

TP0: Introducción al Análisis de Datos

Trabajo práctico orientativo para la materia Sistemas de Inteligencia Artificial
con el objetivo de evaluar la función de captura de un Pokemon.

GIROD JOAQUIN (63512),.....JGIROD@ITBA.EDU.AR
IJAS CHRISTIAN (63555),.....CIJJAS@ITBA.EDU.AR
MAGLIOTTI GIANFRANCO (61172),.....GMAGLIOTTI@ITBA.EDU.AR
FERRUTTI FRANCISCO (62780),.....FFERRUTTI@ITBA.EDU.AR

15 de agosto de 2024

Índice

1. Introducción	2
2. Consigna	3
3. Respuestas Pregunta 1	4
3.1. Ítem a	4
3.2. Ítem b	7
4. Respuestas Pregunta 2	8
4.1. Ítem a	8
4.2. Ítem b	9
4.3. Ítem c	9
4.4. Ítem d	13
4.5. Ítem e	13
5. Conclusión	13
6. Anexo	14

1 Introducción

La franquicia de Pokemon es una de las más conocidas a nivel mundial. La saga de juegos consiste en un mundo con criaturas llamadas Pokemons con poderes de distinto tipo. Los jugadores podrán capturar a estas criaturas utilizando Pokeballs. Sin embargo, la captura de los mismos está sujeta a variables del entorno: el tipo de pokeball y las características del pokemon.

El presente documento detalla el análisis hecho de las probabilidades de captura de los pokemons, haciendo variar distintos parámetros para su captura. Se utilizó Python como herramienta gráfica y de procesamiento de la información y se tuvo en cuenta un grupo reducido de pokemons y pokeballs para el análisis. El flujo de ejecución sigue una patrón de dos etapas, una de carga (escritura de un archivo csv) y y otra de análisis (diagramación con pandas y librerías de representación gráfica de datos). Cada ejercicio se resolvió de esta manera para simplificar la lectura de código y la reutilización del mismo.

2 Consigna

1. Acerca de pokebolas:

a) Ejecutando la función 100 veces, para cada Pokémon en condiciones ideales (HP: 100 %, LVL 100) ¿Cuál es la probabilidad de captura promedio para cada pokebola?

b) ¿Es cierto que algunas pokebolas son más o menos efectivas dependiendo de propiedades intrínsecas de cada Pokémon? Justificar.

Sugerencia: Comparar efectividad (success/total attempts) como proporción de la efectividad de la pokebola básica para cada Pokémon.

2. Acerca del estado del Pokémon:

a) ¿Las condiciones de salud tienen algún efecto sobre la efectividad de la captura? Si es así, ¿Cuál es más o menos efectiva?

b) ¿Cómo afectan los puntos de vida a la efectividad de la captura?

Sugerencia: Elegir uno o dos Pokémon y manteniendo el resto de los parámetros constantes, calcular la probabilidad de captura para distintos HP %.

c) ¿Qué parámetros son los que más afectan la probabilidad de captura?

d) Teniendo en cuenta uno o dos Pokémon distintos: ¿Qué combinación de condiciones (propiedades mutables) y pokebola conviene utilizar para capturarlos?

e) A partir del punto anterior, ¿sería efectiva otra combinación de parámetros teniendo en cuenta un nivel del Pokémon más bajo (o más alto)?

3 Respuestas Pregunta 1

3.1 Ítem a

En primer lugar debemos considerar que contamos con 5 pokemons distintos. Cada uno de ellos cuenta con estadísticas propias e independientes de los demás. A su vez, contamos con 4 tipos de pokeballs distintos, cada una con características que puede afectar a la probabilidad de captura de los pokemon.

Dando lugar a la resolución del primer ítem, observando la Figura 1 notamos a simple vista una gran variedad en las probabilidades de captura para cada pokemon según la pokeball utilizada. Una aclaración previa al gráfico: no consideramos necesario hacer una normalización de las probabilidades ya que este gráfico conduce a una línea de interpretación de los datos que se muestran en las figuras posteriores.

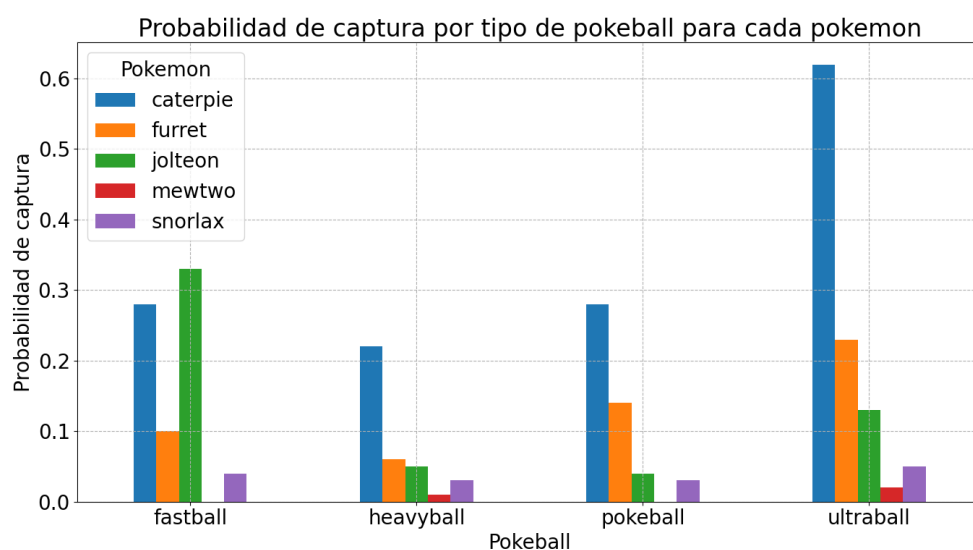


Figura 1: Probabilidad promedio de captura agrupado por pokeball y pokemon para 100 intentos

La primer conclusión que podemos sacar del gráfico la elección de la pokebola es que tiene efecto directo sobre la probabilidad de captura. Como segunda conclusión es que el tamaño de muestreo propuesto por el ejercicio está al borde de ser malo. Se puede ver como la ultraball parece tener muchos positivos para mewtwo a comparación de fastball cuando en realidad la relación probabilística no es así. Por eso decidimos usar una muestra más grande de 10000 (diez mil) catch attempts.

