

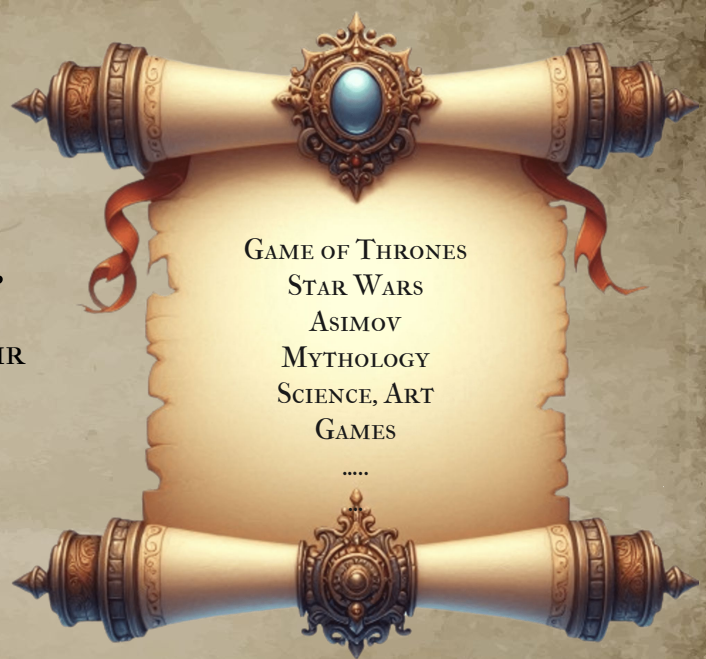


NOMBRES EN VIDEOJUEGOS

EN VIDEOJUEGOS ONLINE COMO **WORLD OF WARCRAFT** , **FINAL FANTASY XIV**,
ELDER SCROLLS, EN JUEGOS NO MMORPGs, COMO **BALDUR'S GATE 3**,
O INCLUSO NOMBRES DE SUMMONER EN **LoL** LOS JUGADORES TIENDEN A ELEGIR
NOMBRES DE **PERSONAJES** MODIFICADOS, BASADOS EN:

- UNIVERSOS FANTASIA / SCI-FI
- MITOLOGIA
- OTROS JUEGOS
- PERSONAJES HISTÓRICOS
- INSULTOS/OFENSIVOS

ES INTERESANTE EL PODER DETECTAR EL ORIGEN O INFLUENCIAS



LORENEXUS

EXTRAE LA ESENCIA DEL ORIGEN DE UN NOMBRE
DETECTA PATRONES FONÉTICOS Y MORFOLÓGICOS

OBJETIVO: CLASIFICAR NOMBRES POR UNIVERSO (ETIQUETA)

- VHAELMYRA (INVENTADO) → **GAME OF THRONES**
- ARAGHURNN → **TOLKIEN**
- ARAHORN → **TOLKIEN** (Y UN POCO PILLO)
- RIHATHRIL (INVENTADO) → **TOLKIEN**
- BRO-TUNJEIN → **MYTHOLOGY**
- FUTHERMUCKER → **OFENSIVO**





NI DATOS, NI ETIQUETAS



FUENTES DE DATOS

WIKIDATA

UNIVERSOS FICTICIOS
PERSONAJES HISTÓRICOS
SPARQL

GODCHECKER

CULTURAS
MITOLÓGICAS

SURGE AI

SLURS/INSULTOS



DATASET
HOMOGENEIZADO

PERSONAJES

NER

LOCALIZACIONES
Y GPE DE LIBROS
EN PDF

PIPELINE (SIMPLIFICADO) DE PROCESADO



NORMALIZAR

LIMPIEZA GENERAL



PROCESADO

CADA FUENTE UN PROCESADO
DIFERENTE.



ETIQUETADO

DINÁMICO
FORMATO FASTTEXT:

__LABEL__ UNIVERSO NOMBRE



??

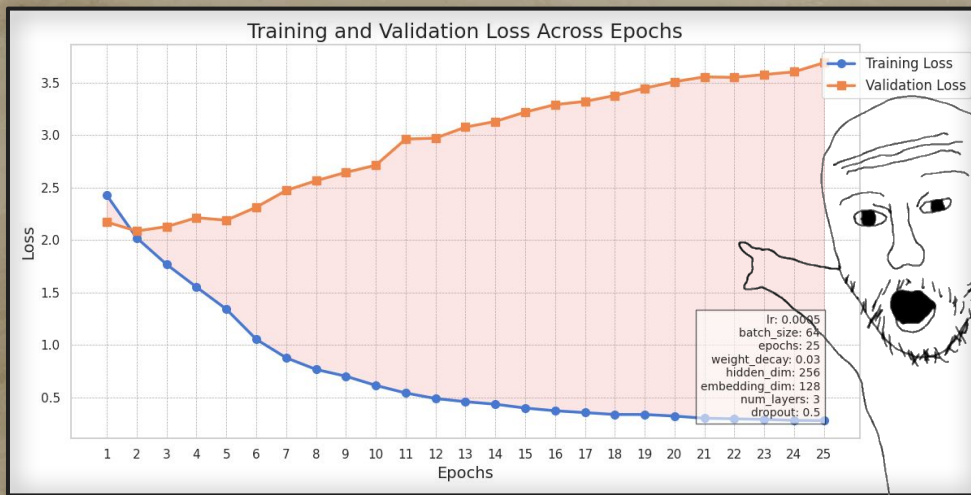




OK.

