

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Administración y Economía

Simulación de juego de ultimátum con inteligencia artificial y sesgos

Julio César Añasco Flores

Bryan Josué Naranjo Navas

Pamela Anahí Proaño Mora

Maria Anastasia Vasileuski Kramskova

Economía

Trabajo de fin de carrera presentado como requisito

para la obtención del título de

Economista

Quito, 20 de mayo de 2021

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ
Colegio de Administración y Economía

HOJA DE CALIFICACIÓN
DE TRABAJO DE FIN DE CARRERA

Comportamiento Artificial

Julio César Añasco Flores
Bryan Josué Naranjo Navas
Pamela Anahí Proaño Mora
Maria Anastasia Vasileuski Kramskova

Nombre del profesor, Título académico:

Carlos Andrés Uribe Terán, Ph.D en
Economía.

Quito, 20 de mayo de 2021

©Derechos de Autor

Por medio del presente documento certifico que he leído la Política de Propiedad Intelectual de la Universidad San Francisco de Quito y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo de investigación quedan sujetos a lo dispuesto en la Política.

Así mismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo de investigación en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Nombres y Apellidos:	Julio César Añasco Flores
Código:	00141586
Cédula de Identidad :	1725077992
Lugar y fecha:	Quito, 19 de mayo de 2021
Nombres y Apellidos:	Bryan Josué Naranjo Navas
Código:	00320604
Cédula de Identidad :	0603461617
Lugar y fecha:	Quito, 19 de mayo de 2021
Nombres y Apellidos:	Pamela Anahí Proaño Mora
Código:	00137379
Cédula de Identidad :	1718015942
Lugar y fecha:	Quito, 19 de mayo de 2021
Nombres y Apellidos:	Maria Anastasia Vasileuski Kramskova
Código:	00203766
Cédula de Identidad :	1718658501
Lugar y fecha:	Quito, 19 de mayo de 2021

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETHeses>.

UNPUBLISHED DOCUMENT

Note: The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETHeses>.

AGRADECIMIENTOS

Julio Añasco

Quiero agradecer a los profesores que permitieron realizar esta investigación, a Santiago José Gangotena por el tiempo que nos dedicó. Y Agradecer a todas las personas que creyeron en mí durante estos años de carrera.

Bryan Naranjo

El presente trabajo de investigación fue realizado bajo la supervisión del PhD. Carlos Uribe Terán y del PhD. Santiago José Gangotena, a quienes expreso mi profundo agradecimiento, por hacer posible la realización de este estudio. Además, agradezco por su tiempo, constancia y dedicación para el éxito de este trabajo. Deseo agradecer a mi familia Naranjo Navas por el apoyo y continuo amor incondicional durante mis años de estudios universitarios. Por ser un claro ejemplo de una familia unida bajo la protección de Dios. Sus

oraciones no han sido en vano y son una bendición para mi vida. Agradezco a mi madre Cumandá Navas, por ser mi mejor amiga y la inspiración que me permite seguir luchando. Gracias madre por tus palabras y acciones de ánimo. Mi padre por ser aquel hombre de Dios que ha seguido adelante a pesar de las dificultades. Mis hermanos Paúl, Andrés y Andrea, quienes han sido mi fortaleza, mis amigos incondicionales y aquellas personas especiales por las cuales mi amor he transmitido. A todos mis amigos que ya no están conmigo en estos momentos. Sé que pronto nos veremos y reiremos otra vez. Sus nombres serán recordados e inmortalizados poco a poco en cada uno de mis escritos. Jonathan, Jaime y Alexander. Agradezco a mi segunda familia Jenks Olson a quienes a pesar del tiempo siempre los tendré en mi corazón. Gracias por sembrar las bases de quien soy ahora Y por último y no menos importante agradezco a Dios por ser tan misericordioso y caballeroso, por esperarme y entenderme. Gracias padre, tú me salvaste.

Pamela Proaño

Este trabajo de investigación, si bien ha requerido de esfuerzo y dedicación por parte de los autores, no hubiese sido posible sin la cooperación de todas y cada una de las personas que mencionaré a continuación. Primero y antes que nada, agradezco a Dios por cada día de vida. Agradezco hoy y siempre a mis padres y a mi abuela materna sin cuyo apoyo, paciencia, amor y crítica constructiva mi contribución en este proyecto no hubiera sido posible. A Santiago Gangotena por su ayuda y conocimientos impartidos. A nuestro tutor, Carlos Uribe por su guía y dirección en la presentación y escritura del documento. A nuestros profesores, Pedro Romero Alemán y Sebastián Oleas Martín por aclarar nuestras dudas en la materia. A mi grupo del proyecto, Julio Añasco, Bryan Naranjo y Anastasia Vasileuski por sus ánimos de siempre aprender, y su constancia en sus responsabilidades del proyecto.

Anastasia Vasileuski

Agradezco a todos los profesores que me ayudaron en cada paso de mi carrera; a la

decana Mónica Rojas por incentivar me en los momentos en los que creí que no podía más; a mi tutor, Carlos Uribe por su apoyo incondicional tanto académico como personal y al profesor Santiago José Gangotena que hizo posible que esta investigación se convierta en un material publicable. Finalmente agradezco a mi madre, Elena Kramskova por entregarme día a día mucho más de lo que parece posible, todos mis logros siempre serán dedicados a ella.

RESUMEN

En esta investigación se desarrollan diseños experimentales del juego del ultimátum con agentes de inteligencia artificial, sesgados y no sesgados. Se utilizó aprendizaje por refuerzo y Bucket Brigade. Mediante simulaciones y comparación de comportamientos se han estudiado las siguientes preguntas: ¿Llega la inteligencia artificial a un equilibrio de subjuego perfecto en el experimento del juego del ultimátum? ¿Cómo se comportaría la Inteligencia Artificial en el experimento del juego del ultimátum si se le incluye un pensamiento sesgado? El resultado principal es que la inteligencia artificial por sí sola no llega a un equilibrio de subjuegos perfecto. Por otro lado, se demostró que los diseños experimentales de agentes con sesgo convergen a un equilibrio rápidamente. Y por último, los agente con sesgo de envidia se comportan igual que los que tienen sesgo de pensamiento justo.

Palabras clave: Aprendizaje por refuerzo, Bucket Brigade, Homoeconomicus, Pensamiento Sesgado, Equilibrio de Nash, Equilibrio de subjuego perfecto y Óptimo de Pareto.

ABSTRACT

In this research we have developed experimental designs of the ultimatum game with supervised agents. These agents have unbiased and biased thinking depending on the case. We used Reinforcement Learning and Bucket Brigade to program the artificial agents. We used simulations and behavior comparison to answer the following questions: Does artificial intelligence reach a perfect subgame equilibrium in the ultimatum game experiment? How would Artificial Intelligence behave in the Ultimatum Game experiment if biased thinking is included in it? This exploratory analysis showed one important result: artificial intelligence by itself doesn't reach a perfect subgame equilibrium. Whereas, the experimental designs with biased thinking agents quickly converge to an equilibrium. Finally, we demonstrated that the agents with envy bias behave the same as the ones with fair minded bias.

Keywords: Reinforcement Learning, Bucket Brigade, Homo-economicus, Biased Thinking, Nash Equilibrium, Perfect Subgame Equilibrium, and Pareto Optimum.

TABLA DE CONTENIDOS

LISTA DE FIGURAS	11
1 INTRODUCCIÓN	13
2 METODOLOGÍA	15
2.1 El juego del ultimátum	15
2.2 Experimento del juego del ultimátum	16
2.2.1 Pseudocódigo del experimento del juego del ultimátum	18
2.3 Comportamiento sesgado	18
2.4 Inteligencia artificial en el diseño experimental	21
3 CONCLUSIONES	23
3.1 Resultados	23
3.1.1 Juego del Ultimátum en Agente sin sesgo con aprendizaje supervisado .	23
3.1.2 Juego del Ultimátum en un Agente con aprendizaje no supervisado sin sesgo	24
3.1.3 Juego del Ultimátum en un Agente Altruista (sesgo) con aprendizaje no supervisado	25
3.1.4 Juego del Ultimátum en un Agente envidioso (sesgo) con aprendizaje no supervisado	26
3.1.5 Juego del Ultimátum en un Agente con pensamiento justo (sesgo) con aprendizaje no supervisado	27
3.2 Conclusiones finales	28
4 REFERENCIAS BIBLIOGRÁFICAS	30
5 ANEXOS	33
5.1 Pseudocódigos	33
5.1.1 Pseudocódigo para la creación de clase de los agentes	33
5.1.2 Pseudocódigo sin sesgo con aprendizaje supervisado	33

5.1.3	Pseudocódigo sin sesgo con aprendizaje no supervisado	34
5.1.4	Pseudocódigo altruismo con aprendizaje no supervisado	34
5.1.5	Pseudocódigo envidia con aprendizaje no supervisado	34
5.1.6	Pseudocódigo pensamiento justo con aprendizaje no supervisado	35
5.2	Histogramas	35
5.3	Resultados del diseño experimental con 100 rondas	39
5.4	Resolución del Diseño experimental con 500 rondas	45
5.5	Media Diseño experimental 1000 rondas	50

LISTA DE FIGURAS

1	Juego del ultimátum original	15
2	Inducción hacia atrás juego del ultimátum	16
3	Diseño experimental del juego del ultimátum	17
4	Diagrama Agente Supervisado sin sesgo	20
5	Diagrama Agente no supervisado sin sesgo	20
6	Diagrama Agente Altruista	20
7	Diagrama Agente envidioso	21
8	Diagrama Agente Justo	21
9	Agente Supervisado sin sesgo 1000 rondas moda	23
10	Agente no supervisado sin sesgo 1000 rondas moda	24
11	Agente Altruista 1000 rondas moda	25
12	Agente envidioso 1000 rondas moda	26
13	Agente Justo 1000 rondas moda	27
14	Agente Supervisado histograma	36
15	Agente sin sesgo histograma	36
16	Agente Altruista histograma	37
17	Agente Envidioso histograma	37
18	Agente Justo histograma	38
19	Agente Supervisado 100 rondas media	39
20	Agente Supervisado 100 rondas moda	40
21	Agente sin sesgo 100 rondas media	40
22	Agente sin sesgo 100 rondas moda	41
23	Agente Altruista 100 rondas media	41
24	Agente Altruista 100 rondas moda	42
25	Agente Envidia 100 rondas media	42
26	Agente Envidia 100 rondas moda	43
27	Agente Justo 100 rondas media	43
28	Agente Justo 100 rondas moda	44

29	Agente Supervisado 500 rondas media	45
30	Agente Supervisado 500 rondas moda	45
31	Agente sin sesgo 500 rondas media	46
32	Agente sin sesgo 500 rondas moda	46
33	Agente Altruista 500 rondas media	47
34	Agente sin sesgo 500 rondas moda	47
35	Agente Envidia 500 rondas media	48
36	Agente Envidia 500 rondas moda	48
37	Agente Justo 500 rondas media	49
38	Agente Justo 500 rondas moda	49
39	Agente Supervisado 1000 rondas media	50
40	Agente sin sesgo 1000 rondas media	50
41	Agente Altruista 1000 rondas media	51
42	Agente Envidioso 1000 rondas media	51
43	Agente Justo 1000 rondas media	52

1. INTRODUCCIÓN

Los economistas Güth, Schmittberger y Schwarze¹ llevaron el juego del ultimátum² a la experimentación por primera vez para testear las respuestas humanas y compararlo con la racionalidad de maximización de beneficios expuesta por la teoría económica. Se concluyó que, los jugadores no tomaban las estrategias que conducían al equilibrio de Nash, por lo que los jugadores no tomaron la combinación de decisiones que más ganancias maximizaban y que mayor beneficios daba. Nuestra motivación parte de aquí para plantear la incógnita de lo que pasaría en un experimento del juego del ultimátum con entes absolutamente racionales, sin distracción ni factores de cansancio, en donde se testea si el rol del supuesto de racionalidad afecta a los resultados. ¿Llega la inteligencia artificial (IA) al equilibrio de subjuego perfecto en el experimento del juego del ultimátum? y ¿Cómo se comportaría la inteligencia artificial en el experimento del juego del ultimátum si es que le incluimos sesgos?

Para contestar estas preguntas de investigación creamos un experimento del juego del ultimátum con distintos tipos de agentes artificiales. Se emplearon dos métodos de aprendizaje y además se creó una distinción entre agentes sesgados y no sesgados. Las distintas metodologías mencionadas permitieron que se estudien los comportamientos que surgen en los jugadores o agentes inteligentes.

La contribución de este estudio a la literatura empírica rodea las ramas de la teoría de juegos, economía experimental, economía del comportamiento e inteligencia artificial. Se aporta al crecimiento de la rama de la IA en el campo de la economía. Los resultados obtenidos sirven como análisis de las variantes de equilibrio en los experimentos del juego del ultimátum y ayudar a profundizar en el estudio del entendimiento del comportamiento humano.

¹Economistas alemanes; primeros en definir un experimento del juego del ultimátum en el año 1982.

²Este juego se remonta hasta la paradoja de San Petersburgo creada por Daniel Bernoulli en 1738, pero no fue hasta 1838 que por primera vez se plantea una solución teórica a este juego por Antoine Augustin Cournot.

Sanfey (2012), con su estudio *Cooperación en redes sociales* definió los sesgos que influyen en los participantes de un experimento del juego del ultimátum y explicó que estos alejan al ser humano del supuesto de *homo economicus*, que se refiere a un supuesto de comportamiento del ser humano.

Samuelson, Keneth y Gale en 1995, especificaron en su estudio *Learning To Be Imperfect* que los experimentos del juego del ultimátum no llegan a un equilibrio teórico porque probablemente este equilibrio está mal planteado. Esto concuerda con un estudio subsecuente realizado por Chang y Maheswaran, en el 2012, llamado *Decision Making with imperfect decision Makers*, en donde concluyen, con el uso de inteligencia artificial, que lo racional para el jugador 2 es poseer pagos más equitativos y no converger al equilibrio de Nash. Nuestro aporte recae en crear distintos métodos de aprendizaje en los agentes artificiales para poder comprobar o refutar estas teorías.

