



Facultad de Ciencias Exactas, Ingeniería y Agrimensura
UNIVERSIDAD NACIONAL DE ROSARIO

Introducción a la Inteligencia Artificial

Trabajo Práctico Lógica Borrosa

Alumnas:

Cipullo, Inés

Sullivan, Katherine

2022

1 Problema

Es innegable el impacto que tienen las redes sociales hoy en día para definir nuestros valores, comportamientos y creencias. La realidad es que las generaciones actuales no solo nos informamos con el contenido que vemos en ellas sino que también nos formamos a través de él.

Partidarios de todas las ideas y creencias se encuentran en las redes sociales. Son demasiados los mensajes que llegan a miles y millones de personas al mismo tiempo y el potencial de eso podría ser tanto esperanzador como peligroso.

Por eso, resulta más que lógico el estar interesado en cómo es que se generan esos mensajes que nos llegan a todos, es decir, cómo es que se genera el contenido viral.

Se desea, entonces, en este trabajo poder predecir la popularidad de videos publicados en la red TikTok por cuentas nuevas, valiéndonos en distintos datos incidentes.

La **hora en la que se publica el video** y la **relevancia de la temática** en la actualidad determinarán la **cantidad de perfiles de prueba** a los que se le mostrará el video en una primera instancia.

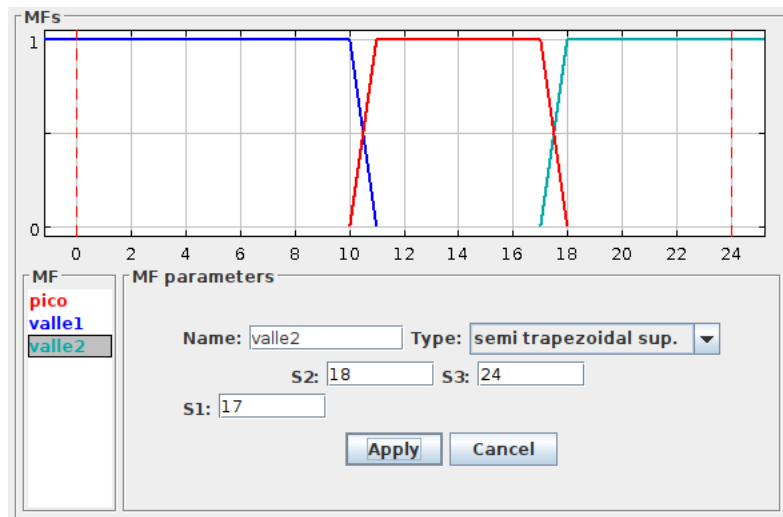
La **cantidad de perfiles de prueba** del video, la **duración del video** y el **tipo de interacción** con el video influirán enormemente en la **popularidad del video**.

2 Modelado

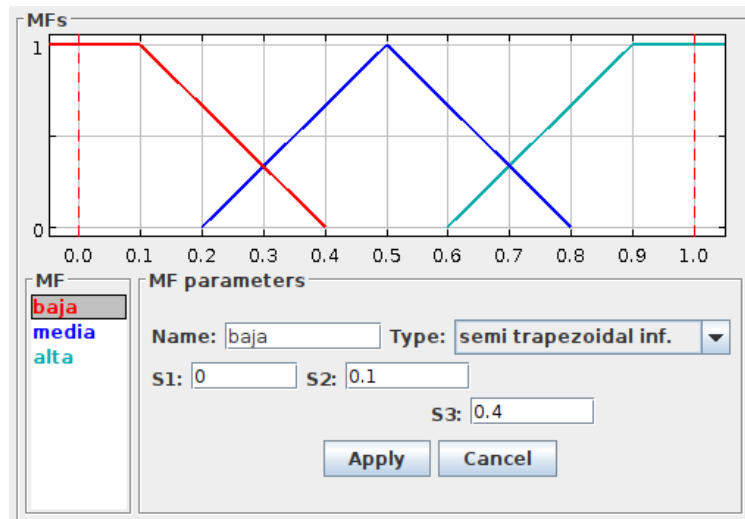
2.1 Variables lingüísticas

Se presentan las variables lingüísticas de entrada y salida de ambas etapas de inferencia con su correspondiente representación de los conjuntos borrosos asociados.

