



¿Qué nos dice la ciudadanía Uruguaya a través de Twitter?

Análisis de los discursos sobre la *Reforma* Educativa 2020-2024

Morella Miraballes

(CI 5 012 515 5)

Universidad de la República

Facultad de Ciencias Sociales

Educación Permanente

R aplicado al análisis cualitativo

Docente: Elina Gómez

Montevideo, Uruguay.

17 de Febrero de 2023



Resumen :

Este trabajo pretende abordar desde una perspectiva sociológica, los discursos que se han llevado a cabo por la ciudadanía Uruguaya a través de Twitter acerca de la Reforma Educativa (2020-2024).

Entendiendo, que vivimos en un mundo atravesado por las tecnologías de la información y la comunicación, en donde hay cada vez más ubicuidad entre el mundo offline y online; me propongo a través de la utilización del Big Data elaborar una tipología de las interacciones existentes en la red.

En definitiva, se pretende contribuir al campo de la sociología de la cultura, centrándome en las Ciencias Sociales Computacionales vinculado al análisis de redes sociales; de una manera reflexiva.

Palabras clave: Educación, Transformación Educativa, Redes sociales, Twitter, Uruguay, Big Data.

--

En esta línea, este trabajo se va a estructurar en cinco puntos principales:

1. Presentación de objetivos
2. Extracción de tweets (web scraping)
3. Limpieza de los datos
4. Presentación de los datos
5. Conclusión

1-Los principales objetivos son :



- Extraer mediante el paquete *rtweet* , aquellos tweets y retweets que traten la temática de interés.
- Realizar la limpieza correspondiente de los datos, eliminando aquellos tweets que no tratan de la Reforma Educativa 2020-2024, así como símbolos y palabras que puedan nublar el análisis.
- Elaborar gráficos, sobre los tweets extraídos con el fin de presentar una vista interactiva de la información. Utilizando el paquete *quanteda*.

2-Extracción de datos

En primer lugar, previo a obtener los datos se definieron aquellos perfiles y/o hashtags relevantes.

Es así que luego de indagar sobre las cuentas que más interactúan sobre la temática y que hashtags eran los más utilizados se definieron dos grupos principales.

Estos son:

A favor de la reforma:

- 1.RobertSilva1971
- 2.ANEP_Uruguay
- 3.TransformaciónEducativa

En contra de la reforma:

- 1.FenapesUruguay
- 2.adur_udelar
- 3.CeipaEnLucha
- 4.QueAvosNoTeDeLoMismo
- 5.TransformaciónDeTerror
- 6.VivaLaEducaciónPública.



Luego de esto, se extrajeron los datos de nueve días atrás (07/02/2023 - 15/02/2023) y se definió un ratio de 500 tweets por perfil y/o hashtags, dado los fines de este trabajo.

Para esto, se cargaron las librerías correspondientes:

- library(httputil)
- library(ggplot2)
- library(rtweet)
- library(dplyr)

En segundo lugar, se obtuvieron las credenciales para acceder a los tweets. Para posteriormente, conectar con *rstudio*.

```
auth=rtweet_app("AAAAAAAAAAAAAAAAAAAAAJCojwEAAAAAcuTZRAzHEuhrop3  
Q7%2Bv5YUvvTg4%3DvTml99TwbQIoW2qUs2zbP4cON8KTc3zHjxXYv8HMH3T*****  
**")  
auth_save(auth,"proyecto")  
auth_as("proyecto")
```

Una vez conectados a R, se comenzó a extraer tanto tweets como retweets de los perfiles de interés:

- RobertSilva<- search_tweets("@RobertSilva1971", n = 500, include_rts = TRUE, retryonratelimit=TRUE)
- ANEP<- search_tweets("@ANEP_Uruguay", n = 500, include_rts = TRUE, retryonratelimit=TRUE)
- FENAPES<- search_tweets("@FenapesUruguay", n = 500, include_rts = TRUE, retryonratelimit=TRUE)
- Adur<- search_tweets("@adur_udelar", n = 500, include_rts = TRUE, retryonratelimit=TRUE)
- Hastagh<- search_tweets("#TransformacionDeTerror", n = 500, include_rts = TRUE, retryonratelimit=TRUE)
- Hastagh2<- search_tweets("#VivaLaEducaciónPública", n = 500, include_rts = TRUE, retryonratelimit=TRUE)



```
- Hastagh3<- search_tweets("#Transformacion Educativa", n = 500, include_rts =  
TRUE,retryonratelimit=TRUE)  
- Ceipa<- search_tweets("@CeipaEnLucha", n = 500, include_rts =  
TRUE,retryonratelimit=TRUE)
```

Por último, todos los tweets extraídos los vamos a "unir" en un único objeto llamado tweets_V1.

```
tweets_V1<- bind_rows(RobertSilva,ANEP,Hastagh2,Hastagh,Hastagh3,Ceipa,FENAPES)  
----
```

3-Limpieza de datos

Para realizar la limpieza de datos, voy a utilizar los siguientes paquetes: *quanteda*, *readtext*, *stringr*, *dplyr*, *quanteda.textstats*, y *quanteda.textplots*.

