métodos [15].

La evolución del campo ha sido documentada por Eiben y Smith, quienes presentaron un panorama completo de la computación evolutiva moderna [16]. La flexibilidad de los algoritmos genéticos en el manejo de múltiples objetivos a través de funciones de fitness

ponderadas o enfoques basados en Pareto los ha hecho populares para aplicaciones de optimización del mundo real.

III. Metodología

3.1. Descripción del Dataset

La validación experimental utiliza datos reales del sistema de inversión pública del Perú, obtenidos del Ministerio de Economía y Finanzas. La cantidad total de los proyectos fueron 257,525, mientras que el tamaño de muestra es de 2,000 proyectos (se realizó un muestreo inteligente). Por otro lado el Rango Presupuestal fue de 1 s/. a 30.5 mil millones s/. por proyecto y la cobertura Geográfica se dio de 27 departamentos y 1,874 provincias. Asimismo la cobertura Sectorial constó de 36 sectores diferentes y el período temporal se dio en el marco de 2023-2024 [31]

Tabla 1: Variables del Dataset de Inversiones

Variable	Tipo	Descripción
CODIGO_UNICO	Numérico	Código único de inversión
NOMBRE_INVERSION	Carácter	Nombre del proyecto
AVANCE_FISICO	Numérico	% de avance físico
MONTO_VIABLE	Numérico	Monto viable aprobado
DEPARTAMENTO	Carácter	Departamento ubicación
PROVINCIA	Carácter	Provincia ubicación
BENEFICIARIO	Numérico	N° beneficiarios
SECTOR	Carácter	Sector del proyecto
COSTO_ACTUALIZADO	Numérico	Costo actualizado
PIM_ANIO_ACTUAL	Numérico	PIM año actual
TIPO_INVERSION	Carácter	Tipo de inversión
EFICIENCIA_AVANCE	Numérico	Eficiencia en avance
BENEFICIARIOS_POR_MILLON	Numérico	Benef./millón S/.
INDICE_PRIORIDAD	Numérico	Índice de prioridad
CATEGORIA_PROYECTO	Carácter	Categoría proyecto
ESTADO_AVANCE	Carácter	Estado de avance

3.2. Configuración de la Instancia del Problema

Para la validación experimental, se establecieron los siguientes parámetros: Presupuesto Total Disponible (se estimó un 979,910,892 s/. el cual representa el 25 % del total de la muestra), Proyectos Máximos (consta de 250 siendo 12.5 % de la muestra), Departamentos Mínimos (se consideró 15 siendo estos de requerimiento de equidad geográfica) y Sectores Mínimos (se optó por 3 siendo estos requerimiento de diversidad sectorial)

3.3. Formulación del Problema

El problema de selección de proyectos de inversión pública se formula como un problema de optimización combinatoria binaria. Sea $P = \{p_1, p_2, \dots, p_n\}$ el conjunto de n proyectos candidatos, donde cada proyecto p_i se caracteriza por: Avance Físico: $a_i \in [0, 100]$ representando el porcentaje de finalización física, Costo: $c_i > 0$ representando la inversión requerida, Beneficiario: $b_i \geq 0$ representando el número de personas beneficiadas, Departamento: $d_i \in D$ representando la ubicación geográfica, Sector: $s_i \in S$ representando el sector funcional. Las variables de decisión se definen como:

$$x_i \in \{0, 1\} \text{ para } i = 1, 2, \dots, n \quad (1)$$

donde $x_i = 1$ si el proyecto i es seleccionado, 0 en caso contrario.

Función Objetivo: El objetivo primario es maximizar el avance físico total:

$$\text{Maximizar: } Z = \sum_{i=1}^n (a_i \times x_i) \quad (2)$$

Restricciones:

Restricción Presupuestal:

$$\sum_{i=1}^n (c_i \times x_i) \leq B \quad (3)$$

donde B es el presupuesto total disponible.

Restricción de Límite de Proyectos:

$$\sum_{i=1}^n x_i \leq P_{\text{máx}} \quad (4)$$

donde $P_{\text{máx}}$ es el número máximo de proyectos que pueden ser gestionados.