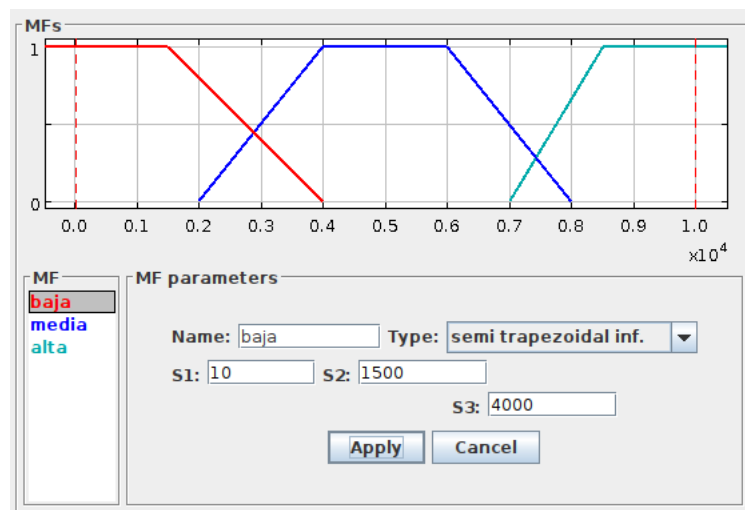
- **Hora en la que se publica el video (H):** determinada por los conjuntos borrosos hora pico y hora valle. Estos están a su vez determinados por la predicción de la cantidad de personas que estarán haciendo uso de la red social en ese momento del día, que será alta en la hora pico y baja en la hora valle. En Argentina se podría considerar hora pico de 11:00 a 17:00 horas.



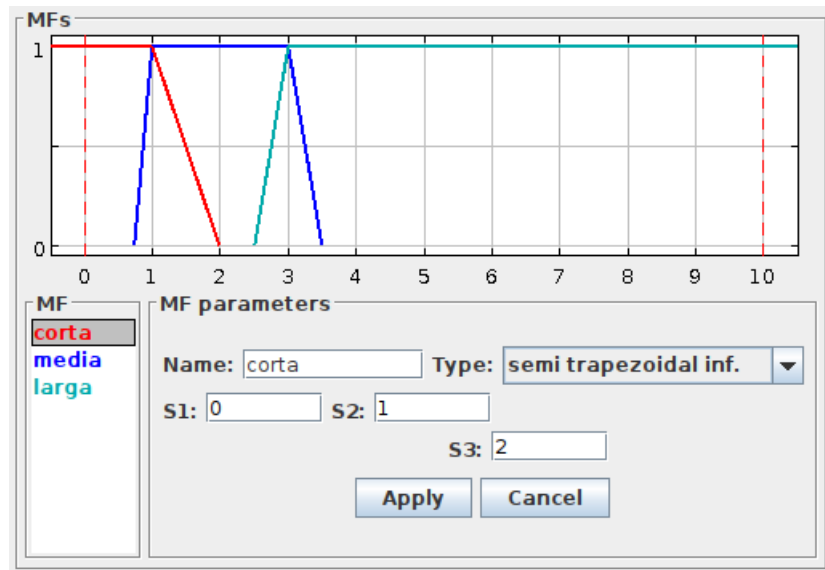
- **Relevancia de la temática en la actualidad (T):** la relevancia de una temática en particular puede ser baja, media o alta y está representada por un índice del 0 al 1.



