



**Universidad del Desarrollo**  
Facultad de Ingeniería

**LA IMPORTANCIA DE LA ESTRATEGIA EN VALORANT**  
Análisis predictivo de los resultados de cada ronda en base a los distintos componentes de la estrategia de los equipos rivales.

**POR: DANIEL ALEJANDRO SMITH DE LA CARRERA**

Capstone project presentado a la Facultad de Ingeniería de la Universidad del Desarrollo para optar al grado académico de Magíster en Data Science

**PROFESOR GUÍA:**

Dr. Cristián Candia, Sr. Alonso Astroza

Diciembre 2022  
SANTIAGO



## AGRADECIMIENTO

Antes de empezar, quiero agradecer la colaboración de los profesores Alonzo Astroza y Cristián Candía quienes me apoyaron en reiteradas ocasiones dándome retroalimentación para poder dirigir este proyecto. Agradezco a mi hermano Javier Smith quien me apoyó en entendimientos sobre redes neuronales complejas, y a todos los profesores del magister de Data Science UDD de quienes aprendí mucho y me permitieron obtener las herramientas y habilidades para el proyecto.

## TABLA DE CONTENIDO

<b>RESUMEN</b> .....	<b>1</b>
<b>1. INTRODUCCIÓN</b> .....	<b>2</b>
<b>2. TRABAJO RELACIONADO</b> .....	<b>¡ERROR! MARCADOR NO DEFINIDO.</b>
<b>3. HIPÓTESIS Y OBJETIVOS</b> .....	<b>4</b>
<b>4. DATOS Y METODOLOGÍA</b> .....	<b>5</b>
4.1. DATOS.....	5
4.2. METODOLOGÍA .....	6
<b>5. RESULTADOS</b> .....	<b>8</b>
<b>6. CONCLUSIÓN Y TRABAJO FUTURO</b> .....	<b>20</b>
<b>BIBLIOGRAFÍA</b> .....	<b>20</b>

## **Resumen**

En este documento analizaremos una base de datos que contiene información de múltiples partidas y rondas de Valorant para realizar un análisis predictivo que logre identificar si un equipo ganará o perderá una determinada ronda en base a la estrategia del equipo contra la estrategia del equipo rival, entre las cuales están la cooperación del equipo (medida con asistencias previas), las armas que utilizarán en la ronda, las decisiones del sitio para plantar la bomba pasadas y como esto influye en el desempeño de la ronda actual y la composición de agentes y roles de cada equipo.

# 1. Introducción

Valorant es un videojuego de disparos competitivo de enfrentamientos 5 contra 5 en un mapa limitado (1).

- Las partidas tienen una duración base de 24 rondas, donde el primero que llegue a las 13 victorias será el equipo ganador de la partida.
- Si se han jugado las 24 rondas y el marcador indica un empate, se juegan 2 rondas adicionales. Si se vuelve a empatar en esas 2 rondas siguen con otras 2 rondas hasta que la diferencia de victorias sea por 2 rondas.
- Durante la partida existirá un equipo atacante y otro defensor, a la mitad de la partida (ronda 13) los roles se invierten.
- Objetivo atacantes: Plantar la bomba y evitar que el equipo defensor la desactive hasta que esta explote.
- Objetivo defensores: Evitar que el equipo atacante plante la bomba y desactivarla antes de que explote en caso de que la planten.
- Existen 19 agentes con distintas habilidades y roles que ayudarán al equipo a idear estrategias y enfrentar los combates.
- Existen 7 mapas jugables. El mapa elegido se decidirá de forma aleatoria al comienzo de la partida.
- Cada partida tiene 2 o 3 sites separados, donde se podrá plantar la bomba.

**Agentes:** Valorant tiene una selección de 19 agentes en los que cada uno tiene su rol específico:

- **Entry Fragger/Duelista:** Su función es entrar primero al site y asesinar a los primeros enemigos para que el resto del equipo pueda entrar tranquilamente.
- **Iniciadores:** Su función es obtener información ya sea para saber en que site el equipo atacante entrará como para saber en que lugar del site está posicionado el equipo defensor. Asimismo tienen la función de ayudar al entry fragger con sus habilidades para que tenga la ventaja en el combate.
- **Centinelas:** Su función es ganar información si el equipo defensor va por la espalda del equipo atacante y dificultarle la entrada al site al equipo atacante con sus habilidades.
- **Controlador/Humos:** Su función es utilizar sus humos para cubrir zonas cuando se ataca, así el entry fragger podrá entrar de manera más segura sin que le disparen de una zona que no está viendo al momento de entrar. Asimismo, en defensa tiene la función de usar estos humos para dificultarle la entrada al equipo atacante.

**Armas:** Existen un total de 17 armas que se pueden comprar al inicio de cada ronda. Cada arma tiene un precio distinto. Asimismo, si un jugador enemigo está muerto, cualquiera puede recoger su arma, lo que es beneficioso para la economía del equipo.

- **Armas de mano:** Armas de poco daño pero económicas, son ideales para el comienzo de las partidas y para rondas de economía.

- **Sub-Rifles:** Armas de daño y precio medio. Generan menos daño a objetivos que se encuentran lejos.
- **Rifles:** Armas de daño y precio alto. En la mayoría de los casos son capaces de matar de 1 disparo en la cabeza.
- **Rifles de francotirador:** Rifles que tienen una mira especial que permite hacer zoom a la visión del jugador, permitiendo disparar de muy largas distancias con mejor precisión.
- **Pesadas:** Armas de daño moderado que se caracterizan por su gran cantidad de balas antes de necesitar recargar.
- **Escopetas:** Armas que permiten disparar múltiples balas en un sector de la mira. Dado que cada bala hace muy poco daño pero en conjunto hacen mucho daño, son muy malas para enfrentamientos a distancia. Se utilizan para enfrentamientos a corta distancia.

### **Insights de Investigación secundaria**

- Los jugadores y entrenadores analizan cada ronda de la partida para adaptar su estrategia para conseguir la victoria (1).
- Los jugadores y entrenadores analizan las partidas del próximo equipo rival y su modo de juego para idear una estrategia ganadora en base a la estrategia predominante de este para así idear una estrategia vencedora (1).
- Las estrategias de cada equipo en los MMOFPS siguen un patrón pero pueden variar dependiendo de las acciones del equipo contrincante (1).
- En los MMOFPS como el CS GO y el Valorant, se emparejan a los jugadores según su ELO (MMR) (2).
- El ordenador del jugador puede influir en la partida. Un jugador con un ordenador que le permita correr Valorant a 144fps puede tener cierta ventaja en reacciones frente a otro que le permita correr Valorant a 30fps (3).
- Muchos jugadores optan por entrenar su AIM para mejorar su precisión y reflejos a la hora de la partida. Una plataforma muy utilizada para esto es AIM Lab, que cuenta con herramientas para mejorar el AIM (4).

## Hipótesis y Objetivos

**Hipótesis:** Dado que en Valorant emparejan a los jugadores con otros de mismo ELO, al comenzar la partida solo se conocen los nombres de los jugadores (salvo excepciones), la composición del equipo, el mapa en el cual están jugando, entonces probabilidad de victoria del equipo en cada ronda será determinada por la estrategia de este por sobre la habilidad de los jugadores, donde los factores que la determinan son: Posicionamiento inicial, Distribución de entradas a un site (atacantes), uso de habilidades en determinados momentos y en determinadas zonas, movimientos de los jugadores en base al movimiento conocido de los atacantes, juego de la economía del equipo, decisión del sitio de plantado de la bomba, armas elegidas, juego en equipo (asistencias), entre otras.