El estudio está organizado de la siguiente manera: la sección 2 de la presente investigación expone la metodología que se utilizó para construir a los participantes del diseño experimental del juego del ultimátum. En ella se explica cómo se definieron los agentes racionales y los agentes sesgados. La sección 3 detalla los resultados conseguidos con cada tipo de agente en las distintas simulaciones. En la sección 4 se proveen las conclusiones de este estudio y se mencionan los posibles temas de investigación que podrían surgir a partir del mismo.

2. METODOLOGÍA

2.1. El juego del ultimátum

En el juego del ultimátum original existen dos jugadores, el primer jugador (J1) posee una cantidad de 1 que debe compartir con el jugador 2 (J2). J1 procede a hacerle una propuesta de cómo repartir este monto definido como $X \in [0, 1]$ y J2 tiene que decidir si acepta esta repartición o no. En caso de rechazar, ambos se quedarán con 0. En caso de aceptar, el J2 obtendrá la repartición ofrecida por J1. Este juego consta de solo una ronda y por ello, de una única oportunidad para tomar la decisión óptima, que es el equilibrio de Nash. La modalidad de este juego se detalla de mejor forma en la figura 1.

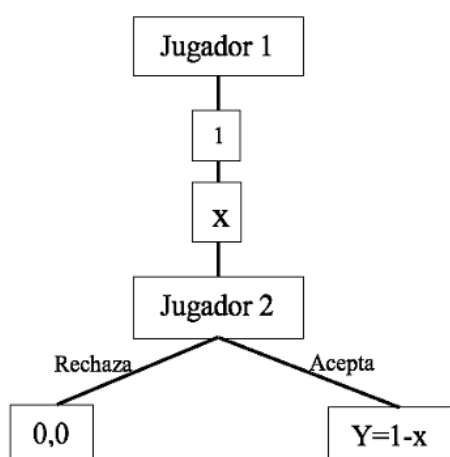


Figura 1: Juego del ultimátum original

La ciencia económica establece que si ambos jugadores actúan de forma racional en este juego, deben converger al equilibrio de Nash (0,9:0,1). "El jugador 1 siempre aproximará la máxima cantidad posible a su favor y ofrecerá la menor cantidad posible al jugador 2"(Gibbons,1992). Se aceptará cualquier valor por encima de 0, porque en otros casos, J2 es indiferente entre aceptar o rechazar. Para llegar al equilibrio teórico en juegos repetidos debe existir un refinamiento, en donde las estrategias de los jugadores deben haber llegado en una ronda al equilibrio de Nash.

En el juego del ultimátum repetido, el jugador 1 distribuye el mismo monto de 1 en N repeticiones. Como en cada ronda la decisión racional sigue siendo la misma, no existe razón alguna para desviarse del equilibrio de Nash cómo cuando hay una sola ronda. Lo más óptimo sigue siendo seleccionar el máximo valor posible siendo jugador 1, y aceptar cualquier valor mayor que 0 siendo jugador 2. Para demostrar esta afirmación se puede utilizar inducción hacia atrás.

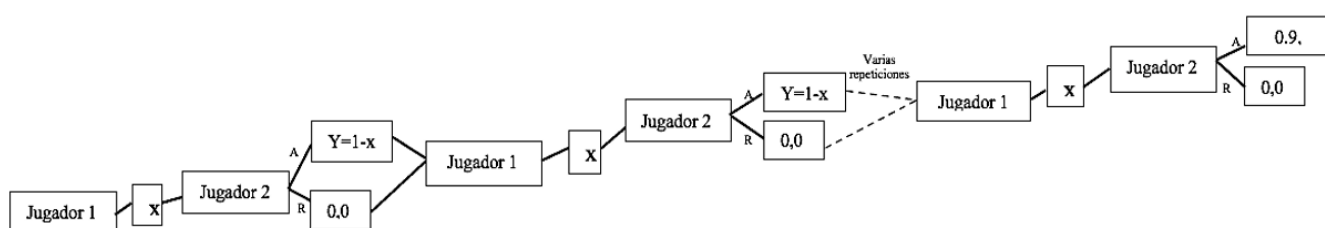


Figura 2: Inducción hacia atrás juego del ultimátum

En la figura 2 se denota que en cada una de las rondas escogidas se procede a la misma decisión racional, por lo que, el proceso de inducción hacia atrás demuestra que el equilibrio de Nash es el mismo para subjuegos. No existe una posible estrategia de gatillo³.

2.2. Experimento del juego del ultimátum

Nuestro experimento está compuesto por 1000 rondas, para asegurar que los participantes lleguen a una convergencia. El rol se asigna aleatoriamente para cada agente y es el mismo durante todo el experimento. Los agentes permanecen con su misma pareja hasta finalizar todas las rondas. Cada vez que una propuesta es aceptada o rechazada, se muestran los pagos y se finaliza la ronda. El valor a repartir entre ambos agentes es de 10 y siempre permanecerá igual. La repartición de 10 unidades será en números enteros para facilitar la extracción de información.

Implementamos 1000 pares de agentes con inteligencia artificial, es decir 2000 jugadores

³Estrategia de gatillo: un jugador coopera siempre y cuando lo haga su rival, pero si se desvía una vez, no podrán volver a cooperar nunca

en total para tener una base de datos amplia que nos permita identificar un patrón. Para la primera ronda, los agentes 1 y 2 prueban un número aleatorio en su intervalo de decisión (1;10) para definir sus estrategias respectivamente. Y el agente 2 aceptará o rechazará comparando su estrategia contra la que le ofrecen, esto con el fin de que ambos puedan entender el funcionamiento del juego. El diagrama que detalla nuestro diseño experimental se observa en la figura 3.

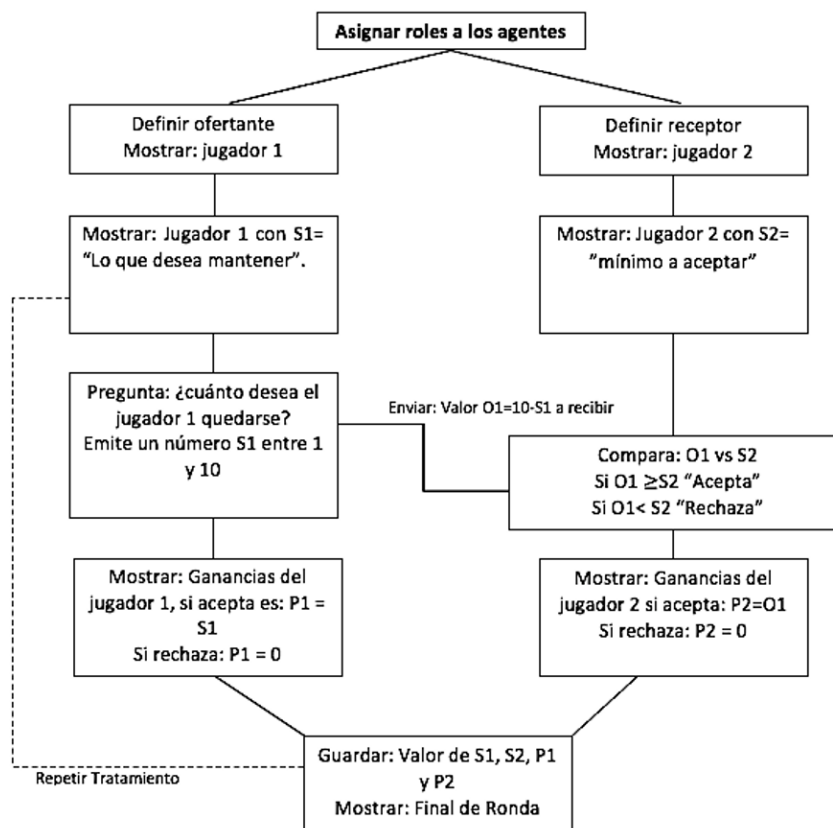


Figura 3: Diseño experimental del juego del ultimátum

Al finalizar este experimento las estrategias de los agentes deberían converger a (9:1), según el equilibrio de Nash en subjuegos perfectos. Sin embargo, de acuerdo a Carrillo (2017), el diseño experimental del juego del ultimátum con seres humanos no concluyen con la mayoría de individuos ofreciendo cantidades por encima del 50 por ciento ⁴.

⁴El estudio de Román Carrillo J. en el año 2017, "Cooperación en redes sociales; el juego del ultimátum" mostró como resultado que, es imposible que exista una experimentación del juego del ultimátum con seres humanos que genere como resultado el equilibrio de Nash. El ser humano no es racional por lo cual no tomará decisiones en función de la maximización su beneficio.

2.2.1. Pseudocódigo del experimento del juego del ultimátum

Para poder programar el entorno experimental, se debe definir dos jugadores para cada juego, agente 1 y agente 2 de una lista de 2000 agentes. Cada jugador sera emparejado y debe cumplir un rol específico, el jugador 1 es el que propone una división de la dotación de 10 y el jugador 2 evaluar y decidir. Función estrategias: cada jugador define su estrategia. Se requiere que los agentes cumplan varias funciones. Función comparación: el jugador 2 compara las estrategias y toma una decisión. Función pagos: Si acepta el jugador 1 obtiene como pago su estrategia y el jugador 2 obtiene 10 - estrategia jugador 1, si no acepta, ambos obtienen 0. Se repite el juego desde que escogen estrategia 1000 veces.

2.3. Comportamiento sesgado

En el año 2012 la investigación de Alan G. Sanfey⁵ mostró que algunos de los sesgos más concurrentes en el experimento del juego del ultimátum son: envidia, pensamiento justo y altruismo.

Procedimos a crear una nueva variable de salida llamada bienestar. El bienestar depende del sesgo que tenga cada uno de los agentes durante el experimento. El propósito es medir qué sucede con las estrategias, los pagos y el bienestar de cada jugador cuando se comportan exclusivamente de forma altruista, envidiosa o justa. Se compararon estos resultados con agentes no supervisados sin sesgo, que son agentes racionales que cumplen con todos los supuestos económicos para que se de un equilibrio perfecto en subjuegos. Y agentes supervisados sin sesgo, que se refiere a agentes racionales que convergen siempre al equilibrio perfecto en subjuegos.

⁵En la investigación de Alan G. se concluyó que el ser humano es cooperador y que está dispuesto a entregar lo justo al momento del reparto. (Román Carrillo J, 2017)

Los comportamientos de los agentes en nuestro experimento se definen de la siguiente forma: a es el pago del jugador 1 y b es el pago del jugador 2.

Si el jugador es altruista y tiene un pago menor al de su contrincante, su bienestar aumenta. Su bienestar es el pago que recibe más el excedente del otro jugador. Por lo tanto, se define de esta manera:

$$\text{Bienestar} = a + (b - a), \text{ si } a < b$$

Si el jugador es envidioso y tiene un pago menor que el de su contrincante, su bienestar disminuye. Su bienestar es el pago que recibe menos el excedente del otro jugador. Por lo tanto:

$$\text{Bienestar} = a - (b - a), \text{ si } a < b$$

Si el jugador es justo y su pago es menor al del otro jugador, su bienestar disminuye. Si su pago es mayor al del otro, sucede lo mismo. A los jugadores les parece injusto tener un pago diferente al otro. Por lo tanto:

$$\text{Bienestar} = a - (b - a), \text{ si } a < b$$

$$\text{Bienestar} = a - (a - b), \text{ si } a > b$$

A continuación, en las figuras 4, 5, 6 y 7 se detalla cómo funcionarán estos sesgos en el experimento definido anteriormente. Es importante mencionar que, cuando existe un comportamiento sesgado, las respuestas de equilibrio esperadas llegan a un punto que se conoce como el Óptimo de Pareto.⁶

⁶El óptimo de Pareto es un equilibrio que se basa en criterios de utilidad: si algo genera provecho, despertará un proceso natural que permitirá alcanzar un punto óptimo. Vilfredo Pareto fue el economista que planteó este equilibrio para determinar dónde se encontraba el punto con el mayor bienestar social alcanzable dadas ciertas circunstancias.

