

# Entrega lógica borrosa - Introducción a la inteligencia artificial

Delfina Martín - Sebastian Morales

Junio 2021

## Descripción del problema

Una cooperativa de diarios y revistas quiere maximizar la cantidad de lectores de los productos que ofrecen. La cooperativa llevó a cabo una encuesta en el último mes que permitió recabar datos que creen pueden servir para aumentar la cantidad de lectores.

La **longitud de la nota** y la **velocidad de escritura** de le periodista afectarán el **tiempo que toma redactar** una nota para publicar.

El **tiempo de redacción** de la nota, la **reputación de le periodista** y la **popularidad de la temática** de la que se escriba influirán enormemente en la **cantidad de lectores** de esa nota.

### Longitud de la nota (L)

La longitud de la nota puede ser corta, media o larga y puede variar de 100 a 600 palabras. Una nota de hasta 200 palabras se considera corta, una de entre 267 a 433 se considera media y una de más de 500 palabras se considera larga.

### Velocidad de escritura (V)

La velocidad de escritura de un periodista puede ser lenta, media o rápida y se mide en cantidad de palabras tipeadas por minuto la cual puede variar de 10 a 70 palabras por minuto. Una persona que escribe entre 10 y 40 palabras por minuto se considera lenta, una que escribe entre 45 y 55 se considera media y una que escribe más de 60 palabras por minuto se considera rápida.

### Tiempo de redacción (T)

El tiempo de redacción puede ser corto, medio o largo y se mide en día y horas. Una nota que se escribe en a lo sumo 2 días tiene un tiempo de redacción corto, una que se escribe entre 2 días; 8 horas y 3 días; 16 horas tiene un tiempo de redacción medio mientras que una que toma más de 4 días tiene un tiempo de redacción largo.

### Reputación de quien escribe la nota (R)

La reputación de le periodista puede ser mala o buena y está indicada por un índice entre 0 y 1 siendo 0 la peor reputación y 1 la máxima. Hasta 0.3 se considera que le periodista tiene una reputación mala mientras que desde 0.7 se considera buena.

### Popularidad de la temática (P)

La popularidad de una temática en particular puede ser baja, media o alta y está representada por un índice el 0 a 1 siendo 0 la popularidad más baja y 1 la más alta. Las categorías de deportes y de política tienen popularidad de 0.9 y 1 respectivamente mientras que Rural tiene 0.1 de popularidad. Economía y finanzas tiene una popularidad de 0.85.

### Cantidad de lectores por nota (C)

La cantidad de lectores de una nota puede ser baja, medio-baja, medio-alta o alta y puede variar entre 0 y 1000 según los resultados arrojados por la encuesta. La cantidad de lectores de una nota es baja si la leen menos de 100 personas, medio-baja si la leen entre 200 y 400 personas, medio-alta si la leen entre 600 y 800 personas y alta si la leen más de 900 personas.

## Resultados de la encuesta - reglas del modelo

Las encuestas realizadas por la cooperativa arrojaron la siguiente información:

1. En relación a la longitud de la nota, la velocidad de escritura y el tiempo de redacción
  - a) Si la longitud de la nota es corta y la velocidad de escritura de quien escribe es rápida entonces el tiempo que toma redactar definitivamente la nota es corto
  - b) Si la longitud de la nota es media y la velocidad de escritura de quien escribe es media entonces el tiempo que toma redactar definitivamente la nota es medio
  - c) Si la longitud de la nota es larga y la velocidad de escritura de quien escribe es lenta entonces el tiempo que toma redactar definitivamente la nota es largo
2. En relación al tiempo de redacción, la reputación del periodista, la popularidad de la temática y la cantidad de lectores
  - a) Si el tiempo de redacción de una nota es corto, la reputación del periodista es buena y la popularidad de la temática es media o alta entonces la cantidad de lectores es alta
  - b) Si el tiempo de redacción de una nota es medio, la reputación del periodista es buena y la popularidad de la temática es media o alta entonces la cantidad de lectores es medio-alta
  - c) Si el tiempo de redacción de una nota es medio, la reputación del periodista es mala y la popularidad de la temática es alta entonces la cantidad de lectores es medio-baja
  - d) Si el tiempo de redacción de una nota es largo, la reputación del periodista es mala y la popularidad de la temática es alta entonces la cantidad de lectores es medio-baja
  - e) Si el tiempo de redacción de una nota es largo, la reputación del periodista es mala y la popularidad de la temática es baja entonces la cantidad de lectores es baja