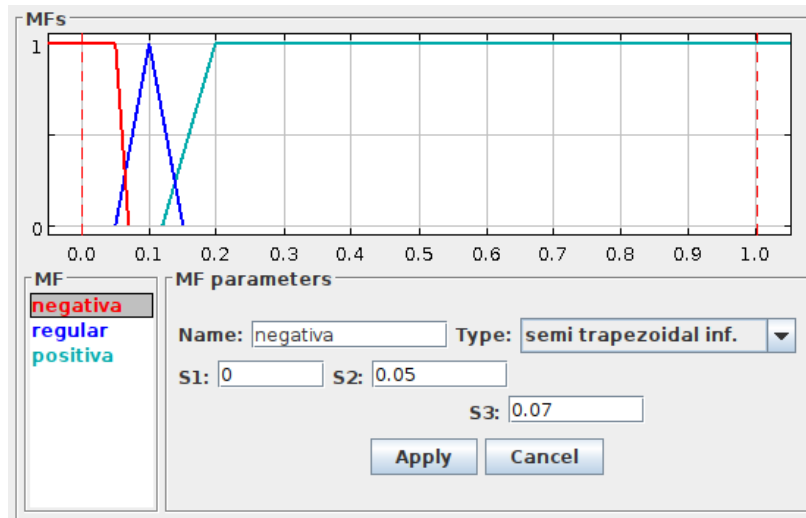
- **Cantidad de perfiles de prueba (C):** hace referencia a la cantidad de perfiles a los que se le muestra el video en una primera instancia, y varía de 10 a 10000. La cantidad de perfiles de prueba es poca si es menor de 1500, es media si está entre 4000 y 6000, y es alta si es igual o mayor a 8500.



- **Duración del video (D):** medido en cantidad de minutos, se puede decir que es corta, media o larga y va desde 0 a 10 minutos, siendo corta de 0 a 1 minuto, media de 1 a 3 minutos y larga si es mayor de 3 minutos.

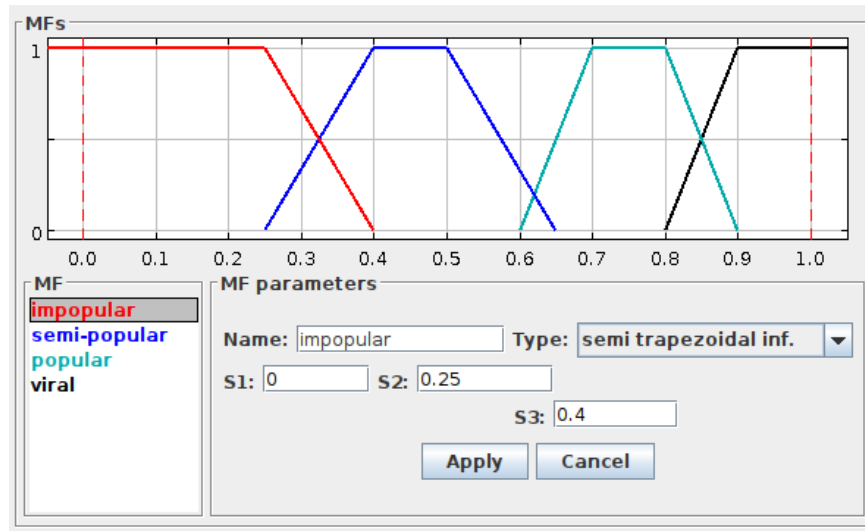


- **Tipo de interacción con el video (I):** medida como el porcentaje de perfiles de prueba que tuvieron una o más interacciones positivas sobre el video. Se considera negativa si es menor al 5%, regular si es alrededor del 10% y positiva si es más del 20%.



- **Popularidad del video (P):** para representar esta variable lingüística se

utilizan 4 conjuntos borrosos definidos a partir de un índice del 0 al 1. Estos son: impopular (de 0 a 0.25), semi-popular (de 0.4 a 0.5), popular (de 0.7 a 0.8) y viral (mayor a 0.9).



2.2 Reglas de inferencia

Planteamos las reglas de inferencia utilizando la aritmética lógica adaptada al lenguaje natural, para presentarlas de forma compacta.

