Notas para la documentación del proceso de creación del módulo de clasificación de violencia

Este documento fue utilizado a modo de bitácora, contiene notas realizadas durante el desarrollo de los modelos. Las denominaciones o numeración de los mismos puede ser diferente a la final encontrada en el repositorio y en el trabajo final. Su propósito principal fue la organización y transmisión de información entre integrantes del proyecto para encontrar puntos de mejora.

## Recopilación de datasets

Noticias, tweets (de la zona de argentina; violentos; que contenían la palabra mujer), frases comunes de google. Frases reales dichas por agresores extraídas del dataset del juzgado número 10 del juzgado de buenos aires.

Los tweets los extrajimos utilizando la API de twitter, y posteriormente hicimos un trabajo de limpieza eliminando: menciones, links, hashtags y tweets realizados por bots o en inglés.

## Adaptación de los exports de Label Studio

A diferencia del entrenamiento del modelo de detección de entidades, no había un formato que fuera directamente exportable y usable por el comando spacy train.

En cambio, al clasificar las frases por categorias, solo se podía exportar el formato json.

Por lo tanto, se tuvo que investigar y armar un script para poder tomar ese json y exportarlo con .spacy, para luego poder usarlo en el entrenamiento.

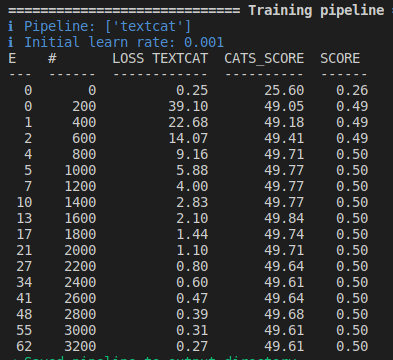
\*nota: Para el entrenamiento y uso de los modelos entrenados con transformers los Transformers tienen que estar en la versión: 4.30

# Modelo Binario:

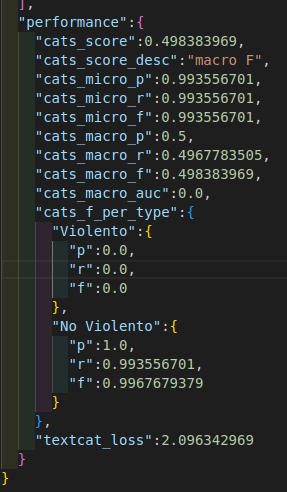
### Entrenamiento 1

El dataset contenía 885 frases violentas y 2586 frases consideradas como no violentas.

Cuando hay un desbalance entre las frases, con mucha mayor proporción de frases no violentas, el entrenamiento da un score de 50



Además, se observa el meta.json, que el entrenamiento solo tiene un score para la categorías “No Violento”. Para la categoría “Violento”, está todo en 0.



### Entrenamiento 2:

Una de las preocupaciones ante este primer entrenamiento es la posible tendencia del modelo a solo tomar como violentos mensajes que digan explícitamente “te voy a matar” dado a que esta es una frase muy común encontrada en el dataset creado.

Dataset:

19 frases se repiten diciendo solo “te voy a matar”

228 frases contienen la frase “te voy a matar”

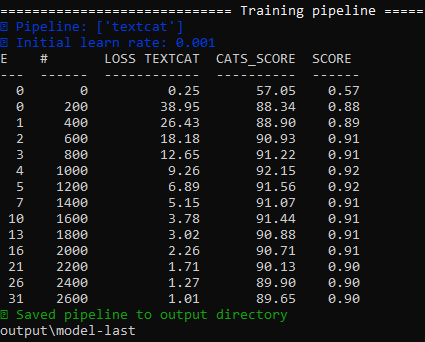
Total: 2266 No violentas y 1823 Violentas (aproximadamente)

De 1823 frases violentas 228 contienen “te voy a matar” lo que representa un 12,5%

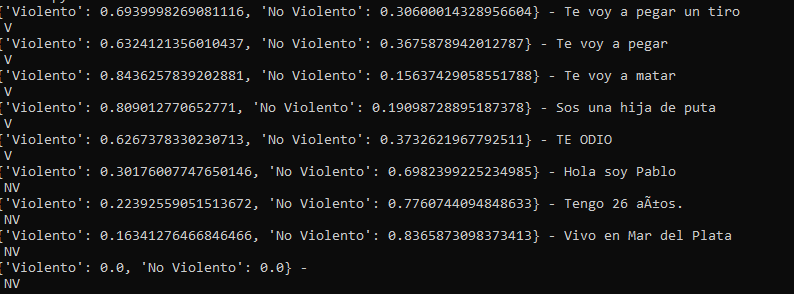
-Elimine las repeticiones de la frase exacta “te voy a matar” (dejando las variaciones) datos-min-1.json)

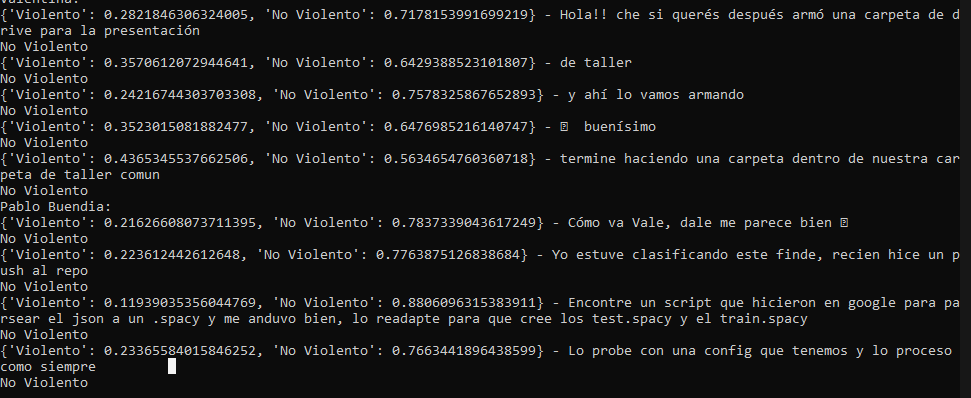
-Eliminé algunos ‘No violentos’ para intentar tener un ratio más parejo.

batch size 800



Primeros resultados:

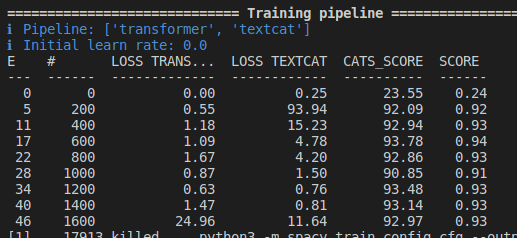




### Entrenamiento 2 con transformers:

Para el entrenamiento 2, con el mismo dataset mejorado ‘datos-min\_2.json’ se cambió la configuración para usar transformers.

Batch size 500



A diferencia del entrenamiento NER, donde había una mejora considerable de 10 puntos en el fscore en los resultados, en este caso el score apenas mejoró 1 punto, de 92 a 93.

Las pruebas sin embargo mostraron algo muy diferente:

### 

Valentina:

{'Violento': 0.0005077447276562452, 'No Violento': 0.9994922876358032} - Hola!! che si querés después armó una carpeta de drive para la presentación

No Violento

{'Violento': 0.002652459079399705, 'No Violento': 0.9973474740982056} - de taller

No Violento

{'Violento': 0.0005447620642371476, 'No Violento': 0.9994552731513977} - y ahí lo vamos armando

No Violento

{'Violento': 0.0005043275305069983, 'No Violento': 0.9994956254959106} - 😄 buenísimo

No Violento

{'Violento': 0.0005220160237513483, 'No Violento': 0.9994780421257019} - termine haciendo una carpeta dentro de nuestra carpeta de taller comun

No Violento

Pablo Buendia:

{'Violento': 0.0005237988661974669, 'No Violento': 0.9994762539863586} - Cómo va Vale, dale me parece bien 👌

No Violento

{'Violento': 0.0004954840405844152, 'No Violento': 0.999504566192627} - Yo estuve clasificando este finde, recien hice un push al repo

No Violento

{'Violento': 0.0005005056736990809, 'No Violento': 0.9994994401931763} - Encontre un script que hicieron en google para parsear el json a un .spacy y me anduvo bien, lo readapte para que cree los test.spacy y el train.spacy

No Violento

{'Violento': 0.0005011692410334945, 'No Violento': 0.9994988441467285} - Lo probe con una config que tenemos y lo proceso como siempre

No Violento

{'Violento': 0.9978233575820923, 'No Violento': 0.0021767013240605593} - Te voy a pegar un tiro

V

{'Violento': 0.9978284239768982, 'No Violento': 0.0021715587936341763} - Te voy a pegar

V

{'Violento': 0.9978395700454712, 'No Violento': 0.0021604441571980715} - Te voy a matar

V

{'Violento': 0.9978491067886353, 'No Violento': 0.0021509474609047174} - Sos una hija de puta

V

{'Violento': 0.9978340268135071, 'No Violento': 0.0021660206839442253} - TE ODIO

V

{'Violento': 0.0005257056327536702, 'No Violento': 0.9994743466377258} - Hola soy Pablo

NV

{'Violento': 0.0006027527269907296, 'No Violento': 0.9993972778320312} - Tengo 26 años.

NV

{'Violento': 0.00056479312479496, 'No Violento': 0.9994351267814636} - Vivo en Mar del Plata

NV

{'Violento': 0.0, 'No Violento': 0.0} -

NV

En las pruebas sin transformers, los resultados no tenían un score demasiado inclinado para ninguno de los dos resultados. Por ejemplo, para la frase

‘Hola soy Pablo’

Tenía valores (0,30 Violento; 0,70 No Violento). Es decir, favorecía la correcta decisión de que no era una frase violenta, pero no definitivamente.

En cambio, en el modelo con transformers, los valores son (0,0005 Violento, 0,9995 No Violento). Correctamente el modelo entrenado con transformers que la frase es completamente común. Así con todas las otras frases.

### Modelo Entrenamiento 3 con transformers

Para el entrenamiento 3, se usó el mismo dataset que en el 2, pero se hizo un reetiquetado de varias frases que antes se consideraban violentas por tener algún insulto o hacer mención a un tema como la muerte, pero no eran violentas hacia un persona o no tenían ese contexto. Básicamente la razón por la que el entrenamiento 2 con transformers percibió como violentas frases como “sos una genia”

### Modelo Entrenamiento 4 con transformers

Se agregaron al dataset 700 ejemplos encontrados de mensajes violentos más variados, usados para el modelo multicategoria. Este entrenamiento llegó a un fscore de 0.99

Frases problemáticas:

“te voy a entrenar al gato/perro/etc.”

“te voy a alimentar al perro”

“te voy a pasear al gato”

“feminazis”

Está haciendo overfitting. Algunas posibles recomendaciones:

-intentar partir del modelo pre entrenado de transformers usando proyectos

- frenar el entrenamiento antes.

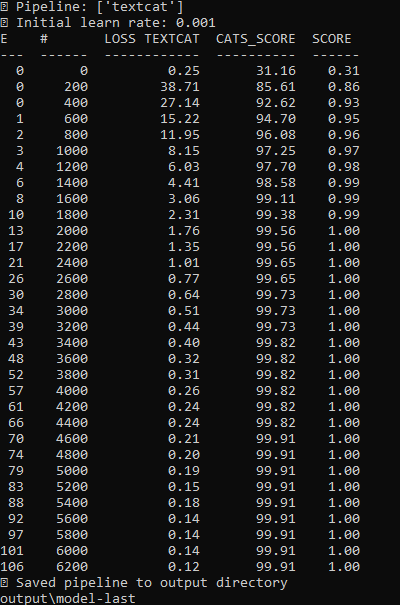
-agregar frases neutras con ‘score neutro’. (esto capaz es complicado, pero es una opción. También se pueden considerar como neutras aquellas que logran una predicción muy cercana a 0,5) -> capaz para modelos entrenados sin los transformers o con el loss más grande.

-aumentar el número de ejemplos

**Posible problema con violencia económica:** Como este tipo de violencia muchas veces no se da en un plano conversacional o depende mucho de un contexto externo, podría ser mejor directamente no incluirla. Algo como “Que compraste con mi tarjeta? me logo el resumen y hay plata que yo no gaste, antes de usar mi tarjeta avisame o preguntame” puede ser marcado como violencia económica cuando en realidad no necesariamente lo es. Si bien frases como “te voy a dejar en la calle” son violentas pueden ser categorizadas (por ejemplo) como posible violencia psicológica. Por ahora da buenos resultados hay que ver si logramos un dataset balanceado.

### Modelo Entrenamiento 4

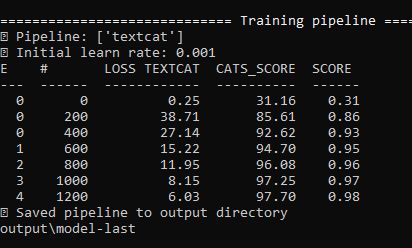
Usando los mismos ejemplos y una configuración estándar (y un batch size de 800) los resultados parecen demasiado buenos. El modelo termina con un score ‘perfecto’ usando solo esos 4500 ejemplos.



Esto indicaría que se está haciendo overfitting, por sobreentrenamiento. Es decir el modelo está aprendiendo demasiado bien los ejemplos del dataset y no funciona tan bien en ejemplos de nuevas situaciones.

Intentamos cambiar las configuraciones default: cargando el modelo es\_core\_news\_md que si tiene el uso de vectores disponible, usando un batch size de 1000. Pero los resultados no variaron.

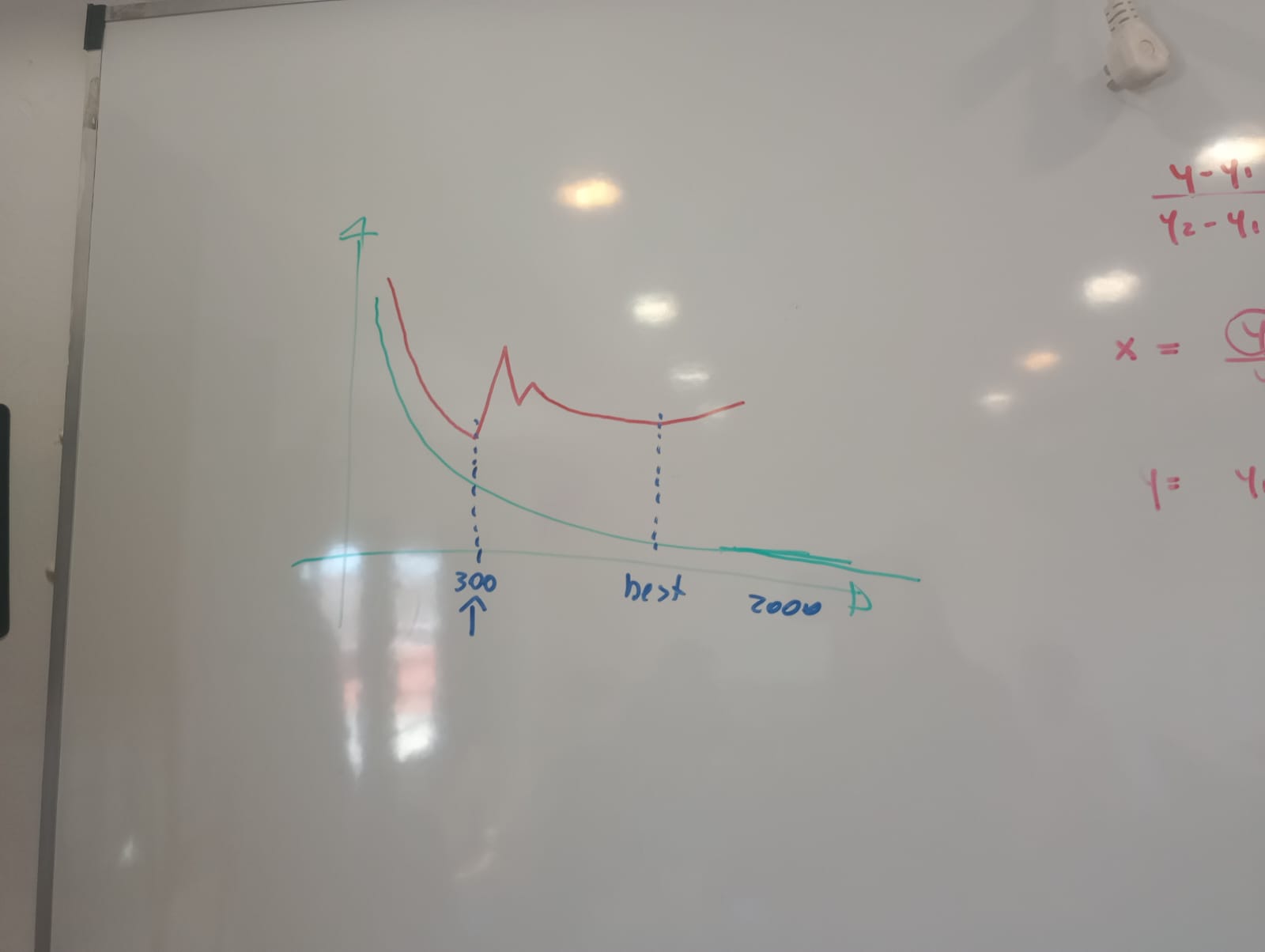
Entonces para evitar el overfitting se decidió hacer early stopping (parar el entrenamiento después de una serie fija de épocas)



Posteriormente debemos testear:

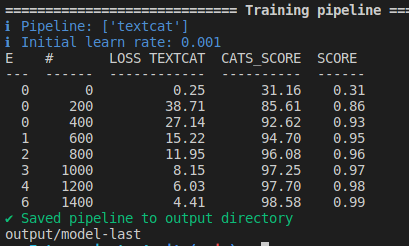
* training parámetro patience: el default es 5000, con max steps 10000 si durante 5000 steps no mejora la métrica corta dice que es el best y si mejora la métrica sigue. Podemos disminuir la patience, esto eventualmente podría hacer que el modelo generalice mejor. (disminuir el overfitting)

gráfico explicando el crecimiento del error marcado en el score(verde) y el error real de validación (rojo)

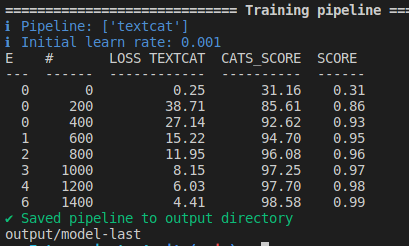


Para la misma configuración y los mismos dataset, se hicieron entrenamientos con distintas configuraciones para el parámetro “patience”:

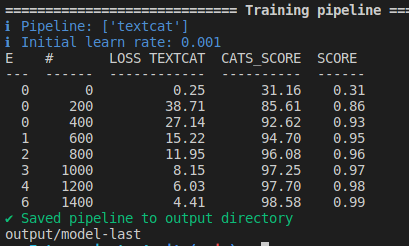
Patience = 500



Patience = 2000



Patience = 10000



No se observan cambios en los score en los entrenamientos. Además, los resultados de las pruebas con los distintos modelos fueron los mismos que los del entrenamiento 4 inicial.

### 

### Entrenamiento 5.05

Durante el entrenamiento 5 se hicieron entrenamientos con dos tipos de datasets:

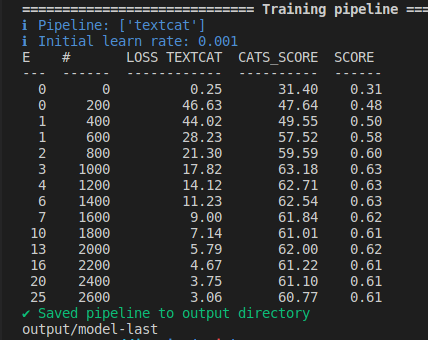
**Un dataset compuesto con frases de Reddit etiquetadas.**

En total,

1570 frases violentas.

674 frases no violentas.

El score máximo fue de 63, con un score para la categoría Violento de 83 y de No Violento de 42.



Pruebas

{'Violento': 0.714601457118988, 'No Violento': 0.2853985130786896} - Te voy a pegar un tiro

V

{'Violento': 0.6660866737365723, 'No Violento': 0.33391332626342773} - Te voy a pegar

V

{'Violento': 0.6946253776550293, 'No Violento': 0.3053746521472931} - Te voy a matar

V

{'Violento': 0.799659252166748, 'No Violento': 0.20034076273441315} - Sos una hija de puta

V

{'Violento': 0.5235807299613953, 'No Violento': 0.47641927003860474} - TE ODIO

V

{'Violento': 0.5672472715377808, 'No Violento': 0.4327527582645416} - Sos una trola

V

{'Violento': 0.646014392375946, 'No Violento': 0.35398557782173157} - Hola soy Pablo

V

{'Violento': 0.518568754196167, 'No Violento': 0.4814312160015106} - Tengo 26 años.