El objetivo de la limpieza es :

1. Eliminar signos de puntuación que puedan nublar los gráficos
2. Eliminar los números y caracteres especiales de los tweets
3. Eliminar *stopwords* con el fin de lograr un análisis exhaustivo

El primer paso, consiste en cargar las librerías a utilizar:

```
library(quanteda)  
library(readtext)  
library(stringr)  
library(dplyr)  
library(quanteda.textstats)  
library(quanteda.textplots)
```

Con el fin, de proceder con la limpieza tenemos que convertir nuestro objeto *tweets* a un document-feature matrix o dfm por sus siglas en inglés.

```
dfm_tweets_v1<- quanteda::dfm(quanteda::tokens(tweets_V1$full_text,  
remove_punct = FALSE,  
remove_numbers = FALSE),  
tolower=FALSE,
```



```
verbose = TRUE) %>%  
quantda::dfm_remove(pattern =  
c(quantda::stopwords("spanish"),tolower(tweets_V1$full_text)),min_nchar=3)%>%  
quantda::dfm_trim(min_termfreq = 4)
```

La primera línea se ejecuta, y al momento de solicitarle que elimine las stopwords, el programa brinda el siguiente error:

```
Error in dfm_select(x, ..., selection = "remove") :  
argument "x" is missing, with no default
```

Asimismo, se prueba hacer la limpieza con la función *kwic* :

```
kwic = quantda::kwic(quantda::tokens(tweets_V1$full_text,  
remove_punct = TRUE,  
remove_numbers = TRUE),  
pattern = quantda::phrase(c("educacion")),  
window = 5)  
datatable(kwic)
```

De todas formas, da error nuevamente:

```
Error in datatable(kwic) : could not find function "datatable"
```

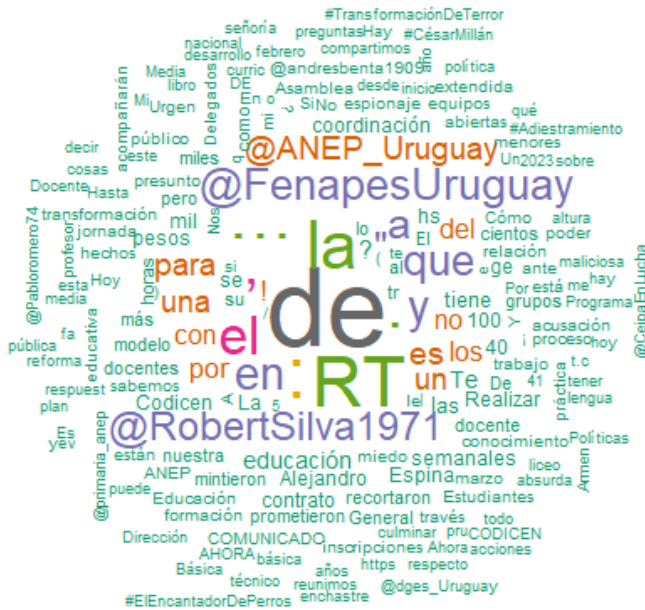
Dado que no se pudo solucionar el error previo a la entrega, se procede a ejecutar los gráficos sin la adecuada limpieza.

3-Presentación de los datos

Para la presentación de los datos, se van a utilizar nuevamente los paquetes del apartado anterior.

En primer lugar utilizando *quantda.textplots* hacemos una nube de palabras :

```
quantda.textplots::textplot_wordcloud(dfm_tweets, min.count = 2,max_words = 200,  
random.order = FALSE,colors = RColorBrewer::brewer.pal(8,"Dark2"),comparison = F)
```



Luego de esto se crear un top de las 20 palabras mas repetidas utilizando *top features*

```
(topfeatures(dfm_tweets_v1,20))
```

Teniendo en cuenta que hay errores en la limpieza, la mayoría de los datos extraídos no son relevantes, a excepción de la mención de los siguientes perfiles:

- Robert Silva
- ANEP
- Fenapes

5- Conclusión

En primer lugar, destacar, como un fenómeno social a estudiar en el futuro el hecho que desde Diciembre hasta la fecha no hubo mayores interacciones entre usuarios/as sobre la *Reforma Educativa* que se va a empezar a llevar a cabo en Marzo de este año y va a traer consigo cambios importantes en la configuración actual.

Destacar que la información extraída, nos acerca a concluir que este fenómeno social posee un ápice personalista a favor del actual presidente del CODICEN e impulsor destacado de



ésta transformación. Por otro lado, a partir de la nube de palabras, se puede visualizar que la discusión por detrás de la propuesta tiene un fuerte componente *político*.

Reconozco que hay muchas cosas que pulir para próximas instancias, sin embargo, me parece importante dejar registro del proceso de trabajo.

En esta línea, este trabajo significó un gran desafío, dado mi inexperiencia con la herramienta. De todas formas, agradezco especialmente a la docente: Elina Gómez por la paciencia, la dedicación y la respuesta casi inmediata a mis dudas.

Este trabajo, pretende ser un borrador y/o un acercamiento al trabajo con el programa y los paquetes para mi trabajo de grado en el curso de este 2023.