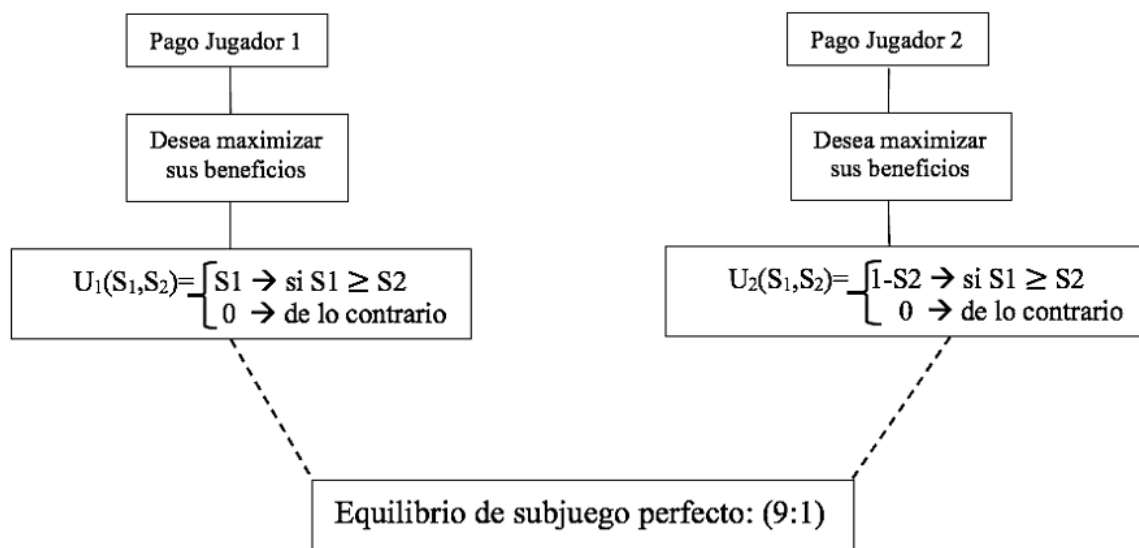


Figura 4: Diagrama Agente Supervisado sin sesgo

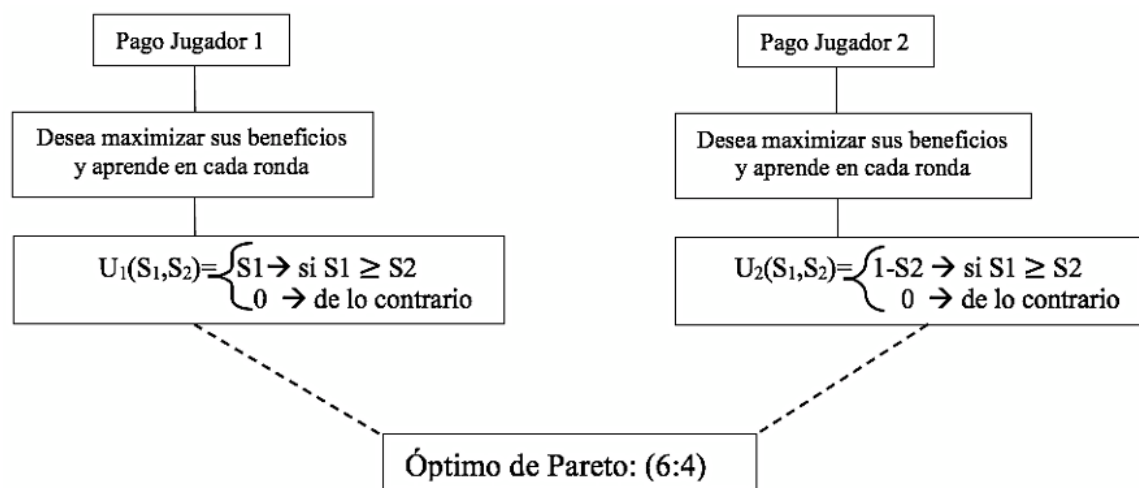


Figura 5: Diagrama Agente no supervisado sin sesgo

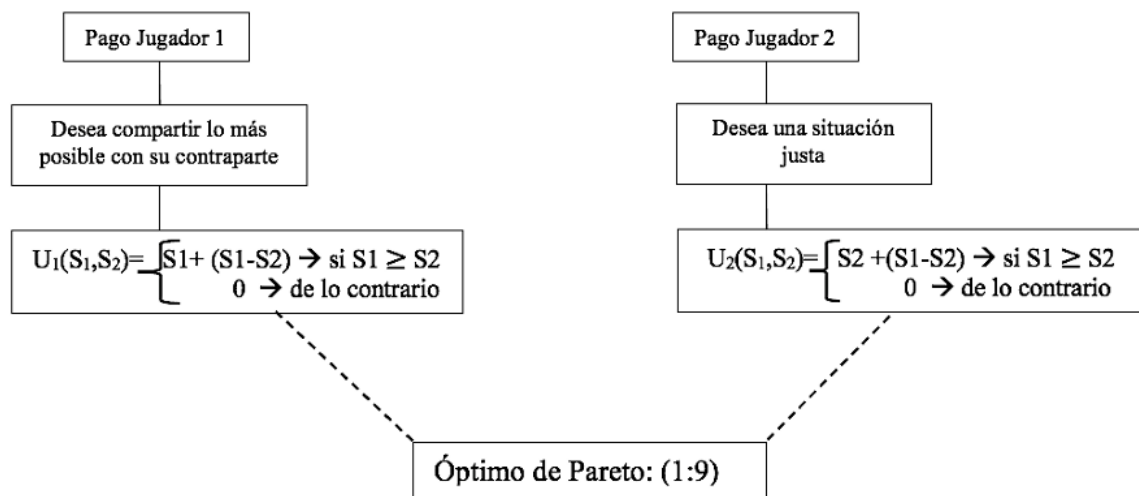


Figura 6: Diagrama Agente Altruista

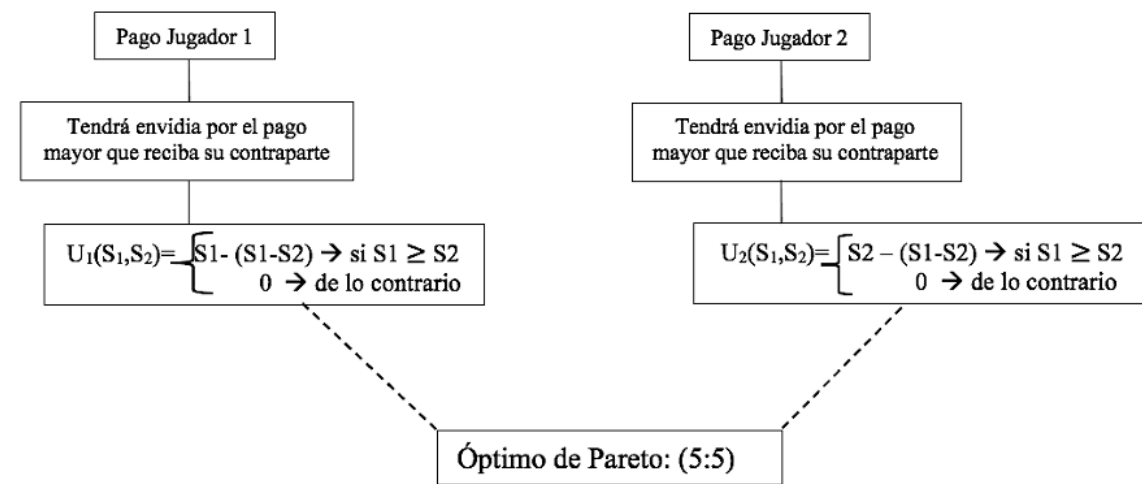


Figura 7: Diagrama Agente envidioso

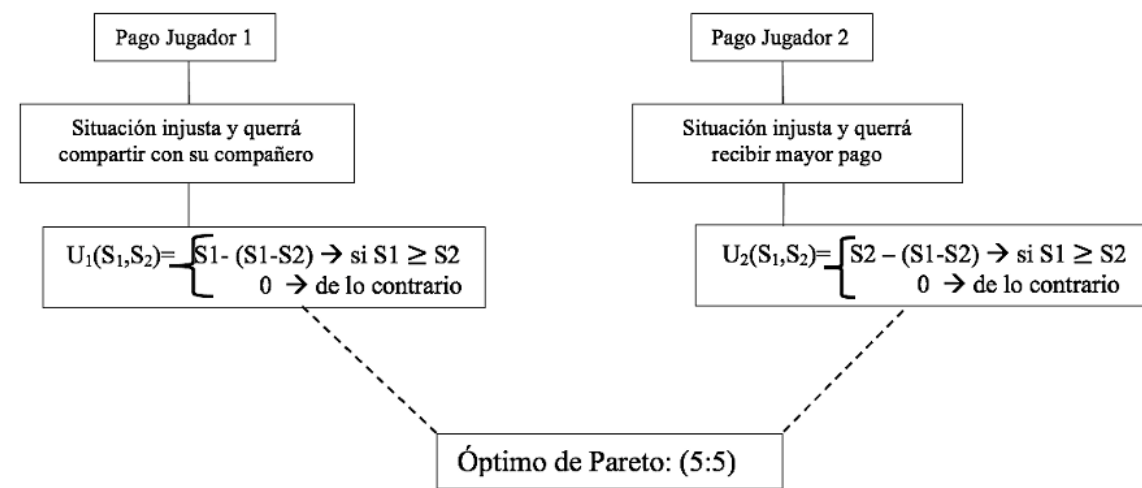


Figura 8: Diagrama Agente Justo

2.4. Inteligencia artificial en el diseño experimental

La inteligencia artificial se define como un conjunto de algoritmos que buscan recrear cierto comportamiento de forma mecánica. En nuestro experimento, utilizamos el procedimiento de Bucket Brigade para generar el entorno experimental. Este permite que los agentes puedan descartar por prueba y error las decisiones que toman porque funciona como una cadena en el algoritmo de programación. Se traslada la información de cada ronda a la siguiente. El agente recibe la información de la estrategia de su anterior ronda y en base a esto construye su estrategia siguiente. Este método tiene cierto factor de aleatoriedad que aparece debido a que en la primera ronda cada jugador escoge un número al azar dentro del intervalo

(1:10). Para estructurar a los jugadores se utilizó programación en base a objetos, el cuál es un método de programación que consiste en la creación de clases que define las características e identidades de cada uno de los agentes.

Se crearon dos tipos de algoritmos totalmente distintos con el fin de extender nuestro análisis. El primer algoritmo siguió el método de Aprendizaje Supervisado, que opera con datos de referencia, también conocidas como las variables de entrada y salida, que el algoritmo debe considerar cómo correctas. Este posee un ajuste, un punto de referencia para que los agentes lleguen al equilibrio de Nash. El segundo algoritmo es el Aprendizaje no Supervisado, que posee variables de entrada pero no de salida, lo que permite que la inteligencia aprenda por su propia cuenta, por acciones anteriores hasta llegar a una convergencia que consideran racional.

Se utilizó aprendizaje supervisado para construir agentes sin sesgo. Estos jugadores saben que el equilibrio de Nash es lo más óptimo y resuelven la forma óptima para llegar a este. También se crearon agentes sin sesgo con aprendizaje no supervisado. Estos agentes desconocen cuál es la forma de maximizar sus beneficios de manera óptima, y solo buscan mejorar su situación actual en base a los pagos que obtienen. Para estos agentes no se asignó una función de ajuste; se esperaría que por ello tomen estrategias gatillo.

El método de aprendizaje no supervisado se utilizó para la creación del resto de agentes, como: altruismo, envidia, pensamiento justo y egoísmo. En este caso, los agentes tienen una variable de ajuste que es su propio sesgo.

3. CONCLUSIONES

3.1. Resultados

3.1.1. Juego del Ultimátum en Agente sin sesgo con aprendizaje supervisado

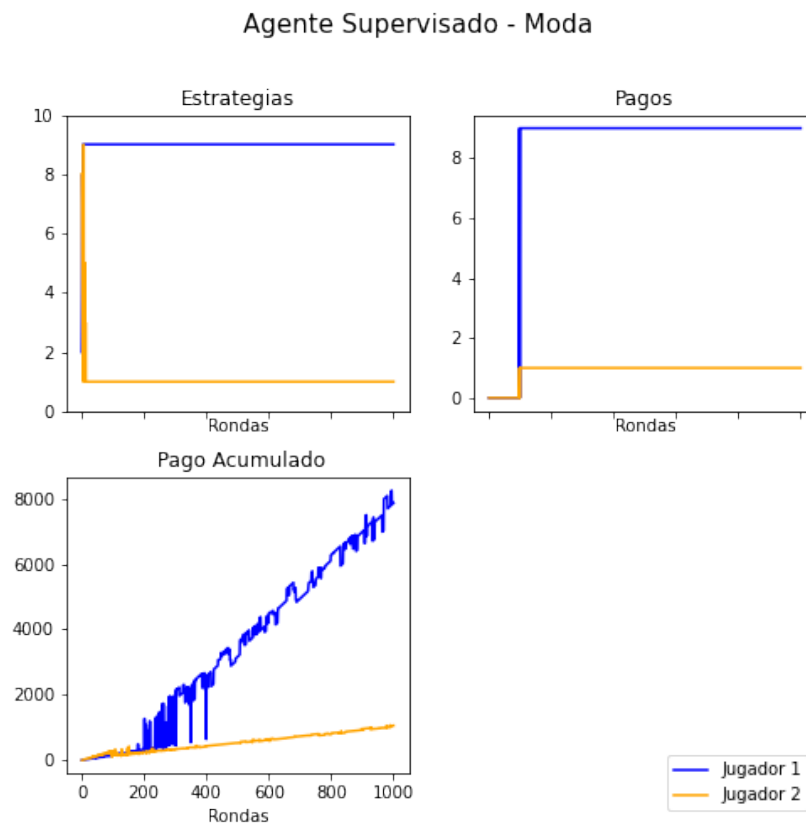


Figura 9: Agente Supervisado sin sesgo 1000 rondas moda

Las estrategias y pagos realizados por el jugador 1 y el jugador 2 convergieron a un equilibrio de subjuego perfecto de (9:1). El jugador 2 tendió a aceptar cualquier oferta mayor a cero, mientras que, el jugador 1 continuó repetidamente ofreciendo lo mínimo posible, tal cual lo estipula la teoría. Aunque existieron rondas en que el jugador 2 rechazó la oferta, esperando un cambio en la estrategia del jugador 1. En este caso, los jugadores terminan ratificando el equilibrio de subjugos perfecto. Se observa en la figura 9.

3.1.2. Juego del Ultimátum en un Agente con aprendizaje no supervisado sin sesgo

A diferencia del diseño experimental anterior, los resultados cambian significativamente. Los agentes no convergen a un equilibrio de subjugos perfecto, sino a un óptimo de Pareto. En esta situación no es posible beneficiar a una persona sin perjudicar a otra. Se llega a un equilibrio que oscila entre (6:4) y (5:5).