## Enunciado

Sobre la información provista en la anterior sección se solicita:

1. Determinar variables lingüísticas de entrada y salida con sus conjuntos borrosos asociados.
2. Analizar si se puede reescribir el modelo en un número de reglas menor admitiendo el uso de más operadores lógicos.
3. Determinar el tiempo de redacción para una nota de longitud de 470 palabras y una velocidad de quien la escribe de 40 palabras por minuto. Indicar qué reglas se dispararon con qué grado de veracidad.
4. Repetir el apartado anterior para una nota de longitud de 240 palabras y una velocidad de redacción de quien la escribe de 63 palabras por minuto.
5. Según el valor crisp obtenido en el apartado 3 y considerando que la reputación de quien escribe la nota es 0 (la mínima) y P 1 (la máxima) estimar la cantidad de lectores de dicha nota. Indicar qué reglas se dispararon con qué grado de veracidad.
6. Según el valor crisp obtenido en el apartado 4 y considerando que la reputación de quien escribe la nota es 0.6 y la temática de la que se escribe es economía y finanzas, estimar la cantidad de lectores de dicha nota. Indicar qué reglas se dispararon con qué grado de veracidad.

# Resolución del ejercicio

## Ejercicio 1 - Variables lingüísticas

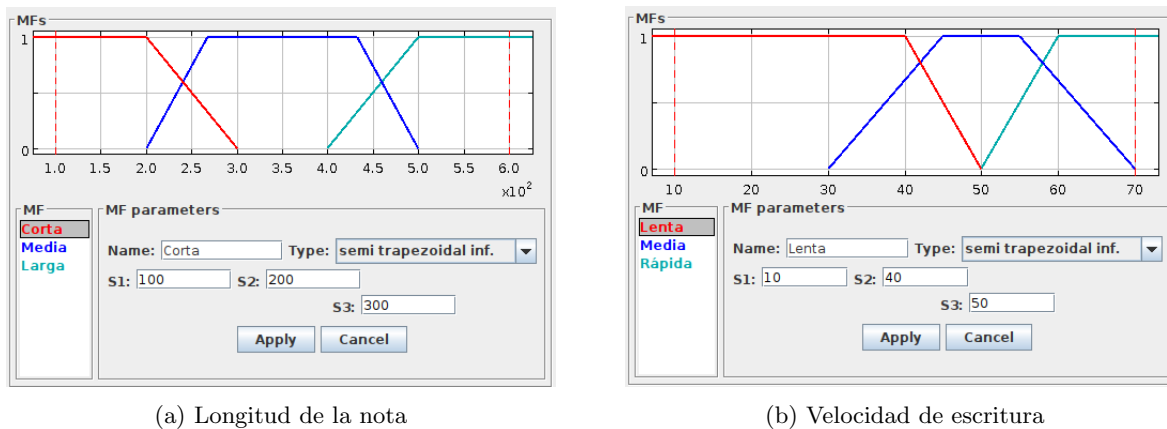
El problema se puede dividir en dos etapas. En la primera intervienen las variables lingüísticas de entrada longitud de una nota y velocidad de escritura de quien redacta la nota. Mientras que la variable de salida de la primera etapa es el tiempo que toma redactar la nota.

En la segunda etapa las variables de entrada son: el tiempo que toma redactar la nota, la reputación de quien escribe la nota y la popularidad de la temática de la que se escribe la nota. La variable de salida de la segunda etapa es la cantidad de lectores de dicha nota.

Los conjuntos borrosos asociados a cada una de las variables se listan a continuación

- longitud de una nota: corta, media y larga
- velocidad de escritura de quien redacta la nota: lenta, media y rápida
- tiempo que toma redactar la nota: corto, medio y largo
- reputación de quien escribe la nota: mala y buena
- popularidad de la temática de la que se escribe la nota: baja, media o alta
- cantidad de lectores de dicha nota: baja, medio-baja, medio-alta y alta

En las figuras (1), (2) (3) y (4) se pueden visualizar las variables de entrada y salida de cada una de las etapas.



