



GOOGLE APP - GAMES

Análisis de Sentimientos...

¿Qué hace a un juego exitoso?

16 diciembre de 2023

Índice

01. Origen de la Info

De qué bases de datos partimos para hacer el análisis.

02. ¿Por qué sólo GAMES?

Limpieza y Análisis
Exploratorio de Datos

03. Frases Características

Análisis de Bigramas usando Wordcloud y conteo de Bigramas comunes.

04. Palabras importantes

Análisis de Palabras Importantes y Relación de Palabras mediante Similaridad del Coseno.

05. Clasificación Sentimientos

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh.

01. Origen de la Info

Para el análisis se usaron 2 bases de datos proporcionadas por Kaggle*, referente a aplicaciones en la Google Play Store. Ambas se consolidaron en una sola para contener la mayor cantidad de información para el modelo final de determinación de sentimientos de los comentarios de los usuarios.

	App	Translated_Review	Sentiment	Sentiment_Polarity	Sentiment_Subjectivity
0	10 Best Foods for You	I like eat delicious food. That's I'm cooking ...	Positive	1.00	0.533333
1	10 Best Foods for You	This help eating healthy exercise regular basis	Positive	0.25	0.288462
2	10 Best Foods for You	NaN	NaN	NaN	NaN
3	10 Best Foods for You	Works great especially going grocery store	Positive	0.40	0.875000
4	10 Best Foods for You	Best idea us	Positive	1.00	0.300000

[googleplaystore_user_reviews.csv](#)

Reseñas "más relevantes" de cada aplicación

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4.1	159	19M	10,000+	Free	0	Everyone	Art & Design	January 7, 2018	1.0.0	4.0.3 and up
1	Coloring book moana	ART_AND_DESIGN	3.9	967	14M	500,000+	Free	0	Everyone	Art & Design;Pretend Play	January 15, 2018	2.0.0	4.0.3 and up
2	U Launcher Lite – FREE Live Cool Themes, Hide ...	ART_AND_DESIGN	4.7	87510	8.7M	5,000,000+	Free	0	Everyone	Art & Design	August 1, 2018	1.2.4	4.0.3 and up
3	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	25M	50,000,000+	Free	0	Teen	Art & Design	June 8, 2018	Varies with device	4.2 and up
4	Pixel Draw - Number Art Coloring Book	ART_AND_DESIGN	4.3	967	2.8M	100,000+	Free	0	Everyone	Art & Design;Creativity	June 20, 2018	1.1	4.4 and up

[googleplaystore.csv](#)

Detalles de las aplicaciones en Google Play

Se eliminaron datos vacíos y duplicados y se aseguró la integridad de la información para, al final, quedarnos con:

**25,653 reviews,
distribuidas en 812 Apps.**

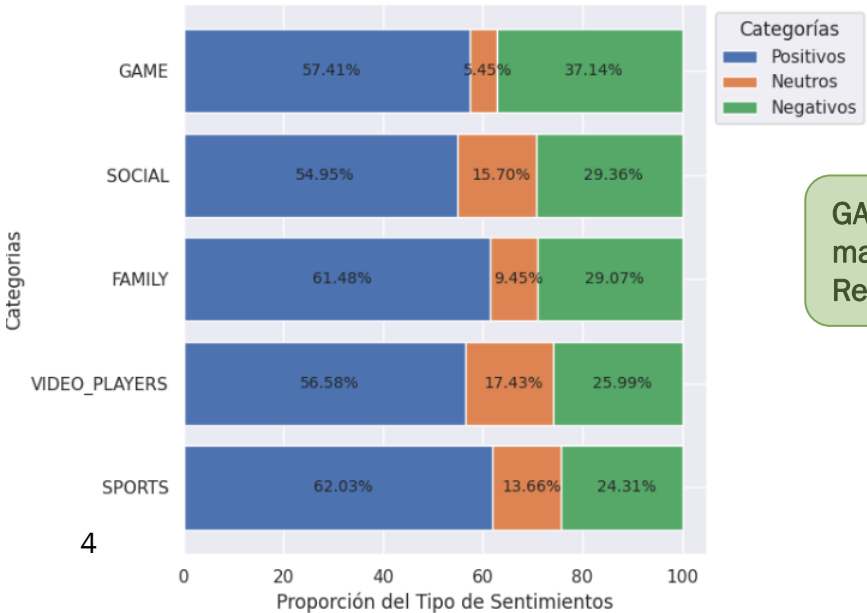
02. ¿Por qué solo GAMES?

