

KAI-U

Sistema Recomendador y Modelo Predictivo

para fomentar el consumo audiovisual

español del anime



Dirigido por:

- Iñaki Fernández Pérez

Presentado por:

- Nicolás Felipe Trujillo Montero
- Jesús Carlos Avecilla de la Herrán

ÍNDICE

- 01 – Introducción
- 02 - Contexto
- 03 - Estado del Arte
- 04 - Proceso ETL
- 05 - Sistema Recomendador
- 06 - Sistema Predictivo
- 07 - Conclusiones
- 08 - Líneas Futuras

Introducción Objeto de estudio

Anime/Manga – Contenido multimedia de origen japonés, en forma de animación o cómic



Merchandising – Venta de productos físicos asociados a una propiedad intelectual

Introducción

Identidad otaku como target comercial

Ensanchamiento demográfico – El público es cada vez más amplio y diverso

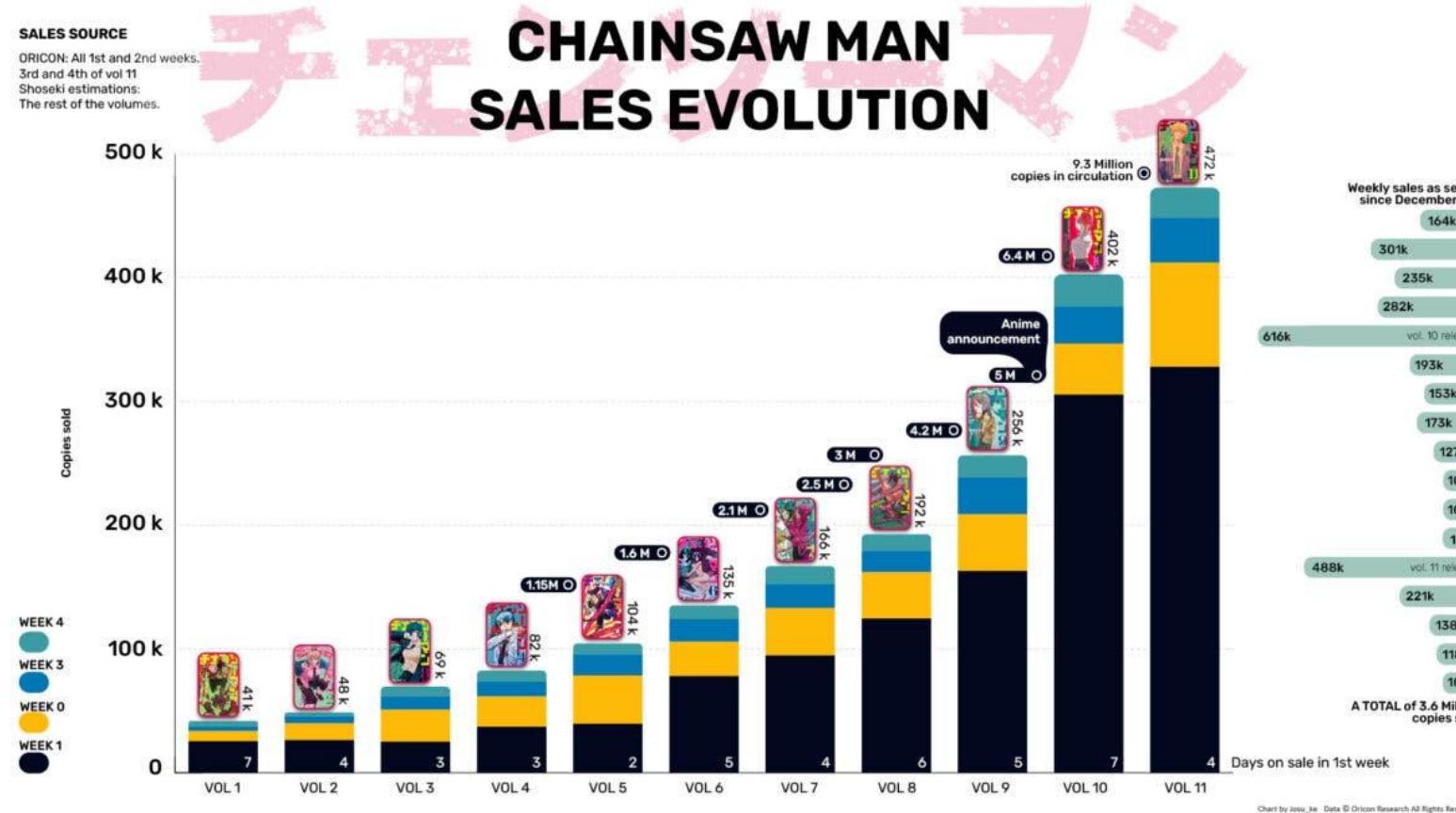


Problema ontológico – Tipología precisa para clasificar recogiendo nuevos matices

