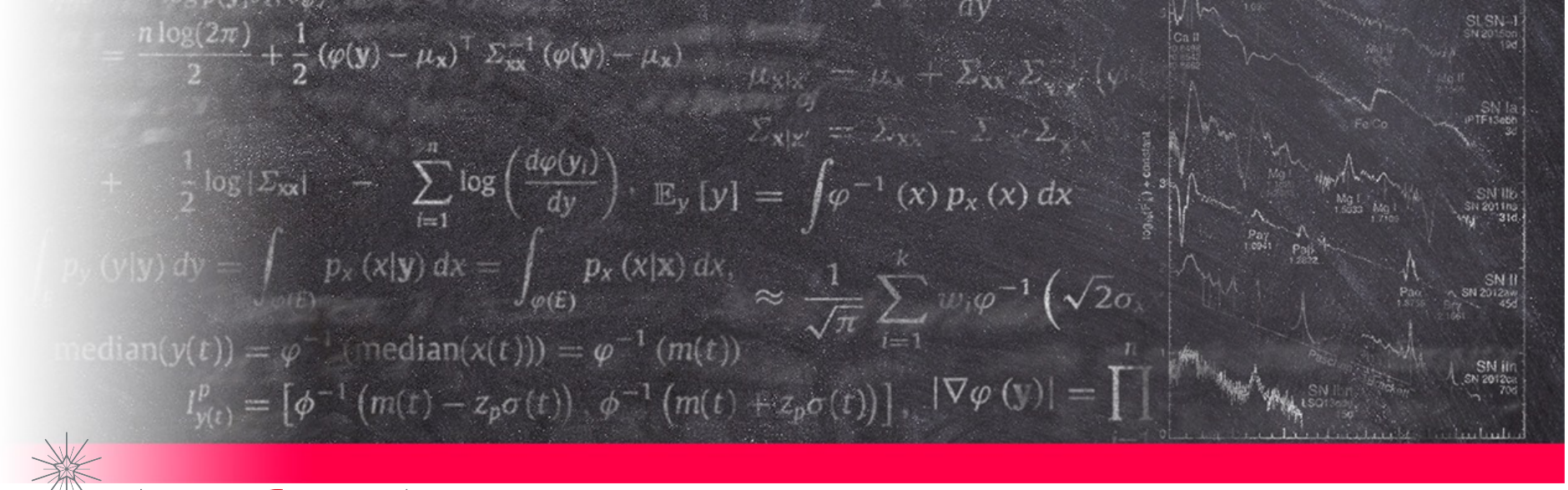


Ingeniería Matemática

FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Procesamiento de datos textuales de EthicApp

Presentación de avances

Camilo Carvajal Reyes

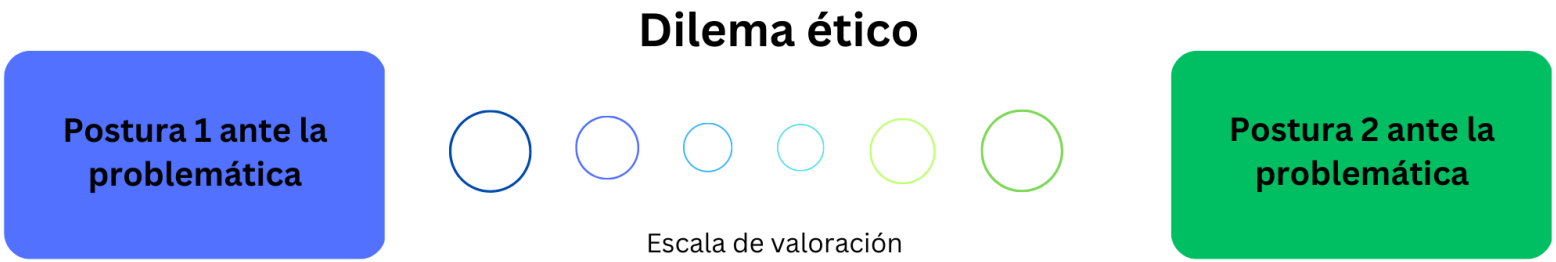
8 de junio, 2023

Índice de contenidos

Índice de contenidos

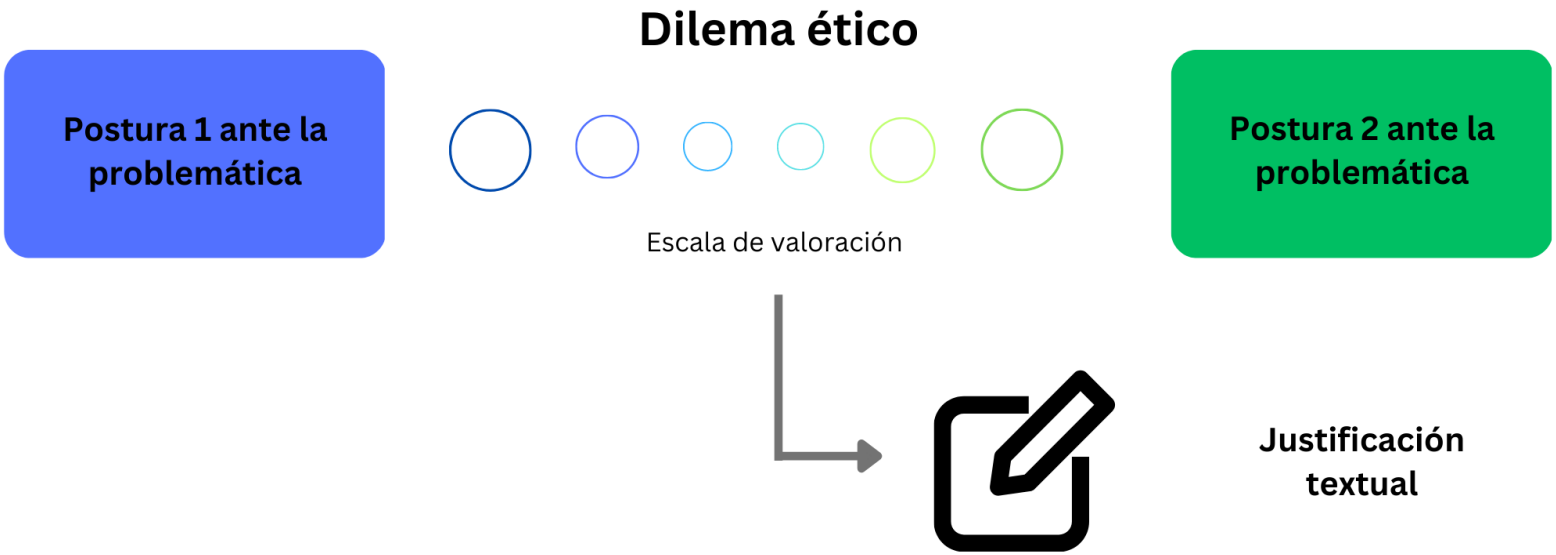
Resumen y motivación

En el marco de actividades ética en cursos iniciales de la FCFM, estudiantes evalúan en una escala de 1 a 6 las respuestas a una **dilema**, a través de la aplicación *EthicApp* [1]. En seguida escriben una **justificación** a tal decisión. Este texto puede contener **información relevante** de la decisión y su estudio es importante para los equipos docentes y el área de ética. Este análisis se dificulta por la **gran cantidad de respuestas**.



Resumen y motivación

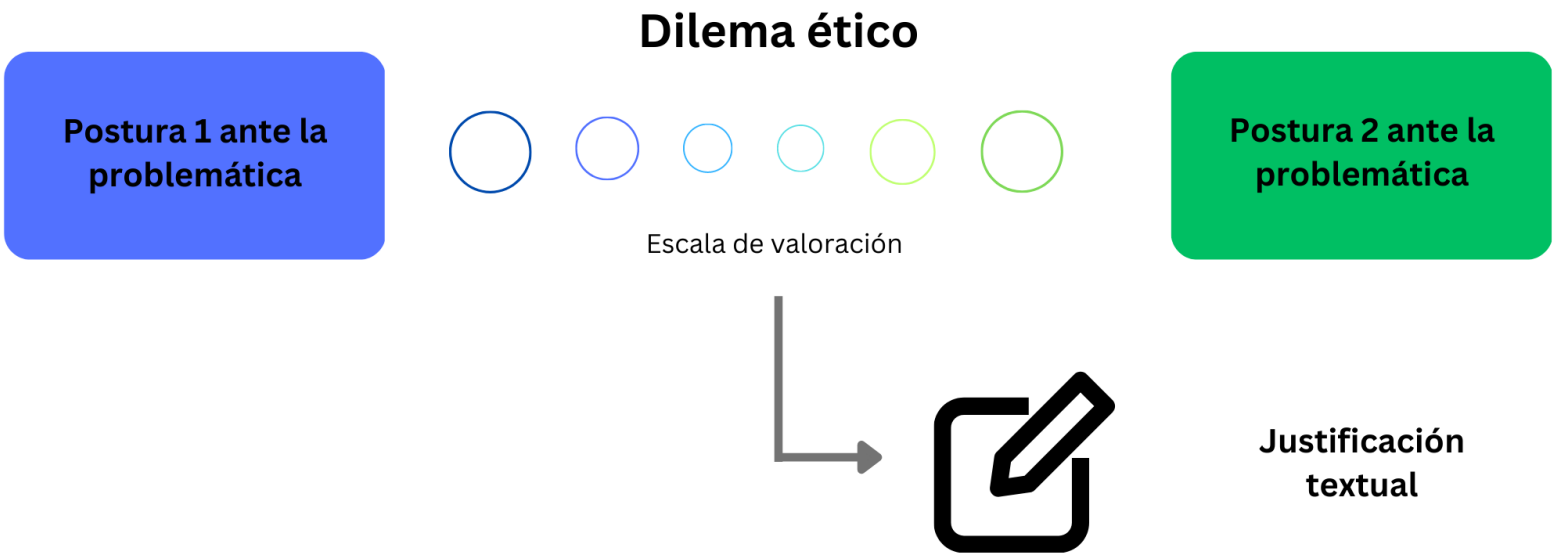
En el marco de actividades ética en cursos iniciales de la FCFM, estudiantes evalúan en una escala de 1 a 6 las respuestas a una **dilema**, a través de la aplicación *EthicApp* [1]. En seguida escriben una **justificación** a tal decisión. Este texto puede contener **información relevante** de la decisión y su estudio es importante para los equipos docentes y el área de ética. Este análisis se dificulta por la **gran cantidad de respuestas**.



Resumen y motivación

Se plantea la utilización de algoritmos de procesamiento de lenguaje natural para:

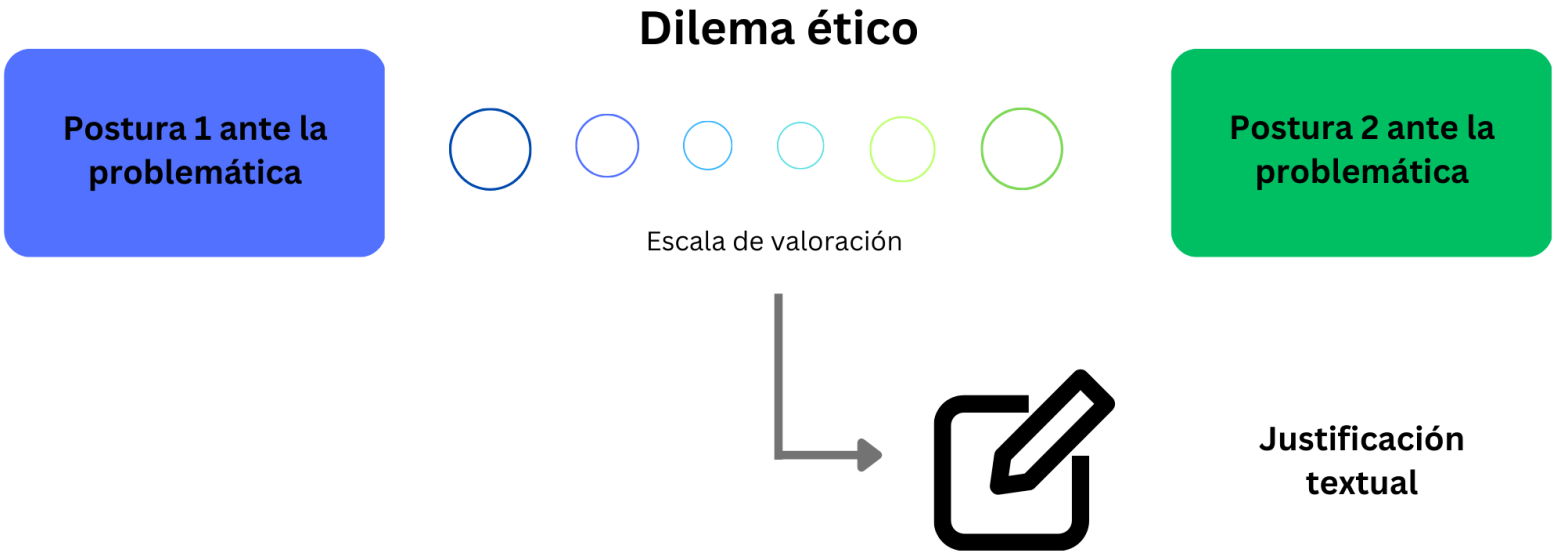
1. evaluar la progresión de competencia ética de los estudiantes con menor inversión humana,
2. estudiar las capacidades de algoritmos de texto de modelar ética.



Ejemplo de caso

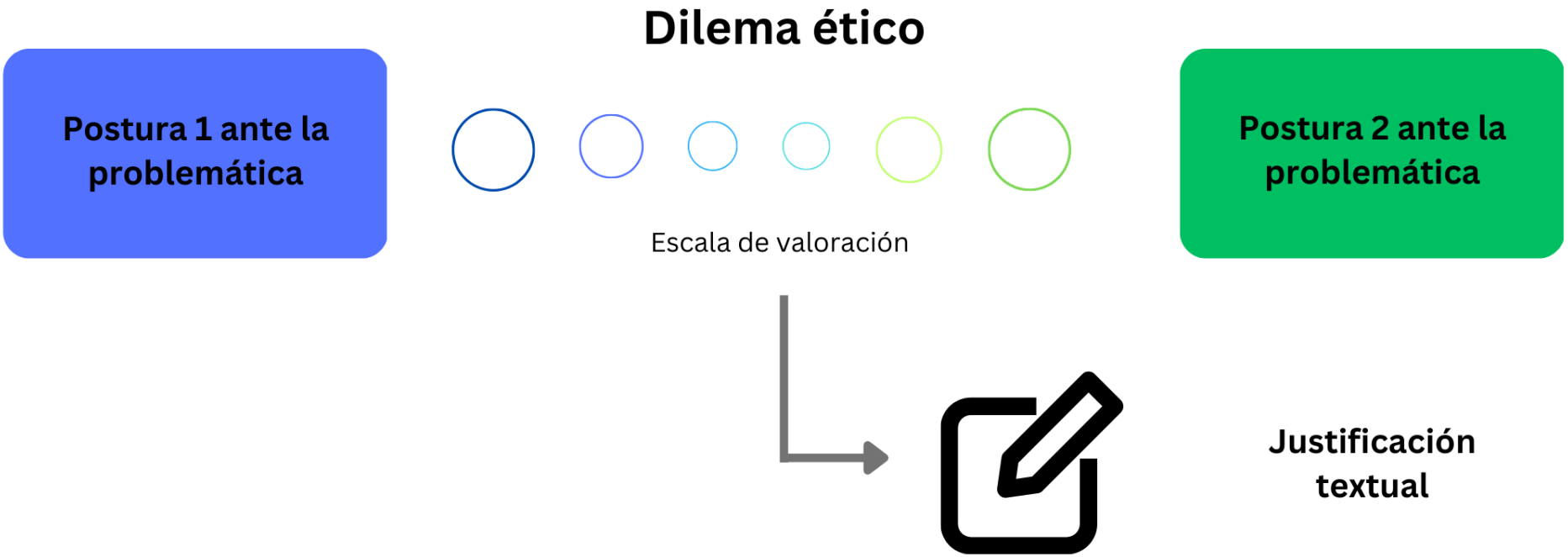
En el último control realizado **Julieta** se ve en la posibilidad de copiar una respuesta que fue compartida en el grupo de WhatsApp de su sección. Julieta en esta situación a la que se ve enfrentada en el control debiera

- 1 Usar la información del grupo de WhatsApp
- 6 No usar la información del grupo de WhatsApp



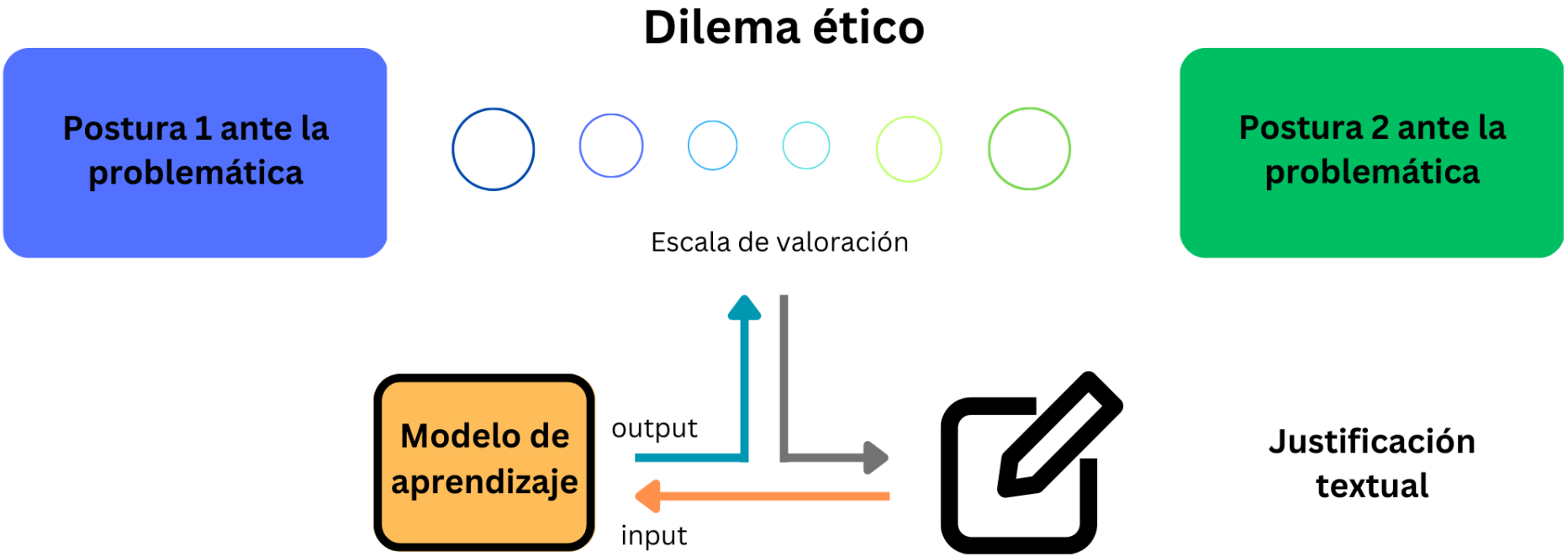
Predicción de respuestas de estudiantes

Se propone la utilización de modelos para la predicción de valoración de la problemática, usando el texto. Hacer esto con modelos interpretables nos dará una idea de que **elementos lingüísticos** se usaron para escoger tal opción. Por otro lado se plantea usar modelos profundos con fines exploratorios.



