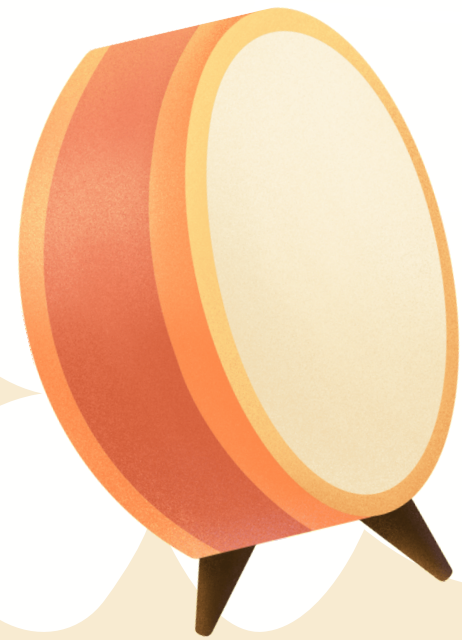


PROYECTO DEL SEMESTRE (VERSIÓN DEFINITIVA)

IAIA

BY JUAN PABLO DAZA PEREIRA
JUAN SEBASTIAN CAMARGO SANCHEZ

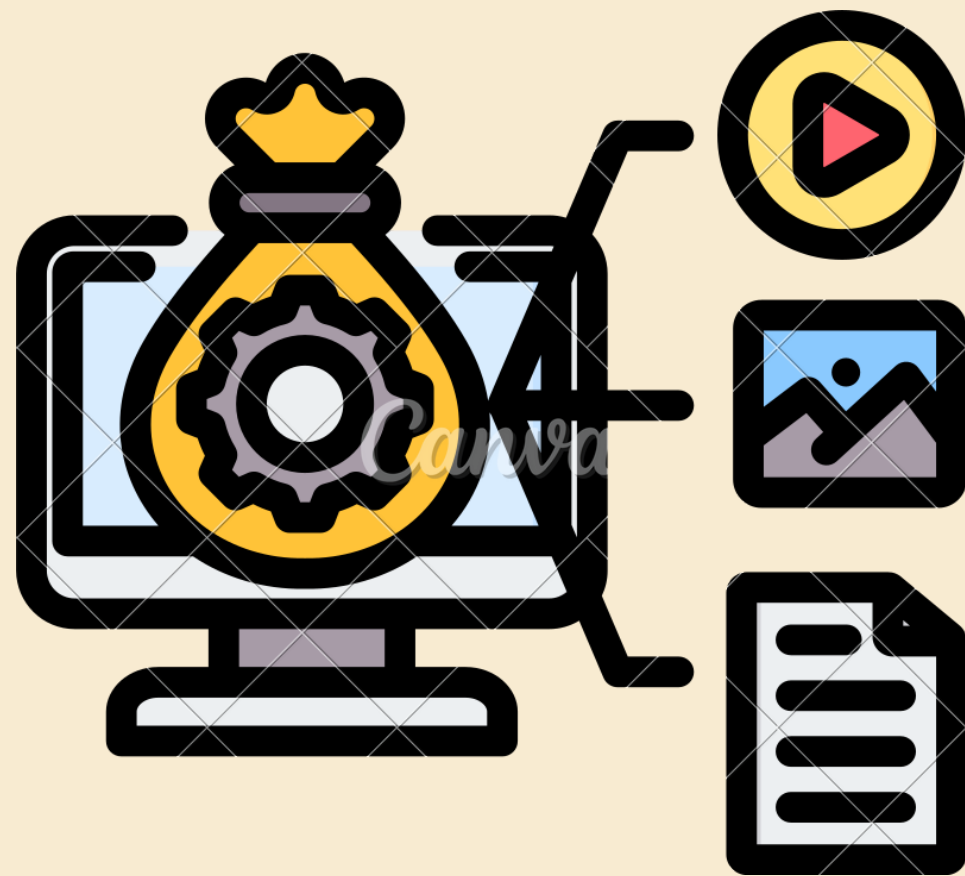




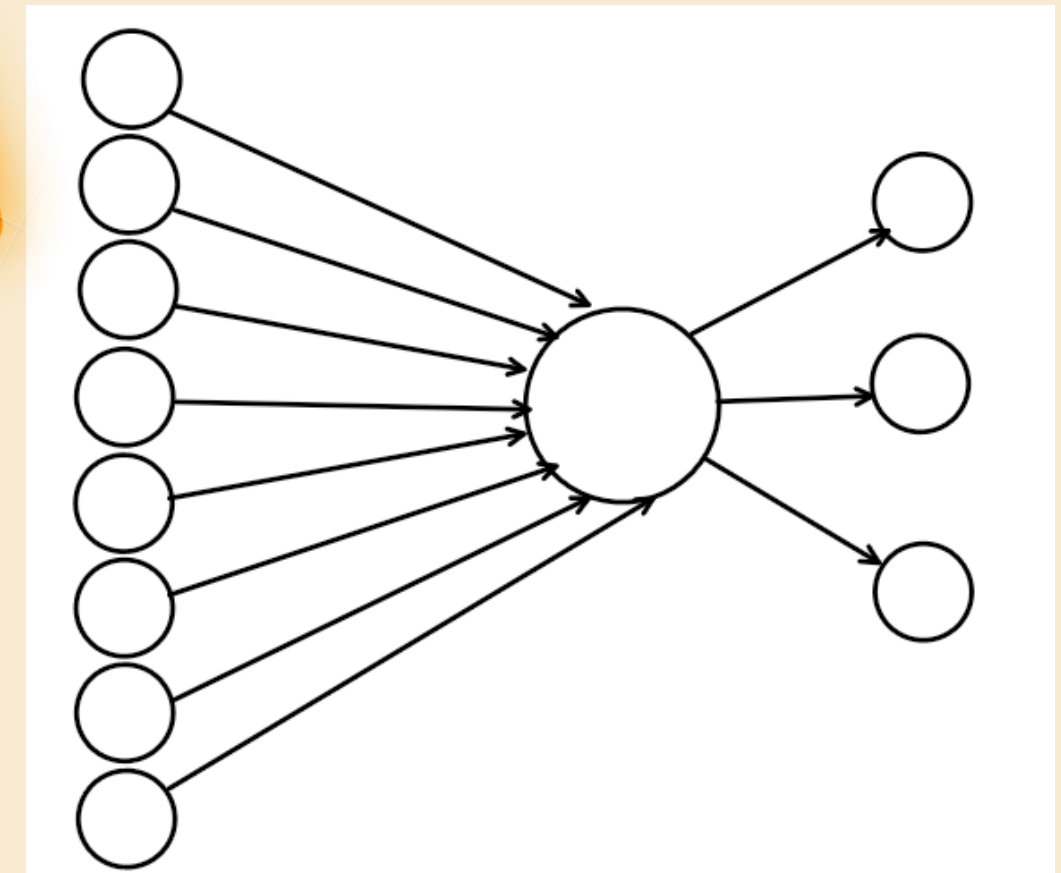
RETE MUSICALE 2.0



¿CUÁL ES EL PROBLEMA?