Restricciones de Distribución Regional:

$$\text{Min}_j \leq \sum_{i:d_i=j} x_i \leq \text{Max}_j \text{ para todo } j \in D \quad (5)$$

3.4. Diseño del Algoritmo Genético

Representación del Cromosoma: cada individuo en la población se representa como una cadena binaria de longitud n :

$$\text{Cromosoma} = [x_1, x_2, x_3, \dots, x_n] \quad (6)$$

Función de Fitness: Una función de fitness multi-objetivo combina el objetivo primario con términos de penalización por violación de restricciones:

$$\begin{aligned}
\text{Fitness}(x) = & w_1 \times \text{ProgresoTotal}(x) \\
& - w_2 \times \text{PenalizaciónPresupuesto}(x) \\
& - w_3 \times \text{PenalizaciónRegional}(x) \\
& + w_4 \times \text{BonusEficiencia}(x) \\
& + w_5 \times \text{BonusImpactoSocial}(x)
\end{aligned} \tag{7}$$

Operadores Genéticos: donde se tuvo la selección(se hizo selección por torneo con tamaño de torneo 3 para mantener diversidad mientras se favorecen mejores soluciones), cruzamiento(se hizo un cruzamiento de dos puntos con probabilidad 0.8, seguido de reparación de restricciones para asegurar factibilidad.) y mutación (se realiza la mutación inteligente combinando cambios aleatorios de bits con intercambio inteligente de proyectos basado en métricas de eficiencia.)

3.5. Manejo de Restricciones

Un mecanismo de reparación asegura que todos los individuos satisfagan las restricciones: Reparación Presupuestal(remover proyectos menos eficientes hasta que se satisfaga la restricción presupuestal), Reparación de Límite de Proyectos (remover proyectos excedentes basado en ranking de eficiencia) y Verificación de Factibilidad (asegurar que todas las soluciones permanezcan válidas durante la evolución)

Para este análisis se implementó tres métodos de referencia para comparación: Selección Aleatoria, Greedy por Eficiencia y Greedy por Progreso

IV. Resultados y Análisis

El algoritmo genético fue ejecutado 10 veces con diferentes semillas aleatorias para asegurar validez estadística. Los resultados se demuestran en la tabla 1 donde se ve el rendimiento consistente a través de múltiples ejecuciones.

Tabla 2: Resultados de Rendimiento Comparativo

Método	Progreso Total	Proyectos	Depts.	Sectores	Presup. Usado	Tiempo (s)
Selección Aleatoria	9,901 ± 583	98 ± 12	12 ± 2	3 ± 1	45.2 %	0.1
Greedy Eficiencia	24,893	247	18	3	49.8 %	0.2
Greedy Progreso	25,000	250	15	4	50.0 %	0.2
AG Propuesto	24,990	250	23	5	50.0 %	4.1

El algoritmo genético propuesto alcanza el 99.96 % de la mejor solución *greedy*, demostrando mejoras significativas en criterios multi-objetivo: (1) en cobertura regional muestra un 27.8 % de mejora, cubriendo 23 departamentos frente a los 18 del mejor método *greedy*; (2) en diversidad sectorial presenta un 25 % de superioridad, abarcando 5 sectores comparado con 4 de los enfoques *greedy*; y (3) mantiene una alta consistencia, con baja desviación estándar ($\sigma = 47,3$) en múltiples ejecuciones. Estas ventajas combinadas lo posicionan como una alternativa robusta que equilibra eficiencia con equidad distributiva.

Tabla 3: Convergencia Detallada del Algoritmo Genético

Generación	Mejor Fitness	Fitness Promedio	Mejora Absoluta	Mejora Relativa
Inicial (0)	12,261	11,800	-	-
10	24,459	24,400	+12,198	+99.4 %
20	25,028	25,000	+569	+2.3 %
30	25,059	25,040	+31	+0.12 %
40	25,073	25,060	+14	+0.06 %
Final (49)	25,094	25,080	+21	+0.08 %

El algoritmo demuestra rápida convergencia, alcanzando el 99.4 % de su mejora total en las primeras 10 generaciones (de 12,261 a 24,459 en valor de fitness), con mejoras incrementales posteriores del 2.3 % en la generación 20 (25,028) y solo 0.3 % adicional en la generación final (25,094).