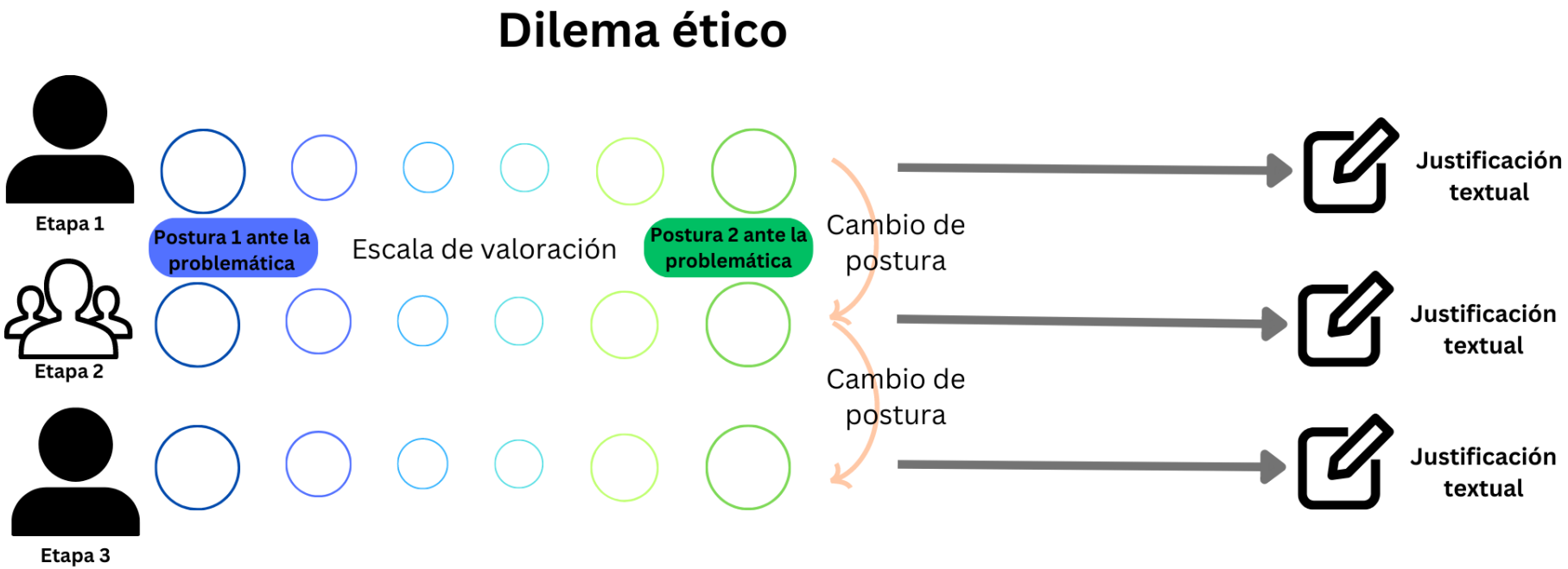
Predicción de respuestas de estudiantes

Se propone la utilización de modelos para la predicción de valoración de la problemática, usando el texto. Hacer esto con modelos interpretables nos dará una idea de que **elementos lingüísticos** se usaron para escoger tal opción. Por otro lado se plantea usar modelos profundos con fines exploratorios.



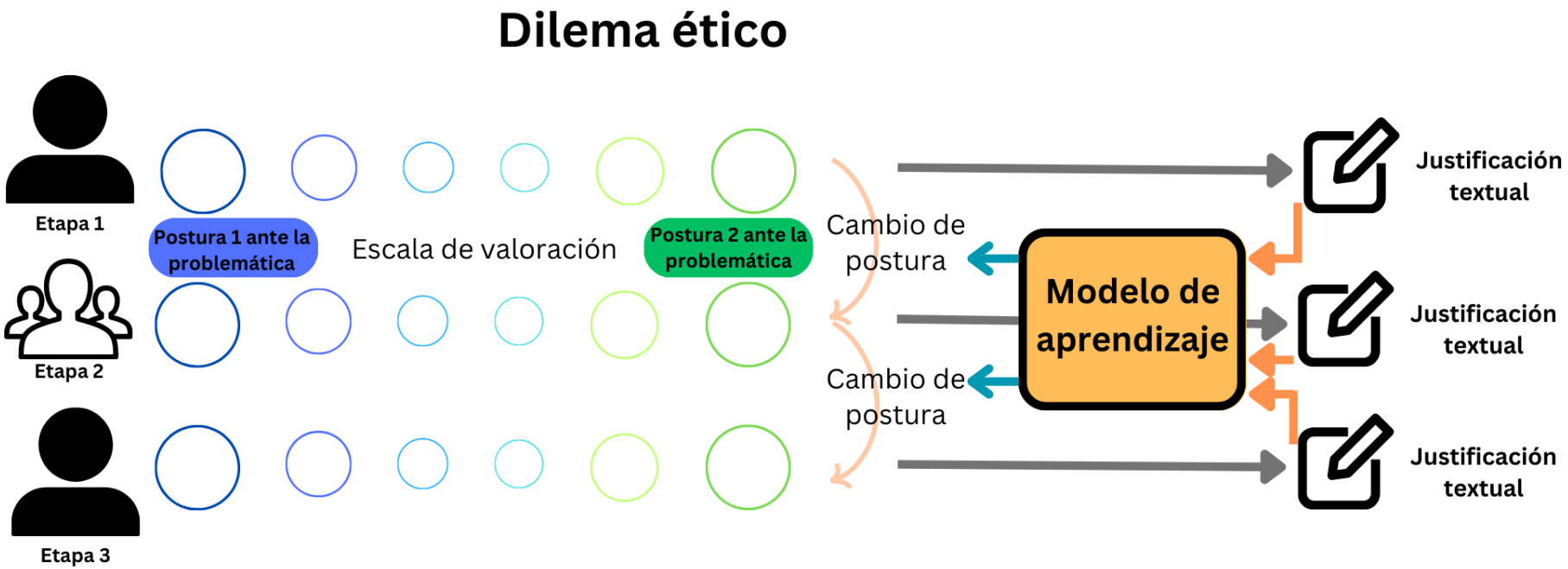
Predicción de cambio en respuesta

Similarmente, queremos usar modelos similares para predecir cambios en las valoraciones de una etapa a otra. Esto nos permite estudiar que **elementos** son **comunes** en un futuro **cambio de postura o valoración**.



Predicción de cambio en respuesta

Similarmente, queremos usar modelos similares para predecir cambios en las valoraciones de una etapa a otra. Esto nos permite estudiar que **elementos** son **comunes** en un futuro **cambio de postura o valoración**.



Índice de contenidos

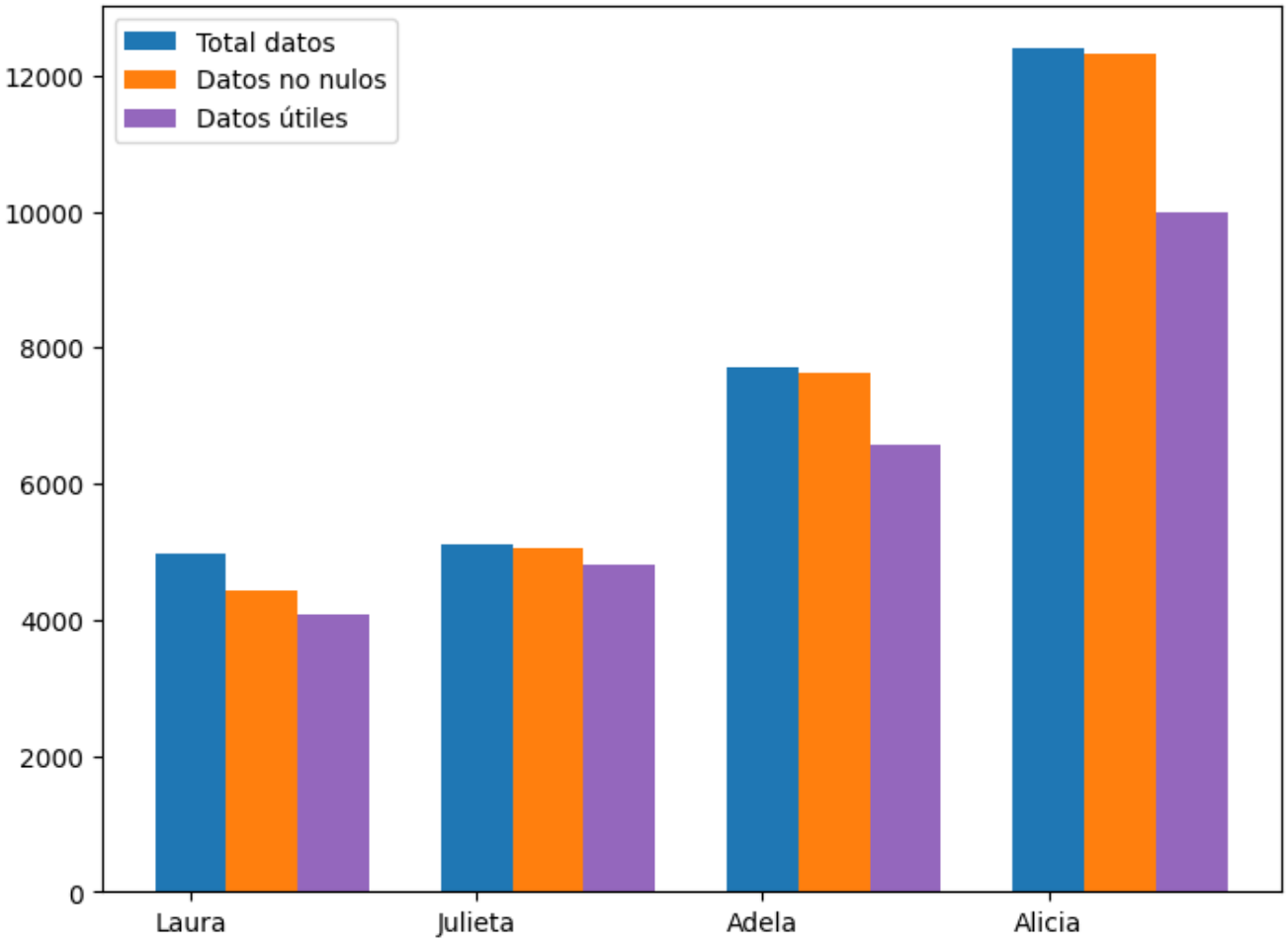
Datos textuales de EthicApp

Cantidad de datos por caso

Caso	Cursos	Cantidad estudiantes	Cantidad grupos
Caso Julieta	1	819	247
Caso Adela	1	237	142
Caso Laura	1	602	335
Caso Alicia	2	1628	549

Datos textuales de EthicApp

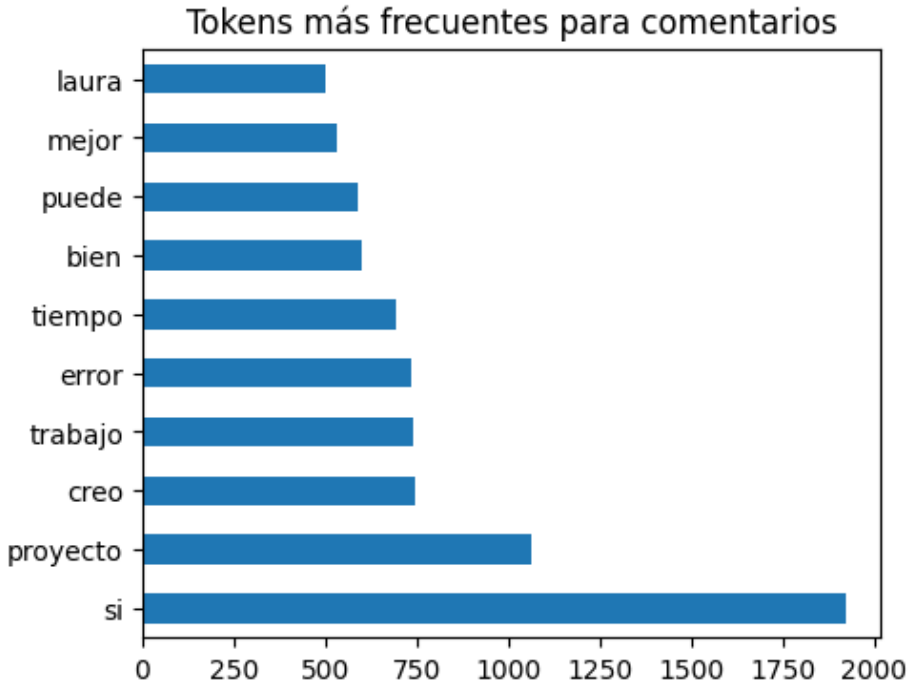
Cantidad de datos por caso



n-gramas comunes caso Laura

Preliminarmente mostramos los 1-gramas y 3-gramas más frecuentes del dataset para el caso Laura.

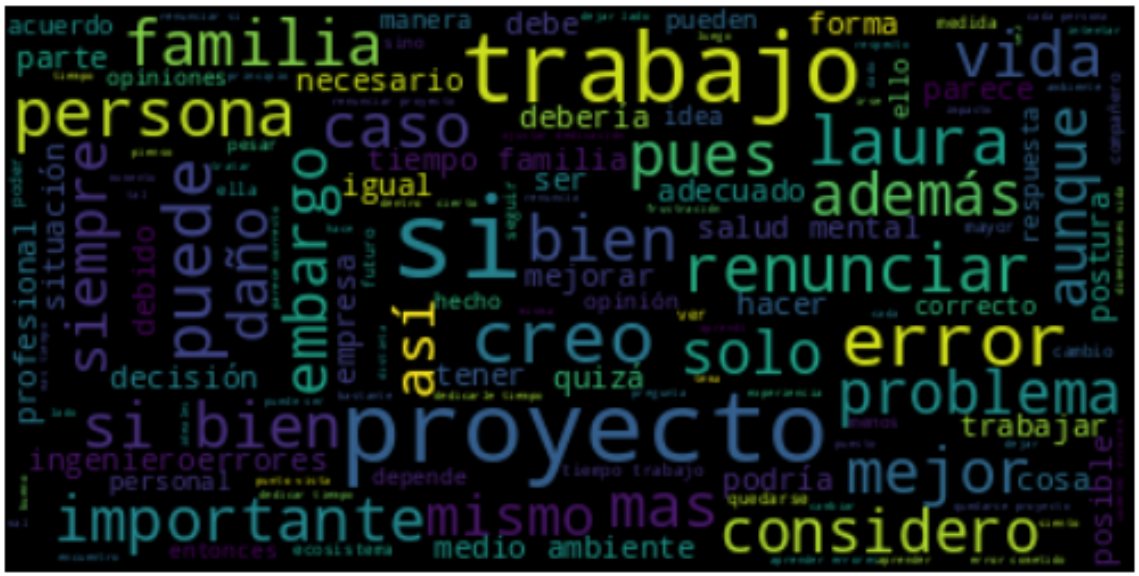
Es adecuado que Laura le dedique paulatinamente más tiempo al trabajo y su desarrollo profesional que a la familia y las otras dimensiones de su vida.



n-gramas comunes caso Laura

Preliminarmente mostramos los 1-gramas y 3-gramas más frecuentes del dataset para el caso Laura.

Es adecuado que Laura le dedique paulatinamente más tiempo al trabajo y su desarrollo profesional que a la familia y las otras dimensiones de su vida.



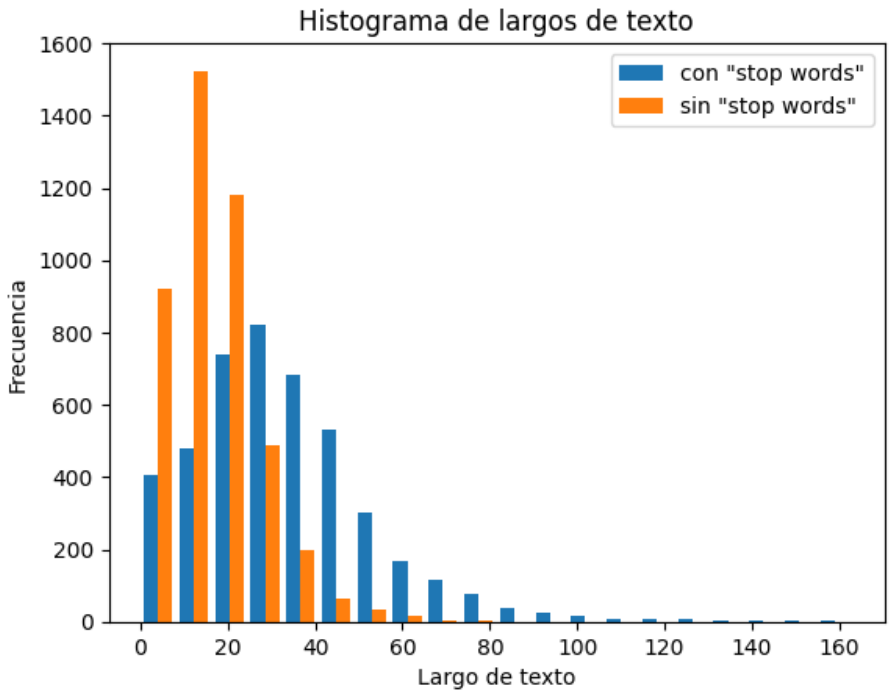
Largos de texto caso Laura

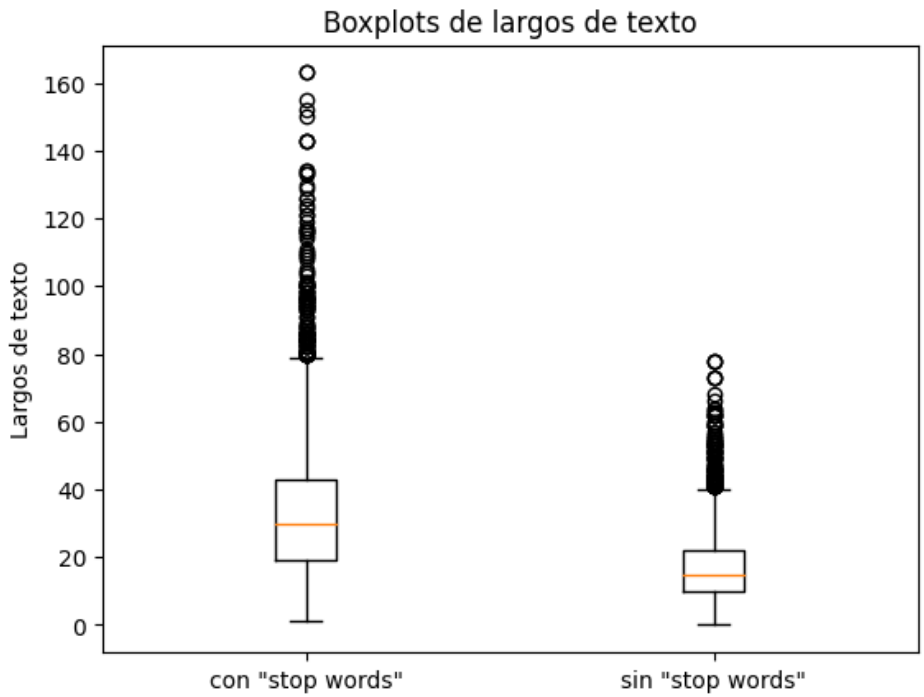
Con *stop words*

Media: 33.042
Desviación estándar: 20.793
Mediana: 30.0
Mínimo: 1
Máximo: 163

Sin *stop words*

Media: 16.658
Desviación estándar: 10.280
Mediana: 15.0
Mínimo: 0
Máximo: 78





n-gramas comunes caso Alicia

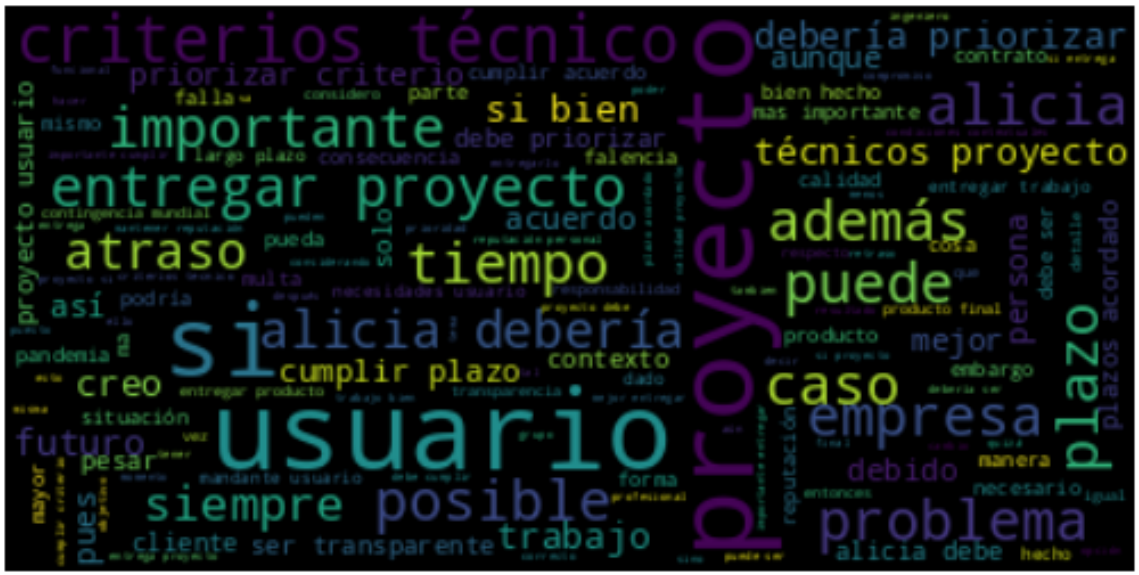
Preliminarmente mostramos los 1-gramas y 3-gramas más frecuentes del dataset para el caso Julieta.