PERO... ¿SON SUFICIENTES
ESTOS DATOS?

SPOILER: NO



TÉCNICAS DE DATA AUGMENTATION



BÁSICAS

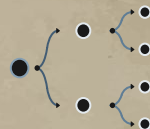
JON SNOW → JONSNOW
→ JON
→ SNOW



FONÉTICAS

OO → U (TATOOINE)
EE → I (BEEKLORD)

MOON VEIL → MUN VEIL





TEXTATTACK

PERMUTACIONES INTERNAS
SWAPS ALEATORIOS,
DUPLICACIONES, INSERCIONES,
BORRADOS DE CARACTERES...

JINX → JNIX
→ JINXX
→ JYNX
→ JNX





~23.0000 NOMBRES

SON TRANSFORMADOS EN

~1.0000.0000 INSTANCIAS

ETIQUETAS Y ESTRATIFICACIÓN



ETIQUETAS

SON DINÁMICAS, PERO MODELO
FINAL TIENE **15 ETIQUETAS**



ESTRATIFICACIÓN

PROPORCIÓN EQUITATIVA

80% TRAIN, 10 % DEV 10% TEST



ESTRATEGIAS

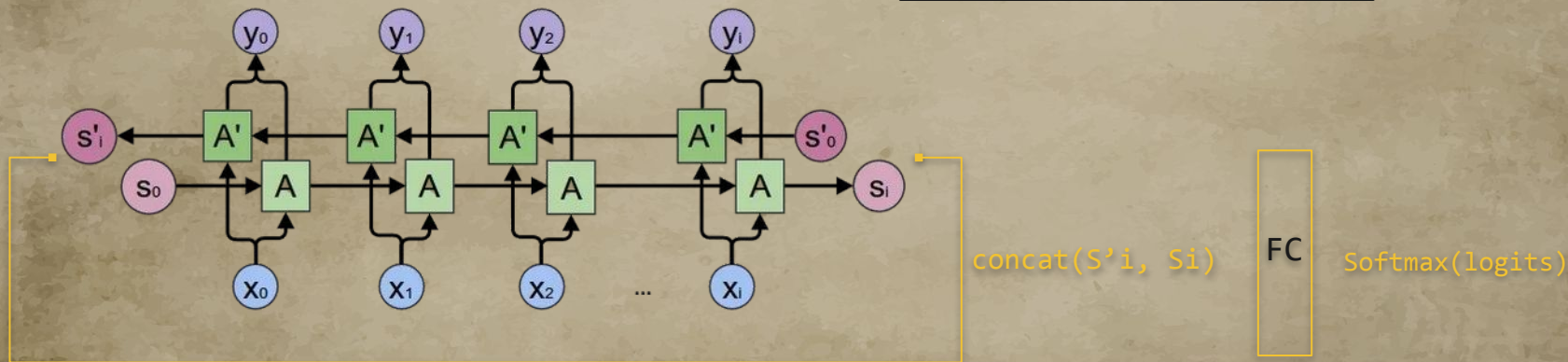
- **TOKENIZACIÓN Y EMBEDDINGS**
A NIVEL DE CARACTER
- **LSTM BIDIRECCIONALES (BiLSTM)**