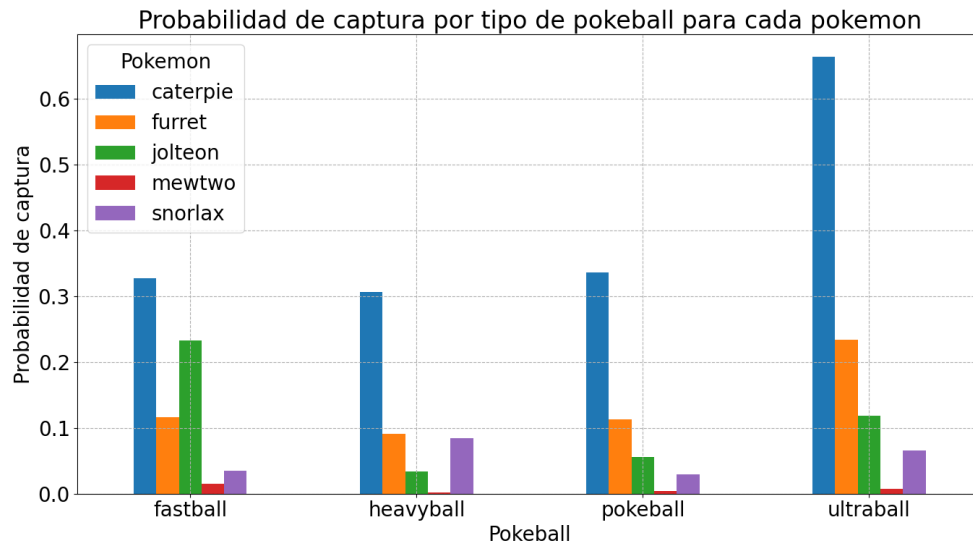


Figura 2: Probabilidad promedio de captura agrupado por pokeball y pokemon para 10000 intentos

Si consideramos el caso de furret, su probabilidad de captura se encuentra al rededor de 10 % para las primeras tres pokeballs, mientras que una ultraball cambia significativamente su probabilidad de captura, aumentando la misma a un 20 %. De manera similar ocurre con caterpie, quién tiene una probabilidad del 30 % de ser capturado para las primeras pokeballs mientras que si se utiliza la ultraball, esa probabilidad supera el 60 %.

En contraposición tenemos el caso de jolteon, que con la fastball aumenta significativamente su probabilidad de captura. Algo similar ocurre con mewtwo, cuyo mayor promedio de captura se alcanza con esta misma pokeball. Podríamos entonces tomar como indicador estos datos y formular una hipótesis, en la que la fastball aumenta la probabilidad de captura de pokemones con alta velocidad, como es el caso de jolteon y mewtwo. Esto mismo será analizado en el ítem b de esta pregunta, así como en los ítem c y d de la pregunta 2.

Para continuar con el análisis del ejercicio, presentamos un boxplot con el fin de visualizar la distribución de los catch attempts para las distintas pokebolas.

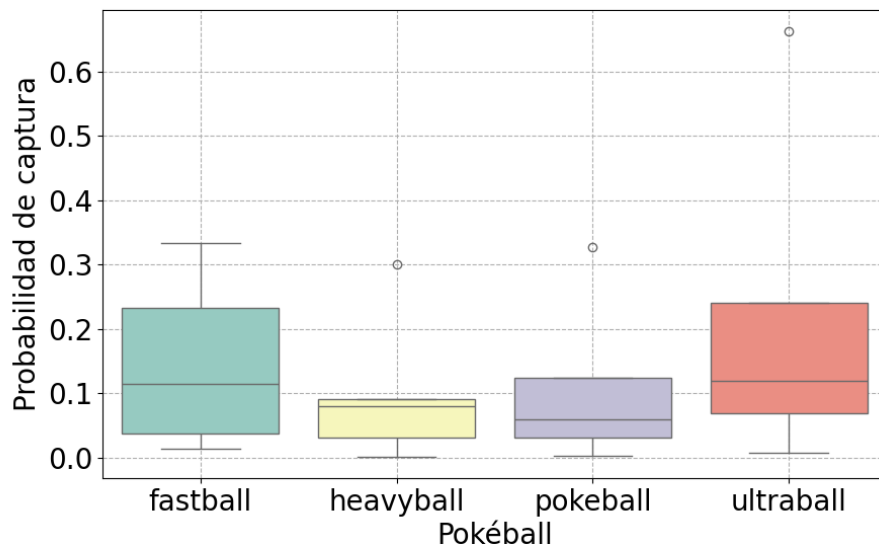


Figura 3: Boxplot de la probabilidad de captura por tipo de Pokeball

Podemos observar la presencia de datos atípicos en la heavyball, pokeball y la presencia de un outlier severo en la ultraball. Esto indica que hay al menos un pokemon que tiene una probabilidad de captura distinta al resto lo cual provoca un desvío de la media y da origen a datos atípicos. Es por esto que se decidió usar boxplot ya que permite una visualización más racional del asunto. Esto último nos sugiere hacer una observación sobre la varianza de las probabilidades dando lugar al gráfico 4 donde finalmente se puede observar el grado de varianza que tienen las pokebolas para los diferentes pokemones.

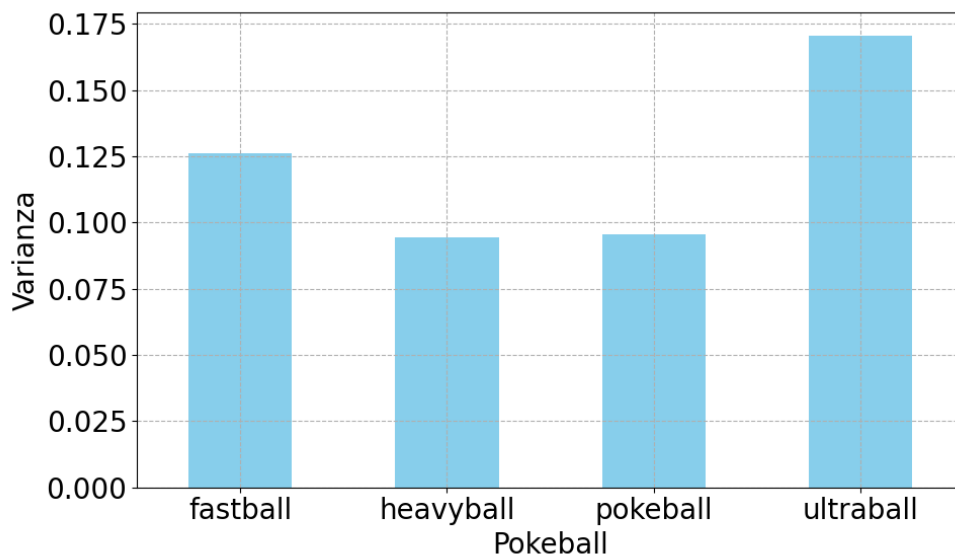


Figura 4: Varianza de las probabilidades de captura para cada pokeball

Se puede concluir que la ultraball presenta la varianza más grande (ya predecible en la figura 3) indicando que su resultado puede depender fuertemente del tipo de pokemon al cual se le esté aplicando. Por eso sin conocer la fórmula de cálculo verdadera se puede decir que no es obvia la elección de ultraball ya que en muchos casos es óptima (hipótesis que se creía verdadera para todos los casos) pero en otros no conviene.