Introducción

Cambio en los patrones de consumo

Distribuidoras nacionales – Costes en localización al mercado europeo

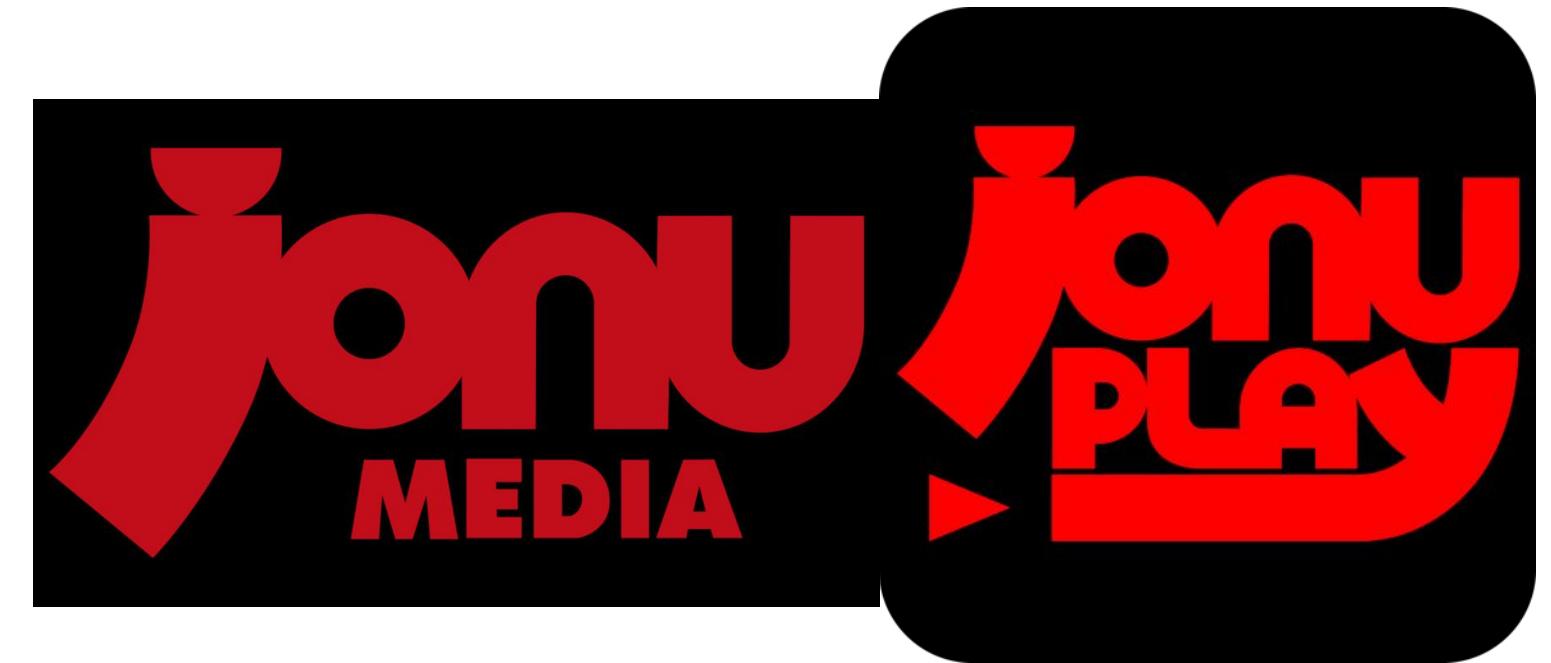


Parametrización del impacto – Aparente éxito de series populares

Introducción

Necesidad de mercado

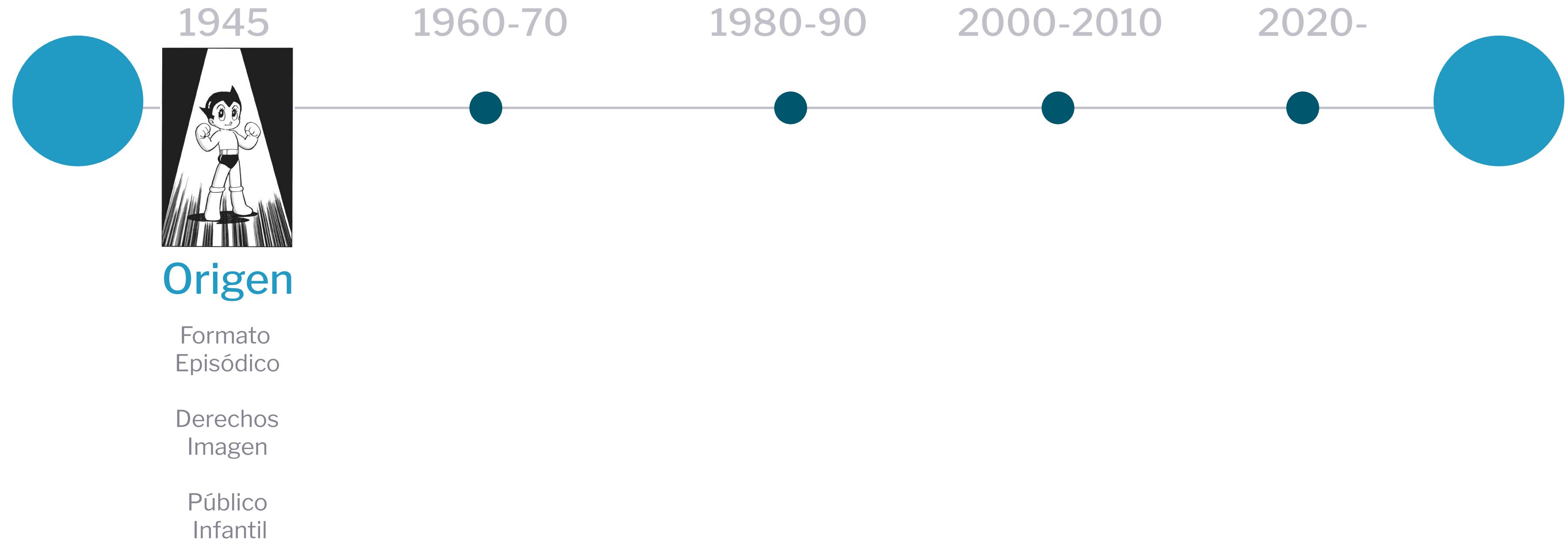
Problema de reach – Llegar a nuevo público objetivo



Problema de engagement – Mantener la atención del público objetivo

Contexto Histórico

Línea Temporal I



550 ¥

Venta Autor Capítulo (1968)
3,38 Euros

2,5M ¥

Producción Real Capítulo (1968)
15,365 Euros

3,13k M ¥

Net Worth Actual (2021)
18,518,965.38 Euros

Contexto Histórico

Línea Temporal II



44k M ¥

Gundam
270 millones de euros

127k M ¥

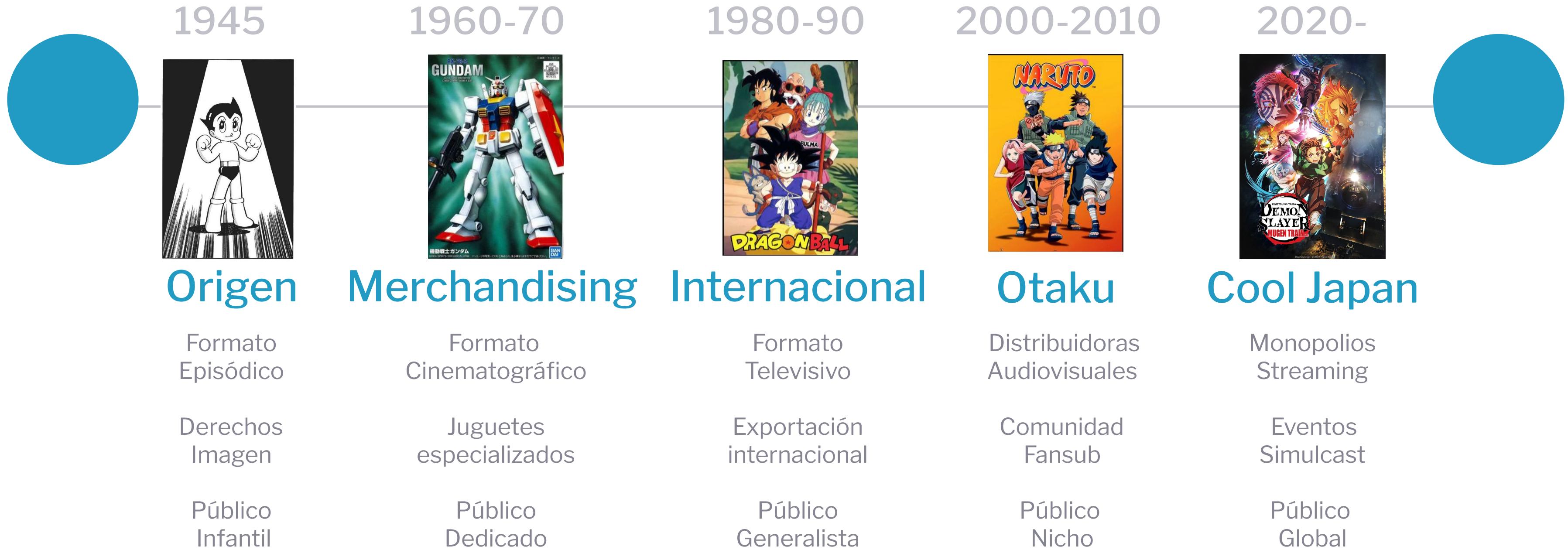
Dragon Ball
780 millones de euros

+250k M ¥

Toei Animation
1536 millones de euros

Contexto Histórico

Línea Temporal



410 M \$

Star Wars (1977)
373 millones de euros

+1k M \$

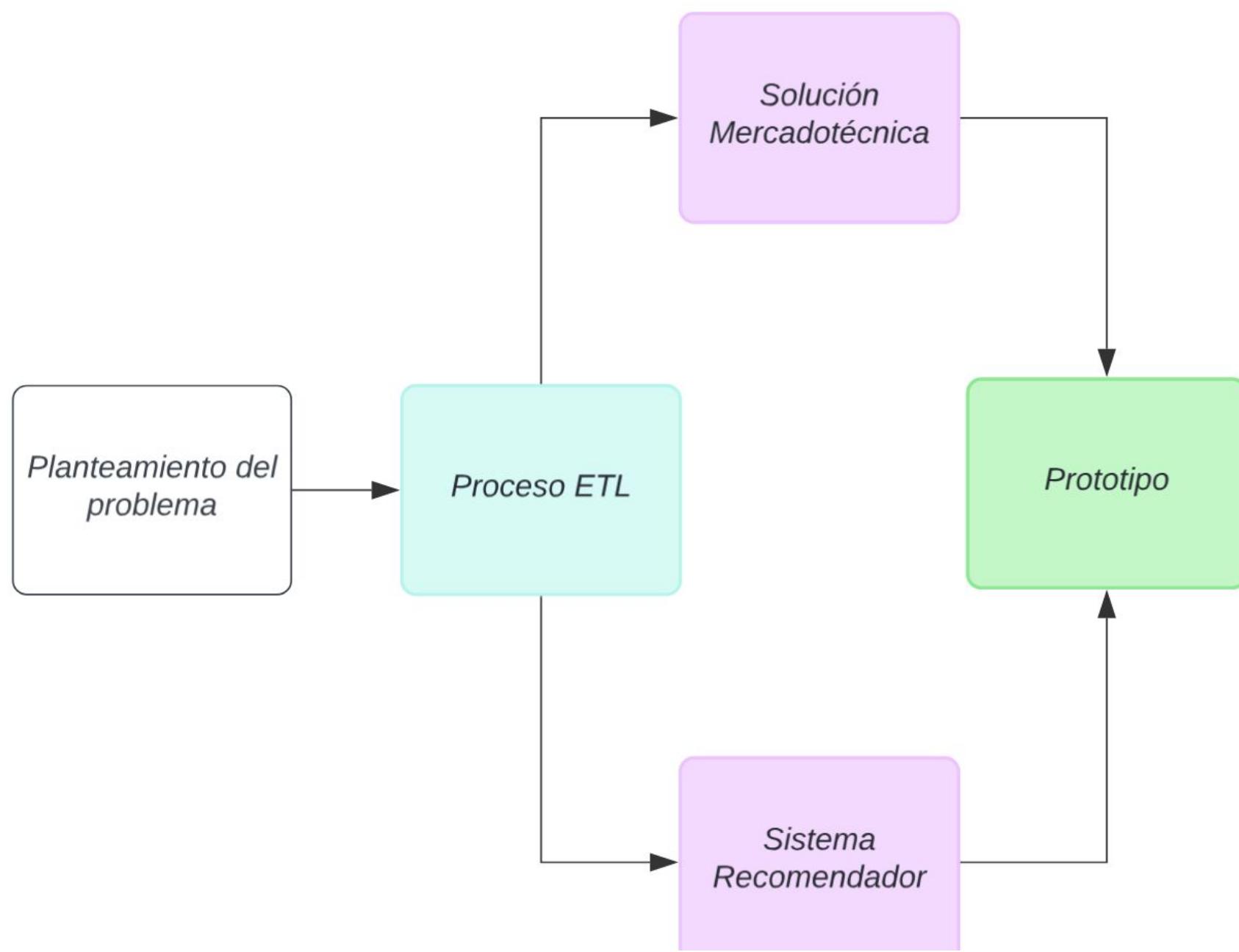
Joker (2019)
911 millones de euros

507 M \$

Demon Slayer (2020)
455 millones de euros

Estado del Arte I

Diagrama



Proceso ETL

Scraping, Manipulación y Transformación de datos
(Extract, Transform & Load)