CATEGORIZACIÓN

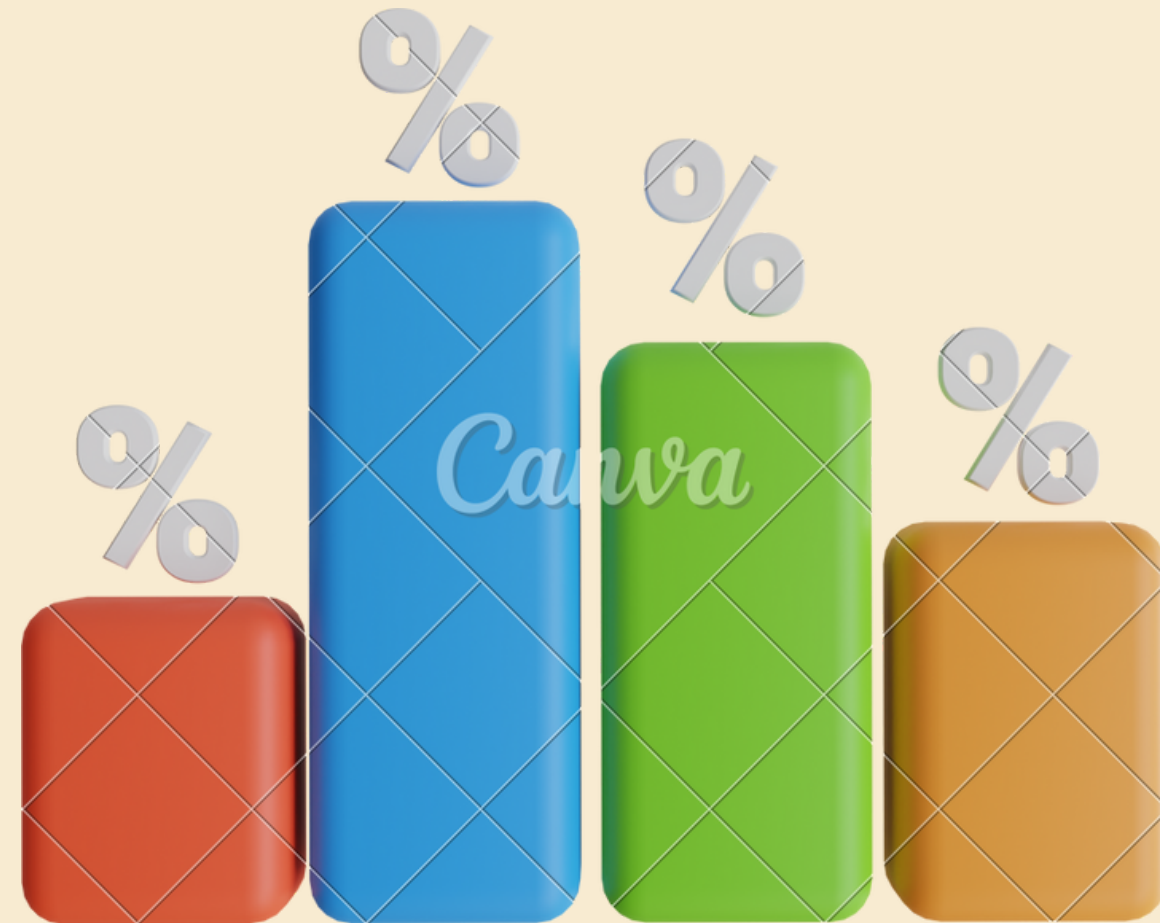


REESTRUCTURACIÓN

