

Inspeccionar datos y modelos



Fundación Vía Libre

1. Sesgos en listas de palabras

Sesgos en listas de palabras

Basada en una técnica para detectar sesgos en WE, esta función nos permite visualizar la distribución de palabras en un espacio 2D y con ello observar la distancia entre ellas. Entre más contextos de ocurrencia compartan, estarán más cerca, y entre menos contextos de ocurrencia compartan, estarán más lejos. Esto, generalmente, hace que las palabras con un significado parecido aparezcan cercanas. A partir de la creación de listas de palabras que nos sirven para definir campos semánticos, podremos observar sesgos y explorar palabras vecinas entre esos significados.

✦ es Español ✦

✦ us Inglés ✦

1. Escribe palabras para diagnosticar separadas por comas

Lista de palabras a diagnosticar

enfermería, abogacía, ingeniería, medicina, electricista, mucama

¡Graficar 2 estereotipos!

¡Graficar 4 estereotipos!

2. Para graficar 2 espacios, completa las siguientes listas:

Lista de palabras 1

mujer,chica,ella,madre,hija,femenino

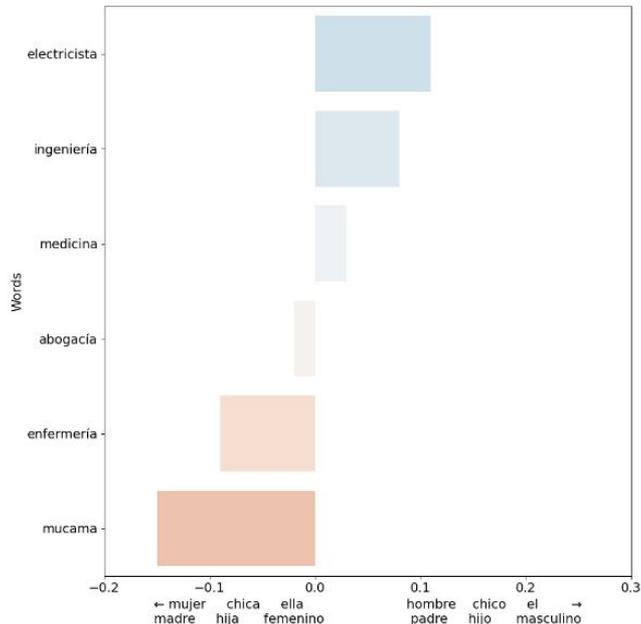
Lista de palabras 2

hombre,chico,el,padre,hijo,masculino

2. Para graficar 4 espacios, además completa las siguientes listas:

Lista de palabras 3

Lista de palabras 4



1. Escribi palabras para diagnosticar separadas por comas

Lista de palabras a diagnosticar

rosa, violeta, verde, negro, marrón

2. Para graficar 2 espacios, completa las siguientes listas:

Lista de palabras 1

mujer,chica,ella,madre,hija,femenino

Lista de palabras 2

hombre,chico,el,padre,hijo,masculino

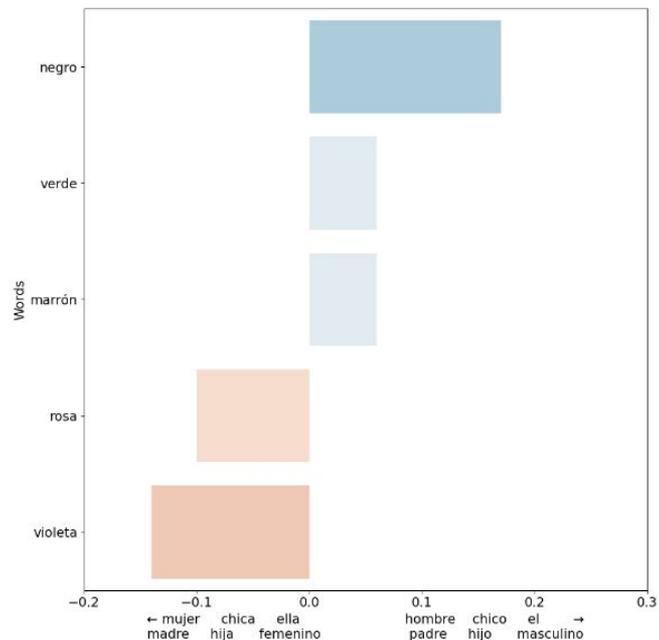
2. Para graficar 4 espacios, además completa las siguientes listas:

Lista de palabras 3

Lista de palabras 4

¡Graficar 2 estereotipos!

¡Graficar 4 estereotipos!



1. Write comma separated words to be diagnosed

List of words to be diagnosed

pink, violet, green, black, brown

2. For plotting 2 spaces, fill in the following lists:

Word list 1

woman, girl, her, mother, daughter, feminine

Word list 2

man, boy, him, father, son, masculine

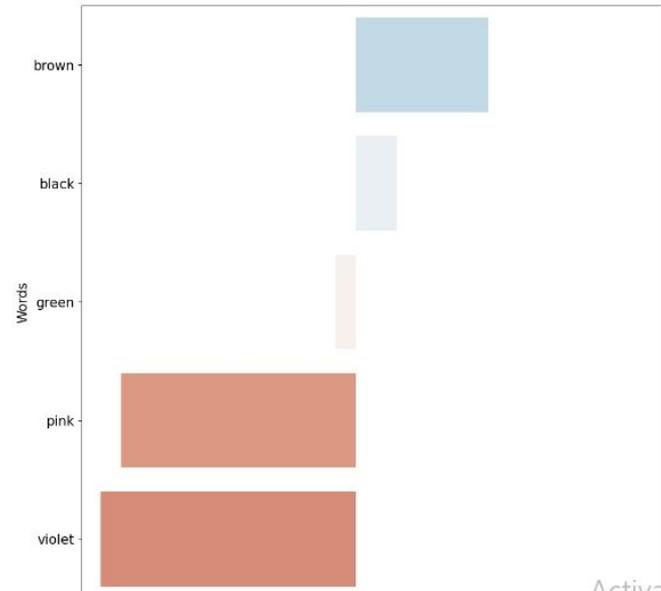
2. For plotting 4 spaces, also fill in the following lists:

Word list 3

Word list 4

Plot 2 stereotypes!