Sistemas/Algoritmos

Solución Mercadotécnica / Modelo Predictivo
Sistema Recomendador

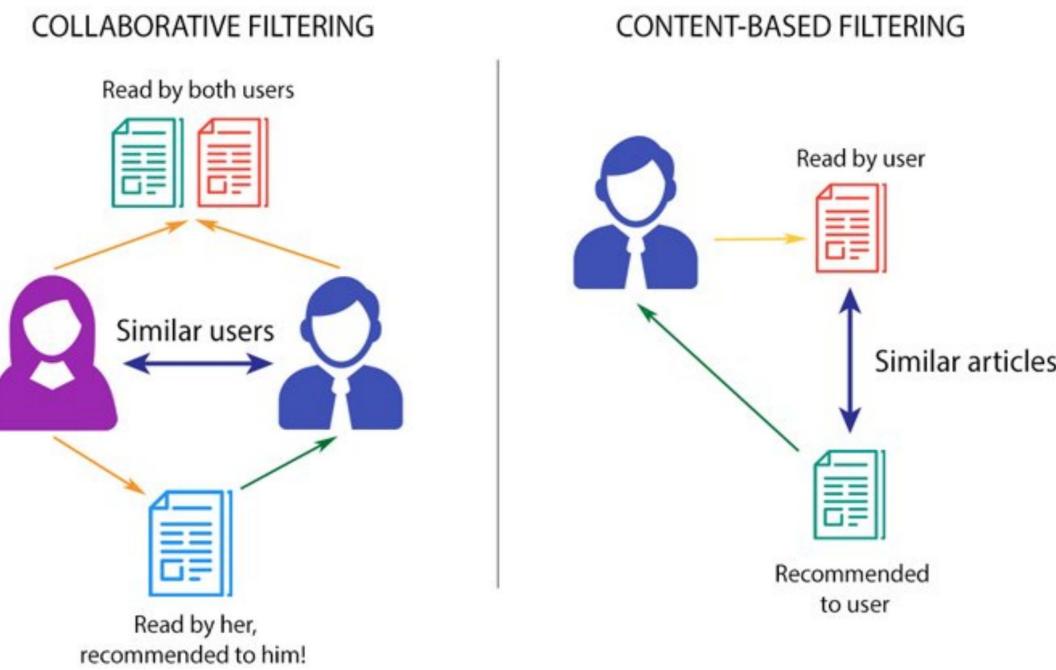
Prototipo

MVP (Minimum Viable Product) para clientes
(No desarrollado)

Estado del Arte II

Sistemas Recomendadores

Preguntas acerca de los
Sistemas Recomendadores



35 % de ventas en Amazon
=
Sistema Recomendador



Estado del Arte II

Sistemas Recomendadores



Las recomendaciones de anime se centran en criterios clásicos basados en demografía, en lugar de las necesidades del usuario

Netflix (Reed Hastings, 2023) confirma que ni la clasificación clásica con demografías, ni las etiquetas arbitrarias son relevantes para recomendar cierto tipo de contenido



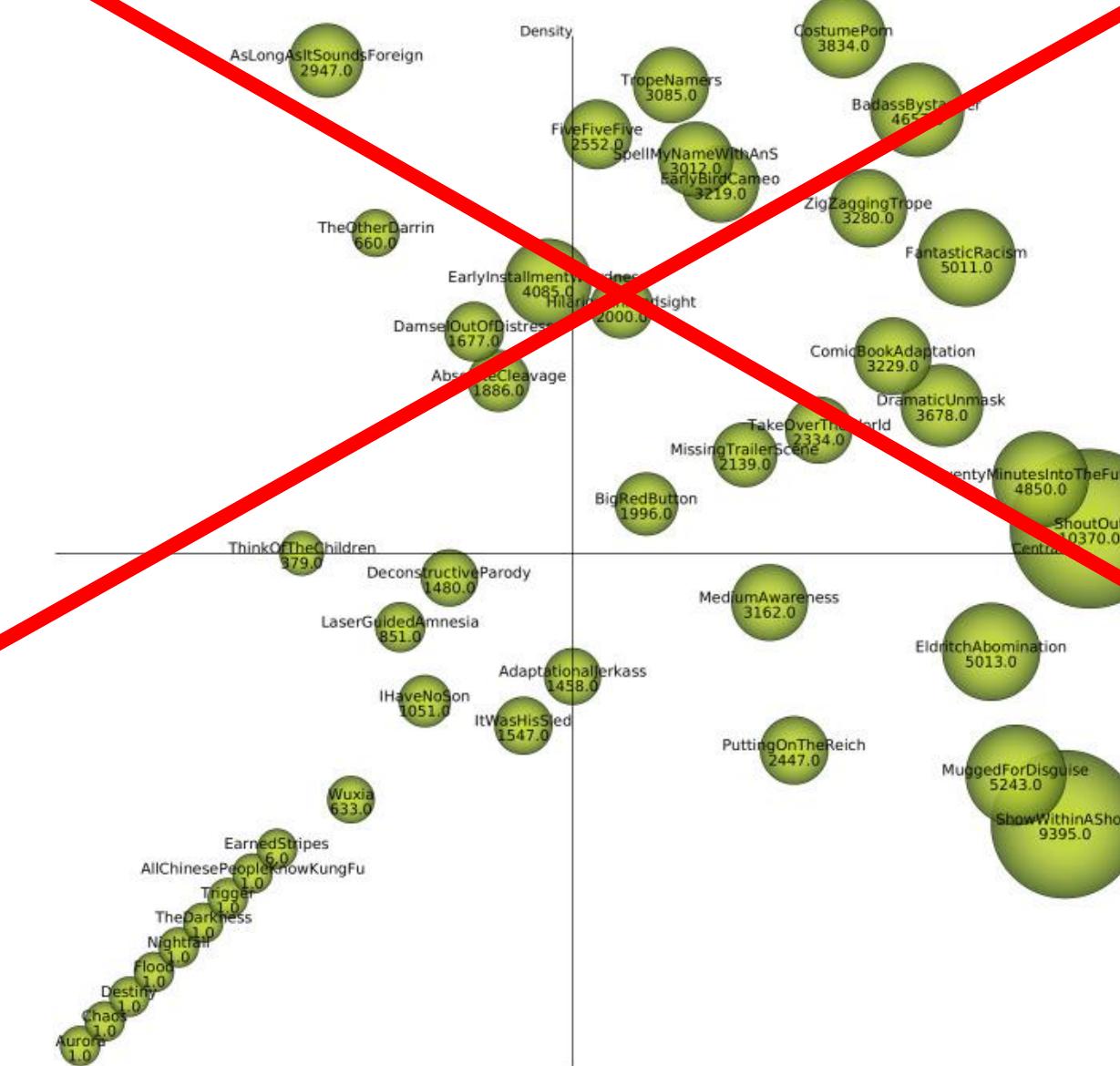
Estado del Arte III

Modelos Predictivos

Las distribuidoras realizan grandes inversiones para localizar productos audiovisuales sin que puedan parametrizar previamente el impacto de los animes que salen cada temporada

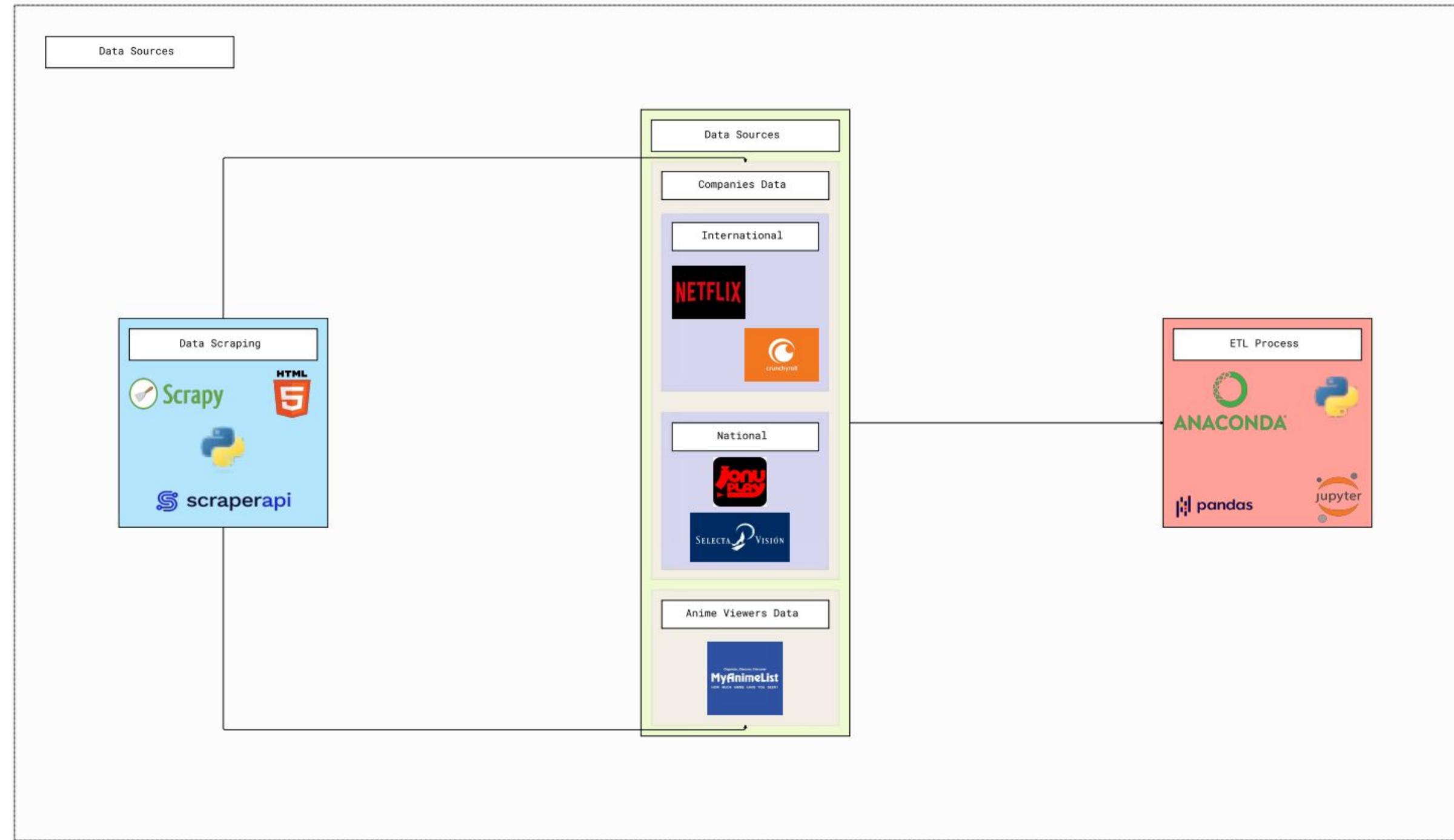
Analysing the conceptual evolution of qualitative marketing research through science mapping analysis (Manuel Cobo)