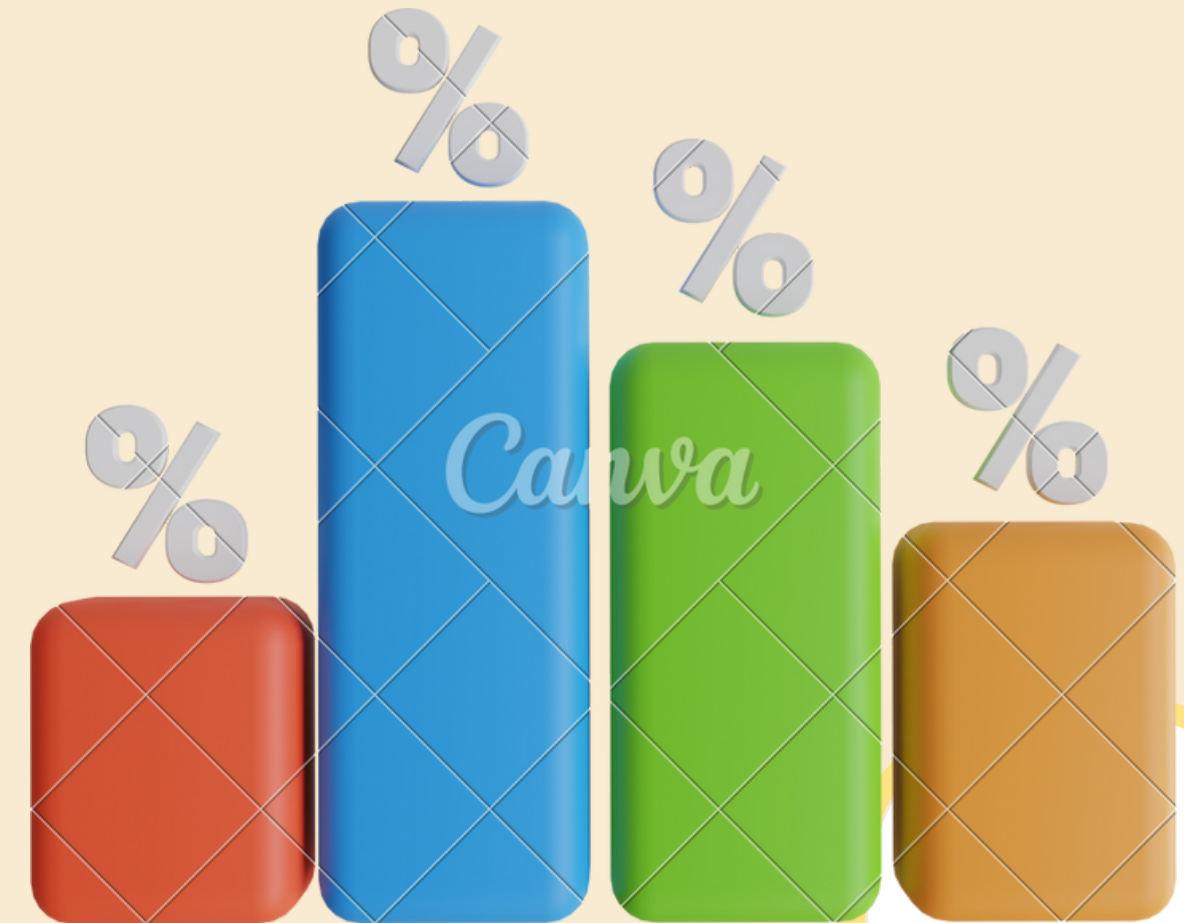
¿POR QUÉ ES IMPORTANTE RESOLVERLO?

