



Algoritmos Genéticos

ITBA - SIA - 2021



Introducción

- El trabajo práctico fue implementado usando Python. Para realizar los gráficos se utilizó la librería Matplotlib sugerida por la cátedra.
- Toda la configuración se realiza desde un archivo yaml, por defecto config.yaml.



Consideraciones

Cada individuo posee una clase, la cual determinará su función de fitness en base a una proporción de su ataque y su defensa.

Estos a su vez se calculan en función a la altura y el itemset que posea el personaje, los genotipos de los individuos.

Los distintos alelos de los genes son los distintos ítems y alturas que los individuos recibieron.



Diversidad

- Que tan variados son los genes? → Como mido que tan variados son?
- Items + Altura → Atributos + Altura
- Si tomo la suma de todos los items no pierdo informacion de que los ítems son variados.
- Suma de atributos + Altura → Desviación estándar / Media
- Desviación estándar → Que tan variados (poco acumulados) son los genes
- Media → Corregir el valor del desvío estándar para independizar de la escala.



Gráficos

Los siguientes gráficos analizaron distintos métodos de selección, mutación y cruce:

- Generaciones de 1000 individuos
- Clase Rogue
- 500 padres seleccionados
- Selección de padres 50% elite 50% universal
- Condición de corte 200 generaciones.
- Método de cruce y mutación uniforme
- Probabilidad de mutación = 0.3
- Método de recombinación: Fill-all



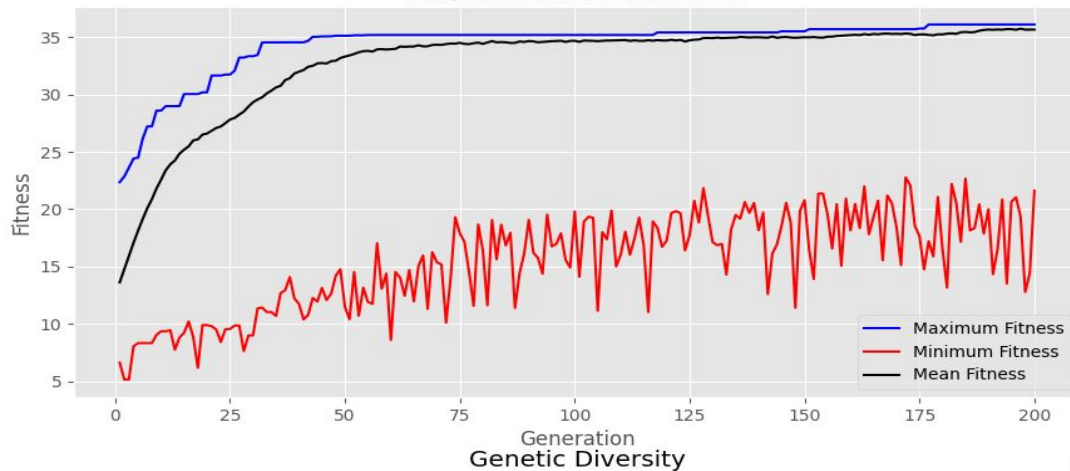
Aclaraciones adicionales

Se seleccionaron de a 500 individuos dado que se consideró que era una cantidad suficiente para no ejercer una gran presión de selección, que pueda ser más significativa que la dada por el método de selección.

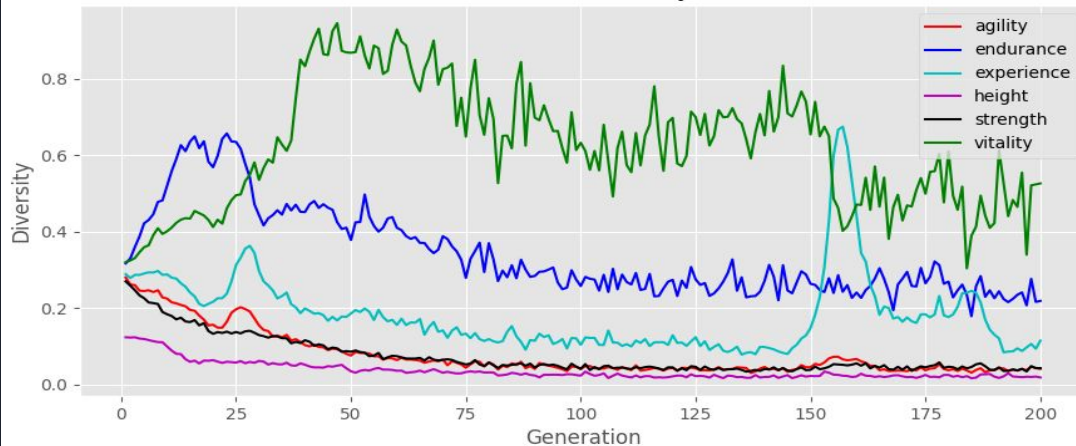
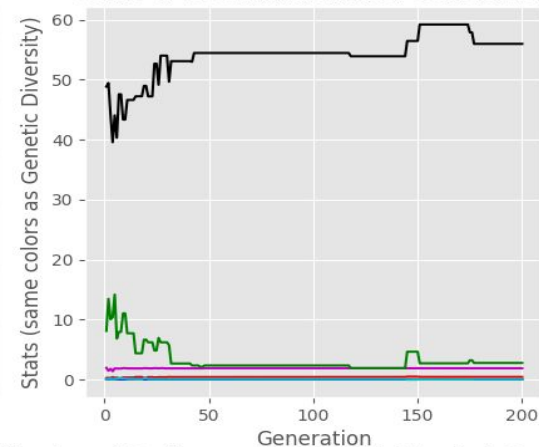
A su vez, si bien la mutación será un factor que afectará a la convergencia prematura y este al depender de la probabilidad será particular en cada caso, al tener todos la misma probabilidad se asume que son comparables. Se usó el método de mutación uniforme dado que es el único que no ejerce una presión adicional en ninguno de los genes y no resulta tan aleatorio como la mutación completa.

Para seleccionar el método de cruce se tomó una vez más el uniforme dado que el hecho de no mantener relación posicional entre alelos favorece a la diversidad.

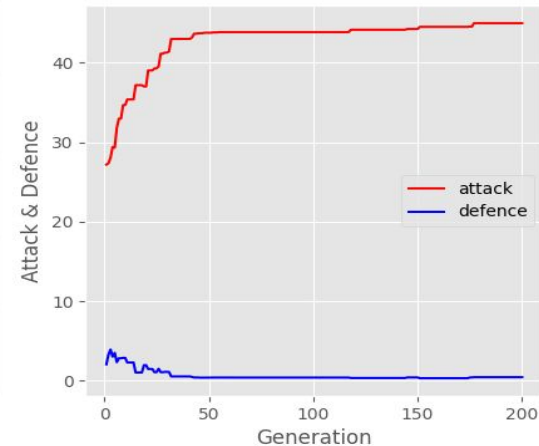
Min, Max and Mean Fitness



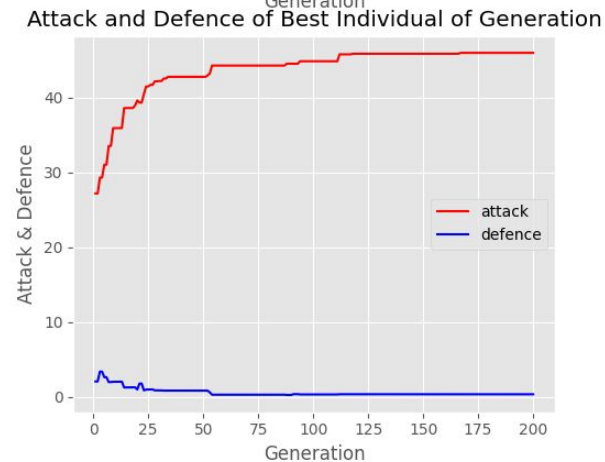
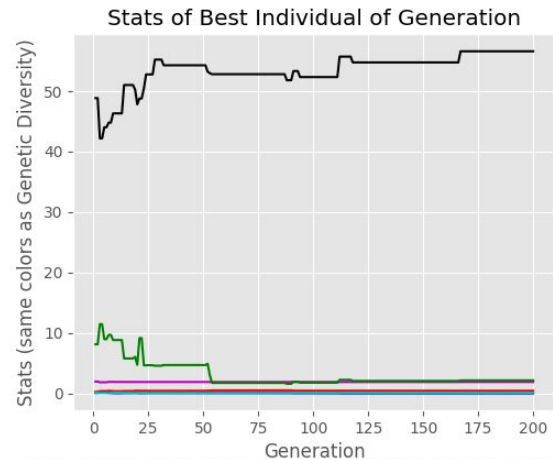
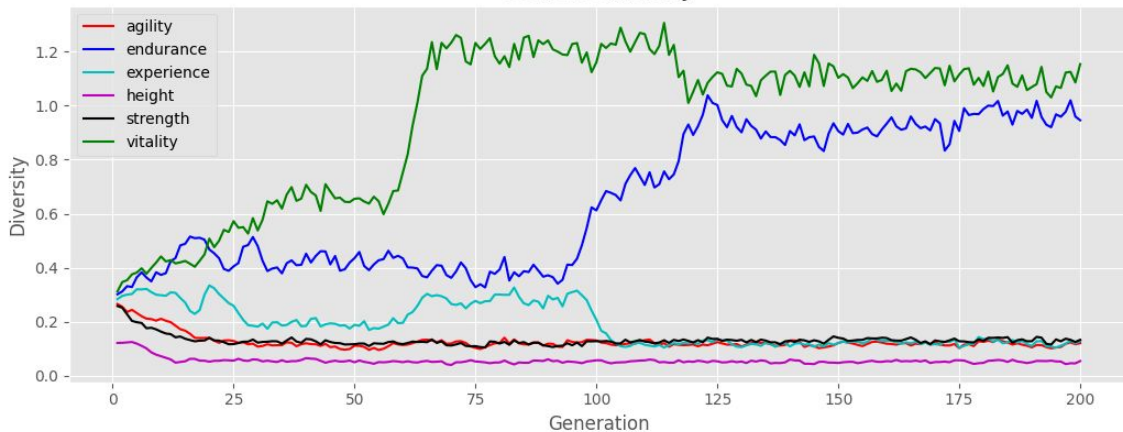
Stats of Best Individual of Generation