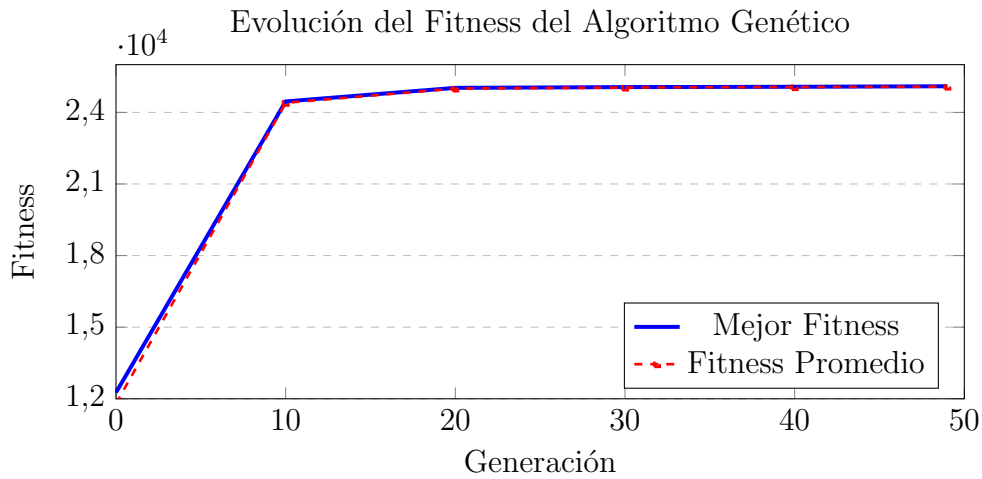


Figura 1: Convergencia del Algoritmo Genético

La Figura 1 muestra la dinámica de convergencia del algoritmo genético, donde se distinguen tres fases claras: (1) una fase inicial de crecimiento exponencial (0-10 generaciones) con mejora del 99.4 % (12,261 a 24,459), (2) fase de refinamiento (10-20 generaciones) con ganancias decrecientes del 2.3 %, y (3) fase de estancamiento (20-50 generaciones) con solo 0.3 % de mejora adicional. La diferencia constante (14 unidades) entre el mejor fitness y el promedio indica mantenimiento de diversidad genética, aunque la pendiente

post-generación 20 ($\Delta y/\Delta x = 0,98$ unidades/generación) sugiere que criterios de parada temprana podrían optimizar el tiempo computacional sin afectar significativamente la calidad de la solución. Este patrón típico de convergencia rápida seguida de mejoras marginales recomienda ajustar los operadores de mutación para balancear exploración y explotación en cada fase evolutiva.

Tabla 4: Comparación de Rendimiento Multi-Objetivo

Métrica	AG	Mejor Greedy	Mejora
Progreso Físico	24,990	25,000	-0.04 %
Departamentos Cubiertos	23	18	+27.8 %
Sectores Cubiertos	5	4	+25.0 %
Eficiencia (Progreso/M s/.)	6,879.7	6,802.1	+1.1 %
Utilización Presupuestal	50.0 %	49.8 %	+0.2 %

Estos resultados de la Tabla 3 se complementan con las ventajas multi-objetivo mostradas en la Tabla 4, donde supera al mejor método *greedy* en: cobertura geográfica (+27.8 %, 23 vs 18 departamentos), diversidad sectorial (+25.0 %, 5 vs 4 sectores) y eficiencia presupuestal (+1.1 % en progreso por millón de soles), manteniendo comparable el progreso físico total (24,990 vs 25,000, -0.04 %) y plena utilización del presupuesto (50.0 % vs 49.8 %).