**Objetivo general:** Determinar un modelo que permita predecir el resultado de una ronda de Valorant utilizando solo componentes estratégico con una precisión sobre 75%.

### Objetivos específicos:

- Unir y limpiar los conjuntos de datos de las 10 tablas.
- Determinar las variables que se utilizarán en el modelo, las cuales corresponderán a variables conocidas antes de que inicie la ronda.
- Adaptar las variables de modo que generen la mejor predicción posible.
- Generar un modelo de predicción del resultado de una ronda de Valorant.
- Evaluar el modelo y concluir.

## 2. Datos y Metodología

### 2.1. Datos

Columnas	Descripción	Tipo de valor
avg_previous_assists	Contiene el promedio acumulado de asistencias anteriores a lo largo de la partida. La ronda 1 parte con 0.0.	Float
weapon_dummies	Contiene un conteo de todas las armas que se utilizarán en la ronda. Cada columna representa un arma.	Dummies (1-0)
agent_dummies	Contiene un conteo de todos los agentes que se utilizarán en la ronda. Cada columna representa un agente.	Dummies (1-0)
role_dummies	Contiene un conteo de todos los roles que se utilizarán en la ronda. Cada columna representa un rol.	Dummies (1-0)
enemy_avg_previous_assists	Contiene el promedio acumulado de asistencias anteriores del equipo enemigo a lo largo de la partida. La ronda 1 parte con 0.0.	Float
enemy_weapon_dummies	Contiene un conteo de todas las armas que el equipo enemigo utilizará en la ronda. Cada columna representa un arma.	Dummies (1-0)
enemy_agent_dummies	Contiene un conteo de todas los agentes que el equipo enemigo utilizará en la ronda. Cada columna representa un agente.	Dummies (1-0)
enemy_role_dummies	Contiene un conteo de todos los roles que el equipo enemigo utilizará en la ronda. Cada columna representa un rol.	Dummies (1-0)
avg_previous_plant_site	Representa un promedio de las veces que el equipo atacante plantó la bomba en cada site.	Float
attack_win	Representa si el atacante ganó la partida o si la perdió.	Dummies (1-0)

Para facilitar la ejecución de las técnicas de modelamiento empleadas, se optó por ver cada ronda desde el punto de vista del atacante, donde se agregaron también las variables de los enemigos (equipo defensor) y donde la victoria de la ronda (1/0) estará dado por la victoria del equipo atacante.

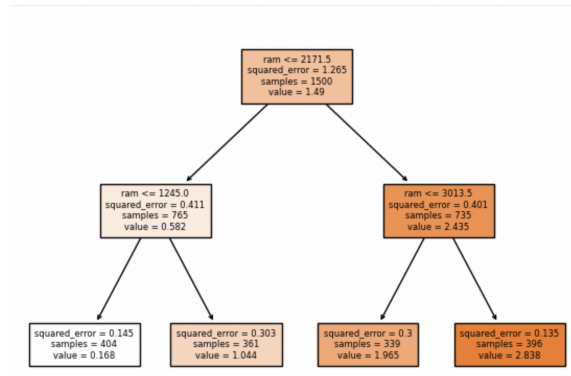
## 2.2. Metodología

Dada la baja cantidad de datos para el modelamiento, para el proyecto se realizó un análisis predictivo con Random Forest y regresión logística que permita predecir la victoria o derrota de un equipo en base a la data recopilada (composición de equipo, asistencias, mapa jugado, sitio en el que la bomba fue plantada, etc.).

### Random Forest:

Random Forest es un método de machine learning basado en ensamblaje de Árboles de decisión (5). Para un modelo de clasificación como el que analizaremos en este documento, Random forest toma todas las clasificaciones en los nodos de salida de todos los árboles y para ese nodo, utiliza la clasificación que haya tenido un mayor porcentaje de aparición frente a las otras.

Un árbol de decisión es una serie de preguntas en forma de árbol, las cuales se realizan sobre un dataset, donde cada pregunta tiene múltiples salidas (más de 1). Cada salida genera una pregunta adicional hasta que la respuesta pueda ser 1 sola en base a la data de la cual se genera el árbol de decisión. Generalmente los árboles de decisión generan 2 salidas, donde las posibles respuestas a la pregunta son Si o No (6).



Asimismo Random forest es un método simple de machine learning donde en la mayoría de los casos no se requiere establecer hiperparámetros para que funcione de buena manera y permite la obtención de la importancia de cada variable en el modelo, siendo eficiente a la hora de elegir las variables que utilizaremos para modelos más complejos (1).

### Modelo de Regresión Logística

El modelo de regresión logístico toma el logaritmo natural de la probabilidad de que un evento pase dadas las variables ingresadas como una función de regresión de predictores, donde los valores de las probabilidades pueden ser entre 0 (no ocurrió el evento) y 1

(Ocurrió el evento). Esto permite realizar de manera simple análisis predictivos para clasificación binaria (7).

Para este modelo se utiliza una función de pesos, parecida a la función lineal, lo que nos permite obtener el peso (positivo o negativo) de cada una de las variables frente al modelo.

### **Métricas de evaluación de los modelos**

Para evaluar el modelo utilizamos:

- **Matriz de confusión:** Nos permite identificar cuantas de las predicciones fueron acertadas y cuantas incorrectas. Para un modelo binario, este puede resultar muy útil e intuitivo ya que podremos tener una matriz de 2x2 evitando visualizaciones complejas (8).
- **Curva ROC:** Nos permite identificar que tan bien puede distinguir el modelo entre las 2 clases utilizando como variables la sensibilidad (verdaderos positivos/total de positivos real) y 1 - especificidad (verdaderos negativos/total de negativos real). Al calcular el área bajo la curva (AUC), nos permite obtener la probabilidad de que el modelo pueda predecir entre clase positiva y negativa de manera correcta (9). Esto permite obtener la probabilidad de predecir el modelo de manera correcta de forma fácil y a su vez visualizar si la sensibilidad y especificidad están equilibradas de tal manera que se llegue a una predicción óptima para el modelo.

### 3. Resultados

#### 3.1. Análisis Exploratorio

Dado que inicialmente contamos con 10 tablas se realizó una búsqueda exhaustiva de los datos que servirían para el modelamiento, tras lo cual, se limpió la data y se unió en 1 solo conjunto de datos que contenga toda la información relevante.

El conjunto de datos cuenta con 128.530 registros de cada jugador en cada ronda de un total de 652 partidas jugadas. Esto representa a su vez 12.853 rondas jugadas.

**Nota:** La cantidad de jugadores por ronda siempre es 10, mientras que la cantidad de rondas por partida puede oscilar dependiendo de si un equipo se rinde o la partida se alarga por rondas extras.