Existió un rechazo más concurrente a las ofertas mínimas. Esta simulación se ha repetido múltiples veces y se ha observado que su distribución de frecuencias se acumula, en forma de campana, en 5 y 6⁷. La inteligencia artificial por sí sola, es decir sin supervisión, llega a un equilibrio distinto del de Nash. La figura 10 muestra los resultados:

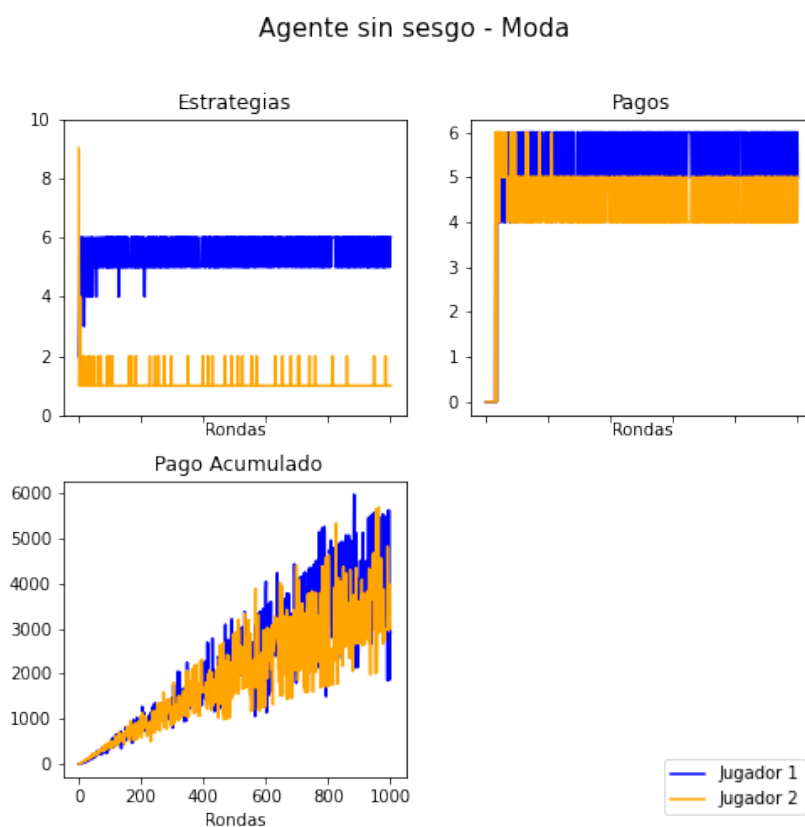


Figura 10: Agente no supervisado sin sesgo 1000 rondas moda

Se dio paso a la aparición de estrategias gatillo. El agente 1 realizó ofertas cada vez

⁷Figura 15

mayores. Hasta encontrar un intervalo de convergencia entre 5 y 6.

3.1.3. Juego del Ultimátum en un Agente Altruista (sesgo) con aprendizaje no supervisado

A diferencia de los dos primeros diseños experimentales, este se caracteriza porque los agentes poseen un sesgo que los condiciona. De acuerdo al algoritmo diseñado, los agentes tendieron a quedarse con lo menos posible para que su contraparte recibiera el mayor beneficio. Como se observa en la figura 11.

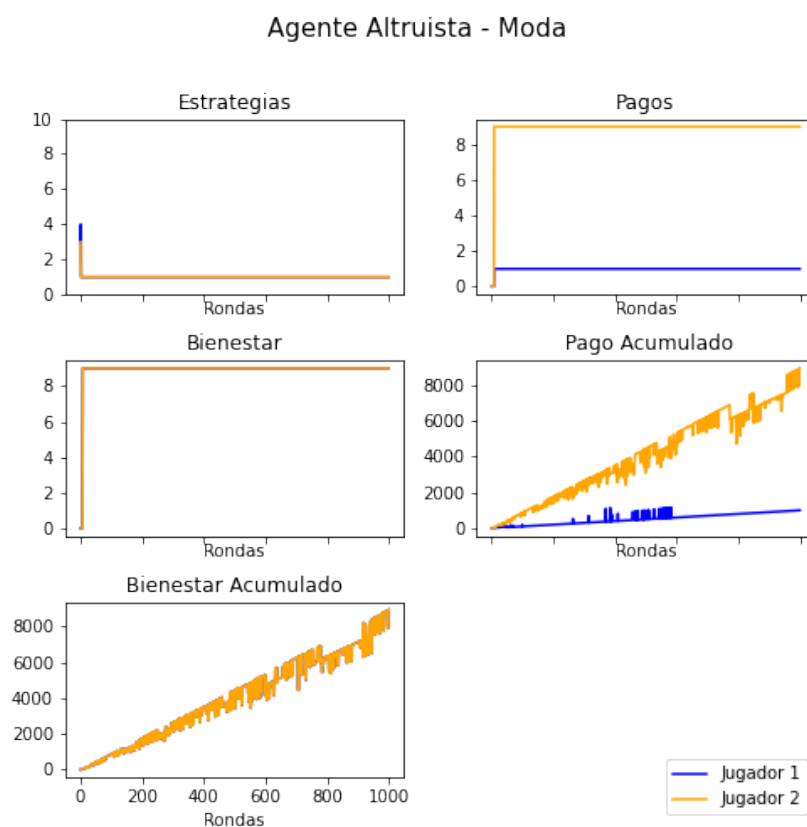


Figura 11: Agente Altruista 1000 rondas moda

En este caso, los resultados convergen a un óptimo de Pareto de (1:9). El agente 2 continuamente rechazó ofertas que considera elevadas por lo que el agente 2, eventualmente terminó entregando ofertas menores. Ser altruista genera un resultado totalmente opuesto al equilibrio de Nash.

3.1.4. Juego del Ultimátum en un Agente envidioso (sesgo) con aprendizaje no supervisado

La posibilidad de mayores recompensas se encontró sesgada debido a la percepción de un ambiente de envidia. Fue más probable generar una estrategia de gatillo entre rondas. La figura 12 muestra los resultados para este caso:

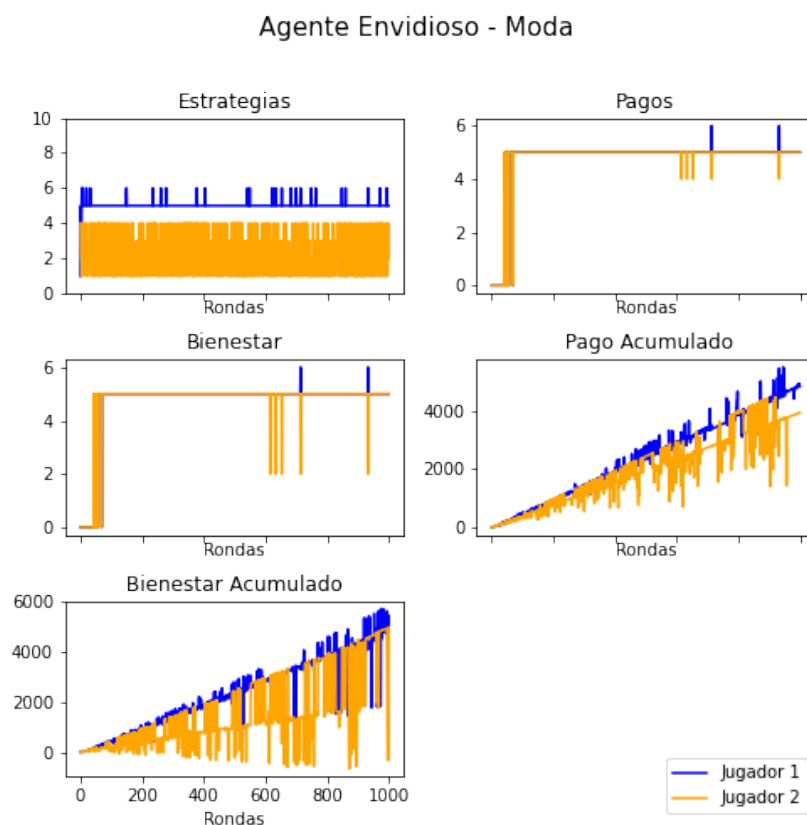


Figura 12: Agente envidioso 1000 rondas moda

Se denota que los agentes convergieron a un óptimo de Pareto de (5:5). El agente 2 continuamente rechazó ofertas que no le producían bienestar, hasta que el agente 1 cambie su estrategia. El agente 1 y el agente 2, eventualmente se conformaron con aquella combinación que no les genere envidia. Llegaron a un punto en el que ninguno sentía envidia del otro.

3.1.5. Juego del Ultimátum en un Agente con pensamiento justo (sesgo) con aprendizaje no supervisado

El pensamiento justo permitió que los jugadores maximicen su bienestar al tener pagos equitativos entre rondas. La figura 13 muestran estos resultados:

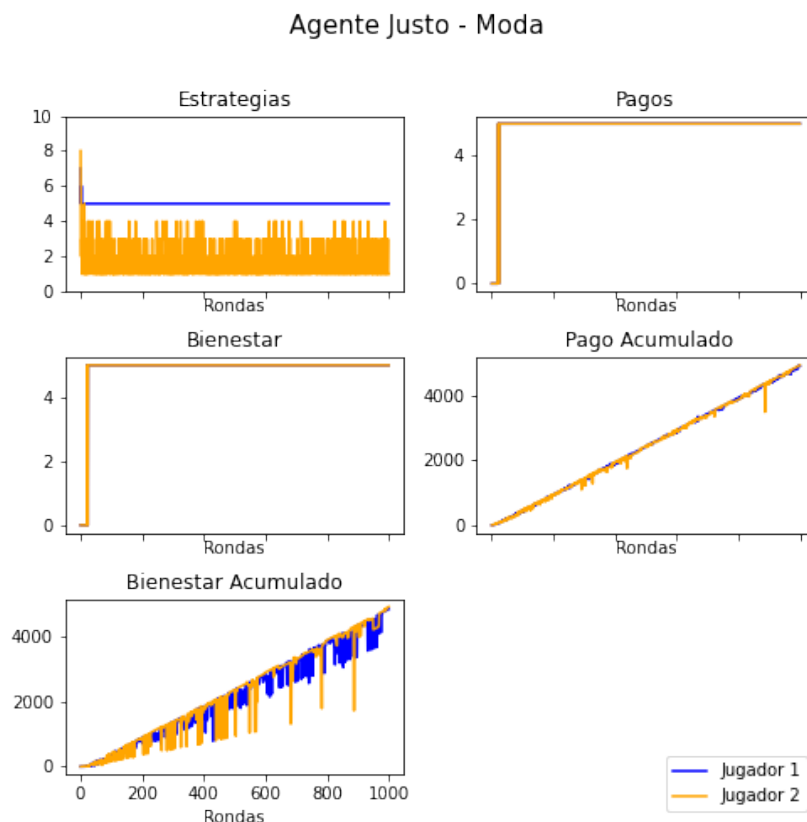


Figura 13: Agente Justo 1000 rondas moda

Llegaron a un óptimo de Pareto de (5:5). El agente 1 y 2 a pesar de no maximizar sus pagos, maximizaron su bienestar cuando sus beneficios fueron iguales. El agente 2 continuamente rechazó oferta mínimas y máximas y el agente 1 entregó ofertas con un pago cada vez más justo.

3.2. Conclusiones finales

Nuestra conclusión más importante, es que, contrario a lo que se esperaba, la inteligencia artificial sin sesgo con aprendizaje no supervisado lleva a los agentes a un equilibrio distinto del de Nash. Como se especifica en el estudio de "Learning To Be Imperfect" Samuelson, Keneth y Gale, esto puede suceder porque los agentes artificiales, como los humanos, aprenden entre rondas. En nuestro experimento, el agente 2 descubre que existen mejores estrategias para maximizar su beneficio, lo que da paso a la aparición de estrategias gatillo que terminan llevando a un equilibrio que oscila entre 5 y 6. Esto demostraría que, la inteligencia artificial por sí sola (sin supervisión) desarrolla estrategias gatillo que llevan a converger a un óptimo de Pareto.

Por otro lado, nuestros resultados demuestran que la inteligencia artificial con sesgos converge rápidamente a un punto de equilibrio. Los agentes artificiales descubren cuál es el punto que maximiza su bienestar dado cierto sesgo. Un caso en particular nos dio un resultado importante: el sesgo de envidia termina llevando a los agentes al mismo equilibrio que el pensamiento justo. Esto sucede debido a que, la combinación de pagos de (5:5) es el único punto en el que los agentes no sienten envidia, por lo que se portan de forma equitativa.

Finalmente, la inteligencia artificial con aprendizaje supervisado converge en cuestión de pocas rondas al equilibrio de Nash (9:1), casi instantáneamente. Esto puede ser explicado con el hecho de que este método se programa con un ajuste que empuja a los agentes al equilibrio de subjugos perfecto. Cómo los agentes saben a que punto llegar, actualizan sus estrategias para alcanzar este punto. Nos acatamos a la publicación de Gale, Samuelson y Keneth al mencionar que el método de aprendizaje de los agentes influye en sus posibles estrategias.

Se podría profundizar este estudio utilizando combinaciones entre distintos tipos de agentes para identificar nuevos comportamientos. También, se podrían agregar grados de sesgo, para medir que pasaría si un agente se comporta, por ejemplo, doblemente envidioso.

Incluso se podría alterar el diseño experimental con un cambio de roles para los agentes a medida que vayan jugando, hacer emparejamientos aleatorios y entre distintos sesgos. Se pueden incluso establecer distintos modelos de equilibrio, macroeconómicos, microeconómicos, etc. con la ayuda de agentes con inteligencia artificial.

4. REFERENCIAS BIBLIOGRÁFICAS

- Aguerrondo, I. (2009). *Conocimiento complejo y competencias educativas*. IBE/UNESCO Working Papers on Curriculum Issues, (8), 7.
- Andersson, O., Galizzi, M., Hoppe, T., Kranz, S., van der Wiel, K. Wengström, E. (2010). *Persuasion in experimental ultimatum games*. *Economics Letters* ,pp 16-18.
- Beshears, J. Gino, F. (2015). *Leaders as decision architects: Structure your organization's work to encourage wise choices*. *Harvard Business Review*, pp 52-62.
- Blink, J., Dorton, I. (2012). *Economics*. Oxford: Oxford University Press.
- Chang, Y.-H., Maheswaran, T. L. (2012). *Decision Making with Imperfect Decision Makers*. The Social Ultimatum Game. En T. V. Guy, M. K. H. Wolpert, pp 136-158.
- De Jong et al (2008). *Learning to Reach Agreement in a Continuous Ultimatum Game*. Obtenido de <https://arxiv.org/pdf/1401.3465.pdf>
- Fernández Álvarez, M. (2014). *Una aplicación de la economía experimental como herramienta didáctica y de investigación: el juego del ultimátum*.
- Gairín, J. (2007). *Competencias para la gestión del conocimiento y el aprendizaje*. Cuadernos de pedagogía, pp 24-27.
- Gale, J., Binmore, K., Samuelson, L. (2021). *Learning to be imperfect: The ultimatum game*.