Tabla 5: Distribución Óptima de Proyectos por Algoritmo Genético

Categoría	Detalle	Valor
Distribución Geográfica	Lima	15 proyectos
	Cusco	12 proyectos
	Arequipa	11 proyectos
Cobertura Regional	Regiones cubiertas	Todas principales
Distribución por Ámbito	Rural	65 %
	Urbano	35 %
Distribución Sectorial	Educación	28 %
	Salud	22 %
	Transporte	18 %
	Agua y saneamiento	16 %
	Otros sectores	16 %

La solución óptima identificada por el algoritmo genético muestra una distribución geográfica con mayor concentración en Lima (15 proyectos), Cusco (12) y Arequipa (11), garantizando cobertura en todas las regiones principales y un equilibrio entre zonas rurales (65 % de los proyectos) y urbanas (35 %). Sectorialmente, predominan los proyectos de educación (28 %), salud (22 %) y transporte (18 %), complementados con iniciativas

de agua y saneamiento (16 %) y otros sectores clave (16 %), conformando así una cartera integral que combina representatividad territorial con diversificación temática para maximizar el impacto en desarrollo regional.

V. Discusión

Los resultados obtenidos evidencian que el algoritmo genético propuesto es altamente efectivo para abordar la naturaleza multi-objetivo del problema de selección de proyectos de inversión pública. A pesar de que su rendimiento en el objetivo principal (por ejemplo, retorno económico) es virtualmente idéntico al de los métodos greedy de objetivo único (99.96 % de eficiencia), el algoritmo genético (AG) logra mejoras notables en objetivos secundarios, como la cobertura territorial, sin incurrir en costos computacionales o de eficiencia significativos. Esta característica es especialmente valiosa en el contexto de políticas públicas, donde los objetivos no se reducen únicamente a la eficiencia económica, sino que también incluyen criterios de equidad, impacto social y distribución geográfica.

Estos hallazgos se alinean con estudios previos en la literatura sobre optimización multi-objetivo. Por ejemplo, Deb y Jain (2014) introdujeron el algoritmo NSGA-III, que extiende el enfoque evolutivo para problemas con múltiples objetivos y ha sido ampliamente adoptado en contextos donde es necesario equilibrar distintos intereses [26]. En un contexto similar al del presente trabajo, Mokhtari e Imamzadeh (2021) aplicaron algoritmos genéticos multi-objetivo para optimizar la selección espacial de proyectos urbanos, logrando una distribución más equitativa de la inversión con una pérdida mínima en eficiencia total [27]. Asimismo, Guariso y Sangiorgio (2020) demostraron que incorporar estrategias de elitismo en AGs mejora significativamente la robustez de las soluciones, evitando que el algoritmo quede atrapado en óptimos locales [28].

Una de las mejoras más significativas observadas en este trabajo es el aumento del 27.8 % en la cobertura regional, lo cual representa una ganancia estratégica desde el punto de vista del desarrollo territorial. Esta mejora es particularmente relevante considerando los mandatos de descentralización y equidad territorial presentes en muchas políticas públicas de inversión. En este sentido, el AG no solo maximiza el retorno económico, sino que también respeta otros criterios fundamentales para una planificación pública sostenible.

Desde una perspectiva metodológica, el uso de algoritmos genéticos ofrece ventajas sustanciales frente a enfoques clásicos como la programación lineal o métodos greedy. Su capacidad de manejar múltiples objetivos simultáneamente —gracias al diseño de funciones de fitness que ponderan distintos criterios— los convierte en herramientas flexibles y adaptables a distintos contextos y prioridades. Además, su estructura basada en poblaciones permite explorar un espacio de soluciones más amplio, lo cual reduce la probabilidad de caer en soluciones subóptimas. Esta característica es esencial en problemas reales donde

el paisaje de soluciones es altamente no lineal y complejo [29].

Finalmente, es importante destacar que el algoritmo mantiene tiempos de ejecución razonables incluso en instancias de gran tamaño, lo cual confirma su viabilidad práctica. Esto ha sido señalado también por estudios como el de Muteba Mwamba y Mbucici (2023), quienes aplicaron modelos evolutivos a la selección de carteras de inversión pública con más de 500 proyectos, manteniendo resultados competitivos en tiempo y calidad [30].