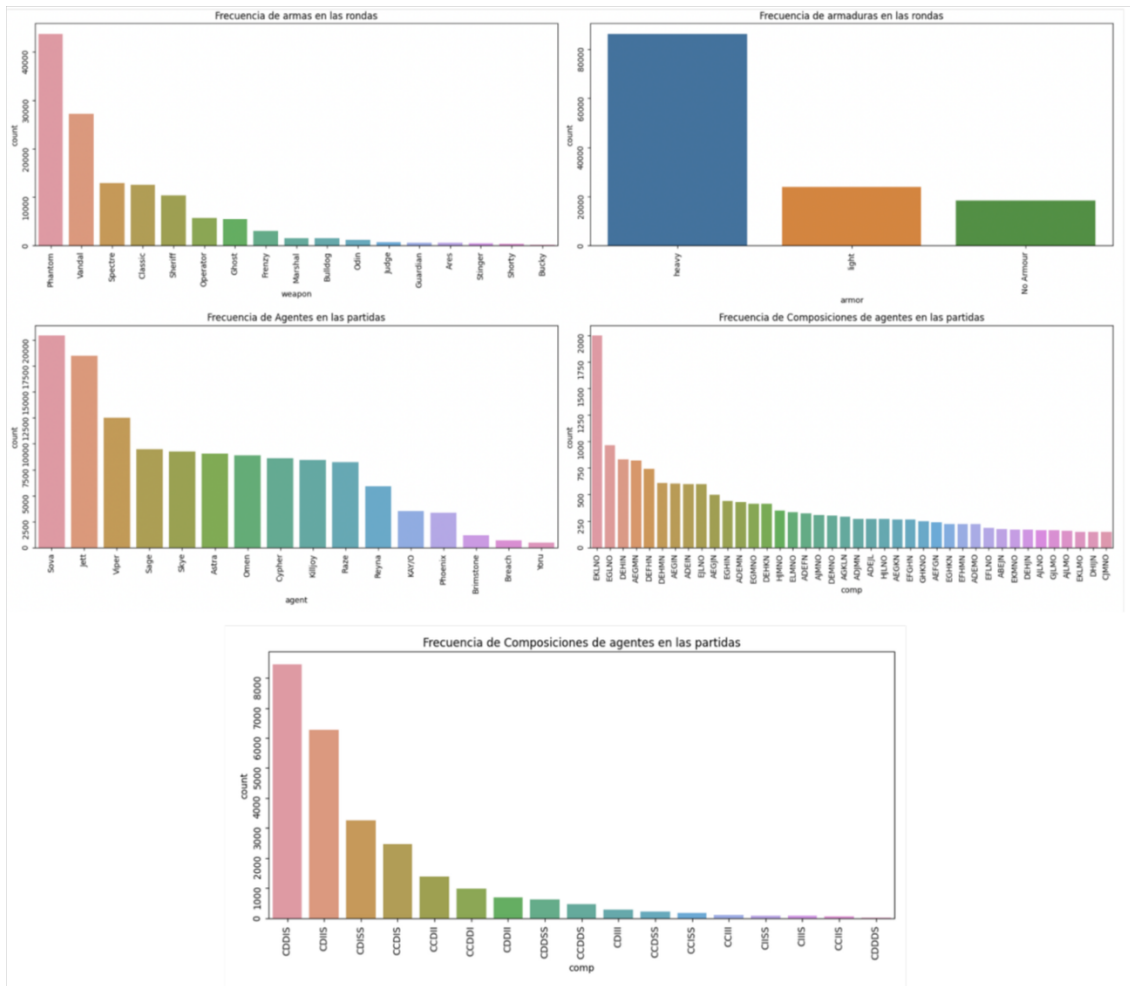
#	Column	Non-Null Count	Dtype	#	Column	Non-Null Count	Dtype	#	Column	Non-Null Count	Dtype
0	roundId	128530 non-null	int64	0	roundNumber_x	12853 non-null	int64	0	roundId	652 non-null	int64
1	roundNumber_x	128530 non-null	int64	1	playerId	12853 non-null	int64	1	roundNumber_x	652 non-null	int64
2	playerId	128530 non-null	int64	2	agent	12853 non-null	int64	2	playerId	652 non-null	int64
3	agent	128530 non-null	object	3	weapon	12853 non-null	int64	3	agent	652 non-null	int64
4	weapon	128530 non-null	object	4	armorId	12853 non-null	int64	4	weapon	652 non-null	int64
5	armorId	128530 non-null	object	5	teamId	12853 non-null	int64	5	armorId	652 non-null	int64
6	teamId	128530 non-null	int64	6	Team	12853 non-null	int64	6	teamId	652 non-null	int64
7	Team	128530 non-null	object	7	matchId	12853 non-null	int64	7	Team	652 non-null	int64
8	matchId	128530 non-null	int64	8	map	12853 non-null	int64	8	map	652 non-null	int64
9	map	128530 non-null	object	9	sitePlant	12853 non-null	int64	9	sitePlant	652 non-null	int64
10	sitePlant	77430 non-null	object	10	Winner	12853 non-null	int64	10	Winner	652 non-null	int64
11	Winner	128530 non-null	object	11	¿Won?	12853 non-null	int64	11	¿Won?	652 non-null	int64
12	¿Won?	128530 non-null	int64	12	killerId	12853 non-null	int64	12	killerId	652 non-null	int64
13	killerId	128530 non-null	int64	13	victimId	12853 non-null	int64	13	victimId	652 non-null	int64
14	victimId	89757 non-null	float64	14	killer_weapon	12853 non-null	int64	14	killer_weapon	652 non-null	int64
15	killer_weapon	89717 non-null	object	15	assistants	12853 non-null	int64	15	assistants	652 non-null	int64
16	assistants	128530 non-null	int64	16	killer_agent	12853 non-null	int64	16	killer_agent	652 non-null	int64
17	killer_agent	89757 non-null	object	17	full_comp	12853 non-null	int64	17	full_comp	652 non-null	int64
18	full_comp	128530 non-null	object	18	role_comp	12853 non-null	int64	18	role_comp	652 non-null	int64
19	role_comp	128530 non-null	object	19	weapon_comp	12853 non-null	int64	19	weapon_comp	652 non-null	int64
20	weapon_comp	128530 non-null	object								

#### Frecuencias de uso de armas, agentes y composiciones:

En estas gráficas se muestra la frecuencia de armas, armaduras, agentes y composiciones (por agente y por roles) de la totalidad del conjunto de datos.

- Se puede apreciar una clara inclinación al uso de la phantom (\$2.900) como arma de 1ra mano, mientras que para las armas de 2da mano se prefiere la sheriff (\$800) y la classic (\$0).
- En armaduras se prefiere la heavy.
- En cuanto a agentes hay una inclinación al uso de Sova (Iniciador), Jett (Duelista), mientras que los que menos se utilizan son Yoru (Duelista), Breach (Iniciador), Brimstone (Controlador), Phoenix (Duelista) y Kay/O (Iniciador). El resto de agentes se encuentra entre el promedio en cuanto a uso.