The Simpsons did it: Exploring the film trope space and its large structure (Manuel Cobo)



Proceso ETL I

Diagrama

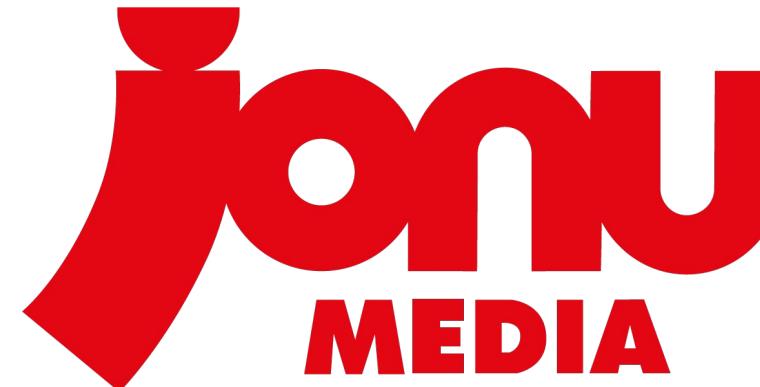
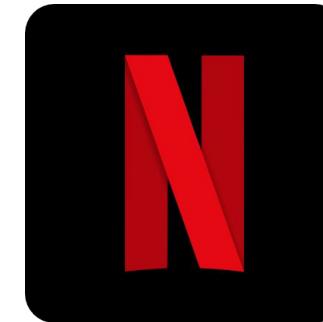


Proceso ETL II

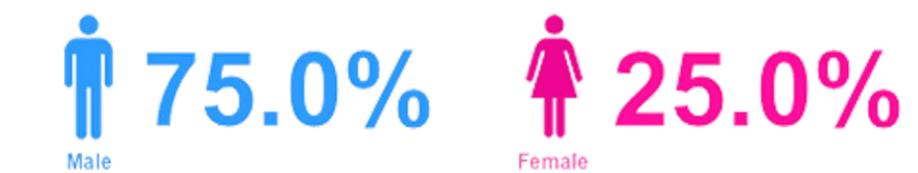
Dataframe Resultante

+40.000 entradas

+16.8M usuarios registrados



Gender



Age

A screenshot of a MyAnimeList page for the anime 'Kami nomi zo Shiru Sekai'.

Kami nomi zo Shiru Sekai
The World God Only Knows

Score: 7.65 (335,607 users)
Ranked #1351 | **Popularity #344** | **Members 630,697**
Fall 2010 | TV | Manglobe

Synopsis:
Keima Katsuragi, known online as the legendary "God of Conquest," can conquer any girl's heart—in dating sim games, at least. In reality, he opts for the two-dimensional world of gaming over real life because he is an unhealthily obsessed otaku of galge games (a type of Japanese video game centered on interactions with attractive girls). When he arrogantly accepts an anonymous offer to prove his supremacy at dating sim games, Keima is misled into aiding a naive and impish demon from hell named Elucia "Elsie" de Lute Ima with her mission: retrieving runaway evil spirits who have escaped from hell and scattered themselves throughout the human world. Keima discovers that the only way to capture these spirits is to conquer what he hates the most: the unpredictable hearts of three-dimensional girls! Shackled to Elsie via a deadly collar, Keima now has his title of "God of Conquest" put to the ultimate test as he is forced to navigate through the hearts of a multitude of real-life girls.

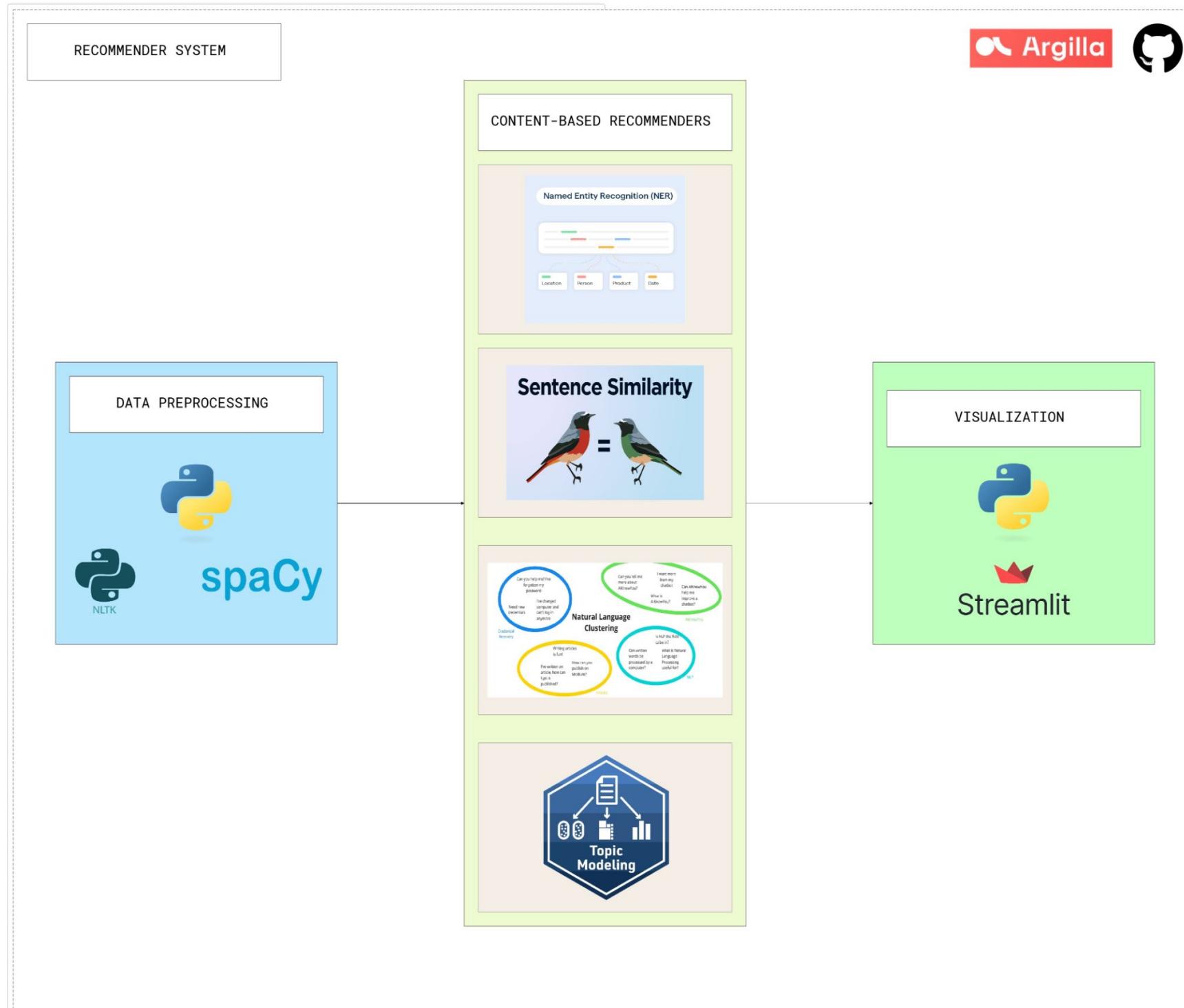
Background:
There was a limited viewing on September 25, 2010. The regular TV broadcasts started on October 7, 2010.