- Guth, W., Schmittberger, R. Schwarze, B. (1983). *An experimental analysis of ultimatum bargaining*. Journal of Economic Behavior and Organization, pp 367-388.
- Ho, T., Lim, N. Camerer, C. (2006). *Modeling the psychology of consumer and firm behavior with behavioral economics*. Journal of Marketing Research, 43(3), pp 307- 331.
- Kyongsik Yun, Dongil Chung, Dongil Chung, Jaeseung Jeong, Jaeseung Jeong (2007). *Emotional Interactions in Human Decision Making using EEG Hyperscanning*. Obtenido de <https://www.researchgate.net/publication/228466519>
- Le Gléau, T., Marjou, X., Tayeb Lemlouma, T., Benoit Radier, B. (2020). *Multi-agents Ultimatum Game with Reinforcement Learning*
<https://hal.archives-ouvertes.fr/hal-02967163/document>
- Michael S. A. (2004). “*The Ultimatum Game, Fairness, and Cooperation among Big Game Hunters.*” In Henrich, Boyd, Bowles, Camerer, Fehr, and Gintis (Eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*, Oxford University Press, pp. 413-435.
- Román Carrillo, J. (2017). *Cooperación en redes sociales: El juego del ultimátum*
- Rouhiainen, L. (2018). *Inteligencia Aartificial*. Obtenido de <https://static0planetadelibroscom.cdnstatics.com/libroscontenidoextra/40/39308Inteligenciaartificial.pdf>
- Sanfey, Alan G., Rilling, James K., Aronson, Jessica A., Nystrom, Leigh E., Cohen,Jonathan D. (2003). “*The Neural Basis of Economic Decision-Making in the Ultimatum Game*”, pp. 1755 - 1758.
- Sanfey, Alan G.(2007). “*Social Decision-Making: Insights from Game Theory and*

Neuroscience Science”, pp 318-598.

Zhong, F., Kimbrough, S. O., Wu, D. (2002). *Cooperative Agent Systems: Artificial Agents Play the Ultimatum Game. Group Decision and Negotiation.*

5. ANEXOS

5.1. Pseudocódigos

5.1.1. Pseudocódigo para la creación de clase de los agentes

- Definir al agente como un objeto con variable de rol, parámetros iniciales: rol, valor quiere quedarse, valor mínimo a aceptar, vector de estrategias rol 1 [1, 10], vector estrategias rol 2[1, 10], pagos, pago acumulado, bienestar y bienestar acumulado.
- Crear función de escoger estrategia: hacer una elección aleatoria del vector correspondiente al rol.1 jugador 2 compara el valor de su compañero y su mínimo a aceptar.
- Función obtención de pagos: Guarda los pagos obtenidos en variable pagos, y acumula en la variable de pago acumulado.
- Función actualización estrategias: agrega n veces la n estrategia en el vector de estrategias correspondiente al rol, siempre y cuando su pago > 0 .
- Función de sesgo correspondiente. Función obtención de bienestar.

A continuación se detallan los casos específicos donde a es el pago del Jugador 1 y b pago del jugador 2.

5.1.2. Pseudocódigo sin sesgo con aprendizaje supervisado

- Creación de los agentes con los parámetros previos.

- Función de actualización de estrategias agrega n veces la estrategia n al vector estrategias rol 1, siempre y cuando su pago ≥ 9 para el jugador 1; agrega n veces la estrategia n al vector estrategias rol 2 siempre y cuando pago ≤ 1 para el jugador 2.
- No tiene función de sesgo ni función de obtención de bienestar.

5.1.3. Pseudocódigo sin sesgo con aprendizaje no supervisado

- Creación de los agentes con los parámetros previos.
- No tiene función de sesgo ni función de obtención de bienestar.

5.1.4. Pseudocódigo altruismo con aprendizaje no supervisado

- Creación de los agentes con los parámetros previos. Función de sesgo altruista: si, $a < b$ obtiene un bienestar $= a + (b - a)$, si no su bienestar $= a$.
- Función actualización de estrategia: actualiza su estrategia n veces la estrategia n que le de un bienestar > 0 .
- Función obtención de bienestar: guarda su bienestar y acumula en la variable bienestar acumulado.

5.1.5. Pseudocódigo envidia con aprendizaje no supervisado

- Creación de los agentes con los parámetros previos.
- Función sesgo envidia: si, $a < b$ obtiene un bienestar $= a - (b - a)$, si no su bienestar $= a$.

- Función de actualización de las estrategias: actualiza su estrategia n veces la estrategia n que le de un bienestar > 0 .
- Función obtención de bienestar: guarda su bienestar y acumula en la variable bienestar acumulado.

5.1.6. Pseudocódigo pensamiento justo con aprendizaje no supervisado

- Creación de los agentes con los parámetros previos.
- Función sesgo pensamiento justo: si, $a < b$ obtiene un bienestar $= a - (b - a)$, o si, $a > b$ obtiene su bienestar $= a - (a - b)$.
- Si no su bienestar $= a$. Función de actualización de las estrategias: actualiza su estrategia n veces la estrategia n que le de un bienestar > 0 .
- Función obtención de bienestar: guarda su bienestar y acumula en la variable bienestar acumulado.

5.2. Histogramas

Se puede observar los cambios de la distribución conforme iban aumentando la rondas. La acumulación de frecuencias nos indican las estrategias que fueron las mas elegidas por los jugadores en la primera, 500 y 1000 rondas.

En la figura 14 existe 1000 rondas y se observa la clara preferencia de 1 para el jugador 2 y 9 para el jugador 1.

En la figura 15 existe 500 rondas y se observa la forma de campana con centro en 5 y 6 para ambos jugadores. Lo mismo sucede en las 1000 rondas.

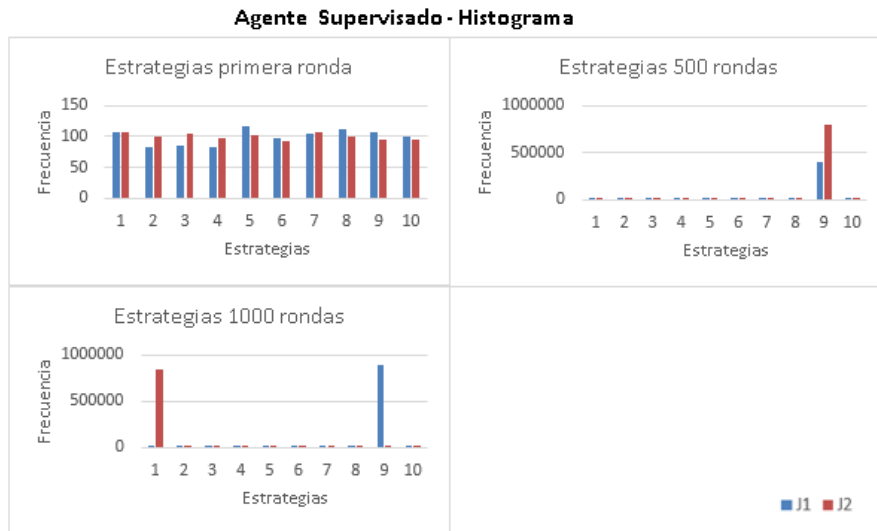


Figura 14: Agente Supervisado histograma

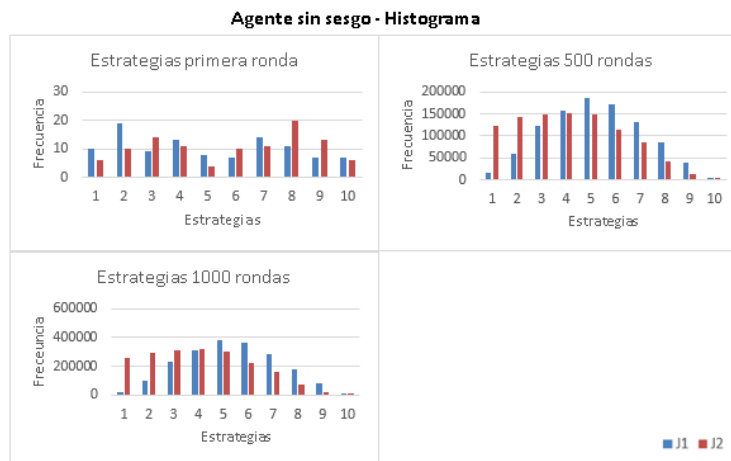


Figura 15: Agente sin sesgo histograma

En la figura 16 existe 500 rondas y se observa una tendencia por preferir estrategias cercanas a 1 para los dos jugadores. Lo mismo sucede en las 1000 rondas, incrementando la tendencia aun más.

En la figura 17 existe 500 rondas y se puede observar como se prefiere para el jugador 1 la estrategia 5 y mayores, pero para el jugador 2 menores a 5 y 6. En las 1000 rondas se incrementa mas este patrón.

En la figura 18 existe 500 rondas y se observa como el jugador 1 tiene una preferencia

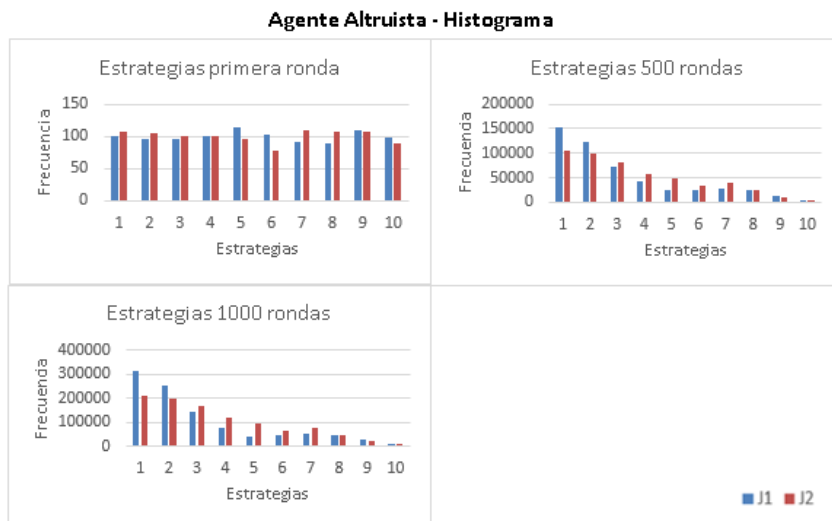


Figura 16: Agente Altruista histograma

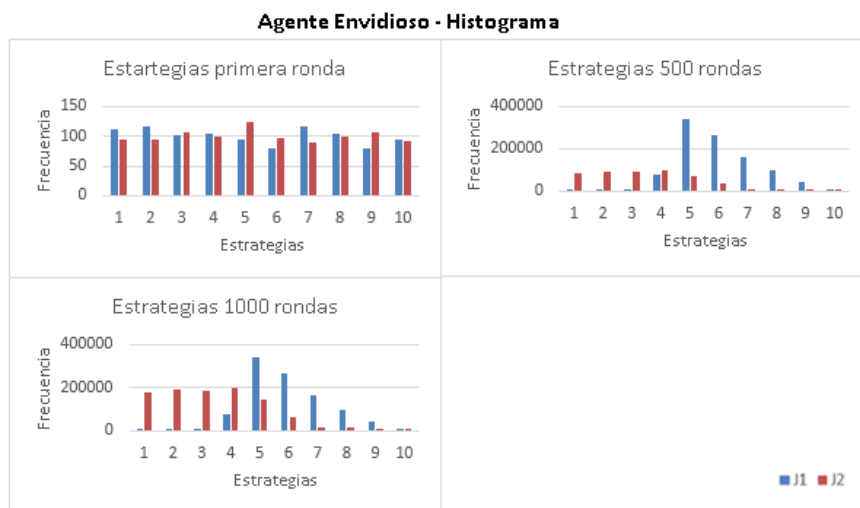


Figura 17: Agente Envidioso histograma

muy alta por la estrategia 5, y el jugador 2 por estrategias menores a 5. En las 1000 rondas esta tendencia se mantiene.

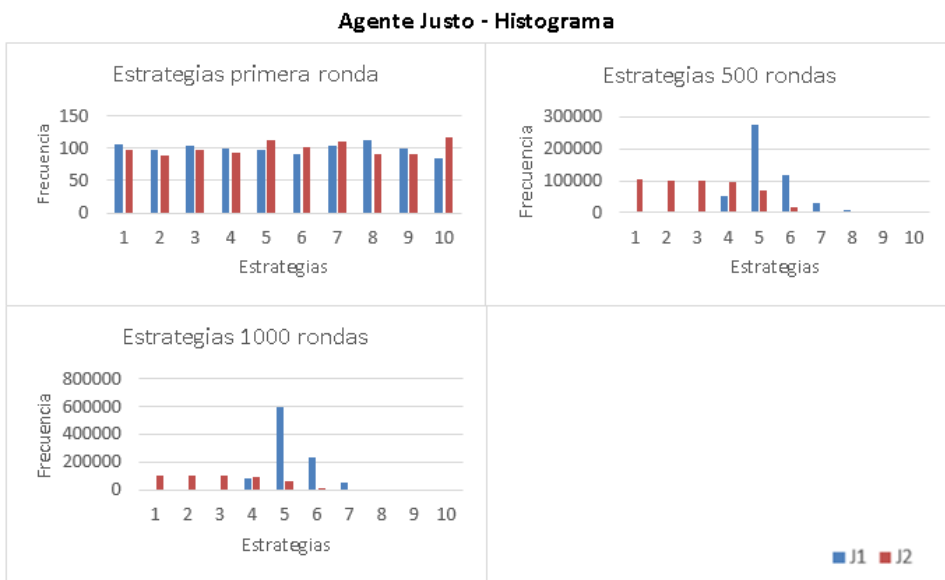


Figura 18: Agente Justo histograma

5.3. Resultados del diseño experimental con 100 rondas

Podemos observar gráficamente a que converge la moda y mediana de cada una de las distintas simulaciones. En distintos puntos de las rondas, para observar mejor el comportamiento de cada jugador.