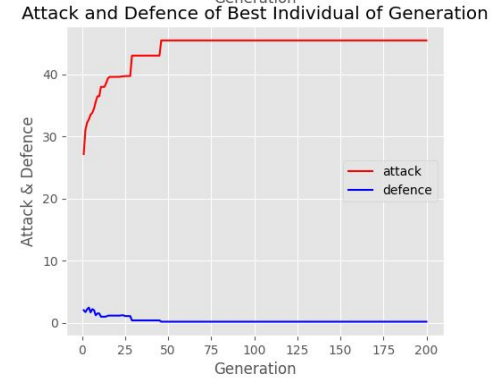
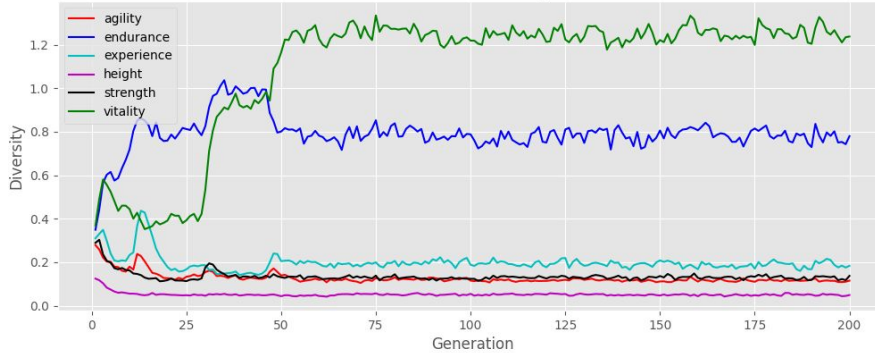
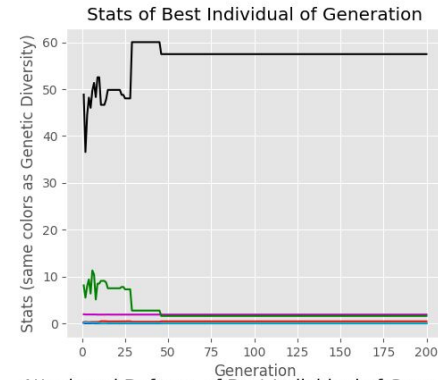
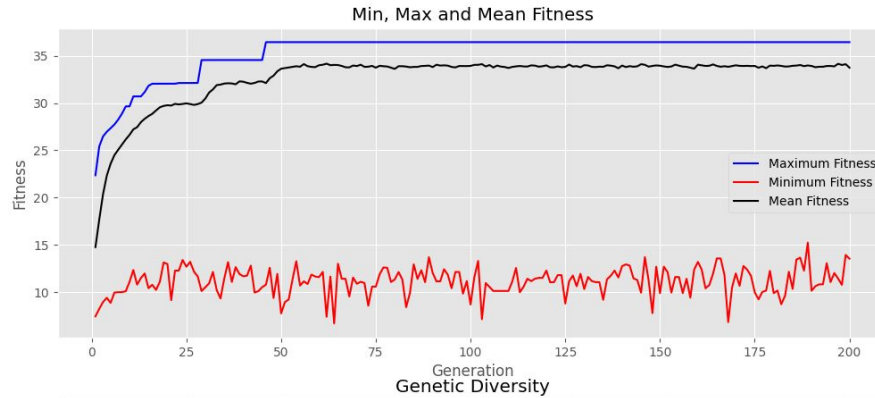
Attack and Defence of Best Individual of Generation



Survivor selection: elite 20%, boltzmann ruleta: universal, $t_0 = 100$, $t_f = 5$, convergencia = 0.05 80%

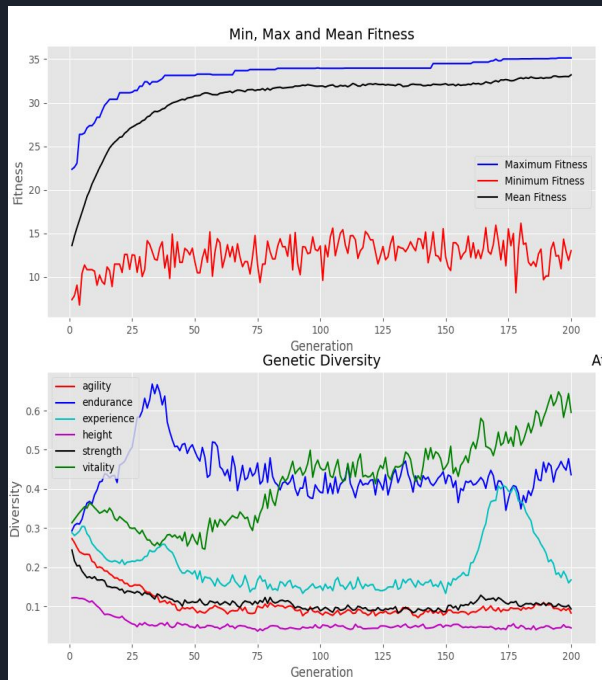


Survivor selection: elite 20%, torneo probabilístico con probabilidad = 0.75 80%

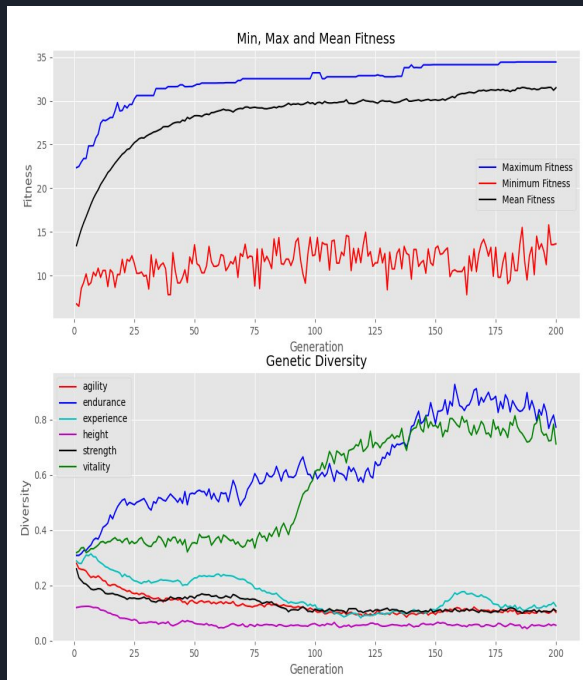


Survivor selection: universal 70%, torneo determinístico con tamaño 7% de la población 30%