3.2 Ítem b

Para determinar la efectividad de las pokeball en función de las propiedades intrínsecas de cada pokemon decidimos aplicar el tamaño de la muestra a 100000 (cien mil) intentos por pokemon y pokeball.

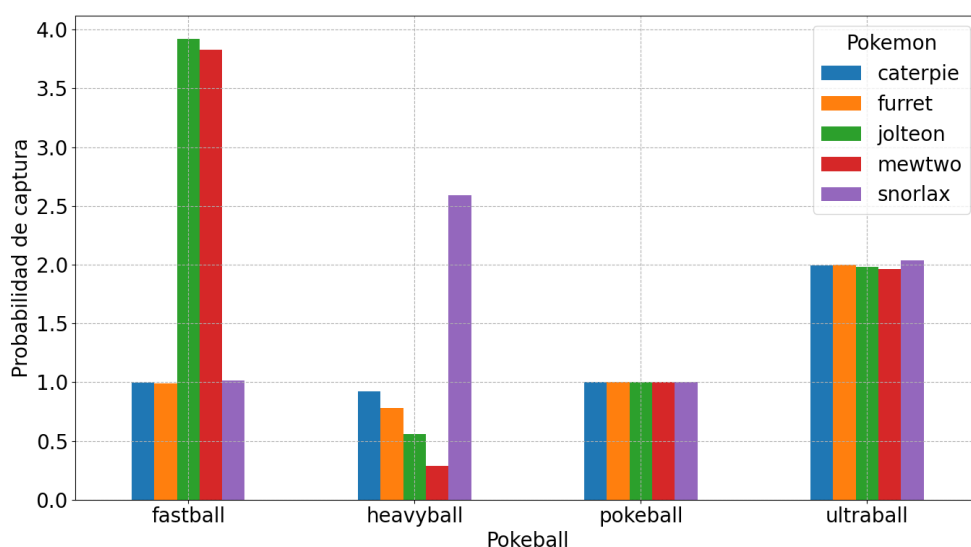


Figura 5: Probabilidad promedio de captura de pokemons, agrupado por pokeball

Además, se normalizaron los promedios obtenidos en función de la Pokeball básica. De esta manera, es posible medir cuánto cambia la efectividad de cada Pokéball al atrapar a cada Pokémon en función de la Pokéball básica. Así, se obtiene una unidad de medida con el fin de comparar los resultados obtenidos y sacar conclusiones más fácilmente.

En base a los resultados obtenidos, podemos afirmar con un alto grado de confianza que es cierto que la efectividad de algunas Pokéball dependen de propiedades intrínsecas de cada Pokemon. En el caso de las Pokéball analizadas, podemos ver dos casos particulares: La Heavyball y la Fastball. Estas dos, a diferencia de la Pokéball básica y la Ultraball que parecen tener modificadores idénticos para todos los Pokémon, muestran grandes diferencias entre las distintas probabilidades promedio de captura.

En primer lugar tenemos la Fastball, cuyo modificador de captura parece cuadruplicarse cuando se intenta capturar a Jolteon o a Mewtwo. Por su nombre, "Fastball", uno podría suponer que este aumento está vinculado con la estadística de velocidad de Pokémon a capturar. Viendo los Pokémon analizados vemos que Jolteon y Mewtwo son los que poseen mayor estadística de velocidad, ambos con 130 puntos. Se hace foco en el hecho de que el incremento en el modificador parece ser binario, en el sentido de que al cumplir cierto requisito o valor mínimo de velocidad la probabilidad de captura aumenta repentinamente. Furret, con una velocidad de 90 puntos, demostro la misma

probabilidad promedio que de los Pokémon mucho mas lentos, como Snorlax cuya velocidad es 30. En cuanto a la Heavyball, esta presenta resultados mucho mas variados a comparación de las demás.

Algo a notar es que esta es la única Pokéball que parece tener menor probabilidad de capturar ciertos Pokémon que una Pokéball básica, como el caso de Mewtwo en el cual la probabilidad de captura promedio resulta ser menos de la mitad.

Tambien se destaca el hecho de que a diferencia de la Fastball se puede ver la posible existencia de mas de dos grupos de modificadores.

Intuitivamente podemos suponer que por su nombre, "Heavyball", el modificador de probabilidad de captura varia segun el peso de cada Pokémon. Esto se correlaciona con los resultados obtenidos, ya que ordenando los Pokémon por peso los Pokémon de menor a mayor la probabilidad relativa de captura aumenta en el mismo sentido.

4 Respuestas Pregunta 2

4.1 Ítem a

Con el fin de determinar si las condiciones de salud tienen algún efecto sobre las probabilidades de la captura, tomamos en consideración a Jolteon para el análisis. Las posibles condiciones de salud consisten en estar envenenado, quemado, paralizado, dormido, congelado o estar libre de cualquiera de las anteriores.

Si observamos la figura 6 donde se puede ver la probabilidad de captura en función del estado del pokemon, podríamos interpretar que la buena salud del pokemon hace más difícil capturarlo. Es decir, que si la salud del pokemon empeora, su probabilidad de captura aumenta.