- Reglas relativas a la primera etapa

R1 H hora pico y T alta o media \rightarrow C alta

R2 H hora pico o valle y T baja \rightarrow C poca

R3 H hora valle y T alta o media \rightarrow C media

- Reglas relativas a la segunda etapa

S1 C alta, D corta e I positiva \rightarrow P viral

S2 C media, D corta o media e I positiva \rightarrow P popular

S3 C alta, D media e I positiva \rightarrow P popular

S4 C media, D larga e I positiva \rightarrow P semi-popular

S5 C poca o media, D media e I regular \rightarrow P semi-popular

S6 I negativa \rightarrow P impopular

En la herramienta **fispro**, algunas de estas operaciones lógicas no se soportan y entonces se plantean, en su lugar, un conjunto de reglas por cada una de ellas. Así es como quedaron planteadas las reglas de inferencia en esta herramienta:

Rule	Active	IF relevancia_temática	AND hora_publicación	THEN cantidad_perfiles_prueba
1	✓	baja	pico	poca
2	✓	baja	valle1	poca
3	✓	baja	valle2	poca
4	✓	media	pico	alta
5	✓	media	valle1	media
6	✓	media	valle2	media
7	✓	alta	pico	alta
8	✓	alta	valle1	media
9	✓	alta	valle2	media

Figure 1: Reglas de la primera etapa

Rule	Active	IF cantidad_perfile...	AND duración_video	AND tipo_interacción	THEN popularid...
1	✓	alta	corta	positiva	viral
2	✓	alta	media	positiva	popular
3	✓	media	corta	positiva	popular
4	✓	media	media	positiva	popular
5	✓	media	larga	positiva	semi-popular
6	✓	media	media	regular	semi-popular
7	✓	baja	media	regular	semi-popular
8	✓			negativa	impopular

Figure 2: Reglas de la segunda etapa

3 Pruebas y ajustes de parámetros

3.1 Pruebas

Definimos 4 casos de prueba según consideramos relevantes para comprobar el correcto funcionamiento del modelo armado. Estos son:

- Ejemplo 1: Se busca determinar la cantidad de perfiles de prueba para un video publicado a las 20:00 sobre la guerra entre Rusia y Ucrania (un topico

considerado como muy relevante actualmente, en particular le asignamos una relevancia de 0.9).

- Ejemplo 2: Se quiere determinar la cantidad de perfiles de prueba para un video publicado a las 17:18 sobre el último partido de la Selección Argentina de Fútbol (considerado como un tema relevante, en particular le asignamos una relevancia de 0.7).
- Ejemplo 3: Según el valor crisp obtenido en el ejemplo 1, determinar la popularidad de dicho video, siendo que dura 2.8 minutos (2 minutos 48 segundos) y en general tuvo una interaccion del 18% por parte de los usuarios.
- Ejemplo 4: Según el valor crisp obtenido del ejemplo 2, determinar la popularidad de dicho video, que recibió una interaccion del 30% y dura 0.9 minutos (54 segundos). También calcular su popularidad si por el contrario durase 2.5 minutos (2 minutos 30 segundos).

3.2 Ajustes de parámetros

Para los casos de prueba que requieren inferir la cantidad de perfiles de prueba a los que mostrarle el video (ejemplos 1 y 2), utilizamos de fuzzificador la T-norma producto algebraico y para los casos de prueba que requieren calcular la popularidad que alcanzaría dicho video (ejemplos 3 y 4), utilizamos la T-norma de Lukasiewicz, diferencia acotada. Ambas operaciones son altamente restrictivas y así fueron elegidas para acercarnos lo más posible al determinismo que hay detrás de la mayoría de los algoritmos de este estilo que son usados por las redes sociales. Para defuzzificar utilizamos, en todos los casos, el método mean max.

4 Resultados

Ejemplo 1

Se busca determinar la cantidad de perfiles de prueba para un video publicado a las 20:00 sobre la guerra entre Rusia y Ucrania (un topico considerado como muy relevante actualmente, en particular le asignamos una relevancia de 0.9).