Related Entries:
Adaptation (Manga)
Kami nomi zo Shiru Sekai
Prequel (OVA)
Kami nomi zo Shiru Sekai: Flag 0

Alternative Titles:
Synonyms: Kamionomi
Japanese: 神のみぞ知るセカイ
Sequel (TV)
Kami nomi zo Shiru Sekai II

Sistema Recomendador I

Diagrama



Named-Entity Recognition

Clasificar elementos temáticos dentro de un texto

Natural Language Processing (NLP)

Sentence Similarity

Semejanza por significado y/o estructura

Clusterización

Agrupar datos similares en conjuntos llamados clusters

Topic Modeling

Temas abstractos en una colección de documentos

Sistema Recomendador II

Snippet de código

Use Case: Inazuma Eleven

```
main_anime_title = "Inazuma Eleven"
main_anime_synopsis = base_df[base_df["JP_Title"] == main_anime_title]["Synopsis"].item()

model_type = "Multi Task Large Model"
model = model_multitask_large

labels = ["Soccer", "Sport", "Friendship"]

animes_wt_synopsis = {}

for row in range(0, base_df.shape[0]):

    title = base_df.iloc[row]["JP_Title"]
    synopsis = base_df.iloc[row]["Synopsis"]

    animes_wt_synopsis[title] = synopsis

animes_more_related = recommender_system_NER(main_anime_title, main_anime_synopsis, animes_wt_synopsis, labels, model_type, model)

scores_sort = list(animes_more_related.keys())
scores_sort.sort(reverse=True)

for score in scores_sort:
    print(score, animes_more_related[score])

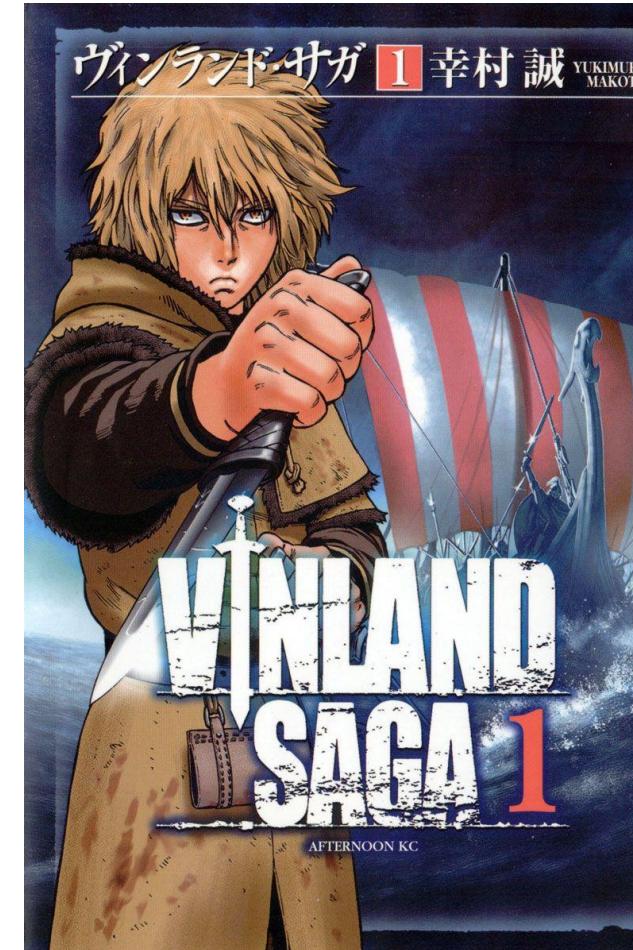
100.0 ['Captain Tsubasa']
66.66666666666666 ['Haikyuu!!', 'Blue Lock']
33.33333333333333 ['Tennis no Ouji-sama', 'Clannad', 'Año Hi Mita Hana no Namae wo Bokutachi wa Mada Shiranai.', 'Black Clover']
0.0 ['Fullmetal Alchemist', 'Gintama', 'K-On!', 'Violet Evergarden']
```

Demo usando la primera técnica

Anotaciones en Gliner sobre series de fútbol

Sistema Recomendador III

Caso de Uso - NER



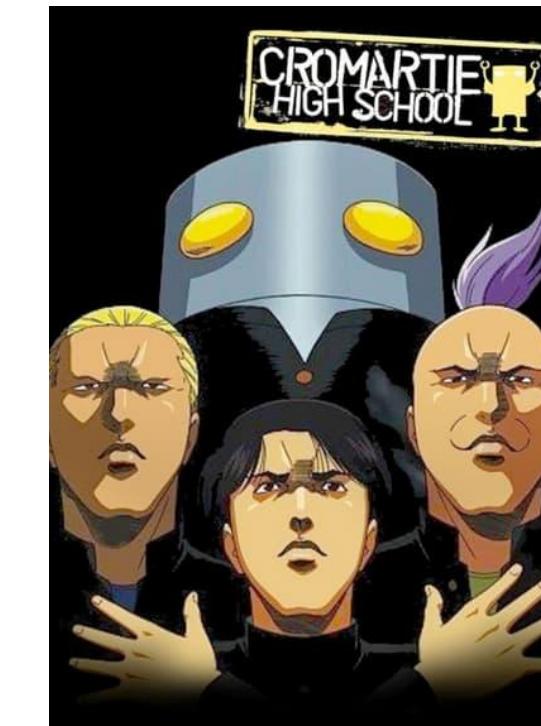
Sistema Recomendador III

Caso de Uso - Clusterizacion



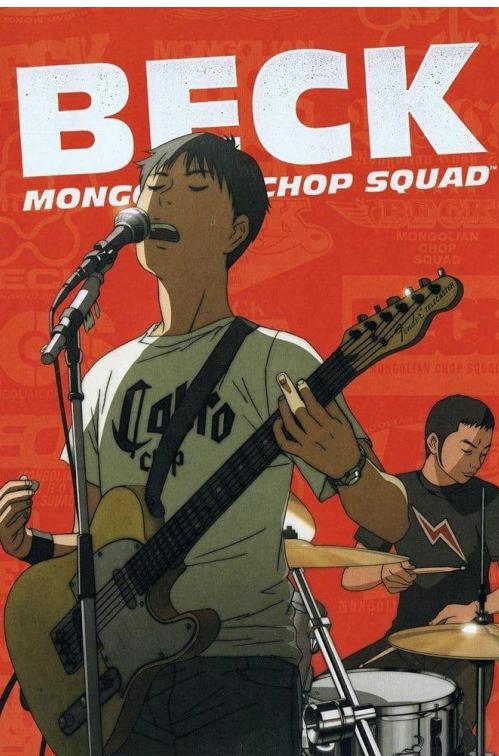
Sistema Recomendador III

Casos de Uso - Text Similarity



Sistema Recomendador III

Casos de Uso - Topic Modeling



Sistema Recomendador IV

Evaluación

Primer Caso de Uso - Named Entity Recognition (NER)

57 respuestas



Figura 5.7: (a) Named Entity Recognition

Tercer caso de uso - Text Similarity

57 respuestas



Figura 5.9: (c) Text Similarity

Segundo Caso de Uso - Clusterización

57 respuestas



Figura 5.8: (b) Clusterización

Cuarto caso de uso - Topic Modeling

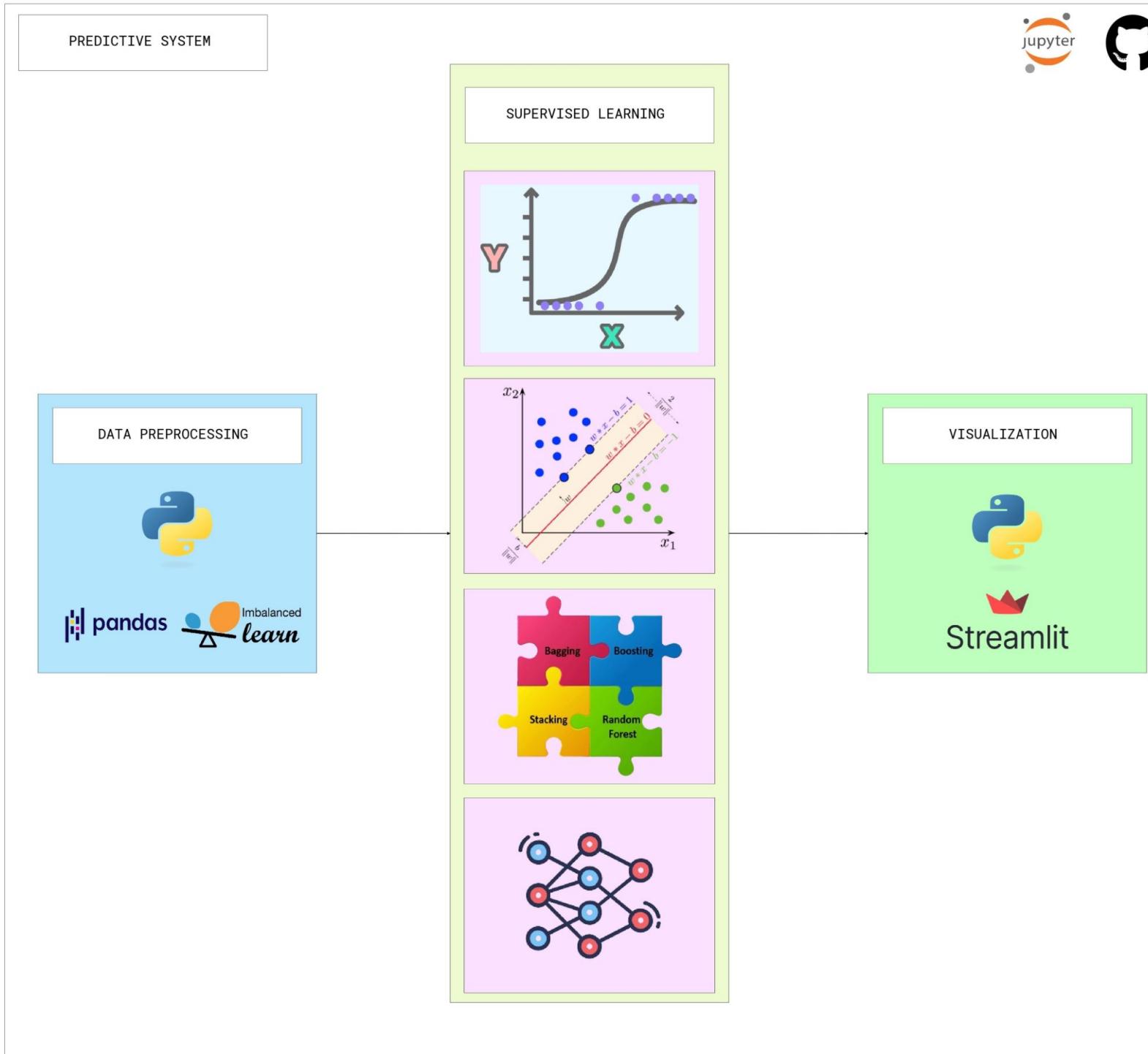
57 respuestas



Figura 5.10: (d) Topic Modeling

Modelo Predictivo I

Diagrama



Regresión Logística

Utilizar la particularidad de que la imagen de la función es entre 0 y 1 para realizar la clasificación