INGENIERÍA DE
CARACTERÍSTICAS

PRUEBAS



PRECISIÓN DEL 65%

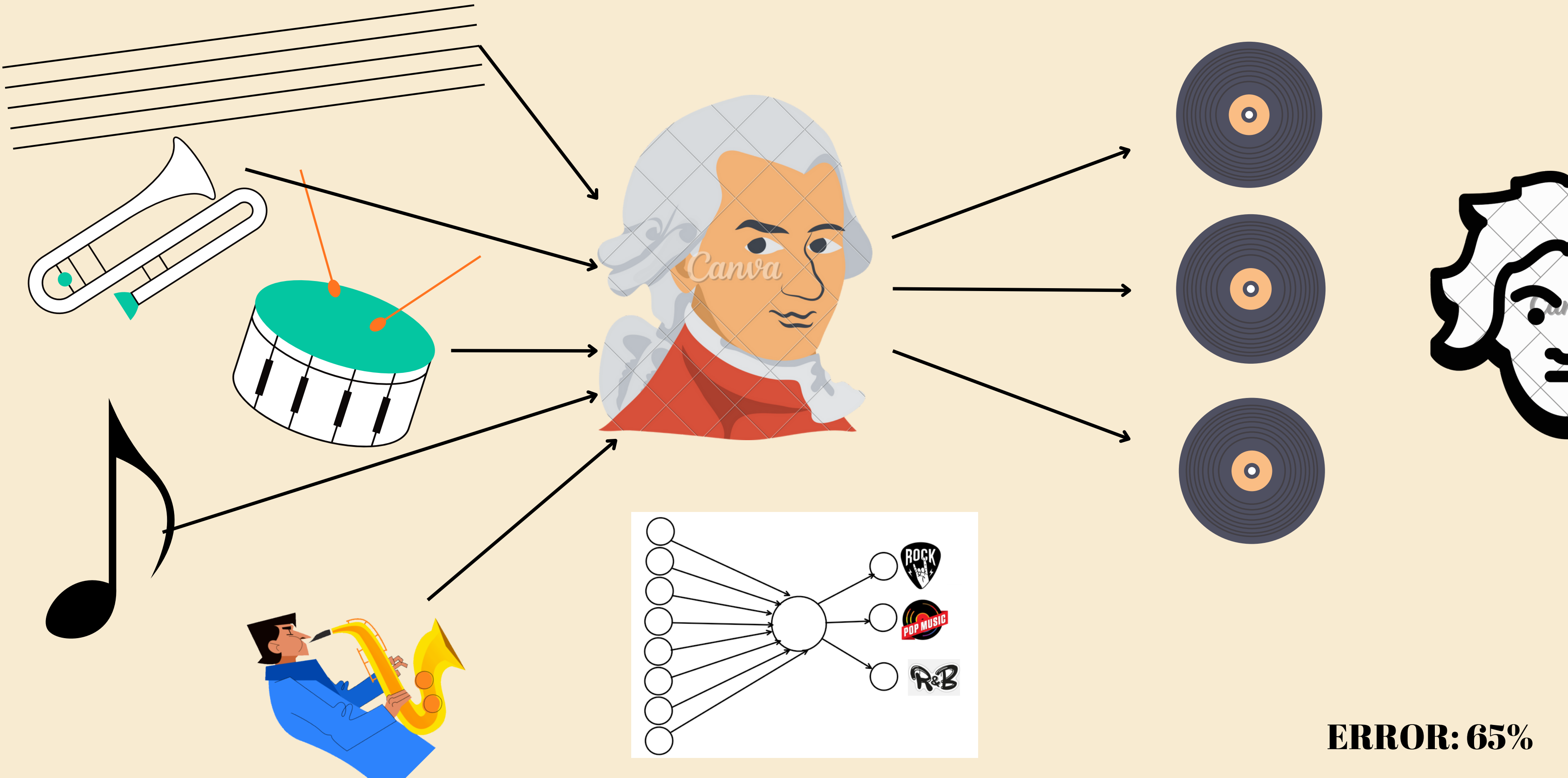


PRECISIÓN DEL 75%

¿QUÉ SE HA HECHO AL RESPECTO? (LIMITACIONES)