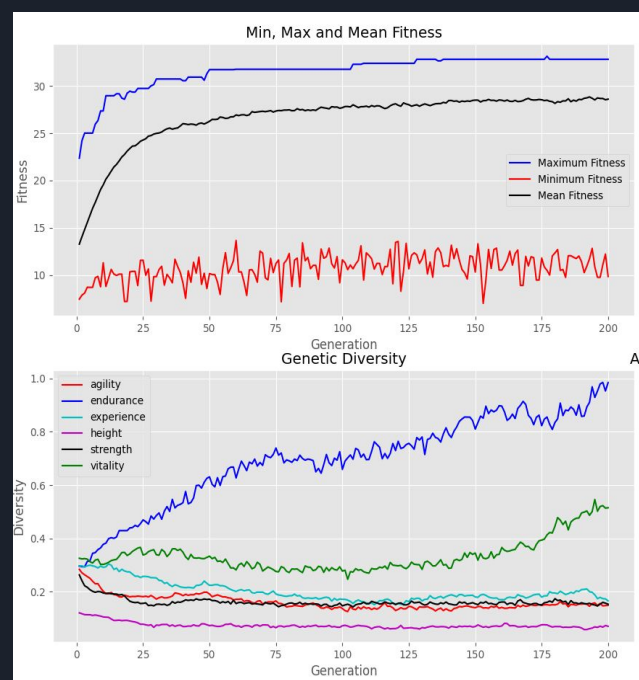
Universal 0.3 - Ranking 0.7



Universal 0.5 - Ranking 0.5



Universal 0.7 - Ranking 0.3



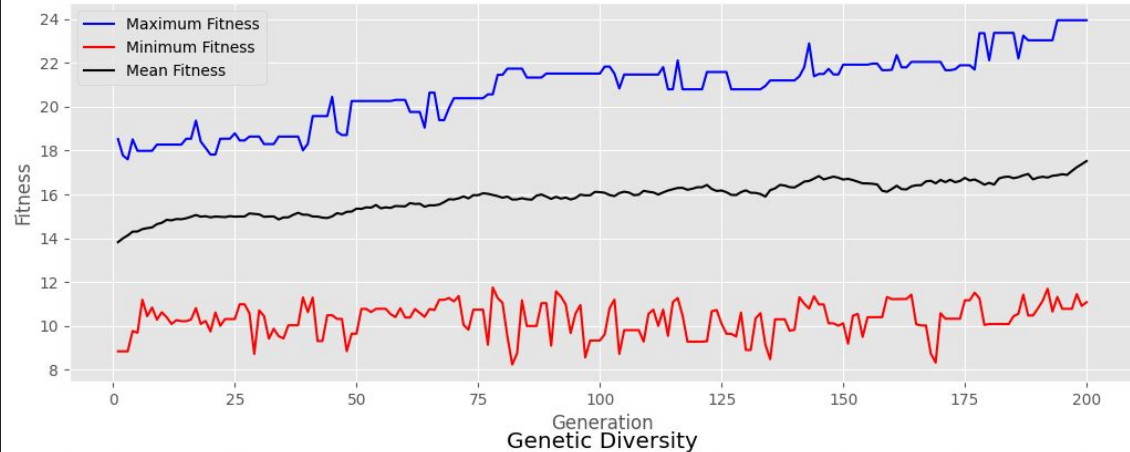


Gráficos

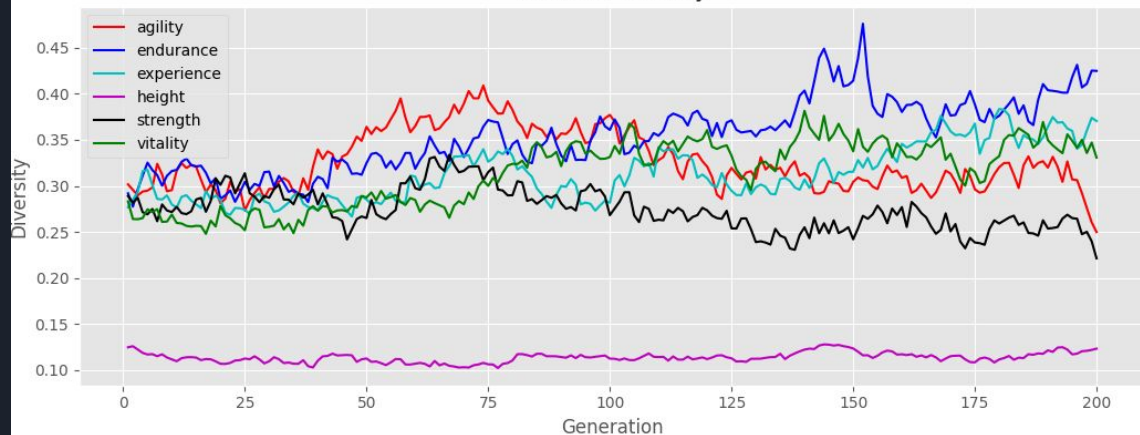
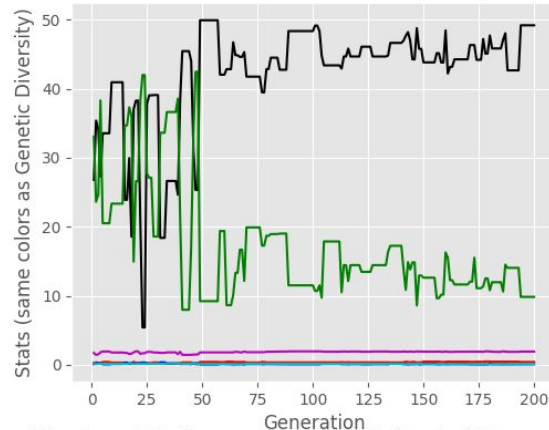
Los siguientes gráficos analizaron distintos métodos de selección, mutación y cruce:

- Generaciones de 1000 individuos
- Clase Guerrero
- 500 padres seleccionados
- Condición de corte 200 generaciones.
- Método de cruce y mutación uniforme
- Selección de sobrevivientes 50% elite 50% universal
- Probabilidad de mutación = 0.3
- Método de recombinación: Fill-all