V

{'Violento': 0.43270543217658997, 'No Violento': 0.5672945380210876} - Vivo en Mar del Plata

NV

{'Violento': 0.5760480165481567, 'No Violento': 0.4239519536495209} - Sos una genia

V

{'Violento': 0.4828243553638458, 'No Violento': 0.5171756148338318} - Vamos a cenar a algun lado

NV

{'Violento': 0.6613060832023621, 'No Violento': 0.33869391679763794} - Llegue re tarde me quiero matar

V

{'Violento': 0.6507004499435425, 'No Violento': 0.3492995798587799} - Que pelotuda que soy

V

{'Violento': 0.5458166599273682, 'No Violento': 0.45418334007263184} - jajajaj si llega a pasar eso me mato

V

{'Violento': 0.5895697474479675, 'No Violento': 0.41043031215667725} - no tengo la más puta idea de que puede ser

V

{'Violento': 0.8040476441383362, 'No Violento': 0.1959524005651474} - Como mierda hacen los jugadores de aldosivi para entrenar con el olor a pescado este de mdq la puta madre

V

{'Violento': 0.46916618943214417, 'No Violento': 0.5308338403701782} - jajaja seguí hablando estupideces y vas a terminar mal

NV

{'Violento': 0.6584547758102417, 'No Violento': 0.3415452241897583} - jajaj te voy a hacer mierda

V

{'Violento': 0.38069894909858704, 'No Violento': 0.6193010807037354} - no te van a encontrar más

NV

{'Violento': 0.48342856764793396, 'No Violento': 0.5165714621543884} - no apareces

NV

{'Violento': 0.5552108287811279, 'No Violento': 0.4447891414165497} - tene mucho cuidado conmigo

V

{'Violento': 0.3719924986362457, 'No Violento': 0.6280075311660767} - vamos a coger quieras o no

NV

{'Violento': 0.4620818495750427, 'No Violento': 0.5379180908203125} - volve a la cocina que es para lo que servis

NV

{'Violento': 0.721910834312439, 'No Violento': 0.2780891954898834} - te voy a matar al perro

V

{'Violento': 0.5530200004577637, 'No Violento': 0.44697993993759155} - voy a sacar a pasear al perro

V

{'Violento': 0.0, 'No Violento': 0.0} -

NV

### Entrenamiento 5

**Un dataset compuesto con frases de Reddit, más un agregado de frases no violentas provenientes del entrenamiento 4.**

En total,

1542 frases violentas.

1499 frases no violentas.

El score máximo fue de 80, con un score para la categoría Violento de 81 y de No Violento de 79.

El agregado de las frases no violentas produjo un aumento en el score de No Violento, pero no produjo un aumento en el score de la categoría de Violento.

# 

Pruebas:

{'Violento': 0.42888975143432617, 'No Violento': 0.5711102485656738} - Te voy a pegar un tiro

NV

{'Violento': 0.3905152976512909, 'No Violento': 0.6094847321510315} - Te voy a pegar

NV

{'Violento': 0.46474918723106384, 'No Violento': 0.5352507829666138} - Te voy a matar

NV

{'Violento': 0.754928469657898, 'No Violento': 0.24507156014442444} - Sos una hija de puta

V

{'Violento': 0.36966317892074585, 'No Violento': 0.6303368210792542} - TE ODIO

NV

{'Violento': 0.24088029563426971, 'No Violento': 0.7591196894645691} - Sos una trola

NV

{'Violento': 0.4957963526248932, 'No Violento': 0.5042036771774292} - Hola soy Pablo

NV

{'Violento': 0.21997730433940887, 'No Violento': 0.7800227403640747} - Tengo 26 años.

NV

{'Violento': 0.34195542335510254, 'No Violento': 0.6580445766448975} - Vivo en Mar del Plata

NV

{'Violento': 0.4594712555408478, 'No Violento': 0.5405287146568298} - Sos una genia

NV

{'Violento': 0.389926552772522, 'No Violento': 0.6100735068321228} - Vamos a cenar a algun lado

NV

{'Violento': 0.34199008345603943, 'No Violento': 0.658009946346283} - Llegue re tarde me quiero matar

NV

{'Violento': 0.42511624097824097, 'No Violento': 0.574883759021759} - Que pelotuda que soy

NV

{'Violento': 0.27919653058052063, 'No Violento': 0.720803439617157} - jajajaj si llega a pasar eso me mato

NV

{'Violento': 0.5588929653167725, 'No Violento': 0.44110700488090515} - no tengo la más puta idea de que puede ser

V

{'Violento': 0.666582465171814, 'No Violento': 0.33341753482818604} - Como mierda hacen los jugadores de aldosivi para entrenar con el olor a pescado este de mdq la puta madre

V

{'Violento': 0.5147814750671387, 'No Violento': 0.48521849513053894} - jajaja seguí hablando estupideces y vas a terminar mal

V

{'Violento': 0.6147348880767822, 'No Violento': 0.38526517152786255} - jajaj te voy a hacer mierda

V

{'Violento': 0.40960457921028137, 'No Violento': 0.590395450592041} - no te van a encontrar más

NV

{'Violento': 0.40117037296295166, 'No Violento': 0.5988296270370483} - no apareces

NV

{'Violento': 0.5104177594184875, 'No Violento': 0.48958224058151245} - tene mucho cuidado conmigo

V

{'Violento': 0.28357312083244324, 'No Violento': 0.7164269089698792} - vamos a coger quieras o no

NV

{'Violento': 0.3660862445831299, 'No Violento': 0.6339138150215149} - volve a la cocina que es para lo que servis

NV

{'Violento': 0.7289817929267883, 'No Violento': 0.2710181474685669} - te voy a matar al perro

V

{'Violento': 0.44240307807922363, 'No Violento': 0.5575969815254211} - voy a sacar a pasear al perro

NV

{'Violento': 0.0, 'No Violento': 0.0} -

NV

### Entrenamiento 5 con frases de reddit y frases entrenamiento 4 completo

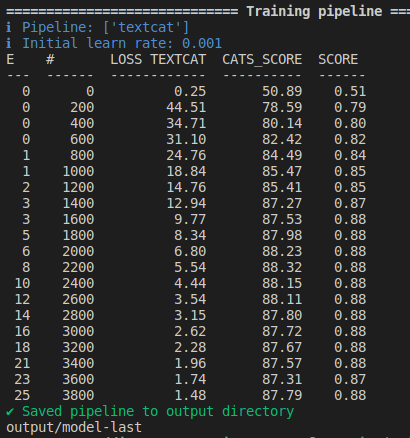
Todo el dataset 4 y el dataset 5.1 (+ frases no violentas si es necesario balancear)

En total,

2930 frases violentas.

3680 frases no violentas.

El score obtenido fue de 88, con un score de la categoria “Violento” de 87 y de la categoria “No Violento” de 89.



{'Violento': 0.8973874449729919, 'No Violento': 0.10261257737874985} - Te voy a pegar un tiro

V

{'Violento': 0.8290730714797974, 'No Violento': 0.17092694342136383} - Te voy a pegar

V

{'Violento': 0.8951576948165894, 'No Violento': 0.10484223812818527} - Te voy a matar

V

{'Violento': 0.8970983624458313, 'No Violento': 0.10290168225765228} - Sos una hija de puta

V

{'Violento': 0.6973831057548523, 'No Violento': 0.3026168644428253} - TE ODIO

V

{'Violento': 0.6853758096694946, 'No Violento': 0.314624160528183} - Sos una trola

V

{'Violento': 0.2710515260696411, 'No Violento': 0.7289484739303589} - Hola soy Pablo

NV

{'Violento': 0.14982788264751434, 'No Violento': 0.8501721620559692} - Tengo 26 años.

NV

{'Violento': 0.15734048187732697, 'No Violento': 0.8426595330238342} - Vivo en Mar del Plata

NV

{'Violento': 0.5004762411117554, 'No Violento': 0.49952375888824463} - Sos una genia

V

{'Violento': 0.20408889651298523, 'No Violento': 0.7959110736846924} - Vamos a cenar a algun lado

NV

{'Violento': 0.35641178488731384, 'No Violento': 0.6435881853103638} - Llegue re tarde me quiero matar

NV

{'Violento': 0.49018731713294983, 'No Violento': 0.5098127126693726} - Que pelotuda que soy

NV

{'Violento': 0.6313979625701904, 'No Violento': 0.36860203742980957} - jajajaj si llega a pasar eso me mato

V

{'Violento': 0.506307065486908, 'No Violento': 0.49369290471076965} - no tengo la más puta idea de que puede ser

V

{'Violento': 0.8543970584869385, 'No Violento': 0.14560295641422272} - Como mierda hacen los jugadores de aldosivi para entrenar con el olor a pescado este de mdq la puta madre

V

{'Violento': 0.22459916770458221, 'No Violento': 0.7754008173942566} - jajaja seguí hablando estupideces y vas a terminar mal

NV

{'Violento': 0.8729429244995117, 'No Violento': 0.1270570158958435} - jajaj te voy a hacer mierda

V

{'Violento': 0.31925398111343384, 'No Violento': 0.6807459592819214} - no te van a encontrar más

NV

{'Violento': 0.3286680281162262, 'No Violento': 0.6713319420814514} - no apareces

NV

{'Violento': 0.14690254628658295, 'No Violento': 0.8530974388122559} - tene mucho cuidado conmigo

NV

{'Violento': 0.3931216597557068, 'No Violento': 0.606878399848938} - vamos a coger quieras o no

NV

{'Violento': 0.3312560021877289, 'No Violento': 0.6687439680099487} - volve a la cocina que es para lo que servis

NV

{'Violento': 0.9488353729248047, 'No Violento': 0.051164645701646805} - te voy a matar al perro

V

{'Violento': 0.7155808210372925, 'No Violento': 0.2844192087650299} - voy a sacar a pasear al perro

V

{'Violento': 0.0, 'No Violento': 0.0} -

NV



Usando directamente V > No V hay muchos falsos positivos.



### Entrenamiento 7

Tomamos el modelo anterior y pasamos este sobre los archivos de reddit de 2014 (menos lo que no este en español y no sean memes conocidos)

Todo el dataset 4 + el dataset de frases de 2013 etiquetados por nosotros + lo revisado de 2014 + frases adicionales no violentas para balancear el modelo.

nota: testear en un entrenamiento anterior se habia encontrado un par de frases con fuertes connotaciones de violencia sexual y no las marcaba como violentas y notar falsos positivos ‘el hijo de maria’, la formula ‘hijo de’ (..)’ da violento casi siempre (notar también que un falso positivo es menos indeseable que un falso negativo, debido a que son mucho más faciles de detectar, y sus consecuencias son más livianas)

### Entrenamiento 7 normalizado

Los entrenamientos anteriores probaron que había un número de problemas en la clasificación de textos de manera binaria.

Para mostrar estos problemas, se utilizó un dataset de prueba especial para el modelo entrenado en el entrenamiento 5.4

El dataset está compuesto por las siguientes frases:

| Frases No Violentas Simples | Frases Violentas Simples | Frases No Violentas Ambiguas | Frases Violentas Ambiguas |
| --- | --- | --- | --- |
| Hola, soy Pablo. | Te voy a matar. | Hola, te voy a buscar. | Hola, si te metés conmigo te voy a buscar. |
| Vivo en Mar del Plata. | Te voy a violar | Se donde vivís. | Guarda conmigo, porque yo se dónde vivís. |
| Sos una genia | Sos una puta. | Ese es mi perro. | Sos una perra. |
| Te voy a ayudar. | Te voy a pegar un tiro. | Te voy a dar una mano con el parcial. | Te voy a dejar en la calle. |
| Voy a sacar a pasear al perro. | Voy a pegarte un balazo en la cabeza. | Me voy para casa. | Voy a ir a tu casa con un fierro. |
| Ayer perdí plata en el casino. | Hija de puta me hiciste perder plata. | Solo necesito un poco más de plata | Solo me haces perder plata |
| ¿Tenes fuego? | Te voy a prender fuego. | Mandale un fueguito. | Te prendo fuego. |
| Andá a trabajar | Andá a trabajar negra de mierda | Que noche negra para el fútbol | Que negra boluda que sos. |
| ¿Cogemos? | Vamos a coger quieras o no. | No quiero coger hoy. | Que incogible que sos. |
| ¿Rompiste el celular? | Rompes todo y yo tengo que pagarlo | Te rompí el plato sin querer. | Si yo quiero te puedo romper la casa. |
| Tomé mucho frio en la garganta anoche. | Mira este cuchillo lo voy a meter en tu garganta | Ese cuchillo viene con el kit. | Ese cuchillo te lo puedo meter bien adentro tuyo. |

| Clasificación | Score Violento | Score No Violento | Frase |
| --- | --- | --- | --- |
| No Violento | 0.32 | 0.68 | Hola, soy Pablo. |
| Violento | 0.91 | 0.09 | Te voy a matar. |
| Violento | 0.77 | 0.23 | Hola, te voy a buscar. |
| Violento | 0.87 | 0.13 | Hola, si te metés conmigo te voy a buscar. |
| No Violento | 0.18 | 0.82 | Vivo en Mar del Plata. |
| Violento | 0.8 | 0.2 | Te voy a violar |
| No Violento | 0.17 | 0.83 | Se donde vivís. |
| No Violento | 0.25 | 0.75 | Guarda conmigo, porque yo se dónde vivís. |
| Violento | 0.5 | 0.5 | Sos una genia |
| Violento | 0.86 | 0.14 | Sos una puta. |
| No Violento | 0.39 | 0.61 | Ese es mi perro. |
| Violento | 0.75 | 0.25 | Sos una perra. |
| Violento | 0.53 | 0.47 | Te voy a ayudar. |
| Violento | 0.91 | 0.09 | Te voy a pegar un tiro. |
| No Violento | 0.47 | 0.53 | Te voy a dar una mano con el parcial. |
| Violento | 0.64 | 0.36 | Te voy a dejar en la calle. |
| No Violento | 0.43 | 0.57 | Voy a sacar a pasear al perro. |
| No Violento | 0.21 | 0.79 | Voy a pegarte un balazo en la cabeza. |
| No Violento | 0.11 | 0.89 | Me voy para casa. |
| No Violento | 0.17 | 0.83 | Voy a ir a tu casa con un fierro. |
| No Violento | 0.11 | 0.89 | Ayer perdí plata en el casino. |
| Violento | 0.53 | 0.47 | Hija de puta me hiciste perder plata. |
| No Violento | 0.07 | 0.93 | Solo necesito un poco más de plata |
| No Violento | 0.13 | 0.87 | Solo me haces perder plata |
| No Violento | 0.21 | 0.79 | ¿Tenes fuego? |
| Violento | 0.74 | 0.26 | Te voy a prender fuego. |
| No Violento | 0.4 | 0.6 | Mandale un fueguito. |
| No Violento | 0.46 | 0.54 | Te prendo fuego. |
| No Violento | 0.18 | 0.82 | Andá a trabajar |
| Violento | 0.8 | 0.2 | Andá a trabajar negra de mierda |
| No Violento | 0.28 | 0.72 | Que noche negra para el fútbol |
| Violento | 0.63 | 0.37 | Que negra boluda que sos. |
| No Violento | 0.22 | 0.78 | ¿Cogemos? |
| No Violento | 0.37 | 0.63 | Vamos a coger quieras o no. |
| No Violento | 0.21 | 0.79 | No quiero coger hoy. |
| No Violento | 0.38 | 0.62 | Que incogible que sos. |
| No Violento | 0.17 | 0.83 | ¿Rompiste el celular? |
| No Violento | 0.19 | 0.81 | Rompes todo y yo tengo que pagarlo |
| No Violento | 0.2 | 0.8 | Te rompí el plato sin querer. |
| No Violento | 0.17 | 0.83 | Si yo quiero te puedo romper la casa. |
| No Violento | 0.32 | 0.68 | Tomé mucho frio en la garganta anoche. |
| Violento | 0.75 | 0.25 | Mira este cuchillo lo voy a meter en tu garganta |
| No Violento | 0.34 | 0.66 | Ese cuchillo viene con el kit. |
| Violento | 0.69 | 0.31 | Ese cuchillo te lo puedo meter bien adentro tuyo. |

Los problemas principales de estas pruebas se pueden dividir en:

1. Las frases del tipo “Te voy a” o “Sos una” tienden a ser clasficadas como violentas para cualquier combinación. Por ejemplo con la frase “Te voy a ayudar” es clasificada como “Violento”. Esto es porque muchos de los ejemplos de frases violentas comienzan de esa manera, por lo que el clasificador tiende a añadir un componente violento a todas estas frases.
2. Existen algunas frases violentas típicas “te voy a prendo fuego”, que todavía no son reconocidas cuando se las cambia ligeramente “te prendo fuego”.
3. Algunas frases no violentas que tienen alguna palabra muy parecida a la de una frase violenta se les asigna un score “Violento” demasiado alto. Por ejemplo, la frase no violenta “Ese es mi perro” tiene un score de violento de 0.39, porque la palabra “perro” es muy parecida a “perra”, usada en la frase violenta “Sos una perra”.
4. Muchas frases violentas pero que no incluyen insultos normalmente, por ejemplo del tipo de violencia económica (“Solo me haces perder plata”) o psicológica (“Guarda conmigo, porque sé donde vivís”) no son detectadas correctamente como violentas por falta de ejemplos en el dataset.

Por tal motivo, para solventar cada uno de estos cuatro problemas se modificó el dataset del entrenamiento 5.4 en las siguientes etapas:

Etapa 1: Limpieza de signos y normalización de frases.

1- Se hizo una limpieza de las frases. Se removieron los siguientes signos de puntuación:  
& gt

;

.

\”

\/

/

\

[

]

\*

‘

-

|

(

)

!

:

,

»

¡

!

+

…

..

…

🤩🤤😭😉🥰❤️❤😮‍💨🔥😍💋😍💘🥴🙄😅😩⭐🤯😂😘✨🤭😤😬💕🙌📉🤰🙂🥲🥳🌚🥂🥺👏🏻💔😓🙃🤷‍♀️😆👀🧐📸❤‍🩹🔫🫀💝🧠🫀👍😊🫶🏻♥️🏆🇦🇷🙏💪👉😣😖🤮😜😥😔💖🤪🤢😶🤦🏻‍♀️😏😵‍💫😝⛈️🌅🫣🤦‍♀️💀🤍😑🥵🤒😪💗😳🫢😁🧉💞❣😟❤🚬💜💬🤔😇😋🥱🤗😌😃💪🤑😐🇶🇦🇧🇷😎🔵⚪️🔴😚😡👎🤏🏻🦭🧡👼🤷🏻‍♀️

Se eliminaron todos los espacios excesivos y se dejó uno solo

Se eliminaron todos los signos de pregunta excesivos y se dejó uno solo

Se eliminaron todas las tildes de acentuación de las letras

Se pasó todo a minúscula

No se añadieron frases nuevas.