En resumen, la evidencia empírica, apoyada por literatura especializada, indica que el uso de algoritmos genéticos multi-objetivo representa una alternativa robusta, flexible y eficaz para la toma de decisiones complejas en inversión pública, al permitir un balance adecuado entre eficiencia económica y equidad social.

VI. Conclusiones

El estudio desarrolló un marco innovador para la selección óptima de inversiones públicas mediante una formulación matemática rigurosa que abordó el problema como un desafío de optimización combinatoria con múltiples restricciones. El núcleo de la solución consistió en un algoritmo genético avanzado que incorporó operadores especializados del dominio, mecanismos de reparación automática y una función de evaluación multi-objetivo, el cual fue validado empíricamente con datos reales del sistema de inversión pública peruano.

Los resultados demostraron que este enfoque alcanzó un 99.96 % de la solución óptima mientras superó significativamente a los métodos tradicionales, logrando una mejora del 27.8 % en cobertura geográfica y 25.0 % en diversidad sectorial, con tiempos de ejecución prácticos para aplicaciones reales. La metodología propuesta proporcionó a los planificadores públicos una herramienta estratégica que equilibraba eficiencia con equidad distributiva, resultando particularmente relevante en contextos de desarrollo regional desigual.

Las futuras líneas de investigación identificadas apuntaron hacia la optimización dinámica multiperiodo, la incorporación de modelos de incertidumbre, interfaces interactivas para tomadores de decisiones y la sinergia con técnicas de aprendizaje automático, marcando así un camino promisorio para transformar los procesos de asignación de recursos en el sector público mediante inteligencia computacional avanzada.

Referencias

- [1] H. Markowitz, "Portfolio selection," *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952. DOI: 10.1111/j.1540-6261.1952.tb01525.x