Ante problemas y retrasos por contingencia mundial, Alicia debería priorizar los plazos o los criterios técnicos?

n-gramas comunes caso Alicia

Preliminarmente mostramos los 1-gramas y 3-gramas más frecuentes del dataset para el caso Julieta.

Ante problemas y retrasos por contingencia mundial, Alicia debería priorizar los plazos o los criterios técnicos?



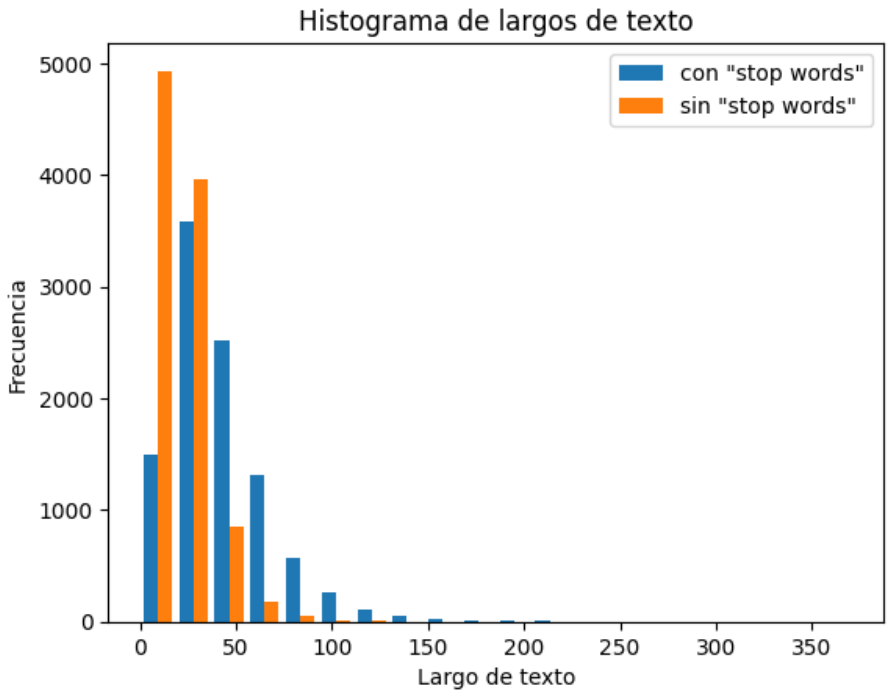
Largos de texto caso Alicia

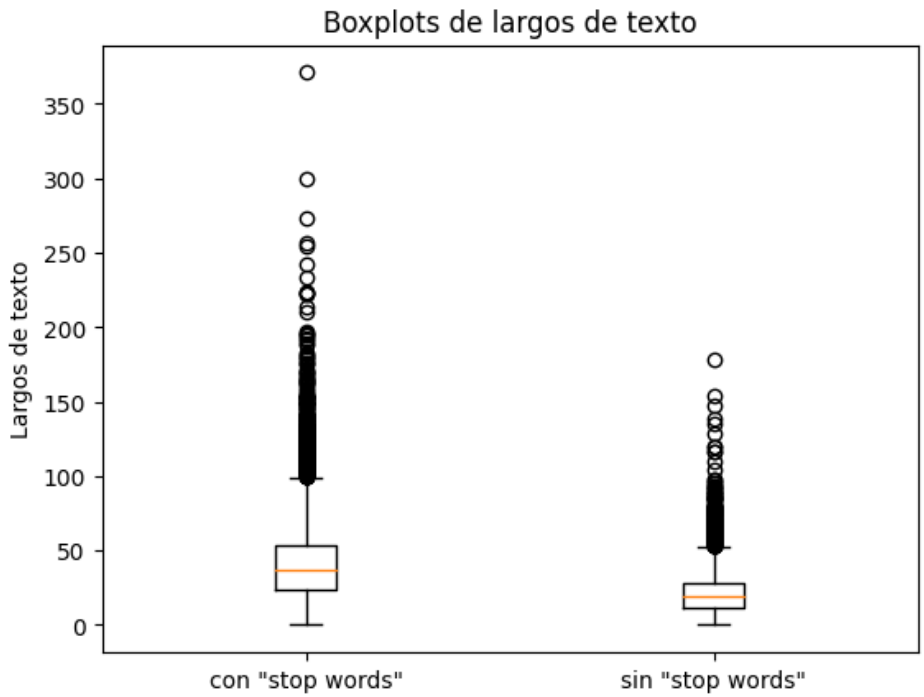
Con *stop words*

Media: 42.591
Desviación estándar: 27.023
Mediana: 37.0
Mínimo: 1
Máximo: 371

Sin *stop words*

Media: 21.601
Desviación estándar: 13.610
Mediana: 19.0
Mínimo: 0
Máximo: 178

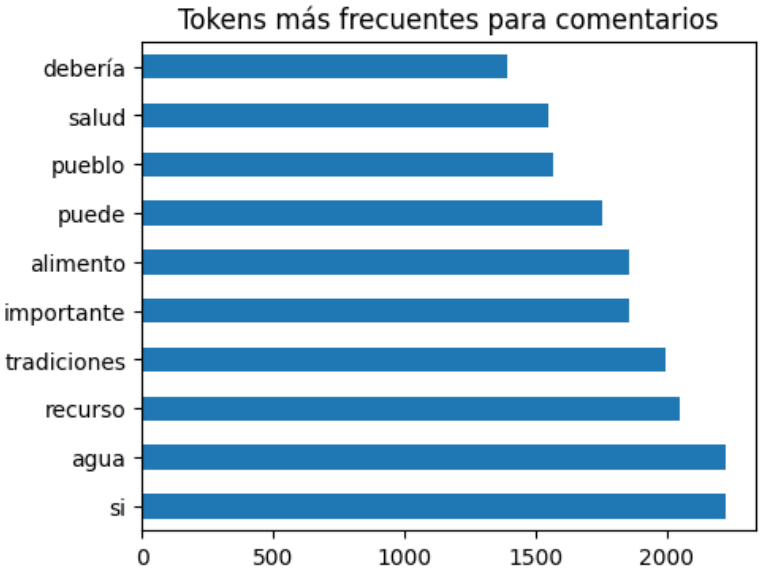


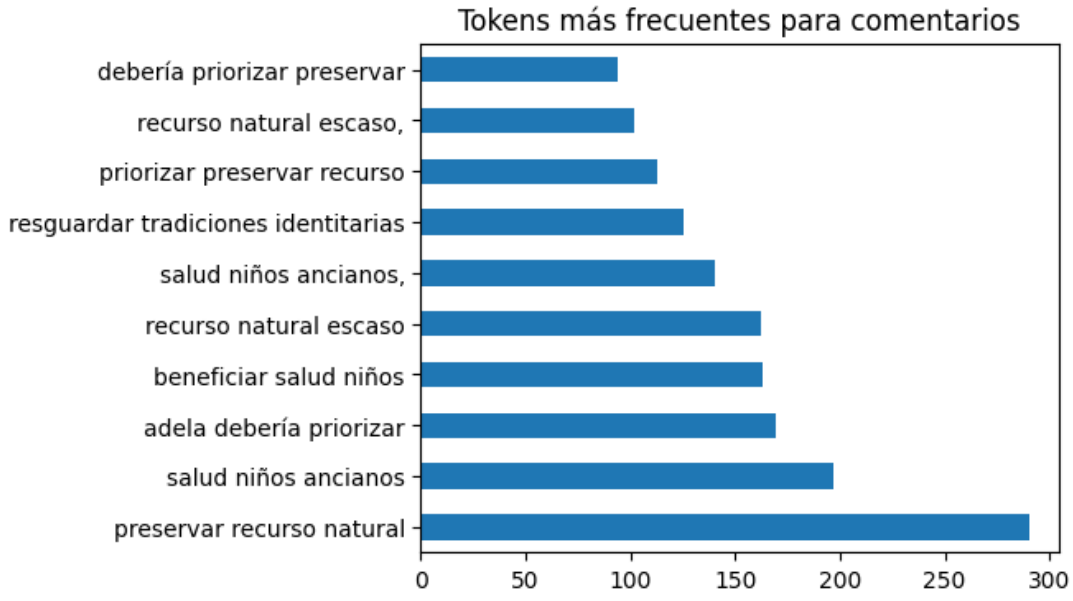


n-gramas comunes caso Adela

Preliminarmente mostramos los 1-gramas y 3-gramas más frecuentes del dataset para el caso Laura.

Adela es una ingeniera de una startup que busca generar un nuevo alimento en beneficio de niñas/os y personas de tercera edad. No obstante, la producción de este alimento consume una cantidad importante de agua en un sector de escasez

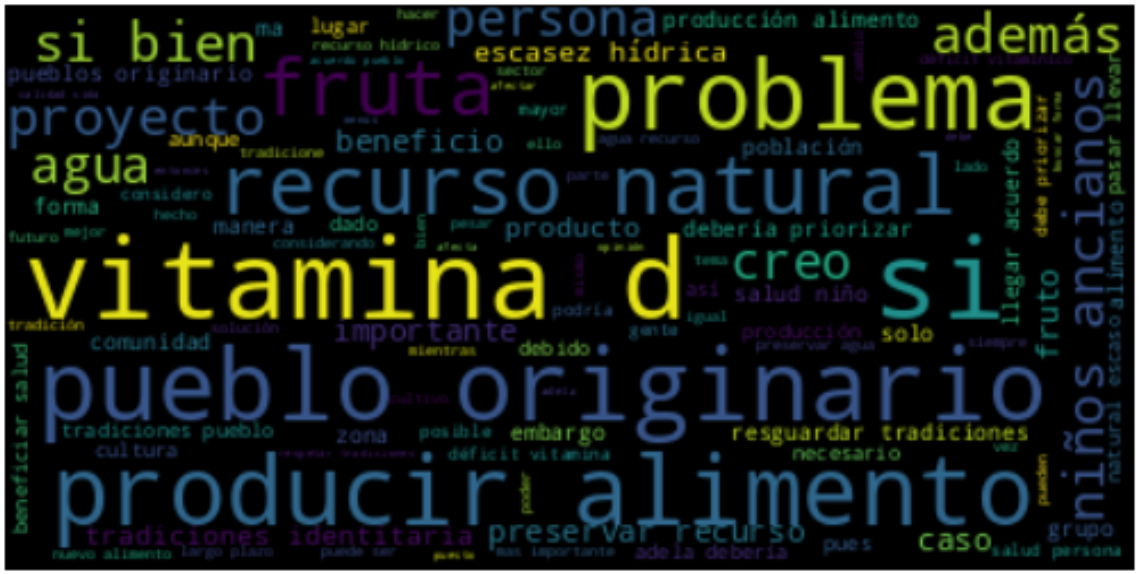




n-gramas comunes caso Adela

Preliminarmente mostramos los 1-gramas y 3-gramas más frecuentes del dataset para el caso Laura.

Adela es una ingeniera de una startup que busca generar un nuevo alimento en beneficio de niñas/os y personas de tercera edad. No obstante, la producción de este alimento consume una cantidad importante de agua en un sector de escasez.



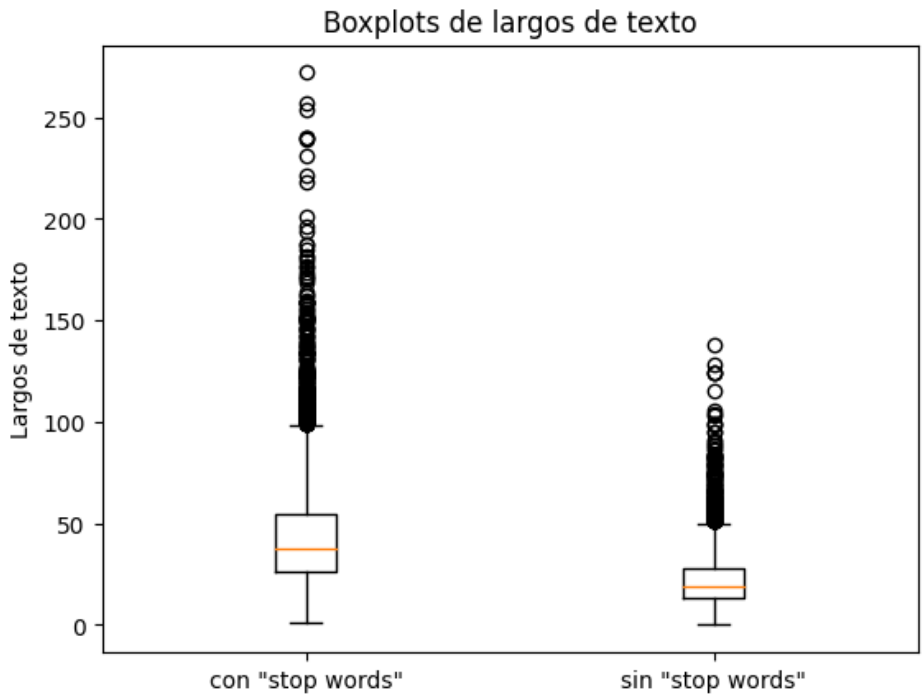
Largos de texto caso Adela

Con *stop words*

Media: 43.341
Desviación estándar: 25.996
Mediana: 38.0
Mínimo: 1
Máximo: 272

Sin *stop words*

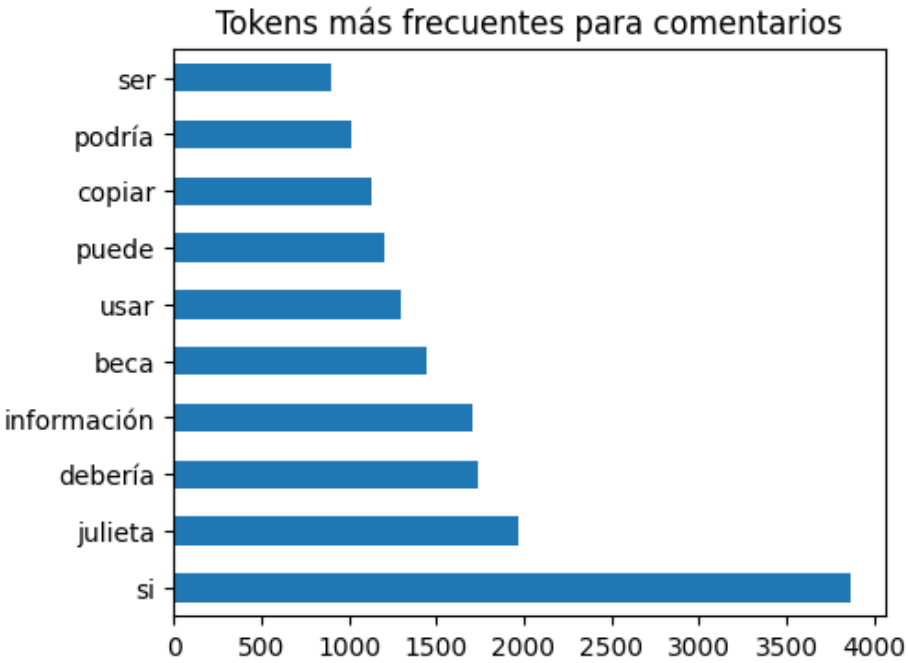
Media: 21.833
Desviación estándar: 13.021
Mediana: 19.0
Mínimo: 0
Máximo: 138



n-gramas comunes caso Julieta

Preliminarmente mostramos los 1-gramas y 3-gramas más frecuentes del dataset para el caso Julieta.

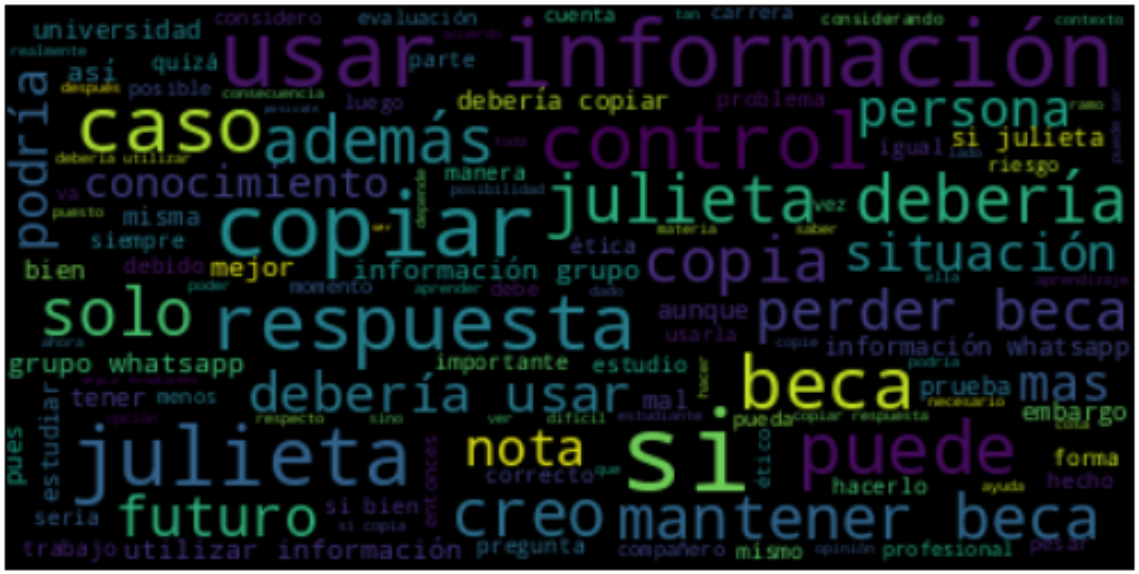
En el último control realizado Julieta se ve en la posibilidad de copiar una respuesta que fue compartida en el grupo de WhatsApp de su sección.



n-gramas comunes caso Julieta

Preliminarmente mostramos los 1-gramas y 3-gramas más frecuentes del dataset para el caso Julieta.

En el último control realizado Julieta se ve en la posibilidad de copiar una respuesta que fue compartida en el grupo de WhatsApp de su sección.