Despues de esto se volvio a entrenar el modelo,

| Clasificacion | Score Violento | Score No Violento | Frase |
| --- | --- | --- | --- |
| No Violento | 0.15 | 0.85 | Hola, soy Pablo. |
| Violento | 0.82 | 0.18 | Te voy a matar. |
| No Violento | 0.47 | 0.53 | Hola, te voy a buscar. |
| Violento | 0.61 | 0.39 | Hola, si te metés conmigo te voy a buscar. |
| No Violento | 0.22 | 0.78 | Vivo en Mar del Plata. |
| Violento | 0.7 | 0.3 | Te voy a violar |
| No Violento | 0.23 | 0.77 | Se donde vivís. |
| No Violento | 0.17 | 0.83 | Guarda conmigo, porque yo se dónde vivís. |
| No Violento | 0.27 | 0.73 | Sos una genia |
| Violento | 0.68 | 0.32 | Sos una puta. |
| No Violento | 0.24 | 0.76 | Ese es mi perro. |
| Violento | 0.54 | 0.46 | Sos una perra. |
| No Violento | 0.27 | 0.73 | Te voy a ayudar. |
| Violento | 0.82 | 0.18 | Te voy a pegar un tiro. |
| No Violento | 0.28 | 0.72 | Te voy a dar una mano con el parcial. |
| No Violento | 0.44 | 0.56 | Te voy a dejar en la calle. |
| No Violento | 0.39 | 0.61 | Voy a sacar a pasear al perro. |
| No Violento | 0.21 | 0.79 | Voy a pegarte un balazo en la cabeza. |
| No Violento | 0.21 | 0.79 | Me voy para casa. |
| No Violento | 0.23 | 0.77 | Voy a ir a tu casa con un fierro. |
| No Violento | 0.15 | 0.85 | Ayer perdí plata en el casino. |
| No Violento | 0.39 | 0.61 | Hija de puta me hiciste perder plata. |
| No Violento | 0.05 | 0.95 | Solo necesito un poco más de plata |
| No Violento | 0.17 | 0.83 | Solo me haces perder plata |
| No Violento | 0.37 | 0.63 | ¿Tenes fuego? |
| Violento | 0.73 | 0.27 | Te voy a prender fuego. |
| No Violento | 0.29 | 0.71 | Mandale un fueguito. |
| No Violento | 0.47 | 0.53 | Te prendo fuego. |
| No Violento | 0.22 | 0.78 | Andá a trabajar |
| Violento | 0.86 | 0.14 | Andá a trabajar negra de mierda |
| No Violento | 0.45 | 0.55 | Que noche negra para el fútbol |
| Violento | 0.75 | 0.25 | Que negra boluda que sos. |
| No Violento | 0.19 | 0.81 | ¿Cogemos? |
| No Violento | 0.39 | 0.61 | Vamos a coger quieras o no. |
| No Violento | 0.16 | 0.84 | No quiero coger hoy. |
| No Violento | 0.43 | 0.57 | Que incogible que sos. |
| No Violento | 0.15 | 0.85 | ¿Rompiste el celular? |
| No Violento | 0.12 | 0.88 | Rompes todo y yo tengo que pagarlo |
| No Violento | 0.19 | 0.81 | Te rompí el plato sin querer. |
| No Violento | 0.14 | 0.86 | Si yo quiero te puedo romper la casa. |
| No Violento | 0.09 | 0.91 | Tomé mucho frio en la garganta anoche. |
| Violento | 0.8 | 0.2 | Mira este cuchillo lo voy a meter en tu garganta |
| No Violento | 0.31 | 0.69 | Ese cuchillo viene con el kit. |
| Violento | 0.84 | 0.16 | Ese cuchillo te lo puedo meter bien adentro tuyo. |

Se observan mejoras en los problemas descriptos en los puntos 1 y 3. Por el contrario, se muestra un empeoramiento del problema descripto en el punto 4 y 2.

Etapa 2:

Se agregaron variaciones de las frases tanto violentas como no violentas. Por ejemplo para evitar el overfitting con los insultos de “te voy a”, se agregaron mas ejemplos inofensivos “te voy a ayudar con…”.

Se reemplazaron expresiones e insultos típicos de otros países como México por expresiones argentinas. Por ejemplo se reemplazo “pinche” con “puta” o “forro”.

Se descompusieron frases que eran muy largas en componentes violentos y no violentos, o se eliminaron completamente para reducir la ambiguedad. Se hizo un corte para remover todas las frases con una longitud mayor a 200 caracteres, ya que podian confundir al clasificador.

Se removieron frases duplicadas y se mejoró el script para garantizar un mejor balance entre frases violentas y no violentas para los datasets de test y train.

| Clasificacion | Score Violento | Score No Violento | Frase |
| --- | --- | --- | --- |
| No Violento | 0.33 | 0.67 | Hola, soy Pablo. |
| Violento | 0.68 | 0.32 | Te voy a matar. |
| No Violento | 0.49 | 0.51 | Hola, te voy a buscar. |
| Violento | 0.7 | 0.3 | Hola, si te metés conmigo te voy a buscar. |
| No Violento | 0.22 | 0.78 | Vivo en Mar del Plata. |
| Violento | 0.58 | 0.42 | Te voy a violar |
| No Violento | 0.39 | 0.61 | Se donde vivís. |
| No Violento | 0.34 | 0.66 | Guarda conmigo, porque yo se dónde vivís. |
| No Violento | 0.43 | 0.57 | Sos una genia |
| Violento | 0.62 | 0.38 | Sos una puta. |
| No Violento | 0.35 | 0.65 | Ese es mi perro. |
| Violento | 0.51 | 0.49 | Sos una perra. |
| No Violento | 0.38 | 0.62 | Te voy a ayudar. |
| Violento | 0.67 | 0.33 | Te voy a pegar un tiro. |
| No Violento | 0.36 | 0.64 | Te voy a dar una mano con el parcial. |
| No Violento | 0.45 | 0.55 | Te voy a dejar en la calle. |
| No Violento | 0.36 | 0.64 | Voy a sacar a pasear al perro. |
| No Violento | 0.43 | 0.57 | Voy a pegarte un balazo en la cabeza. |
| No Violento | 0.42 | 0.58 | Me voy para casa. |
| No Violento | 0.43 | 0.57 | Voy a ir a tu casa con un fierro. |
| No Violento | 0.29 | 0.71 | Ayer perdí plata en el casino. |
| Violento | 0.52 | 0.48 | Hija de puta me hiciste perder plata. |
| No Violento | 0.27 | 0.73 | Solo necesito un poco más de plata |
| Violento | 0.55 | 0.45 | Solo me haces perder plata |
| No Violento | 0.34 | 0.66 | ¿Tenes fuego? |
| Violento | 0.62 | 0.38 | Te voy a prender fuego. |
| No Violento | 0.42 | 0.58 | Mandale un fueguito. |
| Violento | 0.52 | 0.48 | Te prendo fuego. |
| No Violento | 0.36 | 0.64 | Andá a trabajar |
| Violento | 0.73 | 0.27 | Andá a trabajar negra de mierda |
| No Violento | 0.45 | 0.55 | Que noche negra para el fútbol |
| Violento | 0.59 | 0.41 | Que negra boluda que sos. |
| No Violento | 0.28 | 0.72 | ¿Cogemos? |
| Violento | 0.5 | 0.5 | Vamos a coger quieras o no. |
| No Violento | 0.25 | 0.75 | No quiero coger hoy. |
| Violento | 0.57 | 0.43 | Que incogible que sos. |
| No Violento | 0.25 | 0.75 | ¿Rompiste el celular? |
| No Violento | 0.26 | 0.74 | Rompes todo y yo tengo que pagarlo |
| No Violento | 0.29 | 0.71 | Te rompí el plato sin querer. |
| No Violento | 0.29 | 0.71 | Si yo quiero te puedo romper la casa. |
| No Violento | 0.27 | 0.73 | Tomé mucho frio en la garganta anoche. |
| Violento | 0.75 | 0.25 | Mira este cuchillo lo voy a meter en tu garganta |
| No Violento | 0.39 | 0.61 | Ese cuchillo viene con el kit. |
| Violento | 0.7 | 0.3 | Ese cuchillo te lo puedo meter bien adentro tuyo. |

Esta vez se puede ver una mejoría general en la percepción del clasificador para los puntos 1, 2 y 3. Es decir, es mejor diferenciando las frases violentas de las no violentas. Sin embargo todavía no hay una mejora clara del punto 4. Es decir todavía no detecta correctamente frases violentas (mayoritariamente del tipo económica o psicológica) que no tienen un insulto explícito.

Etapa 3:

Durante la etapa 3 se busca solucionar el último problema mediante el agregado de frases de violencia económica y psicológica del dataset del clasificador multi-categoría. Para evitar un overfitting también se añadieron frases no violentas de aspecto parecido. Por ejemplo se agregaron frases del tipo: “sos una genia”, “sos increible”, “sos la mejor amigaa”, para evitar que las frases del tipo “sos una..” sean automáticamente consideradas violentas.

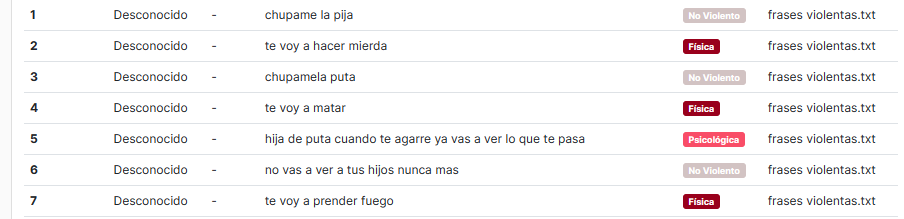
En total, se añadieron alrededor de 300 frases nuevas.

El dataset final se encuentra compuesto de 4072 frases.

| Clasificacion | Score Violento | Score No Violento | Frase |
| --- | --- | --- | --- |
| No Violento | 0.4 | 0.6 | Hola, soy Pablo. |
| Violento | 0.7 | 0.3 | Te voy a matar. |
| Violento | 0.61 | 0.39 | Hola, te voy a buscar. |
| Violento | 0.83 | 0.17 | Hola, si te metés conmigo te voy a buscar. |
| No Violento | 0.25 | 0.75 | Vivo en Mar del Plata. |
| Violento | 0.68 | 0.32 | Te voy a violar |
| No Violento | 0.39 | 0.61 | Se donde vivís. |
| Violento | 0.54 | 0.46 | Guarda conmigo, porque yo se dónde vivís. |
| No Violento | 0.35 | 0.65 | Sos una genia |
| Violento | 0.62 | 0.38 | Sos una puta. |
| No Violento | 0.33 | 0.67 | Ese es mi perro. |
| Violento | 0.53 | 0.47 | Sos una perra. |
| No Violento | 0.39 | 0.61 | Te voy a ayudar. |
| Violento | 0.71 | 0.29 | Te voy a pegar un tiro. |
| No Violento | 0.35 | 0.65 | Te voy a dar una mano con el parcial. |
| Violento | 0.5 | 0.5 | Te voy a dejar en la calle. |
| No Violento | 0.37 | 0.63 | Voy a sacar a pasear al perro. |
| Violento | 0.65 | 0.35 | Voy a pegarte un balazo en la cabeza. |
| No Violento | 0.41 | 0.59 | Me voy para casa. |
| Violento | 0.52 | 0.48 | Voy a ir a tu casa con un fierro. |
| No Violento | 0.43 | 0.57 | Ayer perdí plata en el casino. |
| Violento | 0.66 | 0.34 | Hija de puta me hiciste perder plata. |
| No Violento | 0.33 | 0.67 | Solo necesito un poco más de plata |
| Violento | 0.53 | 0.47 | Solo me haces perder plata |
| No Violento | 0.4 | 0.6 | ¿Tenes fuego? |
| Violento | 0.72 | 0.28 | Te voy a prender fuego. |
| No Violento | 0.36 | 0.64 | Mandale un fueguito. |
| Violento | 0.6 | 0.4 | Te prendo fuego. |
| No Violento | 0.42 | 0.58 | Andá a trabajar |
| Violento | 0.83 | 0.17 | Andá a trabajar negra de mierda |
| No Violento | 0.47 | 0.53 | Que noche negra para el fútbol |
| Violento | 0.64 | 0.36 | Que negra boluda que sos. |
| No Violento | 0.26 | 0.74 | ¿Cogemos? |
| Violento | 0.72 | 0.28 | Vamos a coger quieras o no. |
| No Violento | 0.37 | 0.63 | No quiero coger hoy. |
| Violento | 0.63 | 0.37 | Que incogible que sos. |
| No Violento | 0.2 | 0.8 | ¿Rompiste el celular? |
| No Violento | 0.31 | 0.69 | Rompes todo y yo tengo que pagarlo |
| No Violento | 0.32 | 0.68 | Te rompí el plato sin querer. |
| No Violento | 0.39 | 0.61 | Si yo quiero te puedo romper la casa. |
| No Violento | 0.19 | 0.81 | Tomé mucho frio en la garganta anoche. |
| Violento | 0.8 | 0.2 | Mira este cuchillo lo voy a meter en tu garganta |
| No Violento | 0.39 | 0.61 | Ese cuchillo viene con el kit. |
| Violento | 0.75 | 0.25 | Ese cuchillo te lo puedo meter bien adentro tuyo. |

Se nota una mejora aunque no absoluta del punto 4. Como contraparte, existe una frase no violenta que ahora es considerada violenta. Pero como resultado final durante esta etapa se puede ver como el modelo es mucho más robusto, ya que los errores pasaron de 12 a 3.

nota: probando con frases usadas a modo de ejemplo en el django no dio ningún falso positivo. Pero no detecto algunas como violentas. De estas podemos afirmar que sería importante que detecte variaciones de “chupame la p\*ja” como posiblemente violentas



# Etapa 4:

# Pasarlo a minuscula mejora la deteccion del contexto o permite clasificar mejor las frases que vienen de distintas fuentes (ejemplo muchas frases de chats de Whatsapp tienen mayuscula al principio pero si son de una interfaz web no). Un problema que se encontro es que si se prueban con frases con mayusculas en cualquier lugar tiende a detectarlas como no violentas porque no las conoce.

# Violento 0.79 0.21 sos una hija de puta

# No Violento 0.46 0.54 SOS UNA HIJA DE PUTA

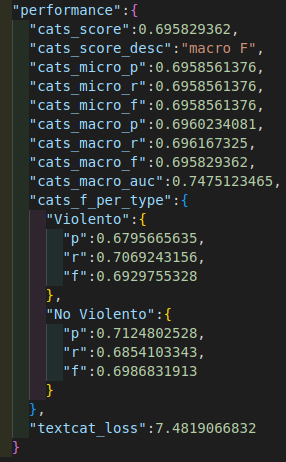
# No Violento 0.49 0.51 sOS unA hIJa dE Puta

# Para mejorar esto se tienen dos posibilidades: agregar un dataset especifico con frases violentas en mayusculas, o hacer un script agregue mayusculas random en todas las frases para que el clasificador entienda que no necesariamente modifica el contexto de la frase (o sea, que no porque la frase "TE CAISTE ANOCHE?" esta todo en mayusculas tiene que ser mas violenta que "te voy a matar").

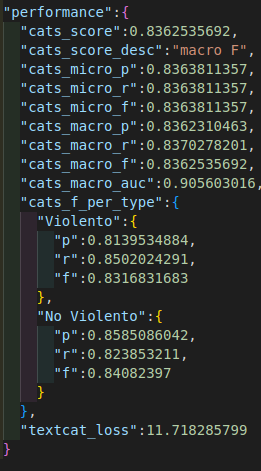
Durante esta etapa se agregaron 100 nuevas frases de violencia económica, psicológica y física. Además se agregaron frases no violentas provenientes del sitio de noticias Telám y frases sintácticamente parecidas a frases violentas.

Para poder mejorar la clasificación de las frases con mayúsculas, se creó un script para poder agregar mayúsculas de manera aleatoria a algunas letras de las frases.

Posteriormente se hizo un entrenamiento donde los score fueron de 69-70.

****

Por el contrario el entrenamiento solo con minúsculas dio un score de 83



Es decir el score bajo 14 puntos con respecto al dataset solo con minúsculas.

Además, las pruebas demostraron un considerable peor desempeño de los casos ambiguos.

(Para las siguientes pruebas se considera una frase violenta a un score de ‘Violento’ mayor o igual a 0,7)

En el caso del dataset con mayusculas random hay errores de todos los tipos, casi la mitad de las frases fueron mal clasificadas:

| Clasificacion | Score Violento | Score No Violento | Frase |
| --- | --- | --- | --- |
| No Violento | 0.34 | 0.66 | Hola, soy Pablo. |
| Violento | 0.84 | 0.16 | Te voy a matar. |
| No Violento | 0.61 | 0.39 | Hola, te voy a buscar. |
| No Violento | 0.68 | 0.32 | Hola, si te metés conmigo te voy a buscar. |
| No Violento | 0.27 | 0.73 | Vivo en Mar del Plata. |
| Violento | 0.71 | 0.29 | Te voy a violar |
| No Violento | 0.51 | 0.49 | Se donde vivís. |
| No Violento | 0.4 | 0.6 | Guarda conmigo, porque yo se dónde vivís. |
| No Violento | 0.47 | 0.53 | Sos una genia |
| No Violento | 0.6 | 0.4 | Sos una puta. |
| No Violento | 0.58 | 0.42 | Ese es mi perro. |
| No Violento | 0.5 | 0.5 | Sos una perra. |
| Violento | 0.71 | 0.29 | Te voy a ayudar. |
| Violento | 0.7 | 0.3 | Te voy a pegar un tiro. |
| No Violento | 0.49 | 0.51 | Te voy a dar una mano con el parcial. |
| Violento | 0.71 | 0.29 | Te voy a dejar en la calle. |
| Violento | 0.75 | 0.25 | Voy a sacar a pasear al perro. |
| No Violento | 0.38 | 0.62 | Voy a pegarte un balazo en la cabeza. |
| No Violento | 0.51 | 0.49 | Me voy para casa. |
| Violento | 0.72 | 0.28 | Voy a ir a tu casa con un fierro. |
| No Violento | 0.34 | 0.66 | Ayer perdí plata en el casino. |
| No Violento | 0.62 | 0.38 | Hija de puta me hiciste perder plata. |
| No Violento | 0.41 | 0.59 | Solo necesito un poco más de plata |
| No Violento | 0.56 | 0.44 | Solo me haces perder plata |
| No Violento | 0.36 | 0.64 | ¿Tenes fuego? |
| Violento | 0.71 | 0.29 | Te voy a prender fuego. |
| No Violento | 0.47 | 0.53 | Mandale un fueguito. |
| No Violento | 0.59 | 0.41 | Te prendo fuego. |
| No Violento | 0.5 | 0.5 | Andá a trabajar |
| Violento | 0.77 | 0.23 | Andá a trabajar negra de mierda |
| No Violento | 0.43 | 0.57 | Que noche negra para el fútbol |
| No Violento | 0.64 | 0.36 | Que negra boluda que sos. |
| No Violento | 0.28 | 0.72 | ¿Cogemos? |
| No Violento | 0.63 | 0.37 | Vamos a coger quieras o no. |
| No Violento | 0.52 | 0.48 | No quiero coger hoy. |
| No Violento | 0.54 | 0.46 | Que incogible que sos. |
| No Violento | 0.18 | 0.82 | ¿Rompiste el celular? |
| No Violento | 0.36 | 0.64 | Rompes todo y yo tengo que pagarlo |
| No Violento | 0.47 | 0.53 | Te rompí el plato sin querer. |
| Violento | 0.78 | 0.22 | Si yo quiero te puedo romper la casa. |
| No Violento | 0.44 | 0.56 | Tomé mucho frio en la garganta anoche. |
| Violento | 0.83 | 0.17 | Mira este cuchillo lo voy a meter en tu garganta |
| No Violento | 0.34 | 0.66 | Ese cuchillo viene con el kit. |
| Violento | 0.73 | 0.27 | Ese cuchillo te lo puedo meter bien adentro tuyo. |

Mientras que para el dataset solo con minusculas solo hay un error:

| Clasificacion | Score Violento | Score No Violento | Frase |
| --- | --- | --- | --- |
| No Violento | 0.44 | 0.56 | hola, soy pablo. |
| Violento | 0.88 | 0.12 | te voy a matar. |
| No Violento | 0.68 | 0.32 | hola, te voy a buscar. |
| Violento | 0.91 | 0.09 | hola, si te metés conmigo te voy a buscar. |
| No Violento | 0.37 | 0.63 | vivo en mar del plata. |
| Violento | 0.83 | 0.17 | te voy a violar |
| No Violento | 0.34 | 0.66 | se donde vivís. |
| Violento | 0.84 | 0.16 | guarda conmigo, porque yo se dónde vivís. |
| No Violento | 0.47 | 0.53 | sos una genia |
| Violento | 0.83 | 0.17 | sos una puta. |
| No Violento | 0.4 | 0.6 | ese es mi perro. |
| Violento | 0.77 | 0.23 | sos una perra. |
| No Violento | 0.55 | 0.45 | te voy a ayudar. |
| Violento | 0.89 | 0.11 | te voy a pegar un tiro. |
| No Violento | 0.44 | 0.56 | te voy a dar una mano con el parcial. |
| Violento | 0.8 | 0.2 | te voy a dejar en la calle. |
| No Violento | 0.58 | 0.42 | voy a sacar a pasear al perro. |
| Violento | 0.79 | 0.21 | voy a pegarte un balazo en la cabeza. |
| No Violento | 0.52 | 0.48 | me voy para casa. |
| Violento | 0.79 | 0.21 | voy a ir a tu casa con un fierro. |
| No Violento | 0.45 | 0.55 | ayer perdí plata en el casino. |
| Violento | 0.85 | 0.15 | hija de puta me hiciste perder plata. |
| No Violento | 0.4 | 0.6 | solo necesito un poco más de plata |
| Violento | 0.8 | 0.2 | solo me haces perder plata |
| No Violento | 0.45 | 0.55 | ¿tenes fuego? |
| Violento | 0.87 | 0.13 | te voy a prender fuego. |
| No Violento | 0.44 | 0.56 | mandale un fueguito. |
| Violento | 0.73 | 0.27 | te prendo fuego. |
| No Violento | 0.32 | 0.68 | andá a trabajar |
| Violento | 0.81 | 0.19 | andá a trabajar negra de mierda |
| No Violento | 0.48 | 0.52 | que noche negra para el fútbol |
| Violento | 0.83 | 0.17 | que negra boluda que sos. |
| No Violento | 0.3 | 0.7 | ¿cogemos? |
| Violento | 0.8 | 0.2 | vamos a coger quieras o no. |
| No Violento | 0.37 | 0.63 | no quiero coger hoy. |
| Violento | 0.74 | 0.26 | que incogible que sos. |
| No Violento | 0.18 | 0.82 | ¿rompiste el celular? |
| No Violento | 0.39 | 0.61 | rompes todo y yo tengo que pagarlo |
| No Violento | 0.7 | 0.3 | te rompí el plato sin querer. |
| Violento | 0.72 | 0.28 | si yo quiero te puedo romper la casa. |
| No Violento | 0.2 | 0.8 | tomé mucho frio en la garganta anoche. |
| Violento | 0.86 | 0.14 | mira este cuchillo lo voy a meter en tu garganta |
| No Violento | 0.39 | 0.61 | ese cuchillo viene con el kit. |
| Violento | 0.93 | 0.07 | ese cuchillo te lo puedo meter bien adentro tuyo. |

Las principales conclusiones de esto son:

* Las mayúsculas añaden una dimensión adicional al entrenamiento del modelo, porque se consideran frases distintas aunque solo haya una letra mayuscula en toda la frase.
* Las frases poseen mayusculas en lugares distintos de acuerdo a la fuente de donde vienen (celuar, interfaz web).
* La presencia de mayusculas no tiene correlación directa con la violencia de la frase. Por ejemplo, la frase “FELIZ CUMPLE” no tiene violencia comparada con “te voy a matar”.

