

Clasificación de niveles de Super Mario Maker 2 por su dificultad

Patricio Eugenio Cantú Treviño

Noviembre 2025

1. Introducción

Creatividad ilimitada es una de las mejores descripciones que le puedes dar al juego Super Mario Maker 2, ya que en este juego se le permite a los jugadores crear sus propios niveles utilizando las mecánicas de la franquicia de Super Mario, y compartirlos por internet por el resto del mundo.

Cada nivel publicado le es asignado una de 4 dificultades: Fácil, Normal, Experto y Súper Experto. No es completamente claro que es lo que define la dificultad de cada nivel, puede tener cierta relación con el ratio de éxito promedio que se tienen los jugadores al jugar el nivel, pero esta no es una división exacta.

2. Descripción de los datos

Se trabajara con un conjunto de datos [4] que contiene información de niveles de Super Mario Maker 2 de hasta el año 2022. El conjunto de datos [4] contiene mas de 26 millones de muestras, pero para este análisis se usara una muestra aleatoria de 10,000 muestras.

En el cuadro 1 se ven las variables que se usarán para el análisis y su descripción, las variables no numéricas o categóricas (a excepción de `difficulty`) fueron eliminadas.

También añadí la variable calculada `unique_clear_rate`, la cual se define como $\min\{\frac{\text{clears}}{\text{unique_players_and_versus}}, 1\}$. Añadí esta variable calculada con la creencia de que el porcentaje de completación del nivel por jugador sería un buen estimador para el nivel de dificultad.

3. Antecedentes

Furia and Mocci realizaron un análisis de regresión para predecir la dificultad de los niveles de Super Mario Maker 2 [2], ellos se enfocaron en predecir el `clear_rate`, y el uso del procesamiento de lenguaje natural.

Mientras que Drachen et al. realizaron agrupamiento a jugadores de World of Warcraft segun sus comportamientos [1].

Cuadro 1: Descripciones de las variables del conjunto de datos.

variable	descripción
difficulty	(Categórica) El nivel de dificultad del nivel
uploaded	(Entero) Cuando fue publicado el nivel
created	(Entero) Cuando fue creado el nivel
game_version	(Entero) La versión del juego en la que se creó el nivel
world_record	(Milisegundo) El tiempo récord de compleción
upload_time	(Milisegundo) El tiempo de compleción del nivel del creador
upload_attempts	(Entero) Intentos del creador para publicar el nivel
num_comments	(Entero) Cantidad de comentarios en el nivel
timer	(Segundo) Tiempo límite para completar el nivel
autoscroll_speed	(Entero) Nivel de movimiento automático de la cámara
clears	(Entero) Cantidad de intentos exitosos
attempts	(Entero) Cantidad de intentos totales
clear_rate	(Decimal) Ratio de éxito en las compleciones
plays	(Entero) Cantidad de sesiones de juego del nivel
versus_matches	(Entero) Cantidad de partidas competitivas
coop_matches	(Entero) Cantidad de partidas cooperativas
likes	(Entero) Cantidad de me gusta
boos	(Entero) Cantidad de no me gusta
unique_players_and_versus	(Entero) Cantidad de jugadores que han jugado el nivel

Cuadro 2: Estadísticos descriptivos de las variables.