Reviews por Categorías



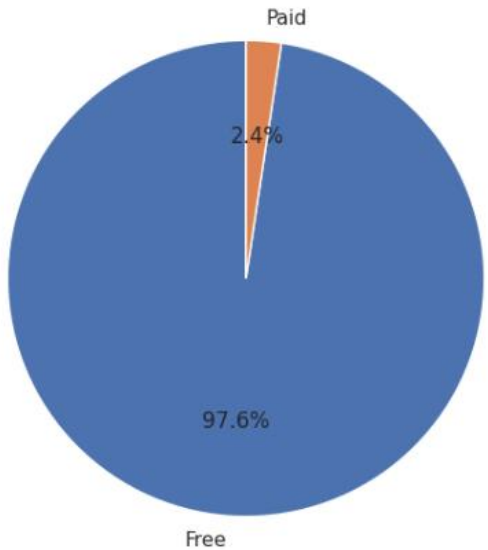
GAMES es la categoría más importante, dado que tiene mayor cantidad de Reviews. (>10%)

Proporción del tipo de Sentimientos por Categorías



GAMES tiene la mayor proporción de Reviews Negativas.

Distribución de Apps Free vs Paid en "GAMES"

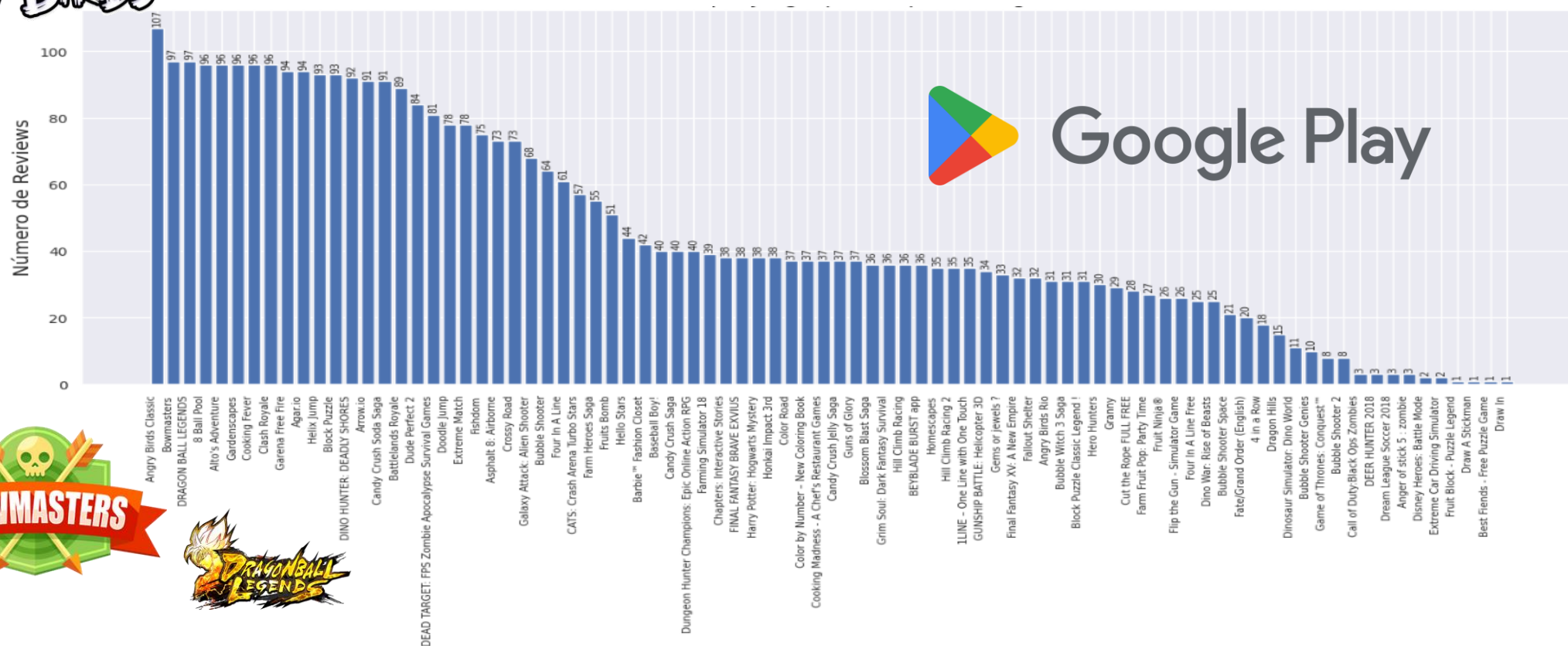


La mayor parte (97.6%) de las Apps de GAMES son gratis.