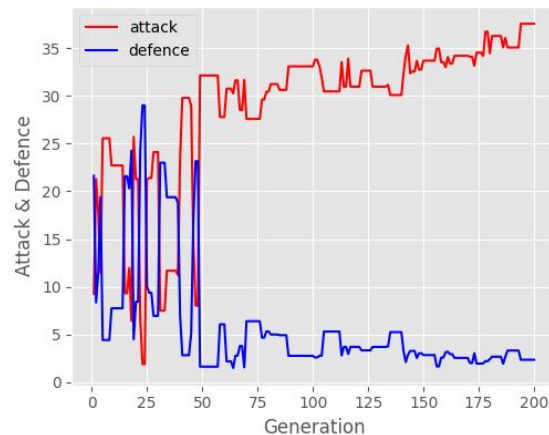
Min, Max and Mean Fitness



Stats of Best Individual of Generation

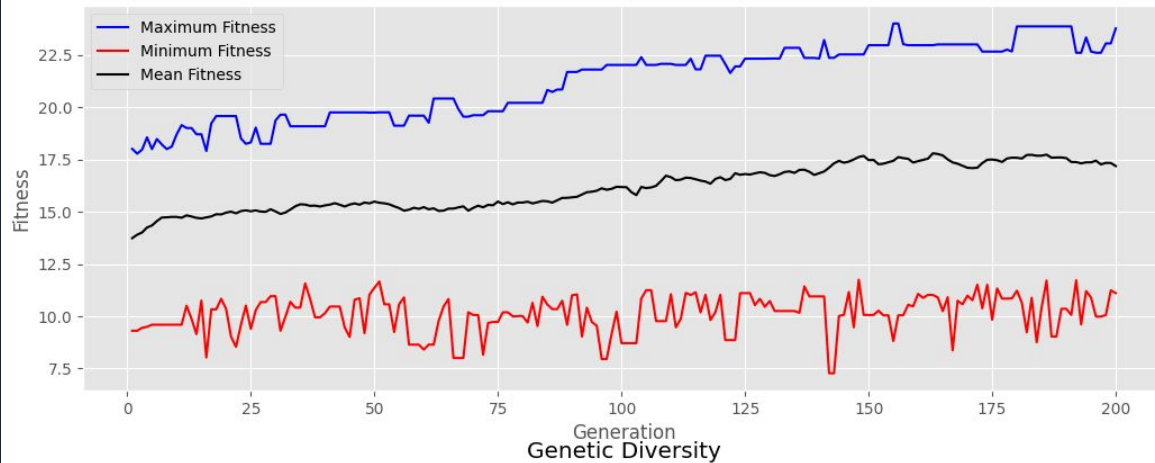


Attack and Defence of Best Individual of Generation

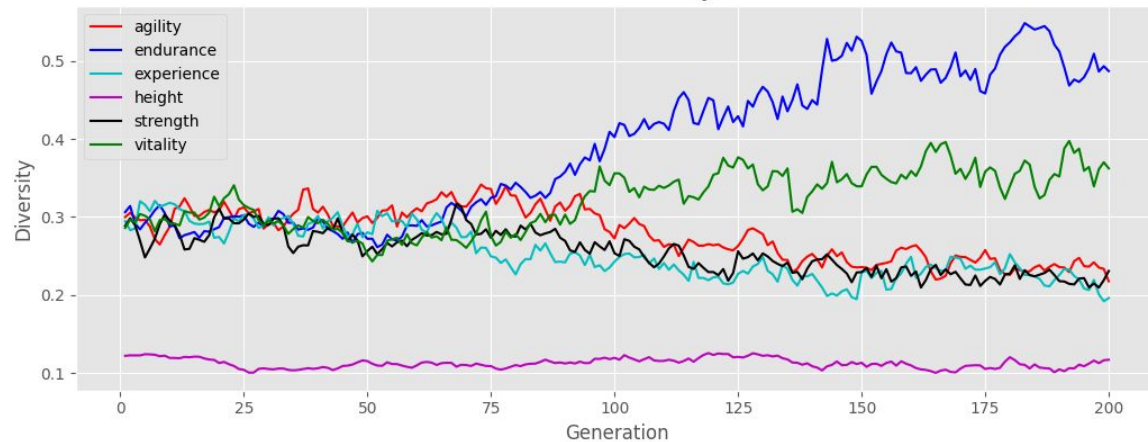
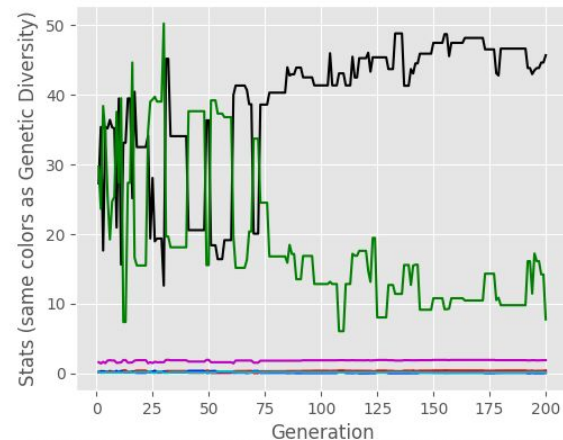


Selección: ruleta 10%, universal 90%

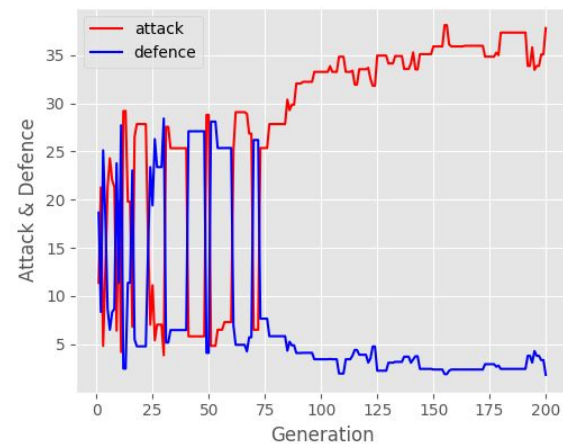
Min, Max and Mean Fitness



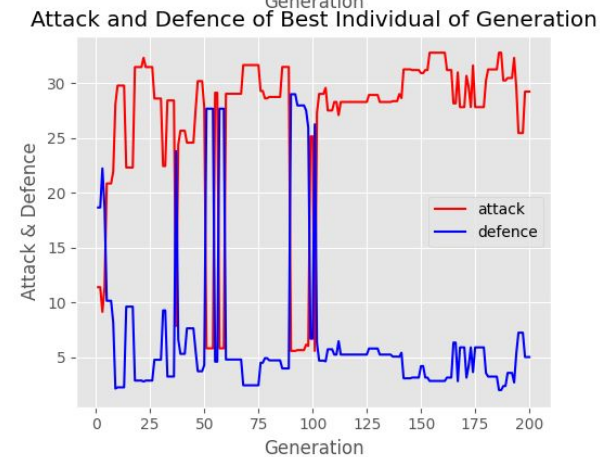
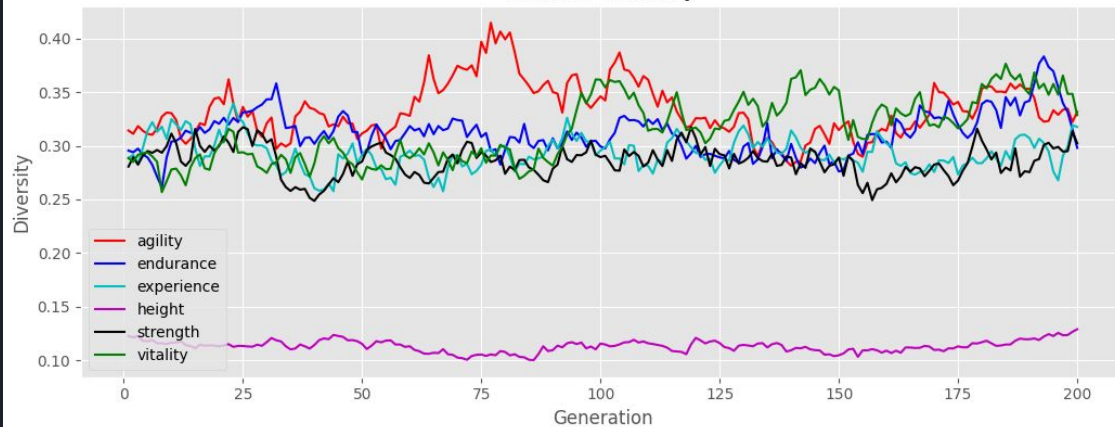
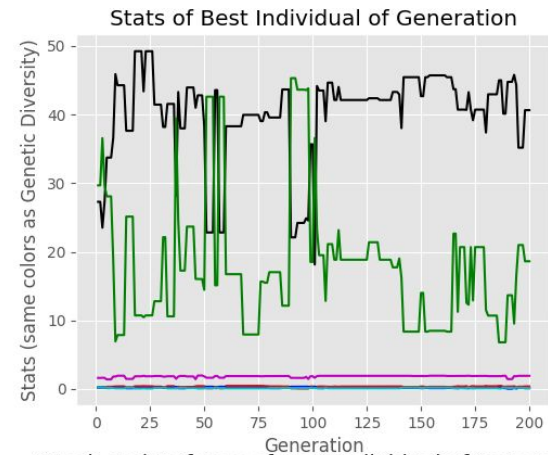
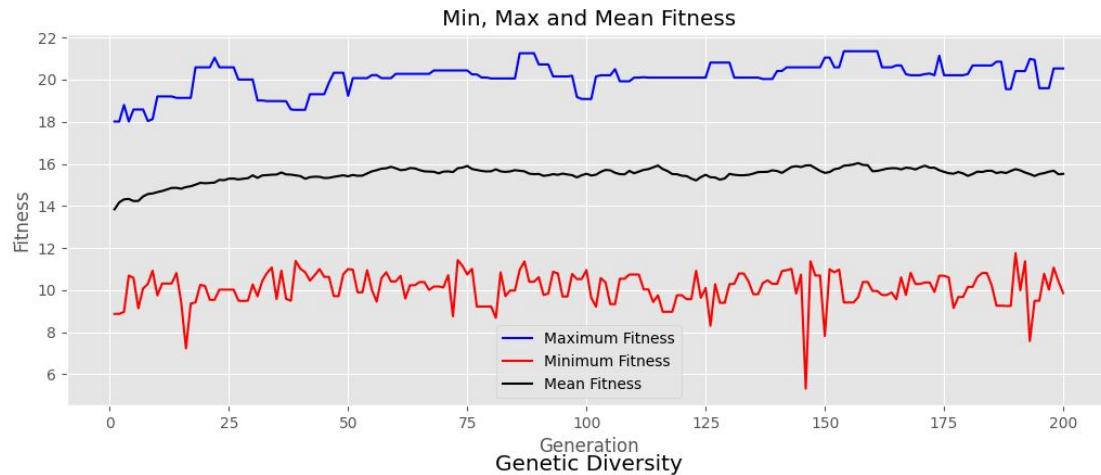
Stats of Best Individual of Generation