Inferencia del caso de estudio en la herramienta fispro:

Rules	relevancia_temática	hora_publicación	cantidad_perfiles_prueba
	0.9	20	5500
1			
2			
3			
4			
5			
6			
7			
8			
9			

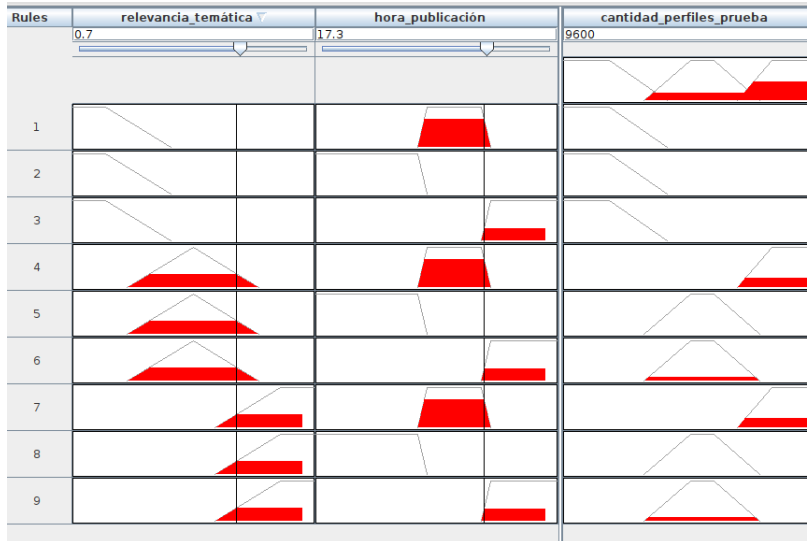
La regla que se activó en este caso fue **R3**, como se esperaba. La cantidad de perfiles de prueba que se obtuvieron fue de 5500, valor que pertenece al conjunto borroso **cantidad de perfiles de prueba media** con un grado de pertenencia de 1.

Ejemplo 2

Se quiere determinar la cantidad de perfiles de prueba para un video publicado a las 17:18 sobre el último partido de la Selección Argentina de Fútbol (considerado

como un tema relevante, en particular le asignamos una relevancia de 0.7).

Inferencia del caso de estudio en la herramienta fispro:

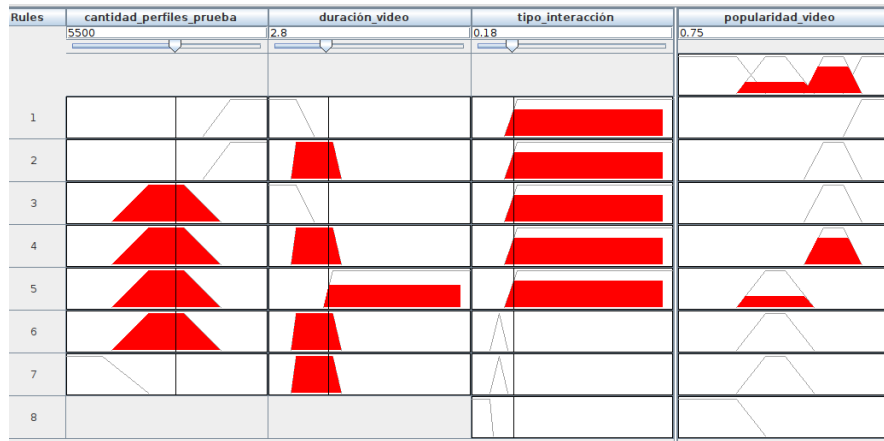


Las reglas que se activaron en este caso fueron **R1** y **R3**, pero la regla **R1** en mayor medida, como era de esperar. La cantidad de perfiles de prueba que se obtuvo fue de 9600, valor que pertenece al conjunto borroso **cantidad de perfiles de prueba alta** con un grado de pertenencia de 1.

Ejemplo 3

Según el valor crisp obtenido en el ejemplo 1, determinar la popularidad de dicho video, siendo que dura 2.8 minutos (2 minutos 48 segundos) y en general tuvo una interaccion del 18% por parte de los usuarios.