Los mejores resultados se encontraron cuando no solo las frases del dataset estaban todas en minsuculas y cuando las frases de prueba también se pasaban a minúscula.

### Evaluación final de modelos entrenamientos

Para poder evaluar de manera uniforme y completa el rendimiento de todos los modelos se armó un dataset de prueba extendido, compuesto por 200 frases.

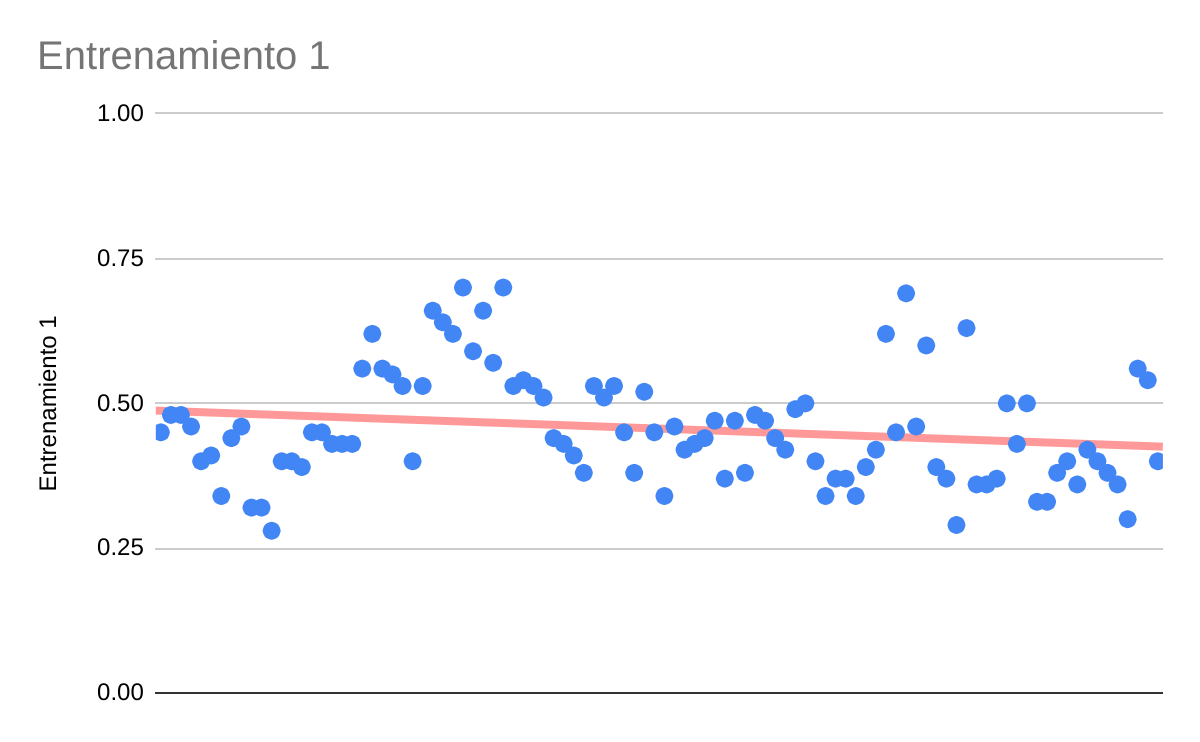
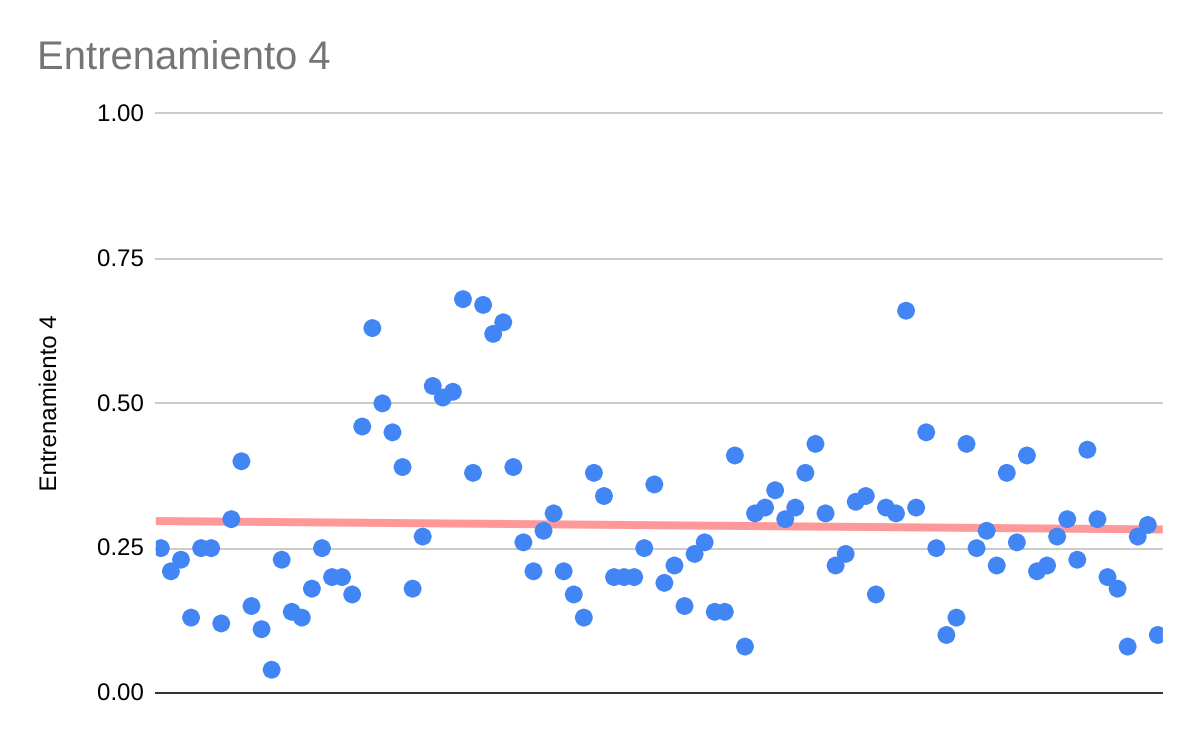
Para que una frase sea considerada violenta, el score mínimo es de 70.

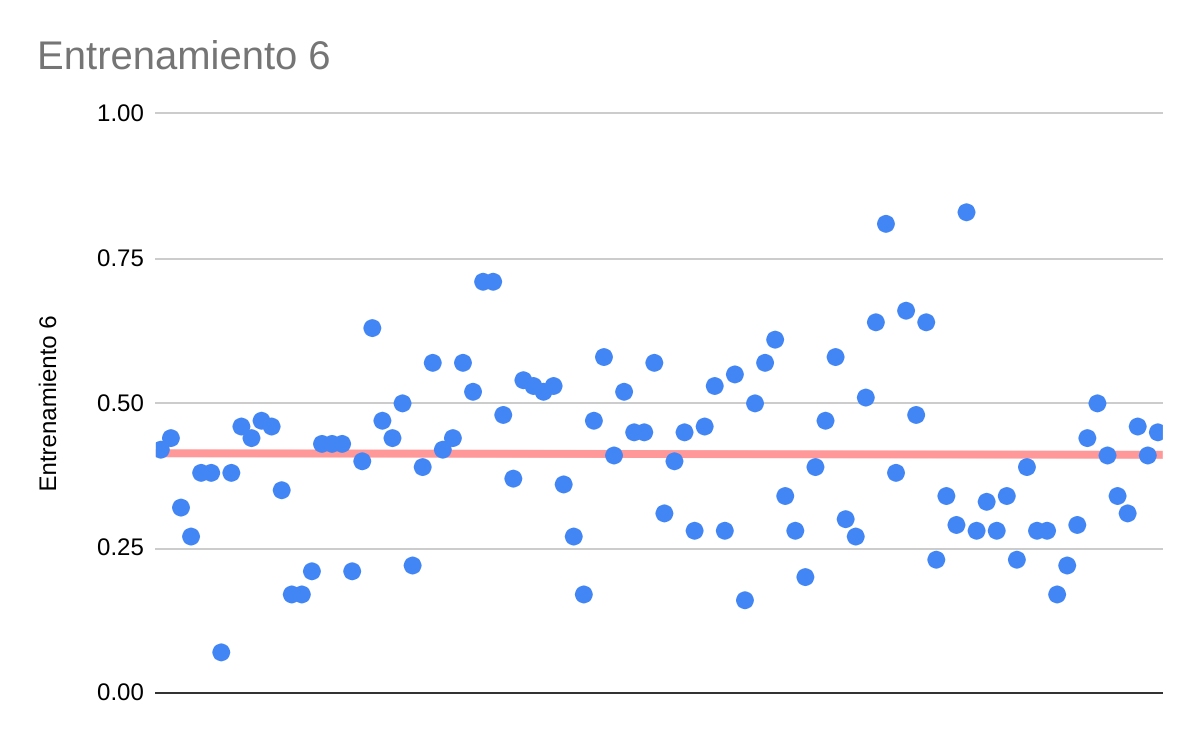
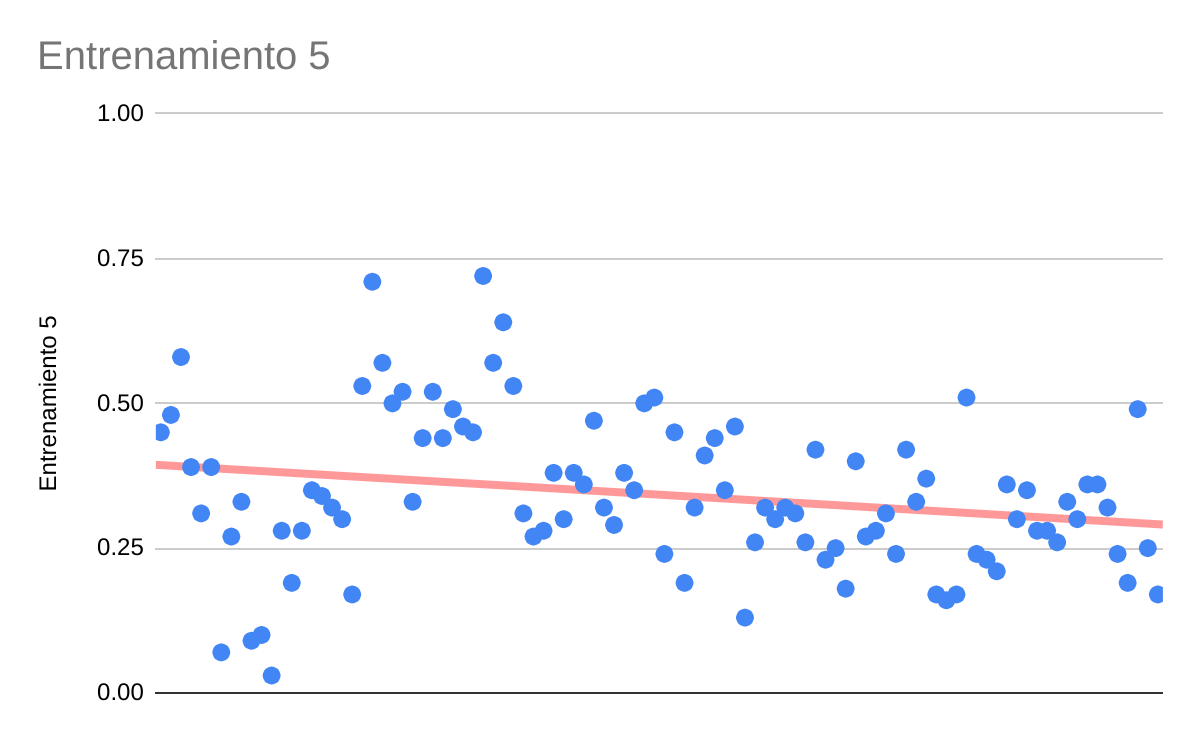
Los entrenamientos 2 y 3 se realizaron con transformers, de ahí la diferencia en el score.