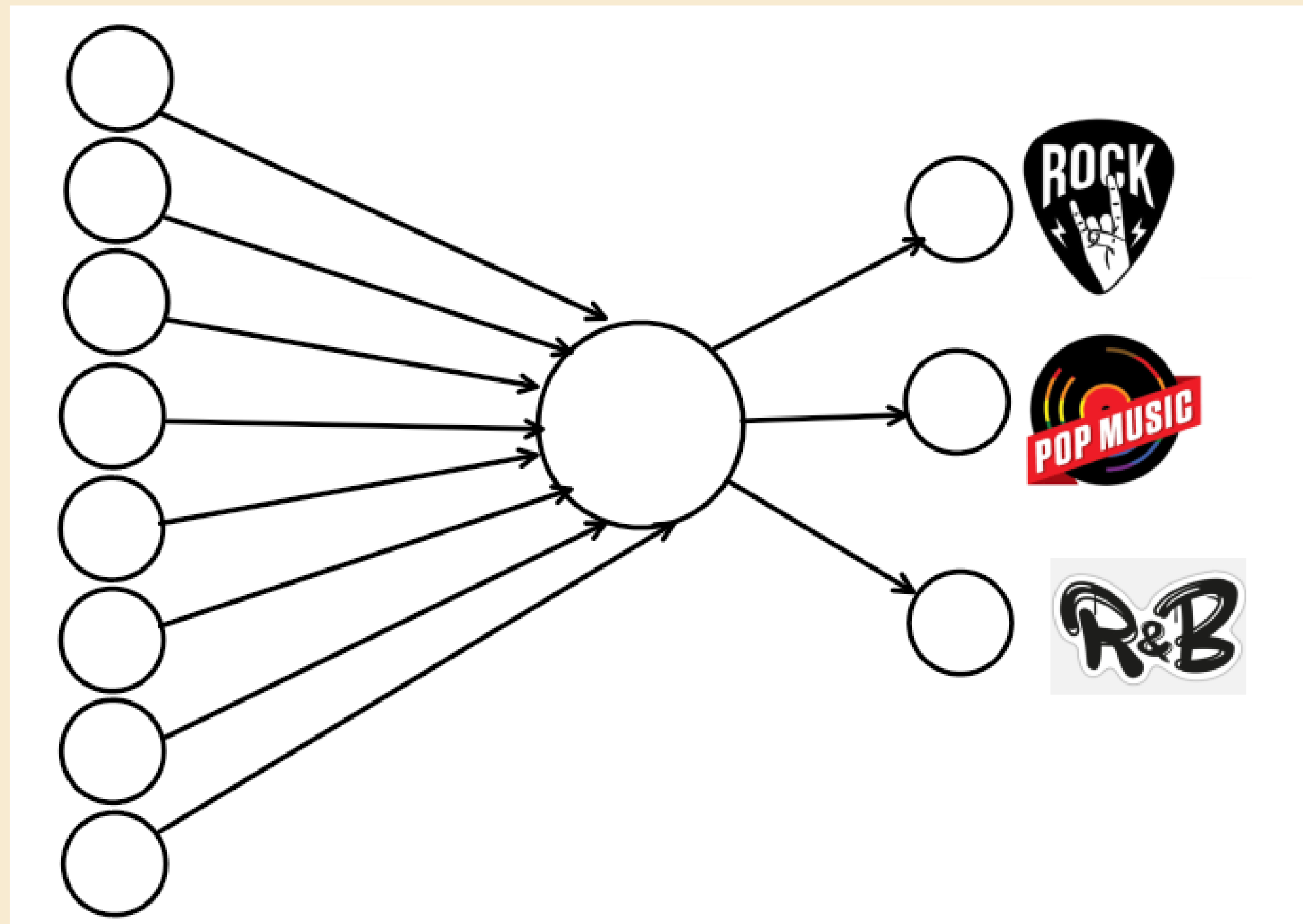


PRIMERA VERSIÓN DE LA RETE MUSICAL



ERROR: 65%

¿CUÁL FUE LA IDEA DE SOLUCIÓN?



REDUCCIÓN DE LAS ENTRADAS POR MEDIO DE REINGENIERÍA

¿CUÁL FUE LA IDEA DE SOLUCIÓN?



INGENIERIA DE SELECCION

```
1 import pandas as pd
2 from sklearn.ensemble import RandomForestClassifier
3
4 # Load your dataset (replace 'your_dataset.csv' with your actual file name)
5 df = pd.read_csv('/content/sample_data/spotify_songs.csv')
6
7 # Separate features and target variable
8 X = df.drop(['track_id', 'track_name', 'track_artist', 'track_popularity', 'track_album_id', 'track_album_name'],
9 y = df['playlist_genre'] # Target variable column
10
11 # Create a random forest classifier
12 clf = RandomForestClassifier(random_state=42)
13
14 # Fit the classifier to your data
15 clf.fit(X, y)
16
17 # Get feature importances from the trained model
18 feature_importances = clf.feature_importances_
19
20 # Create a DataFrame to display feature importances
21 feature_importance_df = pd.DataFrame({
22     'Feature': X.columns,
23     'Importance': feature_importances
24 })
25
26 # Sort the DataFrame by importance in descending order
27 feature_importance_df = feature_importance_df.sort_values(by='Importance', ascending=False)
28
29 # Print the feature importance DataFrame
30 print("Feature Importance:")
31 print(feature_importance_df)
```

SKLEAR

Feature Importance:

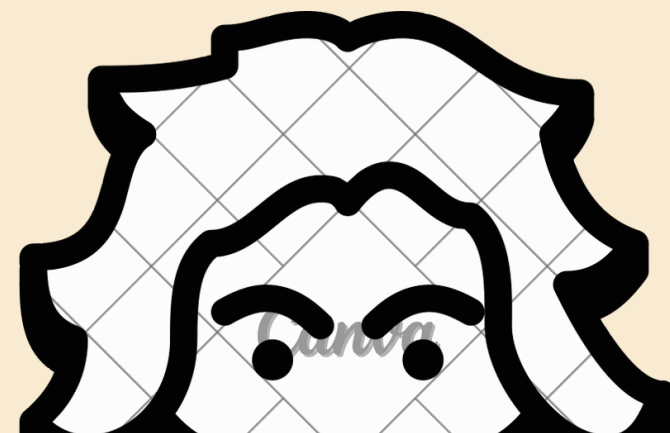
	Feature	Importance
10	tempo	0.122339
5	speechiness	0.117409
0	danceability	0.113161
1	energy	0.096658
11	duration_ms	0.090055
9	valence	0.089920
6	acousticness	0.089616
3	loudness	0.084848
7	instrumentalness	0.073664
8	liveness	0.065454
2	key	0.043470
4	mode	0.013405

¿CUÁL FUE LA SOLUCIÓN?

CON RESPECTO A LOS RESULTADOS DE LA INGENIERIA DE SELECCION:

```
Feature Importance:
      Feature  Importance
10      tempo    0.122339
5    speechiness  0.117409
0    danceability 0.113161
1      energy    0.096658
11   duration_ms 0.090055
9      valence   0.089920
6    acousticness 0.089616
3      loudness  0.084848
7  instrumentalness 0.073664
8      liveness  0.065454
2          key    0.043470
4          mode   0.013405
```