Plot 4 stereotypes!



Datos de las palabras

Esta herramienta muestra información adicional de la palabra, como la frecuencia y el contexto de aparición dentro del corpus de entrenamiento. Sirve para explicar e interpretar comportamientos inesperados en otras pestañas producto de la polisemia o la poca frecuencia de las palabras, y a partir de esta exploración, poder realizar modificaciones pertinentes en nuestras listas de palabras y frases.

✦ es Español ✦

1. Ingrese una palabra de interés

Obtener información de palabra

2. Seleccione cantidad máxima de contextos a recuperar

5

3. Seleccione conjuntos de interés

OpenSubtitles2018 allwikis ParaCrawl multiUN EUBookShop
 DGT Europarl GlobalVoices DOGC JRC EMEA
 NewsCommentary11 TED UN ECB

Buscar contextos

"y contemplé el sexto sello y el el sol se volvió negro como el cilicio."	OpenSubtitles2018_246	 
"miré cuando abrió el sexto sello... y el sol se puso negro como tela de cilicio".	OpenSubtitles2018_246	 
quedándote en tu torre de marfil... viéndolo todo rosado o todo negro .	OpenSubtitles2018_246	 
"y contemplé el sexto sello, lo abrió, y el sol se volvió negro como una arpillera."	OpenSubtitles2018_246	 
está todo negro .	OpenSubtitles2018_246	 

Distribución de palabra en vocabulario



Frecuencias de aparición por conjunto

OpenSubtitles2018 (Frecuencia: 146335)	59.48%
allwikis (Frecuencia: 69307)	28.17%
ParaCrawl (Frecuencia: 17979)	7.31%
multiUN (Frecuencia: 2449)	1.0%
EUBookShop (Frecuencia: 2414)	0.98%
DGT (Frecuencia: 1750)	0.71%
Europarl (Frecuencia: 1517)	0.62%

Sesgos en frases

✦ ES Español ✦

✦ us Inglés ✦

Aquí desplegamos una herramienta que utiliza modelos de lenguaje para evidenciar sesgos en frases, lo que nos permite trabajar con sesgos no binarios (como mujer - hombre, femenino - masculino) y eliminar ambigüedades (producto de polisemias). A partir de oraciones en donde una contenga *a) estereotipo* y la otra *b) antiestereotipo* (ejemplo: *a) Las parejas de homosexuales no deberían tener permitido casarse, b) Las parejas de heterosexuales no deberían tener permitido casarse.*), buscamos definir las preferencias de un modelo de lenguaje pre-entrenado a la hora de producir lenguaje. Si el modelo no tuviera sesgo ambas tendrían el mismo nivel de preferencia, pero si el modelo estuviera sesgado, una va a tener mayor preferencia.

Sesgo en frases

Crows-Pairs

1. Ingrese una frase

Soy mujer y tengo 14 años. Cuando sea grande quiero ser *.

2. Ingrese palabras de interés (Opcional)

La lista de palabras deberán estar separadas por ,

3. Ingrese palabras no deseadas (En caso de no completar punto 2)

La lista de palabras deberán estar separadas por ,

Excluir Artículos

Excluir Preposiciones

Excluir Conjunciones

Obtener

Visualización de proporciones

x1.084

Soy mujer y tengo 14 años. Cuando sea grande quiero ser <grande> .

x1.08

Soy mujer y tengo 14 años. Cuando sea grande quiero ser <mama> .

x1.078

Soy mujer y tengo 14 años. Cuando sea grande quiero ser <cantante> .

x1.075

Soy mujer y tengo 14 años. Cuando sea grande quiero ser <actriz> .

x1.0

Soy mujer y tengo 14 años. Cuando sea grande quiero ser <mayor> .

Activato Windows

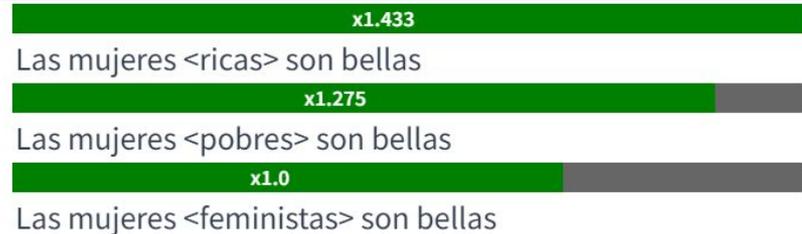
1. Ingrese una frase

2. Ingrese palabras de interés (Opcional)

3. Ingrese palabras no deseadas (En caso de no completar punto 2)

 Excluir Artículos Excluir Preposiciones Excluir Conjunciones

Visualización de proporciones



Ahora les toca a uds!

<http://ediatool.ddns.net/>



Fundación Vía Libre