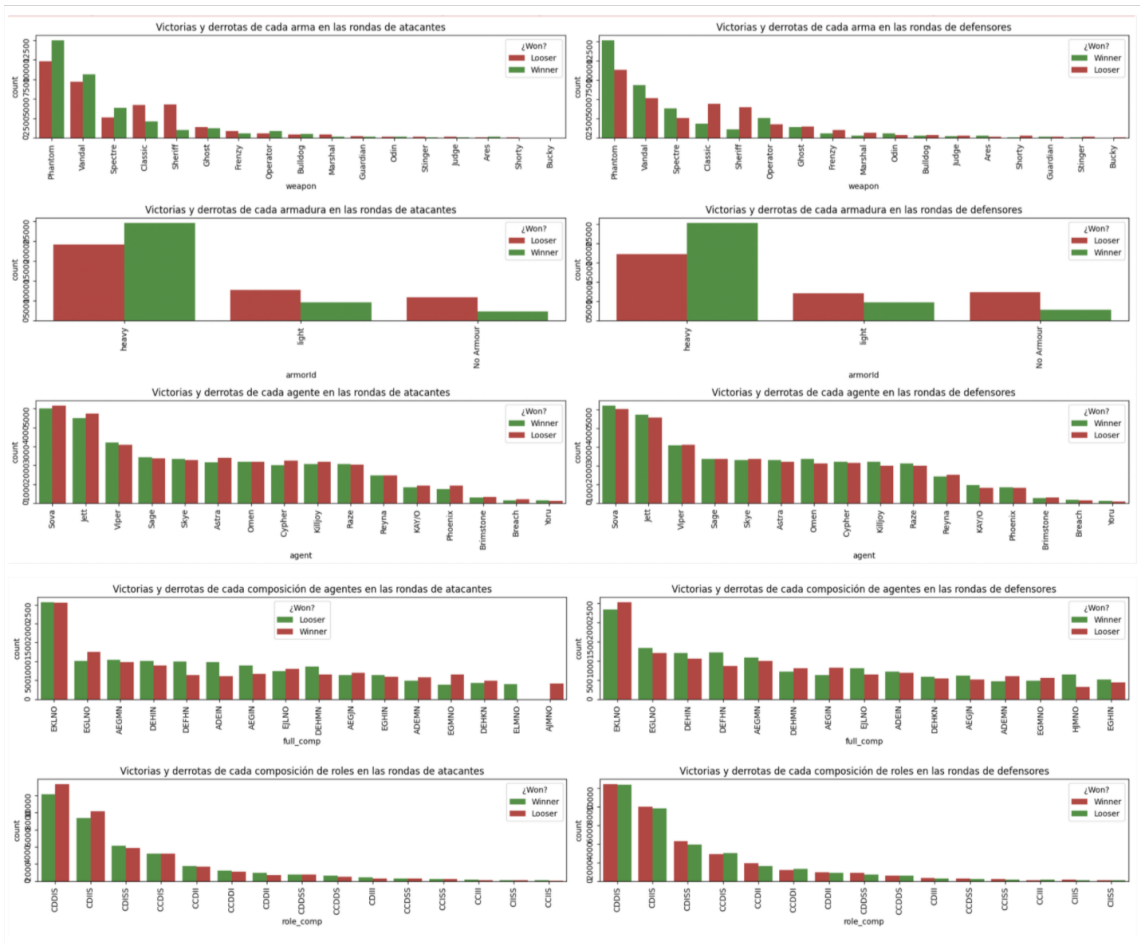
- La composición de agentes más utilizada es EKLNO, la cual está conformada por Jett (Duelista), Reyna (Duelista), Sage (Centinela), Sova (Iniciador) y Viper (Controlador).
- Asimismo la composición más utilizada es CDDIS (Controlador, Duelista, Duelista, Iniciador, Centinela), lo cual cuadra con la composición de agentes más utilizada.



## Ratio de victorias de cada variable

Esta gráfica representa las victorias y derrotas de cada arma, armadura, agente y composición, tanto para equipos atacantes como defensores.

Aparte de la predominancia de la phantom como arma y el uso de armadura heavy y el bajo ratio de victorias de la classic y sheriff, no se puede divisar una predominancia de las otras variables en cuanto al mismo concepto.

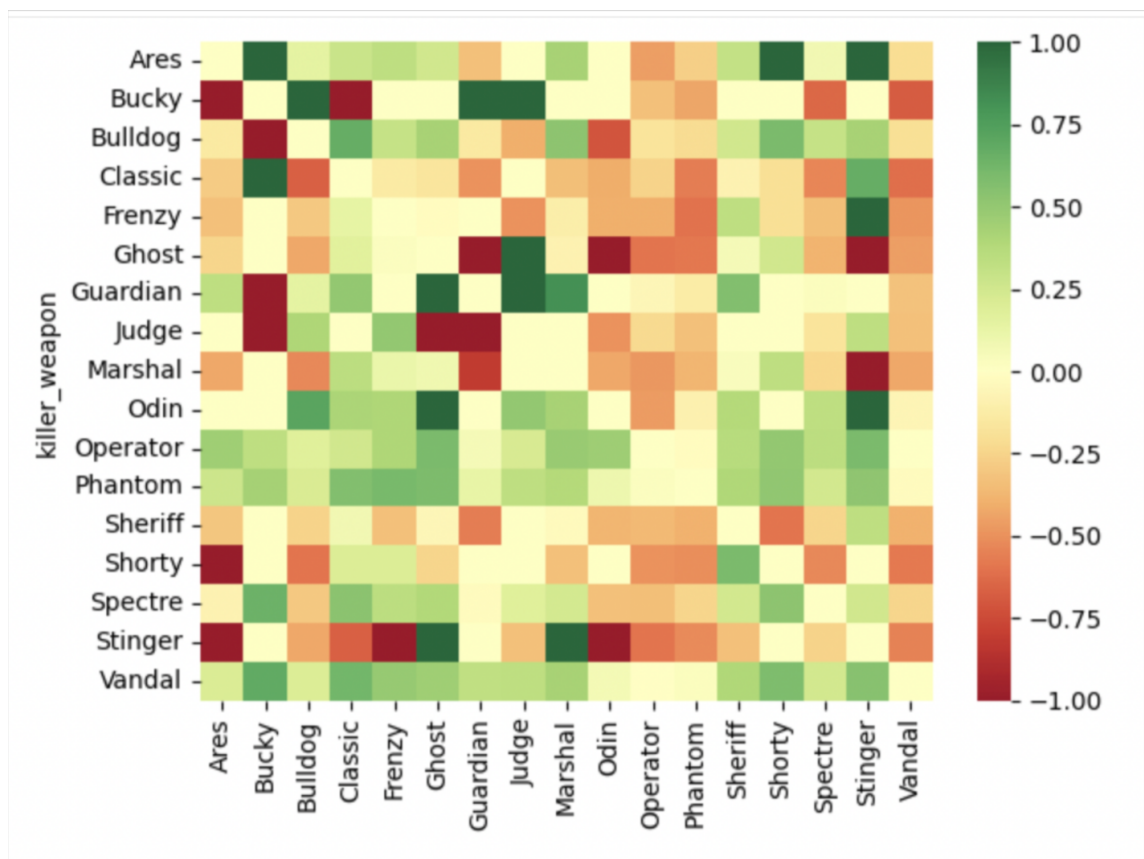


### Ratio de victorias de cada arma contra otra:

Para esta matriz se estudiaron todas las victorias y derrotas en enfrentamiento que tuvo cada arma.

En el eje Y podemos identificar el arma observada, mientras que en el eje X podemos observar el arma enemiga.

La matriz muestra patrones marcados de los enfrentamientos, por lo que visualizamos que el arma utilizada contra el arma enemiga puede ser determinante para la victoria en el combate.