**SE SELECCIONARON LAS 3 ENTRADAS MÁS IMPORTANTES
Y SE IMPLEMENTO LA RED NEURONAL**



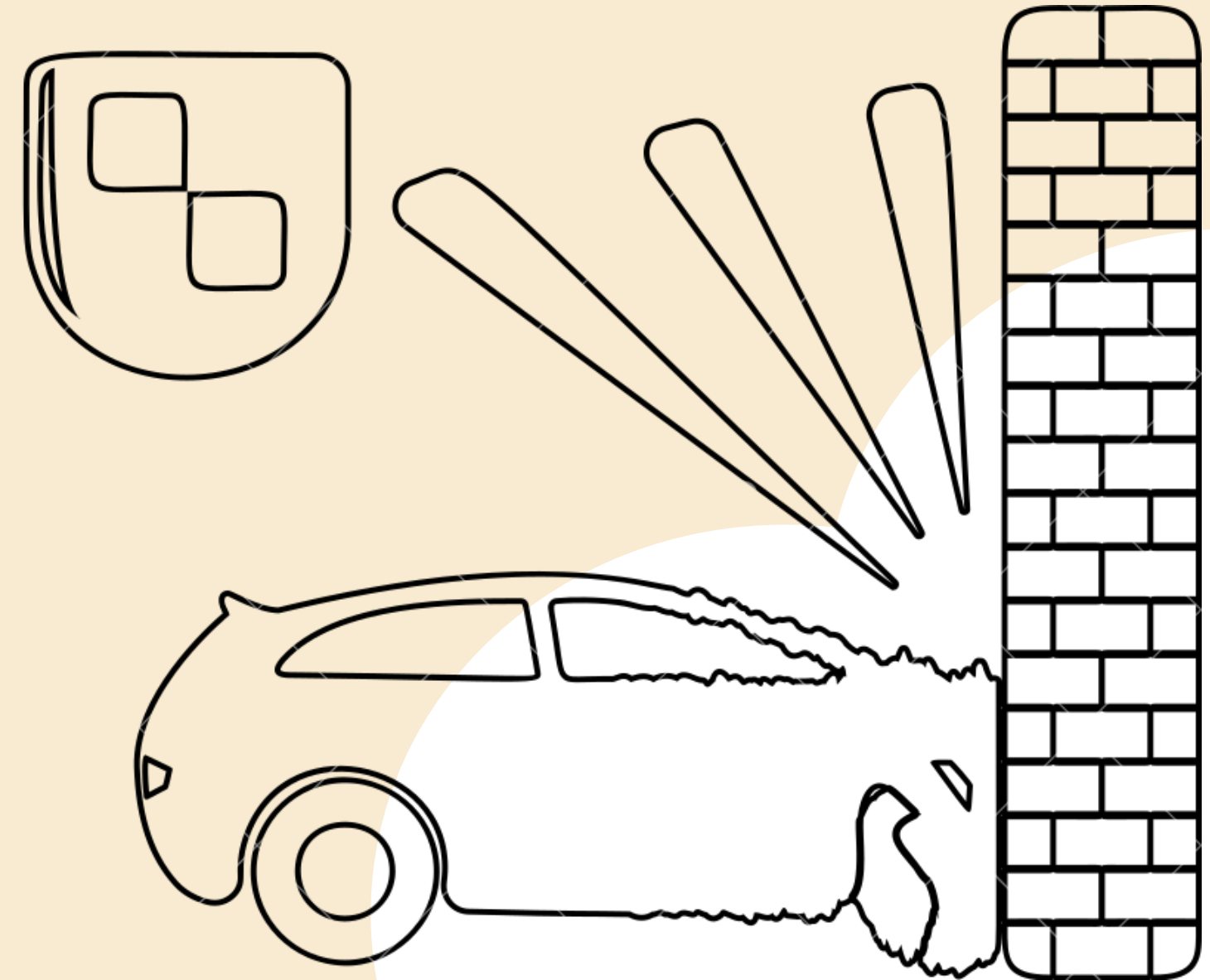
¿QUÉ RESULTADOS HAN OBTENIDO?

```
38 parameters1=train(allGenresInput, Y, [3,4,3], 0.1, 10000, 5000)
(3, 12715)
(3, 12715)
Cost after iteration 0: 2.0824554538748408
Cost after iteration 5000: 1.9078366035720582
Cost after iteration 10000: 1.907916996442692
```

```
38 parameters1=train(allGenresInput, Y, [3,4,3], 0.1, 10000, 5000)
(3, 12715)
(3, 12715)
Cost after iteration 0: 2.0824554538748408
Cost after iteration 5000: 1.9078366035720582
Cost after iteration 10000: 1.907916996442692
```

```
38 parameters1=train(allGenresInput, Y, [3,2,4,2,5,3], 0.1, 10000, 5000)
(3, 12715)
(3, 12715)
Cost after iteration 0: 2.086466593154675
Cost after iteration 5000: 1.9079136286867258
```

```
38 parameters1=train(allGenresInput, Y, [3,12,3], 0.1, 10000, 5000)
(3, 12715)
(3, 12715)
Cost after iteration 0: 2.0815460917848605
Cost after iteration 5000: 1.9060936679529776
Cost after iteration 10000: 1.9073687277981004
```



ERROR: 65%

¿QUÉ PODEMOS APRENDER DE LOS RESULTADOS?

Ya que ninguna de las redes tuvo alguna mejora significativa con respecto a nuestro anterior proyecto y abandonamos la posibilidad de generar redes independientes para cada género. Creemos que deberíamos darle es un enfoque a lo que se aprendió sobre la ingeniería de selección de datos y a identificar problemas para una red. En este caso fueron de selección de entrada, arquitectura e incluso en la elección misma de la base de datos de nuestra red.



¿CUÁLES SON LAS CONTRIBUCIONES PRINCIPALES?



Clasificación de géneros musicales.



Investigación en música computacional.



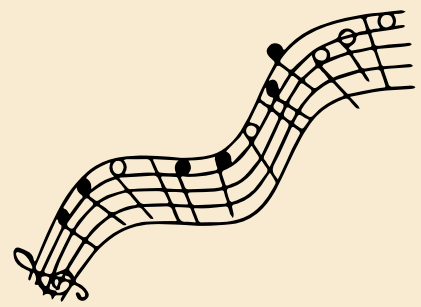
Análisis musical.



¿CUÁLES SON LAS LINEAS DE TRABAJO FUTURO?



Aplicación para más géneros de música



Recomendación de música personalizada



Análisis de tendencias musicales



Mejora en la organización de bibliotecas de música





MUCHAS GRACIAS

¿PREGUNTAS?