(a) Longitud de la nota

(b) Velocidad de escritura

Figura 1: Funciones de pertenencia de los conjuntos borrosos asociados a las variables de entrada de la primera etapa

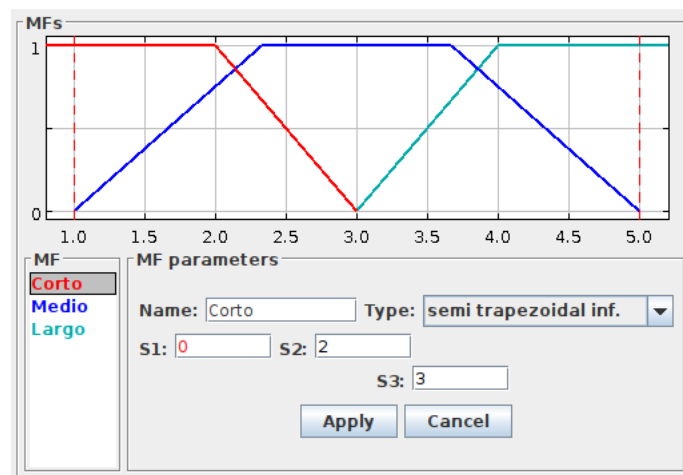
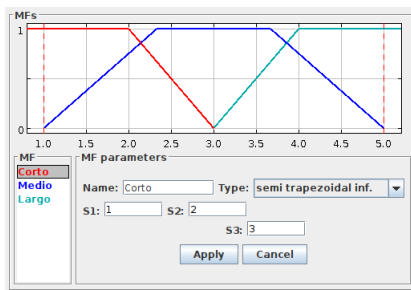
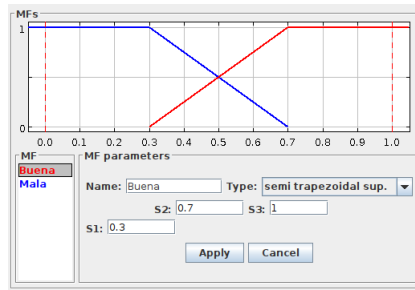


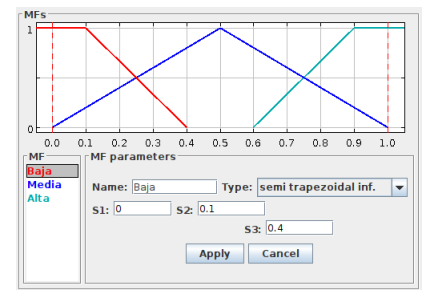
Figura 2: Tiempo que toma redactar una nota - Variable de salida de la primera etapa



(a) Tiempo de redacción



(b) Reputación del periodista



(c) Popularidad de la temática

Figura 3: Funciones de pertenencia de los conjuntos borrosos asociados a las variables de entrada de la segunda etapa

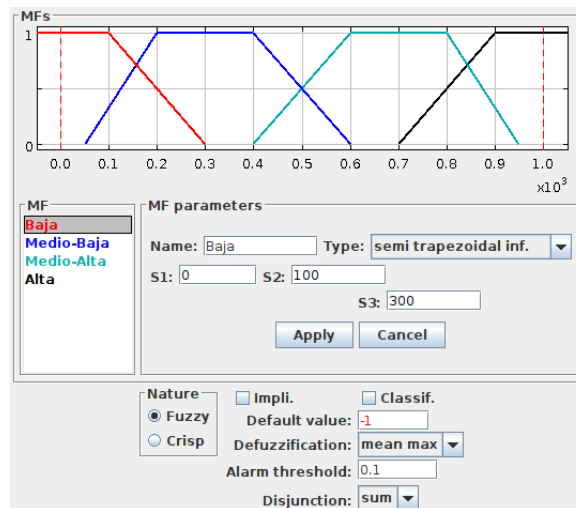


Figura 4: Cantidad de lectores por nota - Variable de salida de la segunda etapa

## Ejercicio 2 - Reglas nuevas basadas en el modelo

El listado de reglas escritas de forma más compacta haciendo uso de operadores lógicos se puede ver a continuación

- Reglas relativas a la primera etapa
  - L corta y V rápida  $\rightarrow$  T corto
  - L media y V media  $\rightarrow$  T medio
  - L larga y V lenta  $\rightarrow$  T largo
- Reglas relativas a la segunda etapa
  - T corto, R buena y P media o alta  $\rightarrow$  C alta
  - T medio, R buena y P media o alta  $\rightarrow$  C medio-alta
  - T medio o largo, R mala y P alta  $\rightarrow$  C medio-baja
  - T largo, R mala y P baja  $\rightarrow$  C baja

Como la herramienta fispro no permite agregar reglas que incluyen el operador de disyunción, reformulamos las reglas para satisfacer estas necesidades.