Largos de texto caso Julieta

Con *stop words*

Media: 43.652

Desviación estándar: 36.528

Mediana: 35.0

Mínimo: 1

Máximo: 528

Sin *stop words*

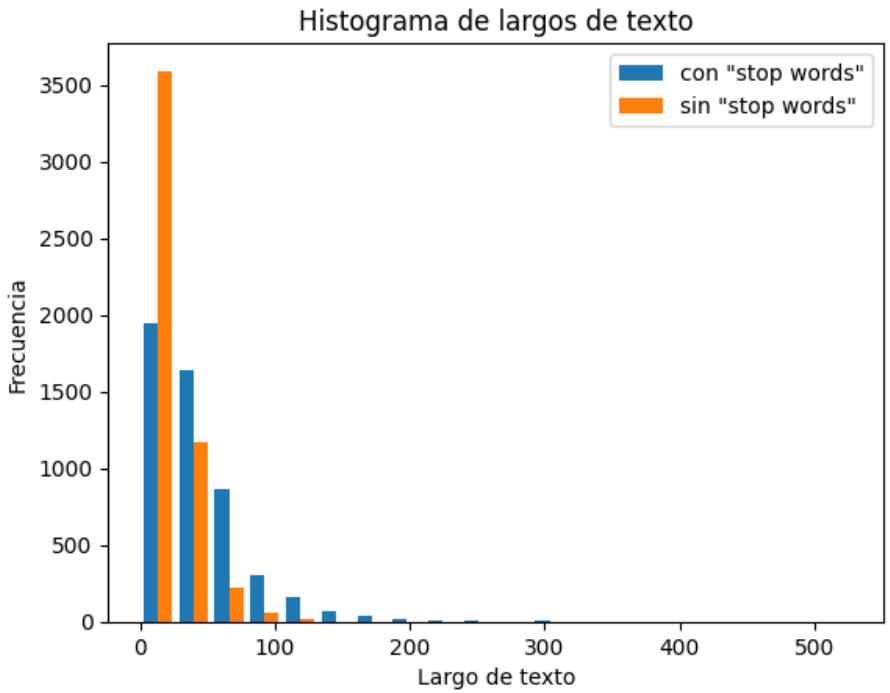
Media: 22.077

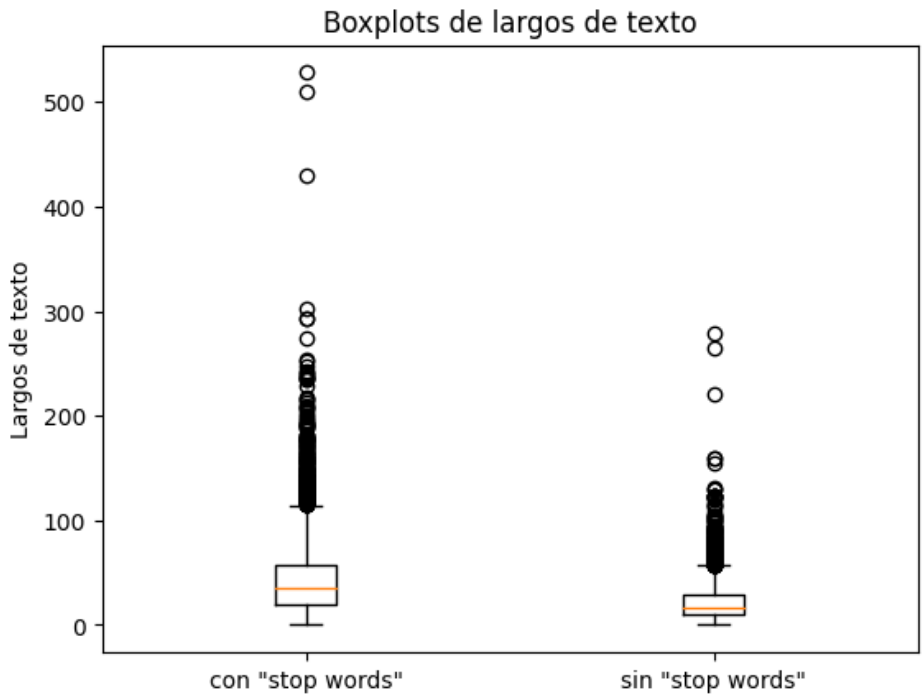
Desviación estándar: 18.345

Mediana: 17.0

Mínimo: 0

Máximo: 279

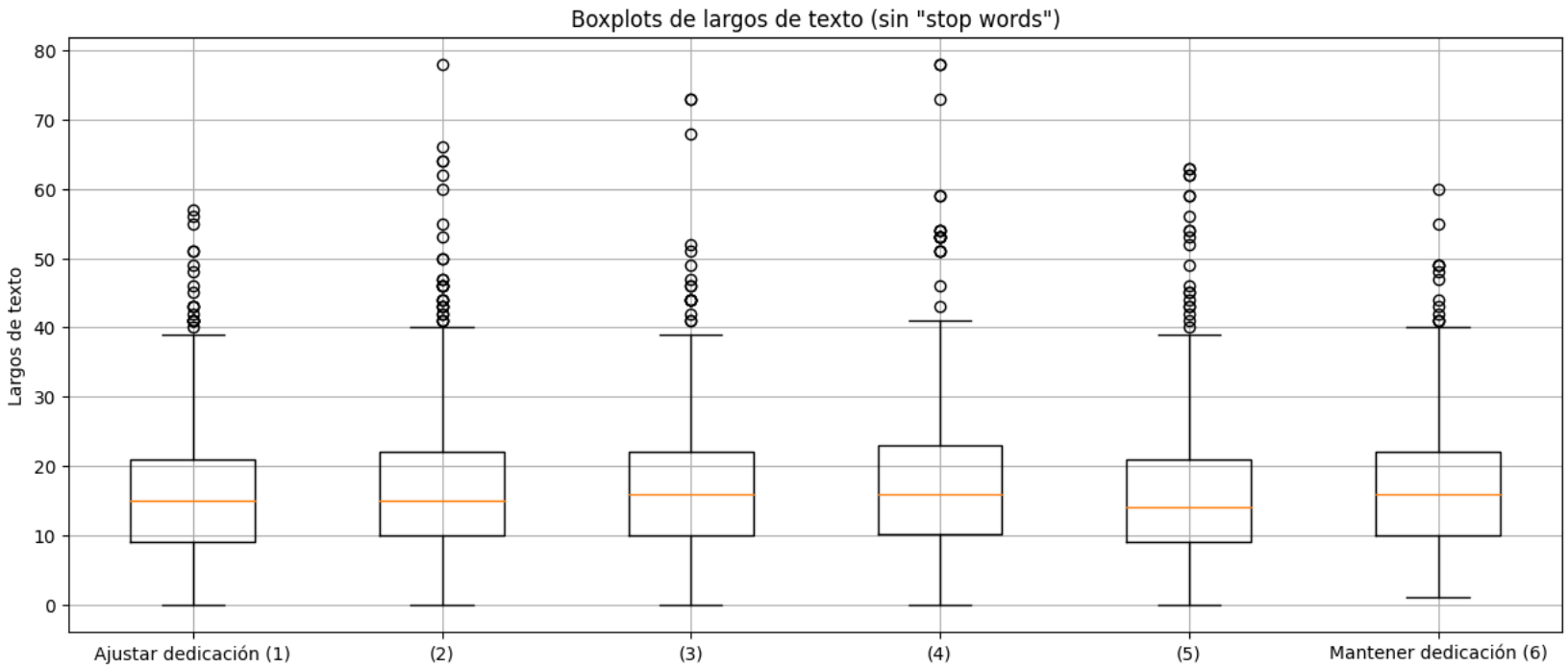




Largos de texto por opción caso Laura

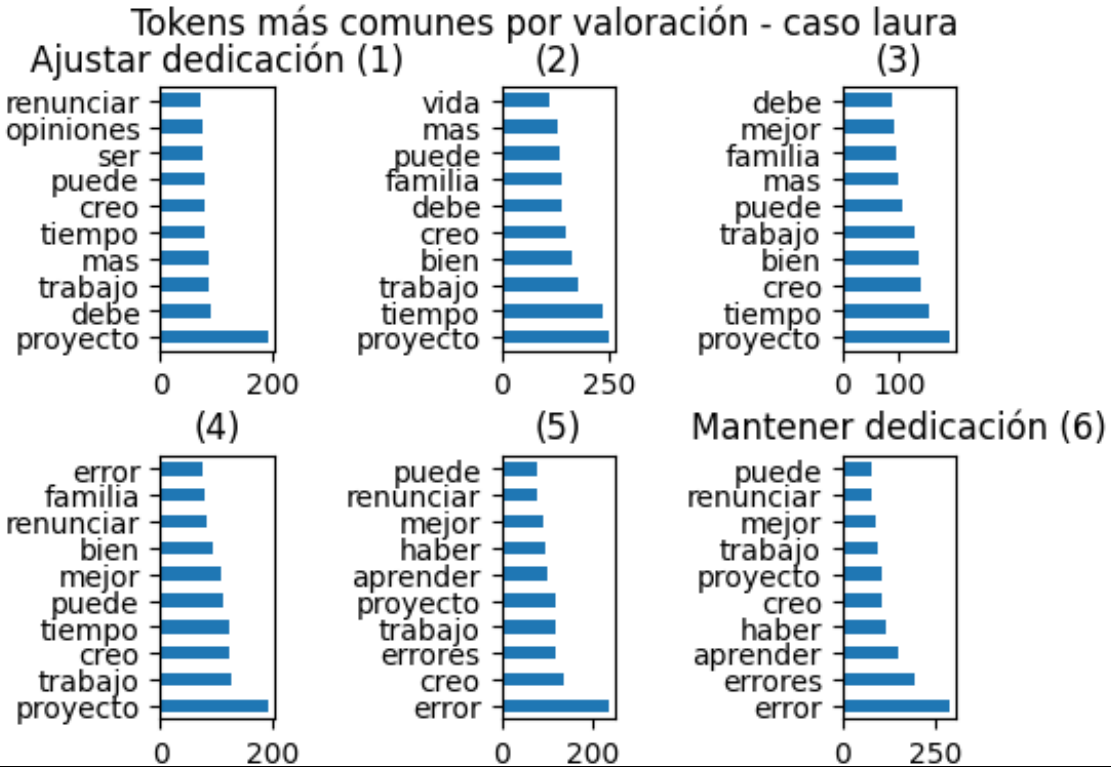
Largos de texto por opción.

Es adecuado que Laura le dedique paulatinamente más tiempo al trabajo y su desarrollo profesional que a la familia y las otras dimensiones de su vida.



Palabras frecuentes por opción caso Laura

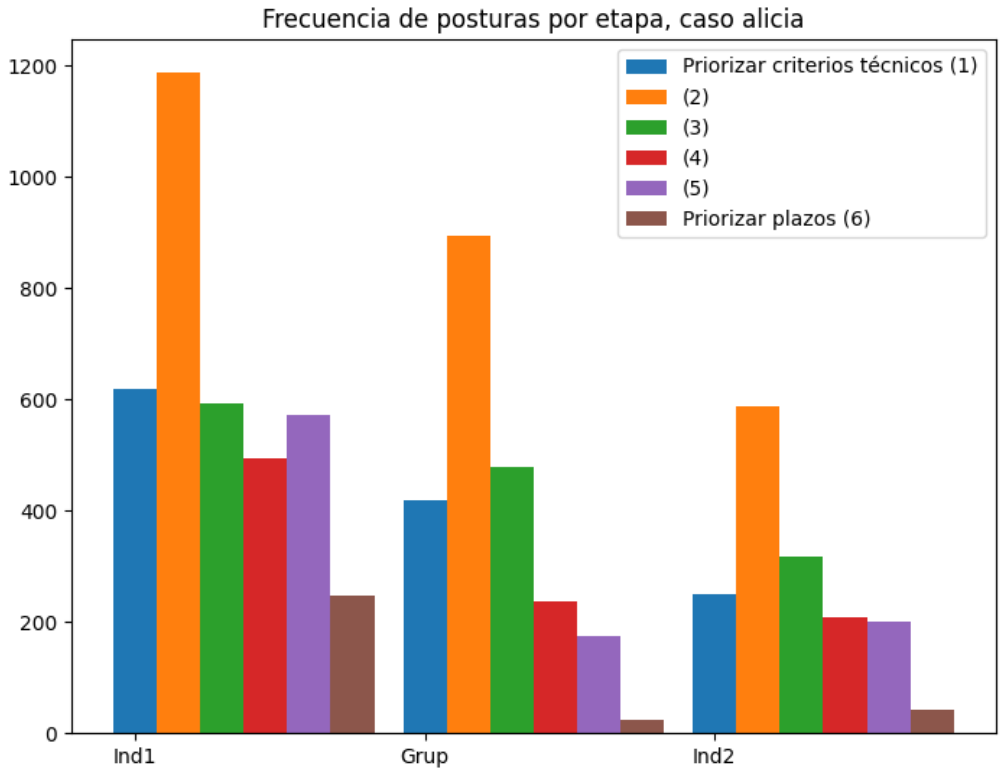
Largos de texto por opción. *Es adecuado que Laura le dedique paulatinamente más tiempo al trabajo y su desarrollo profesional que a la familia y las otras dimensiones de su vida.*



Frecuencias de posturas caso Alicia

Porcentaje de elección de estudiantes por opción y etapa.

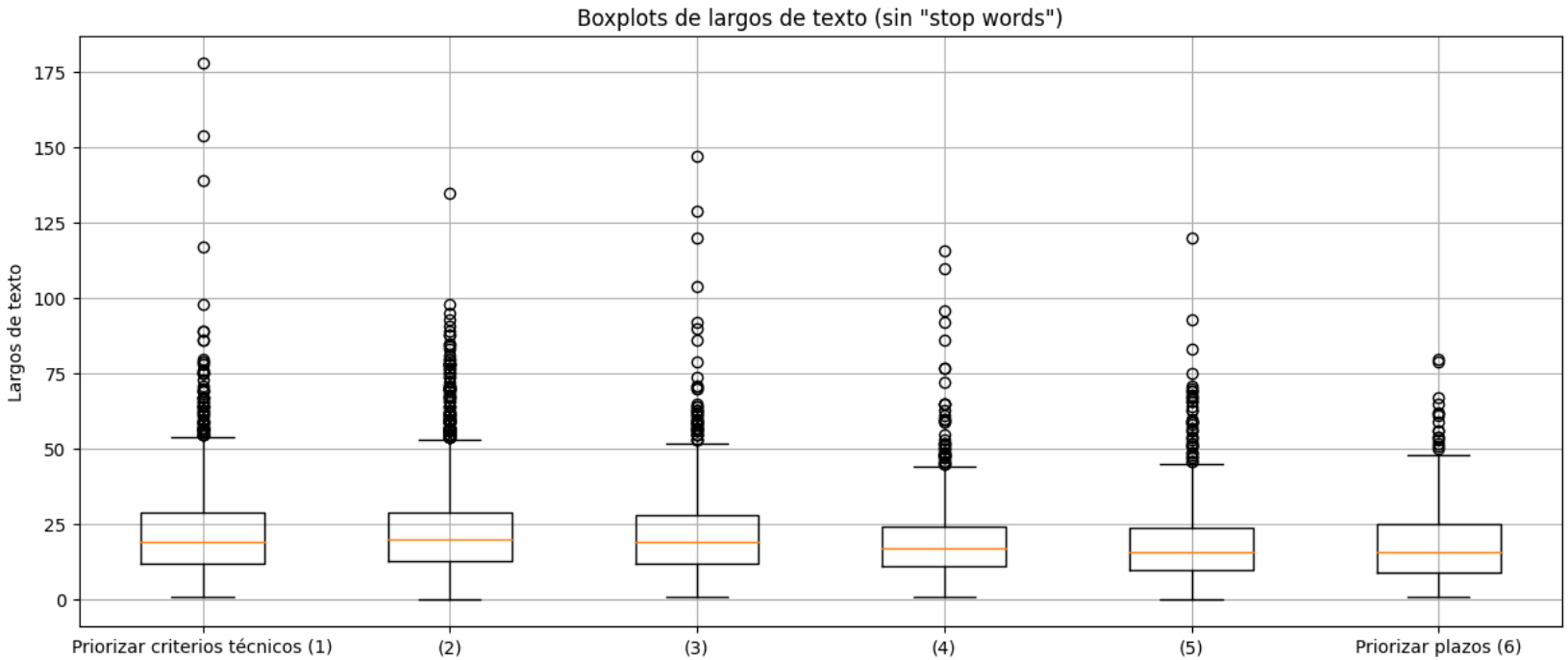
Ante problemas y retrasos por contingencia mundial, Alicia debería priorizar los plazos o los criterios técnicos?



Largos de texto por opción caso Alicia

Largos de texto por opción.

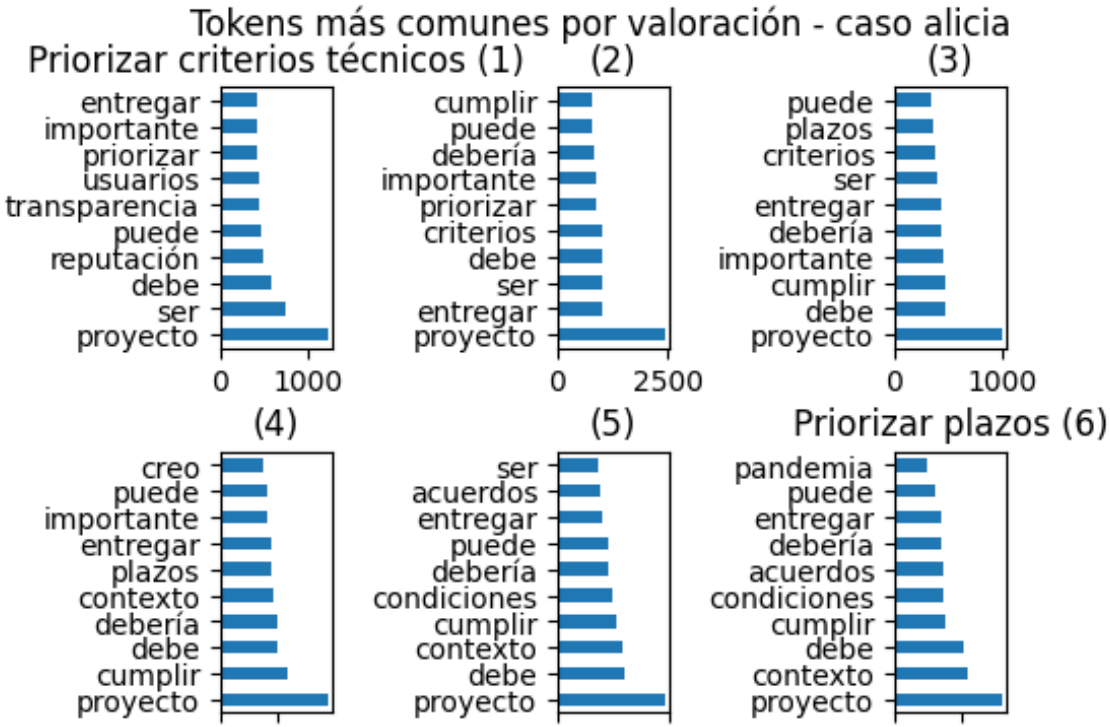
Ante problemas y retrasos por contingencia mundial, Alicia debería priorizar los plazos o los criterios técnicos?



Palabras frecuentes por opción caso Alicia

Largos de texto por opción.