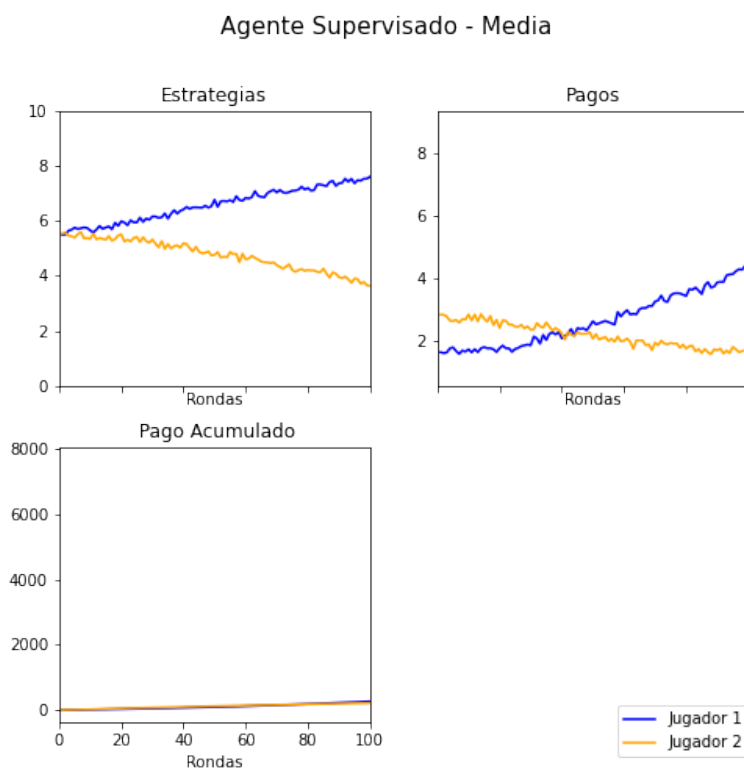


Figura 19: Agente Supervisado 100 rondas media

Agente Supervisado - Moda

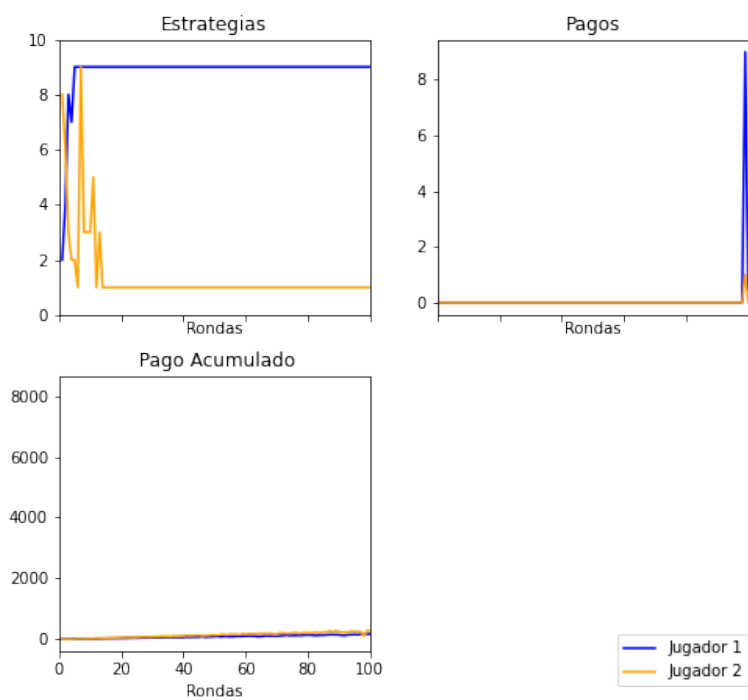


Figura 20: Agente Supervisado 100 rondas moda

Agente sin sesgo - Media

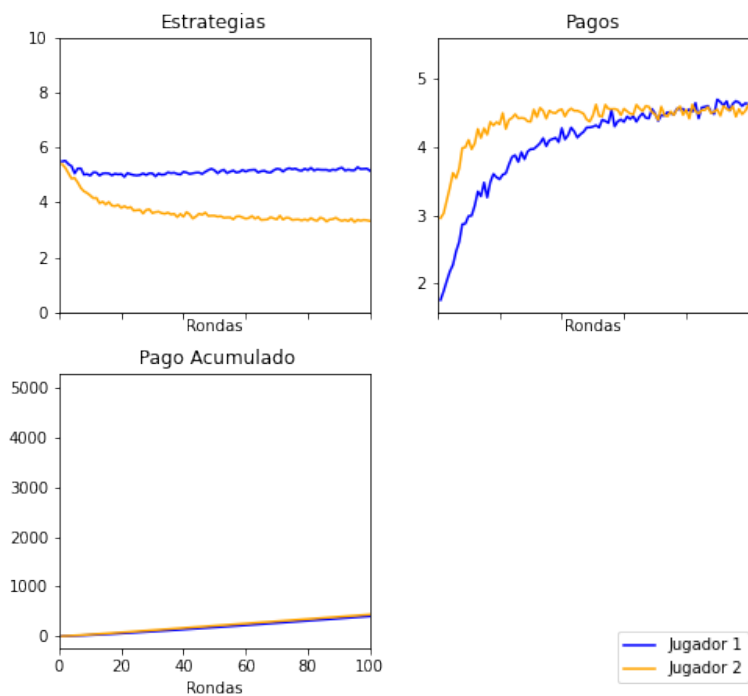


Figura 21: Agente sin sesgo 100 rondas media

Agente sin sesgo - Moda

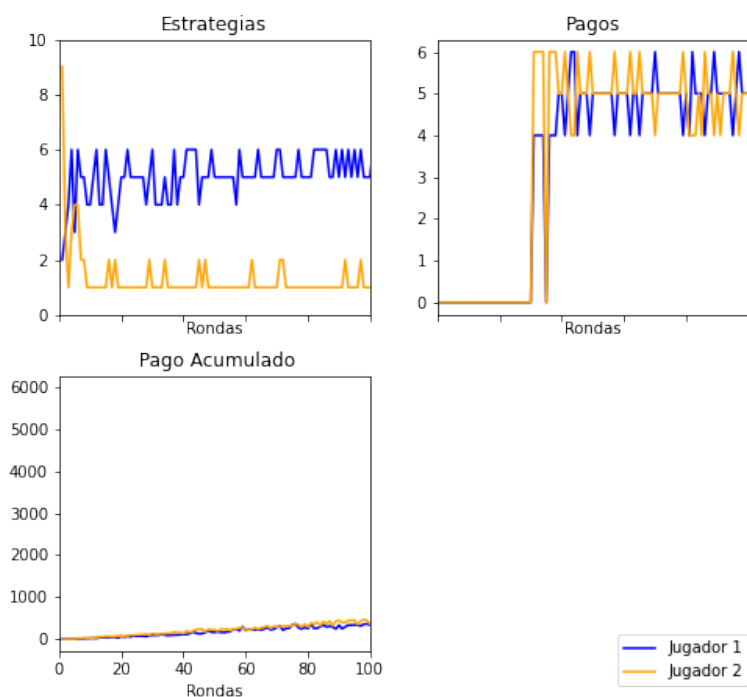


Figura 22: Agente sin sesgo 100 rondas moda

Agente Altruista - Media

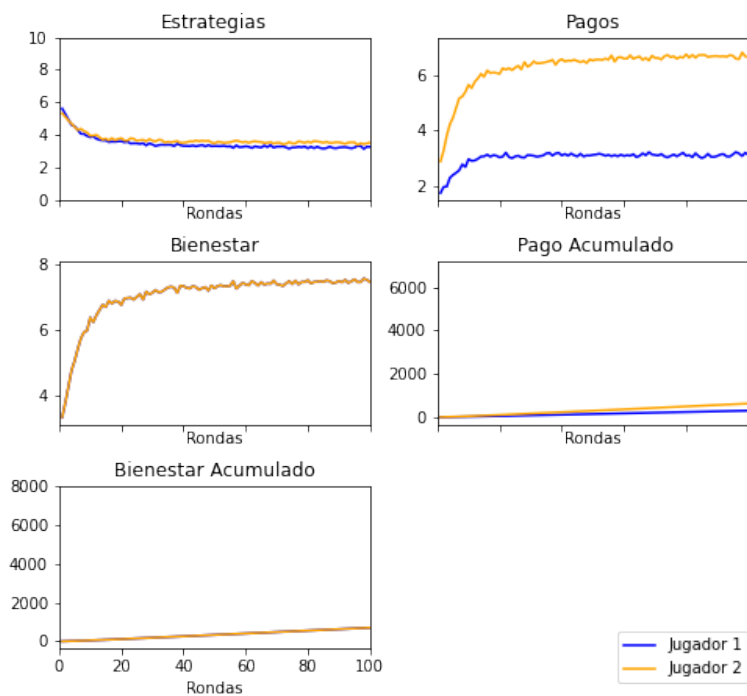


Figura 23: Agente Altruista 100 rondas media

Agente Altruista - Moda

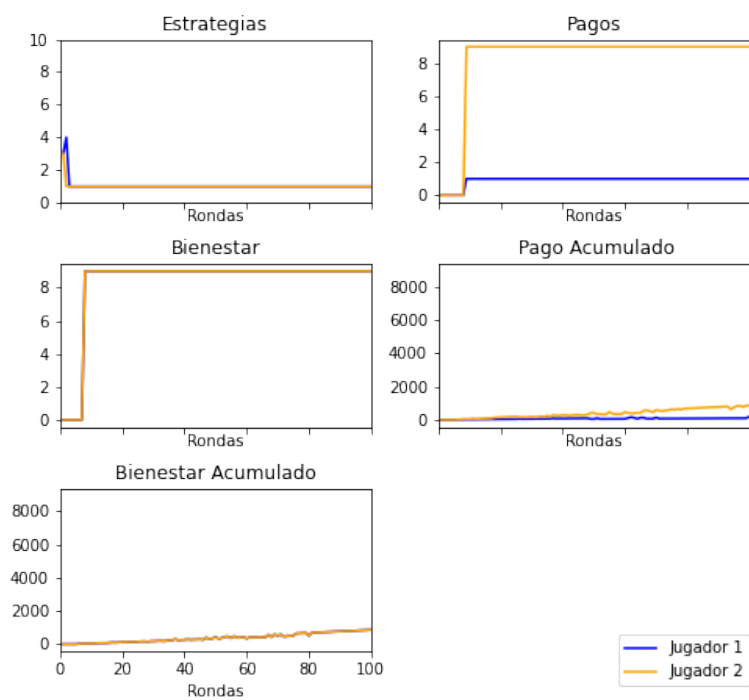


Figura 24: Agente Altruista 100 rondas moda

Agente Envidioso - Media

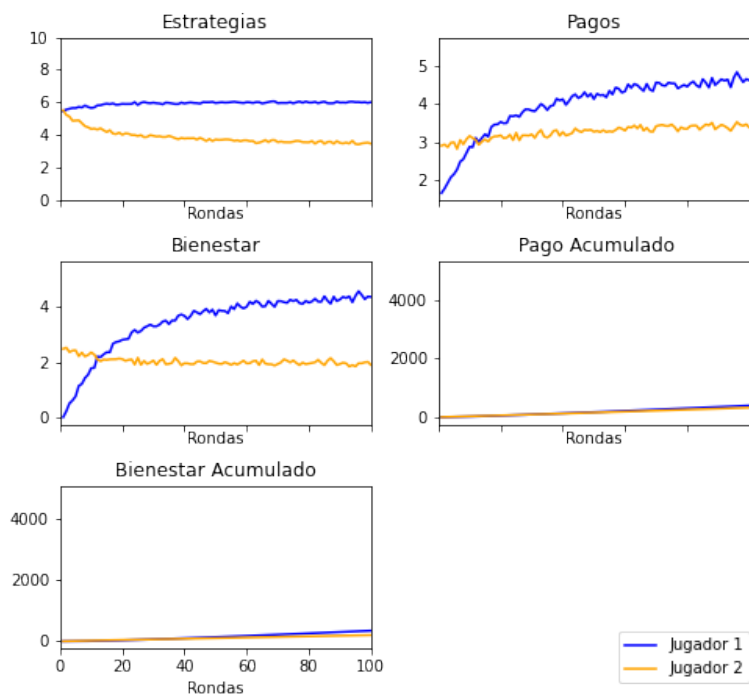


Figura 25: Agente Envidia 100 rondas media

Agente Envidioso - Moda

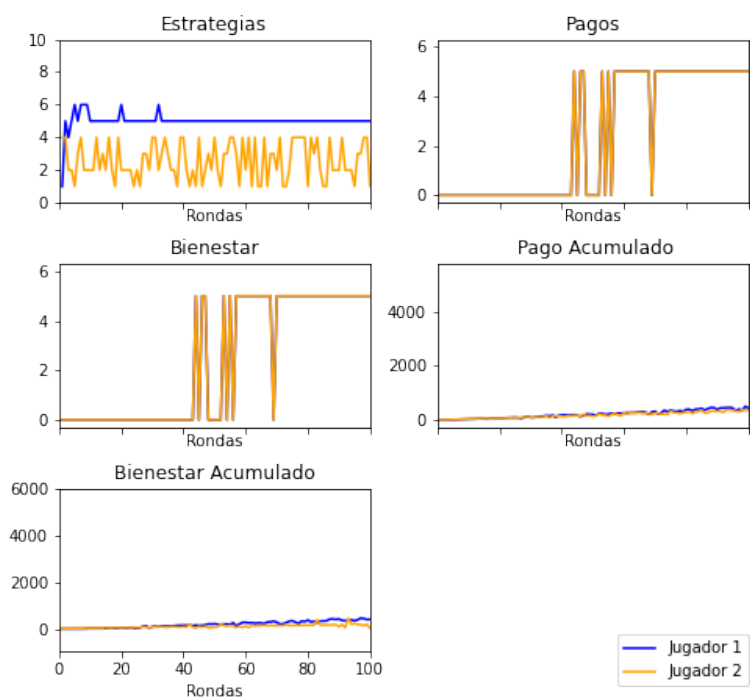


Figura 26: Agente Envidia 100 rondas moda