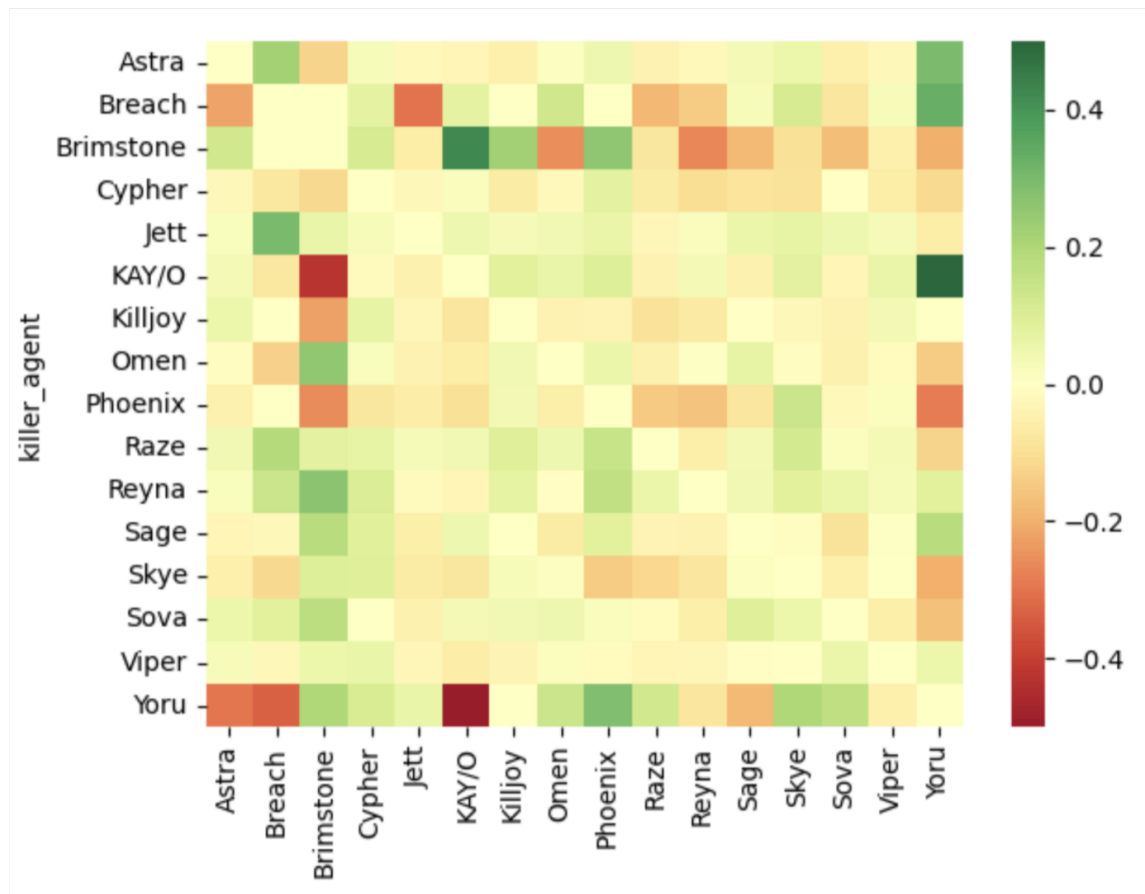


### Ratio de victorias de cada agente contra otro:

Para esta matriz se estudiaron todas las victorias y derrotas en enfrentamiento que tuvo cada agente contra otro.

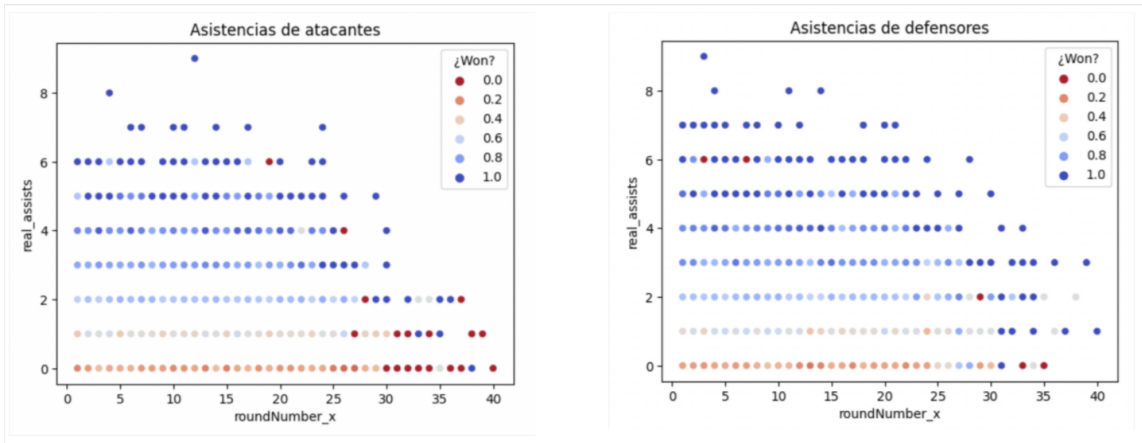
En el eje Y podemos identificar el agente observado, mientras que en el eje X podemos observar el agente enemigo.

La matriz no representa un mayor ratio de victorias en enfrentamientos entre agentes, salvo algunas excepciones.

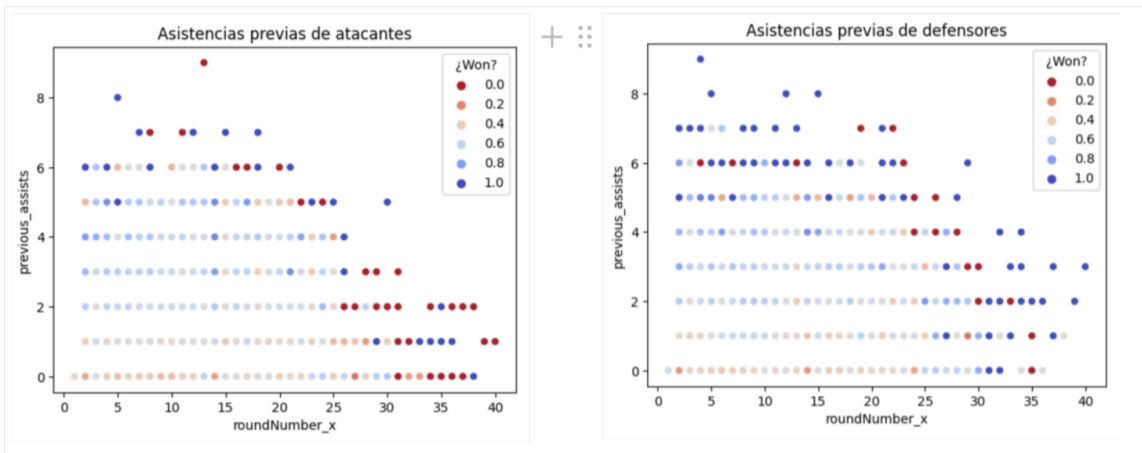


### Ratio de victorias en base a las asistencias en cada ronda:

El gráfico muestra como la cantidad de asistencias en 1 ronda a un mayor ratio de victoria. Esto es analizado la misma ronda en la que se consiguieron esas asistencias.