Máquina de Soporte Vectorial

Usar la SVM para poder dividir el espacio vectorial en dos dimensiones de manera (no) lineal

Machine Learning (Problema de Clasificación Binaria)

Ensemble

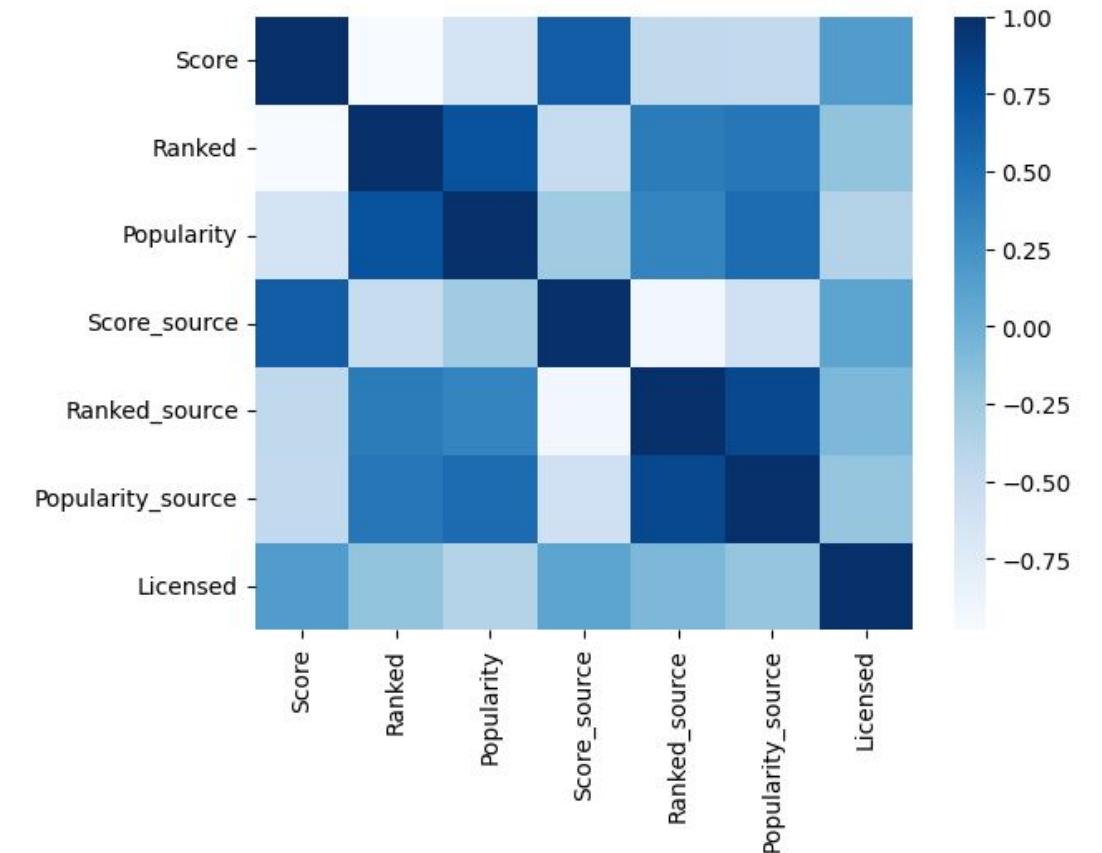
Aprovechar el potencial de utilizar varios modelos para conseguir mejores resultados

Redes Neuronales

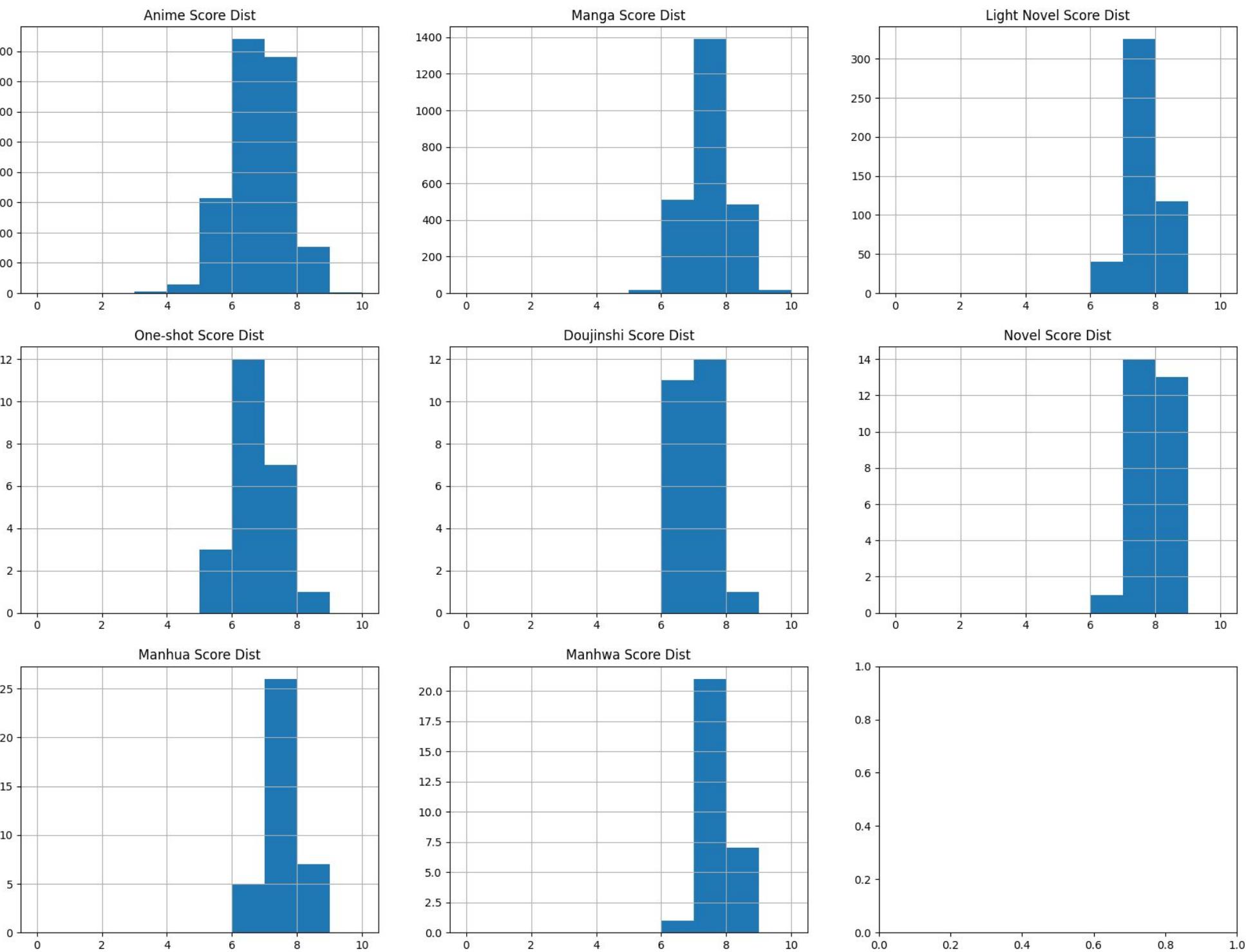
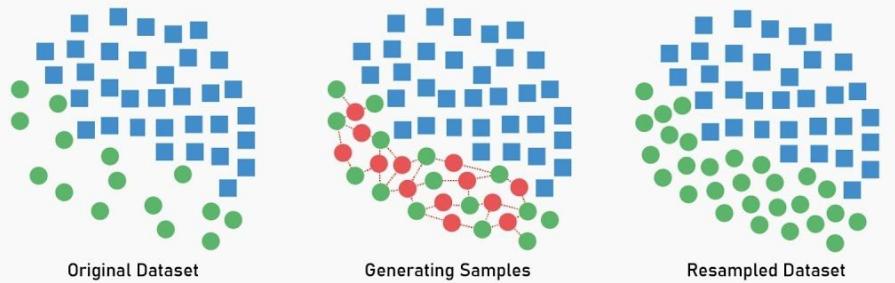
Facilitar el reconocimiento de patrones gracias a la alta capacidad de detección que tienen las Redes Neuronales (No desarrollado)