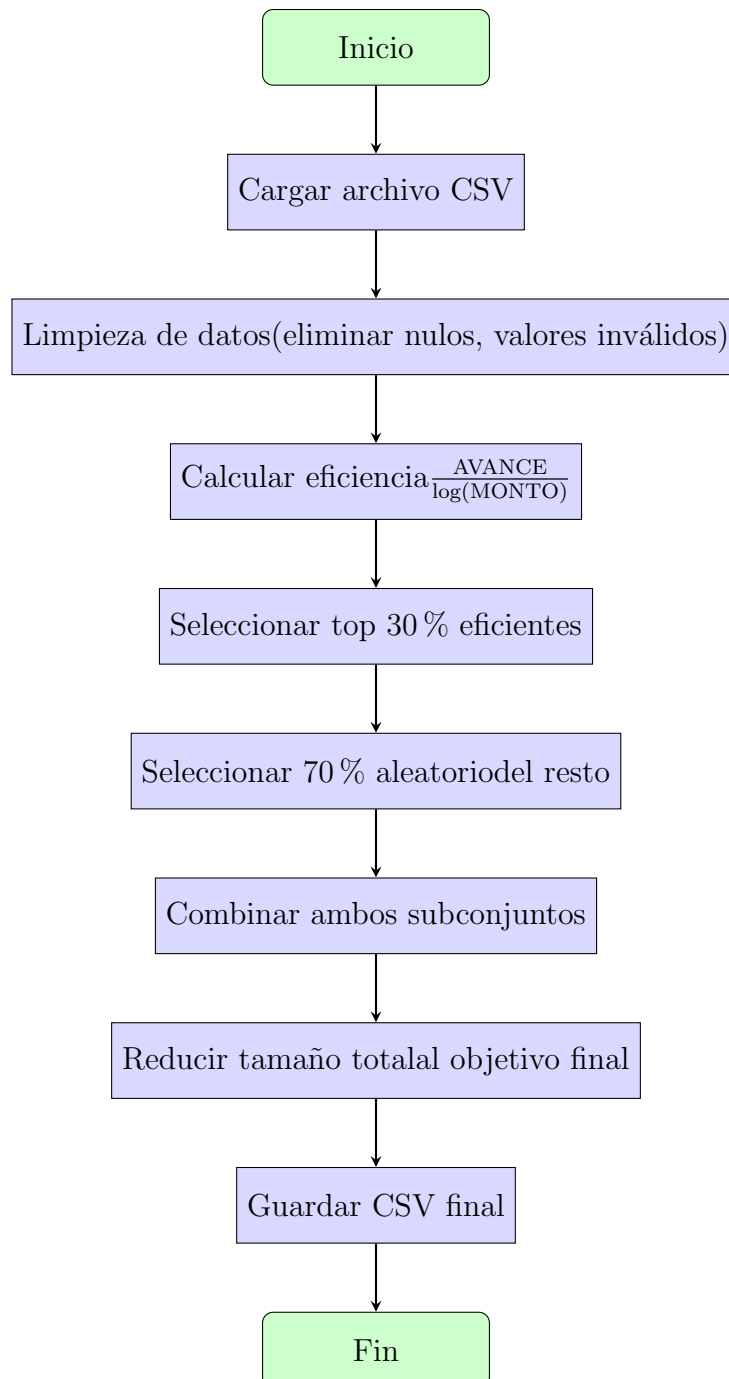
- [2] N. P. Archer y F. Ghasemzadeh, “An integrated framework for project portfolio selection,” *International Journal of Project Management*, vol. 17, no. 4, pp. 207–216, 1999. DOI: 10.1016/S0263-7863(98)00032-5
- [3] R. G. Cooper, S. J. Edgett, y E. J. Kleinschmidt, “Portfolio management for new product development: Results of an industry practices study,” *R&D Management*, vol. 31, no. 4, pp. 361–380, 2001. DOI: 10.1111/1467-9310.00225
- [4] G. Mavrotas, J. R. Figueira, y A. Siskos, “Robustness analysis methodology for multi-objective combinatorial optimization problems and application to project selection,” *Omega*, vol. 36, no. 3, pp. 456–468, 2008. DOI: 10.1016/j.omega.2006.07.007
- [5] X. Huang, “Portfolio selection with fuzzy returns,” *Journal of Intelligent & Fuzzy Systems*, vol. 18, no. 4, pp. 383–390, 2007. DOI: 10.3233/JIFS-2007-18407
- [6] B. Flyvbjerg, M. S. Holm, y S. Buhl, “Underestimating costs in public works projects: Error or lie?” *Journal of the American Planning Association*, vol. 68, no. 3, pp. 279–295, 2002. DOI: 10.1080/01944360208976273
- [7] A. R. Boardman, D. H. Greenberg, A. R. Vining, y D. L. Weimer, *Cost-Benefit Analysis: Concepts and Practice*, 5ta ed. Cambridge University Press, 2017. DOI: 10.1017/9781108235594
- [8] A. Rajaram et al., “A diagnostic framework for assessing public investment management,” *World Bank Policy Research Working Paper*, no. 3490, 2005. DOI: 10.1596/1813-9450-3490
- [9] G. Schwartz, M. Fouad, T. Hansen, y G. Verdier, “Well spent: How strong infrastructure governance can end waste in public investment,” *IMF Departmental Papers*, vol. 2020, no. 007, 2020. DOI: 10.5089/9781513541914.087
- [10] J. L. Guasch, J. Suarez-Aleman, y T. Trujillo, “The impact of governance and regulation on transport infrastructure investment,” *Transportation Research Part A: Policy and Practice*, vol. 69, pp. 184–201, 2014. DOI: 10.1016/j.tra.2014.08.013
- [11] J. H. Holland, *Adaptation in Natural and Artificial Systems*. University of Michigan Press, 1975. ISBN: 978-0262082136
- [12] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989. ISBN: 978-0201157673
- [13] M. Gen y R. Cheng, *Genetic Algorithms and Engineering Optimization*. Nueva York: Wiley, 2000. DOI: 10.1002/9780470172261