Attack and Defence of Best Individual of Generation



Selección: ruleta 90%, universal 10%



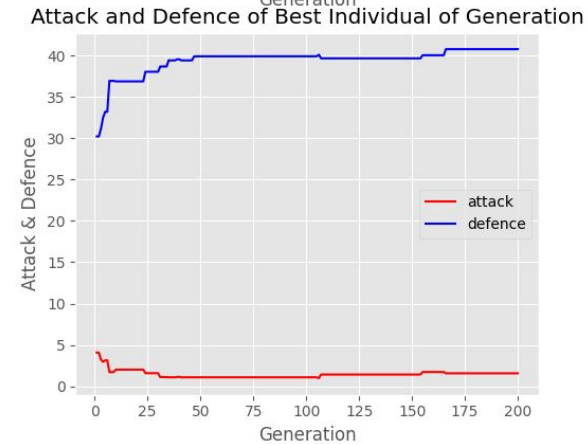
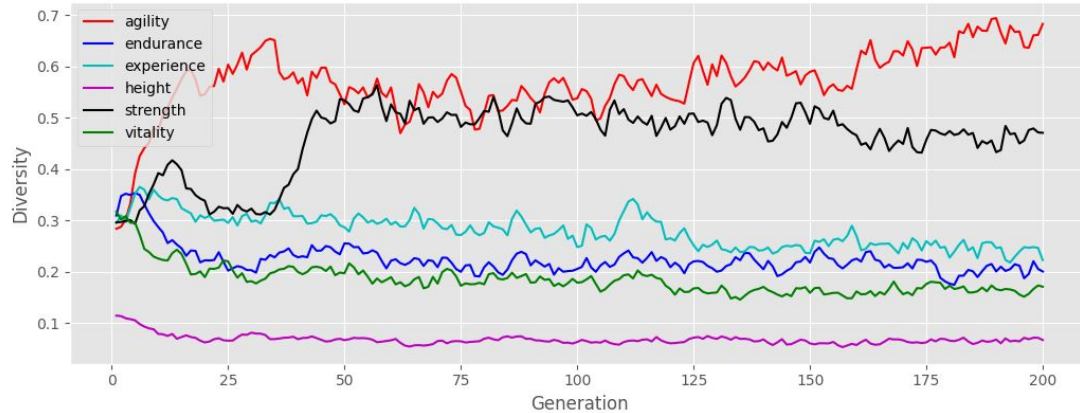
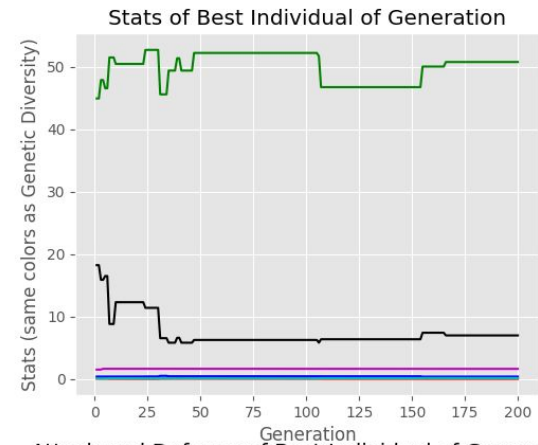
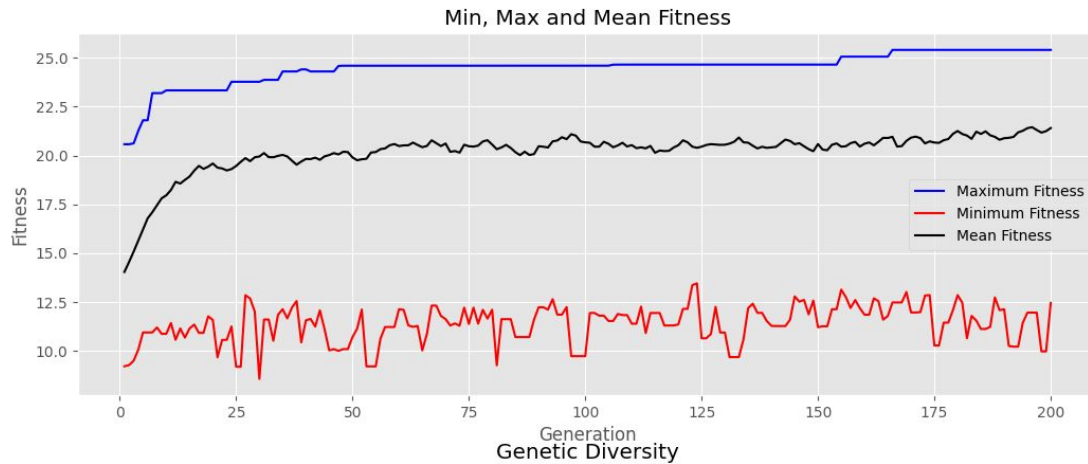
Selección: ruleta 50%, universal 50%



Conclusiones

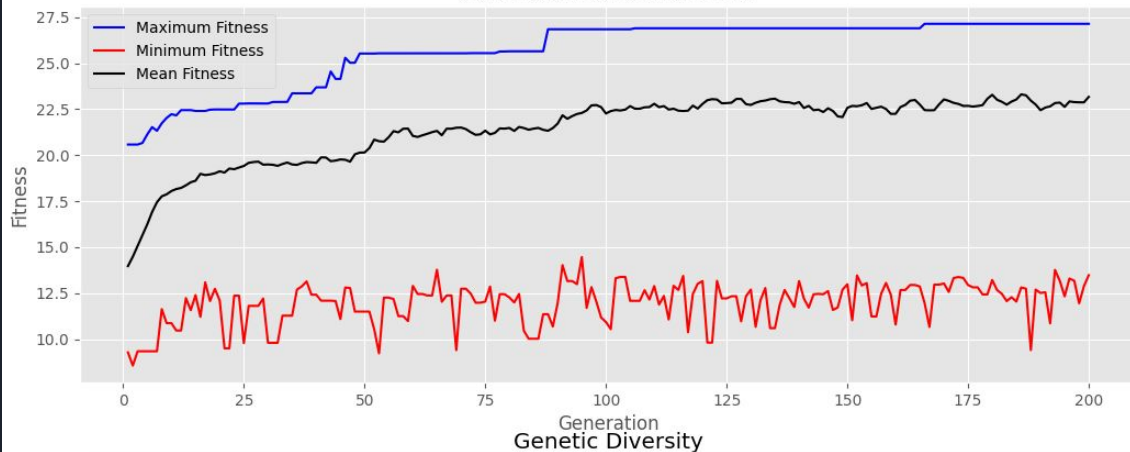
Como se puede ver en el gráfico del caso 1, todas las estadísticas presentan una gran diversidad, lo que se debe a la selección uniforme de números del método universal, la cual reduce la presión de selección propia del método de ruleta. Conforme se le da más espacio a este método, hay menor cantidad de estadísticas que presentan una alta diversidad. Esto se deberá (tomando en cuenta a lo mencionado anteriormente de cómo se tiende a maximizar o fuerza o defensa) los que presenten una mayor fuerza o defensa serán los de mejor fitness, y consecuentemente tendrán más posibilidad de ser elegidos.

En contraste, el caso 2, siendo más elitista, logra obtener un mejor resultado que los demás, sin embargo, comparado con el selector universal, no es una mejora significativa y puede estar muy ligado a las probabilidades de los números que se eligieron para la ruleta.

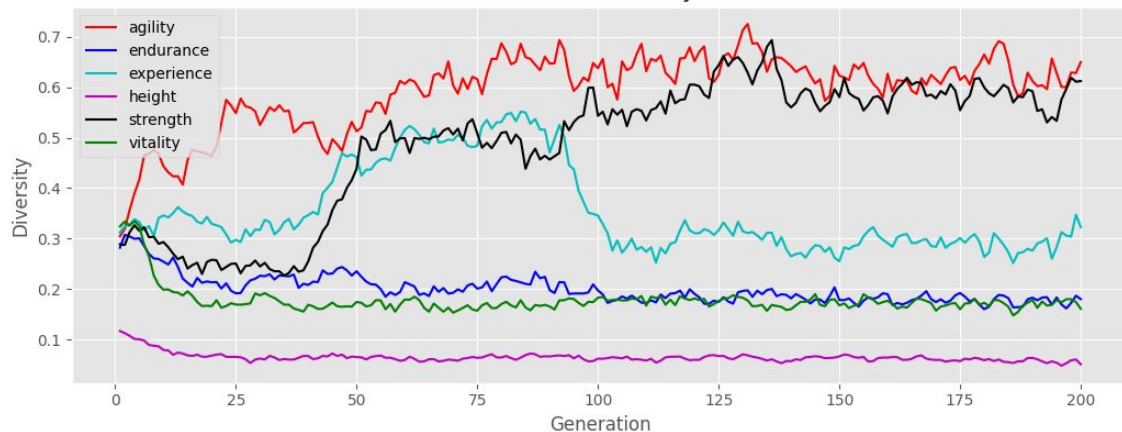
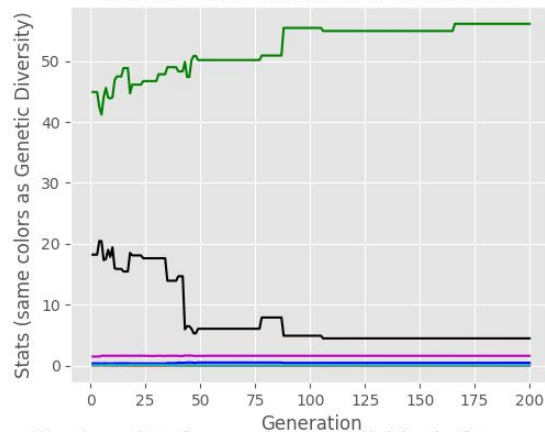