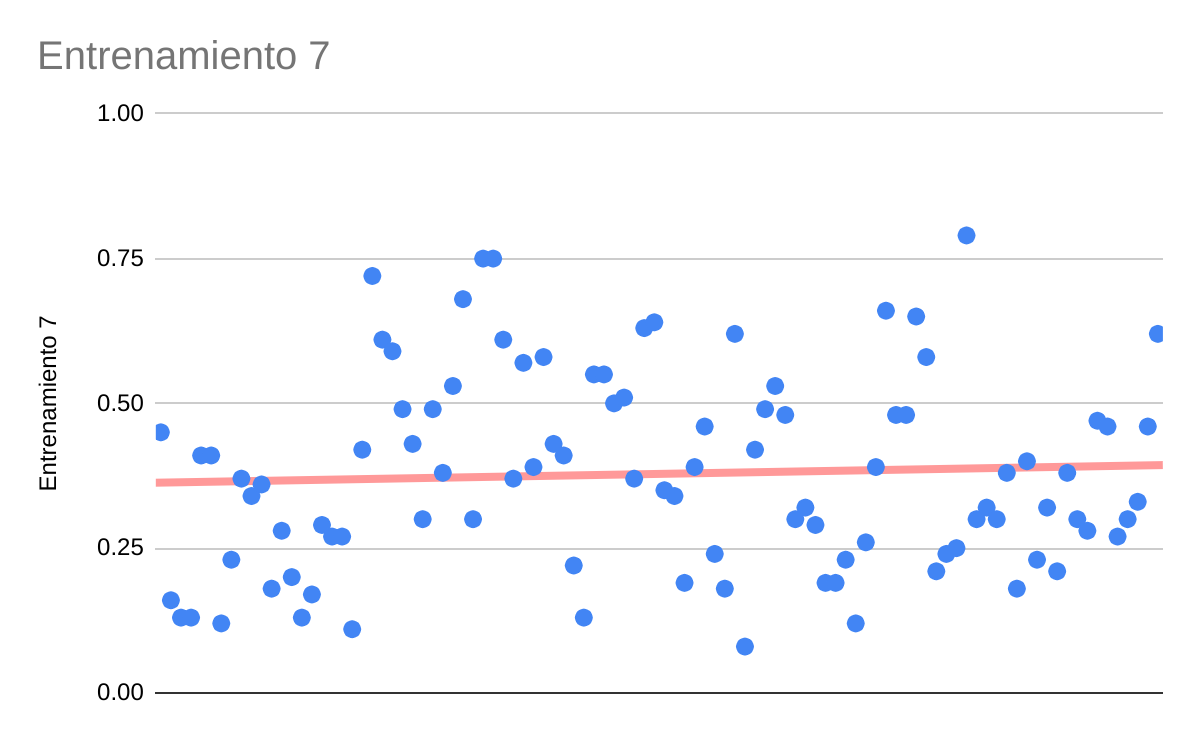
| Frase | Entrenamiento 1 | Entrenamiento 2 | Entrenamiento 3 | Entrenamiento 4 | Entrenamiento 5 | Entrenamiento 6 | Entrenamiento 7 |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Hola | 0.45 | 0 | 0 | 0.25 | 0.45 | 0.42 | 0.45 |
| Hola, soy Pablo. | 0.48 | 0 | 0 | 0.21 | 0.48 | 0.44 | 0.16 |
| Hola, mi nombre es Valentina. | 0.48 | 0 | 0 | 0.23 | 0.58 | 0.32 | 0.13 |
| Hola, como estas, mi nombre es Lautaro | 0.46 | 0 | 0 | 0.13 | 0.39 | 0.27 | 0.13 |
| Me levante | 0.4 | 0 | 0 | 0.25 | 0.31 | 0.38 | 0.41 |
| Me levante a las 8 | 0.41 | 0 | 0 | 0.25 | 0.39 | 0.38 | 0.41 |
| Me levante a las 8 de la mañana para llegar temprano al trabajo | 0.34 | 0 | 0 | 0.12 | 0.07 | 0.07 | 0.12 |
| Vivo | 0.44 | 0 | 0 | 0.3 | 0.27 | 0.38 | 0.23 |
| Vivo aca | 0.46 | 0.03 | 0 | 0.4 | 0.33 | 0.46 | 0.37 |
| Vivo aca en Mar del Plata | 0.32 | 0 | 0 | 0.15 | 0.09 | 0.44 | 0.34 |
| Vivo aca en Mar del Plata con mi familia | 0.32 | 0 | 0 | 0.11 | 0.1 | 0.47 | 0.36 |
| Vivo aca en Mar del Plata con mi familia en una casa | 0.28 | 0 | 0 | 0.04 | 0.03 | 0.46 | 0.18 |
| Tengo | 0.4 | 0 | 0 | 0.23 | 0.28 | 0.35 | 0.28 |
| Tengo dos hermanos | 0.4 | 0 | 0 | 0.14 | 0.19 | 0.17 | 0.2 |
| Tengo dos hermanos y un perro | 0.39 | 0 | 0 | 0.13 | 0.28 | 0.17 | 0.13 |
| Tengo dos hermanos y un perro negro | 0.45 | 0 | 0 | 0.18 | 0.35 | 0.21 | 0.17 |
| Esa es | 0.45 | 0 | 0 | 0.25 | 0.34 | 0.43 | 0.29 |
| Esa es mi gata | 0.43 | 0 | 0 | 0.2 | 0.32 | 0.43 | 0.27 |
| Esa es mi gata llamada | 0.43 | 0 | 0 | 0.2 | 0.3 | 0.43 | 0.27 |
| Esa es mi gata llamada Andromeda y tiene 10 años | 0.43 | 0 | 0 | 0.17 | 0.17 | 0.21 | 0.11 |
| Te busco? | 0.56 | 0.57 | 0 | 0.46 | 0.53 | 0.4 | 0.42 |
| Te voy a buscar? | 0.62 | 1 | 0 | 0.63 | 0.71 | 0.63 | 0.72 |
| Te voy a buscar con el auto? | 0.56 | 0.71 | 1 | 0.5 | 0.57 | 0.47 | 0.61 |
| Te voy a pasar a buscar con el auto? | 0.55 | 0.73 | 0.47 | 0.45 | 0.5 | 0.44 | 0.59 |
| Te ayudo? | 0.53 | 0 | 0 | 0.39 | 0.52 | 0.5 | 0.49 |
| Queres ayuda? | 0.4 | 0 | 0 | 0.18 | 0.33 | 0.22 | 0.43 |
| Te doy una mano? | 0.53 | 0 | 0 | 0.27 | 0.44 | 0.39 | 0.3 |
| Te voy a dar una mano | 0.66 | 0.1 | 0 | 0.53 | 0.52 | 0.57 | 0.49 |
| Te voy a ayudar con el parcial | 0.64 | 0 | 0 | 0.51 | 0.44 | 0.42 | 0.38 |
| Te voy a comprar la entrada | 0.62 | 1 | 0 | 0.52 | 0.49 | 0.44 | 0.53 |
| Te voy a llevar al recital | 0.7 | 0 | 0 | 0.68 | 0.46 | 0.57 | 0.68 |
| Te voy a sacar un turno con el medico | 0.59 | 0.19 | 0 | 0.38 | 0.45 | 0.52 | 0.3 |
| Te voy a liberar un cupo | 0.66 | 0.65 | 0 | 0.67 | 0.72 | 0.71 | 0.75 |
| Te voy a necesitar para un partido | 0.57 | 0.12 | 0 | 0.62 | 0.57 | 0.71 | 0.75 |
| Te voy a agregar al grupo | 0.7 | 0 | 0 | 0.64 | 0.64 | 0.48 | 0.61 |
| Te llamo? | 0.53 | 0 | 0 | 0.39 | 0.53 | 0.37 | 0.37 |
| Te puedo llamar? | 0.54 | 0 | 0 | 0.26 | 0.31 | 0.54 | 0.57 |
| Te puedo hacer una llamada? | 0.53 | 0 | 0 | 0.21 | 0.27 | 0.53 | 0.39 |
| Te puedo llamar para coordinar? | 0.51 | 0 | 0 | 0.28 | 0.28 | 0.52 | 0.58 |
| Hablamos | 0.44 | 0 | 0 | 0.31 | 0.38 | 0.53 | 0.43 |
| Hablamos de mucho | 0.43 | 0 | 0 | 0.21 | 0.3 | 0.36 | 0.41 |
| Hablamos de muchas cosas | 0.41 | 0 | 0 | 0.17 | 0.38 | 0.27 | 0.22 |
| Hablamos de muchas cosas en la reunion | 0.38 | 0 | 0 | 0.13 | 0.36 | 0.17 | 0.13 |
| Sos una genia | 0.53 | 1 | 0 | 0.38 | 0.47 | 0.47 | 0.55 |
| Sos una santa | 0.51 | 1 | 0 | 0.34 | 0.32 | 0.58 | 0.55 |
| Sos una genia amiga | 0.53 | 0 | 0 | 0.2 | 0.29 | 0.41 | 0.5 |
| Sos mi mejor amiga | 0.45 | 1 | 0 | 0.2 | 0.38 | 0.52 | 0.51 |
| Sos la mejor de mis amigas | 0.38 | 0.02 | 0 | 0.2 | 0.35 | 0.45 | 0.37 |
| Sos la hija de mi mejor amiga | 0.52 | 1 | 0 | 0.25 | 0.5 | 0.45 | 0.63 |
| Soy el hijo de Rodriguez | 0.45 | 0 | 0.02 | 0.36 | 0.51 | 0.57 | 0.64 |
| Soy el amigo del hijo de Rodriguez | 0.34 | 0 | 0 | 0.19 | 0.24 | 0.31 | 0.35 |
| Ese es mi perro | 0.46 | 0 | 0 | 0.22 | 0.45 | 0.4 | 0.34 |
| Esa es mi casa | 0.42 | 1 | 0 | 0.15 | 0.19 | 0.45 | 0.19 |
| Ese es mi nuevo celular | 0.43 | 0 | 0 | 0.24 | 0.32 | 0.28 | 0.39 |
| Ese es tu antiguo auto | 0.44 | 0.05 | 0 | 0.26 | 0.41 | 0.46 | 0.46 |
| Esa es mi vieja escuela | 0.47 | 0 | 0 | 0.14 | 0.44 | 0.53 | 0.24 |
| Ese es el camino a tomar | 0.37 | 0 | 0 | 0.14 | 0.35 | 0.28 | 0.18 |
| Eso es lo que le dije ayer | 0.47 | 0 | 0 | 0.41 | 0.46 | 0.55 | 0.62 |
| Ese es el ultimo turno que quedaba | 0.38 | 0 | 0 | 0.08 | 0.13 | 0.16 | 0.08 |
| Me voy de casa | 0.48 | 1 | 0 | 0.31 | 0.26 | 0.5 | 0.42 |
| Me voy a mi casa | 0.47 | 1 | 0 | 0.32 | 0.32 | 0.57 | 0.49 |
| Me voy para tu casa | 0.44 | 1 | 0 | 0.35 | 0.3 | 0.61 | 0.53 |
| Me voy a comprar un auto | 0.42 | 0 | 0 | 0.3 | 0.32 | 0.34 | 0.48 |
| Me voy de viaje | 0.49 | 0 | 0 | 0.32 | 0.31 | 0.28 | 0.3 |
| Me voy de viaje a la capital | 0.5 | 0 | 0 | 0.38 | 0.26 | 0.2 | 0.32 |
| Vamos a jugar | 0.4 | 0 | 0 | 0.43 | 0.42 | 0.39 | 0.29 |
| Vamos a ver un recital | 0.34 | 0 | 0 | 0.31 | 0.23 | 0.47 | 0.19 |
| Vamos a cocinar un rico asado | 0.37 | 0 | 0 | 0.22 | 0.25 | 0.58 | 0.19 |
| Vamos a comprar una ferrari | 0.37 | 0 | 0 | 0.24 | 0.18 | 0.3 | 0.23 |
| Vamos a jugar un partido de futbol | 0.34 | 0 | 0 | 0.33 | 0.4 | 0.27 | 0.12 |
| Vamos a limpiar la puerta de madera | 0.39 | 0 | 0 | 0.34 | 0.27 | 0.51 | 0.26 |
| Lleva abrigo o vas a tomar frio | 0.42 | 0.37 | 0 | 0.17 | 0.28 | 0.64 | 0.39 |
| Recorda ponerte protector solar o te vas a quemar la piel | 0.62 | 0 | 0 | 0.32 | 0.31 | 0.81 | 0.66 |
| Avisa o se va a olvidar | 0.45 | 1 | 0 | 0.31 | 0.24 | 0.38 | 0.48 |
| Acordate o te van a ganar de mano | 0.69 | 0 | 0 | 0.66 | 0.42 | 0.66 | 0.48 |
| Si no le avisas se va a olvidar | 0.46 | 1 | 0.59 | 0.32 | 0.33 | 0.48 | 0.65 |
| Si no te abrigas vas a pasar frio | 0.6 | 1 | 0 | 0.45 | 0.37 | 0.64 | 0.58 |
| Si no me encontras en la fiesta mandame un mensaje | 0.39 | 0 | 0 | 0.25 | 0.17 | 0.23 | 0.21 |
| Cuidado con la lluvia | 0.37 | 0 | 0 | 0.1 | 0.16 | 0.34 | 0.24 |
| Guarda con el sol del mediodia | 0.29 | 0 | 0 | 0.13 | 0.17 | 0.29 | 0.25 |
| Lleva paraguas o te vas a mojar | 0.63 | 1 | 1 | 0.43 | 0.51 | 0.83 | 0.79 |
| Me duele la pierna | 0.36 | 0 | 0 | 0.25 | 0.24 | 0.28 | 0.3 |
| Me duele la garganta | 0.36 | 0 | 0 | 0.28 | 0.23 | 0.33 | 0.32 |
| Me duele la cabeza | 0.37 | 0 | 0 | 0.22 | 0.21 | 0.28 | 0.3 |
| Te duele la pierna? | 0.5 | 0 | 0 | 0.38 | 0.36 | 0.34 | 0.38 |
| Te duele el brazo? | 0.43 | 0 | 0 | 0.26 | 0.3 | 0.23 | 0.18 |
| Te duele la garganta? | 0.5 | 0 | 0 | 0.41 | 0.35 | 0.39 | 0.4 |
| Tome mucho frio en la garganta anoche | 0.33 | 0 | 0 | 0.21 | 0.28 | 0.28 | 0.23 |
| Tome mucha cerveza ayer | 0.33 | 0 | 0 | 0.22 | 0.28 | 0.28 | 0.32 |
| Se me rompio el celular | 0.38 | 0 | 1 | 0.27 | 0.26 | 0.17 | 0.21 |
| Se me cayo el celular | 0.4 | 0 | 0.5 | 0.3 | 0.33 | 0.22 | 0.38 |
| Se me torcio el tobillo | 0.36 | 0 | 0 | 0.23 | 0.3 | 0.29 | 0.3 |
| Se fueron a nadar | 0.42 | 0 | 0 | 0.42 | 0.36 | 0.44 | 0.28 |
| Se enfermo mi tortuga | 0.4 | 0.01 | 0 | 0.3 | 0.36 | 0.5 | 0.47 |
| Subio el pez por la ventana | 0.38 | 0 | 0 | 0.2 | 0.32 | 0.41 | 0.46 |
| Volo la madera por el vidrio | 0.36 | 0 | 0 | 0.18 | 0.24 | 0.34 | 0.27 |
| Aunque los niños tuvieron dos abrigos no limpiaron el recital | 0.3 | 0 | 0 | 0.08 | 0.19 | 0.31 | 0.3 |
| Te va a caminar la plaza cuando caiga el sol | 0.56 | 0 | 0 | 0.27 | 0.49 | 0.46 | 0.33 |
| Espero que te vaya muy bien | 0.54 | 0 | 0 | 0.29 | 0.25 | 0.41 | 0.46 |
| Deseo que tengas muchos exitos en tu nuevo trabajo | 0.4 | 0 | 0 | 0.1 | 0.17 | 0.45 | 0.62 |
| Hija de puta | 0.57 | 1 | 1 | 0.77 | 0.85 | 0.71 | 0.85 |
| Negra de mierda | 0.81 | 1 | 1 | 0.91 | 0.9 | 0.88 | 0.92 |
| Zorra malcogida | 0.41 | 1 | 1 | 0.29 | 0.46 | 0.58 | 0.78 |
| La puta que te pario | 0.65 | 1 | 1 | 0.83 | 0.91 | 0.91 | 0.95 |
| Mogolica | 0.44 | 1 | 1 | 0.33 | 0.34 | 0.56 | 0.64 |
| Pedazo de mierda | 0.71 | 1 | 1 | 0.85 | 0.87 | 0.83 | 0.87 |
| Pedazo de retardada | 0.48 | 1 | 1 | 0.46 | 0.61 | 0.6 | 0.61 |
| La concha de tu madre | 0.52 | 1 | 1 | 0.62 | 0.66 | 0.66 | 0.75 |
| Perra de mierda | 0.69 | 1 | 1 | 0.76 | 0.87 | 0.85 | 0.89 |
| Putita | 0.42 | 1 | 1 | 0.29 | 0.4 | 0.6 | 0.66 |
| Te voy a matar | 0.81 | 1 | 1 | 0.91 | 0.9 | 0.88 | 0.91 |
| Te voy a asesinar | 0.71 | 1 | 1 | 0.79 | 0.78 | 0.78 | 0.79 |
| Te voy a pegar un tiro | 0.65 | 1 | 1 | 0.77 | 0.91 | 0.89 | 0.95 |
| Te voy a degollar | 0.68 | 1 | 1 | 0.69 | 0.69 | 0.71 | 0.78 |
| Te voy a fusilar | 0.68 | 1 | 1 | 0.69 | 0.66 | 0.71 | 0.76 |
| Te voy a asfixiar | 0.68 | 1 | 1 | 0.69 | 0.66 | 0.71 | 0.76 |
| Te voy a ahorcar | 0.68 | 1 | 1 | 0.69 | 0.68 | 0.71 | 0.76 |
| Te voy a violar | 0.68 | 1 | 1 | 0.69 | 0.81 | 0.83 | 0.87 |
| Te voy a coger todo lo que yo quiera | 0.74 | 1 | 1 | 0.81 | 0.77 | 0.87 | 0.96 |
| Te voy a quemar viva | 0.67 | 1 | 1 | 0.68 | 0.71 | 0.82 | 0.9 |
| Te voy a prender fuego | 0.68 | 1 | 1 | 0.73 | 0.72 | 0.87 | 0.94 |
| Voy a pegarte | 0.53 | 1 | 1 | 0.51 | 0.53 | 0.68 | 0.79 |
| Voy a matarte | 0.57 | 1 | 1 | 0.61 | 0.76 | 0.68 | 0.84 |
| Voy a pegarte un tiro | 0.5 | 1 | 1 | 0.59 | 0.67 | 0.77 | 0.87 |
| Voy a matarte de un tiro | 0.56 | 1 | 1 | 0.66 | 0.84 | 0.73 | 0.9 |
| Voy a matarte con mi pistola | 0.56 | 1 | 1 | 0.49 | 0.82 | 0.72 | 0.89 |
| Voy a reventarte la cabeza | 0.55 | 1 | 1 | 0.5 | 0.54 | 0.6 | 0.78 |
| Voy a revertarte de un tiro | 0.52 | 1 | 1 | 0.56 | 0.65 | 0.65 | 0.8 |
| Voy a prenderte fuego | 0.53 | 1 | 0 | 0.51 | 0.55 | 0.72 | 0.79 |
| Voy a quemarte viva | 0.52 | 1 | 1 | 0.47 | 0.48 | 0.64 | 0.82 |
| Voy a romperte toda la casa | 0.57 | 1 | 1 | 0.54 | 0.32 | 0.61 | 0.6 |
| Voy a matar a toda tu familia | 0.78 | 1 | 1 | 0.91 | 0.9 | 0.91 | 0.9 |
| Voy a romperte el culo | 0.56 | 1 | 1 | 0.48 | 0.82 | 0.81 | 0.94 |
| Te voy a dejar en la calle | 0.69 | 1 | 1 | 0.63 | 0.63 | 0.8 | 0.86 |
| Te voy a dejar sin un peso | 0.66 | 1 | 0 | 0.72 | 0.64 | 0.81 | 0.85 |
| Te voy a echar de la casa | 0.67 | 1 | 1 | 0.56 | 0.43 | 0.76 | 0.71 |
| Te voy a sacar al nene | 0.7 | 1 | 0.27 | 0.74 | 0.7 | 0.83 | 0.79 |
| Te voy a sacar todo lo que tenes | 0.71 | 1 | 0 | 0.65 | 0.58 | 0.79 | 0.71 |
| Te voy a hacer sufir | 0.65 | 1 | 0 | 0.63 | 0.6 | 0.73 | 0.75 |
| Te voy a cagar a palos | 0.74 | 1 | 1 | 0.76 | 0.77 | 0.88 | 0.95 |
| Te voy a degollar con este cuchillo | 0.65 | 1 | 1 | 0.74 | 0.8 | 0.76 | 0.87 |
| No vas a vivir tranquila | 0.47 | 1 | 0 | 0.41 | 0.52 | 0.67 | 0.72 |
| No vas a ser feliz nunca mas | 0.37 | 0.02 | 0 | 0.12 | 0.24 | 0.44 | 0.4 |
| Si te metes conmigo la vas a pasar mal | 0.64 | 1 | 1 | 0.34 | 0.42 | 0.91 | 0.94 |
| Si te metes conmigo te voy a hacer mucho daño | 0.78 | 1 | 1 | 0.68 | 0.68 | 0.94 | 0.97 |
| Si te metes conmigo me las vas a pagar | 0.56 | 1 | 1 | 0.35 | 0.54 | 0.93 | 0.97 |
| Guarda conmigo se donde vivis | 0.37 | 1 | 0 | 0.23 | 0.21 | 0.79 | 0.84 |
| Guarda conmigo te puedo hacer cagar | 0.61 | 1 | 1 | 0.43 | 0.36 | 0.93 | 0.86 |
| Tene cuidado conmigo porque no sabes lo que te espera | 0.63 | 1 | 0 | 0.18 | 0.28 | 0.96 | 0.92 |
| Cuidadito conmigo que yo me puedo poner violento | 0.45 | 0 | 0 | 0.1 | 0.18 | 0.66 | 0.31 |
| Vas a morir cuando yo salga | 0.57 | 1 | 0.91 | 0.27 | 0.18 | 0.73 | 0.89 |
| Vas a saber que es el dolor cuando yo salga | 0.56 | 1 | 0 | 0.12 | 0.15 | 0.66 | 0.7 |
| Vas a sufrir cuando yo salga | 0.6 | 1 | 0 | 0.44 | 0.49 | 0.91 | 0.97 |
| Sos una estupida | 0.53 | 1 | 1 | 0.38 | 0.49 | 0.72 | 0.74 |
| No sabes hacer nada bien | 0.41 | 0 | 0 | 0.23 | 0.25 | 0.55 | 0.69 |
| A vos no te sale hacer nada | 0.65 | 1 | 0 | 0.56 | 0.59 | 0.7 | 0.75 |
| Sos una buena para nada | 0.47 | 1 | 0 | 0.25 | 0.2 | 0.37 | 0.37 |
| Siempre arruinas todo | 0.49 | 1 | 0 | 0.34 | 0.51 | 0.41 | 0.55 |
| Yo te puedo coger cuando quiera | 0.68 | 1 | 0 | 0.48 | 0.56 | 0.85 | 0.85 |
| Yo te puedo violar cuando quiera | 0.66 | 1 | 1 | 0.41 | 0.56 | 0.84 | 0.77 |
| Yo te puedo tocar todo lo que quiera | 0.66 | 1 | 0 | 0.5 | 0.41 | 0.77 | 0.73 |
| Aca mando yo y vos tenes que obedecer | 0.54 | 1 | 0 | 0.47 | 0.51 | 0.66 | 0.92 |
| Yo pago por todo y vos solo gastas | 0.5 | 0.01 | 0 | 0.27 | 0.26 | 0.6 | 0.77 |
| Voy a ir a tu casa con un fierro | 0.4 | 1 | 0.87 | 0.33 | 0.41 | 0.79 | 0.46 |
| Voy a ir a tu casa con un bidon de nafta | 0.42 | 0.12 | 0 | 0.26 | 0.38 | 0.65 | 0.51 |
| Voy a ir a tu casa y te tiro la puerta abajo | 0.58 | 1 | 1 | 0.72 | 0.68 | 0.95 | 0.85 |
| Dame al nene o te prendo fuego | 0.65 | 1 | 1 | 0.68 | 0.55 | 0.8 | 0.89 |
| Dame a la nena o te degollo | 0.69 | 1 | 1 | 0.67 | 0.67 | 0.74 | 0.75 |
| Si yo quiero te rompo todo | 0.63 | 1 | 0.66 | 0.4 | 0.22 | 0.62 | 0.75 |
| Si yo quiero te mato ahora mismo | 0.76 | 1 | 1 | 0.79 | 0.89 | 0.71 | 0.74 |
| Si yo quiero te meto este cuchillo en la panza | 0.59 | 1 | 1 | 0.34 | 0.58 | 0.77 | 0.87 |
| Que incogible que sos | 0.48 | 0 | 0 | 0.36 | 0.43 | 0.74 | 0.8 |
| Vamos a coger quieras o no | 0.43 | 1 | 0.88 | 0.43 | 0.39 | 0.8 | 0.82 |
| Espero que te mueras | 0.6 | 1 | 1 | 0.49 | 0.4 | 0.67 | 0.64 |
| Espero que sufras | 0.42 | 0 | 0 | 0.29 | 0.25 | 0.42 | 0.35 |
| Espero que mueras dolorosamente | 0.45 | 0.01 | 0 | 0.31 | 0.28 | 0.54 | 0.51 |
| Te deseo lo peor para tu vida | 0.58 | 0.04 | 0 | 0.7 | 0.69 | 0.72 | 0.88 |
| No mereces nada en esta vida | 0.32 | 0 | 0 | 0.21 | 0.21 | 0.36 | 0.66 |
| Vos sos una basura humana porque me cagaste | 0.6 | 1 | 1 | 0.48 | 0.62 | 0.93 | 0.88 |
| Si te veo con otro te mato | 0.82 | 1 | 1 | 0.93 | 0.96 | 0.87 | 0.85 |
| Si te veo con un macho te voy a hacer cagar | 0.79 | 1 | 1 | 0.89 | 0.88 | 0.95 | 0.9 |
| Vos me estas cagando con otro ya vas a ver | 0.37 | 1 | 1 | 0.24 | 0.56 | 0.69 | 0.89 |
| Vos te estas cogiendo a otro te voy a castigar | 0.81 | 1 | 1 | 0.85 | 0.88 | 0.92 | 0.97 |
| Me engañas con otro y te mato | 0.7 | 1 | 1 | 0.84 | 0.93 | 0.72 | 0.74 |
| Te rompo las piernas y te tiro en una zanja | 0.68 | 1 | 1 | 0.74 | 0.74 | 0.88 | 0.96 |
| Te violo y despues te mato | 0.88 | 1 | 1 | 0.95 | 0.96 | 0.88 | 0.86 |
| Te mato y despues me mato yo | 0.91 | 1 | 1 | 0.98 | 0.99 | 0.89 | 0.89 |
| Voy a ir a la carcel pero no importa te mato antes | 0.75 | 1 | 1 | 0.89 | 0.95 | 0.92 | 0.91 |
| Te hago desaparecer y tu familia nunca te va a encontrar | 0.74 | 1 | 0 | 0.59 | 0.48 | 0.92 | 0.73 |
| Guarda por donde caminas porque me podes encontrar | 0.45 | 0.88 | 0 | 0.28 | 0.17 | 0.53 | 0.51 |
| Voy a pasarte este cuchillo por la garganta | 0.52 | 1 | 1 | 0.62 | 0.68 | 0.72 | 0.88 |
| Vas a saber lo que es el dolor | 0.49 | 1 | 0 | 0.14 | 0.25 | 0.54 | 0.53 |
| Chupame la pija | 0.46 | 1 | 1 | 0.46 | 0.72 | 0.82 | 0.9 |
| O cogemos o te moris | 0.68 | 1 | 1 | 0.59 | 0.4 | 0.73 | 0.8 |
| O me chupas la pija o te violo | 0.67 | 1 | 1 | 0.53 | 0.81 | 0.91 | 0.96 |
| Te voy a romper la concha de un pijazo | 0.76 | 1 | 1 | 0.86 | 0.91 | 0.88 | 0.96 |
| A vos lo que te hace falta es una buena cogida | 0.59 | 1 | 0 | 0.2 | 0.39 | 0.65 | 0.78 |
| Vas a morir me voy a asegurar de eso | 0.62 | 1 | 0.9 | 0.42 | 0.51 | 0.65 | 0.94 |
| Saca la denuncia o ya vas a ver | 0.46 | 1 | 0 | 0.33 | 0.43 | 0.75 | 0.78 |
| Ni se te ocurra denunciarme te puede ir muy mal | 0.58 | 0.88 | 0 | 0.22 | 0.17 | 0.54 | 0.44 |
| **Errores NO VIOLENTO** | 2 | 16 | 3 | 0 | 2 | 4 | 4 |
| **Errores VIOLENTO** | 80 | 10 | 32 | 73 | 68 | 32 | 18 |
| **Errores Totales** | 82 | 26 | 35 | 73 | 70 | 36 | 22 |

Se hicieron diagramas para poder visualizar mejor el desempeño general de los score para cada entrenamiento.