	minimo	Q1	mediana	Q3	maximo	media	devs_std
uploaded	1621048585.000000	1621060323.250000	1621072574.000000	1621085646.750000	162108964.000000	1621073536.052400	15445.848509
created	1039796700.000000	1620836505.000000	1621075800.000000	1621103220.000000	2868488580.000000	1618842678.642000	40335462.936095
difficulty	0.000000	0.000000	1.000000	1.000000	3.000000	0.878800	0.862430
game_version	0.000000	5.000000	5.000000	5.000000	5.000000	4.917800	0.482977
world_record	-1.000000	9056.250000	19070.000000	39415.750000	1835882.000000	33711.396300	53235.787054
upload_time	433.000000	19295.000000	37502.500000	69960.500000	499590.000000	52789.483700	49968.051316
upload_attempts	1.000000	1.000000	1.000000	3.000000	217.000000	3.894600	9.416770
num_comments	0.000000	0.000000	0.000000	1.000000	872.000000	1.061500	13.508959
timer	10.000000	300.000000	300.000000	300.000000	500.000000	294.961000	113.816053
autoscroll_speed	0.000000	0.000000	0.000000	0.000000	2.000000	0.011000	0.138134
clears	0.000000	3.000000	7.000000	19.000000	13480.000000	23.151400	224.180811
attempts	0.000000	14.000000	35.000000	80.000000	166769.000000	131.536800	2034.905108
clear_rate	0.000000	11.538462	29.561460	54.545455	100.000000	36.136169	29.164521
plays	1.000000	6.000000	14.000000	31.000000	14135.000000	35.647000	279.984010
versus_matches	0.000000	0.000000	0.000000	0.000000	5148.000000	1.750700	70.389901
coop_matches	0.000000	0.000000	0.000000	1.000000	297.000000	0.995300	7.180239
likes	0.000000	0.000000	1.000000	3.000000	3126.000000	4.551400	55.901848
boos	0.000000	0.000000	1.000000	1.000000	641.000000	1.411500	10.454449
unique_players_and_versus	1.000000	6.000000	14.000000	30.000000	10938.000000	32.685700	231.207906
unique_clear_rate	0.000000	0.363636	0.666667	0.976744	1.000000	0.628823	0.316359

En este análisis me enfoco en predecir la categoría de dificultad a la que cada nivel pertenece, según como los jugadores han interactuado con el. Utilizo técnicas que ya se han usado con conjuntos de datos similares, pero con un enfoque diferente hacia la clasificación.

4. Metodología

4.1. Determinación número de grupos

Me decanté por el método de silueta para seleccionar la cantidad de k grupos debido su objetividad mas sencilla de comprender. El método de silueta simplemente elige la cantidad de grupos que maximice el coeficiente de silueta, el cual se define como:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

Donde $a(i)$ es la distancia promedio del punto i a los demás puntos de su propio grupo, y $b(i)$ es la distancia promedio del punto i a los puntos del grupo vecino mas cercano.

4.2. Agrupamiento

Drachen et al. obtuvieron resultados interpretables al usar el análisis de arquetipos en sus datos del comportamiento de jugadores en un videojuego [1]. Fue debido a que yo también estoy utilizando datos generados por usuarios en un videojuego que yo me decante por usar el análisis de arquetipos.

4.2.1. Análisis de arquetipos

El análisis de arquetipos forma grupos con los valores mas extremos de los mismos. El análisis de arquetipos busca minimizar la formula: [3]

$$\min \|X - CZ\|$$

Donde $X \in \mathbb{R}^{N \times d}$ es la matriz con los datos, $Z \in \mathbb{R}^{k \times d}$ es la matriz con los arquetipos, $C \in \mathbb{R}^{N \times k}$ es la matriz de coeficientes de X y Z , N es la cantidad de muestras, d es la cantidad de dimensiones de las muestras y k es la cantidad de arquetipos.

4.3. Clasificación

4.3.1. Arbol de decisión

Los árboles de decisión clasifican datos dividiendo el espacio de atributos X en regiones R_m . Cada región se asigna a una clase según la proporción de observaciones:

$$p_{mk} = \frac{1}{N_m} \sum_{x_i \in R_m} I(y_i = k)$$

La predicción elige la clase más frecuente en la región:

$$k(m) = \arg \max_k p_{mk}$$

Para evaluar la calidad de las divisiones se usan medidas de impureza como el índice Gini:

$$\text{Gini}(R_m) = \sum_{k=1}^K p_{mk}(1 - p_{mk})$$

5. Resultados

5.1. Agrupamiento

Utilizando las variables seleccionadas `upload_attempts` y `clear_rate`, obtenemos que según el método de silueta, 3 es la cantidad de grupos ideal, así como se ve en la figura 1.