Solo por curiosidad...



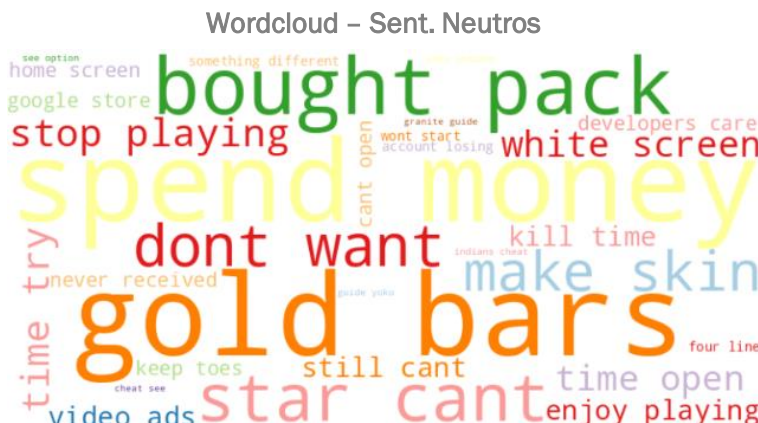
Número de Reviews por Juego que compone categoría "GAMES"



Google Play

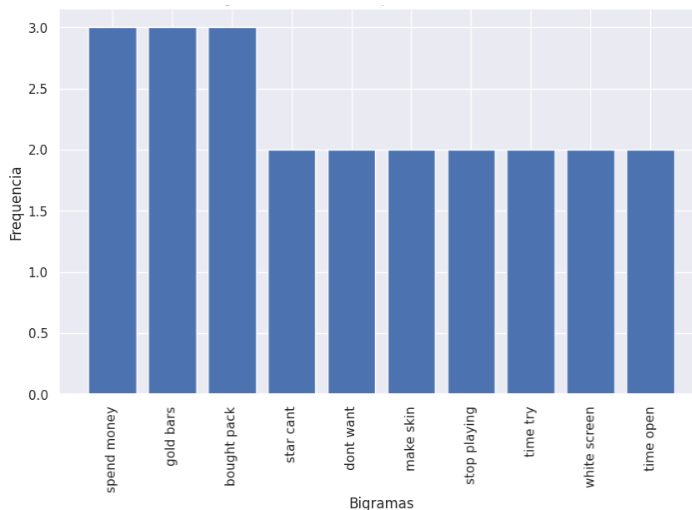


03. Frases características por Sentimiento



Comentarios comunes:

- El juego es una **pérdida de tiempo**.
- A lo **largo del tiempo**... que he usado la app
- **Dinero Real** ~ Niveles
- Publicidad **realmente molesta** (o)
- **No desperdicias**... (tiempo, dinero...)
- Problemas con la **última actualización**.