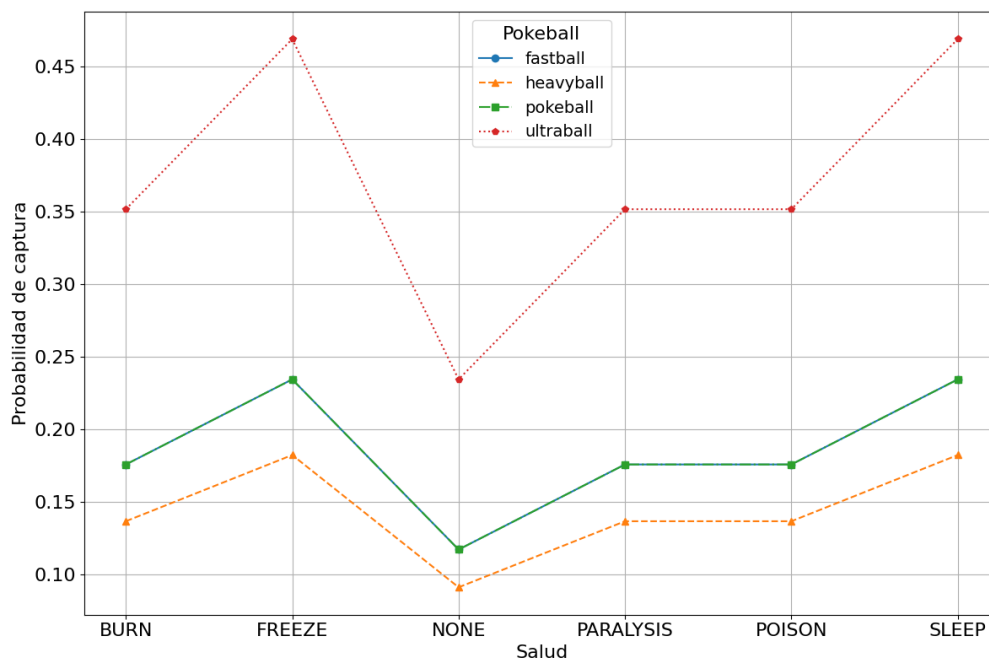


Figura 6: Probabilidad de captura de Jolteon en función del estado de su salud

Al momento de determinar cuál es el estado de salud más efectivo para capturar la pokemon, podemos afirmar que se obtienen resultados similares, independientemente

de la pokeball utilizada, para los casos en los que se encuentra dormido o congelado. En ambos, si consideramos una pokeball en particular, la probabilidad aumenta a un porcentaje similar.

Vale la pena destacar que la conclusión mencionada se condice con lo que ocurre en los juegos de la franquicia. Pues se recomienda inducir estos efectos a fin de facilitar su captura.

4.2 Ítem b

Ahora vamos a analizar si los puntos de vida afectan a las probabilidades de captura. En este caso, decidimos variar el parámetro noise para obtener una variación y así poder estimar el error que se está cometiendo. A continuación presentamos el gráfico de la probabilidad de captura de jolteon en función de su vida.

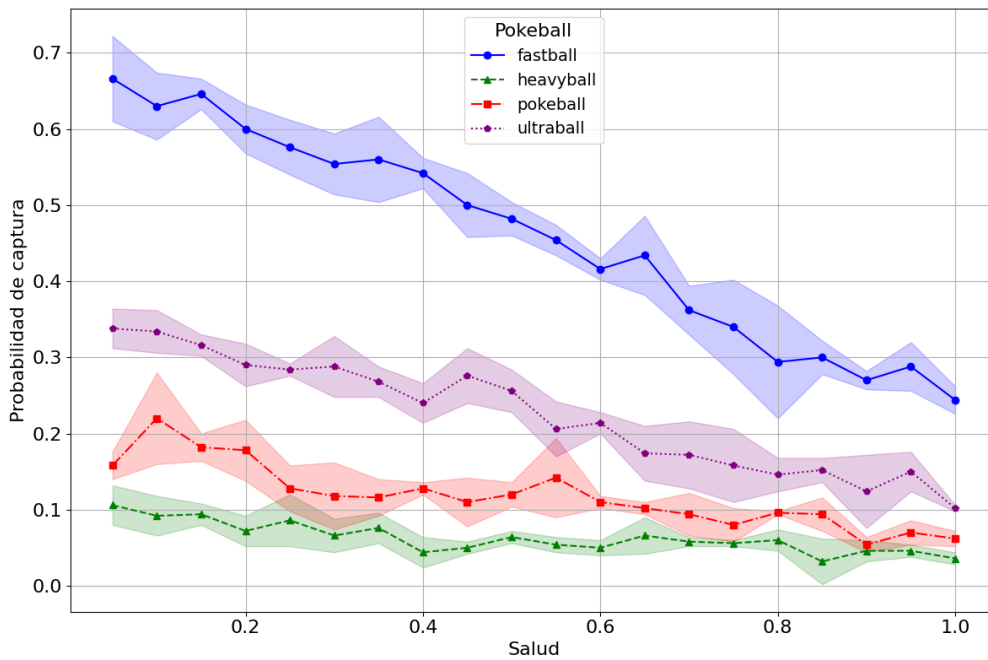


Figura 7: Probabilidad de captura de Jolteon en función de su vida actual

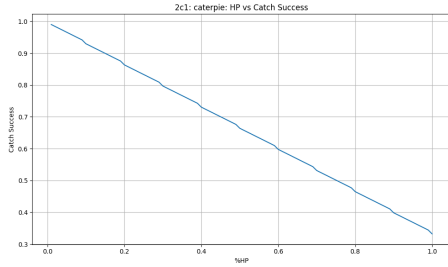
Se puede observar una tendencia lineal decreciente a medida que aumenta la vida del pokemon. Es decir, pareciera que a mayor vida actual, disminuyera su probabilidad de captura. Si analizamos el error que se obtiene con el parámetro noise al 20 % podemos afirmar que el error medido es tal que no afecta a las conclusiones extraídas. Además, esta conclusión se condice con lo que ocurre en los juegos de la franquicia, ya que para capturar a un pokemon se recomienda primero bajar su vida lo más posible, sin que llegue a cero (en cuyo caso, el pokemon muere), para luego capturarlo.

4.3 Ítem c

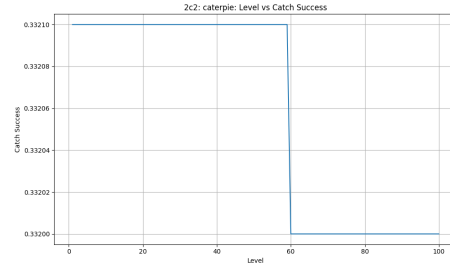
Abordar la respuesta de esta parte vamos a emplear varias observaciones.

Para empezar, discutimos la forma correcta de analizar los parámetros todos en conjunto. Esto fue difícil ya que debemos considerar los factores como variables que dependen de las otras variables para definir su efecto sobre la probabilidad de captura.

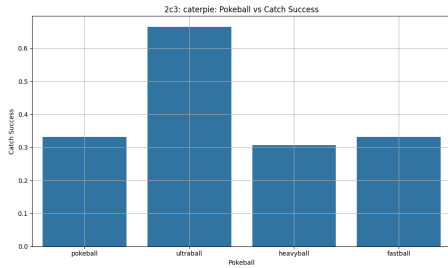
Es por esto que propusimos observar como afectan cada uno de los factores sobre la probabilidad de captura (todas las imágenes a continuación se pueden ver mejor en el anexo y tienen como factor común mostrar la probabilidad de captura en el eje y):



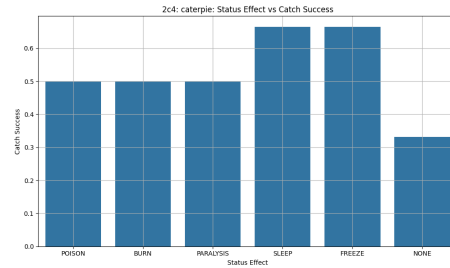
(a) Efecto de %HP sobre la probabilidad de captura.