Selección: torneo probabilístico con $p=1/2$ 50%, torneo determinístico N=50 50%

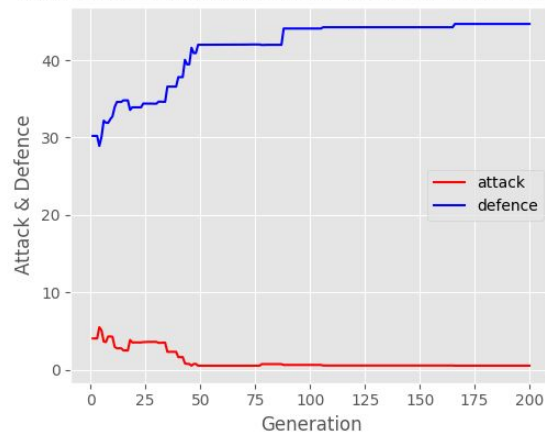
Min, Max and Mean Fitness



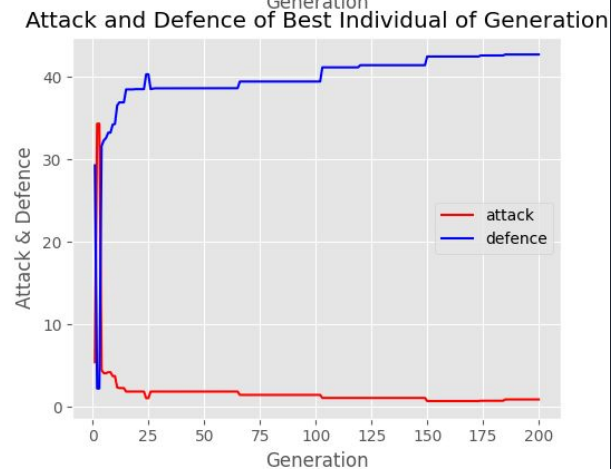
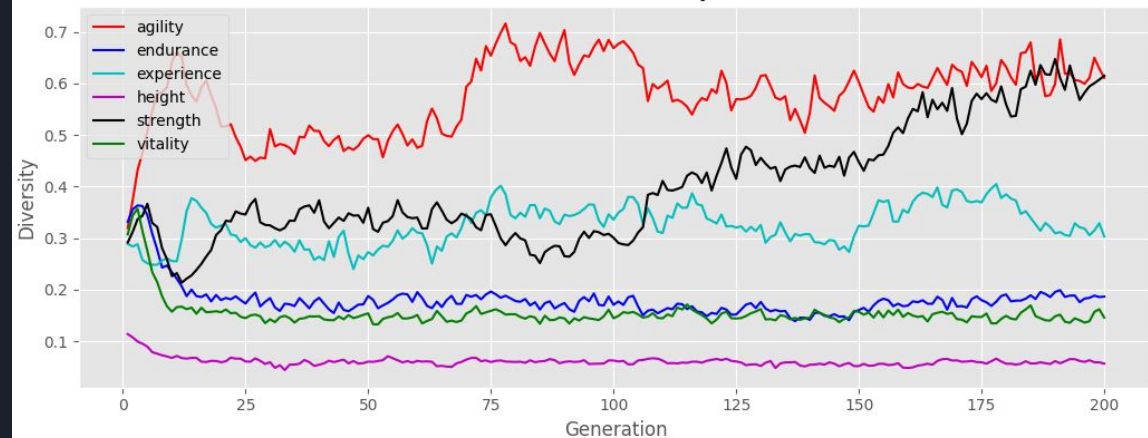
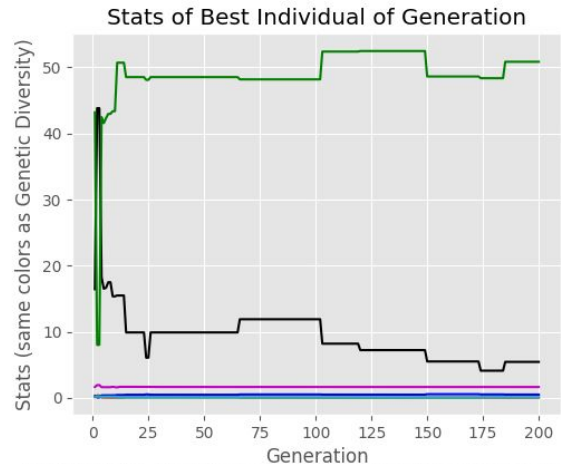
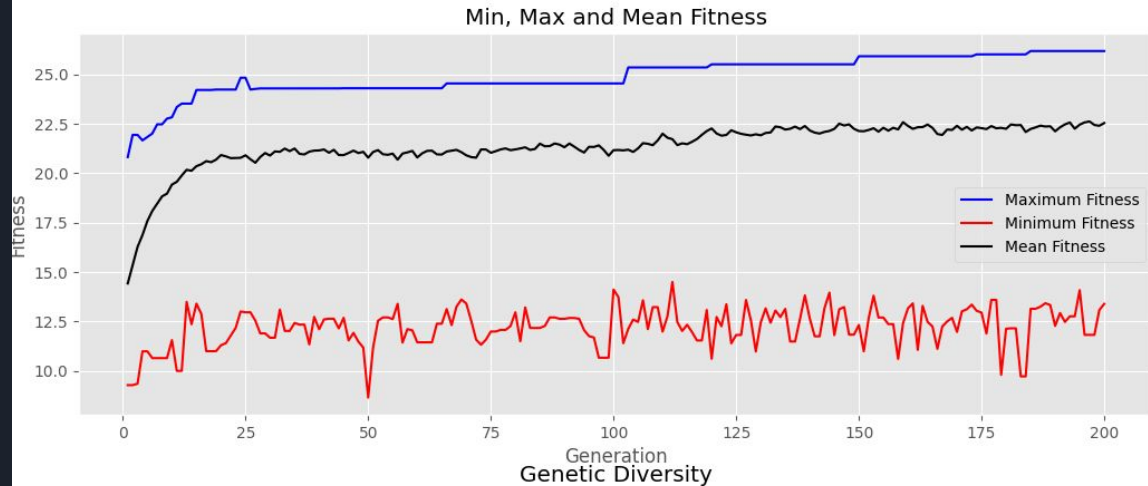
Stats of Best Individual of Generation



Attack and Defence of Best Individual of Generation



Selección: torneo probabilístico con $p=0.99$ 50%, torneo determinístico N=50 50%



Selección: torneo probabilístico con $p = \frac{1}{2}$ 10%, torneo determinístico $N = 50$ 90%



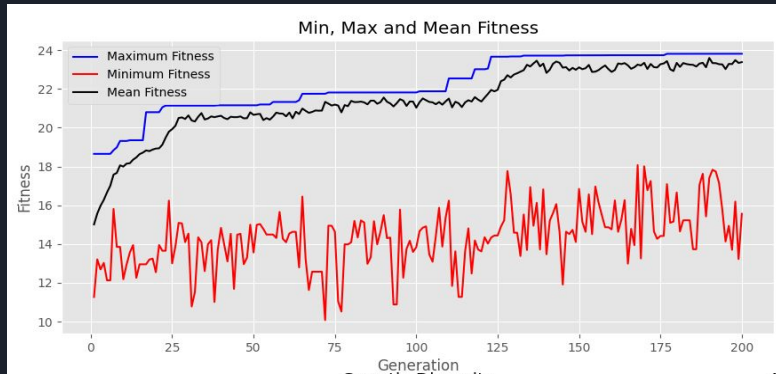
Conclusiones

Como puede comprobarse, el primer caso, en el cual la probabilidad es un medio, resulta ser el que presenta el peor resultado, dado que hay una probabilidad considerable de que se elija al peor individuo y esto termina resultando en un mejor individuo con menor fitness que los otros dos casos.

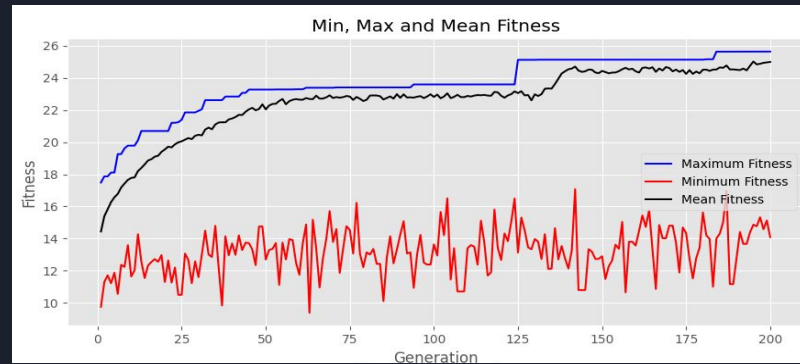
En contraste, el segundo caso, con una probabilidad muy alta, si ejerce una presión de selección importante lo cual, si bien por la naturaleza de la selección por torneos, resulta menor que el método elite, si se muestra una diferencia considerable con el primer caso. Esto termina obteniendo un mejor individuo con la fitness más alta de los tres casos

Por último, en el tercer caso, se puede apreciar como, luego de las primeras generaciones, rápidamente se obtienen mejores individuos, dado que N es el 10% de la generación, se tiene un método elitista, que permite que la curva de fitness media resulte muy vertical en las primeras generaciones, y luego termina ascendiendo asintóticamente.