Para los entrenamientos 1, 4, 5, 6 y 7 en la categoría “No Violento”







Para los entrenamientos 1, 4, 5, 6 y 7 en la categoría “Violento”

# ChartChartChartChartChart

# 

# 

# Medicion de performance (para las 200 frases):

--- Entrenamiento 1 - 0.25 segundos ---

--- Entrenamiento 2 - 6.94 segundos ---

--- Entrenamiento 3 - 6.93 segundos ---

--- Entrenamiento 4 - 0.55 segundos ---

--- Entrenamiento 5 - 0.54 segundos ---

--- Entrenamiento 6 - 0.54 segundos ---

--- Entrenamiento 7 - 0.53 segundos ---

# 

# Modelo Multicategoria:

### Consideraciones previas al entrenamiento

Notas previas al armado del dataset:

**Se puede plantear un modelo que depende del modelo anterior para detectar mensajes violentos y simplemente categoriza dentro de estos a qué categorías se asemejan más. (se decidió por este)** Por otro lado se podría también plantear un modelo que es completamente independiente del anterior y que toma ‘no violento’ como otra categoría posible. Sería conveniente implementar ambos modelos para comparar sus resultados? (**No)**

Además el modelo anterior puede ser usado para encontrar frases violentas en grandes volúmenes de datos (por ejemplo tweets o posts de redes) para así encontrar más datos de entrenamiento para este modelo.

El dataset lo podemos armar con todas las frases juntas y después ir viendo de qué categoría falta según según el etiquetado (usando filtros de label studio si se puede y sino haciendo un script cortito).

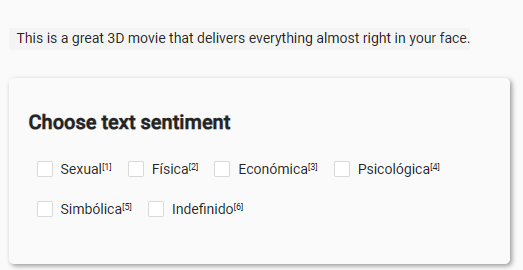
+tweets google juzgado

El dataset usado anteriormente parece no contener muchos ejemplos de las categorías: violencia sexual y violencia económica.

Dudas para reunión:

Notas posteriores a las primeras frases etiquetadas:

* Sería mejor que en lugar de No violento diga ambiguo?? En ese caso se trabajan con los ‘violentos’ ya detectados en el modelo anterior. : **En lugar de No violento se decidió por Indefinido**
* Definir los criterios más claramente, dar ejemplos de cada frase de cada tipo de violencia. **La definición de cada tipo de violencia está en una hoja de excel del dataset**
* Buscar más ejemplos de violencia económica y sexual. **Empezamos por las frases que tenemos y despues vemos que falta agregar al dataset.**
* Una categoría que parece que va a ser muy difícil de detectar si no definimos muy bien qué es y buscamos muchos ejemplos representativos es la violencia simbólica, porque requiere que el modelo detecte cuando se está haciendo presente la dinámica de poder del hombre frente a la mujer (estereotipos, roles de género) Tal vez deberíamos generar frases para el entrenamiento de forma artificial si no encontramos muchos ejemplos en internet o son todos muy parecidos. (Si tenemos 10 ejemplos marcados como violencia simbólica y 7 de ellos dicen alguna variedad de ‘callate y anda a la cocina’ solo va a detectar esas frases) **Por ahora consideramos algunos insultos raciales como violencia simbólica, esto puede no ser lo mejor y deba cambiar en el futuro. (si bien es violencia simbólica no necesariamente aplica a la violencia simbólica hacia la mujer.) Consideramos violencia simbólica a uso de estereotipos negativos o insultos que usan la diferencia de género para establecer una jerarquía de poder (‘tr\*la, anda a la cocina, etc.)**
* permitir la selección múltiple? **No, podría confundir al modelo**

****

### 

### Entrenamiento 1

Se seleccionaron para etiquetar 2086 frases.

De estas frases, para cada categoría se etiquetaron:

Sexual: 29 frases.

Física: 426 frases

Económica: 16 frases.

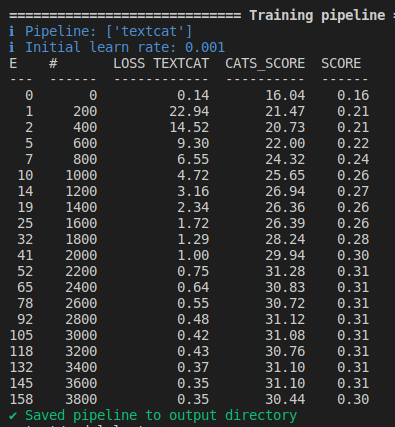
Psicológica: 244 frases.

Simbólica: 144 frases.

Indefinido: 258 frases.

En total, el dataset se compuso de 1117 frases, el resto se descartó.

Posteriormente se hizo un entrenamiento con una configuración típica y con un pipeline ‘es\_core\_news\_sm’.



Se obtuvo como mejor rendimiento un modelo con score de 31.

Este score es un promedio de todas las categorías. Cada categoría tuvo un score muy diferente:

| Score categoría | |
| --- | --- |
| Sexual | 23 |
| Física | 80 |
| Económica | 0 |
| Psicológica | 44 |
| Simbólica | 21 |
| Indefinido | 17 |

Se hicieron pruebas con el siguiente dataset

--Física

Te voy a matar hija de puta

La puta que te pario, te voy a volar la cabeza de un tiro

Dame al nene o te mato

Te voy a romper la cabeza si te vuelvo a ver

Te cago a palos

Te voy a acuchillar

Te voy a reventar

tres tiros te voy a pegar

Te arranco los dientes

--Psicológica

Vos tenes la culpa de lo te voy a hacer si te vuelvo a ver hija de puta

Sos una mierda malcogida no te quiere nadie

Si no volves conmigo no vas a estar con nadie

--Económica

Negra de mierda deja de comprarte ropa con mi plata

Si volves a llegar tarde no te doy mas plata

Puta en casa mando yo sino queres te vas sin nada a la calle

--Sexual

Me vas a chupar la pija aunque no quieras

Hacete coger la concha de tu hermana

Puta de mierda te voy a violar

soy tu marido y te toco todo lo que quiero

--Simbólica

A vos no te da la cabeza para entender

que rapida que sos para conseguir un macho hija de puta

Tendrias que estar limpiando porque yo te lo ordene

--------------------------------------------------------------------------

*Tendrían que clasificarse como ‘Física’*

Te voy a matar hija de puta

Sexual : 0.09 - Física : 0.829 - Económica : 0.001 - Psicológica : 0.039 - Simbólica : 0.008 - Indefinido : 0.033 -

La puta que te pario, te voy a volar la cabeza de un tiro

Sexual : 0.021 - Física : 0.565 - Económica : 0.0 - Psicológica : 0.02 - Simbólica : 0.004 - Indefinido : 0.39 -

Dame al nene o te mato

Sexual : 0.003 - Física : 0.961 - Económica : 0.002 - Psicológica : 0.012 - Simbólica : 0.001 - Indefinido : 0.022 -

Te voy a romper la cabeza si te vuelvo a ver

Sexual : 0.002 - Física : 0.981 - Económica : 0.0 - Psicológica : 0.004 - Simbólica : 0.001 - Indefinido : 0.013 -

Te cago a palos

Sexual : 0.063 - Física : 0.8 - Económica : 0.009 - Psicológica : 0.028 - Simbólica : 0.044 - Indefinido : 0.057 -

Te voy a acuchillar

Sexual : 0.186 - Física : 0.648 - Económica : 0.005 - Psicológica : 0.051 - Simbólica : 0.036 - Indefinido : 0.074 -

Te voy a reventar

Sexual : 0.08 - Física : 0.855 - Económica : 0.003 - Psicológica : 0.022 - Simbólica : 0.014 - Indefinido : 0.027 -

tres tiros te voy a pegar

Sexual : 0.091 - Física : 0.738 - Económica : 0.006 - Psicológica : 0.053 - Simbólica : 0.023 - Indefinido : 0.09 -

Te arranco los dientes

Sexual : 0.034 - Física : 0.848 - Económica : 0.01 - Psicológica : 0.031 - Simbólica : 0.023 - Indefinido : 0.054 -

*Tendrían que clasificarse como ‘Psícológica’*

Vos tenes la culpa de lo te voy a hacer si te vuelvo a ver hija de puta

Sexual : 0.016 - Física : 0.614 - Económica : 0.0 - Psicológica : 0.147 - Simbólica : 0.001 - Indefinido : 0.222 -

Sos una mierda malcogida no te quiere nadie

Sexual : 0.013 - Física : 0.192 - Económica : 0.009 - Psicológica : 0.173 - Simbólica : 0.095 - Indefinido : 0.519 -

Si no volves conmigo no vas a estar con nadie

Sexual : 0.069 - Física : 0.199 - Económica : 0.01 - Psicológica : 0.185 - Simbólica : 0.086 - Indefinido : 0.45 -

*Tendrían que clasificarse como ‘Económica’*

Negra de mierda deja de comprarte ropa con mi plata

Sexual : 0.011 - Física : 0.204 - Económica : 0.011 - Psicológica : 0.022 - Simbólica : 0.625 - Indefinido : 0.127 -

Si volves a llegar tarde no te doy mas plata

Sexual : 0.023 - Física : 0.726 - Económica : 0.007 - Psicológica : 0.112 - Simbólica : 0.035 - Indefinido : 0.096 -

Puta en casa mando yo sino queres te vas sin nada a la calle

Sexual : 0.019 - Física : 0.256 - Económica : 0.003 - Psicológica : 0.088 - Simbólica : 0.245 - Indefinido : 0.389 -

*Tendrían que clasificarse como ‘Sexual’*

Me vas a chupar la pija aunque no quieras

Sexual : 0.429 - Física : 0.182 - Económica : 0.012 - Psicológica : 0.181 - Simbólica : 0.025 - Indefinido : 0.172 -

Hacete coger la concha de tu hermana

Sexual : 0.168 - Física : 0.06 - Económica : 0.016 - Psicológica : 0.08 - Simbólica : 0.644 - Indefinido : 0.032 -

Puta de mierda te voy a violar

Sexual : 0.087 - Física : 0.213 - Económica : 0.002 - Psicológica : 0.059 - Simbólica : 0.176 - Indefinido : 0.463 -

soy tu marido y te toco todo lo que quiero

Sexual : 0.019 - Física : 0.098 - Económica : 0.013 - Psicológica : 0.232 - Simbólica : 0.34 - Indefinido : 0.296 -

*Tendrían que clasificarse como ‘Simbólica’*

A vos no te da la cabeza para entender

Sexual : 0.007 - Física : 0.735 - Económica : 0.002 - Psicológica : 0.021 - Simbólica : 0.027 - Indefinido : 0.209 -

que rapida que sos para conseguir un macho hija de puta

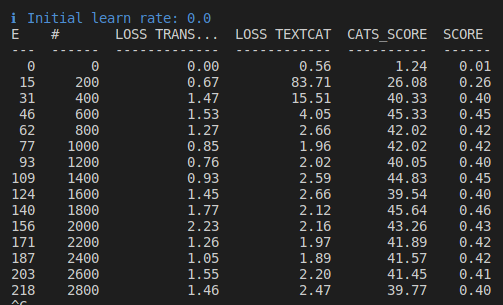
Sexual : 0.025 - Física : 0.037 - Económica : 0.044 - Psicológica : 0.246 - Simbólica : 0.235 - Indefinido : 0.413 -

Tendrias que estar limpiando porque yo te lo ordene

Sexual : 0.009 - Física : 0.181 - Económica : 0.006 - Psicológica : 0.077 - Simbólica : 0.239 - Indefinido : 0.488 -

### Entrenamiento 1 con transformers

Se repitió el entrenamiento con el mismo dataset, pero se utilizaron transformers esta vez.



Se obtuvo como mejor rendimiento un modelo con score de 46.

Este score es un promedio de todas las categorías. Cada categoría tuvo un score muy diferente:

| Score categoría | |
| --- | --- |
| Sexual | 52 |
| Física | 87 |
| Económica | 22 |
| Psicológica | 37 |
| Simbólica | 40 |
| Indefinido | 34 |

Se hicieron pruebas con el mismo dataset que en el entrenamiento 1. Se obtuvieron los siguientes resultados:

*Tendrían que clasificarse como ‘Física’*

Te voy a matar hija de puta

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

La puta que te pario, te voy a volar la cabeza de un tiro

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Dame al nene o te mato

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Te voy a romper la cabeza si te vuelvo a ver

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Te cago a palos

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Te voy a acuchillar

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Te voy a reventar

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

tres tiros te voy a pegar

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Te arranco los dientes

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

*Tendrían que clasificarse como ‘Psícológica’*

Vos tenes la culpa de lo te voy a hacer si te vuelvo a ver hija de puta

Sexual : 0.0 - Física : 0.004 - Económica : 0.0 - Psicológica : 0.995 - Simbólica : 0.0 - Indefinido : 0.0 -

Sos una mierda malcogida no te quiere nadie

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.999 - Indefinido : 0.001 -

Si no volves conmigo no vas a estar con nadie

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 1.0 -

*Tendrían que clasificarse como ‘Económica’*

Negra de mierda deja de comprarte ropa con mi plata

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 - Indefinido : 0.0 -

Si volves a llegar tarde no te doy mas plata

Sexual : 0.0 - Física : 0.03 - Económica : 0.004 - Psicológica : 0.619 - Simbólica : 0.001 - Indefinido : 0.346 -

Puta en casa mando yo sino queres te vas sin nada a la calle

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 - Indefinido : 0.0 -

*Tendrían que clasificarse como ‘Sexual’*

Me vas a chupar la pija aunque no quieras

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Hacete coger la concha de tu hermana

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Puta de mierda te voy a violar

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

soy tu marido y te toco todo lo que quiero

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 - Indefinido : 0.0 -

*Tendrían que clasificarse como ‘Simbólica’*

A vos no te da la cabeza para entender

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.952 - Simbólica : 0.0 - Indefinido : 0.047 -

que rapida que sos para conseguir un macho hija de puta

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 1.0 -

Tendrias que estar limpiando porque yo te lo ordene

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 1.0 -

### Conclusiones del entrenamiento

* Excepto para la categoría ‘Física’, el modelo de reconocimiento multicategoría tiene muchos fallos para detectar correctamente el tipo de violencia.
* Esto se atribuye principalmente a la falta de ejemplos adecuados para casi todos los tipos de violencia.
* Especialmente hay una gran falta de ejemplos en las categorías ‘Económica’ y ‘Sexual’
* En la categoría ‘Física’, el resultado es adecuado. Incluso el modelo reconoce frases que no contienen un insulto o un mención al cuerpo de la mujer (‘tres tiros te voy a pegar’ fue reconocido correctamente).
* El entrenamiento con transformers aumenta considerablemente el score del modelo y de cada categoría. Pero esto no produce un cambio claro en las pruebas. Solo mejoran levemente los resultados de las categoría ‘Sexual’ y ‘Psicológica’.

### 

### Notas post-entrenamiento

Debido a los resultados del entrenamiento y al proceso de etiquetado se decidió que los datos del dataset no eran lo suficientemente representativos como para formar un modelo sólido. Por lo tanto vamos a crear un nuevo dataset.

Creación del nuevo dataset:

1. Generamos un documento para tener en un lugar unos criterios de etiquetado por categoría mejor definidos. (definicion de cada categoria y ejemplos)
2. Creamos excel con las frases que pudimos etiquetar en el dataset pasado-
3. De esas frases podemos encontrar las más representativas e intentar aplicar augmentations para hacer el volumen de la muestra más grande.

Para lograr esto es necesario indagar sobre la aplicación de data augmentation dado que no tenemos conocimientos en este campo.

Uso de chatgpt para encontrar algunas frases de ejemplo que pueden ser modificadas con tecnicas de data agumentation (como el uso de sinónimos, mezclar las palabras o agregar palabras, etc.) No es posible pedir ejemplos de forma directa:

Con esta técnica pudimos generar 130 frases base.

Despues de consultar con bruno y hacer un par de pruebas se llegó a la conclusión de que el uso de data augmentation para texto es bastante complejo y podría traer problemas de overfitting.

Por lo tanto se vuelven a buscar datos reales para etiquetar

**(esto es una idea, todavía no lo aplicamos)**

**link dataset de reddit:**

[**https://huggingface.co/datasets/fddemarco/pushshift-reddit-comments/tree/main**](https://huggingface.co/datasets/fddemarco/pushshift-reddit-comments/tree/main)

Los archivos son muy grandes. Se pueden ir descargando de a uno y haciendo el filtro.

filtrar: subreddit argentina, (también se puede filtrar palabras que buscamos ‘puta’, ‘trola’ ‘mujer’, ‘mujeres’ etc.)

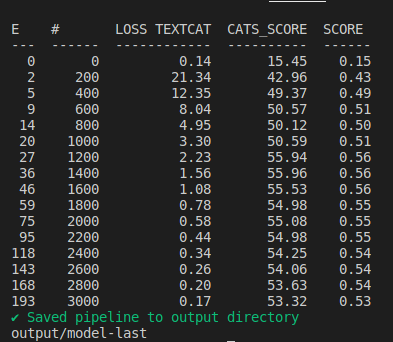
**nota: Deberíamos chequear que el modelo binario identifique las frases de violencia económica (y simbólica) como violentas. Porque como tienden a tener menos insultos, capaz que nunca pasan ese primer filtro. Estaría bueno probar agregando todas las frases que tenemos en ese tipo de violencia como ‘Violentas’ y reentrenar a ver como sale el modelo.**

### Entrenamiento 2

Se hizo un entrenamiento adicional, con un dataset más limitado.

Este dataset se compone de frases seleccionadas del dataset del entrenamiento 1. Además, se agregaron frases simples, y versiones modificadas de estas frases para que sean más parecidas a las formas de expresión habituales en Argentina.