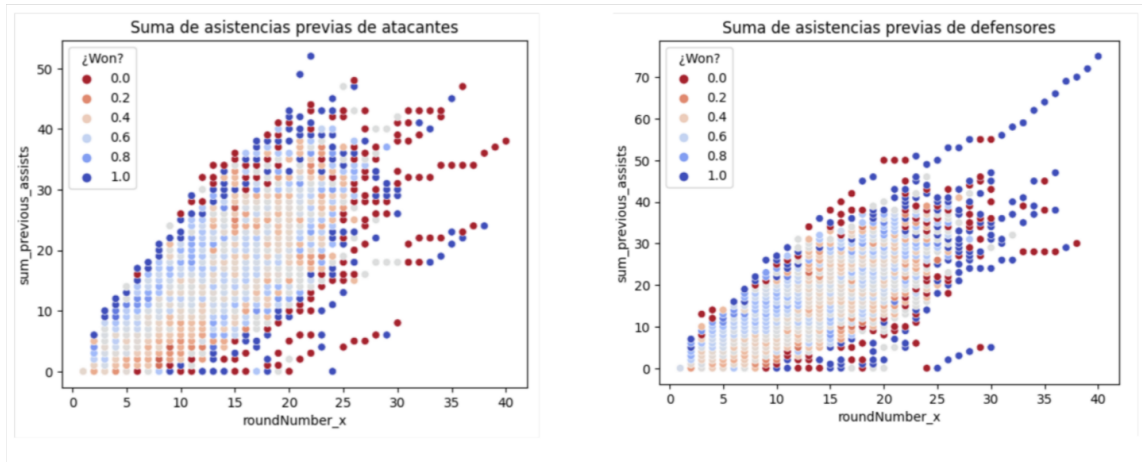


Asimismo, si analizamos las asistencias previas, se puede notar nuevamente una tendencia en la que, mientras mayor cantidad de asistencias previas, mayor es el índice de victoria, por lo que afirmamos que, en cierta medida, las asistencias previas repercuten en la partida actual y pueden ser determinantes de la estrategia del equipo.

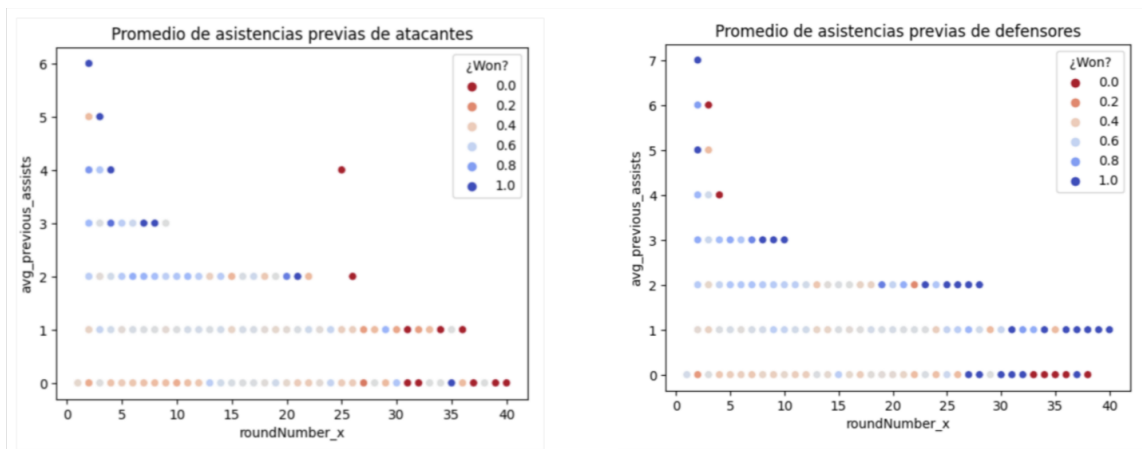


Dado que el juego tiene más de 2 rondas, se optó por sumar la cantidad de asistencias para analizar una nueva gráfica que corresponde a la evolución del ratio de victorias en cada

partida durante cada ronda vs las asistencias previas del equipo en cada ronda. Se puede divisar que hay una leve tendencia a ganar la ronda cuando en las rondas anteriores hay más asistencias.



Por último, se analiza el ratio de victorias en cada ronda con respecto al promedio de asistencias pasadas. En este gráfico ya vemos una clara relación entre las asistencias previas y el ratio de victorias, por lo que utilizaremos el promedio de asistencias previas para el modelo a evaluar. Las asistencias fueron aproximadas al número entero más cercano para facilitar el armado del gráfico.



## 3.2. Modelamiento:

Para el modelamiento se utilizaron 2 técnicas: Random Forest y Regresión Logística. En ambos se utilizó un conjunto de testeo del 30%, lo cual nos permite evaluar el desempeño de los datos.

Para **Random Forest**, establecemos los siguientes hiperparámetros:

- `n_estimators` (300): La cantidad de iteraciones óptima para el modelo.
- `min_samples_split` (800): Dado que las primeras iteraciones del modelo tenían un alto grado de overfitting, establecemos esta variable para corregirlo.

Para la **regresión logística**, establecemos los siguientes hiperparámetros:

- `solver` (liblinear): Solver óptimo para conjuntos de datos más pequeños
- `multi_class` (ovr): Sirve para problemas binarios, mientras que multinomial sirve para problemas de multi clase.
- `C` (0.1): Lo utilizamos para una mayor regularización.

### 3.3. Evaluación del modelo

En la evaluación, la regresión logística presento levemente mejores resultados, por lo que utilizaremos esta como referencia para evaluar el modelo.

#### Random Forest:

El modelo de Random Forest obtuvo un puntaje de precisión del 64,65%.

Como se puede observar en la tabla de importancia de variables, las armas son las que predominan en cuanto a la importancia, seguido por el promedio de asistencias y por último el promedio de veces que se planta la bomba en cada zona, mientras que los agentes y los roles que cumplen no muestran relevancia para el modelo.

Sheriff	0.167026	E_Ares	0.001866	Controladores	0.000400
E_Sheriff	0.145980	E_Stinger	0.001537	KAY/O	0.000399
E_Classic	0.105922	Cypher	0.001471	E_Astra	0.000395
E_Phantom	0.092403	Shorty	0.001408	Killjoy	0.000377
Phantom	0.065521	E_Reyna	0.001368	Reyna	0.000356
Classic	0.065211	Raze	0.001327	E_Yoru	0.000349
E_Operator	0.036263	E_Guardian	0.001301	Sova	0.000321
E_Spectre	0.029613	E_Breach	0.001020	Breach	0.000302
E_Vandal	0.027278	E_Cypher	0.001004	Duelistas	0.000281
Spectre	0.026780	Phoenix	0.000982	E_Sova	0.000275
E_Shorty	0.026288	Sage	0.000930	E_Brimstone	0.000263
Vandal	0.024979	E_Viper	0.000928	E_Phoenix	0.000242
avg_N	0.024619	E_Skye	0.000811	E_Raze	0.000211
avg_assists	0.019331	Iniciadores	0.000701	E_Controladores	0.000202
enemy_avg_assists	0.015327	E_Centinelas	0.000701	Skye	0.000197
E_Frenzy	0.014853	Judge	0.000694	Bucky	0.000181
avg_A	0.014127	E_Iniciadores	0.000690	Yoru	0.000144
E_Marshal	0.009133	Ares	0.000690	Brimstone	0.000137
Marshal	0.008130	Viper	0.000644		
Frenzy	0.007615	E_Bucky	0.000622		
E_Ghost	0.006678	E_Jett	0.000616		
avg_B	0.006443	E_Judge	0.000613		
Ghost	0.005632	E_KAY/O	0.000585		
Operator	0.004815	Astra	0.000574		
E_Odin	0.003662	Odin	0.000567		
Centinelas	0.003354	E_Killjoy	0.000555		
Bulldog	0.003010	Guardian	0.000544		
avg_C	0.002222	Omen	0.000528		
E_Omen	0.002141	E_Duelistas	0.000493		
Stinger	0.002057	Jett	0.000481		

## Regresión Logística:

El modelo de regresión logística obtuvo un puntaje de precisión del 65,32%.