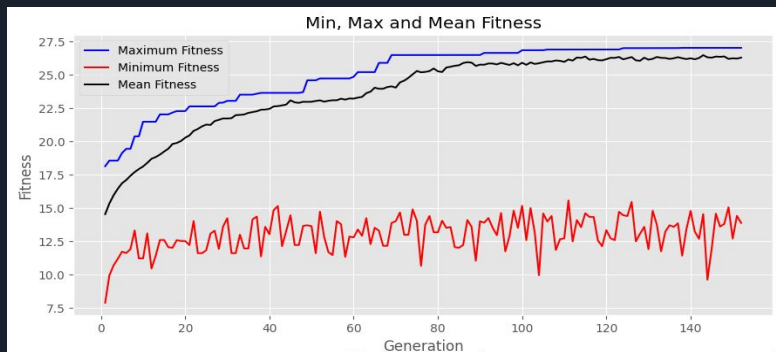
Generación: 100 - Hijos: 50



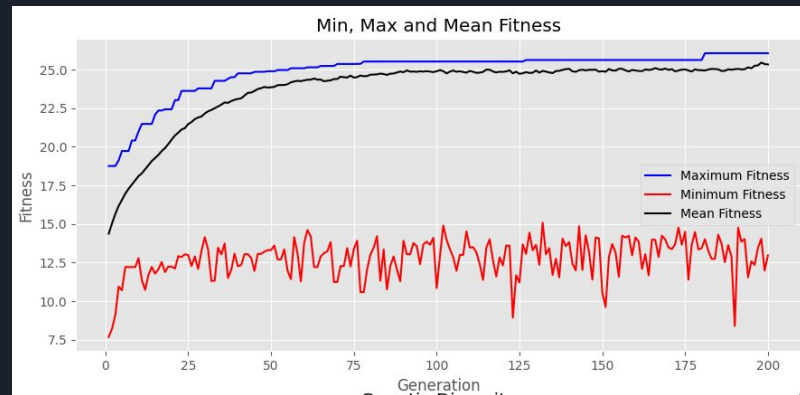
Generación: 250 - Hijos: 125



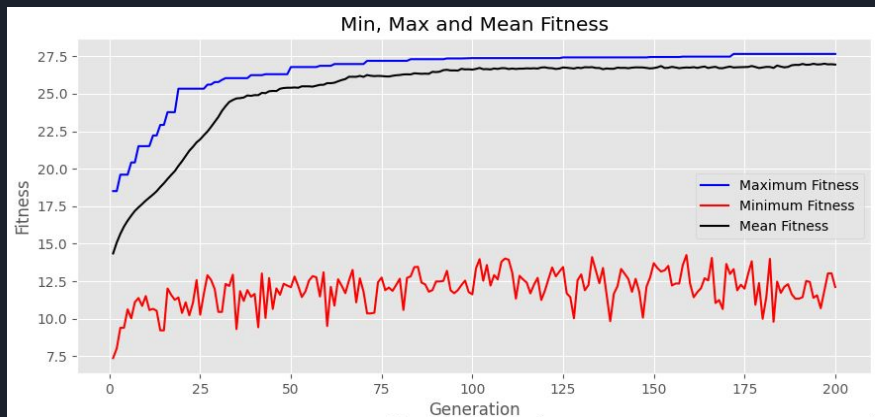
Generación: 500 - Hijos: 250



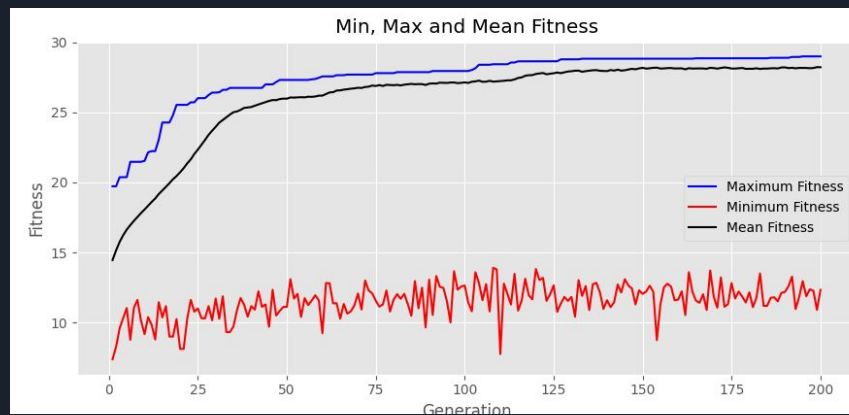
Generación: 1000 - Hijos: 500



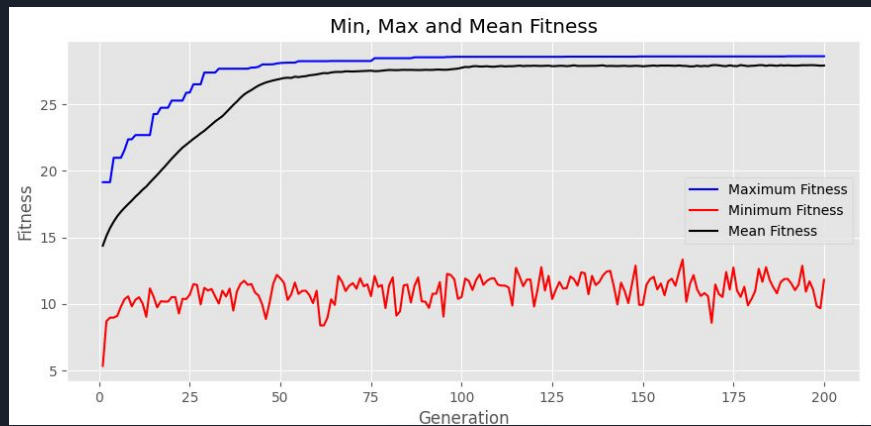
Generación: 2500- Hijos: 1250



Generación: 5000 - Hijos: 2500



Generación: 10000 - Hijos: 5000



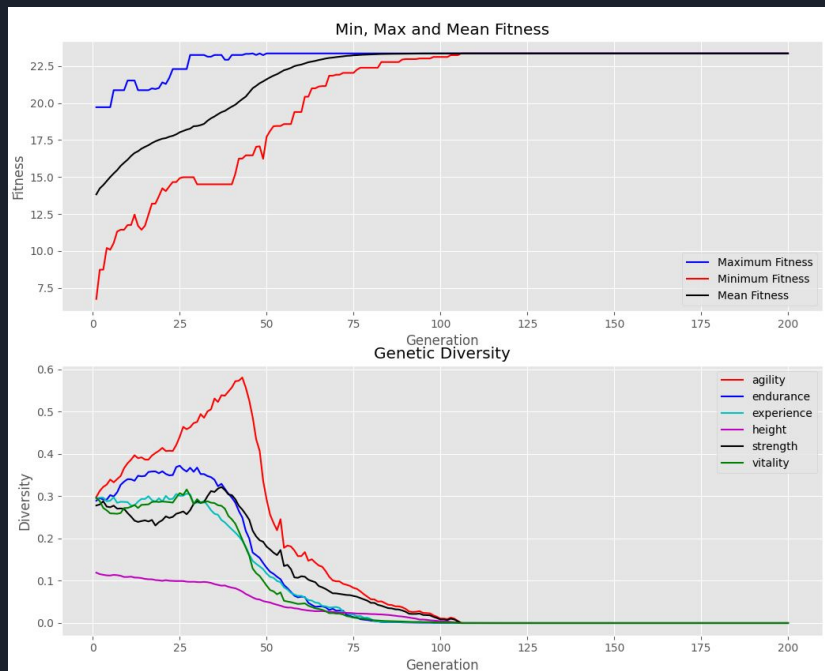


Cambios en el tamaño de población

Analizando el comportamiento de los fitness al correr la misma prueba variando únicamente el tamaño de población y la cantidad de hijos generada, se observa que los **resultados tienden a ser similares pero se alcanzan en menos generaciones.**

En base a esto concluimos que pueden ser usados como valores que se complementan (tamaño de población y cantidad de generaciones) y elegir el que sea más conveniente para el análisis actual. Ej: Si se quiere maximizar la performance, mayor población y menor generaciones sería mejor. En el caso de querer más granularidad en el progreso, sería mejor una menor población y más generaciones.