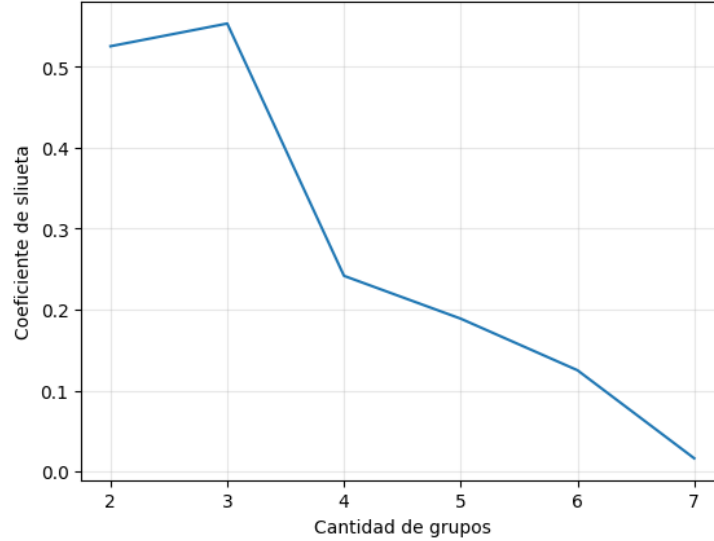


Figura 1: Resultados del método de silueta.

En la figura 2 y el cuadro 3 se encuentran los resultados del agrupamiento del análisis de arquetipos con $k = 3$.

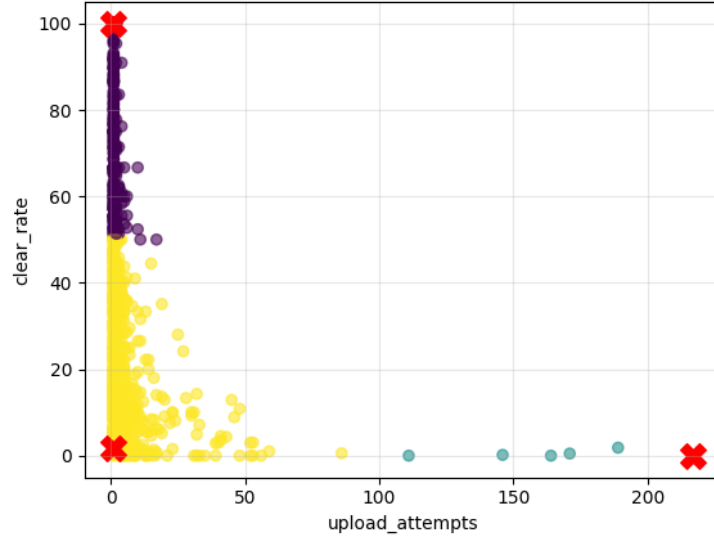


Figura 2: Grupos generados por el análisis de arquetipos.

Cuadro 3: Posiciones de los centroides/arquetipos generados.

k	upload_attempts	clear_rate
0	1.0055	99.9971
1	217	0
2	1.014	1.71

5.2. Clasificación

Con las variables `upload_attempts` y `clear_rate` se generaron múltiples árboles de decisión con distintas profundidades máximas, los cuales obtuvieron distintos niveles de precisión que se ven en la figura 1.

De los árboles de decisión generados, el que tiene una profundidad máxima de 7 fue el que generó los mejores resultados, con una precisión del 82.5 %. La figura 4 muestra el resultado del árbol de decisión, y la figura 5 muestra la matriz de discriminación de las clasificaciones del árbol de decisión.

6. Discusión

6.1. Agrupamiento

La literatura de los datos decía que la cantidad optima de grupos debería de ser 4, debido a que esa es la cantidad de dificultades que queremos clasificar. Pero el método de silueta arrojó que la cantidad optima es de 3 grupos, y por

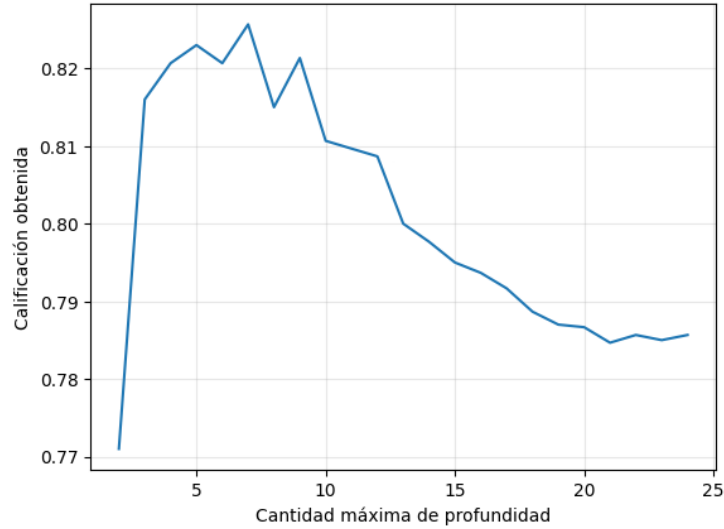


Figura 3: Calificación del árbol de decisión por la cantidad máxima de profundidad.

una diferencia substancial.