- Reglas relativas a la primera etapa
  - R1** L corta y V rápida  $\rightarrow$  T corto
  - R2** L media y V media  $\rightarrow$  T medio
  - R3** L larga y V lenta  $\rightarrow$  T largo
- Reglas relativas a la segunda etapa
  - S1** T corto, R buena y P media  $\rightarrow$  C alta
  - S2** T corto, R buena y P alta  $\rightarrow$  C alta
  - S3** T medio, R buena y P media  $\rightarrow$  C medio-alta
  - S4** T medio, R buena y P alta  $\rightarrow$  C medio-alta
  - S5** T medio, R mala y P alta  $\rightarrow$  C medio-baja
  - S6** T largo, R mala y P alta  $\rightarrow$  C medio-baja
  - S7** T largo, R mala y P baja  $\rightarrow$  C baja

En la figura (5) se pueden ver las reglas tal cual las enumeramos previamente.

Rule	Active	IF L	AND V	THEN T
1	<input checked="" type="checkbox"/>	Corta	Rápida	Corto
2	<input checked="" type="checkbox"/>	Media	Media	Medio
3	<input checked="" type="checkbox"/>	Larga	Lenta	Largo

Rule	Active	IF R	AND P	AND T	THEN C
1	<input checked="" type="checkbox"/>	Buena	Media	Corto	Alta
2	<input checked="" type="checkbox"/>	Buena	Alta	Corto	Alta
3	<input checked="" type="checkbox"/>	Buena	Media	Medio	Medio-Alta
4	<input checked="" type="checkbox"/>	Buena	Alta	Medio	Medio-Alta
5	<input checked="" type="checkbox"/>	Mala	Alta	Medio	Medio-Baja
6	<input checked="" type="checkbox"/>	Mala	Alta	Largo	Medio-Baja
7	<input checked="" type="checkbox"/>	Mala	Baja	Largo	Baja

Figura 5: Reglas cargadas en la herramienta fispro

### Ejercicio 3 - Inferir el tiempo de redacción

Dada una nota de 470 palabras de largo (L media, L larga) y considerando la velocidad de la persona que la redacta de 40 palabras por minuto (V lenta, V media) pudimos inferir con la herramienta fispro que la redacción de dicha nota tomará 4.85 días, es decir, 4 días y 20 horas (T medio, T largo) en completarse.

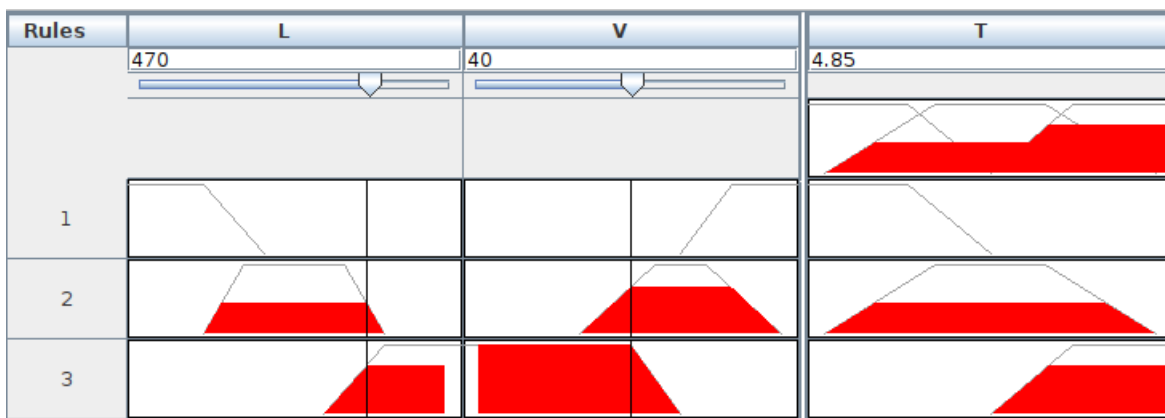


Figura 6: Inferencia del caso de estudio en la herramienta fispro

**Nota** Para inferir utilizamos de *borrosificador* la Norma mínimo y para defusificar utilizamos el método mean max.

Las reglas que se activaron con este caso de estudio son las reglas **R2** y **R3** con valores de veracidad de

Regla	Longitud de la nota	Velocidad de redacción	Veracidad según el mínimo
R2	Media - $\approx 0.5$	Media - $\approx 0.7$	Medio - 0.5
R3	Larga - $\approx 0.7$	Lenta - 1	Largo - 0.7

El tiempo de redacción obtenido, 4 días 20 horas, pertenece al conjunto borroso TiempoRedacción Largo con un grado de pertenencia de 1 y al conjunto borroso TiempoRedacción medio con un grado de pertenencia  $\approx 0.1$ . Es decir, el grado de pertenencia a TiempoRedacción largo es mayor al grado de pertenencia TiempoRedacción medio. Este resultado además, se condice con que el grado de veracidad de la regla **R3** es mayor al grado de veracidad de **R2**. Teniendo todo esto en cuenta, el valor obtenido como resultado fue el esperado.