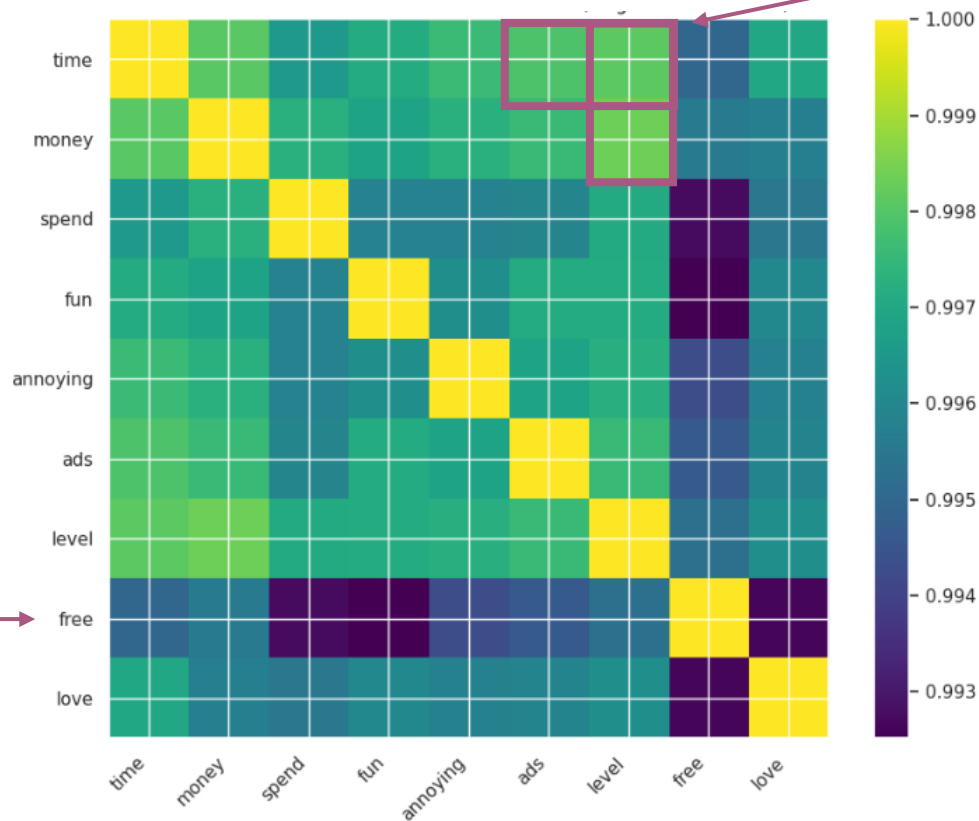
Comentarios comunes:

- El juego es realmente bueno y divertido
- Buen pasatiempo
- Lo había jugado alguna vez...
- Bueno para pasar el tiempo
- Power ups (potenciador)

6 En todos los casos aparece “Spend Money” como uno de los Bigramas fundamentales (En sentimientos positivo y negativo aparece en segundo lugar del conteo y en neutro en el primero).

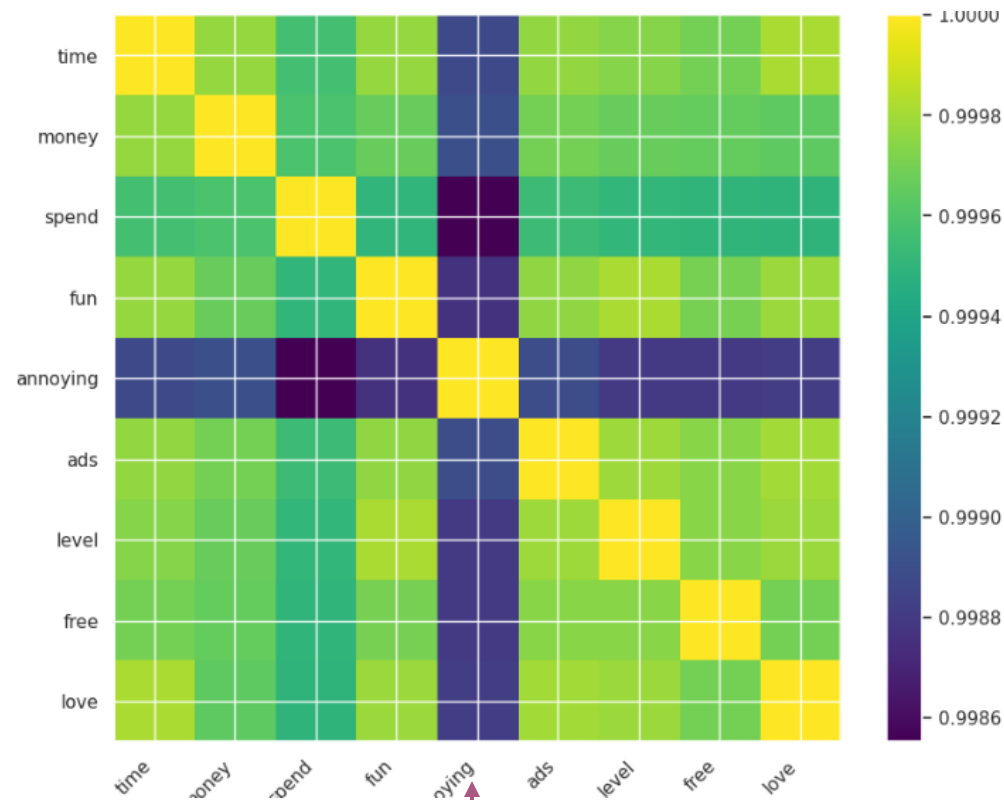
04. Palabras importantes.