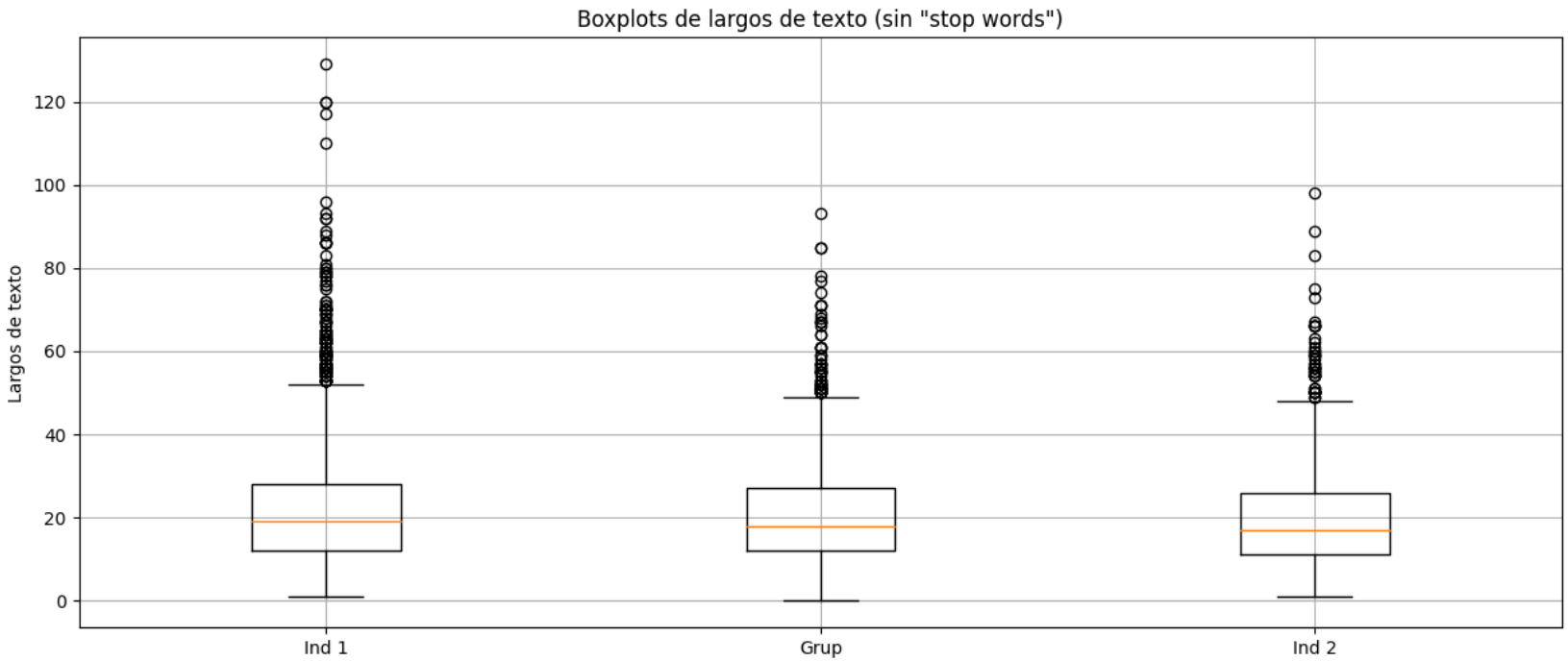
Ante problemas y retrasos por contingencia mundial, Alicia debería priorizar los plazos o los criterios técnicos?



Largos de texto por etapa caso Alicia

Largos de texto por etapa.

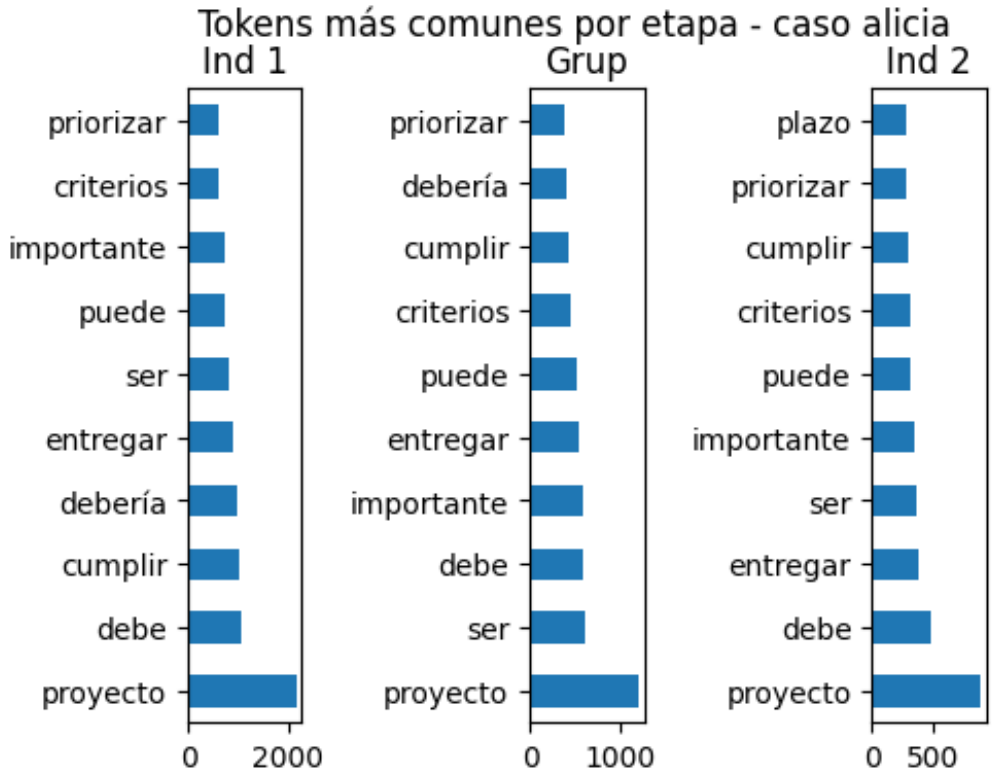
Ante problemas y retrasos por contingencia mundial, Alicia debería priorizar los plazos o los criterios técnicos?



Palabras frecuentes por etapa caso Alicia

Largos de texto por etapa.

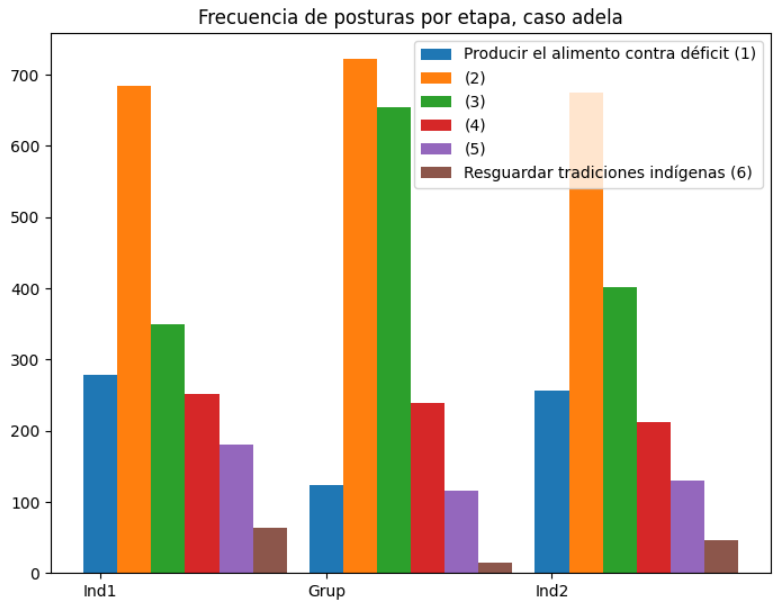
Ante problemas y retrasos por contingencia mundial, Alicia debería priorizar los plazos o los criterios técnicos?



Frecuencias de posturas caso Adela

Porcentaje de elección de estudiantes por opción y etapa.

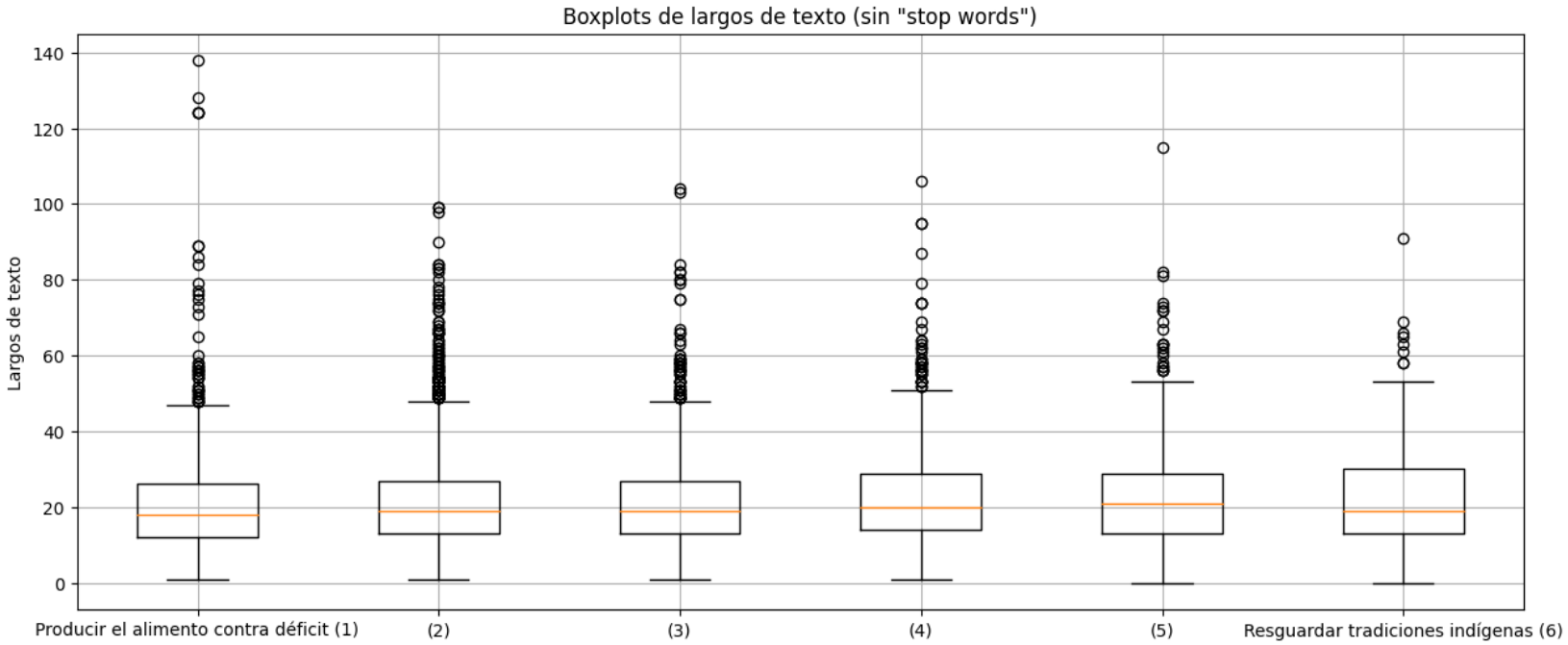
Adela es una ingeniera de una startup que busca generar un nuevo alimento en beneficio de niñas/os y personas de tercera edad. No obstante, la producción de este alimento consume una cantidad importante de agua en un sector de escasez.



Largos de texto por opción caso Adela

Largos de texto por opción.

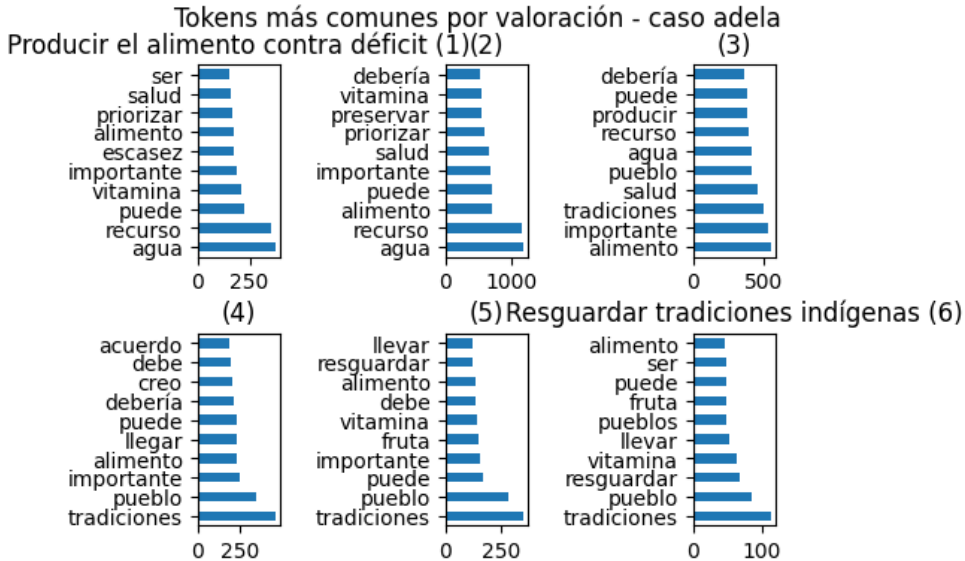
Adela es una ingeniera de una startup que busca generar un nuevo alimento en beneficio de niñas/os y personas de tercera edad. No obstante, la producción de este alimento consume una cantidad importante de agua en un sector de escasez.



Palabras frecuentes por opción caso Adela

Largos de texto por opción.

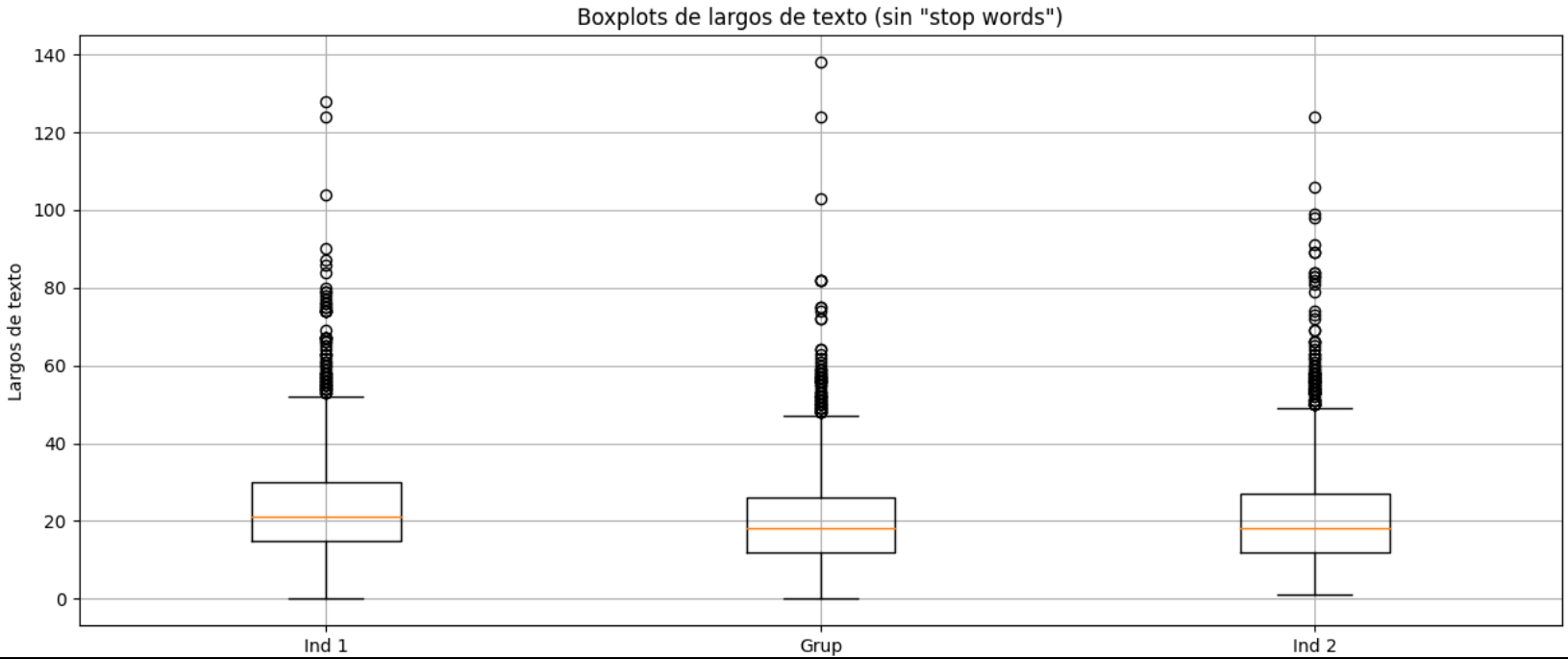
Adela es una ingeniera de una startup que busca generar un nuevo alimento en beneficio de niñas/os y personas de tercera edad. No obstante, la producción de este alimento consume una cantidad importante de agua en un sector de escasez.



Largos de texto por etapa caso Adela

Largos de texto por etapa.

Adela es una ingeniera de una startup que busca generar un nuevo alimento en beneficio de niñas/os y personas de tercera edad. No obstante, la producción de este alimento consume una cantidad importante de agua en un sector de escasez.



Palabras frecuentes por etapa caso Adela

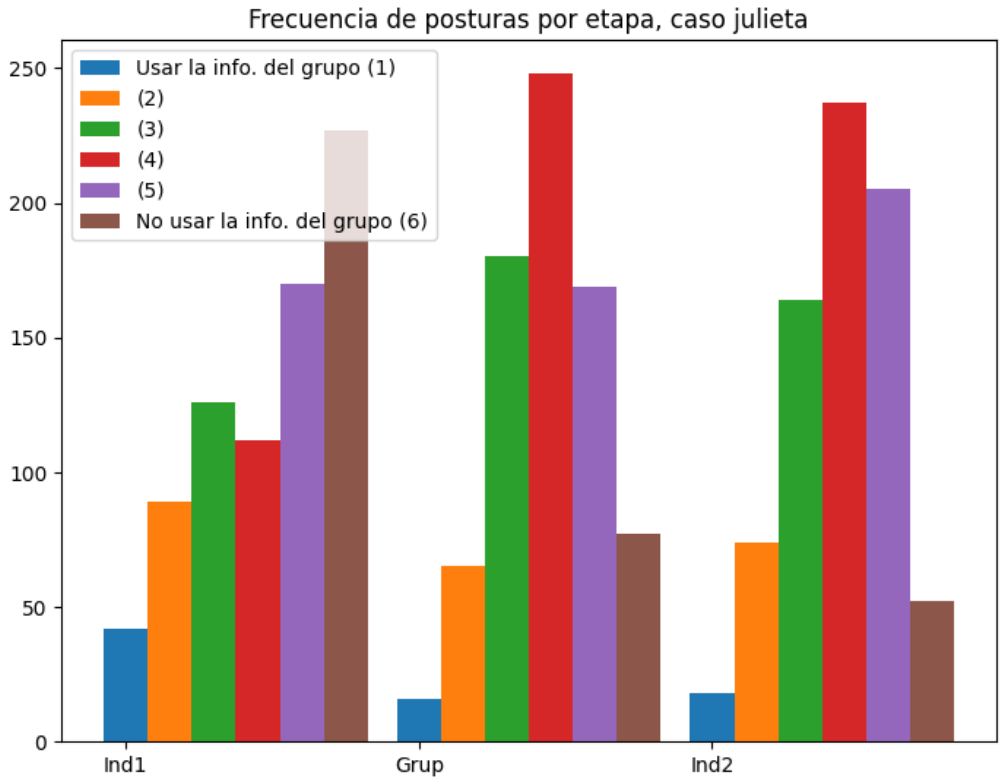
Largos de texto por etapa.

Adela es una ingeniera de una startup que busca generar un nuevo alimento en beneficio de niñas/os y personas de tercera edad. No obstante, la producción de este alimento consume una cantidad importante de agua en un sector de escasez.

Frecuencias de posturas caso Julieta

Porcentaje de elección de estudiantes por opción y etapa.