## Ejercicio 4 - Inferir el tiempo de redacción

Dada una nota de 240 palabras de largo (L corta, L media) y considerando la velocidad de la persona que la redacta de 63 palabras por minuto (V media, V rápida) pudimos inferir con la herramienta fispro que la redacción de dicha nota tomará 1.2 días, es decir, 1 día y 5 horas (T corto, T medio) en completarse.

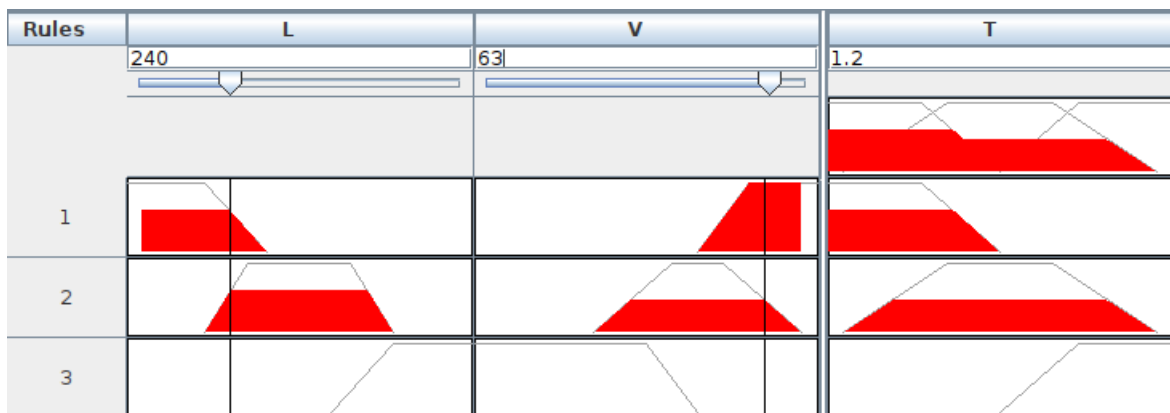


Figura 7: Inferencia del caso de estudio en la herramienta fispro

**Nota** Para inferir utilizamos de *borrosificador* la Norma mínimo y para defusificar utilizamos el método mean max.

Las reglas que se activaron con este caso de estudio son las reglas **R1** y **R2** con valores de veracidad de

Regla	Longitud de la nota	Velocidad de redacción	Veracidad según el mínimo
R1	Corta - $\approx 0.6$	Rápida - 1	Corto - $\approx 0.6$
R2	Media - $\approx 0.6$	Media - $\approx 0.45$	Medio - $\approx 0.45$

El tiempo de redacción obtenido, 1 día y 5 horas, pertenece al conjunto borroso TiempoRedacción Corto con un grado de pertenencia de 1 y al conjunto borroso TiempoRedacción Medio con un grado de pertenencia  $\approx 0.2$ . Es decir, el grado de pertenencia a TiempoRedacción medio es mayor al grado de pertenencia TiempoRedacción corto. Este resultado además, se condice con que el grado de veracidad de la regla **R1** es mayor al grado de veracidad de **R2**. Teniendo todo esto en cuenta, el valor obtenido como resultado fue el esperado.



Ejercicio 5 - Inferir la cantidad de lectores

Dado el tiempo de redacción de 4 días y 20 horas (4.85 días; T medio, T largo), 0 de reputación de quien escribe la nota (R mala) y considerando 1 de popularidad de la temática (P alta), pudimos inferir con la herramienta fispro que la cantidad de lectores de dicha nota será de 300 personas (C media-baja)