Agente Justo - Media

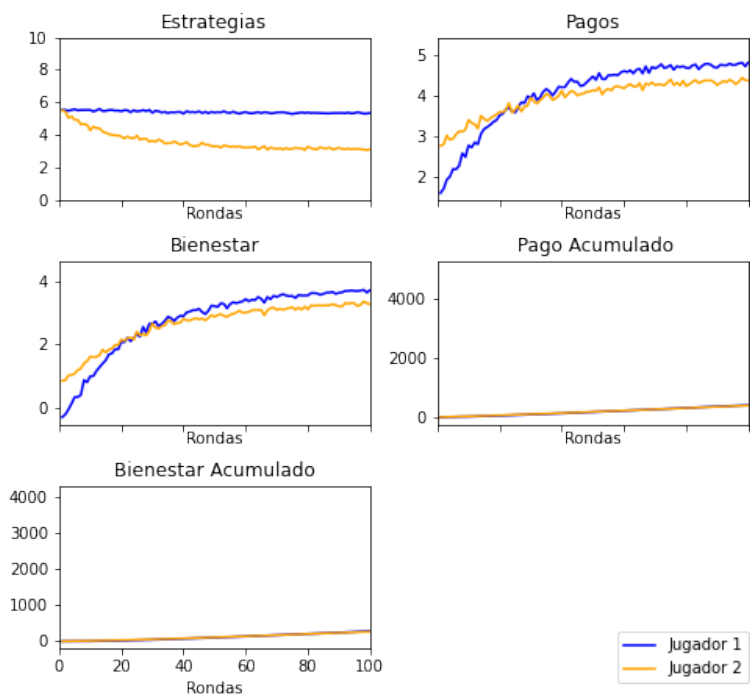


Figura 27: Agente Justo 100 rondas media



Figura 28: Agente Justo 100 rondas moda

5.4. Resolución del Diseño experimental con 500 rondas

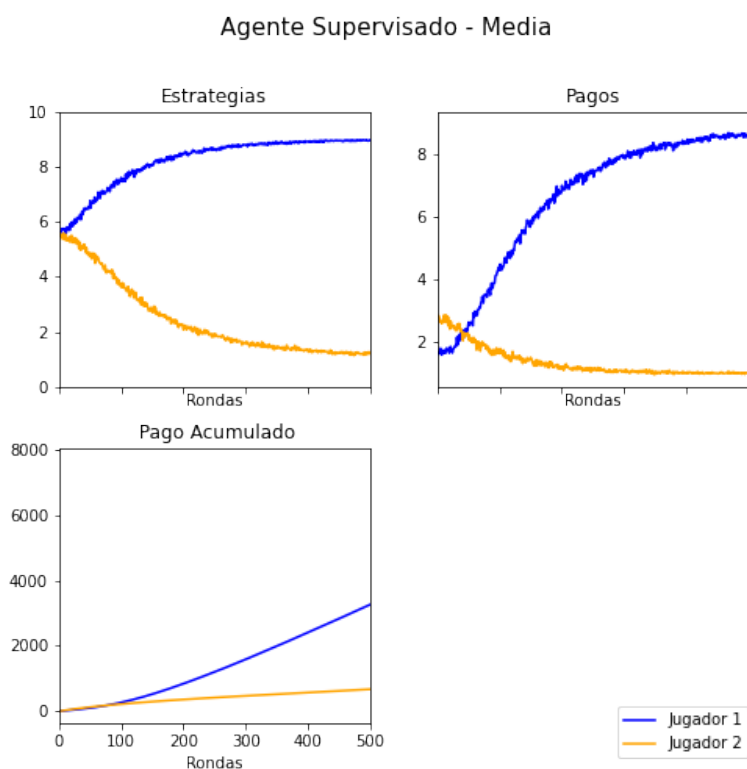


Figura 29: Agente Supervisado 500 rondas media

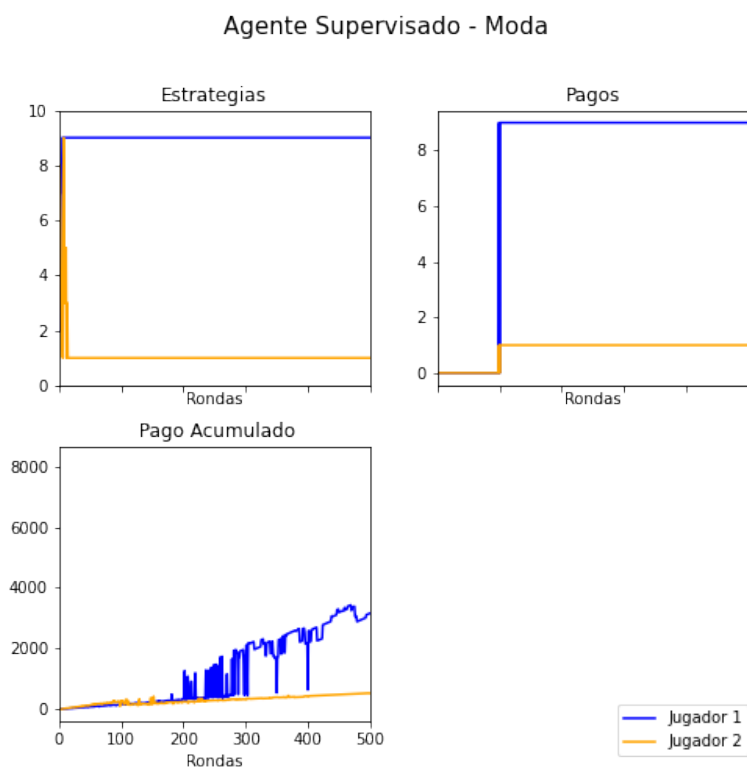


Figura 30: Agente Supervisado 500 rondas moda

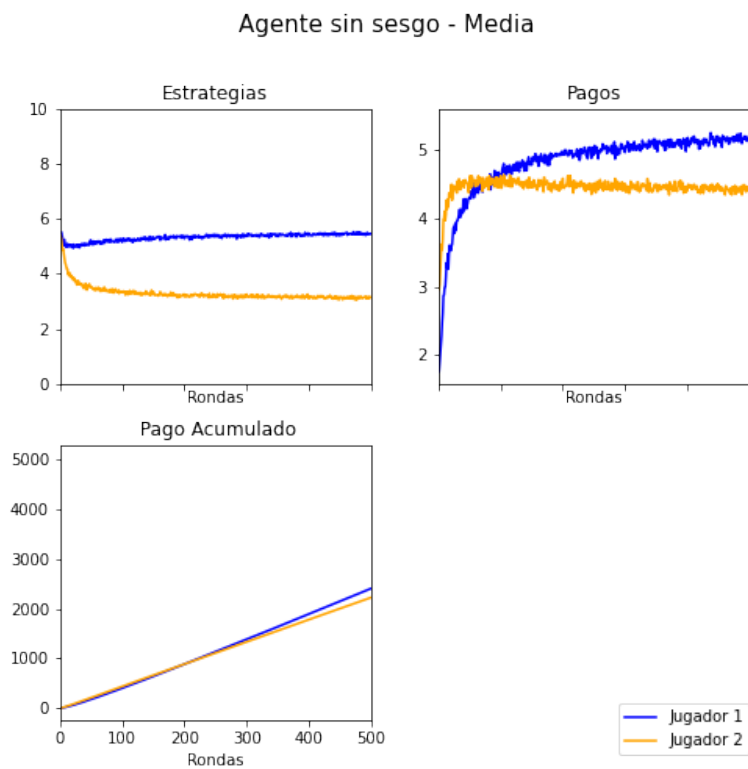


Figura 31: Agente sin sesgo 500 rondas media



Figura 32: Agente sin sesgo 500 rondas moda

Agente Altruista - Media

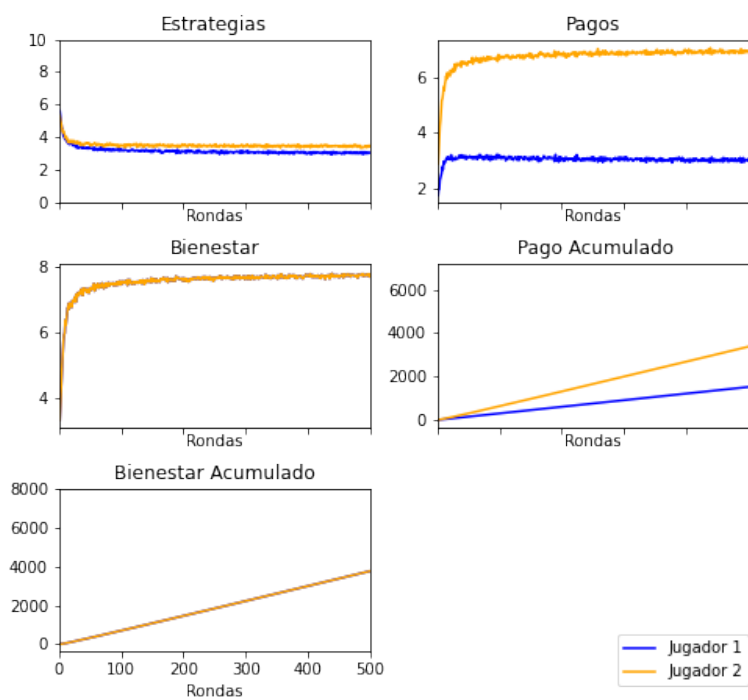


Figura 33: Agente Altruista 500 rondas media

Agente Altruista - Moda

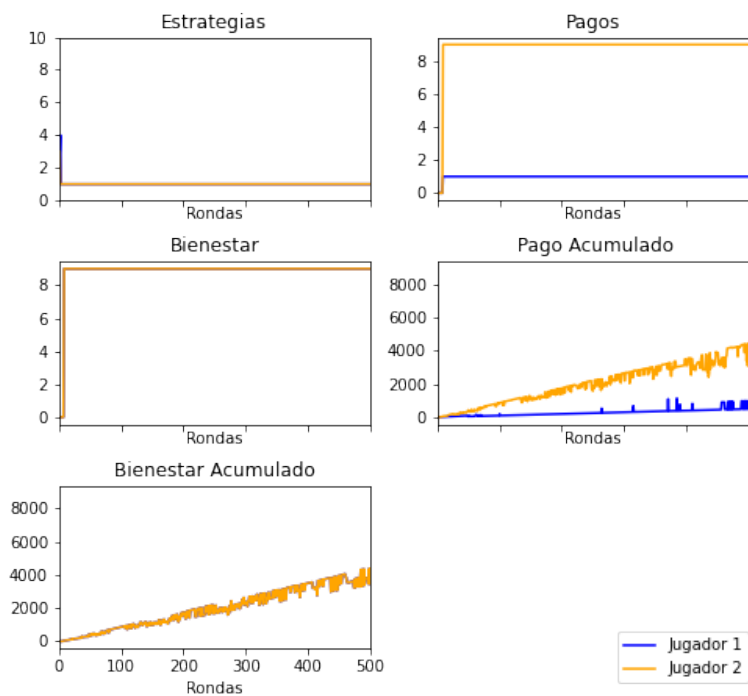


Figura 34: Agente sin sesgo 500 rondas moda

Agente Envidioso - Media

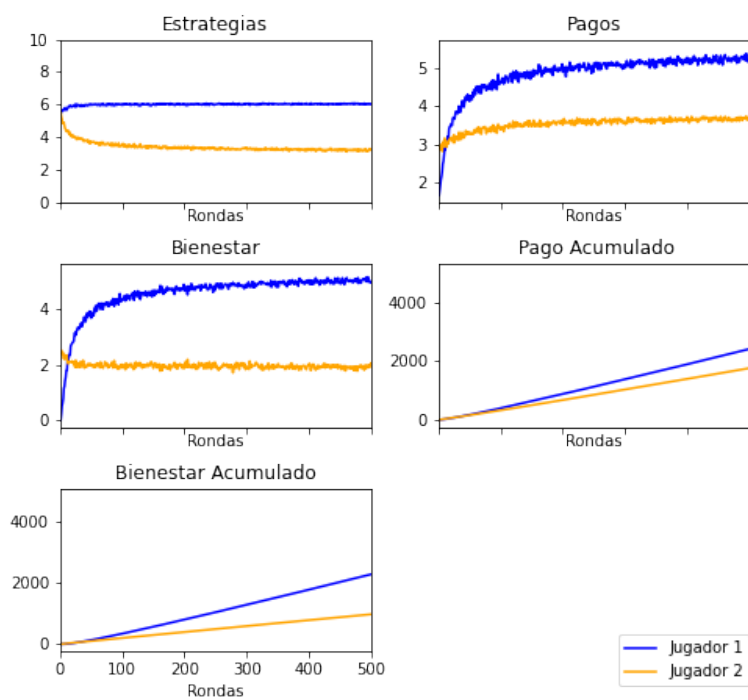


Figura 35: Agente Envidia 500 rondas media