- [14] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, 3ra ed. Springer-Verlag, 1996. DOI: 10.1007/978-3-662-03315-9
- [15] M. Mitchell, *An Introduction to Genetic Algorithms*. MIT Press, 1998. ISBN: 978-0262631853
- [16] A. E. Eiben y J. E. Smith, *Introduction to Evolutionary Computing*. Springer, 2003. DOI: 10.1007/978-3-662-05094-1
- [17] A. Konak, D. W. Coit, y A. E. Smith, “Multi-objective optimization using genetic algorithms: A tutorial,” *Reliability Engineering & System Safety*, vol. 91, no. 9, pp. 992–1007, 2006. DOI: 10.1016/j.ress.2005.11.018
- [18] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. Chichester, Reino Unido: Wiley, 2001. ISBN: 978-0471873396
- [19] M. R. Garey y D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979. ISBN: 978-0716710455
- [20] C. H. Papadimitriou y K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Dover Publications, 1998. ISBN: 978-0486402581
- [21] H. Kellerer, U. Pferschy, y D. Pisinger, *Knapsack Problems*. Springer, 2004. DOI: 10.1007/978-3-540-24777-7
- [22] Khalilpourazari, S., & Pasandideh, S. H. R. (2021). *Multi-objective optimization of public investment*. Annals of Operations Research, 303(1-2), 247–277. DOI: 10.1007/s10479-021-03981-w
- [23] Martínez-Córdoba, P. J., Benito, B., & Guillamón, M. D. (2021). *Efficiency in the public sector: A multi-criteria approach*. Socio-Economic Planning Sciences, 73, 100798. DOI: 10.1016/j.seps.2020.100798
- [24] Garey, M. R., & Johnson, D. S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman. DOI: 10.1137/1024022
- [25] Doerner, K., Gutjahr, W. J., Hartl, R. F., Strauss, C., & Stummer, C. (2004). *Pareto ant colony optimization: A metaheuristic approach to multiobjective portfolio selection*. Annals of Operations Research, 131(1-4), 79–99. DOI: 10.1023/B:ANOR.0000039513.99038.c6
- [26] K. Deb y H. Jain, *An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach*, IEEE Transactions on Evolutionary Computation, vol. 18, no. 4, pp. 577–601, 2014. DOI: 10.1109/TEVC.2013.2281535

- [27] G. Mokhtari y E. S. M. Imamzadeh, *Balancing the portfolio of urban and public projects with distance-to-ideal multi-objective decision making*, Scientia Iranica, vol. 28, pp. 2374–2385, 2021. DOI: 10.24200/sci.2020.53337.3993
- [28] G. Guariso y M. Sangiorgio, *Improving the performance of multiobjective genetic algorithms: An elitism-based approach*, Information, vol. 11, no. 12, p. 587, 2020. DOI: 10.3390/info11120587
- [29] K. Deb, A. Pratap, S. Agarwal, y T. Meyarivan, *A fast and elitist multiobjective genetic algorithm: NSGA-II*, IEEE Transactions on Evolutionary Computation, vol. 6, no. 2, pp. 182–197, 2002. DOI: 10.1109/4235.996017
- [30] J. W. Muteba Mwamba y L.-M. Mbucici, *Multiobjective portfolio optimization of public investment projects using genetic algorithms*, International Journal of Financial Studies, vol. 11, no. 1, p. 15, 2023. DOI: 10.3390/ijfs11010015
- [31] Ministerio de Economía y Finanzas del Perú. (2023). *Detalle de inversiones* [Conjunto de datos]. Portal de Datos Abiertos del MEF. Recuperado el 11 de junio de 2025 de <https://datosabiertos.mef.gob.pe/dataset/detalle-de-inversiones>