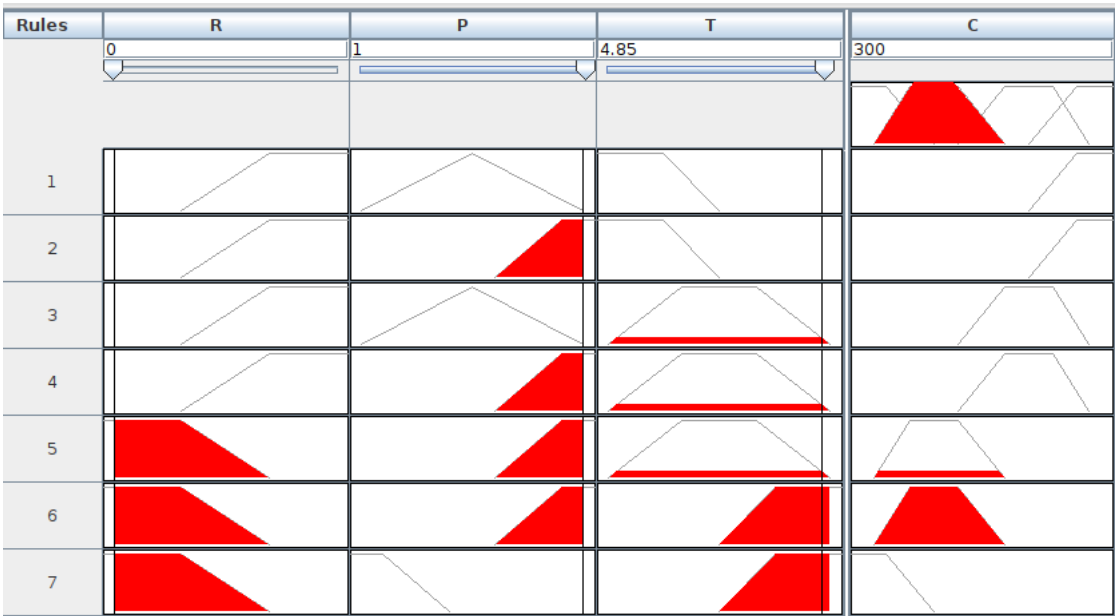


Figura 8: Inferencia del caso de estudio en la herramienta fispro

**Nota** Para inferir utilizamos de *borrosificador* la Norma Lukasciewicz y para defusificar utilizamos el método mean max.

Las reglas que se activaron con este caso de estudio son las reglas **S5** y **S6** con valores de veracidad de

Regla	Tiempo de redacción	Reputación de le periodista	Popularidad de la temática	Veracidad según Lukasciewicz
S5	Medio - $\approx 0.1$	Mala - 1	Alta - 1	Medio-baja - $\approx 0.1^1$
S6	Largo - 1	Mala - 1	Alta - 1	Medio-baja - $1^2$

La cantidad de lectores obtenida, 300 lectores, pertenece al conjunto borroso CantidadLectores Medio-baja con un grado de pertenencia de 1. Este resultado además, se condice con que el grado de veracidad de la regla **R5** y **R6** de 1. Teniendo todo esto en cuenta, el valor obtenido como resultado fue el esperado.

<sup>1</sup> $\max\{0, \max\{0, 0.1 + 1 - 1\} + 1 - 1\}$   
<sup>2</sup> $\max\{0, \max\{0, 1 + 1 - 1\} + 1 - 1\}$

## Ejercicio 6 - Inferir la cantidad de lectores

Dado el tiempo de redacción de 1 día 17 horas (1.7 días; T corto, T medio), 0.6 de reputación de quien escribe la nota (R mala, R buena) y considerando 0.85 de popularidad de la temática (P media, P alta), pudimos inferir con la herramienta fispro que la cantidad de lectores de dicha nota será de 913 personas (C media-alta, C alta)