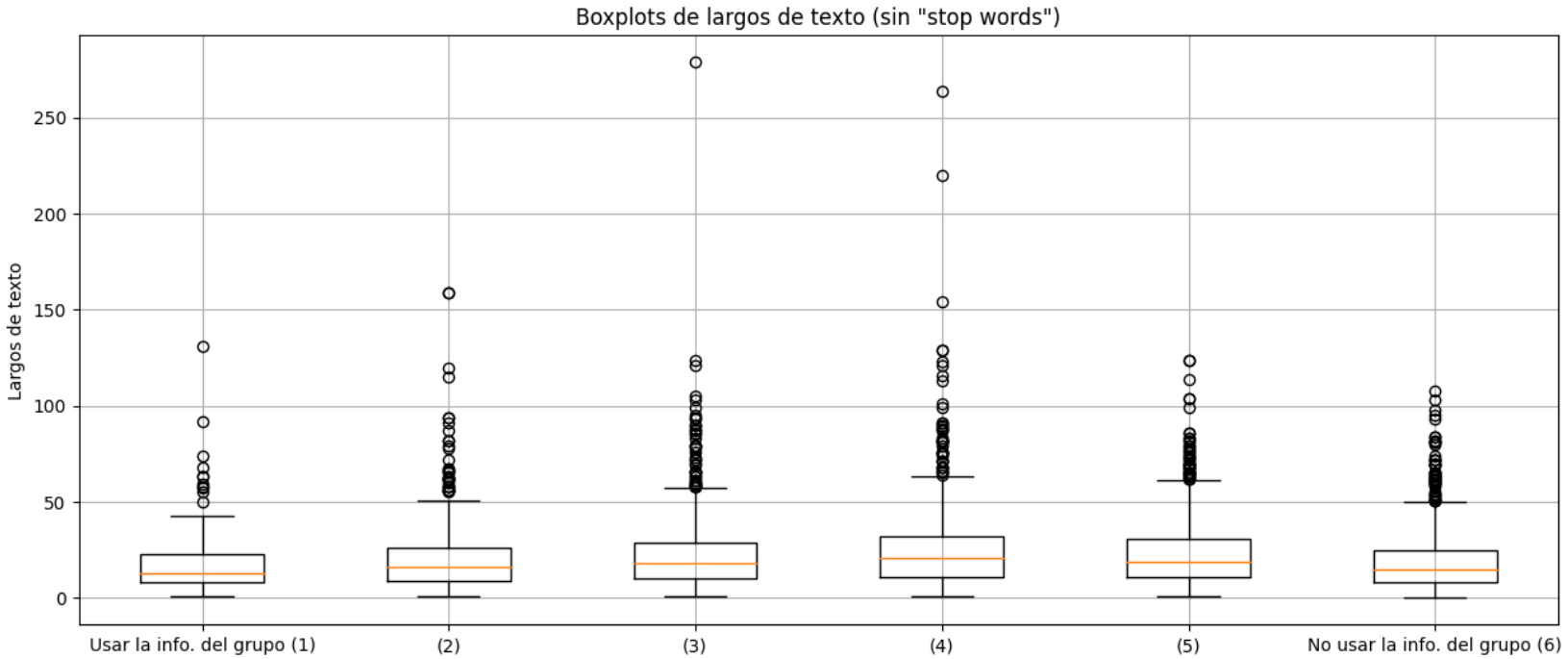
En el último control realizado Julieta se ve en la posibilidad de copiar una respuesta que fue compartida en el grupo de WhatsApp de su sección.



Largos de texto por opción caso Julieta

Largos de texto por opción.

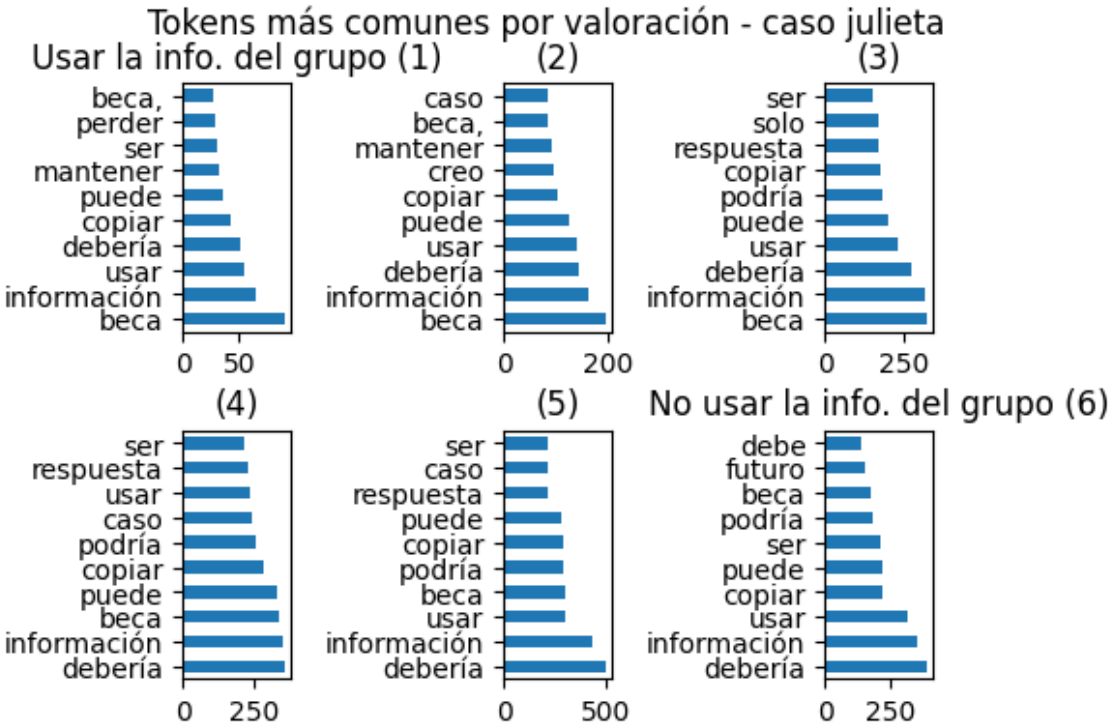
En el último control realizado Julieta se ve en la posibilidad de copiar una respuesta que fue compartida en el grupo de WhatsApp de su sección.



Palabras frecuentes por opción caso Julieta

Largos de texto por opción.

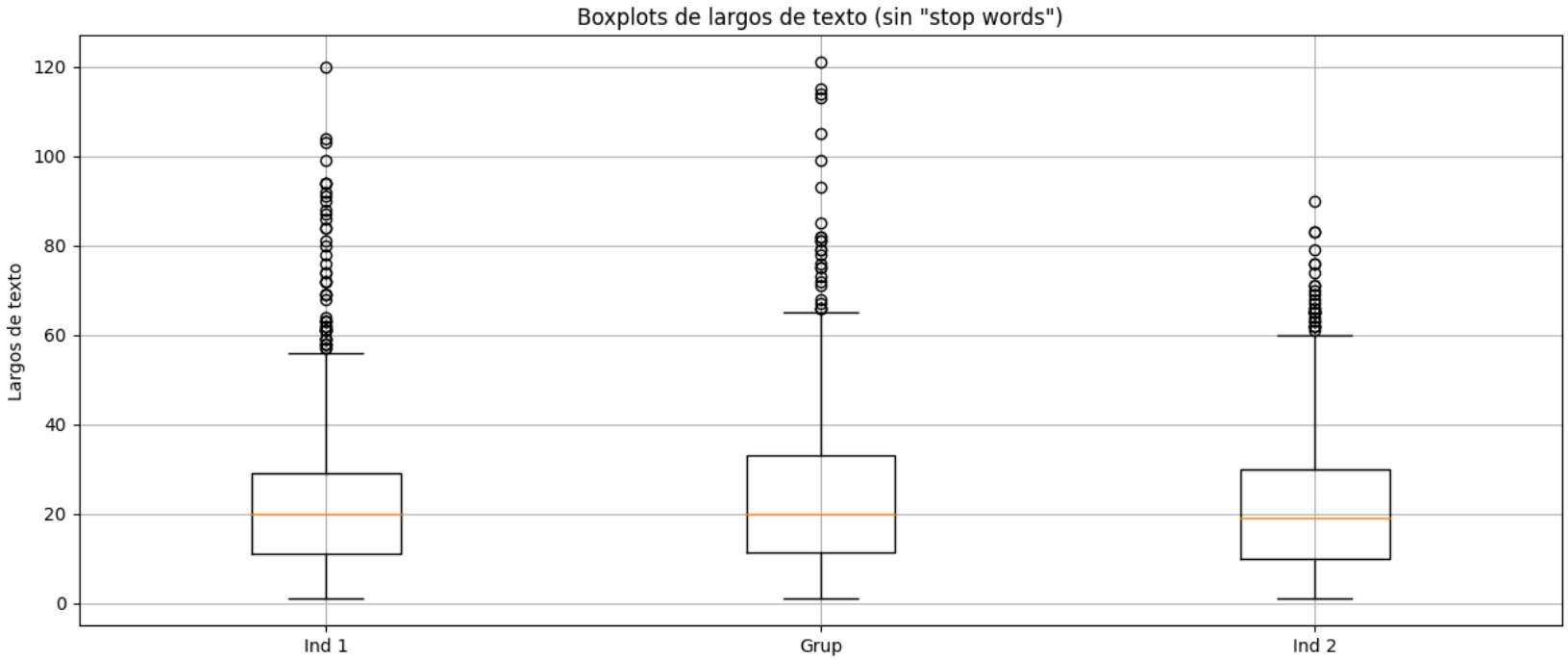
En el último control realizado Julieta se ve en la posibilidad de copiar una respuesta que fue compartida en el grupo de WhatsApp de su sección.



Largos de texto por etapa caso Julieta

Largos de texto por etapa.

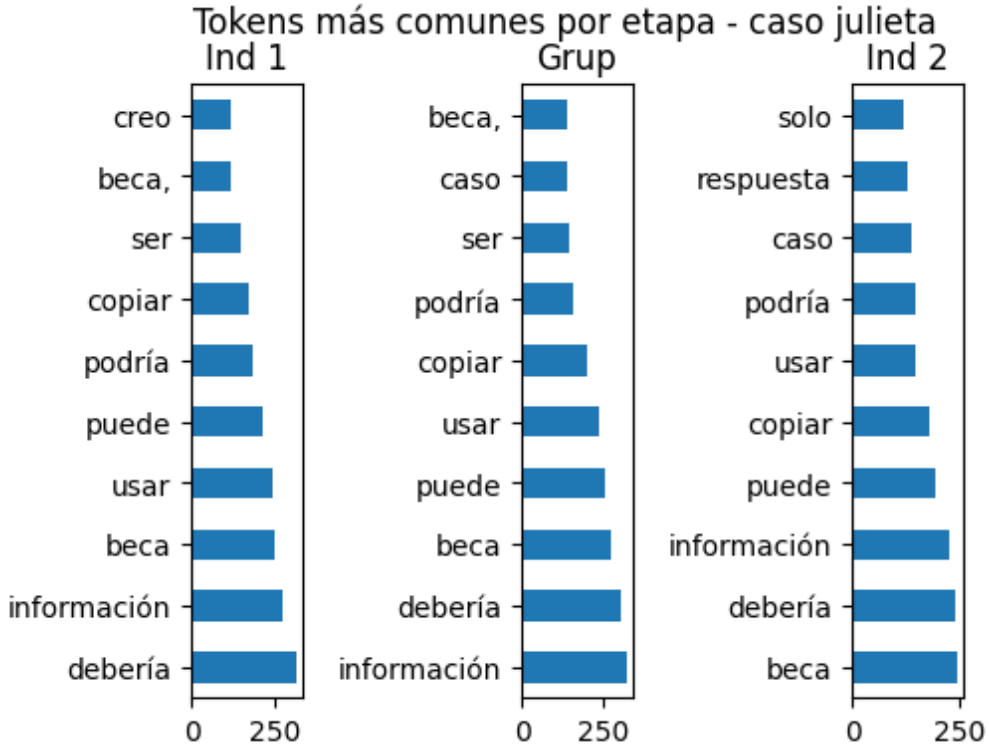
En el último control realizado Julieta se ve en la posibilidad de copiar una respuesta que fue compartida en el grupo de WhatsApp de su sección.



Palabras frecuentes por etapa caso Julieta

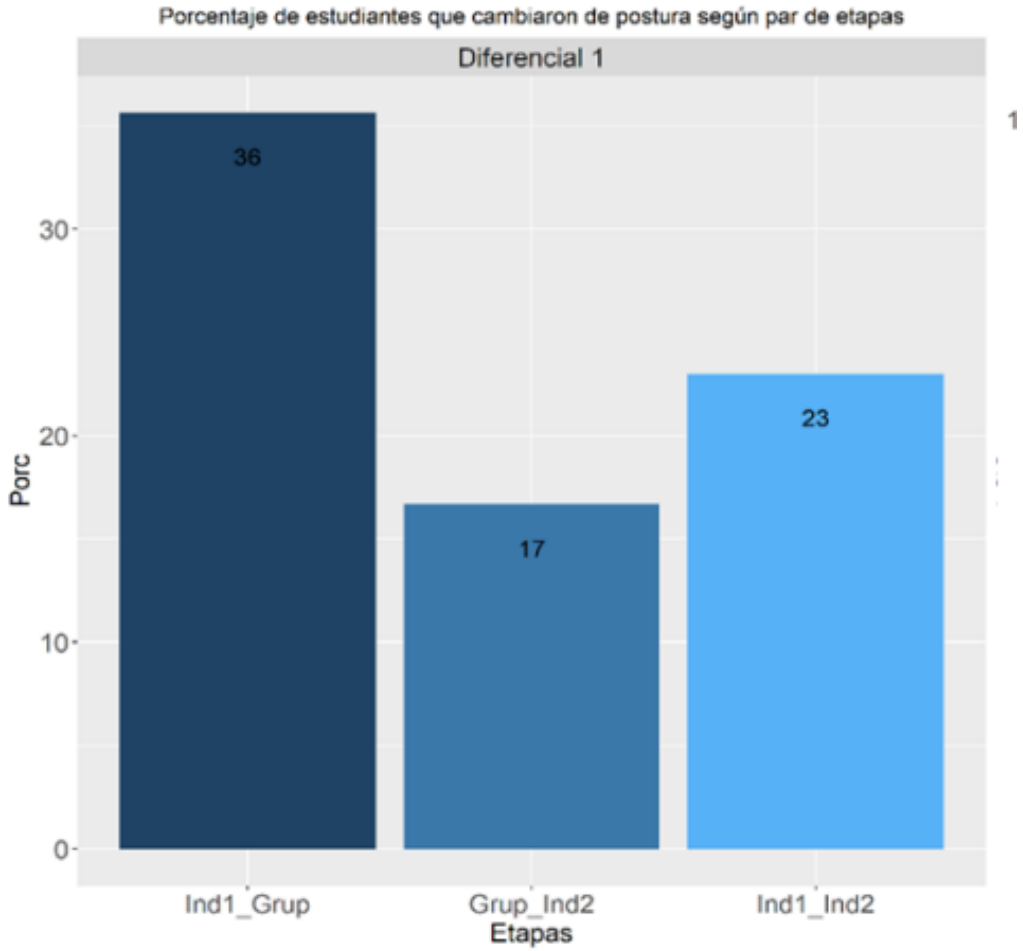
Largos de texto por etapa.

En el último control realizado Julieta se ve en la posibilidad de copiar una respuesta que fue compartida en el grupo de WhatsApp de su sección.



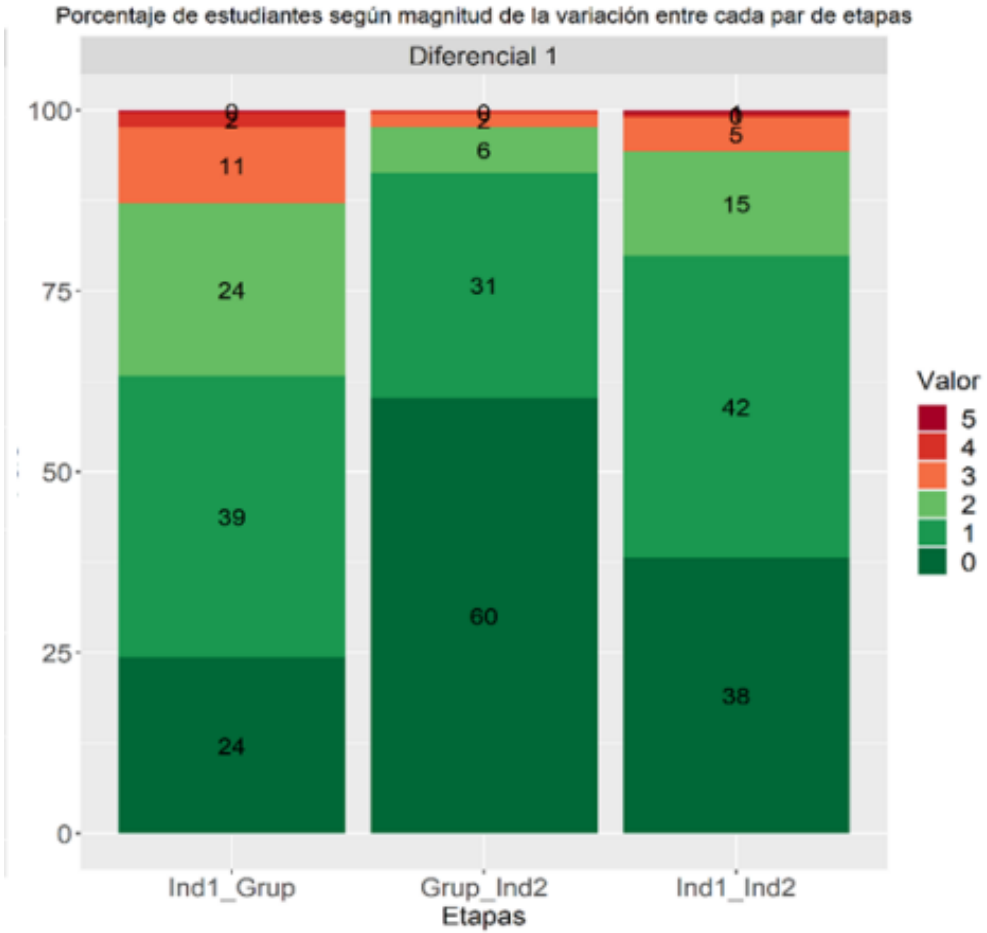
Cambios de postura Julieta

Porcentaje de cambios de postura [2].



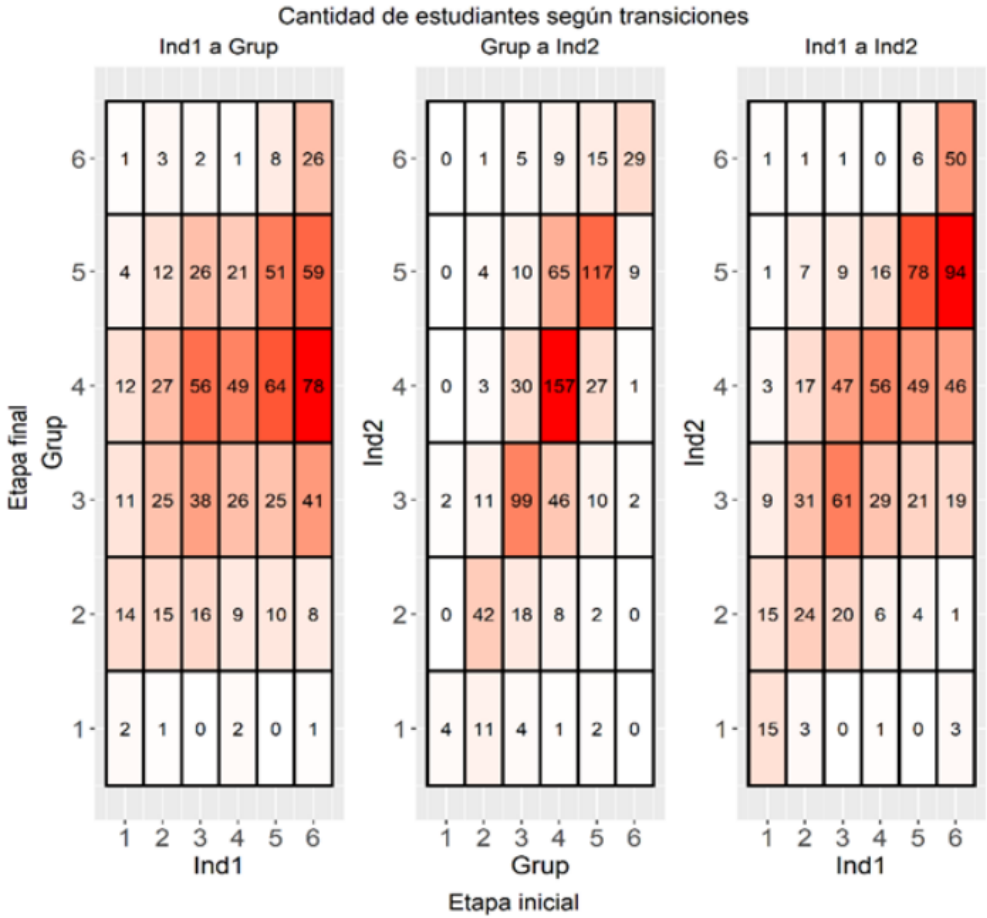
Cambios de postura Julieta

Magnitud de cambio de postura por par de etapas [2].