El entrenamiento arrojó los siguientes resultados:



El mejor modelo arrojó los siguientes scores

| Score categoría | |
| --- | --- |
| Sexual | 33 |
| Física | 85 |
| Económica | 76 |
| Psicológica | 72 |
| Simbólica | 68 |
| Indefinido | 0 |

Se hicieron pruebas con el mismo dataset que en el entrenamiento 1. Se obtuvieron los siguientes resultados:

*Tendrían que clasificarse como ‘Física’*

Te voy a matar hija de puta

Sexual : 0.035 - Física : 0.892 - Económica : 0.022 - Psicológica : 0.045 - Simbólica : 0.006 - Indefinido : 0.001 -

La puta que te pario, te voy a volar la cabeza de un tiro

Sexual : 0.012 - Física : 0.935 - Económica : 0.029 - Psicológica : 0.016 - Simbólica : 0.007 - Indefinido : 0.0 -

Dame al nene o te mato

Sexual : 0.028 - Física : 0.686 - Económica : 0.131 - Psicológica : 0.113 - Simbólica : 0.03 - Indefinido : 0.012 -

Te voy a romper la cabeza si te vuelvo a ver

Sexual : 0.013 - Física : 0.924 - Económica : 0.03 - Psicológica : 0.032 - Simbólica : 0.001 - Indefinido : 0.0 -

Te cago a palos

Sexual : 0.055 - Física : 0.693 - Económica : 0.08 - Psicológica : 0.127 - Simbólica : 0.03 - Indefinido : 0.016 -

Te voy a acuchillar

Sexual : 0.11 - Física : 0.433 - Económica : 0.198 - Psicológica : 0.221 - Simbólica : 0.026 - Indefinido : 0.013 -

Te voy a reventar

Sexual : 0.071 - Física : 0.667 - Económica : 0.111 - Psicológica : 0.122 - Simbólica : 0.018 - Indefinido : 0.011 -

tres tiros te voy a pegar

Sexual : 0.027 - Física : 0.921 - Económica : 0.016 - Psicológica : 0.025 - Simbólica : 0.008 - Indefinido : 0.003 -

Te arranco los dientes

Sexual : 0.112 - Física : 0.329 - Económica : 0.165 - Psicológica : 0.266 - Simbólica : 0.088 - Indefinido : 0.041 -

*Tendrían que clasificarse como ‘Psícológica’*

Vos tenes la culpa de lo te voy a hacer si te vuelvo a ver hija de puta

Sexual : 0.007 - Física : 0.648 - Económica : 0.024 - Psicológica : 0.317 - Simbólica : 0.003 - Indefinido : 0.0 -

Sos una mierda malcogida no te quiere nadie

Sexual : 0.002 - Física : 0.015 - Económica : 0.004 - Psicológica : 0.938 - Simbólica : 0.041 - Indefinido : 0.001 -

Si no volves conmigo no vas a estar con nadie

Sexual : 0.004 - Física : 0.1 - Económica : 0.014 - Psicológica : 0.669 - Simbólica : 0.211 - Indefinido : 0.001 -

*Tendrían que clasificarse como ‘Económica’*

Negra de mierda deja de comprarte ropa con mi plata

Sexual : 0.03 - Física : 0.088 - Económica : 0.662 - Psicológica : 0.104 - Simbólica : 0.114 - Indefinido : 0.001 -

Si volves a llegar tarde no te doy mas plata

Sexual : 0.003 - Física : 0.028 - Económica : 0.92 - Psicológica : 0.031 - Simbólica : 0.016 - Indefinido : 0.001 -

Puta en casa mando yo sino queres te vas sin nada a la calle

Sexual : 0.001 - Física : 0.019 - Económica : 0.883 - Psicológica : 0.062 - Simbólica : 0.035 - Indefinido : 0.0 -

*Tendrían que clasificarse como ‘Sexual’*

Me vas a chupar la pija aunque no quieras

Sexual : 0.377 - Física : 0.069 - Económica : 0.097 - Psicológica : 0.406 - Simbólica : 0.047 - Indefinido : 0.004 -

Hacete coger la concha de tu hermana

Sexual : 0.233 - Física : 0.109 - Económica : 0.209 - Psicológica : 0.226 - Simbólica : 0.21 - Indefinido : 0.013 -

Puta de mierda te voy a violar

Sexual : 0.123 - Física : 0.44 - Económica : 0.11 - Psicológica : 0.251 - Simbólica : 0.073 - Indefinido : 0.002 -

soy tu marido y te toco todo lo que quiero

Sexual : 0.676 - Física : 0.047 - Económica : 0.136 - Psicológica : 0.085 - Simbólica : 0.054 - Indefinido : 0.002 -

*Tendrían que clasificarse como ‘Simbólica’*

A vos no te da la cabeza para entender

Sexual : 0.026 - Física : 0.539 - Económica : 0.149 - Psicológica : 0.166 - Simbólica : 0.116 - Indefinido : 0.003 -

que rapida que sos para conseguir un macho hija de puta

Sexual : 0.011 - Física : 0.06 - Económica : 0.1 - Psicológica : 0.2 - Simbólica : 0.628 - Indefinido : 0.0 -

Tendrias que estar limpiando porque yo te lo ordene

Sexual : 0.019 - Física : 0.114 - Económica : 0.235 - Psicológica : 0.21 - Simbólica : 0.418 - Indefinido : 0.004 -

### Entrenamiento 2 con transformers

Se utilizó el mismo dataset pero se hizo un entrenamiento con transformers. Los resultados fueron

El mejor modelo arrojó los siguientes scores

| Score categoría | |
| --- | --- |
| Sexual | 88 |
| Física | 93 |
| Económica | 93 |
| Psicológica | 83 |
| Simbólica | 79 |
| Indefinido | 0 |

*Tendrían que clasificarse como ‘Física’*

Te voy a matar hija de puta

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

La puta que te pario, te voy a volar la cabeza de un tiro

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Dame al nene o te mato

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Te voy a romper la cabeza si te vuelvo a ver

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Te cago a palos

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Te voy a acuchillar

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Te voy a reventar

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

tres tiros te voy a pegar

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Te arranco los dientes

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

*Tendrían que clasificarse como ‘Psícológica’*

Vos tenes la culpa de lo te voy a hacer si te vuelvo a ver hija de puta

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Sos una mierda malcogida no te quiere nadie

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Si no volves conmigo no vas a estar con nadie

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 - Indefinido : 0.0 -

*Tendrían que clasificarse como ‘Económica’*

Negra de mierda deja de comprarte ropa con mi plata

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Si volves a llegar tarde no te doy mas plata

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Puta en casa mando yo sino queres te vas sin nada a la calle

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

*Tendrían que clasificarse como ‘Sexual’*

Me vas a chupar la pija aunque no quieras

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Hacete coger la concha de tu hermana

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 - Indefinido : 0.0 -

Puta de mierda te voy a violar

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

soy tu marido y te toco todo lo que quiero

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 - Indefinido : 0.0 -

*Tendrían que clasificarse como ‘Simbólica’*

A vos no te da la cabeza para entender

Sexual : 0.0 - Física : 0.144 - Económica : 0.011 - Psicológica : 0.844 - Simbólica : 0.001 - Indefinido : 0.001 -

que rapida que sos para conseguir un macho hija de puta

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.093 - Simbólica : 0.906 - Indefinido : 0.0 -

Tendrias que estar limpiando porque yo te lo ordene

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 - Indefinido : 0.0 -

### Conclusiones del entrenamiento

* El entrenamiento con un dataset más limitado y mejor definido dio mejores resultados.
* El incremento de ejemplos en las categorías “económica” y el decremento en las categorías “psicológica” y “simbólica” de mayor calidad tuvieron un efecto positivo en los score.
* Las pruebas también fueron mucho más consistentes que en el entrenamiento 1. Hubo muchos menos errores en las categorías “psicológica” y “económica”.
* Los entrenamientos con transformers mejoraron aún más estos resultados. En las pruebas de texto los errores pasaron de 12/21 frases en el entrenamiento 1 sin transformers a 2/21 en el entrenamiento 2 con transformers.
* Se dejaron afuera los ejemplos y las frases de categoría “indefinido”. Se tiene que revisar mejor las condiciones de clasificación para esta categoría.

### Notas pre entrenamiento 3

Hicimos búsquedas a mano usando las opciones de búsqueda avanzada de twitter. y busquedas de conversaciones reales encontradas en artículos blogs y noticias.

La categoría de indefinido es la más problemática en sentido de definir qué datos son de esa categoría y cuáles no (cómo es muy amplia prevemos que vamos a tener problemas.

### Entrenamiento 3

Se armó un dataset de frases:

Económica: 199 frases.

Física: 305 frases.

Psicológica: 408 frases.

Sexual: 255 frases.

Simbólica: 230 frases.

Se hizo un entrenamiento y se obtuvieron los siguientes scores:

| Score categoría | |
| --- | --- |
| Sexual | 67 |
| Física | 83 |
| Económica | 80 |
| Psicológica | 69 |
| Simbólica | 68 |

Se armó un dataset de prueba extendido, para probar nuevas frases con mayores ambiguedades. Los resultados fueron los siguientes:

--Física

Te voy a matar hija de puta

Sexual : 0.05 - Física : 0.903 - Económica : 0.014 - Psicológica : 0.029 - Simbólica : 0.004 -

La puta que te pario, te voy a volar la cabeza de un tiro

Sexual : 0.302 - Física : 0.645 - Económica : 0.024 - Psicológica : 0.025 - Simbólica : 0.003 -

Dame al nene o te mato

Sexual : 0.022 - Física : 0.837 - Económica : 0.051 - Psicológica : 0.073 - Simbólica : 0.018 -

Te voy a romper la cabeza si te vuelvo a ver

Sexual : 0.077 - Física : 0.895 - Económica : 0.011 - Psicológica : 0.016 - Simbólica : 0.0 -

Te cago a palos

Sexual : 0.075 - Física : 0.675 - Económica : 0.076 - Psicológica : 0.09 - Simbólica : 0.085 -

Te voy a acuchillar

Sexual : 0.157 - Física : 0.537 - Económica : 0.136 - Psicológica : 0.151 - Simbólica : 0.02 -

Te voy a reventar

Sexual : 0.072 - Física : 0.775 - Económica : 0.06 - Psicológica : 0.075 - Simbólica : 0.019 -

tres tiros te voy a pegar

Sexual : 0.02 - Física : 0.951 - Económica : 0.009 - Psicológica : 0.014 - Simbólica : 0.006 -

Te arranco los dientes

Sexual : 0.152 - Física : 0.384 - Económica : 0.166 - Psicológica : 0.222 - Simbólica : 0.076 -

Mañana amaneces muerta y tirada en una zanja

Sexual : 0.18 - Física : 0.212 - Económica : 0.158 - Psicológica : 0.272 - Simbólica : 0.179 -

A tus hijos los voy a matar frente tuyo

Sexual : 0.076 - Física : 0.675 - Económica : 0.016 - Psicológica : 0.178 - Simbólica : 0.056 -

20000 tiros en la cabeza y chau, no jodes mas

Sexual : 0.096 - Física : 0.13 - Económica : 0.276 - Psicológica : 0.319 - Simbólica : 0.179 -

quien queres que mate primero? tu vieja o tu hermana?

Sexual : 0.108 - Física : 0.027 - Económica : 0.012 - Psicológica : 0.836 - Simbólica : 0.017 -

la policia no te va a proteger, ya estas muerta

Sexual : 0.097 - Física : 0.148 - Económica : 0.013 - Psicológica : 0.724 - Simbólica : 0.018 -

deja de denunciarme porque voy a ir con un fierro a balearte la casa

Sexual : 0.033 - Física : 0.65 - Económica : 0.13 - Psicológica : 0.167 - Simbólica : 0.019 -

queres que te arranque la piel de la cara

Sexual : 0.256 - Física : 0.365 - Económica : 0.168 - Psicológica : 0.126 - Simbólica : 0.084 -

un balazo en la concha y listo, no hablas mas de mi

Sexual : 0.095 - Física : 0.032 - Económica : 0.512 - Psicológica : 0.226 - Simbólica : 0.134 -

prendo fuego tu casa con un bidon con vos adentro

Sexual : 0.006 - Física : 0.911 - Económica : 0.032 - Psicológica : 0.023 - Simbólica : 0.028 -

te van a encontrar carbonizada en el rio

Sexual : 0.42 - Física : 0.077 - Económica : 0.082 - Psicológica : 0.331 - Simbólica : 0.089 -

voy a cortarte en pedazos con mi machete para que aprendas

Sexual : 0.199 - Física : 0.383 - Económica : 0.276 - Psicológica : 0.069 - Simbólica : 0.072 -

mira este cuchillo lo voy a meter en tu garganta

Sexual : 0.181 - Física : 0.447 - Económica : 0.128 - Psicológica : 0.224 - Simbólica : 0.019 -

despedite de tu familia porque ya estoy yendo a matarte

Sexual : 0.016 - Física : 0.053 - Económica : 0.005 - Psicológica : 0.923 - Simbólica : 0.003 -

a mi me van a meter en la carcel pero antes vas a morir

Sexual : 0.243 - Física : 0.358 - Económica : 0.089 - Psicológica : 0.273 - Simbólica : 0.036 -

una bala en la cabeza para vos y luego una para mi

Sexual : 0.112 - Física : 0.173 - Económica : 0.252 - Psicológica : 0.069 - Simbólica : 0.394 -

--Psicológica

Vos tenes la culpa de lo te voy a hacer si te vuelvo a ver hija de puta

Sexual : 0.023 - Física : 0.573 - Económica : 0.044 - Psicológica : 0.358 - Simbólica : 0.002 -

Sos una mierda malcogida no te quiere nadie

Sexual : 0.007 - Física : 0.006 - Económica : 0.005 - Psicológica : 0.925 - Simbólica : 0.057 -

Si no volves conmigo no vas a estar con nadie

Sexual : 0.006 - Física : 0.091 - Económica : 0.01 - Psicológica : 0.566 - Simbólica : 0.326 -

No vales nada, te mereces que te trate como a una mierda

Sexual : 0.03 - Física : 0.048 - Económica : 0.131 - Psicológica : 0.744 - Simbólica : 0.046 -

Parece que le haces mas caso a tus amigas que a mi

Sexual : 0.027 - Física : 0.018 - Económica : 0.143 - Psicológica : 0.029 - Simbólica : 0.783 -

¿Por qué no te matas y te ahorras el sufrimiento de vivir?

Sexual : 0.17 - Física : 0.023 - Económica : 0.225 - Psicológica : 0.449 - Simbólica : 0.133 -

que asco que das, quien te va a coger así?

Sexual : 0.176 - Física : 0.008 - Económica : 0.007 - Psicológica : 0.787 - Simbólica : 0.022 -

que idea horrible que tuviste, nunca se te ocurre algo bueno

Sexual : 0.021 - Física : 0.027 - Económica : 0.22 - Psicológica : 0.675 - Simbólica : 0.058 -

los dos sabemos que esto no te sale, mejor ni lo intentes porque la vas a cagar

Sexual : 0.028 - Física : 0.096 - Económica : 0.059 - Psicológica : 0.77 - Simbólica : 0.048 -

si vos sos así de estupida para que te esforzas?

Sexual : 0.066 - Física : 0.012 - Económica : 0.121 - Psicológica : 0.211 - Simbólica : 0.591 -

enferma mogólica, te mereces lo que te voy a hacer

Sexual : 0.06 - Física : 0.261 - Económica : 0.055 - Psicológica : 0.61 - Simbólica : 0.014 -

con quien carajo estas hablando a esta hora?

Sexual : 0.031 - Física : 0.097 - Económica : 0.015 - Psicológica : 0.834 - Simbólica : 0.023 -

te voy a hacer la vida imposible hija de puta

Sexual : 0.209 - Física : 0.416 - Económica : 0.038 - Psicológica : 0.325 - Simbólica : 0.013 -

suicidate, sos una basura

Sexual : 0.046 - Física : 0.026 - Económica : 0.081 - Psicológica : 0.39 - Simbólica : 0.457 -

te vas a quedar sola toda tu vida

Sexual : 0.163 - Física : 0.3 - Económica : 0.035 - Psicológica : 0.492 - Simbólica : 0.009 -

te deseo una muerte dolorosa hija de puta

Sexual : 0.305 - Física : 0.201 - Económica : 0.09 - Psicológica : 0.258 - Simbólica : 0.147 -

--Económica

Negra de mierda deja de comprarte ropa con mi plata

Sexual : 0.026 - Física : 0.029 - Económica : 0.882 - Psicológica : 0.033 - Simbólica : 0.03 -

Si volves a llegar tarde no te doy mas plata

Sexual : 0.009 - Física : 0.012 - Económica : 0.961 - Psicológica : 0.009 - Simbólica : 0.009 -

Puta en casa mando yo sino queres te vas sin nada a la calle

Sexual : 0.002 - Física : 0.01 - Económica : 0.678 - Psicológica : 0.307 - Simbólica : 0.004 -

queres volver a mi casa a dormir empeza a responderme antes

Sexual : 0.026 - Física : 0.554 - Económica : 0.242 - Psicológica : 0.162 - Simbólica : 0.016 -

quien carajo te crees que sos para comprarte eso sin mi permiso

Sexual : 0.164 - Física : 0.12 - Económica : 0.164 - Psicológica : 0.381 - Simbólica : 0.171 -

otra top mas? yo no te doy un centavo mas

Sexual : 0.001 - Física : 0.001 - Económica : 0.967 - Psicológica : 0.009 - Simbólica : 0.023 -

listo, no te voy a dar nada porque te vas a ir a coger con un negro

Sexual : 0.136 - Física : 0.246 - Económica : 0.409 - Psicológica : 0.202 - Simbólica : 0.007 -

y a vos quien te dio permiso de usar mi tarjeta putita?