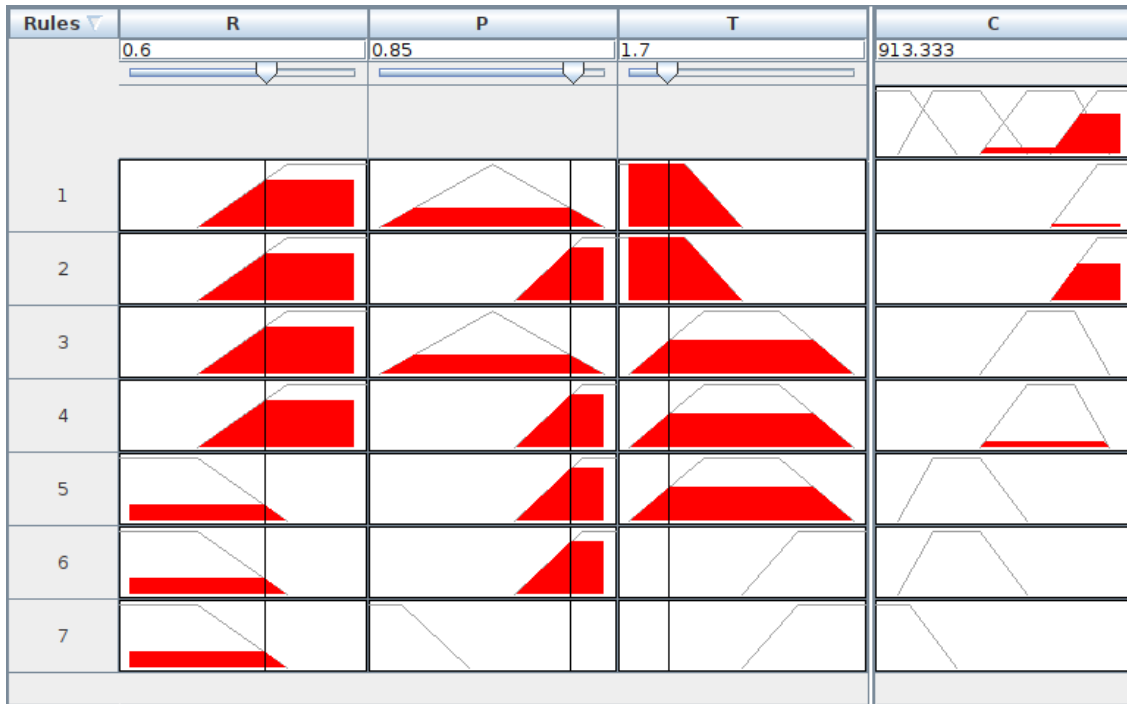


Figura 9: Inferencia del caso de estudio en la herramienta fispro

**Nota** Para inferir utilizamos de *borrosificador* la Norma Lukasciewicz y para defusificar utilizamos el método mean max.

Las reglas que se activaron con este caso de estudio son las reglas **S1**, **S2** y **S4** con valores de veracidad de

Regla	Tiempo de redacción	Reputación de le periodista	Popularidad de la temática	Veracidad según Lukasciewicz
S1	Corto - 1	Buena - $\approx 0.8$	Media - $\approx 0.25$	Alta - $\approx 0.05^3$
S2	Corto - 1	Buena - $\approx 0.8$	Alta - 1	Alta - $\approx 0.55^4$
S4	Medio - $\approx 0.5$	Buena - $\approx 0.8$	Alta - $\approx 0.75$	Medio-alta - $\approx 0^5$

La cantidad de lectores obtenida, 913 personas, pertenece al conjunto borroso CantidadLectores Medio-alto con un grado de pertenencia de  $\approx 0.3$  y al conjunto borroso CantidadLectores Alto con un grado de pertenencia de 1. Es decir, el grado de pertenencia a CantidadLectores Alto es mayor al grado de pertenencia CantidadLectores Medio. Este resultado además, se condice con que el grado de veracidad de la regla **S2** es mayor al grado de veracidad de **S1** y **S4**. Teniendo todo esto en cuenta, el valor obtenido como resultado fue el esperado.

<sup>1</sup> $\max\{0, \max\{0, 1 + 0.8 - 1\} + 0.25 - 1\}$

<sup>2</sup> $\max\{0, \max\{0, 1 + 0.8 - 1\} + 0.75 - 1\}$

<sup>3</sup> $\max\{0, \max\{0, 0.5 + 0.8 - 1\} + 0.25 - 1\}$

## Ejercicio 7 - Inferir la cantidad de lectores

Dado el tiempo de redacción de 4 días (T medio, T largo), 0.6 de reputación de quien escribe la nota (R mala, R buena) y considerando 0.5 de popularidad de la temática (P media), pudimos inferir con la herramienta fispro que la cantidad de lectores de dicha nota será de 687 personas (C media-alta, C alta).