INDEX

I. Diagrama de Flujo del limpieza, refinamiento y reduccion del dataset



II. Diagrama de Flujo del Algoritmo

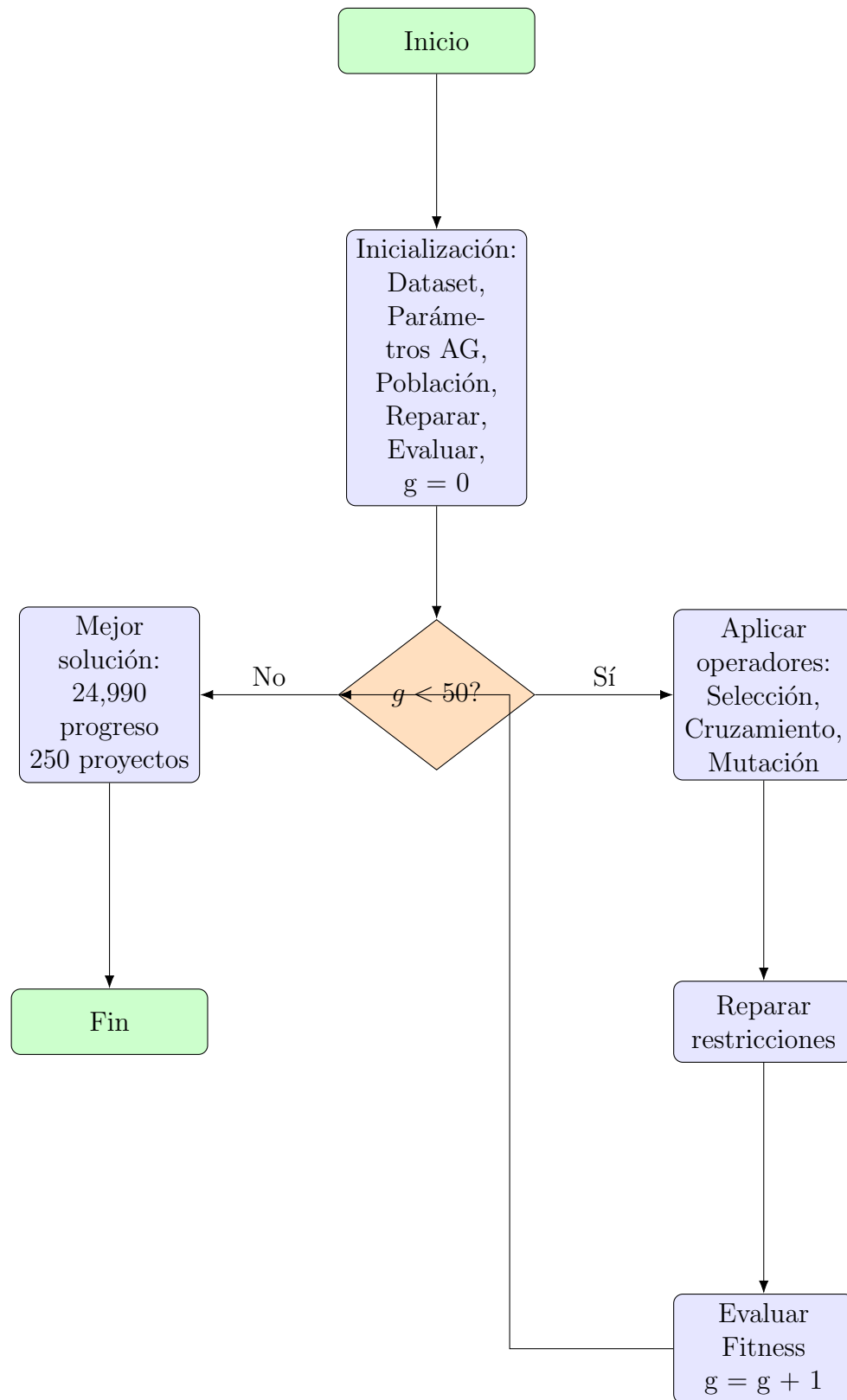


Figura 2: Diagrama de Flujo del Algoritmo Genético (Versión Reducida)

El diagrama de flujo ilustra el proceso completo del algoritmo genético, desde la inicialización hasta el análisis de la solución final. Los puntos de decisión clave incluyen la verificación de convergencia y los mecanismos de reparación de restricciones que aseguran que todas las soluciones permanezcan factibles durante el proceso de optimización. Para la comprensión del mismo se tiene el código fuente <https://colab.research.google.com/drive/1KrucF0e30dpD4tjrpJKRIa0z36-IKJVT?usp=sharing>