(b) Efecto del nivel sobre la probabilidad de captura.



(c) Efecto de la Pokeball sobre la probabilidad de captura.



(d) Efecto del estado de salud sobre la probabilidad de captura.

Figura 8: Efectos sobre la probabilidad de captura para Caterpie. Ver anexo para mayor claridad de imagen (figuras 12, 13, 14, 15)

Por supuesto que estos mismos gráficos fueron reproducidos para todos los pokeballs. Nuestra idea principal fue analizar la diferencia máxima que hay entre el valor más alto de la probabilidad y el valor más bajo para cada uno de los factores y dependiendo de cuál sea la variable con mayor diferencial se consideraría que es el que mayor efecto tiene sobre la probabilidad de captura. En esta parte concluimos que a lo largo de los pokemons, el factor más relevante para la probabilidad de captura es la pokebola.

Consideramos que nuestro análisis estaba muy restringido casos particulares que muy fácilmente podría cambiar por algún caso distinto que contradiría lo dicho. Por eso decidimos "ponernos creativos" probar otras técnicas de visualización de la información.

Una idea inicial es crear un heatmap con seaborn donde sobre un eje tenemos la vida y sobre el otro el nivel. Esto nos permite ver bien la relación pero seguimos limitados a mantener una gran cantidad de gráficos ya que también tuvimos un gráfico por combinación de pokemon, status y pokeball (un total de $\#Pokemon \cdot \#Status \cdot \#Pokeball = 5 \cdot 6 \cdot 4 = 120$ gráficos). Aún así nos pareció divertido mostrar la relación y se presenta a continuación:

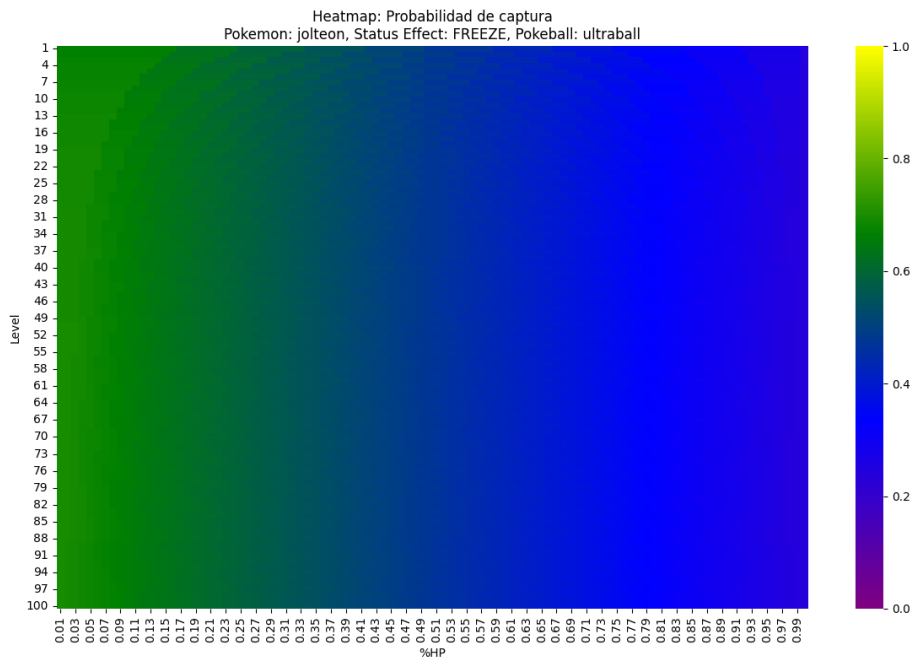
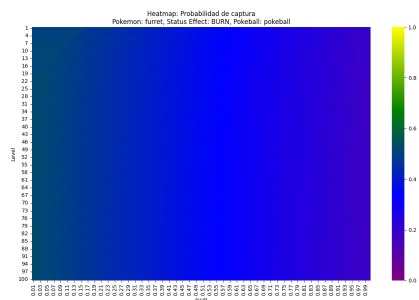


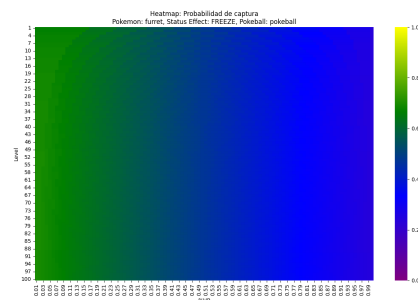
Figura 9: Probabilidad de captura de Jolteon en función del HP y el nivel.

El gráfico en particular sugiere un leve cambio (verticalmente) de probabilidad (color) para casi todas las instancias probadas. Dándose visibilidad que el nivel afecta muy poco sobre la probabilidad de captura. Aún así se produce un patrón interesante que parece circular que muestra cierto cambio en base al nivel (leve).

Comparando estos gráficos podemos ver como manteniendo los factores pokemon y pokeball cambia bastante el gráfico dependiendo del estado de salud.



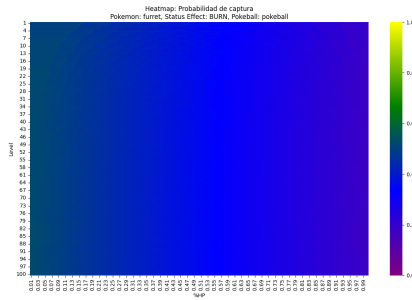
(a) Furret (nivel vs HP) con pokeball en estado BURN



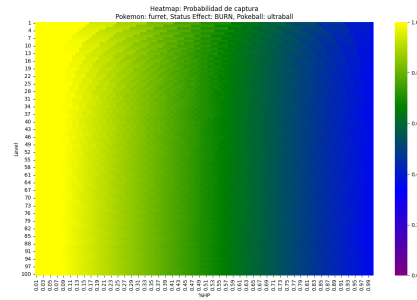
(b) Furret (nivel vs HP) con pokeball en estado FREEZE

Figura 10: Efecto del estado de salud sobre la probabilidad de captura. Ver anexo para mayor claridad de imagen (figuras 16, 17)

cambiando la pokebola y manteniendo lo demás parámetros constante existen cambios más fuertes (mirar figura 11b) que en la figura 10b sobre las probabilidades de captura.