Como se puede ver en la tabla de coeficientes de la regresión, gran parte de las armas, ya sea del equipo atacante o del defensor, tienen los mayores coeficientes. De esto se puede decir que el arma utilizada es más importante que el agente o la composición. Asimismo, se puede ver que el sitio para plantar la bomba en A tiene un mayor peso que plantar la bomba en otros sites, que existen algunos agentes que tienen un mayor peso en el resultado de la ronda (ej: Omen) y que el promedio de asistencias previas del equipo atacante tienen peso positivo sobre la victoria de la ronda mientras que las asistencias previas del equipo enemigo tienen un peso negativo, lo que nos da a entender que el trabajo en equipo y la coordinación entre 2 jugadores para enfrentarse a un enemigo repercute positivamente en el desempeño del equipo durante la ronda.

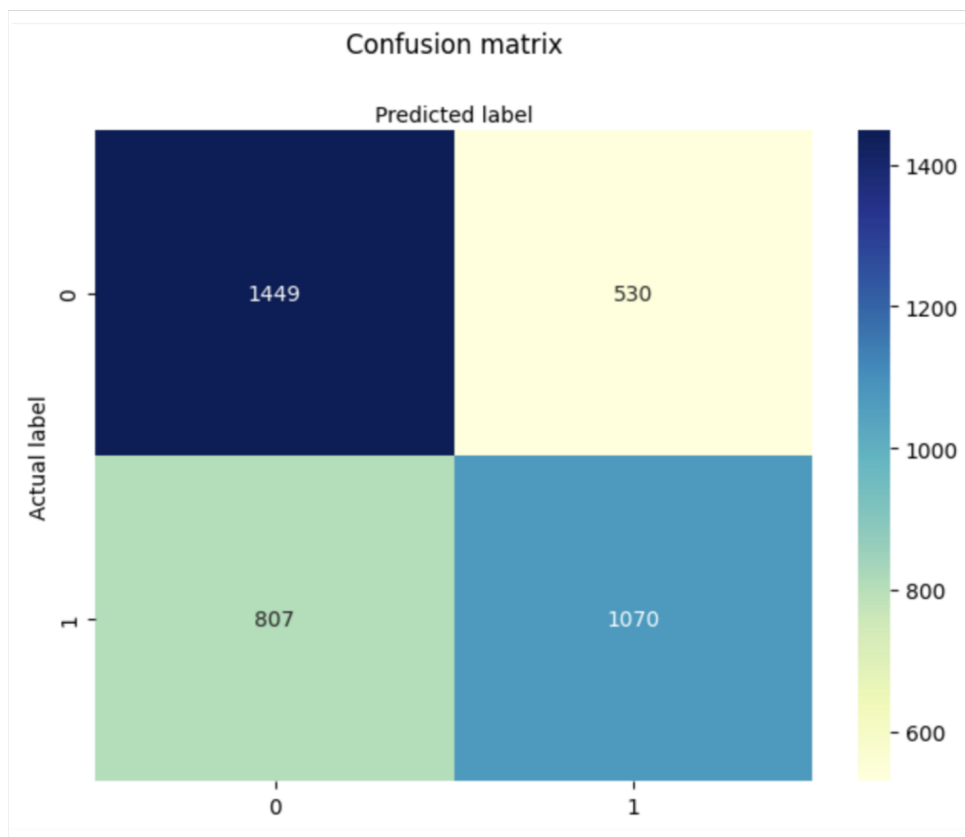
Por otro lado, podemos ver algunas armas del equipo atacante que tienen pesos neutros (cerca de 0) mientras que si la misma arma la tiene el equipo defensor, tienen un peso negativo sobre la victoria del atacante. Esto puede estar relacionado a una falta de data o bien que los equipos atacantes no optan por utilizar estas armas.

E_Classic	0.277104	E_Astra	0.018068	Cypher	-0.050900
Phantom	0.212914	Yoru	0.018057	E_Raze	-0.059168
Operator	0.200384	Guardian	0.017874	enemy_avg_assists	-0.061037
E_Shorty	0.179940	E_Jett	0.012500	Shorty	-0.062803
E_Ghost	0.174841	E_Viper	0.012056	E_Omen	-0.075999
Spectre	0.167539	E_Phoenix	0.010249	Phoenix	-0.080780
E_Frenzy	0.160852	E_Sova	0.009694	E_KAY/O	-0.089848
Vandal	0.153992	E_Duelistas	0.002368	Frenzy	-0.105226
E_Marshal	0.136033	Killjoy	0.002344	Brimstone	-0.109059
E_Sheriff	0.120939	E_Controladores	-0.001078	Marshal	-0.117376
avg_A	0.103559	Odin	-0.002682	E_Odin	-0.170946
avg_assists	0.094933	E_Centinelas	-0.003331	Sheriff	-0.176127
Omen	0.093555	E_Judge	-0.005787	Ghost	-0.180113
Bulldog	0.092445	Controladores	-0.005918	E_Spectre	-0.188521
E_Skye	0.088895	Breach	-0.008272	E_Operator	-0.204171
E_Reyna	0.083312	Centinelas	-0.008320	E_Vandal	-0.226467
Ares	0.075403	E_Iniciadores	-0.010560	E_Phantom	-0.241975
E_Cypher	0.046499	E_Sage	-0.012871	Classic	-0.281391
avg_B	0.045566	E_Bulldog	-0.015024		
Skye	0.044889	Duelistas	-0.016635		
E_Brimstone	0.044798	Stinger	-0.018835		
E_Bucky	0.043501	E_Breach	-0.019301		
Sage	0.040237	E_Guardian	-0.022667		
Raze	0.040180	Bucky	-0.023735		
Judge	0.035139	Astra	-0.024569		
avg_C	0.034479	Reyna	-0.026318		
Viper	0.034155	avg_N	-0.030164		
Jett	0.032225	E_Killjoy	-0.036959		
E_Stinger	0.020489	KAY/O	-0.037096		
Sova	0.018752	E_Yoru	-0.044524		

Dado que el modelo de regresión logística presento mejores resultados, evaluaremos este como modelo principal, para lo utilizaremos la matriz de confusión y la curva ROC, lo que nos permitirá identificar de manera más precisa.