Este resultado probablemente se deba a que hay dos o mas grupos que están muy cercanos, y cuya varianza los transpone bastante.

7. Conclusiones

En este análisis se encontró una muy buena aproximación para la clasificación de la dificultad de los niveles de Super Mario Maker 2. Pero es importante notar que al modelo se le complica detectar los niveles de dificultad Experto, ya que esta tiene muchas similitudes con los niveles Normales.

Referencias

- [1] A. Drachen, C. Thureau, R. Sifa, and C. Bauckhage. A comparison of methods for player clustering via behavioral telemetry. *arXiv preprint arXiv:1407.3950*, 2014.
- [2] C. A. Furia and A. Mocci. What makes a level hard in super mario maker 2? In *2025 IEEE Conference on Games (CoG)*, page 1–8. IEEE, Aug. 2025. doi: 10.1109/cog64752.2025.11114357. URL <http://dx.doi.org/10.1109/cog64752.2025.11114357>.
- [3] E. Gomede. Deciphering Data Extremes: A Practitioner’s Guide to Archetypal Analysis, 4 2024.

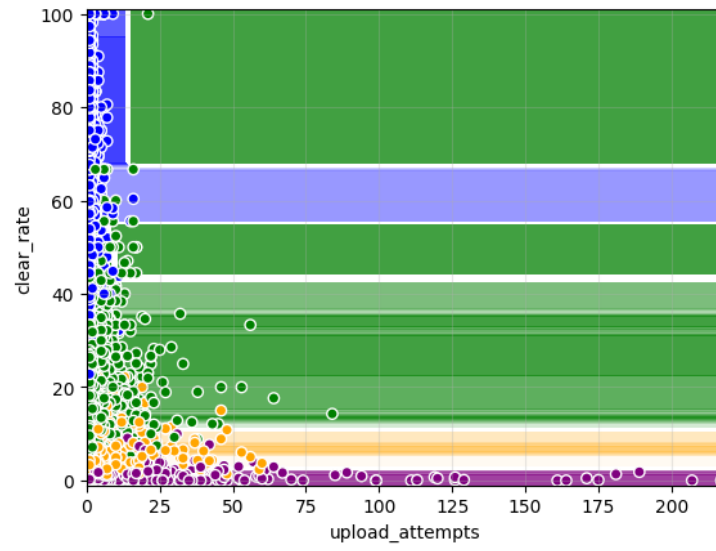


Figura 4: Dispersión de los datos con las áreas del árbol de decisión.

- [4] TheGreatRambler. mm2_level. https://huggingface.co/datasets/TheGreatRambler/mm2_level, 2022.

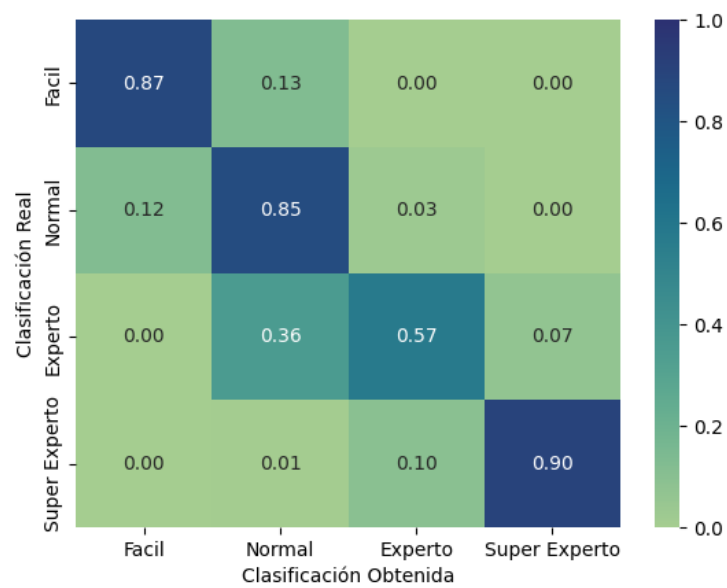


Figura 5: Matriz de discriminación con las clasificaciones del árbol de decisión.