(a) Furret (nivel vs HP) con pokeball en estado BURN



(b) Furret (nivel vs HP) con ultraball en estado BURN

Figura 11: Efecto del estado de salud sobre la probabilidad de captura. Ver anexo para mayor claridad de imagen (figuras 16, 18)

Nuevamente reforzando la hipótesis inicial que la pokebola es el factor que más afecta la probabilidad.

Como último análisis, a modo de experimentación, decidimos investigar métodos para evaluar el efecto de una variable individual sobre la variable respuesta (probabilidad de captura). La probabilidad de captura está dada por

$$P(\text{captura}) = \text{pokemon} + \text{pokebola} + \text{HP} + \text{level} + \text{status}$$

donde las variables son las que dan efecto sobre la probabilidad. Se decidió probar un análisis de la varianza¹ a.k.a. ANOVA² y con una breve introducción en python de como usarlas técnicas³ pudimos producir la siguiente tabla que nos da la respuesta definitiva:

Source	Sum qd	df	F	PR(¿F)
C(Level)	4.097794	99	2.144237e+00	3.031961e-10
C(HP_percent)	6479.656004	99	3.390584e+03	0.000000e+00
C(Status_Effect)	3619.766658	5	3.753020e+04	0.000000e+00
C(Pokeball)	5918.190489	3	1.021940e+05	0.000000e+00
C(Pokemon)	105207.229431	4	1.362522e+06	0.000000e+00
Residual	23160.454511	1199789	NaN	NaN

Cuadro 1: ANOVA para análisis de probabilidad de captura.

La tabla de salida nos da un cuadro completo con variables que permiten ver el nivel de participación de cada uno de los parámetros sobre la probabilidad de captura. Una variable que no fue analizada en los otros gráficos es la elección del pokemon que para sorpresa de nadie es el que más afecta. Vamos a descartarlo porque parece más interesante el análisis de los demás. Aquí vemos que la elección de la Pokeball es el factor más influyente sobre la probabilidad (en acuerdo con las hipótesis previas). Luego se puede ver en orden ascendente por impacto sobre la probabilidad de captura (de mas impacto a menos: estado, hp, nivel).

¹Hicimos este análisis sólo porque es el tp0 y supusimos que no había daño en jugar libremente con diferentes técnicas y modelos; aunque en próximos tps preguntaríamos antes de usar un modelo matemático de la galera.

²Análisis de la varianza - https://es.wikipedia.org/wiki/An%C3%A1lisis_de_la_varianza

³Tutorial ANOVA - <https://www.reneshbedre.com/blog/anova.html>

4.4 Ítem d

Vamos a considerar el caso de Mewtwo. Este forma parte de un grupo de pokemon conocidos como "legendarios", los cuales se distinguen del resto por sus estadísticas notablemente superiores y muy bajas probabilidades de captura. Esto se condice con lo obtenido experimentalmente en el ítem 1a, visible en la figura 2. Para evaluar la mejor combinación de captura se presupone que el jugador o entrenador esta equipado para el final del juego. Es decir, que posee al menos una Pokeball de cada tipo y que también es capaz de infligir cualquier tipo de estado. Teniendo estas condiciones en cuenta, el objetivo es simple: Mejorar la probabilidad de captura con los mayores multiplicadores posibles. En cuanto a la Pokeball, se opta por utilizar la Fastball debido a la elevada velocidad de Mewtwo. Como se puede ver en la figura 5 esta obtuvo los mejores resultados de todas las Pokeballs en capturar a este pokemon específico. En cuanto a la condicion a infligir, se tienen dos opciones que maximizan las probabilidades, 'FREEZE' y 'SLEEP'. Ambos estados duplican las mismas si se compara contra un pokemon con estado 'NONE', como se puede ver en las figuras 6.

4.5 Ítem e

Las combinaciones de parámetros de captura vistas en el ítem anterior mantendrían su efectividad independientemente del nivel. Como se ve en ítems anteriores, el nivel del Pokemon a capturar no influye de manera significativa en la probabilidad final de captura (ver 13). Los estados que un pokemon puede sufrir son independientes del nivel en sí, y las Pokeballs analizadas tienen multiplicadores constantes, como los de la Pokeball básica y la Ultraball, o sus multiplicadores dependen de propiedades intrínsecas del Pokemon, como el peso y estadística base de velocidad para la Heavyball y la Fastball respectivamente.

5 Conclusión

Pokemon es una de las franquicias más conocidas de la historia con una basta variedad de pokemons y pokeballs. En este trabajo se consideró un reducido grupo de ellos, con el fin de analizar sus propiedades y realizar un análisis estadístico. De esta manera se llegó a varias conclusiones sobre las probabilidades de captura de los pokemons y cómo las distintas variables afectan a la misma.

Podemos concluir que definitivamente hay una relación entre las propiedades intrínsecas del pokemon (como su peso y velocidad) y su probabilidad de captura en base a la pokeball utilizada. Siendo la fastball la más efectiva para pokemones de alta velocidad y la heavyball la más efectiva para pokemones con un peso elevado.

También podemos concluir que la salud afecta de manera lineal prácticamente a la probabilidad de captura de los pokemones, resultando en que un menor porcentaje de vida aumenta las probabilidades de captura, pero también conlleva riesgos (como el de matar al pokemon o que el mismo se escape).

En definitiva, este trabajo reveló que para capturar a un pokemon puede resultar más eficaz utilizar elementos que alteren la salud del pokemon y ciertas pokeballs en específico para maximizar su probabilidad de captura. Así como también, que no es única esta combinación de factores, sino que hay más de una manera de aumentar la probabilidad de captura (por ejemplo, dormirlo o congelarlo son dos maneras de afectar el estado de salud del pokemon para aumentar de igual manera su probabilidad de captura).

6 Anexo

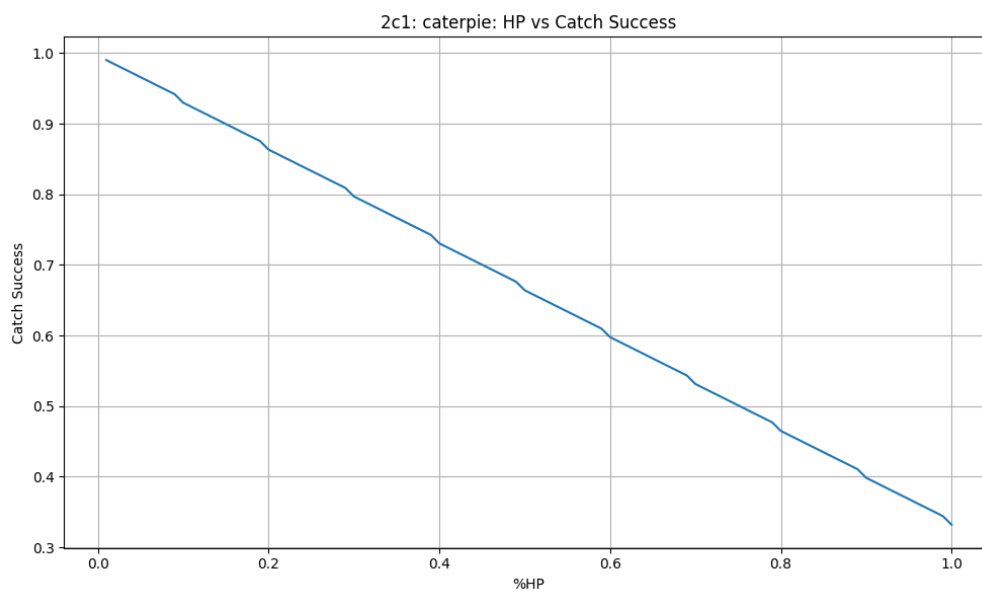


Figura 12: Efecto de %HP sobre la probabilidad de captura.