Cambios de postura Julieta

Detalle de cambios de postura por par de etapas [2].



Índice de contenidos

Clasificación con Naive-Bayes

Naive-Bayes:
Modelo de clasificación que asigna a cada elemento (palabra) una probabilidad de pertenecer a una clase. Las probabilidades se suman para la predicción final.[3]

Naive Bayes

Los métodos Naive-Bayes son una familia de algoritmos supervisados basados en aplicación del teorema de Bayes

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

con la hipótesis "naive" de independencia condicional entre pares de características dado el valor de la variable objetivo (clase). Consideremos vectores n-dimensionales, cada uno correspondiente a un documento sobre un vocabulario de largo n. Dado un documento $x = (x_1, \ldots, x_n)$, nos gustaría etiquetarlos con alguna clase $C_k \in \{C_1, \ldots, C_k\}$, Por el teorema de Bayes tenemos $\forall k \in \{1, \ldots, K\}$:

$$\begin{aligned}\mathbb{P}(C_k|x_1, \ldots, x_n) &= \frac{\mathbb{P}(C_k)\mathbb{P}(x_1, \ldots, x_n|C_k)}{\mathbb{P}(x_1, \ldots, x_n)} \\ &\propto \mathbb{P}(C_k)\mathbb{P}(x_1, \ldots, x_n|C_k)\end{aligned}$$

Naive Bayes

El término $\mathbb{P}(C_k)$ puede ser estimado por la frecuencia de la clase C_k en los datos. Para calcular la verosimilitud $\mathbb{P}(x_1, \ldots, x_n | C_k)$ primero asumimos una distribución de probabilidad (por ejemplo Gaussiana o Multinomial). En el caso de clasificación binaria ($k \in \{0, 1\}$)) usando una distribución de Bernoulli, tenemos para cada característica j (palabra/token):

$$\mathbb{P}(x_j | C_k) = \mathbb{P}(j | C_k) x_j + (1 - \mathbb{P}(j | C_k))(1 - x_j)$$

$\mathbb{P}(j | C_k)$ puede ser estimada también, tomándolo la proporción de documentos que contienen la palabra j entre las realizaciones de la clase C_k.

Pre-procesamiento

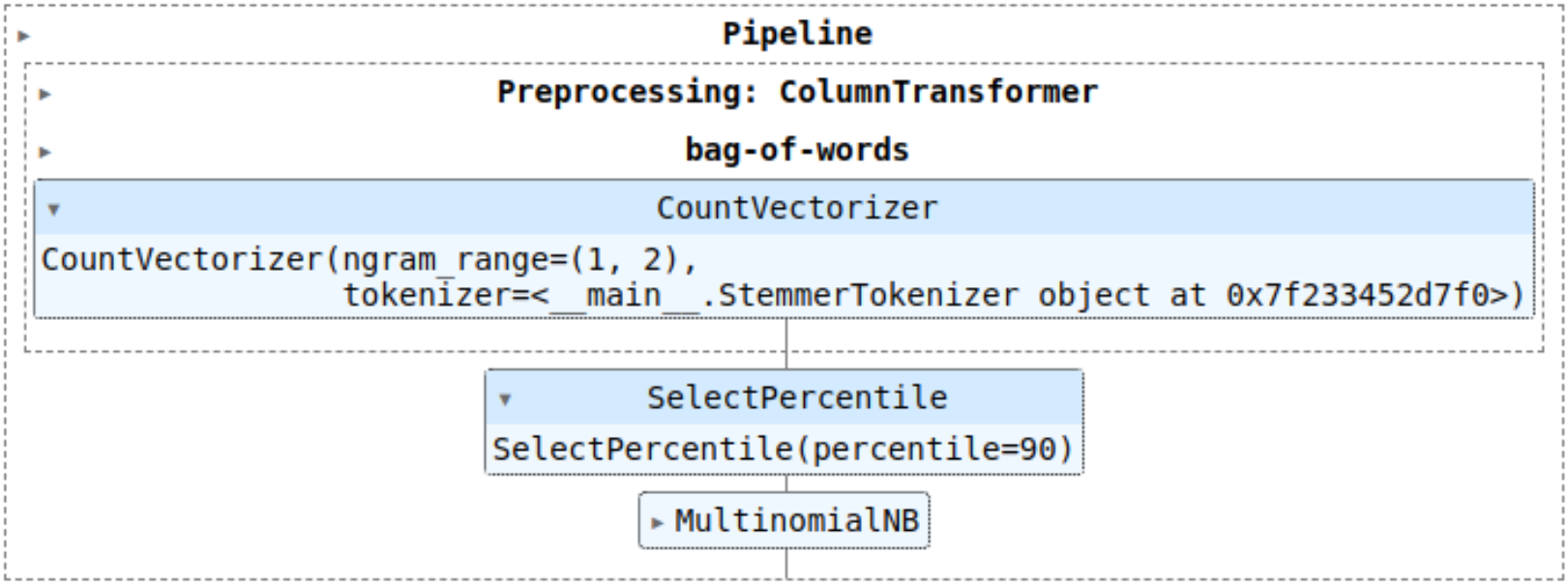
Pasos para vectorizar texto:

1. Todo a minúscula
2. Remover *stop-words*
3. *Stemming*

Pre-procesamiento

Pasos para vectorizar texto:

1. Todo a minúscula
2. Remover *stop-words*
3. *Stemming*
4. Clasificador



Resultados preliminares Laura

Resultados de clasificación.

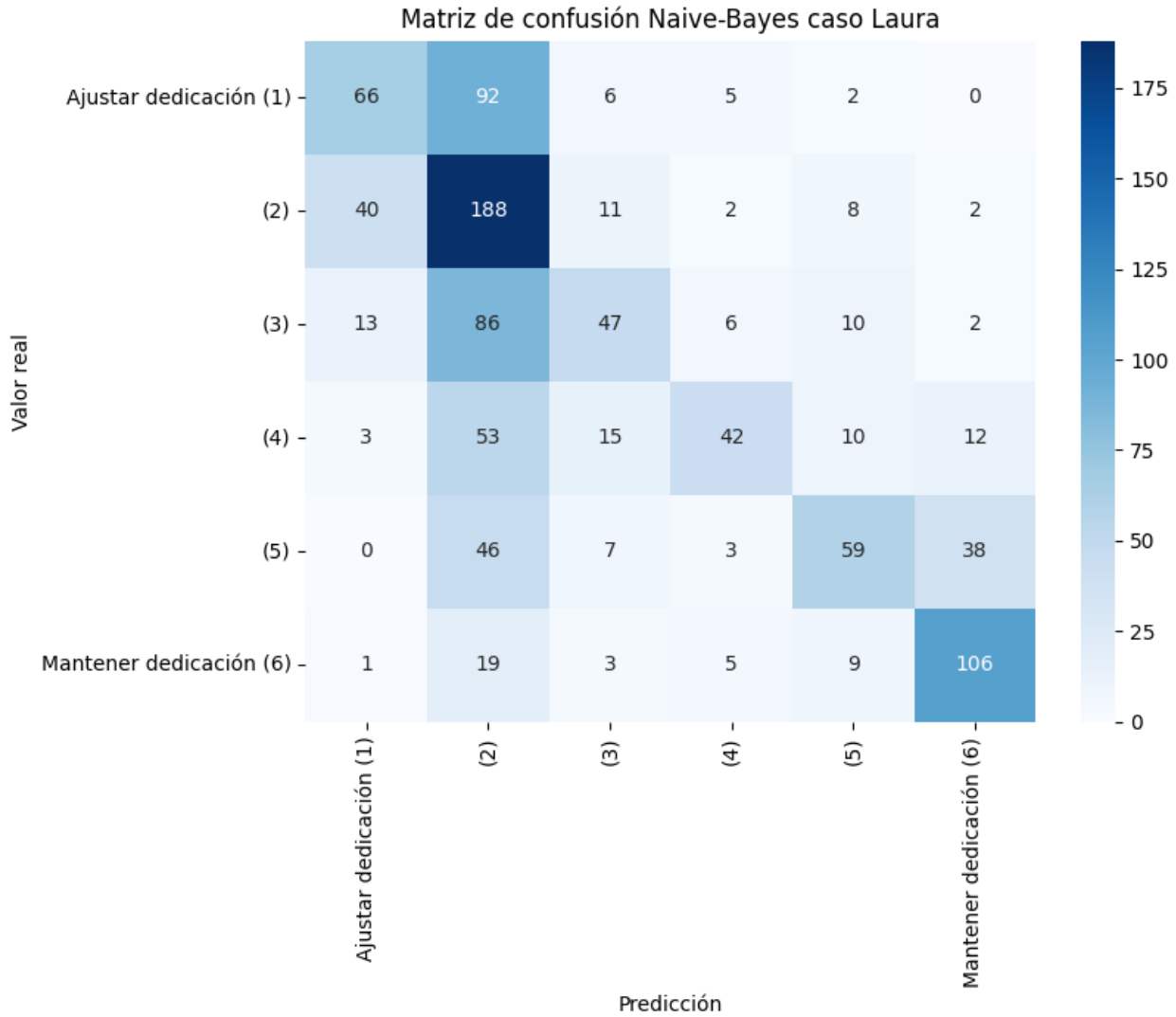
Es adecuado que Laura le dedique paulatinamente más tiempo al trabajo y su desarrollo profesional que a la familia y las otras dimensiones de su vida.

(1) Ajustar dedicación — (6) Mantener dedicación

Resultados clasificador Naive-Bayes multinomial

	precision	recall	f1-score	support
1	0.54	0.39	0.45	171
2	0.39	0.75	0.51	251
3	0.53	0.29	0.37	164
4	0.67	0.31	0.42	135
5	0.60	0.39	0.47	153
6	0.66	0.74	0.70	143
accuracy			0.50	1017
macro avg	0.56	0.48	0.49	1017
weighted avg	0.54	0.50	0.49	1017

Resultados preliminares Laura



Resultados preliminares Laura

Versión binaria: pasamos las elecciones “indecisas” a “fuertes”.

Resultados clasificador Naive-Bayes multinomial (binario)

	precision	recall	f1-score	support
1	0.80	0.90	0.85	586
6	0.83	0.69	0.76	431
accuracy			0.81	1017
macro avg	0.82	0.80	0.80	1017
weighted avg	0.81	0.81	0.81	1017

Resultados preliminares Alicia

Resultados de clasificación.

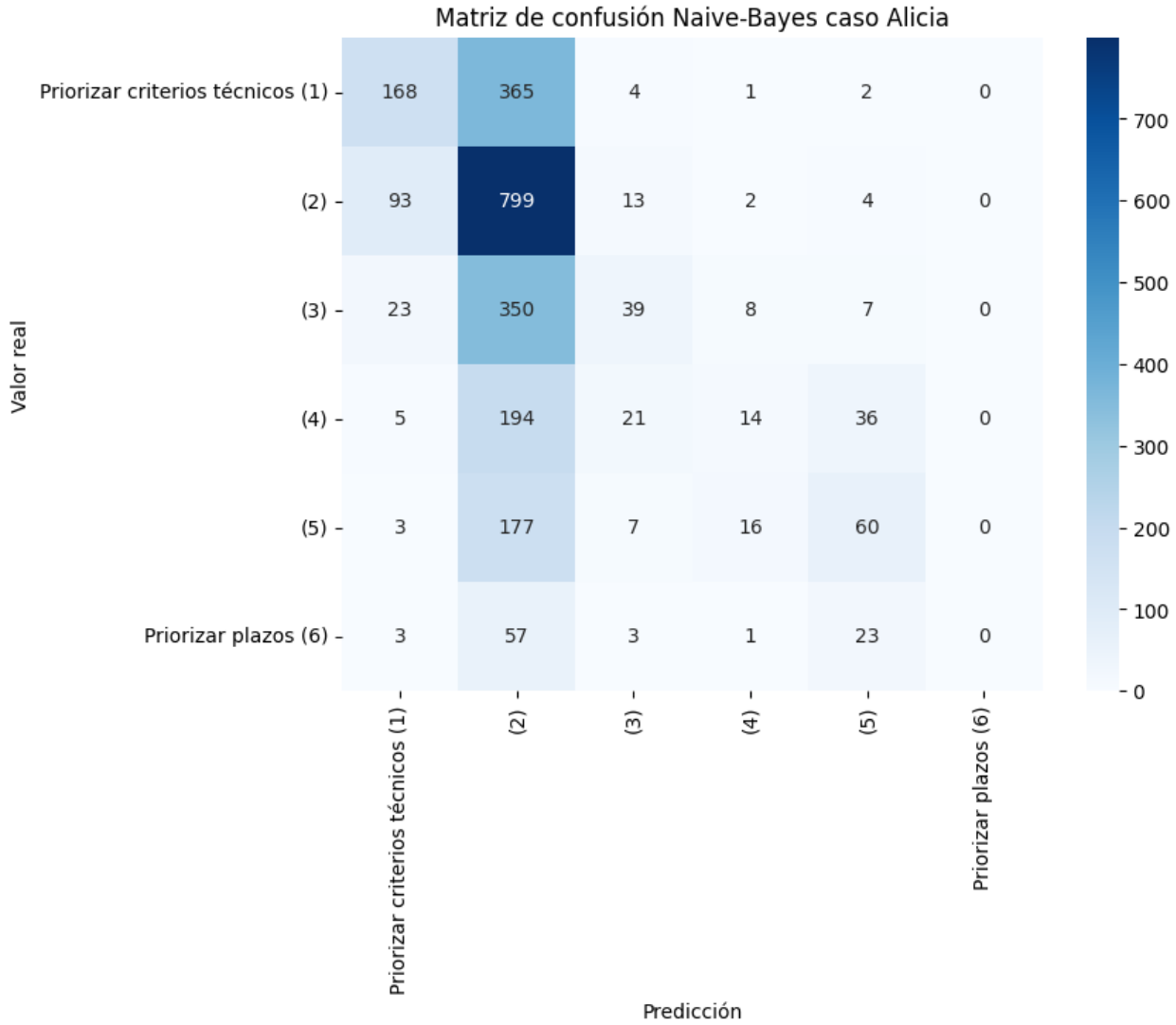
Ante problemas y retrasos por contingencia mundial, Alicia debería priorizar los plazos o los criterios técnicos?

(1) Priorizar criterios técnicos — (6) Priorizar plazos

Resultados clasificador Naive-Bayes multinomial

	precision	recall	f1-score	support
1	0.59	0.31	0.41	540
2	0.42	0.89	0.57	911
3	0.48	0.11	0.17	427
4	0.43	0.07	0.13	270
5	0.44	0.24	0.31	263
6	1.00	0.01	0.02	87
accuracy			0.44	2498
macro avg	0.56	0.27	0.27	2498
weighted avg	0.49	0.44	0.37	2498

Resultados preliminares Alicia



Resultados preliminares Alicia

Versión binaria: pasamos las elecciones “indecisas” a “fuertes”.

Resultados clasificador Naive-Bayes multinomial (binario)

	precision	recall	f1-score	support
1	0.85	0.96	0.90	1878
6	0.79	0.48	0.60	620
accuracy			0.84	2498
macro avg	0.82	0.72	0.75	2498
weighted avg	0.83	0.84	0.83	2498

Resultados preliminares Adela

Resultados de clasificación.

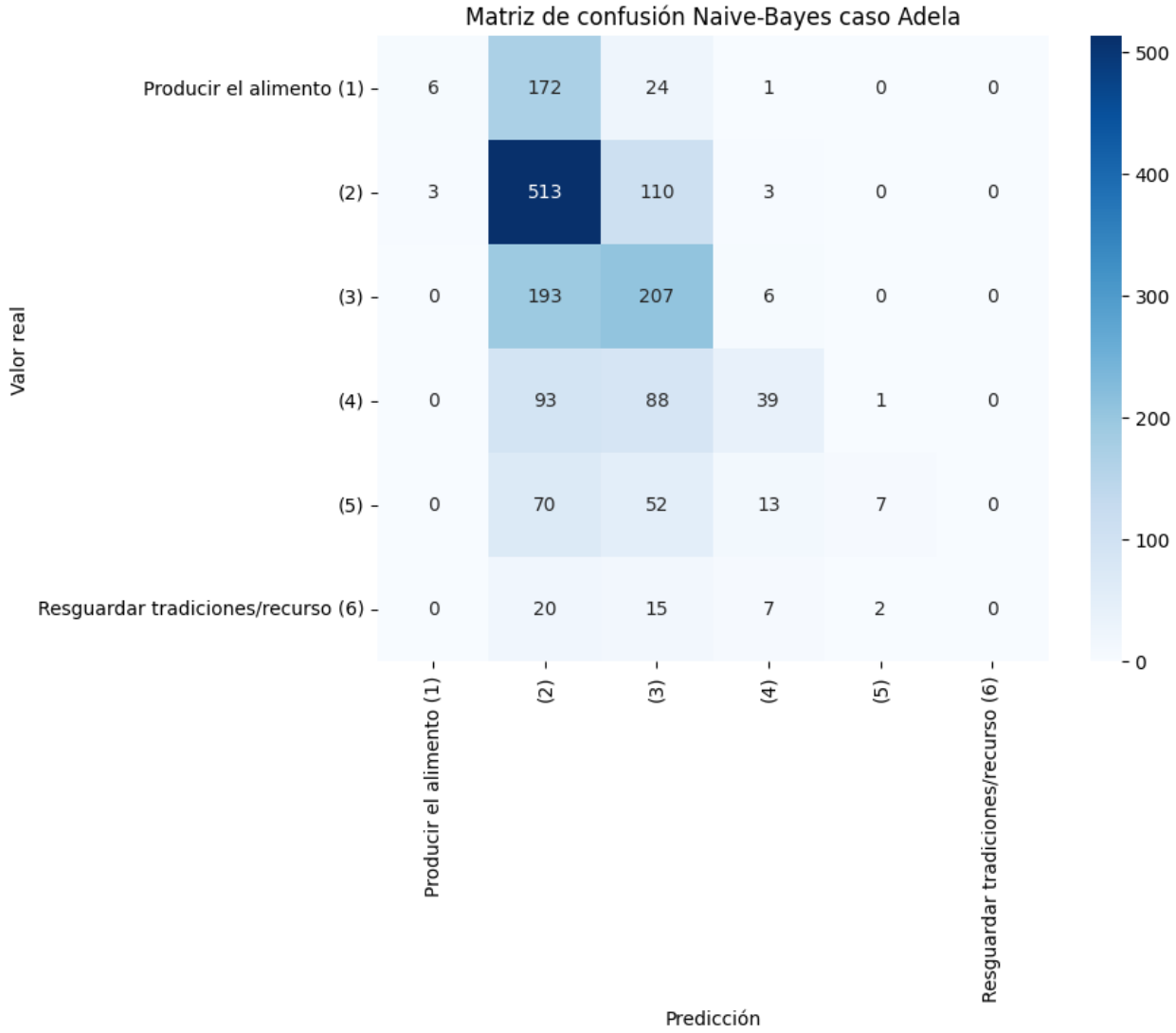
Adela busca generar un nuevo alimento con vitaminas. No obstante, la producción de este alimento consume una cantidad importante de agua en un sector de escasez.