Similitud Coseno entre Vectores de Palabras- Sent. Negativos



Tiempo y dinero está relacionado con ads y level, siendo los que tienen mayor relación en los sentimientos negativos.

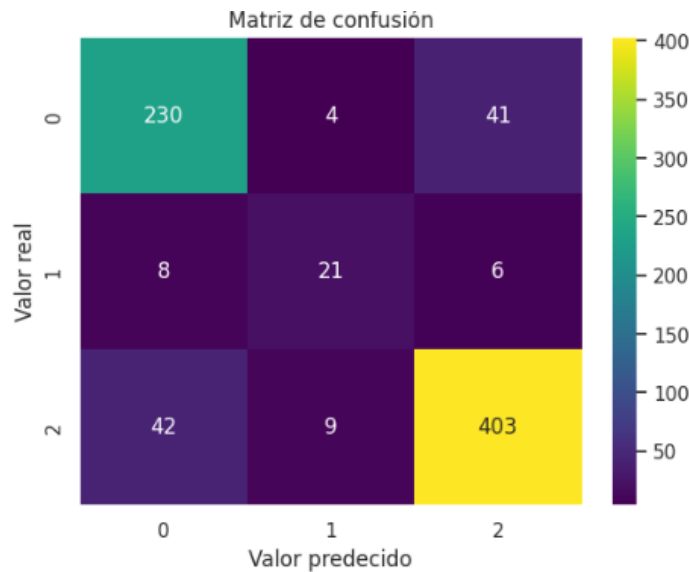
Similitud Coseno entre Vectores de Palabras- Sent. Positivos



Es la menos positiva, sobre todo cuando se relaciona con gasto.

05 Clasificación de Sentimientos.

TPOT (Tree-based Pipeline Optimization Tool) es una herramienta de optimización automática que busca automáticamente el mejor modelo y la mejor configuración de hiperparámetros para un problema específico. Con él se entrenó un modelo que da los siguientes resultados. (Utiliza Random y Grid Search) -> Algoritmo Genético.



Matriz de Precisión

```
1 print(classification_report(tpot_pred,y_test))
```

	precision	recall	f1-score	support
0	0.84	0.82	0.83	280
1	0.60	0.62	0.61	34
2	0.89	0.90	0.89	450
accuracy			0.86	764
macro avg	0.77	0.78	0.78	764
weighted avg	0.86	0.86	0.86	764

Ejemplos:

```
review_test = "AB for me has been really good in terms of learning, in all aspects, but there are some parts that I wish that could be changed. \n In the newest update, they removed the bar telling you how long of that specific achievement you need to advance to the next level \n (i.e. early riser achievement) so I wish that could be added back so I can keep track of my progress. \n But overall, 4 stars."
```

```
review_test = remove_punctuation(review_test)\n# Debemos hacer que nuestro texto de test pase por el mismo proceso que todo el dataset de train\nreview_convert = vectorizer.transform([review_test])\nprediction_test = lr_model.predict(review_convert)
```

```
array([2])
```



```
1 review_test = "I have had an account for over 8 years, and the last few updates have made this app less usable and less helpful. \n 2 The audio does not pick up at times even when you're in a quiet room, there is no explanation for why stuff is wrong sometimes. \n 3 The animations they added to show progress on daily quests are just drawn out for unknown reasons. \n 4 But lucky for all of us, rather than addressing these issues, they just added an avatar creator feature that is just so weird and also feels unnecessary." \n 5 \n 6 review_test = remove_punctuation(review_test)\n 7 # Debemos hacer que nuestro texto de test pase por el mismo proceso que todo el dataset de train\n 8 review_convert = vectorizer.transform([review_test])\n 9 prediction_test = lr_model.predict(review_convert)\n10 prediction_test
```

```
array([0])
```


Conclusiones

A través de los comentarios de los usuarios se pueden encontrar elementos que definen que un Juego en la App de Google Play sea exitoso o no.

El modelo encontrado permite clasificar correctamente los sentimientos de los usuarios para alimentar una base de datos que detecte el efecto de los cambios de sentimientos de los mismos ante una modificación mayor (new release) para saber si fueron favorables o no.

También podría servir para saber si hay errores que están comenzando a ocurrir en tiempo real (degradación) y su efecto en la percepción del juego.





Gracias

Luis Eduardo Candaudap Camacho
candaudap@gmail.com



El año pasado