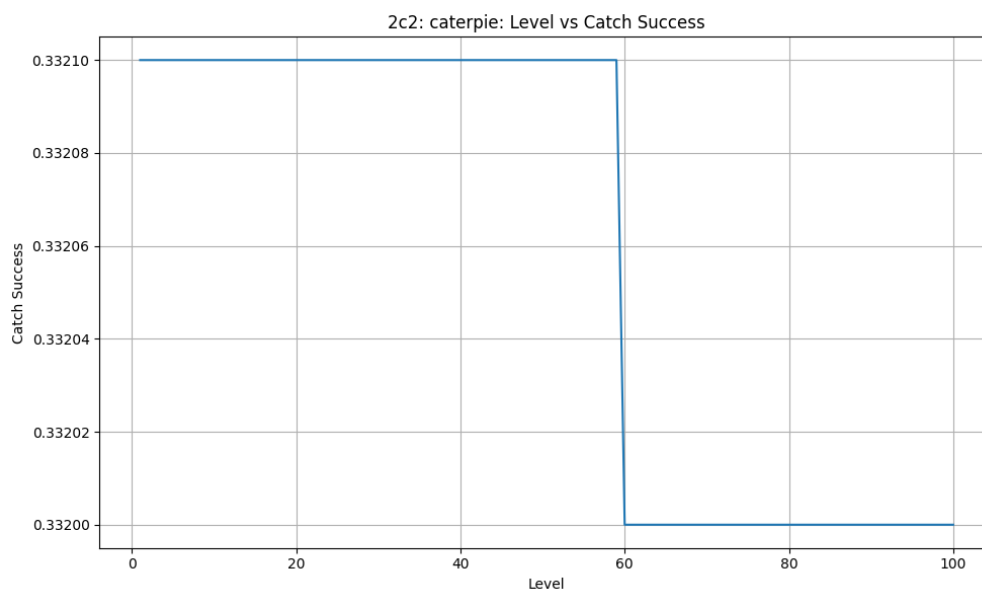


Figura 13: Efecto del nivel sobre la probabilidad de captura.

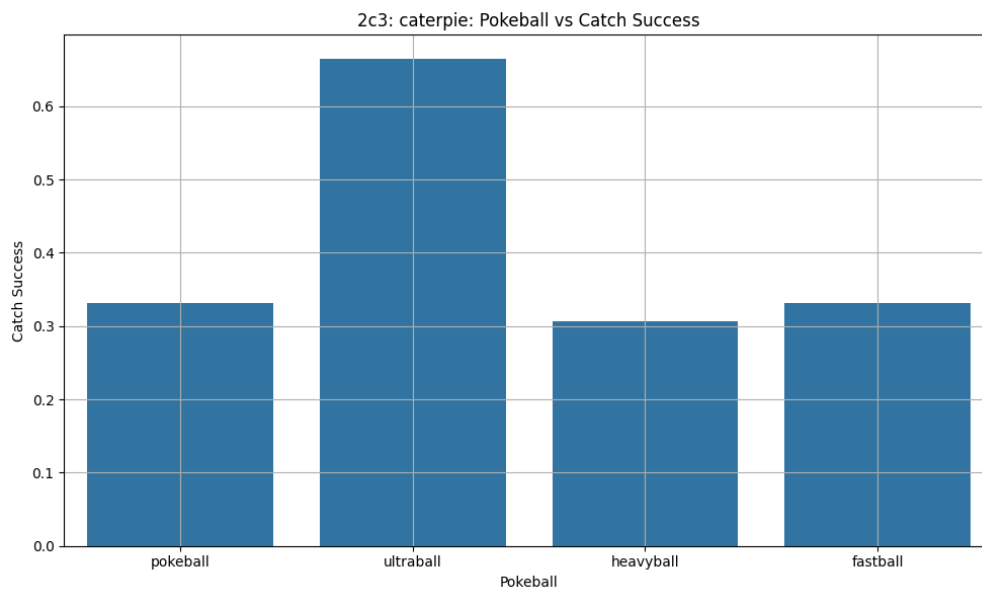


Figura 14: Efecto de la Pokeball sobre la probabilidad de captura.

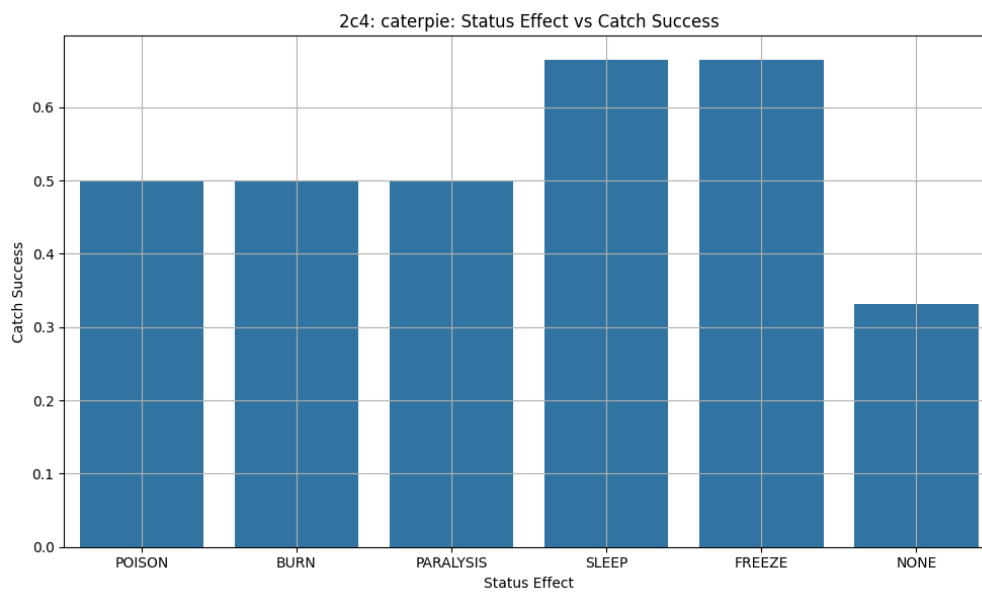


Figura 15: Efecto del estado de salud sobre la probabilidad de captura.