### Matriz de confusión

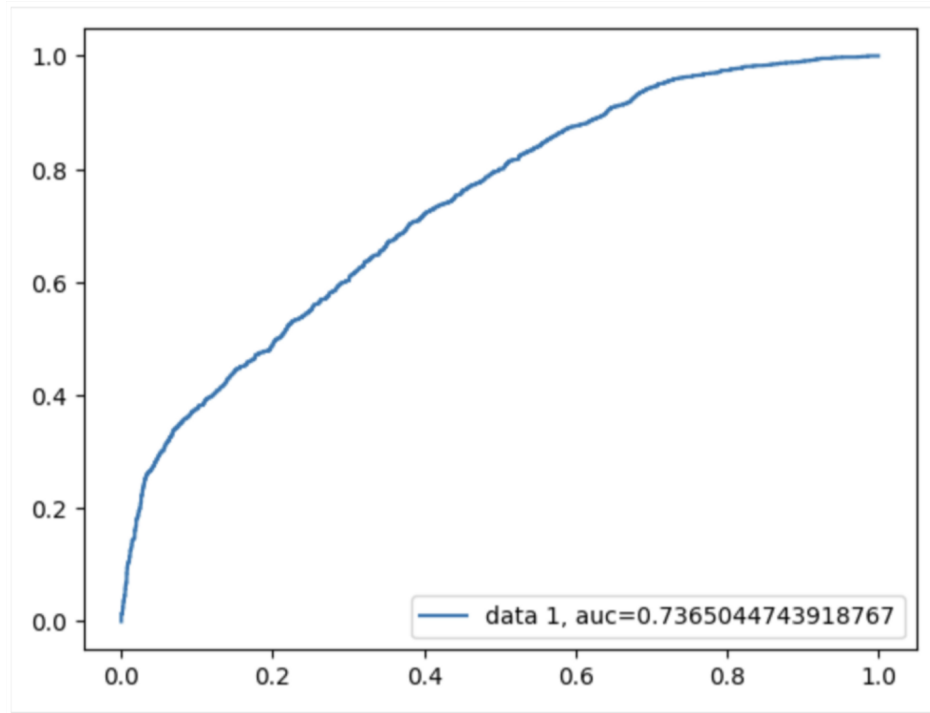
Como se aprecia en la matriz de confusión, pudimos predecir sobre 2500 rondas del conjunto de testeo de manera positiva, sin embargo hay otras 1000 rondas que no pudimos predecir asertivamente.



### Curva ROC

Con la curva ROC, podemos evaluar el modelo entre el falso positivo y el verdadero positivo.

Como podemos apreciar, el AUC score es de 0,74 (siendo 1 un clasificador perfecto y 0.5 el peor clasificador), por lo que apreciamos que el modelo es medianamente bueno y es capaz de predecir el resultado de la partida de manera correcta un 74% de las veces.



## 4. Conclusiones

Dado que el modelo arroja una precisión del 65,32% y que el AUC ratio es de un 0,74, y tomando en cuenta que de las variables estratégicas planteadas en la hipótesis que efectivamente resultaron significativas en el modelo, podemos concluir que la estrategia logra influir en el resultado de la ronda, tomando como referencia 3 variables principales: armas utilizadas en ataque y defensa, sitio de plantación de la bomba y asistencias previas.

Sin embargo dado que el modelo arroja una precisión del 65,32%, inferior al esperado en el objetivo general (75%), no podemos decir que la estrategia sea determinante de una ronda.

Dicho esto rechazamos la hipótesis nula dado que no podemos explicar la mayoría de las victorias o derrotas de rondas a partir de la estrategia, sin embargo, se plantean mejoras para el modelo para trabajos futuros.

### **Mejoras para el modelo:**

- Actualmente contamos con un registro pequeño de rondas (12.853), por lo que contar con más data podría significar una mejora para el modelo.
- Como se comenta en la introducción, durante la partida hay otras variables que influyen en el resultado (posicionamiento inicial durante la ronda, uso de habilidades en rondas pasadas, uso de habilidades en conjunto, interrupción del entrenador entre rondas, etc.). Estas variables no las pudimos obtener de la data, por lo que se quedaron afuera del modelo, pero sería interesante poder analizarlas para ver si llegamos a una mejor precisión.
- Ya con más data, podemos armar un modelo más robusto basado en redes neuronales recurrentes, lo que no utilizamos dada la poca cantidad de data con la que contamos, sin embargo estudios avalan que modelos más simples son mejores en datos tabulares.

Una vez contemos con todas estas variables será posible generar un modelo más robusto y poder validar definitivamente la hipótesis.

## 5. Bibliografía

1. **Adele Cutler, D. Richard Cutler, John R. Stevens.** Random Forsts. [En línea] 2011. <https://www.researchgate.net/file.PostFileLoader.html?id=5940c7b040485459f416c743&assetKey=AS:505020374319105@1497417648285>.
2. **Guides, Esports.** *CSGO RANKS » THE CSGO RANKING SYSTEM EXPLAINED.* [En línea] 2022. <https://www.esports.net/wiki/guides/csgo-ranks-explained/>.
3. **Kajal T. Claypool, Mark Claypool.** On frame rate and player performance in first person shooter games. [En línea] 2007. <http://web.cs.wpi.edu/~claypool/papers/fr/fulltext.pdf>.
4. **Yogi Tri Prasetyo, Cheselle Jan Roldan.** Evaluating The Effects of Aim Lab Training on Filipino Valorant Players' Shooting Accuracy. [En línea] 2021. [https://www.researchgate.net/profile/Yogi-Prasetyo/publication/351911867\\_Evaluating\\_The\\_Effects\\_of\\_Aim\\_Lab\\_Training\\_on\\_Filipino\\_Valorant\\_Players'\\_Shooting\\_Accuracy/links/60b261d5299bf1f6d5847b1e/Evaluating-The-Effects-of-Aim-Lab-Training-on-Filipino-Valor](https://www.researchgate.net/profile/Yogi-Prasetyo/publication/351911867_Evaluating_The_Effects_of_Aim_Lab_Training_on_Filipino_Valorant_Players'_Shooting_Accuracy/links/60b261d5299bf1f6d5847b1e/Evaluating-The-Effects-of-Aim-Lab-Training-on-Filipino-Valor).
5. **Kaitlin Kirasich, Trace Smith, Bivin Sadler.** Random Forest vs Logistic Regression: Binary Classification for Heterogeneous Datasets. [En línea] 2018. <https://scholar.smu.edu/cgi/viewcontent.cgi?article=1041&context=datasciencereview>.
6. **Carl Kingsford, Steven L Salzberg.** [En línea] 2011. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2701298/>.
7. **LaValley, Michael P.** [En línea] 2008. <https://www.ahajournals.org/doi/full/10.1161/CIRCULATIONAHA.106.682658>.
8. **Dalianis, H.** Evaluation Metrics and Evaluation. [En línea] 2018. [https://link.springer.com/content/pdf/10.1007/978-3-319-78503-5\\_6.pdf](https://link.springer.com/content/pdf/10.1007/978-3-319-78503-5_6.pdf).
9. **Radiology, Korean Journal of.** Receiver Operating Characteristic (ROC) Curve: Practical Review for Radiologists. [En línea] 2004. <https://synapse.koreamed.org/articles/1027596>.
10. **Megan Smith, Stephen Lee-Urban, Hector Muñoz-Avila.** RETALIATE: Learning Winning Policies in First-Person Shooter Games . [En línea] [https://faculty.cc.gatech.edu/~surban6/pubs/IAAI02SmithM\\_errata.pdf](https://faculty.cc.gatech.edu/~surban6/pubs/IAAI02SmithM_errata.pdf).