Inferencia del caso de estudio en la herramienta fispro:

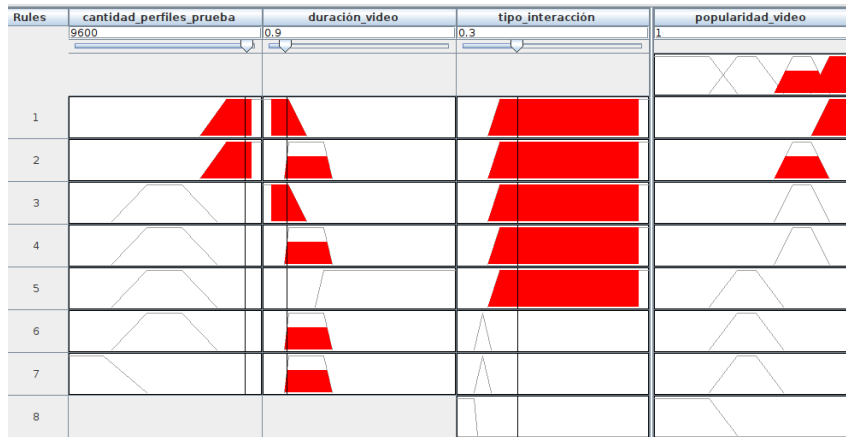


Las reglas que se activaron en este caso fueron **S2** y **S4**, pero la regla **S2** en mayor medida, como era de esperar. La popularidad del video predecida fue de 0.75, valor que pertenece al conjunto borroso **popularidad del video popular** con un grado de pertenencia de 1.

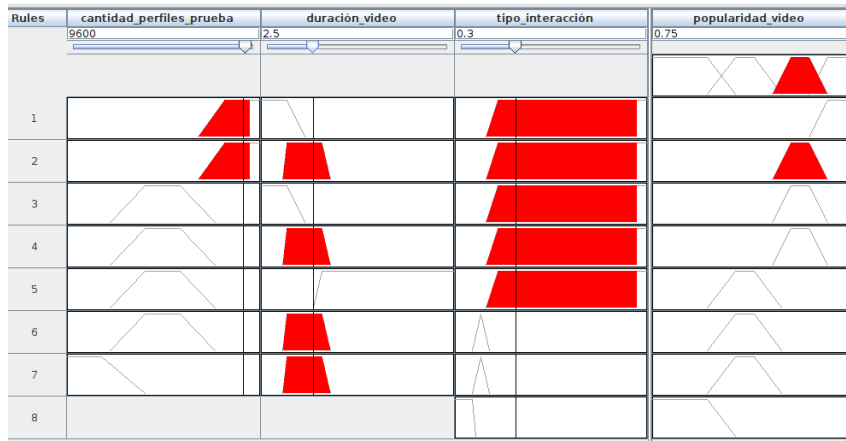
Ejemplo 4

Según el valor crisp obtenido del ejemplo 2, determinar la popularidad de dicho video, que recibió una interaccion del 30% y dura 0.9 minutos (54 segundos). También calcular su popularidad si por el contrario durase 2.5 minutos (2 minutos 30 segundos).

Inferencia del caso de estudio en la herramienta fispro:



(a) Situación 1



(b) Situación 2

En la primera situación, las reglas que se activaron fueron **S1** y **S3**, pero la regla **S1** en mayor medida, como se esperaba. La popularidad del video predecida fue de 1, valor que pertenece al conjunto borroso **popularidad del video viral** con un grado de pertenencia de 1.

En la segunda situación, la regla que se activó fue solo **S3** y en este caso la popularidad del video predecida fue de 0.75, valor que pertenece al conjunto borroso **popularidad del video popular** con un grado de pertenencia de 1.

De estos resultados en particular, se puede deducir claramente que la duración del video tiene una importante incidencia en la determinación de la popularidad de dicho video.

5 Conclusiones

La motivación para plantear este problema fue la de pensar y modelar los factores influyentes a la hora de determinar el grado de viralidad de un video. Esto devino principalmente de la lectura de distintos estudios relacionados, destacándose el caso de #TulsaFlop: A Case Study of Algorithmically-Influenced Collective Action on TikTok, que analiza estadísticamente distintos vídeos de activismo social y su performance en TikTok. Creemos que el modelo resulto propicio y nos permitió verificar nuestras intuiciones, mostrándonos como algunos pequeños cambios pueden afectar sustancialmente el desempeño de un vídeo y otros pueden no hacerlo en igual medida.