(1) Producir el alimento contra el déficit — (6) Resguardar tradiciones/recursos

Resultados clasificador Naive-Bayes multinomial

	precision	recall	f1-score	support
1	0.75	0.04	0.08	203
2	0.49	0.86	0.62	629
3	0.42	0.47	0.44	406
4	0.47	0.12	0.19	221
5	0.39	0.05	0.09	142
6	0.00	0.00	0.00	44
accuracy			0.47	1645
macro avg	0.42	0.26	0.24	1645
weighted avg	0.48	0.47	0.39	1645

Resultados preliminares Adela



Resultados preliminares Adela

Versión binaria: pasamos las elecciones “indecisas” a “fuertes”.

Resultados clasificador Naive-Bayes multinomial (binario)

	precision	recall	f1-score	support
1	0.83	0.94	0.89	1239
6	0.71	0.43	0.54	406
accuracy			0.82	1645
macro avg	0.77	0.69	0.71	1645
weighted avg	0.80	0.82	0.80	1645

Resultados preliminares Julieta

Resultados de clasificación.

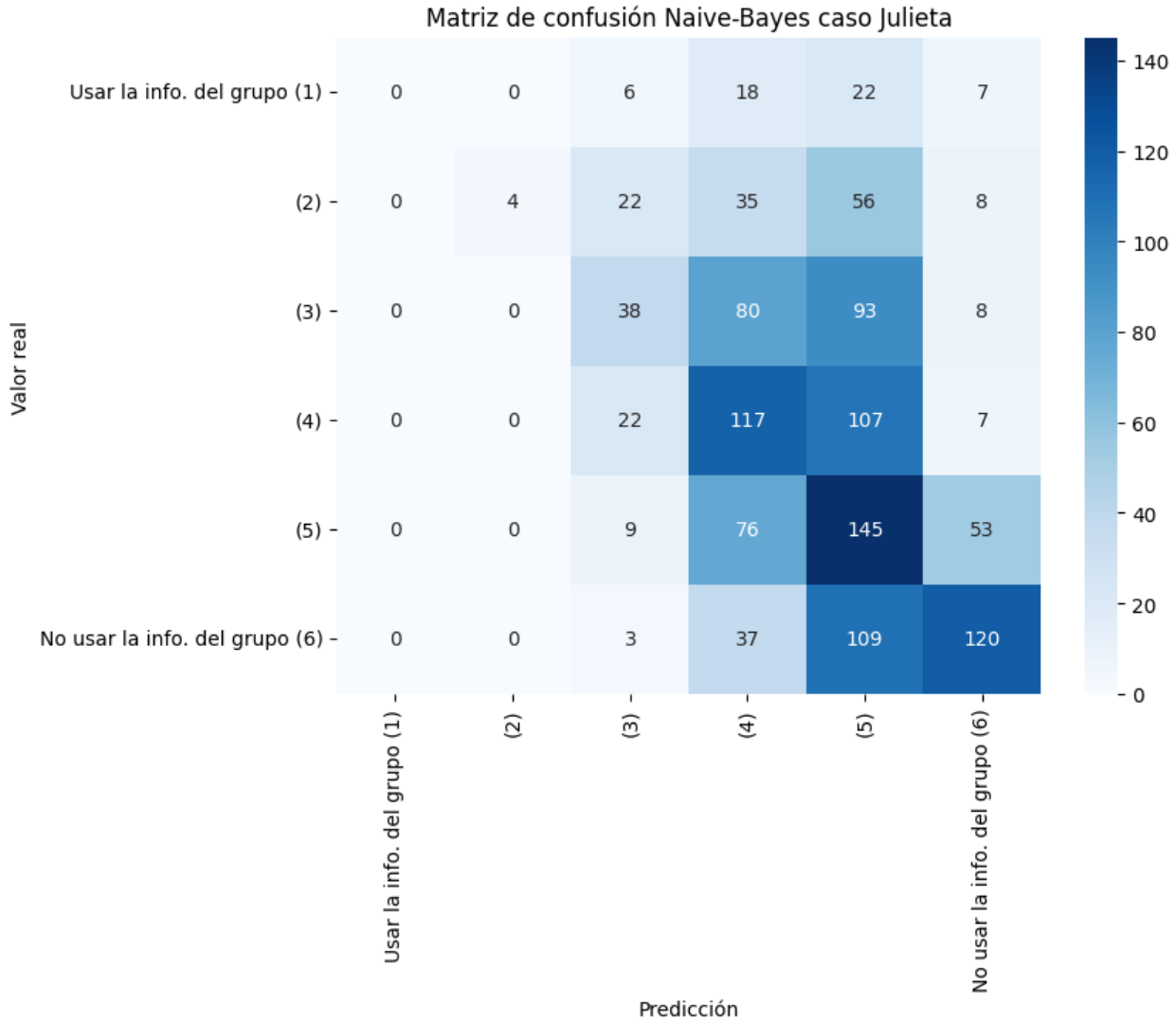
En el último control realizado Julieta se ve en la posibilidad de copiar una respuesta que fue compartida en el grupo de WhatsApp de su sección.

(1) Usar la info del grupo — (6) No usar la info del grupo

Resultados clasificador Naive-Bayes multinomial

	precision	recall	f1-score	support
1	0.00	0.00	0.00	53
2	1.00	0.03	0.06	125
3	0.38	0.17	0.24	219
4	0.32	0.46	0.38	253
5	0.27	0.51	0.36	283
6	0.59	0.45	0.51	269
accuracy			0.35	1202
macro avg	0.43	0.27	0.26	1202
weighted avg	0.44	0.35	0.33	1202

Resultados preliminares Julieta



Resultados preliminares Julieta

Versión binaria: pasamos las elecciones “indecisas” a “fuertes”.

Resultados clasificador Naive-Bayes multinomial (binario)

	precision	recall	f1-score	support
1	0.74	0.26	0.38	398
6	0.72	0.96	0.82	804
accuracy			0.72	1202
macro avg	0.73	0.61	0.60	1202
weighted avg	0.73	0.72	0.68	1202

Índice de contenidos

Interpretación de modelos

Con el modelo de base y otros clasificadores

- Ordenar tokens según probabilidad por clase (Naive-Bayes)
- Verificar que variables son las que afectan más el output (shap values)
- Hacer una clasificación más gruesa (binaria o sacando los extremos)
- Búsqueda de grilla para optimizar parámetros
- Oversampling de clases menos frecuentes para entrenamiento

Se considerarán desde luego modelos de regresión.

Modelos con interpretabilidad

- **Topic modelling** (Latent Dirichlet allocation - LDA)
  Es una técnica que agrupa de manera no supervisada los textos. Genera una distribución palabra - tópico latente (oculto) y tópico - palabra.[4]

Modelos con interpretabilidad

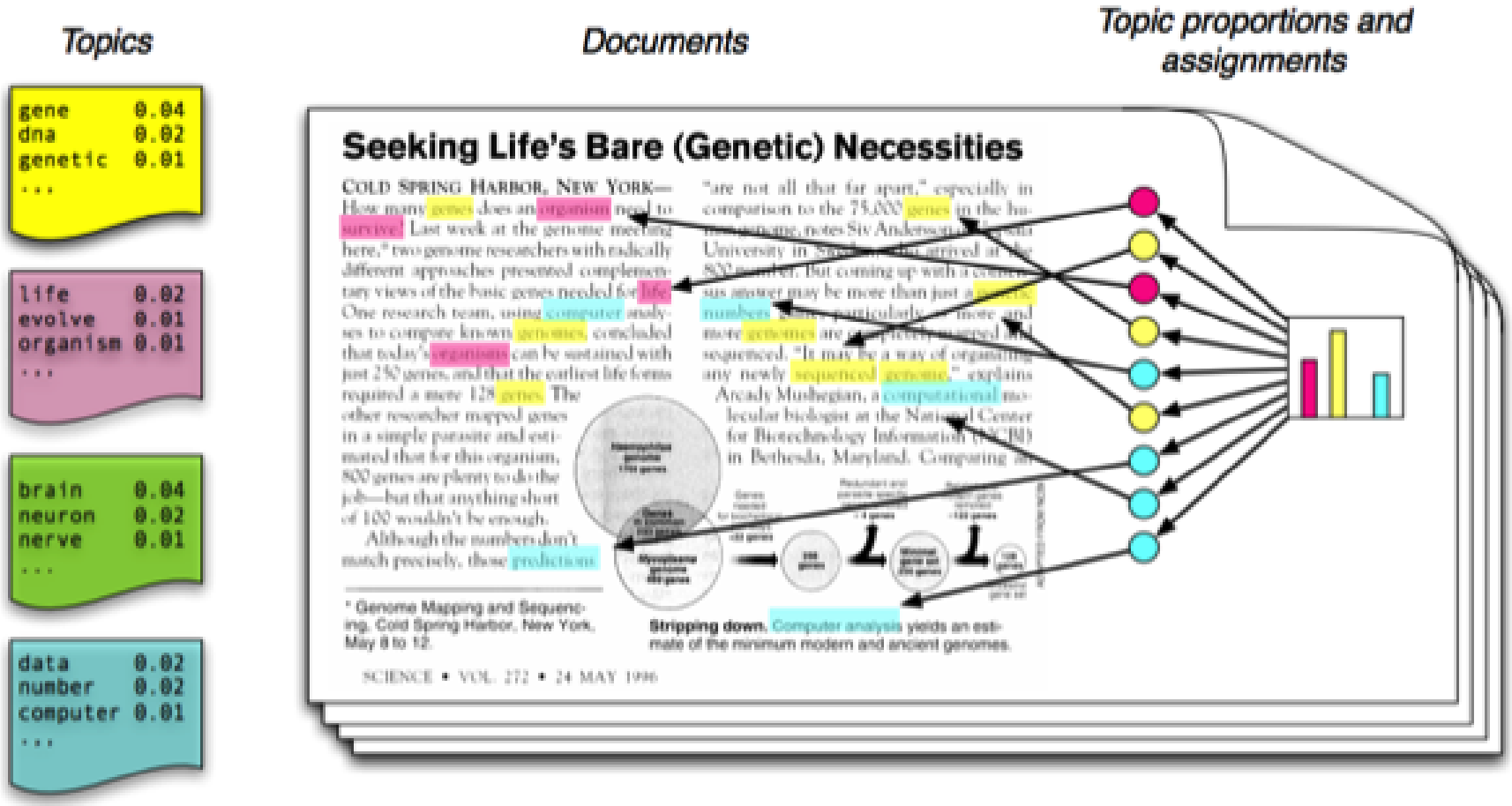


Figure source: Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77-84.

Modelos con interpretabilidad

- **Topic modelling** (Latent Dirichlet allocation - LDA)
  Es una técnica que agrupa de manera no supervisada los textos. Genera una distribución palabra - tópico latente (oculto) y tópico - palabra.[4]
- **Naive-Bayes**:
  Modelo de clasificación que asigna a cada elemento (palabra) una probabilidad de pertenecer a una clase. Las probabilidades se suman para la predicción final.[3]

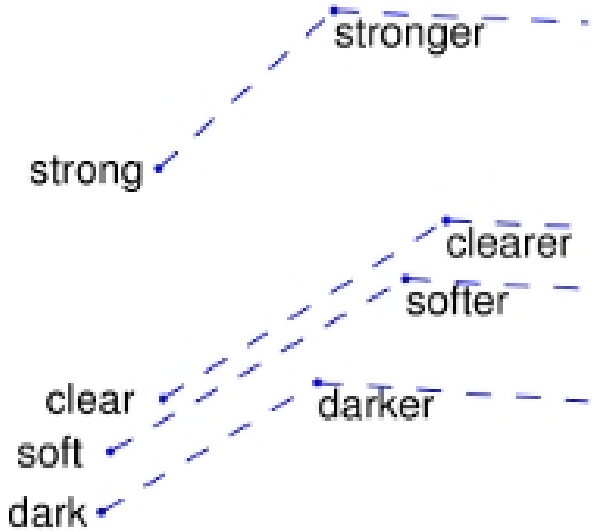
Modelos de aprendizaje profundo

Word embeddings: modelos que permiten vectorizar palabras, basados en su co-ocurrencia.

Algunos ejemplos:

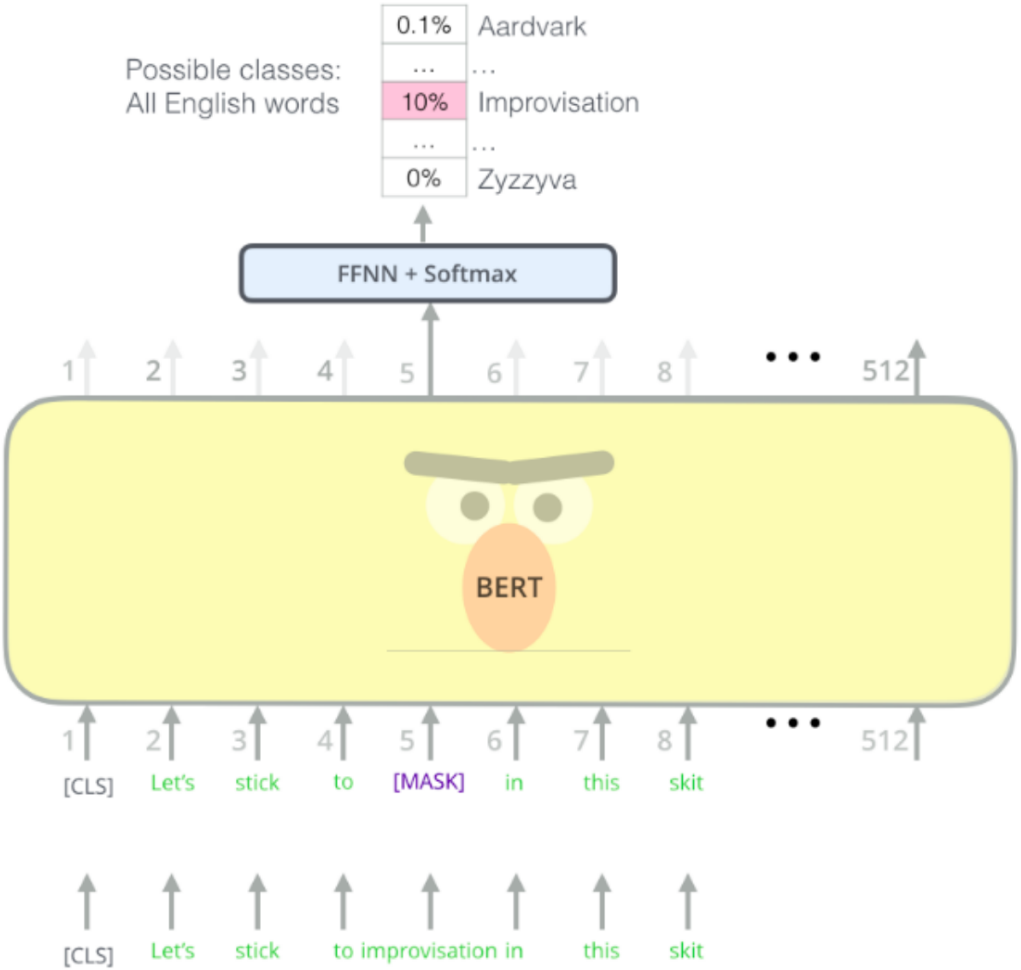
- Word2vec
- GloVe
- Td-idf

Se pueden combinar con diversos modelos de clasificación.



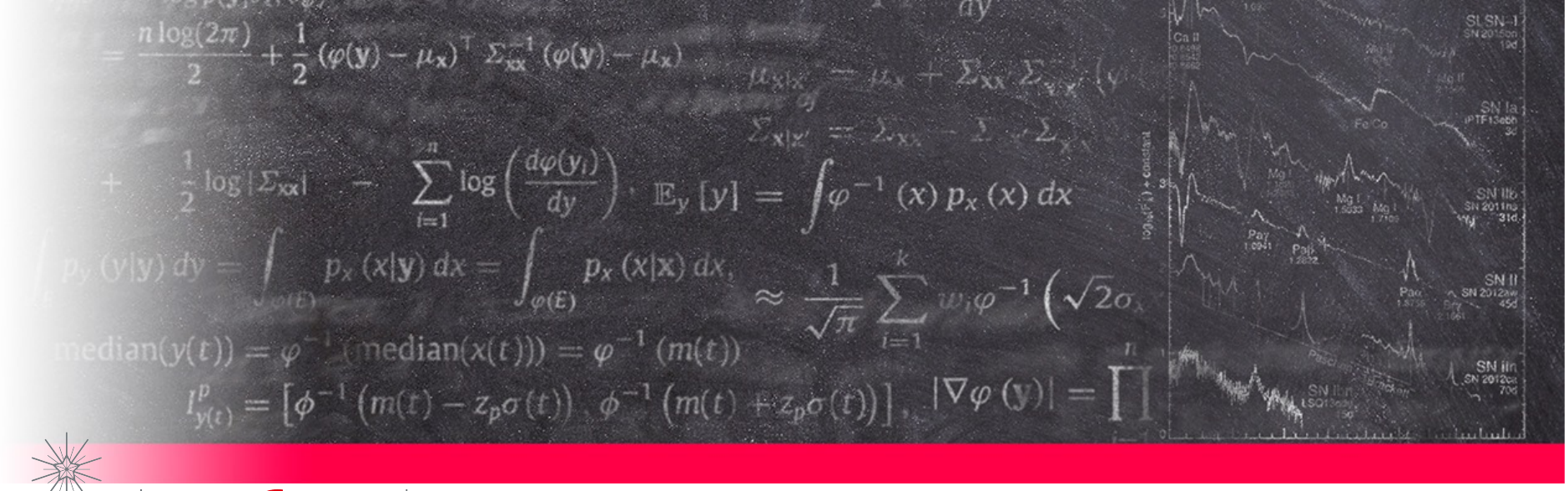
Modelos de aprendizaje profundo

BETO/BERT: modelo profundo basado en la arquitectura Transformers (como es el caso de ChatGPT). Son pre-entrenados en grandes corpuses de texto y se pueden usar en variadas tareas de NLP. Permiten vectorizar palabras y texto.



Referencias

- Alvarez, C., Zurita, G., Hasbún, B., Peñafiel, S., Pezoa, Á., Alvarez, C., Zurita, G., Hasbún, B., Peñafiel, S., Pezoa, Á. (2021). A Social Platform for Fostering Ethical Education through Role-Playing. In Factoring Ethics in Technology, Policy Making, Regulation and AI. IntechOpen. https://doi.org/10.5772/intechopen.96602
- Ramírez Rivas, P., Guerrero, S., Cerda Maureira, J., Ross, J. P., Flores Mandeville, G. (2022). La formación ética canalizada mediante la tecnología. Experiencia y resultados preliminares del uso de la herramienta web Ethicapp. XXXIV Congreso Chileno de Educación en Ingeniería.
- Metsis, V., Androutsopoulos, I., Paliouras, G. (2006, July). Spam filtering with naive bayes-which naive bayes?. In CEAS (Vol. 17, pp. 28-69).
- Blei, D. M., Ng, A. Y., Jordan, M. I. (2003). Latent dirichlet allocation. The Journal of Machine Learning Research, 3(null), 993–1022.
- Devlin, J., Chang, M.-W., Lee, K., Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 4171–4186. https://doi.org/10.18653/v1/N19-1423



Ingeniería Matemática
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Procesamiento de datos textuales de EthicApp

Presentación de avances

Camilo Carvajal Reyes

8 de junio, 2023