Sexual : 0.285 - Física : 0.037 - Económica : 0.135 - Psicológica : 0.424 - Simbólica : 0.119 -

hija de puta para que mierda nos seguis endeudando

Sexual : 0.169 - Física : 0.111 - Económica : 0.142 - Psicológica : 0.432 - Simbólica : 0.147 -

este es mi hogar, y aca se hace lo que yo digo

Sexual : 0.024 - Física : 0.002 - Económica : 0.495 - Psicológica : 0.083 - Simbólica : 0.396 -

yo te doy todo y vos te rascas la concha, mantenida de mierda

Sexual : 0.015 - Física : 0.028 - Económica : 0.874 - Psicológica : 0.078 - Simbólica : 0.005 -

vos no trabajas, estas en casa cuidando al nene nada mas

Sexual : 0.002 - Física : 0.01 - Económica : 0.405 - Psicológica : 0.249 - Simbólica : 0.334 -

seguis usando la plata para tus boludeces

Sexual : 0.04 - Física : 0.078 - Económica : 0.719 - Psicológica : 0.103 - Simbólica : 0.06 -

pedazo de negra parasita, rompes todo y yo tengo que pagarlo

Sexual : 0.089 - Física : 0.046 - Económica : 0.068 - Psicológica : 0.322 - Simbólica : 0.474 -

no me voy a bancar que sigas usando la tarjeta sin mi autorizacion

Sexual : 0.079 - Física : 0.075 - Económica : 0.675 - Psicológica : 0.143 - Simbólica : 0.028 -

conchuda agradece que seguis viviendo bajo mi techo

Sexual : 0.094 - Física : 0.161 - Económica : 0.285 - Psicológica : 0.179 - Simbólica : 0.282 -

no te voy a mandar mas la cuota para el nene

Sexual : 0.126 - Física : 0.231 - Económica : 0.534 - Psicológica : 0.044 - Simbólica : 0.065 -

mejor que no sepas cuanto tengo en la cuenta porque me lo vas a sacar

Sexual : 0.005 - Física : 0.001 - Económica : 0.125 - Psicológica : 0.848 - Simbólica : 0.022 -

cuando cobres dame la plata que la voy a manejar yo

Sexual : 0.011 - Física : 0.034 - Económica : 0.922 - Psicológica : 0.026 - Simbólica : 0.007 -

--Sexual

Me vas a chupar la pija aunque no quieras

Sexual : 0.684 - Física : 0.037 - Económica : 0.021 - Psicológica : 0.234 - Simbólica : 0.023 -

Hacete coger la concha de tu hermana

Sexual : 0.378 - Física : 0.082 - Económica : 0.024 - Psicológica : 0.156 - Simbólica : 0.36 -

Puta de mierda te voy a violar

Sexual : 0.819 - Física : 0.058 - Económica : 0.024 - Psicológica : 0.085 - Simbólica : 0.013 -

soy tu marido y te toco todo lo que quiero

Sexual : 0.855 - Física : 0.02 - Económica : 0.053 - Psicológica : 0.053 - Simbólica : 0.019 -

voy a entrar a tu casa para reventarte la concha

Sexual : 0.095 - Física : 0.483 - Económica : 0.042 - Psicológica : 0.344 - Simbólica : 0.037 -

que ganas de secuestrarte y romperte el orto sin que nadie se entere

Sexual : 0.706 - Física : 0.075 - Económica : 0.042 - Psicológica : 0.089 - Simbólica : 0.088 -

ayer te segui mientras iba al trabajo y desee hacerte mia

Sexual : 0.444 - Física : 0.035 - Económica : 0.031 - Psicológica : 0.197 - Simbólica : 0.292 -

si te llego a encontrar en la calle te violo

Sexual : 0.431 - Física : 0.109 - Económica : 0.179 - Psicológica : 0.265 - Simbólica : 0.016 -

te vamos a coger entre todos, cuando terminemos te tiramos a una zanja

Sexual : 0.267 - Física : 0.223 - Económica : 0.041 - Psicológica : 0.411 - Simbólica : 0.058 -

sino te sentas en mi pija te degollo

Sexual : 0.697 - Física : 0.082 - Económica : 0.12 - Psicológica : 0.069 - Simbólica : 0.031 -

o me mandas una foto tuya desnuda o te voy a buscar

Sexual : 0.343 - Física : 0.093 - Económica : 0.021 - Psicológica : 0.5 - Simbólica : 0.044 -

podemos garchar por las buenas o por las malas, vos elegis

Sexual : 0.316 - Física : 0.017 - Económica : 0.019 - Psicológica : 0.359 - Simbólica : 0.289 -

aunque no quieras te voy a hacer mia esta noche

Sexual : 0.125 - Física : 0.25 - Económica : 0.126 - Psicológica : 0.264 - Simbólica : 0.235 -

cuando salgas de tu casa te voy a agarrar y te la meto bien adentro

Sexual : 0.295 - Física : 0.371 - Económica : 0.029 - Psicológica : 0.295 - Simbólica : 0.011 -

jajaj aunque llames a la policia te voy a dejar embarazada hoy

Sexual : 0.268 - Física : 0.236 - Económica : 0.346 - Psicológica : 0.139 - Simbólica : 0.012 -

cuidate porque si te cruzo te reviento a pijazos

Sexual : 0.065 - Física : 0.772 - Económica : 0.038 - Psicológica : 0.094 - Simbólica : 0.031 -

aunque tengas novio te voy a garchar hoy

Sexual : 0.428 - Física : 0.19 - Económica : 0.119 - Psicológica : 0.208 - Simbólica : 0.056 -

soy tu novio, si yo quiero ponerte en 4 no me lo vas a impedir

Sexual : 0.069 - Física : 0.008 - Económica : 0.108 - Psicológica : 0.659 - Simbólica : 0.156 -

te amo, quiero que seas mia hija de puta

Sexual : 0.374 - Física : 0.062 - Económica : 0.051 - Psicológica : 0.353 - Simbólica : 0.159 -

zorra malparida, ya te voy a llevar a mi cama para cogerte

Sexual : 0.861 - Física : 0.062 - Económica : 0.016 - Psicológica : 0.05 - Simbólica : 0.012 -

te voy a callar metiendote la pija en la boca

Sexual : 0.911 - Física : 0.056 - Económica : 0.021 - Psicológica : 0.006 - Simbólica : 0.006 -

--Simbólica

como sos una mina a vos no te da la cabeza para entender

Sexual : 0.014 - Física : 0.241 - Económica : 0.056 - Psicológica : 0.067 - Simbólica : 0.623 -

que rapida que sos para conseguir un macho hija de puta

Sexual : 0.035 - Física : 0.042 - Económica : 0.33 - Psicológica : 0.098 - Simbólica : 0.496 -

Tendrias que estar limpiando porque sos mujer

Sexual : 0.021 - Física : 0.033 - Económica : 0.017 - Psicológica : 0.059 - Simbólica : 0.87 -

todas las mujeres son igual de putas

Sexual : 0.114 - Física : 0.02 - Económica : 0.02 - Psicológica : 0.06 - Simbólica : 0.786 -

ustedes se hacen las santas y despues se bajan la bombacha rapidito

Sexual : 0.104 - Física : 0.015 - Económica : 0.031 - Psicológica : 0.147 - Simbólica : 0.702 -

claro sos mujer, no te sale tan facil

Sexual : 0.041 - Física : 0.07 - Económica : 0.022 - Psicológica : 0.204 - Simbólica : 0.663 -

el rol tuyo es cocinar y limpiar la casa

Sexual : 0.064 - Física : 0.007 - Económica : 0.042 - Psicológica : 0.123 - Simbólica : 0.764 -

apenas ven una pija ya quieren chuparla

Sexual : 0.427 - Física : 0.048 - Económica : 0.045 - Psicológica : 0.25 - Simbólica : 0.231 -

a mi hija la voy a criar para que sea una buena ama de casa

Sexual : 0.202 - Física : 0.161 - Económica : 0.183 - Psicológica : 0.315 - Simbólica : 0.139 -

no jugues futbol, ese no es un deporte de minas

Sexual : 0.113 - Física : 0.021 - Económica : 0.182 - Psicológica : 0.191 - Simbólica : 0.494 -

deberias ser mas obediente conmigo, soy tu marido

Sexual : 0.054 - Física : 0.016 - Económica : 0.021 - Psicológica : 0.73 - Simbólica : 0.178 -

tu tarea como mujer es satisfacerme

Sexual : 0.204 - Física : 0.042 - Económica : 0.015 - Psicológica : 0.078 - Simbólica : 0.661 -

yo pense que eras diferente pero sos una facil

Sexual : 0.02 - Física : 0.01 - Económica : 0.063 - Psicológica : 0.381 - Simbólica : 0.527 -

no engordes mas, tenes que ser flaquita o ningun hombre te va a mirar

Sexual : 0.072 - Física : 0.098 - Económica : 0.03 - Psicológica : 0.396 - Simbólica : 0.403 -

que emocional que sos, deja de mariconear

Sexual : 0.061 - Física : 0.031 - Económica : 0.059 - Psicológica : 0.167 - Simbólica : 0.681 -

dale te abris de piernas todo el tiempo

Sexual : 0.657 - Física : 0.092 - Económica : 0.06 - Psicológica : 0.031 - Simbólica : 0.16 -

anda a limpiar antes de que tu marido llegue

Sexual : 0.095 - Física : 0.046 - Económica : 0.033 - Psicológica : 0.033 - Simbólica : 0.792 -

uy tenias que ser vieja

Sexual : 0.145 - Física : 0.091 - Económica : 0.041 - Psicológica : 0.484 - Simbólica : 0.238 -

eres muy inteligente para ser mujer

Sexual : 0.064 - Física : 0.061 - Económica : 0.004 - Psicológica : 0.084 - Simbólica : 0.787 -

aprende a comportarte como señorita

Sexual : 0.192 - Física : 0.192 - Económica : 0.102 - Psicológica : 0.271 - Simbólica : 0.241 -

sos una cualquiera, le mostraste las tetas a todo el barrio

Sexual : 0.498 - Física : 0.072 - Económica : 0.046 - Psicológica : 0.114 - Simbólica : 0.27 -

vos solo sabes chupar la pija, te sale natural

Sexual : 0.883 - Física : 0.01 - Económica : 0.01 - Psicológica : 0.043 - Simbólica : 0.054 -

cuando vienen mis amigos no tenes que opinar, anda a la cocina

Sexual : 0.008 - Física : 0.003 - Económica : 0.089 - Psicológica : 0.096 - Simbólica : 0.805 -

y la comida? vos tenes mas facilidad porque sos mina

Sexual : 0.006 - Física : 0.002 - Económica : 0.08 - Psicológica : 0.054 - Simbólica : 0.857 -

todas las minas son igual de jodidas como vos

Sexual : 0.094 - Física : 0.029 - Económica : 0.038 - Psicológica : 0.13 - Simbólica : 0.709 -

te haces la machita pero tenes concha igual que todas

Sexual : 0.116 - Física : 0.053 - Económica : 0.185 - Psicológica : 0.338 - Simbólica : 0.308 -

### Entrenamiento 3 con transformers

Para el mismo entrenamiento con el dataset del entrenamiento 3, se usaron transformers. Los scores fueron los siguientes:

| Score categoría | |
| --- | --- |
| Sexual | 83 |
| Física | 89 |
| Económica | 92 |
| Psicológica | 80 |
| Simbólica | 76 |

Para el mismo dataset de prueba, los resultados son:

--Física

Te voy a matar hija de puta

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

La puta que te pario, te voy a volar la cabeza de un tiro

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

Dame al nene o te mato

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

Te voy a romper la cabeza si te vuelvo a ver

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

Te cago a palos

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

Te voy a acuchillar

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

Te voy a reventar

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

tres tiros te voy a pegar

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

Te arranco los dientes

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

Mañana amaneces muerta y tirada en una zanja

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

A tus hijos los voy a matar frente tuyo

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

20000 tiros en la cabeza y chau, no jodes mas

Sexual : 0.0 - Física : 0.998 - Económica : 0.0 - Psicológica : 0.002 - Simbólica : 0.0 -

quien queres que mate primero? tu vieja o tu hermana?

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

la policia no te va a proteger, ya estas muerta

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

deja de denunciarme porque voy a ir con un fierro a balearte la casa

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

queres que te arranque la piel de la cara

Sexual : 0.001 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.999 -

un balazo en la concha y listo, no hablas mas de mi

Sexual : 0.0 - Física : 0.999 - Económica : 0.0 - Psicológica : 0.001 - Simbólica : 0.0 -

prendo fuego tu casa con un bidon con vos adentro

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

te van a encontrar carbonizada en el rio

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.999 -

voy a cortarte en pedazos con mi machete para que aprendas

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

mira este cuchillo lo voy a meter en tu garganta

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

despedite de tu familia porque ya estoy yendo a matarte

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

a mi me van a meter en la carcel pero antes vas a morir

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

una bala en la cabeza para vos y luego una para mi

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

--Psicológica

Vos tenes la culpa de lo te voy a hacer si te vuelvo a ver hija de puta

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

Sos una mierda malcogida no te quiere nadie

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

Si no volves conmigo no vas a estar con nadie

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

No vales nada, te mereces que te trate como a una mierda

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

Parece que le haces mas caso a tus amigas que a mi

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

¿Por qué no te matas y te ahorras el sufrimiento de vivir?

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

que asco que das, quien te va a coger así?

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

que idea horrible que tuviste, nunca se te ocurre algo bueno

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

los dos sabemos que esto no te sale, mejor ni lo intentes porque la vas a cagar

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

si vos sos así de estupida para que te esforzas?

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.058 - Simbólica : 0.942 -

enferma mogólica, te mereces lo que te voy a hacer

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

con quien carajo estas hablando a esta hora?

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

te voy a hacer la vida imposible hija de puta

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

suicidate, sos una basura

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

te vas a quedar sola toda tu vida

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

te deseo una muerte dolorosa hija de puta

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

--Económica

Negra de mierda deja de comprarte ropa con mi plata

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

Si volves a llegar tarde no te doy mas plata

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

Puta en casa mando yo sino queres te vas sin nada a la calle

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

queres volver a mi casa a dormir empeza a responderme antes

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

quien carajo te crees que sos para comprarte eso sin mi permiso

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

otra top mas? yo no te doy un centavo mas

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

listo, no te voy a dar nada porque te vas a ir a coger con un negro

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

y a vos quien te dio permiso de usar mi tarjeta putita?

Sexual : 0.0 - Física : 0.0 - Económica : 0.608 - Psicológica : 0.0 - Simbólica : 0.392 -

hija de puta para que mierda nos seguis endeudando

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

este es mi hogar, y aca se hace lo que yo digo

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

yo te doy todo y vos te rascas la concha, mantenida de mierda

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

vos no trabajas, estas en casa cuidando al nene nada mas

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

seguis usando la plata para tus boludeces

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

pedazo de negra parasita, rompes todo y yo tengo que pagarlo

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

no me voy a bancar que sigas usando la tarjeta sin mi autorizacion

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

conchuda agradece que seguis viviendo bajo mi techo

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

no te voy a mandar mas la cuota para el nene

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

mejor que no sepas cuanto tengo en la cuenta porque me lo vas a sacar

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

cuando cobres dame la plata que la voy a manejar yo

Sexual : 0.0 - Física : 0.0 - Económica : 1.0 - Psicológica : 0.0 - Simbólica : 0.0 -

--Sexual

Me vas a chupar la pija aunque no quieras

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

Hacete coger la concha de tu hermana

Sexual : 0.001 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.01 - Simbólica : 0.989 -

Puta de mierda te voy a violar

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

soy tu marido y te toco todo lo que quiero

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

voy a entrar a tu casa para reventarte la concha

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

que ganas de secuestrarte y romperte el orto sin que nadie se entere

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

ayer te segui mientras iba al trabajo y desee hacerte mia

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

si te llego a encontrar en la calle te violo

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

te vamos a coger entre todos, cuando terminemos te tiramos a una zanja

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

sino te sentas en mi pija te degollo

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

o me mandas una foto tuya desnuda o te voy a buscar

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

podemos garchar por las buenas o por las malas, vos elegis

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

aunque no quieras te voy a hacer mia esta noche

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

cuando salgas de tu casa te voy a agarrar y te la meto bien adentro

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

jajaj aunque llames a la policia te voy a dejar embarazada hoy

Sexual : 0.996 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.003 -

cuidate porque si te cruzo te reviento a pijazos

Sexual : 0.0 - Física : 1.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

aunque tengas novio te voy a garchar hoy

Sexual : 0.0 - Física : 0.973 - Económica : 0.0 - Psicológica : 0.027 - Simbólica : 0.0 -

soy tu novio, si yo quiero ponerte en 4 no me lo vas a impedir

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.07 - Simbólica : 0.93 -

te amo, quiero que seas mia hija de puta

Sexual : 0.266 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.002 - Simbólica : 0.733 -

zorra malparida, ya te voy a llevar a mi cama para cogerte

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

te voy a callar metiendote la pija en la boca

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

--Simbólica

como sos una mina a vos no te da la cabeza para entender

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.861 - Simbólica : 0.138 -

que rapida que sos para conseguir un macho hija de puta

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

Tendrias que estar limpiando porque sos mujer

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

todas las mujeres son igual de putas

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

ustedes se hacen las santas y despues se bajan la bombacha rapidito

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

claro sos mujer, no te sale tan facil

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

el rol tuyo es cocinar y limpiar la casa

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

apenas ven una pija ya quieren chuparla

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

a mi hija la voy a criar para que sea una buena ama de casa

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

no jugues futbol, ese no es un deporte de minas

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.816 - Simbólica : 0.184 -

deberias ser mas obediente conmigo, soy tu marido

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

tu tarea como mujer es satisfacerme

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

yo pense que eras diferente pero sos una facil

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

no engordes mas, tenes que ser flaquita o ningun hombre te va a mirar

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

que emocional que sos, deja de mariconear

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

dale te abris de piernas todo el tiempo

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

anda a limpiar antes de que tu marido llegue

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

uy tenias que ser vieja

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 1.0 - Simbólica : 0.0 -

eres muy inteligente para ser mujer

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

aprende a comportarte como señorita

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

sos una cualquiera, le mostraste las tetas a todo el barrio

Sexual : 0.369 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.63 -

vos solo sabes chupar la pija, te sale natural

Sexual : 1.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 0.0 -

cuando vienen mis amigos no tenes que opinar, anda a la cocina

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

y la comida? vos tenes mas facilidad porque sos mina

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

todas las minas son igual de jodidas como vos

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

te haces la machita pero tenes concha igual que todas

Sexual : 0.0 - Física : 0.0 - Económica : 0.0 - Psicológica : 0.0 - Simbólica : 1.0 -

### Entrenamiento 3: con categoría de Indefinido

Estos entrenamientos se hicieron sin transformers con el propósito de comparar el modelo con y sin la categoría de indefinido)

| Score categoría | |
| --- | --- |
| Sexual | 67 |
| Física | 81 |
| Económica | 77 |
| Psicológica | 70 |
| Simbólica | 60 |
| Indefinido | 51 |

### Entrenamiento 3: sin categoría de Indefinido

Estos entrenamientos se hicieron sin transformers con el propósito de comparar el modelo con y sin la categoría de indefinido)

| Score categoría | |
| --- | --- |
| Sexual | 66 |
| Física | 80 |
| Económica | 81 |
| Psicológica | 68 |
| Simbólica | 68 |

### Entrenamiento 3: con funcionalidad de proyectos de spacy

Capaz hay que chequear qué configuraciones se pueden cambiar, está prácticamente con el default. Creo que el usar o no projects no hace la diferencia en el entrenamiento pero nos va a ayudar a organizar la información o hacer más fácil la puesta en producción de los modelos.

(sin indefinido)

| Score categoría | |
| --- | --- |
| Sexual | 75 |
| Física | 81 |
| Económica | 75 |
| Psicológica | 67 |
| Simbólica | 65 |

### Extracción de Reddit

Empezamos con la creación de un módulo de exploración del dataset de reddit para lograr aumentar el número de ejemplos

Por cada archivo de aproximadamente 4GB se lee cada post y se detecta si está en español. (Detecta bastante bien, aunque algunas frases en inglés pasan) de cada 1GB aproximadamente 1MB es extraído. Por cada MB extraído en español aproximadamente 50 KB son detectados como violentos.

Al extraer hay problemas con algunos textos en inglés que pasan igual o algunos que consisten en la misma palabra siendo repetida una y otra vez.

Se hicieron algunas pruebas para mejorar la performance del análisis de los textos con el modelo de detección de violencia binario entrenado con transformers.

Usando la función docs = nlp.pipe(textList) para procesar los datos en batches.

Si bien capaz no se encuentran muchos datos útiles para el entrenamiento en el dataset de reddit, es muy útil para encontrar falencias en el modelo de detección de violencia binario. Por ahora lo que más se observa son algunos falsos positivos.

Por ejemplo en posts como: (muchos de estos se repiten, parecen ser ‘copypastas’ comunes que hay en reddit o comentarios de bots)

r e t a r d

b r a v o

l - o - l - a , lola

la mafia .

les miserables

no mames . . . por que se queda así mexicai ? ? ? ? truenan demasiados ccohetes o que onda ?

no homo

la bamba

re - re - re - re - re - re - re - re - re - re - re - repos

re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - re - repost

madre de dios ! ! és el pollo diablo ! !

que se lo ponga los pantalones vaqueros , resista al fuego cabrón . no mames y no haga más pendejadas .

a chinchilla

También se ven muchos posts que dicen ‘me vale verga’, ‘a tomar por culo’, ‘pinche pendejo’ que se marcan como violentos. No se si eso sería incorrecto, creo que la diferencia viene en que el modelo está entrenado con datos de argentina.

Trate de revisar un archivo a mano pero lo más rápido sería usar label studio (lo que si en lugar de violento , no violento, marcar los violentos, y hacer directamente skip a aquellas frases sin sentido o que se repiten varias veces, para no hacer a los ‘no violentos’ muy parecidos entre sí.

quedaron 10701 frases detectadas como ‘violentas’ (muchas de esas de forma incorrecta)

**Se decidió que el dataset de reddit se va a usar para mejorar el modelo de clasificación binario. Así si el clasificador multicategoría tiene algunas falencias (principalmente en la detección de violencia simbólica) no perjudica los resultados finales a los que se busca llegar. Es decir, donde el clasificador multicategoria no puede distinguir ‘traeme un tenedor de la cocina’ de ‘tu lugar es en la cocina’ y diría que ambos son ‘ violencia simbólica’ el clasificador binario es un buen filtro para distinguir entre ellos y solo quedarse con el material que le interesa.**

Notas:

Los modelos entrenados con transformers corren mucho más lento, por lo tanto se van a intentar entrenar modelos clásicos.

Se programaron filtrados para el dataset de reddit.

Estrategia de folding (usualmente usado para pocos datos cuando quiero entrenar con todos mis datos etiquetados): separar dataset en num de folds entrenar con todos los folds menos 1 y comparar con ese, e ir fold por fold.

-Bruno hizo una optimizacion de la extraccion de frases para que corriera más rapido.