Probabilidad de mutación $p = 0$



Simulation Seed: tobias

Total Simulation Generations: 200

Best Character Found on Generation: 46

Best warrior from Simulation: Character(type='warrior', height=1.605617570750096, fitness=23.34819286421461)

Probabilidad de mutación $p = 0.3$



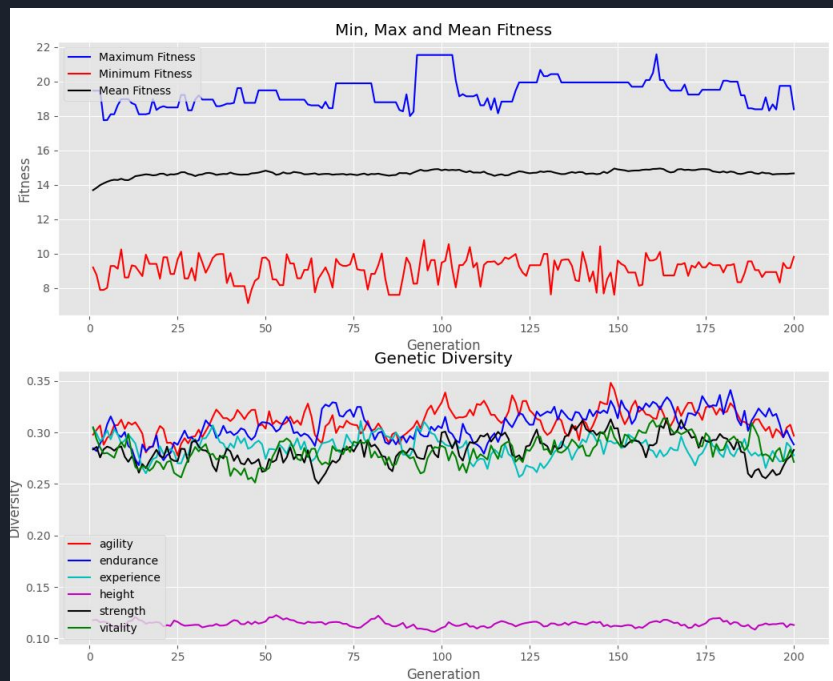
Simulation Seed: tobias

Total Simulation Generations: 200

Best Character Found on Generation: 196

Best warrior from Simulation: Character(type='warrior', height=1.6259463946061223, fitness=24.59626837239013)

Probabilidad de mutación $p = 0.6$



Simulation Seed: tobias

Total Simulation Generations: 200

Best Character Found on Generation: 161

Best warrior from Simulation: Character(type='warrior', height=1.8898441174794323, fitness=21.589343243792133)

Probabilidad de mutación $p = 0.9$



Simulation Seed: tobias

Total Simulation Generations: 200

Best Character Found on Generation: 176

Best warrior from Simulation: Character(type='warrior', height=1.4336686936955418, fitness=20.194533913925493)



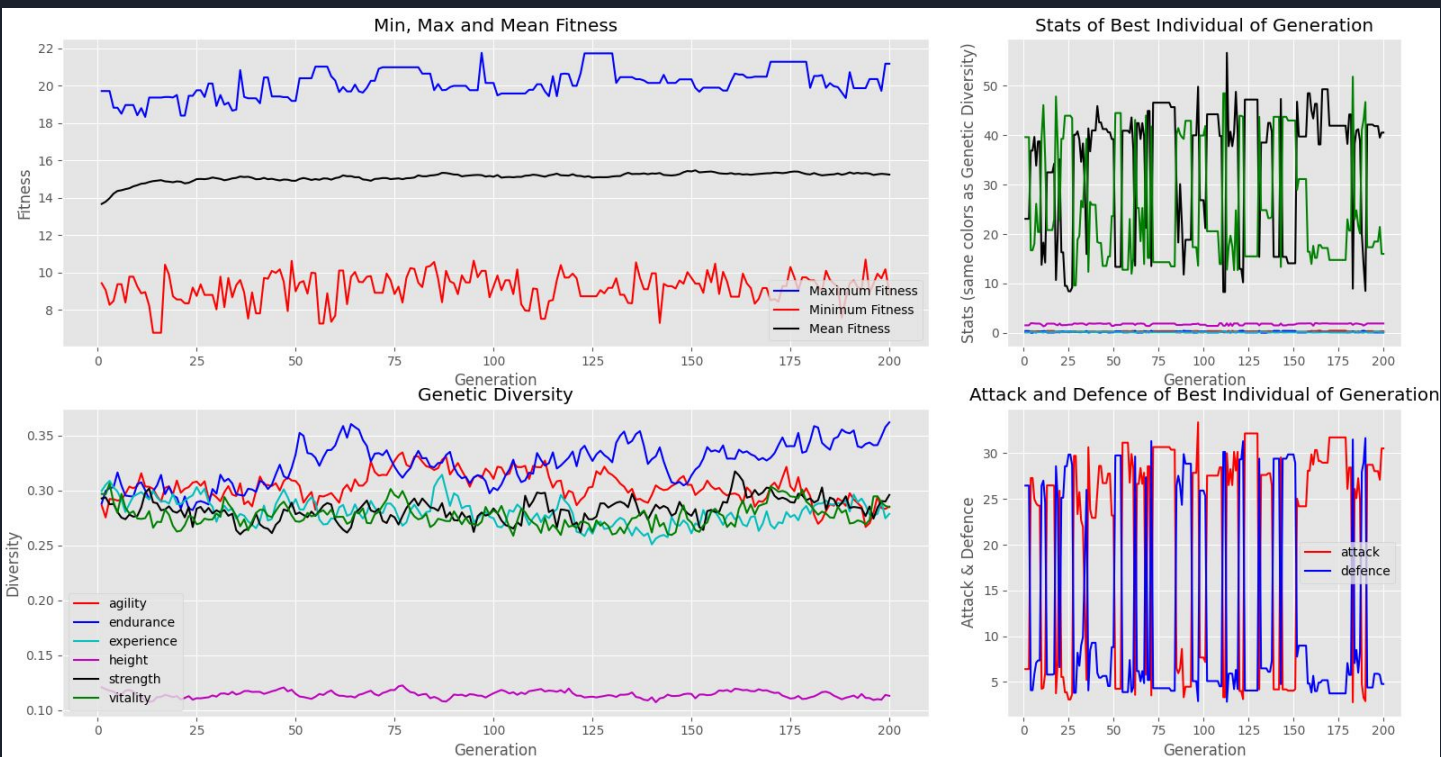
Cambios en la probabilidad de mutación

Configuración: Warrior - Ranking/Roulette - Cruza Uniforme - Fill all

Analizando el comportamiento de los fitness al correr la misma prueba variando únicamente la probabilidad de mutación, se comprueba cómo esta afecta de manera significativa la diversidad. En el primer caso, todos los individuos terminan siendo prácticamente iguales. En el segundo, se ve como las estadísticas principales no presentan tanta diversidad pero las demás si, mientras que en los otros dos casos todas son completamente diversas.

En base a esto concluimos que si bien es importante tener una posibilidad de mutación para aumentar la diversidad y que los individuos no terminen alcanzando un máximo local, esta no debe ser demasiado alta porque sino la población no podrá mejorar ya que constantemente están cambiando de manera aleatoria sus individuos. Esto solo hará que el fitness medio se mantenga constante, pero el máximo no crezca considerablemente.

Probabilidad de cruza $p = 0.3$, método de recombinación fill-parent



Simulation Seed: tobias

Total Simulation Generations: 200

Best Character Found on Generation: 97

Best warrior from Simulation: Character(type='warrior', height=1.8936615189936887, fitness=21.758938811860325)



Cambios en el método de recombinación

Configuración: Warrior - Ranking/Roulette - Probabilidad de mutación uniforme $p = 0.3$ - Cruza Uniforme - Fill parent

Contrastando este caso con el ejemplo similar, pero modificando el método de recombinación, se comprueba como este incrementa considerablemente la diversidad de las estadísticas. Esto se debe a que se eligen más hijos por lo que hay menor posibilidad de repetición en cada generación. Sin embargo, en este caso particular no resulta beneficioso ya que termina habiendo demasiada diversidad, por lo que una vez más no se logra obtener un individuo destacado, sino que se mantiene siempre dentro de una media estable. En conclusión, este método trae más diversidad y habrá que considerar esto a la hora de elegir la configuración general.



Sobre el análisis

- Una condición de corte fija como la de generaciones permite hacer un mejor análisis comparativo entre distintas estrategias. Una vez elegida la estrategia, una condición como convergencia de fitness permite buscar el mejor individuo sin limitar la estrategia y sin hacer desperdicio de recursos.
- La elección de gráficos y datos para analizar el comportamiento de cada experimento es difícil y depende tanto de nuestras hipótesis como de las estrategias utilizadas.
- Poblaciones mayores a 1000 comienzan a tener impacto notable en la performance. Aunque esto puede compensarse reduciendo la cantidad de generaciones.