Modelo Predictivo II

EDA + Feature Engineering



SMOTE
HANDLE IMBALANCED DATASET
Synthetic Minority Oversampling Technique



Clasificar en función de las entregas anteriores del anime

| Modelo Predictivo II

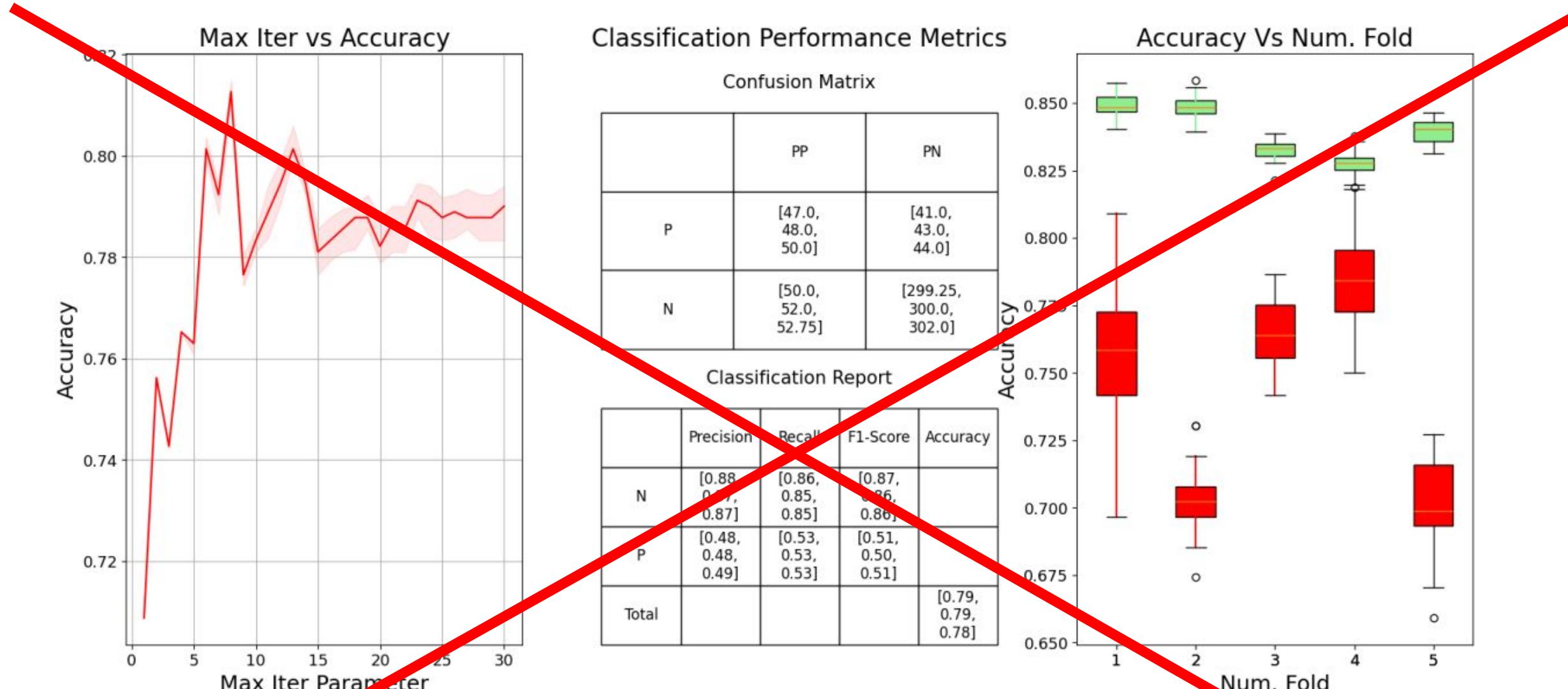
¿Objetivo/Prioridad en la Evaluación?

Maximizar las métricas de evaluación sin llegar a un sobreajuste, pero priorizando siempre la clasificación que se adapte a la etiqueta positiva (1)

Modelo Predictivo IV

Evaluación I

Modelo Lógico



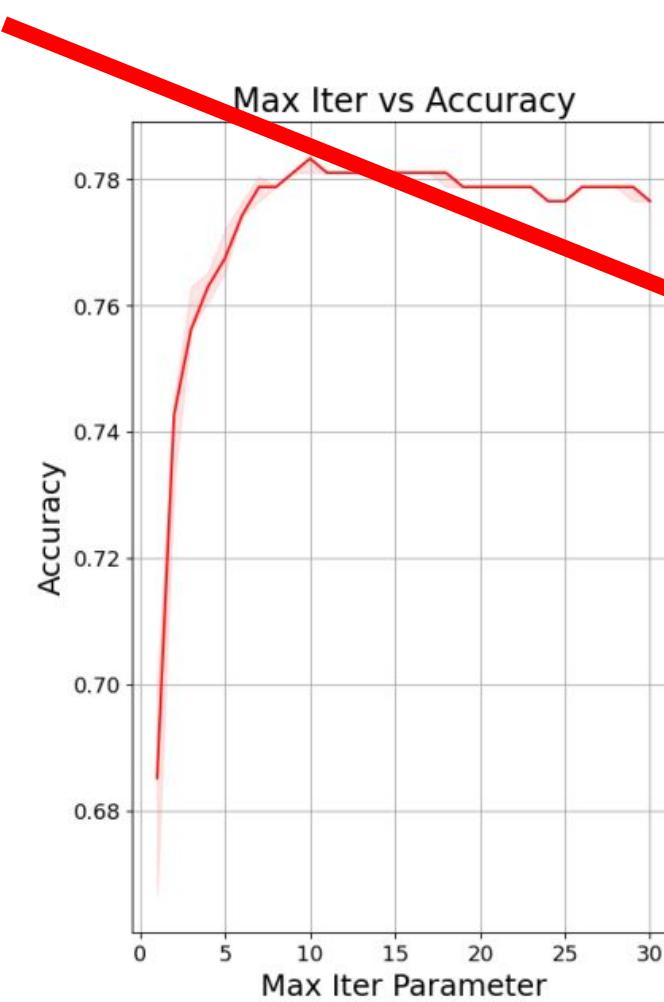
Decision Tree (Vanilla) with 30 different seeds

Figura 6.12: Ejecución del clasificador (*DecisionTreeClassifier*) con 30 ejecuciones con distintas *seeds* y con 30 iteraciones máximas

Modelo Predictivo IV

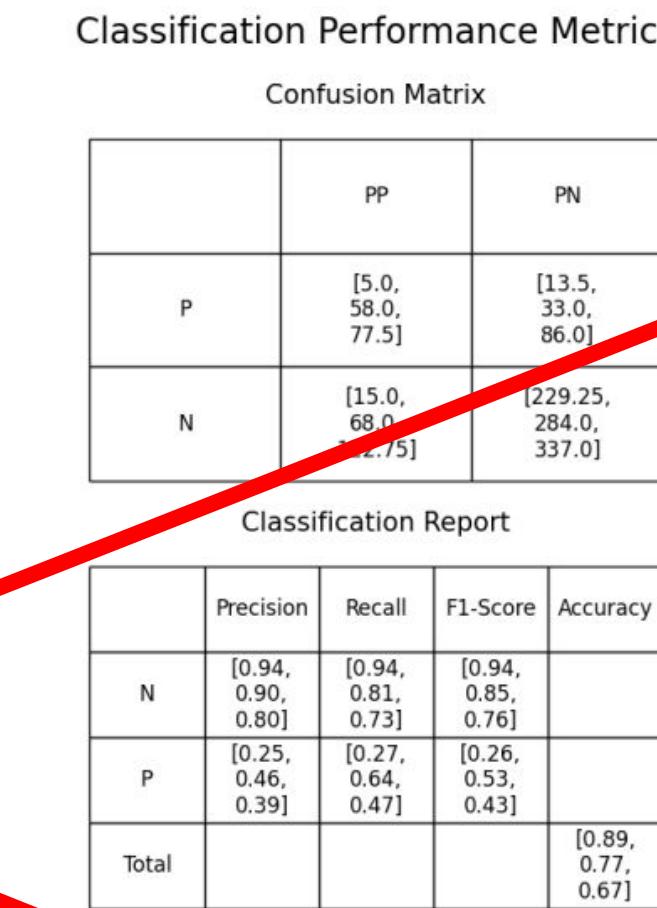
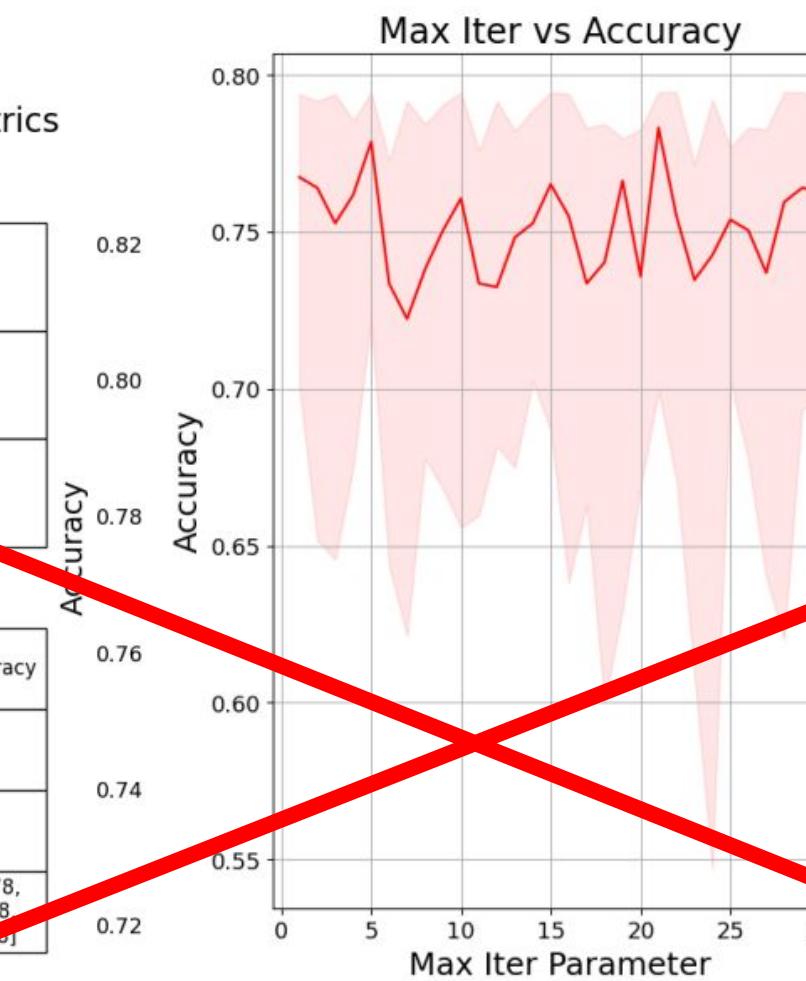
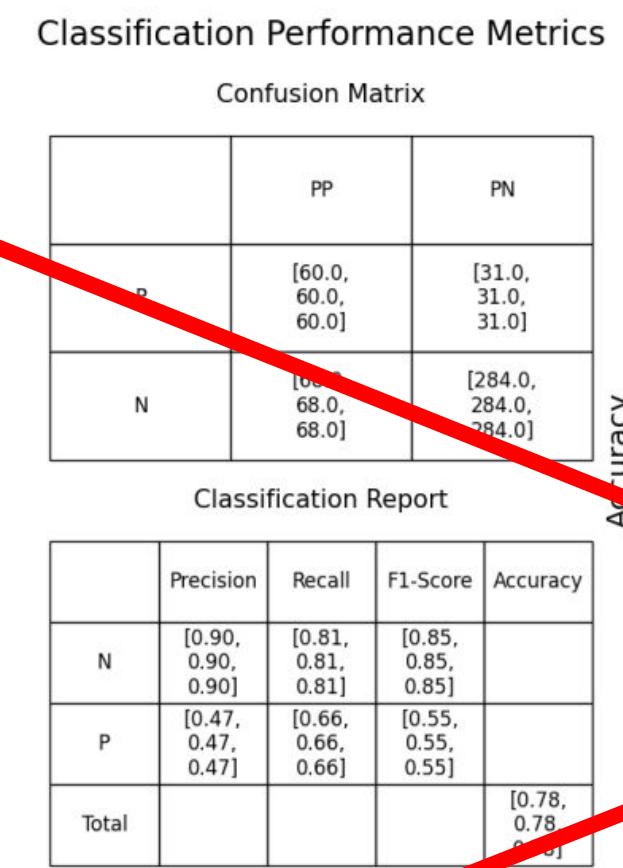
Evaluación II