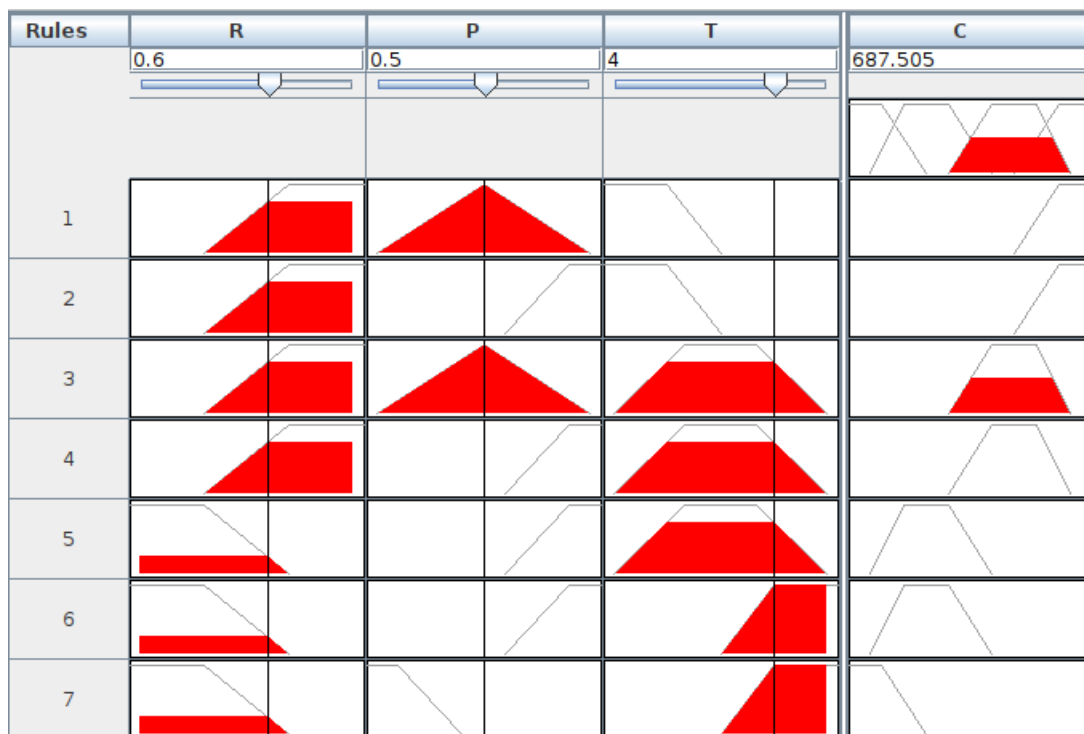


Figura 10: Inferencia del caso de estudio en la herramienta fispro

**Nota** Para inferir utilizamos de *borrosificador* la Norma Lukasciewicz y para defusificar utilizamos el método mean max.

La regla que se activó con este caso de estudio es **S3** con valor de veracidad de

Regla	Tiempo de redacción	Reputación de le periodista	Popularidad de la temática	Veracidad según Lukasciewicz
S3	Medio - $\approx 0.75$	Buena - $\approx 0.75$	Media - 1	Medio-alta - $\approx 0.5^6$

La cantidad de lectores obtenida, 687 lectores, pertenece al conjunto borroso CantidadLectores Medio-alta con un grado de pertenencia de 1, lo cual se condice con un alto grado de veracidad de la regla **S3** (considerando lo ajustado de la norma de Lukasciewicz).

<sup>6</sup> $\max\{0, \max\{0, 0.75 + 0.75 - 1\} + 1 - 1\}$

## Conclusiones

El modelo resultó completo y práctico en términos de los resultados que pudimos obtener a partir de él. Apartado por apartado los resultados fueron verificando nuestra intuición. Además, éstos se correspondieron con las medidas (borrosas) que utilizamos para ganar confianza sobre nuestro modelo.

Para los apartados 3 y 4 que pedían inferir el tiempo de redacción de una nota, usamos la norma mínimo dado que es la operación menos restrictiva. Creemos que elegir una operación que no subestime el tiempo de una nota se alinea con los ideales que podría tener una cooperativa: genera un entorno de trabajo amigable en el que los trabajadores no se sienten presionados a cumplir plazos muy cortos de redacción.

Para los apartados, 5, 6 y 7, en contraste, decidimos operar con la norma de Lukasiewicz, norma más restrictiva que el mínimo, para no sobreestimar la cantidad de lectores que comprará el ejemplar con la nota. Pensamos que tiene sentido que una cooperativa no pueda imprimir más de lo que le permite un presupuesto acotado, por lo que ajustar la estimación de cantidad de lectores va a actuar en su favor.

Como único comentario extra, en la teoría la norma de Lukasiewicz estaba solamente para dos variables de entrada mientras que en nuestro trabajo, tuvimos que aplicar reglas con 3 variables de entrada. Lo que decidimos en este caso es extender la norma para 3 variables de entrada de la siguiente manera:

$$L(L(V1, V2), V3)$$

donde  $L$  es la norma de Lukasiewicz.

Los resultados que obtuvimos a partir de la extensión de esta regla van de la mano con los valores que esperábamos.