FLAIR

LIBRERÍA DE USO SENCILLO

PYTORCH, FROM SCRATCH

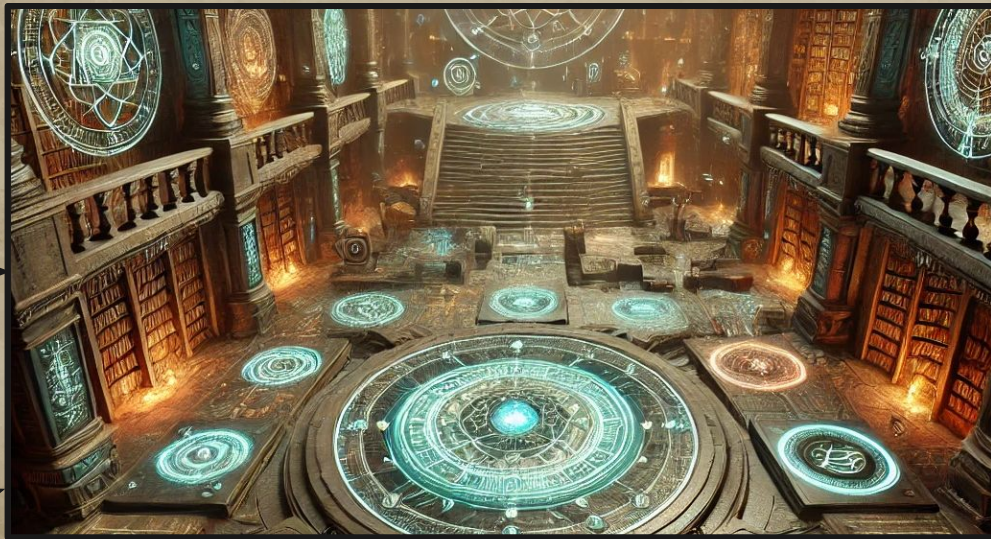
IMPLEMENTACIÓN GRANULAR
PADDINGS, VOCAB DINÁMICO..ETC
CROSS ENTROPY PONDERADA



TRAINING GROUNDS

GRID
HYPERPARAMS

MODELS



WINNER

LOGS EXHAUSTIVOS

MÉTRICAS, GRÁFICAS

MATRICES CONFUSIÓN

SISTEMA AUTOMATIZADO. GANADOR AL FINALIZAR.

ARENA: FLAIR VS PYTORCH

5.2 Combatiente A: Flair

B	Ep	WD	H	E	L	D	F1
32	5	0.03	64	32	1	0.5	0.3227
32	7	0.03	128	64	2	0.4	0.4905
64	7	0.03	128	64	2	0.3	0.6013
32	10	0.02	192	96	2	0.3	0.6880
32	12	0.02	256	128	3	0.3	0.7214

Table 1: Ver log completo: [FlairModel Log](#)

5.3 Combatiente B: Pytorch

B	Ep	WD	H	E	L	D	F1
32	5	0.03	64	32	1	0.5	0.62
32	7	0.03	128	64	2	0.4	0.89
64	7	0.03	128	64	2	0.3	0.90
32	10	0.02	192	96	2	0.3	0.92
32	12	0.02	256	128	3	0.3	0.94

Table 2: Ver log completo: [PytorchModel Log](#)

Final Summary of Best Results:

Model: PytorchModel

Best Accuracy: 0.9585

Best Hyperparameters: {'lr': 0.0005, 'batch_size': 32, 'epochs': 12, 'weight_decay': 0.02, 'hidden_dim': 256, 'embeddings_dim': 128, 'num_layers': 3, 'dropout': 0.3}

Best Classification Report:

	precision	recall	f1-score	support
Historical	0.99	0.98	0.98	34828
Tolkien	0.98	0.93	0.91	4457
Warcraft	0.96	0.94	0.95	413
Offensive	0.95	0.96	0.96	2819
StarWars	0.96	0.92	0.94	3376
Naruto	0.92	0.98	0.95	567
HarryPotter	0.91	0.91	0.91	2740
Mythology	0.98	0.93	0.92	6983
GameofThrones	0.93	0.96	0.94	3574
FinalFantasy	0.95	0.96	0.96	285
DoctorWho	0.96	0.91	0.94	623
TheWitcher	0.94	0.96	0.95	387
ForgottenRealms	0.93	0.86	0.89	502
DragonBall	0.89	0.95	0.92	423
Discworld	0.99	0.91	0.95	174
accuracy			0.96	62151
macro avg	0.94	0.94	0.94	62151
weighted avg	0.96	0.96	0.96	62151

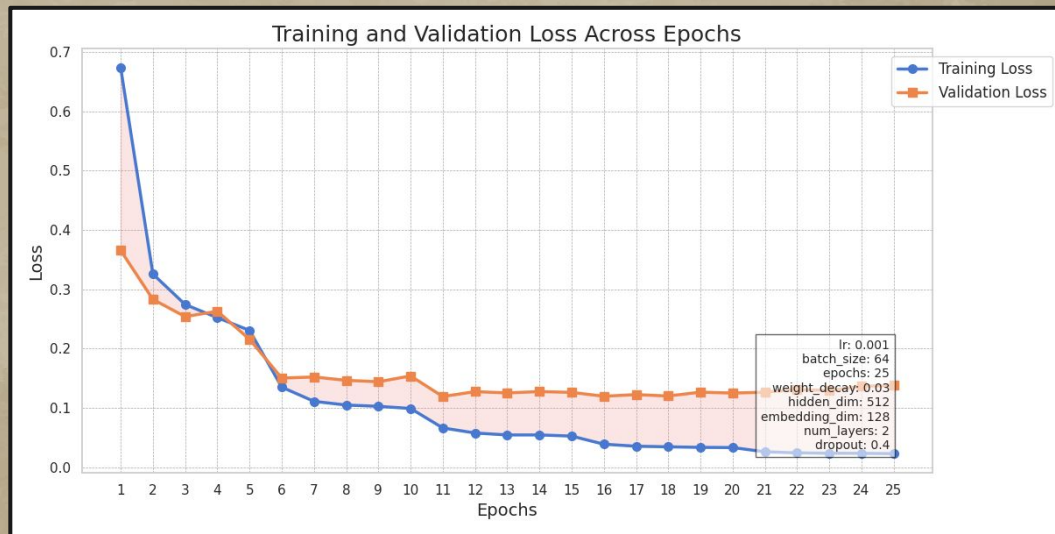
MEJOR MODELO:

LORENEXUSPYTORCH_V1.0

ALTA REGULARIZACIÓN + AUGMENTATION

HIDDEN Y EMBEDDING SIZE: MEDIO
ADAM WEIGHT DECAY + STEP LEARN. RATE
TRAIN ~900K INSTANCIAS
DEV/TEST ~110K

F1-SCORE (MACRO): **0.96**



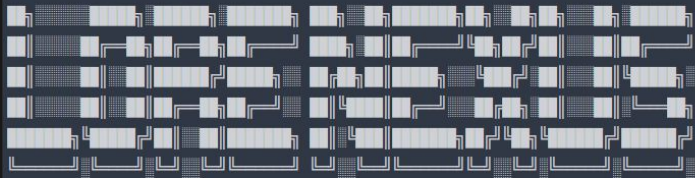
CLI APP

```
Loading model from /home/basajaun/workspace/university/cuarto/pln/lore-nexus/models/pytorch/checkpoints/LoreNexusPytorch_v1.0.pth...
```

```
Loaded Hyperparameters: embedding_dim=128, hidden_dim=512, num_layers=2, dropout=0.4
```

```
LoreNexus (BiLSTM-Pytorch) loaded for CLI predictions:
```

```
Model: /home/basajaun/workspace/university/cuarto/pln/lore-nexus/models/pytorch/checkpoints/LoreNexusPytorch_v1.0.pth
```



```
■ The Lore Nexus is ready to unveil the mysteries of any name you provide ■
```

Choose an option:

1 - Unveil hidden lore from a name 🐼 (Reveal the story and secrets behind)

2 - Exit the Lore Nexus

Your choice: 1

Enter a name to uncover its Lore: Rihathril

```
◄ Ancient Lore Archive: Secrets of: Rihathril ►
```

```
◇ Prediction 1: Tolkien with confidence 0.9993 ■
```

```
Prediction 2: GameofThrones with confidence 0.0002
```

```
Prediction 3: StarWars with confidence 0.0002
```

```
Prediction 4: Historical with confidence 0.0002
```

¿QUIZÁ POR **BELLATRIX** ?

```
◄ Ancient Lore Archive: Secrets of: Napolionix ►
```

```
◇ Prediction 1: Historical with confidence 0.9817 ■
```

```
Prediction 2: HarryPotter with confidence 0.0124
```

```
Prediction 3: StarWars with confidence 0.0055
```

```
Prediction 4: Tolkien with confidence 0.0003
```

```
◄ Ancient Lore Archive: Secrets of: Valyseris ►
```

```
◇ Prediction 1: GameofThrones with confidence 0.8071 ■
```

```
Prediction 2: Historical with confidence 0.1756
```

```
Prediction 3: Discworld with confidence 0.0054
```

```
Prediction 4: Mythology with confidence 0.0054
```


MUESTRAS INFERIDAS EN CLI

Nombre derivado	Ground Truth	Etiqueta	Predicción
Lorecraft	Lovecraft	Historical	Historical (0.87), Game of Thrones (0.11)
Harry Potta	Harry Potter	Harry Potter	Harry Potter (1.00)
Tattuinne	Tatooine	Star Wars	Star Wars (0.99)
Bhoromirian	Boromir	Tolkien	Tolkien (0.94), Historical (0.03)
Dragomir	Inventado	Tolkien	Tolkien (1.00)
Gandalfuck	Gandalf + Fuck	Tolkien, Offensive	Tolkien (0.74), Offensive (0.25)
Hitler	Hitler	Historical, Offensive	Historical (0.89), Offensive (0.09)
Hittler	Hitler	Historical, Offensive	Offensive (0.89), Historical (0.10)
Riverssong	River Song	Doctor Who	Doctor Who (1.00)
Ghaladirel	Galadriel	Tolkien	Tolkien (1.00)
Orhootxiimarru	Orochimaru	Naruto	Naruto (1.00)
Futhermucker	-	Offensive	Offensive (1.00)
Utwatxd	-	Offensive	Offensive (1.00)





MAY THE FORCE BE WITH YOU,
YOUNG PADAWANS