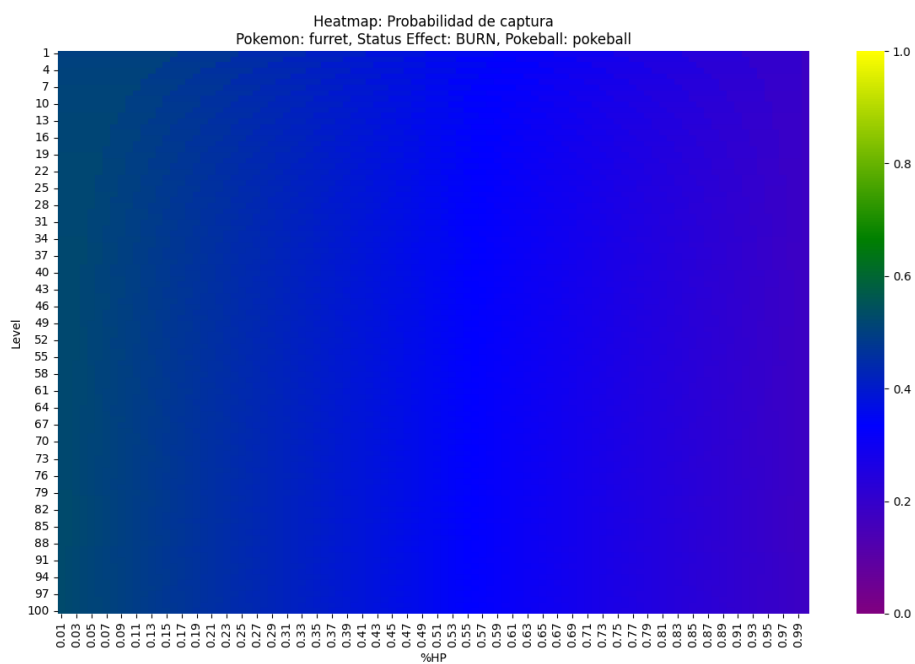


Figura 16: Furret (nivel vs HP) con pokeball en estado BURN

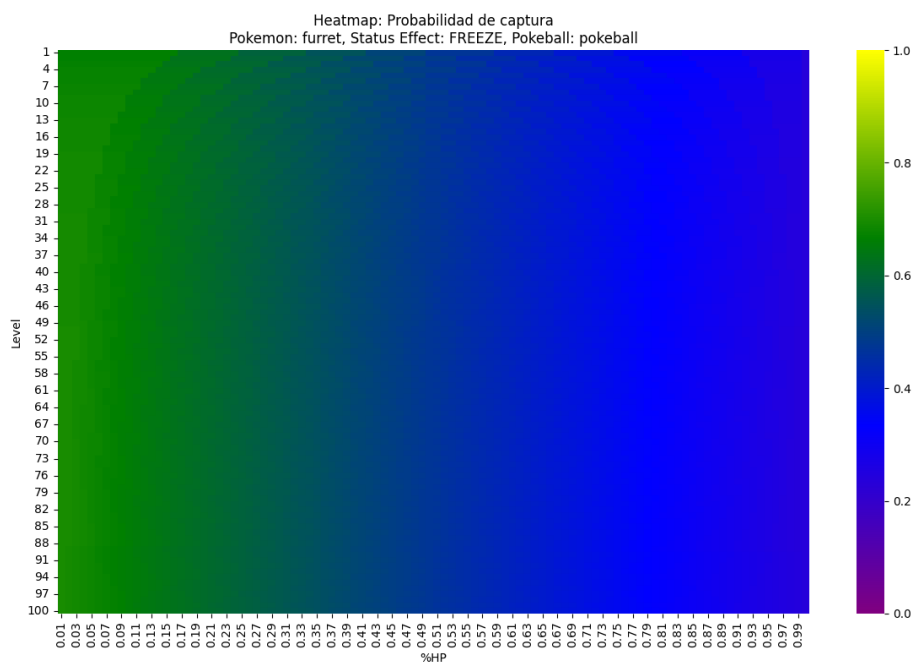


Figura 17: Furret (nivel vs HP) con pokeball en estado FREEZE

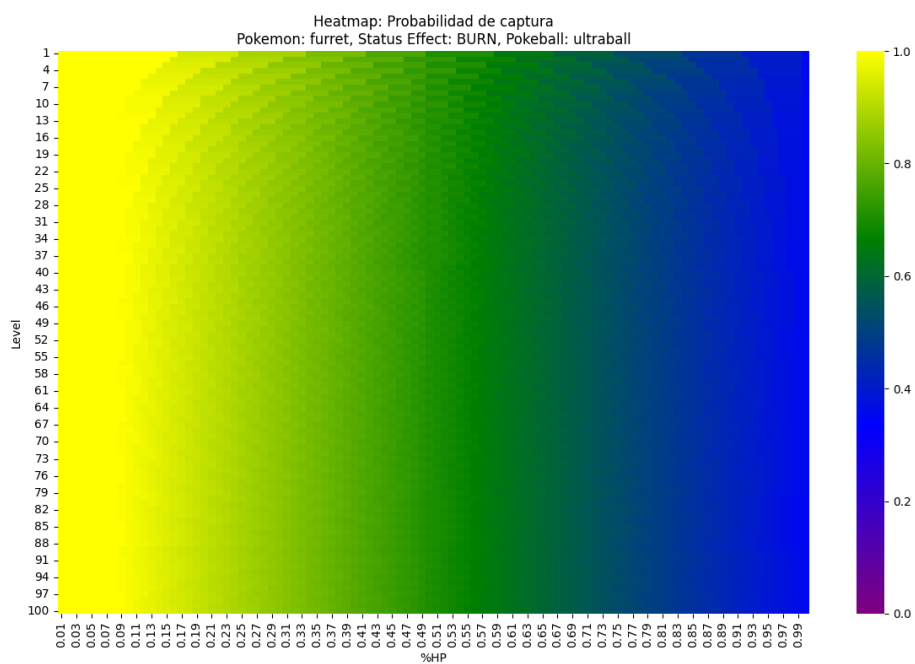


Figura 18: Furret (nivel vs HP) con ultraball en estado BURN