Agente Envidioso - Moda

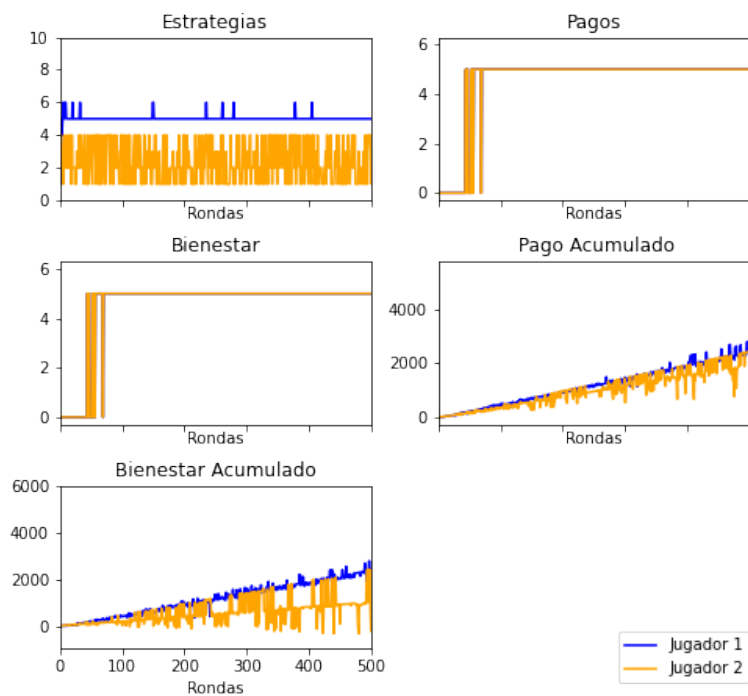


Figura 36: Agente Envidia 500 rondas moda

Agente Justo - Media

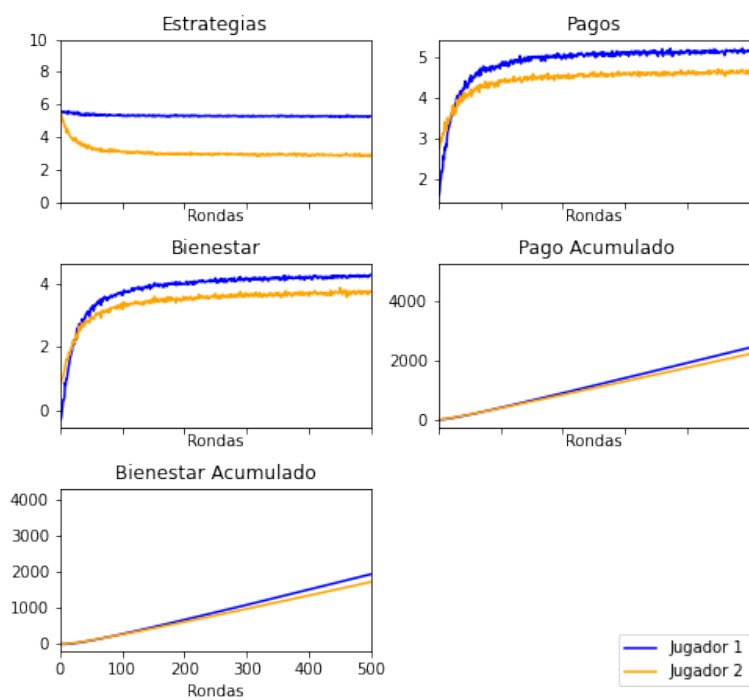


Figura 37: Agente Justo 500 rondas media

Agente Justo - Moda

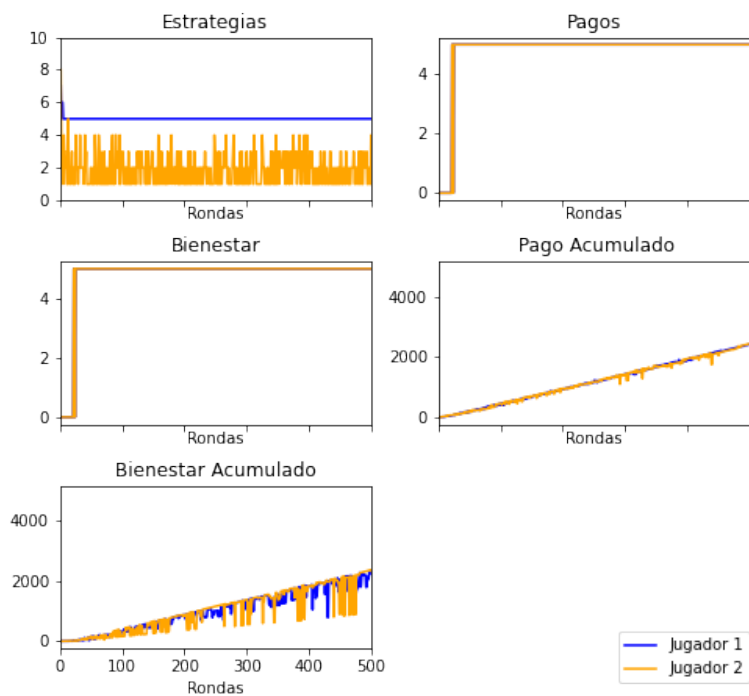


Figura 38: Agente Justo 500 rondas moda

5.5. Media Diseño experimental 1000 rondas

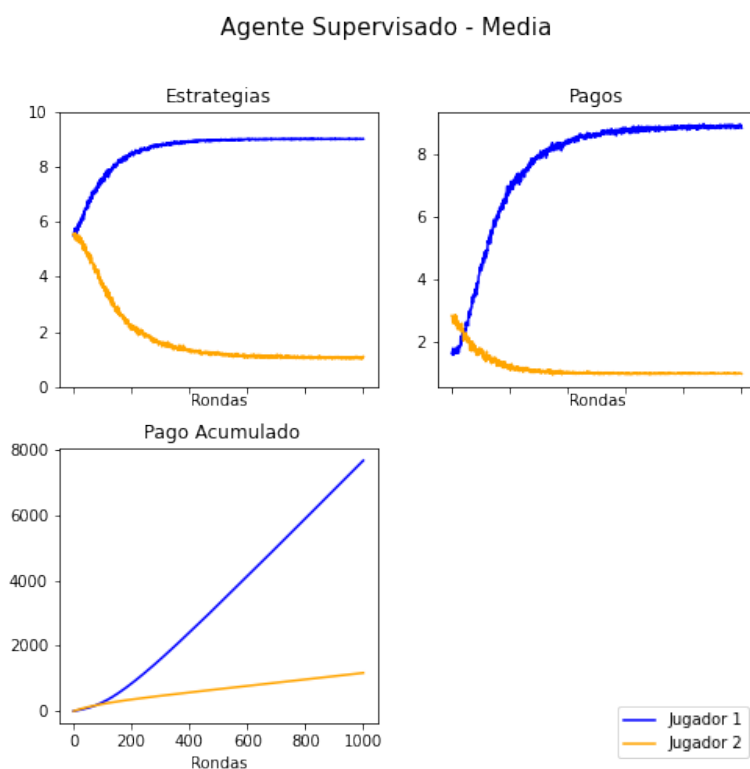


Figura 39: Agente Supervisado 1000 rondas media

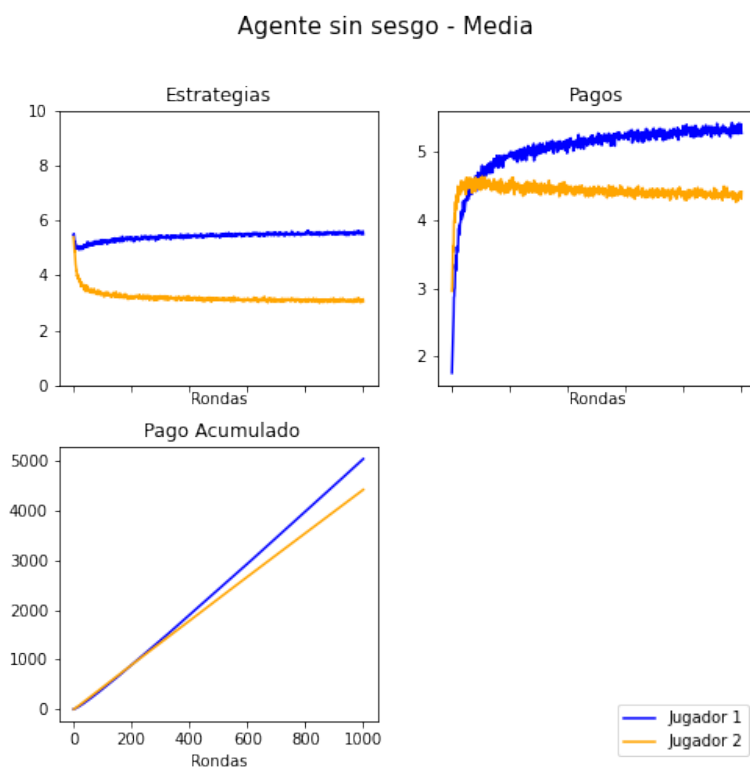


Figura 40: Agente sin sesgo 1000 rondas media

Agente Altruista - Media

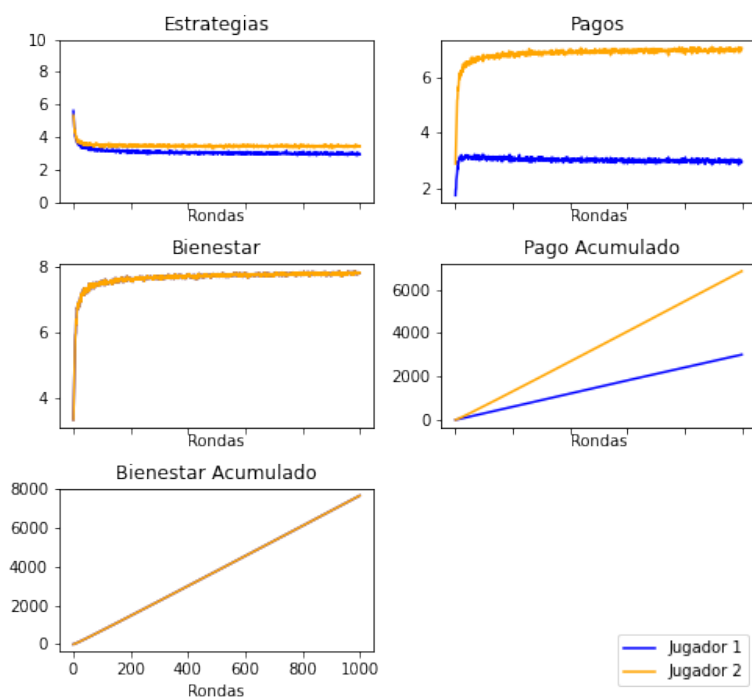


Figura 41: Agente Altruista 1000 rondas media

Agente Envidioso - Media

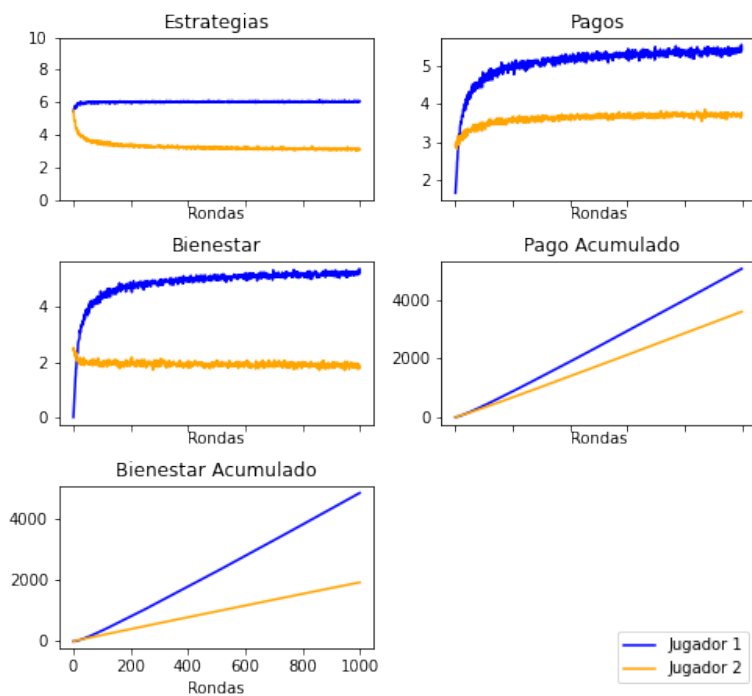


Figura 42: Agente Envidioso 1000 rondas media

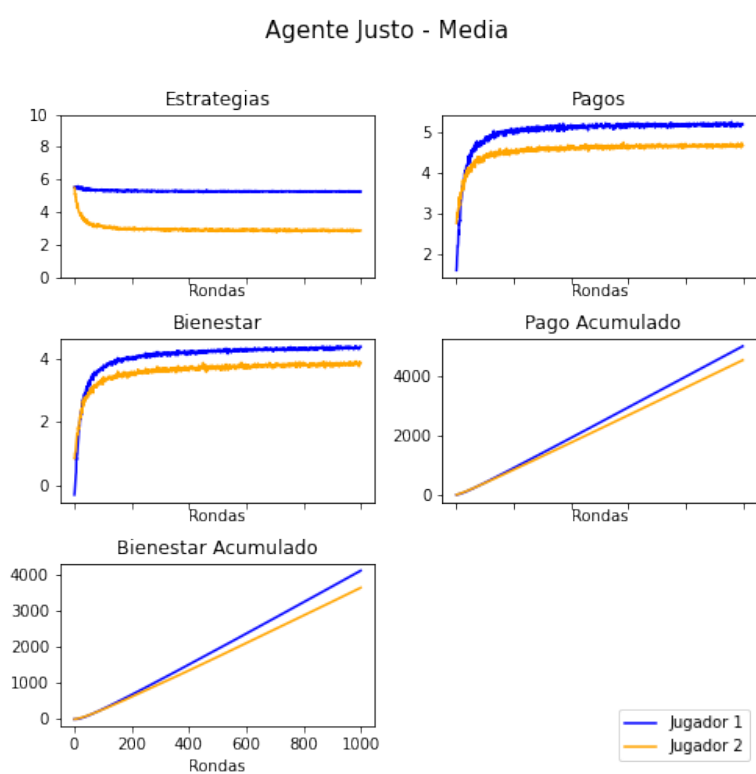


Figura 43: Agente Justo 1000 rondas media