Modelo Geométrico



Logistic Regression (Vanilla) with 30 different

Figura 6.15: Ejecución del clasificador (*Logistic Regression*) con 30 ejecuciones con distintas seeds y con 30 iteraciones máximas



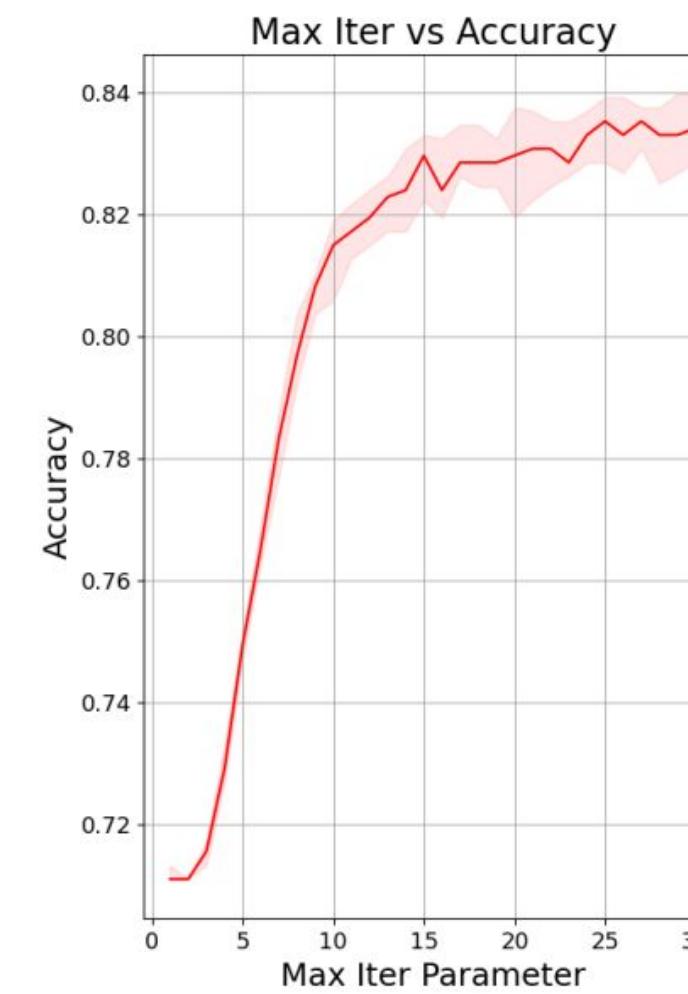
Linear Support Vector Machine (Vanilla) with 30 different seeds

Figura 6.17: Ejecución del clasificador (*LinearSVC*) con 30 ejecuciones con distintas seeds y con 30 iteraciones máximas

Modelo Predictivo IV

Evaluación III

Modelo Mixto



Classification Performance Metrics

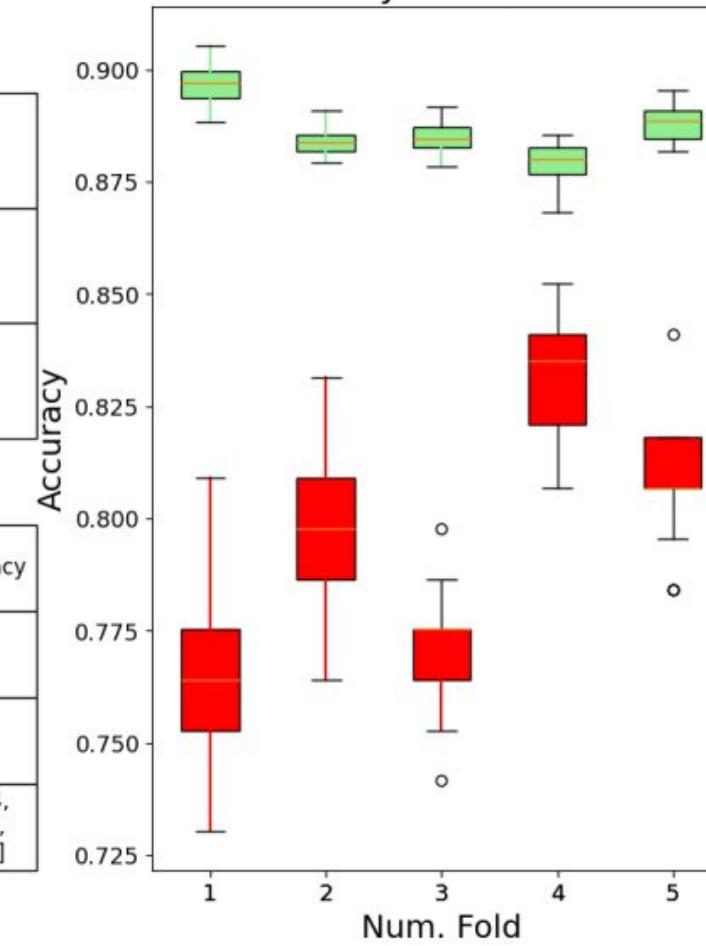
Confusion Matrix

	PP	PN
P	[51.0, 52.0, 53.75]	[37.25, 39.0, 40.0]
N	[33.0, 35.0, 37.0]	[315.0, 317.0, 319.0]

Classification Report

	Precision	Recall	F1-Score	Accuracy
N	[0.89, 0.89, 0.89]	[0.91, 0.90, 0.90]	[0.90, 0.90, 0.89]	
P	[0.61, 0.60, 0.59]	[0.58, 0.57, 0.57]	[0.59, 0.58, 0.58]	
Total				[0.84, 0.83, 0.83]

Accuracy Vs Num. Fold



Random Forest (Vanilla) with 30 different seeds

Figura 6.18: Ejecución del clasificador (*RandomForestClassifier*) con 30 ejecuciones con distintas *seeds* y con 30 iteraciones máximas

CONCLUSIÓN DAFO



DEBILIDAD

- Recomendación existente
- Investigadores noveles
- Escucha activa de los clientes



AMENAZA

- Branding negativo hacia la IA
- Face-value de los clientes
- Coste reputacional



FORTALEZA

- ✓ Competencia sobre NLP/ML
 - ✓ Conocimiento sobre el sector
 - ✓ Nicho de mercado
- ✓ Creación de nuevas plataformas
 - ✓ Competencia entre clientes
 - ✓ Crecimiento de distribuidoras



OPORTUNIDAD

LÍNEA 1

Investigación y Ampliación del desarrollo
Text Similarity, Topic Modeling, NNN, ...

LÍNEA 2

Creación del Prototipo
Possible demo para MVP

LÍNEA 3

Comunicación con Potenciales Clientes
Buscar inversores para el producto

Reflexión

Cita



“El anime no es solo una forma de entretenimiento, es una ventana a universos llenos de posibilidades que inspiran a generaciones”

(Masashi Kishimoto, creador de Naruto)

Muchas gracias por su
atención

¿Alguna pregunta?